The National Academy of Sciences, India
is completing its fifty years
of existence
on the 23rd October 1980.
To celebrate the occasion
the Council of the Academy organised
a Golden Jubilee Session
and published this
Golden Jubilee Commemoration
Volume.

Articles contained in this Volume
have been written by
some of the most
distinguished scholars
in their respective fields,
covering a wide range of scientific studies
from bio-chemistry
to nuclear physics.

There are twenty-seven studies
contributed by
over forty scholars of national
and international fame.

GOLDEN JUBILEE COMMEMORATION VOLUME

1980

NAYA PROKASH

GOLDEN JUBILEE COMMEMORATION VOLUME 1980

THE NATIONAL ACADEMY OF SCIENCES, INDIA
ALLAHABAD
INDIA

THE NATIONAL ACADEMY OF SCIENCES
INDIA

GOLDEN JUBILEE COMMEMORATION VOLUME

1980

Editor
U. S. SRIVASTAVA, F.N.A., F.N.A.Sc.

NAYA PROKASH : CALCUTTA SIX

© National Academy of Sciences, India
October 1980

Printed and Published by
B. Mitra
NAYA PROKASH
206 Bidhan Sarani
Calcutta-700 006

Under the auspices of the
National Academy of Sciences, India
5 Lajpatrai Road
Allahabad-211 002

Price : Rupees two hundred forty only

FOREWORD

Professor Meghnad Saha, F.R.S., whose name will always be written in letters of gold in the history of Science in India, was the first scientist in the country to realise the necessity of establishing an organisation of Indian scientists with the objective not only of bringing them together on a common platform where they could discuss their scientific work and through which they could cultivate and promote science in the country, but more importantly, of organizing them to take a greater and more concrete interest in scientific problems of national importance. In October 1929, he issued an emphatic appeal to start an Academy of Science in the country.

The appeal evoked immediate response and during the session of the Indian Science Congress held at Allahabad in January 1930, a provisional committee was formed to frame the constitution, rules and byelaws of the first such Academy in India. On the 4th December, 1930, the Academy was formally registered with the name, The Academy of Sciences of United Provinces of Agra and Oudh, mainly because of the interest taken by the then Governor of U.P., Sir Malcolm Hailey. The Memorandum of Association of the Academy was signed by Prof. Meghnad Saha, Prof. N. R. Dhar, Prof. D. R. Bhattacharya and Prof. A. C. Banerji of Allahabad University and Prof. K. N. Bahl, Prof. P. S. McMahon and Prof. Chaudhary Wali Mohammad of Lucknow University. In 1931, the Academy had 57 Ordinary Members and 19 Fellows.

It was soon realized that the Academy could hardly fulfil its national aspirations with the name of a province tagged on to its name and in 1934, it was renamed as The National Academy of Sciences, India. Today The National Academy of Sciences, India is the oldest Science Academy in the country and one of the two cooperating Academies of the Indian National Science Academy, New Delhi. Its Members and Fellows, numbering nearly two thousand hail from all parts of India and include many of the most distinguished scientists of the country, besides some foreign ones.

This year, the Academy completes fifty years of its service to science in India and has reasons to feel proud of its record.

The Council of the Academy decided that apart from celebrating the Golden Jubilee of the Academy in a befitting manner by

holding a Golden Jubilee Session and embarking upon several new projects, the occasion should be commemorated by publishing a Golden Jubilee Commemoration Volume which may, in keeping with the broad-based interests of the Academy, contain review articles covering a wide area.

As readers will find, the articles contained in the Volume have been written by some of the most distinguished scholars in their respective fields. I need hardly stress the fact, that but for their enthusiastic response to our invitation to contribute to the Volume and their immense cooperation in many ways, it would have been impossible to bring it out. On behalf of the Academy, I express my deep sense of gratitude to them.

I also express sincere thanks of the Academy to Sri Barindra Mitra of Naya Prokash, Calcutta for undertaking its printing in the face of heavy odds and to Mrs Mitra for personally seeing the work through to completion.

U. S. Srivastava
Editor and Convener
Golden Jubilee Commemoration
Volume Committee

Allahabad
October 20, 1980

Prof. Meghnad Saha
(6th October 1893 - 16th February 1956)

Prof. Nil Ratan Dhar
(Born 2nd January 1892)

CONTENTS

COMPUTERIZATION OF LOGIC **G. N. Ramachandran**	1–42
SPIROPLASMAS OF PLANTS AND ANIMALS : AN OVERVIEW **Karl Maramorosch**	43–62
GENETICS OF HUMAN SERUM LACTATE DEHYDROGENASE —A REVIEW **Bijaya Kumar Swain, Archana Sharma and Geeta Talukder**	63–96
THE ROLE OF LIGHT IN CONTROL OF PLANT GROWTH AND DEVELOPMENT **S. C. Maheshwari, Vipin Rustogi, Kusum Malhotra and Sudhir K. Sopory**	97–128
WAVE FORCES ON OFFSHORE STRUCTURES **S. Narasimhan**	129–180
A FUNCTIONAL APPROACH TO THE ORIGIN OF LIFE PROBLEM **Krishna Bahadur, S. Ranganayaki, Clair Folsome and Adolph Smith**	181–198
MEMBRANE MODIFYING LINEAR POLYPEPTIDE ANTIBIOTICS **R. Nagaraj and P. Balaram**	199–218
WILD LIFE RESOURCES OF INDIA **Madhav Gadgil**	219–246
ON THE CARRIAGE AND RELEASE OF STORES FROM AIRCRAFT **Sridhar M. Ramachandra**	247–270
ALGAL HYDROBIOLOGY IN INDIA **Braj Nandan Prasad and Yashpal Singh**	271–300
COMPUTERS, TOPOLOGY AND DRUG DESIGN **E. V. Krishnamurthy**	301–310
SOME ASPECTS OF ACETYLCHOLINESTERASES IN INSECTS **K. N. Mehrotra**	311–320
C_4-PHOTOSYNTHESIS : AN ADAPTATION TO TROPICAL ENVIRONMENTS **V. S. Rama Das and I. Madhusudana Rao**	321–332

STABILITY OF PARALLEL FLOW THROUGH A POROUS MEDIUM USING BRINKMANN MODEL — 333-348
N. Rudraiah

SYNOVIAL JOINT LUBRICATION : A REVIEW — 349–368
Peeyush Chandra

BICARBONATE STIMULATION OF ELECTRON FLOW IN THYLAKOIDS — 369–378
Doug Jordan and Govindjee

FEMALE GENITAL ANATOMY AND THE MORPHOGENESIS OF FOETAL MEMBRANES OF CHIROPTERA AND THEIR BEARING ON THE PHYLOGENETIC RELATIONSHIPS OF THE GROUP — 379–428
A. Gopalakrishna and K. B. Karim

SOLID-STATE BATTERY—PROSPECTS AND LIMITATIONS — 429–462
Suresh Chandra and Rakesh C. Agrawal

RADIOISOTOPES IN MEDICINE — 463–494
R. S. Mani

A REVIEW ON QUARANTINE FOR SEED — 495–530
Paul Neergaard

THYROID IN REPTILES — 531–540
J. P. Thapliyal

USE OF POLYGRAPH AS AN AID IN TRACING AFFINITIES IN POLYGONACEAE — 541–550
A. K. Sharma and S. N. Sahay

SCALES IN ATMOSPHERIC SCIENCE — 551–572
M. P. Singh

THE LIVING MUTAGENS — 573–606
G. K. Manna

PERSPECTIVES IN THE DEVELOPMENT OF RADIATION CHEMICAL CONCEPTS AND MECHANISMS — 607–680
K. N. Rao and P. N. Moorthy

BIOTECHNOLOGICAL APPLICATIONS OF PLANT TISSUE AND CELL CULTURES : PROBLEMS AND PROSPECTS — 681–708
M. S. Chadha and M. R. Heble

REPRODUCTIVE ENDOCRINOLOGY OF DECAPOD CRUSTACEA — 709–746
R. Nagabhushanam, G. K. Kulkarni and R. Sarojini

THE NATIONAL ACADEMY
OF SCIENCES
INDIA

GOLDEN JUBILEE
COMMEMORATION
VOLUME : 1980

Computerization of Logic*

G. N. Ramachandran**

(*Dedicated to Professor Sir C. V. Raman from whom the author learnt the methods of Scientific research)

**Mathematical Philosophy Group, Indian Institute of Science, Bangalore 560 012

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

Computerization of Logic

1. *Introduction*

Looking at the title of this paper, it would appear at first sight, that it is superfluous to say anything on this subject, because all computers work on the basis of electronic circuits that correspond to logical connectives. In fact, the computer circuits themselves are known by the name of "logic circuits". Therefore, I shall, first of all, explain the exact scope of the study reported in this paper—namely that of obtaining a computerization of the logical processes used in arguments and proofs.

Logic may be defined as the "Science of Reasoning", and one may look at the rules of logic, exactly as a physicist would look at the rules governing, let us say, the motion of a particle under the influence of other bodies. Thus, the best method of listing and stating the laws and rules of logic is to examine the origins of our thinking processes, and then to state them in the form of what may be called "laws", "principles", or "rules". In fact, this is what is done in many standard books on mathematical logic, and a reasonably satisfactory formulation of logical principles is available therein (see for example Refs. 1, 2, 3, 4). However, when we attempted to convert these formulae for practical application in an automatic machine, either *via* simple electronic circuits in an analog computer, or *via* algorithms making up simple programs in a digital computer, we found that the rules available in the literature had to be restated in a different form so as to become capable of being readily manipulated in such a practical manner.

One may give an analogy here. We know that all the laws and theorems of geometry are contained in the well-known formulation of Euclid, as given in his famous "Elements". However, when it comes to practical problems in engineering or physics, no one uses the rigorous formalism of Euclidean geometry. Instead, we employ Cartesian coordinate geometry, and work out our problems in terms of analytical techniques employing these. What we have proposed, and partially demonstrated, is a completely analogous procedure, but applied to methods of solving practical problems in logic. The presentation, in terms of axioms, etc., available in books on mathematical logic are self-contained, and mostly sufficient, and the formulations therein

are satisfactory, but they can be applied only by an expert who knows the subject very well. On the other hand, we have shown that these logical principles can be reformulated in such a form that they can be used for checking arguments and deductions, in a manner that can be carried out by a relatively inexperienced person. In this technique, the various steps in any argument are first represented by a simple graph drawn on paper (see Fig. 2 of Section 4) interconnecting the various statements (called "terms") that are relevant, which we denote as "variables". These are related by logical connectives, which we term as "relations" or "operations", and they are the connecting links between the variables, representing the terms, in the graph. The various steps in this graph are then implemented on our analog computer, or in the digital computer program, and they are then executed to give the output sought for. In fact, we found that the electronic circuits used for the design of the analog computer also turn out to be, probably, the most satisfactory basis for computerizing the same steps on a digital computer, and working out the necessary algorithms *via* various logical functions representing the connectives.

This process employs a mathematical technique which is rather interesting. It is based on the "Theory of Relations", which is used to represent the properties of logical connectives and a simple relation like

"**a** is equal to **b**" (1a)

or

"**a** implies **b**" (1b)

These can be expressed in terms of relations E or Y, connecting **a** and **b**, as mathematical equations

$$\mathbf{a}\ E = \mathbf{b}\ ; \qquad \mathbf{a}\ Y = \mathbf{b} \tag{2}$$

Naturally, one is tempted to ask the question—"What is the corresponding relation in reverse ?". It turns out that this question is answered very simply in our formulation. In fact, it leads to a large expansion of the scope of classical logic. In the two example given above, the corresponding reverse relations are correspondingly :

"**b** is equal to **a**", or $\mathbf{b}\ \overleftarrow{E} = \mathbf{a}$, leading to $\mathbf{b}\ E = \mathbf{a}$ (2a)

"**b** implicates **a**" or $\mathbf{b}\ \overleftarrow{Y} = \mathbf{a}$, leading to $\mathbf{b}\ V = \mathbf{a}$ (2b)

where the reversed E (or $\overleftarrow{E}$) is the same as the forward E itself, and the reverse of the relation "imply" (Y) is "implicates" (V). At this moment, we are not giving any meaning or significance, to the words 'implies' and "implica-

tes", apart from what is given in normal language—"If **a** implies **b**, then **b** should implicate **a**". In any case, it will be clear that every relation has, in principle, a reverse relation, and the mathematical representation of logic will be complete only if one takes into account the full set composed of both the direct and reverse connectives of all types.

The main result that comes out when this is done is, that it is not enough to have only two states,—namely true (T) and false (F)—for an atomic statement, or term, in logical arguments. It is found that two more states are needed, which may be denoted by the words "doubtful" (D) and "contradictory" (X). The full arguments for this is given in Section 3. When both the introduction of the additional two states leading to four possible states for a term, and the introduction of a reverse connective for every forward connective, are made, the necessary apparatus for carrying out the representation of any logical argument by means of mathematical symbols, or entities, becomes complete, for all purposes. The aim of this paper is to give a very brief outline of the ideas underlying these developments (given in Sections 2, 3, and 4), followed by a description of an analogue computer designed to simulate logical arguments (Section 5), in which the final conclusions (in terms of truth values) of a connected argument can be worked out, or the validity of a connected set of statements can be checked, and, if it is invalid, the exact source wherein the invalidity of the argument arises can be detected and indicated. In Section 6, we shall indicate very briefly how these ideas can be converted into computer programs, so that even more complicated problems in logic, than those that can be manipulated upon in the analogue computer, can be worked out and the logical steps can be automatically checked, the results verified, and inconsistencies displayed. Section 7 gives a preliminary account of the types of logic circuits required for First Order Quantified Predicate Logic.

2. *States and Connective Operators in Classical Logic*

(a) *Equivalence, negation and similarity*

We shall begin the mathematical part of this paper by stating the basic rules of classical logic in a form that will be capable of being generalised for the later discussions. Thus, symbols in lower case Roman letters, like **a**, **b**, etc., will be used to represent the "terms" in a logical sentence. The only "states", or truth values, that can occur in classical logic are either T or F. If **a** is in

the state T, then it will be represented by a_T; similarly for a_F, b_T, b_F, etc. It is obvious that the logical connective in a sentence can be represented as a relation between the states of two or more terms. Taking, for example, the simplest of such connectives—namely "equivalence"—we say that "**a** is equivalent to **b**" by writing it as an equation :

$$\mathbf{a}=\mathbf{b} \tag{3}$$

We are not using the symbol $\equiv$ for equivalence in this expression, following the same practice as what is usually followed in algebra, where many equations that are really identities are often expressed by using a simple "equal to" (=), rather than the symbol for an identity ($\equiv$).

The next simplest relation that we can think of is that of "negation". When we talk of **b** being the negation of **a**, we express it in symbols, in either of the two forms given in (4) below :

$$\neg\mathbf{a}=\mathbf{b}, \text{ or } \mathbf{a}\neq\mathbf{b} \tag{4}$$

The significance of $\neg\mathbf{a}$ is clear—namely that, if **a** happens to be true, then $\neg\mathbf{a}$ is false, and *vice versa*. By equating $\neg\mathbf{a}$ with **b**, **b** will be true when **a** is false, and **b** will be false when **a** is true. This is the *unary relation* between **a** and **b** corresponding to the *logical operation of negation*.

The above are in the nature of definitions, and we only use an obvious notation to represent ideas which everyone uses, and they require no special justification.

We now come to the next possible type of connection that can occur—namely a *binary relation*. Consider the relation between **a** and **b** when the equation (3) itself is true. We may designate this operation by *S*, signifying the relation of "similarity", and write (3) as :

$$\mathbf{a}\ S\ \mathbf{b}=\mathbf{c}, \qquad \mathbf{c}=\mathrm{T} \tag{5}$$

Then, Eq. (4), corresponding to **a** being different from **b**, may be written in the form

$$\mathbf{a}\ S\ \mathbf{b}=\mathbf{d}, \qquad \mathbf{d}=\mathrm{F} \tag{6a}$$

We may also use *SN* as the symbol for the relation of "non-similarity" and write (4) as

$$\mathbf{a}\ SN\ \mathbf{b}=\mathbf{e}, \qquad \mathbf{e}=\mathrm{T} \tag{6b}$$

In the above Eqs. (5) and (6), the resultant given by the relation between **a** and **b** is expressed by another logical entity **c**, **d**, or **e**, as the case may be, which itself can have truth values of either T or F. In particular, the truth values of these, as given in (5) and (6), correspond to the unary relations "equal to", and "not equal to", in (3) and (4) respectively.

We may also enquire as to what is the relation between the two relations S and SN, which occur in (5) and (6). Without going into details, the following two Eqs. (7a) and (7b) are obvious

$$(\mathbf{a}\ SN\ \mathbf{b}) = \neg\,(\mathbf{a}\ S\ \mathbf{b}) \tag{7a}$$

or

$$\mathbf{e} = \neg\,\mathbf{d} \tag{7b}$$

where **e** and **d** are as defined in (6a) and (6b). We thus have an interesting interrelationship between the binary relations represented by the symbols S and SN, and the symbol $\neg$ giving the unary relation N (see below) in (4). Because of this, the relation SN is superfluous as a fundamental relation, and its effect can be obtained by saying that, when there is non-similarity, the relation **a** S **b** has the truth value F. This is obviously what is implied by appending the symbol N to S.

We have purposely used different names "similar" and "non-similar" given by the symbol S and SN in (5) and (6), to distinguish them from the ideas of "equivalence", and of "non-equivalence" or "negation relation", that can exist between **a** and **b**, as given in Eqs. (3) and (4). In fact, the corresponding unary relation in (3) and (4) can also be written as follows, using the symbols E and N, for equivalence and negation :

$$\mathbf{a}\ E = \mathbf{b} \tag{8a}$$

$$\mathbf{a}\ N = \mathbf{b}' \tag{8b}$$

The Eqs. (8a) and (8b) symbolically represent that the operation of equivalence (E), or negation (N), when applied to **a** gives the term **b** or **b**′, as the case may be, just as the operation of checking for similarity (S), or non-similarity (SN), between **a** and **b**, leads to the two terms **c** and **e** in (5) and (6b).

We have thus briefly introduced the use of the mathematical idea of a relation to represent the logical concept of a connective between two terms, say **a** and **b**. We shall use this formalism extensively in what follows. Actually, the equations given above, starting from (3) and going up to (8), can all be written in terms of matrices. This is discussed in detail in the book on Practical Logic being written by the author (Ref. 5). We shall make references to this book occasionally, although it is expected to be published only about a year from now. However, we shall try to make the argument in this paper as complete as possible, and independent of the exact methods adopted in the detailed treatment of the book. However, of necessity, most of the derivations given below will have to be brief and only the main results will be given, with the intermediate steps mostly omitted.

(b) *Conjunction and disjunction*

We now come to the very common connectives used in mathematical logic—namely "and" (A) and "or" (O)—usually given also the names of "conjunction" and "disjunction" respectively. In order to formulate the traditional approaches in our system, we shall consider the corresponding relations given by (9) and (10) :

$$\mathbf{a}\ A\ \mathbf{b} = \mathbf{c} \tag{9}$$

and

$$\mathbf{a}\ O\ \mathbf{b} = \mathbf{d} \tag{10}$$

The simplest way of defining the operators A and O is that they correspond to the two truth tables, connecting the truth values of **c** and **d** with those of a and b, as given in Table 1(a) and 1(b) respectively.

(a) "and", A			(b) "or", O		
a / **b**	T	F	**a** / **b**	T	F
T	T	F	T	T	T
F	F	F	F	T	F

Table 1

Truth Tables for the connectives "and" and "or"

Thus, **a** A **b** is *true* only if *both* **a** and **b** are *true*, while **a** O **b** is *false* only if *both* **a** and **b** are *false*. Hence the two connectives have some sort of a dual relationship between each other. It is extremely interesting that this feature is also exhibited by the operations of forming the product, and the sum, in the simple Boolean algebra composed of two elements 0 and 1. The two tables giving the product and the sum of two Boolean variables a and b, are given below as Tables 2(a) and (2b).

(a) *Product*			(b) *Sum*		
a / b	1	0	a / b	1	0
1	1	0	1	1	1
0	0	0	0	1	0

Table 2

Tables showing the results of the operations of Boolean multiplication and addition

It is readily seen that the two pairs of tables, in Table 1 and Table 2, have a complete one-to-one correspondence with one another, if we make the following correspondences :

$$\begin{aligned} &T \longmapsto 1, \qquad F \longmapsto 0 && \text{(a)} \\ &\text{Conjunction} \longmapsto \text{Boolean product} && \text{(b)} \\ &\text{Disjunction} \longmapsto \text{Boolean sum} && \text{(c)} \\ &\neg\, a \longmapsto \sim a && \text{(d)} \end{aligned} \qquad (11)$$

In this, $\sim a$ corresponds to the complement of a. If a is 1, $\sim a$ is 0, and if a is 0, $\sim a$ is 1. In fact, this formal similarity between the connectives in logic and the operations of relating two Boolean elements either by multiplication, by addition, and by complementation has been fully utilised in the monograph on Boolean algebra by Halmos (Ref. 6; see also Appendix B of the book by Copi, Ref. 3).

As shown by Halmos, there is also a homology between Boolean algebra and the formalism of set theory. If all these are taken into account, we obtain at once the well-known De Morgan's law connecting the negation of the operators "and" and "or", with the "or" and "and", respectively, of the negated variables themselves. These are given in (12) and (13) below :

$$\neg(\mathbf{a}\ A\ \mathbf{b}) \Longleftrightarrow (\neg\mathbf{a})\ O\ (\neg\mathbf{b}) \qquad (12)$$

$$\neg(\mathbf{a}\ O\ \mathbf{b}) \Longleftrightarrow (\neg\mathbf{a})\ A\ (\neg\mathbf{b}) \qquad (13)$$

Here, we use the symbol $\Longleftrightarrow$ to emphasize the logical equivalence, or identity, of the left and right hand sides in the two equations.

Therefore, we can define the effective connectives which occur on the left hand sides of (12) and (13) by the names "nand" and "nor". If we wish to refer to these connectives explicitly, we shall use the symbols *AN* and *ON* for these and denote the left hand sides of (12) as $\mathbf{a}\ AN\ \mathbf{b}$ and $\mathbf{a}\ ON\ \mathbf{b}$, and read them as "**a** and **b**, negated", and "**a** or **b**, negated", as the case may be.

(c) *Implication*

For all practical purposes, the operators defined above are sufficient to discuss any problem in classical logic, and all of them are capable of being written as relational operators, which are combinations of negation, conjunction, and disjunction. (In fact it has been shown that just one operator, "nand" is, in principle, sufficient to obtain the effects of all the other operators, when it is suitably combined with itself and with the true and false states of any term). However, right from the ancient ages, logic has always dealt with ideas of "deduction"—i.e., the obtaining of a new logical result, or truth

value of a new term, in terms of the logic of an old result, expressed in terms of the truth value of a known statement. Although, in most books on logic, implication is treated as a binary operation, we shall find the need to give it also a unary representation, as will be seen below.

Taking the standard binary representation of implication and using the symbol *I* for implication, we can write it in different equivalent ways. The equivalences between these different ways of expressing the logical content of a statement of implication, using the operations of implication, conjunction and disjunction, are shown in (14) :

$$\mathbf{a}\ I\ \mathbf{b} \Longleftrightarrow \neg\mathbf{a}\ O\ \mathbf{b} \Longleftrightarrow \neg(\mathbf{a}\ A\ \neg\mathbf{b}) \qquad (14)$$

These are well-known and are discussed thoroughly in every textbook on logic, so that we do not give any derivation, or explanation, of these. However, in order to show how ancient these ideas are, we shall quote the following sentences from the Buddhist text *Kathavathu* (of the early centuries BC) which gives the different ways in which an implication can be stated. The relevant sentence is the following :

> "If my first statement (**a**) was true, then you should consent to my second (**b**). You are wrong to say that **a** is true and **b** is not true. If **b** is not true, then **a** is not true···."

As is readily seen, the equivalence $\mathbf{a}\ I\ \mathbf{b} \Longleftrightarrow \neg(\mathbf{a}\ A\ \neg\mathbf{b})$ in (14) is clearly stated in this ancient text. What is even more interesting is that it states also the equivalence (15), which is given below :

$$\mathbf{a}\ I\ \mathbf{b} \Longleftrightarrow (\neg\mathbf{b})\ I\ (\neg\mathbf{a}) \qquad (15)$$

It may be worthwhile showing how simply the equivalence of the two expressions in (15) may be deduced from Eq. (14). For this, we start with

$$\mathbf{a}\ I\ \mathbf{b} \Longleftrightarrow \neg(\mathbf{a}\ A\ \neg\mathbf{b}) \qquad (16a)$$

in (14). Since the operation of "and" and "or" are completely commutative (as can also be verified from the truth tables given in Table 1), we can write the equivalences :

$$\neg(\mathbf{a}\ A\ \neg\mathbf{b}) \Longleftrightarrow \neg(\neg\mathbf{b}\ A\ \mathbf{a}) \Longleftrightarrow \neg(\neg\mathbf{b}\ A\ \neg(\neg\mathbf{a})) \qquad (16b)$$

Now inverting the process in (16a), we see that

$$\neg(\neg\mathbf{b}\ A\ \neg(\neg\mathbf{a})) \Longleftrightarrow (\neg\mathbf{b})\ I\ (\neg\mathbf{a}) \qquad (16c)$$

Hence, what looks so obvious in common language—namely that the statement, "If what I say is true, then what you say must be true", is the same as, "If what you say is false, then what I say must be false"—is really so because the two *conditional* statements are both related in the same way with a common *definite* statement, namely, "It cannot be true that what I say is true,

and what you say is false". It is very interesting to note that the statements given in the proof of (15), *via* (16a), (16b) and (16c), are all contained in the quotation from *Kathavathu*.

It should be mentioned that the implication, as given in the sentences stated above, are all of the form "if···then", (that is in the form of a conditional statement), while conjunction and disjunction, to which it is related, involve statements which are definite. The fact that, by making a suitably designed definite statement, we can obtain, as its equivalent, a conditional "if ···then" statement, is a matter that requires deep thought. Even if we are not able to fully subscribe to the nuances, or consequences, of the above equivalences with reference to any particular situation in a real argument, we may always take (14) and (15) to refer to the *definition* of "implication" as given therein. Thus, in mathematical logic, an implication of **b** by **a** is *defined* as being given by the corresponding expression in (14), which involves a simple straightforward statement having either an A or an O.

To make this definition clear, we give below in Table 3(a), the truth table for **c**, given by **a** I **b** = **c**.

(a) **a** implies **b**

a \ b	T	F
T	T	F
F	T	T

(b) **a** implicates **b**

a \ b	T	F
T	T	T
F	F	T

Table 3
Truth tables for "implies" and "implicates"

From this, we see that the implication of **b** by **a** has the truth value F only if **a** is T and **b** is F, which is what is meant by (14) and (16a). Hence, if **a** is T, then **b** can only be T ; or **a** being T is a *sufficient* condition for **b** to be true.

Just as we can talk of **a** being a *sufficient* condition for **b** to be true, we can also think of **a** being a *necessary* condition for **b** to be true. In that case we say that "**a** implicates **b**". The corresponding truth table is shown in Table 3(b), and it is readily seen that this condition requires that **b** will be false only if **a** is false. If **a** is true, then nothing can be deduced from this. This is, in a sense, the opposite of the sufficient condition in "**a** implies **b**". We shall not discuss these further now, as they will come up again for discussion in Section 3(e).

(d) *Unary operation of "if"*

All the above results are acceptable, and obvious ; but the place where there is some difficulty in appreciating the range and scope of the operation "implies" is when **a** is F. In that case, both b_T and b_F are possible, and either occurrence corresponds to the truth value T for **c** = **a** I **b**. Suppose now we make a truth table for the *unary relation* "if **a**, then **b**", it will be as shown below in Table 4(a) :

(*a*) *"If...then"* (*Y*) (*b*) *"only if"* (*V*)

a	b: Logical State	b: Boolean Representation
T	T	1
F	T or F (≡D)	1 or 0 (=1/2 say)

a	b: Logical State	b: Boolean Representation
T	T or F (≡D)	1 or 0 (=1/2 say)
F	T	0

Table 4
Truth tables of the unary operations "if" and "only if"
For the names of the symbols Y and V see the next subsection 2(e).

The first row of Table 4(a) is straightforward, and it is the significant part of the logic of an implication statement—namely "if a_T, then b_T". The second row leads to uncertainty—namely, "if a_F, then either b_T or b_F is possible", or *nothing definite can be said of* **b**. We shall call this as the "doubtful" state D (b_D) of **b**, and it may be taken to have a "truth value" of $\frac{1}{2}(1+0)=\frac{1}{2}$, midway between 1 for T and 0 for D. This is a new logical state that has necessarily to be introduced if we include the interpretation of "if···then" as a unary relation. We will see later (Sections 3(d) and (e)) that this D state comes in also when we consider the reverse of the binary relations A and O. Before we discuss this state D in some detail we may first consider the reverse of the unary relation "if···then".

(e) *Reverse Unary Implication*

The method of symbolizing the unary operation, "implies", is best understood by nothing that, just like the statement "**a** negated is equal to **b**", the logical sentence, "**a** implies **b**", also relates **a** to **b** by the operation of "unary implication" acting on **a** to give the term **b**. In order to distinguish this unary operator from the binary operator I in (14), (15) and (16), we shall use the symbol Y for this. (In this, Y stands for the first letter of the

Sanskrit word *Yadi*, which means "if"). Thus, we have for the statement "**a** implies **b**", or "if **a**, then **b**", the mathematical representation

$$\mathbf{a}\ Y = \mathbf{b} \tag{17}$$

similar to $\mathbf{a}\ E = \mathbf{b}$, and $\mathbf{a}\ N = \mathbf{b}$, which we have already considered.

Naturally, as in the case of the unary operators E and N, we can enquire as to what is the operator in the reverse direction, for Y? If we denote this reverse operator by $\overleftarrow{Y}$, then we can say that (17) yields the statement (18), in the reverse direction :

$$\mathbf{b}\ \overleftarrow{Y} = \mathbf{a} \tag{18}$$

Let us find out what is the nature of $\overleftarrow{Y}$. In ordinary language, the answer to this is pretty clear. It can be stated in several forms

(a) **b** implicates **a**
(b) **b** is implied by **a** (18a)
(c) Only if **b** is true, is **a** true.

Thus, just as the word 'if" represents the operator Y, the simple words "only if" represent $\overleftarrow{Y}$. We may signify this again by means of a simple Sanskrit word "*avasya*" (meaning "necessary" or "required") and use the first consonant V in this name, for the operator representing the reversal of Y. Hence (18) takes the form $\mathbf{b}\ V = \mathbf{a}$, so that we have the following equivalence connecting the two operators Y and V :

$$\mathbf{a}\ Y = \mathbf{b} \Longleftrightarrow \mathbf{b}\ V = \mathbf{a} \tag{19}$$

The mathematical significance of V is seen by just transposing the truth table given above for Y, by interchanging **a** and **b**, as is shown in Table 3(b). This immediately gives the truth table for the unary operator V, "only if", which is shown in Table 4(b).

The two unary operators Y and V, respectively, are the logical connectives used for the condition of "sufficiency" and "necessity" in mathematical arguments—Y is used when the consequent (**b**) is true, *if* the antecedent (**a**) is true, and the connective V is used when the consequent (**b**) is true *only if* the antecedent (**a**) is true. Following standard practice, we shall designate the operators Y and V for the unary "if" and "only if" also by a forward arrow ($\Rightarrow$) and a reverse arrow ($\Leftarrow$), respectively, where needed.

In this way, we have covered almost all the commonly used connectives in classical logic. The most interesting new concept that has been introduced is that of the doubtful state, as shown in Table 4. It is seen that, for

$Y = \mathbf{b}$, a_T leads to b_T, while a_F gives only the doubtful state b_D, so that the term **b** is no longer useful in producing a definite consequence. This is the case for the necessary condition (Y) standing for the unary connective "if". On the other hand, when $\mathbf{a}\ V = \mathbf{b}$ holds, a_F leads to b_F, but a_T gives only b_D, which is of no consequence, which occur for the sufficient condition (V) in the unary connective "only if". As shown above, the necessary and sufficient conditions are related by the fact that "**a** is sufficient for **b**" is the same as "**b** is necessary for **a**", and *vice versa*.

3. *New Connectives in the Logic of Statements*

(a) *The connective for consistency—"with"*

In general, the above connectives are sufficient for writing down any general argument in the form of a connected set of equations written in symbolic logic. However, there are certain unstated premises which have been taken into account while working with these sentences. We shall consider below some of these explicitly and indicate their importance and need. We shall also give these connectives suitable names and list their properties briefly.

Suppose we have two different lines of argument, one starting with the term **a** and the other starting with the term **b**, both finally ending up giving some truth value for the term **c**. Let us conveniently call these as $t(\mathbf{c}_a)$ and $t(\mathbf{c}_b)$, where the symbol t stands for the name "truth value", and t has a value 1 when the term is true, and $t = 0$ when it is false. Situations of this type are very common in mathematics. However, it may happen that theorem **c** is one for which a proof of its validity (state c_T) is available, starting from the premise **a**, while the premise **b** does not imply **c**. In other words, if **b** is true, the truth of **c** does not follow uniquely, and **c** may be either true or false. Under these circumstances, the common procedure is to take the theorem to be *valid*, since it has been derived from a valid statement **a**, (which has not been questioned by anything following from **b**), and therefore, the truth value of the theorem **c** is deduced to be T.

We have already given the name D for the logical state of a statement which is either true or false. Therefore, it follows that, in this type of argument, one line of argument gives a positive result T, while the other gives the doubtful state D ; but the exact result T prevails when the two are put together. Obviously a similar result will prevail even if one line of argument gives the truth value F for **c**, and the other gives the truth value D. Once again, the

definite result F will prevail over the indefinite D. On the other hand, if the two lines of approach give, for the same theorem **c**, opposite truth values, —namely $t(c_a)=$ T and $t(c_b)=$ F respectively—then, there is clearly a *contradiction*. That is to say, it is impossible for **c** to be T and F at the same time. A procedure of defining a suitable logical connective, for which such a superposition of truth values T and F will lead to a contradiction, while the superposition of T and D, or of F and D, will lead only to the definite state (T or F, as the case may be) has been devised by us. We denote this connective by the name "with" and use the symbol W to indicate this. Thus, if $\mathbf{c}_1$ and $\mathbf{c}_2$ are the two results for **c**, (say arising from the premises **a** and **b**), then the final result for **c** is

$$\mathbf{c}=\mathbf{c}_1 \text{ "with" } \mathbf{c}_2=\mathbf{c}_1 W \mathbf{c}_2 \tag{20}$$

The operator W has, obviously, the following properties. If both $\mathbf{c}_1$ and $\mathbf{c}_2$ are T, then $\mathbf{c}_1 W \mathbf{c}_2$ is also T. If both $\mathbf{c}_1$ and $\mathbf{c}_2$ are F, then $\mathbf{c}_1 W \mathbf{c}_2$ is F. But if one is T and the other is F, then $\mathbf{c}_1 W \mathbf{c}_2$ is a contradiction, which we shall indicate by the symbol X. One more possibility is that, if either $\mathbf{c}_1$ or $\mathbf{c}_2$ has a "primary" state (T or F), while the other one is the "doubtful" state D, then the primary state will prevail and the result is T or F, as the case may be, and D does not matter. All these are conveniently given in the truth table below for the operator W.

$\mathbf{c}_1$	$\mathbf{c}_2$	$\mathbf{c}=\mathbf{c}_1 W \mathbf{c}_2$
T	T	T
F	F	F
T	F	X
F	T	
D	T	T
T	D	
D	F	F
F	D	

Table 5

Truth table for the connective "with" (W)

Although the state D as such is not explicitly there in classical logic, we have indicated how it could arise as a result of an implication, for which the input statement is false. In such a case, the implication gives nothing about the output statement **c**, leading to the latter being in a D state. On the other hand, if **c** is implied by some other true statement **a**, then it also becomes true, because $a_T\ Y=c_T$.

It should be mentioned that the introduction of the connective "with" in our formulation is not really a new idea, but only a precisely formulated statement of an idea that is generally used by everybody (including authors of treatises on mathematical logic), in working out the results of arguments. Thus, under certain conditions, an argument is usually said to be invalid, because it leads to both truth (c_T) and falsehood (c_F) at the same time for the same term (**c**). We say that, in such a case, the final output **c** is X ($=c_T W c_F$), and therefore the arguments that led to c_X are inconsistent. We shall show later how this result can be put to use, for reversing an argument and finding out the source of the incompatibility or inconsistency. In fact, it could lead finally to a very useful, positive, statement, such as the final step in the proof of a theorem by the technique of *reductio ad absurdum*.

(b) *The connective operator for "unconnectedness"*

In the last subsection, we had considered a situation in which the arguments leading to a statement **c** were mutually contradictory. We shall now consider another situation in which a string of arguments, leading from a premise **a** to a statement **c**, may be such that nothing definite can be stated about **c**, whatever be the input state of **a**, be it T, F or D. (By necessity, X cannot be a valid input in any step, for it will only lead to the state X also in the output—see Section 5 for more details).

Suppose we have two statements

$$\mathbf{a} \Rightarrow \mathbf{b}, \quad \neg \mathbf{b} \Rightarrow \mathbf{c} \tag{21}$$

The corresponding truth table for **a**, **b** and **c** are shown below in Table 6. This truth table represents the outcome **c** corresponding to the three possible states T, F, D of **a** and is for the situation in which $\mathbf{a} \Rightarrow \mathbf{b}$ is followed by $\neg \mathbf{b} \Rightarrow \mathbf{c}$, which may be written as shown in the heading of Table 6.

a	**b**	$\neg$**b**	**c**
T	T	F	D
F	D	D	D
D	D	D	D

Table 6

Truth table for the connection between **a** and **c** for $(\mathbf{a} \Rightarrow \mathbf{b})$ and $(\neg \mathbf{b} \Rightarrow \mathbf{c})$.*

* It leads to **a**)(**c**, or **a** $CN=$**c** (see below).

We see that the consequence of the two unary relations $a \Rightarrow b$ and $\neg b \Rightarrow c$, when taken together, is that, although the state of **a** can be given precisely, the state of **c** is always the doubtful state D, whatever may be the input state of **a**, (T, D or F). This means that **a** is unconnected to **c**, and we may express this symbolically by the symbol **a**)(**c**. Thus we have Eq. (22) below, namely,

$$(a \Rightarrow b) \text{ and } (\neg b \Rightarrow c) \Longleftrightarrow (a)(c) \qquad (22)$$

The symbol for unconnectedness has two brackets pointing in opposite ways, indicating that **a** has nothing to do with **c**. We shall also use the symbol *CN* ("not connected", with the symbol *C* for "connected") for this logical relation, and we shall express it as a unary operator *CN* in Eq. (23)

$$(a)(c) \Longleftrightarrow (a \text{ CN} = c) \qquad (23)$$

whose truth table consists of the first and fourth columns of Table 6.

(c) *Reverse "or"*

We shall now consider another operator which can also lead to the state D for the output, under suitable conditions. This is the reverse binary operator, "reverse or" ($\overleftarrow{O}$). This is defined as follows :

Definition :

$$\text{If } a \; O \; b = c, \text{ then } b = c \; \overleftarrow{O} \; a \qquad (24)$$

It is very easy to write the truth table for this in terms of the truth table for the forward "or", *O*. We shall use essentially only the classical truth values *T* and *F* for the input ; but we shall find that the output statement becomes doubtful, under some conditions, and even contradictory under others.

c	a	b
T	T	D
T	F	T
F	T	X
F	F	F

Table 7

Truth table for the "reverse or", $c \; \overleftarrow{O} \; a$

This table is extremely interesting, in that all the four states we had discussed above, T, F, D and X, occur as a result of the "reverse or" connection between **c** and **b**, for different input states. Obviously, if **a** is true, **c** will be true irrespective of whether **b** is true or false, since "**a** or **b**" becomes T in both cases. Hence the state D for **b** in the first row. In the same way, if **c** is false and **a** is true, the situation is impossible, because **c** can only be false if both **a** and **b** are false. Hence the output X for in the third row. The other two cases are obvious. If **c** is T, and **a** is F, **b** has to be T, in order to make **c** true. On the other hand, if **c** is F and **a** is F, once again **b** has to be F, in order to make **c** false. It is interesting that, for the particular case of $a = F$, the truth value of **c** carries over into the truth value of $b = c \; \overleftarrow{O} \; a$.

(d) *Reverse "and"*

Exactly as we defined the operator "reverse or", we can also define a similar reverse binary operator namely "reverse and" ($\overleftarrow{A}$), which has the following definition :

Definition :

$$\text{If } a \; A \; b = c, \text{ then } b = c \; \overleftarrow{A} \; a \qquad (25)$$

We shall not work out the details of this connective, but give below straightaway its truth table in Table 8 which is similar to Table 7 for $\overleftarrow{O}$.

c	a	b
T	T	T
T	F	X
F	T	F
F	F	D

Table 8

Truth table for the "reverse and", $c \; \overleftarrow{A} \; a$

This table is very similar to Table 7, in that all the four states T, F, D, X, arise as a result of the "reverse and" operation, just as for the "reverse

or" operation. It will be noticed that, in this case, X comes out for the combination (T, F) of **c** and **a**, while for $\overleftarrow{O}$ this particular forbidden combination is (F, T) for **c** and **a**. An interesting consequence, which is observed for both "reverse and" and "reverse or", is that only half the number of input states containing the primary states, T and F, are useful, in that they can be carried over for making further deductions. In the other two cases, either the doubtful state D comes in, with consequent loss of definite information, or the contradictory state X arises, indicating that the previous steps in the argument which led to **b** are inconsistent, or that at least one of them is in error.

(e) *Reverse "if"*

In view of the similarity of the reverse operations for "and" and "or", and since we know that an implication $\mathbf{a} \Longrightarrow \mathbf{b}$ is representable either as an "or" ($\neg\mathbf{a}\ O\ \mathbf{b}$), or as an "and" ($\neg(\mathbf{a}\ A \neg)\mathbf{b}$), it is reasonable to expect that the operation obtained by reversing the "if" connective will also have properties similar to the reversal of "and" and "or". In the light of this, the corresponding tables for "reverse if" ($\overleftarrow{I}$), and "reverse only if" ($\overleftarrow{J}$), are given below in Tables 9(a) and 9(b).

c	a	b
T	T	T
T	F	D
F	T	F
F	F	X

(a) "Reverse if", $\mathbf{c}\ \overleftarrow{I}\ \mathbf{a}$

c	a	b
T	T	D
T	F	F
F	T	X
F	F	T

(b) "Reverse only if", $\mathbf{c}\ \overleftarrow{J}\ \mathbf{a}$

Table 9

Truth tables for the reverse operators $\overleftarrow{I}$ and $\overleftarrow{J}$

These tables are interesting in that whether the implication I itself is true or false, there is a consequence for each combination of the truth value of the states of **c** and of **a**, leading to one of the four possible types of states

for **b**. The same is true of J. Therefore, it looks very reasonable that, for all purposes, one can operate with the connectives I, or J, and their inverses $\overleftarrow{I}$ and $\overleftarrow{J}$ for working out everything in logic. The essential principle behind this is based on the fact that the two vital connectives "and" and "or" are both expressible in a suitable form by using I or J. As a matter of fact, both I and J are not required for this purpose and one of them, say I, is sufficient. Yet, for convenience, we have included both I and J because it is readily seen that an equivalence $\mathbf{a}\ E\ \mathbf{b}$ is the same as the simultaneous truths of $\mathbf{a}\ I\ \mathbf{b}$ and $\mathbf{a}\ J\ \mathbf{b}$. Hence,

$$\text{if } \mathbf{a} \Longrightarrow \mathbf{b} \text{ and } \mathbf{a} \Longleftarrow \mathbf{b}, \text{ then } \mathbf{a} \Longleftrightarrow \mathbf{b} \qquad (26)$$

As a consequence, every one of the operators which we have discussed above, becomes expressible in terms of combinations of I and J.

(f) *Negation of Y and V*

It is readily verified by comparing Table 9(a) with Table 3(a) that the first two rows of Table 9(a) correspond to the consequences for the implication $\mathbf{a}\ Y = \mathbf{b}$. In other words, if the binary I is true as from **a** to **b** ($\mathbf{a} \Longrightarrow \mathbf{b}$), then the relation between **a** and **b** is itself given by the operator Y. Therefore, it is reasonable to ask what is the consequence of the third and the fourth rows in Table 9(a). These correspond to the binary implication I being *false*, and then it follows that a_T leads to b_F and a_F leads to X, a contradiction. Obviously, we can represent the *unary* relation between **a** and **b** in such a case by the operator standing for the "negation of implication" (YN). Thus, we have completely represented the "reverse I" operator by means of two unary operators Y and YN (we shall not consider the consequences for J and the unary operators V and VN, since they follow the same lines as for $\overleftarrow{I}$, Y and YN. Thus, $\overleftarrow{J}$ only essentially duplicates the information contents of V and VN).

Therefore we shall summarise this particular section by saying that the truth, or falsity, of an implication in the binary form can be represented effectively by the unary operators Y and YN. The exact relationships between these are given in (27) below :

$$[(\mathbf{a}\ I\ \mathbf{b})] = \mathrm{T} \Longleftrightarrow (\mathbf{a}\ Y = \mathbf{b}), \text{ or } (\mathbf{a} \Longrightarrow \mathbf{b}) \qquad (27a)$$

$$[(\mathbf{a}\ I\ \mathbf{b})] = \mathrm{F} \Longleftrightarrow (\mathbf{a}\ YN = \mathbf{b}), \text{ or } (\mathbf{a} \not\Longrightarrow \mathbf{b}) \qquad (27b)$$

We shall come back to Y and YN, while dealing with the electronic representation of these operators, in Section 5.

4. *Logical Structures of Arguments*

(a) *Tabular and graphical representations*

We are now in a position to discuss how one proceeds for making an analysis of a connected argument consisting of a number of individual statements. Without going into too much detail, we may say that any such argument will have a number of "input statements", which are then connected with one another by logical connectives, and thus lead to "intermediate steps" of the argument, which again lead either to further intermediate steps, or to the final "output statements". Thus we may indicate the block diagram of an argument as shown in Fig. 1.

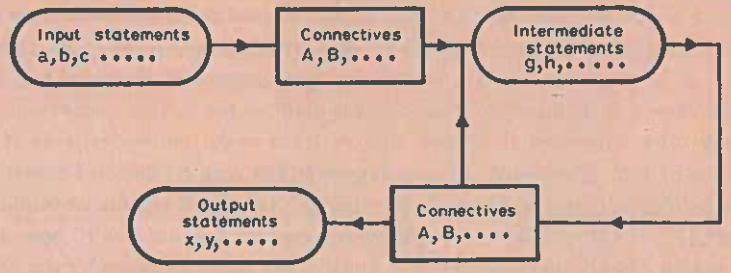


Fig. 1. Schematic outline of the logical structure of an argument

However, this figure is a highly simplified form of the interconnections between different statements, as it actually occurs in an argument. We shall now consider how the schematic outline in Fig. 1 really looks like when the full argument is displayed in a graphical form in a typical case. It consists of eight steps, as shown in Table 10. Without going into the sense

of such an argument, we have merely stated the various steps in our symbolic language, and these are represented in a graphical form in Fig. 2.

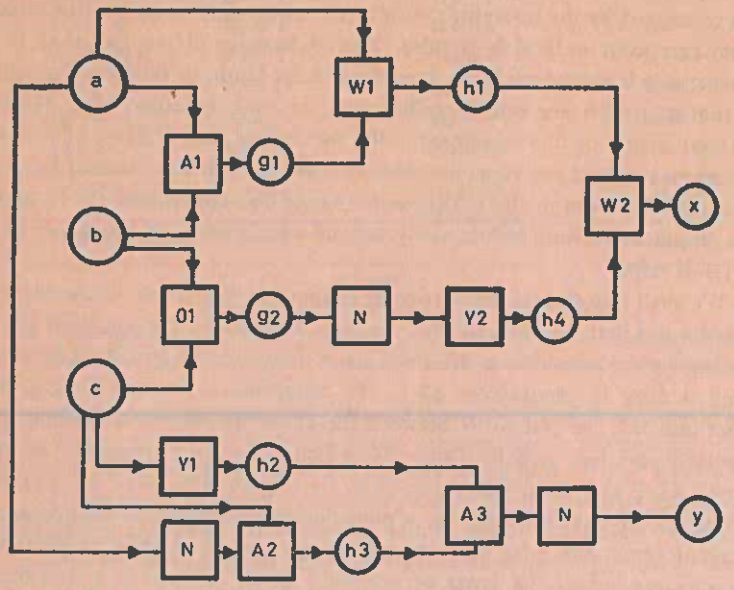


Fig. 2. Graphical representation of the terms and connectives of the argument given in symbols in Table 10

	(a) *Logical statements*	(b) *Sequence of logical operations*
1.	**a** and **b**=**g1** ;	**a** *A1* **b**=**g1**
2.	**b** or **c**=**g2** ;	**b** *O1* **c**=**g2**
3.	**a** with **g1**=**h1** ;	**a** *W1* **g1**=**h1**
4.	**c** implies **h2** ;	**c** *Y1*=**h2**
5.	(not **a**) and **c**=**h3** ;	(**a** *N*) *A2* **c**=**h3**
6.	(not **g2**) implies **h4** ;	(**g** *N*) *Y2*=**h4**
7.	**h1** with **h4**=**x** ;	**h1** *W2* **h4**=**x** (output)
8.	not (**h2** and **h3**)=**y** ;	(**h2** *A3* **h3**) *N*=**y** (output)

Table 10

Details of the steps in an argument giving the outputs x, y from the inputs a, b, c

It will be readily seen that the graphical representation gives, at a glance, the nature of the whole argument. For instance, it is seen that **g1** and **a** must not be inconsistent with each other for **h1** to exist, since these two are connected by the operator "with". It is also seen that such an inconsistency can occur only if **b** is false. This is because of the fact that, if **b** is true, **a** *A* **b** is true, or false, according to the truth, or falsity, of **a** itself, so that **g1** and **a** are both together true, or both together false. Hence, without analysing the argument fully, we can see that, if there should be an inconsistency at the stage proceeding towards **h1**, **b** has necessarily to be false (see also Section 6(b)). This will indicate how convenient it is to make the graphical drawing before analysing the consistency, or otherwise of a series of steps.

We shall not discuss the particular example in Fig. 2 any further except to point out that, for convenience, we have represented the negation of **a**, leading to the connective *A2*, itself as a unary connective *N*. The same thing is done in going from **g2** to **h4**. Since the negation of **g2** is the input into *Y2*, we put an *N* between the circle **g2** and the rectangle *Y2*. Similarly, the last step in Table 10, which involves a negation of the "and" *A3*, is also simply represented by putting an *N* between *A3* and the output **y**. (We have no space for a more detailed discussion of the operators like not (and), not (or), etc., whose properties can readily be deduced by pursuing further the types of approach given in Section 3. The book being written by the author (Ref. 5) will give full details about all these).

(b) *Results in a practical example*

It will be obvious, from Table 10 and Fig. 2, that, in general an argument stated in what is commonly understood as "propositional calculus" of logic, can always be represented by a *graphic diagram* similar to Fig. 2, and can also be represented in terms of a series of *logical equations* employing logical operators like *A*, *O*, *Y*, *W*, etc. In terms of the truth tables we have already given in Sec. 2 and 3, the working out of the final results, namely the nature of the output statements **x** and **y**, is only a matter of detail, if the states of **a**, **b** and **c** are given. In Table 11 below, we have given the truth values of the intermediate statements **g1**, **g2**, **h1**, **h2**, **h3**, **h4**, and of the final results **x** and **y**, for three sets of input truth values of **a**, **b** and **c**.

Step	Input 1 a=T, b=T, c=F	Input 2 a=T, b=F, c=T	Input 3 a=F, b=T, c=T
1. **a** *A1* **b**	**g1**=T	**g1**=F	**g1**=F
2. **b** *O1* **c**	**g2**=T	**g2**=T	**g2**=T
3. **a** *W1* **g1**	**h1**=T	**h1**=X†	**h1**=F
4. **c** *Y1*	**h2**=D	**h2**=T	**h2**=T
5. (**a** *N*) *A2* **c**	**h3**=F	**h3**=F	**h3**=T
6. (**g2** *N*) *Y2*	**h4**=D	**h4**=D	**h4**=D
7. **h1** *W2* **h4**	**x**=T	**x** is not derivable	**x**=F
8. (**h2** *A* **h3**) *N*	**y**=T	**y**=T	**y**=F

Table 11

Practical results for the eight steps in Table 10*

* Note that the conclusions regarding (**x**, **y**) are different for the three different inputs for **a**, **b** and **c**. In the second case, **x** is not derivable, since one of the inputs in step 7 has the state X.

† The origin of the inconsistency in step 3 shows that **a** and **g1** are mutually inconsistent. Tracing this backward to step 1, we see that, in this argument, no conclusion can be derived for **h1** and for those arising from **h1**, if **a**=T and **b**=F.

The above example will show how elementary is logic, and how one can go along working out consequences in any *practical problem* routinely, if the steps are written down properly. (As we have already stated in Section 1, it is not our intention to discuss how theorems are deduced or proved in logic, or how new general results are derived. These are in the realm of "pure logic", as distinct from "applied logic" or "practical logic" which we treat here).

It will also be obvious that, just as a computer can solve any practical problem in mathematics, particularly those in applied mathematics and physics, to give the results in numerical form, a computer is obviously the ideal choice for practical logic also, provided we program the steps, as stated in Table 10, or as shown in the graph as in Fig. 2, for implementation on the computer concerned. We have done this both for an analogue computer which closely simulates the graphical representation, as well as on a digital computer, by representing the steps such as those listed in Table 10 by means of Fortran statements. The next two Sections 5 and 6 will deal with each of these approaches, in outline, and indicate the essential ideas and techniques that have been used for these purposes.

(c) *Connected and unconnected statements*

This is the occasion for discussing some interesting logical tricks that can be used for proving theorems, or propositions. The commonest method of proof consists of a series of deductions, arranged in the form of a string. For example, we can have the sequence

$$a1 \Rightarrow a2 \Rightarrow a3 \Rightarrow b1 \Rightarrow \cdots \Rightarrow x \qquad (28)$$

In such a case, it follows that, if the statement **a1** is true, then **a2** is T, so that **a3** is T, and so on, leading finally to the result that **x** is true. In other words, if the consequent in any particular step in the implication, is the antecedent of the next step of the implication, then a connected sequence of such implications can be summarized by a single statement

$$a1 \Rightarrow x \qquad (29)$$

This is the most general form of deduction where, for instance, each of the statements **a1**, **a2**, **a3**, ··· itself need not be an elementary statement that something is true or something is false, but can itself be composed of a sequence of terms in an argument, containing all the connectives like "and", "or", "if", etc., finally leading to a truth value T or F. Therefore, in essence, we see that any proof, starting from an initial set of input statements, and consisting of a series of deductions, can, in general, finally lead to a definite answer.

We have cautiously used here the word "can" in talking about the possibility of there being a definite result regarding the finally required quantity, namely **x**. To show how this can fail, consider the following simple sequence of two implications :

$$a \Rightarrow b, \qquad \neg b \Rightarrow x \qquad (30)$$

If we write out the truth table of the two implications put together, we obtain what is shown Table 12.

a	b	¬ b	x
T	T	F	D
F	D	D	D
D	D	D	D

Table 12

Truth values for a pair of unconnected statements

$a \Rightarrow b$ and $\neg b \Rightarrow x$

It will be seen from Table 12 that, whatever may be the input state of **a**, namely T, F or D, the output state **x** is always D. This means to say that nothing positive can be stated about **x**. Consequently, the second implication in (30) may be said to be unconnected to the first implication, in the sense that it cannot lead to the onward propagation of the argument from the first step $a \Rightarrow b$, *via* this second step, namely $\neg b \Rightarrow x$. We may, in fact, generalise this and say that if some statement implies **g** at an intermediate step in an argument, then nothing whatsoever can be deduced if the next step is an implication with $\neg g$ as antecedent. Although this looks very simple, such a result of general interest does not appear to have been mentioned in most books on logic.

(d) *Reductio ad absurdum proofs*

Having seen that there are conditions under which a conclusion can definitely be reached, and that there are also conditions under which a definite conclusion does not follow (although all the steps in the argument are not contradictory), we shall now consider a third possibility, namely that of a contradiction, similar to the one in Table 11 which led to no result for **x**. We did not, however, use the contradictions in Table 11 to work out any useful result. We shall now show that under certain conditions, a contradiction can lead to extremely useful results, such as the definite proof of a theorem. This method of proving a theorem is given the general name *reductio ad absurdum* —reducing the problem to an absurdity if we assume that the theorem is not true, and hence demonstrating the contrary result, namely that the theorem is fully valid.

To illustrate this, suppose that we wish to prove that a Theorem **a** is true by this technique. The steps are as follows :

(i) Suppose that **a** is F ;
(ii) Prove that $\neg a \Rightarrow x$ (x_1, say), where (31a)
(iii) We *know* that **x** is false (x_2, say)

We can then show that the argument in (31a) leads to the conclusion (31b), namely that

(iv) Theorem **a** is necessarily true. (31b)

The sequence of logical steps for showing this are as follows :

(a) From (i), $\neg \mathbf{a}$ is T
(b) Hence, from (ii), $\mathbf{x}_1$ is T
(c) Testing for the consistency of (b) and (iii), we get $\mathbf{x}_1\ W\ \mathbf{x}_2 = \mathrm{X}$
(d) Since (iii) is well-established, the contradiction X shows that $\mathbf{x}_1$ is invalid, and $\mathbf{x}_1 = \mathrm{F}$
(e) Writing (ii) in an equivalent form $\neg \mathbf{x} \Rightarrow \mathbf{a}$, and putting in $\mathbf{x} = \mathrm{F}$ from (d), we obtain **a** is TRUE.

(32)

Thus, by proving that the contrary of the required statement **a** leads to a false conclusion, we deduce that **a** itself is true. In establishing such a contradiction with a known fact (as shown in (32c)), this fact (**x**) can be anything whatsoever, so long as its truth is unquestionable. This is why the *reductio ad absurdum* method of proof is so powerful particularly in relation to the proofs of theorems in which the validity has to be proved for *every* value of some parameter.

We shall illustrate this by the example of Fermat's last Theorem, which conjectures that

$$x^n + y^n = z^n \tag{33}$$

has no solution in integers x, y, z, for all integral values of $n \geqslant 3$. Suppose we call the equation (33) as Fermat's equation (**f**, say), and we wish to prove that **f** is false for *all* $n > 2$. Although this result has not yet been proved, the most likely technique that would lead to success in proving it is probably as follows.

Assume that **f** is true (and hence, that the Theorem is false) for some general $n > 2$, and work out its cosequences. If *any one* of the consequences disagrees with any known, or well-established result, say in number theory, algebra, or any other branch of mathematics, then necessarily **f** is false and Fermat's last theorem is true.

To indicate the power of this method, we can show that, if **f** is true,

$$2^{p-1} - 1 \equiv 0 \pmod{p^2} \quad (\mathbf{e} \text{ say}) \tag{34}$$

where p is a prime. We can show very simply that the validity of **f** (Eq. 33) for all integers n follows from its validity for all prime numbers p. Also, (34) is true for the first time for $p = 1063$, and the next p satisfying $\mathbf{e} = \mathrm{T}$ is greater than 2000. Hence, at one stroke, from the equation, $\mathbf{f} \Rightarrow \mathbf{e}$, we have shown

that **f** is not true, and that Fermat's conjecture is valid, for all p less than 2000, except $p = 1063$.

The point we wish to emphasize is that we *do not show* that Fermat's Theorem is "*true*" *for each of these primes*. We only show that it is "*not invalid*" *for all of them*. The essential steps in making this possible are (32c) and (32e), namely $\mathbf{x}_1\ W\ \mathbf{x}_2 = \mathrm{X}$ and

$$(\neg \mathbf{a} \Rightarrow \mathbf{x}) \equiv (\neg \mathbf{x} \Rightarrow \mathbf{a}) \tag{35}$$

5. *Analog Machine for SNS Logic—ESNY*

(a) *Representation of T, F, D, and X in electronic circuits*

Electronic circuits for representing the two truth values T and F of a statement, by the presence or absence of a voltage (which may be denoted by the Boolean numbers 1 and 0), and the combination of such states *via* the common operations of NOT, AND, and OR (of classical logic), which can be represented by suitably designed electronic gates, is well-know in computer science. We shall not explain these any further, but we shall indicate how the same types of circuits can be employed, in different types of combinations to give us a representation of the *four* truth values T, F, D and X, and also for obtaining a representation of the properties for the large number of logical connectives, including in particular W, A, O, I, J, etc., which occur in our formulation.

The essential new idea in doing all these, is the use of *two* electrical lines, which may be represented by the symbols α and β, to describe the truth value of any statement in our logic. Since the formulation of logic that we have developed here includes, in particular, the "doubtful state" (D), we shall give this the name "SNS logic", where the letters SNS stand for the words "Syad-Nyaya System" (*Syad* = may be, and *Nyaya* = logic, in *Sanskrit*). In the SNS system, the two lines α and β can each have the Boolean values 1 and 0. Therefore there are four combinations of these that are possible—namely (1 0), (0 1), (1 1), and (0 0). It is readily verified that the α-line may be identified with truth and the β-line with falsity. Then, the two-element Boolean vector (1 0) represents complete truth and no falsity, and therefore can be used to correspond to the state T in our logic. In the same way, the two-element Boolean vector (0 1) will correspond to the state F in our system.

On the other hand, the vector (1 1) means that there is *both* truth and falsity present; in other words, we do not know whether the state represents complete truth or complete falsehood. It could be either of them, and is

therefore a candidate for representing the *doubtful* state D. In each particular instance of testing the argument, this state D may give one state (say T), or may give the other (F), unless the doubt is cleared by combining this state (D), with another statement giving a definite T or F state, by means of the operator *W*. Therefore, the Boolean vector (1 1) is a very good choice for the state of doubt D.

In the same way, if we take (0 0), we see that there is *neither* truth, *nor* falsity, in it. In other words, it is an *impossible* state for a logical statement, because it is neither true nor false, which may be interpreted to mean that the logical state does not exist. In our logic, we have taken that any statement is either T or F. Therefore, if (0 0) arises in a particular step in a logical argument, we can identify this with the *contradictory* state X, from its above properties. Thus, we have, in effect, the representation of the four states T, F, D, X given by the four two-element Boolean vectors given in (36):

$$T=(1\ 0),\quad F=(0\ 1),\quad D=(1\ 1),\quad X=(0\ 0) \tag{36}$$

Just as we considered for the Boolean scalars 0 and 1, we may also talk of two ways of combining states of the type $(\alpha\ \beta)$—namely by Boolean addition and Boolean multiplication. These are best defined by the following equations (37a) and (37b), representing the two operations applied to two Boolean vectors $(\alpha_1\ \beta_1)$ and $(\alpha_2\ \beta_2)$:

$$\text{Addition}: (\alpha_1\ \beta_1) \oplus (\alpha_2\ \beta_2) = (\alpha_1 \oplus \alpha_2, \beta_1 \oplus \beta_2) \tag{37a}$$

$$\text{Multiplication}: (\alpha_1\ \beta_1) \otimes (\alpha_2\ \beta_2) = (\alpha_1 \otimes \alpha_2, \beta_1 \otimes \beta_2) \tag{37b}$$

If this is done, it is then readily verified that the following equations hold for D and X, in terms of T and F:

$$T \oplus F = (1\ 0) \oplus (0\ 1) = (1\ 1) = D \tag{38a}$$

$$T \otimes F = (1\ 0) \otimes (0\ 1) = (0\ 0) = X \tag{38b}$$

As will be clear from the next subsection, these two Boolean operations applied to $(\alpha\ \beta)$ correspond to the connectives *U* and *W* applied to logical terms. Hence,

$$T\ U\ F = D \text{ and } T\ W\ F = X \tag{38c}$$

(b) *Implementation of binary logical connectives*

Having seen how truth values can be represented by two electrical lines α and β, each having states 1 or 0 corresponding to the existence or non-existence of a voltage, we shall consider how these can be employed for electronically implementing the various types of connectives we have discussed above. It is found that the use of two lines, rather than one, is extremely

convenient for representing the four connectives *A*, *O*, *W* and *U* (see below for the definition of the connective *U*). Without going into details, we indicate, in Fig. 3, the ways in which these four connectives are implemented having a pair of standard NAND gates suitably connected for each by of the lines α and β. The existence of a symmetry between these circuits is also obvious and this is best indicated in terms of symbols AND and OR for one-line circuits, as shown in (39) below:

$$(a_1\ A\ a_2 = a_3)\ (\alpha_1 \text{ AND } \alpha_2 = \alpha_3);\ (\beta_1 \text{ OR } \beta_2 = \beta_3) \tag{39a}$$

$$(a_1\ O\ a_2 = a_3)\ (\alpha_1 \text{ OR } \alpha_2 = \alpha_3);\ (\beta_1 \text{ AND } \beta_2 = \beta_3) \tag{39b}$$

$$(a_1\ W\ a_2 = a_3)\ (\alpha_1 \text{ AND } \alpha_2 = \alpha_3);\ (\beta_1 \text{ AND } \beta_2 = \beta_3) \tag{39c}$$

$$(a_1\ U\ a_2 = a_3)\ (\alpha_1 \text{ OR } \alpha_2 = \alpha_3);\ (\beta_1 \text{ OR } \beta_2 = \beta_3) \tag{39d}$$

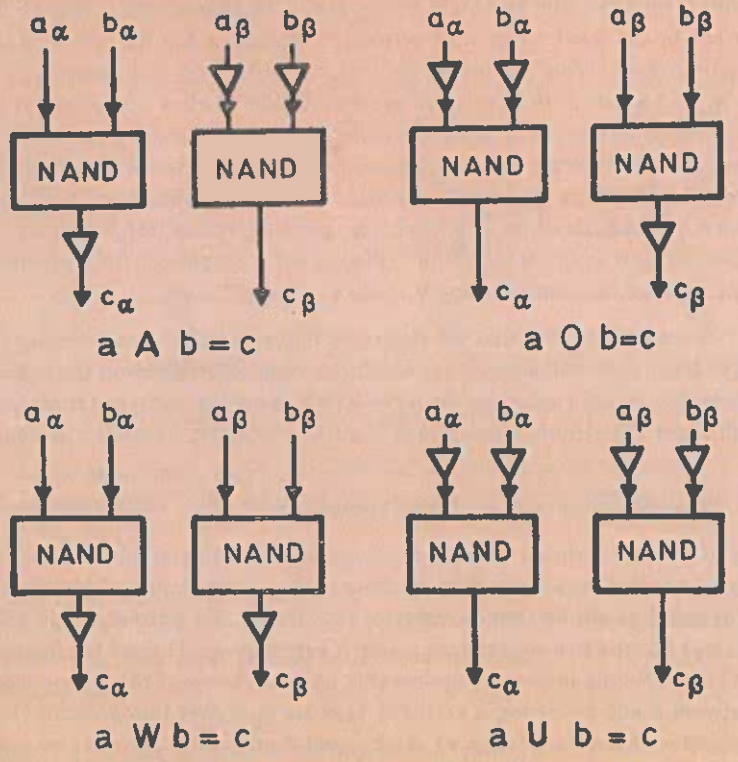


Fig. 3. SNS logic circuits for the four operators *A*, *O*, *W*, *U*.

It is readily verified that the first two rows of (39) faithfully implement the operations of "and" (A) and "or" (O) of SNS logic, as given in Table 1. So also, it is seen that the third row (39c) implements the first four rows of the truth Table 5, of Section 3, for the operator "with" (W). Consequently, it can also be verified that the remaining four rows of Table 5, for W, will also come out to be correct for Eq. (39c) and for the corresponding circuit shown in Fig. 3(c).

Fig. 3(d) and Eq. (39d) represents a new connective, which obviously arises from the symmetry of the relations between A, O and W. Just as is the case for W, it is easy to see that this operator U also puts together information of a logical nature arising from two different sources $\mathbf{x}_1$ and $\mathbf{x}_2$ for the same term $\mathbf{x}$ in order to obtain $\mathbf{x}(=\mathbf{x}_1 U \mathbf{x}_2)$. The connective U may be described by the word "unanimity operator", because it has the following properties. Only if *both* $\mathbf{x}_1$ and $\mathbf{x}_2$ are true, will $\mathbf{x}$ be true, and similarly, only if $\mathbf{x}_1$ and $\mathbf{x}_2$ are *both* false, will $\mathbf{x}$ be false. Under all other conditions, $\mathbf{x}$ will be doubtful. It is readily seen, therefore, that this operator gives a definite output (T or F) when the two inputs agree, and will give the indefinite output (D) when they disagree in any way. The short word "upon", which also starts with a letter U, (e.g., $\mathbf{x}_1$ upon $\mathbf{x}_2$ gives $\mathbf{x}$), is used for describing this operator, just as the word "with" (W) is used to represent the consistency operator for the combination "$\mathbf{y}_1$ with $\mathbf{y}_2$ gives $\mathbf{y}$".

In view of the fact that the electronic implementation has been done of SNS logic, the analog computer which has been constructed on the basis of these circuits has been given the name ESNY being the shortened form of the full name "Electronic Syad-Nyaya Yantra"—(*Yantra, Sanskrit* = machine).

(c) *Implementation of unary logical connectives*

If we use two lines, α and β, to represent the truth and falsity components of a logical term as mentioned above, the simple circuits shown in Fig. 4(a) and 4(b) will be seen to represent the circuits, for getting $(b_\alpha\ b_\beta)$, given $(a_\alpha\ a_\beta)$, for the two connectives E and N respectively. The one for negation (N) is interesting in that it requires only an interchange of the lines α and β, between **a** and **b**. Hence, a vector (1 1) again goes over into a vector (1 1), so that we have the relation $a_D\ N = b_D$, and doubt is not removed by negation. This looks reasonable, since, if **a** can be either T or F, $\mathbf{b}(=\neg\,\mathbf{a})$

also will be either F or T, which means the same thing as the state of **a** itself.

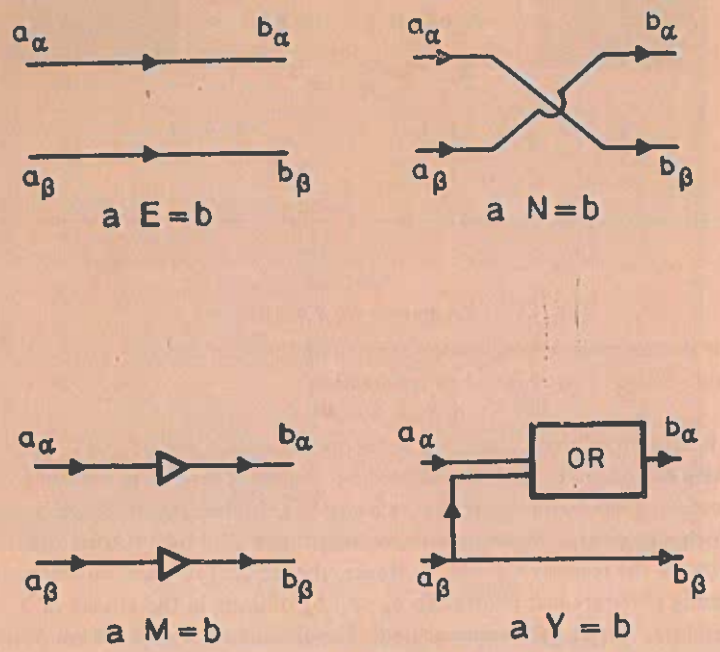


Fig. 4. SNS logic circuits for the unary operators E, M, N, Y.

On the other hand, we can devise a second way of negating **a**, namely by "complementation" (M), which is best understood with reference to Fig. 4. Here, we have

$$\sim a_\alpha = b_\alpha \text{ and } \sim a_\beta = b_\beta \tag{40}$$

so that (1 0) becomes (0 1), and (0 1) becomes (1 0). Hence, in so far as the states T and F are concerned, the operators N and M behave identically. However the effect on (1 1) of M is different from that of N, since it leads to (0 0). Hence,

$$a_D\ M = b_X \tag{41}$$

or, in other words, M converts a completely doubtful state D ($\equiv$T$\oplus$F) into the contradictory, impossible, state X ($=$ T$\otimes$F).

This effect of M becomes useful while dealing with the unary operators Y and YN. The circuit for Y is shown in Fig. 4(d), and it is readily seen to correspond to its truth table given below in Table 13(a).

(a) a Y=b,		(b) a YN=b	
a	b	a	b
T	T	T	F
F	D	F	X
D	D	D	X

Table 13

Truth tables for Y and YN

The corresponding table for the negated implication operator Y is seen, from Table 12(b), to correspond to the relation

$$\mathbf{a}\ Y = \mathbf{a}\ YM = \mathbf{b} \tag{42}$$

It is to be noted that, unlike the other operators like A, O, U, W, I, etc., for which the result for the corresponding negated operator is obtained by a succeeding operation by N (e.g., $\mathbf{a}O = \mathbf{a}ON$), in the case of Y and V alone, which are unary, the negation is representable only by a further operation by M on the results of Y and V. Hence, the circuit for Y can be obtained by putting inverters just before the b_α and b_β outputs in the circuit of Y, and therefore, it is not shown separately. The circuits for V and VN are obtainable similarly, by remembering that $\mathbf{a}\ V = \mathbf{b}$ is equivalent to $(\neg\mathbf{a})Y = (\neg\mathbf{b})$.

(d) *Operation of ESNY*

There is only one more bit of information that has to be mentioned about the design of this analog computer ESNY. Since the first two unary connectives, namely E and N, involve only wires connecting between the terminals at either end, which are either direct (a_α to b_α, and a_β to b_β) for E, and (a_α to b_β and a_β to b_α) for N, they are implemented by having wires of suitable length which can connect the relevant modules. Therefore, the introduction of N anywhere is very simple and requires no special facilities in this computer. On the other hand, in order to implement the complementation module M, the circuit shown in Fig. 4(c) is used and a number of M-modules have been made available.

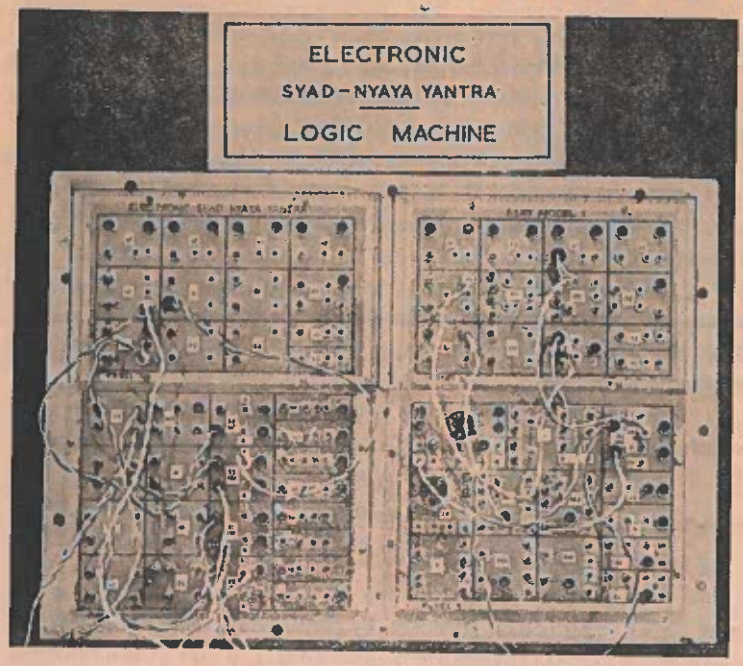


Fig. 5. ESNY II wired for a simple problem

Fig. 5 shows a photograph of the computer ESNY (Model II as constructed in our laboratory) with connections made between the various terms and operators representing a simple problem. There are green and red lamps to indicate T and F states respectively in the modules representing all input states, intermediate states, and output states. If both of them are on, it indicates the occurrence of the doubtful state D. If both are off, it represents the state X, in which case an alarm would sound and also a yellow lamp would flash intermittently in the module concerned. Thus, the existence of contradictions are immediately displayed by the computer. However, even though a contradiction may arise at one step in the argument, it can go through without affecting the rest of the problem, which can be correctly implemented in those parts of the problem in which there is no contradiction. However, where a contradictory state is itself an input for a connective the output becomes again an X for all operators. Thus, in general, if the problem

has no contradictions, there will be no alarm sounded and the answers for the truth values of the outputs, say x_1, x_2, and x_3 can be seen by looking at the lamps on the relevant modules, when the appropriate input states, say a_1, a_2, b_1, b_2 are fed in.

If there is contradiction somewhere, the problem is more complicated. However, in the case of those outputs which *do not show the contradictory state*, their indication (T, F or D) can be taken to be definitely valid corresponding to the inputs that are put in. Thereafter, it is possible to analyse the region showing contradiction (on the lines indicated in Sec. 4), and then reverse the argument in order to obtain information about where exactly the contradiction arises. We shall not say more about this, as this has several ramifications, not all of which have been fully examined as yet.

It should be mentioned that we have not described at all the representation of the reverse operators on our analog computer; but it is only a matter of detail and will be discussed in full in other publications.

(e) *Actual problems solved using ESNY*

(i) *Lewis Carroll Problem from Suppes' book*

We shall illustrate how our SNS system works, and how easy it is to apply ESNY for practical purposes, by considering a simple example (taken from Lewis Carroll) discussed by Suppes (Ref. 1, p. 61). It is as follows :

> "No ducks are willing to waltz. No officers are unwilling to waltz. All my poultry are ducks. Therefore, none of my poultry are officers". (43)

These statements can be symbolized, using d = duck, w = willing to waltz, p = poultry and o = officers. Then the three input statements are

(*1*) $d \Rightarrow \neg w$; (*2*) $o \Rightarrow w$; (*3*) $p \Rightarrow d$

and we wish to prove that these lead to

(*4*) $p \Rightarrow \neg o$

To solve these, we first notice that there is a chain of two arguments (*3*) and (*1*) composed of unary implications leading from p to $\neg w$. However, (*2*) has w as a consequent, while b is the antecedent. Hence, we reverse this to yield $w \Leftarrow o$, which we know is equivalent to

(*5*) $\neg w \Rightarrow \neg o$

Now, (*3*) and (*1*), followed by (*5*), gives *an unbroken chain of connected implications* and, in the two intermediate steps (*1*) and (*5*), the input is exactly

the same as the output of the previous step. Therefore, from the considerations stated in Section 4(c), we see that they lead indeed to (*4*), namely $p \Rightarrow \neg o$, the desired result. This one-step proof is to be contrasted with more than six steps required in the proof given in Suppes' book.

If we used ESNY, we would compare the results of (*1*), (*5*) and (*3*) in succession, and of (*4*) alone, in giving the truth value of o, corresponding to those of p that are possible, namely T, F, and D. The corresponding logical graph is shown in Fig. 6, and the states of all the relevant terms, as read from ESNY, are given in Table 14. Here, $p \Rightarrow d \Rightarrow \neg w \Rightarrow \neg o$ gives the state x (of o), while $p \Rightarrow \neg o$ gives y (of o), and the test for the agreement between the two is given by $z = x\ G\ y$, where G is the agreement operator, which is not discussed in this paper. If z = T, then it means that the two agree, and the uniform appearance of the state T for z shows that we have proved the relation posed in the problem.

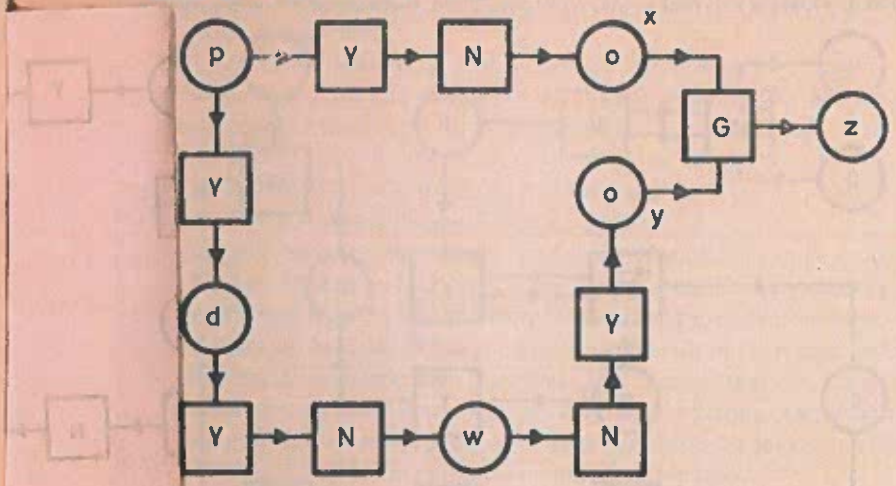


Fig. 6. Logic circuit graphics of the Lewis Carroll problem

p	d	w	x	y	z = x G y
T	T	F	F	F	T
F	D	D	D	D	T
D	D	D	D	D	T

Table 14

Table showing the truth values of all relevant terms in the Lewis Carroll problem

(ii) *More complicated problem from Stoll's book*

The superiority of the techniques in our SNS logic is brought out most forcefully by the following example taken from Stoll's well-known book on logic (Ref. 2, Ex. 4.3, p. 184). The problem as possed is as follows :

Show that the following four relations

(*1*) $(\mathbf{w}\ O\ \mathbf{p})\ Y = \mathbf{i}$

(*2*) $\mathbf{i}\ Y = (\mathbf{c}\ O\ \mathbf{s})$

(*3*) $\mathbf{s}\ Y = \mathbf{i}$

(*4*) $\neg\,\mathbf{c}\ A\ \neg\,\mathbf{u} = \mathrm{T}$

together predict that w will be always false. Stoll gives a proof based on 13 steps.

We shall solve the problem, using its graphic representation, as given in Fig. 7, employing the reverse operators $\overleftarrow{O1}$, $\overleftarrow{O2}$ and $\overleftarrow{A}$ for representing (*1*), (*2*) and (*4*), since this makes w an output. Then, there are only three possible

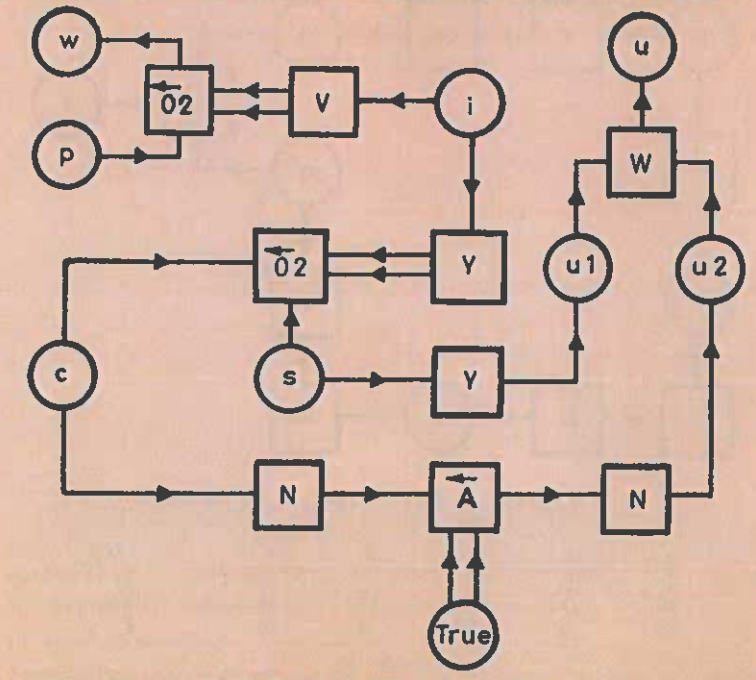


Fig. 7. Graph of Problem in Stoll's book

independent inputs p, i and c, and the condition to be fulfilled is that there should be no X's coming out anywhere in the whole circuit. (Note that a double input line is used for the reversed truth value of the "and" or "or", respectively).

Although an argued solution of the problem can be given using this graph, we shall first indicate how it was solved using ESNY II. The connections between the various modules were made, according to Fig. 7. Then, p, i and c were made T and F, (8 possibilities in all), and the siren sounded for seven of the eight combinations, showing that they are contradictory. The eighth case with p = F, i = F and c = F, was free of contradictions, for which the other variables were found to be w = F, s = D and u = F.

In other words, not only could we prove that the result required to be proved, namely that w = F, is the only possible truth value of w, but we could also prove that the equations (*1*), (*2*), (*3*), (*4*) of Stoll's problem necessarily require that

(*5*) All the variables w, p, i, c, u are false, while s may be T or F.

Having found this result using the analog computer, it was possible to devise a proof in three steps to give these results. It is as follows :

(i) If c = T, u2 = X, and hence c = F

(ii) If c = F and i = T, u = X, so that i = F and s = D

(iii) If i = F, necessarily w = F and p = F.

It is indeed remarkable that six times as much information can be deduced from the set of the given four input equations by using SNS logic, with just three deduction steps (i), (ii) and (iii), than the only useful information w = F, as deduced by Stoll. What is more, the whole argument is seen to depend only on three independent variables. It is believed that our algebraic method of writing down everything in the form of equations *may have the capability of working out all possible consequences* from the logical information that is put in. Studies in this direction are in progress.

6. *Fortran Computer Programs for SNS Logic*

(a) *Description of the program NYAYA*

Since all the concepts of propositional logic, including the reverse connectives, could be implemented on ESNY, we proceeded to convert the ideas involved in the design of the various modules into a Fortran program for working out the logic of a connected set of statements on a digital computer. As would have been clear, the logic circuits are all based on elementary

logic chips for the one-element Boolean algebra (with states 0 and 1) of the connectives EQU, NOT, AND and OR. These are available as logical functions in Fortran, and therefore the various subroutines for the SNS connectives were written using them.

First, some simple systems were considered, incorporating the SNS connectives *E*, *N*, *A*, *O*, *U*, *W*, *I*, *M* and *Y*, and the first program so written, NYAYA 1 (*nyaya*, *Sanskrit* = logic), was tested and found to work satisfactorily. Then, the complete range of connectives, discussed in Sections 2, 3 and 4, were incorporated into the program, and a revised version NYAYA 2 was written.

The operators included in NYAYA 2 are the following, (this is not a complete list) :

(a) Unary connectives *E*, *N*, *M*, *Y*, *V*, and their negations, reverses, and negated reverses ;

(b) Binary connectives *A*, *O*, *W*, *U*, *I*, *J*, *G* and their negatives, reverses, and negated reverses.

In the set (a), the reverses are all contained in the original set itself, while for (b), they are different.

The Fortran symbology developed is as follows :

(i) The logical constans T, F, D, X are called, respectively

ZTR, ZFL, ZDB, ZXC (44a)

All logical variables start with the letter V. Examples are

VA, VX1, VCITY, VC(J) (44b)

(ii) The unary logical connectives are, in an obvious symbology,

SE, SN, SM, SY, SV (45)

where S is universally used as the first letter for all logical functions. Thus,

a *N* **b** is written VB = SN(VA) (46a)

a *YN* **b** is written VB = SN(SY(VA)) (46b)

(iii) The names of binary connectives also start with S, but have three letters, the second being F for forward, and R for reverse, operators. The third letter is the logical symbol for the operator, namely *A*, *O*, *I*, etc. The list contains the following :

SFA, SFO, SFW, SFU, SFI, SFJ, SFG (47a)

and those obtained by replacing F by R, such as

SRA, SRO, SRI, SRJ, SRG (47b)

These occur in statements as follows :

a *A* **b** = **c** is written VC = SFA(VA, VB) (48a)

c $\overleftarrow{I}$ **a** = **b** is written VB = SRI(VC, VA) (48b)

a *ON* **b** = **d** is written VD = SN(SFO(VA, VB)) (48c)

e $\overleftarrow{J}$ *N* **a** = **b** is written VB = SN(SRJ(VE, VA)) (48d)

Using these, the essential logical steps in the Fortran program of the problem given in Table 10 will be as follows :

```
VG1 = SFA (VA, VB)
VG2 = SFO (VB, VC)
VH1 = SFW (VA, VG1)
VH2 = SY (VC)
VH3 = SFA (SN (VA), VC)
VH4 = SY (SN (VG2))
VX  = SFW (VH1, VH4)
VY  = SN (SFA (VH2, VH3))
```

(49)

In order to illustrate the case of using the computer, we may mention that the program written for this problem, including input and output statements, feeding in of the eight possible combinations of T and F states for **a**, **b** and **c**, and printing the output in tabular form, as shown in Table 15 below, required only 34 lines. Of course, the relevant subroutines had to be taken from the main program NYAYA2.

TABLE 15. COMPLETE OUTPUT FOR THE PROBLEM POSED IN TABLE 10

A	B	C	G1	G2	H1	H2	H3	H4	X	Y
T	T	T	T	T	T	T	F	D	T	T
T	T	F	T	T	T	D	F	D	T	T
T	F	T	F	T	X	T	F	D	XX	T
T	F	F	F	F	X	D	F	T	XX	T
F	T	T	F	T	F	T	T	D	F	F
F	T	F	F	T	F	D	F	D	F	T
F	F	T	F	T	F	T	T	D	F	F
F	F	F	F	F	F	D	F	T	X	T

(b). *Application of the program NYAYA 2*

It is readily verified that lines 2, 3 and 5 of Table 14 agree completely with the data worked out by hand, and listed in Table 11. It is to be noted that we have used the symbol X in Table 14 to represent the contradictory state when it first appears, and XX for a second appearance of X with an X as input. Thus, corresponding to the case of line 7 of input 2 of table 11, we find an XX in Table 14, in line 3, showing that the output **x** is not determinable.

On examining Table 14, the following additional consequences come out :

(a) On lines 3 and 4 of this table, corresponding to the inputs, T, F, T and T, F, F, the output **x** is found to be not determinable, and on looking back, the cause in both cases is found to be the same, namely the occurrence of the contradictory state X for **h1**. The latter again arises because of the fact that **a** = T and **b** = F, in both the examples. Therefore what we conclude is that, in this problem, the combination **a** = T and **b** = F is impossible and that if **b** = F, **a** cannot be T, and *vice versa*.

(b). There is another contradiction that arises for the last combination F, F, F where it occurs only for **x** arising from the last but one step, **h1** with **h4** = **x**. It is rather interesting that **h1** and **h4** depend on all three of the inputs **a**, **b** and **c** and it is not possible to say that either **a** alone, or **b** alone, or **c** alone, has led to this contradiction. It has occurred because of all the three having the false truth values.

(c). It is interesting that, although **x** has an *impossible* truth value for three cases out of the eight possible ones, the output **y** has a *definite* truth value in every one of the eight cases. This is rather interesting, and brings out the fact that a number of logical relations may lead to a definite consequence for one conclusion (**y** in our case), although they may be incompatible in other ways. We shall not discuss this aspect more deeply here, as these logical peculiarities are reserved for discussion in a separate paper.

7. *Extension of SNS Concepts to First Order Predicate Logic*

The extension of the use of two lines corresponding to the α and β components of the truth value of a term in propositional logic, to the case of first order predicate logic (which contains quantified statements), is fairly obvious, though not simple. We have only made an introductory survey of

this problem and it appears that all sentences in quantified logic can be specified by using six elements α, β, γ, δ, ϵ, ζ, to represent them. Of these (α, β) will correspond to the ideas already discussed above. The next three elements (γ, δ, ϵ) will represent the nature of the quantifier such as "for all" ($\forall$), "for none" (ϕ), "there exists" ($\exists$), and "not for all" ($\neg\forall$). Actually it is convenient to add one more symbol ζ, which is a one-element Boolean variable having possible values 0 and 1, to represent the *not* that qualifies the quantifier. This "not" is distinguished from the not of the SNS variable, by using the symbol $\neg$. Thus we will read $\neg\exists$ as "not there exists".

It would appear that, on combining all the four new components γ, δ, ϵ, ζ, (each of which can have two possible Boolean states 0 and 1), we should get 16 possible states. However, the set of eight possible states obtained from the vector (γ, δ, ϵ) itself contains the negations of these obtained by the operation of ζ, and only eight of them are distinct, although the use of ζ in the algebra seems to be indicated (see below). Again, of these eight, conventional quantified logic uses only four, which may be listed as follows :

$$\forall,\ \neg\forall,\ \phi,\ \neg\phi\ (\equiv\exists) \tag{50}$$

What is even more interesting is that, if we include in addition the negation N($\neg$), of SNS logic, for the propositional term, then all these four can be interrelated and the four possible quantifiers in (50) can be shown to correspond to the four in (51), all of which are obtained from $\forall$ alone (see Ref. 5, in this connection) :

$$\forall,\ \neg\forall,\ \forall\neg,\ \neg\forall\neg \tag{51}$$

On examining the interrelationship between these four basic quantified states, the interesting possibility comes out that they form a representation of a group of order 4. Each of four suitably defined logical operators $\mathscr{E}$, $\mathscr{L}$, $\mathscr{M}$, $\mathscr{N}$ which form a group of order 4, transform the four states in (50), or in (51), into one another, and the effect of each of these on the six elements forming the quantified vector $\{(\alpha\ \beta)\ (\gamma\ \delta\ \epsilon)\ (\zeta)\}$ can be worked out. The group involved is isomorphous with the famous *Vierergruppe* of Klein, or the crystallographic point group D_2. These properties, and their possible application for a simple representation of quantified first order predicate logical statements, are being examined. It appears that the six elements of the quantified logic vector can also be represented by electronic circuits similar to those implemented on ESNY. Such an extended system has been constructed in part, and has been given the name QUANT. It is undergoing examination.

8. *Acknowledgements*

The work reported in this paper was initiated about two years ago, and I wish to acknowledge very much the assistance provided by Dr. (Mrs.) Vandana Shiva in the early development of our approach towards SNS logic. Afterwards, the design and construction of ESNY was competently done by my colleagues Mr. T. A. Thanaraj, and Mr. R. E. C. Johnson, to both of whom I am most grateful for their help. The extension of the ideas incorporated in ESNY for developing the programs NYAYA 1 and NYAYA 2 was done almost completely by Mr. Thanarj under my general guidance. Detailed papers regarding each of these contributions will be written by the participants in this project in due course.

I wish to record my grateful thanks to Prof. Satish Dhawan, Director of our institute, who gave me full freedom to carry on our work, in this rather recondite field, by providing the necessary facilities in the Institute.

This paper comprises MATHPHIL Report No. 8, March 1980.

REFERENCES

1. Suppes, P. "Introduction to Logic", 1957 New York, Van Nostrand.
2. Stoll, R. R. "Set Theory and Logic", 1961 San Francisco, W. H. Freeman.
3. Copi, I. M. "Symbolic Logic", Fourth Edition 1973 New York, Macmillan.
4. Church, A. "Logic", Article in Encyclopaedia Britannica, 1966 Chicago, p 209.
5. Ramachandran, G. N. "Principles of Practical Logic", 1980 (in preparation).
6. Halmos, P. R. "Lectures on Boolean Algebras", 1974 New York, Springer-Verlag.

Spiroplasmas Of Plants And Animals : An Overview

Karl Maramorosch*

*Waksman Institute of Microbiology, Rutgers—The State University
P. O. Box 759, Piscataway, New Jersey 08854, USA

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

What are spiroplasmas ?

Spiroplasmas belong to the class Mollicutes, or mycoplasmas. They are among the smallest prokaryotic microorganisms, passing through filters that retain bacteria. They lack a cell wall, having, instead, a pliable unit-type membrane. Spiroplasmas approach in size the largest viruses, such as pox viruses and the Mahrburg virus, but they differ from viruses by their ability to grow in artificial, cell-free media, and by containing both DNA and RNA. Spiroplasmas can infect vertebrate and invertebrate animals and plants. Certain species require alternating insect and plant hosts for survival in nature. Other spiroplasmas have been found in the nectar of flowers and at present these microorganisms are not known to cause disease. The trivial term "mycoplasma" has been used to refer to all members of the class Mollicutes. However, in the strict sense of the Latin term, members of the family Mycoplasmataceae are not known to cause diseases of plants or of insects. The only properly isolated and cultured plant pathogenic mycoplasmas belong to the family Spiroplasmataceae.

Spiroplasmas are usually associated with insects. Known vectors of plant spiroplasmas belong to leafhoppers, planthoppers, and probably psyllids. A report of spiroplasmas in aphids has not been confirmed. The finding of spiroplasmas in ticks by Clark[8,9] does not imply that ticks can transmit these agents. Furthermore, the ticks in which spiroplasmas were found, feed on plant-eating rabbits. The possibility of an association of Clark's agent with plants cannot be ruled out. A spiroplasma, not yet cultured, but classified as a member of this group on the basis of its morphology, has been incriminated in the abnormal sex ratio of fruitflies. The androcidal spiroplasma kills all males in four neotropical species of Drosophila[72].

Historical background

Important discoveries are often made accidentally and not necessarily through carefully planned experiments. One such accidental discovery made in the spring of 1967 resulted in an exciting new development in microbiology, plant pathology, entomology, and animal pathology. Several plant pathologists in Professor Hidefumi Asuyama's laboratory in Tokyo were working on plant virus diseases[87]. Y. Doi was assigned to study the mulberry dwarf disease and his task was to find and characterize morphologically the

virus believed to cause the dwarf disease. Mulberry trees are of economic importance to the silk industry as the leaves provide the food for growing silkworm caterpillars. Despite repeated attempts in which several thousand thin sections of diseased material were studied through the electron microscope, Doi was unable to find particles that would resemble viruses. A fortunate incident occurred one night when Kaoru Koshimizu, a microbiologist working with thin sections of cultured *Mycoplasma gallisepticum* in the Department of Veterinary Microbiology of the Faculty of Agriculture in Tokyo, happened to see the electron micrographs prepared by Doi. Koshimizu pointed out the striking resemblance of the bodies in the phloem of diseased mulberry plants to those of *M. gallisepticum*. Doi, who until then never heard the term "mycoplasma" was very skeptical at first, saying that plants do not have diseases caused by these microorganisms and that mulberry dwarf is a virus disease. When Koshimizu asked whether tetracyclines had been tested on mulberry trees, Doi replied that tetracyclines would not affect a virus. At this point, Koshimizu suggested trying tetracyclines to see whether the microorganisms that appeared to be similar to those he was familiar with, would respond to tetracycline antibiotics. The next morning Doi reported the conversation of Professor Asuyama, who, in turn, contacted the chairman of veterinary microbiology and obtained from him samples of several tetracycline antibiotics. When these were applied through roots to diseased mulberry trees, the plants began to recover within 3 weeks and continued to develop normal leaves as long as tetracyclines were supplied. It was not possible, however, to obtain a complete cure and when larger amounts of tetracycline were applied, the effects proved toxic to plants. Three other diseases that cause stunting, proliferation and a witches' broom effect were tested by Asuyama and coworkers and in all instances they found mycoplasma-like microorganisms in the phloem. Similarly to mulberry dwarfed trees, the plants with other yellows-type diseases reacted temporarily to tetracycline chemotherapy. The results of these tests were reported in the summer of 1967 at the Annual Meeting of the Phytopathological Society of Japan in Tokyo. The conclusions drawn from these experiments had a profound effect but, ironically, Koshimizu's role in this breakthrough has not been mentioned by his colleagues. They felt that the impression ought to be given that the findings were the result of carefully-planned experiments !

I had the good fortune to find out about the Japanese report in the fall of 1967 and to have in a glasshouse several yellows-type diseases, such as aster

yellows and corn stunt, as well as stock colonies of their respective insect vectors. I was able to confirm the findings of Asuyama's group with respect to the American plant diseases. In a paper, presented in January 1968 at the New York Academy of Sciences these results, obained by me and my associates, Eishiro Shikata and Robert Granados, were announced and widely distributed.[45] Consequently, the same findings were made in France, Germany, Great Britain, Australia, India, China, and the Soviet Union. At first, there was the expectation that tetracyclines might provide a permanent cure for some, if not for all, yellows-diseased plants. The Japanese made a concentrated effort, coordinating the work at numerous experiment stations and employing nearly 150 scientists. They concluded that none of the available tetracyclines could cure yellows-affected plants permanently. Eventually, others arrived at the same conclusions, but in some instances it took several years and very large expenditures, as in the case of coconut lethal yellowing in Florida, before tetracycline chemotherapy became discredited and abandoned. It is still of practical uses in treating pear trees with pear decline in commercial orchards in the United States.[51] In the meantime, the search goes on for other compounds that might destroy the plant mycoplasma agents permanently rather than cause temporary remission.

Cultivation on cell-free media

Unfortunately the plant pathological literature of the past decade contains numerous unconfirmed reports of cultivation of "plant mycoplasmas". This is understandable in a new branch of science, where there was lack of knowledge of specialized techniques in preparing media and handling of mycoplasmas. The most common source of error can be ascribed to contamination of the media, either from without by *Mycoplasma hominis*, or from within by using animal sera contaminated with *M. argenini*. Mycoplasma experts have offered their assistance and anyone who isolates a presumptive plant mycoplasma can avail himself of this free assistance, by sending the culture to a recognized authority. Despite this excellent collaboration, a few workers have refused to submit their so-called plant mycoplasmas to outside scrutiny. Their findings could not be confirmed elsewhere and their work had to be ignored by the scientific community. The minimum requirement, according to international rules of mycoplasmologists is to provide evidence—by outside experts—that the new isolate is antigenically different from species sharing the same biological properties and habitat. The Mycoplasma Committee on the Taxonomy of Mycoplasmatales requires now obligatory serological tests, growth or metabolism inhibition, together with the less specific double-immunodiffusion or complement fixation test. Recent books and reviews provide detailed information on the biology of spiroplasmas and mycoplasma-like agents[2,20,36,37,40,41,45,71].

Culture media

No single medium has been devised for isolation of different spiroplasmas. In fact, divergent media have been used successfully by different groups to isolate the same corn stunt spiroplasma[6,11,30,73]. The finding by Davis of spiroplasmas "in search for disease", in the nectar of tulip tree flowers, indicated that these microbes are quite common, but that they have been overlooked in the past[18].

Citrus stubborn disease and its spiroplasma

There were two notable exceptions among the earlier attempts to culture plant mycoplasmas. Two groups working independently but communicating freely their findings, were able to isolate a new agent from the citrus stubborn disease. One group, headed by George Calavan at Riverside, California, and the second, headed by Jose Bove in Bordeaux, France, obtained a pure culture of a microorganism, that was different from any previously-known mycoplasmas. In 1971 the French group published their findings[61] and shortly thereafter the report from Riverside appeared[18]. Yet, it took until 1974 before definitive evidence was obtained that the cultured microorganism is the causal agent of the stubborn disease. In the meantime, it was found that this agent is not a member of the Family Mycoplasmataceae, but that it belong to a group that was new to science.

Corn stunt

The corn stunt disease is of little importance in the United States. Since 1949 I had maintained pathogen-free cultures of *Dalbulus maidis* and since 1955 also of *D. elimatus*, the first known leafhopper vectors of the corn-stunt agent. Since 1968 we had no difficulty in finding the mycoplasma-like microorganisms in diseased corn plants and we published electron micrographs showing the movement of these infectious agents from one phloem cell to another. We also found the same agents in the salivary gland of insect vectors. The results

were communicated in the Proceedings of the US National Academy of Sciences in 1968. A few years later Dr. Robert E. Davis at the US Department of Agriculture in Beltsville became interested in our reports. After obtaining our diseased material and stock insects, he started various tests but having no electron microscopy training, he wondered whether there could be found some other way to ascertain the presence of the "plant mycoplasmas". It so happened that at this time Shmul Razin of the Hadassah Medical Center in Jerusalem visited Davis. When asked whether an electron microscope is essential to detect mycoplasmas, Razin, one of the world's leading mycoplasmatologists, pointed out that these tiny microbes can also be seen with a good light microscope, using phase contrast or dark-field techniques. It never occurred to plant pathologists to try light microscopy, as there was a general belief that the respective diseases were of viral origin, and it was known that viruses could not be seen that way.

The same day Davis checked crude sap from corn-stunt-diseased plants. To his amazement he found a large number of rapidly-moving spiral forms. No spiral forms could be found in healthy corn extracts. I must admit that, at first, I did not believe that one can see these spirals, but soon I found that they do occur in fairly large quantities and that Davis certainly was correct in his observation. Davis named the microorganisms spiroplasmas.[15] When the cultured agents of citrus stubborn were examined by dark field microscopy, the same spiral forms were detected. Consequently, the Latin, scientific name of the cultured organism has now become *Spiroplasma citri* Saglio *et al.* No spiral forms are produced by members of the Mycoplasmataceae or Acholeplasmataceae. The new plant pathogenic spiroplasmas were separated by taxonomists and tentatively grouped into a new family, Spiroplasmataceae.

In the following years the successful cultivation of the corn stunt spiroplasma was announced by the US Department of Agriculture workers[73] and by Rutgers University scientists[6].

Cactus spiroplasma

In 1975 a spiroplasma was isolated in our laboratory from the ornamental cactus *Opuntia tuna monstrosa*[25]. This spiroplasma grew on the same medium as *S. citri*, but it was difficult to isolate it from the diseased plant. Only 2 of 14 tubes inoculated with the plant extract contained spiroplasmas. The other tubes remained sterile. The antibiotic spectrum of the cactus isolate proved similar to that of *S. citri*[26]. A rod-shaped particle, believed to be a viral particle,

was constantly associated with the cultured spiroplasma. Morphologically this rod-shaped virus resembled the rod-shaped virus described by Cole[11] as one of three viruses associated with *S. citri*. No bacteriophage particles were detected. The cactus spiroplasma has been found recently to provide the first, and so far the only indicator strain for the bacteriophages of *S. citri* (Alan Liss, personal communication). When *O. tuna monstrosa* plants were treated via roots with various tetracycline compounds, temporary recovery was striking. The plants changed into the normal *Opuntia tuna* cactus. This change lasted from 8 to 24 months. Afterwards, the plants reverted to the *monstrosa* type. When thin sections of these plants were examined by electron microscopy, typical MLOs were again found in sieve elements.

Aster yellows spiroplasma

Numerous attempts have been made since 1967 to culture the agent of aster yellows. These attempts, in our laboratoy as well as elsewhere, met at first with failure.[64] In 1976 we modified our procedure and used as source material lettuce, *Lactica sativa*, with aster yellows. The disease is called "white heart of lettuce" in *L. sativa*. Extracted sap contained numerous spiral forms, as noted under darkfield microscopy.[28] When this sap was filtered and growth medium added, spiroplasmas grew in 14 of the 15 tubes prepared. Newly isolated spiroplasma was used as inoculum and *M. fascifrons*, the sixspotted aster leafhopper, was inoculated by abdominal injections. A small number, approximately 5% of the injected leafhoppers were rendered infectious and, when placed on lettuce seedlings, transmitted the spiroplasma causing typical aster yellows disease. The spiroplasma has been deposited in the American Type Culture Collection.[35] Upon subculturing, infectivity was lost. The spiroplasma was found constantly associated with a rod-shaped virus. Serological and growth inhibition tests, as well as hybridization tests have been performed in collaboration with Prof. E. Freundt in Denmark. A striking resemblance with *S. citri* and a pronounced difference with the corn stunt spiroplasma makes the distinction of the spiroplasmas complicated. PAGE (polyacrylamide gel electrophoresis), performed in our laboratory by Padhi *et al.*[59] permitted the clear distinction between the corn stunt and the aster yellows spiroplasma, but not the distinction between the aster and *S. citri* isolates. Additional criteria will be required to separate the various isolates and species.

The isolation of the aster yellows spiroplasma revealed important points. The same agent, until 1967 believed to be a virus, causes numerous economic plant diseases, such as carrot bolting, white heart of lettuce, spinach yellows, potato purple top, onion yellows, and many others. Observations of phloem cells of infected plants permitted the detection of "MLOs," (mycoplasma-like organisms), but gave no indication that the causative agent is a spiroplasma (Fig. 1). Yet, upon isolation and cultivation in broth media, typical spiral forms appeared. In older cultures on agar spiral forms were also produced.[43] Until now, the only crude plant sap in which we were able to find spiral forms was the sap from diseased lettuce. When we examined carefully electron micrographs from *Nicotiana rustica* phloem, we were able to detect elongated, twisted MLO's but these were extremely rare. This indicated that the formation of spiral forms by the aster yellows spiroplasma requires certain, as yet undetermined conditions. Aster yellows is worldwide in its distribution, and in recent years it has been described from Kashmir.[59] It does not occur in tropical areas.

Spiroplasmas pathogenic to honeybees and fruit flies

In 1977 Clark discovered very small, helical organisms associated with a disease of honeybees derived from Florida. Subsequently spiroplasma-infected bees have also been collected in nature in Hawaii, Maryland, New York, New Jersey, and Louisiana. Davis *et al.*[14] were able to isolate and culture the spiroplasma and prove that it was the causal agent of the fatal bee disease. The microorganism proved infectious to fruit flies and other insects. The natural mode of acquisition of this spiroplasma is not known, however, although it has been suspected that bees might acquire it from flower nectar. This hypothesis led Davis to search for spiroplasmas in the nectar of tulip trees (Liriodendron). Numerous spiroplasmas were detected in the nectar, but they were different from the honeybee pathogens, and, in fact, were not pathogenic to plants or insects.[18]

In 1957 Malogolowkin and Poulson described a sex ratio phenomenon in several Drosophila species and reported that the disease agent can be transmitted by needle inoculation, using insect hemolymph as inoculum. Although no proper culture medium is known until now for the propagation of the microorganism, electron micrographs have revealed that it is a member of the Spiroplasmataceae[72]. The sex ratio spiroplasma kills the embryo

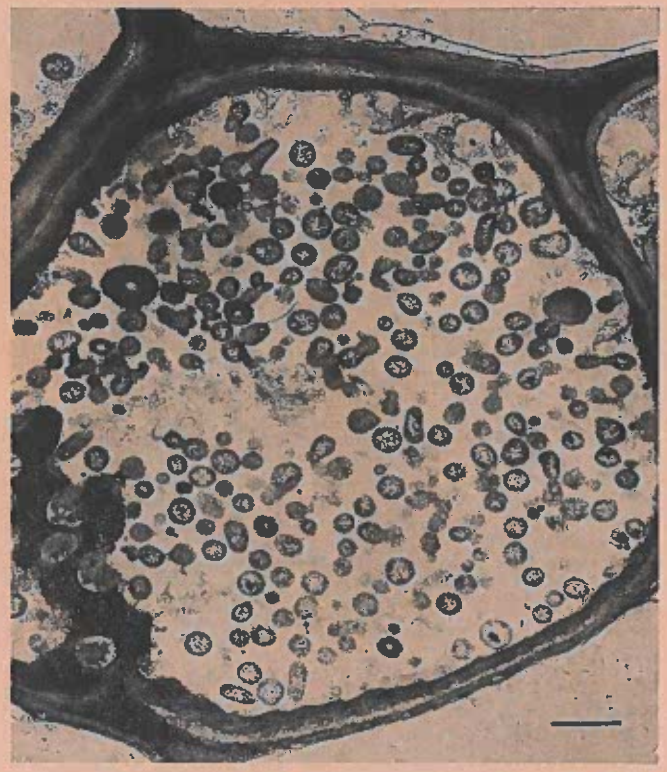


Fig. 1. Cross-section through a phloem cell of a China aster (*Callistephus chinensis* Nees) plant infected with aster yellows spiroplasma. Note that the pathogens are rounded, not spiral. Magnification ×21,000. Original electron micrograph by F. Kondo, Waksman Institute of Microbiology.

Fig. 2. Higher magnification of aster yellows spiroplasmas in a China aster phloem cell. The microorganisms are bounded by a unit membrane and contain ribosomes and thin nuclear strands. Magnification ×67,000. EM by F. Kondo.

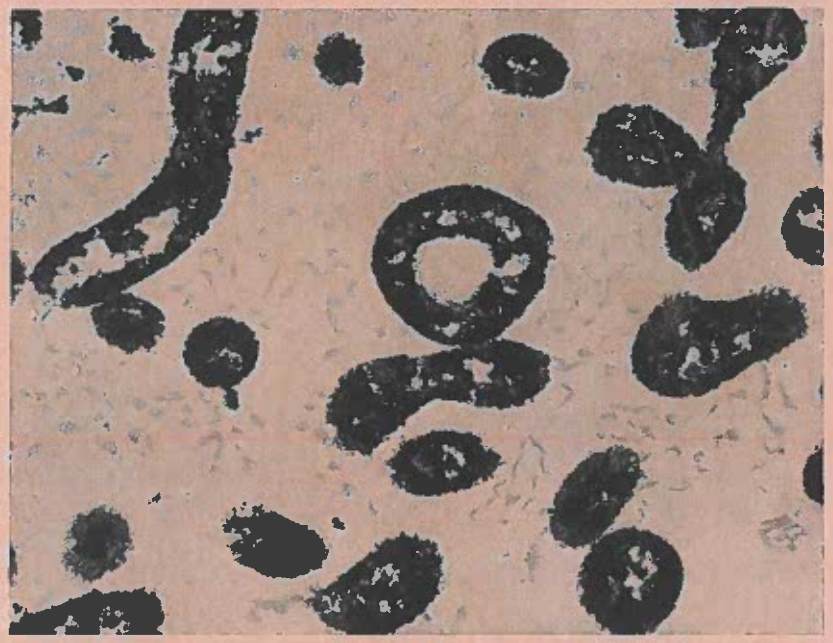


Fig. 3. Cultured aster yellows spiroplasma : rounded, elongated, and ring structures in broth culture. Original EM by D. M. Phillips, Rockefeller University, New York.

Fig. 4. Cultured aster yellows spiroplasma, negatively stained with phosphotungstic acid. Magnification ×45,000. EM by F. Kondo.

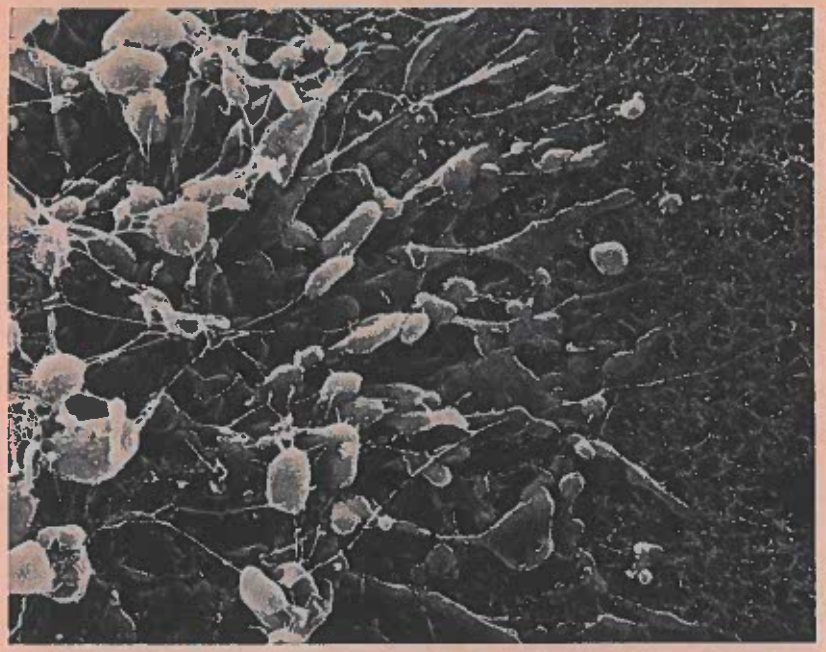


Fig. 5. Aster yellows spiroplasma growing on agar. Note rounded bodies and mycelium-like thin protrusions. Scanning electron micrograph by D. M. Phillips.

Fig. 6. Older culture of aster yellows spiroplasma on agar, with twisted, cork-screw forms. EM by D. M. Phillips.

Fig. 7. Aster yellows spiroplasma with rod-shaped virus. Arrows point to particles adhering to cell membrane and to a virus inclusion within the deteriorating cell. ×80,000. EM by F. Kondo.

or later stages of male fruit flies, but does not affect the female flies. Thus, it is androcidal.

A different spiroplasma, which does not alter the sex ratio, has been found recently in Japan in *Drosophila hydei* (Oishi, personal communication). The spiroplasma is usually associated with a virus, similar or identical to the virus reported from the sex ratio spiroplasma. When the virus is injected experimentally into the hemocoele of spiroplasma-carrying female fruit flies, the spiroplasmas are totally destroyed. If this is done in flies infected with the sex ratio microorganism, the normal sex ratio is restored and males are again produced.

Spiroplasmas pathogenic to higher animals

Clark[8,9] isolated a pathogenic agent from rabbit tickets, *Haemaphysalis leporispalustris*. This agent is now known to be a member of the family Spiroplasmataceae[5]. Upon intracerebral inoculation into newborn mice and rats, this spiroplasma multiplies and produces cataracts and chronic infection of the brain of the inoculated mammals[8,9]. The spiroplasma can be grown in embryonated chicken eggs, killing the embryos within a few days. Broth and agar cultures produce spiroplasmas that are morphologically similar to *S. citri*, but the optimum temperature for cultivation is 37°C, while *S. citri* and the aster yellows spiroplasma grow best at 25–29°C, and they cannot grow at 37°C. In fact, 32°C is critical for the aster yellows spiroplasma and at this and higher temperatures the plant pathogen is destroyed completely. The optimum growth at 37°C is also the normal temperature of the human body. This finding, coupled with the ability to multiply in various tissues of rodents, might have far reaching implications. The etiology of human cataracts is unknown, as is the etiology of several other human diseases. Could spiroplasmas be implicated? There are no "standard" media for the isolation of spiroplasmas and the cataract agent could not be isolated on conventional mycoplasma media. It is, therefore, conceivable that not only additional insect diseases, but also animal and human diseases caused by spiroplasmas might be uncovered in the future.

Morphology of spiroplasmas

The morphology and fine structure of spiroplasmas has been studied by phase contrast and darkfield microscopy, as well as by transmission and sca-

nning electron microscopy. The standard use of cacodylate buffer in the preparation of material for electron microscopy has resulted in the distortion of spiroplasmas[12]. In broth cultures the formation of spiral forms depends on several factors, such as the age of the culture, the composition of the medium, and the species of spiroplasma. One variant of *S. citri* does not produce spirals. The aster yellows spiroplasma ordinarily produces spirals in the sap of diseased lettuce (*Lactica sativa L.*) and in broth culture, but not in diseased aster plants (*Callistephus chinensis* Nees). The corn stunt spiroplasma occurs in spiral forms in the sap of diseased corn plants, infected with the Rio Grande stunt, but not in plants infected with the Mesa Central strain. Yet, there is interference ("cross-protection") between these two strains, as demonstrated several years ago[14]. Because of these intriguing morphological characteristics and the failure to detect spiral forms in certain host plants infected with these agents, several plant pathologists and entomologists proposed to separate mycoplasmas detected in plants and insects into spiroplasmas proper and MLOs, that is, mycoplasma-like organisms. The latter group comprises mycoplasmas that have been observed by electron microscopy, but not yet isolated and cultured. This artificial division has resulted in confusing statements that MLOs are different from spiroplasmas and, like the Rickettsiae, might not be cultivable. This reasoning overlooks the fact that until 1972–73 all spiroplasmas known today have been known as MLO's and only gradually nearly a dozen have been isolated in cell-free media.

Thin sections of cultured spiroplasmas, as well as spiroplasmas within phloem cells of diseased plants or in salivary glands of insect vectors reveal a simple ultrastructure. A cell membrane surrounds the microorganisms (Figs. 1 and 2). Inside the cell of spiroplasmas there are ribosomes, containing RNA, and the characteristic prokaryotic genome, a single coil of DNA, visualized in sectioned material as separate strands in the nuclear region. The morphology of cultured forms as well as of spiroplasmas within the phloem of diseased plants has been studied extensively by electron microscopy. The salient features of these agents are illustrated in Figs. 1-6. The existence of ring forms in spiroplasmas, illustrated in Fig. 3, is reminiscent of ring forms that are predominant in Eperythrozoon, a genus in the family Bartonellaceae, order Rickettsiales; however, Eperythrozoon is a walled bacterium, while spiroplasmas have no cell wall. Furthermore, the guanine+cytosine ratio (G+C) is lower than the ratio known for bacteria. The genome size of spiroplasmas is less than one-half of the size recorded for most bacteria.

Negative staining with phosphotungstic acid reveals the typical spiral structure (Fig. 4).

Scanning electron micrographs of cultures of spiroplasma grown on agar are shown in Figs. 5 and 6. Rounded bodies are predominantly found at first, when cultures begin to grow on semi-solid agar media. In older cultures twisted, spiral forms are formed as well. This has been found in aster yellows spiroplasma and it revealed that the ability to form spirals is an inherent characteristic of these microorganisms, dependent, to some extent, on the substrate medium[8,6].

Other as yet uncultured yellows disease agents (MLOs)

At the turn of the century a mysterious and devastating disease of coconut palms appeared in Cuba. Within 5 years, it destroyed and killed practically all coconut palms on the largest island in the Carribean. The epiphytotic disease in Cuba had profound political and economic consequences. Instead of coconut palms that no longer could be grown there, soon large areas were converted to sugarcane plantations. The economy of the island changed and the sugar barons became very wealthy, while the vast majority of the population became more and more impoverished. The conditions for a successful take-over by Castro's regime were thus provided.

What caused the destruction of coconut palms in Cuba ? Was the disease new ? If not, where did it come from ? What was the infectious agent of the palm disease ? Studies carried out over the years were inconclusive. A description of an earlier outbreak on Great Caiman Island indicated that the same, or a very similar disease struck in the Carribean Islands some 100 years earlier. A bacterium was found in dying palms but it was not possible to reproduce the disease using a pure culture of this bacterium.

Several years later the same disease appeared on the north-western shores of Jamaica. It was called the West End Budrot, as only palms on the west end of the island seemed to become affected. Although no association with a virus could be established, experts were convinced that the disease was due to a virus, because no fungi or bacteria were incriminated this time.

In 1963 the Food and Agriculture Organization of the United Nations requested that I make a worldwide survey of coconut diseases of uncertain etiology. Among others I visited several Carribean Islands, Key West, Florida, and in West Africa Ghana and Togo. I found that the same disease,

under different names, occurred in America and in Africa. The English name, now generally accepted, Lethal Yellowing Disease of coconut palms, is synonymous with the Cape St. Paul. Wilt in Ghana, Kaincope disease in Togo, and Kribe disease in East Africa. The causative agent eluded detection until 1972, when Biljana Plavsic, a Yugoslav botanist, on a Fulbright-Hays Fellowship in my laboratory examined material collected by me and Peter Hunt in Jamaica. In ultrathin sections of the phloem cells of diseased palm inflorescences she detected microorganisms that closely resembled mycoplasmas[54,55].

Attempts by several groups of plant pathologists and entomologists, assisted by mycoplasmologists and bacteriologists, have been made during the past decade to isolate and culture the lethal yellowing disease pathogen. Until now no success has been reported. The search for a vector, intensely pursued in Florida and Jamaica, has indicated a leafhopper carrier[66] but final confirmation of the findings is still lacking.

Is lethal yellowing caused by a spiroplasma or by a mycoplasma-like organism, an "MLO"? The answer to this question must await the successful cultivation of the pathogen. Let us remember that all presently known plant-pathogenic spiroplasmas were, at first, classified as MLOs and those that have been properly cultivated are now recognized as spiroplasmas. No single culture medium exists for all different spiroplasmas and the isolation of a new member of this group still depends on trial and error.

The status of the causative agents of yellows-type diseases in India

Approximately 100 different plant diseases are now known around the world to be associated with mycoplasma-like agents. Only a few of those agents have been characterized properly as spiroplasmas and the remainder remains grouped as MLOs. Mycoplasmas have been observed by electron microscopy in thin sections of sieve elements, and, in a few instances where vectors were known and carefully studied, in the cells of these vectors. In addition, dodder used to transmit the disease agent from plant to plant, also has been found to contain MLOs. Nine MLO diseases studied in India in detail are listed in the 1977 book by S. P. Raychaudhuri and T. K. Nariani, entitled "Virus and Mycoplasma Diseases of Plants in India."[59] The first of these diseases is of special scientific as well as economic interest.

Sandal spike. This disease has received careful attention since it was first described in India nearly 90 years ago. Coleman in his classic studies called attention at the other yellows-type diseases in forests where sandal spike occurred and his suggestion that the disease affects numerous other plant species showed his great vision and observational powers. In 1932 Narasimhan published his work in the journal Phytopathology, calling attention to the fact that the host plants onto which sandal grafts itself also become diseased with a similar disease. This finding opened the way to the artificial transmission of other yellows disease agents, using various species of dodder (cuscuta), that can graft themselves onto unrelated plant species. In 1969 Varma *et al.*[70], Hull *et al.*[34], and Dijkstra and Ie[17] demonstrated the presence of mycoplasma-like bodies in spiked plants. When the agent was transmitted via cleft graft, mycoplasma-like organisms were detected in cross-sections of sieve elements[54]. A recent survey by Muniyappa and Subba Rao of sandal forests, carried over the past 3 years, confirmed that the disease is most severe in the southern part of Karnataka State and that further north sandal has remained healthy. This provides a lead for those who are trying to discover the insect vectors of sandal spike mycoplasma. Certain species of leaf hoppers have been incriminated but none confirmed. Cultivation attempts of the infectious agents have been carried out in several laboratories but claims of successful cultivation have remained unconfirmed.

Practically all MLOs and spiroplasmas that cause plant diseases can be transmitted to *Vinca rosea*, either by dodder or by specific insect vectors. *Vinca* can harbor the spiroplasmas of aster yellows, corn stunt, citrus stubborn as well as those of poorly defined MLOs. If and when the sandal spike agent is cultured in cell-free media, it will be posible to test its susceptibility *in vitro* to chemotherapeutic agents. Until now, such tests were possible only by using infected plants *in vivo*. The *in vivo* tests are less accurate in assessing the chemicals. The ultimate solution to eliminate losses due to sandal spike disease will probably come from a combined effort in which resistant sandal trees will be grown, and where the insect vector and its alternate hosts will be prevented from close contact with sandal trees. In addition, inexpensive chemotherapeutic compounds might be found to treat affected trees.

Rice yellow dwarf. This disease, known from Orissa, Bihar, West Bengal, Andhra Pradesh and Delhi[68] is similar to the rice yellow dwarf of Japan, known since 1967 as a mycoplasma disease. Association with MLOs of plants

in India has been described in 1972[44,55]. The leafhopper vector, *Nephotettix impicticeps* (Ishihara) can be controlled by systemic insecticides, such as Furadan. Breeding for resistance to this vector appears to be the most promising control approach[58]. The inheritance of resistance has been studied first by Athwal *et al.* (1971), and more recently by Siwi and Khush[68]. Work is now in progress to obtain varieties with better levels of resistance to the green leafhopper[17].

Cotton small leaf. This disease, also called cotton stenosis, has been shown associated with mycoplasma[5]. It has been described from the Punjab, Madras, and Maharashtra[68]. The insect vector has not yet been discovered. With the increased cotton production, especially the introduction of high yielding hybrid cotton in India, this MLO disease will again attract the attention of plant pathologists.

Potato purple top. In 1973, Nagaich and Giri demonstrated the presence of MLOs in diseased potato plants[50]. The leafhopper *Orosius albicinctus* Distant has been described as the vector. Purple top of potato has been described from other countries as well and there is strong evidence that it is caused by strains of aster yellows spiroplasma. Since the latter has been cultured in cell-free media, similar tests could be carried out with the purple top agent, by transmitting it via dodder to alternate hosts that would be more suitable for isolation procedures.

Little leaf of Brinjal. Brinjal, *Solanum melongena*, also known as egg-plant, is affected by a phyllody disease and the causal agent has been shown to be transmitted by *Hishimonus phycitis* more than 40 years ago[65]. In 1969 Varma *et al.* demonstrated by electron microscopy the presence of MLOs in diseased plants[70]. Whether potato purple top and brinjal little leaf are caused by related or identical strains of the same mycoplasma has not been established, but since the vectors differ, most likely the MLO's differ as well.

Phyllody of sesamum. According to Raychaudhuri and Nariani[59] the disease is known throughout India, wherever sesamum is grown. The vector, *O. albicinctus*, also transmits the MLO of potato purple top. This might indicate that the same MLO is the causative agent of both diseases.

Apple rubbery wood. This disease of apple trees in Himachal Pradesh has been found associated with MLOs in 1971 by Beakbane *et al.* No vectors have been found as yet and since the agent is graft-transmissible certified healthy stock is recommended in apple orchards.

In addition to the above listed diseases, a large number of other yellows-type diseases has been described from various areas, including Jaipur[23,47]. In some, the presence of MLOs has been confirmed[42]. The difficulty in ascertaining the mycoplasma etiology by electron microscopy can now be overcome, because simple staining with Diene's stain and light microscopy observation of the stained material can be used[16].

White leaf disease of Bermudagrass. A previously undescribed yellows-type disease of Bermudagrass was noted near Bangalore, India by Muniyappa *et al.* High rates of disease incidence were observed near rice and sugarcane fields. Electron microscopy of thin sections carried out by Muniyappa *et al.*[48] revealed the presence of mycoplasmas and the etiology of the disease was further confirmed by hot water therapy and tetracycline chemotherapy. The occurrence of white leaf disease in almost all of the rice and sugarcane-growing areas of India makes this disease a potentially important entity. As pointed out by Ghosh and Raychaudhuri[18] Bermudagrass may be a significant source of yellows-type mycoplasmas infecting rice and sugarcane. The wide distribution of this disease indicates the presence of an efficient vector. However, the vector has not yet been determined. In Taiwan, where this disease also occurs in Bermudagrass, the vector has remained unknown until now. A spiroplasma has been isolated and cultured from Bermudagrass white leaf diseased material by Chen *et al.*[7] and Raju and Chen[57] but the spiroplasma could not be transmitted to healthy plants because the vector remained unknown.

Current status of citrus greening. A few years ago, it was generally assumed that citrus greening is caused by a mycoplasma agent. Work by Lafleche and Bove[29] and by Ghosh *et al.*[19] demonstrated the presence of microorganisms in the phloem of diseased plants, and it was assumed that the bodies were MLOs. More recently, in India Govindu and Naidu (personal communication) and in France Saglio[60] described the morphology of the greening agent and established that it possesses a cell wall instead of a unit-type membrance. Therefore, the greening agent is not a member of the Class Mollicutes.

Present and future control approaches

With very few exceptions, MLO and spiroplasma diseases of plants cannot be controlled by controlling mycoplasma vectors. The reason is simple —we are unable to eliminate 100% of the vectors and can only reduce the ve-

ctor population. Mere decrease is inadequate to prevent further spread of the disease agents. This is particularly obvious in tropical and subtropical climates. For instance, the green leafhopper that carries the MLO of rice yellow dwarf can have 10 generations in a single year and thus it cannot be destroyed by a single, or even by repeated insecticide treatments. New insecticides might be more efficient but their cost is high. In temperate zones and in instances where a vector has but one or two generations per year, as is the case with the vectors of pear decline, cranberry false blossom and blueberry stunt in the United States of America, chemical treatment of pear trees, or use of insecticides in blueberry and cranberry plantings has proved economical. But usually chemotherapy is much too costly. Neither rice, sugarcane or potatoes can be treated in the field with tetracycline drugs, even though these drugs affect directly the causative spiroplasmas or MLOs. Tetracyclines work only when introduced via roots or injected into the stems of trees, but not when sprayed onto leaves or poured on the soil. More important, drugs intended for human use should not be spread in the environment, as this might result in creating resistant strains of human pathogens.

There are other possible approaches to control plant mycoplasma diseases. Breeding for resistance to the MLOs and spiroplasmas is the standard approach. Breeding for resistance to insect vectors is another approach, and this can be incorporated into the breeding program along with the breeding for resistance to insect pests[80, 46].

Other approaches are still only on the drawing board and it might take several years before they can be tested, but some will be mentioned here. Many years ago Kunkel[28] demonstrated that one strain of aster yellow can "protect" against infection by another strain. While working in Kunkel's laboratory at the Rockefeller University, I found a few years later (1958) that the same principle of "cross-protection" applies to two strains of the corn stunt agent[84]. Thus, in two spiroplasma diseases, long before the existence of spiroplasma was known, the phenomenon of strain interference has been recognized. This principle could be applied to protect plants from infection with a severe, wild strain by inoculating them preventively with a mild strain. In plant pathology this principle took hold only recently, when Dutch scientists patented a mild strain of tobacco mosaic virus for use as a preventive vaccine of tomato seedlings. The mild inoculum is commercially sold. It is more difficult to inoculate plants with MLOs and spiroplasmas, because these agents cannot be transmitted mechanically to plants. The need for proper insect vec-

tors complicates the actual application of the method, but in instances where the plant is a tree, such as a sandal tree or a coconut palm, the method could be tried.

Another approach to the control of mycoplasma diseases in even more speculative and at present in the realm of basic, rather than applied, research. The citrus stubborn spiroplasma, *S. citri,* is known to be affected by three different viruses : two bacteriophages and a rod-shaped virus[11]. The aster yellows spiroplasma is affected by a rod-shaped virus[88] (Fig. 7). The rod-shaped virus can destroy the spiroplasma in cultures, and evidence from electron micrographs of diseased phloem elements indicates that viral destruction also might occur *in vivo.* If we could manipulate the balance between the spiroplasma viruses and the spiroplasmas, it would seem conceivable to use this for the *in vivo* destruction of the infectious mycoplasmas and to cure the plants. This might sound today like science fiction, but there is already one instance known, in which virus infection of a spiroplasma causes a cure *in vivo.* This instance is not known from plants but from the androcidal spiroplasma infection of fruit flies. Poulson and Oishi[56] demonstrated that the natural virus infection of Drosophila spiroplasma can proceed so far as to eliminate all spiroplasmas from the body of a fly. Consequently, instead of producing only female offspring, as is the case in infected flies, both males and females are produced, at a 50/50 ratio.

If methods were developed to produce large amounts of aster yellows or citrus stubborn spiroplasma viruses and their introduction into living infected plants were perfected, one could expect a similar cure as in the case of spiroplasma-infected Drosophila.

While we are speculating on future approaches, one more of the as yet strictly theoretical means of control can be outlined. Spiroplasmas are not entirely harmless to their insect vectors. In 1956-1960 a study of salivary glands of *Macrosteles fascifrons,* the vector of the aster yellows spiroplasma, showed severe cytopathic changes in infected insects[81, 82]. One strain of the corn stunt spiroplasma has been found to have a deleterious effect on the corn leafhopper vectors[21]. Once these agents are obtained in culture, standard bacteriological methods can be used to induce chemically or physically mutations. Some mutants can be expected to be more harmful to the vectors than the original isolate, and through such manipulation and selection one could expect to develop strains that would kill leafhopper vectors before

these vectors get an opportunity to transmit the spiroplasmas to susceptible plants.

The above-mentioned approaches are still highly speculative. In the meantime, we have to rely on available methods of control, primarily on the search for resistant varieties—resistant not only to the spiroplasmas but also to their vectors. In the future, most likely an integrated control approach will prove the most practical and economical. Control will eventually be achieved through a combination of basic and applied research and through close international cooperation among scientists.

REFERENCES

1. Athwal, D. S., Pathak, M. D., Bacalangoo, E. H., and Pura, C. D. (1971) Crop. Sci. 11 : 747.
2. Barile, M. F., and Razin, S. (1979) eds. The Mycoplasmas I : Cell Biology. Academic Press, New York, 547 pp.
3. Beakbane, B., Mishra, M. D., Posnette, A. F., and Slater, C. H. W. (1971) J. Gen. Microbiol. 66 : 55.
4. Brinton, L. P., and Burgdorfer, W. (1976) Int. J. System. Bact. 26 : 554. —
5. Capoor, S. P., Pande, P. K., and Sinha, R. C. (1972) Hindustan Antibiotics Bull. 15 : 40.
6. Chen, T. A., and Liao, C. H. (1975) Science 188 : 1015.
7. Chen, T. A., Su, H. J., Raju, B. C., and Huang, W. C. (1977) Proc. Am. Phytopathol. Soc. 4 : 171.
8. Clark, H. F. (1964) J. Infect. Dis. 114 : 476.
9. Clark, H. F. (1974) Prog. Med. Virol. 18 : 307.
10. Clark, T. B. (1977) J. Invert. Pathol. 29 : 112.
11. Cole, R. M. (1977) In : The Atlas of Insect and Plant Viruses, K. Maramorosch, ed. Academic Press, New York : 451.
12. Cole, R. M,, Tully, J. G., Popkin, T. J., and Bove, J. M. (1973) J. Bacteriol. 115 : 367.
13. Davis, R. E. (1978) Can. J. Microbiol. 24 : 954.
14. Davis, R. E., Lee, I. M., and Basciano, L. K. (1979) Can. J. Microbiol. 25 : 861.
15. Davis, R. E., and Worley, J. F. (1973) Phytopathology 63 : 403.
16. Deeley, J., Stevens, W. A., and Fox, R. T. V. (1979) Phytopathology 69 : 1169.
17. Dijkstra, J., and Ie, T. S. Neth. (1969) J. Plant Pathol. 75 : 374.
18. Fudl-Allah, E. A., Calavan, E. C., and Igwegbe, E. C. K. (1972) Phytopathology 62 : 729.
17. Gallun, R. L., and Khush, G. S. (1980) In : Breeding Plants Resistant to Insects. F. G. Maxwell and P. R. Jenninfs, eds. Wiley-Interscience, New York : 64.

18. Ghosh, S. K., and Raychaudhuri, S. P. (1974) In : Current Trends in Plant Pathology. S. P. Raychaudhuri and J. P. Varma, eds. Lucknow Univ. Press, India : 112.
19. Ghosh, S. K., Raychaudhuri, S. P., Varma, A., and Nariani, T. K. (1971) Curr. Sci. 40 : 299.
20. Granados, R. R., Maramorosch, K., and Shikata, E. (1968) Proc. Nat. Acad. Sci. (U. S.) 60 : 841.
21. Granados, R. R., and Meehan, D. J. (1975) J. Invert. Pathology 26 : 313. 1975.
22. Haggis, G. H., and Sinha, R. C. (1978) Phytopathology 68 : 677.
23. Hegde, R. K., Siddaramaiah, A. L., and Krishna Prasad, K. S. (1977) Plant Dis. Reptr. (U. S.) 61 : 588.
24. Hull, R., Horne, R. W., and Nayar, R. M. (1969) Nature (London) 234 : 1121.
25. Kondo, F., McIntosh, A. H., Padhi, S. B., and Maramorosch, K. (1976) Proc. 34th Electron Microscopy Soc. Amer. : 56.
26. Kondo, F., and Maramorosch, K. (1977) Abstracts Ann. Meeting Amer. Soc. Microbiol. : G19.
27. Kondo, F., Maramorosch, K., McIntosh, A. H., and Varney, E. H. (1977) Proc. Amer. Phytopathol. Soc. 4 : 190.
28. Kunkel, L. O. (1955) Adv. Virus Research 3 : 251.
29. Lafleche, D., and Bove, J. M. (1970) C. R. Acad. Sci., Paris, 270 : 1915.
30. Liao, C. H., and Chen, T. A. (1975) Proc. Am. Phytopathol. Soc. 2 : 100.
31. Littau, V. C., and Maramorosch, K. (1960) Virology 10 : 483.
32. Littau, V. C., and Maramorosch, K. (1956) Virology 2 : 128.
33. Malogolowkin, C., and Poulson, D. F. (1957) Science 126 : 32.
34. Maramorosch, K. (1958) Virology 6 : 448.
35. Maramorosch, K. (1975) Proc. Indian Natl. Sci. Acad. 41 (B) : 298.
36. Maramorosch, K. (1976) In : Encyclopedia of Plant Physiology. IV. P. H. Williams and R. Haytefuss, eds. Springer : 150.
37. Maramorosch, K. (1979 a) In : Plant Disease 4. J. G. Horsfall and E. B. Cowling. eds. Academic Press, New York : 203.
38. Maramorosch, K. (1979b) Abstracts Ann. Meeting Amer. Soc. Microbiol. : G20.
39. Maramorosch, K. (1980) In: Breeding Plants Resistant to Insects. F. G. Maxwell and P. R. Jennings, eds. Wiley-Interscience, New York : 138.
40. Maramorosch, K., Granados, R. R., and Hirumi, H. (1970) Adv. Virus Research 16 : 135.
41. Maramorosch, K., Hirumi, H., Kimura, M., and Bird, J. (1975) Ann. N. Y. Acad. Sci. 266 : 276.
42. Maramorosch, K., Misra, S., and Kondo, F. (1977) Marcelia 40 : 225.
43. Maramorosch, K., and Phillips, D. M. (1978) Zentr. Bakteriol. Parasit. infekt. u. Hygiene 241 : 224.
44. Maramorosch, K., Plavsic-Banjac, B., John, V. T., and Raychaudhuri, S. P. (1972) Phytopathology 62 : 776.
45. Maramorosch, K., Shikata, E., and Granados, R. R. (1968) Trans. N. Y. Acad. Sci. 30 : 841.
46. Maxwell, F. G., and Jennings, P. R., eds. (1980) Breeding Plants Resistant to Insects. Wiley-Interscience, New York 683 pp.
47. Misra, S., Manju Behl and Sharma, S. (1979) Abstracts All India Workshop on Mycoplasma, Bangalore : 6.
48. Muniyappa, V., Raju, B. S., and Nyland, G. (1979) Plant Disease Reptr. 63 : 1072.

49. Muniyappa, V., and Subba Rao. (Personal Communication). 1980.
50. Nagaich, B. B., and Giri, B. K. (1973) Amer. Potato J. 50 : 79.
51. Nyland, G., and Moller, W. J. (1973) Plant Disease Reptr. (U. S.) 57 : 634.
52. Padhi, S. B., McIntosh, A. H., and Maramorosch, K. (1977) Phytopathol. Z. 90 : 268.
53. Pathak, M. P., and Saxena, R. C. (1980) In : Breeding Plants Resistant to Insects. F. G. Maxwell and P. R. Jennings, eds. Wiley-Interscience, New York : 422.
54. Plavsic-Banjac, B., Maramorosch, K., Raychaudhuri, S.P., Chenulu, V.V., Varma, A., and Ghosh, S. K. (1973a) FAO Plant Prot. Bull. 21 : 25.
55. Plavsic-Banjac, B., Maramorosch, K., John, V. T., and Raychaudhuri, S. P. (1973b) FAO Plant Prot. Bull. 21 : 1.
56. Poulson, D. F., and Oishi, K. (1975) Int. Virol. 3 : 145.
57. Raju, B. C., and Chen, T. A. (1977) Proc. Am. Phytopathol. Soc. 4 : 195. —
58. Raychaudhuri, S P., Mishra, M. D., and Ghosh, A. (1967) Plant Dis. Reptr. 51 : 1040.
59. Raychaudhuri, S. P., and Nariani, T. K. (1977) Virus and Mycoplasma Diseases of Plants in India. 102 p. Mohan Primlani, Oxford and IBH Publ. Co., New Delhi.
60. Saglio, P. H. M. (1979) Plant Dis. Reptr. (U.S.) 63 : 293.
61. Saglio, P., Lafleche, D., Bonisol, C., and Bove, J. M. C. R. Hebd. (1971) Seances Acad. Sci. Ser. D 272 : 1387.
62. Saxena, K. N., ed. Mycoplasmal Diseases. (1975) INSA Bull. 52, Indian National Sci. Academy, New Delhi, Proc. Indian Natn. Sci. Acad. 41(B).
63. Siwi, B. H., and Khush, G. S. (1977) Crop. Sci. 17 : 17.
64. Skowronski, B. S., McIntosh, A. H., and Maramorosch, K. (1974) Plant Disease Reptr. (U.S.) 58 : 797.
65. Thomas, K. M., and Krishnaswami, C. S. (1939) Proc. Indian Acad. Sci 10(B) 201.
66. Tsai, James H. In : The Mycoplasmas III. R. F. Whitcomb and J. G. Tully (1979) eds. Academic Press, New York : 265.
67. Tully, J. G., Whitcomb, R. F., Clark, H. F., and Williamson, D. L. (1977) Science 195 : 892.
68. Uppal, B. N., Capoor, S. P., and Raychaudhuri, S. P. (1944) Curr. Sci. 13 : 284.
69. Varma, A., Asha Sang, S. K., Ghosh, S. K., Raychaudhuri, S. P., Chenulu, V.V., and Nam Prakash. (1974) Cur. Sci. 43 : 349.
70. Varma, A., Chenulu, V. V., Raychaudhuri, S. P., Prakash, N., and Rao, P. S. (1969) Indian Phytopathology 22 : 289.
71. Whitcomb, R. F., and Tully, J. G. eds. (1979) The Mycoplasmas III. Plant and Insect Mycoplasmas. Academic Press, New York, 351 pp.
72. Williamson, D.L., and Poulson, Donald F. (1979) In : The Mycoplasmas III. R.F. Whitcomb and J. G. Tully, eds. Academic Press, New York : 175.
73. Williamson, D. L., and Whitcomb, R. F. (1975) Science 188 : 1018.

Genetics Of Human Serum Lactate Dehydrogenase —A Review

Bijaya Kumar Swain[1], Archana Sharma[1] and Geeta Talukder[2]

[1]Human Genetics Unit, Department of Botany
University of Calcutta
35 Ballygunge Circular Road, Calcutta-700 019

[2]Department of Pathology and Bacteriology
Institute of Post-Graduate Medical Education and Research
Calcutta-700 020

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

Introduction

The use of changes in enzyme activities in the diagnosis of clinical conditions is receiving progressively greater attention since the observation that the enzyme contents of extracellular body fluids can furnish important data on the localization, nature and development of certain pathological processes. In this context, the enzyme lactate dehydrogenase (LDH) occupies a special position due to the fact that its activity was found to increase in the sera of patients suffering from a variety of disorders of the heart and liver and different types of malignancy.

It occurs in a series of polymorphic forms which give rise to five identifiable fractions following electrophoretic analysis[353] which are identified serially as LDH-1 to LDH-5[159,171,230,311,330]. LDH occupies a key position in carbohydrate metabolism at the point where both aerobic and anaerobic decompositions meet and the concentrations of different fractions (1 to 5) vary with the tissue concerned. The principal function of the enzyme is to regulate the intracellular ratio NAD to $NADH_2$.

More than 14 different genetically determined variants of both A and B subunits of human LDH have been described in a number of populations from all over the world[142,190,303] but the frequencies of such variants were always low. In India, the different types of genetic variants have been located from populations all over the country. They have often been given trivial names as Calcutta-1, Calcutta-2, Madras-1 and Delhi-1[8,25,63,193,195].

Nomenclature of LDH

The five different LDH isozymes are identified serially as 1 to 5, the fastest fraction being referred to as LDH-1. This system has now been accepted by most workers[159,230,311,330,346].

Richterich *et al.*[233] has suggested naming the LDH after the corresponding protein fractions–α_1 LDH, α_2 LDH, β LDH and γ LDH. This system, however, indicates only four out of the five fractions. Since the fractions do not necessarily occur at the same location under different methods of fractionation any particular nomenclature implicitly presupposes knowledge of the method applied. Zondag[353], following agar gel electrophoresis at a pH 8·36, and Wieme[330] identified the fractions as :

LDH-1 : fastest, most anodic fraction situated between albumin and α_1-globulin.

LDH-2 : coincident with the α_2-globulin fraction.

LDH-3 : coincides with β-globulin fraction.

LDH-4 : coincides with fast γ-globulin fraction.

LDH-5 : The slowest, most cathodic fraction coinciding with the slow moving γ-globulin fraction.

Physico-chemical Properties

LDH is the prototype of an oligomeric, intracellular enzyme[217]. The isoenzymes differ from one another in a variety of ways, including affinity for the substrates[209,221,307] behaviour towards coenzyme analogues[130,209,301] sensitivity to inhibitors[39,75,224,325], Wieland and pH optima[306] and thermostability[223,226,250]. In many physico-chemical parameters, some isoenzymes are also remarkably alike, as for example, in sedimentation and diffusion coefficients; partial specific volume, molecular weight and frictional ratio.

A distinguishing feature of LDH isoenzymes is their inhibition by substrate. Even before separating their crude homogenates into isoenzymes, Kaplan and Ciotti[181] suspected heterogeneity in the enzyme because the extracts of one tissue of an organism varied in affinities for substrate and coenzyme from extracts of other tissues of the same organism. Pfleiderer and Wachsmuth[216] observed LDH-5 to be resistant to the inhibitory effects of lactate on LDH-1. They therefore suggested that LDH-5 predominated in anaerobic tissues, like liver and skeletal muscle, due to its capacity to function in the presence of high lactate concentrations while LDH-1 predominated in aerobic tissues, like brain and heart, where lactate accumulation was minimal. The inhibition of LDH-1 to a greater extent than LDH-5 by a given concentration of pyruvate in tissues most dependent on anaerobic metabolism has been used to interprete the variable isoenzyme patterns of different human skeletal muscles[128] and also of identical muscles from different species[180]. These species differences in isoenzyme pattern have been correlated with evolutionary alteration in the physiological role of a specific muscle.

LDH serves to regulate the intracellular ratio of NAD to $NADH_2$. The ratio is maintained above one, both by the establishment of an enzyme-dependent equilibrium constant for the reaction which is ten times more favourable for NAD than for $NADH_2$ formation as well as by the greater affinity of LDH for pyruvate (approximately $KM^{10^{-5}}$) than for lactate (approximately $KM^{10^{-3}}$). The isoenzymes vary in affinity for each of these two subs-

trates. Therefore, due to its specific intracellular location each isoenzyme may develop slightly different ratio of NAD to $NADH_2$ within various areas of the cell.

"Nothing dehydrogenase"

Sometimes a band is observed at the top of the separating gel (at the charging point), erroneously called as LDH-5 by Konda[150]. Dietz and Lubrano[69] showed the band to persist in the absence of lactate from the staining solution. They called it "nothing dehydrogenase", attributing the diffuse staining to protein-SH group. It was however, considered as an alcohol dehydrogenase by Shaw and Koen[258] and protein-SH group by Katz and Kalow[134].

Subunit structure of LDH

Markert and Appella[172] proved that LDH molecule consists of a number of smaller polypeptide chains, established by dissociation experiments with guanidine hydrochloride and urea. These chemicals inactivated the LDH molecule and dissociated it into four polypeptide chains of equal molecular weight (34,000). These chains correspond to about one quarter of the intact enzyme molecule (about 135,000 to 140,000). On the basis of their charge they can be distinguished into two types, later called *A* or *M* (muscle) and *B* or *H* (heart). The two structural genes *a* and *b* may be located on two different operons and are thus under the control of two different operator genes. The separation provides an explanation for the fact that the *a* and *b* structural genes in mammalian tissues rarely seem to produce *A* and *B* subunits at equal rates. Most of the human tissues exhibit preponderance of either isoenzyme 1 or isoenzyme 5, but rarely isoenzyme 3, instead of 1:4:6:4:1 as expected if equal numbers of *A* and *B* subunits were available intracellularly for random recombination in groups of four. The above ratio was accepted by Markert[171] and subsequently by several others[49,182,333]. Markert and Ursprung[174] assumed that two different genes are concerned in the production of *A* and another for *B* subunit. Evidences in support of the hypothesis for LDH structure include the observation that the number of peptides resulting from complete trypsin digestion is approximately one-fourth of the number of arginine plus lysine residues found on total amino acid analysis[9,171]. Immunochemical evidence[45,46,49,173,206,221,222,334] shows that antibody made to LDH-1 will not cross-react with LDH-5 but will cross-react in diminishing amounts with LDH-2, 3 and 4 in quantities

predicted from the proposed structure of the five isoenzymes. Similarly antibodies to LDH-5 will not cross-react with LDH-1, but will react with LDH-2, 3 and 4 to the extent expected[49,183,173,198].

The concept, that *A* and *B* subunits are controlled by two gene loci *a* and *b*, gained additional support from the tetrameric nature of the isoenzyme structure[37,151,199,257]. The subunit structure received the first support for the human beings from Boyer *et al.*[37] In the course of investigation into LDH red cell isoenzyme composition in 50 Whites, 300 American Negroes, 100 Papuans and 200 Nigerians, a variant was discovered in a healthy 25 year old Nigerian male. After starch gel electrophoresis the LDH-1 fraction appeared to be split-up into five, LDH-2 into four, LDH-3 into three and LDH-4 into two. Thus the total number of isozymes in this variant is fifteen. Boyer *et al.*[37] assumed a mutation in the B gene to explain this phenomenon.

Variations of LDH isoenzymes in tissues and organs

LDH isozymes occur in cytoplasm of human and animal cells in concentrations varying from tissue to tissue. Wróblewski[337] observed human tissues to differ greatly in their isozyme compositions as well. Each tissue is characterised by a more or less specific LDH pattern. Such differences, first observed in animal materials, were later confirmed in human tissues[221,222,328,339]. These findings have since then been extended both qualitatively and quantitatively to include different human tissues by different authors[2,27,109,205,218].

In spite of variations in different tissues, the isozyme pattern invariably presents one single maxima which agrees with the tetramer structure of LDH molecule and the two monomers concerned. On the basis of the LDH isozymes they contain, Zondag[358] divided human tissues into three main categories :

Group I—Heart muscle group with maximum LDH-1 activity.
Group II—Liver group with maximum LDH-5 activity.
Group III—General group with maximum LDH-2 and LDH-3 or LDH-4 activity.

Group I consists of heart muscle, kidney, brain and erythrocytes having high percentage of LDH-1 and LDH-2 and low percentage of LDH-4 and LDH-5. Vesell *et al.*[311] and Vesell and Bearn[308,309] have stressed that the erythrocytes are the only cells in the human body in which 4 LDH fractions have been demonstrated. Wieme[329] pointed out that LDH-5 is lacking in

erythrocytes and Vesell *et al.*[311] suggested a correlation between this fact and the absence of nucleus in these cells. The presence of LDH-5 in the haemolysate of red cells having nucleus (reticulocytes) may indicate also the association of LDH-5 with nucleus.

However, the activity of LDH-5 in relation to the total activity is so low that occasionally it may defy detection. But even on the assumption that the ripe erythrocytes do not contain LDH-5 whereas the reticulocytes do, this need not mean that this fraction is either localised in the cell nucleus or is connected with it. Since the loss of nucleus leads ultimately to the cessation of protein synthesis, the absence of LDH-5 in anucleate erythrocytes may be explained by the fact that as it is the most labile isoenzyme, it is lost first. Though Starkweather and Schoch[265], Vesell and Bearn[307], Wieme[328,329] and Zondag[353] could not detect LDH-5 fraction in erythrocytes, yet pfleiderer and Wachsmuth[218,219] located it in certain erythrocytes. It may be possible however that the blood of some of the 9 patients investigated by the latter two authors might have contained many reticulocytes.

The reported absence of LDH-5 in leucocytes[265,307] may be attributed to alternate freezing and thawing of cell suspensions leading to disintegration of leucocytes. Subsequently Vesell *et al.*[311] and Zondag[353] reported all the five fractions from the leucocytes.

A second exception—the adult human testis and sperm contains a sixth LDH isozyme migrating electrophoretically between LDH-3 and LDH-4[52,54]. It is termed as an *x* or *c* band. It appears to be related to spermatogenesis since it is not found in the pre-pubertal testis and forms about 80% of the spermatozoal LDH. LDH-*x* is a tetramer, composed of four identical *x* subunits which differ in amino acid composition from the H and M subunits[349]. The *x* subunit appears to form hybrid isozyme with other subunits, like LDH-1. It has a relatively high affinity for 2-oxobutyrate, is more sensitive to inhibition by oxalate and more resistant to urea[264,382].

Principle underlying separation of LDH isozymes

Following starch or acrylamide gel electrophoresis, on incubation in a buffered substrate (pH 8·1), the enzyme activity becomes visible as five quite distinct, more or less intensely coloured bands. The intensity is a measure of the enzyme activity at that point of the gel, which can be determined densitometrically for each fraction.

$$\text{Lactate/Lactic acid} + \text{NAD} \rightarrow \text{Pyruvate} + \text{Reduced NAD} \quad (1)$$

$$\text{Reduced NAD} + \text{PMS} \rightarrow \text{NAD} + \text{Reduced PMS} \quad (2)$$

$$\text{Reduced PMS} + \text{NBT} \rightarrow \underset{\text{(Colourless)}}{\text{PMS}} + \underset{\text{(purple colour)}}{\text{Reduced NBT}} \quad (3)$$

The enzyme reaction (1) is *via* a second reaction (2) being linked to the actual colour reaction (3). Starting from lactate and NAD, this first results in the formation of pyruvate and $NADH_2$. PMS functions as a hydrogen transporter between the $NADH_2$ that is being formed and the tetrazolium salt NBT. Finally when the latter is reduced, the intensely coloured, water insoluble NBT formazan is produced. The principle of this colour reaction is functional and is based on the demonstration of $NADH_2$ using lactate as a substrate. The reaction is specific for the enzyme capable of reducing lactate with the acid of NAD (formation of $NADH_2$).

Conditions producing alterations in the isozyme pattern

I. *Tissue Culture Condition*

The pattern of the isozymes resulting from recombination of LDH–1 and LDH–5 remains unchanged despite of the addition of various substrates and coenzymes. However, certain alterations in the cellular environment have been effective in causing shifts of the isozyme pattern. The environments of the cell in tissue culture undergoing malignant proliferation and of the cells in the tissues exposed to hormonal stimulation exhibit such alterations.

The cells from a variety of tissues, when grown *in vitro* with Eagle's medium and 10% calf serum, exhibit shifts in isozyme pattern[90,220,309,311]. In culture, under the conditions described by them, all the tissues of a given organism assume a typical uniform pattern, characterised by prominent cathode bands and diminishing or disappearing anodal bands. This observation shows a regularity of the pattern displayed by a variety of human tissues in culture over differing periods. Earlier Tsao[289] had shown that isozymes of several different enzymes existed in cells grown in culture. Alterations in isozyme patterns of cells in culture may be interpreted in the light of the hypothesis for the genetic control of the LDH subunits and its corollary, that tissues maintain different isozyme patterns *in vivo* because of epigenetic conditions leading to relative dominance of either the *a* and *b* genetic loci. In culture, cells from different environments are exposed to identical conditions and they res-

pond in identical fashion by developing a uniform enzyme pattern, suggesting predominance of the *a* over *b* genetic locus. The rapid rate of change in the isozyme pattern is attributed to dedifferentiation or modulation under the simplified cultural condition or outgrowth of a particular cell population favoured by the *in vitro* environment. The latter possibility clearly obviates the necessity of a genetic explanation for the shift in the isozyme pattern.

In culture, LDH–5 predominates, being the isozyme most resistant to anaerobic conditions and high lactate concentrations. By raising the oxygen tension of cells grown in culture Cahn[48] prevented this shift towards LDH–5. Wilson[355] noted the effect of copper on LDH in human endometrium culture leading to a marked inhibition of LDH isozymes containing the 'M' subunit.

Safronova *et al.*[246] observed an increase of the LDH-M polypeptide in embryonic cell lines in comparison with initial tissues. In certain long term cultures, Nitowsky and Soderman[310] have shown retention of the tissue enzyme pattern although most of the other cells in culture exhibited preponderance of LDH-5.

II. *Extraneous factors*

Vesell and Brody[310] observed that factors affecting LDH isozyme pattern include extraction procedure followed, storage of the extract, activity of the material applied to the gel, conditions during electrophoresis and concentrations of the reagents in the incubating medium. Hill[111] found a higher level of LDH activity in the sera of normal children as compared to adults and also its inhibition by mercuric chloride and p-chloromercuribenzoite. Fendley *et al.*[78] have noted a significant difference in *t* and *f* tests on the effect of various dilutants on serum LDH isozymes.

The numerous reports on serum and plasma LDH activity are conflicting. Wróblewski[337] found no significant differences between the two. Bowers[56], Cohen and Larson[68], Kingsley *et al.*[139] observed a higher activity in serum. Still others, like Neely *et al.*[204] found lower LDH activity in serum. Bais *et al.*[11], Kluge and Haug[148] and Rothwell *et al.*[239] attribute these differences in activity to the presence of platelets in the plasma. They advocated centrifugation (3000xg for 15 minutes) for obtaining platelet-free plasma. Bais *et al.*[11] prepared platelet-rich plasma by centrifuging at 1000 xg for 10 minutes and platelet-poor plasma at 3000 xg for 10 minutes and observed a decrease in plasma LDH activity in the latter. LDH activity is found to be inversely proportional to centrifugation speed and to the

number of platelets present[239]. It increases with the concentration of the detergent in the buffer[11], due to lysis of the platelets, releasing the enzyme[116].

III. *Sulphydryl compounds*

Gershbein and Raikoff[91] observed, in presence of penicillamine and other sulphydryl compounds, changes in the respective zymogram with drastic decrease in the total enzyme content. Incubation with penicillamine for one hour at 25° C leads to absence of LDH-5 and alteration of Rf of the remaining isozymes. Schmitt *et al.*[258] have noted inhibition of LDH subunit A with reagents like $AgNO_3$.

IV. *Buffer*

LDH catalyzes the reversible conversion of lactate to pyruvate and NAD^+ to NADH. It is initiated with lactate and the amount of inactivation depends on the buffer used. Both LDH-1 and LDH-5 had higher L→P activity in tris than in sodium pyrophosphate buffer[43]. Higher concentrations of glycine and pyrophosphate were inhibitory with increasing buffer concentrations[44]. Buhl *et al. a*, *b*[43,44] showed that commercial grade of lactic acid led to low LDH activities. Babson and Arndt[10] observed variable results for LDH with commercial NAD^+ preparations. Vesell[302] noted that tissue homogenization, enzyme purification and dilution may alter the isozyme pattern.

V. *Hormonal stimulation and imbalance* play an important role in maintaining the integrity of the enzyme pattern. Richterich *et al.*[234], followed by Dawson *et al.*[66] and Goodfriend and Kaplan[96], have shown that during pregnancy there is a shift in isozyme pattern towards LDH-5. Administration of estradiol at *a* and *b* genetic loci causes differential stimulation of the genes. This conclusion was drawn from the observation that puromycin administration following hormone treatment resulted in a proportionate decline in the concentration of A and B subunits and preferential synthesis of A subunits could be blocked by prior administration of actinomycin D[96]. Testosterone and progesterone, however, resulted in a peripheral increase of both A and B.

VI. *Effect of temperature*

Zondag[353] concluded that LDH activity can be determined best at room temperature. In other temperatures a decline is observed, being far more in low temperatures (– 20 °C).

Quick freezing of the cell suspension, followed by quick thawing frequently used for destruction of cells, entails the risk of loss of activity in the 4th and 5th fractions. Thus Vesell[301] originally found a LDH isozyme pattern in leucocytes showing the highest activity in the middle fraction. The real distribution established for this pattern corresponds to the order LDH-1 < LDH-2 < LDH-3 < LDH-4 < LDH-5[21]. The rate of inactivation of the labile fraction further depends on pH, buffer and concentration of the homogenate[353]. The addition of NAD prevents inactivation, at low temperature, of LDH-4 and LDH-5 which may be explained by the subunit structure of Appella and Markert[9]. Markert and Appella[172,173] suggest that to each monomer a molecule of NAD is attached under certain conditions with A-subunit. These show a greater tendency to become detached than those allied to the B-subunit. Zondag[353] suggested that all A-subunits lose their NAD molecules so that the protection afforded by active sulphydryl groups of these subunits against inactivation is lost. Its exposure to low temperature can actually cause inactivation of monomers deprived of their NAD-molecule. This may serve as an explanation for the changes found in the isozyme pattern. The activity of molecules consisting of more subunits than a single one is evidently proportional to the number of active subunits. On this basis the protective effect of NAD against inactivation is comprehensible. The excess of NAD enables the 4 subunits to bind a new molecule of NAD and restore the balance.

The behaviour of LDH isozyme in serum is different from tissue homogenates. Complete inactivation of LDH-4 and LDH-5 is not observed before keeping at −20°C. Though NAD is entirely lacking in serum, a greater sensitivity to cold inactivation might be expected due to the high protein content of serum which exercises a protective influence on the isozyme molecule. The LDH activity, particularly LDH-5, is better preserved at room temperature than at −20°C or 4°C. Knudsen *et al.*[140] have noted the influence of temperature on LDH isozymes as well.

These evidences indicate that LDH-5 is the least heat stable and LDH-1 the most while other fractions are intermediate in their stability. However, the instabilities of LDH-5 and to certain extent LDH-4 are related to both increase and decrease in temperatures. Juul[126], without however separating the isozymes, found that LDH in plasma kept much better at 4°C for 120 days with citrate as an anticoagulant while it lost considerable activity at −20°C

in one day. Vesell and Yielding[312] confirmed that LDH-5 loses activity both in tissue homogenates and in serum at room temperature and even more so at 37°C, suggesting that stability is pH dependent and all fractions resist heat better at acid pH below 7·0. Zondag[353] and Roman[287], on the other hand, found temperature to have quite opposite effects. Dietz and Lubrano[69,70] observed little difference in stability between storage of serum at 4°C or at 25°C. Homogenates keep much better[13] than sera. Pure LDH-5 can be kept in a saturated ammonium sulphate solution at 4°C for a very long time[287].

Wróblewski and Gregory[338,339] observed LDH-1 and LDH-2 to be much less affected by heat than other fractions. Dubach[73] and Wust *et al.*[343,344] found that 3 LDH classes can be differentiated by preincubation of the sample for 30 minutes. Strandjord *et al.*[269] also distinguished between heat stable and heat labile LDH. Bell[17] and Latner and Skillen[156] separated LDH-1+2 from other fractions also by heat stability at 60°C and called the ratio of heat stable to total LDH the "heat stability index". They proposed it as a diagnostic test. Karl and Peters[129] heated tissue extracts on sera with saline for one hour at 58°C and found no difference in the remaining activity of heat stable muscle. Roman[287], after comparing various methods, found that the most promising is the separation according to heat stability. Heating to two different temperatures (55°C and 60°C) in the presence of buffered substance is most successful for the estimation of LDH fraction in sera for diagnostic purposes. The temperatures proposed by most other authors are however too high and the application of lower temperature will avoid unwanted inactivation.

Swain *et al.*[270] found a gradual decrease of enzymic activity (from LDH-4 to LDH-3 to LDH-2 to LDH-1) with alternate thawing and freezing of the same serum sample for a number of days and even complete loss of LDH-3 and LDH-2. These data confirm the report of Zondag[353].

Inheritance of LDH

Assignment of LDH genes to chromosome

Somatic cell hybrid clones between mouse and human cell lines[104,242,318,319] were examined for human lactic acid dehydrogenase-A (EC 1.1.1.27) and lactic acid dehydrogenase-B (EC 1.1.1.27) by Ruddle and Nichols[248] and Tischfield *et al.*[287] (see Chen *et al.*[52]) by starch and acrylamide gel electro-

phoresis. Human chromosomes were identified by Q[50], or G[278] and C-banding[51] methods.

Klebe *et al.*[146] and Santachiara *et al.*[251] reported that pep-B and LDH-B are syntenic (linked to the same chromosome). The human and mouse forms of LDH-B possess identical electrophoretic mobilities and cannot be resolved quantitatively. In some combinations, the human LDH-B is expressed more strongly and thus can be scored quantitatively. In another series of hybrids the human parent is heterozygous for an electrophoretic variant for LDH-B and thus the human variant form can be distinguished. Chen *et al.*[52], Van Someren *et al.*[295] Weil *et al.*[317], and Shows and McAlpine[260] confirmed the correlation between pep-B and LDH-B, supporting the assignment of the structural genes for these two enzymes to chromosome 12.

The possibility that linkage involves regulatory, rather than structural genes, is not ruled out. The structural pep-B locus may be linked to a gene, the presence of which leads to an increase in mouse LDH-B activity in the hybrids and may also be necessary for human LDH-B expression[251]. Genetic regulation of LDH-B activity has been demonstrated in mouse erythrocyte[261]. Ruddle *et al.*[244] and other authors have suggested that LDH-B and pep-B are asyntenic and have assigned LDH-B to the short arm of chromosome 12[85,135,177,196,282]. Rethoré *et al.*[231], from studies on rearrangement of the short arm of chromosome 12 in three patients, concluded that the LDH-B locus is placed at the middle between 12 p 12·1 and 12 p 12·2. It was later confirmed by Shows and McAlpine.[260]

Tetraploidisation has been regarded as a step in the origin of human karyotype. Morphologic similarities of chromosomes 11 and 12 are consistent with the idea of tetraploidisation, which is also supported by localisation of LDH-B on the short arm of chromosome 12.[181]

During the survey of eleven independently derived human-mouse hybrids, two populations were found to have both mouse and human LDH-A[88]. On being analysed for enzyme and chromosome complements these two populations have shown a constant correlation between the presence of autosome C-11 and the A subunit of LDH, indicating the possible assignment of LDH-A subunit gene to C-11 chromosome.

Francke *et al.*[88] have studied the quantitative expression of LDH-A from the haemolysates of two patients with different but overlapping interstitial deletions of the short arm of the chromosome 11. The gene for LDH-A has

been assigned to chromosome 11[88] and specially to its short arm[81,82,97,98]. Relative deficiency of A or B subunits, due to deletion of respective chromosomal segment, has been detected electrophoretically by quantitative comparison of LDH isozymes, as for example, LDH-B deficiency by Mayeda *et al.*[177]

Francke *et al.*[83] studied the expression of LDH activity in erythrocytes from four unrelated individuals with del 11 p karyotypes. LDH-A activity was normal in three and reduced in one. They concluded that the structural gene for LDH-A is located in that part of 11 p which is deleted in the latter patient. In the former it is within band 11 p 12 (region p 1203→p 1208). A less likely explanation would be heterozygosity for a mutation, causing LDH-A deficiency unrelated to chromosomal deletion. An additional proof of their explanation is that haemoglobin electrophoresis in the above case indicated that she was heterozygous $\beta S/\beta A$, as also her mother. This result excludes the β-globulin gene from region p 11·1→1203 of the short arm of chromosome 11.

Linkage between LDH-A and LDH-B

Clonal differences in the activities of total LDH-B are taken to indicate that the genes coding for A and B subunits of LDH are unlinked[88,116,196]. As electrophoretic mobilities of mouse and man B subunits are identical the conclusion was only tentative. Other possibilities include linkage of LDH-A and LDH-B structural genes and regulation by other unlinked gene(s) of the synthetic rate of mouse and human LDH-B subunits. This suggestion is supported by evidence for presence of genes which control the expression of LDH-B in mouse erythrocyte. Ruddle *et al.*[244] and Shows and Ruddle[261] argued that the problem might be solved using either a mouse or human structural variant of the LDH-B subunits. They have described a human LDH-B variant, LDH-B (var), and its segregation in clonal populations of man/mouse hybrids, which provides strong evidence for the absence of linkage between LDH-A and LDH-B genes, also supported by Santachiara *et al.*[251]. Van Someren *et al.*[296] have claimed that the genes GPT-B and GPT-C controlling GPT (glutamic pyruvic transaminase) are located on two different chromosomes and are linked to LDH-B and LDH-A respectively.

The location of LDH-A on chromosome 11 and LDH-B on chromosome 12 is interesting since the two chromosomes are similar in their overall length, arm ratio and quinacrine banding patterns. LDH-A and B subunits are similar

in their molecular weight and amino acid constitution. This observation is therefore compatible with the hypothesis[54] advocating a polyploid event in the evolution of primate karyotype.

Conjugation of LDH Isozyme with Immunoglobulin

Several workers have reported abnormal electrophoretic patterns of serum LDH isozymes, in which one or more LDH isozyme fraction was absent or electrophoretically abnormal and anomalous LDH bands were present[24, 87, 187, 140, 197, 259]. In some cases abnormalities were caused by the formation of a complex between one or more of the LDH isozymes with immunoglobulins[24, 87, 187, 197, 259]. Biewenga[20,21] in gel filtration experiments observed abnormal LDH fractions to have a higher molecular weight. The possible causes of the formation of LDH-IgA complex are abnormalities in the structure of either LDH or IgA and the linkage could be caused by the presence of a binding agent.

Another factor may be the presence of abnormal structural gene of subunits H and M of LDH isoenzyme[140]. It would lead to a hereditary defect, altering synthesis of one or other of the subunits. Biewenga[21], following immunoelectrophoresis of the serum of a patient with abnormal LDH isoenzymes, concluded that the abnormalities were due to the presence of LDH-IgG complex. The anomaly in the structure of IgG was the probable cause of the complex formation. Biewenga and Feltkamp[22] reported another patient with LDH-IgG_s complex.

Sera with LDH-IgA complex appeared to have more or less pronounced abnormalities in the electrophoretic mobilities of the LDH fractions as compared to normal sera. These sera could be divided into two groups[22] characterised by the absence of LDH-2 fractions and by the presence of only minor abnormalities.

Other Abnormal LDH Isoenzyme

Isoenzyme fractions, additional to the normal five bands, have been reported in the sera of patients with malignancy[254]; with malignant gliomas[40] with carcinoma prostrate[158] and liver diseases[55, 83, 115, 166, 167]. Hepatitis-B antigen[15, 56, 314] laryngeal carcinoma with hyperthyroidism[281] chronic myeloid leukaemia and acute lymphatic leukaemia[279] have also shown variant types. Such extra components appear to be due to the formation in the tumour of an abnormal subunit which subsequently hybridizes with a normal subunit to give a series of isozymes with abnormal electrophoretic mobilities. Clear evidence for such polymorphism of the M subunit was obtained during the study of a patient with liver metastases exhibiting a complete set of isozymes with unusual electrophoretic properties[86]. Lubrano *et al.*[166] have suggested several causes for such bands, including : (i) the presence of a reducing substance in the serum analogous to the "nothing dehydrogenase" ascribed to alcohol dehydrogenase[251], (ii) the presence of bilirubin or bile salts in high concentrations and (iii) abnormal binding of NAD or NADH to LDH, since alteration of LDH isozyme migration has been reported following the addition of NAD or NADH[152] and (iv) binding of LDH to an abnormal protein as reported earlier[24].

An anomalous band was reported by Baxi *et al.*[15] Damle and Talavdkar[56] and Vyas *et al.*[314] in the sera of patients with positive hepatitis-B (Australia HBs Ag) antigen. It was absent in HBs Ag negative sera and disappeared from HBs Ag positive sera when neutralised with HBs Ab (HBs antibody). Apparently this anomalous LDH band was either a part of HBs Ag molecule or intimately associated with it and was termed as LDH-HBs Ag complex.

The extra 'T' band reported by Tanaka *et al.*[281] did not form any complex with immunoglobulin. It has no genetic abnormalities since the LDH isozymes in red cell were normal. It, however, disappeared after removal of cancer tissue from the patient.

Forbes *et al.*[80] have reported duplication of LDH-1 in a patient receiving multiple transfusion. Meaney *et al.*[152] observed a transient variant of serum LDH in a normal individual with only one zone of LDH activity. After five weeks this reverted to the normal pattern and remained unchanged on two widely separated occasions. This variant did not form any complex with immunoglobulins and displayed no hereditary abnormal LDH bands. Even incubation prior to electrophoresis did not change the isozyme pattern. This variant may not be due to electrophoretic interaction but may involve bands of much stronger nature.

LDH and Human Disease

The fact that there are more diseases than just one that can cause an elevated or decreased LDH activity in serum tends to reduce its differential

diagnostic value. Yet the important position that enzyme determination has come to occupy in recent years is mainly due to its striking success in the diagnosis of the diseases of the heart or more particularly myocardial infarction and cancer.

Myocardial infraction

LDH isoenzyme is assumed to behave in a way similar to serum glutamate oxaloacetate transaminase (SGOT), which is increased after myocardial infarction. Wróblewski *et al.*[341] found an increased serum LDH level in experimentally induced myocardial infarction in dogs as well as in patients suffering from the disease. Subsequently Wróblewski[337] found a correlation between increase in serum LDH and size of the infarct, later confirmed by Ruegsegger *et al.*[245]. They further established that a temporary reduction in the blood supply to the cardiac muscle (experimental coronary insufficiency) was not associated with a rise of LDH in the serum as long as the ischaemia did not result in histologically observable necrosis, later supported by others as well[270,348]. White[324] arrived at the conclusion that after myocardial infarction serum levels of LDH remain elevated for periods considerably longer than other enzymes. Rodney and Shapiro[236], King and Waind[138] and Hazelhorst and Fockert[101] applied serum LDH determination to establish or reject unequivocally the presence of acute myocardial infarction and necrosis.

The elevated level may be caused by pathological conditions or lesion, either in the heart muscle or in the liver tissue or a combination of both. The diagnosis of myocardial infarction turns on the fractions of LDH-1 and LDH-2. Vesell and Bearn[305] demonstrated an elevated level of these fractions in two patients with myocardial infarction, subsequently confirmed by others[12,30,72,73,103,107,108,109,165,268,269,270,328,329,338,339,340,344,345,346,355] Zondag[358] later showed decisively that LDH isoenzyme might be a valuable tool in the diagnosis of myocardial infarction. Kurashvili and Dzhikiya[154] and Warburton *et al.*[316] have noted an increase in the activity of LDH-1 and a decrease in LDH-2, LDH-3 and LDH-4.

Swain *et al.*[277] have noted a rise both LDH-1 and LDH-2 fractions in patients suffering from cardiac diseases. In one case of myocardial infarction there was a 6th band with rise in LDH-1 and LDH-2 fractions.

LDH and Cancer

The extensive clinical literature dealing with serum and plasma LDH activity gives a confusing picture in malignancy, including leukaemia. Hill[110,112] and Hill and Levi[118] reported a rise in LDH activity in neoplastic diseases, subsequently confirmed by several authors[7,68,76,89,122,123,157,158,160]. Others also noted a rise in the LDH-5 fraction in different types of cancer[79,123,183,185].

Zimmerman and Weinstein[348] found even the normal value to be raised. Zondag[351] and Zondag and Klein[354] noted an elevation of LDH-3 whereas Goldman *et al.*[95] found constant shift to LDH-5. Zuppinger *et al.*[355] observed no constant pattern while Garbe *et al.*[88], noted LDH-5 : LDH-1 ratio to be increased.

Lührs and Negelein[168], Tamura and Oda[280] and Denis and Prout[68] confirmed that serum LDH is known to rise following extension of malignant neoplasm. Hsieh *et al.*[120] found an elevated value in cases where malignant nature of the disease is established. They also reported normal values in carcinoma bowel, stomach, kidney and increased values in leukaemia, malignant lymphoma and liver metastasis. The level of LDH and the nature of tissue involved in the malignant growth, as indicated by the published data are frequently contradictory[119,229,321,342].

It seems plausible that the degree of malignancy as well as its size plays a considerable role in elevation of LDH activity. Brindley and Francis[38] succeeded in establishing a correlation between the measurement of tumour mass and LDH activity. It however does not hold good for patients in the final stages of the disease.

The special significance of the presence of liver metastasis is often admitted[238,321,342]. Evidently metastases is not an essential pre-requisite for the elevation of LDH activity, though it considerably enhances the possibilities. However, solitary tumours, without secondary growth, may induce elevated LDH activity and conversely a metastasised carcinoma may be concordant with normal LDH activity[358].

Regarding the distribution of LDH isozyme in the malignant diseases, reported a deviating isozyme pattern in the form of an increased second fraction in leukaemia. Zondag[352] confirmed it from an observation of the increasing activity of LDH-1 and LDH-2 in different types of malignant

diseases. Swain *et al.* (unpublished) also noted increase in LDH-1 and LDH-2 fractions in different types of carcinoma.

After the report of high serum LDH activity in acute leukaemia and chronic myeloid leukaemia, several studies on the subject have been published[19, 80, 119, 121, 147, 176, 279, 313, 322]. Khmelev *et al.*[136] have shown LDH-1 was greatly increased while LDH-3 and LDH-4 were present in reduced amounts. Neely *et al.*[204] noted an elevated level of LDH activity in sickle cell anaemia. Wro'blewski and Gregory[336] showed a marked preponderance of the fast moving isozymes (LDH-1 and LDH-2) with LDH-1 activity greater than LDH-2 in pernicious anaemia. Levi *et al.*[168], on the other hand, suggested that serum LDH level has no significance in the diagnosis of leukaemia.

Froewlich[85] recorded an unknown LDH distribution pattern in advanced stages of leukaemia and in 6 cases with malignant tumour, named as type *c*. Surinov and Menzhinskaya[275] have worked on the isozymes of LDH, esterase and alkaline phosphatase in some tumours which did not change from normal even after chemotherapy and radiation. Wilson *et al.*[335] have observed that treatment with vitamin B_{12} not only reduced the greater activity but also brought about a reversal of LDH of the LDH-1/LDH-2 isoenzyme distribution in megaloblastic anaemia. Stott[267] compared the findings in the red cells in 27 cases of megaloblastic anaemia with those of normal and found similar abnormal isoenzyme patterns in the serum and red cells of the patients. With iron deficiency anaemia or haemolytic anaemia the abnormal isoenzyme pattern is not apparent. Zondag[353] also confirmed the presence of such abnormal patterns in megaloblastic anaemia and haemolytic anaemia.

Plasma LDH

A number of studies deal with the status of plasma LDH isoenzymes in various malignant and non-malignant conditions[288, 352, 353, 354]. No clear-cut pattern emerges from these observations specific to malignancy alone. However, Wieme *et al.*[331] have claimed that plasma LDH-3 was of value during the therapy of bronchogenic carcinoma. Van Gennip *et al.*[294], from their studies on serum LDH isoenzyme distribution of children with malignant diseases, found a significantly decreased LDH-3 isoenzyme fraction in many patients. The LDH-4 and LDH-5 were also reduced, which they have explained through inhibition of the synthesis of LDH in the tumour cell by prednisone. Inhibition of the M-monomer is less probable, since LDH does not decrease at all in all patients receiving prednisone.

Massey *et al.*[176] reported that plasma LDH-5 was useful as a biochemical parameter during the treatment of adeno-carcinoma of pancreas. Damle and Talavadekar[55], in the isoenzyme composition of sera and effusions from cases of malignant conditions, cirrhosis of liver and pulmonary tuberculosis, observed significant increases in LDH-2, 3 and 4 fractions. Talageri[278] and Talageri *et al.*[279] reported a rise in total LDH activity and isoenzyme fractions in oral cancer, osteosarcoma, chronic myeloid leukaemia and acute lymphatic leukaemia. The possible mechanisms implicated in the elevation of the enzyme include (i) over production of the enzyme by the tumour ; (ii) tumour blockage of the duct system through which the enzyme passes, (iii) changes in the permeability of cell allowing leakage of soluble enzyme into circulation, (iv) status of kidney function where homeostasis mechanism involves urinary excretion of the enzyme, (v) possible presence of abnormal conditions of endogenous and exogenous inhibitors or activators [254, 5, 5]. Talageri[278] found that only 62% of oral cancer cases showed increased values. The tentative results of these anomalies are (i) differing concentration of activities at various sites of oral cavities from which tumours have arisen ; (ii) different rates of release of the enzyme in terms of passage through the cell membrane and (iii) different fates of the liberated enzyme in the serum, primarily in relation to activation or removal. He also reported an increase in LDH-1 and LDH-2 in myocardial infarction, angina pectoris and ischemia during trachycardia, haemolytic anaemia and malignancy of the gonads, such as, seminoma tests and dysgerminoma of the ovary and increase in LDH-5 fraction in liver injury and other diseases[352]. Masry *et al.*[175] noted a sharp difference between the different fractions in haemolytic diseases.

Other diseases

The great variety of disorders that may lead to an increase in serum LDH might well result in an endless series of erroneous results. The disorders such as angina pectoris, congestive heart failure, paroxysmal tachycardia, pericarditis and pulmonary embolism are usually not attuned with an elevated level of serum LDH as long as there is no question of lesion of heart muscle or other tissues. Wacker *et al* [315] found normal LDH level in angina pectoris in contrast to the elevation recorded by Rodney and Shapiro[286].

Hess and Gehm[106] reported high activity of serum LDH in patients suffering from pernicious anaemia although moderately raised serum LDH level is seen in anaemia, haemorrhage and iron deficiency. Anderson[8] found

occasionally slightly raised serum LDH levels in anaemias. Amelung[2] suggested that appropriate treatment of megaloblastic anaemias with either B_{12} or folic acid produces a dramatic reduction of serum LDH activity. Khemelev *et al.*[136] have worked out the LDH isozyme pattern in leukaemia and Helbig *et al.*[102] in patients with leukaemic lymphocytes. Schmidt *et al.*[252] have worked on different isozymes in diseases of liver and the percentage liver LDH (LDH-5) was in agreement with other workers.

In evaluating the serum LDH activity of patients with malignant diseases the therapy, if any, may be taken into consideration. In patients receiving various therapies like X-ray[98,114,221] chemotherapy[136,221] and surgical treatment[98,114] the initial elevated LDH activity may become normal.

Distribution of LDH in different populations

The enzyme is a tetramer consisting of four polypeptide chains arising through combinations by two genetic loci. All probable combinations of 4 polypeptide chains of two kinds produce five isozymes. Any mutation at either A or B subunit may give rise to a variant. The variants have either a phenotype with normal B with an A variant which may be slow or fast or a phenotype with normal A and B variant which may be again of slow or fast. There are more than 14 different types of variant, found from different populations of the world. They include LDH-Calcutta-1, LDH-Calcutta-2, LDH-Delhi-1, LDH-Madras-1, LDH-Memphes-1, LDH-Memphes-2, LDH-Memphes-3, LDH-Memphes-4, LDH-Nagasaki-1, LDH-Kuala Lumpur-1, LDH-Kuala Lumpur-2, LDH-Kuala Lumpur-3, LDH-Chinese-1, LDH-Chinese-2, LDH-Sugano, LDH-Bulgaria, LDH-Malaya-1. Lie-Injo *et al.*[165] suggested that apparently their LDH-Kuala Lumpur-1 was the same as LDH-Chinese-1 of Blake *et al.*[28]. The similarities between the other forms of these two phenotypes suggest that possibly there may be some overlapping.

In most of the populations studied, LDH is monomorphic except the Luo population of Nayanda in Kenya[105] and LDH-Calcutta-1 in India which is polymorphic. The rate of incidence of total LDH variants found in the human samples outside India is little less than 1%. The overall frequency of LDH variants in India is the highest in the world[63,240,241]. In Europeans it is sporadic but is more frequent in Negroes.

After comparing the variants found in India with those observed elsewhere it may be summarised that (i) LDH-Mem-1 and LDH-Mem-2, LDH-Cal-1, LDH-Cal-2 are faster 'A' subunit variants, of which LDH-Mem-2 is still faster than LDH-Mem-1. LDH-Cal-2 is faster than LDH-Cal-1. (ii) LDH-Mem-4 is the only slow 'A' subunit variant without any parallel in India. (iii) LDH-Mad-1, LDH-Del-1 and LDH-Mem-3 and also Bulgarian variant are all slower 'B' subunit variants. It is likely that LDH-Mad-1 is slightly faster than LDH-Del-1.

The geographical distribution of different variant patterns is given in Table 1.

Isoenzyme variants as markers of population movement in man :

The first genetic variant of LDH in man was discovered in an itinerant Yoruba peddler in West Africa [87] and varieties of both *A* and *B* loci were described by a number of workers, the highest combined frequency approaching 1 percent, being in Black Americans[151]. Variants in most population are reported sporadically. For example, in Caucasians tests of over 4,000 randomly sample individuals in United States, England, Iceland and Sweden only seven persons with B variant were detected[303,304,188]. A exception to this low frequency was the discovery of six persons with an *A* subunit variation amongst 406 members of the Muripan clans of the Enga in the Western High lands of Papua, New Guinea[97]. The situation in India is very different. The first report of Das *et al.*[63,64] listed 10 'Calcutta-1' variants in 614 Bengalies living in Calcutta, a variant frequency of 1·63 percent. Since then Das *et al.*[65] have tested many parts of India.

In a few of the series no variant were detected, but in the great majority 'Calcutta-1' was present with a frequency ranging from under 1% to a maximum of 4% Das *et al.*[61] have suggested that persons from lower caste had a relatively higher frequency of 'Calcutta-1'. Kirk[142] has postulated that the gene LDH Cal-1 has arisen as a mutant allele in tribal populations, predating the Aryan invasions and that the incorporation of tribal elements into Hindu Society had introduced the gene into the lower castes. Gradually it had permeated even into high caste Brahmin groups. Such a pattern of gene flow could be contrary of course to what happened with another specific marker HbS. It occurs in very high frequencies in many tribal populations in India, but only rarely in Hindu populations[1,247]. In some populations like Kaora in Bengal, probably derived from tribal people, LDH-Cal-1 is present but not

Table 1 : WORLD DISTRIBUTION OF IDH VARIANTS

Place	Population	Authors	No. of samples	Normal	Phenotypes with Normal 'B' and an 'A' Variant										Not tested	Phenotypes with Normal 'A' and a 'B' Variant							No. of variants	All variants %
					Cal-1 Fast	Cal-2 Fast	Mem-1 Fast	Mem-2 Fast	Mem-4 Slow	Kua-1 Fast	Kua-2 Slow	Chi-1	Chi-2	Mal-1		Mem-3 Slow	Mad-1 Slow	Del-1 Slow	Bull Slow	Naga-1 Fast	Sugano Slow	Kua-3 Slow		
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
ASIA																								
India																								
West Bengal	Bengalis (Hindu)	Das *et al.* 1970a	524	516	8																		8	1·53
	Non-Hindu	Das 1970b	90	88	2																		2	2·22
	Bengalis (Hindu Upper Caste)	Das 1972	493	491	2																		2	0·41
	Bengalis (Hindu Lower Caste)	Das *et al.* 1972a	937	920	17																		17	1·81
	Bengalis (Calcutta)	Das *et al.* 1970	269	264	5																		5	1·86
	Bengalis (Mahishyas)	Mukherjee *et al.* 1973	100	99	1																		1	1·00
	Bengali (Muslims)	Mukherjee *et al.* 1974	123	119	3	1																	4	3·24
	Mahishyas	do	100	99	1																		1	1·00
Bihar	Oran (Tribe)	do	131	128	3																		3	2·24
	Munda (Christian)	Mukherjee *et al.* 1973	230	226	4																		4	1·60
	Munda (Non-Christian)	do	142	140	2																		2	1·40
		Mukherjee *et al.* 1975	131	128	3																		3	2·24
Andhra Pradesh	Savara	Rao *et al.* 1978a	62	62																			0	0·0
	Jatapu	do 1978b	157	157																			0	0·0
	Kolams	Ramesh *et al.* 1979	218	218																			0	0·0
Tamil Nadu	Tamils (Hindu)	Ananthakrishnan *et al.* 1970	696	691	5												2						7	1·01
	Tamils (Non-Hindu)	do	21	21																			0	0·0
	Tamils (Brahmin Caste)	do	323	320	3																		3	0·93
	Tamils (Naikar Caste)	Ananthakrishnan *et al.* 1970	168	163	5																		5	2·98
	Tamils (Nadar)	do	145	139	6																		6	4·14
	Tamils (Reddiar Caste)	do	148	148																			0	0·0
	Tamils (Other Caste)	do	365	352	11												2						13	4·07
Kerala	Kadar	Saha *et al.* 1974	213	206	7																		7	3·29
	Nayar	Saha *et al.* 1976	142	139	3																		3	2·1
	Izhavan	do	68	66	2																		2	2·1
	Muslims	do	125	124	1																		1	0·8
	Irula	do	175	174	1																		1	0·6
	Kurumba	do	43	40	3																		3	6·9
Maharashtra	Dange	Das *et al.* 1973	68	67	1																		1	1·47
	Halmat	do	11	11																			0	0·0
	Hande	do	98	97	1																		1	1·02
	Hatekar	do	239	235	4																		4	1·67
	Hatikanan	do	28	28																			0	0·0
	Kutekar	—	2	2																			0	0·0
	Mandhe	—	169	163	3														3				6	3·55
	Sanger	—	60	60																			0	0·0
	Sheger	—	100	98	2																		2	2·00

Table 1 (Contd.)

(2)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Maharashtra (Contd.)	Unikanan	—	48	48																			0	0·0
	Zande	—	153	148	5																		5	3·26
	Marathas	Sharma and Baxi 1978	96	93	3																		3	3·1
	Other Hindus	do	88	88																			0	0·0
	Muslims	—	27	27																			0	0·0
	Christians	—	9	9																			0	0·0
	Marathis	Blake *et al.* 1976	504	495	9																		9	1·78
	Gujratis	Blake *et al.* 1970	501	496	5																		5	0·998
	Parsis	Undevia *et al.* 1972	418	403	15																		15	3·58
	Iranis	—	48	48																			0	0·0
	Vania Soni	Undevia *et al.* 1978	267	266	1																		1	0·374
	FROM PLACENTA																							
	Nava Budh	Mukherjee *et al.* 1978	144	139	3	2																	5	3·47
	Brahmins	—	49	48		1																	1	2·04
	Marathies	—	209	208		1																	1	0·48
	Low Castes	—	41	41																			0	0·0
	Middle Class	—	47	46		1																	1	2·13
	Muslims	—	55	53	2																		2	3·64
	Others	—	99	89	1																		1	1·11
Madhya Pradesh	Hindu	Papiha and Chhaparwal 1973	176	176																			0	0·0
	Muslims	do	163	161	2																		2	1·22
Delhi	Arora (Hindu Caste)	Blake *et al.* 1971	71	71																			0	0·0
	Khatri	do	78	77	1																		1	1·28
	Rajput	do	41	40	1																		1	2·44
	Brahmin	do	61	60	1																		1	1·64
	Hindu Sch. Caste	do	85	83	2																		2	2·35
	Hindu Miscellaneous	do	32	32																			0	0·0
	Non-Hindu ,,	do	61	60	1																		1	1·61
	Vaish (Hindu)	do	63	62	1																		1	1·59
	Rajput	Das *et al.* 1974 a, b	95	94	1																		1	1·05
	Jat	do	104	103	1																		1	0·96
	Gujar	do	100	99	1																		1	1·00
	Ahir	do	101	99	1													1					2	0·99
	Chamar	do	100	98	2																		2	2·0
	Punjabi	do	182	182																			0	0·0
Harayana	Muslim	Papiha *et al.* 1976	not stated		1																		not stated	
	Harayanvis (Jat)	do	not stated																				not stated	
Punjab	Punjabis	Papiha *et al.* 1972	140	140																			0	0·0
	Brahmin	Singh *et al.* 1974	106	106																			0	0·0
	Arora	Singh *et al.* 1974	103	103																			0	0·0
	Khatri	do	132	132																			0	0·0
	Jat	do	159	159																			0	0·0
Nepal	Nepalese	Sunderland *et al.* 1979	207	207																			0	0·0
Bangladesh	Bengali Muslim	Papiha *et al.* 1975	200	199	1																		1	0·5
Japan	Japanese from Hiroshima	Ueda *et al.* 1977	2652																				1	0·24
	Japanese from Nagasaki	do	1376	4027																1				
	Japanese	Neel *et al.* 1978	4028	4027																1			1	0·24
	Japanese	Sugano *et al.* 1973																			1		1	

Table 1 (Contd.)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
Malaysia	Chinese	Lie-Injo *et al.* 1973	458	457						1													1	0·21
	Malays	do	471	469	2																		2	0·42
	Indians	do	906	892	10												4						14	1·03
Malaysia (cord blood)	Chinest	do	557	554						1	1											1	3	1·54
	Malays	—	559	557						1														0·53
	Indians	—	283	279	2						1						1						2	0·357
																	1						4	1·41
Singapore	Malays	Blake *et al.* 1973 a, b	259	259																				
		do	378	377								1												0·26
Sumatra	Batak	Lie-Injo *et al.* 1974																						
	Minangkabau	—																						
	Atjehenese	—	502	502																				
	Chinese	—																					0	0·0
	Other Indonesians	—																						
Philippines	Nigritos	Omoto *et al.* 1978	129	129																				
	Tagalog	—	253	252																			0	0·00
	Visayan	—	125	125												1							1	0·39
	Ifugao	—	97	97																			0	0·0
																							0	0·0
Iran	Turkie	Farhud *et al.* 1973																						
	Turkoman	Kirk *et al.* 1975			1																			0
	Caspian Littoral	Kirk *et al.* 1977	463	462	1																		1	0·21
Turkey	Cypriots	Davidson *et al.* 1965	245	243																			2	0·816
OCEANIA																								
New Guinea	Melanesia (English speaking people)	Sinnett *et al.* 1970	416	410					6														6	1·44
	Melanesia (English speaking people)	Boyer *et al.* 1963	100	100																			0	0·0
	do	Blake *et al.* 1969	408	402					6														6	1·47
	Manus Islanders	Malcolm *et al.* 1972	183	183																			0	0·0
New Zealand	Samoans	Booth *et al.* 1977	77	77																			0	0·0
	Non-Samoans	—	24	24																			0	0·0
West Indies	Pestle Mainland	Basu *et al.* 1976	55	55																			0	0·0
Western	Feuralep	Neel *et al.* 1976	744	738	5																		5	0·67
Carolina	Attol	Blake *et al.* 1973	382	381	1																		1	0·26
Micronesia																								
Marina	Micronesia	Tashian 1965	238	238																			0	0·0
Island		Neel *et al.* 1976	372	368	4																		4	1·07
AUSTRALIA	Malag	Kirk *et al.* 1969	595	595																			0	0·0
	Australian aborigine	Kirk *et al.* 1971	1500	1500																			0	0·0
	Fuyuge speakers	Woodfield *et al.* 1974	173	173																			0	0·0

Table 1 (Contd.)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
EUROPE																								
Germany																								
Mainz	Psoriatic patients	Ananthakrishnan *et al.* 1974	120	120																			0	0·0
	Germans	—	3000	3000																			0	0·0
Ireland	Irish	Tills 1977	1800	1797											3 slow								3	0·16
Iceland	Icelanders	Nilsson and Erikson 1972	1000	996					4														4	0·16
	Icelanders	Mourant *et al.* 1978	1000	999											1								1	0·1
Sweden	Psoriatic patient	Beckman *et al.* 1977 Mourant *et al.* 1968	1000	1000																			0	0·0
	Swedish	Nilsson and Erikson 1972	4171	4171																			0	0·0
Finland	Finnish	do	1767	1767					2														0	0·113
	Finland Swedish	—	873	873																			0	0·0
	Laps	—	1776	1776																			0	0·0
Greenland	Maris	Nilsson and Erikson 1972	317	317																			0	0·0
	Eskimos	—	153	153																			0	0·0
Bulgaria	Bulgarians	Ananthakrishnan *et al.* 1972a	138	137															1				1	0 72
		Ananthakrishnan *et al.* 1972b	138	137															1				1	0·72
Britain		Davidson *et al.* 1965	1015	1013									2 A variant										2	0·19
	Whites (New York)	Vesell and Bearn 1962	80	80																			0	0·0
	Negroes (New York)	do	280	280																			0	0·0
	Whites	Boyer *et al.* 1963	50	50																			0	0·0
	American Negroes	do	300	300																			0	0·0
	Negroes	Tashian 1965	95	94											1								1	1·05
	Indians (USA)	—	284	284																			0	0·0
	Micronesians (Marina Islanders)	—	238	238																			0	0·0
	Whites	—	28	28																			0	0·0
Memphis & Tennessee	Negroes	Kraus and Neely 1964	610	603			3	2								2							7	1·15
	Whites	Vesell 1965	600	599											1 Fast								1	0·16
	Whites	Kraus and Neely 1964	330	329					1														1	0·30
	Negro patients (Wash. DC)	Vesell 1965	600	597			3																3	0·50
SOUTH AMERICA	Xavante	Tashian *et al.* 1967	84	84																			0	0·0
		Neel 1964	74	74																			0	0·0
	Baniwa	Mohrenweiss 1979	377	377																			0	0·0
MIDDLE AMERICA																								
Dominica	Carib	Harvey *et al.* 1964	126	126																			0	0·0
SOUTH AFRICA	Bantu	Kirk *et al.* 1971b	304	304																			0	0·0
	Nubians	Bertin *et al.* 1978	154	154																			0	0·0

in higher frequency than in many other Indian populations. By contrast one of the highest frequencies has been observed among Parsis in Bombay[201].

In another group of migrants of Persia, the Iranis, who arrived within the past 200 years Undevia *et al.*[203] have sampled a mere 48 cases without finding a single 'Calcutta-1' variant. However, in testing samples in Caspian littoral Kirk[142] found only one example of 'Calcutta-1' in more than one thousand persons tested. It suggests that the LDH Cal-1 gene may be present outside India in Middle-East but the search for its centre of distribution has so far not been completely successful.

Acknowledgement—The authors are thankful to Professor A. K. Sharma, Professor and Head of the Department of Botany for the facilities provided to carry out the work and to the University Grants Commission, New Delhi for the award of teacher fellowship to B. K. S.

REFERENCES

1. Ajmani, M., Sharma, A. and Talukder, G. (1978) *Acta Genet. Med. Gemellol.*, 27 : 11.
2. Amelung, D. (1961) *Dtsch. Med. Wochschr.*, 86 : 171.
3. Ananthakrishnan, R., Blake, N.M., Kirk, R.L. and Baxi, A.J. (1970) *Med. J. Australia*, 2 : 787.
4. Ananthakrishnan, R., Tsacheva, L. and Walter, H. (1972a) *Hum. Hered.*, 22 : 271.
5. Ananthakrishnan, R., Walter, H. and Tsacheva, L. (1972b) *Humangenetik*, 15 : 186.
6. Ananthakrishnan, R., Walter, H., Hoede, N., Morsches, B. and Holzmann, H. (1974) *Hum. Hered.*, 24 : 53.
7. Ananthanarayana, P.H. and Ramakrishnan, S. (1978) *Ind. J. Med. Res.*, 68 : 455.
8. Anderson, N. (1964) *Scand. J. Haematol.* 1 : 212.
9. Appella, E. and Market, C.L. (1961) *Biochim. Biophys. Res. Commun.* 6 : 171.
10. Babson, A.L. and Arndt, E.G. (1970) *Clin. Chem.*, 16 : 254.
11. Bais, R., Prior, M.P. and Edwards, J.B. (1977) *Clin. Chem.*, 22 : 1056.
12. Barnett, H. (1962) *Lancet*, 2 : 199.
13. Barnett, H. (1964) *J. Clin. Pathol.*, 17 : 567.
14. Basu, A., Namboodiri, K.K., Weitkamp, L.R., Brown, W.H., Pollitzer, W.S. and Spivey, W.A. (1976) *Hum. Biol.*, 48 : 245.

15. Baxi, A.J., Bapat, J.P., Damle, S.R., Talavadekar, R.V., Rajpal, R.M. and Dave, J.K. (1975) *Vox Sang.*, 31 : 70.
16. Beckman, G., Beckman, L., Cedergren, B. and Liden, S. (1977) *Hum. Hered.*, 27 : 405.
17. Bell, R.E. (1963) *Am. J. Clin. Pathol.*, 40 : 216.
18. Bertin, T., Harris, J.E., Ferrell, R.E. and Schull, W.J. (1978) *Hum. Hered.*, 28 : 66.
19. Bierman, H.R.,Hill, B.R., Reinhardt, L. and Emory, E. (1957) *Cancer Res.*, 17 : 660.
20. Biewenga, J. (1972) *Clin. Chim. Acta*, 40 : 407.
21. Biewenga, J. (1973) *Clin. Chim. Acta*, 47 : 139.
22. Biewenga, J. and Feltkamp, T.E.W. (1975a) *Clin. Chim. Acta*, 58 : 239.
23. Biewenga, J. and Feltkamp, T.E.W. (1975b) *Clin. Chim. Acta*, 64 : 101.
24. Biewenga, J. and Thijs, L.G. (1970) *Clin. Chim. Acta*, 27 : 293.
25. Blake, N.M., Kirk, R.L. and Baxi, A.J. (1970) *Hum. Hered.*, 20 : 409.
26. Blake, N.M., Kirk, R.L., McDermid, E.M., Omoto, K. and Ahuja, Y.R. (1971) *Hum. Hered.*, 21 : 440.
27. Blake, N.M., Kirk, R.L., Pryke, E. and Sinnett, P. (1969) *Science*, 163 : 701.
28. Blake, N.M., McDermid, E.M., Kirk, R.L., Ong, Y.W. and Simsons, M.J. (1973a) *Singapore Med. J.*, 14 : 2.
29. Blake, N.M., Omoto, K., Kirk, R.L. and Gajdusek, D.C. (1973b) *Am. J. Hum. Genet.*, 25 : 413.
30. Blanchaer, M.C. (1961) *Clin. Chim. Acta*, 6 : 272.
31. Blanchaer, M.C., Green, P.T., Maclean, J.P. and Hollenberg, M.J. (1958) *Blood*, 13 : 245.
32. Blanco, A. and Zinkhan, W.H. (1963) *Science*, 139 : 601.
33. Boone, C., Chen, T.R. and Ruddle, F.H. (1972) *Proc. Nat. Acad. Sci. (USA)*, 69 : 510.
34. Booth, P.B., Taoagali, J.L., Kirk, R.L. and Blake, N.M. (1977) *Hum. Hered.*, 27 : 412.
35. Bootsma, D. and Ruddle, F.H. (1978) *Cytogenet. Cell. Genet.*, 22 : 74.
36. Bowers, G.N. (1963) Lactic dehydrogenase : *In. Standard methods of clinical chemistry* 4, *ed.* D. Seligson, Academic Press, New York, p. 103.
37. Boyer, S.H., Fainer, D.C. and Watson-Williams, E.J. (1963) *Science*, 141 : 642.
38. Brindley, C.O. and Francis, F.L. (1963) *Cancer Res.*, 23 : 112.
39. Brody, I.A. (1964) *Nature*, 201 : 685.
40. Buckell, M. and Barnes, G.K. (1968) *Brit. J. Cancer*, 22 : 237.
41. Buhl, S.N., Copeland, W.H., Fasce, C.F. Jr., Kowalski, P., Richards, A.H. and Vanderlinde, R.E. (1977a) *Clin. Chem.*, 23 : 1000.
42. Buhl, S.N., Jackson, K.Y., Lubiniski, R. and Vanderlinde, R.E. (1977b) *Clin. Chem.*, 23 : 1289.
43. Buhl, S.N., Jackson, K.Y., Lubiniski, R. and Vanderlinde, R.E. (1976a) *Clin. Chem.*, 22 : 1098.

44. Buhl, S.N., Jackson, K.Y., Lubiniski, R. and Vanderlinde, R.E. (1976b) *Clin Chem.*, **22** : 1872.
45. Burd, J.F. and Usategui-Gomez, M. (1973) *Biochim. Biophys. Acta*, **310** : 238.
46. Burd, J.F., Usategui-Gomez, M., Fernandez de Castro, A., Mathatre, N.S. and Yeager, F.M. (1973) *Clin. Chem. Acta*, **46** : 205.
47. Cabaud, P.G. and Wróblewski, F. (1958) *Am. J. Clin. Pathol.*, **30** : 234.
48. Cahn, R.D. (1964) *Develop. Biol.*, **9** : 327.
49. Cahn, R.D., Kaplan, N.O., Levine, L. and Zwilling, E. (1962) *Science*, **136** : 962.
50. Caspersson, T., Lomakka, G. and Zech, L. (1971) *Hereditas*, **67** : 89.
51. Chen, T.R. and Ruddle, F.H. (1971) *Chromosoma*, **34** : 51.
52. Chen, T.R., McMoris, F.A., Creagan, R., Rucciuti, F., Tischfield, J. and Ruddle, F.H. (1973) *Am. J. Hum. Genet.*, **25** : 200.
53. Cohen, L. and Larson, L. (1966) *New Engl. J. Med.*, **273** : 465.
54. Comings, D.E. (1972) *Nature New Biol.*, **238** : 455.
55. Damle, S.R. and Talavadekar, R.V. (1974) *Ind. J. Cancer*, **11** : 419.
56. Damle, S.R., Talavadekar, R.V., Rajpal, R.M., Bapat, J.P. and Baxi, A.J. (1975) *Ind. J. Med. Res.*, **63** : 1363.
57. Das, S.K., Mukherjee, B.N., Malhotra, K.C. and Majumdar, P.P. (1978) *Ann. Hum. Biol.*, **5** : 25.
58. Das, S.R., Malhotra, K.C., Mukherjee, B.N. and Das, S.K. (1973) *Jap. J. Hum., Genet.*, **18** : 305.
59. Das, S.R., Malhotra, K.C., Mukherjee, B.N. and Das, S.K. (1974a) *India. Jap. J. Hum. Genet.*, **18** : 401.
60. Das, S.R., Mukherjee, B.N., Das, S.K., Roy, M. and Chhatui, S.S. (1974b) *Hum. Hered.*, **24** : 24.
61. Das, S.R., Mukherjee, B.N. and Das, S.K. (1972a) *Humangenetik*, **14** : 151.
62. Das, S.R., Mukherjee, B.N., Das, S.K. and Malhotra, K.C. (1972b) *Hum. Hered.*, **22** : 264.
63. Das, S.R., Mukherjee, B.N., Das, S.K., Ananthakrishnan, R., Blake, N.M. and Kirk, R.L. (1970a) *Humangenetik*, **9** : 107.
64. Das, S.R., Mukherjee, B.N., Das, S.K., Blake, N.M. and Kirk, R.L. (1970b) *Indian J. Med. Res.*, **58** : 866.
65. Davidson, R.G., Fildes, R.A., Glen-Bott, A.M., Harris, H. and Robson, E.A. (1965) *Ann. Hum. Genet.*, **29** : 5.
66. Dawson, D.M., Goodfriend, T.L. and Kaplan, W.D. (1964) *Science*, **143** : 929.
67. Deal, W.C., Rutter, W.J., Massey, V., Van Holde, K.E. (1963) *Biochem Biophys Res. Commun.*, **10** : 49.
68. Denis, L.J. and Prout, G.R. Jr. (1963) *Invest Urol.*, **1** : 102.
69. Dietz, A.A. and Lubrano, T. (1967) *Anal. Biochem.*, **20** : 246.
70. Dietz, A.A., Lubrano, T. and Rubinstein, H.M. (1970) *Clin. Chim. Acta*, **27** : 225.
71. Dioguardi, N., Agnostone, A., Fiorelli, G. and Lumanto, B. (1963) *J. Lab. Clin. Med.*, **61** : 713.

72. Dubach, U.C. (1961) *Helv. Med. Acta*, **28** : 469.
73. Dubach, U.C. (1962) *Schweiz. Med. Ws.*, **92** : 1436.
74. El-Hassan, A.M., Godber, M.G., Kope'c, A.C., Mourant, A.E., Tills, D. and Lehman, H. (1968) *Man*, **3** : 272.
75. Emerson, P.M., Wilkinson, J.H. and Withycombe, W.A. (1964) *Nature*, **202** : 1337.
76. Erikson, R.J. (1961) *J. Am. Med. Assoc.*, **176** : 794.
77. Farhud, D.D., Ananthakrishnan, R., Walter, H. and Loser, J. (1973) *Hum. Hered.*, **23** : 263.
78. Fendley, T.W., Hochholzer, J.M. and Frings, C.S. (1973) *Clin. Chem.*, **19** : 1079.
79. Flocks, R.H. and Schmidt, J.D. (1972) *J. Surg. Oncol.*, **4** : 161.
80. Forbes, C.D., King, I. and McNicol, G.P. (1971) *Clin. Chem.*, **17** : 948.
81. Francke, U. and Busby, N. (1974) Assignments of the human genes for lactate dehydrogenase-A and thymidine kinase to specific chromosomal regions. *Rotterdam Conference* 143.
82. Francke, U. and Busby, N. (1975) *Cytogenet. Cell Genet.*, **14** : 313.
83. Francke, U., George, D.L., Brown, M.G. and Riccardi, V.M. (1977) *Cytogenet. Cell Genet.*, **19** : 197.
84. Fritz, P.J. and Jacobson, K.B. (1963) *Science*, **140** : 64.
85. Froewlich, C. (1971) *Aerztl. Forsch.*, **25** : 277 *Ibid* : 316, *Ibid* : 396.
86. Fujimato, Y., Najarian, I. and Wilkinson, J.H. (1968) *Enzymol. Biol. Clin.*, **9** : 124.
87. Ganrot, P.O. (1967) *Experientia*, **23** : 593.
88. Garbe, E., Randoux, A., Vielillard, A., Borel, J.P. and Cattan, A. (1971) *Rev. Eur. Etud. Clin. Biol.*, **16** : 482.
89. Gerhardt, W., Clausen, J., Christensen, E. and Riishede, J. (1963) *Acta Neurol. Scand.*, **39** : 85.
90. German, J.L. III, Evans, V.J. Cortner, J.A. and Westfall, B.B. (1964) *J. Natl. Cancer Inst.*, **32** : 681.
91. Gershbein, L.L. and Raikoff, K.G. (1977) *Clin. Chem.* **23** : 229.
92. Ghosh, A.K., Kirk, R.L., Joshi, S.R. and Bhatia, H.M. (1977) *Hum. Hered.*, **27** : 225.
93. Gold, J.A. (1961) *Diseases Chest*, **39** : 62.
94. Goldberg, E. (1963) *Science*, **139** : 602.
95. Goldman, R.D., Kaplan, N.O. and Hall, T.C. (1964) *Cancer Res.*, **24** : 389.
96. Goodfriend, T.L. and Kaplan, N.O. (1964) *J. Biol. Chem.*, **239** : 130.
97. Grzeschik, K.H. (1974) Assignment of genes to human chromosomes 1 and 11. *Rotterdam Conference*, p. 172.
98. Grzeschik, K.H. (1975) *Cytogenet. Cell Genet.*, **14** : 342.
99. Harris, J.I. and Perham, R.N. (1963) *Biochem. J.*, **89** : 60.
100. Harvey, R.G., Godber, M.J., Kope'c, A.C., Mourant, A.E. and Tills, D. (1969) *Hum. Biol.*, **41** : 342.
101. Hazelhorst, H. and Fockert, J.A. De (1959) *Folia Medica Neerlandica*, **2** : 79.

102. Helbig, W., Sommer, H.J., Parthum, K., Fleck, E. and Janke, H. (1971) *Folia Haematol.* (Leipzig), 95 : 347.
103. Helm, H.J. van der, Zondag, H.A., Hartog, H.A.Ph. and Kooi, W.M. van der (1962) *Clin. Chim. Acta*, 7 : 540.
104. Herbschleb-Voogt, E., Monteba-van Heuvel, M., Wijen, L.M.M., Westerveld, A., Pearson, P.L. and Meera Khan, P. (1978) *Cytogenet. Cell Genet.*, 22 : 482.
105. Herzog, P., Bohatová, J. and Drdová, A. (1970) *Am. J. Hum. Genet.*, 22 : 287.
106. Hess, B. and Gehm, E. (1955) *Klin Wochschr.*, 33 : 91.
107. Hess, B. and Walter, S.I. (1960) *Klin. Wochschr.*, 38 : 1080.
108. Hess, B. and Walter, S.I. (1961a) *Klin. Wochschr.*, 39 : 213.
109. Hess, B. and Walter, S.I. (1961b) *Ann. N.Y. Acad. Sci.*, 94 : 890.
110. Hill, B.R. (1955) *Proc. Am. A. Cancer Res.*, 2 : 24.
111. Hill, B.R. (1956) *Cancer Res.*, 16 : 460.
112. Hill, B.R. (1958) *Ann. N.Y. Acad Sci.*, 75 : 304.
113. Hill, B.R. and Levi, C. (1954) *Cancer Res.*, 14 : 513.
114. Hoch-Ligeti, C. (1962) *Cancer*, 15 : 816.
115. Hoenigova, J. and Hoenig, V. (1967) *Clin. Chim. Acta*, 18 : 312.
116. Holmsen, H., Day, H.J. and Primentel, M.A. (1969) *Biochim. Biophys. Acta*, 186 : 244.
117. Homolsky, M.W. and Kaplan, N.O. (1961) *Circulation*, 23 : 102.
118. Hopkins, D.A. and Harris, H. (1971) *Ann. Rev. Genet.*, 5 : 5.
119. Hsieh, K.M. and Blumenthal, H.T. (1956) *Proc. Soc. Exp. Biol. Med.*, 91 : 626.
120. Hsieh, K.M., Suntzeff, V. and Cowdry, E.V. (1956) *Cancer Res.*, 16 : 237.
121. Hule, V. (1967) *Clin. Chim. Acta*, 17 : 349.
122. Ishibe, T. (1971) *Invest. Urol.*, 9 : 104.
123. Ishibe, T., Mori, K., Usui, T., Nihara, H. and Yasuishi, K. (1975) *Urol. Int.*, 30 : 437.
124. Jenkins, T. and Nurse, G.T. (1976) *Prog. Med. Genet.*, 1 : 211.
125. Jongsma, A.P.M., Ios, W.R.T. and Hagemeijer, A. (1974) In : *New Haven Conference* (1973) : Third International Workshop on HumanGene Mapping. Birth Defects : Original article series 10 : 106. The National Foundation, New York.
126. Juul, P. (1967) *Clin. Chem.*, 13 : 416.
127. Kägi, J.H.R. and Vallee, B.L. (1960) *J. Biol. Chem.*, 235 : 3188.
128. Kar, N.C. and Pearson, C.M. (1963) *Proc. Nat. Acad. Sci.* (USA), 50 : 995.
129. Karl, W.F. and Peters, Th. Jr. (1967) *Am. J. Clin. Pathol.*, 47 : 171.
130. Kaplan, N.O. (1964) *Brookhaven Symp. Biol.*, 17 : 131.
131. Kaplan, N.O. and Ciotti, M.M. (1961) *Ann. N.Y. Acad. Sci.*, 94 : 701.
132. Kaplan, N.O. and Cahn, R.D. (1962) *Proc. Natl. Acad. Sci.*, 48 : 2123.
133. Kaplan, N.O. and White, S. (1963) *Ann. N.Y. Acad. Sci.*, 103 : 835.
134. Katz, A.M. and Kalow, W. (1965) *Can. J. Biochem.*, 43 : 1653.
135. Kessler, F.J. and Blumenberg, F.W. (1961) *Arzneimittel Forschung*, 11 : 188.
136. Khmelev, G.N. Yurkov, Y.A. and Evdokimova, A.I. (1969) *Vop. Okhr. Materin. Det.*, 14 : 38.

137. Kindmark, C.O. (1969) *Scand. J. Clin. Lab. Invest.*, 24 : 49.
138. King, J. and Waind, A.P.B. (1960) *Brit. Med. J.* 2 : 1361.
139. Kingsley, D.P.E., Cook, J. and Varban, A.E. (1965) *Clin. Chim. Acta*, 12 : 489.
140. Kitamura, M., Iijima, N., Hashimoto, F. and Hiratsuka, A. (1971) *Clin. Chim. Acta*, 34 : 419.
141. Kirk, R.L. (1974) Proc. Conf. Ind. Soc. Hum. Genet. I : *Human Population Genetics in India*, Orient Longman Ltd., Bombay.
142. Kirk, R.L. (1975) In : *Isoenzyme : genetics and evolution* 4 : *ed.* C.L. Markert, Academic Press, New York, p. 169.
143. Kirk, R.L., Blake, N.M., Moodie, P.M. and Tibbs, G.J. (1971a) *Hum. Biol. Oceania*, 1 : 54.
144. Kirk, R.L., Blake, N.M. and Vos, G.H. (1971b) *S. Afr. Med. J.*, 45 : 69.
145. Kirk, R.L., Keats, B., Blake, N.M., McDermid, E.M., Ala, F., Karimi, M., Nickbin, B., Shabazi, H. and Kamet, J. (1977) *Am. J. Phys. Anthropol.*, 46 : 377.
146. Klebe, R.J., Chen, T.R. and Ruddle, F.H. (1970) *J. Cell Biol.*, 45 : 74.
147. Klima, R., Rettenbacher-Däubner, H. and Rieder, H. (1962) *Wein. Klin. Wochschr.*, 74 : 90.
148. Kluge, A. and Haug, H. (1970) *Med. Welt.*, 46 : 1981.
149. Knudsen, F.U., Knudsen, H.E. and Gormsen, J. (1970) *Anal. Biochem.*, 36 : 192.
150. Konda, H. (1966) *Disc electrophoresis information service*, Canal Industrial Corp., Bethesda, Abstr. No. 382.
151. Kraus, A.P. and Neely, C.L. (1964) *Science*, 145 : 595.
152. Kreutzer, H.H. and Jacobs, P. (1965) *Clin. Chim. Acta*, 11 : 184.
153. Kreutzer, H.H., Jacobs, P. and Francke, C. (1965) *Clin. Chim. Acta*, 11 : 159.
154. Kurashvili, R.B. and Dzhikiya, L.V. (1974) *Kardiologia*, 14 : 124.
155. Latner, A.L. and Skillen, A.W. (1961) *Lancet*, 2 : 1286.
156. Latner, A.L. and Skillen, A.W. (1963) *Proc. Assc. Clin. Biochem.*, 2 : 100.
157. Laurova, L., Cizkova, J., Jandova, A. and Skoda, V. (1972), *Sb. Lek.*, 74 : 188.
158. Laurova, L., Jandova, A., Skoda, V. and Pezlarova, J. (1974) *Neoplasma*, 21 : 63.
159. Lawrence, S.H. (1962) *Lancet*, 2 : 559.
160. Lee, Y.T., Chandor, S.B. and Weiner, J.M. (1976) *Med. Pediatr. Oncol.*, 2 : 397.
161. Lefevre-Witier, P. and Vergnes, H. (1977) *Hum. Hered.*, 27 : 454.
162. Lehmann, H., Ala, F., Hedeyat, S., Montazemi, K., Nejad, H.K., Lightman, S.K., Ada, C., Mourant, A.E., Teedale, P. and Tills, D. (1973) *Phil. Trans. S.B.*, 266 : 195.
163. Levi, J.A., Speden, J.B., Vincent, P.C. and Gunz, F.W. (1973) *Cancer*, 31 : 939.
164. Lie-Injo, L.E., Kosasih, E.N. and Tann, G. (1974) *Humangenetik*, 22 : 331.
165. Lie-Injo, L.E., Lopez, C.G. and Ganesan, J. (1973) *Hum. Hered.*, 23 : 487.
166. Lubrano, T., Dietz, A.A. and Rubinstein, H.M. (1971a) *Clin. Chem.*, 17 : 657.
167. Lubrano, T., Dietz, A.A. and Rubinstein, H.M. (1971b) *Clin. Chem.*, 17 : 882.
168. Lührs, W. and Negelein, E. (1956) *Klin. Eschr.*, 34 : 148.
169. Malcolm, L.A., Woodfield, D.G., Blake, N.M., Kirk, R.L. and McDermid, E.M. (1972) *Hum. Hered.*, 22 : 305.

170. Markert, C.L. (1961) Isozymes in kidney development. In : Proceedings of the thirteenth annual conference of the kidney. Hereditary, developmental and immunologic aspects of kidney diseases. blz. 54. North-Western University Press, Evanston, Illinois, USA.
171. Markert, C.L. (1963) *Science*, 140 : 1329.
172. Markert, C.L. and Appella, E. (1961) *Ann. N.Y. Acad. Sci.*, 94 : 678.
173. Markert, C.L. and Appella, E. (1963) *Ann. N.Y. Acad. Sci.*, 103 : 915.
174. Markert, C.L. and Ursprung, H. (1962) *Develop. Biol.*, 5 : 363.
175. Masry, B.E., Berzy, I. and Nova'k, M. (1973) *Haematologia*, 7 : 405.
176. Massey, W.H., Dennis, D.L., Fletcher, W.S. and Wood, D.C. (1971) *Am. J. Surg.*, 122 : 209.
177. Mayeda, K., Weiss, L., Lindahl, R. and Dully, M. (1974) *Am. J. Hum. Genet.*, 26 : 59.
178. Mayr, K. and Schweitzer, F. (1962) *Wien. Med. Wochschr.*, 112 : 505.
179. McKenzie, D. and Henderson, A.R. (1974a) *Clin. Chim. Acta*, 51 : 53.
180. McKenzie, D. and Henderson, A.R. (1974b) *Clin. Chim. Acta*, 50 : 147.
181. McKusick, V.A. and Ruddle, F.H. (1977) *Science*, 196 : 390.
182. Meaney, R.F. Cavey, W.F. and Pollard, A.C. (1976) *Clin. Chim. Acta*, 73 : 127.
183. Milunicova, A., Jandiva, A., Laurova, L., Novotna, J. and Skoda, V. (1969a) *Neoplasma*, 16 : 303.
184. Milunicova, A., Jandova, A., Laurova, L., Novotna, J. and Skoda, V. (1969b) *Neoplasma*, 16 : 311.
185. Milunicova, A., Jandova, A. and Skoda, V. (1971) *Neoplasma*, 18 : 219.
186. Modiano, G. (1969) Cited in *The distribution of human blood groups and other polymorphisms*. Ed. A.E. Mourant, A.C. Kope'c and K. Domaniewska-Sobczak, 2nd ed. Oxford, London, 1976, p. 772.
187. Mohrenweiser, H., Neel, J.V., Mestrainer, M.A., Salzano, F.M., Migliazza, E., Simoes, A.L. and Yoshihara, C.M. (1979) *Am. J. Phys. Anthropol.*, 50 : 237.
188. Mourant, A.E., Beckman, L., Beckman, G., Nilsson, L.-O. and Tills, D. (1968) *Acta Genet. Stat. Med.*, 18 : 553.
189. Mourant, A.E., Godber, M.J., Kope'c, A.C., Tills, D. and Woodhead, B.G. (1976a) *Proc. R. Soc. Lond., Ser. B.*, 194 : 17.
190. Mourant, A.E., Kope'c, A.C. and Domaniewska-Sobczak, K. (1976b) *The distribution of human blood groups and other polymorphisms*, 2nd ed., Oxford, London.
191. Mukherjee, B.N. and Das, S.R. (1974) Distribution of some polymorphic enzyme group systems in India. 1, In : *Human Population Genetics in India*, Orient Longman, New Delhi, p. 21.
192. Mukherjee, B.N., Das, S.K., Basak, S.N. and Kellermann, G. (1974) *Hum. Biol.*, 46 : 425.
193. Mukherjee, B.N., Das, S.K. and Dash Sharma, P. (1973) *Humangenetik*, 19 : 285.
194. Mukherjee, B.N., Das, S.K. and Dash Sharma, P. (1975) *Ann. Hum. Biol.*, 2 : 201.
195. Mukherjee, B.N., Das, S.K., Malhotra, K.C., Kate, S.L., Mutalik, G.S., Sainani, G.S. and Bhidya, S. (1978) *Ann. Hum. Biol.*, 5 : 435.

196. Nabholz, M., Miggiano, V. and Bodmer, W. (1969) *Nature*, 233 : 358.
197. Nagamine, M. (1972) *Clin. Chim. Acta*, 36 : 139.
198. Nance, G.W., (1963) *Ann. N.Y. Acad. Sci.*, 103 : 980.
199. Nance, W.E., Claflin, A. and Smithies, O. (1963) *Science*, 142 : 1075.
200. Neel, J.V., (1978) *Ann. Rev. Genet.*, 12 : 365.
201. Neel, J.V., Ferrell, R.E. and Canaro, R.A. (1976) *Am. J. Hum. Genet.*, 28 : 262.
202. Neel, J.V., Salzano, F.M., Junqueira, F., Keiter, F. and Maybury-Lewis, D. (1964) *Am. J. Hum Genet*, 16 : 52.
203. Neel, J.V., Ueda, N., Satoh, C., Ferrell, R.E., Tanis, R.J. and Hammilton, H.B. (1978) *Ann. Hum. Genet.*, 41 : 429.
204. Neely, C.L., Wajma, T., Karus, A.P., Diggs, L.W. and Barreas, L. (1969) *Am. J. Clin. Pathol.*, 52 : 167.
205. Nilsson, L.-O. and Erikson, A.W. (1972) *Hum. Hered.*, 22 : 372.
206. Nisselbaum, J.S. and Bodansky, O. (1959) *J. Biol. Chem.*, 234 : 3276.
207. Nisselbaum, J.S. and Bodansky, O. (1961a) *J. Biol. Chem.*, 236 : 323.
208. Nisselbaum, J.S. and Bodansky, O. (1961b) *J. Biol. Chem.*, 236 : 401.
209. Nisselbaum, J.S. Packer, D.E. and Bodansky, O. (1964) *J. Biol. Chem.*, 239 : 2830.
210. Nitowsky, H.M. and Soderman, D.D. (1964) *Exp. Cell. Res.*, 33 : 562.
211. Nurse, G.T. and Jenkins, T. (1977) *Ann. Hum. Biol.*, 4 : 465.
212. Omoto, K., Misawa, S. Harada, S. Sumpaico, J.S. Medado P.M. and Ogonuki, H. (1978) *Am. J. Hum. Genet.*, 30 : 190.
213. Papiha, S.S. and Chhaparwal, B.C. (1973) *9th Int. Cong. Anthrop. and Ethno. Sc.*, Chicago, USA.
214. Papiha, S.S. Roberts, D.F., Wig., N.N. and Singh, S. (1972) *Am. J. Phys. Anthropol.*, 37 : 293.
215. Papiha, S.S., Roberts, D.F., Ali., S.G.M. and Islam, M.M. (1975) *Human-genetik*, 28 : 285.
216. Papiha, S.S., Roberts, D.F., and Gulati, P.D. (1976) *Hum. Biol.*, 48 : 323.
217. Pfleiderer, G. (1978) *Naturwissenschaften*, 65 : 397.
218. Pfleiderer, G. and Wachsmuth, E.D. (1961a) *Biochem. Zs.*, 334 : 185.
219. Pfleiderer, G. and Wachsmuth, E.D. (1961b) *Klin. Wschr.*, 39 : 352.
220. Philip, J. and Vesell, E.S. (1962) *Proc. Soc. Exp. Biol. Med.*, 110 : 582.
221. Plagemann, P.G.W., Gregory, K.F. and Wróblewski, F. (1960a) *J. Biol. Chem.*, 235 : 2282
222. Plagemann, P.G.W., Gregory, K.F. and Wróblewski, F. (1960b) *J. Biol. Chem.*, 235 : 2288.
223. Plagemann, P.G.W., Gregory, K.F. and Wróblewski, F. (1961) *Biochem., Z.* 334 : 37.
224. Plummer, D.T. and Wilkinson, J.H. (1961) *Biochem. J.* 81 : 38.
225. Pollitzer, W.S., Ruckengel D., Tashian, R., Shreffler, D.C., Leyshow, W.C., Namboodiri, K. and Elston, R.C. (1970) *Am. J. Phys. Anthropol.*, 32 : 65.

226. Ramesh, A., Murty, J.S. and Blake, N.M. (1979) *Hum. Hered.*, **29** : 147.
227. Rao, P·M., Blake, N.M. and Veeraju, P. (1978a) *Hum .Hered.*, **28** : 122.
228. Rao, Y.N., Singh, G.P., Chakravorty, M. and Khanna, N.N. (1978b) *Ind. J. Cancer*, **15** : 39.
229. Rehn, J. and Köhnlein, E. H. (1960) *Med. Welt.*, **16** : 841.
230. Ressler, N. and Joseph, R. (1962) *J. Lab. Clin. Med.*, 60 : 349.
231. Rethore, M.-O Junien, C., Malpuech, G., Baccichetti, C., Tenconi, R., Kaplan, J.-C., de Remeuf, J. and Lejeune, J. (1976) *Ann. Genet.*, **19** : 140.
232. Rethore', M.-O., Kaplan, J.C., Junien, C., Cruvellier, J., Dutrillaux, B., Aurias, A., Carpentier, S., Lafourcade, J., and Lejeune, J. (1975) *Ann. Genet.*, **18** : 81.
233. Richterich, R., Burger, A., and Weber, H. (1962) *Helv. Physiol. Acta*, 20 : C78.
234. Richterich, R., Schafroth, P. and Aebi, H. (1963) *Clin. Chim. Acta*, **8** : 178.
235. Roberts, D.F., Papiha, S.S. Creen, C.K., Chhaparwal, B.C. and Mehta, S. (1974) *Ann. Hum. Biol.*, **1** : 159.
236. Rodney, M.B. and Shapiro, M. (1959) *Am. Heart. J.*, **58** : 807.
237. Roman, W. (1969) *Enzymologia*, **36** : 189.
238. Rose, A., West, M. and Zimmerman, H.J. (1961) *Cancer*, **14** : 726.
239. Rothwell, D.J., Jendrzejczak, B., Becker, M. and Doumas, B.T. (1976) *Clin. Chem.*, **22** : 1024.
240. Roychoudhury, A. (1980) *Indian Society of Human Genetics*, Hyderabad—8th Feb. to 10th Feb. 1980.
241. Roychoudhury, A.K. (1975) *Trans. Bose Res. Inst.* (Calcutta), **38** : 9.
242. Ruddle, F.H. (1972) Linkage analysis using somatic cell hybrids. In : *Advances in Human Genetics* **3** : *ed.* H. Harris and K. Hirschhorn, Plenum, New York, p. 173.
243. Ruddle, F.H. and Nichols, E.A. (1971) *in vitro*, **7** : 120.
244. Ruddle, F.H., Chapman, V.M., Chen, T.R. and Klebe, R.J. (1970) *Nature*, **227** : 251.
245. Ruegsegger, P., Nydick, I., Freiman, A. and Ladue, J.S. (1959) *Circulation Res.*, **7** : 4.
246. Safronova, E.E., Gubernieva, L.M., Volkova, N.M., Mamafo, V.B. and Annenkov, G.A. (1978) *Biokhimiya*, **43** : 268.
247. Saha, N. and Banerjee, B. (1973) *Acta Genet. Med. Gemellol.*, **22** : 117.
248. Saha, N., Kirk, R.L., Shanbhag, S., Joshi, S.H. and Bhatia, H.M. (1974) *Hum. Hered.*, **24** : 198.
249. Saha, N., Kirk, R.L., Shanbhag, S., Joshi, S.R. and Bhatia, H.M. (1976) *Hum. Hered.*, **26** : 175.
250. Saha, N., Samuel, A.P.W., Omer, A., Ahmed, M.A., Hussein, A.A. and Gaddoura. E.N. (1978) *Ann. Hum. Biol.*, **5** : 569.
251. Santachiara, S.A., Nabholz, M., Miggiano, V., Darlington, A.J. and Bodmer, W. (1970) *Nature*, **227** : 248.
252. Schmidt, E., Schmidt, W. and Otto, P. (1967) *Clin. Chim. Acta*, **15** : 283.
253. Schmitt, J., Ritter, H. and Schmidt, K. (1975) *Humangenetik*, **28** : 239.
254. Schwartz, M.K. (1973) *Clin. Chem.* **19** : 10.

255. Schwartz, M.K. (1977) *Ann. Clin. Lab. Sci.*, **7** : 99.
256. Sharma, S.K. and Baxi. A.J. (1978) *Ind. J. Med. Res.*, 68 : 656.
257. Shaw, C.R. and Barto, E. (1963) *Proc. Nat. Acad. Sci.* (USA), **50** : 211.
258. Shaw, C.R. and Koen, A.L. (1965) *J. Histochem. Cytochem.*, **13** : 431.
259. Shoda, T., Ueda, S. and Yamada, O. (1975) *Kowasaki Med. J.*, **1** : 187.
260. Shows, T.B. and McAlpine, P.J. (1978) *Cytogenet. Cell Genet.*, **22** : 132.
261. Shows, T.B. and Ruddle, F.H. (1968) *Proc. Nat. Acad. Sci.* (USA), 61 : 574.
262. Singh, S., Sareen, K.N. and Goedde, H.W. (1974) *Humangenetik*, **22** : 133.
263. Sinnett, P., Blake, N.M., Kirk, R.L., Lai, L.Y.C. and Walsh, R.J., (1970) *Archaeol. Phys. Anthropol. Oceania*, **5** : 236.
264. Stambaugh, R. and Buckley, J. (1967) *J. Biol. Chem.*, **242** : 4053.
265. Starkweather, W.H. and Schoch, H.K. (1962) *Biochim. Biophys. Acta*, 62 : 440.
266. Stellwegen, E. and Schachman, H.K. (1962) *Biochemistry*, (Wash.), **1** : 1056.
267. Stott, A. (1973) *Ann. Clin. Biochem.*, **10** : 72.
268. Strandjord, P.E. and Clayson, K.J. (1966) *Postgrad. Med.*, A 60.
269. Strandjord, P.E., Clayson, K.J. and Freier, E.F. (1962) *J. Am. Med. Assoc.*, **182** : 1099.
270. Strandjord, P.E., Thomas, K.E. and White, L.P. (1959) *J. Clin. Invest.*, **38** : 2111.
271. Sugano, G., Kano, S., Saga, T., Irihisa, T., Iwagaki, H. and Nazaki, U. (1973) *Igakuno, Ayumi.*, **87** : 621.
272. Sukernik, R.I., Karaphet, T.M. and Osipova, L.P. (1978) *Hum. Hered.*, **28** : 321.
273. Sumner, A.T., Evans, H.J. and Buckland, R.A. (1971) *Nature New Biol.* **232** : 31.
274. Sunderland, E., Sawhney, K.S. and Bethell, I.L. (1979) *Hum. Hered.*, **29** : 14.
275. Surinov, B.P. and Menzhinskaya, G.V. (1975) *Vopr. Onkol. (Leningr.)* **21** : 23.
276. Swain, B.K., Sharma, A. and Talukder, G. (1980a) Effect of thawing and freezing on the activity of lactic acid dehydrogenase (In press).
277. Swain, B.K., Talukder, G. and Sharma, A. (1980b) A case of an abnormal LDH isoenzyme band from a patient with myocardial infarction (In press).
278. Talageri, V.R. (1977) *Ind. J. Cancer*, **14** : 50.
279. Talageri, V.R., Nadkarni, J.S. and Goollerkeri, M.P. (1977) *Ind. J. Cancer*, **14** : 42.
280. Tamura, Z. and Oda, T. (1970) *ed. Serum Enzyme*, Igaka-Shoin Co. Ltd., Tokyo.
281. Tanaka, F., Amino, N., Hayashi, C., Miyai, K. and Kumahara, Y. (1976) *Clin. Chim. Acta*, **68** : 235.
282. Tanis, R.J., Neel, J.V., Dovey, H. and Morrow, T. (1973) *Am. J. Hum. Genet.*, **25** : 655.
283. Tanis, R.J., Neel, J.V. and Torres de Arauz, R. (1977) *Am. J. Hun. Genet.*, **29** : 419.
284. Tashian, R.E. (1965) *In : Progress in Medical Genetics*, **4** : (1965) *ed.* A.G. Steinberg, and A.G. Bearn, William Heineman Medical Books Ltd., London, p. 128.
285. Tashian, R.E. Brewer, G.J., Lehmann, H., Davie, D.A. and Rucknagel, D. (1967) *Am. J. Hum. Genet.*, **19** : 524.
286. Tills, D. (1977) *Ann. Hum. Biol.*, **4** : 35.
287. Tischfield, J., Bernhard, H.P. and Ruddle, F.H. (1973) *In* : Chen *et al.* (1973) *Am. J. Hum. Genet.*, **25** : 200.

288. Trujillo, N.P., Nutter, D. and Evans, J.M. (1967) *Arch. Int. Med.*, **119** : 333.
289. Tsao, M.V. (1960) *Arch. Biochem.*, **90** : 234.
290. Ueda, N., Sato, C., Tanis, R.-J., Ferrell, R.E. Kishimoto, S., Neel, J.V., Hamilton, H.B. and Baba, K. (1977) *Ann. Hum. Genet.*, **41** : 43.
291. Undevia, J.V. (1969) *Population genetica of the Parsis : Comparison of genetical characteristics of the present Parsi population with its ancestral and affiliated groups.* Ph.D. Thesis, Bombay University.
292. Undevia, J.V., Balakrishnan, V., Kirk, R.L., Blake, N.M., Saha, N. and McDermid, E.M. (1978) *Hum. Hered.*, **28** : 104.
293. Undevia, J.V., Blake, N.M., Kirk, R.L. and McDermid, E.M.(1972) *Hum. Hered.*, **22** : 274.
294. Van Gennip, A.H., Tabak-Van Gorcum, J.A., Taminiau, J.A.J.M. and Wadman, S.K. (1975) *Clin. Chim. Acta*, **58** : 85.
295. Van Someren, H., Meera Khan, P., Westerveld, A. and Bootsma, D. (1972) *Nature (New Biol.)*, **240** : 221.
296. Van Someren, H., Meera Khan, P., Westerveld, A. and Bootsma, D. (1974) *Nature* **249** : 279
297. Vergnes, H. and Cabannes, R. (1976). *Ann. Hum. Biol.*, **3** : 423.
298. Vergnes, H. and Gherardi, M. (1971) *Ann. Genet.*, **14** : 199.
299. Vergnes, H., Quilici, J.C., and Constans, J. (1976) *Ann. Hum. Biol.*, **3** : 577.
300. Vergnes, H., Quilici, J.C., Gherardi, M. and Bejarano, G. (1976) *Hum. Hered.*, **26** : 252.
301. Vesell, E.S. (1961) *Ann. N.Y. Acad. Sci.*, **94** : 877.
302. Vesell, E.S. (1962) *Nature*, **195** : 497.
303. Vesell, E.S. (1965a) In : *Progress in Medical Genetics, ed.* A.G. Steinberg and A.G. Bearn, William Heinemann Medical Books Ltd., London, p. 128.
304. Vesell, E.S. (1965b) *Proc. Natl. Acad. Sci.*, **54** : 111.
305. Vesell, E.S. and Bearn, A.G. (1957) *Proc. Soc. Exp. Biol.*, **94** : 96.
306. Vesell, E.S. and Bearn, A.G. (1958) *Ann. N.Y. Acad. Sci.*, **75** : 286.
307. Vesell, E.S. and Bearn, A.G. (1961) *J. Clin. Invest.* **40** : 580.
308. Vesell, E.S. and Bearn, A.G. (1962a) *Proc. Soc. Exp. Biol. Med.*, **111** : 100.
309. Vesell, E.S. and Bearn, A.G. (1962b) *J. Gen. Physiol.*, **45** : 553.
310. Vesell, E.S. and Brody, I.A. (1964) *Ann. N.Y. Acad. Sci.*, **121** : 544.
311. Vesell, E.S., Philip, J. and Bearn, A.G. (1962) *J. Exp. Med.*, **116** : 796.
312. Vesell, E.S. and Yielding, K.L. (1968) *Ann. N.Y. Acad. Sci.*, **151** : 678.
313. Vetter, K. (1961) *Folia Haematologica*, **6** : 80.
314. Vyas, G.N., Peterson, D.L., Townsend, R.M., Damle, S.R. and Magnius, L.O. (1977) *Science*, **198** : 1068.
315. Wacker, W.E.C., Ulmer, D.D. and Vallee, B.L. (1956) *New. Eng. J. Med.*, **255** : 449.
316. Warburton, F.G., Allan, E., Laing G.S. and Naddecar, J. (1967) *Nature*, **215** : 287.
317. Weil, D., Van Cong, N., Rebourcet, R. and Frezal, J. (1978) *Hum. Genet.*, **42** : 181.

318. Weiss, M.C. and Green, H. (1967) *Proc. Nat. Acad. Sci.* (USA), **58** : 1104.
319. Weiss, M.C., Ephrussi, B. and Scaletta, L.J. (1968) *Proc. Nat. Acad. Sci.* (USA), **59** : 1132.
320. Weitkamp, L.R. and Neel, J.V. (1970) *Am. J. Hum. Genet.*, **22** : 533.
321. West, M. and Zimmerman, H.J. (1958) *A.M.A. Arch. Int. Med.*, **102** : 103.
322. West, M., Heller, P. and Zimmerman, H.J. (1958) *Am. J. Med. Sci.*, **235** : 689.
323. Westhead, E.W., Butler, L. and Boyer, P.D. (1963) *Biochemistry* (Wash.), **2** : 927.
324. White, L.P. (1956) *New Eng. J. Med.*, **255** : 984.
325. Wieland, Th. and Pfleiderer, G. (1957) *Biochem. Zs.*, **329** : 112.
326. Wieland, Th. and Pfleiderer, G. (1961) *Ann. N.Y. Acad. Sci.*, **94** : 691.
327. Wieland, Th., Pfleiderer, G., Haupt, I. and Worner, W. (1959) *Biochem, Z.*, **332** : 1.
328. Wieme, R.J. (1959a) *Clin. Chim. Acta*, **4** : 46.
329. Wieme, R.J. (1959b) *Clin. Chim. Acta*, **4** : 321.
330. Wieme, R.J. (1962) *Lancet*, **1** : 70.
331. Wieme, R.J., VanHove, W.Z. and Van der Straelen, E.M. (1968) *Ann. N.Y. Acad. Sci.*, **151** : 213.
332. Wilkinson, J.H. and Withycombe, W.A. (1965) *Biochem. J.*, **97** : 663.
333. Wilson, A.C. Cahn, R.D. and Kaplan, N.O. (1963) *Nature*, **197** : 331.
334. Wilson, A.C. Kaplan, N.O. Levine, L., Pesce, A., Reichlin, M. and Allison, W.S. (1964) *Fed. Proc.* **23** : 1258.
335. Wilson, E.W. (1977) *Contraception*, **16** : 367.
336. Woodfield, D.G., Scragg, R.F.R., Blake, N.M., Kirk, R.L. and McDermid, E.M. (1974) *Hum. Hered.*, **24** : 507.
337. Wro'blewski, F. (1957) *Am. Heart. J.*, **54** : 219.
338. Wro'blewski, F. and Gregory, K.F. (1961a) In : *Proceedings of the fourth International Congress on Clinical Chemistry*, Edinburgh, 1960, E & S Livingstone Ltd., Edinburgh.
339. Wro'blewski, F. and Gregory, K.F. (1961b) *Ann. N.Y. Acad. Sci.*, **94** : 912.
340. Wro'blewski, F., Rose, C. and Gregory, K.F. (1960) *New Eng. J. Med.*, **263** : 531.
341. Wro'blewski, F., Ruegsegger, P. and Ladue, J.S. (1956) *Science*, **123** : 1122.
342 Wu, C.P. and Juli, L.S. (1962) *Gastroenterology*, **42** : 580.
343. Wüst, H., Arnold, E. and Schön, H. (1965) *Klin. Ws.*, **43** : 500.
344. Wüst, H., Berg, G. and Schon, H. (1960) *Verh. Dtsch. Ges. Inn. Med.*, **66** : 216.
345. Wüst, H., Schön, H. and Berg, G. (1962) *Klin. Ws.*, **40** : 1169.
346. Yakulis, V.J., Gibson, C.W. and Heller, P. (1962) *Am. J. Clin. Pathol*, **38** : 378.
347. Yielding, K.L. and Tomkins, G.M. (1960) *Proc. Nat. Acad. Sci.* (USA) **46** : 1483.
348. Zimmerman, H.J. and Weinstein, H.G. (1956) *J. Lab. Clin. Med.*, **48** : 607.
349. Zinkham, W.H., Blanco, A. and Kupchyk, L. (1963) *Science*, **142** : 1303.
350. Zondag, H.A. (1963) *Science*, **142** : 965.
351. Zondag, H.A. (1964) *Rhode Island Med. J.*, **47** : 273.

352. Zondag, H.A. (1965) In : *Enzymes in Clinical Chemistry, ed.* Ruyssen-Vandendriessche, Elsevier Pub. Co., Amsterdam.
353. Zondag, H.A. (1967) *Determination and diagnostic significance of lactate dehydrogenase isoenzymes*, Van Gorcum and Comp. N.V. Assen.
354. Zondag, H.A. and Klein, F. (1968) *Ann. N.Y. Acad. Sci.*, 151 : 578.
355. Zuppinger, K., Richterich, R. and Rossi, E. (1962) *Schweiz. Med. Wochschr.*, 92 : 169 ; Ibid, 92 : 198.

The Role of Light in Control of Plant Growth and Development

S. C. Maheshwari[1], Vipin Rustogi[1], Kusum Malhotra[1]
and
Sudhir K. Sopory[2]

[1]Department of Botany, University of Delhi, Delhi
and
[2]School of Life Sciences, Jawaharlal Nehru University, New Delhi

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

INTRODUCTION

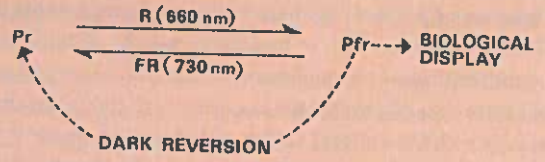
The complex relationship between plants and sunlight has been under investigation for many years. Of the various interactions between plants and light, the best understood is photosynthesis, i.e. photophosphorylation and concomitant carbon dioxide fixation. However, light exerts a control and a powerful one—also overgrowth and differentiation in plants quite independent of photosynthesis. These developmental processes can be broadly classified into two major categories, namely, (a) *photoperiodic*—in which periodic stimuli may control flowering, leaf-fall, seed germination, onset and breaking of dormancy and a number of other morphogenic processes in all of which some form of a time-measuring device operates, and (b) *non-photoperiodic*—in which we may group a number of other processes such as phototropism where directional stimuli elicit bending movements of plant organs and appropriate orientation of plant parts to intercept maximal sunlight.

In this article our main attention is on the first category of processes controlled by a photoreceptor called phytochrome, because these responses are not only the most intensively investigated but also work on its mechanism of action has been under investigation in our own laboratory for some years. Nevertheless, the other photomorphogenic phenomena under category (b) are extremely important, too, and during the last 2-3 years, a great deal of evidence has accumulated to show that some of these are under control of a new blue light absorbing pigment, attention to which is invited in the concluding section of this article.

History of discovery and isolation of phytochrome

To take up phytochrome first, our understanding of the control of plant growth and differentiation by light, i.e. photomorphogenesis, began with the pioneering work of Garner and Allard [41, 42] who first enunciated the principles of photoperiodism. That red light is the effective region of the solar spectrum was shown by the neoclassical work of Flint and McAlister[33]. But, precise location of the active bands and involvement of a reversible photoreaction, i.e. antagonistic effects of R (red) and FR (far-red) light in the control of seed germination and flowering and the equivalence of FR light to darkness in terms of its physiological effect, was first established by

the group of Borthwick and Hendricks at the U.S. Department of Agriculture at Beltsville, Maryland, based on their action spectra studies[7,8]. The term 'phytochrome' was coined in the same laboratory when the pigment was detected spectrophotometrically both *in vivo* and *in vitro*[18]. This was made possible by designing an appropriate spectrophotometer for detecting the small changes in absorbance, at 660 and 730 nm, against a large background absorption. Phytochrome was postulated to exist in two photoreversible forms, the R absorbing (Pr) and FR absorbing (Pfr) form, whose relationship can schematically be represented as follows :



Pr, the prevailing stable form in darkness, is considered to be inactive and Pfr is considered to be biologically active form and responsible for inducing diverse and multiple phytochrome-mediated changes at the molecular, cellular and morphological levels. Pfr can revert to Pr form in dark (dark reversion), although more slowly, and this forms the basis for time-measuring device of photoperiodic reactions in many plants.

Molecular properties

The study of isolation and biochemical characterization of phytochrome is undoubtedly one of the most brilliant accomplishments of modern biology and we summarize below the major findings.

Early in the work, based on its physiological properties, phytochrome was postulated to be a protein[58] and indeed partially purified from maize by ammonium sulphate precipitation[18]. Subsequently, purification of phytochrome from etiolated oat seedlings was undertaken by employing gel filtration[186]. Since then, the chromoprotein has been purified (to near homogeneity) from oat itself by several groups[12, 77, 107, 123]. Similarly, highly purified phytochrome preparations have also been obtained from rye by certain groups[20, 21, 110] and from pea[155]. Although initial reports indicated that phytochrome obtained from oat was substantially different from that of rye, but later the difference was found to be due to an artifact, i.e. partial

destruction by endoproteases in oat[12]. The most suitable procedure which yields undegraded phytochrome is the one described by Briggs and co-workers[110] for rye tissue. This protocol has been derived from earlier procedures after a few necessary modifications to minimize proteolysis and also utilizes sequential column chromatography with brushite, diethylaminoethyl-cellulose, hydroxyapatite, Bio-Gel A and Sephadex G-200. Recently, phytochrome has also been purified through immunoaffinity chromatography[57].

Biochemical properties

Mainly because of proteolysis there had been considerable controversy between different laboratories over molecular weight, subunit structure and amino acid composition of phytochrome. However, there seems to be general agreement now that phytochrome is a dimer of 120,000 dalton subunits as determined by sodium dodecyl sulfate gel electrophoresis[72,107,140]. In the electron microscope the subunits appears to be dumb-bell shaped[140]. Amino acid analysis of phytochrome indicates a high content of polar as well as non-polar amino acids suggesting that the pigment is capable of undergoing both hydrophilic and hydrophobic interactions. The optical data of Tobin and Briggs[150] indicated that the tertiary structure of the protein moiety is composed of 20% α-helix, 30% β-helix and 50% random coil.

Conformational changes during phototransformation

The two forms of phytochrome, i.e. Pr and Pfr differ in their absorption spectra and photomorphogenic actions, and therefore, it was quite natural to assume that they must be undergoing some conformational changes during phototransformation. Butler *et al.*[17] provided evidence for minor conformational alterations based on denaturation by various agents such as urea, trypsin and pronase, N-ethyl maleimide and paramercuribenzoate. It has been found that the two forms (i.e. Pr and Pfr) were differentially sensitive to spectral loss, in general the Pfr being more susceptible. Roux and Hillman[122] studied the effect of glutaraldehyde on phytochrome and also showed that the Pr and Pfr forms are differentially labile. Conformational changes on phototransformation are also supported by the differential velocity sedimentation studies of Hopkins (*cited in* Briggs and Rice[12]) who showed that Pr had a slightly larger sedimentation coefficient (5·1 S) than Pfr (5·0 S). Differences between Pr and Pfr have been reported also by more

modern spectroscopic studies—circular dichroism and optical rotatory dispersion[2,56,76].

Chromophore

As early as 1950, on the basis of action spectra studies Parker and co-workers postulated that phytochrome might be containing a phycocyanin-like pigment as chromophore. Later, Siegelman and Hendricks[157] reported that the chromophore, a bilitriene, resembled algal pigments. There has been considerable progress in our understanding of the nature of chromophore. However, whether the two subunits of the 240,000 dalton molecule of phytochrome have two chromophores[80] or one chromophore each[106] remains unresolved.

The chromophore was shown to be attached to the peptide leu-arg-ala-pro-his-(ser, cys)-his-leu-glu-try[85]. Two possible alternatives had been proposed by Borthwick *et al.*[59] to explain photochemical nature of reversible absorption changes between Pr and Pfr, e.g. (a) *cis-trans* isomerization and (b) tautomerism. Many studies have been done since then employing low temperature, circular dichroism spectroscopy and other techniques[14,46,126,148]. It appears that the tetrapyrrole chain is extended in the Pr as well as the Pfr forms, but phototransformation may be accompanied by a more condensed cyclic intermediate and both a *cis-trans* isomerization and a proton exchange with the protein might be involved in this rearrangement.

Other photochemical aspects—Existence of intermediates

After the capture of a quantum of light energy by Pr, in about 10^{-17} seconds, the chromophore molecule undergoes rearrangement and this culminates in the formation of a metastable Pfr form in about 4 s[107]. On the other hand, reconversion of Pfr to Pr form after photon absorption is complete within 10^{-2}s. The relatively long time taken by these reactions after photon absorption is suggestive of a complex pathway in both directions. A number of intermediates appearing during the reversible phototransformation of the two forms have been identified and characterised spectrophotometrically. These intermediates were first detected *in vitro*, using flash photolysis and low-temperature spectrophotometric techniques. In flash photolysis, the effects of short (about 5 μs) pulse of irradiation on the absorption changes at a particular wavelength are recorded by an oscilloscope[80,109]. On the other hand, in low-temperature spectrophotometry, advantage is taken of

the fact that irradiation at very low temperature followed by gradual warming results in sequential appearance of different intermediates of phytochrome—because of slower reactions, under such circumstances, ordinary recording arrangements are adequate[22,108]. Irradiation at low temperature to whole plants, i.e. *in vivo* has also revealed a number of intermediates[78]. Another method of detecting intermediates is pigment cycling in which the sample is irradiated with light that excites both Pr and Pfr resulting in accumulation of relatively more stable intermediates which again are detectable by conventional equipment and at ordinary temperature[31].

Analysis of data obtained by the application of these methods indicates that the two phototransformation pathways involve different intermediates, of which four have been identified in the Pr to Pfr path and two in the Pfr to Pr path[78]. An interesting point with regard to intermediates is that under certain conditions, these may constitute a large fraction of total phytochrome pool—if actinic beam comparable in intensity to that of natural sunlight is used—suggesting that perhaps one of these intermediates may be physiologically active[72].

Biologically active form of phytochrome

In etiolated tissue phytochrome is present in the Pr form. A measurable effect on the enlargement of etiolated pea leaves has been observed on irradiation with red light of very low energy (10^{-14} Einstein sq cm^{-1}) which is sufficient to convert just 0·004% of the phytochrome to Pfr form[52]. One of the conclusions derived from this fact by earlier workers was that Pfr may be the active form of phytochrome. This conclusion or assumption has neither been seriously questioned nor unequivocally substantiated. In general, the responses to red light—below the saturation but above a certain threshold level—are directly proportional to the logarithm of the incident light. But in certain instances the attempts to correlate quantitatively the spectrophotometrically detectable Pfr with biological responses have failed completely. Hillman[55] coined the term 'paradoxes' for such situations and proposed the concept of 'active' and 'bulk' (spectrophotometrically detectable but physiologically inactive) phytochrome to rationalize these paradoxes. Several workers have taken recourse to kinetic studies to discover whether there are sub-species of phytochrome molecules, but, no definitive evidence has yet been found.

Nevertheless, some difficulties have arisen in the literature in accepting

the classical view of Hendricks and Borthwick for phytochrome action, i.e. of Pfr being the sole active form. Many workers have noted that prolonged irradiation by monochromatic light has exactly an opposite physiological effect than obtained by short pulses. Such phenomena have been investigated in detail by Mohr and his co-workers[95]. Thus, although it is well known that short duration of red light inhibits stem elongation and far-red reverses it—prolonged irradiation by far-red light leads to the same effect as red light—the action maximum, however, shifts to 713 from 737 nm. It is important to remind ourselves at this juncture that in nature probably the latter situation, i.e. of prolonged irradiation, prevails. One explanation of these unusual responses is that one requires an excited species of Pfr, Pfr*, or it may be an intermediate in the Pr→Pfr conversion process, rather than Pfr itself. Another view is that a dynamic cycling reaction between Pr and Pfr may be more important than mere static conversion to one or the other form[60]. Yu[162] has also argued along somewhat similar lines, i.e. in favour of some revision of the original Hendricks and Borthwick proposal.

Mode of action—The case for phytochrome control of gene activity

Elucidation of the nature and molecular properties of phytochrome has been only the first step along the difficult road of understanding the mode of its action.

The extraordinary complexity of phytochrome responses was already evident in the fifties from the diversity of its morphological and physiological effects. These range from the most classical example of flowering and seed germination in a number of plants to initiation of leaf-fall, unbending of the hook in seedlings, stem elongation or expansion of leaves[58]. To consider only one of them in detail, e.g. flowering, even in the most suitable systems—e.g. in the Lemnaceae[47], 3 or 4 days are taken for its consummation and in many plants it may take weeks. Of course, this is not surprising since flower primordia initiation and development of its various parts, as corolla, calyx and sex organs, involves not only photoperception, but also translocation of some kind of stimulus to the growing points of plants, followed by initiation of cell divisions, and finally organogenesis. All these steps require profound changes in cellular metabolism. And, similarly, other morphogenic responses too were indicative of complex changes involved in metabolic activity of the cell. Therefore, it had been realized long ago that there must be

a primary amplification mechanism to link the first step, i.e. formation of Pfr with the manifestation of these responses.

The earliest idea about the primary mechanism of phytochrome action is the one proposed by Hendricks and Borthwick[58] that it might be acting as an enzyme. Somewhat later, based on the stimulative effect of light on anthocyanin synthesis in discs of apple skin—in darkness instead of anthocyanins, volatile esters are formed—it was suggested that phytochrome may be an enzyme responsible for converting acetate to acetyl Co-A, as the latter serves as an essential intermediate in synthesis and utilization of fat, in cellular respiration and in anthocyanin synthesis[10]. However, this suggestion was based on minimal evidence and since phytochrome effects are so diverse, this idea later has never been taken or considered too seriously. Furthermore, anthocyanin synthesis itself begins after a lag phase, which should not have occurred if pigment itself was the suspected enzyme.

Control of enzyme synthesis—De Novo synthesis

The first serious theory about mechanism of action of phytochrome is the differential control of gene activity proposed by Hans Mohr[94] which we discuss below. Mohr and co-workers showed that Pfr formation led to considerable changes in rate of protein and RNA synthesis[95,157]. Similarly, anthocyanin biosynthesis also depended on fresh RNA synthesis as inferred from the studies using actinomycin D, the well-known inhibitor for transcription. A requirement of synthesis of new RNA was also found obligatory for a positive photoresponse in enlargement of cotyledons in mustard seedlings. About the same time, Durst and Mohr[28] reported phytochrome control of activity of an enzyme, phenylalanine ammonia-lyase in mustard seedlings. Earlier phytochrome control of another enzyme, $NADP^+$ dependent glyceraldehyde-3-phosphate dehydrogenase, had been demonstrated[86,87]. In any case, since 1966, the activity of a large number of enzymes has been shown to be under the control of phytochrome[181]. Some of the important enzymes are adenylate kinase[15], α-amylase[26], ascorbic acid oxidase[25,151], glycolate oxidase[75], isocitrate dehydrogenase[44], lipoxygenase[99,144], nitrate reductase[64,159], peroxidase[138], ribonuclease II[1], transketolase[82] and δ-amino laevulinate dehydratase[5].

Perhaps the most intensively studied enzyme system is phenylalanine ammonia-lyase in mustard cotyledons by Mohr and co-workers. This enzyme starts accumulating in response to a brief red exposure, with a lag of 40 min,

and a peak activity at about 20 hours[132]. The experiments with the usual inhibitors of RNA and protein synthesis demonstrate that synthesis of new enzyme molecules is involved. More recent studies using heavy isotopes have provided further evidence in support of the idea that *de novo* synthesis of proteins is the basis for red-induced increase in enzyme activity[181].

Mohr and his students have carried out studies on several other enzymes, in most of which similar *de novo* synthesis of enzyme molecules is a constant feature. The phytochrome regulation of peroxidase activity has also been shown in mustard cotyledons [138]. *De novo* synthesis has been also implicated in the Pfr-mediated induction of soluble fraction of ascorbic acid oxidase in mustard cotyledons by other workers based on comparative density labeling experiments[4].

While some enzymes are stimulated by light, others are repressed. Oelze-Karow and Mohr [98] showed that in case of lipoxygenase, phytochrome results in inhibition of enzyme activity during a particular sensitive phase in mustard seedlings, which has been interpreted by them as a consequence of repression of synthesis of this enzyme at the genetic level. Changes in enzyme synthesis are thus implicated in phytochrome action and it appears that transcriptional changes are involved at some or the other stage.

Some observations on nitrate reductase and peroxidase

Considerable work has been done on the enzyme nitrate reductase, after Jones and Sheard[64] first showed the involvement of phytochrome in light-mediated increase in enzyme activity in etiolated pea buds. An example of such an increase by light is shown by our own work on maize and wheat. In maize a marked increase in enzyme activity was observed within 4 hours after a 5 min R exposure[117]. The increase in enzyme activity was completely inhibited by a subsequent 5 min FR exposure (Fig. 1). A strong inhibition in increase of enzyme activity by both tungstate and inhibitors of RNA and protein synthesis is suggestive of a phytochrome-mediated *de novo* synthesis of enzyme protein. Similar results have been obtained also in wheat[152]. However, a problem in both cases is that there is a distinct lag phase of 1-2 hours, and we do not know what is happening during this time.

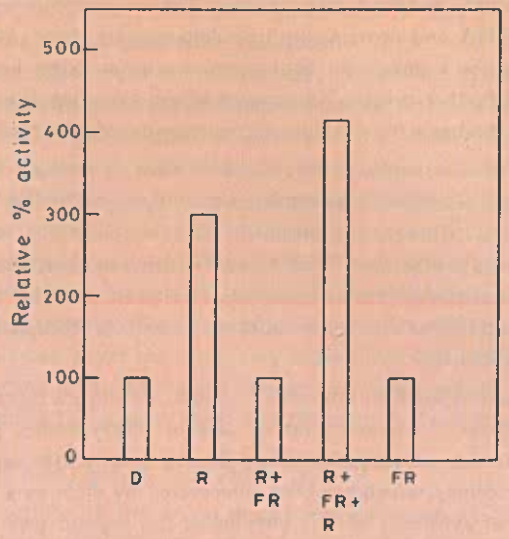


Fig. 1. Photomodulation of nitrate reductase activity in maize seedlings by brief *in vivo* exposures of red and far-red light

Work on phytochrome regulation of peroxidase activity in maize leaves was also initiated in our laboratory[185]. A 5 min R exposure to etiolated leaves resulted in significant increase in its activity, and this increase in enzyme activity could be reversed by subsequent far-red light, if given immediately after red (Fig. 2). Detailed studies using inhibitors for transcription (actinomycin D) and translation (chloramphenicol, puromycin and cycloheximide indicate that the Pfr-mediated increase is not due to increase in mRNA or translation of mRNA on 70S ribosomes—as it was insensitive to both actinomycin D and chloramphenicol. However, the increase could well be due to fresh synthesis of enzyme molecules on the cytoplasmic 80S ribosomes, as the puromycin nullified the increase due to Pfr. Surprisingly, cycloheximide resulted in enhancement of enzyme activity. Again, there is some difficulty in interpreting the results. Possibly, synthesis of an inhibitor may be blocked and Pfr may cause an increase in peroxidase activity not by direct stimulation of synthesis but by inhibiting the degradation of peroxidase.

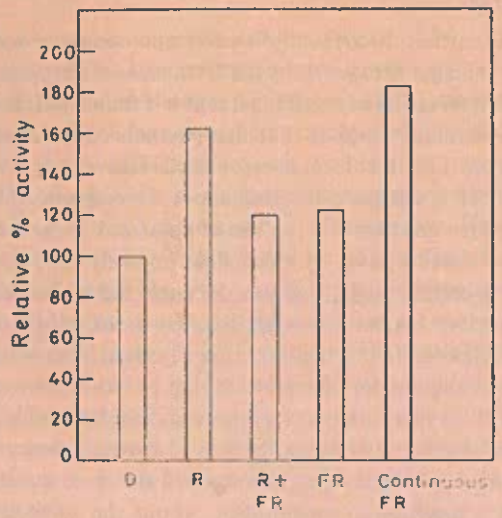


Fig. 2. Photomodulation of peroxidase activity in maize seedlings by brief *in vivo* exposures of red and far-red light

Mode of action—The case for membrane-based control of phytochrome reactions

Although differential gene regulation has been the earliest seriously elaborated theory for explaining the mechanism of action of phytochrome, in most metabolic changes, e.g. increase in enzyme activity discovered in the early phase of phytochrome research, it should be noted that in almost all instances there is a considerable lag period between time of onset of irradiation and actual response. It is not only possible, but very probable therefore that some other primary action(s) may have preceded phytochrome regulation of gene expression and consequent modulation of activity of the enzymes.

It is evident that best chance of identifying the primary action is to investigate such changes that take the minimum of time lag after the photoconversion of phytochrome. And, indeed, much new ground has been broken in recent years by the discovery of a number of so-called 'rapid' or 'quick' responses to red and far-red light.

Rapid responses

Perhaps the earliest discovered of these responses is the reorientation of chloroplast of the alga *Mougeotia* by the German botanist, Werner Haupt[49]. By microbeam irradiation of single algal cells in a filament to R and FR light, Haupt and co-workers[51] showed that, first phytochrome is located in plasma membrane rather than in chloroplasts or nuclei since irradiation of only the periphery of a cell is adequate to bring about the response. Further confirmation of this idea has come by the use of polarized R and FR light. To achieve the maximal response by either light not only the polarizer needs to be oriented at a certain angle, but also the angle has to be changed by 90° in order to achieve maximal reversibility. Obviously, this means that the pigment molecules reside in a highly oriented fashion in membranes and they undergo some change in conformation during phototransformation.

In the last 10-15 years many more responses have been added to this class. Etzold[29], who observed directional growth of filamentous germ tubes of the fern *Dryopteris* in polarized light, also reached the same conclusion, i.e. the photoreceptor is localized in membranes. About the same time came the discovery of phytochrome control of leaf movement of *Mimosa pudica* by Fondeville *et al.*[34] These effects are detectable within five min of irradiation and are thus too quick to involve the various steps of gene activation. On the other hand, since these movements are known to be due to turgor changes in motor pulvinal cells, they provide strong support to the idea of a change in permeability of membranes to ions. It is on the basis of these studies that Hendricks and Borthwick[54] formally proposed the theory of phytochrome control of membrane function. Strong supporting evidence for a membrane based action of phytochrome has come in the last decade also from the work of Galston and Satter[40] on *Albizzia julibrissin* and *Samanea saman*, wherein, too, pulvinal movements similar to those of *Mimosa* occur in response to R and FR irradiation.

Changes in bioelectric potentials

The fastest known phytochrome response, however, occurring at the physico-chemical level, is of a change in bioelectric potentials across membranes, e.g. root tips[58] and in etiolated coleoptiles[97]. A 10-second R exposure to oat coleoptiles, after a lag phase of only 15 s, results in development of 5 to 10 mV of potential and which is FR reversible. A change in surface potential had also been inferred earlier from the well-known reversible 'Tanada

Effect'[146], i.e. R/FR adhesion of root tips to a negatively charged glass surface. This response is detectable within only 30 s of Pfr formation.

The changes in bioelectric potentials must obviously be due to changes in ion transport across membranes. One is not quite sure as to which ion is responsible, but, for pulvinal movements bringing about closure of leaflets in *Albizzia julibrissin* and *Samanea saman*, Galston and co-workers[40] have collected an impressive body of evidence (employing electron probe microanalysis) that changes in K^+ content underlie the response. However, Haupt's group has recently provided evidence that for chloroplast reorientation flux in Ca^{2+} ion may be more important in at least the alga, *Mougeotia*[51]. It may be pointed out that all these rapid responses are insensitive to inhibitors of protein and RNA synthesis. Thus, the idea that repression and derepression of genes might be the primary mode of action now seems difficult to accept in light of the numerous rapid responses known. On the other hand it is logical to ascribe these effects to changes in the membrane properties and, thus, it is inferred that Pfr functions by modulating one or more of membrane functions.

Control of other membrane properties and specific proposals for nature of changes in membranes

As regards more detailed knowledge about the changes in membrane properties, as an alternative amplification mechanism to differential gene activation, a large number of experimental observations have been made during the last decade. Reversible changes in pelletable phytochrome indicating migration of phytochrome to and from membranes depending on whether phytochrome is present in Pr or Pfr form (these are discussed in greater detail in the following section), a R/FR reversible change in conductance when phytochrome is included in model lipid membranes[124], and accumulation of Ca^{2+} in the presence of ATP in isolated membrane fractions[90], all seem to strongly support the view that Pfr—the active form in the membranes—can affect membrane properties.

In terms of functional properties, there are two hypotheses by which involvement of membranes as the primary site of action has been explained. According to one hypothesis, phytochrome may directly modulate a function like transport through membranes. Regarding the mechanism by which phytochrome could regulate it, no definite conclusion can be drawn. The control could either be specific or general. In 1970, Smith[139] favoured the

former idea, i.e. transport of a single specific metabolite across the membrane which could result in changes in cell metabolism. On the basis of their work on pulvinal movements, Satter and Galston[129] suggested that phytochrome might be controlling a specific electrogenic ion pump (for K^+ as discussed above). Subsequent work, however, has shown the involvement of several diverse molecules—evidence which goes against the former idea indicating strongly that pigment may be controlling membrane properties in a more 'general' manner. Thus, phytochrome has been shown to control release of gibberellins in isolated etioplasts preparation[30], changes in transport of NAD^+ (in isolated mitochondrial preparation, Furuya and Manabe[86], and accumulation of Ca^{2+} (as discussed above). Smith[141] now favours a more flexible scheme, i.e. Pfr formation may control transport of more than one substance across the membrane.

According to the other hypothesis phytochrome may function through formation of some kind of a secondary messenger—akin to that present in animal cells. In animal systems, cAMP—a secondary messenger—results in a cascade of reactions regulating control of several metabolic activities. It is quite possible that formation of a secondary messenger in phytochrome-mediated responses too may in turn regulate diverse cellular activities. Way back in 1970 Jaffe[59], first proposed the idea that a neurotransmitter acetylcholine (ACh), might be serving as secondary messenger, since R exposure within 4 min caused a rapid efflux of ACh which was FR reversible. However, with the reports contradicting the possible involvement of ACh by Mohr and his co-workers[70] and Tanada[147], it seemed quite unlikely that this controversial compound, ACh, might be acting as secondary messenger. And, in 1975, Janistyn[60] and Drumm reported rapid changes in contents of yet another important compound, cAMP in mustard seedlings under continuous far-red light. It seems that photomodulation of enzyme resulting in changes of cAMP levels *may be* one of the primary actions of Pfr and also that cAMP might be acting as a secondary messenger.

In conclusion, at this point, it can merely be said with certainty that phytochrome does alter membrane properties or function but to determine in what specific manner these are accomplished—through membrane transport of one or more substances or through secondary messenger or both—remains a major task for future.

Localization—Binding to Membranes

We wish in this section to discuss localization, since phytochrome owes its action to alteration of some property of the membrane. Thus, its localization in the membrane, and some changes as a result of R and FR treatment should be possible to demonstrate experimentally. There has been several studies on phytochrome distribution, which can be grouped under two heads : (a) *in situ*, i.e. at the cellular level by both microspectrophotometric (Galston, 1968) and immunocytochemical methods[110] and (b) cell fractionation, i.e. in cell-free extracts.

In situ localization

Direct microspectrophotometric investigations on the intracellular localization of phytochrome have been attempted, but because of low pigment concentration and unsuitable optical properties of the tissue—such as light scattering and optical screening by other interfering pigments—it has been of limited use[88]. Galston[88] employed this technique to localize the pigment in cells of etiolated oat and pea seedlings. He interpreted the results to as indicative of an association of phytochrome with the nucleus. Marmé and Schäfer[92] using saturating doses of linearly polarized red and far-red irradiations also attempted to localize phytochrome spectrophotometrically. Corn coleoptiles were irradiated with polarized red and far-red light vibrating parallel or normal to the longitudinal axis. Both the irradiations convert 15-20% more phytochrome when vibrating parallel to the axis. They concluded that at least a fraction of the phytochrome population is located at or associated with the plasma membrane and could probably be oriented along a screw line around the cell cylinder.

Recently, Pratt and co-workers[111], by using peroxidase labelled antibodies against phytochrome, have made considerable progress in understanding of intracellular distribution of phytochrome. They have shown that the Pr form possesses no discernible, unique association with any single organelle or membrane. The results are consistent with the possibility that the Pr is largely present as a soluble protein in the cytoplasm of etiolated plant tissues. On the contrary, Pr upon photoconversion to Pfr form appears to migrate—from a relatively uniform distribution to—discrete but as yet unidentified areas. The movement is very rapid and detectable within only 2 min at 3 °C. However, the immunochemical assays do not answer

whether the detected antigens are biologically active and also do not discriminate between Pr and Pfr.

Thus, immunocytochemical studies seem to rule out a unique association of phytochrome with any one subcellular component. However, they do show that phytochrome, upon photoconversion, does change its intracellular distribution.

Cell fractionation studies

The standard extraction and fractionation procedures have also been employed to investigate possible association of the pigment with a subcellular fraction(s). The pigment has been shown by numerous workers to be associated with membranous sedimentable material, at least in the Pfr form, although the exact identity of membranous material remains unsettled. It has been claimed to be associated with mitochondria[48,82,83], plasmalemma[91], endoplasmic reticulum[160] and etioplasts[80]. One must admit that the whole technique of fractional centrifugation has some serious limitations like the purity of subcellular fractions and co-precipitation of phytochrome with the organelles[98]. Nevertheless, the first point to consider is whether in various plant extracts appreciable levels of phytochrome are really bound to sedimentable material.

Indeed, two different types of subcellular associations have been observed. The first is the small proportion of the total phytochrome pool that is always associated with particulate material as reported by Rubinstein *et al.*[125], Evans and Smith[80] and Furuya and Manabe[86]. This type of association has not received much attention because, firstly, the small quantity of phytochrome is too near the limit of spectral assay. Secondly, it can arise due to contamination and its significance is therefore difficult to interpret[107]. In contrast, a second type of phytochrome pelletability—induced by irradiation—involves a large proportion of the phytochrome pool. It can be easily measured spectrophotometrically and seems to be of some significance.

Pelletability induced by in vivo irradiations

The latter type of phytochrome association with particulate material, just referred to above, was first reported by the Freiburg and Harvard groups[89,116]. They demonstrated that an appreciable amount of pigment

binds to particulate material in extracts obtained from red-irradiated corn coleoptiles and squash hypocotyls.

After these initial reports, phytochrome pelletability has been investigated intensively in a number of laboratories and in at least 10 other genera, namely pea[83], oat[45], mustard, sunflower, radish, soybean, bean, lupine, rye[112] and barley[121], enhancement in pelletability similar to that reported for maize and squash has been shown following R exposure of tissue. In the best cases, enhancement up to 10-fold and involving as much as 80% of the total phytochrome pool has been demonstrated.

The plants which have been studied in greater detail are maize, oat, pea and squash. An examination of the requirements for observing pelletability induced by *in vivo* exposures[45,112] indicates that it is relatively insensitive to extraction conditions, which is in contrast to the situation wherein pelletability is induced by *in vitro* exposures (which is discussed later). However, the presence of a divalent cation (Mg^{2+} or Ca^{2+}) in extraction medium is important for an association of phytochrome with the particulate material; addition of these ions *after* extraction is not effective in increasing binding although it results in aggregation of subcellular debris[112].

Before going into the discussion of experiments on pelletability induced by *in vitro* exposures, however, we would like to invite attention to one important fact. Reversal of R-induced pelletability (in the normal protocol a 3 min R exposure is followed by a 5 min FR exposure and then the tissue is immediately homogenized) has been achieved only in squash, and, here too success is only partial[6]. In the other two systems, namely oat[112] and maize[78,115] no reversibility could be achieved by the normal protocol. However, if long dark periods are interposed between FR irradiation and homogenization and centrifugation, some release of phytochrome could be obtained[112,116]. The release process appears to follow a first order kinetics with at $\frac{1}{2}$ between 25 and 50 min in these materials. Anyhow, the failure to obtain R/FR reversibility easily does raise questions about the physiological significance of the binding.

Pelletability induced by in vitro irradiations

In their work on squash, Marmé *et al.*[89] showed that brief irradiations of buffer extracts of dark-grown squash hypocotyl hooks with red light also resulted in Pfr binding to a particulate fraction. They showed that a subsequent FR exposure led to the release of bound molecules, although only

partially, and that too at low concentration of Mg^{2+} ions. This latter observation, i.e. of *in vitro* modulation of pelletability has generated perhaps the maximal interest because if changes in membranes can be shown *in vitro*, this would open up an interesting facet in the biochemical investigation of phytochrome action.

In general, however, results of *in vitro* studies in other systems have been far more disappointing than *in vivo*. To date enhancement of pelletability by *in vitro* R exposure has been demonstrated in only four other plants: maize[113], pea[161], mustard[112] and barley[121,142]. Nevertheless, squash remains the only plant which has been investigated in detail and where partial reversibility by FR has been demonstrated[91,114]. In comparison to *in vivo* experiments, the *in vitro*-induced pelletability is more susceptible to pH and ionic conditions. At pH 7·0, red light irradiation of squash extract shows up to 40% pelletability compared to only 10% or so in unirradiated control[89], but it can be increased to 80% by lowering the pH of extraction buffer to 6·5. The optimal concentration of Mg^{2+} for binding is the same, i.e. 10 mM for both *in vivo* and *in vitro*-induced pelletability, although the range in which binding can be observed in the latter case is more restricted.

The question of physiological significance of pelletability and relevance of present work on cabbage and cauliflower

The failure to obtain far-red reversal of red light-induced pelletability whether *in vivo* or *in vitro* has raised serious doubts in the literature concerning the real significance of the pelletability data discussed in earlier pages (*see* also Pratt[107]). Since far-red completely reverts the physiological effects mediated by red light, theoretically, one should expect similar results in pelletability experiments. This is obviously not the case and specificity of such phytochrome-membrane interactions has therefore been questioned. In squash, pelletability of phytochrome in a large measure has been shown to be due to electrostatic adsorption to 31S degraded ribonucleoprotein (RNP) particles[114]. Either one must postulate that phytochrome could act through its binding to RNP particles (rather than membranes) or one must ignore such interactions altogether. The results of *in vitro* studies are also particularly disappointing because frequently FR exposure after R increases the percentage of bound or pelletable phytochrome rather than decreasing it. Such paradoxical results have been obtained even by *in vivo* exposure in

certain plants, e.g. maize[116] and pea[88], and so far no satisfactory explanation is available for resolving this difficulty.

It is in this connection that our studies on cabbage and cauliflower are of special interest. We first describe our results obtained by *in vivo* exposures. A 5 min exposure of R to tissues (*in vivo*) resulted in about 4 to 5-fold increase in pelletable phytochrome, in both cabbage and cauliflower[81]. The enhanced phytochrome binding to the presumed membrane fraction could be completely reserved by a subsequent FR exposure given after first R. And, interestingly, a red exposure given immediately even after the first R/FR cycle was equally effective in increasing phytochrome pelletability.

As to the results obtained *in vitro*, i.e. by exposures of buffer extracts to red light, of the two systems being investigated in our laboratory—cabbage and cauliflower—in the latter we did not obtain any enhancement in phytochrome pelletability. However, in cabbage, a red exposure to extracts resulted in 4 to 5-fold increase in pelletable phytochrome. Interestingly, the effect of red light was completely reversed by a subsequent far-red treatment (Fig. 3). To our surprise, the binding effect could also be brought about by a terminal red exposure after the first R/FR cycle[81,127].

These results with cabbage therefore constitute the first example of an experimental system wherein photomodulation of phytochrome binding has been demonstrated subsequent to both *in vivo* and *in vitro* light treatments. They strongly suggest that the binding of phytochrome—at least of the Pfr form—to membranes is not artifactual and may have some real significance in terms of its physiological action. Indeed, phytochrome in both forms may normally reside in the membranes. We interpret our release data by FR irradiation to show that phytochrome does change its conformation and the binding may now be relatively more loose—perhaps it may assume a more peripheral position. The release, although demonstrated *in vitro* by our experiments, need not necessarily mean that Pfr molecules are altogether released from the membranes—this could also happen because of severity of the grinding process or/and altered ionic environment. Of course, the total release of Pfr form is supported by the results of Pratt and co-workers also (*see* Pratt[117]). However, it may be instructive to recall in this connection the studies of Haupt[60] on *Mougeotia* as well as of Marmé and Schäfer[92] on corn coleoptiles favouring the idea that both forms of phytochrome may be associated with membrane because the polarized light is more effective in bringing

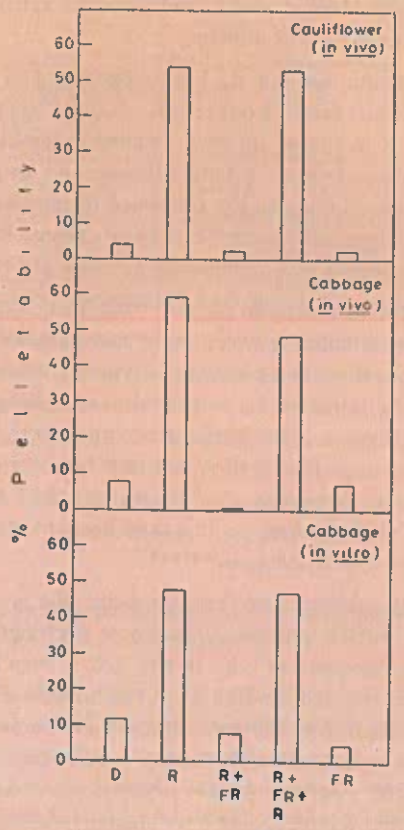


Fig. 3. Photomodulation of phytochrome pelletability in cauliflower and cabbage shoots by brief *in vivo* and *in vitro* exposures of red and far-red light

about phototransformation of phytochrome in both directions, i.e. from Pr to Pfr as well as Pfr to Pr.

Can there be a unifying hypothesis for phytochrome action ?

If we accept the idea that phytochrome resides in the membranes, as seems likely in view of the mounting evidence, a question naturally arises as to whether there is any possibility of reconciling the membrane hypothesis

with that of regulation of gene activity. Perhaps, at this point one should take into account also the recent evidence for direct activation of several enzymes and some of the important ones being, NAD kinase[149], nitrate reductase[62], peroxidase[102], ATPase[66,67], adenylate kinase and hexokinase[68]. Kinetic and other studies show that in these cases, the effects are too rapid following exposures to light and also in several cases density labeling experiments fail to provide evidence for *de novo* protein synthesis. There is, thus, a pressing need for construction of a hypothesis for a basic amplification mechanism that can explain all these diverse actions of the pigment.

Let us take the hypothesis of direct modulation of gene activity, first. This could be accommodated easily within the framework of membrane, i.e. if phytochrome is present in the nuclear membrane. Nevertheless, since RNA polymerase—as far as we know—is localized in the nucleoplasm rather than in the membrane, the possibility of a direct modulation of gene activity in *sensu stricto* seems remote, but that of a second messenger seem very likely. At present one cannot make a guess with any degree of certainty as to the identity of the second messenger. But this could be a simple ion like Ca^{2+} as proposed by Haupt and Weisenheel[51], Weisenheel and Ruppert[158] and Dreyer and Weisenheel[34] or acetylcholine[60], cAMP[140] or some other organic molecule (Fig. 4). Such molecules could then modulate the activity of different enzymes within the cell or in the nucleus.

We now look at the mechanism of activation of enzyme such as those listed on the preceding page. In case of NAD kinase[161], it had been reported that phytochrome co-purifies with the enzyme. It is just possible that there may be some form of an association. The mechanism of the enzyme modulation can easily be explained if it occurs in the membrane. The same holds true for nitrate reductase. The question of intracellular localization of nitrate reductase has been debated and it is known that although at the standard pH of the extraction buffer (i.e. 7·0) it is soluble, at lower pH values the same enzyme is largely found in a sedimentable fraction. However, if nitrate reductase is really present in the cytoplasm, then the existence of a second messenger analogous to cAMP seems likely. Indeed, release of a second messenger immediately following photoconversion of phytochrome molecules to the Pfr form would explain a large number of responses to light in plants.

Of course, if the enzyme resides in the membrane itself, e.g. ATPase or peroxidase or adenylate kinase, then there is no conceptual difficulty in

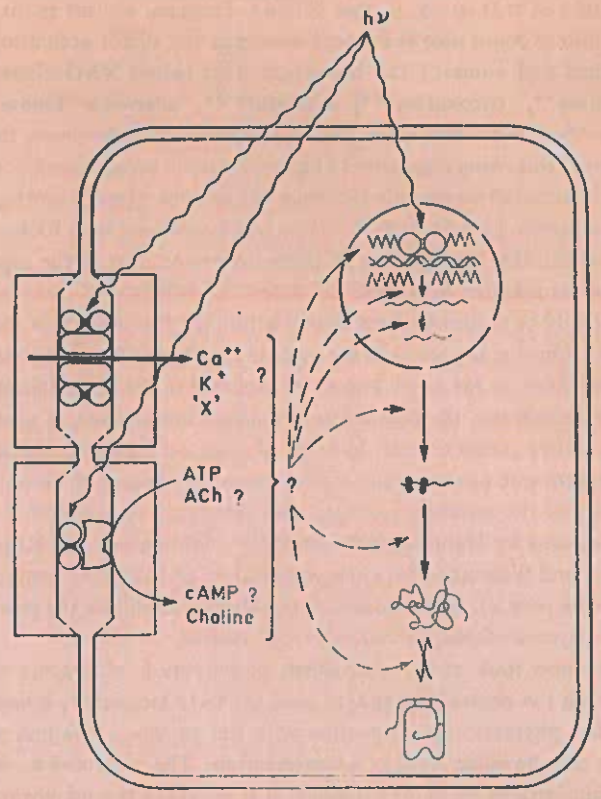


Fig. 4. A hypothetical model of phytochrome action to show its control through both, changes in membrane properties and in gene activity. It should, however, be noted that phytochrome pigment may be localized in the nuclear membrane rather than being associated with chromosomal proteins, in addition to be an integral part of other membranes, e.g. plasma membrane

explaining the photomodulation of these enzymes. Either, the phytochrome molecules are directly juxtaposed to the enzyme molecules or the conformational change during Pfr formation is such that it leads to a general purturbation extending laterally in the entire membrane.

The blue light photoreceptor and its mechanism of action

In recent years yet another pigment system is being involved in the control of growth and morphogenesis. This pigment system senses blue light (400-450 nm) and unlike phytochrome does not show reversibility. Its occurrence in lower organisms like *Neurospora crassa* and *Dictyostelium discoideum* is well established and the pigment has been found to be a flavin[96,103]. Absorption of blue light by the flavin brings about the reduction of a cytochrome which gets reoxidised in darkness. In *Dictyostelium* the photoreducible cytochrome also has been purified and characterized—the molecular weight and absorption pattern indicate that it is a *b* type cytochrome[84]. Unfortunately, in *Dictyostelium* a physiological role cannot yet be assigned for this new photoreceptor, although in *Neurospora* blue light is known to control conidiation by disturbing in some manner the circadian rhythm of conidiation. The pseudoplasmodia of *Dictyostelium* do show phototactic movement, but the pigment responsible seems to be a different one—called "phototaxin" has a porphyrin group with absorption peak extending into the green region[104].

In the higher (eukaryotic) green plants, too, an increasing number of biological responses are turning out to be mediated by blue light. The earliest discovered response is phototropism which has been investigated for more than 40 years. However, many other responses have been discovered since then, some of which are anthocyanin synthesis[23,27], callus growth and shoot initiation[154,156] inhibition of hypocotyl growth[57,66,67], inhibition in flowering[100], and retardation in flower opening[128], chloroplast development[79,130], stomatal opening[85], and unrolling of rice leaves[71,145]. At the molecular level, blue light has been reported to divert flow of carbon in photosynthesis towards amino acids[153], enhance photorespiration[105] and activate nitrate reductase[3,64].

However, the nature of the participating pigment(s) mediating the response is not yet known with certainty. A number of attempts have been made to identify the molecule involved and until recently, in all cases, the method for the investigation has been to match the action spectra of the phenomenon under investigation with the absorption spectra of the putative photoreceptor pigments. To take up phototropism first, long ago Johnston[68] reported the occurrence of two peaks in the blue region when he studied the action spectra for phototropism in *Avena*. This observation indicated

that the pigment absorbing blue light must be a yellow compound. Therefore, flavins and carotenoids were considered to be the most likely candidates. Wald and DuBuy[154] proposed that as the *Avena* coleoptile has a large amount of carotenoids concentrated at the tip, they must be the blue light absorbing pigment. Later Bünning[18] also supported the same. This view remained uncontradicted till Galston's work[39] who suggested riboflavin as the possible receptor while working with pea hypocotyls and from here on the long-standing controversy began as to whether the carotenoids or the flavins are the blue light photoreceptors.

The reason why carotenoids and flavonoids both were considered as possible candidates is because both of them show almost similar absorption spectra with minor differences. β-carotene in organic solvents shows three bands in the visible region but no clear peak in the near-UV, whereas flavins in aqueous solutions show one band in the blue region and another at about 370 nm in the near-UV-region. In recent years action spectra of the physiological responses have been extended towards the ultraviolet region, and now, it has become almost certain that blue light photoreceptor is a flavin and the photoresponsive pigment is cytochrome. An important point in favour of flavin acting as the receptor is its property to mediate redox changes (which carotenoids do not have) and its close biochemical association with cyt *b*. However, specifically, which flavin is the photoreceptor and which cytochrome is associated with it is still an unanswered question.

Lately, blue light has been found to affect several other phenomena. Mohr[93] discovered a particularly striking effect on morphogenesis of fern gametophyte which in dark or red light grows as a long filament. The typical two-dimensional growth begins only after blue light is given. Action spectra indicated that flavin is the photoreceptor. Seibert *et al.*[154], who studied growth and shoot initiation in tobacco callus found that the action spectrum resembled the absorption spectra of both flavonoid(s) and carotenoid(s). But, in experiments where they used differential inhibitors of carotenoid and flavonoid synthesis, they have shown that flavin is the blue light photoreceptor. When 2-(4-chlorophenylthio) triethylamine hydrochloride—an inhibitor of carotenoid synthesis—was provided to the growth medium and the cultures incubated in light, the callus grew normally. On the other hand, when the medium was supplemented with an inhibitor of flavoprotein—quinacrine-hydrochloride or lumichrome, the callus failed to grow indicating that flavin is the blue light absorbing pigment.

The suggestion that flavin(s) is the blue light photoreceptor and cytochrome as the photoresponsive pigment is also supported by Briggs and co-workers who have followed the approach of difference spectroscopy. They have reported similarities in the blue light mediated absorbance changes in the membrane fractions of corn [10, 61] and flavin mediated reduction of cytochrome in artificial systems reported by Schmidt and Butler[130]. The idea of a flavoprotein acting as the photoreceptor finds further support from their studies on inhibitors of flavin reduction, like iodide ion and phenylacetic acid. When these inhibitors where added to the extract *in vitro*, the light induced cytochrome reduction was abolished. Inhibitors of mitochondrial electron transport did not affect the response indicating thereby that the cytochrome involved in blue light responses does not belong to the mitochondrial electron transport chain[19].

A critical examination of the available data indicates, therefore, that in higher plants possibly a flavoprotein(s) is the blue light photoreceptor. However, it must be noted that the conclusion has largely been drawn on the basis of action spectra and inhibitor studies. As a matter of fact, there is as yet no good evidence of a correlation between the flavin mediated cytochrome reduction and a recognized physiological response. Further work is required in this area in order to have a total picture of the biology of the blue light response.

So far our preoccupation was with the pigment. But, let us now ask how does blue light work ? This is a very interesting question but unfortunately at this moment we can do little more than speculate. Briggs and co-workers long time ago reported that *Avena* coleoptiles showed a dramatic decrease in the electric potential difference between coleoptile tip and base of the plant within one second of illumination with blue light[11], similar fast changes in biopotentials have also been reported recently by Hartmann with the hypocotyl hooks of *Phaseolus vulgaris*[48]. The rapidity of the response indicates that the pigment might be associated with the membrane. Skarr and Johnson[138] are also of the opinion that blue light pigment is localized in the membrane. These reports seem to show that for both the blue light photoreceptor (s) and the phytochrome, membrane probably is the primary site of action. But the question is what happens after photoreception. Further, does all this work indicate that blue light works through phytochrome or independently of phytochrome ? As in phototropism, blue light must have some effects of its own. However, Tanada[148] has reported that blue light

increases phytochrome content in mung bean hooks, which means there are interacting effects too. How this comes about is not clear but the author has cited three possibilities, i.e. blue light may (a) increase phytochrome synthesis, or (b) bring about the release of 'bound' phytochrome or (c) reverse the degradation of phytochrome. The implication, therefore, is that blue light in some cases works through phytochrome. In milo (*Sorghum vulgare*) too, for anthocyanin synthesis phytochrome can act only after blue light has been shone[97]. These workers have referred to this mechanism as the Obligate Sequential Mechanism.

In contrast, however, other workers think that the two pigments work independently. Recently Gaba and Black[97] reported that inhibition of *Cucumis sativus* hypocotyl (de-etiolated) growth by blue light and red light follow different kinetics. With blue light, inhibition starts within 5 min and reaches a maximum after 35 minutes ; but, when red light is shone on the hypocotyl, inhibition starts only after 5-6 hours and reaches a peak at 10 hours. This would mean that the two pigments have evolved to control independent functions.

Conclusion and Prospects

It is evident that a great deal of progress has been made in understanding control of growth and morphogenesis by light, since the first pigment, i.e. phytochrome, was detected some 20 years ago. Although the detailed mechanism is not known the possibilities are now considerably narrowed and we can design experiments to find out the probable mechanism. If phytochrome modulates the activity of enzymes within the membranes, it would be interesting to prepare liposomes artificially and incorporate both phytochrome and enzymes within the bilipid membranes and study the detailed mechanism of photomodulation. It seems extremely important also to extend the work on peroxidase and ATPase where *in vitro* modulation of activity has already been shown in response to light treatment. The study of bioelectric potentials and identification of ions responsible for development of such potentials is another area of future work. The possibility of phytochrome regulating one or more ion-pumps directly also needs to be looked into more closely.

It is equally likely that phytochrome may modulate the activity of some enzyme such as adenyl cyclase (may be even guanyl cyclase exists in plants as in vertebrate photoreceptors) or acetylcholinesterase and thus effect the

release of a second messenger. The intriguing result of Smith and co-workers on the release of gibberellins as a result of phytochrome transformations also needs further investigation.

An almost totally untouched facet of phytochrome action is also its relationship to circadian rhythms.

Finally, the detailed investigations having just begun on the blue light photoreceptor, a host of questions have arisen and which need to be answered. Why is there an additional pigment? What is its mechanism of action? Is there any relationship between this pigment and phytochrome action? How does reduction of cytochrome bring about a physiological response ? All these questions offer almost limitless and challenging opportunity of future investigations.

REFERENCES

1. Acton, G. J. and Schopfer, P. (1974) *Biochem. J.*, **142** : 449.
2. Anderson, G. P., Jenner, E. L. and Mumford, F. E. (1970) *Biochim. Biophys. Acta*, **221** : 69.
3. Aparicio, P. J., Roldan, J. M. and Calero, F. (1976) *Biochem. Biophys. Res. Commun.*, **70** : 1071.
4. Attridge, T. H. (1974) *Biochim. Biophys. Acta*, **362** : 258.
5. Balange', A. P. and Rollin, P. (1979) *Physiol. Ve'g.*, **17** : 153.
6. Boisard, J., Marme', D. and Briggs, W. R. (1974) *Pl. Physiol.*, **54** : 272.
7. Borthwick, H. A., Hendricks, S. B. and Parker, M. W. (1952a) *Proc. Natl. Acad. Sci. USA*, **38** : 929.
8. Borthwick, H. A., Hendricks, S. B., Parker, M. W., Toole, E. H. and Toole, V. K. (1952b) *Proc. Natl. Acad. Sci. USA*, **38** : 662.
9. Borthwick, H. A., Hendricks, S. B., Toole, E. H. and Toole, V. K. (1954) *Bot. Gaz.*, **115** : 205.
10. Brain, R. D., Freeberg, J. A., Weiss, C. V. and Briggs, W. R. (1977) *Pl. Physiol.*, **59** : 948.
11. Briggs, W. R. (1976) *In* : Smith, H. (Ed.), Light and Plant Development, p. 7, Butterworths, London.
12. Briggs, W. R. and Rice, H. V. (1972) *Ann. Rev. Pl. Physiol.*, **23** : 293.
13. Bünning, E. (1937) *Planta*, **27** : 583.
14. Burke, M. J., Pratt, D. C. and Moscowitz, A. (1972) *Biochemistry*, **11** : 4025.
15. Butler, L. C. and Bennett, V. (1969) *Pl. Physiol.*, **44** : 1285.

16. Butler, W. L. and Downs, R. J. (1960) *Sci. Am.*, 203 : 56.
17. Butler, W. L., Hendricks, S. B. and Siegelman, H. W. (1964) *Photochem. Photobiol.*, 3 : 521.
18. Butler, W. L., Norris, K. H., Siegelman, H. W. and Hendricks, S. B. (1959) *Proc. Natl. Acad. Sci. USA*, 45 : 1703.
19. Caubergs, R. J., Goldsmith, M. H. M., Briggs, W. R. (1978) Carnegie Inst., Washington Year Book, p. 356.
20. Correll, D. L., Edwards, J. L., Klein, W. H. and Shropshire, W. Jr. (1968a) *Biochim. Biophys. Acta*, 168 : 36.
21. Correll, D. L., Steers, E., Jr., Towe, K. M. and Shropshire, W., Jr. (1968b) *Biochim. Biophys. Acta*, 168 : 46.
22. Cross, D. R., Linschitz, H., Kasche, V. and Tenenbaum, J. (1968) *Proc. Natl. Acad. Sci. USA*, 61 : 1095.
23. Downs, R. J. and Siegelman, H. W. (1963) *Pl. Physiol.*, 38 : 25.
24. Dreyer, E. M. and Weisenheel, M. H. (1979) *Planta*, 146 : 31.
25. Drumm, H., Brüning, K. and Mohr, H. (1972) *Planta*, 106 : 259.
26. Drumm, H., Elchinger, I., Möller, J., Peter, K. and Mohr, H. (1971) *Planta*, 99 : 265.
27. Drumm, H. and Mohr, H. (1978) *Photochem. Photobiol.*, 27 : 241.
28. Durst, F. and Mohr, H. (1966) *Naturwissenschaften*, 53 : 531.
29. Etzold, H. (1965) *Planta*, 64 : 254.
30. Evans, A. and Smith, H. (1976) *Proc. Natl. Acad. Sci. USA*, 73 : 138.
31. Everett, M. S. and Briggs, W. R. (1970) *Pl. Physiol.*, 45 : 679.
32. Feierabend, J. and Pirson, A. (1966) *Z. Pflanzenphysiol.*, 55 : 235.
33. Flint, L. H. and McAlister, E. D. (1935) *Smithsonian Misc. Coll.*, 94 : 1.
34. Fondeville, J. C., Borthwick, H. A. and Hendricks, S. B. (1966) *Planta*, 69 : 357.
35. Fry, K. T. and Mumford, F. E. (1971) *Biochem. Biophys. Res. Commun.* 45 : 1466.
36. Furuya, M. and Manabe, K. (1976) *In* : Smith, H. (Ed.), Light and Plant Development, p. 143, Butterworths, London.
37. Gaba, V. and Black, M. (1979) *Nature*, 278 : 51.
38. Galston, A. W. (1968) *Proc. Natl. Acad. Sci. USA*, 61 : 454.
39. Galston, A. W. and Baker, R. S. (1949) *Am. J. Bot.*, 36 : 773.
40. Galston, A. W. and Satter, R. L. (1976) *In* : Smith, H. (Ed.), Light and Plant Development. p. 159, Butterworths, London.
41. Garner, W. W. and Allard, H. A. (1920) *Jour. Agric. Res.*, 18 : 553.
42. Garner, W. W. and Allard, H. A. (1923) *Jour. Agric. Res.*, 23 : 871.
43. Gordon, S. A. (1961) *In* : Christensen, B. C. and Buchmen, B. (Eds.), Progress in Photobiology, p. 441, Elsevier, Amsterdam.
44. Graham, D., Grieve, A. M., Smillie, R. M. (1971) *Phytochemistry*, 10 : 2905.
45. Grombein, S., Rüdiger, W., Pratt, L. H. and Marmé, D. (1975a) *Pl. Sci. Lett.*, 5 : 275.
46. Grombein, S., Rüdiger, W. and Zimmerman, H. (1975b) *Hoppe-Seyler's Z. Physiol. Chem.*, 356 : 1709.

47. Gupta, S. and Maheshwari, S. C. (1970) *Pl. Cell Physiol.*, 11 : 83.
48. Hartmann, E. (1975) *Physiol. Plant.*, 33 : 266.
49. Haupt, W. (1959) *Planta*, 53 : 484.
50. Haupt, W. (1970) *Physiol. Vég.*, 8 : 551.
51. Haupt, W. and Weisenheel, M. H. (1976) *In* : Smith, H. (Ed.), Light and Plant Development, p. 63, Butterworths, London.
52. Hendricks, S. B. (1960) *Cold Springer Harbor Symp. Quant. Biol.*, 25 : 245.
53. Hendricks, S. B. and Borthwick, H. A. (1955) *In* : Rudnick, D. (Ed.), Aspects of Synthesis and Order in Growth, p. 149, Princeton Univ. Press.
54. Hendricks, S. B. and Borthwick, H. A. (1967) *Proc. Natl. Acad. Sci. USA*, 58 : 2125.
55. Hillman, W. S. (1967) *Ann. Rev. Pl. Physiol.*, 18 : 301.
56. Hopkins, D. W. and Butler, W. L. (1970) *Pl. Physiol.*, 45 : 567.
57. Hunt, R. E. and Pratt, L. H. (1979) *Pl. Physiol.*, 61 : 332.
58. Jaffe, M. J. (1968) *Science*, 162 : 1016.
59. Jaffe, M. J. (1970) *Pl. Physiol.*, 46 : 768.
60. Janistyn, B. and Drumm, H. (1975) *Planta*, 125 : 81.
61. Jesaitis, A. J., Heners, P. R., Hertel, R. and Briggs, W. R. (1977) *Pl. Physiol.* 59 : 941.
62. Johnson, C. B. (1976) *Planta*, 128 : 127.
63. Johnston, E. S. (1934) *Smithsonian Misc. Coll.*, 92 : 1.
64. Jones, R. W. and Sheard, R. W. (1972) *Nature New Biol*, 238 : 221.
65. Jones, H. W. and Sheard, R. W. (1974) *Can. J. Bot.*, 52 : 1433.
66. Jose, A. M. (1977a) *Planta*, 136 : 125.
67. Jose, A. M. (1977b) *Planta*, 137 : 203.
68. Jose, A. M. and Schäfer, E. (1979) *Planta*, 146 : 75.
69. Jose, A. M. and Vince-Prue, D. (1978) *Photochem. Photobiol.*, 27 : 209.
70. Kasemir, H. and Mohr, H. (1972) *Pl. Physiol.*, 49 : 453.
71. Katsura, N. and Inada, K. (1979) *Pl. Cell Physiol.*, 20 : 1071.
72. Kendrick, R. E. and Frankland, B. (1976) *In* : Phytochrome and Plant Growth, Edward Arnold Limited.
73. Kendrick, R. E. and Spruit, C. J. P. (1973) *Photochem. Photobiol.*, 18 : 139.
74. Kendrick, R. E. and Spruit, C. J. P. (1977) *Photochem. Photobiol.*, 26 : 201.
75. Klein, A. O. (1969) *Pl. Physiol.*, 44 : 897.
76. Kroes, H. H. (1970) *Meded. Landbouwhogesch. Wageningen*, 70 : 1.
77. Kroes, H. H., Van Rooijen, A., Geers, J. M. and Greuell, E. H. M. (1969) *Biochim. Biophys. Acta*, 175 : 409.
78. Lehman, U. and Schäfer, E. (1978) *Photochem. Photobiol.*, 27 : 767.
79. Lichtenthaler, H. K., Buschmann, C. and Rahmsdorf, U. (1979) *In* : Senger, H. (Ed.), International Conference on Effects of Blue Light on Plants and Microorganisms, p. 66.
80. Linschitz, H. and Kasche, V. (1967) *Proc. Natl. Acad. Sci. USA*, 58 : 1059.
81. Maheshwari, S. C., Rustogi, V. and Sihag, R. K. 1980 VIII Int. Cong. Photobiol., France (*Submitted*).

82. Manabe, K. and Furuya, M. (1974) *Pl. Physiol.*, **53** : 343.
83. Manabe, K. and Furuya, M. (1975) *Planta*, **123** : 207.
84. Manabe, K. and Poff, K. L. (1978) *Pl. Physiol.*, **61** : 961.
85. Mansfield, T. A. and Meidner, H. (1966) *J. Exp. Bot.*, **17** : 510.
86. Marcus, A. (1960) *Pl. Physiol.*, **35** : 126.
87. Margulies, M. M. (1965) *Pl. Physiol.*, **40** : 57.
88. Marmé, D. (1977) *Ann. Rev. Pl.Physiol.*, **28** : 173.
89. Marme, D., Boisard, J. and Briggs, W. R. (1973) *Proc. Natl. Acad. Sci. USA*, **70** : 3861.
90. Marmé, D. and Gross, J. (1976) VII Int. Cong. Photobiol., Rome.
91. Marmé, D., Mackenzie, J.M., Jr., Boisard, J. and Briggs, W.R. (1974) *Pl. Physiol.*, **54** : 263.
92. Marmé, D. and Schäfer, E. (1972) *Z. Pflanzenphysiol.*, **67** : 192.
93. Mohr, H. (1956) *Planta*, **47** : 127
94. Mohr, H. (1966) *Photochem. Photobiol.*, **5** : 469.
95. Mohr, H. (1972) Lectures on Photomorphogenesis, Springer-Verlag, Berlin.
96. Muñoz, V. and Butler, W. L. (1975) *Pl. Physiol.*, **55** : 421.
97. Newman, I. A. and Briggs, W. R. (1972) *Pl. Physiol.*, **50** : 687.
98. Oelze-Karow, H. and Mohr, H. (1970) *Z. Naturforsch.*, **25** : 1282.
99. Oelze-Karow, H., Schopfer, P. and Mohr, H. (1970) *Proc. Natl. Acad. Sci. USA*, **65** : 51.
100. Ohatani, T. and Ishiguri, Y. (1979) *Physiol. Plant.*, **47** : 255.
101. Parker, M. W., Hendricks, S. B. and Borthwick, H. A. (1950) *Bot. Gaz.*, **111** : 242.
102. Penel, C., Greppin, H. and Boisard, J. (1976) *Pl. Sci. Lett.*, **6** : 117.
103. Poff, K. L. and Butler, W. L. (1975) *Pl. Physiol.*, **55** : 427.
104. Poff, K. L., Loomis, W. F. and Butler, W. L. (1974) *J. Biol. Chem.*, **249** : 2164.
105. Poskuta, J. (1968) *Experientia*, **24** : 796.
106. Pratt, L. H. (1975) *Photochem. Photobiol.*, **21** : 99.
107. Pratt, L. H. (1978) *Photochem. Photobiol.*, **27** : 81.
108. Pratt, L. H. and Butler, W. L. (1968) *Photochem. Photobiol.*, **8** : 477.
109. Pratt, L. H. and Butler, W. L. (1970) *Photochem. Photobiol.*, **11** : 361.
110. Pratt, L. H. and Coleman, R. A. (1971) *Proc. Natl. Acad Sci. USA*, **68** : 2431.
111. Pratt, L. H., Coleman, R. A. and Mackenzie, J. M., Jr. (1976) *In* : Smith, H. (Ed.), Light and Plant Development, p. 75, Butterworths, London.
112. Pratt, L. H. and Marmé, D. (1976) *Pl. Physiol.*, **58** : 686.
113. Quail, P. H. (1974) *Planta*, **118** : 357.
114. Quail, P. H. (1975) *Planta*, **123** : 235.
115. Quail, P. H. and Briggs, W. R. (1978) *Pl. Physiol.*, **62** : 773.
116. Quail, P. H., Marmé, D. and Schäfer, E. (1973) *Nature New Biol.*, **245** : 189.
117. Rao, V. M., Datta, N., Sopory, S. K. and Guha-Mukherjee, S. (1980) *Physiol. Plant* (*in press*).
118. Rice, H. V. and Briggs, W. R. (1973) *Pl. Physiol.*, **51** : 927.

119. Rice, H. V., Briggs, W. R. and Jackson-White, C. J. (1973) *Pl. Physiol.* **51** : 917.
120. Richter, G. (1979) *In* : Senger, H. (Ed.), International Conference on Effects of Blue Light on Plants and Microorganisms, p. 60.
121. Roth-Bejerano, N. and Kendrick, R. E. (1979) *Physiol. Plant.*, **47** : 67.
122. Roux, S. J. and Hillman, W. S. (1969) *Arch. Biochem. Biophys.*, **131** : 423.
123. Roux, S. J., Lisansky, S. G. and Stoker, B. M. (1975) *Physiol. Plant.* **35** : 85.
124. Roux, S. J. and Yguerabide, J. (1973) *Proc. Natl. Acad. Sci. USA*, **70** : 762.
125. Rubinstein, B., Drury, K.S. and Park, R. B. (1969) *Pl. Physiol.*, **44** : 105.
126. Rüdiger, W. (1972) *In*: Mitrakos, K. and Shropshire, W., Jr. (Eds.), Phytochrome, p. 129, Academic Press, London.
127. Rustogi, V., Sihag, R. K. and Maheshwari, S. C. (1980) *Planta* (*in press*).
128. Saito, M. and Yamaki, T. (1967) *Nature*, **214** : 1027.
129. Satter, R. L. and Galston, A. W. (1973) *Bioscience*, **23** : 407.
130. Schmidt, W. and Butler, W. L. (1976) *Photochem. Photobiol.*, **24** : 77.
131. Schopfer, P. (1977) *Ann. Rev. Pl. Physiol.*, **28** : 223.
132. Schopfer, P. and Mohr, H. (1972) *Pl. Physiol.*, **49** : 8.
133. Schopfer, P. and Plachy, C. (1973) *Z. Naturforsch.*, **28** : 296.
134. Seibert, M., Wetherbee, P. J. and Donald, D. (1975) *Pl. Physiol.*, **56** : 130.
135. Sharma, R., Sopory, S. K. and Guha-Mukherjee, S. (1976) *Pl. Sci. Lett.*, **6** : 69.
136. Siegelman, H. W. and Firer, E. M. (1964) *Biochemistry*, **3** : 418.
137. Siegelman, H. W. and Hendricks, S. B. (1965) *Fed. Proc.*, **24** : 863.
138. Skarr, H. and Johnson, A. (1978) *Physiol. Plant.*, **43** : 390.
139. Smith, H. (1970) *Nature*, **227** : 665.
140. Smith, H. (1975) Phytochrome and Photomorphogenesis, p. 54, McGraw-Hill Book Company.
141. Smith, H. (1976) *In* : Smith, H. (Ed.), Light and Plant Development, p. 493, Butterworths, London.
142. Smith, H., Evans, A. and Hilton, J. R. (1978) *Planta*, **141** : 71.
143. Song, P. S. and Chae, Q. (1979) *Photochem. Photobiol.*, **30** : 117.
144. Surrey, K. (1967) *Pl. Physiol.*, **42** : 421.
145. Takahashi, N. (1964) *Bull. Inst. Agric. Res. Tohoku Univ.*, **15** : 185.
146. Tanada, T. (1968) *Proc. Natl. Acad. Sci. USA*, **59** : 376.
147. Tanada, T. (1972) *Pl. Physiol.*, **49** : 860.
148. Tanada, T. (1977) *Physiol. Plant.*, **41** : 146.
149. Tezuka, T. and Yamamoto, Y. (1974) *Pl. Physiol.*, **53** : 717.
150. Tobin, E. M. and Briggs, W. R. (1973) *Photochem. Photobiol.*, **18** : 487.
151. Van Poucke, M., Barthe, F. and Mohr, H. (1969) *Naturwissenschaften*, **56** : 417.
152. Vijayaraghavan, S. J., Sopory, S. K. and Guha-Mukherjee, S. (1979) *Pl. Cell Physiol.*, **20** : 1251.
153. Voskresenskaya, N. P., Grishina, G. S., Chmora, S. N. and Poyarkova, N. M. (1970) *Can. J. Bot.*, **48** : 1251.
154. Wald, G. and DuBuy, H. G. (1936) *Science*, **84** : 247.
155. Walker, T. S. and Bailey, J. L. (1970) *Biochem. J.*, **120** : 613.

156. Ward, H. B. and Vance, B. D. (1968) *J. Exp. Bot.*, 19 : 119.
157. Weidner, M., Jakobs, M. and Mohr, H. (1965) *Z. Naturforsch.*, 20 : 689.
158. Weisenheel, M. H. and Ruppert, H. K. (1977) *Planta*, 137 : 225.
159. Whitelam, G. C., Johnson, C. B. and Smith, H. (1979) *Photochem. Photobiol.*, 30 : 589.
160. Williamson, F. A., Moore', D. J. and Jaffe, M. J. (1975) *Pl. Physiol.*, 56 : 738.
161. Yamamoto, K. T. and Furuya, M. (1975) *Planta*, 127 : 177.
162. Yu, R. (1977) *J. Theor. Biol.*, 69 : 581.

Wave Forces on Offshore Structures

S. Narasimhan

Indian Institute of Technology, Bombay

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

1. Introduction

The consumption of petroleum products in India has gone up steeply from about 4 million tonnes in 1951 to about 22·7 million tonnes in 1974. The rate of crude production onshore, though it has registered a rise from a bare 270,000 tonnes in 1951 to 9 million tonnes in 1976, has continuously lagged behind consumption. Since it is likely that onshore production may level off at 9 million tonnes per annum, it is indeed extremely fortunate for the country to have found the giant oil reserves of the Bombay High which has an estimated reserve of 800 million tonnes from which a production rate of 18 million tonnes per annum is anticipated. Recent investigations have revealed potential oil fields at other locations such as the offshore areas near Ratnagiri, Godavari basin and Andamans.

Until recently, harbours and coastal morphology formed the main topics associated with conventional coastal engineering. In recent times man's interest in working at sea has increased rapidly in view of the hope to utilise, to the fullest extent, the food, mineral and oil potential of the sea in near-shore and offshore regions to meet the ever-increasing demand. Especially, this offshore venture, so critical to the nation's economy, is perhaps the most exciting challenge to the scientists, engineers and technologists alike in the recent times. The planning, design, erection and operation of our off-shore facility (situated in a hostile and complicated environment of winds, waves, tides and currents and located in great depths and founded on marine deposits)call for an inter-disciplinary association of physical oceanographers, geologists, marine scientists, and mechanical, petroleum and civil engineers in the wide spectrum of activities linked with off-shore engineering and technology. In the design of an off-shore structure, perhaps the least understood is the evaluation of the environmental loads on the structures, and in this article an attempt is made to review the methods adopted for describing wave forces.

2. Wave Climate

Wind-generated waves are the most common type of waves in the ocean, natural lakes and artificial reservoirs though other agencies like earthquake and submarine landslide may also cause waves. One of the striking features

of these wind-generated wave systems is their irregularity or randomness of occurrence both with respect to space and time. They also appear to move in groups and these group movements are also irregular. Although the waves appear to be travelling in some sort of an average direction, each wave (or groups of waves) appears to be travelling in slightly varying directions different from the one either preceding or succeeding it. The engineer often finds himself in the complex task of defining the wave climate of the region and describe the statistical properties of the waves so that he can prescribe the characteristics of the design wave in terms of its height, period and energy. A design wave spectrum will describe more appropriately the random components of the wave system from which the engineer can derive many useful design parameters for any coastal or offshore structure. The wave climate can be obtained in a number of ways like hindcasting by means of synoptic charts or by the use of various forecasting relationships like the SMB or PNJ methods. These methods rely on the assumed form of wind and wave characteristics and are approximate. Perhaps the most accurate and reliable technique available today is the measurement and analysis of actual ocean waves observed and recorded at a given site.

3. Deterministic and Non-Deterministic Methods

The study of ocean waves and wave loading is normally made by either a deterministic or a non-deterministic approach. The deterministic approach follows the classical convention of assuming that the waves are regular and periodic. Some of the theories followed are Airy's linear wave theory, Stokes' non-linear higher order theory, Cnoidal theory and Dean's stream function theory. All these theories deal with an ideal fluid and the flow kinematics and subsequent force fields are found by obtaining a solution for velocity potential, φ or the stream function, ψ by solving the boundary value problems. These methods give unique solutions.

As observed earlier, the sea surface is random and fixed solutions given by deterministic theories (though they attempt to replace the random sea by an equivalent wave system in the overall sense containing regular waves) are poor approximations in detail and hence unsuitable for design purposes. Recourse then has to be made to non-deterministic or probabilistic techniques to describe the random wave characteristics over all its range of frequencies. Statistical descriptions of the general wave climate are also possible.

4. Deterministic Wave Theories

4.1. In the linear wave theory, we assume that the amplitude, a (half the wave height, H) of the wave is very small compared to the wavelength ($a \ll L$), and allows us the approximation that the outward normal to the free surface is not very much different from the vertical direction, z. The velocity at the free surface is expressed as a reasonable approximation as,

$$\frac{\partial \eta}{\partial t} = \frac{\partial \varphi}{\partial z} \tag{1}$$

Conditions of continuity for the incompressible fluid and irrotationality give us

$$\frac{\partial^2 \varphi}{\partial x^2} + \frac{\partial^2 \varphi}{\partial z^2} = 0 \tag{2}$$

which is supported by the Bernoulli equation of energy at the free surface, $z = \eta$

$$\frac{\partial \varphi}{\partial t} + g\eta = 0 \tag{3}$$

Since the pressure at the free surface is zero and we may omit the nonlinear term $V^2/2$ in the energy equation as a good approximation. Boundary condition at the level bottom of the ocean gives us

$$\frac{\partial \varphi}{\partial z} = 0 \tag{4}$$

The equation to the free surface is assumed periodic with x and t

$$\eta = a \cos (kx - \omega t) \tag{5}$$

where $k = 2\pi/L$ is the wave number and ω denotes the angular frequency ($\omega = 2\pi/T = 2\pi f$), T is the wave period and f, the frequency. (Definition sketch in Fig. 1.)

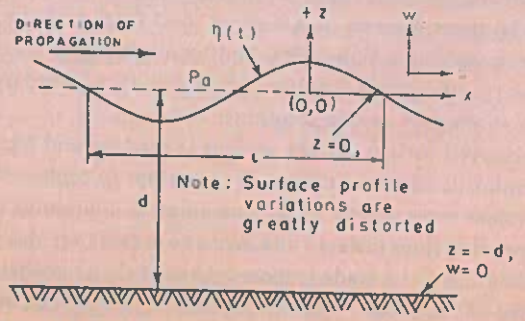


FIG. 1(a) DEFINITION SKETCH: REGULAR WAVE

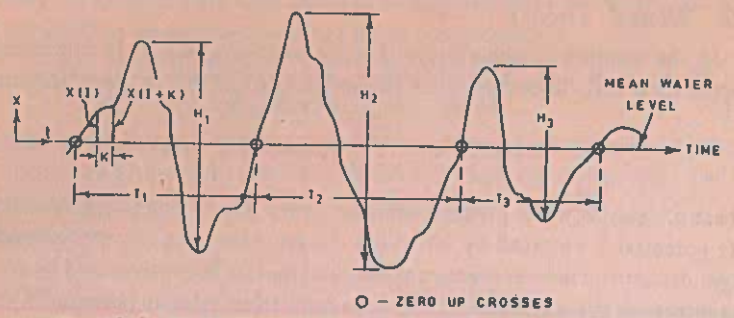


FIG. 1(b) DEFINITION SKETCH: RANDOM WAVES

Equations (1) to (5) are used to derive the value for velocity potential, φ

$$\varphi = \frac{ag}{\omega} \frac{\cosh k(z+d)}{\sinh kd} \sin (kx - \omega t) \tag{6}$$

from which expressions for velocity and acceleration components in the horizontal and vertical direction may be obtained.

$$u = \frac{\partial \varphi}{\partial x} = a\omega \frac{\cosh (kz+kd)}{\sinh kd} \cos (kx - \omega t) \tag{7}$$

$$w = \frac{\partial \varphi}{\partial z} = a\omega \frac{\sinh (kz+kd)}{\sinh kd} \sin (kx - \omega t) \tag{8}$$

$$\frac{\partial u}{\partial t} = +a\omega^2 \frac{\cosh (kz+kd)}{\sinh kd} \sin (kx - \omega t) \tag{9}$$

$$\frac{\partial w}{\partial t} = -a\omega^2 \frac{\sinh (kz+kd)}{\sinh kd} \cos (kx - \omega t) \tag{10}$$

The pressure at any point is a sum of hydrostatic and hydrodynamic pressures.

$$p = -\rho gz + \rho ga \frac{\cosh k(z+d)}{\cosh kd} \cos (kx - \omega t) \tag{11}$$

The celerity of the small amplitude wave is

$$c = \sqrt{\frac{gL}{2\pi} \tanh \frac{2\pi d}{L}} \tag{12}$$

The expressions given above by the linear theory have been used fairly widely in determining the wave forces since the expressions, though approximate, are simple to use.

4.2. *Stokes' Theory*

In the nonlinear higher order theories of Stokes, $\eta(x, t)$ is expressed in a series form with the coefficient of the series as functions of wave amplitude

$$\eta(x, t)=a_0+\sum_{n=1}^{\infty} a_n \cos n(kx-\omega t) \tag{13}$$

where a_n denotes wave profile coefficients that can be computed. The velocity potential is obtained by satisfying the boundary conditions that can be expanded into an infinite series containing its partial derivatives and by applying successive approximations. The wave profile and velocity potential φ given by the second order theory ($n=2$ in eqn. (13)) is obtained with an additional harmonic function to the linear wave

$$\eta=a\cos(kx-\omega t)+\frac{\pi a^2}{2L}\frac{\cosh kd}{\sinh^3 kd}(2+2\cosh 2kd)\cos 2(kx-\omega t) \tag{14}$$

$$\varphi=\frac{ac\cosh k(z+d)}{\sinh kd}\sin(kx-\omega t)+\frac{3}{4}\frac{\pi a^2 c}{L}\frac{\cosh 2k(z+d)}{\sinh^4 kd}\sin 2(kx-\omega t) \tag{15}$$

Consideration of another additional harmonic function in eqn. (13) yields the third order theory in deep water with its free surface given by

$$\eta=\tfrac{1}{2}ka^2+a\cos(kx-\omega t)+\tfrac{1}{2}ka^2\cos 2(kx-\omega t) -\tfrac{3}{8}k^2a^3\cos 3(kx-\omega t) \tag{16}$$

$$\varphi=\frac{CL}{2\pi}[F_1\cosh k(z+d)\sin(kx-\omega t)+\tfrac{1}{2}F_2\cosh 2k(z+d)\times \sin 2(kx-\omega t)+\tfrac{1}{3}F_3\cosh 3k(z+d)\sin 3(kx-\omega t)] \tag{17}$$

where

$$F_1=\frac{2\pi a}{L}\frac{1}{\sinh kd}-\left(\frac{2\pi a}{L}\right)^2\frac{(1+5\cosh^2 kd)\cosh^2 kd}{8\sinh^5 kd}$$

$$F_2=\frac{3}{4}\left(\frac{2\pi a}{L}\right)^2\frac{1}{\sinh^4 kd}$$

$$F_3=\frac{3}{64}\left(\frac{2\pi a}{L}\right)^3\left(\frac{11-2\cosh 2kd}{\sinh^7 kd}\right)$$

4.3. *Stokes' Fifth-Order Theory*

As can be seen from the expressions for flow kinematics in the second and third order theories, as the order of the theory increases so does the number of coefficients involved. Skjelbreia and Hendickson[22] have tabulated these coefficients, the use of which makes the fifth-order theory simple and

reliable. Writing $\theta=k(x-ct)$ the free surface profile to the fifth-order can be obtained by expanding the series up to 5 harmonics :

$$k\eta(t)=\lambda\cos\theta+(\lambda^2B_{22}+\lambda^4B_{24})\cos 2\theta+(\lambda^3B_{33}+\lambda^5B_{35})\cos 3\theta +\lambda^4B_{44}\cos 4\theta+\lambda^5B_{55}\cos 5\theta \tag{18}$$

The coefficients λ and B are functions of hyperbolic sines and cosines. In using the fifth-order theory for design purposes, the design wave has to be specified in terms of H, d and T. Then d/L and λ are calculated by the relationships

$$\frac{\pi H}{d}=\frac{L}{d}[\lambda+\lambda^3B_{33}+\lambda^5(B_{35}+B_{55})] \tag{19}$$

$$\frac{d}{L_0}=\frac{d}{L}\tanh\frac{2\pi d}{L}(1+\lambda^2C_1+\lambda^4C_2) \tag{20}$$

where $L_0=gT^2/2\pi$ which can be determined by design wave dimensions. B_{33}, B_{35}, B_{55}, C_1 and C_2 are functions of d/L only. Knowing d/L and λ we can make use of the tables to calculate the various flow kinematics. The expression for φ is

$$\frac{k\varphi}{c}=(\lambda A_{11}+\lambda^3A_{13}+\lambda^5A_{15})\cosh k(d+z)\sin\theta+ (\lambda^2A_{22}+\lambda^4A_{24})\cosh 2k(d+z)\sin 2\theta+ \lambda^4A_{44}\cosh 4k(d+z)\sin 4\theta+ \lambda^5A_{55}\cosh 5k(d+z)\sin 5\theta \tag{21}$$

4.4. *Cnoidal Theory*

The linear theory is valid in all depths of water but the Stokes' higher order solutions become divergent in shallow water when $d/L < 1/10$. In shallower water, Cnoidal theory is satisfactory. It is generally used when d/L ratio is between 1/10 and 1/50. For $d/L < 1/50$ solitary wave theory is used which is a limiting case of the Cnoidal wave theory. In these theories, the wave profile is expressed in terms of Jacobian elliptic function 'cn u'

$$y_t=y_t+H\,\mathrm{cn}^2\left[2K(k)\left(\frac{x}{L}-\frac{t}{T}\right),k\right] \tag{22}$$

where cn is one of the Jacobian elliptic functions, y_s is the vertical distance from the ocean bed to the free surface, y_t is the distance between the wave trough and the bottom, $K(k)$ is the complete elliptic integral of the first kind of the modulus K. When $k=1$, the period becomes infinite in which case we have the limiting case of a solitary wave.

The celerity of the solitary wave then becomes

$$c = \sqrt{g y_t}\left(1 + \frac{H}{2y_t}\right) \tag{23}$$

Another limiting case is the linear theory for which $k \to 0$ and $E(k)/K(k) = 1$ and

$$c = \left[gd\left(1 - \frac{4\pi^2 d^2}{3L^2}\right)\right]^{\frac{1}{2}} \tag{24}$$

which is an approximate form of eqn. (12).

A full description of the method with expressions for wavelength, velocity, acceleration and pressures, etc., are given by Wiegel[26].

4.5. *Dean's Stream Function Theory*

Dean[7] developed the numerical theory in terms of the stream function ψ rather than the velocity potential, φ which was used in all the earlier theories. The stream function is expressed in the form of a series

$$\psi = \sum_{n=1}^{M} A_n \sinh nkz \cos nkx \tag{25}$$

The series can be expanded to any order and thus becomes increasingly applicable to a very wide range of ocean problems. The coefficients A_n can be calculated numerically. For given values of wave height H, period T and depth d we treat ψ, formulated by the expression (25), as the unknown which is then found by computing the coefficients A_1, A_2, etc., such that the error in the dynamic boundary condition for a number of points on the free surface is minimised. The free surface is modelled as a streamline of constant stream-function value so that its tangent at any point becomes the direction of the velocity vector and the kinematic boundary condition is exactly satisfied. This is the main advantage of the stream function theory over the other theories.

4.6. *Range of Relative Validity of Various Theories*

No precise statements may be made of the range of validity of various water wave theories. The choice between the first or higher order theories lies in the magnitude of the wave heights in relation to the wavelength. A typical chart is shown in Fig. 2 showing the various regions of applicability of the different theories depending on the value of d/H and d/L. In

general, the linear theory is useful in cases of deep water, short-crested waves, broad-banded nature of the wave spectra, swell-dominated conditions, as well as for flexible structures for which dynamic characteristics cannot be overlooked. Nonl-inear theories on the other hand are used in cases of shallow water, directional sea with larger waves and narrow-banded spectra. Dean[8] has shown that his stream function theory which may be extended to any order provides a better fit for almost all wave conditions and involves least error in satisfying various boundary conditions when compared with other theories. However, the linear theory, though less accurate than the higher order ones, has been used extensively in assessing the wave characteristics in ocean engineering. In the design spectrum method of determining

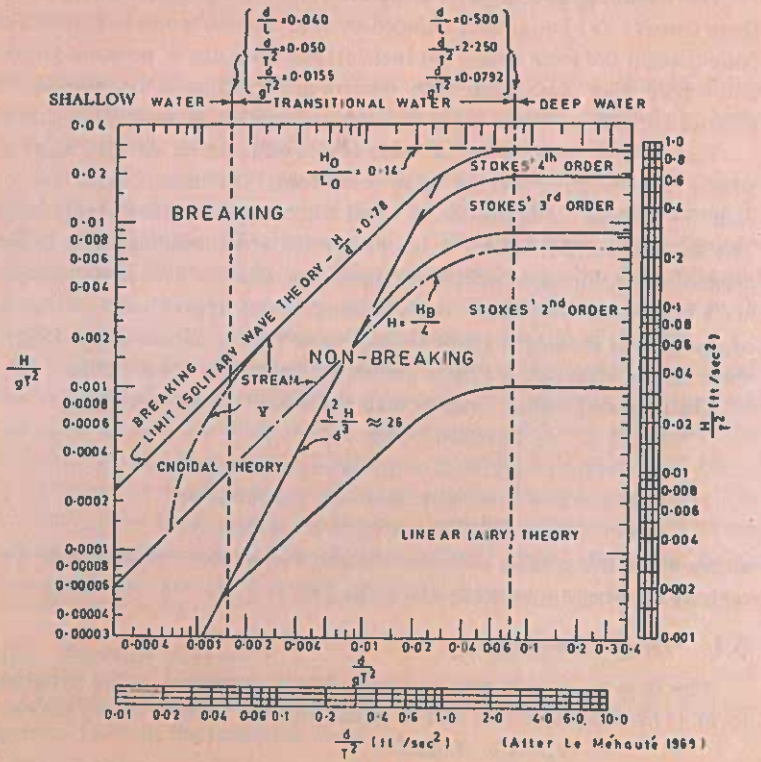


FIG.2 REGIONS OF VALIDITY FOR VARIOUS WAVE THEORIES

the wave forces, the linear theory is used for defining the velocity and acceleration terms. Its popularity, to be sure, is essentially due to its simplicity in both formulation and solution.

5. General Method of Computing Wave Forces

Offshore platforms either rest on a number of small diameter circular piles driven into the ocean bed or on large cylinders (condeep-gravity platforms) resting on the sea bottom. Even floating platforms, tension legs adopt circular geometry not only from the criterion of strength, but also to meet the changing directional approach of waves. The attention is therefore mainly focussed on circular cylinders subjected to wave action.

The wave loads on offshore structures may be classified as arising out of three causes : (a) Drag force induced by skin resistance and flow separation (skin friction and form drag), (b) Inertial drag force due to pressure gradient and added mass associated with relative acceleration of the ambient fluid and (c) Diffraction force due to scattering of incident waves by the structure.

The relative importance of these forces will depend on the nature and size of the structure and the wave conditions. The first of these is significant for tubular components of small diameter in waves of large height. The second (inertial loads) is important for structural members of large dimensions. Diffraction effects are significant when the sectional dimensions are a substantial fraction of the wavelength. An approximate delineation of the various loading regimes are shown in Fig. 3, expressed in terms of wave height, structure diameter and depth below the water surface.

The following values roughly map the various force regimes :

$D/L > 1$ — Pure reflection
$D/L > 0{\cdot}2$ — Diffraction increasingly predominant
$D/W > 0{\cdot}2$ — Inertia increasingly predominant
$D/W < 0{\cdot}2$ — Drag increasingly predominant

where W is the orbital diameter at free surface ($W = H/\tanh kd$). These regimes are diagrammatically shown in Fig. 3.

5.1. *Drag Component*

The drag on a body kept in steady flow is expressed in the Newtonian form. (This force is in line and in phase with the velocity of the stream) :

$$F_D = C_D \cdot A \cdot \tfrac{1}{2} \rho U^2 \qquad (26)$$

where C_D is the drag coefficient, A refers to a characteristic projected area

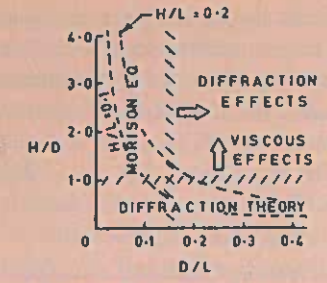


FIG. 3 RANGE OF APPLICATIONS OF WAVE FORCE THEORIES (Zienkiwicz et al-1978)

of the body, ρ is the mass density of the fluid and U the undisturbed free stream velocity. The value of C_D is dependent on the shape of the body, Reynolds number $\rho UD/\mu$ and roughness of the surface.

The force defined above takes into account both the skin-friction drag and the form drag but in situations where the body pierces the free surface and the surface profile is altered by the presence of the body, a wave drag component may also arise. When a pile is inclined at an angle α and to the horizontal, the total force may be resolved into two components, F_N and F_T in the normal and tangential directions with respect to the axis of the pile.

It is assumed that the same form of the drag equation is valid for unsteady flow also though this assumption is debatable. It is obvious that the wake formation behind the body during one direction of flow is going to affect the flow pattern and pressure distribution when the flow is reversed. Hence, the value of C_D varies over a wide range (unlike the case for steady flow where it can be satisfactorily predicted). The value of the velocity to be used may be chosen from any of the theories outlined earlier depending on the wave conditions and the accuracy required.

5.2. *Inertial Force*

Morison[18] was perhaps the first to suggest the widely accepted form for the total force in the following form

$$F = C_D A \tfrac{1}{2} \rho |U| U + C_M (\rho \forall) \frac{dU}{dt} \qquad (27)$$

where C_M is termed as the inertial force coefficient and $\forall$ is the volume of the fluid displaced by the body. More explicitly the inertial force may be expressed as a sum of two forces, namely, Froude-Krylov force and added mass force. The former is the force required to accelerate the fluid particles in the volume of the cylinder in the absence of the cylinder. This acceleration is achieved by a pressure gradient across the cylinder which would constitute a force. The added mass force arises due to the fact that a certain volume of the fluid around the cylinder tends to behave as one with the cylinder. This volume must be proportional to the displaced volume and hence we can write $C_M = C_F + C_a$ where $C_F \approx 1{\cdot}0$ and C_a, the coefficient of added mass found theoretically for the cylinder is also equal to unity. However, in oscillating flow, the effective shape of the body may be changed by the

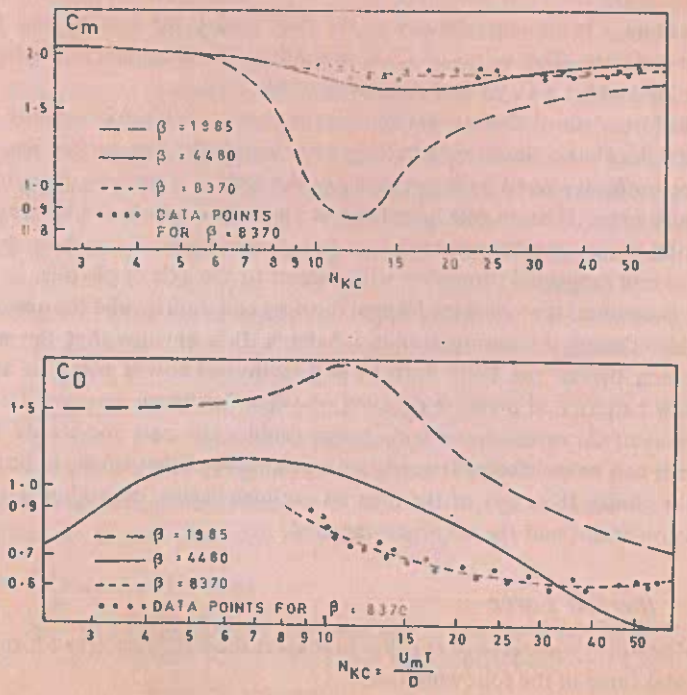


FIG. 4 VARIATION OF C_m AND C_D (Sarpakaya-1976)

formation of the wakes on either side of the body. Hence, the coefficient C_M is dependent on the shape of the body, the relative value of the maximum horizontal traverse of the fluid particle to the diameter, U_mT/D (known as the Keulegan-Carpenter number), roughness of the surface and a boundary layer parameter $\beta = D^2/\nu T$ which is a form of Reynolds number. It may be noted that $\sqrt{\beta}$ is a measure of the ratio of the cylinder diameter to the ratio of the boundary layer thickness which in turn governs the point of separation along with roughness parameter. It has been observed that for small values of K-C number, C_M is close to 2·0 and for large values of β, $C_M \approx 2{\cdot}0$ for all values of K-C numbers. For large values of K-C numbers (conditions representing quasi-steady flow) the function has a minimum value just as was found by Roshko for steady flow past a cylinder. The variation in the value of C_M as a function of the K-C number and β is shown in Fig. 4 after Sarpakaya[20].

5.3. *Morison's Equation—Total Drag*

Combining the drag and inertial forces, the total drag was given by Morison *et al.*[15] though the algebraic addition of two forces is somewhat arbitrary. The total drag force on a vertical cylinder of elementary height, Δz is

$$\Delta F = \left[C_D \cdot D \cdot \tfrac{1}{2}\rho |u| u + C_M \cdot \rho \pi D^2/4 \frac{\partial u}{\partial t}\right] \Delta z \tag{28}$$

Another assumption is made that the total acceleration is nearly equal to the temporal acceleration.

Using one of the deterministic theories, say, linear theory, the force becomes

$$F = \left[C_D \cdot D \cdot \tfrac{1}{2}\rho a^2 \omega^2 \frac{\cosh^2 k(z+d)}{\sinh^2 kd} \left|\cos(kx-\omega t)\right| \cos(kx-\omega t) + C_M \cdot \frac{\pi d^2}{4} \rho \left(\frac{-a\omega^2 \cosh k(z+d)}{\sinh kd} \sin(kx-\omega t)\right)\right] \Delta z \tag{29}$$

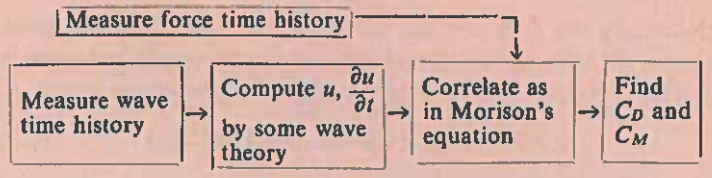
It may be seen that the drag force component and inertia component are 90° out of phase with each other and, hence, the maximum force on the pile at any given location can be obtained by equating $\partial F/\partial t = 0$ and finding the phase angle which gives the maximum force. The total force on the cylinder is found by integrating the above equation for z varying from $-d$ to η. The total moments about the ocean bed may be obtained by integrating $\Delta F\,(z+d)$ between the same limits.

5.4. *General Outline of Computation of Wave Forces*

The major steps to be taken in the computation of wave loads using Morison's equation are :

(1) Values of C_D and C_M are obtained using field and laboratory data
(2) Velocity, u and acceleration, $\partial u/\partial t$ are found by a suitable wave theory (Airy's, Stokes', or stream function)
(3) Forces on each member element are obtained by Morison's equation (or any other form) and summed up over the length of the member below the water surface to get the total force on the entire structure
(4) Allowances are made for the asymmetry of crest and trough above MWL as well as the direction of wave attack.

Determination of force coefficients C_D and C_M is always a major problem in computing the wave forces. A number of field and laboratory studies have been made for this purpose. The general procedure for doing so is as follows :



Considerable scatter is observed in values of C_D and C_M. This is attributed to many factors like presence of current, wake geometry and turbulence, arbitrary variations in waves, directionality of sea, idealization used in various wave theories, equipment inaccuracies and vibrations. Once we obtain a set of C_D, C_M values at the site by above procedure, there are two methods at our disposal to integrate these values and get the final fixed values. They can be classified as deterministic and statistical methods. In the deterministic method, regular and unidirectional waves are assumed and we have (1) trough-crest linear theory fitting, (2) numerical wave theory fitting and (3) fitting by linear fitter method.

The statistical fitting method is naturally very appropriate in confused sea conditions and they assume validity of linear theory and of linear superposition principle. They basically involve a probability law for the characterising statistic like the spectral density. In this method there are three groups : (1) method of force moment, (2) covariance and spectral least square fitting

and (3) frequency dependent cross-density estimates. Out of all these methods the second method is found to give good and reliable results.[8]

5.5. *Limitations of Morison's Equation*

The Morison's equation is simple and direct and most of the wave force analysis made so far in the literature is based upon it. It has provided fair agreement as far as the laboratory tests are concerned. But there are many objections to the use of this theory such as : (a) The drag term used in it is for the conditions of steady flow which is assumed to be true for the oscillating periodic flow as well. But the flow immediately behind the cylinder may not maintain the steady state ; (b) It is questionable whether the drag and inertia force expression can be used together as use of one violates assumption of the other. The linear combination of these terms is also intuitional ; (c) Certain features of natural conditions like energy spreading, multidirectionality, effect of upstream turbulence, past history of flow which can significantly alter the force structure are unaccounted for in Morison's equation ; (d) Large scatter is observed in C_D and C_M values which cause considerable variations in computed wave forces ; (e) The wave kinematics are calculated by means of some wave theory which may have its own limitations ; (f) The increasing tendency to build large monolithic structures demand analysis of diffraction forces which are not considered in this equation ; (g) The large structures give rise to resonant condition with incident waves and thus long-term loading and fatigue analysis become important than the unique design wave concept implied by Morison's equation ; (h) The large building costs of the offshore structures give rise to more loss of economy as factors of safety kept with analysis based on Morison's equation are large.

However, in absence of any other universally accepted formulations, Morison's equation continues to be used for design purposes.

6. Wave Forces on Large Diameter Bodies

For large monolithic structures (condeep platforms) now being developed for offshore platforms in the North Sea and elsewhere, the diameter of the bodies may be larger than 1/5 of the wavelength. In this regime the disturbance pressures on the body generate a significant amount of radiated wave motion often referred to as the scattered wave component. The problem is basically to determine the potential function describing the fluid motion

which satisfies all boundary conditions. The total potential φ is assumed as a linear sum of two components $\varphi_i+\varphi_s$, where the subscripts i and s refer to the known incident and the unknown scattered wave components. φ_s is generated by the disturbances pressures on the body surface and may be compounded by a set of component potentials φ_r originating from an array of pulsating pressure points or sources of strength m_r distributed over the surface such that

$$\varphi_s=\sum_r^R m_r \varphi_r \tag{30}$$

Pertinent equations are :

Laplace's equation

$$\nabla^2 \varphi(r, \theta, y, t)=0 \qquad \text{for } r \geqslant a \tag{31}$$

Boundary condition at the ocean bottom

$$\frac{\partial \varphi}{\partial z}=0 \qquad \text{at } z=-d \tag{32}$$

Bernoulli theorem

$$\frac{\partial \varphi}{\partial t}+g\eta+\frac{1}{2}\left(\frac{\partial \varphi}{\partial r}\right)^2+\left(\frac{\partial \varphi}{\partial z}\right)^2+\left(\frac{1}{r}\frac{\partial \varphi}{\partial \theta}\right)^2=0 \tag{33}$$

at the free surface $z=\eta$ and $r \geqslant a$

Boundary condition at the free surface

$$\frac{\partial \varphi}{\partial t}+\frac{\partial \varphi}{\partial r}\frac{\partial \eta}{\partial r}+\frac{1}{r^2}\frac{\partial \varphi}{\partial \theta}=\frac{\partial \varphi}{\partial z} \qquad \text{at } z=\eta,\ r \geqslant a \tag{34}$$

Boundary condition on the cylinder

$$\frac{\partial \varphi}{\partial r}=0 \qquad \text{at } r=a \quad -h \leqslant z \leqslant \eta \tag{35}$$

φ, η and σ can be expressed in power series of a perturbation parameter $\varepsilon=a_i k$ in the form

$$\left.\begin{aligned}\varphi&=\varepsilon\varphi_1+\varepsilon^2\varphi_2+\varepsilon^3\varphi_3\\ \eta&=\varepsilon\eta_1+\varepsilon^2\eta_2+\varepsilon^3\eta_3\\ \sigma&=\sigma_0+\varepsilon\sigma_1+\varepsilon^2\sigma_2+\varepsilon^3\sigma_3\end{aligned}\right\} \tag{36}$$

where a_i is the incident wave amplitude, k is the wave number $=\frac{2\pi}{L}$ and σ_0 is the linear theory frequency parameter. McCamy and Fuchs[14] found that the horizontal force and moment about the base are given by

$$F_h=\frac{2\rho g H}{k}\frac{\cosh k(d+z)}{\cosh kd}A\left(\frac{D}{L}\right)\cos(\sigma t-\varkappa) \tag{37}$$

and

$$M_0=\frac{2\rho g H}{k^3}B(kd)A\left(\frac{D}{L}\right)\cos(\sigma t-\varkappa) \tag{38}$$

where

$$A\left(\frac{D}{L}\right)=\left[J'^2_1\left(\frac{\pi D}{L}\right)+Y'^2_1\left(\frac{\pi D}{L}\right)\right]^{-\frac{1}{2}} \tag{39}$$

where J'_1 and Y'_1 are derivatives with respect to r of the Bessel functions of the first and second kind respectively

$$B(kd)=\frac{1-\cosh kd+kd\sinh kd}{\cosh kd} \tag{40}$$

$$\tan\varkappa=J'_1\left(\frac{\pi D}{L}\right)\Big/Y'_1\left(\frac{\pi D}{L}\right) \tag{41}$$

Chakrabarti introduced a diffraction coefficient, C_h, defined as the total horizontal force divided by the Froude-Krylov force. When D/L is less than 0·2, C_h will be nearly equal to the coefficient C_M as defined by the added mass. This is true for a fixed cylinder.

7. Morison's Equation for Moving Bodies

When the cylinder itself moves with some other acceleration $\dot{u}_b$ the actual mass of the cylinder must be taken into account.

A body with volume $\forall$, and a known mass coefficient $C_M=1+C_a$ and actual mass M is being accelerated with respect to the fixed axis at the rate $\dot{u}_b$ in a fluid with ambient acceleration with respect to the same axes at the rate of $\dot{u}$. If F_e is the external force, F_k, Froude-Krylov force due to undisturbed pressure field and F_d is the force due to pressure gradient required to sustain the acceleration, we get

$$M\dot{u}_b=F_e+F_k+F_d=F_e+\rho\forall\,\dot{u}+C_a\rho\forall\,(\dot{u}-\dot{u}_b) \tag{42}$$

With $C_a=C_M-1$, we get

$$F_e=C_M\rho\forall\,(\dot{u}_b-\dot{u})+(M-\rho\forall)\,\dot{u}_b \tag{43}$$

The above form of equation is applicable for all cases of bodies, floating or submerged, fixed or moving in steady or accelerating flow. For rigid fixed bodies, $\dot{u}_b=0$ and for floating bodies $M=\rho\forall$ with the result that the second term disappears.

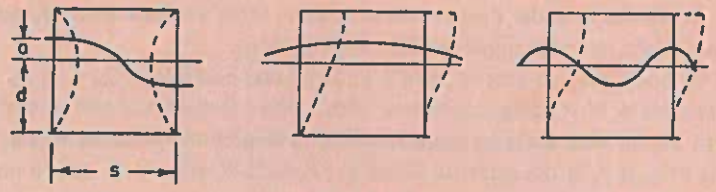
Finally we have the total force given by the equation

$$F=C_D A\cdot\tfrac{1}{2}\rho\,|u-u_b|\,(u-u_b)+C_M\rho\forall\,(\dot{u}_b-\dot{u})+(M-\rho\forall)\,\dot{u}_b \tag{44}$$

8. Groups of Cylinders

When the cylinders are widely spaced, there may not be any influence of one cylinder on the other. But the load on the entire structure requires careful analysis. For illustration, we may discuss the loading on an unbraced

4 column structure. There are three possible types of critical loading conditions depending on the values of wavelength and spacing between adjacent cylinders in line with the direction of the wave. In the first case, when the spacing between the columns is half of the wavelength with a crest occurring at one column and a trough occurring at the other, the net force experienced by the platform is a minimum since the two forces oppose each other. But the deformations in each pile (Fig. 5a) are large and may lead to high stresses. When two nodes occur at each of the pile the net force is zero. In the second case, when $s \ll L$ as shown in (Fig. 5b), the force on the platform is a maximum, the two forces acting in the same direction whether a crest or trough passes over the two piles simultaneously. In the last case ($s/L = 1{\cdot}0$) when two successive crest of troughs pass over the adjacent piles at the same time the force is again significant, the normal forces on the piles being in phase (Fig. 5c).



(a) $s/\lambda = 0{\cdot}5$ (b) $s << \lambda$ (c) $s/\lambda = 1{\cdot}0$

FIG. 5 GROUPS OF CYLINDERS

8.1. *The Effect of Neighbouring Cylinders*

For two cylinders with their joining axis perpendicular to the direction of flow, Landweber observed that if K, the ratio of clear spacing to diameter is less than 0·45, then one single vortex street is formed. If $K > 0{\cdot}45$ but not in the range of 1·5, double vortex streets are formed in which vortices from outer edge of one of the cylinders is in phase with the vortices from the inner edge of the other. If $K = 1{\cdot}5$, vortices from other edge are in phase and vortices from inner edge are in phase. In the former case, the cylinders will vibrate in phase and in the latter the cylinders will vibrate 180° out of phase.

Laird found that when the test cylinder was in the wake of its neighbours, the drag force may increase by 75 percent or decrease to less than (−15 percent) depending on the spacing. If it is directly behind a cylinder the average

drag forces are decreased but fluctuations in drag increase until the interference effect disappears for spacings beyond $10D$. The lift forces and its fluctuations, however, increase considerably at $S = 2D$ and decreases progressively for larger spacings ; when the test cylinder is outside the wake of the neighbours, the drag increases. Lift forces as much as 80 percent of the single cylinder drag force were observed. Drag force fluctuations are attributed to the eddies produced by the neighbouring cylinders.

Laird found the effect of spacing of piles exposed to ocean waves by oscillating a group of 24 vertical $\frac{1}{8}$" cylinders with a 13' amplitude in the horizontal direction. Drag on the group increased with spacing from $D_0/3$ at cylinder contact to 0·9 D_0 at 4 diameter spacing and ultimately to D_0 at 8 diameter spacing (D_0 is defined as 24 times the normal drag on an individual cylinder in uniform velocity). Drag fluctuations increases from $D_0/2$ to D_0 as R varied from 5000 to 10000. At small spacings transverse forces decreased from D_0 to $D_0/3$ as R increased from 5000 to 15000.

9. Effect of Currents

A current moving in a direction opposite to the wave direction can cause steepening of waves and cause the wave to break. If this happens on the structural element, significant impact forces result. The second interference effect is that the current velocities will considerably modify the orbital velocities which are used in Morison's equation for the estimate of the force. In deep water the surface orbital velocity component in the horizontal direction is $\pi H/T$ and if u_c is the current velocity, the drag is increased by a factor

$$1 + \frac{2u_c T}{\pi H} + \left[\frac{u_c T}{\pi H}\right]^2$$

The current can also modify the wake and eddy structure (transverse current can completely sweep away the eddies clear of the structures) such that the coefficients of drag obtained earlier are no longer valid.

10. Fluid-Structure Interaction

10.1. *Steady Flow Around a Stationary Cylinder*

The flow patterns around a stationary cylinder at various Reynolds numbers have been investigated widely both analytically and experimentally. Some of the main results are summarised in the following sections.

When the Reynolds number is less than 3, the flow follows the Stokian law and the laminar creeping motion remains attached to the cylinder. In the range of Reynolds numbers between 5 and 10, a pair of fixed symmetrical vortices are formed in the wake of the cylinder. The fixed vortex pair is separated from the main stream by vortex layers which are a continuation of the laminar boundary layer after its separation and these vortex layers join together again below the vortex pair. The symmetrical formation is unstable to asymmetric disturbances and, as R increases, the vortices elongate in the direction of flow. A Foppl vortex is shed at $R \approx 20$ by a cross-flow that may be induced by a transverse displacement of the cylinder.

In the neighbourhood of the Reynolds number of 90 one of the fixed vortices breaks away from the cylinder. This causes an asymmetry in the wake pressure which results in the other vortex leaving the cylinder and the process repeats itself. Alternate vortex shedding takes place and a Karman vortex street is formed in the wake. Kovasnay and Tritton found that the wake flow behind the fixed vortex pair is unstable at $R \approx 40$ and is the cause for subsequent vortex shedding but the spacing of vortices in the street caused by wake instability appears to be different from that found by Karman, $b/a = 0{\cdot}281$, where a is the longitudinal spacing between successive vortices on the same row and b is the transverse spacing between the two rows.

As R is further increased (150–300), the free vortex layers springing from the separation point becomes turbulent before breaking away into the street. The transition to turbulence depends on initial turbulence and R.

For $R < 150$, the vortex layers roll up into vortices and the flow within is laminar. The transition to turbulence occurs in the range of $150 < R < 300$ and according to Roshko, at a distance 50 D from the cylinder, the development of a turbulent wake occurs devoid of any evidence of upstream periodicity. Beyond $R > 300$, the vortex shedding is irregular and the vortices decay rapidly. This decay of energy of vortices is essentially similar up to $R = 10000$. The irregular range appears to be fully established at a Reynolds number far lower than that required to bring transition of the free vortex layers to the separation point. The transition point on the free vortex layers has been found to occur at $1{\cdot}4\ D$ for $R = 3500$ and $0{\cdot}7\ D$ for $R = 8500$.

The cause which transforms the wake into a vortex street is explained due to the fact that the two vortex layers comprised of the two laminar layers after its separation from the cylinder are disturbed by a downstream asymmetric disturbance. While this is true at higher Reynolds numbers,

the folding up of vortices at low Reynolds number (say $R = 40$) is essentially due to the instability of the line wake. At the critical value of $R = 90$, the eddies oscillate with vortex street frequency and the flow in the attached eddies remain the same. But at a higher Reynolds numbers, a fully developed vortex exists right from the start and the fluid moves continually into the street.

Roshko introduced the width b^* between free shear layers ($b = \lambda b^*$) and then finds $C_D D/D^*$ as a pure Strouhal function of U_S. A new Strouhal number is also introduced as

$$S^* = \frac{nD^*}{U_S} = \frac{S}{\varepsilon}\frac{D^*}{D} \tag{45}$$

for similarity in wakes. Roshko also showed that a splitter plate of length $5D$ attached to the cylinder completely prevented alternate shedding and reduced the drag by 50 percent. Also a small splitter plate, as it is moved down along the centre line of the wake causes the frequency of shedding to decrease until a minimum is reached with the leading edge of the splitter plate located at a distance $4D$ from the cylinder. Any further movement will only cause an abrupt increase in the frequency.

10.2. *Strouhal Number*

It has been explained that beyond the critical Reynolds number of 90, alternate vortices are shed. The frequency of shedding is given by

$$f = \frac{SU}{d} \text{ or } S = \frac{fd}{U} \tag{46}$$

but is really a function of the Reynolds number. The variation of C_D and $1/S$ is shown in Fig. 6a. The reciprocal of S is plotted to show the similarity in variation between C_D and $1/S$.

For a stationary cylinder, in the subcritical region $10^3 < R < 10^5$, $C_D \approx 1{\cdot}1$ and $S \approx 0{\cdot}20$ or $1/S = 5{\cdot}0$. In the critical region (at $R = 0{\cdot}6 \times 10^6$), a reduction in C_D and $1/S$ occurs simultaneously with $C_D = 0{\cdot}40$ and $1/S = 2{\cdot}2$. Further increase in R increases $1/S$ to $3{\cdot}6$ and C_D to $0{\cdot}8$ at $R = 4 \times 10^6$. A general relationship linking S and R is of the form in steady flow system :

$$\left[1 - \frac{19{\cdot}7}{R}\right] S = \frac{fD}{U} \tag{47}$$

for subcritical Reynolds numbers.

10.3. *Flexible Cylinders in Steady Flow*

A flexible cylinder responds to the alternating transverse force (due to the vortex shedding) and vibrates. The dynamic response to the fluctuating load is a function of the natural frequency of the system and the Strouhal frequency of vortex shedding. There are considerable deviations in the drag and lift when the cylinder is free to vibrate compared to those for the stationary cylinder. For example, Gerrard reports that the coefficient of lift increases gradually from $R=4\times10^3$ and reaches a maximum value at $R=7\times10^4$ where the value is 80 times the value at the central value of R. Other results have shown that the maximum lift occurs at a lower R approximately around $R=4\times10^4$ beyond which the lift coefficient decreases. For a flexible cylinder located in steady flow it has been found that if the natural frequency of the cylinder is in the neighbourhood of the Strouhal frequency then the cylinder motion forces the vortex shedding to take place at the cylinder vibration frequency and not at Strouhal frequency. This is frequently referred to as 'lock-in' phenomena.

When the predicted Strouhal frequency is far from the natural frequency, the vortex shedding drives the cylinder at Strouhal frequency. As the Strouhal frequency approaches the natural frequency of the system, the commonly observed synchronisation or lock-in of the vortex shedding and natural frequency occurs. If the inferred shedding frequency is very close to the natural frequency, then a 'resonant lock-in' occurs. Under resonant conditions, the vibration is said to be self-excited and the amplitude would tend to infinity. But equilibrium is, however, reached at large but finite amplitudes of oscillations when the energy lost in dissipation is equal to the energy input into the system of flow. The dissipation occurs from internal friction, radiation, transmission to supports and viscous losses in the fluid.

The Strouhal frequency for a vibrating cylinder is not the same as that for a stationary cylinder at the same Reynolds number. For example, for a suspended cable in tidal currents it was found that Strouhal frequency was between 0·16 and 0·18 in the Reynolds number range of 360 to 850. The relation between S (or $1/S$) and R for a flexible cylinder is given in Fig. 6b. It may be seen that $S=0{\cdot}16$ in the subcritical range. There is no sudden variation in S as R changes from subcritical to supercritical. There is, however, a gradual increase in S between $R=2\times10^5$ and 10^6 and at $R=10^6$, $S=0{\cdot}2$; $R=2\times10^6$, $S=0{\cdot}22$. For resonant lock-in it was seen that the amplitude of

vibration was about $\pm0{\cdot}7$ times the diameter but non-resonant lock-in showed changes in amplitude with a broad spectrum in oscillation with a strong peak at natural frequency. The average amplitude was $\pm0{\cdot}2$ times the diameter of the cables. When Strouhal frequency is very different from natural frequency of the cable, amplitude of fluctuations were more severe and the spectrum showed that the peak was at the Strouhal frequency with a broader band width than in the two other previous cases mentioned.

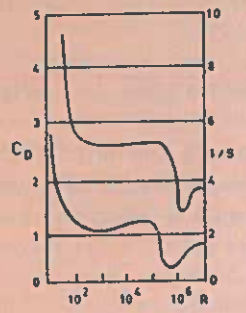


FIG. 6 (a)
MEAN C_D & 1/S Vs. R FOR STATIONARY CYLINDER

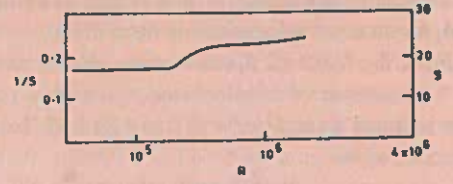


FIG. 6 (b)
APPARENT S Vs. R FOR A CYLINDER FREE TO OSCILLATE

As mentioned earlier, the oscillating cylinder sheds an eddy at the end of each oscillation cycle and the normal width of the wake is enhanced by the total amplitude of vibration. The maximum force is also exerted at this point of eddy shedding and corresponds to a forced linear vibration. This is true only if the amplitude is not large and the velocity of vibration is small. This model will fail for large displacements induced near resonance.

To repeat what has already been said, if n is very different from f_s, then only small amplitude low frequency oscillations occur. If n is close to f_s, large amplitude vibrations occur and nonlinear forces arein troduced. The additional energy required to maintain nonlinear vibrations come at the expense of wake pressure. Low wake pressures exist at resonant conditions.

10.4. *In-Line Oscillations—Variation in Drag*

During the construction of an oil jetty at Immryham, UK, in-line oscillations were found up to $0{\cdot}6\times10^6$. Velocity at which these in-line oscillations

were found was approximately one-fourth the velocity at which cross-flow oscillations are induced. Hence, data from stationary cylinders would be misleading. In-line fundamental mode oscillations occurred when U/nD was greater than 12, where n is the natural frequency in water, U the flow velocity and D the diameter, and K_S (a stability parameter) is less than 1·2.

$$K_S=\frac{2\,me\,\delta}{\rho D^2},\quad me=\frac{\int_0^l my^2\,dx}{\int_0^h y^2\,dx} \tag{48}$$

y = deflection, m = mass, l = total length of cylinder, h = depth, ρ = density of fluid, δ = structural logarithmic decrement.

For $1{\cdot}2 < U/nD < 2{\cdot}5$, vortices are shed symmetrically and for $2{\cdot}7 < U/nD < 3{\cdot}8$ vortices are shed alternately. Cross-flow fundamental mode oscillations were induced when $U/nD < 4{\cdot}5$ and $K_S < 17$. Instability in the second mode occurred when

$$\frac{U}{fd} > 1{\cdot}2 \text{ and } K_S < 1{\cdot}2$$

In-line oscillations were observed by Mandini at low velocities since in-line force variation is the same regardless of which side the vortex is located, the fluctuating drag occurs at two times at the frequency of the lift force. A cylinder flexible enough to have transverse oscillations with amplitudes more than 0·5 times the diameter of cylinder, while performing large amplitude oscillations in water tend to oscillate transversely at the eddy shedding frequency and to vibrate at twice the eddy-shedding frequency in the direction of motion. More stiffly supported cylinders tend to vibrate in both modes at the natural frequency of the cylinder. The cylinders tend to follow a crescent path. It is thus seen that the movement of the flexible cylinder is truly three-dimensional and nonlinear.

10.5. *Oscillations in Forces*

Laird found that for a single cantilevered flexible square cylinder, the wake width was nearly twice its normal width. The mean coefficient of drag was large for small amplitudes of vibration and decreases rapidly with increasing velocity to a normal value just before the wake capture. The lift force frequency, f_L, tends to have two values which are (1) less than the eddy shedding frequency f_s or (2) just less than natural frequency n up to the

velocity of capture (lock-in) and there is one value beyond it. The frequency of drag oscillations f_{D_2} in the direction of flow tends rapidly to $2f_L$. At a substantially larger velocity, f_L is even greater than n. For still larger velocities, barring close approach to other natural frequencies of the system f_L becomes rapidly equal to n. These changes in frequency apparently coincide with apparent resonant effects in C_{LR}, the coefficient of lift response based on the amplitude of transverse oscillations. The dragwise oscillation coefficient C_{D_2} also goes through an apparent resonance. Laird found an anomaly that C_D starts with large values (nearly 3 times the steady drag) for small amplitudes of vibration (2 percent of diameter) and decreases with increase in value of U and returns to the normal value just before wake capture. At resonance, C_D is largely associated with large transverse oscillations. Beyond resonance C_D is again normal. Under near resonant conditions the transverse oscillations were nearly 5 times the value at lower frequencies.

Laird found that when the eddy shedding frequency is close to the natural frequency, an oscillating cylinder, moved to-and-fro by a carriage, apparent resonant effects occurred in the lift and drag forces with C_D and C_L being 4 to 5 times those obtained for a stationary cylinder.

The wake near the cylinder for a flow with $R < R_{cr}$ has a clear periodic structure, with a dominant frequency (at each side of the cylinder) expressed by the Strouhal number. But the wake flow with $R > R_{cr}$ is much more turbulent and it is necessary to speak of a power spectrum rather than of a single dominant frequency in the subcritical range. The amplitude of the lift and drag coefficients are given in the table below.

Author	*Reynolds Number*	*Amplitudes of fluctuating forces* C'_L *and* C'_D	
		Lift	*Drag*
Schwabe	735	0·45	$C_D = 1{\cdot}09$
Bingham, Weiner, Griffith	76,000	0·90	—
McGregor	45,000 to 130,000	0·60	C'_D Decreasing from 0·076 to 0·044 $C_D = 0{\cdot}76$
Phillips	40 to 160	0·76	$C'_D = 0{\cdot}076$
Drescher	104 to $1{\cdot}13 \times 10^5$	0·6 to 1·3	0·06 to 0·30 ; C_D varied between 0·95 to 1·2

At higher Reynolds numbers approaching the transitional regime, the fluctuating forces are nonharmonic. In the supercritical range, Fung found that the root-mean-square value of the lift coefficient was about 0·12 with peak values reaching 0·3 in the range of Reynolds numbers between $0{\cdot}4\times 10^6$ to $1{\cdot}4\times 10^6$. In the same range C_D (mean value) was 0·25 and the root-mean-square value of the fluctuating drag coefficient was about 0 045.

The nondimensionalised power spectrum $P(S)$ plotted as a function of Strouhal frequency, S, shows that the power spectrum is not a function of Reynolds number. Fung oscillated the cylinder and found that the lift coefficient does not alter very much for $R > 600{,}000$.

$$\frac{f_e D}{U} < 0{\cdot}12 \text{ and } \frac{a}{d} < 1/24 \quad (a = \text{amplitude of forced oscillation})$$

McGregor discusses the formation of partial wakes in unsteady oscillating flow with flow separation. In such cases, the forces attributable to drag and virtual mass is quite complex. At moments of maximum oscillation and zero velocity, there is no wake and the pattern of flow closely resembles that the irrotational flow without separation. Subsequently, as the velocity increases a wake forms, if the flow were to persist in one direction for a long enough time, a wake would grow until it became similar to the wake obtained for a steady flow. If not, a wake is partially formed and destroyed, the process being repeated in each half cycle.

The size of wake affects both the drag and the virtual mass. The larger the wake becomes, the higher the pressure within the wake and the smaller the coefficient of drag. Also a larger wake causes more disturbance to the surrounding flow so that the coefficient of virtual mass is larger. Thus, the changes in the wake have profound influence on the resistance. The dependence of both C_D and C_M on the wake geometry is obvious. Hogben assumes that the oscillating wake model is an elliptical cylinder as shown in sketch (Fig. 6) at different times of the cycle.

By a simplified analysis, he shows that

$$C_M = \frac{\sqrt{N^2 C_D/\pi^3 + \frac{1}{4}} + \frac{3}{8}}{\sqrt{N^2 C_D/\pi^3 + \frac{1}{4}} + \frac{1}{2}} \tag{49}$$

where $N = \pi U_0/a\omega$, a = radius of cylinder.

C_M thus calculated has been compared with Sarpakaya's result (assuming C_D) and the agreement is qualitatively good. Sarpakaya and Garrison found that the relationship between C_D and C_M was dependent on l/d, where l is

the distance travelled by the cylinder from its rest position. McKnown and Keulegan investigated the case of cylinder located at the nodal point of a standing wave so that the motion was purely horizontal at the cylinder location. C_D and C_M were found with the ratio T/T_e where T is the wave period and T_e is the period of a pair of eddies shedding in steady flow at a velocity characteristic of the unsteady flow (U_{max}). They found that if $T/T_e < 0{\cdot}1$, separation and eddy formation were relatively unimportant with the inertial effects being approximately those for classical unseparated flow and if $T/T_e > 10$, the motion was quasi-steady.

The significance of Strouhal number in wave motion past cylinders is found as follows. If

$$U = \frac{\pi H}{T}\cos(2\pi t/T) \quad (\text{at } X=0) \tag{50}$$

the average velocity may be taken as $V_{av} = \frac{\pi H}{2T}$

For at least one eddy to have time to form it is necessary with

$$\frac{f_e D}{V} = S, \quad f_e = \frac{VS}{D} = \frac{\pi H}{2T}\frac{S}{D}, \quad T = \frac{1}{f_e} = \frac{2TD}{\pi HS} \tag{51}$$

or if

$$S = 0{\cdot}2 \quad H > 10\frac{D}{\pi}, \text{ or } \frac{U_{max}T}{D} = \frac{\pi H}{T} > 10$$

In fact, experiments by Bidde have shown that Keulegan-Carpenter number is more important in establishing relationship between C_D and C_M than Reynolds number. With K-C number, $\frac{U_{max}T}{D} < 2$ there was no separation. With K-C number between 2 and 3, a small separation was observed without a clear eddy developing. With $N_{KC} = 3$, the first eddy is shed. Asymmetry starts and lift force begins to be non-zero. When $3 < N_{KC} < 4$, more than 2 eddies are shed within half cycle and Karman vortex sheet is observed. At $5 < N_{KC} < 7$ wake becomes turbulent and additional eddies are caused when swept back. Frequency of lift force is twice the wave frequency. When $N_{KC} > 7$, extremely turbulent wake is observed (Fig. 7).

The ratio of maximum lift was found to be approximately 40 percent (maximum value 60 percent) of the longitudinal force (Fig. 8). At these high values of N_{KC}, there is no single predominant frequency of eddy shedding but a spectrum is observed just as in case of supercritical Reynolds number but there was a strong predominance for the lift to have a frequency at twice the wave frequency.

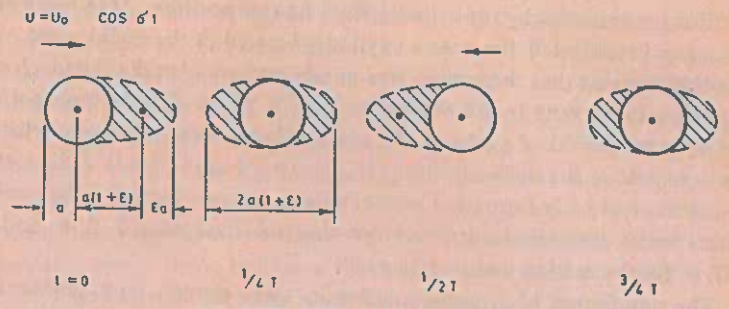


FIG 7 OSCILLATING WAKE (Hogben – 1976)

Bishop and Hassan observed from their experiments for a stationary cylinder that C_L increased with Reynolds number varying from 0·3 at $R=$ 3,600 to 0·65 at $R=11{,}000$. In the synchronisation range, composite forms are obtained. A regular beat phenomena is observed of addition of two forces one occurring at Strouhal frequency and the other at natural frequency. This is observed for both lift and drag forces. Mention must be made here again that Strouhal frequency for a vibrating cylinder is not the same as that for a stationary cylinder.

Bishop and Hassan have presented results for oscillating cylinders in moving streams in the general form

$$F(t)=\varphi\left(\frac{a}{d},\ \frac{Ud}{\nu},\ \frac{fd}{U}\right)\frac{1}{2}\rho U^2 A \qquad (52)$$

Analysis of the cylinder as a system having two degrees of freedom shows great mathematical complexity and difficulty in describing the flow about the cylinder. The forcing function for each mode is an irregular function of time and of deflection in each direction.

11. Non-Deterministic Analysis of Ocean Waves

The non-deterministic methods can be further grouped into two classes as 'statistical' and 'spectral' methods.

In the statistical method, individual waves are formed by 'zero cross method' in which the 'wave' is defined as the portion of water surface profile which is in between two successive upcrosses of Mean Water Level (MWL) as shown in Fig. 1(b). Then, using these individual waves, characterized by wave height, H, and wave period, T, as units of analysis all other gross

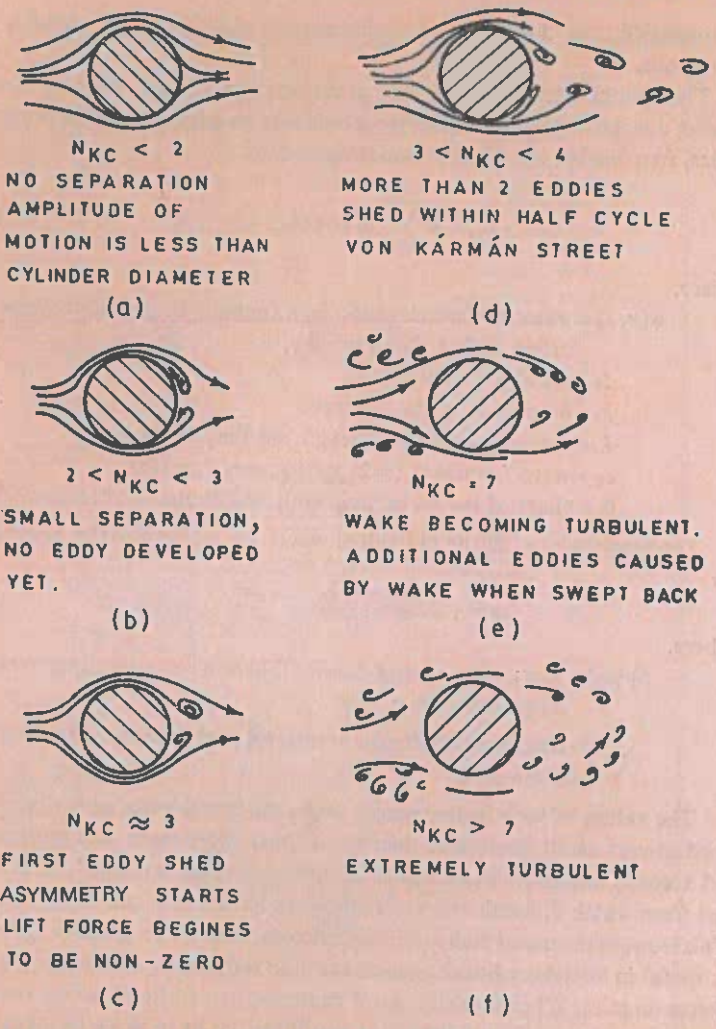


FIG. 8 WAKE CHARACTERISTICS
(Wiegel-1980)

characteristics of the waves at a site, such as significant wave height, are found out.

The spectral method consists of characterizing the sea state by superposing a large number of linear progressive waves each with different amplitudes, frequencies, length and random phase, as

$$\eta(x, t)=\sum_{j=0}^{N} a_j \cos(k_j x-\omega_j t+\theta_j) \tag{53}$$

where,

$\eta(x, t)$ = water surface elevation as a function of position x from the origin and at time instant t,

N = number of wavelets

a_j = amplitude of the j th wave

k_j = wave number ($=2\pi$/length) of the j th wave

ω_j = wave frequency ($=2\pi \times$ frequency f in Hz)

θ_j = phase of the wave, uniformly distributed over interval $(0, 2\pi)$

The amplitudes a_j of the individual waves are related to the 'spectrum' by

$$a_j=\sqrt{2\, S_\eta(\omega_j)\, \Delta\omega} \tag{54}$$

where,

$S_\eta(\omega_j)$ = spectral or energy density function corresponding to radian frequency ω_j and

$\Delta\omega$ = frequency resolution or interval used in computing the above function.

The values of such instantaneous water surface fluctuations, $\eta(x, t)$, are read at very small, successive interval of time, Δt (which may typically be 0·1 second) and using these values as units, 'spectral moments' are calculated from which different characteristic wave parameters can be determined. This requires the use of high-speed digital computers. The spectral technique is useful in analyzing linear phenomena like refraction, diffraction, inertia forces on piles. When the principle of superposition of linear waves becomes inadequate due to predominance of nonlinearities as in wave breaking and overtopping, recourse is generally made to statistical analysis. Spectrally found parameters are naturally more meaningful as all records are analyzed in the same fashion. Spectral analysis is also found to be more economical than the statistical one.

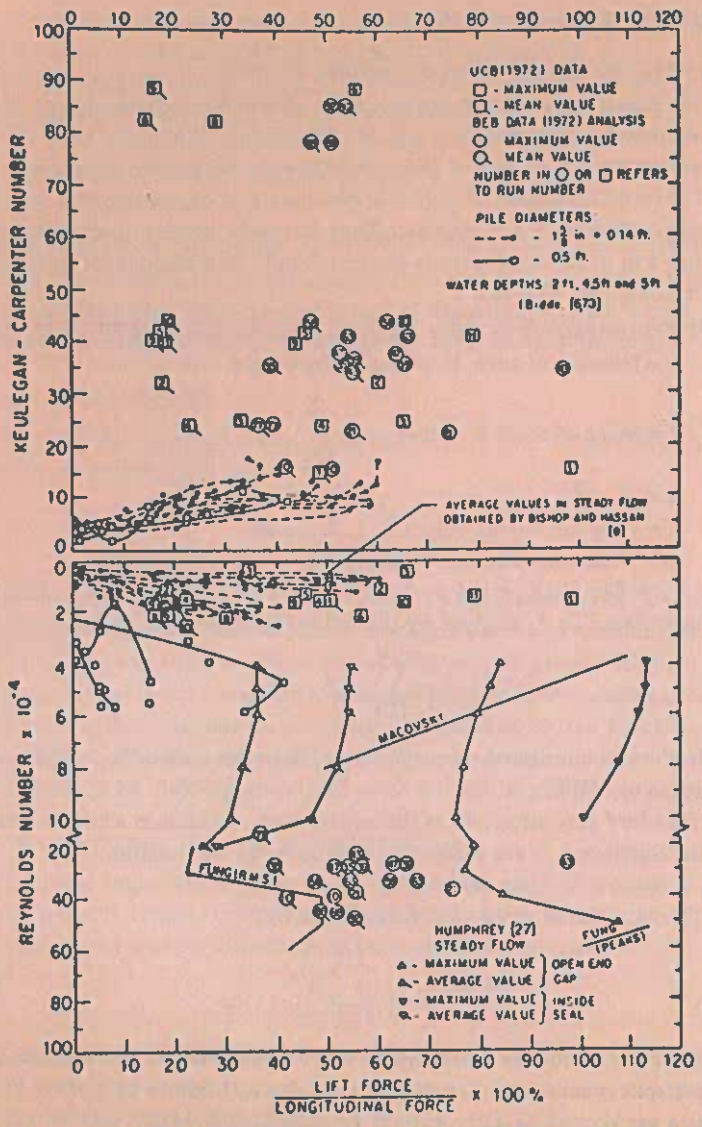


FIG. 9 VARIATION OF N_{KC} AND R_e WITH RATIO OF LIFT TO LONGITUDINAL FORCE (Wiegel - 1980)

11.1. *The Spectral Analysis*

Variance, Standard Deviation and Central Moments

The ocean wave records are generally obtained continuously for about twenty minutes followed by a gap of three hours. For every such record, we can compute a variety of characteristic wave parameters as under.

The formulae employed to obtain various statistics, assume that the wave process is stationary and ergodic. They basically involve integration over a range and hence are generally converted into their discrete forms for their use in computers as below :

If Y_i = Instantaneous water surface fluctuation above MWL at ith time,
N = Number of such Y_i values in the record

$\overline{Y}$ = Mean of such Y_i values $= \frac{1}{N}\sum_{i=1}^{N} Y_i$

$X_i = Y_i - \overline{Y}$

K = Time lag having values 0, 1, 2, ..., M

M = Maximum number of lag values

Δt = Sampling period or digitizing interval used to read Y_i values.

Then 'variance', σ_x^2, is given by the following expression

$$\sigma_x^2 = \frac{1}{N-1}\sum_{i=1}^{N} X_i^2 \tag{55}$$

It gives us an idea of the magnitude of the water surface fluctuations with reference to MWL.

'Standard Deviation', σ_x, is the square root of variance which is related to the 'Significant Wave Height' in the record by the relation

$$H_i = 4\sigma_x \tag{56}$$

Third central moment, $MX3$, is defined by

$$MX3 = \frac{1}{N\sigma_x^3}\sum_{i=1}^{N} X_i^3 \tag{57}$$

It tells us about the symmetry of the X_i data values. Its value is zero for perfectly symmetrical distribution and takes up negative or positive values as data are skewed to left or right respectively. A better way to get the information about skewness is to compute 'skewness parameter', $\beta_1 = \mu_3^2/\mu_2^3$

where μ_3 = third central moment and μ_2 = second central moment or variance. For the Gaussian normal distribution, which is symmetrical, $\beta_1 = 0$.

Fourth central moment, $MX4$, which is the measure of the tendency of the distribution to form a peak, also called the peakedness, is defined by

$$MX4 = \frac{1}{N(\sigma_x^2)^2}\sum_{i=1}^{N} X_i^4 \tag{58}$$

The peakedness parameter $\beta_2 = MX4/\mu_2^2$ has a value '3' for Gaussian Normal Distribution. β_1 and β_2 together enable us to know how far the distribution of the data values resemble that of Gaussian normal, which is a very important assumption made in deriving various wave properties.

Autocorrelation Function

Autocorrelation Function for time lag k is given by the average lagged product of neighbouring values, i.e.

$$R(k) = \frac{1}{N-K}\sum_{i=1}^{N-K} X_i X_{i+k} \tag{59}$$

For random data, as in the present case, it generally shows an oscillatory decay with an approximate tendency towards sinusoidal variation when plotted against various time lag values, K (Fig. 10). It gives additional information about wave irregularity as well as dominant frequencies present in the wave system. It may be seen that the autocorrelation function for zero time lag is the variance itself, i.e. $R(0) = \sigma_x^2$. Autocorrelation coefficient is obtained by dividing every $R(k)$ value by $R(0)$ as $C(k) = R(k)/R(0)$

The Spectral Density Function and Energy Spectra

The most important statistic of interest in any random process is the 'Power Spectral Density Function'. Basically, its value for frequency f is expressed by the Fourier transform of autocorrelation function

$$E(f) = \int_{-\infty}^{\infty} R(k) e^{-i\omega\pi fk}\, dk \tag{60}$$

In its discrete form, this can be expressed by

$$E(f) = 2\Delta t\left[R(0) + 2\sum_{k=1}^{M} R(k)\cos 2\pi fk\right] \tag{61}$$

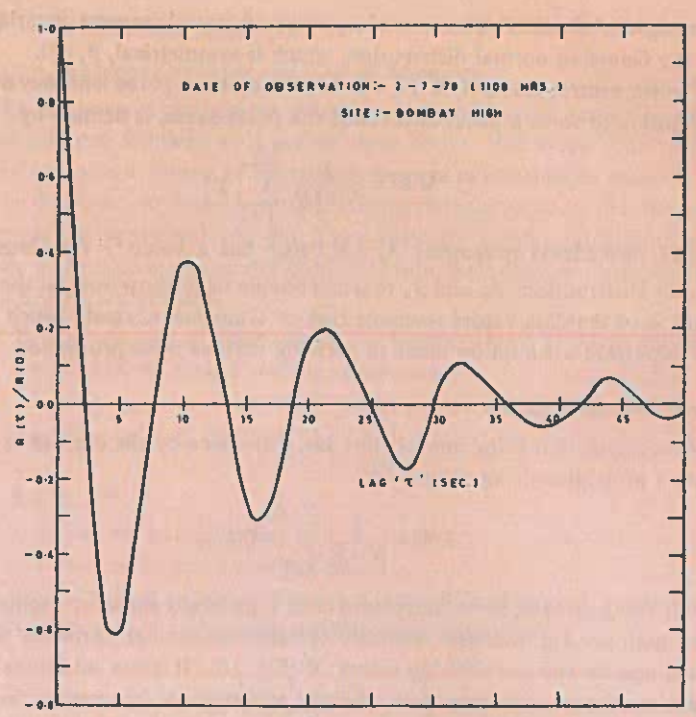


FIG.10 AUTO-CORRELOGRAM

The physical significance of the spectral density function, $E(f)$, is as follows.

The amplitudes a_j in eqn. (54) describing the free surface are distributed such that for any frequency interval $\Delta\omega$

$$\sum_{\omega_j}^{\omega_j+\Delta\omega} \frac{1}{2}a_j^2 = E(\omega_j)\,\Delta\omega \qquad (62)$$

Kinsman[11] has shown that the average potential energy of the wave system is proportional to the variance of the water surface elevations and variance is one-half of the sum of squares of all coefficients obtained as left-hand side of preceding equation. $E(f)$ is thus proportional to the energy of the wave system corresponding to frequency f.

When the above spectral (or energy) density function is plotted against corresponding frequencies, the resulting graph is called the 'Energy Spectrum'. One such typical observed spectrum is shown in Fig. 11. The area under the curve gives the total energy of the wave system. The so-called 'Directional Spectrum' gives the distribution of the energy density with respect to frequency and direction.

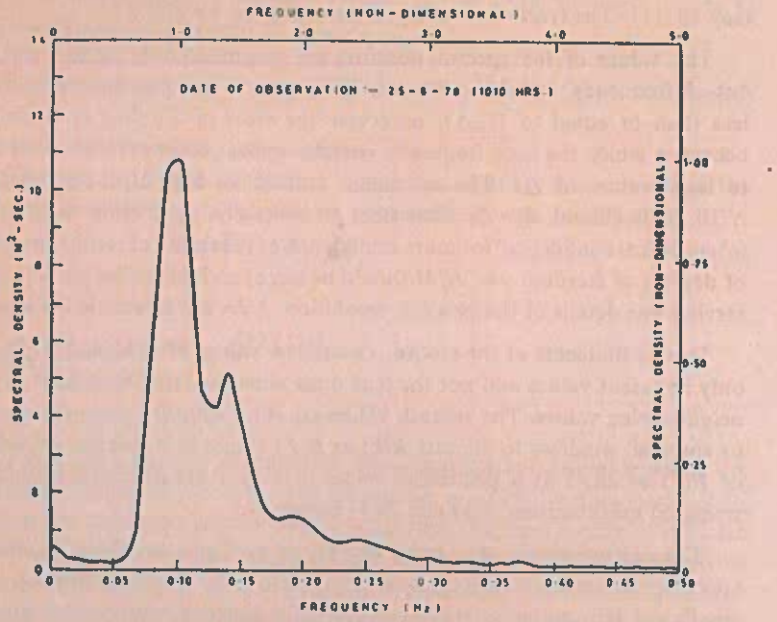


FIG.11 WAVE ENERGY SPECTRUM

Most computations associated with spectral analysis become simple if the wave spectra are assumed to be narrow banded which means that most of the energy is confined to a narrow band of frequencies. Mathematically, the spectra can be considered narrow banded if the spectral width parameter ε, defined later, is greater than 0·5. The variation of $\eta(t)$ for such a narrow band process is nearly sinusoidal with gradual change in amplitudes.

The method of finding the spectral density function by means of Fourier Transform of autocorrelation function, as discussed above, is called the covariance method. This takes relatively large computer time which is highly undesirable when vast data are under consideration. A comparatively faster method to compute the spectral density is the Fast Fourier Transform Method which achieves reduction in time by limiting the redundancy of the same complex potentials appearing several times in the record. This is achieved by writing the number of data point $N=2^k$ where k is an integer (say 10, 11). The fractional saving in time is given by k/N.[8]

The values of the spectral densities are computed only up to a certain cut-off frequency, called the 'Nyquist Frequency'. This frequency should be less than or equal to $1/2\triangle t$, otherwise the error of 'Folding or Aliazing' occurs in which the high frequency energies appear at lower frequencies due to large values of $\triangle t$. The maximum number of lag, M, is restricted to $N/10$. This should also be consistent to mutually conflicting criteria of (*a*) statistical confidence (for more confidence or reliability of results, number of degrees of freedom $\gamma=2N/M$ should be large) and (*b*) resolution (for preserving fine details of the process, resolution $\triangle f=2N/M$ should be small).

Due to finiteness of the record, computed values of $R(k)$ and $E(f)$ are only apparent values and not the true ones showing large variations in the neighbouring values. The smooth values are then obtained by applying 'lag or spectral' windows to the raw $R(k)$ or $E(f)$ values in which the net values of $R(k)$ or $E(f)$ at a particular values of k or f are obtained by adding weighted neighbouring $R(k)$ and $E(f)$ estimates.

The energy spectra have got a number of applications. Once the energy spectrum is known, a variety of information can be obtained from it. The significant information obtained is that of frequency composition present in the wave system and dominant frequencies present. The total energy of the wave system as well as the root-mean-square value of the water surface fluctuations is obtained by computing the area under the spectral curve. By applying suitable transfer functions to the wave surface spectral densities, we can get the response spectra useful for the design purpose. Knowing the spectrum, we can generate the random sea in the laboratory. The numerical integration of the energy spectra gives us important design parameters such as significant wave height, average spectral wave period, spectral width parameter and so on as discussed below.

The Spectral Moments and Design Parameters

The nth spectral moment is given by

$$m_n=\int_0^\infty f^n E(f)\,df \tag{63}$$

The zeroth moment $m_0=\int_0^\infty E(f)\,df$ gives area under the spectral density curve or total energy of the wave system. This is useful to obtain values of the significant wave height, as $H_s=4\sqrt{m_0}$.

The second spectral moment is given by $m_2=\int_0^\infty f^2E(f)\,df$ which yields the characteristic average wave period, $T_{av}=\sqrt{m_0/m_2}$

The spectral width parameter, which is indicative of the width of the energy spectrum over which there is substantial power, is defined by

$$\varepsilon=\sqrt{1-m_2^2/m_0\,m_4} \tag{64}$$

where

$$m_4=\int_0^\infty f^4E(f)\,df$$

11.2. *Theoretical Spectra*

There are a number of standard theoretical spectra available which are generally compared with observed spectra so as to evolve the spectrum of best fit which in turn is useful while calculating response spectra of offshore structures as well as providing a basis for model prototype confirmity in laboratory simulation studies. A better organization of observed spectra is also another advantage from such a comparison with the theoretical spectra. Some of the important theoretical spectra are as follows.

Pierson-Muskowitz Spectrum

Using Kitaigorodski's similarity hypothesis that when plotted in a certain non-dimensional form, plots of observed energy spectra must have the same shape, Pierson and Muskowitz analyzed data from North Atlantic Sea. Assuming that the spectral density, $S(f)$ depends upon four variables of frequency, f, gravity, g, friction velocity, U', fetch, F, i.e.

$$S(f)=\varphi\,(f, g, U', F) \tag{65}$$

and using a curve fitting approach he gave the expression for spectral density as[18] :

$$E(\omega)=\alpha g^2\omega^{-5}\exp\left[-\beta(\omega_0/\omega)^4\right] \tag{66}$$

where $\alpha=0{\cdot}0081$, $\beta=0{\cdot}74$, $\omega_0=g/U$

This is a standard spectrum for designing of offshore structures. However, it has a disadvantage that it is developed for fully developed sea conditions. The fully developed sea is produced when a wind of unlimited duration blows over unlimited fetch length. In fully developed seas, all wind energy input is dissipated by breaking. The spectrum in such a sea state consists of all frequencies each being capable of carrying maximum energy. The P-M spectrum may be compared with the Bretschneider[4] spectrum :

$$E(\omega)=\frac{m_0}{\omega_0}\left(\frac{\omega_0}{\omega}\right)^5\exp\left[-1{\cdot}25\left(\frac{\omega_0}{\omega}\right)^4\right] \tag{67}$$

JONSWAP Spectrum

In order to consider the most practical case of partially developed seas, Hesselman and others[9], based on data of North Sea, proposed another widely used 'JONSWAP' spectrum which is defined by multiplying P-M spectrum by a function of peak enhancement factor, γ as follows :

$$E(\omega)=\alpha g^2\omega^{-5}\exp\left[-1{\cdot}25\left(\frac{\omega_0}{\omega}\right)^4\right]\gamma^{\exp\left[-(\omega-\omega_0)^2/2\sigma^2\omega_0^2\right]} \tag{68}$$

where α and ω_0 are functions of U and F

$$\gamma=\frac{\text{Maximum spectral energy}}{\text{Max. energy corresponding to P-M spectra}}$$

$=3{\cdot}3$ average

$\sigma=0{\cdot}08$ average

Scott's Spectrum

Scott based his analysis on Darbyshire's data[21] and gave expression for energy spectrum as follows :

$$E(\omega)=3{\cdot}424\,m_0\exp-\left[\frac{(\omega-\omega_0)^2}{0{\cdot}065\,(\omega-\omega_0+0{\cdot}26)}\right]^{\frac{1}{2}} \tag{69}$$

for $-0{\cdot}26<(\omega-\omega_0)<1{\cdot}65$

This is found to give a better fit for the spectra observed off Mangalore[6] and Bombay[16] in India.

Scott-Wiegel Spectrum

Wiegel[27] observed that the constants 3·424 and 0·065 in the above equation of Scott's spectrum are changing with variance or total energy of the wave system. He generalized the Scott's spectrum as

$$E(\omega)=A\,m_0\exp-\left[\frac{(\omega-\omega_0)^2}{B\,(\omega-\omega_0+0{\cdot}26)}\right]^{\frac{1}{2}} \tag{70}$$

Wiegel has given values of the constants A and B for various levels of significant wave height (which is indicative of energy of the wave as $H_s=4\sqrt{m_0}$).

11.3. *Directional Spectrum*

Wave components with most energy in the vicinity of the spectral peak (which is also representative of wave energy of significant waves) travel in the same direction as the mean wind direction and there is relatively little spread in the direction (say $\pm 30°$ to $40°$) for these energetic waves. The wave spectra described earlier ignores the specific contributions of wave components from different directions but only highlights the net result at the point of reference.

However, for certain engineering design or operation problems in the ocean environment, one needs to know the wave period and direction in addition to the wave heights. A directional spectrum is essentially a representation of these quantities for a composite wave spectrum.

If $S(\omega,\theta)$ = directional spectrum of frequency f and

$$\text{at direction } \theta=\sum_{\omega}^{\omega+\delta\omega}\sum_{\theta}^{\theta+\delta\theta}\frac{1}{2}a_i^2=\frac{2k^2}{\omega}E(k,\theta) \tag{71}$$

then the expression for the water level elevation $\eta(x, y, t)$ being viewed as a result of superposition of linear waves is

$$\eta(x,y,t)=\int_0^{\infty}\int_0^{2\pi}\sqrt{4\,S(\omega,\theta)}\cos\left(kx\cos\theta+ky\sin\theta-\omega t+\varphi\right)d\theta\,d\omega \tag{72}$$

where x and y are horizontal reference axes, k and φ are the wave number and phase angles respectively.

One of the models used for directional spectra makes use of the circular normal distribution function

$$S(f,\theta)=S(f)\exp\left[\frac{-k\cos(\theta_0-\theta_m)}{2\pi I_0(k)}\right] \tag{73}$$

where θ_m is the mean direction, $I_0(k)$ is the modified Bessel function of the first kind of zero order for imaginary argument.

Thus it can be seen that the directional wave spectra give the distribution of wave energy density with respect to frequency and direction. It specifies wave climate more completely than representations like significant wave heights, predominant wave period and direction. Among many engineering problems where use of directional spectra themselves are input parameters for correct design are (1) study of refraction, diffraction of ocean waves, (2) vibration and 3-dimensional analysis of structures with torsional loads, (3) prediction of response of ships and floating drilling vessels to sea conditions, (4) study of littoral transport, (5) ocean outfall sewers and (6) wave forecasting and hindcasting.

The measurements of directional spectra incorporate use of arrays of pressure transducers (or wave staffs), buoys or remote sensing techniques. Line arrays are used in case of waves coming roughly from one direction, otherwise polygonal arrays are used. Buoy measurements of surface elevations at a single point are associated with measurements of surface slopes or velocities along two axes simultaneously by current meters. Remote sensing methods involve use of visible light optical technique and radar backscatter techniques using radiowaves.

11.4. *The Statistical Analysis*

The Short-term Wave Height Statistics

The short-term analysis is restricted to a single record observed for a short interval of time for which the sea conditions are assumed to be stationary and are such that the wave properties can be studied around mean values.

Longuet-Higgins[18] assumed that (i) the distributions of instantaneous water surface fluctuations is Gaussian normal, (ii) the wave energy is confined to a narrow range of frequencies and (iii) ocean waves are results of superposition of many sinusoidal components in random phase. Under

these assumptions, he stated that the pdf of individual wave heights, H, ($=2\times$amplitude) is given by a typical 'Rayleigh Distribution'

$$P(H)=\frac{2H}{H_{rms}}e^{-(H/H_{rms})^2} \tag{74}$$

where H_{rms} = root-mean-square wave height of the record.

The distribution function for any numerical wave height being equal to or less than H_1, can be derived from this.

$$P(H_1)=\int_0^{H_1}\frac{2H}{H_{rms}}e^{-(H/H_{rms})^2}\,dH \tag{75}$$

$$=1-e^{-(H_1/H_{rms})^2} \tag{76}$$

From this the probability that the wave height exceeds any particular value H_1 is

$$1-P(H_1)=e^{-H_1^2/H_{rms}^2} \tag{77}$$

Thus, knowing H_{rms} from the record we can find the probability that a particular wave height will exceed a given value. Based upon fitting of theoretical random distribution curves as above to the observed wave height data, Longuet-Higgins gave typical relationships between other characteristic wave heights as follows :

$$\begin{aligned}&\text{Mean wave height, } H_m=0{\cdot}885H_{rms}\\&\text{Average of highest 33·3 percent waves, } H_s=1{\cdot}416H_{rms}\\&\text{Average of highest 10 percent waves, } H_{1/10}=1{\cdot}8H_{rms}\\&\text{Maximum wave height, } H_{max}=2{\cdot}172H_{rms}\end{aligned} \tag{78}$$

Knowing any one typical wave height, we can compute other wave heights by the preceding relations. Longuet-Higgins, using above formulations, has given tables to find out probable maximum wave amplitude, prob (a_{max}), if number of waves, N, that occurred at a site is known. These tables give values of N against prob (a_{max})/$\bar{a}$ ratio (where $\bar{a}$ = mean wave amplitude $=H_m/2$). Ratios of average highest $1/p$th wave to $\bar{a}$ are also presented.

Tucker's Method

Tucker's method enables us to determine the characteristic wave heights in a record by considering only a few larger amplitudes. This method, though somewhat approximate, is simple and quick and generally followed when rapid calculations are necessary.

Assuming narrow band of frequencies, Tucker[25] gave expressions for H_{rms} in a record as

$$H_{rms} = \sqrt{2}H_1(2\theta)^{-\frac{1}{2}}(1+0{\cdot}289\,\theta^{-1}-0{\cdot}247\,\theta^{-2})^{-1}$$

also

$$H_{rms} = \sqrt{2}H_2(2\theta)^{-\frac{1}{2}}(1-0{\cdot}289\,\theta^{-1}-0{\cdot}103\,\theta^{-2})^{-1} \tag{79}$$

where

$H_1 = A + C$, $H_2 = B + D$, $\theta = \log N_z$,

A = height of the highest crest above MWL

B = height of the second highest crest above MWL

C = depth of lowest trough below MWL

D = depth of second lowest trough below MWL

N_z = number of zero-up-crosses in the record.

Long-Term Wave Height Statistics

The above methods describe the wave conditions for a short interval of time normally of 3 hours. The long-term knowledge of the wave heights is obtained by analyzing and estimating records over one or more years and is aimed at predicting extreme wave conditions. The most useful statistic obtained from such analysis is the 'Design Wave' which is one that is exceeded once in N years, where value of N may be 1, 10 or 100 years.

For any reliable long-term prediction, wave records must be available for a long length of time. For the prediction of long-term wave height, observed probabilities of occurrences are compared with theoretical long-term distributions and use of the best fit distribution is made to carry out the linear extrapolation for getting the design waves with a specified return period. The available long-term distribution functions are mainly four, viz. (1) Log normal, (2) Exponential, (3) Weibull and (4) Frechet. Probability distribution for all of them, in general, can be given by

$$p(H) = \exp\,(-H/A)^n \tag{80}$$

where A and n are constants. n is arbitrary for Frechet distribution, any positive integer for Weibull distribution and '1' for exponential distribution. Use of all four distribution functions is recommended for getting a linear trend.

Tickell and Holmes[24] recommended that for getting the long-term wave height distribution at a site, the long term data should be summarized by bivariate histogram of observed H_s and $\bar{T}_z$ values, observed for at least a

year. For every pair of H_s and $\bar{T}_z$ the extremal short-term Rayleigh distribution of wave height is

$$p\left(\frac{H}{H_{s_i}, \bar{T}_{z_j}}\right) = 1{\cdot}0 - \exp\,(-2H^2/H_{s_i}^2) \tag{81}$$

The long-term wave height distribution, $p_{LT}(H')$ for any height H' is obtained by summing up above equation with all H_s and $\bar{T}_z$ values in the histogram.

$$\therefore \quad p_{LT}(H') = \text{prob}\,(H \leq H')$$

$$= 1{\cdot}0 - \frac{\left\{\sum_i \sum_j \exp\,[-2(H'/H_{s_i})^2]\, n_{ij}/\bar{T}_{zj}\right\}}{\bar{N}_z} \tag{82}$$

where n_{ij} = time interval for which

$$H_{s_i} - \frac{\Delta H_s}{2} \leq H_s \leq H_{s_i} + \frac{\Delta H_s}{2}$$

and

$$T_{zj} - \frac{\Delta \bar{T}_z}{2} \leq \bar{T}_z \leq \bar{T}_{zj} + \frac{\Delta \bar{T}_z}{2}$$

and

$\bar{N}_z$ = average number of waves/unit time

$$= \sum_i \sum_j n_{ij}/\bar{T}_{zj}$$

This derived distribution $p_{LT}(H')$ for various H' values is fitted to either Weibull distribution given by

$$p(H) = 1 - \exp\left[-\left(\frac{H-\alpha}{\beta}\right)\gamma\right] \tag{83}$$

or to Gumbel distribution,

$$p(H) = \exp\left\{-\exp\left[-\left(\frac{H-\delta}{\xi}\right)\right]\right\} \tag{84}$$

(where α, β, γ, δ, ξ, are parameters found from data) and then extrapolated to get the design wave with a given probability of exceedence.

12. Random Wave Forces

12.1 *Spectral Wave Load Description*

Application of Morison's Equation: Morison's equation[15] may be rewritten as

$$F = K_i \dot{u} + K_d |u| u \tag{85}$$

where $K_i = C_M \rho \pi D^2/4$ and $K_d = C_D \rho D/2$

In order to apply this equation statistically, we have to assume that, in the short term, long crested sea is statistically stationary and that the probability distribution of surface fluctuations $\eta(t)$ is Gaussian normal. Since u and $\dot{u}$ are linear functions of $\eta(t)$ in the linear theory, their individual distributions are also Gaussian normal (but u^2 may not follow the Gaussian distribution) and their linear combination in turn gives the force values also distributed in a Gaussian normal way. The probability density of force is given by

$$p(f)=\frac{1}{\sqrt{2\pi}\sigma_f}\, e^{-\frac{1}{2}(f/\sigma_f)^2} \tag{86}$$

(where σ_f = standard deviation for the force variable f).

However, in field this is found to be true only in case of inertia-dominated conditions.

From the force spectrum, we have the knowledge of the force densities versus the frequencies present in the wave system. Square root of the area under the spectral density curve gives the r.m.s. value of the forces. By applying proper transfer functions to the spectral density values we can get the response spectra. The knowledge of various wave force levels associated with various known probabilities of occurrence obtained is very useful in designing the ocean structure. Various probability laws can be obtained from the force spectrum.

Fundamental Non-linear Spectral Representation

The force spectral density by definition is the Fourier transform of the force covariance function. Borgman[1,2] gives expressions for force covariance function,

$$R_{FF}(\tau)=K_d^2\,\sigma_{u\,u}^4\,G[R_{uu}(\tau)/\sigma_{u\,u}^2]+K_i^2\,R_{AA}(\tau) \tag{87}$$

where $R_{uu}(\tau)$ and $R_{AA}(\tau)$ are the covariance functions of the horizontal component of water particle velocity, $u(t)$ and local acceleration, $A=\partial u/\partial t$. The variance $\sigma_{u\,u}^2$ of the horizontal water particle velocity is given by

$$\sigma_{u\,u}^2=2\int_0^\infty S_{uu}(f)\,df=R_{uu}(0) \tag{88}$$

and

$G(r)=G(R_{uu}(\tau)/\sigma_{u\,u}^2)$ is given by the expression

$$G(r)=[(2+4r^2)\arcsin r+6r\sqrt{1-r^2}]/\pi$$
$$=\frac{1}{\pi}\left[8r+\frac{4r^3}{3}+\frac{r^5}{15}+\frac{r^7}{70}+\frac{5r^9}{1008}+\cdots\right] \tag{89}$$

and the series converges rapidly for $0\leq r\leq 1$. It has been found that for $r=1$, $G_1(r)$ taking only the first term differs from $G(r)$ by only 15 percent and the cubic approximation $G_3(r)=\frac{1}{\pi}(8r+4r^3/3)$ differed from $G(r)$ by only 1·1 percent. Substituting the series in eqn. (87), we have

$$R_{FF}(\tau)=\frac{K_d^2\,\sigma_{u\,u}^4}{\pi}\left[\frac{8R_{uu}(\tau)}{\sigma_{u\,u}^2}+\cdots\right]+K_i^2\,R_{AA}(\tau) \tag{90}$$

Its Fourier transform gives the desired spectral density of the force

$$S_{FF}(f)=\frac{K_d^2\,\sigma_{u\,u}^4}{\pi}\left[\frac{8S_{uu}(f)}{\sigma_{u\,u}^2}+\cdots\right]+K_i^2\,S_{AA}(f) \tag{91}$$

Linearization

Retaining only the first term, the series expansion in eqn. (89) becomes $G(r)=8r/\pi$ which is equivalent to writing $u|u|=u\,\sigma_{uu}\sqrt{8/\pi}$. Then we have

$$S_{FF}(f)=\frac{8K_d^2\,\sigma_{u\,u}^2}{\pi}S_{uu}(f)+K_i^2\,S_{AA}(f) \tag{92}$$

Fig. 12 shows a typical plot of the force spectrum evaluated using this linearization along with that of corresponding wave spectrum.

This linearization is found to be fairly accurate when drag forces are negligible, i.e., in inertia-dominated situations. It has been found to have many advantages also, e.g., the probability density function of force becomes Gaussian and hence probability density of peak force and force range can be found by Rayleigh density.

Transfer Functions

The spectral densities of velocities (S_{uu}) and accelerations (S_{AA}) in eqn. (92) are found by applying transfer functions to the spectral density of water surface elevations ($S_{\eta\eta}$). The principle involved in evaluating transfer functions is as follows.

Assuming that the short-term sea surface for a long crested sea is statistically stationary with Gaussian probability distribution, then using first order wave theory (eqns. 7 and 9)

$$S_{uu}(\omega)=\left[\frac{\omega^2\cosh^2(kz+kd)}{\sinh^2 kd}\right]S_{\eta\eta}(\omega) \tag{93}$$

and

$$S_{AA}(\omega)=\left[\frac{\omega^4\cosh^2(kz+kd)}{\sinh^2 kd}\right]S_{\eta\eta}(\omega) \tag{94}$$

where quantities in the brackets are the respective transfer functions T_{uu} and T_{AA} respectively. Change of cosh $k(z+d)$ into sinh $k(z+d)$ gives transfer

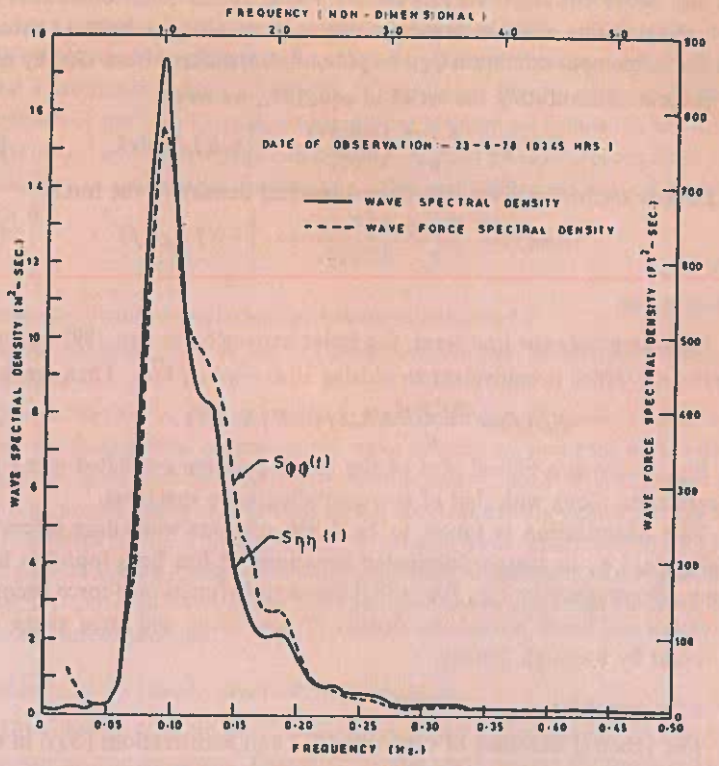


FIG.12 WAVE AND FORCE SPECTRA

functions in the preceding expressions in the vertical direction. It must also be remembered that the covariance functions and spectral density functions are Fourier transforms of each other with the result

$$R_{uu}(\tau)=\int_{-\infty}^{\infty} S_{uu}(f)\, e^{i2\pi f\tau}\, df \tag{95}$$

and

$$R_{AA}(\tau)=\int_{-\infty}^{\infty} S_{AA}(f)\, e^{i2\pi f\tau}\, df \tag{96}$$

Borgman[1] found that transfer function for velocity was more of less constant with frequency variation whereas that for acceleration considerably

changed with frequency. This can be seen from their expressions also. Paape[17] performed experiments in the laboratory with irregular waves and observed that transfer functions for total force varied significantly but uniformly with frequency for large diameter bodies. The difference between transfer function values found theoretically as well as experimentally for small diameter bodies (involving higher non-linear terms) was, however, considerable.

Research work is currently going on to evaluate transfer functions considering free stream effects and discontinuous loadings to which members near MWL are subjected.

12.2. *Probabilistic Wave Load Description*

The spectral description of wave loads is oriented around the 'variance' of the wave force process. The probabilistic models for wave load description, which includes a variety of probability density and distribution functions for short as well as long terms are very useful in getting the design wave load values with a known probability of occurrence. There are three formulations available in this regard, namely, semi-deterministic, nonlinear and linear models.

Semi-Deterministic Method

The first 'semi-deterministic' model uses both deterministic as well as non-deterministic analysis techniques. It uses the individual wave-by-wave approach. The individual wave height and period values are used to generate flow kinematics by Stokes' fifth-order theory which in turn are put in Morison's equation to get peak force during one wave cycle. This is again weighted by the probability of occurrence of the sea state to get the peak distribution function. Thus this method gives estimation of peak distribution. Tickell *et al.*[28] have found that this method made good predictions of peak loads after proper selection of wave height and period values.

Linearization

With Borgman's linearization of the drag term indicated earlier we can write the variance of observed forces as

$$\sigma_{FP}^2=\frac{8}{\pi} K_D^2\, \sigma_{uu}^4+K_i^2\, \sigma_{AA}^2 \tag{97}$$

The value of σ_f can be put into the expression for probability density function for force in eqn. (86) to get other probability distribution like short-term and long-term, cumulative as well as peak distributions. This linearization is valid in inertia-dominated flow conditions with negligible drag forces.

Nonlinear Models

(*a*) *Short-term Models* : The probability density function for force becomes non-Gaussian if the squared drag term in the Morison's equation is to be taken into account. The expression for it is obtained by considering the joint probability of occurrence of u and $\partial u/\partial t$. Thus,

$$p(F)=\frac{1}{2\pi K_i \sigma_{uu} \sigma_{AA}} \int_{-\infty}^{\infty} \exp - \left[\frac{\frac{u^2}{\sigma^2_{uu}}+\frac{(\partial u/\partial t)^2}{\sigma^2_{AA}}}{2}\right] du \qquad (98)$$

where F is the horizontal force per unit length and is a general force variable. Sometimes we are interested in the peak values of F as in problems pertaining to designing by fatigue failure loads.

The distribution of peak values of the force involves consideration of the trivariate process given by $p\left(f, \frac{\partial f}{\partial t}, \frac{\partial^2 f}{\partial t^2}\right)$ where $\frac{\partial f}{\partial t}$ and $\frac{\partial^2 f}{\partial t^2}$ are the first two time differentials and the resulting expression is extremely complicated to solve. The expression becomes simple if the force spectrum is assumed to be narrow banded and its time history to be independent of $\frac{\partial f}{\partial t}$. Then

$$p(F)=1{\cdot}0-\frac{p(f=F)}{p(f=0)} \qquad (99)$$

These assumptions are valid for negligible drag conditions.

(b) *Long-term Models* : The expressions (98) and (99) represent short-term conditions characterised by one set of H_s and $\overline{T_z}$ values. The knowledge of long-term distributions (which is required in problems pertaining to designing by first excursion failure) is obtained by weighting above force densities by the probabilities of occurrence of the sea state and integrating these quantities over the histogram of H_s and $\overline{T_z}$ values.

$$p_{LT}(F)=\int_0^{\infty}\int_0^{\infty} p\left(\frac{F}{H_s, \overline{T_z}}\right) p(H_s, \overline{T_z})\, dH_s\, d\overline{T_z} \qquad (100)$$

Long-term probability of peak forces, as against above individual forces F, involves division by proper zero-cross period in order to convert time into number of peaks.

Tickell *et al.*[23,24] have found that the long-term distribution of peak force can be rapidly and satisfactorily computed from linearized and narrow band assumptions as seen above, provided drag is negligible. When drag is significant, linearized methods underestimate peak forces.

12.3 *Statistics of the Response*

The behaviour of the structure towards the incident ocean waves is termed as 'Response'. The response of a structure in a long-crested (i.e., unidirectional) sea can be described by a Response Spectrum as

$$\Phi(\omega)=E(\omega)\, T^2(\omega) \qquad (101)$$

where $\Phi(\omega)$ = response spectrum (which may typically be a force or pressure spectrum) for frequency ω,
$E(\omega)$ = sea surface spectrum, and
$T(\omega)$ = transfer function which is output of the linear system to unit input.

This is a scalar response spectrum. The directional response spectrum for multi-directional seas is obtained by

$$\Phi(\omega, \theta)=E(\omega, \theta)\cdot T^2(\omega, \theta) \qquad (102)$$

where $E(\omega, \theta)$ = directional wave spectrum for angle of incidence θ.

Integration of this directional spectrum through all directions gives the scalar spectrum $\Phi(\omega)$ in equation (101), i.e.

$$\Phi(\omega)=\int_{-\pi}^{\pi} \Phi(\omega, \theta)\, d\theta \qquad (103)$$

The transfer functions can be derived theoretically or may be found by model experiments.

Short-term Response Distribution

If linear relationship between response amplitudes X and incident wave amplitudes a is assumed, then the distribution of X becomes similar to that of a, i.e., the short-term distribution of X can be given by Rayleigh distribution as follows :

$$P(X)=1-\exp\left[-\frac{X^2}{Ex}\right] \qquad (104)$$

where

$P(X)$ = cumulative distribution of X

$Ex = 2\sigma_x^2$ where σ_x^2 = variance of the continuous random variable X or area under the response spectrum

$$= \sum_{\text{all } \omega} \varphi(\omega)\, \Delta\omega = \sum_{\text{all } \omega} E(\omega)\, T^2(\omega)\, \Delta\omega \qquad (105)$$

Analogous to short-term wave statistics, most probable largest response for given N responses is

$$X_{\max} = \sqrt{Ex}\ \ln N \qquad (106)$$

It can be seen from the above equations that knowledge of Ex is sufficient to describe short-term response of the structure.

Long-term Response Distribution

The long-term distribution of response variable X is obtained by weighting each short-term distribution given by eqn. (104) by corresponding probability of occurrence of the sea state. A better approximation to it can also be provided by Weibull distribution,

$$p_{LT}(X) = 1 - \exp\left[-(A/b)^k\right] \qquad (107)$$

where

$$A = (X/a)^2 \qquad (108)$$

a, b, k are the parameters of the distribution determined by the measured data. Brunn[5] has discussed at length the method to determine them.

REFERENCES

1. Borgman, L. E. (1965) 'The spectral density for ocean waves', Univ. of California, Berkeley.
2. Borgman, L. E. (1972) 'Statistical methods for ocean waves and wave forces', Recent Advances in Hydroscience, Vol. 8, Academic Press.
3. Borgman, L. E. (1976) 'Statistical analysis of sea waves in coastal wave hydrodynamics', M.I.T. Lecture Notes, Summer Course, Ch. 2.
4. Bretschneider, C. L. (1963) 'A one-dimensional gravity wave spectrum', Proc. Conf. on Ocean Wave Spectra.

5. Brunn, P. (1976) 'Port Engineering', Ch. 3, Gulf Pub. Co., Houston, Texas.
6. Dattatri, J. (1978) 'Analysis of regular and irregular waves', Ph.D. Thesis, IIT Madras.
7. Dean, R. G. (1965) 'Stream function representation of nonlinear ocean waves', *Jour. of Geophy. Res.* 70, No. 18.
8. Dean, R. G. (1967) 'Relative validities of water wave theories', Proc. First Coastal Engg. Conf., San Francisco.
9. Hesselman, K. *et al.* (1973) 'Measurement of wind wave growth', Deutsches Hydrographisches Inst., Hamburg.
10. Hogben, N. (1974) 'Fluid loading on offshore structures—a state of art appraisal', Maritime Technology Monograph, Royal Inst. of Naval Architects, also in BOSS, 1976.
11. Kinsman (1965) 'Wind Waves', Prentice Hall, Inc.
12. Longuet-Higgins (1952), 'On the statistical distribution of heights of sea waves', *Jour. of Marine Res.*, IX-3 : 245-268.
13. Longuet-Higgins (1952) 'The statistical analysis of a random moving surface', *Phil. Trans. Royal Soc. London*, A, 249 : 321-387.
14. MacCamy, R. C. and Fuchs, R. A. (1954) 'Wave forces on piles—a diffraction theory', Beach Erosion Board, Tech. Memo. 69.
15. Morison *et al.* (1950) 'The force exerted by surface waves on piles', *Petroleum Trans.*, 189, No. 7P 2846 : 149-154.
16. Narasimhan, S. and Deo, M. C. (1979) 'Probabilistic analysis of ocean waves', Project Report No. I, IIT Bombay. Also 'Spectral analysis of ocean waves—a study', Conf. on Civil Engg. in Oceans IV, San Fransisco, Sept. 1979.
17. Paape, A. (1968) 'Wave forces on piles in relation to wave energy spectra', Proc. Conf. on C.E., London, Sept. 1968, Ch. 61.
18. Pierson, W. J. and Muskowitz (1964) 'A proposed spectral form for fully developed seas', *Jour. of Geophy. Res.*, 69, 24 : 5181-5190.
19. Pierson, W. J. and Holmes, P. (1975) 'Irregular wave forces on a pile', Journal of WH Div., *Proc. ASCE*, Nov. 1965.
20. Sarpakaya, T. (1976) 'Vortex shedding and resistance in harmonic flow about smooth and rough circular cylinders', BOSS, 1976, Troudheim.
21. Scott, J. R. (1965) 'A sea spectrum for model tests and long term ship prediction', *Jour. of Ship Res., SNAME*, 9(3) : 145-152.
22. Skelbreia, L. and Hendricson, J. (1960) 'Fifth order gravity wave theory', Proc. Coastal Engg. Conf. : 184-196.
23. Tickell, R. G. *et al.* (1976), 'Long term wave loading on offshore structures', Proc. I.C.E., 2 : 61.
24. Tickell, R. G. and Holmes, P. (1978) 'Approaches to fluid loading', Ch. 2 in 'Numerical Methods in Offshore Engineering' by O. C. Zienciwicz, *et al.*, John Wiley and Sons.
25. Tucker, M. J. (1963) 'Analysis of records of sea waves', Proc. Inst. of C.E., London, 26 : 305.

26. Wiegel, R. L. (1964) 'Oceanographical Engineering', Prentice Hall Inc.
27. Wiegel, R. L. (1980) 'Design Wave', Lecture Notes for Short Course on Small Harbour Engineering, IIT Bombay.
28. Zienciwicz, O. C. *et al.* (1978) 'Numerical methods in offshore engineering', John Wiley and Sons.

A Functional Approach to the Origin of Life Problem

Krishna Bahadur[1], S. Ranganayaki[1], Clair Folsome[2] and Adolph Smith[3]

[1]Chemistry Department, University of Allahabad
[2]Laboratory for Primordial Biology, University of Hawaii, Honolulu, HI 96822
[3]Physics Department, Concordia University, Montreal, now at Ames Research Center, 239-12, Moffett Field, NASA, California, 94035

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

1. The Problem in Origin of Life Studies

To explain how life began is surely one of the most fundamental aspects of biological research, yet today we see studies in origin of life generally as an esoteric field away from the main stream. Why ? For one thing there has been little progress in the field in the past decade. During the 1950's with the announcement of the Miller-Urey experiment (1953) origin of life became a much discussed subject which found its way into standard textbooks. During the succeeding decade a variety of presumed prebiotic experiments in which various energy sources such as light, electric discharges, and acoustic shock waves were carried out. A wide variety of simple biochemicals were found to be products in these experiments. Some work was done on polymers but most of the work centered about identifying and listing the smaller organic molecules such as amino acids and nucleic acid bases formed from simple precursors.

The gap between a collection of biochemicals and the properties which are intrinsic to the simplest living organism remained just as large after two decades. There has also been the objection that origin of life work is meaningless because it is largely a matter of conjecture as to what the conditions on a prebiological earth were : hence the relevance of any simulation experiment was suspect.

We want to question the fundamental assumptions of present-day work and to suggest an alternative simplifying approach which bridges the gap between biochemicals and simple cell functions. Experimental results which demonstrate the utility of a functional approach to the origin of life problem will be presented.

2. Modern Ideas on Origin of Life

Modern ideas on the origin of life come from the theories of Oparin[25] and Haldane[23]. They assume a primitive, reducing atmosphere (methane, ammonia, hydrogen and water) so that complex molecules could arise. The reducing atmosphere was thought necessary so that organic molecules would not be oxidized. With sources of high energy in this atmosphere (electric discharge, UV, etc.) it was thought that a variety of organic molecules would be produced. Since there were no organisms to consume these molecules, they accumulated in the primitive oceans to form a primordial "soup". There is a great difference of opinion on the concentration of the supposed soup

from nearly zero[45] or an amino acid concentration of 10^{-7} M[19] to one gram per liter.

At this point, the heterotrophic hypothesis, originally advanced by Horowitz[26] is invoked. By some chance a primitive organism is assumed to have arisen out of the soup mixture and this primitive organism used the soup components for further metabolism and growth. Based upon the heterotrophic concept, Oparin adduces his concept of coacervates which arise from complicated molecules and then absorb various compounds from the environment. In this way, the coacervates are conceived as evolving into primitive cells.

Fox has synthesised microspheres starting from heated dry amino acids and has formed some sort of cell-like macromolecular structures and proteinoids. However, his experimental conditions seem not to be geologically plausible and the arbitrary properties he lists as being lifelike are not from any one microsphere preparation but rather from many different microspheres produced under different conditions. In a summary of the prevailing heterotrophic hypothesis, Keosian[27] writes :

> "Life at its origin is envisioned by some as a simplified version of a present-day 'primitive' heterotroph-simplified not in the reduction of metabolites or metabolic pathways, but in the substitution of these by simpler substances and simpler reactions supposedly performing the same functions. One need only scan the claims of chemical evolution over the past two decades to see the trend. There seems to be a general belief in an abiotic mechanism in a prebiotic soup. However, the harmonious coordination of separate reactions into a smoothly functioning, interconnected metabolic network in even the simplest heterotroph is the result of untold millions of years of evolution. Turning that evolutionary clock backwards is not just a matter of substituting simple substances for complex ones or simple reactions for complex pathways. The concept that all parts of the first living thing pre-existed, and that its formation was simply a matter of spontaneous generation therefrom is a mathematical absurdity, not probability. All present approaches to a solution of the problem of the origin of life are either irrelevant or lead into a blind alley. Therein lies the crisis."

We present a conceptual and experimental approach which is based on wellknown catalytic properties of minerals in cell-sized domains, uses an

autotrophic model, and has already yielded results which narrow the gap between biochemicals and the functions manifested by living cells.

3. The Autotrophic Hypothesis

The place of autotrophy in origin of life theory has been given scant attention. Geochemists have objected to the concept of a primitive soup on several grounds. One major objection has been that the "primitive soup" should have resulted in formation of large areas containing organic sediments but no such areas have been found. In a recent review, Brooks and Shaw[17] conclude :

"We can conclude with some certainty that :

(a) There never was any substantial amount of 'primitive soup' on Earth when ancient Precambrian sediments were formed ; and that

(b) If such a 'soup' ever existed it was only for a brief period of time. If we subtract the idea of a substantial amount of 'primitive soup' and a long period of time from the basic concept of Chemical Evolution Theory, there is very little time left."

If there were not any 'primitive soup' we must assume that the first organisms were autotrophic, that is, they obtained their carbon from CO_2 and their nitrogen from atmospheric N_2. Cairns-Smith[18] proposed that the first organisms were clay-like. The minerals of the clay would catalyze the formation of new materials and act as templates for replication.

About the chemical nature of the living system Pirie[85] holds that proteins are necessary for present-day cells because they are enzymes and catalyze many biochemical reactions. But many metallic ions also show enzyme-like activity and it is possible that to begin with these might have been performing the work of protein enzymes. Bernal[14] was of the opinion that there might have been organisms which had only inorganic catalysts in the place of proteins and such organisms might have been sluggish but could have certainly performed all the functions of life. According to Smirnova[49] many inorganic substances have physical properties commonly found in organic substances and it is quite possible to conceive organisms made of these inorganic substances only. Bernal[15] writes "unless it is desired to push back the doctrine of special creation to the creation of enzymes and co-enzymes (there is a school that would take one of these, namely the co-enzymes in the polymerised form as nucleic acid) as the beginning of life, unless then we are prepared to take such an easy way out, we must assume that before

there were enzymes to carry out the catalytic reactions in metabolism there were some other agents that did it, not so well, but sufficiently well for the slow time of the origin of life." These evolve to produce protein-nucleic acid cellular life which we now observe on the earth. Bahadur[3,6] has suggested a probable locale for these forms of life. These microstructures could have shown the functional properties of living system as clay-like models.

One of the important aspects of the problem of origin of life is to study the physico-chemical factors which brought the molecules which formed the earliest living systems together, kept them held together and arranged them in some specific pattern that could show the properties of biological order.

A recent review of autotrophic origin by Hartman[34] uses a scheme in which clays, transition metals, disulfide, dithiols, ultraviolet radiation and cyanide ion from a primitive metabolic system which uses CO_2 and N_2 fixation and which could evolve in a simple environment. He looks at intermediary metabolism, and examines the central role of the citric acid cycle, and concludes that the main problem of starting the cycle is one of carbon dioxide fixation. He considers that nitrogen fixation could have been accomplished by molybdenum complexes in the primitive organisms if a strong reductant were available. It has been observed that the mixtures of jeewanu which start with simple substances have almost all the acids of the Krebs' cycle[44]. It has been suggested that Krebs' cycle was perhaps the most elementary form of metabolism.

4. The Functional Approach to Origin of Life

We present here an approach to origin of life which emphasises the matter-energy flows which are fundamental to all living organisms and their ecologies. In the course of the many prebiotic experiments, microstructures have appeared but generally these structures and their properties have been ignored. The reason for this has not been a case of oversight but rather is in line with the mainstream of procedure in biochemistry. A major activity of biochemists has been the separation, identification and purification of various biochemicals involved in each stage of metabolism of a structured once-functioning system.

In the first Miller-Urey experiments polymeric material in the form of microstructures was observed. Yet surprisingly these structures were not studied systematically until recently[20]. In the 1963's, a series of articles by

Bahadur and coworkers[1,3,5,7,9] on microstructures formed from transition metals, ammonium phosphate, biological minerals and formaldehyde appeared. These structures were called "Jeewanu" (a Sanskrit word for particles of life). Various catalytic properties of these particles were reported but these were largely ignored.

Jeewanu are photochemically synthesized particles formed in sterilized aqueous mixtures containing simple organic source as formaldehyde, inorganic nitrogen and minerals commonly found in cells. These particles have definite boundary wall and intricate internal structures[52,58]. They are very similar to the present-day cell in chemical composition and differ from the common microorganisms that they cannot be grown on any known bacterial culture medium[2,5,7]. These particles multiply by budding and the small buds grow to maturity size and bud again. Wherever organic materials and necessary inorganic substances are present, sunlight synthesized amino acids, peptides, sugars and such other biochemicals and the transitional metal coordination complexes with ligand as these biochemicals with the help of surface of semi-conductor could organise in the form of microstructures and form jeewanu[2,20,30,44,46,47,54,55,56,59]. These jeewanu were capable of adaptability and so these evolved into the present-day cellular life[7].

Bahadur and Ranganayaki[9] photochemically produced another variety of these particles by the interaction of ammonium molybdate, diammonium hydrogen phosphate, minerals commonly found in cells and formaldehyde in aqueous mixtures which have 31·21% of molybdenum.

In the jeewanu the properties of growth, multiplication and metabolic activity are observed in a natural way, once the experiments are set and no specific chemicals are needed for any specific property[1,2,5,7].

These particles with high molybdenum content have a number of amino acids in free form and in combined form as peptides, and sugars as ribose, deoxyribose, fructose and glucose. They have distinct boundary wall and intricate internal structure. The internal structures can be clearly viewed under high magnification. The particles under phase-contrast microscope reveal the boundary wall and the internal structure of the particles clearly[10]. These particles multiply by budding (Bahadur and Ranganayaki 1970).

On separation of the particles from the mixture, extraction with chloroform : methanol : : 80 : 20 in a soxhlet yields a viscous yellow liquid. This contains an ethyl alcohol soluble compound which on chromatography gives the test for phospholipid[47]. On hydrolysis of the particles with perchloric

acid or formic acid in a sealed tube the products give the tests for nucleic acid bases as adenine, guanine, cytosine, thymine and uracil[39,58]. If the particles are kept in 1*N* sodium hydroxide for 24 hours, filtered and the filtrate acidified with dilute acetic acid a white precipitate is obtained which on subsequent hydrolysis gives the test of nucleosides[40].

The particles can be fixed with chromic acid and subsequently stained with gentian violet and then eosin. In the central portion chromatin-like blue structures are seen and the portion outside the central zone stains red with eosin, like cytoplasm[8]. It has been reported that the materials of the particles on digestion with hydrochloric acid show strong optical activity[9].

Briggs[16] reported esterase and phosphatase activity in these particles. Bahadur and Ranganayaki[9] reported urease-like, ATP-ase like and peroxidase-like activity in these photochemically formed microstructures. Jeewanu are sensitive to the presence of antibiotics in the irradiative mixture[11,36]. Presence of sulphur drugs also affect their growth[36].

The factors which are responsible for the natural formation of objects with a definite morphology have been discussed by Bahadur[3,5,7]. The molecules of different chemicals are brought together and are first held by coacervate forming factors. If this contains macromolecules with low M.W., molecules attached by various intermolecular forces, such as van der Waals forces, hydrogen bonding, hydrophobic bonding, molecular bonding and others, and also many other molecules adsorbed, absorbed and held together by electrostatic forces, on crystallisation, a highly deformed crystal is formed and the whole thing results in a molecular mesh having wide gaps and passages through which small environmental molecules have specific permeability. The various gaps contain different molecules held in this deformed crystal structure and remain active chemically and also catalytically. The whole structure attempts to acquire a spatio-energetic pattern representing the state of minimum energy and it results in biological looking structures. The outer material forms a boundary wall and the aggregate appears to have an intricate internal structure. If this structure is present in an appropriate environment, containing molecules which can form its body material, and if it has a source of energy, which may preferably be obtained by irradiation, the environmental molecules enter the aggregate through the appropriate passages in the boundary wall from the outer material of the aggregate, interact and finally result in the material from which the aggregate is formed[3,4,5,7].

All such abiogenic morphological structure may not be able to show the

properties of biological order, but those which happened to be in the environment to allow the flow of energy appropriate for them could show these properties. Of such innumerable particles those which depended on such materials which were continuously being formed in the mixture—say by photochemical process—continued their living activity and the rest ceased their functions soon after the supply of the necessary molecules was finished[3,6,7].

Many such microstructures are observed in ores and rocks and are reported as silicate particles, found in sedimentary rocks and are described as microfossils and observed in carbonaceous chondrites and mentioned as "Organised Elements". Many of these may be models of the earliest structures in which life was expressed.

The presence of transitional metal in these microstructures which might easily form coordination complexes with ligands as amino acid, peptide, nucleic acid bases could make these structures extremely reactive for a number of chemical transformations.

In 1971, microstructures resulting from ultraviolet irradiation of ammonium thiocyanate and ferrous sulphate were reported to show pH changes upon further irradiation[51].

Earlier microstructures have been produced by the action of ultraviolet light on aqueous mixture of ammonium thiocyanate and formaldehyde[45,49,50]. This mixture was modified by introducing a mineral solution containing sodium chloride, potassium sulphate, calcium acetate, magnesium sulphate and potassium dihydrogen phosphate. In addition to this mineral solution the concentration of which was as usually used in microbial culture excess of potassium hydrogen phosphate and calcium acetate was added to provide a more inorganic base. In another mixture ferric chloride was added. Three sets of two such mixtures were prepared. One set of iron and non-iron containing mixtures was exposed to. artificial electric light from a 100-watt electric bulb, another was covered with several folds of thick black cloth as the dark set and the third was similarly covered with black cloth and kept in a lead chamber of 1 inch thick solid lead walls. More microstructures were formed in the mixtures exposed to light. The mixture showed the formation of amino acids and peptides. More synthesis of amino acid was observed in the exposed mixture, lesser in the mixture kept in dark and least in the mixtures kept in lead chamber. Ferric chloride containing mixtures showed formation of more amino acids than the mixtures without

ferric chloride. In each mixture cell-like microstructure formation was observed which showed the presence of amino acids and peptides.

The experiments indicate that for the formation of microstructures chemical reactions as well as irradiation are helpful and the coordination complexes of transitional elements with ligands as amino acids and peptides as suggested by Beck[13] together with the catalytic functions of the insoluble compounds of semi-conductor elements with their increasing surface as suggested by Tsigdinos and Swanson[67], were effective in these processes.

In an attempt to study the "obvious", the simplest microstructures in prebiotic experiments, we consider globally the metabolism of the simplest living cells, the blue-green algae. We can view the flows into a blue-green algae as Figure 1.

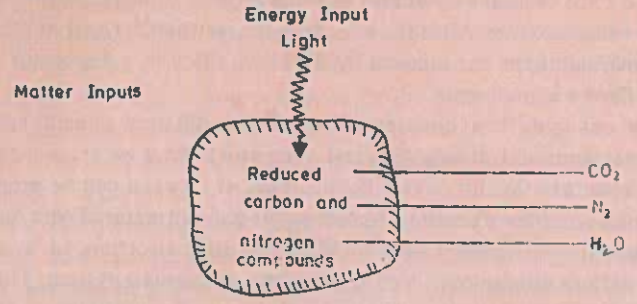


Fig. 1

Global matter-energy flows in a blue-green algae.

The global overview suggested several catalytic properties to search for in the microstructures, i.e. nitrogenase activity, carbon dioxide reduction, water photolysis. Since transition metal complexes have been intensively studied for their catalytic properties in promoting the above reactions, we decided that the microstructures studied by Bahadur would be a good starting point since they contain transition metals closely associated with organic compounds.

5. Experimental Results

Molybdenum complexes with amino acid and small peptide ligands have been reported by Schrauzer[48] to have nitrogenase activity, so this function was selected for initial testing.

As jeewanu are synthesised photochemically a search of the functional properties of splitting of water in sunlight and utilisation of hydrogen thus produced is obvious. A typical jeewanu mixture is prepared as follows : 1 vol of 4% w/v of ammonium molybdate solution, 2 vol of 3% w/v of diammonium hydrogen phosphate solution and 1 vol of mineral solution prepared by dissolving 20 ml of each of sodium chloride, potassium sulphate, calcium acetate, magnesium sulphate, manganese sulphate and potassium dihydrogen phosphate and 50 mg of ferrous sulphate in 100 ml of distilled water are mixed together and the mixture is cotton plugged and sterilised. After cooling 1 vol of 36% of formaldehyde is added in the mixture aseptically and the mixture is exposed to sunlight under sterilised conditions for 15 hr giving five hr exposure each day for three days. The mixture is shaken once each day by gentle whirling motion to disperse the sediments in the whole mixture. After the exposure the jeewanu formed in the mixture are separated from the mixture by filtration, dried in a desiccator and used for different experiments.

By changing the concentrations of the different constituents of the mixtures and introducing different semi-conductors or transitional metal soluble salts in the mixtures different types of jeewanu can be prepared.

These samples of jeewanu were found to contain material with ferredoxin-like activity[8,7]. Jeewanu have ability to transfer electrons to hydrogenase from sodium dithionite as well as from the chloroplast system. The activity of these compounds are listed in the Table No. 1. The particles are very stable in air at room temperature.

Type and concentration of mediator	Micromoles H_2 evolved per hour	
	per mg chlorophyll 25° 10000 lux	100 m M Dithionite 30°
J I, 1 mg/ml.	4·1	2·4
J II, 1 mg/ml.	1·0	1·3
J III, 1 mg/ml.	1·0	1·0
J IV, 1 mg/ml.	1·8	1·7
15 m M Spinach ferredoxin	40	Not determined
1·25 m methyl viologen	0	50

Table No. 1

H_2 Evolution by various samples of Jeewanu[(J)] with the hydrogenase from *C. pasteurianum*

Ferredoxin is essential for many biological electron transfer processes and all the living cells have ferredoxin in them. A system of chloroplast with its intact membrane and chlorophyll, ferredoxin and hydrogenase can split water molecule on exposure to sunlight in many algae. Jeewanu can be used in this system in the place of ferredoxin under laboratory conditions[39]. Systems containing light harvesting chloroplast membranes and electron mediators to hydrogenase, to produce hodrogen from water, are under study for utilising such systems for energy conversion and storage using sunlight as the primary energy source. Ferredoxin, cytochrome C_3 and NADH are the known physiological electron carriers to hydrogenase. In some instances these carriers can be substituted by viologen dyes. Jeewanu can mediate electron transfer from sodium dithionite or illuminated chloroplasts to *C. pasteurianum* hydrogenase as indicated in Fig. 2

Jeewanu mixtures show reversible photochemical electron transfer[7]. Thus the mixture in which jeewanu is produced, on exposure to sunlight becomes blue in colour. This blue colour is formed by the reduction of molybdenum from colourless(ic) stage to blue coloured(ous) stage. If this blue mixture is covered with black cloth and kept in dark the mixture becomes colourless. The process of making the mixture coloured and colourless can be repeated again and again by keeping the mixture in light and dark respectively. Many ferredoxins have hydrogenase activity also and so the possibility of splitting water by jeewanu in presence of light was investigated. It was observed that so far jeewanu are present in the environmental medium in which they are produced there is no apparent evolution of gas from the mixture. However, if the particles are separated from the environmental medium in which they are produced and put in distilled water and the mixture vigorously shaken and kept in sunlight bubbles of gas start coming out of the mixture after about 15 min of exposure. This evolution of gas continues for about three hours and then gradually stops. If the mixtures are given a second thorough shaking and exposed the evolution of gas continues for the next one and half hours but the yield is poor. After this the gas evolution stops and no amount of shaking and exposure produces any gas. But if this mixture is kept overnight the particles are revived and next day again the mixture liberates gas as on the first day. It is observed that jeewanu are slowly destroyed when kept in distilled water and result in a colloidal suspension. However, they retain their shape for a long time if kept in very dilute mineral solution. So for the study of the photolytic evolution of gas

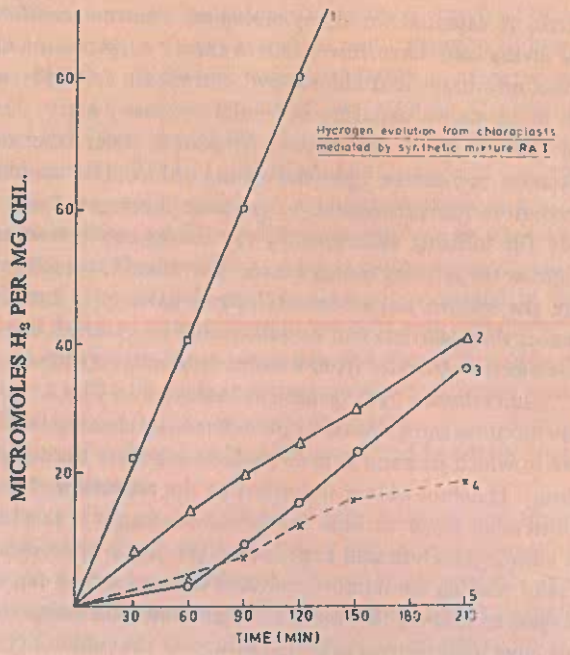


Fig 2

The system contained *C. Pasteurianum* hydrogenase and the components mentioned under Fig. except Ferredoxin. (1) with 30 *N* Moles of FD/2 ML, (2) with 30 *N* Mole FD+4 MG RA I in 2 ML, (3) with 2 MG RA I in 2 ML, (4) with 4 MG RA I in 2 ML, (5 A) with no light and (B) no mediator in light.

from water usually dilute mineral solution is used. The results of these experiments will be published elsewhere.

It has been observed that the jeewanu water mixture evolves gas for about 20 to 30 days and then slowly the evolution stops. Then no treatment produces gas. The evolution of gas by jeewanu is optimum if a certain period of exposure is given in the formation of jeewanu. For the mixture described here the optimum exposure period is 24 hr exposure to sunlight. It is done by giving 4 days' exposure of six hours each day. If the exposure period is considerably shorter or longer, then the particles showed less evolution of gas. The evolution of gas decreases with the aging of the particles and is maximum when freshly prepared.

The gas evolved by the photolysis of water was collected and was found to be a mixture of hydrogen and oxygen. Greater amount of gas was obtained when some air was kept over the irradiated aqueous mixture of jeewanu and water and was smaller when the air above the mixture is evacuated. An atmosphere of nitrogen above the irradiated mixture produces gas which contains more than one volume of oxygen for 2 volumes of hydrogen produced. It was because some of the hydrogen is used up for fixation of nitrogen. If a slow current of the gas is maintained through the irradiated mixture with the same gas circulating again and again though the mixture with a pyrogallol-alkali mixture in the circuit to remove the oxygen formed, the liberation of gas continues for a long time and the circulating gas goes on increasing in volume.

It was further observed that jeewanu and water mixture have nitrogenase-like activity. That is, if acetylene is kept over the aqueous mixture of jeewanu and water containing a little mineral solution and glucose, and the mixture is exposed to artificial light from a Xenon lamp, acetylene slowly gets converted into ethylene and acetylene ethylene ratio slowly increases with increasing period of exposure[52] as indicated in Fig. 3. On passing nitrogen through a mixture of jeewanu and water, molecular nitrogen is fixed and more fixation is observed when only nitrogen is passed through the mixture and lesser when a mixture of nitrogen, oxygen of the composition of air is passed through[12,42].

It was further observed that if carbon dioxide is passed through a mixture of jeewanu, mineral solution and water and the outcoming gas is passed through distilled water, this water starts showing the properties of unsaturated organic compounds collected in it as indicated by its ability to decolourise cold aqueous permanganate solution and bromine water. The unsaturation increases with increasing period of irradiation.[41] It was observed that if sodium bicarbonate, jeewanu, water and mineral solution is exposed to sunlight the bicarbonate carbon gets converted into organic carbon.[41] Carbon dioxide fixation using ^{14}C containing HCO_3^- was investigated and it was observed that the bicarbonate carbon is converted[58] to organic carbon.

That in these processes of nitrogen and carbon dioxide fixation it is the hydrogen of the water molecule which is obtained by the splitting of water molecule and not the hydrogen from any organic material present in jeewanu was investigated using D_2O.[58] They selected the reaction of nitrogen fixation and created acetylene atmosphere over the jeewanu, water mixture and

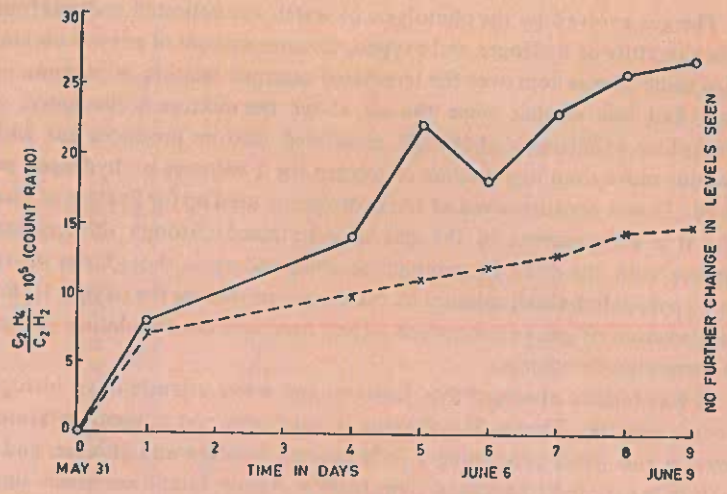


Fig. 3

Acetylene to Ethylene reduction by Jeewanu prepared from Glycine+L-Cysteine+1 : 2 : 1 : 1+Fe Xenon irradiation
1 mg of particles in 1 ml 0·1 M glucose solution.
Control of particles in dark showed O
Control of acetylene+glucose showed O
O particle suspension about 40 cm from Xenon lamp.
X particle suspension about 50 cm from Xenon lamp.
No stirring

observed that the hydrogen of the ethylene formed was mostly deuterium indicating photolytic decomposition of water and utilisation of the liberated hydrogen for its reduction of acetylene.

The particles produced in the ammonium thiocyanate, formaldehyde mixture having additional mineral solution as described earlier show the formation of red cell-like structure if ammonium molybdate was added in the mixture within a few hours of exposure to sunlight. Red coloured microstructures are formed on exposure to sunlight.[4,8] These particles also indicate liberation of gas when the particles are separated from their environment and a mixture of the particles and water is exposed to sunlight.

Conclusion

The earlier living systems might have been primitive autotrophs. There was not enough organic material on the earth during the emergence of life

on the earth. The atmosphere was, in all probability, composed of nitrogen, carbon dioxide and water vapours and solar energy was the chief source of energy available. Some ores of semi-conducting elements had ability of breaking water molecules absorbed on their surface into hydrogen and oxygen by the UV radiation of the sun. Some of these could fix molecular nitrogen and carbon dioxide by the hydrogen thus produced. The oxygen simultaneously produced made the atmosphere less and less permeable to UV. Some of these reduced nitrogen and organic molecules formed by the reduction of carbon dioxide combined and formed compounds as amino acid, peptides and nucleic acid bases. The transitional metal coordination complexes utilising these biochemicals as ligands acted as catalysts for a number of reactions.

The mixture of semi-conductor element insoluble compounds and the coordination complexes formed systems with lots of inorganic surface and little organic materials. Some of these developed the ability of splitting water with visible light and utilising the hydrogen thus produced for the reduction of atmospheric nitrogen and carbon dioxide and the resultant product acted as the energy material for further development. The flow of energy from sunlight to these biochemicals stabilised these earliest microstructures and they developed and evolved with the changing environment and these were the first primitive autotrophs.

The ability of some of the microstructures as jeewanu with perfected property of splitting water molecules with sunlight and fixation of nitrogen and carbon dioxide had advantage over the inorganic minerals because of the increasing concentration of oxygen in the atmosphere and decreasing permeability of UV and could show the other functional properties as growth, multiplication and metabolism. The earliest living system might have been more rich in inorganic material. With evolution and perfection of biological function the organisms increased in organic material and decreased in inorganic constituents though a certain minimum is yet essential and maintained.

REFERENCES

1. Bahadur, K., Ranganayaki, S., Verma, H. C., Srivastava, R. B., Agarwal, K. M. L., Pandey, R. S., Saxena, I., Malviya, A. N., Kumar, V., Perti, O. N. and Pathak, H. D. (1963) *Vijnana Parishad Anusandhan Patrika*, 6 : 63.

2. Bahadur, K., Ranganayaki, S., Verma, H. C., Srivastava, R. B., Agarwal, K. M. L., Pandey, R. S., Saxena, I., Malviya, A. N., Kumar, V., Perti, O. N. and Pathak, H.D. (1964) *Zbl. Bakt.*, **117** (2) : 567.
3. Bahadur, K. (1964) *Zbl. Bakt.*, **118** (2) : 671.
4. Bahadur, K., Perti, O. N. and Pathak, H. D. (1965) *Agra Univ. J. Res. (Sci.)*, **13**, (II) : 197.
5. Bahadur, K. (1966) "Jeewanu, The Protocell", Ramnarainlal Beni Prasad, Allahabad.
6. Bahadur, K., Ranganayaki, S., Kumar, A. and Srivastava, P. (1966) *Zbl. Bakt.*, **120** (2) : 740.
7. Bahadur, K. (1967) *Zbl. Bakt.*, **121** (2) : 291.
8. Bahadur, K. and Gupta, J. L. (1972) *Zbl. Bakt.*, **127** (2) : 643.
9. Bahadur, K. and Ranganayaki, S. (1970) *J. Brit. Interplanetary Soc.*, **23** : 813.
10. Bahadur, K. (1975) *Zbl. Bakt.*, **130**, (II) : 211.
11. Bahadur, K., Ranganayaki, S., Singh, Y. P. and Kumar, S. (1975) *Revista do Instituto de Antibioticos, Recife*, **15**, N ($\frac{1}{3}$) : 33.
12. Bahadur *et al.* (1979) (*To be published elsewhere*).
13. Beck, M. (1979) *Kemiai Kozlemenyck*, **50** : 223.
14. Bernal, J. D. (1959) Proc. First Int. Symp. on "The Origin of Life on the Earth", Moscow, 1957, Pergamon Press, New York, 78.
15. Bernal, J. D. (1961) Oceanography Am. Association for the Advancement of Science, 95.
16. Briggs, M. H. (1965) *Spaceflight*, 7 (4) : 129.
17. Brooks, J. and Shaw, G. (1978) In Origin of Life, Ed. H. Noda, Center for Academic Publications, Japan, 597.
18. Cairns-Smith, A. G. (1966) *J. Theoret. Biol.* **10** : 53.
19. Dose, K. (1975) *Biosystems*, **6** : 224.
20. Folsome, C. (1975) *Precambrian Research*, **2** : 263.
21. Fox, S. and Dose, K. (1972) Molecular Evolution and the Origin of Life, W. H. Freeman & Co., San Francisco.
22. Gupta, J. L. (1976) Cytological and Histochemical Study of Abiogenic Cell-like Molecular Association, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.
23. Haldane, J. B. S. (1929) The Origin of Life, Rationalist Annual.
24. Hartman, H. (1975) *J. Mol. Evol.* **4** : 359.
25. Herrara, A. L. (1942) *Science*, **96** : 14.
26. Horowitz, N. (1945) *Proc. Nat. Acad. Sci. USA*, **31** : 153.
27. Keosian, J. (1978) In Origin of Life, Ed. H. Noda, Center for Academic Publications, Japan, 569.
28. Kumar, A. (1966) Study of Abiogenesis of Amino Acids, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.
29. Kumar, S. (1976) Abiogenesis of Nucleic Acid Bases, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.

30. Maurya, H. K. (1977) Studies in the Abiogenesis of Amino Acids, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.
31. Miller, S. L. (1953) *Science*, **117** : 528.
32. Narain, K. (1974) Studies in the Abiogenesis and Biopoesis, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.
33. Oparin, A. I. (1929) Proiskhozdenie Zhizni, Moscovsky, Robotchii, Moscow.
34. Orgel, L. E. (1972) Origin of Life, Molecules and Natural Selection, Wiley, New York.
35. Pirie, N. W. (1959) Proc. First Int. Symposium on "The Origin of Life on the Earth", Moscow, 19-24 August, 1957, Pergamon Press, New York, 184-185.
36. Pradhan, D. N. (1977) Effect of Antibiotics and Sulpha drugs on the formation of abiogenic microstructures, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.
37. Rao, K. K., Morris, P. and Hall, D. O. (1978) "Presented at the Workshop Meeting on Hydrogenases : Their catalytic activity, structure and function", held at Gottingen, August.
38. Raina, V. (1973) Studies in Some Aspects of Abiogenesis of Organic Compounds of Biochemical Interest, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.
39. Rao, K. K., Adams, M. W. W., Morris, P. and Hall, D. O., Ranganayaki, S. and Bahadur, K. (1978) Abstract presented at the BASE Symposium, Madurai, India, December.
40. Ranganayaki, S., Raina, V. and Bahadur, K. (1972) *J. Brit. Inter. Planet. Soc.*, **25** (5) : 279.
41. Ranganayaki, S. and Kumar, S. (1979) (*To be published elsewhere*).
42. Ranganayaki, S. *et al.* (1979) (*To be published elsewhere*).
43. Schrauzer, G. N. (1979) *J. Amer. Chem. Soc.*, **101**, 925.
44. Sharma, K. (1977) Studies in the Abiogenesis and Biopoesis, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.
45. Sillen, L. G. (1965) *Arkiv Kemi*, **24** : 431.
46. Singh, R. C. (1973) Studies in the Abiogenesis of Enzyme-like Material, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.
47. Singh, Y. P. (1975) Studies in Abiogenesis of Phospholipids, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.
48. Smirnova, A. Ya. (1959) Proc. First Int. Symposium on "The Origin of Life on the Earth", 19-24 August, 1957, Pergamon Press, New York, 184.
49. Smith, A. E., Silver, J. J. and Steinmann, G. (1968) *Experientia*, **24** : 36.
50. Smith, A. E., Steinmann, G. and Galand, C. (1969) *Experientia*, **25** (3) : 255.
51. Smith, A. E., Raab, K. and Ekpaha-Mensah, T. (1971) *Experientia*, **27** : 648.
52. Smith, A. E. (1979) (*To be published elsewhere*).
53. Smith, A. E. and Folsome, C. (1979) (*To be published elsewhere*).
54. Srivastava, P. (1967) Studies in the Origin and Evolution of Enzymes, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.

55. Srivastava, V. K. (1970) Studies in Chemical Transformations of Abiogenic Molecular Associations, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.
56. Srivastava, M. (1977) Studies in Abiogenesis of Amino Acids, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.
57. Tsigdinos, G. A. and Swanson, W. W. (1978) *Ind. Eng. Chem. Res. Dev.*, **17** (3) : 208.
58. Upadhyaya, G. (1977) Cytological and Histochemical Studies of Self-sustaining Coacervates, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.
59. Verma, M. L. (1970) Studies in Abiogenesis of Molecular Association, D. Phil. Thesis, Chemistry Department, University of Allahabad, India.

Membrane Modifying Linear Polypeptide Antibiotics

R. Nagaraj and P. Balaram

Molecular Biophysics Unit, Indian Institute of Science, Bangalore-560 012, India

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

INTRODUCTION

Peptides possessing antibiotic activity isolated from microbial sources have been the subject of intensive structural and biological investigation over the past two decades. Perhaps, the discovery and widespread use of penicillin, a molecule biosynthetically derived from a tripeptide precursor[87], as a strong antibacterial agent[22], has provided the necessary impetus for the detailed study of microbial peptides. While many of these peptides have not been used clinically, they show unique metal binding properties and often possess the ability to modify the electrical properties or ion permeabilities of artificial lipid membranes[73,82]. Hence, these peptides have been used extensively to study transmembrane ion transport processes in model and natural systems like mitochondria, chloroplasts and plasma membranes[73,82,83].

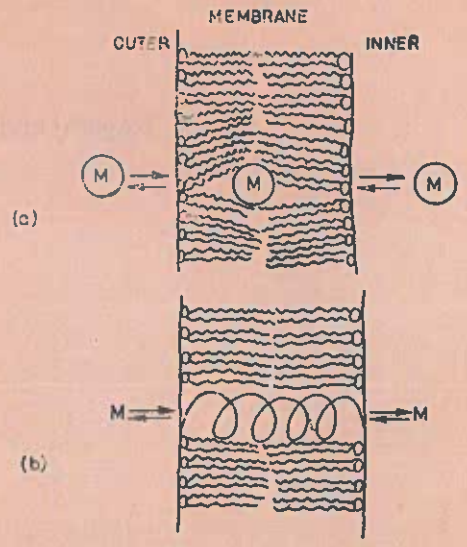


Fig. 1. Passage of metal ions through membranes modulated by (i) carriers ; (ii) channel formers

Peptide ionophores may be broadly classified into two groups. These are : (1) Cyclic peptide ionophores, which function by selective chelation of metal ions and whose metal complexes are lipophilic enough to diffuse through the lipid bilayer[73,82,83] (Fig. 1a). Examples in this category

include the depsipeptides valinomycin and enniatin. (2) Linear acyclic polypeptides which may function as transmembrane channels for ion transport (Fig. 1b). The best studied class of molecules in this category are the linear gramicidins A, B, C (Fig. 2). Another class of linear polypeptide ionophores with unique membrane properties are the microbial peptides containing α-aminoisobutyric acid (H_2N-C-$(CH_3)_2$-COOH, Aib). These include alamethicin[56,76], suzukacillin[48], antiamoebins[74], emerimicins[75], trichotoxins[14] and hypelcins[25]. These peptides have aroused considerable interest in recent years because the electrical properties of lipid bilayers, in their presence, resemble those of excitable membranes isolated from muscle and nervous tissue[9,10,65].

Gramicidin A. HCO-Val-Gly-Ala-D-Leu-Ala-D-Val-Val-D-Val-Trp-D-Leu-Trp-D-Leu-Trp-D-Leu-Trp-NH-CH_2-CH_2-OH

Gramicidin B. HCO-Val-Gly-Ala-D-Leu-Ala-D-Val-Val-D-Val-Trp-D-Leu-Phe-D-Leu-Trp-D-Leu-Trp-D-Leu-Trp-NH-CH_2-CH_2-OH

Gramicidin C. HCO-Val-Gly-Ala-D-Leu-Ala-D-Val-Val-D-Val-Trp-D-Leu-Tyr-D-Leu-Trp-D-Leu-Trp-D-Leu-Trp-NH-CH_2-CH_2-OH

Fig. 2. Sequences of gramicidins A, B and C

The structure and properties of the linear peptide antibiotics form the subject-matter of this review. Since several reports have appeared on the chemistry and biology of the linear gramicidins[85,94,100,101,103] only a brief survey of these peptides is given and greater emphasis is laid on Aib containing ionophores.

Linear Gramicidins

Sequence

The gramicidins (A, B and C) were isolated from *baccilus brevis* by Hotchkiss *et al.* in 1941[38]. Presumably, the absence of acidic and basic functions lead to the postulation of a cyclic structure for the gramicidins (A, B and C) by earlier workers. The linear sequences shown in Fig. 2 were arrived at by Sarges and Witkop[94] by Edman degradation. The presence of the amino terminal formyl group was detected by them after careful hydrolysis experiments, coupled with ^{1}H NMR analysis of the hydrolysates.

Biosynthesis

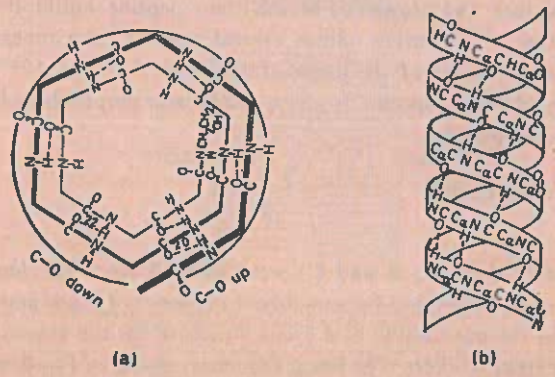
Studies by Bauer *et al.*[3] indicate that the linear gramicidins are biosynthesised non-ribosomally as in the case of gramicidin S and tyrocidine, on a multienzyme complex. A recent report by Akashi *et al.*[1] indicates that the synthesis is initiated by For-Val.

Synthesis

Gramicidin A was synthesised by Sarges and Witkop[95] by classical solution phase methods. The polypeptide has also been subsequently synthesised by solid phase methods[99].

Conformation

Essentially two types of models have been proposed for gramicidin A. Urry[100] has proposed a helical structure involving 20 and 22 membered hydrogen bonded rings. A novel feature of this helix, denoted as π^6(LD) is that the peptide C=O bonds alternate in direction, with orientations primarily parallel and antiparallel to the helix axis (Fig. 3a). On going from N-H to C=O, the L residues in π^6(LD) helices would have their C=O groups parallel to the direction of propagation of the helix and in the D residues, antiparallel to this direction. In 3_{10} or α-helical structures, the C=O bond vectors are all parallel to the helix axis vector. The values of the dihedral angles ϕ and ψ in such a structure would be $-120°$, $+105°$, for



(a) (b)

Fig. 3. (a) Schematic representation of hydrogen bonding pattern in π^6(LD) helix. Ref. 101 ; (b) Schematic diagram of antiparallel double helical gramicidin dimer. Ref. 103.

an L-residue and $+105°$, $-120°$ for a D-residue. According to Urry, a channel obtained by head to head dimerisation of a π^6(LD) helix would have a length of 30 Å and diameter of 4 Å and would account for the ion selectivity and conductance properties observed for gramicidin A. A similar model has been proposed by Ramachandran and Chandrasekaran[85] on the basis of theoretical calculations. In the other model by Veatch *et al.*[103] (Fig. 3b), a double helical structure with intermolecular hydrogen bonds has been proposed. However, no conclusive evidence has yet been obtained by NMR[28,36], IR[103] or X-ray crystallographic[47] studies on gramicidin A favouring the proposed models. A very recent crystal structure analysis of the linear oligopeptide Boc-(L-Val-D-Val)$_4$-OMe has in fact yielded a double stranded left-handed helix, with antiparallel chains[5]. This suggests that for polypeptides with alternating L and D residues the double helical model should be seriously considered.

Membrane properties

The conductance properties of natural and artificial membranes modified by gramicidin A have been studied extensively[2,17,34,37,40,64,98,104,105]. These studies indicate that gramicidin modulates the passage of cations by forming channels rather than by a carrier mechanism, as in the case of valinomycin and enniatin. The conductance-concentration relationship indicate that the conductance channel is a dimer. Recent excitation energy-transfer studies on fluorescent labelled gramicidin A in lipid systems[104] seem to substantiate the conductance mechanism of gramicidin A.

Aib containing antibiotics

Sequence

The sequences of Aib containing microbial peptides are illustrated in Fig. 4. The first Aib containing peptide antibiotic to be studied in detail was alamethicin[77]. The peptide was isolated from the fungus *Trichoderma viride* by Meyer and Reusser[61]. On the basis of amino acid analysis, they reported the composition as (Glu)$_2$(Al)a$_2$(Pro)$_2$(Val)$_2$(Gly)$_1$(Leu)$_1$(Aib)$_7$. The presence of Aib was noted earlier by Kenner and Sheppard[44] in an antibiotic isolated from *Paecilomyces*. The chirality of the other amino acids was established by isolation of the individual amino acids and determination of their optical rotation[61].

The primary sequence of alamethicin was first suggested by Payne *et al*[77]. They observed, as in the case of gramicidin A, that degradative enzymes like chymotrypsin, pepsin, papain and subtilisin did not act on alamethicin. Small fragments were obtained by acid hydrolysis under different conditions and sequenced by the Edman-dansyl method. These studies established the following :

(a) Alamethicin has one free carboxyl group, two amide residues per molecule and no free amino group.
(b) All the peptide linkages are α since the residues could be removed by Edman degradation.
(c) Residues 6 and 18 are amidated, hence the positions of Gln was ascertained.

On this basis the cyclic structure shown in Fig. 4 was suggested.

McMullen[59], on careful examination, found that the alamethicin used in the studies by Payne *et al.*[77] and Meyer and Reusser[61] was a mixture of three components. Martin and Williams[56] used the major component in their investigation of the sequence by NMR methods. On the basis of careful hydrolysis experiments coupled to detailed 270 MHz ^{1}H NMR studies, these authors argued for the presence of amino terminal Ac-Aib and carboxy terminal phenylalaninol (the alcohol derived from phenylalanine). Hence they suggested a linear sequence shown in Fig. 4.

The peptides suzukacillin[43] and trichotoxin[14] were isolated from different strains of *Trichoderma viride*. The emerimicins[74] and antiamoebins[75] were isolated from *Emericellopsis microspora* and *Emericellopsis poonesis*. Similar peptides, the hypelcins[25] have been isolated from *Hypocrea peltata*. Recently, a crystalline peptide having the amino acid composition of alamethicin has been isolated from a microorganism in soil[18]. As in the case of alamethicin, these peptides were found to be heterogenous and the major components were used for sequence determination.

The sequence for suzukacillin was proposed by Jung *et al.*[43], using a combination of gas chromatographic and mass spectral analysis of the fragments obtained by controlled hydrolysis. The chirality of the amino acids was established by gas chromatography using columns packed with optically active substrates. Although the presence of phenylalaninol could be detected by gas chromatography and ^{13}C NMR, these authors were unable to obtain evidence for the presence of Ac-Aib as the amino terminal moiety or the

Alamethicin : Pro-Aib-Ala-Aib-Ala-Gln-Aib-Val(-Aib)
Ref. 77 Glu-Aib-Aib-Val-Pro-Aib-Leu-Gly
Gln

Alamethicin : Ac-Aib-Pro-Aib-Ala-Aib-Ala-Gln-Aib-
Ref. 56 Val-Aib-Gly-Leu-Aib-Pro-Val-Aib-Aib-
Glu-Gln
Phol

Alamethicin I : Ac-Aib-Pro-Aib-Ala-Aib-*Ala*-Gln-Aib-
Ref. 76 Val-Aib-Gly-Leu-Aib-Pro-Val-Aib-Aib-
Glu-Gln-Phol

Alamethicin II : Ac-Aib-Pro-Aib-Ala-Aib-*Aib*-Gln-Aib-
Ref. 76 Val-Aib-Gly-Leu-Aib-Pro-Val-Aib-Aib-
Glu-Gln-Phol

Suzukacillin A : Ac-Aib-Pro-Val-Aib-Val-Ala-Aib-Ala-
Ref. 43 Aib-Aib-Gln-Aib-Leu-Aib-Gly-Leu-Aib-
Pro-Val-Aib-Aib-Glu-Gln-Phol

Antiamoebin : Ac-Phe-Aib-Aib-Aib-Iva-Gly-Leu-Aib-
Ref. 74 Aib-Hyp-Gln-Iva-Hyp-Aib-Pro-Phol

Emerimicin III : Ac-Phe-Aib-Aib-Aib-Val-Gly-Leu-Aib-
Ref. 75 Aib-Hyp-Gln-Iva-Hyp-*Ala*-Phol

Emerimicin IV : Ac-Phe-Aib-Aib-Aib-Val-Gly-Leu-Aib-
Ref. 75 Aib-Hyp-Gln-Iva-Hyp-*Aib*-Phol

Trichotoxin
A-40 : Ac-Aib-Gly-Aib-Ala-Aib-Glu-Aib-Aib-
Ref. 14 Aib-Ala-Aib-Aib-Pro-Leu-Aib-Iva-Gln-
Valol

Hypelcin A : Ac-Aib-Pro-Aib-Ala-Aib-Aib-Gln-Leu-
Ref. 25 Aib-Gly-Aib-Aib-Aib-Pro-Val-Aib-Aib-
Gln-Gln-Leuol

Fig. 4. Sequences of Aib containing microbial peptides

linkage of phenylalaninol. They argued for the presence of the Glu-Gln-Phol sequence on the basis of a common biosynthetic mechanism for alamethicin, gramicidin A and suzukacillin. Trichotoxin was sequenced by Bruckner *et al.*[14] by analysis of fragments, obtained after acidolysis with trifluoroacetic acid, using chemical and mass spectral techniques as in the case of suzukacillin. The hypelcin[25] sequence reported recently has also been similarly derived.

A major contribution in the sequence determination of Aib containing microbial peptides has been made by Rinehart and his coworkers[74,75,76,92],

using gas chromatography, high resolution electron impact mass spectrometry (GC-HREIMS) and high resolution field desorption mass spectrometry. The presence of Ac-Aib and C-terminal phenylalaninol could be unambiguously established. A reinvestigation of the major component of alamethicin originally used by Martin and Williams[56] revealed the presence of two polypeptides[76]. It was proposed that in both polypeptides the C-terminal fragment is Glu-Gln-Phol, with the γ-carboxyl group of Glu being free. Further, the two polypeptides differ in the residue at position 6. The sequence with Ala at position 6 is denoted as alamethicin I while the Aib (6) polypeptide has been designated as alamethicin II. These authors have also suggested use of the term "peptaibophol" to describe peptide antibiotics containing Aib and phenylalaninol. However, this term may not be sufficiently general since the recently described Aib antibiotics, trichotoxin A-40 and hypelcin A, contain valinol and leucinol respectively and do not possess a phenylalaninol residue.

While there has been some confusion surrounding the alamethicin sequence in the literature, a further difficulty has been the recent demonstration that natural "alamethicin" contains as many as 12 components, by high pressure liquid chromatography[58]. It is, however, likely that the major component possesses the sequence of alamethicin I proposed by Pandey *et al.*[76].

Biosynthesis

Information regarding (i) the site of synthesis, (ii) the manner in which the peptide chain is built up, (iii) the synthesis of individual components in the biosynthesis of Aib containing peptides have been provided by the biosynthetic studies on alamethicin and suzukacillin by Reusser[90], Ooka and Takeda[70], Rindfleish and Kleinkauf[91] and Mohr and Kleinkauf[92]. These studies indicate that the synthesis takes place non-ribosomally on a multienzyme primer, as in the case of gramicidin S, tyrocidin and the linear gramicidins[48]. The multienzyme complex, "Alamethicin synthetase" having molecular weight $\sim$ 480,000 has been isolated[91]. *In vivo* and *in vitro* studies indicate that the synthesis of alamethicin is initiated by the formation of Ac-Aib and terminated with the linking of phenylalaninol. In the other Aib containing antibiotics for which biosynthetic studies have not yet been reported, the site and mode of initiation of the synthesis are likely to be similar to alamethicin.

Synthesis

Solid phase and solution phase techniques have been followed in the synthetic approaches to alamethicin. Gisin *et al.*[27] have reported the synthesis of the linear sequence proposed by Martin and Williams[56] (Fig. 4), by the solid phase method. The synthetic alamethicin obtained, however, was not identical to natural alamethicin, by thin layer chromatography. Although the synthetic peptide did show the characteristics of a transmembrane channel former, there were quantitative differences in the electrical properties of lipid bilayers, when compared to natural alamethicin. Titration of the new compounds revealed that the pKa's of the acidic group in natural and synthetic alamethicin were different. On the basis of a study of the model compounds Boc-Glu-γ-benzylester and Boc-Glu-α-benzyl ester it was then presumed that the carboxyl group might be γ rather than α in natural alamethicin. Subsequently Gisin *et al.*[26] synthesised the linear sequence with Phol attached to Gln (19) instead of γ-carboxyl group of Glu (18). The synthetic alamethicin now had a similar pKa value and chromatographic properties as natural alamethicin. On assay of the peptide for antibiotic activity it was found to be 60-80% as potent as natural alamethicin. In bilayer lipid membranes, the specific activity of this synthetic alamethicin was 20-30% of natural alamethicin. It was therefore suggested that these differences might be due to the heterogenous nature of natural alamethicin. However, this observation seems contradictory to the recent observations by Marshall and Balasubramanian (personal communication) that the major component of alamethicin (corresponding to the sequence alamethicin I, Fig. 4) does not show antibacterial activity and only low activity in lipid bilayer membranes.

Marshall *et al.*[55] first attempted the synthesis of the cyclic structure proposed by Payne *et al.*[77] by solution phase methods. These authors attempted cyclisation of the linear 1–17 peptide. The products so obtained did not have ionophore activity. This was attributed to deletion sequences and rearrangement during cyclisation. Subsequently, Marshall and Balasubramanian[53] have reported the synthesis of alamethicin I by solution phase methods. The synthetic peptide obtained did not show any antibacterial activity as in the case of the major component of alamethicin. Urban *et al.*[99] have reported the synthesis of alamethicin by the four component condensation method[99]. Details of the procedure have not yet been published. The synthesis of alamethicin I has been achieved recently in the authors' laboratory by solution phase methods[60]. Mayr *et al.*[58] have reported the synthesis of the

amino terminal 1–11 fragment of alamethicin bound to polyoxyethylene (POE) (molecular mass 10,000) as a solubilising support. Oekonomopulos and Jung[71] have very recently reported the synthesis of a nona-decapeptide analog of alamethicin by solution phase methods.

Membrane properties

The observation, of Mueller and Rudin[65], that the electrical properties of bilayer lipid membranes in the presence of alamethicin was remarkably similar to those of excitable biological membranes, initiated detailed studies on alamethicin-lipid interactions. ^{1}H NMR and electron microscopic studies[35,49,50] suggest that alamethicin molecules penetrate into lipid vesicles and the multilamellar state of lipid molecules is disrupted. Fusion of lipid vesicles in presence of alamethicin has also been observed. The lytic action of alamethicin on erythrocytes and leukocytes can thus be explained on the basis of the strong lipophilic nature of the peptide. However, the aspect of alamethicin-lipid interaction which has attracted considerable attention is the study of electrical properties of lipid bilayers in presence of alamethicin.

Studies on alamethicin modified electrical conductance across lipid bilayers, by a number of workers support a mechanism for ion transport that involves the formation of channels or pores, as in the case of gramicidin A. Measurements of the dependence of conductance on concentration, suggest that 5-8 molecules are involved in an aggregate and that the conductance is cationic. Further ion selectivity has not been observed. In spite of numerous reports, a clear picture regarding the molecular structure of the pores and the manner in which these pores modulate the passage of ions through the bilayer, giving rise to a large voltage dependent conductance, has not yet emerged. Attempts to derive a molecular mechanism for the effect of alamethicin on membranes have suffered from the early reports assigning an incorrect cyclic structure to the ionophore, which were described in the preceding section. For example, Bauman and Mueller[4] and Boheim[7] have suggested that the alamethicin molecules is adsorbed to the membrane interface as a monomer and is in equilibrium with regard to cation complexation. Two methods by which the pores could then form are: (1) The monomer may be pulled into the membrane under the action of a potential difference and then form oligomers, by random association. Fluctuations in pore size may occur by uptake or release of single monomers by the pores.

(2) Adjacent monomers at the interface may aggregate and then be pulled into the membrane. Gordon and Haydon[30-32] have proposed an aggregate of parallel pores which all show similar properties. The conductance properties are thought to result from the central hole which is created by the aggregate of surrounding pores and from interactions between them. Hall[33] has suggested that the conducting pore, with different states, is created by conformational changes of the oligomer. In a recent report, Boheim and Kolb[8] have suggested that pore formation starts from a pre-aggregate of 6 alamethicin molecules at the membrane interface. During pore nucleation, two or three molecules are inserted into the membrane, each transferring an equivalent of gating charge across the membrane and the hexameric pre-aggregate disappears after pore nucleation. However, Fringeli and Fringeli[23] argue that the alamethicin aggregates are located in the membrane itself and not at the interface. These conclusions are based on a study of pore formation in lipid membranes by Infrared Attenuated Total Reflection Spectroscopy[24]. While detailed studies on the interaction of the other Aib containing microbial peptides with membranes have not appeared, preliminary reports indicate[9,10,76] that they possess membrane-modifying properties similar to alamethicin.

In a recent report from the authors' laboratory[67], the use of fluorescent labelled amino terminal fragments of emerimicin in studying peptide-lipid interaction has been explored. While fluorescent probes have been extensively used in studies of membrane and lipid structure and dynamics[12,84], relatively by few investigations have been reported on peptide-lipid interactions. In the case of membrane active peptides, the use of fluorescence has been restricted to valinomycin[79] and linear gramicidins[102,104]. This may, in part, be due to the fact that introduction of fluorescent labels in many of these peptides would be difficult due to lack of reactive groups. The study of lipid-peptide interaction using fluorescent labelled fragments of emerimicin demonstrates the potential use of such an approach, in studies of membrane active peptides related to alamethicin. Recently reports have appeared wherein, the ability of alamethicin to form channels and act as a detergent has been used to study sidedness of natural membrane systems like lymphocyte plasma membranes[11] and membranes from cardiac muscle cells[40].

Secondary structure

The interest in the development of structure-function correlations in

the Aib antibiotics provides a strong impetus for the study of the stereochemistry of Aib containing peptides. The steric hindrance introduced by α-alkylation of α-amino acids was first noted during synthetic investigations on Aib containing peptides[58]. The stereochemically allowed conformations for an Aib residue in the model compounds Ac-Aib-$NHCH_3$ have been analysed theoretically by Marshall and Bosshard[64], Burgess and Leach[15] and Pletnev *et al.*[78] Their calculations suggest that there are primarily only two small regions in the ϕ, ψ plane free from steric overlap for the Aib residue. These are the left-handed and right-handed α-helical conformations with ϕ, $\psi = 50°$, $50°$ and $-50°$, $-50°$ respectively (Fig. 5a). The stereochemically allowed regions for Gly, L-Ala and D-Ala[66] residues are shown in Fig. 5b-d. It can be seen that introduction of methyl groups reduces the "steric overlap free" space in L-Ala, D-Ala and Aib residues. Comparing the ϕ, ψ diagrams of L-Ala and D-Ala residues (Fig. 5c, d) the only areas in common are the left-handed and right-handed α-helical regions. Thus, these are expected to be the only sterically allowed regions for an Aib residue, having two methyl groups at C^{α}.

Until 1976 there were relatively few experimental studies on the conformation of Aib containing antibiotics related to alamethicin and these were largely restricted to CD studies[42,60]. The CD spectrum of alamethicin is typical of an α-helical structure. Three bands at 221 nm (−), 207 nm (−) and 190 nm (+) are observed in a wide variety of solvents. Suzukacillin[43] and trichotoxin[14] also show similar CD behaviour indicating a rigid helical structure. Studies on the conformation of alamethicin by theoretical methods have been reported[16,39]. However, the models proposed for the structure of alamethicin on the basis of these studies may not be valid as the incorrect cyclic structure[77] was used in the calculations.

The alamethicin sequence shown in Fig. 4 consists of large fragments of alternating L-amino acids and Aib residues. The points where there is a departure from strict alteration in the case of alamethicin I is at residues 6 and 7 (Ala-Gln), residues 11 and 12 (Gly-Leu) and at the C-terminal end. However, L-Pro restricts the ϕ value to $-60° \pm 15°$. Aib residues are capable of adopting conformations with positive ϕ, ψ values allowed for D-residues. It is therefore interesting to examine whether the π(LD) helical structure proposed for gramicidin A[100] is possible for Aib containing sequences. As mentioned in the discussion on the conformation of gramicidin A, $\phi = 105°$ and $\psi = -120°$ are necessary for the D-residues in the π^6 (LD) helix. From

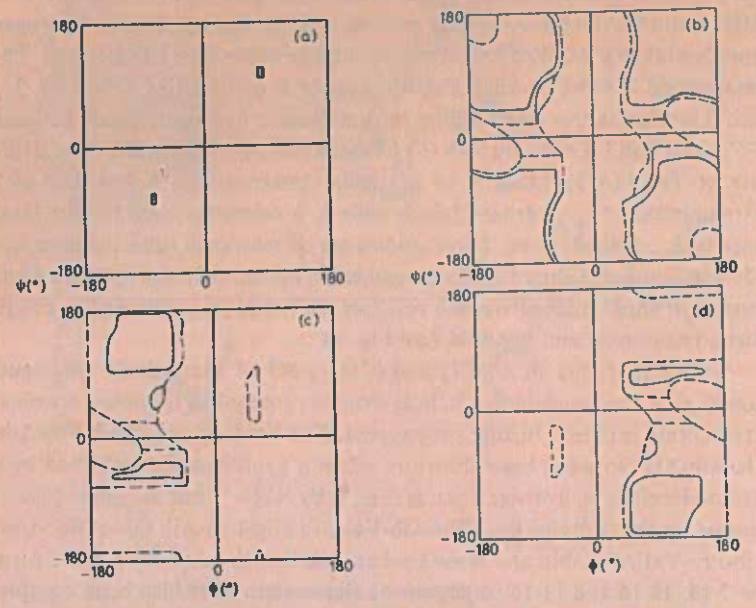


Fig. 5. ϕ, ψ plot for (a) Aib, (b) Gly, (c) L-Ala, (d) D-Ala residues. Fully allowed conformations are outlined by a solid line and partially allowed ones by a dotted line

the earlier description of the Aib steric maps it is seen that these values are stereochemically disallowed for Aib. Alamethicin and related Aib peptides are therefore likely to adopt structures entirely different from those proposed for gramicidin A.

Conformations of model Aib peptides and alamethicin fragments

Studies in this laboratory using NMR[80,98], IR[88,89] and X-ray methods[80,81,96,97] have established the tendency of Aib residues to adopt conformations favouring the formation of β-turns. The stereochemical constraints introduced by the gem dimethyl groups of the Aib residue allow well-defined conformations in solution, even for small acyclic peptides. Further, excellent agreement is obtained between spectroscopic studies in solution and X-ray diffraction in the solid state. The solid state conformations of some

representative Aib containing peptides are illustrated in Fig. 6. Z-Aib-Pro-NHMe and Z-Aib-Aib-Ala-OMe possess a single intramolecular hydrogen bond, stabilising Aib-Pro and Aib-Aib β-turns of the type III category. The tetrapeptide Z-Aib-Pro-Aib-Ala-OMe adopts a consecutive β-turn or 3_{10} helical conformation stabilised by intramolecular hydrogen bonds between Z-CO and Aib (3) NH and Aib (1) CO and Ala (4) NH groups. The structure of Tosyl-$(Aib)_5$-OMe is of particular interest as it is stabilised by 3 intramolecular C_{10} hydrogen bonds with ϕ, ψ values for each residue lying close to 3_{10} helical values. These studies are of relevance since the tetrapeptide constitutes the amino terminal segment of alamethicin and hypelcin while clusters of three consecutive Aib residues are found in antiamoebin, emerimicin, trichotoxin and hypelcin (*see* Fig. 4).

^1H NMR studies in $CDCl_3$ and $(CD_3)_2SO$ of alamethicin fragments suggest that the incipient 3_{10} helical structure formed in the amino terminal tetrapeptide is indeed further propagated. The hexapeptide Z-Aib-Pro-Aib-Ala-Aib-Ala-OMe has been shown to adopt a conformation stabilised by 4 intramolecular C_{10} hydrogen bonds (Fig. 7) by NMR[66] and IR techniques[88]. Studies on the peptides Boc-Gln-Aib-Val-Aib-Gly-Leu-Aib-OMe, Boc-Leu-Aib-Pro-Val-Aib-OMe and Boc-Gly-Leu-Aib-Pro-Val-Aib-OMe which form the 7-13, 12-16 and 11-16 fragments of alamethicin have also been completed[66,88,89]. The 7-13 heptapeptide shows the presence of 5 solvent shielded peptide NH groups in $(CD_3)_2SO$ suggesting that the molecule indeed adopts a highly folded helical conformation. The presence of Pro at the central position in the 12-16 pentapeptide and 11-16 hexapeptide leads to the possibility of both $4 \rightarrow 1$ and $5 \rightarrow 1$ intramolecular hydrogen bonds corresponding to the 3_{10} and α-helical structures, respectively. Using an analysis of the CO stretching vibrations in the 1600-1700 cm^{-1} region it has been suggested that in these sequences there does exist evidence for C_{13} $(5 \rightarrow 1)$ hydrogen bonds involving the Boc CO and Val NH in the pentapeptide and Gly CO and Val NH in the hexapeptide. These results would suggest that the 3_{10} helical fold in alamethicin may expand to a single turn of an α-helix at residues 12-15.

Evidence for folded structures in alamethicin fragments also comes from ^{13}C NMR and CD studies. Pro C^{α} and Aib C^{α} ^{13}C chemical shifts have been found to be dependent on secondary structure. These resonances shift about 4 ppm downfield on going from Z-Aib-Pro-OMe to alamethicin suggesting that the chemical shifts are sensitive to helical folding[66]. Early

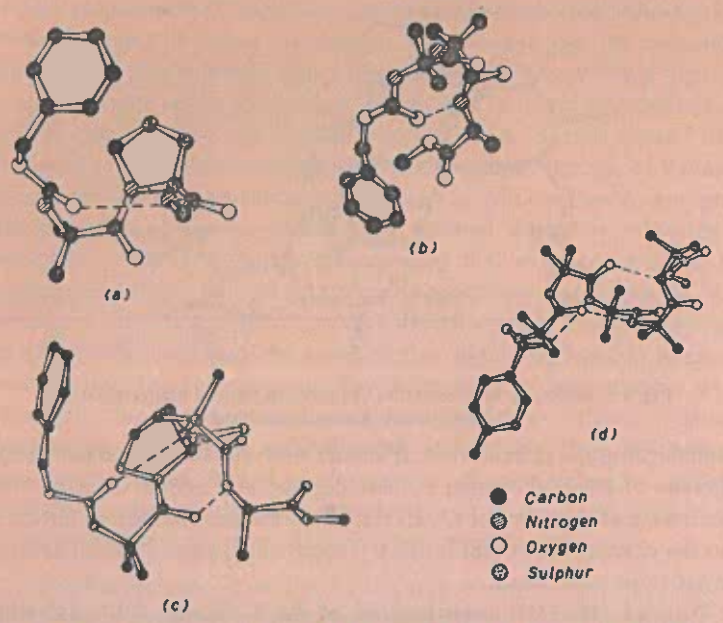


Fig. 6. Crystal structures of (a) Z-Aib-Pro-NHMe Ref. 80 ; (b) Z-Aib-Aib-Ala-OMe Ref. 81 ; (c) Z-Aib-Pro-Aib-Ala-OMe Ref. 68 ; (d) Tos-$(Aib)_5$-OMe Ref. 97

CD studies on alamethicin suggested 40% α-helical content with the spectrum being dependent on solvent[42]. The CD spectra of amino terminal and central synthetic fragments (1–13, 1–17, 1–20, 7–13, 11–17, 11–16) show two negative bands at ~207 nm and 220 nm in trifluorethanol, suggestive of helical folding in these peptides. The tendency of Aib residues to facilitate the formation of helical structures has also been observed in polyoxyethylene (POE) bound Aib-Ala oligomers by ^{13}C NMR[51] and CD[88]. The amino terminal undecapeptide fragment of alamethicin bound to POE also exhibits a helical conformation[88]. A very recent study of the hexadecapeptide Boc-$(Aib\text{-}Ala)_5$-Pro-Ala-Aib-Aib-Glu(γ-OBz)-Gln-OME and the nonadecapeptide analog of alamethicin, possessing membrane modifying properties, Boc-$(Aib\text{-}Ala)_5$-Gly-Ala-Aib-Pro-Ala-Aib-Aib-Glu(γ-OBz)-Gln (γ-OMe) has established an α-helix content of 32 and 45% respectively in ethanol[72]. This may be compared with the value of 42% reported for

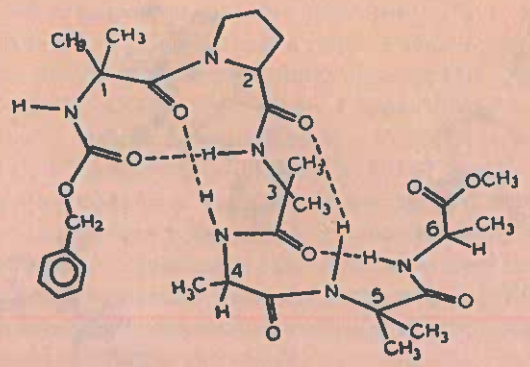


Fig. 7. Schematic representation of hydrogen bonded conformation of Z-Aib-Pro-Aib-Ala-Aib-Ala-OMe

alamethicin in the same solvent. It should, however, be stressed that detailed estimates of α-helical content in these peptides are subject to considerable uncertainty as the effect of C^{α} alkylation on the peptide Cotton effects and also the possible differences in the CD spectra of 3_{10} and α-helical structures are yet to be established.

Detailed ^1H NMR investigations of the 1–13 and 1–17 alamethicin fragments, synthesised in this laboratory, have so far not been undertaken due to extensive overlap of NH resonances at 270 MHz[66]. However, evidence for folded structures in these peptides has been obtained from CD, ^{13}C NMR and IR studies. In particular, IR studies, of the 1–13 fragment in chloroform solution, suggest that a large fraction of the NH groups are involved in intramolecular hydrogen bonding. Evidence for this comes from the very large intensity of the ν_{NH} (hydrogen bonded) at 3330 cm^{-1} relative to the ν_{NH} (free) at 3440 cm^{-1}. While quantitative studies have been reported for smaller fragments, such measurements have yet to be carried out on large fragments and alamethicin itself[66].

The accumulated body of evidence on model Aib containing peptides a nd alamethicin fragments suggests that the polypeptide is likely to have an initial 3_{10} helical fold with an expansion to a single turn of an α-helix at residues 12–16. This conformation is stabilised by a system of $4 \rightarrow 1$ intramolecular hydrogen bonds, with the possibility of $5 \rightarrow 1$ hydrogen bonds near Pro-14. Such a folding pattern is favoured by the presence of eight Aib residues and also the strategically located Pro residues at positions 2

and 14. Studies on the C-terminal tripeptide fragment Boc-Glu-Gln-Phol (unpublished results) favour a random conformation for the molecule in dimethylsulfoxide solution. This tripeptide is likely to provide a negatively charged, flexible tail in alamethicin, which may facilitate orientation in the lipid bilayer. The Gly-Leu segment may serve as a "flexible hinge" in the molecule, a feature that would permit conformational changes as a function of environment. The constraints introduced by Aib residues do not permit the formation of structures with large internal diameters, allowing the passage of ions. It is therefore noteworthy that evidence for alamethicin aggregates as the functional membrane channel have been reported in the literature[8]. Further structure-function correlations must therefore necessarily require detailed studies of the mode of alamethicin aggregation in aqueous solution and in lipid dispersions. The presence of a large number of Aib residues in other membrane active polypeptides like suzukacillin, trichotoxin, emerimicins, antiamoebin and hypelcins, suggests that these molecules may also favour highly helical conformations. A comprehensive molecular mechanism rationalising the membrane activity of these polypeptides must await detailed investigations of their conformation.

Acknowledgement—Research in the area of peptide synthesis and conformation in this laboratory has been supported by grants from the Department of Science and Technology. R.N. thanks the CSIR for award of a fellowship.

REFERENCES

1. Akashi, K., Kuboto, K. and Kurahashi, K. (1977) *J. Biochem. Tokyo*, **81** : 269.
2. Bamberg, E. and Lauger, P. (1973) *J. Membrane Biol.*, **11** : 177.
3. Bauer, K., Roskoski, R. Jr., Kleinkauf, H. and Lipmann, F. (1972) *Biochemistry*, **11** : 3266.
4. Baumann, G. and Mueller, P. (1974) *J. Supramolec. Struct.*, **2** : 538.
5. Benedetti, E., Blasio, B. D., Pedone, C., Lorenzi, G. P., Tomasic, L. and Gramlich, V. (1979) *Nature*, **282** : 630.
6. Besch, Jr., H. C., Jones, L. R., Fleming, J. W. and Watanabe, A. M. (1977) *J. Biol. Chem.*, **252** : 7905.
7. Boheim, G. (1974) *J. Membrane Biol.*, **19** : 277.
8. Boheim, G. and Kolb, H. A. (1978) *J. Membrane Biol.*, **38** : 99.
9. Boheim, G., Janko, K., Leibfritz, D., Ooka, T., Konig, W. A. and Jung, G. (1976) *Biochim. Biophys. Acta*, **433** : 182.
10. Boheim, G., Irmscher, G. and Jung, G. (1978) *Biochim. Biophys. Acta*, **507** : 485.

11. Bonnafous, J. C., Dornand, J. and Mani, J. C. (1979) *Biochem. Biophys. Res. Commun.*, **86** : 536.
12. Brand, L. and Gohlke, J. R. (1972) *Ann. Rev. Biochem.*, **41** : 843.
13. Brewer, D., Hanson, A. W., Shaw, I. M., Taylor, A. and Jones, G. A. (1979) *Experientia*, **35** : 294.
14. Bruckner, H., Konig, W. A., Greiner, M. and Jung, G. (1979) *Angew. Chem. Internat. Edn.* **18** : 476.
15. Burgess, A. W. and Leach, S. J. (1973a) *Biopolymers*, **12** : 2599.
16. Burgess, A. W. and Leach, S. J. (1973b) *Biopolymers*, **12** : 2691.
17. Chappel, J. B. and Crojts, A. R. (1965) *Biochem. J.*, **95** : 393.
18. Chelack, W. S. and Petkau, A. (1973) *J. Lipid Res.*, **14** : 255.
19. Cherry, R. J., Chapman, D. and Graham, D. E. (1972) *J. Membrane Biol.*, **7** : 325.
20. Dubischar, N., Diehn, B. and Jung, G. (1973) Abstr. 24 Int. Congr. Pure Appl. Chem., Hamburg : 178.
21. Eisenberg, M., Hall, J. E. and Mead, C. A. (1973) *J. Membrane Biol.* **14** : 143.
22. Fleming, A. (1929) *Brit. J. Exp. Path.* **10** : 226.
23. Fringeli, U. P. and Fringeli, M. (1979) *Proc. Natl. Acad. Sci. USA*, **76** : 3852.
24. Fringeli, U. P. (1977) *Z. Naturforsch*, **32C**, 20.
25. Fujita, T., Takaishi, Y. and Shinomoto, T. (1979) *J. Chem. Soc. Chem. Commun.* 413.
26. Gisin, B. F., Kobayashi, S., Davis, D. G. and Hall, J. E. (1977) *In* "Peptides", Eds. M. Goodman and J. Meienhofer, Wiley, New York : 215.
27. Gisin, B. F., Kobayashi, S. and Hall, J. E. (1977) *Proc. Natl. Acad. Sci. USA*, **74** : 115.
28. Glickson, J. D., Meyers, D. F., Settine, J. M. and Urry, D. W. (1972) *Biochemistry*, **11** : 477.
29. Gordon, L. G. M. (1974) *In* "Drugs and Transport", Ed., B. A. Callingham, University Park Press, London : 251.
30. Gordon, L. G. M. and Haydon, D. A. (1972) *Biochim. Biophys. Acta*, **255** : 1014.
31. Gordon, L. G. M. and Haydon, D. A. (1975) *Phil. Trans. R. Soc. London*, **B270** : 433.
32. Gordon, L. G. M. and Haydon, D. A. (1976) *Biochim. Biophys. Acta*, **436** : 541.
33. Hall, J. E. (1975) *Biophys. J.*, **15** : 934.
34. Harris, E. J. and Pressman, B. C. (1967) *Nature*, **216** : 918.
35. Hauser, H., Finer, E. G. and Chapman, D. (1970) *J. Mol. Biol.*, **53** : 419.
36. Heitz, F., Spach, G., Cary, P. D. and Robinson, C. C. (1979) *Eur. J. Biochem.* **93** : 135.
37. Hladky, S. B. and Haydon, D. A. (1970) *Nature*, **225** : 451.
38. Hotchkiss, R. D. and Dubos, R. J. (1941) *J. Biol. Chem.*, **141** : 155.
39. Johnson, M. E. and Wu, T. T. (1976) *J. Theor. Biol.*, **60** : 183.
40. Jones, L. R., Besch, Jr., H. R. and Watanabe, A. M. (1977) *J. Biol. Chem.*, **252** : 3315.
41. Jung, G., Leibfritz, D., Ottnad, M., Dubischar, N. and Probst, H. (1974) *Hoppe Seyler's Z. Physiol. Chem.*, **355** : 1213.
42. Jung, G., Dubischar, N. and Leibfritz, D. (1975) *Eur. J. Biochem.*, **54** : 395.
43. Jung, G., Konig, W. A., Leibfritz, D., Ooka, T., Janko, K. and Boheim, G. (1976) *Biochim. Biophys. Acta*, **433** : 164.

44. Kenner, G. W. and Sheppard, R. C. (1958) *Nature*, **181** : 48.
45. Kolb, H. A. and Boheim, G. (1977) *J. Membrane Biol.*, **38** : 541.
46. Kolb, H. A., Lauger, P. and Bamberg, E. (1975) *J. Membrane Bio.*, **20** : 133.
47. Koeppe, R. E., Hodgson, K. D. and Stryer, L. (1978) *J. Mol. Biol.*, **121** : 41.
48. Kurahoshi, K. (1974) *Ann. Rev. Biochem.*, **43** : 445.
49. Lau, A. L. Y. and Chan, S. I. (1974) *Biochemistry*, **13** : 4942.
50. Lau, A. L. Y. and Chan, S. I. (1975) *Proc. Natl. Acad. Sci. USA* **72** : 2170.
51. Leibfritz, D., Mayr, W., Oekonomopubos, R. and Jung, G. (1978) *Tetrahedron*, **34** : 2045.
52. Leplawy, M. T., Jones, Kenner, G. W. and Sheppard, R. C. (1960) *Tetrahedron*, **11** : 39.
53. Marshall, G. R. and Balasubramanian, T. M. (1979) *In* "Sixth American Peptide Symposium, Abstract No. A-111.
54. Marshall, G. R. and Bosshard, H. E. (1972) *Circ. Res. Suppl.* II, **30** : 143.
55. Marshall, G. R., Bosshard, H. E., Kendrick, N. U. E., Turk, J., Balasubramanian, T. M., Cobb, S. M. H., Moore, M., Leduc, L. and Needleman, P. (1976) *In* "Peptides", Ed. Loffet : 361.
56. Martin, D. R. and Williams, R. J. P. (1976) *Biochem. J.* **153** : 181.
57. Mauro, A., Nanavati, R. P. and Heyer, E. (1972) *Proc. Natl. Acad. Sci. USA*, **69** : 3742.
58. Mayr, W., Oekonomopulos, R. and Jung, G. (1979) *Biopolymers*, **18** : 425.
59. McMullen, A. I. (1970) *Biochem., J.* **119** : 10P.
60. McMullen, A.I., Marlborough, O.I. and Bayley, P. M. (1971) *FEBS Lett.*, **16** : 278.
61. Meyer, G. E. and Reusser, E. (1967) *Experientia*, **23** : 85.
62. Mohr, H. and Kleinkauf, H. (1978) *Biochim. Biophys. Acta* **526** : 375.
63. Mueller, P. (1976) *In* "Horizons *In* Biochemistry and Biophysics", Eds. E. Quagliariello, F. Palmieri and T. P. Singer, Vol. 2 : 230.
64. Mueller, P. and Rudin, D. O. (1967) *Biochem. Biophys. Res. Commun.* **26** : 398.
65. Mueller, P. and Rudin, O. (1968) *Nature* **217** : 713.
66. Nagaraj, R. (1980) Ph.D. Thesis, Indian Institute of Science, Bangalore, India.
67. Nagaraj, R. and Balaram, P. (1979) *Biochem. Biophys. Res. Commun.*, **89** : 1041.
68. Nagaraj, R., Shamala, N. and Balaram, P. (1979) *J. Am. Chem. Soc.*, **101** : 16.
69. Noda, K. and Gross, E. (1972) *In* "Chemistry and Biology of Peptides" Proc. Third Amer. Peptide Symposium, Ed. J. Meienhofer, Ann. Arbor. Michigan : 241.
70. Ooka, T. and Takeda, I. (1972) *Agric. Biol. Chem.*, **30** : 700.
71. Oekonomopulos, R. and Jung, G. (1979) *Liebigs. Ann. Chem.*, **8** : 1151.
72. Oekonomopulos, R. and Jung, G. (1980) *Biopolymers*, **19** : 203.
73. Ovchinnikov, Y. A., Ivanov, V. T. and Schkrob, A. M. (1974) *In* "Membrane-Active Complexones". Elsevier, Amsterdam, Oxford and New York.
74. Pandey, R. C., Carter Cook, Jr., J. and Rinehart, Jr., K. L. (1977) *J. Am. Chem. Soc.* **99** : 5203.
75. Pandey, R. C., Carter Cook, Jr., J. and Rinehart, Jr., K. L. (1977) *Ibid.*, **99** : 5205.
76. Pandey, R. C., Carter Cook, Jr., J. and Rinehart, Jr., K. L. (1977) *Ibid.* **99** : 8469.
77. Payne, J. W., Jakes, R. and Hartley, B. S. (1970) *Biochem. J.* **117** : 757.

78. Pletnev, V. Z., Gromov, E. P. and Popov, E. M. (1973) *Khim. Prir. Soedin.*, 9 : 224.
79. Pohl, G. W., Knoll, W., Gisin, B. F. and Stark, G. (1976) *Biophys. Struct. Mech.*, 2 : 119.
80. Prasad, B. V. V., Shamala, N., Nagaraj, R., Chandrasekaran, R. and Balaram, P. (1979) *Biopolymers*, 18 : 1635.
81. Prasad, B. V. V., Shamala, N., Nagaraj, R. and Balaram, P. (1980) *Acta Cryst.* 36B : 107.
82. Pressman, B. C. (1967) *Fed. Proc.*, 27 : 1283.
83. Pressman, B. C. (1976) *Ann. Rev. Biochem.*, 45 : 101.
84. Radda, G. K. (1975) *In* "Methods in Membrane Biology", Vol. 4, Ed. E. D. Korn, Plenum, New York and London : 97.
85. Ramachandran, G. N. and Chandrasekaran, R. (1972) *Ind. J. Biochem. Biophys.*, 9 : 1.
86. Ramachandran, G. N. and Sasisekharan, V. (1967) *Adv. Protein Chem.*, 23 : 284.
87. Ranganathan, D. and Ranganathan, S. (1976) *In* "Art in Biosynthesis. The synthetic chemists challenge", Vol. 1, Academic Press, New York, San Francisco, London.
88. Rao, Ch. P., Nagaraj, R., Rao, C. N. R. and Balaram, P. (1979) *FEBS Lett.*, 100 : 244.
89. Rao, Ch. P., Nagaraj, R., Rao, C. N. R. and Balaram, P. (1980) *Biochemistry*, 19 : 423.
90. Reusser, F. (1967) *J. Biol. Chem.*, 242 : 243.
91. Rindfleish, H. and Kleinkauf, H. (1976) *FEBS Lett.*, 62 : 276.
92. Rinehart, K. L., Jr., Cook, J. C., Jr., Meng, H., Olson, K. L. and Pandey, R. C. (1977) *Nature*, 269 : 832.
93. Roy, G. (1975) *J. Membrane Biol.*, 24 : 71.
94. Sarges, R. and Witkop, B. (1965) *J. Am. Chem. Soc.*, 87 : 2011.
95. Sarges, R. and Witkop, B. (1965) *J. Am. Chem. Soc.*, 87 : 2020.
96. Shamala, N., Nagaraj, R. and Balaram, P. (1977) *Biochem. Biophys. Res. Commun.*, 79 : 292.
97. Shamala, N., Nagaraj, R. and Balaram, P. (1978) *J. Chem. Soc. Chem. Commun.* : 996.
98. Tosteson, D. C., Andreoli, T. E., Tieffenberg, M. and Cook, P. (1968) *J. Gen. Physiol.*, 51 : 3738.
99. Urban, R., Marquarding, D. and Ugi, I. (1979) Proceedings of the Sixth American Peptide Symposium, Abstract No. A-10.
100. Urry, D. W. (1972) *Proc. Natl. Acad. Sci. USA*, 69 : 1061.
101. Urry, D. W., Goodall, M. C., Glickson, J. D. and Meyers, D. F. (1971) *Proc. Natl. Acad. Sci. USA*, 68 : 1907.
102. Veatch, W. R. and Stryer, L. (1977) *J. Mol. Biol.*, 113 : 89.
103. Veatch, W. R., Fossel, E. T. and Blout, E. R. (1974) *Biochemistry*, 13 : 5219.
104. Veatch, W. R., Mathies, R., Eisanberg, M. and Stryer, L. (1975) *J. Mol. Biol.*, 99 : 75.
105. Zingsheim, H. P. and Neher, E. (1974) *Biophys. Chem.*, 2 : 197.

Wild Life Resources of India

Madhav Gadgil

Centre for Theoretical Studies
Indian Institute of Science
Bangalore-560 012

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

यावत् भूमंडलम् धत्त समृगवनकाननम् ।
तावत् तिष्ठति मेदिन्याम् संतति: पुत्रपौत्रिकी ।।

So long as the earth bears woods and forests full of wild animals
will it continue to support the progeny of man.

Introduction

Ecologically and biogeographically the Indian subcontinent is one of the most fascinating regions in the world. On this subcontinent, we have a rainless desert and the rainiest place on earth, the hottest place on earth and the permanently snowbound heights of Himalayas. The great variety of ecological conditions thus prevailing allows this landmass to support a great diversity of wild life, including hot desert forms like the wild ass† and cold desert forms like the Tibetan antelope, animals of open scrubland like the blackbuck and of grassy swamps like the rhinoceros, animals of deciduous forest like the gaur and of tropical rain forest like the liontailed macaque.

The unique biogeographic position of the Indian subcontinent adds further interest to its wild life. The region lies at the confluence of Ethiopian, Palaearctic and Indo-Malayan faunas and possesses a number of interesting components from each of these realms as well as several peculiarly indigenous genera. The Ethiopian element is represented by the chinkara, the hyena and the ratel, the Palaearctic by lynx, wolf, markhor, urial and hangul, the Chinese by red panda and musk deer and the Indo-Malayan by the hoolock gibbon, the elephant, the goat-antelopes and the mouse-deer. The genera restricted to India include *Melursus* (sloth bear), *Antilope* (blackbuck), *Tetracerus* (chousinga or four-horned antelope) and *Boselaphus* (nilgai).

At one time our wild life was as remarkable for its abundance as it is even now for its variety, with tigers haunting the forests in hundreds, and antelopes the open plains in thousands upon thousands. Over the centuries it has played an important role in our economy and culture—as a source of protein for the masses, as items of trade, and as an inspiration for a rich religious lore. The wild life has gradually declined over the millenia, and very rapidly since the British took over the country. At last we have reached a critical stage at which most, if not all of our wild life appears on the verge

† See appendices for scientific names.

of being totally wiped out. Only bold and imaginative action in conservation and management will enable us to save this rich heritage from an irreversible loss at this critical juncture.

The Fauna

The Indian subcontinent may be divided into the following eight regions for our purpose† (see Fig. 1) :

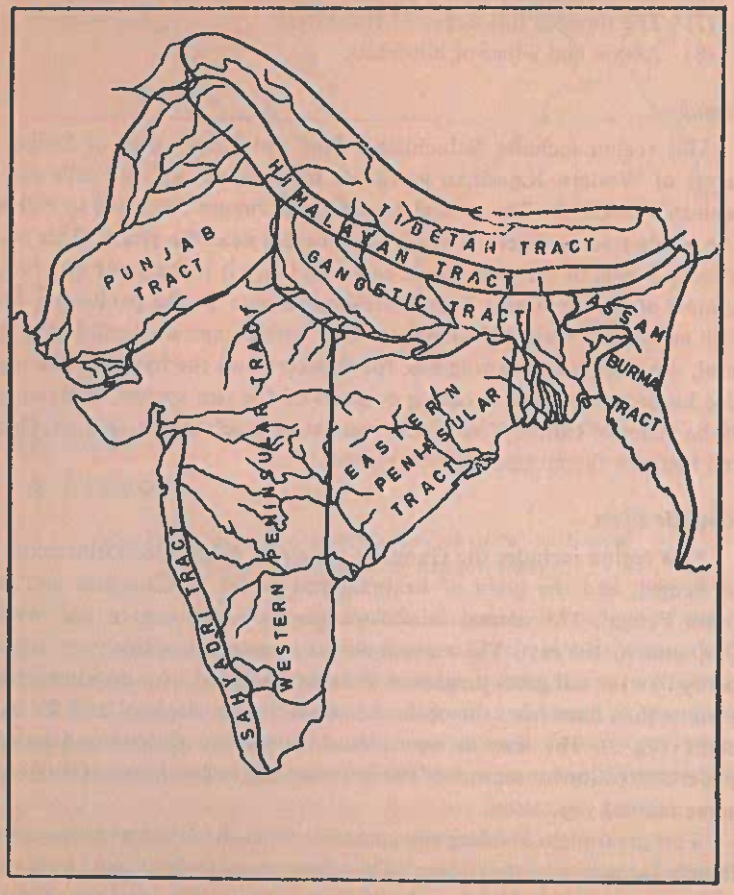


Fig. 1. Biogeographic regions of the Indian subcontinent (After Blanford 1901)

† For reasons of expediency, this review will be mainly restricted to larger mammals.

(1) Punjab and the adjoining semi-arid and desert regions
(2) The alluvial plains of Ganges and lower Brahmaputra
(3) The western peninsular tract excluding the Western Ghats
(4) The eastern peninsular tract
(5) The Western Ghats or Sahyadris
(6) Ladakh and adjacent regions
(7) The forested hill-slopes of Himalayas
(8) Assam and adjacent hill-tracts.

Punjab

This region includes Baluchistan, Sind and Punjab west of Delhi, the desert of Western Rajasthan as far as the base of Aravalli hills and the peninsula of Cutch. The annual precipitation ranges from 100 to 600 mm. The whole tract is desert or semi-desert except near the rivers. This region is really a part of the palaearctic realm. Its fauna is rather poor and lacking in most of the prominent Indo-Malayan elements in the peninsular fauna such as langurs, loris and elephant. The notable species include the ovid urial, the wild goat, the wild ass, the chinkara and the hyaena. The region also harbours the great breeding colonies of the two species of flamingoes in the Rann of Cutch. The Indus river fauna is related to that of Ganges and includes the mugger and the gharial.

Gangetic Plain

This region includes the Gangetic plain and delta from Delhi to the Bay of Bengal, and the plain of Brahmaputra as far as Goalpara and all of lower Bengal. The annual rainfall ranges from 600 mm in the west to 2000 mm in the east. The natural climax vegetation of this tract was probably like the tall grass jungles of Terai interspersed with deciduous forest. It must then have been the optimal habitat for the elephant and the rhinoceros (Fig. 2). The tract is now almost completely cleared and has been under cultivation for centuries. The Brahmaputra valley, however, still retains some natural vegetation.

This great plain dividing the peninsula from the Himalayas has a fauna largely belonging to the former. Thus characteristic peninsular forms such as hyaena, ratel, blackbuck and nilgai occur in suitable localities. The most characteristic animals of the Gangetic plain are the hog deer and the hispid hare or Assam rabbit, although both species extend into Assam. The hispid

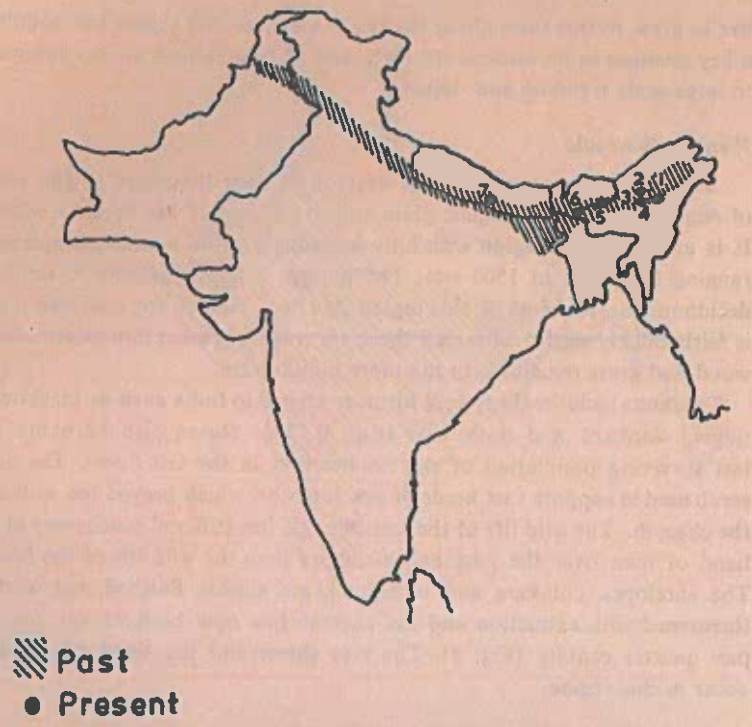


Fig. 2. The past and present distribution of the one-horned rhinoceros in India (After Seshadri 1969)

hare is probably extinct by now, and the hog deer almost gone from the plains. The most extensive natural habitat still left in this region is the Sunderbans, a swampy mangrove forest at the mouth of the river Ganges. Most of the Sunderbans are now in Bangladesh, and those on the Indian side are considerably reduced. They do, however, still possess some interesting wild life including the tiger. Notable freshwater animals of Ganges include the gharial and the Gangetic dolphin ; both are threatened species.

Although most of the wild life has thus been nearly wiped out of this thickly settled region, there remain substantial populations of the two monkeys, the rhesus macaque and the hanuman langur in rural as well as urban areas. These monkeys, unmolested by man because of religious taboos,

live in great troops throughout the tract. Of these, the rhesus has acquired a key position in biomedical research, and its populations are declining due to large-scale trapping and export.

Western Peninsula

This region lies roughly to the west of 80° east longitude to the south of Aravalli hills and Gangetic plain and to the east of the Western Ghats. It is an undulating region with hilly outcrops with an annual precipitation ranging from 400 to 1500 mm. The natural climax vegetation is scrub or deciduous forest. Most of this region has been cleared for cultivation and is fairly thickly settled, although there are tracts of rather thin forest, brush-wood and grass remaining in the more hilly terrain.

The fauna includes the typical forms restricted to India such as blackbuck, nilgai, chinkara and sloth bear (Fig. 3). This region also harbours the last surviving population of the Asiatic lion in the Gir forest. The open scrub used to support vast herds of antelopes on which preyed the wolf and the cheetah. The wild life of the scrub jungle has suffered much more at the hand of man over the past half-a-century than the wild life of the forest. The antelopes, chinkara and blackbuck are almost finished, the wolf is threatened with extinction and the cheetah has now been extinct for the past quarter century (Fig. 4). The tree shrew and the slender loris also occur in this region.

Eastern Peninsula

This region lies between the Gangetic plain to the north and the river Krishna to the south, and extends from the Bay of Bengal to the longitude 80° east. It is a hilly country with an average yearly rainfall of 1250 mm. Being relatively sparsely populated, it still retains much forest cover. The natural climax vegetation is a deciduous forest, dominated over much of the region by sal (*Shorea robusta*).

This tract has many forest species with Indo-Malayan affinities such as the elephant, the gaur and the sambar as well as typical peninsular forms such as the blackbuck, nilgai and sloth bear. A number of tigers still survive in this region, as do small populations of two other threatened species, namely wild buffalo and the hard-ground race of barasinga (Fig. 5). Interestingly enough, the giant squirrel occurs here as on the Western Ghats. A number of bird genera are distributed over the eastern peninsula, the Western Ghats

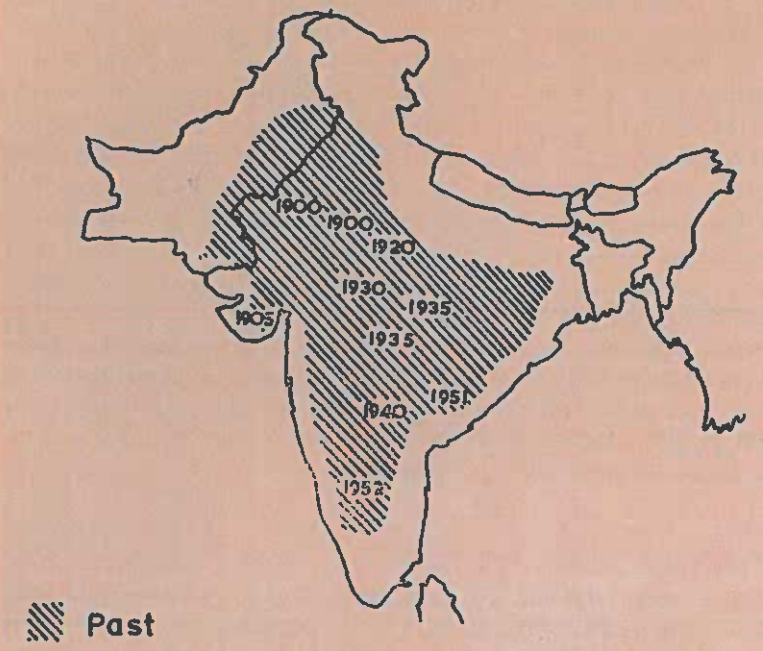


Fig. 4. The past distribution of the cheetah, with dates of the last known record (After Seshadri 1969)

and Himalayas. This suggests that there was once continuous distribution of many forest forms from Himalayas to the Western Ghats, and that they were probably exterminated from intervening areas by the clearance of forest.

Western Ghats

The hill range of Western Ghats runs north-south from river Tapti to the southern tip of the peninsula. Along with the narrow coastal strip of Konkan and Malabar, this is a region of heavy rainfall ranging from 2500 to over 7000 mm a year. Although the coastal strip has been cleared for cultivation, as also the hills over much of Maharashtra and Kerala, other hill tracts still retain substantial forest cover.

This region has a rich fauna, remarkable for the fact that some species are entirely restricted to this tract, and that related groups often occur only on the Western Ghats and the Himalayas. Forms peculiar to the Western Ghats include the Nilgiri langur, the lion-tailed macaque, the Nilgiri tahr and the Malabar civet, all of which are now on the list of endangered species. The Nilgiri langur and the lion-tailed macaque do not enjoy on the ghats the protection enjoyed elsewhere by monkeys as sacred animals. In fact they are believed to possess many medicinal properties and the black monkey soup concocted from them is a favourite beverage. The Nilgiri tahr is notable for the fact that it is one amongst the only two species of the genus *Hemitragus*, the other, the Himalayan tahr occurring in the Himalayas. The Western Ghats harbour extensive populations of forest species with Indo-Malayan affinities such as the gaur, the elephant and the sambar (Fig. 6). The region has a particularly rich reptilian and amphibian fauna, with an entire family, Uropeltidae, and seven genera of reptiles as well as five genera of amphibians occurring nowhere else in the world.

Ladakh

The Gilgit-Ladakh tract is a continuation of the Tibetan plateau lying mostly above the tree-line. The fauna is entirely that of Central Asia, thus belonging to the Palaearctic realm. The notable mammals include wild goats and sheep of genera *Ovis* and *Capra*, the Tibetan wild ass, the Tibetan antelope and the musk deer. The musk deer has been nearly exterminated through persecution for the musk produced by males. The only hope for its conservation lies in establishment of musk deer ranches, where musk may be extracted without killing the animals.

Himalayas

This region extending from Hazra and Kashmir to Mishmi bills includes the forested slopes of the Himalayas from the base to 3000 to 4500 meters in altitude. The rainfall ranges from 750 to 3500 mm gradually increasing to the east. The eastern Himalayas are better clad with forest cover and are more tropical in nature. A rich fauna still persists in the Himalayas, although this region too has suffered heavily from deforestation and destruction of wild life.

The fauna is predominantly Indo-Malayan, with Palaearctic elements such as hangul or Kashmir stag restricted to the west. *Ailurus* (red panda)

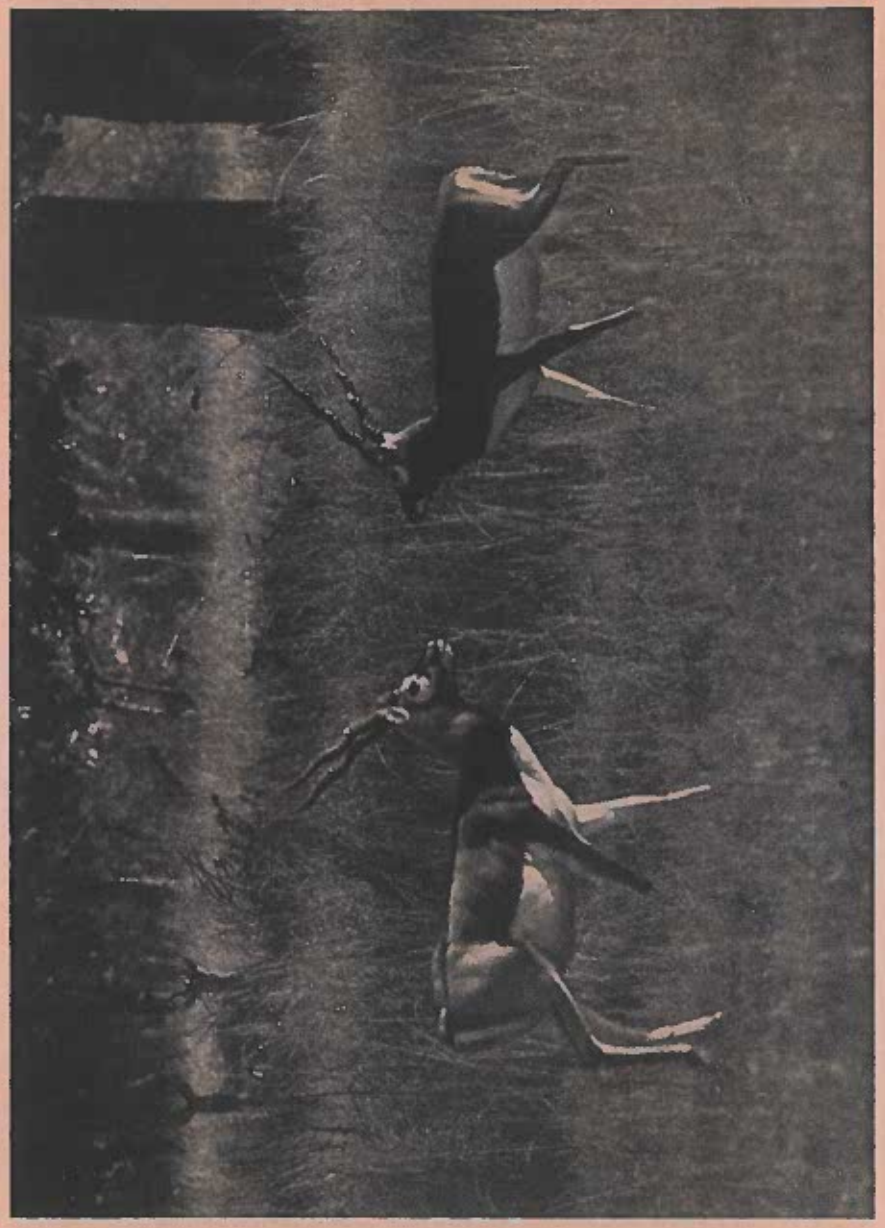


Fig. 3. Blackbuck (Photo : B. N. S. Deo)

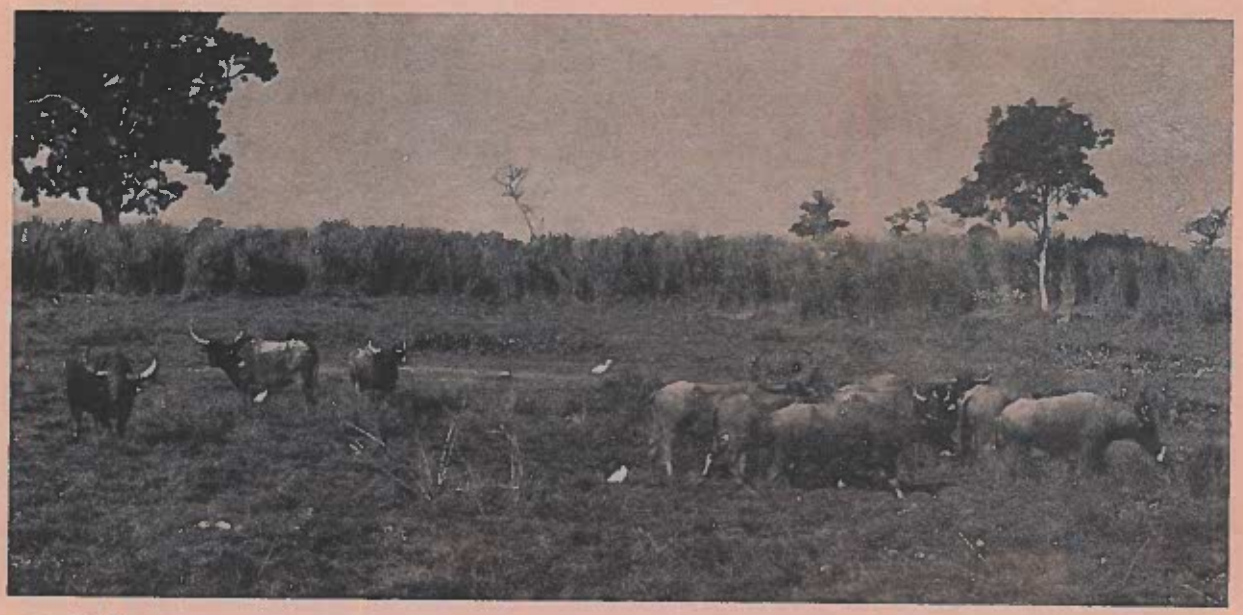


Fig. 5. A herd of wild buffaloes (Photo : B. N. S. Deo)

Fig. 6. A herd of gaur at water (Photo : B. N. S. Deo)

Fig. 7. The one-horned rhinoceros from Kaziranga Sanctuary, Assam
(Photo : B. N. S. Deo)

is the only genus almost totally confined to the Himalayas. Other notable Himalayan forms include the snow leopard and clouded leopard, both threatened with extinction, the spotted linsang, the brown bear, the Himalayan black bear, the giant flying squirrel, the Himalayan tahr and the goat-antelopes-serow, goral and takin.

Assam

Assam, Garo, Khasi and Naga hills, Manipur, Chittagong and Arrakan, the hill ranges to the west of Irrawaddy river and the upper Brahmaputra valley, constitute a single natural region. This is a hilly area of heavy rainfall, the average annual precipitation exceeding 2500 mm. The region is sparsely populated and still well-wooded in many parts. The many tribes inhabiting these regions, however, have no taboos against killing any animal whatsoever, and the once rich wild life is now very much depleted because of tribal hunting.

The wild life of this region, with very strong Indo-Malayan affinities, includes many remarkable species such as hoolock gibbon—which is the only Indian ape, the slow loris, the golden langur, the capped langur, the ferret-badger, the one-horned rhinoceros, the wild buffalo, the swamp deer, the serow, the takin, the thamin, the pygmy hog, the clouded leopard and the marbled cat. A number of these species are endangered and survive only in a few protected places such as Kaziranga and Manas sanctuaries (Fig. 7).

Man and the Beasts

Man the hunter

The story of the interaction of man with the beasts has many facets and goes far back into the past. The human species originated as group-hunters of the enormous herds of large mammals that roamed the African plains, and has been responsible, as hunters, for the extermination of a number of mammalian species. The extermination of several species of large mammals by the first wave of humans invading the Americas about 12,000 years ago is now well-documented. Fossil evidence shows that the Indian fauna included the hippopotamus and baboons about 10,000 years ago, and suggests that the old stone age man was responsible for their extermination from India. Stone age rock paintings of Central India show numerous scenes of hunt (Fig. 8). Mahabharata describes how Arjuna ran into conflict with

Shiva while hunting a wild boar. Kautilya's *Arthashastra* advises princes against excessive indulgence in hunting, and Barni's *Tarikh-i-Firoz Shahi* describes how Sultan Alauddin was attacked by his nephew while absorbed in a hunt halfway between Delhi and Ranthambhor. There is thus abundant evidence that hunting is in human blood and was followed avidly by all classes of Indians throughout the historical period. It was no doubt, at least in part, responsible for the extermination of much of the wild life even during the historical period. Thus rhinoceros occurs on a Mohenjodaro seal (Fig. 9), and was hunted by Emperor Babar in bush country near Indus, but was probably exterminated from Punjab some time soon after that. But the real onslaught and the final extermination of our wild life began only with the introduction of modern fire-arms by the British. The British soldiers, planters and administrators wiped out the royal game of the Indian lion over Northern India during the second half of nineteenth century. The nabab of Junagarh had to pretend that there were only a few lions left in his Gir forest to protect them from every Governor, Viceroy and General who wanted his share of the booty (Fig. 10). The advent of the jeep during the Second World War threw open the remotest corner of forest to even the laziest hunter and the last thirty years have witnessed the final plunge of our wild life towards ultimate extinction.

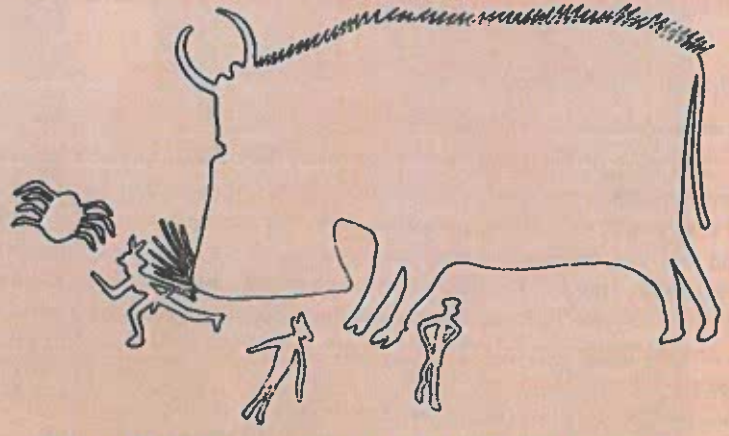


Fig. 8. 'The terrible gaur'. A stone age rock painting from Bhimbetka (after Wakankar and Brooks 1976)

Fig. 9. Rhinoceros from a Mohenjodaro seal (after Gee 1964)

Conflicts with Wild Animals

Apart from hunting for meat, and also no doubt to satisfy a primaeval urge, man has hunted, and continues to hunt wild animals because of a variety of conflicts with them. The most acute conflict is with animals that kill man, and in India tigers, panthers and wolves have occasionally turned man-eaters (Fig. 11). The tigers of Sunderbans are particularly famous for this and have regularly killed humans over the centuries. The larger herbivores such as elephants, gaur and wild pig have also attacked and killed men when they felt threatened. The elephants are particularly apt to retaliate in this fashion when shot at. These so-called rogue elephants have become a serious menace in places where cultivation has recently encroached on former elephant territory as in many places on the Western Ghats.

A number of carnivores prey on man's livestock and poultry. The lions of Gir, for example, depend on the domestic buffalo for much of their sustenance and the buffalo-keepers in turn sometimes accept such a loss as a

Fig. 10. The past (hatched) and present (dot) distribution of the Indian lion with dates of last known record (after Seshadri 1969)

necessary price for the privilege of utilizing the good grazing available in the Gir forest. Much more frequently, however, the offending predator is killed off. As early as the time of writing of Kautilya's *Arthashastra* tigers were killed by setting out poisoned meat when they attacked cattle, and today the vastly more effective poisons such as folidol are employed to the same end.

Just as man's livestock are an easy target for wild predators, his crops offer excellent grazing for wild herbivores. Elephants, gaur, blackbuck and many species of deer, and above all wild pigs come in conflict with cultivators in this fashion. Our own field studies in North Kanara district of Karnataka have shown how herd after herd of elephants has been decimated because they raided paddy fields which were newly established in the midst of elephant forests in the last half century (Fig. 12).

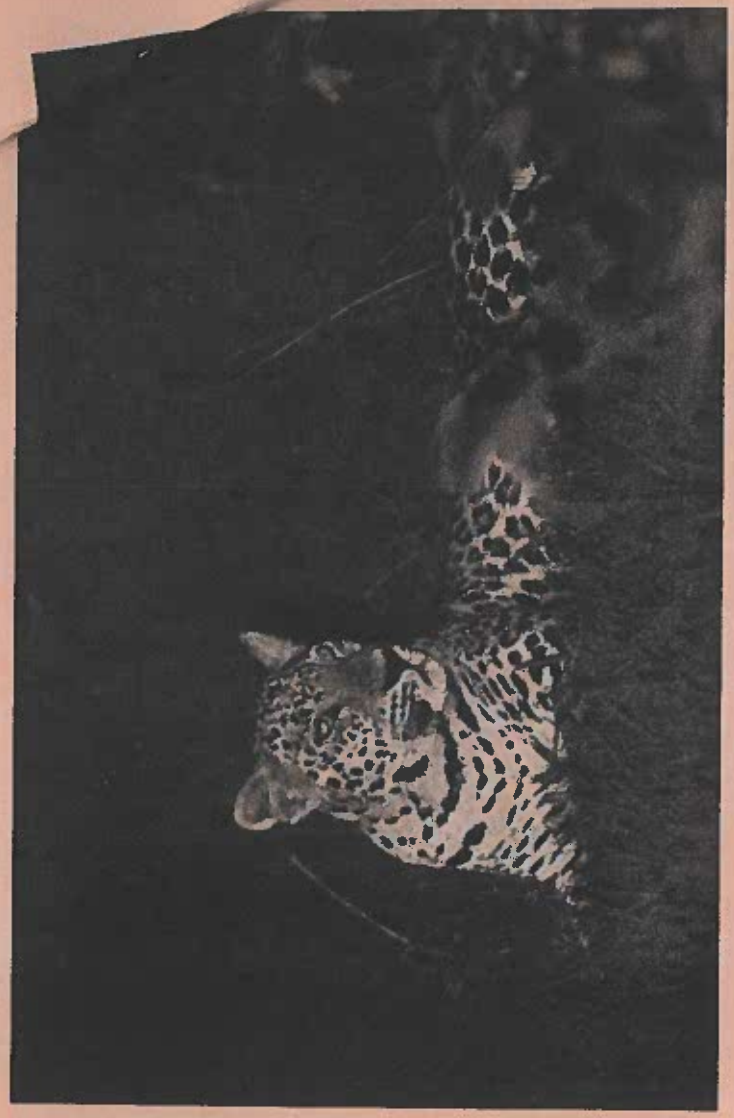


Fig. 11. A panther (Photo : B. N. S. Deo)

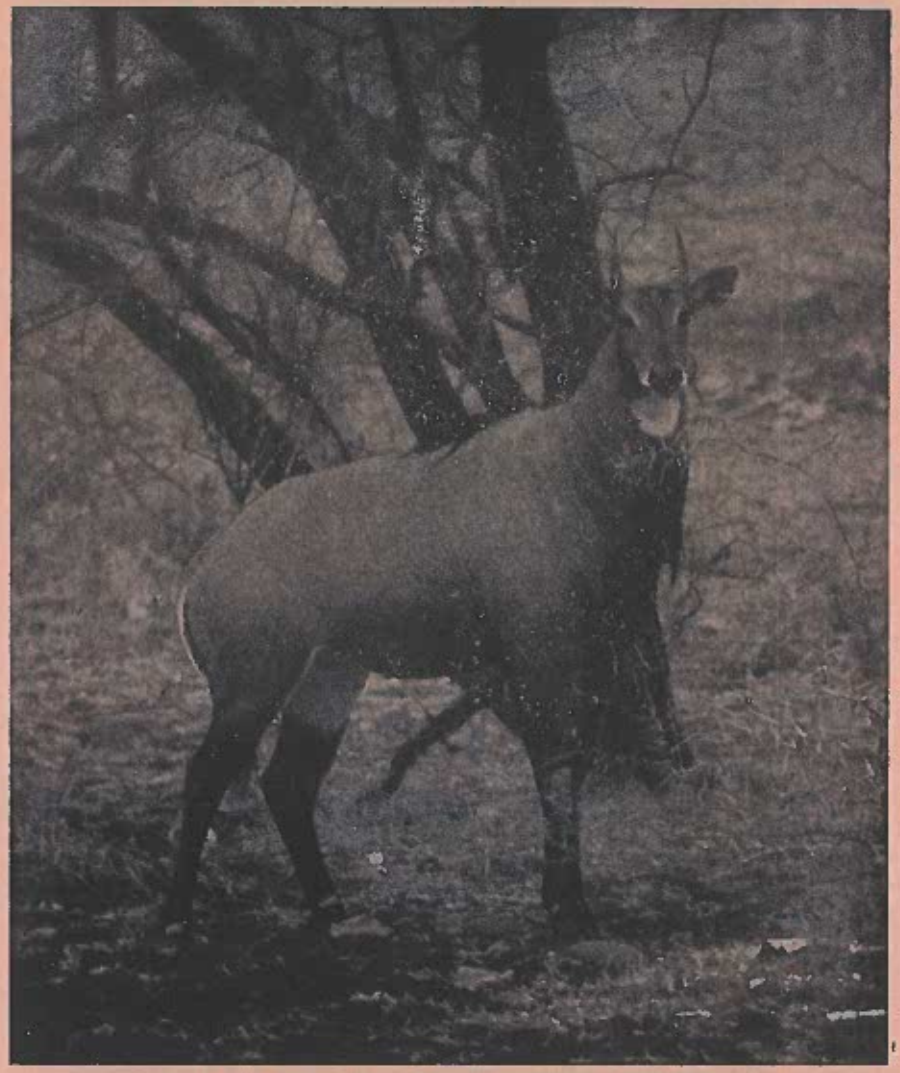


Fig. 13. The Nilgai—one of our most promising species for wild life ranching (Photo : B. N. S. Deo)

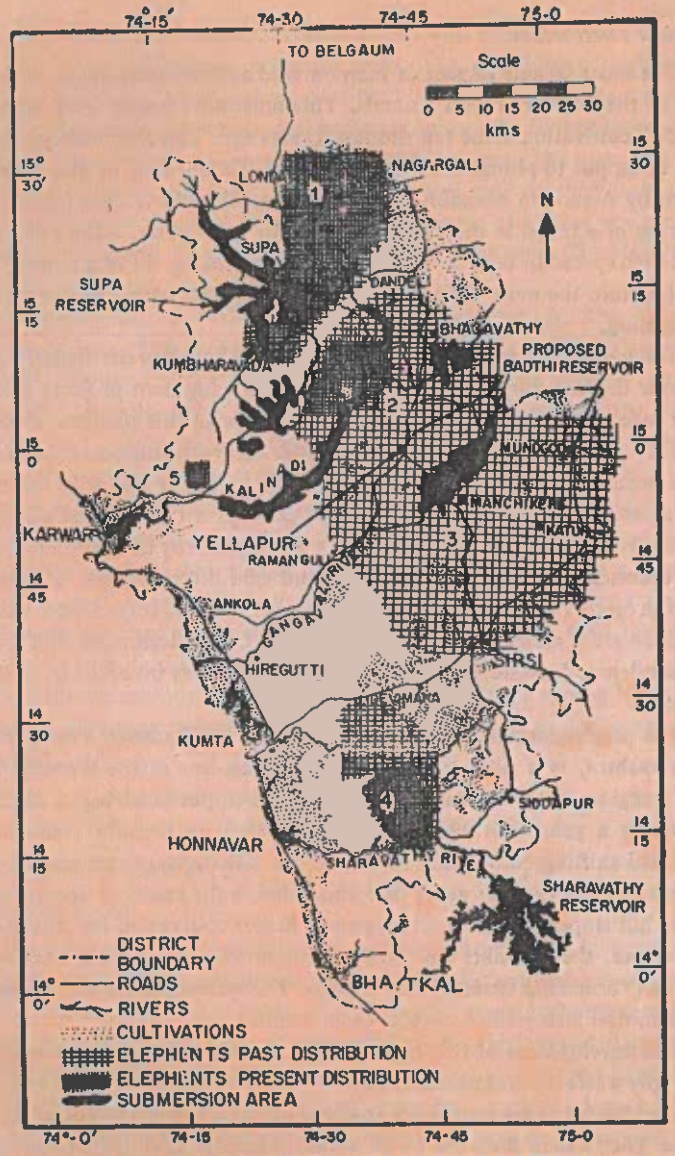


Fig. 12. The past and present distribution of the ELEPHANT in the district of North Kanara (after Nair and Gadgil, 1978)

Habitat Destruction

The most serious impact of man on wild animals arises from his destruction of the habitat of wild animals. This must have begun ever since man took to cultivation some ten thousand years ago. The river valleys were the first to be put to plough and the episode of the burning of the Khandava forest by Arjuna in Mahabharata times is apparently a description of the clearing of a forest in the Gangetic plain for cultivation. Man's demand on land then spread to drier and more undulating tracts, till finally now he has reached into the most difficult terrain for the construction of reservoirs and for mining.

The pace of destruction of the wild life habitats has particularly picked up over the last thirty years, and the many development projects taken up after independene have been a significant factor in this process. These projects for the development of water or other natural resources could of course have been made vehicles of the development of forest and wild life resources as well. Unfortunately, the projects always act to provide an entry to private vested interests ready to reap a quick profit by the rapid exploitation and eventual destruction of the forest and wild life resources. The development projects have invariably served to wipe out our forests and wild life. As these are a major and a continuing cause of the decline of wild life since ndependence, I would like to illustrate the processes involved by a typical example.

The catchment basin of Mula river dammed at Panshet, west of Pune in Maharashtra, is a hilly tract close to the crest-line of the Western Ghats. It is a region of narrow valleys and steep hill-slopes receiving a rainfall of 4000 mm a year. Till 1950, it was populated by Kunabi peasants who practised shifting cultivation on the lower hill-slopes, grew paddy in the valleys, and hunted the game that abounded in the reserved forests on the upper hill slopes. While practising their *Kumri* cultivation for millets on the hill slopes, the peasants left intact the economically valuable trees such as harada (*Terminalia chebula*) and mango. This protected the hill-slopes from erosion, and also yielded regular cash income.

The development of this region began in 1955 when a dam was planned to supply water for the industrial city of Pune and for the sugarcane growers of Pune district. This previously inaccessible region was connected by good roads. The trees in the area which would submerge had to be removed, and timber merchants from Pune moved into reap the harvest. But the timber

so available was small compared to that on the vast hill-slopes in the watershed owned by the peasants. Now, the elementary principles of watershed conservation dictate that maximum possible tree-cover should be retained in the catchment area. However, the engineers joined the timber merchants to persuade the peasants to deforest these vast hill-slopes. The peasants were told that they would be relocated, and might as well cash on their tree crop. Forced thus the peasants sold off this huge amount of timber at abysmal prices ranging from 25 paise to 2 rupees for an entire tree. The dam was built, collapsed in 1962, was rebuilt in 1970.

As of today, very few families have been relocated. All the villages remain occupied as before, but now without their paddy fields which have submerged. Deprived of the tree-cover, the hill-slopes have become rapidly eroded under the onslaught of the heavy rains. This has depressed the yields of the shifting cultivation for millets on the slopes. The peasants have been forced to encroach into the reserved forests on the upper hill-slopes which have been destroyed. With the forest cover gone, the wild life has almost vanished. Slaves of their old habits, the peasants are continually on lookout for game. But where they used to feast once or twice a week on a sambar or a barking deer, they can now get only an occasional hare. There are landslides now on the bare slopes which once used to teem with wild life, and the reservoir is no doubt silting up rapidly.

This is not an exceptional case. The Western Ghats are dotted with such examples all along their length, and no doubt other parts of India too are no different.

Wild Life as a Resource

As a Protein Source

The peasants of the Panshet valley, as narrated earlier, used to get a deer or two every week till modern development overtook them. In fact our field work on the Western Ghats shows that both the tribals and peasants used to depend heavily on wild meat as a source of protein throughout this region. Barking deer, chital, sambar, mouse deer, porcupine, blacknaped hare, wild pig, gaur, field mice, peafowl, jungle fowl, spurfowl, quails and partridges, pigeons and doves, teals and ducks, monitor lizard, snakes and crocodiles ; a staggering diversity of wild animals served them as a regular supply of

meat. The scrub and forest surrounding the villages were in fact veritable game ranches.

As a matter of fact, as has been pointed out for Africa, the cropping of natural animal communities at moderate rates is an excellent way of utilizing land to which intensive inputs are not possible. The natural animal communities of any region are so evolved that the total community leaves none of the vegetational resources unexploited. The diversified natural animal communities acting as a whole, can maintain a far greater production of animal flesh than the goat, sheep, cattle and buffalo which more or less covers the total range of livestock in India. The wild animals are also hardier and more resistant to draught and disease. If we wish to base animal production, not on an intensively managed animal husbandry, but on an exploitation of free range grazing, then a community of wild animals can perform far better. This is why the supply of animal flesh has come down so drastically with the vanishing of wild life in the Panshet region mentioned above.

Man has yet to explore fully the potentialities of wild life ranching, but the few experiments tried out in Rhodesia and US have been quite successful. We have a number of species such as wild pig, chital, sambar, nilgai and blackbuck which are very promising in this connection (Fig. 13). In fact some very successful game ranches in Texas are based on Indian species such as nilgai and blackbuck. It would be very much worth while to channelise some of the large investment going into wild life conservation to developing the technology of wild life ranching. The problems of management of such ranches would no doubt be rather thorny. Nevertheless, they are not insurmountable and when they are resolved, wild life ranching will go a long way towards ameliorating the problems of protein malnutrition, particularly in hilly and undulating countryside with large tracts of uncultivable waste land or forest land.

Ivory, Musk and Monkeys

Many products of wild animals, apart from meat, have served a variety of purposes since time immemorial. Kautilya's *Arthashastra* mentions, amongst others, hide, bone, teeth, horns, antlers, hooves, tails, bile, feathers, musk and ivory, as products of wild animals that find their way in the Government warehouses. It also describes a number of precious skins coming from Himalayas. Ivory, peacock feathers and musk have been important

items in India's foreign trade since ancient times. Today, India's wild life based trade includes ivory articles, peacock feathers, munias, hill mynas and many other live birds, and above all the rhesus macaque. The rhesus monkey of North India is now a key experimental animal in biomedical research. Some two million of these have been exported from India over the last four decades earning us crores in foreign exchange. This export in such large numbers has been possible because of the protection of vast populations of monkeys of our countryside by the religious tolerance of our masses whose fields and even houses are regularly raided by them. Unfortunately, the people who thus support the monkey populations get nothing in return when the monkeys are trapped and exported.

Wild Life Tourism

Wild life tourism is a major business in certain parts of Africa today, and India, with its rich wild life also has a great potential for development in this direction. Much of our wild life in fact now survives free of all fear only in the vicinity of some tourist lodges in our major wild life sanctuaries. Wild life tourism therefore needs to be continued and further strengthened. However, it should not be solely directed towards the richer classes as it is today. On the contrary we need to emulate the example of Tanzania which is seriously encouraging visits to its great national parks such as Serengetti by rural school children, on whom depends the future of the wild life of any country.

As Gene Banks

The Indian sub-continent and the adjacent lands have been a source of a variety of cultivated plants and domesticated animals. The importance of preserving the genetic diversity of the wild ancestors and relatives of these plants and animals has been realized only recently. We now have a programme for preserving wild rice varieties in Assam for use in future rice breeding programmes whose success will depend on the availability of large pools of genetic diversity to choose from. The same logic applies to domesticated animals and we must maintain stocks of wild animals as gene banks. This particularly applies to our very rich wild fauna of family Bovidae, which includes gaur, mithun, banteng, wild buffalo and wild yak.

Scientific Value

The future development of three exciting areas of modern biology, namely, ecology, animal behaviour and sociobiology depends critically on the availability of populations of higher mammals in their natural setting, i.e. in environments approximating those under which they have evolved. The phenomenon of male infanticide discovered in our Hanuman langur is a particularly striking example of how such observations can lead to major new scientific insights. Our larger social mammals, the chital, the elephant, the hoolock gibbon would all eventually make it possible for us to understand and control human social behaviour. Preservation of their populations under natural conditions is therefore of utmost scientific importance.

Conservation and Management

The Tradition

Wild life, as a resource has rarely been managed in a systematic fashion throughout the course of human history. It has often been conserved due to religious sentiments and for the pleasures of the ruling classes, or has been exploited in a haphazard way. In India, religion has played a critical role in the preservation of populations of monkeys through the length and breadth of the country. The rhesus macaque of North India, thus preserved is now annually earning us crores of rupees in foreign exchange. Wild life has been preserved by strong local taboos as in the vicinity of villages of Vishnois in Rajasthan who believe their clan to have been descended from an antelope. Wild life has also been preserved in India in the hunting preserves of the princes. Kautilya's *Arthashastra* lays down that every prince must set aside two hunting preserves for himself, one near the city, and the other in a wilder locality. The Mughals, the Rajputs and the Marathas all were fond of hunting and maintained their hunting preserves. Such royal hunting preserves served to conserve wild life up to the time of independence, and many have become wild life sanctuaries of today. Dachigam in Kashmir, Gir in Saurashtra, Bharatpur in Rajasthan, Shivpuri in Madhya Pradesh, Similipal in Orissa and Bandipur in Karnataka are just a few examples of this.

The Present Approach

The British Government did nothing for the conservation of our wild life, and in fact, wild life was preserved during the British raj only in the

hunting preserves of princely States. With independence and the dissolution of the princely States the situation seemed desperate in the early years. That famous hunter of man-eating tigers, Jim Corbett, believed that tiger will vanish from India by 1950's. It still survives and in fact the Indian Government has invested several crores in a project to save the tiger. The Government has also created over 150 wild life sanctuaries all over the country. The forest departments are creating wild life wings. All these steps are laudable and have decidedly contributed to the conservation of our rich wild life resources.

There are, however, some serious flaws in this approach. Former Prime Minister, Morarji Desai, made a pointed reference to these weaknesses when Dr. Sálim Ali called on him in early 1977 (Hornbill, April-July 1977, p. 6). As he pointed out, all our wild life conservation effort is directed towards city dwellers whose interest in this cause is largely of an academic nature in as much as they themselves are seldom the direct sufferers of the depredations of the wild life. Nothing, on the other hand, is being done to induct the cooperation of the tribal and the cultivator who are the people in direct contact with the wild animals and often suffer from their depredations.

The current approach at wild life conservation is in fact one of preservation of game in the amusement parks of the privileged classes. Our present-day wild life sanctuaries are places where most of the visitors come to amuse themselves with cards, wine and women and have little interest in the wild life around them. The serious weakness of this elitist approach becomes evident when one considers the problem of urban monkeys. These survive because of the traditional religious protection, and earn valuable foreign exchange. But there is always a hue and cry in our cites to get rid of them. The very same people who argue that an occasional killing of a man by a tiger in Sunderbans or by an elephant in the Annamalais has to be tolerated, believe that monkeys must be shot out because they damage their own ornamental gardens in the cities.

A New Approach

It is obvious that we cannot afford to continue to let the amusement of the rich be the sole basis of our wild life conservation effort. We must instead turn to the genuine economic benefits to be derived from our wild life resources, and see to it that these benefits percolate to the tribals and the peasants who are the real custodians of our wild life.

A major function that the wild life may play in this context is in the

production of protein on scrub and forest land. In any protected locality, such as around the tourist lodge in Bandipur, the rather open and poor forest can be seen to support large concentrations of wild animals. If cropped scientifically on a sustained yield basis, such wild animal populations can be an excellent source of animal protein for our rural population. It was argued above that, in fact, such a function used to be served by the wild animals before they were largely wiped out due to overexploitation and improper management in the last few decades. If only we could set up village co-operatives to supervise the wild life populations around and to ensure that no poaching is allowed, the wild life would flourish again. This in fact happens wherever villagers give spontaneous protection to wild life as is the case with the Vishnois of Rajasthan or Jangams of Maharashtra. The proper harvest could then be made from the wild animals by quietly snaring them at night without causing the animals undue disturbance. No fence is needed to keep in the animals, they will automatically remain in the protected areas near the villages. Admittedly the setting up of a cooperative to manage such a wild life ranch will pose serious problems. But they would certainly be surmounted more easily than problems of educating the villagers to preserve wild life when they do not derive any tangible benefit from it, but do suffer losses from its depredations.

Serious effort needs to go into organizing the monkey and other legitimate wild life trade such as in snake venom so that the people who support this wild life by suffering their depredations get a proper return. Monkeys have definite home ranges so that who supports which troop is quite clear. We should therefore organize village cooperatives to run the monkey trade of the country. Only a certain, properly determined fraction of a troop should be harvested periodically and the proceeds shared amongst those which support that troop. Thus motivated, our monkey resources which are rapidly declining could be assured a long-term conservation and utilization.

We must also ensure that the benefits of wild life tourism flow to the tribals and villagers of the area supporting the wild life. As of now, the local population at best receives a few of the lowest paid jobs in the sanctuaries. All the better paid jobs and profits go to outsiders who have no real feel for or interest in the animals. This ought to change and consideration must be given to the possibility of eventually turning over the wild life tourism to the village and tribal cooperatives of the area concerned, to the people who have a genuine knowledge and understanding of wild life, and

on whose exploitation all our present-day trade and tourism in wild life is ultimately based.

We would also want to protect wild life for the longer term benefits that may accrue from it, whether as gene banks or for scientific investigations or for purely aesthetic reasons. Wherever such preservation without immediate economic benefits is attempted, we must see to it that proper and adequate protection and compensation is given to those who suffer in consequence. Little serious thought is being given today to developing the technology of protection of crops against raiding by wild animals, and no schemes of adequate compensation for loss of crops or livestock exist. Again, it would be very much worthwhile to divert some of the resources going into wild life conservation for these purposes.

The current thinking of our elite on wild life conservation is dominated by the western quasi-religious beliefs in the value of nature conservation. There is no serious thought given at all to the economic benefits derivable from wild life, and no attempt to see to it that these benefits percolate to the tribals and villagers who are the real custodians of our wild life. This has to change.

APPENDIX—I : The distribution and status of the more significant wild life species of the Indian sub-continent. Affinity refers to biogeographical origins and the number I to VIII refer to biogeographic sub-regions. See notes for further details (After Blanford 1901 and Prater 1965)

	Affinity	I	II	III	IV	V	VI	VII	VIII
PRIMATES									
Hoolock (*Hylobates hoolock*)	Oriental								x
Bonnet Macaque (*Macaca radiata*)	Oriental			x	x	x			
Rhesus Macaque (*Macaca mulatta*)	Oriental		x		x	x		x	x
Assamese Macaque (*Macaca assamensis*)	Oriental							x	x
Liontailed Macaque (*Macaca silenus*)+	Oriental					x			
Hanuman Langur (*Presbytis entellus*)	Oriental		x	x	x	x		x	x
Capped Langur (*Presbytis pileatus*)	Oriental								x
Golden Langur (*Presbytis geei*)+	Oriental							x	
Nilgiri Langur (*Presbytis johni*)+	Oriental					x			
Slow Loris (*Nycticebus coucang*)	Oriental								x
Slender Loris (*Loris tardigradus*)	Oriental			x		x			
CARNIVORA									
Tiger (*Panthera tigris*)+	Palaearctic		x	x	x	x		x	x
Lion (*Panthera leo*)+	Ethiopian			x					
Leopard or Panther (*Panthera pardus*)		x	x	x	x	x		x	x
Snow Leopard (*Panthera uncia*)+								x	

National Academy of Sciences, India

Golden Jubilee Commemoration Volume, 1980

APPENDIX—I (Contd.)

	Affinity	I	II	III	IV	V	VI	VII	VIII
CARNIVORA (Contd.)									
Clouded Leopard (*Neofelis nebulosa*)+	Oriental							x	x
Caracal (*Felis Caracal*)	Ethiopian	x	x						
Lynx (*Felis lynx*)	Palaearctic						x		
Spotted Linsang (*Prionodon pardicolor*)	Oriental							x	x
Toddy Cat (*Paradoxurus hermaphroditus*)	Oriental		x	x	x	x		x	x
Mongoose (*Herpestes* sp.)	—	x	x	x	x	x		x	x
Hyaena (*Hyaena hyaena*)	Ethiopian	x	x	x	x	x			
Wolf (*Canis lupus*)+	Palaearctic	x	x	x	x		x	x	
Jackal (*Canis aureus*)	Palaearctic	x	x	x	x	x		x	x
Fox (*Vulpes* sp.)	Palaearctic	x	x	x	x	x	x	x	
Wild Dog (*Cuon alpinus*)+	Palaearctic			x	x	x	x	x	x
Sloth Bear (*Melursus ursinus*)+	Indigenous		x	x	x	x			x
Brown Bear (*Ursus arctos*)	Palaearctic							x	
Himalayan Black Bear (*Selanarctos thibetanus*)+		x						x	x
Red Panda (*Ailurus fulgens*)	Palaearctic							x	
Otter (*Lutra* sp.)		x	x	x	x	x	x	x	x
Marten (*Martes* sp.)	Oriental					x		x	x
Ratel (*Mellivora capencis*)	Ethiopian	x	x	x	x				
INSECTIVORA									
Tree shrew (*Anathana elliotti*)				x	x				
Mole (*Talpa micrura*)								x	x

APPENDIX—I (Contd.)

	Affinity	I	II	III	IV	V	VI	VII	VIII
CHIROPTERA									
Flying Fox (*Pteropus giganteus*)		x	x	x	x	x			x
RODENTIA									
Giant Flying Squirrel (*Petaurista* sp.)			x	x	x	x		x	x
Giant Squirrel (*Ratufa* sp.)				x	x	x		x	x
Porcupine (*Hystrix indica*)		x	x	x	x	x		x	x
LAGOMORPHA									
Indian Hare (*Lepus nigricollis*)		x	x	x	x	x	x	x	x
Hispid Hare (*Caprologus hispidus*)+			x					x	x
Himalayan Mouse Hare (*Ochotona rolylei*)	Palaearctic							x	
PROBOSCIDEA									
Elephant (*Elephas maximus*)+	Oriental				x	x		x	x
PERISSODACTYLA									
Asiatic Wild Ass (*Equus hemionus*)+	Palaearctic	x							
One-horned Rhinoceros (*Rhinoceros unicornis*)+	Oriental							x	x
ARTIODACTYLA									
Gaur (*Bos gaurus*)+	Oriental				x	x			x
Banteng (*Bos banteng*)+	Oriental								x
Yak (*Bos grunniens*)+	Palaearctic						x		

National Academy of Sciences, India

Golden Jubilee Commemoration Volume, 1980

APPENDIX—I (Contd.)

	Affinity	I	II	III	IV	V	VI	VII	VIII
ARTIODACTYLA (Contd.)									
Wild Buffalo (*Bubalus bubalis*)+	Oriental		x		x				x
Wild Sheep (*Ovis* sp.)	Palaearctic	x					x		
Bharal (*Pseudois nayaur*)	Palaearctic						x	x	
Wild Goat (*Capra* sp.)+	Palaearctic	x					x		
Himalayan Tahr (*Hemitragus jemlahicus*)	—							x	
Nilgiri Tahr (*Hemitragus hylocrius*)+	—					x			
Goat-Antelopes (Fam. Rupicaprinae)	Oriental							x	x
Tibetan Antelope (*Pantholops hodgsoni*)	Palaearctic						x		
Chinkara (*Gazella gazella*)	Ethiopian	x	x	x					
Blackbuck (*Antilope cervicapra*)	Indigenous	x	x	x	x				
Four-horned Antelope or Chousinga (*Tetracerus quadricornis*)	Indigenous		x	x	x				
Nilgai (*Boselaphus tragocamelus*)	Indigenous	x	x	x	x				
Hangul, Kashmir Stag (*Cervus elaphus*)+	Palaearctic							x	
Thamin (*Cervus eldi*)+	Oriental								x
Barasingha or Swamp Deer (*Cervus duvaceli*)+	Oriental		x		x				x
Sambar (*Cervus unicolor*)	Oriental		x	x	x	x		x	x
Hogdeer (*Axis porcinus*)	Oriental		x						x
Chital (*Axis axis*)	Indigenous		x	x	x	x			x
Barking Deer, Muntjac (*Muntjacus muntjak*)	Indigenous			x	x	x		x	x
Musk Deer (*Moschus moschiferus*)+	Palaearctic						x		

APPENDIX—I (Concld.)

	Affinity	I	II	III	IV	V	VI	VII	VIII
ARTIODACTYLA (Contd.)									
Mouse Deer (*Tragulus meminna*)	Oriental			x	x	x			
Wild Boar (*Sus scrofa*)		x	x	x	x	x		x	x
Pygmy Hog (*Sus sylvanius*)+									x
PHOLIDATA									
Pangolin (*Manis* sp.)		x	x	x	x	x		x	x

Note :

I = Punjab
II = Gangetic Plain
III = Western Peninsula
IV = Eastern Peninsula
V = Western Ghats
VI = Ladakh
VII = Himalayas
VIII = Assam

\+ Denotes a species listed as a threatened species in the Red Data Book.

X Denotes that the species occurs in a given subregion.

APPENDIX II

Scientific names of species mentioned in the text, but not covered by Appendix I.

Marbled Cat	*Felis marmorata*
Cheetah	*Acinonyx jubatus*
Malabar Ciet	*Viverra megaspila*
Ferret-Badger	*Melogale* sp.
Gangetic Dolphin	*Platanista gangetica*
Markhor	*Capra falconeri*
Urial or Shapu	*Ovis orientalis*
Serow	*Capricornis sumatrensis*
Goral	*Nemorhaedus goral*
Takin	*Budorcas taxicolor*
Flamingo	*Phoenicopterus roseus*
Lesser flemingo	*Phoeniconaias minor*
Mugger	*Crocodilus palustris*
Gharial	*Gavialis gangeticus*

ACKNOWLEDGEMENTS

I am grateful to numerous friends in the academic community, amongst wild life enthusiast and forest officers, and last but not the least amongst villagers and tribals who have taught me much about India's wild animals. I am particularly grateful to my colleagues P. V. K. Nair, H. C. Sharatchandra and S. N. Prasad for help in field-work, for stimulating discussions and much new information. I am thankful to Shri B. N. S. Deo for the use of his magnificent photographs.

REFERENCES

1. Blanford, W. T. (1901) *Philosophical Transactions of Royal Society*, Series B, 194 : 335.
2. Gee, E. P. (1964) 'The Wild life of India', Collins, London : 224.

3. Nair, P. V. K. and M. Gadgil (1978) J. Bombay Nat. Hist. Soc., Salim Ali Commemoration Volume (*in press*).
4. Prater, S. H. (1965) 'The book of Indian animals', Bombay Natural History Society, Bombay : 324.
5. Seshadri, B. (1969) 'The twilight of India's wild life', John Baker, London : 212.
6. Wakankar, V. S. and R. R. Brooks (1976) 'Stone age painting in India', Taraporewala, Bombay : 116.

On The Carriage And Release Of Stores From Aircraft

Sridhar M. Ramachandra

Professor of Aeronautical Engineering
Alfateh University, Tripoli, Libya

Formerly Deputy Chief Design Engineer
Hindustan Aeronautics Limited
Bangalore 560017, India

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

1. Introduction

The advent of military aviation has brought in the need to carry and deliver large quantities of offensive and defensive warload during military operations. Although sometimes this warload is carried internally in the aircraft in special compartments, more often the warload is carried externally on pylons. In this review we shall designate all jettisonable materials carried on board the aircraft externally and/or internally by the generic name STORE. In the early days of military aviation the proportion of externally/internally carried stores formed only a small fraction of the basic aircraft weight. With the continuing improvements and sophistication in the performance and effectiveness of combat aircraft and their inevitable cost escalation, it has become imperative to improve the cost effectiveness of these operations. Consequently, aside from the improved quality and diversity of the stores carried, the quantity of these stores is approaching a significant fraction of the basic aircraft weight. Whereas stores were carried usually under the wing, it is becoming increasingly common to mount them on the fuselage under-surface also together with increasingly strong trends towards their carriage on the wing top surface. Single store installations are yielding place to twin, triple and even multiple store stations through the use of TER and MER systems. Increasing trends are present for tandem, axially staggered and conformal carriage configurations. Conformal carriage would permit supersonic cruise capability besides also minimizing the incremental radar cross-section caused by carriage of the warload thereby enhancing the cost effectiveness and survivability of the weapon system sizeably.

Jettisoning externally carried drop tanks after their exhaustion is essential for crucial missions to augment the combat radius and acceleration capability before and after combat in addition to obtaining sizeable gains in combat performance characteristics.

Stores have to be jettisoned in flight not only during rectilinear flight but during maneuvers as well. Consequently, to overcome the difficulties and limitations imposed by the freedrop method power assisted store ejection is fast replacing the gravity release system. The carriage of air-to-air and air-to-ground missiles for air combat and ground support poses peculiar problems for their firing.

Although stores carriage and release problems originated from the requirements of the military, they are equally important during peace time for

providing relief to stranded populations by air-dropping food, clothing, medicines during periods of natural disasters like earthquakes, floods, tidal waves, volcanic eruptions, etc. Even though such operations have been carried out earlier with internally carried stores, improved encapsulation techniques may be used to carry the relief supplies both internally and externally on military aircraft.

Interesting applications of stores carriage and release have been the launching of the X-15 aircraft from the B-52 aircraft belly and the piggy-back launching of the American space shuttle from the Boeing 747.

Besides airplanes, helicopters have been frequently used as sky-cranes for carrying houses, lumber, steel structures, etc., to otherwise inaccessible areas and span cables across rivers, valleys, canyons, etc. Offshore exploration of fossil fuels, minerals and other natural resources is increasing the role of helicopters in the erection of floating drilling platforms and other similar systems.

Almost every characteristic of an aircraft is affected by the external carriage of stores. In addition to increasing the aircraft weight, the carriage and release of these stores causes considerable changes to the moments of inertia and location of the center of gravity, degrade stability, reduce lift, increase drag often by far more than the store drag itself, and sometimes give rise to flutter, thereby leading to serious operational limitations arising from undesirable flight handling problems. Their external carriage alters the aerodynamic characteristics both during carriage and release leading to significant adverse effects in terms of carriage loads, performance and handling.

Earlier, military aircraft design practices concentrated on the development of a good basic clean aircraft. Externally carried stores and their mounting fixtures were appended thereafter accepting the resulting penalties/limitations in performance and handling. Current approach emphasizes satisfactory operational performance and handling qualities in the ultimate flight configuration of the aircraft in which the carriage and release problems of stores form a natural concomitant. Furthermore, air forces the world over may be expected to demand far higher performance for future generations of combat aircraft posing difficult challenge to the aerodynamicist. A revival of interest in swept forward and other hybrid wing configurations may be expected to achieve this performance augmentation incorporating specially

aeroelastically tailored wings. Wing mounted stores may pose hazards of inboard store movement after-release increasing the adverse aerodynamic interference during carriage for the forward swept wing configuration. It is important to evolve new departures from the conventional under-wing store carriage used on the swept-back wing configurations. A proper understanding of the airframe-store aerodynamic interactions will make a significant contribution to the development of such a low drag weapon system.

A primary requirement for carrying stores on board an aircraft is the ability to jettison them safely over the entire operational envelop with no collision between them after release. The fundamental problem arising in this exercise is the ability to predict safe operational release envelop for flight demonstration during development programs. An additional problem would be the extension of the carriage and safe release envelop of an existing aircraft beyond an existing domain in the event of such a need to carry newer varieties of stores. Further, an accurate prediction of complete flight trajectories from the release point to the impact point on ground or in air would increase the effectiveness of offensive operations.

For a satisfactory performance of the foregoing objectives, the ability to calculate the aerodynamic characteristics of the flight vehicle and the store separately and in combination is crucial as the governing equations of motion of both are intimately related to these characteristics. Theoretical methods, wind tunnel techniques and flight test methods have been employed towards this objective. Studies of the carriage and release characteristics of stores have formed a subject of extensive theoretical and experimental investigations in several countries. However, due to the inherently sensitive nature of the generated information the results of these researches are not very extant in the open literature. Schindel[69] has surveyed the problem of store separation somewhat extensively. The following paper attempts to review the state-of-art in this area relating to the carriage and release problems of the stores as well as the response of the aircraft and project avenues of research for future work. In this paper the review will be restricted to problems arising on fixed wing aircraft only due to its relatively greater importance to the military. Research in this field has been extensive and often of a classified nature. Consequently, a large body of the papers is unfortunately not accessible to the author for inclusion in this survey. Further, even among the published papers many may have escaped attention. For all these a note of apology is expressed here.

2. Trends in Carriage of Stores

Conventionally, stores have been carried under the wings, under the fuselage and at the wing tips, depending on the nature of the stores. In order to overcome the adverse effects of stores carriage under the wings, as mentioned earlier, new locations have been explored using conventional stores. When problems of safe store release are insurmountable using conventional and simple solutions over the required operational envelop, alternative mechanical solutions may be attempted without resorting to power jettisoning devices. These solutions are also useful to extend the required release envelop either with old or new stores types. Thus, Short *et al.*[72] have reported on the use of a pivot mechanism on the F-15 aircraft to obtain an extended jettison envelop. A similar pivot mechanism has also been used on the Soviet MiG-21 aircraft to obtain safe gravity store release over the entire speed range whereby aerodynamic and gravity forces rotate the store about a pivot located in the pylon up to a certain angle at which it is automatically free from the aircraft. Power ejection systems with provision for obtaining graduated thrust differentials between the forward and the rear thrusters permits usage of a wider range of stores when faced with difficulties of safe release using gravity alone. However, avoidance of performance penalties associated with conventional store locations demands study of alternative locations including modifications to store aerodynamic shapes. Conformal carriage, axial stagger and tandem store carriage are the other alternatives. In addition, over-wing carriage may also be attempted for warload augmentation coupled, of course, with power ejection.

The conformal carriage of stores provides an aerodynamically smooth fairing to house an array of individual bomb racks under the fuselage (Fig. 1). Conventional weapons are then carried tangent to the fuselage conformally on adaptors. This method improves both the store separation characteristics and the aircraft performance through a significant drag reduction as also a considerable improvement in handling characteristics. Thus, conformally carried weapon release tests of both conventional and bluff body shapes reported by Smith[75, 76] from the phantom F-4B from Mach 0·6 to 1·6 and over altitudes from 1524 m to 9144 m showed safe and uniform separation with little pitch or yaw motion for both 30° dive release and ripple mode with little lateral dispersion. Supersonic releases showed a small initial pitch up and pitch down as the weapon passed through the aircraft

bow shock wave. Despite its drag reducing abilities, conformal carriage has also its limitations which require attention for alleviation. Thus, when a variety of stores of different lengths need to be carried on a spatially fixed ERU (ejector release unit) installation, the resulting gaps may cause a sizeable rise in the installed drag. Further, stores suited for close packing in conformal carriage are unsuitable for under-wing carriage on conventional carriers due to their high free air drag. Solutions can be found to some of these problems by adjustable under-fuselage ERU positions to suit different store lengths rather than develop stores tailored for such carriage.

Present store carriers have large structural protuberances for holding the store during carriage. These protuberances vitiate the aerodynamic cleanliness of the store installation and contribute excessive drag both during store carriage and after store release. The use of MACE (Minimum Area Crutchless Ejector)-type carriers can reduce the installed drag of single stores by over 25% and has strong possibilities for multiple carriers too.

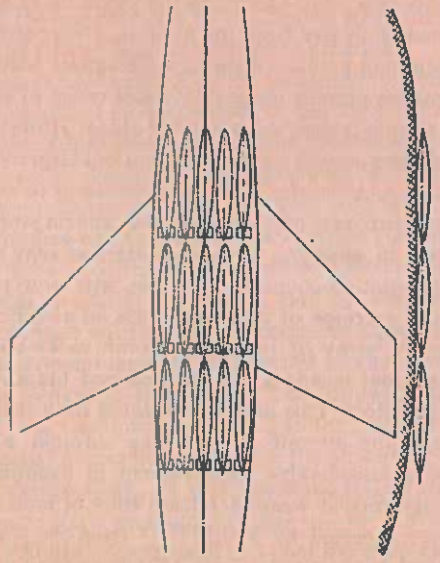


Fig. 1. Conformal stores carriage

Placing one store in the wake of another (Fig. 2) thereby shielding it from the free stream would reduce its drag significantly. The maximum

benefit of this tandem arrangement is obtained by close package of stores with large flat bases or wide wakes and is effective at all speeds. Tandem carriage of stores with relatively high free air drag can give drag savings of up to 25% especially for under-fuselage locations obviating to a large extent problems of center of gravity shifts and ejection velocity variations due to beam flexibility.

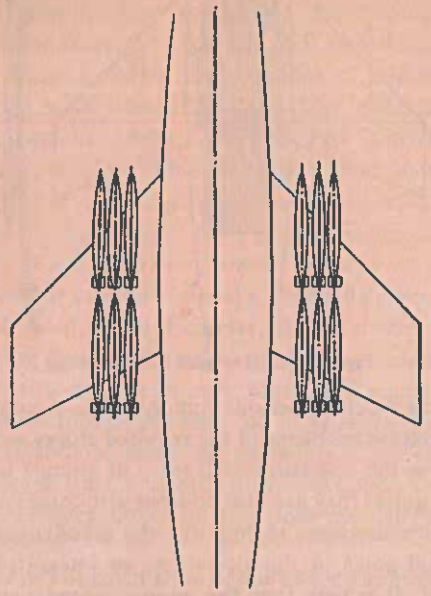


Fig. 2. Tandem triple stores carriage

For drag reduction at transonic speeds with stores of high wave drag, a subtle method would be the axial stagger (Fig. 3) whereby the pressure fields of adjacent stores are so disposed as to promote their mutual cancellation thus delaying the onset of shock waves. Hence, axial stagger is most beneficial at transonic speeds when using stores with a relatively large wave drag component in the total drag. It is possible to combine tandem carriage and axial stagger to realise large benefits when carrying groups of stores.

The introduction of retarded bombs, retard-accelerate bombs, bluff bombs, new missiles and other stores conceived in isolation from the airframe performance bring in serious aerodynamic, structural and flight

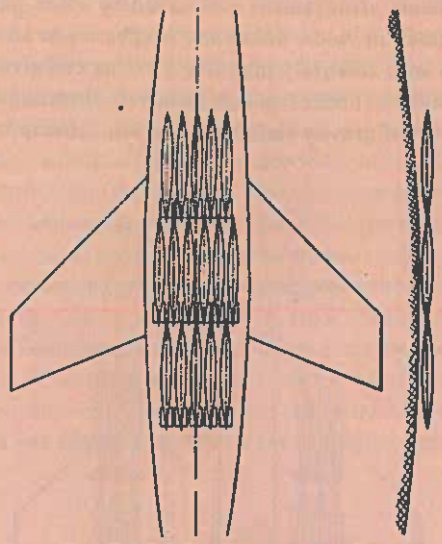


Fig. 3. Axial stagger stores carriage

handling problems requiring serious compromises. Aerodynamic solutions to carriage and release problems of the required stores must continue to be studied to improve the operational efficiency of aircraft with existing store types. However, unless they are also coupled with conscious attempts on the part of the weapons designers to improve the aerodynamic configurations themselves, overall gains in the aircraft as an integrated weapon system would be limited. It is here that the weapon aerodynamicist must work closely with the airframe designer for an optimal weapon system. The process is a long one and its importance cannot be overestimated.

3. Theoretical Methods—Trajectory Prediction

The value of the theoretical methods and wind tunnel methods including computer graphics simulations lies in the enormous savings realized in curtailing the number of hazardous flight drop tests materially, especially when the cost of the missiles and other weaponry is increasing continuously. Nielsen and his associates (refs. 21, 29, 43, 56) have reported extensive computer graphics simulation of the six-degree-of-freedom equations of motion using an elaborate mathematical source/sink modelling of the flow

field for calculating the aircraft-store interactions and store release trajectory prediction. They have obtained store trajectories including the effects of induced velocity fields of air-inlets. Their calculated trajectories show good agreement for the pitch and yaw motions with the results of captive store wind tunnel measurements. Pugh, *et al.*[62] have reported the interim Royal Aircraft Establishment experience with the application of Nielsen's method. They find that frequently, "the outcome of a release is substantially different from the expected and modest changes in initial conditions often required during investigations of release problems of an aircraft-store combination can alter the whole character of the store trajectory". The personality of the method appears to demand an excessive price for increased precision within a range of store attitudes. Although they had no difficulty in obtaining satisfactory source/sink distributions on ogive-cylinder bodies, RAE experienced great difficulty when dealing with a body comprising of a 2 caliber long ogive followed by a 4 caliber long cylinder and a 2 caliber long conical wake. Difficulties were experienced in dealing with blunt nosed bodies. However, limited studies of the accuracy of Nielsen's method have given most encouraging results. Berry[7] gives examples of aircraft-store interference. McKinney *et al.*[48] have reported extensive NASA wind tunnel test data for external stores in the carriage position. Their calculations show the importance of airplane speed, dive angle and load factor on the release trajectories and also show the significant effects of buoyancy forces on the store normal force and pitching moment characteristics.

Analytical methods of calculation of the forces and moments on external stores have been reported by Crose *et al.*[19], Fernandes[23,24], Looney[40] Martin *et al.*[45,46], Walkley[90], Smith[74], Chadwick[9], Sullivan[87] and Serbin[71] all of whom use extensive aerodynamic flow field modelling around the aircraft-store combination. Although detailed flow field studies around the complete airplane-store combination are important for a complete understanding of the aerodynamic interactions to enable reliable prediction of release trajectories, simpler approximate approaches are also essential to plan initial tests for exploratory purposes. Schoch's criterion or Covert's[16] criterion can be used to establish store separation boundaries for preflight planning of store drop tests. Covert[16] has shown that the initial velocity of each point on a store depends not only on the location of the ejection mechanism with respect to the center of mass and store moments of inertia but

also on the elastic properties of the rack and pylon and the ejector thrust-time variation program. He showed through a series solution that the initial acceleration of any point on the store is determined by the rack-mate loads effect (the second order term) and that the flow field non-uniformities affect the third and fourth order terms. Covert showed also that each store separation could be mapped as a point of a velocity-acceleration plane whose abscissa is the initial acceleration and the ordinate is the initial velocity. He found a well-defined locus in this plane dividing the plane into a region of safe separation and a region of unsafe separation. In the general case of store separation from a maneuvering aircraft, safe separation criteria are defined using the relative velocity and relative acceleration between the store and rack at the instant of release. Structural flexibility of the MER (Multiple Ejector Release) reduces the ejection velocity of the store, lowers the ejection pitch rate and induces a rolling moment. Covert showed that a heavy rack imparts the largest fraction of energy for store ejection. Further, short, dense, high drag stores separate better than aerodynamically good light weight stores, as may also be expected intuitively. Considering the pitch plane bending and the store pitch dynamics, Spearman[80] finds that structural flexibility accounts for no more than 10% deviation from the rigid case for the ejection velocity whereas ejection angular velocity is affected more significantly. McKinney *et al.*[48] have given the separation characteristics of external stores configurations. Holloway *et al.*[31] have given separation criteria for densely packed stores carried in bomb bays. Cox *et al.*[17] have given comparisons of predicted and actual flight release trajectories of externally carried stores. Ramachandra[64] has recently given a comparison of the predicted and flight test trajectories of drop tanks using a photographic technique. A statistical method for predicting the launch characteristics of stores has been developed by Sekellick[70] who used a multiple-regression analysis to obtain an empirical relation between the initial pitch maximum and the line integral of the store nose distance from the initial launch position by expressing the configuration geometry and the flow condition. This method may be extended to a composite scoring method to predict hazardous store separations.

Current analytical and experimental methods presuppose a steady aerodynamic condition in the airflow around the store during its free motion. All available trajectory comparisons relate to theoretical and captive model wind tunnel simulations and indicate excellent agreement whereas

Ramachandra's recent studies[64] have shown large disparities. The motion of the store after its release from the aircraft is very complex. Even though the general motion is vertical, it is accompanied, particularly in the case of empty or low density stores, by pronounced tumbling in both the vertical and horizontal planes in a somewhat random manner as observed in the studies of Ramachandra. Store trajectories of externally carried stores have also been studied by Ramachandra, *et al.*[66] for determining the store aerodynamic characteristics using parameter identification techniques on flight test motion picture records of the store separation paths. These studies have shown the necessity of considering the entire range of flow angles from 0–360 degrees in both incidence and yaw for a satisfactory trajectory prediction from the moment of release to its impact with the ground. Also, angular velocities close to 6 radians per second have often been encountered soon after release in the flight tests of drop tanks reported by Ramachandra *et al.*[65]. Furthermore, in the case of power ejected stores, their studies of the trajectories showed sensitiveness to initial changes in the initial store pitch velocity $\dot{\theta}$ similar to Covert's observations[16]. Analytical prediction of the store aerodynamic coefficients over this large range of flow angles is not possible, at present, especially in view of the possible flow separation over the store at these large angles and the rapidity and randomness of the tumbling motion. However, due to the fluid dynamic lag, it may be possible to neglect flow separation even at these high incidences and consider the store motion as in an inviscid fluid but taking into account apparent mass effects. For calculating the aerodynamic coefficients for both the steady and unsteady case, higher order nonlinear terms in $(\alpha, \dot{\alpha})$ representing incidence and incidence rate have to be considered to account for the large variations experienced. Besides, store angular motion of a monotonic rather than an oscillatory motion needs to be considered for calculating the aerodynamic characteristics.

4. Effect of Stores Carriage on Performance and Aircraft Stability

Taisseire[68] and Spearman[79,83] have investigated the effects of carriage of external stores on the stability characteristics of aircraft. Wind tunnel tests on a 60° delta wing model of a fighter showed little change in the longitudinal stability with the addition of pylons and missiles, giving a small destabilizing trend at low Mach numbers to stabilization with increasing Mach numbers. The addition of pylons increased the minimum drag but

reduced the lift dependent drag so that store drag had little effect beyond 10° angle of attack. The maximum lift-drag ratio was reduced by the addition of stores. The static directional stability increased at $M=1{\cdot}6$ but had little benefits at higher Mach numbers. Directional stability, stores on, is maintained up to sufficiently high angles of attack ($\approx 19°$). The effective dihedral increased with stores on up to $M=1{\cdot}9$ and became zero or slightly unfavourable beyond $M=1{\cdot}9$. Wind tunnel tests reported by Whoric[94] on the phantom F-4C showed an aggravation of the tuck-under characteristics of the baseline configuration with the addition of winged external stores. Addition of external stores caused loss of elevator power and elevator effectiveness in the transonic Mach number range so that increased stabilator deflections were required for trimming. The addition of external pylon mounted stores caused a large shift of up to 8% in the neutral point location in the transonic range peaking at $M=0{\cdot}925$ followed generally by a sharp afterward neutral point shift with increase of M to 1·05.

Release of stores from aircraft causes a step change in the aerodynamic and inertia characteristics of the aircraft. This step change imparts both linear and angular accelerations to the aircraft. Ramachandra[88] has investigated this problem and given analytical expressions for the resulting perturbed motions. The effect of angular momentum of the power plant is such as to impart a yaw motion and a roll motion even in the case of symmetric store jettison. He has shown the possibility of determining the installed aerodynamic characteristics of the store from a measurement of the perturbed motions. Pinsker[59] has also given a method of determining the installed drag of stores from a measurement of the aircraft acceleration after a store jettison.

5. Aerodynamic Interference

Due to the close proximity between the aircraft and the stores carried, a satisfactory prediction of their aerodynamic characteristics requires a study of the strong mutual interference field between them. Further, since most problems of carriage and release of stores arise in combat aircraft and particularly at high speeds, it would be important to consider compressibility effects in these interference studies. An extensive review of the wing-body interaction problem has been carried out by Ashley *et al.*[8]. Morino[50-54] and his group have developed the Green's function and finite element methods for studying this problem around complex configurations in both

subsonic and supersonic flow and have developed computer programs for these calculations. Woodward[95,96] has developed computer programs for this purpose in both the subsonic and supersonic domains. Ramachandra *et al.*[66] developed computer programs to calculate the wing-body interference field and used it to design an aerodynamically compensated pitot-static tube located at arbitrary positions. An interference potential concept was developed by Bera[6] to calculate the interaction potential field around a family of wing-body configurations. Measurements of the interference field of some wing-body combinations were made by Siddalingappa and *et al.*[73] in low speed flow.

Ashley *et al.*[8] classify all wing-body interaction theories as finite element methods since these attempt to satisfy the boundary condition at a finite number of stations and hence determine the strengths of a finite number of singularities. Some theories are suitable for the calculation of pressure distributions and load over the aircraft surface so that the forces and moments may be determined by numerical integration. The Woodward[95,96] and the NASA Langley methods compute some or all of the forces from the momentum flux through a far-field control volume. Kernel-function methods have been proposed by Stark[85], Watkins *et al.*[92,93] and Laschka[88,89] for a practical solution of the oscillatory lifting surface problem involving the kernel of the pressure-downwash equation. The pressure loading functions assumed in the kernel function methods are not always suitable and have to be avoided especially when dealing with solutions for non-planar interfering configurations arising in store carriage problems. This is especially so when it is possible to increase the number of unknowns and the complexity of the system of equations which can be solved using current digital computers.

If $\Phi(x, y, z, t)$ be the velocity potential at a point (x, y, z, t) of the flow field, we have the governing differential equation

$$C^2\nabla\Phi-\frac{\partial^2\Phi}{\partial t^2}=\frac{\partial}{\partial t}(\nabla\Phi)^2+(\nabla\Phi)\cdot(\nabla\Phi-\nabla\nabla\Phi) \tag{5.1}$$

where C is the local speed of sound. To deal with the free motion of a store after release from the aircraft we have to slove eqn. (5.1) from which the perturbation velocities and the surface pressures on the aircraft and the store may be obtained as a function of the time t. For the relatively simpler problem of the carriage of the stores we need only to deal with the steady form of eqn. (5.1) written as

$$C^2\nabla^2\Phi=\nabla\Phi\cdot(\nabla\Phi\cdot\nabla\nabla\Phi) \tag{5.2}$$

Most problems of practical importance relate to the homentropic case for which eqn. (5.2) may be written as

$$C_\infty^2 \nabla^2\Phi = \nabla\Phi \cdot (\nabla\Phi \cdot \nabla\nabla\Phi) + \tfrac{1}{2}(\gamma-1)(\nabla\Phi)^2 \nabla^2\Phi \tag{5.3}$$

where C_∞ is given by

$$C_\infty^2 = C^2 + \tfrac{1}{2}(\gamma-1) Q^2 \tag{5.4}$$

with the suffix ∞ indicating the total or stagnation value of the quantity on the given streamline, γ is the ratio of specific heats of the fluid and $\vec{Q}$ is the velocity vector $\vec{Q} = \nabla\Phi$. C_∞ is now constant over the whole flow field. The pressure p at any point may be defined through the pressure coefficient C_p given by

$$C_p = \frac{p-p_\infty}{\frac{1}{2}\rho_\infty U^2} = \frac{2}{\gamma M^2}\left\{\left[1 - \frac{\gamma-1}{C_\infty^2}\left(\Phi_t + \frac{\vec{Q}\cdot\vec{Q} - U^2}{2}\right)^{\frac{\gamma}{\gamma-1}} - 1\right]\right\} \tag{5.5}$$

where U is the free stream velocity. Exact solutions of the nonlinear equation eqn. (5.3) have been found by special methods in a few cases only while numerical solution methods have been used for others. Generally, linearization of the eqn. (5.3) is resorted to using the hodograph transformation in 2-dimensions or power series expansion in terms of some parameter or variable assuming its convergence for sufficiently small values of the variable or parameter. Postulating a slender body enables derivation of a linear equation for the perturbation velocity potential φ which can then be solved with the appropriate boundary conditions for a given problem. For this purpose we may define a slender body as a body whose local thickness is small compared to its length in all directions normal to the free stream. The linearized equation of motion for the perturbation velocity potential φ defined by

$$\Phi = Ux + \varphi(x, y, z, t)$$

satisfies the equation

$$(1-M^2)\varphi_{xx} + \varphi_{yy} + \varphi_{zz} = \frac{2M^2}{U}\varphi_{xt} + \frac{M^2}{U^2}\varphi_{tt} \tag{5.6}$$

The pressure at any point given by the exact equation may be linearized into the form

$$C_p = -\frac{2u}{U} + (M^2-1)\left(\frac{u}{U}\right)^2 - \frac{v^2+w^2}{U^2} \tag{5.7}$$

for steady flow over slender bodies and as

$$C_p = -\frac{2u}{U} - \frac{2}{U^2}\varphi_t \tag{5.8}$$

for steady flow over thin wings.

The perturbation velocity components u, v, w, defined by

$$(u, v, w) = \text{grad}\ \varphi \tag{5.9}$$

satisfy the condition

$$u, v, w \ll U$$

and the perturbations of the state variables p, ρ from the free stream condition p_∞, ρ_∞, are small. To linearize the equations and remove the transonic nonlinear terms in the differential equations two further conditions are used as follows

$$\left.\begin{aligned}(M\delta)^2 &\ll 1\\ (M^2\delta)(\gamma+1)/(M^2-1) &\gg 1\end{aligned}\right\} \tag{5.10}$$

where δ is a perturbation parameter representing the body surface slope. To these equations of motion must be added the Neumann boundary condition that the fluid velocity at the surface equals the surface velocity at every point of the surface. Thus, if the surface immersed in the fluid be described by the equation

$$F(x, y, z, t) = 0 \tag{5.11}$$

the Neumann boundary condition may be written as

$$DF/Dt = (\partial/\partial t + \vec{Q}\cdot\text{grad})\, F = 0 \tag{5.12}$$

which reduces to zero normal velocity at the surface in steady flow. Assuming the perturbation parameter $\delta \ll 1$, the Neumann condition for the surface represented by

$$F = z - z(x, y, t) = 0 \tag{5.13}$$

generally used for a thin wing at small incidence in an unsteady motion and nearly parallel to the x-y plane may be written as

$$(\partial z/\partial x,\ \partial z/\partial y,\ \frac{1}{U}\partial z/\partial t) \ll 1 \tag{5.14}$$

so that the velocity component w normal to the x-y plane becomes

$$w(x, y, z, t) = (\partial z/\partial t) + U(\partial z/\partial x) \tag{5.15}$$

which is generally enforced at the surface $z=0$ and not at the surface itself.

Extensive discussions of the linearized equations are available in the well-known works of Ward, Miles[49] and others for flow over wings and bodies. Instead of dealing with these equations, Takashima[89] has developed a virtual mass method for calculating the interference flow field of strap on multibody configurations commonly used on satellite launch vehicles. Stahara *et al.*[84] have given a rapid predictive method for flow fields of airplanes with external stores in the transonic domain. Chou[11] has extended Borland's[8]

image concept for bodies of elliptical cross-section to include nacelles. He also attempted a generalized image approach so that longitudinal cross-section area variations could be considered and has given methods of calculating the aerodynamic loads on aircraft structures including the stores. Spangler and Mendenhall and Giesing *et al.*[28] have extended the image method to complicated configurations including both nacelle and pylon effects. The Giesing *et al.*[28] method is very general and can determine both the spanwise and chordwise loading of an assemblage of lifting surfaces (wings, tails, pylons, control surfaces) and bodies (fuselage, nacelles, stores, etc.) in or out of ground effect for either steady or oscillatory motion. This method is a combination of the lifting surface theory, the method of images and the slender body theory. The doublet lattice method models the lifting surfaces while the slender body theory models the bodies. The interference between the lifting surface and the bodies is obtained by the use of images and the slender body theory. The lifting surfaces may be of arbitrary planform and dihedral and the bodies may have elliptic cross-section with an arbitrary distribution of width and radius. Vortex lattice methods have been used by Nayfeh *et al.*[4, 5] for calculating the aerodynamics of wings and wing-body combinations. But when dealing with problems of complex configurations of stores carried externally on an aircraft this linearized superpositional representation is inadequate to obtain near field interaction effects either for the store carriage condition or the store release trajectory determination.

6. Experimental Methods

Experimental methods consist principally of three techniques, the captive model method, the flow field method and the drop model method. Coste[15], Furey *et al.*[26], Christophe *et al.*[13] and Landers *et al.*[36] have reported wind tunnel studies of store separation characteristics using the captive model technique. In the captive model method, a scale model of the store is carried on an independent sting near the basic aircraft model in a wind tunnel. The sting carries its own miniaturized balance for measuring the forces and moments on the store model. Using the measured forces and moments the store's linear and angular positions are incremented over a small time interval using the six-degree-of-freedom equations of motion. This process is repeated until the store is out of the interference flow field of the aircraft. Clearly, the larger and finer the number of time intervals used the greater is the accuracy of the store trajectory. Practical limitations

to miniaturization of the sting and the associated balance arising from the requirements of mechanical strength and tunnel size introduce errors in the measured forces and moments on the store model due to sting interference thereby vitiating the usefulness of the results to some extent. Some measurement systems are sophisticated using an on-line computer controlled sting-model position and attitude control system to simulate the entire trajectory in the wind tunnel. The method is most suited to continuous wind tunnels and introduces further complexities and errors in intermittent wind tunnels quite apart from the usual scale effects common to both types of tunnels.

The flow field method or the grid method consists of measuring the perturbation velocity field of the airplane comprising of the magnitude and direction of the velocity at each point. Such a method has been described by Landers, *et al.*[36] using a yawmeter system for measuring the flow field on the captive trajectory. The forces and moments acting on the store at each point are calculated from this flow starting from the initial store location for the corresponding store attitude and location. This method assumes negligible interaction between the store flow field and airplane flow field. Again, the number of points at which the flow field data has to be measured would be large and the resulting computations of the store motion complex. As with the captive model method, the grid method is most convenient for continuous wind tunnels.

The drop model method consists of observing the trajectories of dynamically similar scaled store models in free flight in the wind tunnel. Trajectories may be recorded photographically using high speed motion picture cameras mounted in two perpendicular directions, in the side wall and the top or the bottom wall of the wind tunnel. Using the dynamical similarity relations the wind tunnel trajectory may be scaled to full scale. Apart from the inevitable scale effects on the trajectories, the utility of the results is vitiated by the errors of store model, imperfections in geometry and inertia scaling, store setting and release errors, tunnel flow unsteadiness and wall effects. Besides requiring a large number of carefully prepared store models, practical difficulties are introduced due to the need to stop the tunnel to retrieve the model and replace by another in case of damaged models. Narayana, *et al.*[55] have developed a multiple exposure high speed photographic method at Hindustan Aeronautics Limited by which the trajectories of bodies may be recorded in wind tunnels. The bodies under study in their

tests were scaled geometrically and inertially relative to the full scale so that wind tunnel store trajectory photographs could be related to the full scale. They applied this method to the determination of the release paths of airplane canopies from aircraft models in wind tunnels. This technique has also been applied earlier by McCabe *et al.*[47] for the Harpoon/P-3 Orion separation program in subsonic wind tunnels. In this method, due to the high speeds involved there is added danger of damage to the wind tunnel structure caused by the fly-away store model, especially in high speed tunnels. Further, aerodynamic coefficients obtained from captive model tests or grid method or the drop model tests in wind tunnels are incapable of simulating the unsteady flow conditions due to the random store rotation velocities encountered in free flight.

Brooks[10] gives a comparison of store trajectories using the captive trajectory, flow field survey and the drop model data with predictions and flight test data in both the horizontal and the vertical planes. Spahr[77] provides comparison of predicted store trajectories with the flight drop test results on an F-4D aircraft accompanied by a T-3E chase aircraft. The sophisticated photographic observations were supplemented by telemetry data from the store itself released from ejector TER racks. Differences between the calculated and observed separation trajectories were attributed to variations from cartridge to cartridge thrust and friction from rack to rack giving rise to differences in the linear and angular velocities which have been found by Covert[15] also to be crucial for the safe store separation envelop. Pitch and yaw rate discrepancies between theory and observations were explained by canard lag arising from canard shaft and actuator friction and canard inertia. Nevertheless, vertical motion histories of greatest interest to the pilot were in excellent agreement.

Alders[2] has described in-flight measurement of loads on captive stores and has compared them with data from other sources. Ramachandra[64] has given comparisons between predicted store trajectories with flight test motion pictures obtained from a chase aircraft for stores with and without fins. He has also compared the predicted and flight test data of the linear and angular accelerations of the stores along their trajectory. Although good agreement is obtained for the longitudinal component of the store motion, the downward motion appears to be over-predicted for power ejection and under-predicted for gravity release. Further, the store pitch angle displays a steady monotonic time variation compared to the oscillatory predicted

behaviour. Crombie[18] has described the development of a technique at the British Aircraft Corporation for identifying the coefficients of a missile system using telemetry data during a flight test. Unfortunately, satisfactory solutions were not obtained due to the noise introduced by the recording and processing of the telemetry data confusing the calculation of the performance index. Consequently, the optimized coefficients were not in agreement with the expected values. The general polynomial model used to determine the coefficients was unsatisfactory since the forcing function contained too little information at the low frequency end to permit their accurate determination in the presence of noise. Ramachandra *et al.* photographic method of store aerodynamic data estimation from flight motion pictures of separation trajectories has been reported[65,66].

A novel photo-data analysis system (PDAS) using a superposition of a flight test movie of store separation and the image of a scale model of the store through video-mixing has been reported by Cooper *et al.*[13] and Cooper *et al.*[14], so that aerodynamic influence coefficients can be obtained. They claim high precision data extraction from tests through the use of manifold roster points for image superposition unlike in conventional methods. Photogrammetric methods of analysis of the store separation trajectories has been developed at the Grumman Aircraft Corporation and have been described by Reilly[68].

7. Concluding Remarks

Despite the many difficulties and deficiencies encountered with the current experimental methods, especially those using the wind tunnels, they provide the most reliable data under controlled conditions. Telemetry needs to be used effectively during flight testing to generate information on the state variables and the principal control parameters of the store trajectories using instrumented stores coupled with photographic motion picture data taken from chase aircraft to correlate with theoretical studies. Increased attention to buffet and flutter flight test data during stores carriage is required to understand the aerodynamic influence coefficients of the aircraft-store combination. Some data were generated on this aspect during an aircraft-store integration program recently in India. But the results are of little scientific value due to deficiencies in test instrumentation and program organization.

On the theoretical side, besides studies on the flow field determination from the governing equations of motion, perhaps, Monte-Carlo techniques may profitably be used to obtain store trajectories taking into account the random variations of ejector cartridge thrust, rack friction, release angles and angular velocities, atmospheric turbulence, store geometry and inertia imperfections, fuel sloshing effects in partially filled drop tanks and many other small but intangible parameters in a probabilistic sense from release up to ground impact. A concerted effort would go a long way toward an effective and economical weapon system with appreciable beneficial effects on peaceful applications.

REFERENCES

1. Anonymous, "DATCOM—USAF Stability and Control Methods", (1968).
2. Alders, G. J. (1977) Proc. 4th JTCG Aircraft/Stores Compatibility Symp., Fort Walton Beach, Florida, Oct. 12-14.
3. Ashley, H. and Rodden, W. P. (1972) *Ann. Rev. Fluid Mechanics*, 4 : 431.
4. Asfar, K. R., Mook, D. T. and Nayfeh, A. H. (1978) AIAA Paper No. 78-1205, 11th AIAA Fluid and Plasma Dynamics Conf., Seattle, Washington, July 10-12.
5. Atta, E. H. and Nayfeh, A. H. (1978) AIAA Paper No. 78-1206, AIAA 11th Fluid and Plasma Dynamics Conf., Seattle, Washington, July 10-12.
6. Bera, R. K. (1969) "Wing-body interference of a family of wing-body configurations", M.Tech. Thesis, Dept. of Aero. Engg., Ind. Inst. of Technology, Kanpur, India, Aug.
7. Berry, J. R. (1970) AGARD CP-71-71, Sept.
8. Borland, C. J. (1966) USAFFDL-TR-66-37, Part I.
9. Chadwick, W. R. (1973) AIAA Paper No. 72-971, AIAA 2nd Atmospheric Flight Mechanics Conf., Palo Alto, California, Sept. 11-13.
10. Brooks, D. L. (1971) AIAA Paper No. 71-294, AIAA 6th Aerodynamic Testing Conf. Albuguergue, New Mexico, Mar. 10-12.
11. Chou, D. C. (1966) USAFFDL-TR-66-37, Part II.
12. Christophe, J. and Coste, J. (1970) AGARD Conf. Proc. No. 71 on Aerodynamic Interference, AGARD CP-71-71, AGARD Fluid Dynamics Panel Specialists Meeting, Silver Springs, Maryland, USA, Sept. 28-30.
13. Cooper, G. E. and Kingery, R. W. (1974) 5th Ann. Symp., Soc. Flight Test Engineers, Aug.
14. Cooper, G. E. and Gnagy, J. (1977) Proc. 4th JTCG Aircraft/Stores Compatibility Symp. Fort Walton Beach, Florida, Oct. 12-14.
15. Coste, J. (1977) Proc. 4th JTCG Aircraft/Stores Compatibility Symp. Fort Walton Beach Florida. Oct. 12-14.

16. Covert, E. (1973) Proc. 2nd JTCG Aircraft/Stores Compatibility Symp. Sacramento. Calif. Sept. 18-20.
17. Cox, G. A. and Carr, K. (1975) AGARD Conf. Proc. No. 187 on Flight/Ground Testing Facilities Correlation-Flight Mechanics Panel Symp. at Valloire/Modane. France, June 9-12.
18. Crombie D. (1976) "System using flight trials data". Proc. Conf. on Optimization in Action. Univ. of Bristol. Jan. 1975 Ed. L. C. W. Dixon., Academic Press.
19. Crose, Gorden G. and Bristow, Dean R. (1970) AGARD CP 71-71 Sept.
20. Daniels, P. and Clare, T. A. (1973) *J. Aircraft*, 10 : 511.
21. Dillenius, M. E. E., Goodwin, F. K., Nielsen, J. N. and Dver, Calvin. (1973) Proc. 2nd Aircraft/Stores Compatibility Symp., Sacramento. Calif. Sept. 18-20.
22. Everett, R. N. (1977) Proc. 4th JTCG Aircraft/Stores compatibility Symp. Fort Walton Beach Florida. Oct. 12-14.
23. Fernandes, F. D., NACA-CR-112065-1.
24. Fernandes, F. D. (1970) AGARD CP-71-71. Sept.
25. Foughner, J. T. and Bensin, C. T. (1977) Proc. 4th Aircraft/Stores Compatibility Symp. Fort Walton Beach. Florida., Oct. 12-14.
26. Furev, R. J. and Martin, C. J. (1970) AGARD CP-71-71. Sept.
27. Gardner, D. and Guest, A. L. (1977) Proc. 4th JTCG Aircraft/Stores Compatibility Symp. Fort Walton Beach. Florida., Oct. 12-14.
28. Giesing, J. P., Kalman, T. P. and Rodden, W. I. (1971) USAFFDL-TR-71-5, Parts I & II.
29. Goodwin. F. K. Dillenius. M. E. E. and Nielsen, J. N. (1974) AFFDL-TR-72-82, Oct.
30. Herron, R. D. (1973) Proc. 2nd JTCG Aircraft/Stores Compatibility Symp., Sacramento, Calif., Sept. 18-20.
31. Holloway, R. B., Sutcliffe, Donald L. and Woodward, James D. (1970) AGARD CP-71-71, Sept.
32. Holmberg, C. J. L. (1973) Proc. 2nd JTCG Aircraft/Stores Compatibility Symp., Sacramento, Calif., Sept. 18-20.
33. Kalivretenos, C. A., (1970) AGARD CP-71-71, Sept.
34. King, K. H. (1970) AGARD CP-71-71, Sept.
35. Labrujere, Th. E., (1972) ICAS Paper No. 72-49, 8th Congress of the ICAS, Amsterdam, Netherlands, Aug. 28-Sept. 2.
36. Landers, E. R., Simpson, G. J. and Bishop, G. (1975) W. R. E. TN 1522 (W R&D), Australia, Oct.
37. Larson, E. S., (1976) Proc. 10th Congress of the CAS, Oct. 4-8, Ottawa, Ontario, Canada.
38. Laschka, B. (1969) Pilot paper presented at 29th AGARD Structures and Materials Panel Meeting, Istanbul, Turkey, Sept. 28-Oct.8.
39. Laschka, B. Z. (1963) *Z. Fluggwiss*, 11 : 265.
40. Looney, L. N., Air Force Inst. of Tech. Rep.
41. Maddox, Arthur R., (1973) Proc. 2nd JTCG Aircraft/Stores Compatibility Symp., Sacramento, Calif., Sept. 18-20.

42. Maestri R. R. (1973) Proc. 2nd JTCG Aircraft/Stores Compatibility Symp., Sacramento, Calif., Sept. 18-20.
43. Marnix, F. E. D., Goodwin, F. K., and Nielsen, J. N. (1975) *J. Aircraft*, **12** (10) : 812.
44. Marsden, P., and Haines, A. B. (1967) "Aerodynamic loads on external stores : A review of experimental data and method of prediction", ARCR & M 3503.
45. Martin, Fred W., Saunders, Crady H. and Smith, Charles J. (1975) *J. Aircraft*, **12**, (3) : 151.
46. Martin, F. W. and Burkhalter, J. E., (1977) Proc. 4th JTCG Aircraft/Stores Compatibility Symp., Fort Walton Beach, Florida, Oct. 12-14.
47. McCabe, Earl F. Jr. and Furey, R. J. (1973) Proc. 2nd JTCG Aircraft/Stores Compatibility Symp. , Sacramento,Calif., Sept. 18-20.
48. McKinney, L. W. and Polhamus, E. C. (1966) NASA TN-D-3582, Nov.
49. Miles, J. W. (1959) "Potential theory of unsteady supersonic flow, Cambridge Univ. Press.
50. Morino, L. and Kuo, C. C. (1974) *AIAA J.* **12** : 191.
51. Morino, L. and Tseng, K. (1977) Proc. of CANCAM 77, Canada.
52. Morino, L. (1974) NASA CR-2464.
53. Morino, L., Chen, L. T. and Suciu, E. O. (1975), *AIAA J.*, **13** : 368.
54. Morino, Luigi (1978) Rep. No. ASI-TR-78-45, Aerospace Systems, Inc., Burlington, Mass, 0-803, USA.
55. Narayana, M. S., Aravamudan, M. D., Lokabhiram, K. and Kesavan, N.R. (1977) unpublished work, Hindustan Aeronautics Limited, Bangalore, India.
56. Nielsen, Jack, Goodwin, Fred K. and Dillenius, F. E. (1970) AGARD CP-71-71, Sept.
57. Heckham, D. H. (1977) Proc. 4th JTCG Aircraft/Stores Compatibility Symp., Fort Walton Beach. Florida, Oct. 12-14.
58. Pick, George S. and Dodd, Phillip G. (1973) Iroc. 2nd Aircraft/Stores Compatibility Symp., Sacramento, Calif., Sept. 18-20.
59. Pinsker, N. J. G. (1968) RAE TR 68249, Oct.
60. Pitts, William, Nielson, Jack N. and Kaatari, George E. (1957) NACA TR 1307.
61. Pollock, S. J. and Cooley, Dale E. (1977) Proc. 4th JTCG. Aircraft/Stores Compatibility Symp., Fort Walton Beach, Florida, Oct. 12-14.
62. Pugh, P. G., Peto, J. W., Ilott, G. P. (1977) "Experience of the use of the Nielsen Method of store trajectory prediction within aerodynamics department", Royal Aircraft Establishment, Farhborough, Sept. (private communication).
63. Ramachandra, M. (1977) Proc. 4th Aircraft/Stores Compatibility Symp., Fort Walton Beach, Florida, Oct. 12-14.
64. Ramachandra, S. M. (1978) *J. Aircraft*, **15** (10) also as NTIS Rep. No. N78-24114.
65. Ramachandra, S. M. Sudhakar, K., and Kiran, N. S. (1980) of Aircraft, May.
66. Ramachandra, S. M. and Gupta, S. C. (1971) Defence Project Rep. Dept. of Aero. Engg., Ind. Inst. of Technology, Kanpur, June.
67. Roberts, R. H. *et al.* (1974) Phase V, AEDC-TR-73-87, March.

68. Reilly, T. J. (1977) Proc. 4th JTCG Aircraft/Stores Compatibility Symp., Fort Walton. Beach, Florida, Oct. 12-14.
69. Schindel, L. H. (1975) AGARDOGRAPH No. 202, June.
70. Sekellick, M. (1970) AGARD CP-71-71, Sept.
71. Serbin, H. (1973) *J. Aircraft*, **10** : 123.
72. Short, C. R. and Lautermann, B. J. (1973) Paper 17, 2nd Aircraft/Stores Compatibility Symp., Sacramento, Calif., Sept. 18-20.
73. Siddalingappa, S. R. and Ramachandra, S.M. (1971) Defence Project Report, Dept. of Aero. Engg. Ind. Inst. of Technology, Kanpur, Aug.
74. Smith, C. J. (1975) "Mutual aerodynamic interference effects for two similar bodies," M.S. Thesis, Dept. of Aerospace Engg., Auburn Univ., Ala., Dec.
75. Smith, D. L. (1973) proc. 2nd Aircraft/Stores Compatibility Symp., Sacramento, Calif., Sept. 18-20.
76. Smith, R. E. (1973) Proc. 2nd Aircraft/Stores Compatibility Symp., Sacramento, Calif., Sept. 18-20.
77. Spahr, H. R. (1975) *J. Aircraft*, **12** (10) : 807.
78. Spahr, H. R. (1973) Proc. 2nd Aircraft/Store Compatibility Symp., Sacramento, Calif., Sept. 18-20.
79. Spearman, L. L. (1972) NASA TN D-6775, April.
80. Spearman, L. (1973) Proc. 2nd JTCG Aircaft/Stores Compatibility Symp., Sacramento, Calif., Sept. 18-20.
81. Spearman, L. (1977) Proc. 4th JTCG Aircraft/Stores Compatibility Symp., Fort Walton Beach. Florida, Oct. 12-14.
82. Spearman, L. M. and Monta, William J. (1977) NASA Tech. Memo. 74090, Dec.
83. Speer, Lt. T. E. (1977) Proc. 4th JTCG Aircraft/Stores Compatibility Symp., Fort Walton Beach, Florida, Oct. 12-14.
84. Stahara, S. S., Hemsch, M. J. Perkins, S. C. and Spreiter, J. R. (1977) Proc. 4th JTCG Aircraft/Stores Compatibility Symp., Fort Walton Beach, Florida, Oct. 12-14.
85. Stark, V. J. E. (1958) SAAB Aircraft Co. TN 41.
86. Sudhakar, K., Kiran, N. S. and Ramachandra, S. M. (1977) Paper presented at 29th Ann. Meeting, Aero. Soc. of India, Trivandrum, India, May 13-15.
87. Sullivan, M. B. (1975) AFFDL-TR-73-126.
88. Taisseire, R. (1970) AGARD CP-71-71, Sept.
89. Takashima, K. (1977) Proc. Int. Symp. on Space Technology, Tokyo.
90. Walkley, K. B. (1973) "Aerodynamic interference of wing-pylon-body combinations at low subsonic speed," M.S. Thesis, Dept. of Aerospace Engg., Auburn Univ., Auburn, Alabama, USA, June.

91. Welch, B. H. and Richter, B. J. (1977) Proc. 4th JTCG Aircraft/Stores Compatibility Symp., Fort Walton Beach, Florida, Oct. 12-14.

92. Watkins, C. E., Runyan, H. L. and Woolston, D. S. (1955) NACA TR 1234.

93. Watkins, C. E., Woolston, D. S. and Cunningham, H. J., (1959) NASA TR-R48.

94. Whoric, J. M. (1974) AEDC Rep. TR-74-75, AFATL-TR-74-113, Oct.

95. Woodward, F. A. and Hague, D. S. (1969) General Dynamics Res. Rep. ERR-F-W-867 Fort Worth, Texas.

96. Woodward, F. A., Tenoco, E. N. and Larsen, J. W. (1967) NASA CR-73106, (Part I).

Algal Hydrobiology In India

Braj Nandan Prasad and Yashpal Singh

Department of Botany, Lucknow University, Lucknow

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

Introduction

The phenomenal increase in human population has stimulated an all-round exploitation of natural resources, including the freshwater aquatic resources of the globe. Increased eutrophication of lakes and streams has, however, resulted in problems of taste, odour, toxicity and obnoxious growths caused by algae. In order to understand these problems, it is necessary to have an adequate basic knowledge of algal taxonomy, physiology, cytogenetics, hydrobiology and ecology[24] with special reference to pollution and consequent shifts in the aquatic biodynamics.

While considerable work has been done in temperate algal ecology, studies of a similar nature of Indian bodies of water are still in the cradle. The economics of algal ecology is brought to the fore in problems involving the spoilage of water needed for domestic and industrial utilisation.

The aquatic environment in fresh waters is broadly divisible into lotic (running) and lentic (stationary) categories. On account of considerable differences in the two types of ecosystems, it would be convenient to consider them separately.

I. Lotic Hydrobiology

The hydrobiology of lotic environments in India has been studied, only in respect of some rivers, e.g., Adyar[7], Cooum[38], Chambal[42], Cauvery[77], Hooghly[62], Ganga[1,36,48], Yamuna[9,51], Jhelum[80], Mahanadi[46], Malampuzha[8], Moosi[18,82,83,84,85,86], Sone[39,46], the river Varuna[45], and Gomati[49,78]. Studies on the benthic algae of lotic bodies have been negligible in India and the only notable contributions in this respect are by Venkateswarlu[84] and Singh[78].

The algae are primary producers and occupy the base level in the transfer of energy within natural aquatic systems. The distribution and occurrence of algae in both time and space are related to certain fundamental factors which may, in general, be divided into three categories : physical, chemical and biological. The physical factors most clearly effective in determining algal distribution are temperature and light while the chemical factors include the pH and nutrients including trace elements.

(i) *Physical Factors*

(a) *Temperature*

The temperature regimes of freshwater lotic environments have been investigated by Roy[62], Chakrabarty *et al.*[9] and Venkateswarlu[84]. Venkateswarlu[84] states that high temperature, coupled with maximum sunshine, favours rapid liberation of CO_2 from the $CaHCO_3$ in water. Singh[78] observed that the water temperature in the river Gomati at Lucknow varied from 16 °C—31 °C and sewage effluents do not significantly alter the water temperature characteristics of the water.

(b) *Light*

The amount of incident light penetrating the water is a critical factor and determines the depth limit of growth of certain algae while turbulence controls their growth by effecting the mixing and dispersal of nutrients in water. Singh[78] has observed that a sewage-contaminated stations, water-transparency values are considerably reduced as compared to uncontaminated stations and that the values of water-transparency are maximum in summers and lowest in monsoons while winters are characterised by intermediate values. The water-transparency ranges from 3·81 cm to 35·56 cm[78].

(ii) *Chemical Factors*

(a) *Hydrogen ion concentration*

Venkateswarlu[82] has found that pH is highest in May in Moosi river. Singh[78] has observed the pH of the waters of the river Gomati as ranging from 7-8·7 with maximum values in late winters and early summers and minimum values in monsoons and contamination does not seem to cause a statistically significant difference in pH. Singh[78] has reported the pH to exhibit a positive relationship with bicarbonate, total alkalinity, calcium, magnesium and total hardness. Such relationships have not been established so far in other freshwater lotic habitats of India.

(b) *Alkalinities*

The concentrations of carbonates, bicarbonates, total alkalinity, total hardness, calcium and magnesium in natural and eutrophic waters and their effects on the diversity and quantity of algae have been studied[39,46,78]. Philipose[46] has divided the inland waters of India into three categories based on the values of total alkalinity. Total alkalinity values of more than 60 ppm

indicate the Jhelum river water to be of the hard type[80]. In Gomati, carbonates vary from 0 ppm to 43·3 ppm with the maximum values in winter and late summer and low values in early summer and during the monsoon[78]. The bicarbonate alkalinity varied from 34·16 ppm to 405·04 ppm with maximum values in late winter and early summer and minimum values in monsoon[78]. Sewage-contaminated stations exhibited a significantly higher concentration of bicarbonates[78]. Philipose[46] has recorded higher total alkalinities from the polluted portions of the river Sone in Bihar.

The studies of Singh[78] demonstrate that pH, carbonates, bicarbonates, total alkalinities, excess alkalinity, calcium, magnesium and total hardness form a set of positively related factors.

(c) *Chlorides*

Most workers[1, 9, 36, 43, 45, 51, 73, 82] who have studied the chloride content in river waters record an increase in chlorides with sewage contamination. Indeed Rai[51] believes that excess chlorides in the river Yamuna can be the best indicators of the extent of pollution. Venkateswarlu[82] noted that unpolluted and polluted stations in the river Moosi exhibited an average of 10·63 and 40·75 ppm of chlorides respectively. Agarwal *et al.*[1] have recorded the chloride content at sewage-polluted zones in the Ganges as ranging from 28 ppm to 54 ppm with an average of 45·7 ppm. Singh[78] has also found that contaminated stations have significantly more chlorides. Chlorides are highest during the period December to May and lowest in the monsoon. A summer maximum and a monsoon and winter minimum has been noted[9, 36, 43, 82] and the maximum values in winter and early summer have been attributed to reduced flow[78] and chlorides have been found to be positively correlated with free and saline ammonia and albuminoid ammonia and negatively correlated with nitrate nitrogen[78].

(d) *Dissolved Oxygen*

River water shows seasonal variation in dissolved oxygen content[9, 39, 43, 46, 51, 62, 82] and it has been stated that very low values of dissolved oxygen are usually associated with high organic concentrations (pollution)[73, 82].

Singh[78] has observed dissolved oxygen to exhibit a general periodicity in which the values of dissolved oxygen are lower in the late summer and highest in monsoon. The dissolved oxygen content was negatively correlated to nitrate nitrogen. The same conclusion has been reached by Agarwal *et al.*

(e) *Nitrogen Complex*

Notable studies have been made dealing exclusively with the nitrogen relations inclusive of the ammoniacal and nitrate nitrogen in rivers[1, 39, 46, 62, 82]. Most workers have noted a decrease in nitrate going hand in hand with pollution although Agarwal *et al.*[1] have observed nitrates to increase with pollution. In the polluted waters, the processes of denitrification and decomposition of nitrogenous organic matter are accelerated at higher temperatures, resulting in high yield of free ammonia which in spite of lowered solubility at higher temperatures, tends to accumulate in such waters[82]. An increase in free and saline ammonia contents in waters due to the addition of sewage effluents[78] and a summer maximum for free and saline ammonia[78, 82] have been observed. A direct relationship between free and saline ammonia and albuminoid ammonia has been noted[1, 73, 82]. Singh[78] reports that, at the uncontaminated stations, free and saline ammonia values shows a negative relation to the chloride values. Their content in the river Gomati[78] varied from 0 to 10·08 ppm.

Albuminoid ammonia has been found to increase progressively with pollution[1, 78, 82]. Free and saline ammonia are also positively related to total alkalinity, bicarbonates, excess alkalinity, total hardness and magnesium[78]. Albuminoid ammonia is positively correlated with carbonate, pH, excess alkalinity and calcium. A summer maximum for albuminoid ammonia in the Ganga is reported[1] but in the Gomati, except for a monsoon low, the albuminoid ammonia content does not exhibit any fixed periodicity[78].

The nitrate content of polluted waters do not seem to differ significantly from that of unpolluted waters[78]. An inverse relationship between nitrates and dissolved oxygen has been suggested[1, 78]. A negative correlation has also been noted between nitrates and chlorides[78, 82] and between nitrates and carbonates[78].

(f) *Calcium-Magnesium*

Singh[78] has observed sewage-contaminated stations to exhibit higher total hardness values. The total hardness of the Gomati waters ranged from 80 to 246 ppm and was highest in late winter and late summer with minimum values in the monsoon. In this study, the calcium ranged from 14·42 to 49·6 ppm with the highest values at the sewage-contaminated stations. Magnesium was found to be directly related to sewage contamination and ranged from 8·67 ppm to 51·3 ppm. The lowest values were always

attained in the monsoon and free and saline ammonia was positively related to total hardness and magnesium, albuminoid ammonia to calcium. Dissolved oxygen shows a negative correlation with total hardness and calcium. Chlorides are positively correlated with magnesium.

(iii) *Algal Periodicity*

The environmental factors affecting the waters are generally cyclic and hence, algae exhibit a known degree of constancy in their periodicity and succession, subject to variability of the factors. Some workers have suggested that total algae reach their maximum development during summer and minimum during winter,[9,46] others believe that the density of total algae is maximum during post-monsoon, winter and early summer and is lowest in late summer and monsoons[36,43,62,73,84]. The expression of total algae at sewage-contaminated stations did not differ quantitatively from the total algae at the uncontaminated stations. A high pH has been associated with the phytoplankton and benthic algal maxima[62,83] but Singh[73] states that pH, hardness, nitrates and free and saline ammonia had little effect on the numbers of attached algae. In this study it was seen that the algal periodicity varies as the periodicity of diatoms and total alkalinity.

Most workers who have studied the periodicity and distribution of Cyanophyceae in Indian lotic bodies[9,43,46,49,71,73,84] associate their intensity with high temperature, abundance of organic matter and low dissolved oxygen and hold that the members of this group can serve as indicators of pollution. However, Singh[73] did not find any significant effect of water temperature on the expression of Cyanophyceae and believes that a low dissolved oxygen content in waters contaminated by sewage may be favourable for the growth of Cyanophyceae and Venkateswarlu[84] attaches more importance to the organic matter content and dissolved oxygen in water. pH does not seem to have a significant relationship with the percentage of Cyanophyceae[73]. The studies of Singh[73] reveal some important relationships of the Cyanophyceae—the group is negatively correlated with total hardness, diatoms and Euglenophyceae and it appears that Cyanophyceae which are richest in the sewage-contaminated stations of the river Gomati are supported more by low transparency, high calcium, high free and saline ammonia and high albuminoid ammonia. Venkateswarlu[84] observed a high percentage of Cyanophyceae at stations possessing higher

values of albuminoid ammonia. In the river Gomati, the dominance of either *Oscillatoria* spp. or *Spirulina* spp. may serve to indicate a certain degree of sewage contamination.

The periodicity of Chlorophyceae in relation to seasons and their qualitative and quantitative distribution have been studied[9,62,73,83,84]. Venkateswarlu[83] and Singh[73] have observed that Chlorophyceae are favoured by higher concentrations of dissolved oxygen and low organic matter. Both these authors[73,85] have also observed sterile *Spirogyra* spp. in lotic waters. High atmospheric or water temperature coupled with bright sunshine is considered to be important in the periodicity of Chlorococcales[84]. Chakrabarty *et al.*[9] noted that *Pediastrum* is very common in the river Yamuna and occurs throughout the year, but is more abundant from April to July. Roy[62] has found that during the summer months Chlorococcales and Conjugales dominated in the river Hooghly. Venkateswarlu[84] states that low oxidizable organic matter coupled with higher temperatures favours the growth of Chlorococcales and Desmids, but if the oxidizable organic matter is high, the Cyanophyceae attain dominance. He has also observed the paucity of Chlorococcales to be related to a toxicity due to higher concentrations of nitrates and low dissolved oxygen at the polluted stations. The unpolluted stations exhibit more Desmids, showing a summer peak and a direct relationship with total solids. Singh[73] found Chlorophyceae to be represented mainly by species of the genera *Scenedesmus*, *Hydrodictyon*, *Uronema*, *Stigeoclonium*, *Oedogonium*, *Bulbochaete*, *Cladophora*, *Rhizoclonium*, *Spirogyra* and *Cosmarium*. The class was best represented in the late winter and summer months and lowest in the monsoon with continuing low values in early winter. The periodicity of the Chlorophyceae exhibits correlative trends with the periodicity of albuminoid ammonia, carbonates and magnesium and moves inversely to calcium and nitrate nitrogen content.

Some of the important Indian findings on the ecology of lotic Euglenineae are of Chakrabarty *et al.*[9], Venkateswarlu[84] and Singh[73]. Members of this class are scarce in fast-flowing waters[84]. In the Gomati for instance Euglenophyceae are rarely seen except where water stagnates due to reduced flow[73]. Chakrabarty *et al.*[9] place more emphasis on the temperature, organic matter and albuminoid ammonia as influencing the growth and development of the Euglenineae. In the river Moosi[84] this class is mainly represented by species of *Euglena*, *Phacus* and *Lepocinclis* which are abundant when water temperature ranges from 28°C to 31·5°C. Venkateswarlu[84] and Singh[73]

report very high summer temperatures to be unfavourable for the growth of this class of algae, which is inversely correlated with pH and directly with total iron ; other factors may also be significant. Sewage contamination, characterised by low transparency, low dissolved oxygen and high bicarbonates, total alkalinity, chlorides, total hardness, calcium, magnesium, free and saline ammonia and albuminoid ammonia are favourable to the growth of Euglenophyceae[73].

The seasonal distribution and environmental requirements of freshwater diatoms of lotic habitats have been studied by a few workers[9,23,36,48,46,62,73,84,85]. Benthic diatoms have been discussed by Venkateswarlu[83,84] and Singh[73]. The diatoms in general reach their maximum during the winter months, are low during summers and minimum during monsoon[36,62,73,84] but in the river Mahanadi, Chlorophyceae and Bacillariophyceae attain their maximum members in summers and minimum during monsoon[46]. Pahawa and Mehrotra[43] have observed two peaks in the river Ganga—one in summer and the other in winter, with the summer peak higher than the winter. Most workers have found diatoms to constitute the major bulk of phytoplankton. Water temperature plays a very important role in the determination of diatoms periodicity—low temperatures being associated with high diatom numbers and vice versa[62,73,84]. Apparently low temperature allows nitrates to accumulate in water and these stimulate abundant growth of diatoms in winters. With a rise in temperature, the process of breakdown of nitrates is accelerated leading to their depletion in water. This accounts for the decrease in diatom numbers[62].

The silica content of river water and its relationship with the diatoms has been studied by Iyengar and Venkataraman[23], Roy[62], Chakrabarty *et al.*[9] and Venkateswarlu[84].

Roy[62] has observed that the highest range of silica concentration is associated with both the maxima and minima of phytoplankton mainly composed of diatoms in the river Hooghly whereas Venkateswarlu[84] found that silicates and diatom populations are more or less inversely proportionate in the summer months at all the stations although, at other periods of the year, no relationship could be established. Chakrabarty *et al.*[9] and Venkateswarlu[84] have observed the maximum amount of silica in June.

It has been noted that an increase in nitrates helps in the growth of diatoms[84] but Singh[73] concludes that in the subtropical river Gomati, nitrates have little importance in determining diatom population.

In the river Hooghly, the phytoplankton, chiefly composed of diatoms, is inversely related to the phosphate content of the waters[62].

Dissolved oxygen and diatom populations go hand in hand[62,84] and apparently a higher concentration of dissolved oxygen facilitates the formation of more nitrates and there is a possibility that dissolved oxygen nitrates and the diatom populations are linked together in an ecological complex[84]. Singh,[73] however, contends that dissolved oxygen alone is not operative in bringing about a change in the population of diatoms, but it is the organic matter also which determines the distribution of diatoms and a low organic matter content always favours the growth of this class of algae at optimum temperatures. He has observed that the Cyanophyceae and the Bacillariophyceae are inversely correlated within the benthic algae of the river Gomati. Venkateswarlu[84] has also sampled the phytoplankton of the river Moosi and states that the plankton showed similarities in general both quantitatively and qualitatively to the bottom flora. This is in conformity with the concept that the benthos projects itself into the plankton.

(iv) *Pollution*

It is now well recognised that the most important use to which algae can be put lies in pollution studies. Nevertheless studies on the pollution ecology of freshwater lotic algae in India have been rather meagre. In general, biological pollution results in an increase in total solids, turbidity, pH, carbonates, bicarbonates, total hardness, calcium, magnesium, chlorides, phosphates and ammoniacal nitrogen and a decrease in dissolved oxygen and transparency. Palmer[44] attributed the decrease of calcium in oligotrophic waters to the utilisation of CO_2 through photosynthesis and the conversion thereby of soluble calcium bicarbonate into insoluble calcium carbonate. In eutrophic waters, bacterial action causes the opposite effect, bringing about on increase in calcium level. Rai[61] stresses that excess of chloride in the river Yamuna can be the best indicator of the extent of pollution. Most of the workers have noted a decrease in nitrate with pollution although, Agarwal *et al.*[1] have observed nitrates to increase with pollution. Singh[73] has found no significant variation in nitrates between polluted and unpolluted stations. It is known that ammoniacal nitrogen can be an excellent indicator of water-pollution by sewage. Proteinaceous matter present in human faeces is finally degraded and nitrogen is converted into ammonia by bacterial action. Ammonia is then oxidised to nitrates.

Venkateswarlu[84] suggests that the process of denitrification and decomposition of nitrogenous organic matters are accelerated at higher temperatures resulting in high yields of free ammonia which, in spite of lowered solubility at higher temperatures, tends to accumulate in such waters. Jepson and Greene (cf. Klein[89]) state that nitrates are important in waters deficient in oxygen as anaerobic decomposition is prevented by bacteria making use of this oxygen source. He has also reported instances where nitrates were deliberately added to rivers to assist in the removal of odours in standing and sluggishly moving waters.

The most important effect of organic pollution on a stream is that of enrichment and fertilisation ; the total number of algal species is generally reduced but those which grow increase massively with a corresponding increase in total production[6]. The importance of a biological survey in monitoring water quality is increased by the fact that the qualitative and quantitative composition of aquatic population is determined by recurrent critical conditions even though of short duration. It has also been emphasised that the physicochemical damage to a body of running water is an ephemeral thing. The complex of organisms which grows in a given area is indicative of environmental conditions which have been created during its development.

The phytoplankton community, although important in lentic water, seems to be of little use in monitoring water quality in flowing waters. Round[61] states that the epipelic and epiphytic algae may form excellent indicators of water-pollution. However, the mere absence or presence of a species may not be indicative of environmental conditions. Unless long-term quantitative studies on the comparative species composition of polluted and unpolluted zones is attempted, no concrete conclusions can be drawn to identify biological indicators of pollution.

Studying the periodicity of certain species of algae most commonly met in the river Moosi, Venkateswarlu[85] has established four categories based on the relationship of algae to (1) temperature, oxidizable organic matter and dissolved oxygen (2) current (3) nitrates and (4) total iron. It appears that the dominance of *Oscillatoria chalybea*, *Stigeoclonium tenue*, *Nitzschia palea*, *Navicula pygmaea*, *Pinnularia biceps* var. *amphicephala*, *Cyclotella meneghiniana* and *Achnanthes exigua* may indicate sewage-contaminated waters ; *Gomphonema sphaerophorum* is characteristic of mildly-polluted waters and an abundance of *Achnanthes minutissima* indicates an abundance

of oxygen. Venkateswarlu[86] has later discussed the ecology of 57 taxa of algae in the polluted and unpolluted regions of this river and correlated them with factors most closely associated to pollution.

In subtropical and tropical environments, blue-green algae form a major component of the aquatic flora and it has been emphasised that one may successfully find valuable indicators of pollution within this component. Profuse blue-green algal growth during periods of high temperatures, high organic matter and reduced water flow is indicative of high organic pollution.

Studying the influence of industrial effluents on some lotic algae of the river Gomati, Saxena[65] has recognised five classes of algae on the basis of their tolerance to pollution. Amongst the most tolerant algae she places eleven species of *Oscillatoria*, three of *Phormidium* and two of *Lyngbya*, five species of *Oscillatoria*, one of *Phormidium*, and two of *Lyngbya* are regarded tolerant to medium levels of pollution, and a single species each of *Gloeothece*, *Aphanocapsa*, *Anabaena*, *Scytonema* and *Gloeotrichia*, two each of *Cylindrospermum*, *Oscillatoria* and *Calothrix* and three species of *Lyngbya* have been found distinctive of unpolluted water. From Saxena's[65] observations it appears that the generic variety of blue-green algae is richer in the unpolluted zone where the total number of genera are 9 as compared to the polluted zone where only 3 genera were represented. It, thus, appears reiterated that a reduction in generic variety and an increase in the number of taxa of *Oscillatoria* can indicate high pollution. Singh[78] also found the dominance of *Oscillatoria* spp. or *Spirulina* spp. to indicate pollutants of a biological origin.

Diatoms also form an important constituent of the algal flora. Singh[78] has tabulated the ecological information on various diatom taxa with respect to the quantitative distribution, seasonal distribution, pH status, halobion status and saprobien status in the Gomati river. This is the only study in India which has been computer-processed with the Shannon Index of Species Diversity analysed. All the taxa were apparently alkaliphilous, oligohalobous, rheobiontic, mesothermal and periphytic. He has also concluded that dominance of *Amphora veneta*, *Cyclotella meneghiniana*, *Gomphonema parvulum*, *Melosira crenulata f. tenuis*, *M. granulata*, *Navicula accomoda* var. *accomoda*, *N. cryptocephala* var. *cryptocephala*, *N. pseudocuspidata* var. *rostrata*, *N. rhynchocephala* var. *rhynchocephala*, *Nitzschia communis* var. *abbreviata*, *N. diducta*, *N. palea*, *N. palea* var. *debilis*, *N. palea* var. *fonticola*, *N. stagnorum*, *N. subtilis*, *N. thermalis*, *N. thermalis* var. *intermedia* and *N. thermalis*

var. *minor* is indicative of sewage polluted zones. *Cocconeis placentula, Cymbella affinis, Gomphonema micropus, G. sphaerophorum, Synedra amphicephala* var. *amphicephala, S. ulna* var. *ulna, Navicula heufleri* var. *leptocephala* and *Nitzschia amphibia* indicate waters which are cleaner and characterized by low organic matter and high dissolved oxygen. It has been established that the taxonomic variety of algal associations decreases with stress due to the entry of pollutants. The only study so far on the generic variety of benthic algae is that on diatoms by Prasad and Singh[50] who state that the generic variety of diatoms decreases in polluted zones and that an increase in the number of taxa of the genus *Nitzschia* coupled with a decrease in generic variety indicates organic contamination.

Singh[78] has calculated for the first time in India, within a lotic subtropical river (Gomati) the ratios between the total number of species and importance values of individuals—usually referred to as the 'Species Diversity Indices'. The 'Shannon Index' of Species Diversity ($\bar{H}$) at the various stations ranged from 0·637 to 3·121, when $\bar{H}$ was correlated with the different physicochemical parameters it did not reveal marked differences between the 'species diversity indices' at the different stations and the index was seen to exhibit no fixed periodicity. Although the structure of the diatom community remained unaltered throughout his period of study, the specificity of taxa varied. The worker also concluded that a profuse growth of Cyanophyceae at the polluted stations results in a fall in the number of Bacillariophyceae which exhibit an erratic behaviour. It has been suggested that the qualitative changes in the diatom community can be a more reliable indicator of water quality and that the 'species diversity indices' of benthic diatoms cannot serve as indicators of pollution.

II. Lentic Hydrobiology

(i) *Physical Factors*

(a) *Temperature*

The thermal patterns of a few Indian lakes, reservoirs, tanks and ponds have been described[4,14,19,25,41,57,68,70,76,90,92,94].

Singh [68] found the water temperature in a small lake in Delhi to range from 17-31 °C and closely parallel to the air temperature. Bharati and Kore[4] contend that the water temperature followed the atmospheric temperature only during day time and from 6 p.m. to 3 a.m. it showed a steady

difference with a total shift of 7 °C. The Husain Sagar lake [99] tends to get thermally stratified in March and shows a well-marked stratification in April and May, a thermocline does not form in this lake and its annual heat budget is very high as compared to the head budgets of temperate and arctic lakes. Ganapati [14] has recorded an almost identical summer stagnation for Stanley reservoir. Jana [25] has noted maximum water temperature in summers and minimum in late winters. Gonzalves and Joshi [19], Zafar [90] and Rao[57] have noted a rise in water temperatures towards the end of monsoon and, in the studies of Rao[57], this post-monsoon rise in water temperature even exceeded the water temperature of the preceeding summers. Jana[25] and Sreenivasan[76] are of the opinion that water temperature is never critical in the seasonal periodicity of phytoplankton and by itself it is not a primary factor controlling blue-green algal populations.

Zafar[94] states that waters bearing thick algal populations or high salt content take a much longer time to get heated.

(ii) *Chemical Factors*

(a) *pH*

Philipose[46] has classified the inland waters of India into three categories and states that most of the small, confined waters used for rearing fish possess alkaline waters with the pH 7·4 to 8·4. These waters are fairly rich in nitrates and phosphates with their ratios variable. Inland waters exhibiting acidic characteristics have low or moderate concentrations of nitrates and phosphates with nitrate/phosphate ratio usually higher than unity. Dissolved CO_2 and carbonates have a profound influence on the pH of water[19, 25, 56, 58, 90]. Free CO_2 varies inversely to the pH. There is a direct relationship between carbonates and pH[17, 27]. In the Ramgarh and Suraha Tals,[70] a high alkalinity is always maintained (pH 8 to 10) and undergoes relatively little change from season to season. Alkalinity does not appear to be affected by changes in either oxygen or carbon dioxide content of the waters. Conclusions regarding the influence of pH on the total solid contents differ, while Ganapati[12] found highly alkaline waters to be rich in total solids, Zafar[90] noted that the pH and dissolved solids exhibit an inverse relationship. Rao[55] found that the alkaline ponds are invariably richer than the acidic ones in dissolved solids but highly alkaline ponds

(pH 8·4) or the just alkaline ones (pH 7·5) do not support a very high salt concentration[58]. This has been attributed to the rapid conversion of salts in the highly alkaline lakes into insoluble carbonates by the thick algal population through rapid withdrawal of carbon dioxide. The pH of water changes with changes in other climatological and vegetational factors[39]. In spring, with very little vegetation and low temperature the pH reached 6·9, in the summers it reached a maximum 8·7, hand in hand with the vegetation. pH rises in periods of active growth of *Anabaenopsis raciborskii*[67] but its values are not indicative of controlling blue-green algal environment[76]. In the Yemmekeri Pond at Dharwar, pH ranged from 6·6 to 8·2 with the maximum values in January and minimum in February[3]. Jana[25] has found pond water pH to range from 6·5–8·2, the water was slightly alkaline from summer to early autumn and acidic in late autumn and winter.

(b) *Alkalinities*

Zafar[90] has noticed very low free CO_2 concentrations in the ponds studied by him. The values are always lower than 0·5 mg/L. Bharati and Hosmani[3] state that in a Dharwar pond, the values of free CO_2 were constant during the first three months and reached a maximum of 322·6 mg/L in the month of January and again came down to 93·668 mg/L in February. Singh[70] records the free CO_2 content in the Ramgarh Tal to range from 90·68 ppm to 213·90 ppm with minimum values in March, April and maximum in July and December–January. The value remains high throughout but the December peak is higher than the July peak while in the Suraha Tal, the July peak is more pronounced. Zafar[90,92] states that the basic buffer system in small freshwater bodies revolves round the inorganic carbon in different forms and calcium and that carbonates and pH move parallel to each other and are inversely proportional to the bicarbonate and calcium contents of the waters. The water tends to be more alkaline when it possesses carbonates and much less alkaline when it supports larger quantities of bicarbonates, carbon dioxide and calcium. In winters, the distribution of calcium and bicarbonates is practically uniform in all strata of the Husain Sagar lake and, as the summer approaches, these factors gradually decrease at all the levels whereas carbonates increase appreciably. This is probably due to the escape of carbon dioxide from water at higher temperatures which results in the precipitation of insoluble calcium carbonate[92]. It has been observed[58,92] that the free CO_2 content, bicarbonates and calcium move hand in

hand. The pH of the water appears to be dependent upon the relative quantities of calcium, carbonates, bicarbonates and CO_2. Singh[68] and Jana[25] have observed the bicarbonates to be at a minimum during the rainy season. Singh[68] has observed a maximum in October. Seenayya and Subbaraju[67] have noticed that in periods of active multiplication of *Anabaenopsis raciborskii*, bicarbonates were increasing but when the bloom attained a climax, the bicarbonates showed a distinct decline and the water was rendered more alkaline. It is generally considered that blue-green algal biotopes have a high methyl-orange alkalinity although majority of areas investigated by Sreenivasan[75] demonstrated an inverse relationship between the blue-green algae and methyl-orange alkalinity. Tropical waters exhibit usually high values for bicarbonate alkalinity[25,67,68,70,76]. Rao[58] feels that the concentration of bicarbonates is dependent on the free CO_2. The total hardness content of waters shows an irregular pattern with a minimum during monsoons[68,70] or in the autumn[25] a maximum during early summer to mid-summer[25,68] or in the winters[70]. Singh[70] considers waters with a total hardness value of 55-92 ppm as soft and 36-56 ppm as moderately soft.

(c) *Silica*

The silica content of water of the Ramgarh and Suraha Tal varied between 20 and 30 ppm[70], in a pond at Delhi the silica content varied from 0·03 to 0·09 mg/L,[68] and in a pond in West Bengal from 6·0 to 16·0 ppm. The silica content has been found to exhibit a minimum value in April[70] and September[25] and a maximum in March[70] or February, May and November[68] or in early April[25]. It thus appears that the silica content of natural or artificial lentic waters in India has an irregular periodicity with possibly a late winter early-summer maximum.

(d) *Chlorides*

The chloride content of waters seems to be a direct result of pollution of animal origin[70,90]. It shows an increase in summer[10,25,68] and minimum in winter[68]. The summer increase is due to reduced water levels[19]. Zafar[90] and Rao[58] have observed a direct inter-dependence of sodium and chloride, both being complementary to each other. Sarkar and Rai[64] have noted that in Suraj Kund (Delhi) alkaline waters with appreciable quantities of carbonates have a high content of chlorides.

Rao[68] has studied the seasonal periodicity in the concentrations of Na,

K, Mg, Ca, CO_3, SO_4 and halides (all bicarbonates being converted to carbonates) in freshwater ponds of Hyderabad and it is seen to vary directly with air temperature and to decrease invariably after heavy rains. In the Nacharam pond, Rao[58] has observed an increase in the salinity after early rains and attributes it to the inflow of salts from the neighbouring areas. Gonzalves and Joshi[19] and Rao[58] have observed a reduction in salinity of the water during thick growths of algae, probably because of active absorption of salts.

(e) *Iron*

The quantity of iron in natural waters depends to a great extent on the pH of the water and the behaviour of this element seems to be similar to that of calcium[58,90,92]. Rao[58] found that where the bicarbonates increased independently of calcium, either sodium or iron exhibited a considerable rise in concentration and it is possible that above certain levels of concentrations of calcium, sodium and bicarbonate, the behaviour of iron is relatively indifferent to fluctuations of bicarbonate and dependent on pH. Singh[68] has observed that the total iron content of water in a lake at Delhi increased from April and reached a solitary maximum during the close of monsoon.

(f) *Phosphates*

Phosphates remain low during the winter and early summer months and rise during the rainy season[68,70,92]. Singh[68] has observed a maximum in May and June but Jana[35] obtained maximum values during winter and early summers (0·25 to 0·55 ppm), low values (0·25 ppm) in the remaining months. In a pond near the Golkonda fort, Hyderabad, Seenayya and Subbaraju[67] observed phosphates to be high just prior to the development of a bloom of *Anabaena raciborskii*; their concentration decreased with the multiplication of the alga and completely exhausted when the bloom attained a climax. Zafar[92] has followed the pattern of changes in phosphate concentration in the Husain Sagar lake and states that phosphates occur in traces in winters in all strata, start appearing in February in the bottom strata, increase in March and culminate in a tendency to accumulate in the subsurface strata. Phosphates could not be detected in May. The setting in of the monsoon sees an increase in the phosphate concentration but this increase is less than the increase in summers probably because phosphates are not easily leached from soil by rainwater. Ganapati[14] is of the opinion

that in tropical waters, phosphates are always present in sufficient quantities and they do not seem to constitute a limiting factor.

(g) *Nitrogen Factors*

Ganapati[13,14] has pointed out that tropical waters, particularly unpolluted ones, are poor in nitrates and attributes it to the utilization of the ammoniacal nitrogen as soon as it is formed by the dense growths of algae. This does not leave sufficient free ammonia for its conversion to nitrites or nitrates. Zafar[90] has also observed oxygen deficient tropical habitats to be poorer in nitrates although later he[92] emphasised that, as a freshwater body is not a self-contained unit, one can come across freshwater bodies even in the tropics which are considerably rich in nitrates, depending upon the nature of the catchment area. It has been suggested that the factor which apparently control the fluctuations of nitrates are temperature and rainfall. The monsoon maxima of nitrates have been explained on the assumptions that suspended particles might have accelerated the nitrification by providing a large surface area for bacterial activity[19,58,70,90]. Singh[70] states that the Ramgarh Tal containing 0·088-1·42 ppm of nitrates is richer than the Suraha Tal containing 0·008-0·885 ppm. The main source of nitrate is the drainage water and the nitrates are richest during and immediately after periods of flood and rainfall. Zafar[90] has observed that the nitrates are maximum during the winters and that the nitrates and albuminoid ammonia go hand in hand. In a lake at Delhi, the nitrate content varied from 0·07 to 1·63 ppm[68]. The minimum value was obtained during mid-summer. Seenayya and Subbaraju[67] state that 12·6 ppm nitrates were recorded from a Hyderabad Pond and the nitrates do not exhibit any fluctuation in relation to the bloom of *Anabaena raciborskii*. They increased in waters after the bloom had declined. Bharati and Hosmani[8] have observed highest values for nitrate nitrogen during December and February (886 mg/L) and lowest during November (55·375 mg/L). An inverse relationship has been established between the nitrites and nitrates[68,70].

Tropical bodies of water exhibit low values for nitrites[8,25,67,68,70]. Bharati and Hosmani[8] state that the nitrites in the Yemmekeri Pond show a gradual decrease from 0·1548 mg/L in September until they completely disappear in January and again reappear in February (0·01857 mg/L). Seenayya and Subbaraju[67] have observed 0·0009 ppm nitrite in the waters

of a pond at Hyderabad. No relationship was observed with the fluctuations in the population of *Anabaena raciborskii* although the nitrites declined after the decline of this bloom. Zafar[90] regards nitrites to be positively correlated to free ammonia and albuminoid ammonia and that free ammonia is higher at polluted stations. Kundra *et al.*[35] have also observed free ammonia content to be more at the polluted stations and believe that an increase in the concentration of free ammonia reveals pollution. The free ammonia content exhibits highest values in the rainy season[68,70]. Ganapati[13], Zafar[90] and Rao[58] report free ammonia and nitrates to be inversely related. Rao[56] and Zafar[90] have observed that free ammonia and albuminoid ammonia to go hand in hand.

It is well established that albuminoid nitrogen is one of the best indicators of nitrogeneous organic matter. In a lake in Delhi, the albuminoid nitrogen was maximum in May, when nitrates and nitrites were at the lowest and minimum in September when the ammoniacal nitrogen was highest[68]. Singh[70] has also observed an inverse relationship between albuminoid nitrogen and nitrates. Zafar[90] and Hosmani and Bharati[20] have associated higher quantity of albuminoid ammonia with polluted waters. Bharati and Hosmani[3] have observed the Yemmekeri Pond in Dharwar, exhibiting euglenoid blooms, to possess a range of albuminoid nitrogen extending from 1·36 mg/L to 91·2 mg/L, the maximum being observed in January. No marked periodicity could be established.

(b) *Dissolved Oxygen*

The dissolved oxygen content of natural lentic tropical bodies of water in India is usually low. Temperature seems to play a very important role in the distribution of dissolved oxygen. Periods of high temperature show low oxygen content[19,56,70] although Singh[68] failed to find this relationship in a lake at Delhi. The dissolved oxygen content is high during rains, highest during winters and lowest in June[68,70]. In a pond in West Bengal it ranged from 4-12 ppm[25]. The lower values (4-6·4 ppm) were observed during April and May and the rest of the year was characterised by higher values (7·6-12 ppm). Bharati and Hosmani[3] note a maximum in November although their values never exceeded 5 mg/L.

It is generally believed that high dissolved oxygen content of a sheet of water containing abundant vegetation is invariably accompanied by an increased pH and decreased carbon dioxide, nitrite, nitrate and silica

(cf. Singh[70]). Carbon dioxide, however, exhibits a direct relationship with the dissolved oxygen at the Ramgarh and Suraha Tals and pH does not seem to exhibit a direct relationship with the dissolved oxygen. The values of nitrate, silica and nitrite are independent of the dissolved oxygen values[70] although a deficiency of dissolved oxygen may cause an accumulation of nitrite in freshwaters[58,90]. A low dissolved oxygen content has also been found to coincide with a higher ammonia content[13,90].

(i) *Organic Matter*

The decomposition products of dead organisms, extracellular products secreted by living organisms and decomposed organic matter of catchment areas are the chief sources of organic matter in a freshwater aquatic system. High values of dissolved organic matter during summers and the monsoon have been attributed to the accelerated rate of bacterial decomposition during summers and surface run off during monsoon,[19,54,70,90,92] but, no fixed periodicity in the oxidizable organic matter content of two ponds where the concentrations ranged from 0-6·8 and 0-4·5 mg/L was noted[90]. Singh[68] has reported the highest values from September to November. Rao[58] states that soluble organic matter concentration is directly related to air temperature and it accumulates in large proportions in the water in the summers and early monsoon, decreases after heavy rainfall and reaches a minimum in the winters, although he has also noted an increase after heavy rainfall. Bharati and Hosmani[3] have observed maximum dissolved organic matter in February.

An inverse relationship exists between dissolved oxygen and soluble organic matter[13,58,68,70,90]. The low values of dissolved oxygen along with high values of organic matter may be due to the fact that a large quantity of oxygen is used up in the oxidation of organic matter or the chemical oxidation demand[45].

Gonzalves and Joshi[19] and Rao[54] find a direct relationship between organic matter and albuminoid ammonia although Zafar[90] demonstrates an inverse relationship. Rao[54] and Zafar[90] observed the C/N ratio to be determined mainly by organic matter and Rao[58] states that the higher C/N ratio of the ponds studied by him indicates that the ponds are rich in allochthonous materials (from external sources) and a higher C/N ratio in the monsoon suggests that the allochthonous materials are rich in carbon. Rao[58] and Zafar[90,91] have noted a direct correlation between magnesium

and oxidizable organic matter, suggesting that magnesium plays an important role in the decomposition of organic matter. Zafar[90] has also observed the pH to rise in water with a large concentration of organic matter. The concentration of total iron, however, falls when organic matter content is high. High turbidity[58] has also been observed to have a profound influence on the decomposition of soluble organic matter and the monthly fluctuations of phosphate and soluble organic matter were identical. Rao[58] states that bacterial activity could have been enhanced by phosphorus which can result in an increased conversion of particulate organic matter into soluble form.

(iii) *Algal Periodicity*

The seasonal periodicity and the relation of the phytoplankton and the benthic algae to the physicochemical environment has been extensively surveyed for tropical lentic habitats. It is generally believed that a high pH is related to heavy blooms of phytoplankton[10,25,38,74] but an inverse relation has also been noted[79]. An increase in dissolved oxygen also seems to cause an increase in phytoplankton bloom[10,25,38,79]. Khan and Qayyum[81] have noted that the ratios of bicarbonate/carbonate and calcium/carbonate were indicative of production levels. Singh[70] has observed two plankton peaks in the Ramgarh Tal, one in September and October and the other in December and January and one peak in the Suraha Tal in February. Singh[68] has noted the total phytoplankton to show a striking parallelism with suspended solids. Zafar[91] states that when the basic ratio of the water fluctuates between 0·027 and 1·2, diatoms are dominant and, between 1·2 and 2·0, green algae contribute equally to the bulk of the phytoplankton, between 2·0 and 3·7, Chlorococcales and desmids become dominant. Beyond 3·7 and up to 5·5, the percentage of green algae in the population does not increase further. Kant and Kachroo[29] have noted three peaks of phytoplankton and state that when pH is high the phytoplankton shows its maximum development. Zutshi *et al.*[95] have demonstrated that the Kashmir lakes which range from ultra-oligotrophic to oligotrophic, showing short-lived thermal stratification (cold polymictic), are poor in plankton, with only a few specimens of Chlamydomonas being collected. These lakes showed an increase in productivity from higher to lower altitudes a change from ultra-oligotrophy through oligotrophy and mesotrophy to eutrophy. The low altitude lakes are in a rapid process of eutrophication. Kant and

Kachroo[30] state that in the Dal and Nagin lakes, both of which are eutrophic, the peak for Myxophyceae was in October, Chlorophyceae in August, Bacillariophyceae in January–February, Cryptomonadineae in September-October and Dinophyceae and Euglenineae in August-September. Bacillariophyceae had a secondary peak in October. The Bacillariophyceae always dominated, followed by Chlorophyceae and then the Cyanophyceae. Lakes richer in oxidizable organic matter are usually richer in Cyanophyceae[19,60,88,68]. Singh[70] has, however, noted certain blue-green forms to be indifferent to organic matter content, probably because the amount of dissolved organic matter, remaining at high levels throughout the year, did not limit the production of these forms. Zafar[93] and Rao[60] state that certain blue-greens (Chlorococcales) do prefer oxygen-rich waters with low organic matter. Zafar[93] concludes that it is the carbon rather than the nitrogen of the dissolved organic matter which is more important in influencing the periodicity of the Myxophyceae. Species of *Oscillatoria* dominate in regions of high organic matter and low oxygen content[60]. A low pH with low dissolved oxygen are believed to favour the growth of certain Cyanophycean members[12,60,88] although *Merismopaedia glauca* and *Chroococcus minutus*[93] are favoured by a high dissolved oxygen content. *Merismopaedia tenuissima* exhibited fluctuations[60] coinciding with those of albuminoid ammonia and *Gomphonema aponina* multiplied in larger concentrations of potassium. The dominance of blue-green algae has been correlated with hot summer months usually characterised by bright sunshine and higher atmospheric and water temperatures[14,19,25,28,93] although Singh[70] has observed the lowest values in the summer months of April and May and concludes that temperature is not the factor which limits the production of blue-green algae. Ganapati *et al.*[15] trace the cause of a permanent bloom of blue-green algae in a tank to thermal stratification and diurnal turnover. Vyas and Kumar[86] conclude from their studies on the Indra Sagar pond that no correlation could be obtained between Cyanophycean dominance and phosphate or nitrate. Ganapati[14] states that stagnation, salts of Ca and Mg and large amounts of phosphorus were the factors favouring the blooms of blue-green algae. Sreenivasan[76], however, contradicts this statement. Jayangoudar and Ganapati[28] state that absence of water movements and excessive amounts of phosphates favour blue-green algae. Venkataraman[81] found that the addition of calcium did not alter the yield of *Anacystis montana* although a rise in total iron increased it. He also stated that the

blue-green algae tolerated high sodium. Sreenivasan[76] states that waters with low total hardness, low alkalinity, low chlorides and low salts harboured more blue-greens as compared to waters with high values for these parameters. Microcystis and other blue-green algal blooms occurred in waters with chloride content as low as 4 to 10 and as high as 188–560 ppm.

Higher atmospheric or water temperature is considered to be one of the important factors in the periodicity of Chlorococcales[29,70,93]. Zafar[93] states that the various species of Chlorococcales react diversely to different temperatures whereas Singh[70] reports a second maximum for Chlorococcales in the winters, Kant and Kachroo[29] observed a second peak of Chlorococcalean growth after the rainy season. Well-oxygenated waters have been considered a prerequisite for harbouring a larger number of Chlorococcales[19,59,70,90,91]. Gonzalves and Joshi[19] find a direct relationship of the Chlorococcales with pH and Zafar[93] has observed a deviation from the pH range of 7·5 to 9 to be inimical to the growth of Chlorococcales. Zafar[91] has also correlated low values of calcium with a thick growth of the green algae as a whole and found that higher concentrations of calcium and magnesium disfavour the growth of Chlorococcales whereas those of sodium and potassium seem to promote it[93]. Zafar[91,93] has pointed out the importance of Pearsall's basic ratio in the ecology of diatoms, Chlorococcales and desmids. Chlorococcales become dominant in the phytoplankton when the basic ratio fluctuates between 2·0–3·7. Singh[70] and Zafar[91] state that Chlorococcales prefer waters with a low average for nitrates and phosphates. Green algae usually favour ponds with a high concentration of albuminoid ammonia (0·46 mg/L). Singh[70] states that high free carbon dioxide contents are suitable for a better growth of Chlorococcales. Jana[35] has observed the Chlorophyceae to be dominant in summer, less numerous in winter and completely absent in autumn and concludes that phosphates are directly correlated to green algal populations.

It is generally well known that desmids prefer acidic waters, a high 'Pearsall's basic ratio' and a low nitrate phosphate ratio. Zafar[93] maintains that a higher concentration of dissolved oxygen is favourable for the growth of desmids and that the desmids remain dominant when the 'basic ratio' range fluctuates between 1·2–2·0 and in wider ranges it shows an inverse relationship. Singh[70] and Zafar[93] have observed a higher percentage of desmids during the summers although Singh[70] has observed two peaks, the other in winters. Zafar[93] states that higher temperatures (31°–37°C) during May

and June may inhibit desmid growth. Gonzalves and Joshi[19] state that the total solid content is important in the periodicity of desmids although Zafar[93] states that in the ponds he studied, water concentration is not the factor. Zafar[93] further states that a higher pH, coupled with low calcium concentration, is favourable for the growth of desmids and that low calcium is more important than a low magnesium, sodium and potassium seem to directly effect the growth of desmids.

Zafar[93] has summarised the ecological information on the Peridineae and states that the ecological complexes favouring this group of algae are markedly different from those for Chlorococcales and desmids. A pH range from 7·6–8·3 is favourable for species of Peridineae. Philipose[46] has observed a bloom of *Peridinium inconspicuum* to frequently form dirty-brown patches on the surface layer of turbid waters.

Gonzalves and Joshi[19] regard high temperature, great light intensity and large amount of organic matter as favourable for the growth of *Trachelomonas* spp. and many species of *Euglena*. A rich organic matter content has also been found favourable for the growth of Euglenineae[20,40,60,89]. Singh[70] is, however, of the opinion that a correlation with high organic matter content is not universal and that the findings that Euglenas reach their maximum growth when the nitrate content is low and albuminoid nitrogen is at its maximum do not seem to be supported by the behaviour of *Euglena* spp. in the Ramgarh and Suraha Tals. Zafar[91] attributes the paucity of euglenoid flagellates to very low iron contents and alkaline nature of the waters, while higher concentrations of free carbon dioxide, phosphates and iron and lower concentrations of oxygen considered favour for the growth of the Euglenineae[20,40,70]. Seenayya[66] suggests that dissolved oxygen is high during the bloom because of the photosynthetic activity and also because water temperature around 28 °C is favourable for the growth of Euglenineae. Rao[60] has found higher salinity and higher basic ratio as unfavourable for the development of *Trachelomonas volvocina*. Singh[70] has observed maximum development of Euglenineae in the months of September–February although Rao[60] noticed best development in the early summer in all the ponds he studied.

Iyengar[21] has discussed the ecology of colonial Volvocales of South India. He contended that in rainwater pools harbouring a rich population of *Chlamydomonas* spp., *Gonium pectorale* and *Pandorina morum*, organisms

growing around the sand particles were more healthy and accumulated there possibly due to well-aerated conditions of the thin film of water covering the sand grains. Gonzalves and Joshi[19] relate the periodicity of *Gloeocystis vesiculosa* and *Pandorina morum* to the meteorological factors and salt content of water. Philipose[46] and Rao[59] attach more importance to the water reaction and the concentration of nitrates and phosphates. Rao[59] states that the Volvocales prefer eutrophic waters having high salt content and demonstrate two peaks, one in early summer and the other in early winter. The favourable range of temperature extends from 23-28 °C. An interesting phenomenon in the periodicity of Volvocales was the appearance of one form after the other in close succession. Rao[59] explains this by stating that the extra-cellular substance produced by a particular species possibly served as nutrient for the other.

Temperature has been emphasized as playing an important role in the distribution of diatoms. Singh[70], George[17], and Zafar[93] found the colder months favourable for the multiplication of diatoms. Zafar[94] is of the opinion that the cold preference of diatoms could be due to their heavier bodies being supported better by denser, cooler waters. Rao[60] states that in his studies, the diatoms exhibited no fixed periodicity, some forms developed best in summers, other in winters while some were indifferent to seasonal variations. Jana[25] has observed two peaks in summer (March and late April) and another, the highest in autumn (September). Kant and Kachroo[29] state that bright sunlight, open, medium-shaded areas and associations of aquatic angiosperms are factors stimulating the growth of a number of diatoms. It is generally held that ponds rich in nitrates, phosphates, silicates, calcium, sodium, oxygen, basic ratio, pH and organic matter are better suited for the growth of diatoms in general. A higher percentage of diatoms in the ponds studied by Zafar[91] has been attributed to a higher concentration of phosphates and low organic matter, high calcium (1·61-2·18 mg/L), moderate pH (7 5-8·1) and low sodium content (2·4-3·3 mg/L). Singh[70] has observed an inverse relation between nitrates, phosphates and silicates and *Melosira granulata* in the Ramgarh Tal. The organic matter content was observed to exhibit no relation to diatom numbers. Singh[70] concluds that these factors in the 'Tals' studied by him have very high concentrations throughout the year and, as such, they do not seem to limit the abundant growth of diatoms. Zafar,[93] Jana,[25] and Rao[60] have endorsed this observation and state that silica, nitrate and phosphate, although

otherwise emphasized, do not play any controlling role in determining the periodicity of Bacillariophyceae.

(iv) *Water-Blooms*

Water-blooms may be mixed, i.e. constituted by a number of species or they may be unialgal (pure). They may last a few days only (temporary) or for longer periods (permanent). Interestingly other conditions remaining constant, a single factor may trigger the active multiplication of bloom. Colonial Volvocales, which otherwise grow best in monsoons, form blooms in the winters or even at the peak of summers in South India[21]. These periods were characterised by either turbid or deep and shaded waters. With a high concentration of nitrates and phosphates as in the monsoon, it appears that light is the only important factor which causes the development of the bloom. Philipose[46] believes that the difference between temporary and permanent blooms is that the latter can tolerate a wider range of physicochemical variations and are chiefly constituted by the blue-green algae. The mortality of fishes in bloom-infested waters is a matter of controversy. While some regard it to be due to oxygen depletion, others attribute it to extracellular toxic products. An analysis of the chemical constituents of bloom-forming algae and the surrounding water, failed to reveal any toxic substance[69]. Iyengar[32] states that waters harbouring a bloom of *Microcystis* were alkaline and rich in organic matter and chlorides. Ganapati[12] also noted an association of blooms with high pH, organic matter, free and albuminoid ammonia small quantities of phosphates and little or no nitrites or nitrates which he attributed to the alga utilising them immediately on formation. Singh[69] in addition, finds reducing conditions as favourable. Philipose[47] has observed different *Microcystis* spp. mostly from fish ponds enriched by nitrogen and phosphorus, in ponds receiving regular organic pollution, in temple tanks (South India) and a few lakes and reservoirs in which the infestation was mostly during winter or summer or both. The pH was always on the alkaline side.

(v) *Diurnal Variations*

Studies on diurnal variation in polluted and freshwater bodies are few[4,16,26,40,63,75,77,87]. The pH has been observed to be maximum in the daytime (3 p.m.) and minimum during late night[4,26]. Sreenivasan *et al.*[77]

observed that the maximum amplitude of variation in pH in a day was 1·4 units in October 1967 and the minimum was 0·4 units in February and November 1969. They found no thermal stratification in a shallow pond at Madras, the water temperature was highest between 14 to 18 hours and lowest between 2 and 6 hours. Sreenivasan[75] had also observed similar though lesser fluctuations. Bharati and Kore[4] state that the water temperature followed the atmospheric temperature only during the daytime, and the waters exhibited very little diurnal variations in transparency.

The dissolved oxygen content shows a distinct diurnal variation owing to photosynthetic and tropholytic activities. Maximum dissolved oxygen was observed between 10 and 18 hours and minimum between 2 and 6 hours[4,26,77]. Verma[87] found a direct relationship between pH and dissolved oxygen. Sreenivasan *et al.*[77] have recorded carbon dioxide during late and early hours of the day and, as expected, Bharati and Kore[4] report free CO_2 to move inversely to the dissolved oxygen content. The phosphates were lowest during the daytime and showed an inverse relation to the dissolved oxygen[4]. Sreenivasan *et al.*[77] have given a very comprehensive account of the diurnal variation and state that the peaks of phenolphthalein alkalinity and pH coincide with the lowest level of methyl-orange alkalinity. The phenolphthalein alkalinity was always absent in the late and early hours of the day. Although, silicates exhibit a pronounced seasonal variation, diurnal changes in them were not manifested. Chlorides also behave identically in not exhibiting diurnal variations. Total hardness varied from 84–310 ppm and the maximum amplitude of diurnal variation was recorded in July (18 ppm).

Philipose[46] reported a downward migration of *Trachelomonas volvocina* during day hours and an upward movement during night although, Verma[87] recorded a maximum concentration at the surface at 9. a.m. and a minimum at 6 p.m. Jana[36] has observed this species to retreat at night and surface in the morning. It seems that the migration of this taxon is affected by illumination. Jana[36] has observed *Pediastrum* to exhibit a marked downward nocturnal movement and was most abundant on surface when the solar energy was at its maximum. *Spirogyra* aggregated on the surface in afternoons followed by a midnight sinking. Although *Navicula* exhibited no marked vertical migration, the maximum upward climb was noted in conjunction with sunrise. Bharati and Kore[4] have observed a well-marked diurnal variation in *Spirulina*. It was found to be maximum at 6 p.m. and 6 a.m.

Possibly, light intensity and temperature are the principal causes of variation. *Euglena* and *Phacus* appeared to be phototropic and exhibited a characteristic movement.

(vi) *Pollution*

The economics of algal hydrobiology, is manifest best in the spoilage of water. That certain taxa are specific to a given set of condition and grow best in them has been widely established. Polluted waters harbour a characteristic algal flora and it is possible to get an approximate idea of the chemistry of waters by noting predominating algae. Several studies have been conducted on the pollution ecology of Indian lentic waters[3,5,11,54,57,46,52,58,58,60,71,72]. Alikunhi *et al.*[2] have stated that the mere presence of *Microcystics* spp. is indicative of the fertility of water. Philipose[46] has observed *Nitzschia* sp., *Surinella robusta*, *Scenedesmus quadricauda* var. *longispina*, *S. abundans*, *S. longus* and *Merismopaedia* sp. along with *Navicula* sp., *Oscillatoria* sp. and some members of Volvocales and Chlorococcales to dominate in a tank at Vishakapatnam which is regularly polluted with sewage. Bick *et al.*[5] have emphasised the applicability of the Kolkwitz and Marsson[33] system for monitoring water quality under Indian conditions. Singh and Saxena[71] have studied the composition of the algal flora and its periodicity in raw and stabilized sewage and state that members of the Cyanophyta and Euglenophyta predominate in raw sewage whereas the stabilized sewage harbours a richer percentage of Chlorococcales. Singh *et al.*[72] state that nitrogen fixing forms of blue-green algae were not observed in sewage and the percentage distribution of different groups of algae follows a definite trend in the sewage. The algal flora of raw sewage was characterised by a dominance of blue-green algae. Rana *et al.*[53] state that *Plectonema boryanum* is extremely tolerant to zinc and can be used in the assessment of zinc pollution. Rao[58] is of the opinion that higher values of oxygen/organic matter may indicate a lesser degree of eutrophication and that *Oscillatoria limosa*, *O. chalybea* and *O. grunowiana* indicate a higher degree of eutrophication in the water. Franklin[11] also contended that blue-green algae in general indicated eutrophic waters. Kumar *et al.*[34] also found blue-green algae, flagellates and euglenoids to be most intimately associated with organically rich effluents and green algae such as *Chlorella* to be the dominant forms found in waters rich in nitrogenous compounds. Marathe and Nandkar[57] indicate that the genus *Oscillatoria* is very tolerant to pollution

and can be used as an indicator of the α-meso-saprobic zone of polluted waters. Rai and Kumar[52] have observed that, in polluted waters with high total alkalinity, chlorides, nitrogen and total solids, the diatoms were the commonest algae and the Euglenoids had the maximum density and abundance. No heterocystous form was found growing in the nitrogen rich habitats and this is in accord with the established fact that ammonia inhibits the development of heterocysts. *Oscillatoria* was also observed by him to be very tolerant to pollution.

It appears that the pollution ecology of inland waters of India is in infancy and attempts should be made to study the algal community in relation to pollution on a long-term basis. Structural changes in the community-complex should be studied as it is believed that individual taxa, because of the complexity of factors involved, cannot serve as useful indicators of water quality. It is advocated that community changes, dominant associations, generic and species diversity can be successfully used to monitor water quality, as was done by Singh[78] for lotic environments.

REFERENCES

1. Agarwal, D. K., Gaur, S. D., Tiwari, I. C., Narayanswami, N. and Marwah S. M. (1976) *Indian J. Environ. Hlth.*, 18(3) : 201.
2. Alikunhi, K. H., Chaudhari, H. and Ramachandran, V. (1955) *Indian J. Fisheries*, 21 : 257.
3. Bharati, S. G. and Hosmani, S. P. (1973) *Karnatak Univ. J. Sci.*, 18 : 246.
4. Bharati, S. G. and Kore, A. P. (1975) *Karnatak Univ. J. Sci.*, 20 : 157.
5. Bick, H., Krishnamoorthy, K. P. and Lakshminarayana, J. S. S. (1967) WHO EBL., 1.
6. Blum, J. L. (1956) *Bot. Rev.*, 22(5) : 291.
7. Chacko, P. I. and Ganapati, S. V. (1949) *Indian Geogr. Journal*, 24(3) : 35.
8. Chacko, P. I., Srinivasan, R. and Evangeline, G. (1953) *Ind. Comm. J. Madras*, 8(1) : 118.
9. Chakrabarty, R. D., Roy, P. and Singh, S. B. (1959) *Indian J. Fisheries*, 6(1) : 186.
10. Das, S. M. and Srivastava, V. K. (1956) *Proc. Nat. Acad. Sci India*, 26B : 85.
11. Franklin, T. (1972) First Internat. Symp. on Taxonomy and Biology of Blue-Green Algae. Univ. Madras, Madras, 442.
12. Ganapati, S. V. (1940) *J. Bombay Nat. Hist. Soc.*, 42(1) : 65.
13. Ganapati, S. V. (1943) *Proc. Indian Acad. Sci.* B 17 : 41.
14. Ganapati, S. V. (1960) *Proc. Symp. Algol.*, ICAR., New Delhi, pp. 204.
15. Ganapati, S. V., Chacko, P. I. and Srinivasan, R. (1953) *J. Asiatic. Soc.*, 19 : 149.
16. George, M. G. (1961) *Hydrobiologia*, 18 : 265.
17. George M. G. (1966) *Hydrobiologia*, 27 : 81.
18. Ghousuddin, M. (1934) *J. Osmania Univ.*, 1.

19. Gonzalves, E. and Joshi, D. B. (1946) *J. Bombay Nat. Hist. Soc.*, 46(1) : 154.
20. Hosmani, S. P. and Bharati S. G. (1975) *Karnatak Univ. Jour. Sci.*, 20 : 151.
21. Iyengar, M. O. P. (1933) *J. Linn. Soc. Bot.*, 49 : 323.
22. Iyengar, M. O. P. (1938) Proc. 25th Indian Sc. Congr., Part IV., Discussion, 140.
23. Iyengar, M. O. P. and Venkataraman, G. (1951) *J. Madras Univ.*, 21 : 140.
24. Jackson, D. F. (1964) Algae and Man., Plenum Press, New York.
25. Jana, B. B. (1973) *Int. Revue ges. Hydrobiol.*, 58(1) : 127.
26. Jana, B. B. (1974) *Ekologia Polska*, 22(2) : 287.
27. Jana, B. B. and Sarkar, H. L. (1971) *Hydrobiologia*, 37 : 33.
28. Jayangoudar, I. and Ganapati, S. V. (1962) Symp. Waste Treatment by Oxidation Ponds, CPHERI, Nagpur, 1-5.
29. Kant, S. and Kachroo, P. (1971) *Proc. Nat. Acad. Sci. India*, B. 37 : 163.
30. Kant, S. and Kachroo, P. (1973) *Proc. Indian Nat. Science Academy*, 39B(5) : 632.
31. Khan, J. and Qayyum, A. (1966) *Hydrobiologia*, 28 : 195.
32. Klein, L. (1957) Aspects of River Pollution. Butterworths Scientific Publications, London : 621.
33. Kolkwitz, R. and Marsson, M. (1908). *Ber Deutschen Bot. Gessel.*, 26a : 505. Transl. U. S. J. Publ. Res. Ser., Ed. Keup *et al.*, Fed. Wat. Poll. Cont. Admn. Cincinnati, Ohio, 1967, 47.
34. Kumar, H. D., Bisaria, G. P., Bhandari, L. M., Rana, B. C. and Sharma, V. (1974) *Ind. J. Environmental Hlth.*, 16 : 247.
35. Kundra, R., Nagpal, J. L., Verma, S. R. and Shrivastava, S. K. (1977) *Indian J. Environ. Hlth.*, 19(4) : 329.
36. Lakshminarayana, J. S. S. (1965) *Hydrobiologia.*, 25 : 119.
37. Marathe, K. V. and Nandkar, P. B. (1976) *Indian J. Environ. Hlth.*, 18(4) : 331.
38. Moitra, S. K. and Bhattacharya, B. K. (1965) *Icthyologica*, 4 : 8.
39. Motwani, M. P., Banerjea, S. M. and Karamchandani, S. J. (1956) *Indian J. Fisheries*, 3 : 334.
40. Munnawar, M. (1970) *Hydrobiologia*, 31 : 105.
41. Munnawar, M. and Zafar, A. R. (1967) *Hydrobiologia*, 29 : 140.
42. Olaniya, M. S., Saxena, K. L. and Sharma, H. C. (1976) *Indian J. Environ. Hlth.* 18(3) : 219.
43. Pahwa, D. V. and Mehrotra, S. N. (1966) *Proc. Nat. Acad. Sci. Indian*, B. 36(2) : 157.
44. Palmer, C. M. (1967) Proc. IV Ann. Wat. Quality. Res. Symp. New York State Dep. Hlth., Albamy, New York, 8.
45. Pearsall, W. H. (1923) *J. Ecol.* 11 : 165.
46. Philipose, M. T. (1960) Proc. Symp. Algol., I. C. A. R., New Delhi, 272.
47. Philipose, M. T. (1972) First Internat. Symp. on Taxonomy and Biology of Blue-Green Algae. Univ. Madras, Madras, 455.
48. Prasad, B. N. (1964-65) *B.H.U. Jour. Sc. Res.*, 15 (1) : 142.
49. Prasad, B. N. and Saxena M. (1980) *Indian J. Environ. Hlth. (In Press).*
50. Prasad, B. N. and Singh, Y. (1980) III Indian Geophytol. Conf., Lucknow, Abst., 41.
51. Rai, H. (1974) *Arch. Hydrobiol.*, 73 (3) : 369.
52. Rai, L. C. and Kumar, H. D. (1976) *Nova Hedwigia*, 27 : 805.
53. Rana, B. C., Gopal, T. and Kumar, H. D. (1971) *Environ. Hlth.*, 13 (2) : 138.

*54. Rao, C. B. (1949) On the distribution and periodicity of algae in some small ponds., Ph.D. Thesis, Univ. Lond.
55. Rao, C. B. (1953) *J. Ecol.*, 41 : 62.
56. Rao, C. B. (1955) *J. Ecol.*, 43 : 291.
57. Rao, V. S. (1971) *Hydrobiologia*, 38 : 213.
58. Rao, V. S. (1972) *Hydrobiologia*, 39 : 351.
59. Rao, V. S. (1975) *Hydrobiologia*, 47 : 319.
60. Rao, V. S. (1977) *Hydrobiologia*, 53 : 13.
61. Round, F. E. (1965) Biological Problems in water Pollution. Ed. Tarzwell *et al.*, US Deptt. Hlth., Education and Welfare., Cincinnati, Ohio, 29-33.
62. Roy, H. (1955) *Ecology*, 36 (2) : 169.
63. Saksena, S. B. and Adoni, A. D. (1973) *Hydrobiologia*, 43 : 535.
64. Sarkar, H. L. and Rai, H. (1964) *Trans. Amer. Microsc. Soc.*, 83 : 260.
65. Saxena, M. (1978) Studies on the Influence of Industrial Effluents on some lotic Algae., Ph.D. Thesis, Univ. Lucknow, Lucknow, India.
66. Seenayya, G. (1972) *Hydrobiologia*, 39 : 247.
67. Seenayya, G. and Subbaraju, N. (1972) First Internat. Symp. on Taxonomy and Biology of Blue-Green Algae., Univ. Madras, Madras, 52-57.
68. Singh, M. (1965) *Phykos*, 4 : 61.
69. Singh, R. N. (1955) *Proc. Internat. Assoc. of Theor. and Appl. Limn.*, 12 : 831.
70. Singh, V. P. (1960) Proc. Symp. Algol., I.C.A.R., New Delhi, p. 243.
71. Singh, V. P. and Saxena, P. N. (1969) *Hydrobiologia*, 34 : 503.
72. Singh, V. P., Saxena, P. N., Tewari, A. and Khan, M. A. (1970) *Phykos*, 9 : 57.
73. Singh, Y. (1979) Ecological Studies on River Algae with special Reference to Bacillariophyceae., Ph.D. Thesis., Univ. Luck., Lucknow, India.
74. Sreebuvasan, A. (1963) *Curr. Sci.*, 32 : 130.
75. Sreenivasan, A. (1965) *Hydrobiologia*, 26 : 501.
76. Sreenivasan, A. (1972) First Internat. Symp. on Taxonomy and Biology of Blue-Green Algae. Univ. Madras, Madras, 461.
77. Sreenivasan, A., Soundaraj, R. and Franklin, T. (1973). *Phykos*, 12 (1 & 2) : 86.
78. Sreenivasan, A., Ananthanarayanan, R. and Paramsivm, M. (1977) *Indian J. Environ. Hlth.*, 19 (2) : 141.
79. Srivastava, V. C. and Sahai, R. (1976) *Geobios.*, 3 (6) : 187.
80. Vaas, K. K., Raina, H. S., Zutshi, D. P. and Khan, M. A. (1977) *Geobios.*, 238.
81. Venkataraman, G. S. (1962). *Proc. nat. Inst. Sci. India*, B 20 : 77.
82. Venkateswarlu, V. (1969a) *Hydrobiologia*, 33 : 117.
83. Venkateswarlu, V. (1969b) *Hydrobiologia*, 33 : 352.
84. Venkateswarlu, V. (1969c) *Hydrobiologia*, 34 : 533.
85. Venkateswarlu, V. (1970) *Hydrobiologia*, 35 : 45.
86. Venkateswarlu, V. (1976) *Nova Hedwigia*, 27 : 661.
87. Verma, M. N. (1967) *Hydrobiologia*, 30 : 129.
88. Vyas, L. N. and Kumar, H. D. (1968) *Hydrobiologia*, 31 : 421.
89. Zafar, A. R. (1959) *J. Indian bot. Soc.*, 34 : 430.
90. Zafar, A. R. (1964a) *Hydrobiologia*, 23 : 179.
91. Zafar, A. R. (1964b) *Hydrobiologia*, 24 : 556.
92. Zafar, A. R. (1966) *Phykos*, 5 (1 & 2) : 115.
93. Zafar, A. R. (1967) *Hydrobiologia*, 30 : 96.
94. Zafar, A. R. (1968) Proc. "School on Plant Ecology (1966)", B.H.U. 368.
95. Zutshi, D. P., Kant, V. and Vass, K. K. (1972) *Verh. Internat. Verein. Limnol.*, 18 : 599.

*Not seen in Original.

Computers, Topology and Drug Design

E. V. Krishnamurthy

Indian Institute of Science, Bangalore 560012, India

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

Introduction

An important facet of modern medicinal chemistry is drug design, wherein the chemist has to design a compound with a prescribed biological profile. Although the achievement of this aim still appears to be far in the future, due to the lack of our understanding of biological and disease processes at the molecular level, recently several methods have been developed to relate the molecular structure with drug activity. Two approaches to this problem have emerged and both have been of significant value in the discovery and development of new drug therapy.

The first approach is a biochemical approach in which on the basis of established and/or hypothesized models, one deduces the mechanism of drug action at the molecular level. The other approach, which we would deal with in this article, is primarily a statistical approach in which a variety of combinatorial and taxonomic computational techniques are applied to establish the correlations between the level of a given biological activity and other measurable (or calculable) properties associated with the chemical structure of the molecule. Much of this developing quantitative methodology, as we shall see, depends heavily on the modern computing machinery for statistical and non-numerical data processing.

Quantitative Methodology

A drug is a chemical substance that exerts a reproducibly observable effect on a biological system. This effect may be called biological activity or biological response, and is usually dose-dependent. The dose is the quantity or concentration of the administered compound ; the response is the observed effect, which can be as objective as the percent inhibition of an enzymatic reaction, or a change in the heart rate or as subjective as a change in the mood or posture of a test animal. In view of the enormous complexity of factors that affect the *in vivo* activity of a drug (absorption, transport, tissue distribution, metabolism, etc.) it is not surprising that structure-activity correlation has been most successful with data obtained from less complex biological systems, such as isolated organs, cell-cultures or purified enzymes. In addition, since the experimental variability in biological measurement is much larger than that in most physical or chemical measurements, it becomes necessary to obtain very reliable biological data with a well-defined confidence limit, to obtain meaningful correlations.

In relating biological activity mathematically to other properties of the molecular structure, one tries to construct a linear system of equations of the form

$$\text{Activity} = a_0 + \sum_{i=1}^{m} a_i x_i$$

where x_i are structural properties and the coefficients a_i are determined from analysis. This equation allows prediction of the biological activity when x_i are known.

Procedures for obtaining the m-coefficients $a_1, a_2, \cdots, a_m$ require an experimentally determined biological activity and a group of m-structural properties for each of the n-compounds which are structurally related. Because of the relatively low precision of the biological data and the uncertainty that a linear correlation model is applicable to any structure activity relationship, good practice in drug design requires that $n \gg m$. The number of compounds in excess of the minimum $(n+1)$ required for analysis is called the number of degrees of freedom.

The coefficients a_i are usually obtained by using least squares method or equivalently using generalized inverse[1] methods.

The linear model is an oversimplification to handle more complex drug design problems where the relationship is not linear and several variables are involved, it is necessary to use multivariate statistical analysis, e.g., Cluster analysis, a mathematical technique for classification in which similarity is defined in terms of a distance[2,3,4] between the points representing the objects (i.e., compounds) in multidimensional variable space. For convenience of quick inference, these results can be displayed in a graphic facility for visual pattern recognition.

Quantum Chemical Method

For studying the structure-activity relationship, it becomes necessary to quantitatively represent the molecular structure. The most complete quantitative representation is obtained by the use of quantum mechanics. Molecular Orbitals methods (MO) have been used with success in structure-activity correlation[5,6,7,8] ; these involve enormous amount of computation. Accordingly several alternative approaches which are simpler, have been proposed recently.

The most common of these in terms of practical application is the use of a physical property such as partition coefficient as a surrogate for the

structure[6]. This is also not entirely satisfactory since it does not permit an unambiguous relationship between the salient structure characteristics and the activity.

Accordingly, attempts have been made at a level below the quantum chemical level, to provide quantitative structure descriptors. These methods are :

1. Free-Wilson's additive model
2. Cramer's substructural analysis method
3. Kier's connectivity index method

Of these Free-Wilson's[9,10,11] method is essentially limited to structurally closely related series of compounds. In the substructural analysis method the biological activity of a molecule is related to the combination of contributions of the biological activities of substructural fragments.

The connectivity index method is based on arriving at a numerical index for a compound, based on the topological connectedness of each atom and all its neighbours. We will discuss these methods in the succeeding sections.

Free-Wilson Method

Here one assumes that the biological activity is an *additive* property of the substituents that vary within the series. Analysis based on this model have been able to account for a substantial part of the variation of the biological activity in numerous compounds. In the Free-Wilson method, every substituent is assigned a constant which represents the contribution of that substituent to the overall biological activity of the molecule in which it is present. These substituent constants are evaluated using least-square approach, one for each of the molecule in the series, using the concept :

Biological activity = Mean biological activity + Sum of substituent contributions

Using this approach, fairly good results have been obtained for design of antimalarial drugs using a series of phenanthrene-carbinols.

Topological Methods

(a) *Substructural Analysis (Fragmentation)*

A severe limitation of the Free-Wilson method is that it is restricted to structurally closely related compounds and it is inappropriate for :

(i) Correlation of data where compounds fall into different structural series or into no series at all.

(ii) Prediction of active compounds outside a structural class of established biological interest.

Therefore Cramer *et al.*[12] developed a substructural analysis method ; here, the biological activity of a molecule is related to the combination of contributions of the biological activities of structural components or substructures and their intra- and inter-molecular interactions. Since a very large body of information will be generated (by the resulting combinatorial operations), even with the most modest-sized molecule, we need to make some additional assumption such as :

(a) The probability of given biological activity can be usefully approximated by a first-order analysis of substructural contributions (i.e., ignoring interactions).

(b) The contribution of a given substructure to the probability of activity can be obtained from data on previously tested compounds containing that substructure.

With these two assumptions and a topological structure or substructure coding schemes[13,14,15,16,17], the basic functional groups, rings, chains, inorganic moieties, and other frequent combinations are recognized. Such a substructure coding scheme is naturally based on recognizing the subgraphs in a graph. Algorithms for automatic recognition of these are currently being developed by chemical-information scientists[17]. Having recognized these basic fragments, the next step is to prepare a substructure 'experience table' summarizing the data. A Substructure Activity Frequency (SAF) defined for each substructure as (A/T), the ratio of the number of active compounds (A) containing that substructure to the number of tested compounds (T) containing the substructure, represents the contribution that substructure can make to the probability of a compound being active. The experience table contains n (Say). SAF's corresponding to complete set of n substructures (fragments) previously recognized and coded in the test compounds.

Then we compute for each compound a 'Mean Substructure Activity Frequency' (MSAF), the arithmetic mean of the SAF values of the substructure present in that compound.

An example is given below ; here MSAF is computed for a compound which contains a substructure unique among the 770 compounds tested for antiarthritic/immuno regulatory activity[12]. [Fig. 1]

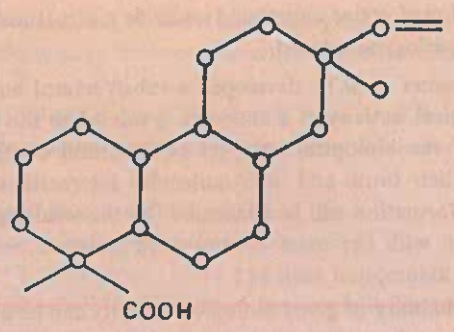


Fig. 1

No.	*Fragment*	*Tested (T)*	*Active (A)*	*SAF (A/T)*
1.	Aliphatic COOH	108	23	·212
2.	Phenanthrene ring system, one unsaturation	1	1	1·000
3.	Carbon chain as functional group	198	47	·237
4.	Methyl chain	309	83	·268
5.	Quarternary carbon	78	18	·230
6.	Three rings	174	45	·258
				2·205

$$\text{MSAF} = \frac{2 \cdot 205}{6} = \cdot 37$$

The 770 compounds can then be marked in descending MSAF value. Since a meaningful correlation would be reflected in a tendency for compounds of higher MSAF value to be active more frequently the 770 marked compounds are partitioned into 10 sets each containing 77 compounds. Those sets with higher MSAF values were found to be active for more frequently than those with low MSAF values.

This technique has been successfully used to predict the probability of certain types of activity.

(b) *Kier's Connectivity Index Method*

We now refer to the appendix, for definitions and terminology, used in the theory of graphs, which forms the basis for this method.

It has been demonstrated empirically by several workers that physico-chemical properties of a molecule are largely dependent upon the topology (or connectivity of atoms) of a molecule and accordingly these can be inferred within a reasonable range, by computing the composite of contributions from all the structural features influencing that property. It is therefore reasonable to suppose that if we could identify and quantify those salient topological features and weigh their contributions, we could arrive at an equation which relates very closely, the structure (as described by molecular connectivity indices) with the value of the property[15].

In an analogous manner, the biological properties can also be related to the molecular connectivity indices, which we shall define below :

Let G be the given hydrogen-suppressed chemical graph with n nodes and m edges ; list out all its connected subgraphs with one, two, three, ..., m edges. Let $d_i (i = 1, 2, \ldots, n)$ be the degree of each vertex (Degree graph) and Δ_i $(i = 1, 2, \cdots, n)$ be the non-hydrogenic valency of the atom depicted by the vertex (Valence graph).

Then the kth order molecular connectivity index is defined by

$$k(x) = \sum_p \left(\prod_r d_r \right)^{-\frac{1}{2}}$$

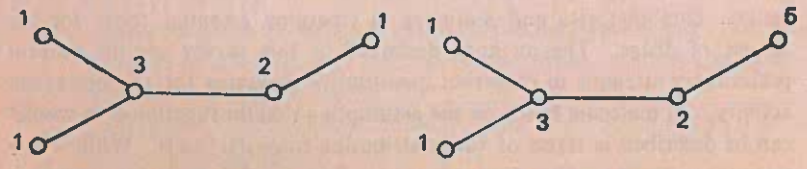
where p is the number of k-edged, connected subgraphs. These can be trees (t), stars (s) or cycles (c) or other suitable classes ; d_r denote the actual degree of nodes.

The kth order valence connectivity index is similarly

$$k(v) = \sum_p \left(\prod_r \Delta_r \right)^{-\frac{1}{2}}$$

Example :

Consider the graphs (Fig. 2)



Degree Graph Valence Graph

Fig. 2

$0\,(X)=4{\cdot}285$	$0\,(V)=3{\cdot}732$
$1\,(X)=2{\cdot}70$	$1\,(V)=1{\cdot}879$
$2\,(X)=1{\cdot}802$	$2\,(V)=1{\cdot}576$
$3_c\,(X)=0{\cdot}816$	$3_c\,(V)=0{\cdot}365$
$3_s\,(X)=0{\cdot}408$	$3_s\,(V)=0{\cdot}408$
$4_c\,(X)=0{\cdot}408$	$4_c\,(V)=0{\cdot}183$

The idea of describing the molecular information structure at an intermediate level rather than the most sophisticated quantum-chemical level is useful as a surrogate as it contains details of the number of atoms, degree of connections, Hetero atom-types, cyclization, etc. Although it is not unique it is a useful screening procedure. It has been shown that a proper set of a molecular indices correlate with many physical and biological properties.

Example :

The following table indicates the vapour toxicity of alcohols against red spider[18]

Compound	*1 (X)*	*Toxicity*
Methanol	1·00	2·8
Ethanol	1·41	3·0
Propanol	1·91	3·3
Isopropanol	1·75	3·3
Isobutanol	2·27	3·72
Butanol	2·4	3·8
Pentanol	2·9	4·1
Isopentanol	2·77	4·1

Concluding Remarks

The essential purpose of our survey here is to indicate the role of computers, combinatorics and statistics as emerging essential tools for the design of drugs. The methods described in this survey are the current preliminary attempts to construct quantitative measures for the biological activity of a molecule based on the assumption that the function of 'a whole' can be described in terms of the contribution from its 'parts'. While these techniques by themselves do not lead to a fool-proof design of drugs, these are very useful as preliminary steps for the selection of a suitable group of

drugs for test. Also these techniques have tremendous applications in the design of cosmetics, flavoroids, hormones and pesticides.

Drug design is still in its infancy and in the absence of rationally founded chemistry, there seems to be no alternative than being guided by the structure-activity correlations.

With improved computer manipulation of chemical structure records, substructural analysis can be improved by emphasising on those molecular features that prove to be of fundamental significance to biological mechanism. Advances in computer and programming technology are beginning to make these feasible.

APPENDIX

(For details see Ref. 19 & 20)

1. *Graph* : It is an object consisting of two sets called its node or vertex or point set and its edge or arc or line set. The vertex set is a finite non-empty set. The edge set may be empty ; but otherwise its elements are two element subsets of the vertex set.
2. *Pictorial representation* : It is customary to represent graphs by drawings, by representing vertices by heavy dots and edges by line joining these dots. The geometry of the lines are unimportant, as they indicate only the adjacency of vertices, indicating only the topological connections.
3. *Chemical graph* : A structural diagram of a molecule represented as a graph by identifying atoms as vertices and imaginary bonds as lines.
4. *Degree of a node* : The total number of arcs or edges connected to a node.
5. *Subgraph* : A graph H is a subgraph of graph G, if the vertex set of H is a subset of the vertex set of G and the edge set of H is a subset of the edge set of G.
6. *Connected graph* : A graph is called connected, if every pair of vertices is joined by a sequence of not necessarily distinct vertices ; otherwise it is said to be disconnected.
7. *Component* : is a connected subgraph that is not contained in a larger connected subgraph.
8. *Tree* : A connected graph having at least two nodes, that has no circuits ; there is a unique set of edges connecting any pair of vertices.

9. *Cycle* : A connected graph having at least three nodes, that has a closed circuit.
10. *Star* : A tree graph in which a single node is adjacent to all the other remaining nodes in the graph, none of them being adjacent.

REFERENCES

1. Krishnamurthy, E. V. and Sen, S. K. (1976) : Computer-Based Numerical Algorithms, Affiliated East-West Press, New Delhi.
2. Krishnamurthy, E. V. (1975) : *J. Ind. Inst. of Sci.*, 57 : 358.
3. Sneath, P. H. A. and Sokal, R. R. (1973) : Numerical Taxonomy, W. H. Freeman and Co., San Fransisco.
4. Hawkes, J. G. (1968) : Chemotaxonomy and Serotaxonomy, Academic Press, New York.
5. Kier, L. B. (1971) : Molecular Orbital Theory In Drug Research, Academic Press, New York.
6. Kier, L. B. and Hall, L. H. (1977) : Molecular Connectivity in Chem. and Drug Res., Academic Press, New York.
7. Purcell, W. P. Bass G. E. and Clayton, J. M. (1973) : Strategy of Drug Design, John Wiley and Sons, New York.
8. Kier, L. B. (1971) : Molecular Orbital Studies in Chemical Pharmacology, Springer Verlag, New York.
9. Burger, A. (1970) : Medicinal Chemistry, Wiley Inter Science, New York.
10. Ariens, E. J., (1971) : Drug Design, Academic Press, New York.
11. Redl, G., Cramer, R. D., and Berkoff, C. E., (1974) : *Chem. Soc. Rev.*, 3 : 273.
12. Cramer, R. D., Redl, G. and Berkoff, C. E. (1976) : Report of Technology Assessment Division, Smith Kline and French Laboratories, Philadelphia, Penn., USA.
13. Lynch, M. F., Harrison, J. M., Town, W. G. and Ash, J. E. (1972) : Computer Handling of Chemical Structural Information, American Elsevier, New York.
14. Davis, C. H., and Rush, J. E. (1975) : Information Retrieval and Documentation in Chemistry, Greenwood Press, West Port, Conn.
15. Krishnamurthy, E. V., Sankar, P. V. and Krishnan, S. (1974) : *J. Chem. DOC*, 14 : 130.
16. Krishnan, S., and Krishnamurthy, E. V. (1976) : *Inf. Proc. and Management*, 12 : 19.
17. Krishnamurthy, E. V. : Technical Report, Ind. Inst. of Sci., Bangalore, India.
18. Kier, L. B. and Hall, L. G. (1977) : *Eur. J. Med. Chem.* 12 : 307.
19. Harary, F. (1969) : Graph Theory, Addison Werley Reading, Mass.
20. Busakker, R. G. and Saaty, T. L., (1965) : Finite Graphs and Networks, McGraw-Hill, New York.

Some Aspects of Acetylcholinesterases In Insects

K. N. Mehrotra

Division of Entomology, Indian Agricultural Research Institute, New Delhi-110 012

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

Acetylcholinesterase (AChE, 3.1.1.7) belongs to a very large group of enzymes called hydrolases. They catalyze the reaction :

$$(CH_3)_3\overset{+}{N}CH_2CH_2COOH \xrightarrow{H_2O} (CH_3)_3\overset{+}{N}CH_2CH_2OH + CH_3COOH$$

Because of the probable importance of AChE in nerve transmission, it has been extensively studied in various organisms[2,18,37,48]. AChE obtained from insects has been of special interest not only because of its importance in nerve transmission but also because of its role in organophosphorus and carbamate poisoning. In insects, its occurrence has been demonstrated in embryonic, post-embryonic developmental stages and adults, biochemically[5,6,28,35,50,51] and histochemically by electronmicroscopy[4,11,46,55,56,62]. It is mainly localised in the nervous tissues, specially in the neuropile region where large numbers of synapses are present and also along the axon membranes. It is also present in the glial folds[11,56]. The concentration of AChE activity in different insect species differs by the order of 20-fold or more, suggesting that various insect species, belonging to different orders of insects, show wide variations in total concentration of AChE activity. Highest activity has been reported from the heads of housefly[5,37] and lowest in many lepidopterous larvae[30,51]. It is now well documented that the transmission of nerve impulse in the central nervous system of insects is cholinergic in nature and that AChE plays a vital role in it[49]. Neuromuscular transmission, however, is not cholinergic in nature. All efforts to detect AChE in the muscles of insects have proved to be futile. Harlow[17] showed the insensitivity of the *Locusta migratoria* nerve-muscle preparation to anticholinesterases. These findings were confirmed by O'Connor *et al.*[41] and Faeder *et al.*[12] Histochemically also AChE could not be detected in the neuromuscular junction of *Rhodnius*[61]. One must, however, mention a report by Booth and Lee[4] who found the presence of AChE in virtually all the muscles of the body in the house cricket, *Acheta domesticus* but they also reported the absence of AChE from neuromuscular junctions of *Periplaneta americana*, *Oncopeltus fasciatus*, *Protoparce sexta* and *Apis mellifera*. With the exception of *A. domesticus*, it appears that cholinergic system must be operative only in the central nervous system of an insect. The evidence is accumulating, however, to suggest that glutamate may act as neuromuscular transmitter in insects[26].

AChE in insects is a family of enzymes with overlapping properties. Much of the emphasis on research on AChE has been directed to search

for the differnces between insect AChE and the comparable enzyme in mammal so that the information could be fruitfully utilized for synthesizing selective insecticides. The insect AChE is able to hydrolyze choline esters like acetylcholine butyrylcholine and acetyl beta-methylcholine but show considerable species variation with respect to substrate specificity, including their ability to hydrolyze non-choline esters. Comparative studies on AChE from different insect sources show that the differences in substrate specificity of the enzyme are sufficient to indicate towards the feasibility of synthesis of organophosphorus and carbamate anticholinesterases showing high selective toxicity[5,6,15,36].

The main features of insect AChE may be summarised here. The enzyme obtained from various insect sources is capable of hydrolyzing esters of choline, but it also invariably hydrolyzes butyrylcholine. This is in contrast to the AChE from vertebrate sources where butyrylcholine is not hydrolyzed by the enzyme. All purified preparations of AChE from housefly[10,60], honeybee, *Apis mellifera*[25], *Blatella germanica*[25] show a certain degree of butyrylcholinesterase activity. Indeed, this is one of the main characteristics of the insect AChE. AChE obtained from aphids appears to be of completely different type. AChE from pear aphid, *Macrosiphum pisi* did not hydrolyze acetyl beta-methylcholine and also did not show inhibition by excess substrate[5]. Later work by Sakai[45] showed that AChE from the pear aphid, *Toxoptora piricola* is similar to that reported by Casida[5] and also hydrolyzed butyrylcholine faster than acetylcholine. Further, Smissaert[53] and Zahavi *et al.*[64] reported that the AChE obtained from *Myzus persicae* was inhibited by dithio-bis-(2-nitrobenzoic acid) and *p*-choloromercuribenzoic acid.

These differences in the catalytic properties are most likely due to the configuration of amino acids at the active site of enzyme because it has been shown by O'Brien[38] that the distance between the anionic and esteratic site in the enzyme from mammals and housefly differs tremendously. By his elegent studies, he showed that the distance between anionic and esteratic site was less than 4·5 Å in mammalian AChE, but in case of AChE from housefly the distance was greater than 4·5 Å but not more than 5·4 Å. Comparative studies on the AChE obtained from rat, spider mite and houseflies showed that there are substantial differences at the anionic sites of the enzyme. The rat and housefly enzyme could hydrolyze acetylcholine as well as its dimethyl-alkyl and diethyl-alkyl analogues, but could not hydrolyze di-*n*-propyl-alkyl and di-*n*-butyl-alkyl analogues. The mite enzyme, however, was

able to hydrolyze all these substrates equally well. These results indicated that for housefly and rat enzymes the anionic site is able to accommodate groups which do not exceed 8·5 Å in diameter. In contrast to this the mite AChE is able to accommodate bigger groups having a diameter of more than 8·4 Å[8,31]. Furthermore, evidences have also been presented by Hellenbrand and Krupka[19,20] to suggest that mechanism of reaction of the enzyme obtained from houseflies show substantial differences in its kinetic properties when compared to that of mammals. Insect enzyme differs in having hydrophobic sites at or near the active sites, which affect the rate of reaction with substrates profoundly.

Since insects have been exposed to anticholinesterase pesticides now for quite some time, they have developed resistance to both organophosphate as well as carbamate insecticides. Selections with anticholinesterases are 'short time evolutionary process'. It is, therefore, not surprising that selections with esterase inhibitors have resulted in a change of the esterase specificity pattern and their catalytic properties. The first clear-cut evidence of the change in AChE of organophosphate resistant spider mite *Tetranychus urtice* was provided by Smissaert[52], who observed a decreased sensitivity to organophosphorus compounds. His results suggested that the deacylation was most probably the rate limiting in AChE hydrolysis. It was also suggested that the AChE from resistant mites had a slightly different position of imidazole residue relative to the serine hydroxyl involved in the reaction. Similar changes in AChE have also been observed in *Boophilus microplus*[24,44,47], *Lucilia cuprina*[48], *Ceratitis capitata*[03], *Nephotetix cincticeps*[16,32], *Musca domestica* [59] and *Anopheles albimanus*[8].

The mutant AChE from housefly resistant to tetrachlorvinphos has been extensively studied so as to elucidate the molecular basis for the insensitivity of the mutant enzyme to inhibitors. The mutant enzyme did not differ significantly from the normal strain as far as its affinity to acetylthiocholine was concerned. The mutant enzyme apparently also had a high turnover rate [59]. A comparison of the substrate specificity of the mutant AChE and normal AChE revealed that they differ very little except that the mutant enzyme had a very little affinity to butyrylcholine. The mutant enzyme, however, showed tremendous insensitivity to tetrachlorvinphos and other organophosphates and carbamates. The insensitivity of the mutant enzyme to organophosphates was due to the fact that the mutant enzyme had a far lower affinity to the insecticide. Amino acid analysis of the purified normal

and 'mutant' AChE showed large changes in the amino acid composition of the two enzymes. The mutant enzyme was deficient in about seven molecules of phenylalanine and ten molecules of tyrosine per subunit of enzyme. Further, it had about eight extra molecules of glutamate. The tryptic hydrolysates showed a single extra peptide in the mutant enzyme, when compared to normal enzyme[40].

On the basis of reaction with anticholinesterases also the AChEs from insects differ tremendously from their counterpart in mammals and also among themselves. Support for this view could be found from the studies of O'Brien[86]; Dauterman and O'Brien[9]; Metcalf and Frederickson[85]; Mehrotra and Phokela[84]; Mehrotra and Lal[58]; Hellenbrand[16]; O'Brien[89]; Dale and Mehrotra[7] and Mehrotra[29]. This suggests that the AChE obtained from insects must be differing significantly at the active catalytic sites from vertebrate sources and also among insects themselves.

Evidence has been accumulating in recent years to show that the anticholinesterase compounds (organophosphates and carbamates) react with cholinesterase in a way precisely analogous to its natural substrate acetylcholine[1]. The reactions of cholinesterase with its substrate acetylcholine, an organophosphate and carbamate are :

$$EH + ACh \underset{k-1}{\overset{k_1}{\rightleftharpoons}} [EH.ACh] \xrightarrow[Ch]{k_2} EA \xrightarrow{k_3} EH + AOH \quad -(1)$$

$$EH + POX \underset{k-1}{\overset{k_1}{\rightleftharpoons}} [EH.POX] \xrightarrow[HOX]{k_2} EP \xrightarrow{k_3} EH + POH \quad -(2)$$

$$EH + COX \underset{k-1}{\overset{k_1}{\rightleftharpoons}} [EH.COX] \xrightarrow[HOX]{k_2} EC \xrightarrow{k_3} EH + COH \quad -(3)$$

where EH is free enzyme, ACh is acetylcholine, POX an organophosphate, and COX a carbamate, EA, EP and EC are acetylated, phosphorylated and

carbamylated enzymes respectively. k_1 is the second-order constant for the formation of Michaelis-complex and k_{-1} and k_2 are the first-order constants for the breakdown of Michaelis-complex of the enzyme with acetylcholine, organophosphate and carbamate respectively. The constant k_3 is the first-order constant for the hydrolytic breakdown of the acetylated, phosphorylated or carbamylated enzyme.

From the above, therefore, it would be seen that the acetylcholine, organophosphates and carbamates react with AChE in an identical way. In all these cases, firstly, there is Michaelis-complex formation, which can be easily measured by *Ka* or affinity constant, k_{-1}/k_1 ; secondly, acetylation, phosphorylation and carbamylation (k_2) of enzyme at the active catalytic site, presumably at the serine OH group ; and thirdly hydrolysis (k_3) or deacetylation, dephosphorylation or decarbamylation of the enzyme to yield the original enzyme. In all the three cases there is a leaving group, we may call HOX (e.g., choline in case of acetylcholine, *p*-nitrophenol in case of paraoxon and 1-naphthol in case of carbaryl) which is given off in the second step. For all three groups of compounds, we now have values k_{-1}/k_1 (which is also termed as *Ka*) and for k_2 and k_3 (Table 1). From the table it would be seen that *Ka* is always very small so that the formation of Michaelis-complex is favoured. The values of *Ka* are $7{\cdot}9\times10^{-4}$ M for acetylcholine[1] ; 10^{-3}–10^{-6} M for various organophosphates and carbamates. For majority of carbamates, however, the values are lower than that for organophosphates. The rate of acetylation, phosphorylation and carbamylation, k_2, however, differ significantly ; with acetylcholine the value of k_2 is around 3×10^{5} min^{-1}, with organophosphates, k_2 is moderately fast, i.e., anywhere between 4 and 130 min^{-1}, depending upon the compound and enzyme source, with carbamates the k_2 values are generally very low, i.e., between 0·2 and 5·0 min^{-1}. Similarly k_3 values are roughly equal to k_2 values in case of acetylcholine. With organophosphates, k_3 is practically zero, consequently phosphorylated enzyme persists. With carbamates, although k_3 is low, yet it is still significant and its value is between 0·01 and 0·4 min^{-1}. These values of k_2 and k_3 in the case of carbamates thus lead to the accumulation of small levels of (EHCOX) and a large level of EC or carbamylated enzyme under normal condition. The overall reaction takes place extremely rapidly and it has been estimated that the turnover of ACh may be as high as 750 m moles ACh hydrolyzed per mg protein per hour. The rate of reaction with organophosphates and carbamates, however, are

comparatively slower[1]. It may, however, be added that the scheme presented above is a very elementary representation of a likely mechanism of its action. The enzyme probably is far more complex and may have a number of extra binding sites which affect its reaction with anticholinesterases[40,48]. Further, it may also be stated that the steric effects generated by organophosphates or carbamates may also influence inhibitory potency either by binding on the enzyme active site or at the hydrophobic sites[21].

In conclusion, it may be stated that AChE is a much more complex enzyme than hitherto thought. It differs in its properties from one insect species to another. Out of nearly a million described species of insects so far, AChE has been studied extensively in only about half a dozen species. Even with these limited studies, it appears that it differs in its histological and subcellular distribution and in its kinetic and physical properties. Further, it now appears that it occurs in multimolecular forms[32,40,49,58,59]. These isoenzymes differ in their kinetic properties and their reaction to anticholinesterases[57]. It appears that we have just begun to explore the differences between the enzymes from various sources. A further intensive study on the enzyme obtained from insect sources may reveal yet unknown differences. This information, when available, will be of great importance not only in better understanding of the biochemistry and physiology of the nervous system but may also provide better insight for the synthesis of selective pesticides.

Unit	$Ka(k_1/k)$ M	k_2 min^{-1}	k_3 min^{-1}
ACh	$10^{-3}-10^{-4}$	3×10^{5}	3×10^{5}
Organophosphates	$10^{-3}-10^{-6}$	4 – 130	practically 0
Carbamates	$10^{-3}-10^{-6}$	0·2 – 5	0·01 – 0·4

Table 1

Comparative values of the parameters of esterases for various types of compounds

The values given above have been compiled from Aldridge and Reiner[1].

REFERENCES

1. Aldridge, W. N. and Reiner, E. (1972) 'Enzyme Inhibitors as substrates', North Holland, Amsterdam.
2. Augustinsson, K. B. (1948) *Acta. Physiol. Scand.*, 15 (suppl. 52) : 1.
3. Ayad, H. and Georghiou, G. P. (1975) *J. Econ. Entomol.*, 68 : 295.
4. Booth, G. M. and Lee, A. H. (1971) *Bull. WHO*, 44 : 91.
5. Casida, J. E. (1955) *Biochem. J.*, 60 : 487.
6. Chadwick, L. E. (1963) *In* 'Handbuch der experimentallen pharmacologie', Vol. 15, edited by G. B. Koelle, Springer-Verlag, Berlin.
7. Dale, D. and Mehrotra, K. N. (1976) *Entomon.*, 1 : 39.
8. Dauterman, W. C. and Mehrotra, K. N. (1963) *J. Insect Physiol.*, 9 : 257.
9. Dauterman, W. C. and O'Brien, R. D. (1964) *J. Ag. Food Chem.*, 12 : 318.
10. Dauterman, W. C., Talens, A. and Van Asperen, K. (1962) *J. Insect Physiol.*, 8 : 1.
11. Dhari, K., Prakash, N. and Mehrotra, K. N. (1976) *Indian J. Exp. Biol.*, 14 : 544.
12. Faeder, I. R., O'Brien, R. O. and Salpeter, M. M. (1970) *J. Exp. Zool.*, 173 : 187.
13. Froede, H. C. and Wilson, I. B. (1971) *In* The Enzymes, Vol. V, Ed. P. D. Boyer. Third edition, Acad. Press, New York and London.
14. Fukuto, T. R. (1978) In Neurotoxicology of Insecticides and Pheromones. Ed. T. Narahashi, Plenum Press, New York.
15. Guiltbault, G. G., Kuan, S. S. and Sadar, M. H. (1970) *J. Ag. Food Chem.*, 18 : 692.
16. Hama, H. and Iwata, T. (1971) *Appl. Entomol. Zool.*, 6 : 183.
17. Harlow, P. A. (1958) *Ann. Appl. Biol.*, 46 : 55.
18. Hellenbrand, K. (1967) *J. Ag. Food Chem.*, 15 : 825.
19. Hellenbrand, K. and Krupka, R. M. (1970) *Biochem.*, 9 : 4665.
20. Hellenbrand, K. and Krupka, R. M. (1974) *Comp. Biochem. Physiol.*, 47-B : 271.
21. Hollingworth, R. M. (1975) *In* 'Pesticide Selectivity'. Ed. J. C. Street, Dekker, New York.
22. Iwata, T. and Hama, H. (1972) *J. Econ. Entomol.*, 65 : 819.
23. Kunkee, R. E. and Zweig, G. (1963) *J. Insect Physiol.*, 9 : 495.
24. Lee, R. M. and Batham, P. (1966) *Entomol. Exp. Appl.*, 9 : 13.
25. Lord, K. A. (1961) *Biochem. J.*, 78 : 483.
26. Lunt, G. G. (1973) *Comp. gen. Pharm.*, 4 : 75.
27. Main, A. R. (1964) *Science*, 144 : 992.
28. Mehrotra, K. N. (1960) *J. Insect Physiol.*, 5 : 129.
29. Mehrotra, K. N. (1978) *In* Improving Crop and Animal Productivity. Ed. Sethi, G. R. and Chathrath, M. S., Oxford and IBH Publishers, New Delhi.
30. Mehrotra, K. N. and Chandra, D. (1974) *Indian J. Expt. Biol.*, 12 : 457.
31. Mehrotra, K. N. and Dauterman, W. C. (1963) *J. Neurochem.*, 10 : 119.
32. Mehrotra, K. N. and Farooqui, T. N. A. (1977) *Indian J. Expt. Biol.*, 15 : 1229.
33. Mehrotra, K. N. and Lal, S. (1970) *Indian J. Biochem.*, 7 : 167.

34. Mehrotra, K. N. and Phokela, A. (1969) *Arch. Internat. Physiol. Biochem.*, 77 : 799.
35. Metcalf, R. L. and Frederickson, H. (1965) *J. Econ. Entomol.*, 58 : 143.
36. Metcalf, R. L., March, R. B. and Maxon, M. G. (1955) *Ann. Ent. Soc. Am.*, 48 : 222.
37. Nachmansohn, D. and Wilson, I. B. (1951) *Advan. Enzymol.*, 12 : 25.
38. O'Brien, R. D. (1963) *J. Ag. Food Chem.*, 11 : 163.
39. O'Brien, R. D. (1967) Insecticides—Action and Metabolism, Acad. Press, New York.
40. O'Brien, R. D. (1979) *In* 'Advances in Pesticide Science'. Ed. Geissbuchler, H., Brooks, G. T. and Kearney, P. C., Pergmon Press, Oxford.
41. O'Connor, A. K., O'Brien, R. D. and Salpeter, M. M. (1965) *J. Insect Physiol.*, 11 : 1351.
42. Plapp, F. W. (1976) *Ann. Rev. Entomol.*, 21 : 179.
43. Rosenberry, T. L. (1975) *Advan. Enzymol.*, 43 : 103.
44. Roulston, W. J., Schnitzerling, H. J. and Schuntner, C. A. (1968) *Aust. J. Biol. Sci.*, 21 : 759.
45. Sakai, M. (1967) *Appl. Entomol. Zool.*, 2 : 111.
46. Salkeld, E. H. (1964) *Can. Ent.*, 96 : 389.
47. Schuntner, C. A. and Roulston, W. J. (1968) *Aust. J. Biol. Sci.*, 21 : 173.
48. Schuntner, C. A., Roulston, W. J. and Schnitzerling, H. J. (1968) *Aust. J. Biol. Sci.*, 21 : 97.
49. Silver, A. (1974) 'Biology of cholinesterases'. American Elsevier Publication, New York.
50. Singh, H. N. and Mehrotra, K. N. (1970) *J. Insect Physiol.*, 16 : 2385.
51. Smallman, B. N. and Mansingh, A. (1969) *Ann. Rev. Entomol.*, 14 : 387.
52. Smissaert, H. R. (1964) *Science*, 143 : 129.
53. Smissaert, H. R. (1976) *Pesticide Biochem. Physiol.*, 6 : 215.
54. Smissaert, H. R., Voerman, S., Oostenbrugge, L. and Renooy, N. (1970). *J. Ag. Food Chem.*, 18 : 63.
55. Smith, D. S. and Treherne, J. E. (1963) *Adv. Insect Physiol.*, 1 : 401.
56. Smith, D. S. and Treherne, J. E. (1965) *J. Cell Biol.*, 20 : 445.
57. Steele, R. W. and Maneckjee, A. (1979) *Pestic. Biochem. Physiol.*, 10 : 322.
58. Steele, R. W. and Smallman, B. N. (1976) *Biochim. Biophys. Acta.*, 445 : 131.
59. Tripathi, R. K. and O'Brien, R. D. (1973) *Pesticide Biochem. Physiol.* 3 : 495.
60. Tripathi, R. K. and O'Brien, R. D. (1977) *Biochim. Biophys. Acta.*, 480 : 382.
61. Wigglesworth, V. B. (1958) *Quart. J. Micros. Sci.*, 99 : 441.

62. Winton, M. Y., Metcalf, R. L. and Fukuto, T. R. (1958) *Ann. Ent. Soc. Am.*, **51** : 436.

63. Zahavi, M. and Tahori, A. S. (1970) *Isr. J. Entomol.*, **5** : 185.

64. Zahavi, M., Tahori, A. S. and Klimer, F. (1972) *Biochim. Biophys Acta.*, **276** : 577.

C_4-Photosynthesis : An Adaptation to Tropical Environments

V. S. Rama Das and I. Madhusudana Rao*

Centre for Photosynthetic Studies
School of Life Sciences
University of Hyderabad
Hyderabad 500 001

*Present address of I. M. R. : Pulse Physiology
International Crops Research Institute
Patancheru P.O.
Andhra Pradesh-502 324, India

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

1. Introduction

Photosynthesis is the most important biochemical process on earth. The improved photosynthetic efficiency holds the key to the future adequacy of our food supplies. Green plants are the primary harvesters of free solar energy. They are biological sun traps[29]. They are net producers of food and energy on a renewable basis[18,29].

The discovery of the C_4-pathway of photosynthesis has undoubtedly stimulated renewed interest in research on the carbon metabolism of photosynthesis[33,43]. This has led to the concept that, as well as the classical Calvin cycle of photosynthesis, there are other metabolic pathways, that allow plants to harvest inorganic carbon from the environment[4,12,31,38].

Since the discovery of C_4-photosynthesis, higher plants have been categorised into C_3-plants, which form 3-phosphoglycerate (PGA) as the first labelled intermediate ; C_4-plants, which form 4-carbon acids as predominant early photosynthetic products, and CAM (Crassulacean Acid Metabolism) plants. At first sight, this suggests that there are fundamental alternative pathways for photosynthetic carbon fixation. However, it is now generally accepted that, in all photoautotrophs, the Calvin cycle (reductive pentose phosphate cycle, C_3-pathway) is the mechanism of light driven CO_2 reduction.

Both the C_4-pathway of photosynthesis, which was originally believed to represent an alternative to the Calvin cycle[32] and CAM are now understood to be auxiliary mechanisms that are linked to the Calvin cycle in order to facilitate the reductive carbon assimilation under less favourable ecological conditions[3,4].

2. Occurrence of C_4-Photosynthesis

One of the primary approaches has been to screen the plant kingdom for the occurrence of C_4-photosynthesis. The first compilation of C_4-plants was made by Downton[19] who later revised it. The revised list of C_4-plants[20] consisted of 485 species of angiosperms in 117 genera belonging to 13 families. C_4-photosynthesis was thought to operate only in herbs and shrubs till Pearcy and Troughton[50] described several tree *Euphorbia* sps. with C_4-pathway. C_4-photosynthesis is now (till April 1977) known to occur in 943 species, 196 genera belonging to 18 families[60]. The families having C_4-plants are

Acanthaceae, *Aizoaceae*, *Amaranthaceae*, *Asclepiadaceae*, *Asteraceae*, *Boraginaceae*, *Capparidaceae*, *Caryophyllaceae*, *Chenopodiaceae*, *Euphorbiaceae*, *Nyctaginaceae*, *Polygalaceae*, *Portulacaceae*, *Scropulariaceae*, and *Zygophyllaceae* (in dicotyledons) ; *Cyperaceae*, *Liliaceae* and *Poaceae* (in monocotyledons).

The following genera are known to have both C_4- and C_3-species : *Aerva*, *Alloteropsis*, *Alternanthera*, *Atriplex*, *Bassia*, *Boerhaavia*, *Chamaesyce*, *Cyperus*, *Euphorbia*, *Flaveria*, *Heliotropium*, *Kallstromia*, *Kochia*, *Mollugo*, *Panicum*, *Scirpus*, *Suaeda* and *Zygophyllum*. Both C_3- and C_4-ecotypes have been noticed to exist within the same species of *Alloteropsis semialata*[23] and *Mollugo verticillata*[72]. Although several C_4- and C_3-species of *Boerhaavia* are reported, Das and Santakumari[16] feel that there are only three valid species of *Boerhaavia*, namely *B. diffusa* (C_4), *B. erecta* (C_4) and *B. Chinensis* (C_3). Based on the photosynthetic characteristics and geographical distribution of the species, they also feel that the C_4-pathway is of comparatively recent origin and may have evolved independently in *B. diffusa* and *B. erecta*.

3. Criteria Used to Detect C_4-Plants

Kranz leaf anatomy is the best known example, showing a chlorenchymatous bundle sheath (usually staining more intensely for starch than mesophyll) and radiating mesophyll tissue. Other characteristics of C_4-plants include a low CO_2 compensation point, no effect of low oxygen on photosynthetic rates, C_4-dicarboxylic acids as the initial products of CO_2 fixation and a lower discrimination against ^{13}C than C_3-plants as shown in Table 1. Hattersley and Watson[25] recently demonstrated that C_4-plants can also be distinguished from C_3-plants by their 'maximum lateral cell count' (2 to 4 chlorenchymatous mesophyll cells intervening between parenchymatous bundle sheath cells of laterally adjacent vascular bundles in C_4-species : more than 4 in C_3-species) and the 'maximum cell distant count' (in C_4-species no chlorenchymatous mesophyll cell is separated from the nearest parenchymatous bundle sheath call by more than one other chlorenchymatous mesophyll cell, i.e., one cell distant criterion).

4. Variations from the Kranz Syndrome

Unconventional kranz anatomy was noticed in leaves of *Arundinella hirta*[11], *Suaeda monoica*[73] and *Glossocordia boswallaea*[17]. The basic pattern of kranz anatomy is maintained both in *S. monoica* and *G. boswallaea* since

they clearly exhibit the three-tire system of innermost vascular bundles surrounded by chlorenchymatous bundle sheath (staining for starch) and outermost mesophyll cells. On the other hand *Panicum hians* and *P. milioides* are C_3-plants with kranz anatomy[5,39,42].

Characteristics	PLANTS		
	C_4	C_3	CAM
(a) Initial fixing enzyme	PEP carboxylase	RuBP carboxylase	both
(b) Kranz anatomy	present	absent	absent
(c) Succulence	no	yes and no	yes
(d) CO_2 compensation point	10 ppm	30–150 ppm	?
(e) Light saturation	full sunlight	40 Klux	10 Klux
(f) Optimum temperature for CO_2 fixation	30–35 °C	15–20 °C	35 °C
(g) Photosynthetic rate (mg CO_2 dm^{-2} hr^{-1})	40–80	15–35	0·5 – 11·0
(h) Maximum growth rate (g dry wt. m^{-2} day^{-1})	4–5	0·5–2	0·015–0·018
(i) Transpiration ratio (g H_2O used for g dry wt. produced)	250–350	450–950	50–125
(j) Photorespiration	very low	active	very low
(k) $^{13}C/^{12}C$ ratio	−9 to −18	−23 to −36	−9 to −36

Table 1
Some distinguishing characteristics of three types of photosynthetic plants

Raghavendra *et al.*[63] reported that a single plant of *Mollugo nudicaulis* may have some leaves with C_3- characteristics as well as some with C_4-characteristics according to their position on the stem. Their observations indicate that there is an imbalance in the development of C_3- and C_4-systems in *M. nudicaulis* leaves. The C_3-component forms first (in younger leaves) and C_4-system gradually develops later (in older leaves).

5. Carbon Metabolism of C_4-Plants

C_4-plants preferentially metabolise either malate or aspartate as the principal C_4 donor acid which is rapidly transported into the bundle sheath and decarboxylated to CO_2[4,12,19]. During steady state photosynthesis in

air, labelled with $^{14}CO_2$, and C_4-acids oxaloacetate, malate and aspartate appear both as initial products and as primary intermediates of the pathway. Malate is formed by the leaves of sugarcane, maize and *Sorghum*, but aspartate is labelled predominantly by leaves of *Amaranthus*, *Atriplex* and *Cynodon*[7,19,30]. An unique feature of an exclusive formation of aspartate during short-term photosynthesis in Elensine coracana was observed by Das and Rathnam[14]. Based on corresponding decarboxylating enzyme, C_4-plants are being classified into NADP-malic enzyme type, PEP-carboxykinase type and NAD-malic enzyme type[28,31,55].

(a) *Movement of labelled photosynthetic intermediates between mesophyll and bundle sheath*

A rapid movement of metabolites is implied in explanations of the C_4-pathway[32]. The distribution of ^{14}C labelled compounds in mesophyll and bundle sheath cells was studied in several C_4-plants[21,33,56]. The light dependence of label movement was reported for sugarcane and *Eleusine coracana*[32,70].

(b) *Possible involvement of ATPase in the movement of C_4-acids*

Karpilov *et al.*[40] indicated the necessity for an active transport of C_4-acids and also concluded that an ATPase might be participating in it. Raghavendra and Das[66] suggested that an ATPase may be participating in active transferred of C_4-acids from the mesophyll into the bundle sheath. The location of the most of the Na^+-K^+ ATPase activity in bundle sheath cells of both *Setaria italica* and *Amaranthus paniculatus* strongly supported this suggestion.

(c) *Enzyme location in mysophyll and bundle sheath cells*

Data now available for the photosynthetic capabilities and enzyme content of mesophyll and bundle sheath cells separated from a large number of species confirm the essentially absolute partitioning of PEP carboxylase in mesophyll cells and RuBP carboxylase, together with Calvin cycle enzymes, in bundle sheath cells[31]. There is also clear evidence that mesophyll chloroplasts, active in several light-dependent transformation of carbon compounds, totally lack RuBP carboxylase and Calvin cycle activity[28,60].

PEP carboxylase in leaves of *Eleusine coracana* has been shown to be partly associated with the envelope membranes of the mesophyll chloroplasts[70]. This evidence indicate that PEP carboxylase in leaves of C_4-plants may be a membrane-associated enzyme[69]. Aspartate, malate, oxaloacetate and malonate were found to be strong allosteric inhibitors of PEP carboxylase[36,45,52,55]. Recently, Rathnam[66] proposed that the high-temperature tolerance of C_4-plants is not only due to the presence of a heat resistant protein synthesizing machinery[46], but also, at least in part, due to the greater thermostability of their primary carboxylating enzyme.

Radiotracer studies indicate that oxaloacetate, formed via PEP carboxylase in mesophyll cells, is rapidly converted into large pools of malate and aspartate. These transformations are attributed to NADP malate dehydrogenase and aspartate aminotransferase, respectively, both being freely reversible reactions[31,60]. In fact, it is now clear that the C_4-acid moving to bundle sheath cells, and the metabolic steps leading to its decarboxylation, differ among C_4-plants. These differences provide the basis for grouping of C_4-plants. The nomenclature used to define these subgroups, 'NADP-ME-type', 'PCK-type' and 'NAD-ME-type', is that suggested by Hatch *et al.*[34]

The metabolism of PGA in both the mesophyll and bundle sheath was indicated by the recovery of NADP triose phosphate dehydrogenase and triose phosphate isomerase. The occurrence of large number of mitochondria in the bundle sheath cells indicates high activities of NAD malic dehydrogenase and cytochrome oxidase with bundle sheath fractions[60].

Rathnam and Das[67] studied the intercellular and intracellular distribution of nitrate assimilating enzymes. Nitrate reductase was found to be localized on the chloroplast envelope membranes. The chloroplastic NADPH-glutamate dehydrogenase was concentrated in the mesophyll cells. The extrachloroplastic NADH-glutamate dehydrogenase was localized in the bundle sheath cells. Based on the results, Rathnam and Das[67] proposed a scheme to account for the nitrate metabolism in the leaves of *Elusine coracana* in relation to its aspartate-type C_4-pathway of photosynthesis.

Recently, Venkataramana and Das[74] reported the predominant localization of photoactive ATP dependent glutamine synthetase in the mesophyll chloroplasts of *Setaria italica*. The localization of leaf glutamine synthetase was thought to be an advantage to the plant from compartmentalization of nitrogen metabolism, particularly in the chloroplasts where it allows the

process to be closely linked to the provision of energy and to carbon fixation.

(d) *Carbon fixation by isolated chloroplasts*

The information on carbon fixation by isolated chloroplasts from C_4-plants was limited and offered conflicting evidences. Baldry *et al.* (1969) observed light stimulated PEP dependent carbon fixation by isolated chloroplasts of sugarcane, later reports described the functioning of only Calvin cycle in maize chloroplasts[36,40]. The chloroplasts from leaves of C_4-plants unresolved into mesophyll and bundle sheath chloroplasts, fixed most of $H^{14}CO_3$ into Calvin cycle intermediates but only a minor part of the labile into C_4-acids. However, the several fold stimulation of carbon fixation by the mesophyll chloroplasts by PEP of pyruvate+ATP indicated the functional PEP carboxylation in them[61,64,65]. The positive response of carbon fixation by the bundle sheath chloroplasts to RuBP and ribose-5-phosphate +ATP is also an evidence for the operation of Calvin cycle in the bundle sheath cells[61].

PEP carboxylase is reported to be either loosely associated with the mesophyll chloroplasts[38] or located in cytoplasma of the mesophyll cells[38]. Raghavendra and Das[61] reported that CO_2 fixation in the mesophyll cells via B-carboxylation is dependent on both chloroplasts and extrachloroplastic PEP carboxylase. The argument that Calvin cycle operates in the mesophyll chloroplasts and only starch synthesis takes place in the bundle sheath chloroplasts[10] was overruled by Raghavendra and Das[61].

(e) *Photochemical characteristics of mesophyll and bundle sheath chloroplasts*

More extensive recent surveys of photochemical properties of chloroplasts isolated from C_4-plants confirm the photosystem II deficiency in the bundle sheath cells of NADP-ME-type species[41,47,62]. However, plants of the NAD-ME-type and PCK-type show normal granal development in both mesophyll and bundle sheath cells[28,34]. A decline in PS II activity of bundle sheath chloroplasts during leaf ontogeny was observed by Raghavendra and Das[62].

The existence of a complete photosynthetic electron transfer chain in mesophyll and bundle sheath chloroplasts of NADP-ME as well as NAD-ME C_4-species was now confirmed[62]. The bundle sheath chloroplasts can

provide ATP at significant rates through cyclic electron flow in NADP-ME C_4-plants, while ATP may come from either cyclic or non-cyclic electron flow in the bundle sheath chloroplasts of NADE-ME C_4-plants[1,54,62,71]. High rates of endogenous photophosphorylation (mediated by ferredoxin, a natural cofactor) by the bundle sheath chloroplasts of C_4-plants supports the above idea[58].

The location of 35 to 52% of leaf chlorophyll in the bundle sheath suggests that a considerable photochemical potential exists with the bundle sheath chloroplasts of C_4-plants[62]. Further, it was also indicated that both the mesophyll and bundle sheath chloroplasts of NAD-ME type C_4-plants possess much more photochemical potential than their counterparts in NADE-ME C_4-species[41,51,53,62].

6. Photorespiration

Illuminated leaves of many species have a rapid rate of respiration (usually considered as CO_2 evolution) that occurs by biochemical reactions differing from that of the better understood 'dark' respiration that takes place mostly in the mitochondria. This photorespiration diminishes net CO_2 uptake greatly because some recently fixed products of photosynthesis are oxidized to CO_2. Photorespiration requires light and occurs best at high levels of O_2 in the atmosphere, is inhibited by high concentrations of CO_2, is very temperature dependent, and has a close connection with the synthesis and further oxidation of glycolic acid[4,37,76].

C_4-plants on the other hand, are characterized by their apparent lack of photorespiration[4,37,68]. Thus the presence or absence of detectable photorespiration serves as a very useful criterion in identifying C_3- and C_4-plants[18]. C_4-plants possess photorespiratory activity and that this activity is localized in the bundle sheath cells[9,60,68]. Bundle sheath cells contain a large number of microbodies and enzymes involved in photorespiration, such as RuBP oxygenase, catalase, glycolate oxidase, hydroxypyruvate reductase (NAD^+) and phosphoglycolate phosphatase[8,25,68]. Decreasing photorespiration by inhibiting the synthesis of glycolic acid or altering its further metabolism by biochemical or genetic methods could bring about large increases (at least 50%) in net photosynthesis and dry weight yield in less efficient species[75].

The magnitude and function of photorespiration in C_4-plants remains unknown. The lack of apparent photorespiration may be responsible for

the high photosynthetic capacity of C_4-plants. Recently, Naidu *et al.*[48] reported that the dark respiration rates of C_4-plants were invariably higher (20 to 30%) than the nearest taxonomically related C_3-species.

7. Ecological Implications

C_4-pathway may confer remarkable ecological adaptations in certain plants with respect to their behaviour and survival under extreme climatic conditions. The C_4-pathway enables the plant to cope with the restrictions imposed on photosynthesis by high oxygen and low CO_2 concentration in the atmosphere[2]. Therefore, the adaptive advantage of the C_4-pathway is particularly great under conditions of high irradiance, high thermal load and low CO_2 concentration.

The higher photosynthetic capacity in C_4-leaves undoubtedly gives these plants advantages in certain ecological situations, though the advantage of the C_4-cycle in agronomic crop production has been questioned[24,27]. One advantage of C_4-plants that has been well established is their greater efficiency in water utilisation, i.e., their ability to produce more dry matter per unit of water transpired. This has been substantiated in several recent experiments[15,16,19]. This increased efficiency by C_4-plants results from higher CO_2 fixation, high stomatal resistance and lower transpiration rates in C_4-species when compared to C_3-species.

This advantage in use of water by C_4-species and some xerophytic characteristics of their leaves have led to the speculation that C_4-species evolved under dry climates[44]. The lack of inhibition of photosynthesis by O_2 in C_4-species and their ability to extract CO_2 from atmosphere low in CO_2 (10 ppm) suggests that C_4-photosynthesis may have evolved during geologic periods of elevated atmospheric O_2 content and/or lower CO_2 content. The well-documented adaptation of C_4-species to conditions of high irradiance and high temperature indicated that C_4-plants may have evolved in response to such an environment.

Recently, Brown[6] put forward the hypothesis that C_4-plants have a greater nitrogen use efficiency (biomass production per unit of nitrogen in the plant) than do the C_3-plants. However, this is apparently due to relatively smaller investment of nitrogen in the photosynthetic carboxylation enzymes of C_4-plants than of C_3-plants. The greater nitrogen use efficiency of C_4-plants may give them an adaptive advantage, particularly in low nitrogen sites. It is, therefore, possible that C_4-plants evolved in the dry tropics in

response to a combination of factors, including high temperature, water stress and nitrogen availability.

Ehleringer[22] investigated the implications of a reduced quantum yield (initial slope of the photosynthetic light response curve) in C_4-plants and temperature dependence of quantum yield in C_3-plants on total canopy primary production using computer simulations. He performed the simulations for three biogeographical or environmental conditions : the great plains region of North America, the Sonoron Desert of North America and shade habitats. He concluded that while the C_4-photosynthetic mechanism is undoubtedly a significant advantage for plants in some environments, the increased energy cost of the C_4-pathway can be a significant liability in other habitats.

8. General Conclusions

The discovery of the C_4-pathway of photosynthesis has given us a deeper understanding of the intricacies of photosynthesis. This discovery also has led to the hypothesis that increased productivity might be achieved by manipulating plants to incorporate some of the C_4-characteristics, such as the efficient utilization of lower concentrations of CO_2, the ability to grow under water stress and high salt concentrations, and the ability to use intense light efficiently. The process of photorespiration may decrease yields up to 50 per cent. This loss arises from recycling of the photosynthetically fixed carbon in the plant so as to re-evolve CO_2 which is thus lost from the plant.

A knowledge of C_4-characteristics of plants and photorespiration may eventually be used for somatic and genetic hybridization leading to the selection of efficient photosynthetic plants.

Acknowledgement

The preparation of this article was supported by a grant from the Science and Engineering Research Council, Department of Science and Technology, Government of India, New Delhi.

References

1. Anderson, J. M., Boardman, N. K. and Spencer, D. (1971) *Biochem. Biophys. Acta.*, 245 : 253.
2. Bjorkman, O. (1973) *In* : Current Topics in Photobiology, Photochemistry and Photophysiology, Ed. A. Giese, Vol. VIII, pp. 1-63, New York, Academic Press.

3. Bjorkman, O. (1976) *In* : CO_2 Metabolism and plant productivity, Ed. R. H. Burris and C. C., Black, pp. 287-310, Baltimore - London - Tokyo, University Park Press.
4. Black, C. C. (1973) *Annu. Rev. Plant Physiol.*, 24 : 253.
5. Brown, R. H. and Brown, W. V. (1975) *Crop. Sci.*, 15 : 681.
6. Brown, R. H. (1978) *Crop. Sci.*, 18 : 93.
7. Chen, T. M., Brown, R. H. and Black, C. C. (1971) *Plant Physiol.*, 47 : 199.
8. Chen, T. M., Dittrich, P., Campbell, W. H. and Black, C. C. (1974) *Arch. Biochem. Biophys.*, 163 : 246.
9. Chollet, R. and Ogren, W. L. (1973) *Plant Physiol.*, 51 : 787.
10. Coombs, J. and Baldry, C. W. (1972) *Nature* 238 : 268.
11. Crookston, R. K. and Moss, D. N. (1973) *Plant Physiol.*, 52 : 397.
12. Das, V. S. R. (1974) *Ind. J. Plant Physiol.*, 18 : 1.
13. Das, V. S. R. (1979) *Reg. J. Energy Heat Mass Transfer*, 1 : 291.
14. Das, V. S. R. and Rathnam, C. K. M. (1974) *In* : Mechanisms of Regulation of Plant Growth, Bulletin 12, Ed. R. L. Bieleski, A. R. Ferguson and M. M. Cresswell, Roy. Soc. N.Z., Wellington, New Zealand.
15. Das, V. S. R. and Santakumari, M. (1977) *Plant Cell Physiol.*, 18 : 935.
16. Das, V. S. R. and Santakumari, M. (1978) *Photosynthetica*, 12 : 418.
17. Das, V. S. R. and Raghavendra, A. S. (1976) *Proc. Ind. Acad. Sci.* (B), 84 : 12.
18. Das, V. S. R. and Rajendrudu, G. (1976) *Proc. Ind. Acad. Sci.* (B), 84 : 148.
19. Downton, W. J. S. (1971) *In* : Photosynthesis and photorespiration, Ed. M. D. Hatch, C. B. Osmond and R. O. Slatyer, pp. 3-17, New York, Wiley Interscience.
20. Downton, W. J. S. (1975) *Photosynthetica*, 9 : 96.
21. Edwards, G. E., Huber, S. C., Ku, S. B., Rathnam, C. K. M., Gutierrez, M. and Mayne, B. C. (1976) *In* : CO_2 Metabolism and plant productivity, pp. 83-112, Ed. R. H. Burris, and C. C. Black, Baltimore - London - Tokyo, University Park Press.
22. Ehleringer. J. R. (1978) *Oecologia*, 31 : 255.
23. Ellis, R. P. (1974) *South Afr. J. Sci.*, 70 : 169.
24. Evans, L. T. (1975) *In* : Crop physiology : some case histories, pp. 327-335, Ed. L. T. Evans, London, Cambridge University Press.
25. Frederick, S. E. and Newcomb, E. H. (1971) *Planta*, 96 : 152.
26. Gibbs, M., Latzko, E., O'Neal, D. and Hew, C. S. (1970) *Biochem. Biophys. Res. Commun.*, 40 : 1356.
27. Gifford, R. M. (1974) *Aust. J. Plant Physiol.*, 1 : 107.
28. Gutierrez, M., Gracen, V. E. and Edwards, G. E. (1974) *Planta*, 119 : 279.
29. Hall, D. O. (1979) *Solar Energy*, 22 : 307.
30. Hatch, M. D. (1971) *In* : Photosynthesis and Photorespiration, pp. 139-152, Ed. M. D. Hatch, C. B. Osmond and R. O. Slatyer, New York, Wiley Interscience.
31. Hatch, M. D. (1976) *In* : CO_2 Metabolism and Plant Productivity, pp. 59-81, Ed. R. H. Burris and C. C. Black, Baltimore, University Park Press.
32. Hatch, M. D. and Slack, C. R. (1966) *Biochem. J.*, 106 : 141.
33. Hatch, M. D. and Kagawa, T. (1973) *Arch. Biochem. Biophys.*, 159 : 842.
34. Hatch, M. D., Kagawa, T. and Craig, S. (1975) *Aust. J. Plant Physiol.*, 2 : 111.
35. Hattersley, P. W. and Watson, L. (1975) *Phytomorphology*, 25 : 325.
36. Huber, S. C. and Edwards, G. E. (1975) *Plant Physiol.*, 55 : 835.
37. Jackson, W. A. and Volk, R. J. (1970) *Annu. Rev. Plant Physiol.*, 21 : 385.
38. Kagawa, T. and Hatch, M. D. (1974) *Aust. J. Plant Physiol.*, 1 : 51.

39. Kanai, R. and Kashiwagi, M. (1975) *Plant cell Physiol.*, **16** : 669.
40. Karpilov, Y. S., Bil, K. Y. and Gukasyn, I. A. (1975) *Fiziol. Rast.*, **22** : 1113.
41. Ku, S. B., Gutierrez, M., Kanai, R. and Edwards, G. E. (1974) *Z. Pflanzenphyiol.*, **72** : 320.
42. Ku, S. B., Edwards, G. E. and Kanai, R. (1976) *Plant cell Physiol.*, **17** : 615.
43. Kortschak, H. P., Hartt, C. E. and Burr, G. (1965) *Plant Physiol.*, **40** : 209.
44. Laetsch, W. M. (1969) *Sci. Prog. Oxf.*, **57** : 323.
45. Lowe, J. and Slack, C. R. (1971) *Biochim. Biophys. Acta.*, **235** : 207.
46. Lyons, J. M. and Raison, J. K. (1970) *Plant Physiol.*, **45** : 386.
47. Mayne, B. C., Dec., A. M. and Edwards, G. E. (1974) *Z. Pflanzenphysiol.*, **74** : 275.
48. Naidu, K. R., Rajendrudu, G. and Das, V. S. R. (1980) *Z. Pflanzenphysiol. (in press.)*
49. O'Neal, D., Hew, C. S., Latzko, E. and Gibbs, M. (1972) *Plant Physiol.*, **49** : 607.
50. Pearcy, R. W. and Troughton, J. H. (1974) *Carnegie Inst. Year Book*, **73** : 809.
51. Polya, G. M. and Osmond, C. B. (1972) *Plant Physiol.*, **49** : 267.
52. Raghavendra, A. S. and Das, V. S. R. (1975) *Biochem. Biophys. Res. Commun.*, **66** : 160.
53. Raghavendra, A. S. and Das, V. S. R. (1976a) *Ind. J. Plant Physiol.*, **19** : 101.
54. Raghavendra, A. S. and Das, V. S. R. (1976b) *Photosynthetica*, **10** : 345.
55. Raghavendra, A. S. and Das, V. S. R. (1976c) *Z. Pflanzenphysiol.*, **78** : 434.
56. Raghavendra, A. S. and Das, V. S. R. (1977a) *J. Exp. Bot.*, **28** : 1169.
57. Raghavendra, A. S. and Das, V. S. R. (1977b) *J. Exp. Bot.*, **29** : 39.
58. Raghavendra, A. S. and Das, V. S. R. (1977c) *Ann. Bot.*, **41** : 667.
59. Raghavendra, A. S. and Das, V. S. R. (1978a) *Photosynthetica*, **12** : 200.
60. Raghavendra, A. S. and Das, V. S. R. (1978b) *Z. Pflanzenphysiol.*, **87** : 379.
61. Raghavendra, A. S. and Das, V. S. R. (1978c) *Photosynthetica*, **12** : 166.
62. Raghavendra, A. S. and Das, V. S. R. (1978d) *Physiol. Plant*, **43** : 107.
63. Raghavendra, A. S., Rajendrudu, G. and Das, V. S. R. (1978) *Nature*, **273** : 143.
64. Rajendrudu, G., Madhusudana Rao, I., Raghavendra, A. S., Das, V. S. R. (1979) *Proc. Ind. Acad. Sci.* (B) **88** : 143.
65. Rao, I. M., Rajendrudu, G. and Das, V. S. R. (1978) Photosynthetic ^{14}Carbon fixation in isolated leaf cells and mesophyll protoplasts of C_3, C_4 and CAM plants. Paper presented at the International Symposium on 'Biological Applications of solar energy (BASE)' held at Madurai, India.
66. Rathnam, C. K. M. (1978) *Planta*, **141** : 289.
67. Rathnam, C. K. M. and Das, V. S. R. (1974a) *Can. J. Bot.*, **52** : 2599.
68. Rathnam, C. K. M. and Das, V. S. R. (1974b) *J. Biochem. Biophys.*, **11** : 303.
69. Rathnam, C. K. M. and Das, V. S. R. (1975a) *Z. Pflanzenphysiol.*, **74** : 377.
70. Rathnam, C. K. M. and Das, V. S. R. (1975b) *Z. Pflanzenphysiol.*, **75** : 360.
71. Rathnam, C. K. M. and Das, V. S. R. (1976) *Aust. J. Plant Physiol.*, **3** : 185.
72. Sayre, R. T. and Kennedy, R. A. (1977) *Planta*, **134** : 257.
73. Shomerz-ilan, A., Beer, S. and Waisal, Y. (1975) *Plant Physiol.*, **56** : 676.
74. Venkataramana, S. and Das, V. S. R. (1979) *Z. Naturforsch.*, **34** : 210.
75. Zelitch, I. (1971) 'Photosynthesis, Photorespiration and Plant Productivity', New York-London : Academic Press.
76. Zelitch, I. (1975) *Annu. Rev. Biochem.*, **44** : 123.

Stability of Parallel Flow Through a Porous Medium Using Brinkmann Model

N. Rudraiah

Department of Mathematics, Central College,
Bangalore University, Bangalore-560001, India

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

1. Introduction

Flow through porous media has attracted considerable research activity in recent years because of its potential applications in soil physics, geohydrology, filtration of solids from liquids, chemical engineering and in biological systems. Much interest has also been evinced in wetting and dewetting of solids by liquids which involve the movement of fluid over the solid particles. The stability of this movement is of immense use in petroleum industry. This is because, the behaviour of fluid in petroleum reservoir rock depends to a large extent on the property of the rock and the stability of this fluid that yields new or additional information on the characteristics of the rock would contribute to a better understanding of the petroleum reservoir performance. Furthermore, the friction force offered by the solid particles to the fluid may reduce its drag in turbulent flows. Therefore, stability of flow through porous media may provide a convenient means of experimentally demonstrating nonlinear effects (i.e., turbulent phenomenon). In ordinary Poiseuille flow it is necessary to use extremely thin fluid to detect this phenomenon[1], but in porous media the friction force is much larger, so the depth of the flow can be greatly increased.

The stability of parallel flow with boundaries in the absence of porous media is a classical problem associated with the name of stability of Poiseuille flow. The copious literature on this problem and its extensions has been reviewed many times, notably by Lin[9], Stuart[14], Drazin and Howard[4], Betchov and Criminale[2], Joseph[8]. The stability of plane Poiseuille flow through porous media has not been given much attention and is investigated in this paper using the Brinkmann model.

When a fluid permeates a porous medium, the actual path of an individual particle of the fluid cannot be followed analytically. The gross effect, as the fluid slowly percolates through the pores of the porous medium must be represented by a macroscopic law applying to masses of fluid large compared with the dimensions of the porous structure of the medium which is the usual Darcy law. In this case, the frictional force offered by the solid particles to the fluid denoted by $(-\mu/k)\vec{q}$, where μ is the viscosity of the fluid, k is the permeability of porous media (which has the dimension of length squared) and $\vec{q}$ is the mean filter velocity rather than the true velocity, replaces the usual viscous term in the Navier-Stokes equations. This form of Darcy law in the case of homogeneous isotropic material is of the potential

nature and is not of much use to discuss many practical problems cited above. This is because in many practical problems the porosity varies due to non-uniform packing. For example, at the bounding surface the void is large due to sparse distribution of particles and away from the surface, say, along the axis, the void will be less due to dense packing. This non-uniform packing gives rise to the bounding surface (Beavers & Joseph[1]) and in that case, one has to use the boundary layer type of equation. This boundary layer type of equation for flow through porous media which involves viscous resistance term in addition to the Darcy resistance was first postulated by Brinkmann[6] and a rigorous theoretical proof using the ensemble averages and point force approximations was given later by Tam[16]. The validity of the Brinkmann model depends on the magnitude of k/h^2, where h is the vertical thickness of the porous media. For example, if a porous medium is made up of spherical particles then the Brinkmann model is valid up to the magnitude of k/h^2 of order 10^{-3} or so. This corresponds to considerably high values of d/h, where d is the diameter of the fillings. It may be of interest to note that for high values of d/h, the assumption of homogeneity of the media may become questionable. As regards the theory, however, it must be pointed out that the microscopic state has been homogenised assuming the macroscopic velocity field and that the theory has been formulated using the permeability, k, which is defined in terms of the macroscopic quantities. In addition, k has to be determined as a property of the unit space per one of the filling particles. Under these approximations, it is not unreasonable to assume that even when d/h becomes comparatively high, the accuracy of the results obtained from the Brinkmann model does not deviate considerably provided that the variation in velocity is not appreciable in the medium.

Therefore in this paper, the linear and nonlinear hydrodynamic stability problem for a plane modified Poiseuille flow through porous media is solved using the Brinkmann model. It should be pointed out that the present work is not an attempt to explain turbulent drag reduction due to the friction force offered by the solid particle to the fluid but is merely intended to shed some light on the general hydrodynamical stability properties of parallel flow through porous media as such. As far as transition from laminar to turbulent flow is concerned, an attempt is made in this paper to study the nonlinear stability of parallel flow through porous media including inertia terms.

The mathematical formulation of the stability problem is given in Section 2. The linear stability is investigated in Section 3. The bound on the amplification factor and the corresponding bound on the Reynolds number are determined using Schwartz's inequality (Rudraiah[11]) and elementary isoperimetric inequalities (Joseph[7]). It is shown that isoperimetric inequalities given better results than the Schwartz's inequality. The nonlinear stability of parallel flows through porous media is discussed in Section 4 using Liapunov direct method. It is shown that the Rayleigh's[10] method of determining the eigenvalues improve the bound on the Reynolds number obtained by the method of calculus of variation.

2. Mathematical Formulation

The problem to be considered is the hydrodynamic stability of a steady parallel flow through porous media bounded by two infinitely large parallel rigid walls at $y=\pm h$. To study this problem the Brinkmann model will be used. In this model the basic equations for homogeneous incompressible viscous fluid through fluid saturated porous media are

$$\frac{\partial \vec{q}}{\partial t}+(\vec{q}\cdot\nabla)\vec{q}=-\frac{1}{\rho}\nabla p-\frac{\nu}{k}\vec{q}+\nu\nabla^2\vec{q}, \tag{2.1}$$

$$\nabla\cdot\vec{q}=0 \tag{2.2}$$

with no-slip conditions

$$\vec{q}=0 \text{ at } y=\pm h$$

Here $\vec{q}\,(u, v, w)$ is the mean filter velocity rather than the true velocity, p is the pressure, ρ is the density, ν is the kinematic viscosity, k is the permeability of the porous media having the dimension of length squared and

$$\nabla^2\equiv\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}$$

The basic flow is governed by

$$\frac{\partial}{\partial t}=0,\quad \frac{\partial}{\partial x}=0,\quad \frac{\partial}{\partial z}=0,\quad v=w=0,\quad u=v_0(y),$$

$$p=p_0 \text{ with } \frac{\partial p_0}{\partial x}=\text{const} \tag{2.3}$$

where $v_0(y)$ satisfies the equation

$$\frac{d^2v_0}{dy^2}-\frac{1}{k}v_0=\frac{1}{\mu}\frac{\partial p}{\partial x} \tag{2.4}$$

Solution of this satisfying the no-slip condition (2.2) is

$$v_0=\frac{ch\sigma-ch\sigma y^*}{ch\sigma-1} \tag{2.5}$$

where v_0 is made dimensionless using the maximum velocity V_m in the channel, $y^*=y/h$ and $\sigma=h/\sqrt{k}$

To study the linear hydrodynamic stability of basic flow (2.5) through porous media, we superimpose on it a small symmetrical disturbance of the form

$$u=u_0(y)+u'(x, y, z, t),\quad v=v'(x, y, z, t),$$
$$w=w'(x, y, z, t) \text{ and } p=p_0+p'(x, y, z, t) \tag{2.6}$$

where the primed quantities denote the perturbed quantities which are assumed to be small compared with the basic quantities.

Linearizing equations (2.1) and (2.2) using (2.6) and making the resulting equations dimensionless using v_m for velocity and h for the distance and assuming the perturbed quantities vary in the form

$$\text{function of } (y)\ \text{Exp}\{i(lx+mz)-ilct\} \tag{2.7}$$

we get

$$il(u_0-c)u+Du_0v=-ilp-\frac{\sigma^2}{R}u+\frac{1}{R}(D^2-a^2)u \tag{2.8}$$

$$il(u_0-c)v=-Dp-\frac{\sigma^2}{R}v+\frac{1}{R}(D^2-a^2)v \tag{2.9}$$

$$il(u_0-c)w=-imp-\frac{\sigma^2}{R}w+\frac{1}{R}(D^2-a^2)w \tag{2.10}$$

$$i(lu+mw)+Du=0 \tag{2.11}$$

where

$$R=\frac{v_m h}{\nu} \text{ is the Reynolds number,}$$

$$a^2=l^2+m^2$$

and

$$D=d/dy$$

The boundary conditions are

$$u=v=w=0 \text{ at } y=\pm 1 \tag{2.12}$$

We now wish to show that this three-dimensional problem can be reduced, by means of Squire's transformation to an equivalent two-dimensional problem. For this purpose we get

$$\tilde{a}u=lu+mw,\quad \tilde{v}=v,\quad \tilde{l}p=ap,\quad \tilde{a}R=lR \tag{2.13}$$

Substituting these new variables into (2.8) to (2.11), we get

$$ia(u_0-c)\tilde{u}+Du_0\tilde{v}=-ia\tilde{p}-\frac{\sigma^2}{R}\tilde{u}+\frac{1}{R}(D^2-a^2)\tilde{u} \tag{2.14}$$

$$ia(u_0-c)\tilde{v}=-D\tilde{p}-\frac{\sigma^2}{R}\tilde{v}+\frac{1}{R}(D^2-a^2)\tilde{v} \tag{2.15}$$

$$ia\tilde{u}+D\tilde{v}=0 \tag{2.16}$$

These equations together with the boundary conditions

$$\tilde{u}=\tilde{v}=0 \quad \text{at} \quad y=\pm 1 \tag{2.17}$$

obviously have the same mathematical structure as equations (2.8) to (2.12) with $m=w=0$ and they thus define the equivalent two-dimensional problem for all σ for the porous parameter σ is invariant with respect to Squire's[1,8] transformation. Thus we have the generalization of Squire's theorem to porous media which states that 'the three-dimensional problem is equivalent to a two-dimensional problem at a small Reynolds number for all values of the porous parameter σ'.

Eliminating $\tilde{p}$ between (2.14) and (2.15) and using (2.16), we get

$$(D^2-a^2)^2\tilde{v}-\sigma^2(D^2-a^2)\tilde{v}=iaR[(u-c)(D^2-a^2)\tilde{v}-(D^2u_0)\tilde{v}] \tag{2.18}$$

which can also be written as

$$(D^2-a^2)^2\tilde{v}=iaR[(u-e)(D^2-a^2)\tilde{v}-(D^2u_0)\tilde{v}] \tag{2.19}$$

where

$$e=c+i\frac{\sigma^2}{aR}=C_r+i\left(C_i+\frac{\sigma^2}{aR}\right)=C_r+iC_i$$

$$C_1=C_i+\lambda/a, \quad \lambda=\sigma^2/R \tag{2.20}$$

The corresponding boundary conditions are

$$\tilde{v}=D\tilde{v}=0 \quad \text{at} \quad y=0 \quad \text{and} \quad 1 \tag{2.21}$$

Equation (2.18) (or 2.19) is the modified Orr-Sommerfeld equation for porous media ; modified in the sense that the term $\sigma^2(D^2-a^2)\tilde{v}$ due to the friction offered by the solid particles appears on the left-hand side of equation (2.18). According to equation (2.19), this is equivalent to saying that the Orr-Sommerfeld equation for porous media is modified by changing the usual amplifiction factor C_i to $C_i+\lambda/a$. From this it is clear that the

porous media affects only the growth of the amplification of waves and not the phase velocity C_r. Thus the singular point of (2.19) in the case of porous media occurs at the value $y=y_e$ such that $v=e$ in contrast to $v=C$ at $y=y_c$ in the absence of porous media. In other words, the critical layer $y=y_e$ in the case of porous media is such that $C_i=-\frac{\sigma^2}{aR}$ and $v-C_r=0$ where $\tilde{v}(y_e)$ is not zero. The structure of this singularity and its mathematical consequences are to be discussed in detail. Work in this direction is under progress and in this paper we present only the sufficient conditions for stability.

We note that, as stated in Section 1, the Brinkmann model is valid only up to the order of $\sigma\approx 10^3$ or so. In that case the term on the left-hand side is of the same order as that of the right-hand side and equation (2.19) is valid. However, for large values of σ, say, of order greater than 10^5, the viscous shear term namely the right-hand side term of (2.19) is small compared to those on the left-hand side and in that case equation (2.19) transforms to

$$(v_0-e)(D^2-a^2)\tilde{v}-D^2u_0\tilde{v}=0$$

i.e.

$$(v_0-e)(D^2-a^2)\tilde{v}-D^2u_0\tilde{v}=\frac{i\sigma^2}{aR}(D^2-a^2)\tilde{v} \tag{2.22}$$

and satisfies the boundary conditions

$$\tilde{v}=0 \quad \text{at} \quad y=\pm 1 \tag{2.23}$$

Equation (2.22) is the stability equation for the stability of parallel flow through porous media governed by the usual Darcy law. Since the Darcy law under isothermal homogeneous conditions is of potential nature and hence equation (2.22) is analogous to the stability equation of inviscid in the absence of porous media. Thus equation (2.22) can be called as Rayleigh-Darcy equation for flow through porous media. We note that in the limit of $\sigma\to\infty$, the right-hand side of (2.22) tends to zero and we will have the usual Rayleigh stability equation namely

$$(v_0-c)(D^2-a^2)\tilde{v}-D^2u_0\tilde{v}=0 \tag{2.24}$$

It is of interest to note that although equations (2.22) and (2.24) are analogous but their properties are quite different. These properties are useful in discussing the singular solutions of (2.22) around the critical

layer $y=y_s$ at which $v-e=0$ as stated earlier. However, in this paper we will be interested in determining the eigenvalues of equation (2.19). This is determined using Schwartz inequalities (Rudraiah[11]), isoperimetric inequalities (Joseph[6]) and Galerkin method.

3. Eigenvalue Bounds

In this section the estimates of the eigenvalues $C(\lambda, R, \sigma)$ belonging to the manifold of solutions of (2.19) are constructed by the application of both Schwartz's inequality (Rudraiah[11]) and elementary isoperimetric inequalities (Joseph[7]).

For this purpose equation (2.19) is multiplied by $\tilde{v}^*$, the complex conjugate of $\tilde{v}$ and integrated with respect to y from -1 to 1 to obtain

$$ia\,Re(I_1^2+a^2 I_0^2)=ia\,RQ+I_2^2+2a^2 I_1^2+a^4 I_0^2 \tag{3.1}$$

where

$$Q=\int_{-1}^{1}\{v_0|D\tilde{v}|^2+a^2|\tilde{v}|^2+D^2 v_0|\tilde{v}|^2\}\,dy$$

$$+\int_{-1}^{1} Dv_0\, v^*\, Dv\, dy$$

$$I_2=\int_{-1}^{1}|D^2\tilde{v}|^2\,dy,\quad I_1^2=\int_{-1}^{1}|D\tilde{v}|^2\,dy,\quad I_0^2=\int_{-1}^{1}|v|^2\,dy$$

$$e=C_r+i(C_i+\sigma^2/aR)$$

Theorem 1 : Let $e(a, R, \sigma)$ be any eigenvalue of (2.19) satisfying (2.21). Then

$$C_i\leq \frac{qI_0I_1-(aR)^{-1}(I_2^2+2a^2I_1^2+a^4I_0^2)}{I_1^2+a^2I_0^2}-\frac{\sigma^2}{aR} \tag{3.2}$$

where

$$q=(v_0^1)_{\max}=\max_{y\in(-1,1)}|v_0^1(y)|$$

Moreover, no amplified disturbances ($C_i>0$) of (2.19) and (2.21) exist if

$$aRq<f(a)\equiv\max(M_1, M_2, M_3) \tag{3.3}$$

where

$$M_1=2a(2a+2+\sigma^2)^{\frac{1}{2}}(a^2+\sigma^2)^{\frac{1}{2}} \tag{3.4}$$

$$M_2=(2a^2+1)^{\frac{1}{2}}(4a^2+1+4a^2\sigma^2)^{\frac{1}{2}} \tag{3.5}$$

$$M_3=2(2a^2+4+\sigma^2)^{\frac{1}{2}}(a^4-4+a^2\sigma^2)^{\frac{1}{2}} \tag{3.6}$$

Proof: Consider functions $v_0(y)$ with two continuous derivatives and we deduce results (3.2) and (3.3) from (3.1) for elements of $\tilde{v}$ of a complex-valued Hilbert space, $\bar{H}$, complete under the norm I_2^2 by the addition of limit points of sequences of four times continuously differentiable functions satisfying (2.21).

Adding to (3.1) its own complex conjugate, we get

$$2(I_2^2+2a^2I_1^2+a^4I_0^2)=-i\sigma R(Q-Q^*)-2aR\left(C_i+\frac{\sigma^2}{aR}\right)$$

$$\times(I_1^2+a^2I_0^2) \tag{3.7}$$

Since

$$|Q-Q^*|=\left|\int_{-1}^{1} Dv_0(\tilde{v}^*D\tilde{v}-\tilde{v}D\tilde{v}^*)\,dy\right|$$

it follows that

$$|Q-Q^*|\leq\int_{-1}^{1}|Dv_0|\;|Dv|\;|v|\,dy$$

If $q=\max_{y\in(-1,1)}|DU_0|$ and applying Schwartz's inequality, we get

$$|Q-Q^*|\leq 2qI_0I_1 \tag{3.8}$$

The estimate (3.2) follows easily from (3.7) and (3.8). The relation (3.2) gives an upper bound for C_i. If aR is small enough so that the upper bound is negative for all functions $\tilde{v}$ in $\bar{H}$, satisfying (2.21), the flow is stable. Thus a sufficient condition for stability is

$$aRq<\min\left[\frac{I_2^2+2a^2I_1^2+a^4I_0^2}{I_0I_1}+\sigma^2\,\frac{I_1^2+a^2I_0^2}{I_0I_1}\right] \tag{3.9}$$

If $\sigma\to 0$ (i.e. impervious case), this tends to the result

$$aRq<\min\left[\frac{I_2^2+2a^2I_1^2+a^4I_0^2}{I_0I_1}\right] \tag{3.10}$$

which coincides with the one given by Synge[15]. Comparison of (3.9) and (3.10) reveals that in the case of porous media, the system will be stable even at large values of σR, for the range of $q\sigma R$ has increased by an amount of

$\frac{\sigma^2(I_1^2+a^2I_0^2)}{I_0I_1}$. Physically this means that the friction force offered by the solid particles to the fluid inhibits the onset of instability.

To obtain the estimate (3.3), consider any two real constants ξ and η such that

$$\int_{-1}^{1}\left(\tilde{v}+\xi\tilde{v}'+\eta\tilde{v}''\right)\left(\tilde{v}^*+\xi\tilde{v}^{*\prime}+\eta\tilde{v}^{*\prime\prime}\right)dy>0 \tag{3.11}$$

From this it follows that

$$\eta^2I_2^2>(\xi\eta-\xi^2+2\eta)I_1^2+(\xi-1)I_0^2 \tag{3.12}$$

Further, multiplying (3.2) by η^2 and using (3.12), we get

$$\eta^2aRC_i(I_1^2+a^2I_0^2)\leq\eta^2aRqI_0I_1-I_1^2(2a^2\eta^2+\xi\eta-\xi^2+2\eta+\sigma^2\eta^2)-I_0^2(a^4\eta^2+\xi-1+a^2\eta^2\sigma^2) \tag{3.13}$$

Stability is assured if the right-hand side of (3.13) is negative definite, which in turn follows from the inequalities

$$(\eta^2aRq)^2<4(2a^2\eta^2+\xi\eta-\xi^2+2\eta+\sigma^2\eta^2)(a^4\eta^2+\xi-1+a^2\eta^2\sigma^2) \tag{3.14}$$

$$2a^2\eta^2+\xi\eta-\xi^2+2\eta+\sigma^2\eta^2>0 \tag{3.15}$$

$$a^4\eta^2+\xi+a^2\eta^2\sigma^2-1>0 \tag{3.16}$$

Thus for any pair of real values of ξ and η satisfying (3.15) and (3.16), the inequality (3.14) gives a sufficient condition for stability. To obtain (3.4), we substitute $\xi=\eta=1$ in (3.14). Similarly to obtain (3.5) and (3.6) we substitute $\xi=\eta=2$ and $\xi=\eta=1/2$ respectively in (3.14).

These inequalities are obtained from Schwartz's inequality. To improve these inequalities we apply following Joseph[7], the elementary isoperimetric inequalities and the results are as follows.

Theorem 2: Let $c(a, R, \sigma)$ be any eigenvalue of (2.19) satisfying (2.21). Then

$$C_i<\frac{q}{2a}-\frac{\pi^2+a^2}{\sigma R}-\frac{\sigma^2}{\sigma R} \tag{3.17}$$

Moreover, no amplified disturbances ($C_i>0$) of (2.19) and (2.27) exist if

$$aRq<g(a)=\max\,[M_4, M_5] \tag{3.18}$$

where

$$M_4=(4{\cdot}73)^2\pi+2^{3/2}a^3+2\sigma^2a \tag{3.19}$$

and

$$M_5=(4{\cdot}73)^2\pi+2\pi\sigma^2+2a\sigma^2 \tag{3.20}$$

Proof: In the real-valued Hilbert space H corresponding to $\bar{H}$ the following inequalities hold

$$I_1^2>\pi^2I_0^2,\quad I_2^2\geq\pi^2I_1^2,\quad I_2^2\geq(4{\cdot}73)^2I_0^2 \tag{3.21}$$

The value $(4{\cdot}73)^2$ is the smallest eigenvalue of a vibrating rod with displacement $\tilde{v}$ satisfying (2.21) (Rayleigh[10], Joseph[7]). Following Joseph[7], we can easily show that inequalities (3.20) are also true in $\bar{H}$. For example, consider

$$\varphi=\alpha+i\beta$$

where $\alpha, \beta\in H$. Then

$$I_1^2(\alpha)\geq\pi^2I_0^2(\alpha),\quad I_1^2(\beta)\geq\pi^2I_0^2(\beta)$$

and by addition

$$I_1^2(\varphi)\geq\pi^2I_0^2(\varphi)$$

Other inequalities in (3.20) follow similarly.

From the inequalities (3.21) it follows that

$$2aI_0I_1\leq I_1^2+a^2I_0^2 \tag{3.22}$$

$$I_2^2+2a^2I_1^2+a^4I_0^2=(I_2^2+a^2I_1^2)+a^2(I_1^2+a^2I_0^2)\geq(\pi^2+a^2)(I_1^2+a^2I_0^2)$$

Therefore

$$\frac{I_2^2+2a^2I_1^2+a^4I_0^2}{I_0I_1}\geq\begin{cases}(4{\cdot}73)^2\pi+\dfrac{2a^2}{I_1I_0}\left(I_1-\dfrac{aI_0}{\sqrt{2}}\right)^2+\dfrac{2aI_0I_1}{\sqrt{2}}\\ (4{\cdot}73)^2\pi+\dfrac{2a^2I_1^2}{I_1I_2}\end{cases} \tag{3.23}$$

$$\geq\begin{Bmatrix}M_1\\M_2\end{Bmatrix}$$

The estimate (3.17) follows easily from (3.2), (3.22) and (3.23). Inequality (3.18) follows from (3·2) by observing that for a set of values of QR the left-hand side of (3.2) is negative and then applying (3.23). Inequalities (3.3) and (3·18) are numerically evaluated for different values of σ and the results are tabulated in the following table.

σ	min $f(a)/a$ from (3.3)	min $g(a)/a$ from (3.18)
0	4·04	42·22
10	242·22	262·96
20	955·77	910·38

For $\sigma=0$ (i.e. impervious case) the values agree closely with those of Joseph[7]. From this table it is clear that the inequality (3.2) gives better values than the inequality (3.18) for small values of σ, whereas for large values of σ inequality (3.18) gives better estimate than (3.2). Since the wave speed C_r

is not affected by the porous parameter σ, the estimates on C_r obtained by Joseph[7] are also true for the porous media, with a suitable change in the basic velocity distribution.

4. Nonlinear Stability Using Liapunov Direct Method

In Section 3 we have discussed analytic linear stability theory which gives no information about the effect of finite disturbances.

In this section we discuss the stability of solutions of the Brinkmann equations for parallel flow through porous media subject to disturbances of finite amplitude. The general approach is to use the Liapunov direct method as developed by Serrin[12] and Joseph[6] which has attracted considerable attention in stability analysis because of its elegances and potential power. The ability to analyse the finite amplitude behaviour of a nonlinear system without even solving nonlinear differential equations is extremely attractive. The main drawback in this method has been the general inability to construct Liapunov functions which define stability regions of sufficient size to be useful in practice. In certain cases, however, the structure of the basic equations gives a clue for constructing Liapunov functions. For example, in the present problem we note that under homogeneous isotropic conditions the nonlinearities in the Brinkmann equations (2.1) are only quadratic which suggests to use only the quadratic Liapunov function. Physically it represents the kinetic energy of the system. In this case some analytical solutions can be obtained without approximation. In particular we note that, since we are dealing with four partial differential equations (2.1) and (2.2) a good deal of mathematical manipulation is required.

4.1 *Universal stability criterion*

The basic equations are (2.1) and (2.2). The boundary conditions are the no-slip on bounding solid surfaces and boundedness in the x-direction because of infinite extent.

Although the Liapunov direct method is applicable even when the basic state is time dependent, we consider the basic state to be steady. The steady state solutions to (2.1) and (2.2) are denoted by $\vec{q}_s$ and p_s, and the perturbation variables are denoted by

$$u=\vec{q}-\vec{q}_s, \quad P=p-p_s \tag{4.1}$$

Because of the no-slip boundary condition $\vec{u}$ vanishes identically on solid

boundaries. Substituting (4.1) into (2.1) and (2.2), we get the difference equations

$$\frac{\partial u}{\partial t}+(\vec{q}_s+\vec{u})\cdot\nabla\vec{u}+\vec{u}\cdot\nabla\vec{q}_s=-\nabla p-\frac{\sigma^2}{R}\vec{u}+\frac{1}{R}\nabla^2\vec{u} \tag{4.2}$$

$$\nabla\cdot\vec{u}=0 \tag{4.3}$$

We define the Liapunov function V by the kinetic energy of the disturbance given by

$$V=\frac{1}{2}\int_\tau \vec{u}\cdot\vec{u}\,d\tau=\frac{1}{2}\int_\tau u^2\,d\tau \tag{4.4}$$

The region τ is bounded by the solid surface and extends one wavelength in any direction of infinite extent. In taking the time derivative we have to consider the changes due to variation of $\vec{u}$ within τ as well as its variation across the boundary S of τ. Therefore

$$\dot{V}=\frac{dv}{dt}=\int_\tau \vec{u}\,\frac{\partial\vec{u}}{\partial t}\cdot d\tau+\frac{1}{2}\int_S u^2\,\vec{q}\cdot\hat{n}\,dS \tag{4.5}$$

where $\hat{n}$ is the outward normal along the boundary. The surface integral in (4.5) vanishes because $\vec{u}$ is zero along solid boundaries and $\vec{u}$ and $\vec{v}$ are periodic in directions of infinite extent. Thus

$$\dot{V}=\int_\tau \vec{u}\cdot\frac{\partial\vec{u}}{\partial t}\,d\tau \tag{4.6}$$

Multiplying (4.2) scalarly by $\vec{u}$ and then substituting in (4.6), we get

$$\dot{V}=\int_\tau \vec{u}\left[\frac{1}{R}\nabla^2\vec{u}-\frac{\sigma^2}{R}\vec{u}-\vec{q}\cdot\nabla\vec{u}-\vec{u}\cdot\nabla\vec{q}_s-\nabla p\right]d\tau \tag{4.7}$$

The stability is ensured if we can find conditions for which $\dot{V}\leqslant 0$

It is useful to transform (4.7) to a suitable form using the vector identities

$$\vec{u}\cdot\nabla^2\vec{u}=\nabla\cdot\nabla\left(\tfrac{1}{2}u^2\right)-\nabla\vec{u}:\nabla\vec{u},$$

$$\vec{u}\cdot(\vec{q}\cdot\nabla)\,\vec{u}=\nabla\cdot\left(\tfrac{1}{2}u^2\vec{q}\right)-\tfrac{1}{2}u^2\nabla\cdot\vec{q},$$

$$\vec{u}\cdot\nabla p=\nabla\cdot(\vec{u}\,p)+p\,\nabla\cdot\vec{u},$$

$$\vec{u}\cdot(\vec{u}\cdot\nabla\,\vec{q}_s)=\vec{u}\cdot\vec{\vec{D}}\cdot\vec{u},$$

$$\vec{\vec{D}}\equiv\tfrac{1}{2}\left[\nabla\vec{q}_s+(\nabla\vec{q}_s)^T\right]$$

where $\vec{\vec{D}}$ is the rate of strain tensor for the basic flow. Since the fluid is incompressible $\nabla\cdot\vec{u}=\nabla\cdot\vec{q}=0$. Then (4.7) can be written as

$$\dot{V}=-\int_\tau\left(\vec{u}\cdot\vec{\vec{D}}\cdot\vec{u}+\frac{\sigma^2}{R}\vec{u}\cdot\vec{u}+\frac{1}{R}\nabla\vec{u}:\nabla\vec{u}\right)d\tau+\int_\tau\nabla\cdot\vec{\vec{\varphi}}\,d\tau \tag{4.8}$$

where

$$\vec{\varphi} = \nabla(\tfrac{1}{2}u^2) + \tfrac{1}{2}u^2\vec{q} + P\vec{u} \tag{4.9}$$

We can use Green's theorem to write

$$\int_\tau \nabla \cdot \vec{\varphi}\, d\tau = \int_S \vec{\varphi} \cdot \hat{n}\, dS$$

and the surface integral vanishes for the reasons stated earlier. Finally we have

$$\dot{V} = -\int_\tau (\vec{u} \cdot \vec{D} \cdot \vec{u} + \frac{\sigma^2}{R}\vec{u} \cdot \vec{u} + {}_R\nabla\vec{u} : \nabla\vec{u})\, d\tau \tag{4.10}$$

which is the Reynolds-Orr equation for porous media.

Now, for Poiseuille flow through porous media given by (2.5)

$$\vec{u} \cdot \vec{D} \cdot \vec{u} = -uvDu_0 \tag{4.11}$$

Let

$$q = \max Du_0 \tag{4.12}$$

Further

$$(|u| - |v|)^2 \leq u^2 + v^2 - 2|u|\,|v| \geqslant 0$$
$$\text{i.e. } 2|u|\,|v| \leq u^2 + v^2 \leq u^2 + v^2 + w^2 = \vec{u} \cdot \vec{u}$$
$$\therefore \int_\tau (\vec{u} \cdot \vec{D} \cdot \vec{u})\, d\tau \leq -q\int_\tau \vec{u} \cdot \vec{u}\, d\tau \tag{4.13}$$

Extension of Serrin's inequality to the rectangular geometry (*see* Appendix), yields

$$\int_\tau \nabla\vec{u} : \nabla\vec{u}\, d\tau \geqslant \frac{\pi^2}{4}\int_\tau \vec{u} \cdot \vec{u}\, d\tau \tag{4.14}$$

Thus (4.10) using (4.4) and (4.14) becomes

$$\dot{V} \leq \frac{2}{R}(qR - \sigma^2 - \pi^2/4)\, V \tag{4.15}$$

Integrating with respect to time, we have

$$V(t) \leq V(0)\, \exp\left\{\frac{2}{R}(qR - \sigma^2 - \pi^2/4)\, t\right\} \tag{4.16}$$

where $V(0)$ is the value of V at $t = 0$. Thus a sufficient condition for stability is

$$qR < \sigma^2 + \pi^2/4 \tag{4.17}$$

which establishes the existence of an open region of stability near the origin of the qR-space for every fixed σ.

Note that for impervious case of $\sigma \to 0$, $q \to 2$, with a suitable definition of R, condition (4.17) reduces to

$$R < \pi^2/8 \tag{4.18}$$

4.2 *Optimization approach*

The universal stability condition (4.16) obtained by Liapunov direct method is quite conservative. This may be due to the fact that we have used very weak bounds in passing from the exact equality (4.10) to the sufficient condition (4.17). The search for the strongest condition, i.e. for the largest region of stability for a given Liapunov function naturally suggests the use of optimization methods.

For this we use the sufficient condition from (4.10) in the form

$$qR\int_\tau uv\, d\tau - \sigma^2\int_\tau \vec{u} \cdot \vec{u}\, d\tau - \int_\tau \nabla\vec{u} : \nabla\vec{u}\, d\tau \leq 0 \tag{4.19}$$

For positive qR this can be written as

$$\frac{1}{qR} \geqslant \frac{\int_\tau uv\, d\tau}{\int_\tau (\nabla\vec{u} : \nabla\vec{u} + \sigma^2\vec{u} \cdot \vec{u})\, d\tau} \tag{4.20}$$

Then a sufficient condition for stability is

$$qR \leq \lambda \tag{4.21}$$

where

$$\frac{1}{\lambda} = \frac{\max\int_\tau uv\, d\tau}{\int_\tau (\nabla\vec{u} : \nabla\vec{u} + \sigma^2\vec{u} \cdot \vec{u})\, d\tau} \tag{4.22}$$

and the maximization is over all functions satisfying the boundary conditions and the continuity equation (4.3). Determination of the maximum of the ratio of integrals in (4.22) is a problem in optimization and we use the calculus of variations. From the calculus of variation we know that the function which causes $\frac{1}{\lambda}$ to take its maximum is a solution of the Euler equation

$$\frac{d^2u}{dy^2} + (\sigma^2 - \lambda)u = 0 \tag{4.23}$$

where $u(\pm 1) = 0$ (4.24)

The eigenfunctions of (4.23) are

$$\sin\{(\sigma^2-\lambda)^{\frac{1}{2}}y\} \text{ with eigenvalues}$$
$$\lambda=\sigma^2+n^2\pi^2$$

The smallest eigenvalue is

$$\lambda_{\min}=\sigma^2+\pi^2 \tag{4.25}$$

It is of interest to note that, if we follow the method of Rayleigh[10] to determine the least eigenvalue of (4.23) satisfying (4.24), we get

$$\lambda_{\min}=(4{\cdot}73)^4\pi^2+\sigma^2 \tag{4.26}$$

It now follows from (4.21) and (4.25) or (4.26) that parallel flow through porous media is monotonically and globally stable, whenever

$$qR<\pi^2+\sigma^2 \tag{4.27}$$

Or, from (4.26)

$$qR<(4{\cdot}73)^4\pi^2+\sigma^2 \tag{4.28}$$

From (4.27) we note that there is about 7·4 fold increase in the critical value of qR when compared to the value given by (4.17).

ACKNOWLEDGEMENT

This work is sponsored by the UGC special assistance programme.

REFERENCES

1. Beavers, G. S. and Joseph, D. D. (1967) *J. Fluid Mech.*, 30 : 197-208.
2. Betchov, R. and Criminale, W. O. (1967) *Phys. Fluids*, 19 : 359-3652.
3. Brinkmann, H. C. (1947) *J. Inst. of Maths. Applies*, 3 : 313.
4. Drazin, P. G. and Howard, L. N. (1966) *Advan. Appl. Mech.*, 9 : 1-89.
5. Joseph, D. D. (1964) *Phys. Fluids*, 7 : 1761-1771.
6. Joseph, D. D. (1965) *Arch. Ratl. Mech. Anal.*, 20 : 59.
7. Joseph, D. D. (1968) *J. Fluid Mech.*, 33 : 617.
8. Joseph, D. D. (1976) Stability of fluid motions I & II Springer Tracts in natural philosophy, 27.
9. Lin, C. C. (1957) *Appl. Mech.*, 2 : 136-148.
10. Rayleigh, Lord (1878) 'Scientific papers', Cambridge Univ. Press, London and New York, 1 : 361-371.
11. Rudraiah, N. (1964) *Appl. Sci. Res.*, B 11 : 105-127.
12. Serrin, J. (1959) *Arch. Ratl. Mech. Anal.*, 3 : 120.
13. Squire, H. B. (1933) *Proc. Roy. Soc.*, A 142 : 621-628.
14. Stuart, J. T. (1960) *J. Fluid Mech.*, 9 : 353-370.
15. Synge, J. L. (1938) Semi-Centennial publication of the American Mathematical Society, (Addresses)Am. Math. Soc. Providence, Rhode Island, 2 : 227-269.
16. Tam (1969) *J. Fluid Mech.*, 38 : 124.

Synovial Joint Lubrication : A Review

Peeyush Chandra

The Mehta Research Institute
of
Mathematics and Mathematical Physics
26, Dilkusha, New Katra, Allahabad-211 002

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

Introduction

In the recent years, physicians, engineers and scientists from other fields have started collaborating to meet the challenges of nature in interpreting its secrets in the medical profession. One such secret to which much inter-disciplinary research is currently directed, is related to synovial joint. Many physiologists and physicians have been conducting experiments and making observations to study the mechanism of synovial joints for last several years, and it is only recently that engineers and other scientists have also started taking interest in this endeavour. The synovial joint of the body, whose main function is to provide trouble-free motion, has long been identified as an ideal bearing system[68]. It supports considerable load with low friction service over a long span of one's lifetime even under variable conditions. The most fascinating aspect of synovial joints is its very low coefficient of friction ('002 or less)[12,22,47,57]. The basic problem related to joint mechanism, as outlined in Gray's Anatomy, is 'to account convincingly in terms of movement, loading and lubrication for their unusual and unique efficiency'. This is of particular interest to engineers who have yet to construct a mechanical bearing which may work as efficiently as the biological bearing works. Thus, the proper understanding of lubrication mechanism in synovial joints may enable them to possibly improve upon the existing mechanical bearings.

Many speculations based on experimental as well as analytical studies have been directed to this problem of synovial joint lubrication. Engineering methodologies have been extensively employed in these investigations. The first such attempt was made in 1932 by an anatomist, MacConaill[61] and the efforts are still continuing. Many excellent reviews on joint lubrication already exist in literature, notably those by Dowson[22], McCutchen[70], Mow and Lai[86], Radin and Paul[98], Swanson[118] and others. The purpose of this review article is twofold: firstly, to review the current ideology of lubrication mechanism of synovial joints and to bring out the contribution of analytical studies in this direction, and secondly, to introduce the biomechanical aspects of the study to those who aspire to probe further in this direction.

Synovial Joint Anatomy

Those joints in the body which allow various degrees of relative motion of the bones are called synovial joints, e.g., hip, knee and finger joints. These

are also known as diarthrodial joints. A synovial joint consists of two or more bones. The load bearing ends of these bones are enclosed in the joint capsule and their opposing ends are covered with a soft glistening tissue, known as articular cartilage. The cartilaginous surfaces enclose between them a joint cavity which is filled with a small quantity of synovial fluid. The whole system is surrounded by ligaments and other soft tissues to provide stability during motion[85,138].

These synovial joints encounter a wide range of operating conditions. They support large loads (4–5 times the body weight)[88] under complicated motions and varying speeds[22]. These loads may be of a fluctuating nature also[97]. Moreover, the load bearing area of cartilages is very small[135], e.g., it is 127 mm^2 for hip joints[110]. It has been suggested that the ability of joints to provide ideal performance in spite of the severe operating conditions is the result of complex interaction between the bearing components, i.e., the cartilage surface and synovial fluid[85]. So, we shall discuss in brief the composition and properties of these two materials.

Articular cartilage is a soft connective tissue which surrounds the articulating ends of bones at synovial joints. The purpose of this articular cartilage is to provide a suitable covering to the bone ends so as to facilitate motion. These are the actual load carrying surfaces and can be regarded as the bearing surfaces while looking at the synovial joint from the tribiological viewpoint. The thickness of the cartilage layer ranges between 1–7 mm and it varies from species to species and also from one joint to another.

The chemical composition of cartilage differs with the age and site of the joint. It can be described as a highly specialized connective tissue mainly composed of firm turgid water gel[9] which is rich in extracellular materials. The ratio of cell mass to organic tissue mass is of the order 0·1 to 0·01[88]. The cells (chondrocytes) are embedded in a matrix of extracellular materials which consists of collagen fibrils, water and chondroitin sulfate. The water contents of cartilage matrix is about 70 per cent (by weight) and the rest of the matrix material is equally distributed (by weight) between collagen fibrils and chondroitin sulfate. The collagen fibrils are bound to protein and provide mechanical strength to the tissue[87]. Two types of chondroitin sulfate, i.e., of type A and C are found in the cartilage which are helpful in regulating consistency and elasticity of the cartilage[58,110,133].

The structure of cartilage matrix has been studied by various workers[15,60,90,128]. The electron microscope techniques have confirmed

that the cartilage can be considered as a layered medium[61]. Although the earlier classification in terms of characteristic direction of fibre orientation[8] is no more valid, still some regularity in structure can be observed[72,124]. The following classification of cartilage structure has been recognised in most of the works.

ZONE I : Superficial Tangential Zone (STZ). This zone is adjacent to the joint cavity. In this zone, the collagen fibrils are densely packed in bundles (with diameter 200–300 Å and with a periodicity 640 Å) and these bundles are arranged tangentially to the surface[128]. The cells in this region are elliptic in shape with the long axis running parallel to the surface. The STZ is approximately 10 per cent of total tissue thickness and is 200-600 μm thick. It has been observed that the most superficial zone is covered with a thin membrane-like structure, surface lamina and the cells in this region (STZ) are basically different from those below.

ZONE II : Intermediate or Transitional zone. It lies below the STZ and consists of coiled collagen fibrils which are of somewhat random nature. The fibre thickness ranges from 300 Å to 600 Å. These S-shaped fibres are arranged in an open meshwork with large spaces between them. The cells in this region are different from those in STZ. They are spheroidal in shape and are equally spaced[72].

ZONE III : Deep or Radiate zone. It is the continuation of Zone II, but the fibres here form a lighter meshwork and are predominantly radial to the cartilage surface. The shape of the cells is spheroidal again but now they are arranged in columnar groups. Each column consists of four to eight cells.

ZONE IV : Calcified zone. The calcified zone consists of only a few cells and is rich in crystals of calcium salts. It is separated from the deep zone by a wavy basophilic line[96]. Below this zone lies the subchondral and cancellous bone.

The articular cartilage surface, contrary to the earlier belief, is not a smooth surface. With the advent of electron microscopes, it has been revealed that the cartilage surface is rough on a microscopic scale[13,84]. The centre line average of cartilage is 1 μm and it increases with age[152]. It is superimposed upon a directionally oriented undulating waveform having amplitudes between 2–5 μm which again increases with age[36,125]. The roughness waves show a periodic pattern with the periodicity of 20-50 μm. The regular wavy patterns of the cartilage surface has also been observed

by Clarke[13,14] and Garden and McGillivery[85]. Sayles *et al.*[103] and Thomas *et al.*[118] have directly measured the cartilage surface roughness with a stylus instrument, and have found that the height distribution is Gaussian with inverse square power spectrum. The measurement on a femoral head has shown roughness of up to 7 μm, slopes of up to 2° and peak radii of up to 0·9 mm[103].

The physical structure of cartilage can be described as a fluid filled, porous and deformable material. The tissue deforms instantaneously on the application of load of short duration and it deforms extensively while loaded for longer periods. The load deformation behaviour is of viscoelastic type and the phenomenon of creep takes place due to extrusion of matrix fluid[39,48,49,131]. In fact the colloidal nature of the cartilage is supposed to be responsible for its viscoelastic behaviour[5]. The dispersed phase of colloidal cartilage matrix is formed by chondroitin sulfate and protein, while the fluid filled in the matrix acts as the dispersant phase[85]. This has suggested that the cartilage can be modelled as a binary mixture consisting of a linear elastic solid matrix with viscous properties and an incompressible fluid[51]. Creep and stress relaxation experiments have been used to determine 'aggregate elastic modulus' and the values have been reported as 0·76 MN/m^2 and 0·87 MN/m^2 respectively[52,86]. The elastic property of the cartilage has been studied in detail[27,46,59]. Fessler[31,32] has reported that the elastic behaviour of cartilage is mainly due to the interaction of collagen and sulfated protein with the interstitial fluid. Experiments have been conducted based either on the indentation tests on the cartilage surfaces[43,54,111] or on the compression tests on section plugs of the cartilage surfaces[46,48,74]. Hori and Mockros[43] have made comparison of various experimental results and have reported the Young's Modulus for normal human cartilage as 7·87 MN/m^2. Johnson[45] has reported its value at zero creep as 20 MN/m^2. The Poisson's ratio of 0·42 for instantaneous loading has been reported by Torzilli[119] which decreases asymptotically to 0·37 for long duration loading.

The cartilage matrix is slightly permeable to the dialysate phase of synovial fluid. The permeability has been found to play an important role during the movement of interstitial fluid[66] as well as in lubrication mechanism[78]. Therefore, it has been studied extensively. The earlier studies[7] were of qualitative nature and only recently the quantitative studies are found to be important[64,68,70,71,74]. McCutchen[74] determined the permeability to

flow normal to the surface by using Darcy's law, as $0{\cdot}58 \times 10^{-15}$ m⁴/N-sec. It has also been shown that the permeability decreases with the depth of the cartilage layer from the surface[25,71,74] and the cartilage becomes more porous with age[105]. McCutchen[74] also determined the tangential permeability coefficient and the value is reported as $0{\cdot}545 \times 10^{-15}$ m⁴/N-sec. However, contradictory experimental evidences have been reported about the tangential permeability[86,91]. The pore diameter of the cartilage is about 60 Å, and the small permeability indicates high resistance to fluid flow.

More details about the chemical composition and its biomechanical properties can be found in the well-written reviews by Mow[85], Hamerman *et al.*[37] Wright *et al.*[184], Torzilli[110] and in 'Adult Articular Cartilage' (Ed. Freeman)[88].

Synovial fluid is the fluid in joint cavity which acts mainly as a lubricant to reduce the friction and wear of cartilage surfaces[62]. It also serves as the source of nutrient for the cartilage cell. Normal synovial fluid is a clear yellowish viscous fluid. In most of the joints very small quantity of synovial fluid is found, e.g. a knee joint contains only 0·2 – 0·5 ml of synovial fluid[102].

Synovial fluid mainly consists of the dialysate phase of blood plasma along with the addition of hyaluronic acid molecules. The hyaluronic acid molecules which were first separated by Meyer *et al.*[92] from the synovial fluid of cattle, are found in varying amount. The average concentration of hyaluronic acid in synovial fluid is 3·5 mg/g. It is a long chained unbranched polymer, known as mucopolysaccharide, with molecular weight[4] of the order 5×10^6. The linear length of a hyaluronate molecule is 10^4 Å and so it cannot pass through the pores of cartilage[94]. This implies that the cartilage matrix contains only the dialysate phase of synovial fluid. Synovial fluid also contains various proteins which are assumed to be associated with the hyaluronic acid[94].

The physical and chemical properties of synovial fluid have been studied in detail[6,16,21,77,129,150]. It is established beyond doubt that it has shear dependent viscosity and exhibits various non-Newtonian behaviours, such as pseudoplasticity[6,8,17,85], elasticity[17,95,128], viscoelasticity[13,95,105]. A number of theoretical models describing the rheological behaviour of synovial fluids have been discussed in a review article by Lai *et al.*[58]. It is a highly viscous fluid at low shear rates and exhibits shear thinning behaviour at high shear rates[6,44]. Its apparent viscosity falls from 100 P at 0·01 sec^{-1} to

10 CP at 100 sec^{-1}. Myers *et al.*[92] have found the value of dynamic elasticity to lie in the range 2–40 dyne/cm² and have noted that it is viscous for slowly moving joints but assumes the property of elasticity as well for rapidly moving joints. While Negami[93] and others[17,128] have pointed out that the elasticity increases with shear rates, Furguson *et al.*[29] have reported that it does not vary to any great extent with shear rates. However, the concentration of hyaluronate has profound effect on its viscoelastic properties[1,29,30]. The hyaluronate concentration plays an important role in increasing the viscosity of synovial fluid and is responsible for its non-Newtonian behaviour.

For detailed information of synovial fluids one may refer to Hamerman *et al.*[37], Wright *et al.*[188] and Scott Balir[104].

Theories of Classical Lubrication

With this background, we come close to the study of synovial joint lubrication. The early realization that nature's bearing system and a mechanical bearing function more or less on similar principles, motivated scientists to explain the joint lubrication mechanism on the basis of the theories of lubrication used for mechanical bearings. A mechanical bearing usually operates under one of the following mechanisms[22] :

(a) Fluid film lubrication
(b) Boundary lubrication
(c) Elastohydrodynamic lubrication
(d) Mixed lubrication.

The term 'fluid film lubrication' is often referred to as the ideal conditions of lubricated contact. It provides high resistance to wear and tear. The two bearing surfaces are kept apart from each other by a thin film of a fluid lubricant (of the order 10^{-4}–10^{-5} cm) which is many times greater than the surface irregularities. The characteristics of the contact zone are governed by physical properties, mainly the viscosity, of the fluid.

The term boundary lubrication is used when the two surfaces are separated by a very thin film of molecular proportion. The bearing characteristics in this case depend upon the chemical properties of the lubricant rather than its viscous property. The presence of long chain polymer compounds in the lubricant is a significant feature of this mechanism. These long chain molecules get absorbed by the surfaces and become strongly bound to them.

Thus they form the absorbed layer of the polymer compound which avoids the direct contact.

Between the above two lubrication mechanisms, we have elastohydrodynamic (EHD) lubrication, which may be considered just another form of hydrodynamic lubrication. In this case, the elastic behaviour of the surfaces plays an important role. When the load is applied, elastic deformations of the bearing surfaces take place, which modify the fluid film as well as the fluid film pressure distribution. It enables the bearing to maintain fluid film for longer intervals. An interesting feature is that the fluid film thickness almost remains constant throughout the loaded region. The increased load just flattens the elastic surfaces without much influencing the film thickness.

The mixed lubrication, obviously, is a mechanism which operates when the conditions of hydrodynamic as well as boundary lubrication prevail. The chemical and physical properties of the lubricant, both, govern the bearing characteristics.

Mechanism of Synovial Joint Lubrication

Now, we are well set to discuss the lubrication mechanism for synovial joints. It may be of interest to quote Radin and Paul[98] in this context—'The low friction of cartilage at all speeds, high or low, certainly suggests that the bearing load is carried largely or wholly by some sort of liquid or liquid-like film at all speeds. The big question is : How is the film created ?' It is the answer to this question which would suggest the lubrication mechanism for the joints and attempts have been made from time to time to find a suitable explanation for it.

The first step in this direction was taken by an anatomist, MacConaill[61]. As hydrodynamic lubrication is the ideal lubricating process, he suggested that the synovial joints function as hydrodynamically lubricated bearings. The proposal was based on the observation of extremely low coefficient of friction and the wedge-shaped geometry of synovial joints. This theory of hydrodynamically lubricated joints, though a crude one, as it was based on inadequate information, was not disputed till 1959. In 1959, Charnley[12] proposed boundary lubrication for synovial joints. He expressed his doubts regarding the existence of fluid film between the cartilage surfaces under high load and low speed conditions. Moreover, it was pointed out that hydrodynamic lubrication does not suit to the intermittent reciprocating

bearings which the synovial joints certainly are. Further, Charnley[12] observed chemical affinity of hyaluronate molecules with the cartilage surface. These observations led him to conclude boundary lubrication for the synovial joints which are lubricated with the molecular layer by hyaluronic acid molecules.

It was only in 1959 that, for the first time, a thought was given to the biological environment of the system as well as to the characteristics of its components. McCutchen[73,74,76,78] considered the significance of the porous nature of the cartilage matrix. He performed *in vitro* tests and observed that the surfaces, which are saturated with a water-like fluid, weep liquid on being pressed against each other. The expelled liquid forms a fluid film between the joint surfaces and enables the joint to support load. This process, which is almost unique to biological conditions, was termed as 'Weeping lubrication'[78]. This mechanism seems to be plausible for the joints but only under the conditions of squeezing and it does not explain satisfactorily when the surfaces are in tangential motion. Radin *et al.*[99] found weeping action significant even under excessive load.

Dintenfass[20] considered the role of cartilage elasticity coupled with the non-Newtonian behaviour of synovial fluid to explain the joint lubrication. He put forward the concept of EHD lubrication for synovial joints, which was later supported by Tanner[117]. Both of them tried to calculate the fluid film thickness. Dintenfass[20] did not succeed much because of the lack of consistent data, but Tanner[117] produced the theoretical value of film thickness as 10^{-5} cm. Dintenfass totally ruled out any possibility of boundary lubrication while Tanner did indicate the slight possibility of boundary lubrication at the asperity contacts. Snidle and Biswas[109] also suggested that spherical contact synovial joints may be elastohydrodynamically lubricant.

In 1967, Dowson[22] realized that one mode of lubrication is insufficient to explain the joint mechanism. He suggested that two or more mechanisms might be effective under different ranges of operating conditions. He observed that in the beginning elastohydrodynamic conditions might be prevalent at low speeds and then the persistent loading would be supported by squeeze film action which may preserve the fluid film for long intervals. Under the extreme conditions of loading, boundary lubrication might be dominating, according to him. This idea of the existence of more than one mode of lubrication, was promptly supported by McCutchen[75,76], who concluded

that weeping lubrication was not the sum total of joint lubrication but was supplemented with boundary lubrication by synovial mucin.

In 1968, the Leeds group, headed by Dowson suggested a new concept of lubrication, 'boosted lubrication'[126], under the conditions of squeezing. The concept makes use of the surface roughness and the porosity of the cartilage and the concentration effects of hyaluronate molecules on synovial fluid viscosity. It may be pointed out here that the hyaluronate molecules, because of their bigger dimensions, cannot pass through the cartilage pores under squeezing. This along with the observation of hyaluronate gel formation[70] on the dry cartilage surfaces, inspired them to suggest that, under squeezing, the hyaluronate molecules get trapped in the roughness grooves while the base fluid would pass through the cartilage matrix. This results into the trapped pools of concentrated synovial fluid (which is rich in hyaluronate concentration) and this process of 'entrapment and enrichment' helps the joint to support load for longer durations[126,127].

It may be worth noting here that the two concepts of lubrication suggested under squeeze action, i.e., weeping[74] lubrication and boosted lubrication[126], are based on the contradictory requirements of fluid transport. While the former requires exudation of the fluid from the cartilage to the cavity, the other one requires a net flow of the base fluid from the fluid film region to the cartilage matrix.

Higginson and Norman[42] have investigated the squeeze film action of synovial joints. They accounted for the elasticity and permeability of the cartilage and the variable viscosity of synovial fluid. They concluded that weeping and filtration mechanism through the cartilage do not make substantial contribution to the lubrication mechanism, and the entrapment of fluid is an elastohydrodynamic phenomenon, which plays an important role in providing squeeze film conditions. Later, Higginson[40] observed that prospects of elastohydrodynamic lubrication are quite good, but the promising mode is squeeze film and not rolling or sliding.

It may be noted that in most of these studies the role of synovial fluid in joint lubrication has been based on its viscous properties. However, now serious doubts have been cast on its lubricating role in the cartilage—cartilage lubrication and low friction, a characteristic of synovial joint, has been observed not to increase when the viscous properties of synovial fluid are changed[58,79,100,105]. Thus, Swanson[113] and Little *et al.*[60] proposed boundary lubrication mechanism for synovial joint. It is suggested that

natural lubrication of the joint depends upon the assumption of formation of adsorbed layer of hyaluronate protein at the surface. This layer, which is formed at least partly due to fat contents of cartilage, acts as boundary lubricant. Swann[113] and Davies *et al.*[18], have strongly supported the boundary lubrication mechanism for joints. They have postulated that multiple layers of lubricating glycoprotein are formed at the surface which act as lubricant and also enable water to move in and out of porous matrix.

Swanson[114] and Wright and Dowson[132] also supported boundary lubrication at least under certain conditions. However, they suggested, like McCutchen[75], that it is supplemented with other mechanisms as well. The squeeze film situation develops at the heel strike, which leads to elastohydrodynamical mechanism together with the boundary lubrication, however, ultimately it is hydrodynamic mechanism which takes over during the free phase of walking[132].

Thus, it may be noticed that considerable controversy has existed for years regarding the joint lubrication mechanism and, the advent of new techniques and interdisciplinary studies has hardly contributed in clearing the confusion. However, renewed and refined investigations have thrown light on the biomechanical properties of cartilage and synovial fluid and on their dynamic interaction. Unless something more is known about the wide-ranging operating conditions, one may not reach the final conclusion. However, our present knowledge suggests that the remarkable performance of synovial joint in the combined result of various mechanisms each valid under different operating conditions[9,70,85,98,114,119].

Analytical Studies

Now, as mentioned in the introduction, we shall discuss the analytical studies which are conducted to throw light on lubrication mechanism of synovial joints. However, a relevant question arises "What is the need of analytical models and how do they help in the investigation of joint lubrication mechanism." This will be clear if we realize the experimental complications involved in the *in vivo* study. It may be pointed out that the cartilage and synovial fluid may not show the normal behaviour when taken out of the biological environment and so the *in vitro* measurements may be a poor reflection of the actual ones. Thus, at this stage the need for analytical models arises. Moreover, such models are also helpful in the study of the

variation of various parameters which may be required in the study of pathological cases. The aim of these analytical studies is to make a physically realistic mathematical model for the joint. These models, consistent with the experimental results, may ultimately lead to the understanding of joint lubrication mechanism. Also, these models have been successfully used to study the movement of interstitial fluid in the cartilage matrix. However, it may be pointed out that it is not possible to simulate all the complicated features of synovial joint in one single mathematical model so as to make it completely comprehensive. So, in most of the cases a simple model is considered and complicated features are introduced one by one. This is continued till either a situation of sufficient complexity is reached or the computational aspects do not allow further sophistication. We shall discuss some analytical models here.

Though Dintenfass[20], Tanner[117] and Dowson[22] had made some numerical calculations for the fluid film thickness using the classical results of elastohydrodynamic lubrication, Fein[28] was probably the first person to give a mathematical model for synovial joint. He used the classical Reynolds equation[101] for hydrodynamic lubrication to calculate the fluid film thickness between two normally approaching circular plates. The sphere-on-flat plate model, elastically equivalent to synovial joint with spherical surfaces, gives a contact region over a circle of Hertzian radius b.

$$b = 1{\cdot}15(WR/E)^{1/3} \tag{1}$$

where W is the applied load, R is relative radius of spherical surface and E is the equivalent elastic modulus. Assuming an elliptic pressure distribution for elastic surfaces (instead of the usual parabolic one for rigid surfaces) and constant viscosity μ of synovial fluid, he obtained the film thickness h at time t as

$$h = 2{\cdot}86\,\frac{\mu^{1/2}W^{1/6}}{t^{1/2}}\left(\frac{R}{E}\right)^{2/3} \tag{2}$$

It is observed that after loading of one hour a fluid film thickness of $5{\cdot}8\times10^{-6}$ cm is maintained under a load of 30 kg even for low viscosity and high elastic modulus. He concluded that the joint is squeeze film lubricated.

Dowson *et al.*[24] have given yet another simple formulation. They considered the viscosity (η) variation with the hyaluronate concentration (c) as follows

$$\eta = \eta_0\left(1+\frac{c}{0{\cdot}16}\right) \tag{3}$$

where η_0 is the base fluid viscosity. A two-dimensional flow was considered and the Reynolds equation was simplified to

$$\frac{d}{dx}\left(\frac{h^3}{\eta}\frac{dp}{dx}\right) = -12W \tag{4}$$

while considering various squeeze film geometries (e.g., for the 'cylinder near a plane' system, (x, r, θ) coordinate is considered, p is the pressure, h is film thickness and W is the squeeze velocity). The flow in porous matrix of cartilage has been shown to be negligible as compared to the flow in fluid film region and so porosity effect has not been considered. The 'boosted' squeeze time has been determined and compared with the normal squeeze time. The ratio of the two depends upon hyaluronate concentration and the analysis supports the concept of 'boosted lubrication'.

Shukla *et al.*[108] presented a mathematical analysis for diseased joints which considered linear relation for viscosity (μ) dependence on concentration (c), i.e.

$$\mu = \mu_0(1+\lambda c) \tag{5}$$

(where μ_0 is base fluid viscosity and λ is constant). This relation is similar to the one used by Dowson *et al.*[24], however, here c is not constant, but depends upon diffusion, biochemical reaction, etc. The concentration is assumed to be given by the following form of diffusion equation

$$D\frac{\partial^2 c}{\partial y^2} - ac = 0$$

(D is coefficient of diffusion and $a > 0$ is the rate of decrease in concentration of hyaluronic acid due to biochemical reactions). Chandra[9] has also considered a model for diseased joints. It has been shown that the squeeze time decreases as hyaluronate concentration decreases.

Ling[56] presented an analytical model to account for the nonlinearity of cartilage. He considered the fluid-filled cartilages as circular porous discs of thickness λ which are approaching each other. The movement of fluid in cartilage was governed by Darcy's law, and the equations governing the pressures ϕ (in cartilage) and Φ (in fluid film) are given by

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\lambda\frac{\partial\phi}{\partial\rho}\right) = \frac{\partial\lambda}{\partial\tau} + \gamma(\rho,\tau) \tag{7}$$

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\lambda_f^3\frac{\partial\Phi}{\partial\rho}\right) = \frac{1}{J}\frac{\partial\lambda_f}{\partial\tau} - 2\gamma(\rho,\tau) \tag{8}$$

(where ρ and λ_f are non-dimensional radial distance, and fluid film thickness respectively, τ is dimensionless time, $\gamma = s(\phi - \Phi)$, s is a dimensionless parameter). It has been shown that the two mechanisms of weeping and boosted lubrication are not exclusive to each other, rather they supplement each other.

Chandra[10] has also considered a similar model where the effects of cartilage porosity and hyaluronate concentration in synovial fluid have been studied. It is shown that the effect of permeability coefficient on squeeze time is not very significant and that the increase in concentration parameters delays the time of squeezing. The analysis supports the boosted lubrication under the conditions of high load and no tangential movements of surfaces.

Snidle and Biswas[109] calculated film thickness formula for spherical contact synovial joints and indicated the importance of classical elasto-hydrodynamics effects in such joints.

Higginson and Norman[41], Dinnar[19], Mansour *et al.*[66] have given more rigorous mathematical models. They incorporated elastic as well as porous behaviour of the cartilage in their analyses. Higginson and Norman[41] considered the elastic deformation of cartilage through the film thickness as

$$h = x + \frac{r^2}{2R} + Ap \qquad (9)$$

(x is the centre line displacement of upper spherical surface, A is flexibility of lower poro-elastic plane surface and R is radius of curvature of upper surface). One-dimensional, nonlinear Reynolds equation was solved numerically for fluid film thickness. The analysis suggested that the permeability of the cartilage does not play a significant role and that the enrichment mechanism due to filtration may not be very effective. However, filtration through surface aggregates[129] has been supported.

Dinnar[19] suggested two-layered model for cartilage, each layer having different permeability and thickness. Mansour *et al.*[66] considered cartilage as a single layer of homogeneous isotropic poro-elastic material which rests on an elastic non-porous material (bond). Using Biot's theory of consolidation a mixed lubrication process has been suggested. Their analysis, unlike Mow and Ling[84] (who also used Biot's theory) did not support either the boosted or weeping mechanism. They also pointed out that a circulatory mode of fluid transport is created due to mechanical pumping effect. The fluid is exuded 'in front of and under the leading half of the advancing load' (which is moving across the surface) and is imbibed 'under the trailing

half and behind the advancing load'. This view has been supported by various analyses[56,69,121]. Torzilli and Mow[120,121] in two papers have viewed articular cartilage as a binary mixture of two mechanically interacting continua, i.e., the solid matrix phase and liquid phase of composed water. It has been shown that the flow process depends upon (i) tissue geometry, (ii) characteristic frequency of loading, and (iii) tissue properties (e.g., permeability, stiffness, etc.). Mansour and Mow[65] considered the non-linearity of cartilage and a three-layered two-phase medium (consisting of liquid and solid phase) is used to model the flow and deformation of cartilage matrix. Kuei *et al.*[53] and Eisenfeld *et al.*[26] have also used the biphasic model for cartilage to study the creep and relaxation behaviour of articular cartilage. The interaction of synovial fluid and cartilage was also discussed by Mow and co-workers[67,89]. Lai and Mow[55] have discussed ultrafiltration of synovial fluid by cartilage and it is pointed out that when a certain critical concentration of macromolecules in the solution is reached, a three-dimensional network formation (gelling) takes place as experimentally observed by Maroudas[67,69]. This process of ultrafiltration is significant in the context of boosted lubrication as well as to explain boundary lubrication. Kenyon[50] has analysed the flow of synovial fluid in cartilage to determine when the surface flow becomes less favourable than a flow predominantly parallel but entirely within the cartilage.

Some analytical studies considering the non-Newtonian behaviour of synovial fluid have also been reported. Shukla and Chandra[107] presented mathematical analysis where synovial fluid was characterized by power law model and cartilage as poro-elastic material with rough surface (the roughness was represented by a superposition of cosine waves with small amplitudes). The modified Reynolds equations governing pressure were derived and the effects of various parameters such as permeability, elasticity and roughness were discussed separately.

Tandon and Jaggi[116] have considered the concentration effect of hyaluronate molecules while presenting a micropolar fluid model for synovial fluid. Tandon and Agrawal[115] have considered viscoelastic model for synovial fluid.

Chandra[11] has accounted for surface roughness of the cartilage in his analysis. Micropolar fluid model is used for suspension behaviour of synovial fluid and stochastic roughness model is presented. It has been shown that

as the fluid looses its micropolar behaviour, squeeze time decreases. A model for roughness is also suggested.

ACKNOWLEDGEMENT

The author wishes to thank the Department of Science and Technology, Government of India for sponsoring the ongoing research project 'Studies in Biomechanics'.

REFERENCES

1. Anadere, I., Chmiel, H. and Laschner, W. (1979) *Biorheology*, 16 : 179.
2. Anderson, C. E. (1962) *AAOS Instructional Course Lecture*, 44 (A).
3. Balazs, E. A. and Gibbs, D. A. (1970) *In* Chemistry and Molecular Biology of the Intercellular Matrix (Ed. E. A. Balazs), Academic Press, 1241.
4. Balazs, E. A., Watson, D., Duff, I. E. and Roseman, S. (1967) *Arth. Rheum.*, 10 : 357.
5. Barnett, C. H., Davies, D. V. and MacConaill, M. A. (1961) Synovial Joints—their Structure and Mechanics, Longman Green, London.
6. Bloch, B. and Dintenfass, L. (1963) *Aust. N. Z. J. of Surg.*, 33 : 108.
7. Brower, T. D., Akadoshi, Y. and Orlic, P. (1962) *J. Bone Jt. Surg.*, 44A, 456.
8. Caygil, J. C. and West, G. H. (1969) *Med. and Biol. Engg.*, 7 : 507.
9. Chandra, Peeyush (1975) Ph.D. Thesis, I.I.T. Kanpur.
10. Chandra, Peeyush (1980) P. L. Bhatnagar Memorial Vol., Nat. Acad. Sci. India, 265.
11. Chandra, Peeyush (1980) Proc. Nat. Acad. Sci. India, Series A (*to appear*).
12. Charnley, J. (1959) *New Scientist*, July 9.
13. Clarke, I. C. (1971) *Ann. Rheum. Dis.*, 30 : 15.
14. Clarke, I. C. (1971) *J. Bone Jt. Surg.*, 53B : 732.
15. Clarke, I. C. (1973) *J. Microscopy*, 97 (3) : 309.
16. Davies, D. V. (1967) *In Proc. Inst. Mech. Engrs.*, 181 (3J) : 25.
17. Davies, D. V. and Palfrey, A. J. (1969) *In* Lubrication and Wear in Joints (Ed. V. Wright), Lippincott Co., 20.
18. Davies, W. H., Lee, S. L. and Sokoloff, L. (1979) *J. Biomech. Eng.*, (*Trans. ASME*) 101 : 185.
19. Dinnar, U. (1976) *Ann. Biomed. Eng.*, 4 : 91.
20. Dintenfass, L. (1963) *J. Bone Jt. Surg.*, 45A : 1241.
21. Dintenfass, L. (1966) *Fed. Proc.*, 25 : 1054.
22. Dowson, D. (1967) *In Proc. Inst. Mech. Engrs.*, 181 (3J) : 45.

23. Dowson, D., Longfield, M. D., Walker, P. G. and Wright, V. (1968) *Proc. Inst. Mech. Engrs.*, 182 (3N) : 70.
24. Dowson, D., Unsworth, A. and Wright, V. (1970) *J. Mech. Eng. Sci.*, 12 (5) : 364.
25. Edwards, J. (1967) *In Proc. Inst. Mech. Engrs.*, 181 (3J) : 16.
26. Eisenfeld, J., Mow, V. C. and Lipshitz, H. (1978) *Math. Biosci.*, 39 : 97.
27. Elmore, S. M., Sokoloff, L., Narris, P. and Carmeci, P. (1963) *J. Appl. Physiol.*, 18 : 393.
28. Fein, R. S. (1967) *In Proc. Inst. Mech. Engrs.*, 181 (3J) : 125.
29. Ferguson, J., Boyle, J. A., McSween, R. N. M. and Jasani, M. K. (1968) *Biorheology*, 5 : 119.
30. Ferguson, J. and Nuki, G. (1971) *Rheol. Acta.*, 10 : 15.
31. Fessler, J. H. (1957) *Nature* (*London*), 179 : 126.
32. Fessler, J. H. (1960) *Biochem. J.*, 76 : 124.
33. Freeman, M. A. R. (1973) Adult Articular Cartilage, Gune and Stratton, N.Y.
34. Gardner, D. L. (1972) *Ann. Rheum. Dis.*, 31 : 235.
35. Gardner, D. L. and McGillivery, I. (1971) *Ann. Rheum. Dis.*, 30 : 3.
36. Goodfellow, J. W. and Bullough, P. G. (1967) *J. Bone Jt. Surg.*, 49B : 175.
37. Hamerman, D., Rosenberg, L. C. and Schubert, M. (1970) *J. Bone Jt. Surg.*, 52B : 725.
38. Hamerman, D. and Schubert, M. (1962) *Am. J. Med.*, 33 : 555.
39. Hayes, W. C. and Mockros, L. F. (1971) *J. Appl. Physiol.*, 3 : 562.
40. Higginson, G. R. (1977) *In Proc. Inst. Mech. Engrs.*, 191 33/77 : 217.
41. Higginson, G. R. and Norman, R. (1974) *J. Mech. Eng. Sci.*, 16 (4) : 250.
42. Higginson, G. R. and Norman, R. (1974) *Phy. Med. Biol.*, 19 (6) : 758.
43. Hori, R. Y. and Mockros, L. F. (1976) *J. Biomech.*, 9 : 259.
44. Jebens, E. H. and Monk-Jones, M. E. (1959) *J. Bone Jt. Surg.*, 41B : 338.
45. Johnson, G. R. (1974) Ph.D. Thesis, University of Leeds.
46. Johnson, G. R., Dowson, D. and Wright, V. (1977) *Int. J. Mech. Sci.*, 19 (5) : 301.
47. Jones, E. S. (1934) *Lancet*, 1 : 1426.
48. Kempson, G. E., Freeman, M. A. R. and Swanson, S. A. V. (1971) *J. Biomech.*, 4 : 239.
49. Kenedi, R. M., Gibson, T., Daly, C. H. and Abraham M. (1966) *Fed. Proc.*, 25 : 1084.
50. Kenyon, D. E. (1980) *J. Biomech.*, 13 : 129.
51. Kuei, S. C. (1978) Ph.D. Thesis, Rensselear Poly. Inst. Troy.
52. Kuei, S. C., Mow, V. C., Lai W. M. and Ancona, M. G. (1978) *Trans. Orthop. Res. Soc.* 24th, 3 : 10.
53. Lai, W. M., Kuei, S. C. and Mow, V. C. (1978) *J. Biomech. Eng.* (*Trans. ASME*), 100 : 169.
54. Lai, W. M., Kuei, S. C. and Mow, V. C. (1977) Proc. 2nd Ann. Eng. Mech. Div. Spec. Conf., 166.
55. Lai, W. M. and Mow, V. C. (1978) *J. Eng. Mech. Div., ASCE*, 104 : 79.
56. Ling, F. F. (1974) *J. Lub. Tech.* (*Trans. ASME*), 96F (3) : 449.
57. Linn, F. C. (1967) *J. Bone Jt. Surg.*, 49 A : 1079.

58. Linn, F. C. and Radin, E. L. (1968) *Arth. Rheum.*, **11** : 674.
59. Linn, F. C. and Sokoloff, L. (1965) *Arth. Rheum* , **8** : 481.
60. Little, T., Freeman, M. A. R. and Swanson, S. A. V. (1969) in Lubrication and Wear in Joints (Ed. V. Wright), Lippincott Co., 110.
61. MacConaill, M. A. (1932) *J. Anat.*, **66** : 210.
62. MacConaill, M. A. (1966) *Lab. Prac.*, **15** (1) : 60.
63. MacConaill, M. A. (1967) *In Proc. Inst. Mech. Engrs.*, **181** (3J) : 1.
64. Mansour, J. M. and Mow, V. C. (1976) *J. Bone Jt. Surg.*, **58A** : 509.
65. Mansour, J. M. and Mow, V. C. (1977) *J. Lub. Tech. (Trans. ASME)*, **99F** : 163.
66. Mansour, J. M., Mow, V. C. and Redler, I. (1973) *In* Proc. New England Conf. on Bioeng , Univ. of Vermount, 183.
67. Maroudas, A. (1967) *In Proc. Inst. Mech. Engrs.*, **181** (3J) : 122.
68. Maroudas, A. (1968) Ph.D. Thesis.
69. Maroudas, A. (1969) *In* Lubrication and Wear in 4 Joints (ed. V. Wright), Lippincott Co., 124.
70. Maroudas, A. and Bullough, P. G. (1968) *Nature*, **219** : 1260.
71. Maroudas, A., Bullough, P. G., Swanson, S. A. V. and Freeman, M. A. R. (1968) *J. Bone Jt. Surg.*, **50B** : 166.
72. McCall, J. (1969) *In* Lubrication and Wear in Joints (Ed. V. Wright), Lippincott Co., 39.
73. McCutchen, C. W. (1959) *Nature*, **184** : 1284.
74. McCutchen, C. W. (1962) *Wear*, **5** : 1.
75. McCutchen, C. W. (1966) *Fed. Proc.*, **25** (3) : 1061.
76. McCutchen, C. W. (1967) *In Proc. Inst. Mech. Engrs.*, **181** (3J) : 55.
77. McCutchen, C. W. (1969) *Clinical Orth.*, **64** : 18.
78. McCutchen, C. W. (1969) *In* Lubrication and Wear in Joints (Ed. V. Wright), Lippincott Co., 117.
79. McCutchen, C. W. (1978) *In* the Joints and Synovial Fluid, **1** (Ed. L. Sokoloff), Academic Press, 407.
80. Meachim, G. and Roy, S. (1969) *J. Bone Jt. Surg.*, **51B** : 529.
81. Meachim, G. and Stockwell, R. A. (1973) *In* Adult Articular Cartilage led M. A. R. Freeman, Gune and Stratton, N.Y. 1.
82. Meyer, K., Smyth, E. M. and Dowson, M. D. (1939) *J. Biol. Chem.*, **128** : 319.
83. Morrison, J. B. (1968) *Bio-med. Eng.*, **3** : 164.
84. Mow, M. C. and Ling, F. F. (1969) *ZAMP*, **20** : 156.
85. Mow, V. C. (1969) *J. Lub. Tech. (Trans ASME)*, **91F** : 320.
86. Mow, V. C. and Lai, M. W. (1979) *Ann. Rev. Fluid Mech.*, **11** : 247.
87. Mow, V. C., Lai, M. W. and Redler, I. (1974) *J. Biomech.*, **7** : 449.
88. Mow, V. C., Lai M. W., Eisenfeld, J. and Redler, I. (1974) *J. Biomech.*, **7** : 457.
89. Mow, V. C., Mansour, J. M. and Redler, I. (1974) *In* Proc. 1974 Symp. Bioengg. (ASME).
90. Muir, H., Bullough, P. and Maroudas, A. (1970) *J. Bone Jt. Surg.*, **52B** : 554.
91. Mulholland, R., Millington, P. F. and Manner, J. (1975) *Ann. Rheum. Dis.*, **34** : 104.

92. Myers, R. R., Nigami, S. and White, R. K. (1966) *Biorheology*, **3** : 197.
93. Negami, S. (1964) Ph.D. Thesis, Lehigh Univ., Penn.
94. Ogston, A. G. and Stanier, J. E. (1950) *Bio Chem. J.*, **46** : 364.
95. Ogston, A. G. and Stanier, J. E. (1953) *J. Physiol.*, **119** : 244.
96. Palfrey, A. J. (1972) *In* Perspective in Biomed. Eng. (Ed. R. M. Kenedi) MacMillan, 139.
97. Paul, J. P. (1969) *In* Lubrication and Wear in Joints (Ed. V. Wright), Lippincott Co., 77.
98. Radin, E. L. and Paul, I. L. (1972) *J. Bone Jt. Surg.*, **54A** : 607.
99. Radin, E. L., Paul, I. L. and Pullock, D. (1970) *Nature*, **226** : 554.
100. Radin, E. L., Paul, I. L., Swann, D. A. and Schottstaead, E.S. (1971) *Ann. Rheum. Dis.*, **30** : 332.
101. Reynolds, O. (1886) *Phillos. Trans. Roy. Soc. London*, **177** : 157.
102. Ropes, M. W. and Bauer, W. (1953) Synovial Fluid Changes in Joint Disease, Harvard Univ. Press, Cambridge.
103. Sayles, R. S., Thomas, T. R., Anderson, J., Haslock, I. and Unsworth, A. (1979) *J. Biomech.*, **12** : 257.
104. Scott Blair, G. W. (1974) An Introduction to Biorheology, Elsevier Scientific Publishing Co., Ch. VIII, 103.
105. Sellers, P. C., Dowson, D. and Wright, V. (1970) at the Conf. on Rheol. in Med. and Pharm., British Soc. Rheol., London.
106. Simmon, W. H. (1971), *J. Biomech.*, **4** : 379.
107. Shukla, J. B. and Chandra P. (1975) at the Physiological Fluid Dynamics Winter School, I.I.T. New Delhi.
108. Shukla, J.B., Chandra, P. and Srivastava, R. C. (1972) Proc. First World Conf. in Industrial Tribology, New Delhi (Paper No. C-4).
109. Snidle, R. W. and Biswas, S. (1977) *In* Biomech. Symposium ASME, AMD **23** : 171.
110. Sokoloff, L. (1963) *Science*, **41** : 1055.
111. Sokoloff, L. (1966) *Fed. Proc.*, **25** : 1089.
112. Swann, D. A. (1978) *In* the Joints and Synovial Fluid (Ed. L. Sokoloff), Academic Press, 407.
113. Swanson, S. A. V. (1973) *In* Adult Articular Cartilage (Ed. M. A. R. Freeman) Gune and Stratton, N.Y.
114 Swanson, S. A. V. (1975) *In* Recent Advances in Orthopaedics (Ed. B. McKibbin) Churchill Livington, London.
115. Tandon, P.N. and Agrawal, Kusum (1979) *Ind. J. Pure* and *Appl. Math.*, **10**(1) : 43.
116. Tandon, P. N. and Jaggi, S. (1977) *Proc. Ind. Acad. Sci.*, **85A** : 144.
117. Tanner, R. I. (1966) *Phy. Med. Biol.*, **11** : 119.
118. Thomas, T. R., Sayles, R. S., Haslock, I. (1980) *J. Biomech. Eng. (Trans. ASME)*, **102** : 50.
119. Torzilli, P. A. (1976) *In* Hand book of Eng. in Med. & Biol., 225.
120. Torzilli, P. A. and Mow, V. C. (1976) *J. Biomech.*, **9** : 541.
121. Torzilli, P. A. and Mow, V. C. (1976) *J. Biomech.*, **9** : 587.

122. Unsworth, A. (1976) Ph.D. Thesis, Univ. of Leeds.
123. Vos, R. and Theyse, F. (1969) *In* Lubrication and Wear in Joints (Ed. V. Wright) Lippincott Co., 29.
124. Walker, P. S. (1969) Ph.D. Thesis, Univ. of Leeds.
125. Walker, P. S. and Hajek, J. V. (1972) *J. Biomech.*, 5 : 581.
126. Walker, P. S., Dowson, D., Longfield, M. D. and Wright, V. (1968) *Ann. Rheum. Dis.*, 27 : 512.
127. Walker, P. S., Sikorski, J., Dowson, D., Longfield, M. D., Wright, V. and Buckley, T. (1969) *Ann. Rheum., Dis.*, 28 : 1.
128. Weiss, C., Rosenberg, L. and Helfet, A. J. (1968) *J. Bone Jt. Surg.*, 50A : 663.
129. White, R. K. (1963) *Clinical Orthop.*, 29 : 213.
130. White, R. K. (1963) *J. Bone Jt. Surg.*, 45A : 1084.
131. Woo, S., Akeson, W. H., Jemmott, G.F., Solomon, N.G. and Lubock, P.(1977) *In* Proc. Biomech. Symp. ASME, AMD 23 : 155.
132. Wright, V. and Dowson, D. (1976) *J. Anat.*, 121 : 107.
133. Wright, V., Dowson, D. and Kerr, J. (1973) in Int. Rev. of Connective Tissue Research, 6 : 105.
134. Wright, V., Dowson, D. and Seller, P. C. (1971) *In* Modern Trends in Rheumatology, Chapter 2.

Bicarbonate Stimulation of Electron Flow in Thylakoids

Doug Jordan* and Govindjee†

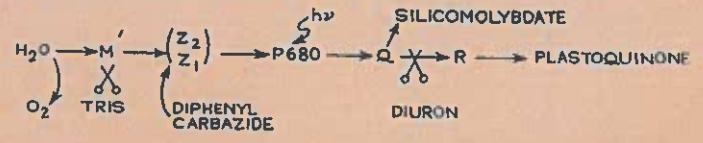
Departments of Physiology & Biophysics,† Botany† and Agronomy,*
University of Illinois, Urbana, IL. 61801 (USA)

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

In addition to being the source of fixed carbon in photosynthetic carbon assimilation, bicarbonate* serves as a reversible *in vitro* activator of three major photosynthetic processes : (I) The carbon dioxide fixing enzyme, RUBP carboxylase, is known to be activated by bicarbonate[12, 13, 15]. The level of bicarbonate required for half maximal activity has been shown to depend on the concentration of the essential metal activator, the pH, and the presence of a number of organophosphate activators. Though bicarbonate has been treated as an essential activator, the absolute requirement for bicarbonate has not been clearly documented. (II) Photophosphorylation rates and the P/e ratio, as studied in broken chloroplasts, are stimulated by the addition of bicarbonate[1, 2, 16, 17]. Bicarbonate enhancement of phosphorylation was most pronounced in a photosystem II** (PS II) independent cyclic phosphorylation system, as induced by pyocyanin. In a similar system, bicarbonate was shown to increase the levels of "energetic intermediates" of ATP synthesis[1, 2]. These effects have been noted to have a seasonal dependence with respect to the source of chloroplasts, and this is thought to be related to the hormonal status of the plant material[1]. A seasonally independent bicarbonate effect has been demonstrated with purified coupling factor I[14], in which bicarbonate stimulates the rates of magnesium and calcium ATPase activities. Bicarbonate is a nonessential activator in these systems since complete removal of bicarbonate does not eliminate phosphorylation ability. In addition, several carboxylic acids are nearly as effective as bicarbonate in stimulating coupling factor I activities[14], so the bicarbonate effect may be more appropriately categorized as a nonspecific

* Except where the question of the activating species is addressed, the word bicarbonate is used without implying that it is the activating species.

** Photosystem II of green plants consists of the following components : a light-harvesting complex that feeds energy to a reaction center chlorophyll *a* labelled as P 680 ; an electron acceptor quinone molecule labelled as Q ; a 2-electron acceptor quinone molecule, labelled as B or R, that accepts electrons from Q; two electron donors[7] Z_1 and Z_2 that donate electrons to P 680$^+$; and an oxygen evolving Mn-containing enzyme (M) that receives electrons from H_2O. The sequence is written as :



activation by carboxylic acids. (III) Bicarbonate activation of the Hill reaction (electron flow from H_2O to an artificial electron acceptor) was the first of these bicarbonate activated systems to be discovered[26, 27, 28] and in recent years a considerable amount of scientific effort has been devoted to understanding this phenomenon. In comparison to photophosphorylation, the Hill reaction is activated by much lower concentrations of bicarbonate, and the requirement for bicarbonate appears to be both specific and absolute. Through the use of artificial electron donors and acceptors combined with several physical methods, the sites of bicarbonate activation in PS II electron transport and on the PS II reaction center have been systematically elucidated. The following paragraphs are devoted to reviewing the experimental evidence* characterizing the effects of bicarbonate on PS II activity and defining some of the questions that await experimentally determined solutions. The terminology used throughout the text is based on a recent review of this topic[5]. The term *thylakoid* refers to membranes isolated from chloroplasts.

The essentiality of bicarbonate for Hill reaction activity may be inferred from studies which show that bicarbonate depletion results in a 90-95% loss in Hill reaction rates[4, 8, 9, 10, 19-23, 30]. The results of recent studies[11, 20] indicate that the residual activity is probably due to incomplete removal of bound bicarbonate. The existence of high affinity and low affinity bicarbonate binding sites in chloroplast fragments was demonstrated. The removal of tightly bound bicarbonate was shown to be a slow process when chloroplast fragments were incubated under low salt, and neutral pH conditions. Removal of labeled bicarbonate was found to be facilitated by washing the chloroplasts with a low pH buffer containing high levels of sodium chloride and sodium formate, which had been found to be effective in depleting chloroplasts of Hill reaction activity in earlier bicarbonate activation studies[3, 4, 8, 10, 19-23, 30]. By washing bicarbonate labeled chloroplasts with neutral pH, low salt buffer, the high affinity bicarbonate sites could be selectively preloaded with labeled [^{14}C] bicarbonate. Chloroplasts were prepared in this fashion and then various treatment washes were applied to samples of the labeled preparation. The treatments removed bicarbonate to varying degrees and Hill reaction activity was found to correlate well with the amount of label which remained bound. These results are important for two reasons.

* This review includes papers published up to 1979.

First, the correlation of bound bicarbonate with Hill activity provides firm physical evidence for bicarbonate activation. Second, the demonstrated effectiveness of the high salt and low pH wash medium in removing the tightly bound bicarbonate activator confirms that the effectiveness of this treatment is related to its ability to remove the effector and not to some sort of ion inhibition.

A previous study had shown that the requirement for salt treatment of chloroplasts was restricted to the bicarbonate depletion medium, and that bicarbonate sensitivity was observable in reaction media with or without high levels of salt[21]. It was also known that the effect of 100 mM formate in the incubation mixture increased the concentration of bicarbonate needed for half saturation of Hill activity by an order of magnitude[10], suggesting a competition by formate for bicarbonate binding sites.

The tight binding and slow exchange of bicarbonate at the activating sites explains the variability of early demonstrations of bicarbonate enhancements. Warburg[26,27,28] and several other researchers[3,24,28,29] had found it necessary to incubate chloroplast preparations for two or three hours, in the presence of bicarbonate traps, in order to observe bicarbonate enhancements of Hill reaction activity. These long-term incubations (some of which were conducted in the light) produced irreversible inactivation of the chloroplasts and probably were not totally effective in removing the activating species under investigation. So, the early reports are only useful for qualitative evaluations.

The universality of the bicarbonate effect can be inferred from the studies by independent laboratories, which used different higher plant and algal chloroplast material in their investigations. This point was emphasized by Stern and Venessland[25] who showed that the degree of bicarbonate enhanced Hill activity was the same in several different species.

In order to address criticisms that suggested that the bicarbonate effects were due to photodestruction of chloroplast activity during prolonged treatments, Good[8] showed that added bicarbonate was effective in restoring Hill activity to bicarbonate depleted chloroplasts, which had been pretreated with or without light. More recent studies have demonstrated the reversibility of each of the observed effects of bicarbonate. Good[8] and Stemler and Govindjee[21] noted that bicarbonate was more effective if the chloroplasts were incubated in the dark prior to measuring Hill reaction activity. Stemler reports that this is due to the preferred binding of bicarbonate to chloroplast

having the two electron acceptor quinone, R, in the oxidized state (as it would be in the dark).

The specificity for bicarbonate appears to be absolute. A number of carboxylic acids and other anions have been shown to be ineffective in raising Hill reaction activity in the absence of bicarbonate[8]. Additionally, bicarbonate has been shown to be a poor effector of the chloride activator, which has been shown to be located on the oxidizing side of PS II[9]. As mentioned earlier, bicarbonate activation of the Hill reaction has been differentiated from the bicarbonate activation of photophosphorylation. The lack of specificity for the activator of photophosphorylation is one major difference. In addition, the Hill reaction bicarbonate enhancement is more pronounced in uncoupled chloroplasts[3,10,29] ; while, bicarbonate stimulation of photophosphorylation is greater in a PS I cyclic phosphorylation system.

The bicarbonate effect on the Hill reaction was obverved to be more pronounced at high light intensities by most investigators[3,6,19,23,20]. The effect at low light intensities was less defined and some of the experimental results are still unexplainable. For example. West and Hill[29] found that a small bicarbonate effect was observed at low light intensities when DCPIP (2, 6 dichlorophenol indophenol) was the Hill oxidant, but no effect of bicarbonate was observed using ferricyanide. Stemler and Govindjee[23] observed the same bicarbonate dependence at low light intensities as at high light intensities when the ferricyanide Hill reaction was measured amperometrically, but when the same reaction was measured polarographically a definite light dependency was evident[23]. The larger bicarbonate effect at high light intensities indicated a bicarbonate requirement for a dark reaction of PS II. Further definition of this effect and convincing evidence for an effect on PS II chemistry was provided by measuring partial Hill reactions and transient intermediates using chemical and physical methods[9,10].

There are three lines of evidence supporting a bicarbonate enhancement of PS II chemistry. First, when oxygen flash yields were measured following widely spaced (1 s) saturating light flashes, bicarbonate depleted chloroplasts yields were approximately half those of control and bicarbonate replenished chloroplasts[23]. This approximated the bicarbonate enhancement of the DCPIP Hill reaction at low light intensities in the same study. Second, Jursinic *et al.*[8], measuring the electron spin resonance signal II_{vf}, corresponding to oxidized Z_2, found that the signal was depressed by 40% in bicarbonate depleted samples. Since the μs rise kinetics in chlorophyll a fluorescence

was shown to be the same in chloroplasts with or without bicarbonate, the rate constant of electron flow from Z to P-680 was taken to be independent of bicarbonate and the ESR signal depression could be assigned to a difference in PS II reaction center activity. Third, the levels of reduced quinones (mainly Q and R at 334 nm) were measured by absorption spectroscopy following a rapid succession of saturating light flashes[19]. Bicarbonate depletion resulted in an initial absorption decrease of 40%, indicating that the PS II reaction centers were inhibited.

The bicarbonate effect on the dark reactions was originally thought to involve the oxidizing side of PS II[21,22,28]. When chloroplasts were heat treated in order to inactivate oxygen evolution capability, bicarbonate produced no effect on the partial Hill reaction going from DPC (diphenyl carbazide) to DCPIP. DPC donates electrons to Z_1, so bicarbonate was thought to activate between the oxygen evolving site and Z_1. A later study by Wydrzynski and Govindjee[30] demonstrated that when alkaline Tris washing was used to inactivate the oxygen evolving system, the rate of DCPIP reduction, using DPC as the donor, was increased by more than twofold when bicarbonate was present. Control chloroplasts, not treated with Tris, showed a tenfold bicarbonate enhancement of DCPIP reduction. The discrepancy in the two studies may be explained by the report[30] that heat treatments irreversibly inactivate the bicarbonate activating site, whereas Tris treatment does not. There is an, as yet, unexplained activation of DCPIP reduction by the electron donor DPC, which must account for the reduced amount of bicarbonate activation seen when it is used.

Khanna *et al.*[10] established the assignment of a major bicarbonate dark effect to the reducing side of PS II by showing that the partial Hill reaction from water to silicomolybdate was not affected by bicarbonate. (Silicomolybdate accepts electrons from Q.) Electron flow from water to DAD (diaminodurene) was limited by bicarbonate depletion (to the same extent as ferricyanide reduction) and it was concluded that electron flow between Q and PQ (plastoquinone pool) was retarded.

Further evidence for a bicarbonate effect past Q was shown by variable chlorophyll a fluorescence measurements. The chlorophyll a florescence transient was shown to be dramatically altered by bicarbonate (reviewed in ref. 5). Bicarbonate depleted chloroplasts exhibit an initial rapid rise in chlorophyll a fluorescence (O → I) followed by a slow second phase (I → P); the first phase gives the appearance of a process as is seen with diuron

poisoning. (Diuron blocks electron flow from Q- to R). Bicarbonate restored the normal biphasic florescence pattern by slowing the initial phase and accelerating the second phase. (For further explanation, see ref. 5.) Wydrzynski and Govindjee[30] showed that Tris treated chloroplasts, supplied with a variety of electron donors, still required bicarbonate in order to restore a normal biphasic florescence transition pattern, thus indicating a bicarbonate requirement on the reducing side of PS II.

The dark decay kinetics of the ESR signal II_{vf} measuring oxidized Z_2, was found to be unchanged by the presence of bicarbonate, indicating that the rate constant of the 10 ns electron flow was not limited from H_2O to Z_2 and that the bolck was on the reducing side of PS II[8].

Chlorophyll a fluorescence yield decay measurements were made as another test for a block after Q. This method followed the fluorescence yield of chloroplast preparations (using a weak nondisturbing light source) after treatment with a short (10 ns), saturating light flash. Bicarbonate depleted samples were found to have fluorescence yield decay half-lives of about 3 ms, while bicarbonate sufficient chloroplast half-lives were about 0.5 ms[8]. Siggel *et al.*[19], using absorption spectrescopy, found that the half life for the reoxidation of Q^- in bicarbonate depleted chloroplasts was extended to about 10 ms from the 0.5 ms half-life observed in bicarbonate sufficient samples. The observed retardation of Q^- oxidation matches the measured extension in S_2 state relaxation times ($S'_2 \rightarrow S_2$) determined by measuring oxygen yield following three light flashes with variable spacing between the first and second flashes[28]. (The S states represent the status of the PS II complex including the oxygen evolving system and Q.) It was reasoned that the oxidation of Q was rate limiting in the conversion of $S'_2 \rightarrow S_2$.

To further explore the effect of bicarbonate depletion on electron transport, diuron induced chlorophyll a fluorescence change was measured as a function of pre-flash number[4]. In this system, bicarbonate sufficient chloroplasts show an oscillating fluorescence pattern with high yields on odd flashes and low yields on even flashes. Bicarbonate depleted chloroplasts had high fluorescence yields without an oscillating pattern. This was interpreted as a block between the R and PQ step. In further experiments without diuron addition, it was shown that the Q^-R^{2-} state could be accumulated by giving three short light flashes to bicarbonate-depleted thylakoids. Measuring chlorophyll a fluorescence, the half-time of the Q^-R^{2-} decay rate was found to be 100 to 200 ms in bicarbonate depleted samples, while it was approximately

1 ms in bicarbonate sufficient chloroplasts, thus indicating a major bottleneck in electron transport alleviated by bicarbonate addition. Siggel *et al.*[9] measured the decay of the reduced quinone pool (R^{2-} and PQ^{2-}) using absorption measurements and found that control chloroplasts had a half-time of decay of 26 ms, while bicarbonate-depleted chloroplasts had a half-time of 100 ms. Thus, in bicarbonate depleted chloroplasts the rate limiting step in electron transport was large enough to account for the overall inhibition of Hill reaction activity.

Evidence cited above has shown that bicarbonate is a reversible activator of the PS II reaction center and of two points within electron transport on the reducing side of PS II. Bicarbonate activation appears to be an essential requirement for the Hill reaction, and specificity for bicarbonate appears to be absolute. Good[3] implied that the active species was the bicarbonate ion because of the synergism of chloride and formate ions with the bicarbonate ions is producing a greater enhancement of Hill reaction activity. More recent evidence[10] has shown that formate acts as a competitive inhibitor with respect to bicarbonate in the activation of the Hill reaction, supporting the contention that the ionic species may be the activator. It would be of interest to know whether formate is binding to the same site. Other carboxylic acids with different pka's may be useful in determining whether the activator site binds ions. Kinetic experiments determining the rate of activation by bicarbonate as compared to carbon dioxide may be difficult since: (I) the pH optimum for the Hill reaction is close to the pKa of carbonic acid ; (II) the rate of binding may be slow so the ionic and gaseous species equilibrate before full activation is achieved ; and (III) the activating site may be buried within the membrane allowing only gaseous carbon dioxide to penetrate. In this case, the activation kinetics would measure the rate of membrane penetration and not the reactivity to the activating site. Sarojini and Govindjee[10] have recently observed that when CO_2 is injected to bicarbonate depleted chloroplasts kept at 5 °C, there is an almost instantaneous stimulation of the Hill reaction, whereas if bicarbonate ion is injected, there is an approximate lag of eight seconds. These results suggest that CO_2 may be the active species that binds to the membrane and that the binding site may be closer to the outer than the inner side of the membrane. Khanna *et al.*[11] have speculated that the effect of CO_2 on the photochemical properties of PS II occurs via CO_2 binding to a polypeptide of the PS II complex which functions as an electron carrier on the reducing

side of PS II. This suggestion is consistent with their observation on the alteration of the binding affinity of ^{14}C-atrazine to the thylakoids upon CO_2-depletion ; atrazine is a PS II inhibitor and has been suggested to bind to the "R-protein" located on the outer (not the inner) side of thylakoid membrane.

Since the activating site binds bicarbonate tightly, it lends itself to active site directed reagent and bicarbonate protection studies. [^{14}C] labeled cyanate is available commercially and could prove useful in purifying a labeled activating site. Another possible method for labeling the site permanently is to bind [^{14}C] bicarbonate to the high affinity sites and stabalize the attachment with a methylating agent, such as diazomethane. The labeled moieties could then be purified by standard methods.

Whether bicarbonate activation of PS II activity is important *in vivo* is difficult to judge. The tight binding of bicarbonate would seem to assure full activation under most conditions. Yet, the binding affinity is known to be inhibited by high pH and by formate ions. Further evidence with regard to other modulators of bicarbonate binding will shed light on the possible importance of bicarbonate activation of the Hill reaction *in vivo*.

The preparation of this short review was supported by financial assistance from USDA (5901-40) and NSF (24532).

REFERENCES

1. Batra, P.O. and Jagendorf, A. T. (1965) *Plant Physiol.*, 40 : 1074-1079.
2. Cohen, W. S. and Jagendorf, A. T. (1974) *Plant Physiol.*, 53 : 220-223.
3. Good, N. E. (1963) *Plant Physiol.*, 38 : 298-304.
4. Govindjee, Pulles, M. P. J., Govindjee, R., Van Gorkom, H. J. and Duysens, L. N. M. (1976) *Biochim. Biophys. Acta.*, 449 : 602-605.
5. Govindjee and Van Rensen, J. J. S. (1978) *Biochim. Biophys. Acta.*, 505 : 183-213.
6. Izawa, S. (1962) *Plant Cell Physiol.*, 3 : 221-227.
7. Jursinic, P. and Govindjee (1977) *Biochim. Biophys. Acta.*, 461 : 253-267.
8. Jursinic, P., Warden, J. and Govindjee (1976) *Biochim. Biophys. Acta.*, 440 : 322-330.
9. Kelly, P.M. and Izawa, S. (1978) *Biochim. Biophys. Acta.*, 502 : 198-210.
10. Khanna, R., Govindjee and Wydrzynski, T. (1977) *Biochim. Biophys. Acta.*, 462 : 208-214.

11. Khanna, R., Pfister, K., Van Rensen, J. J. S., Keresztes, A. and Govindjee (1980) *Biochim. Biophys. Acta.*, (*in press*).
12. Laing, W. A., Ogren, W. L. and Hageman, R. H. (1974) *Plant Physiol.*, 54 : 678-685.
13. Lorimer, G. H., Badger, M. R. and Andrews, T. J. (1976) *Biochim.*, 15 : 529-536.
14. Nelson, N., Nelson, H. and Racker, E. (1972) *J. Biol. Chem.*, 247 : 6506-6510.
15. Pon, N. G., Rabin, B. R. and Calvin, M. (1963) *Biochem. Z.*, 388 : 7-19.
16. Punnett, T. (1965) *Plant Physiol.*, 40 : 1283-1284.
17. Punnett, T. and Iyer, R. V. (1964) *J. Biol. Chem.*, 239 : 2335-2339.
18. Sarojini, G. and Govindjee (1980) *Proc. 5th International Congress on Photosynthesis, Greece ; Biochim. Biophys Acta.*, (*in press*).
19. Siggel, U., Khanna, R , Renger, G. and Govindjee (1977) *Biochim. Biophys. Acta.*, 462 : 196-207.
20. Stemler, A. (1977) *Biochim. Biophys. Acta.*, 460 : 511-522.
21. Stemler, A. and Govindjee (1973) *Plant Physiol.*, 52 : 119-123.
22. Stemler, A. and Govindjee (1974) *Plant Cell Physiol.*, 15 : 533-544.
23. Stemler, A., Babcock, G. T. and Govindjee (1974) *Proc. Natl. Acad. Sci.*, 71 : 4679-4683.
24. Stern, B. K., and Vennesland, B. (1960) *J. Biol. Chem.*, 235 : PC 51-53.
25. Stern, B. K. and Vennesland, B. (1962) *J. Biol. Chem.*, 237 : 596-602.
26. Warburg, O. and Krippahl, G. (1958) *Z. Naturforsch.*, 13b : 509-514.
27. Warburg, O. and Krippahl, G. (1960) *Z. Naturforsch.*, 15b : 367-369.
28. Warburg, O., Krippahl, G., Gewitz, H. S. and Volker, W. (1959) *Z. Naturforsch.*, 14b : 712-724.
29. West, J. and Hill, R. (1967) *Plant Physiol.*, 42 : 819-826.
30. Wydrzynski, T. and Govindjee (1975) *Biochim. Biophys. Acta.*, 387 : 403-408.

Note added in proof :

At the proof stage, the present authors became aware of additional new research in the area of this review ; these include : (a) Stemler, A. (1980) *Plant Physiol.*, 65 : 1160-1165 ; (b) Vermaas, W. F. J. and Van Rensen, J. J. S. (1980), Proc. 5th International Congress on Photosynthesis, Greece ; (c) Stemler, A. (1980), Proc. 5th International Congress on Photosynthesis Research ; and (d) Van Rensen, J. J. S. and Kramer, H. J. M. (1979) *Plant Sci. Lett.*, 17 : 21-27. The readers should consult these publications for newer progress in this field.

Female Genital Anatomy and the Morphogenesis of Foetal Membranes of Chiroptera and their Bearing on the Phylogenetic Relationships of the Group

A. Gopalakrishna and K. B. Karim

Department of Zoology, Institute of Science, Nagpur

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

I. Introduction

The bats form one of the largest groups of living mammals including seventeen extant families[114] incorporating about 180 genera[127]. Their unique morphological adaptations suited to their aerial life and their peculiar posture are unmatched by any other mammal and have necessitated in their inclusion in a separate taxonomic Order—Chiroptera. However, it has been recognized in recent years that many mammals, which have similar morphological characters, have been shown to be phylogenetically unrelated, and their apparent morphological similarities are due to convergent adaptations to meet the demands of a common environmental stress or a common mode of life. In the absence of adequate palaeontological record, and because of the extreme morphological specializations of the bats, one is compelled to rely on the conservative anatomical and embryological characters for understanding the phylogenetic relationships of these animals.

The present review is intended to present the available data on the female genital characters and embryology of the bats with a view to understanding the phyletic position of this great Order and the interrelationships of the various families included in this Order.

II. Female Genitalia

A. *Morphology*

The details of the anatomy of the female genital structures of many species of bats have been studied so far[22,24,39,58,59,62,63,85,86,109,128,130,132,133]. From these studies it is evident that no other single mammalian Order presents such a diversity in the construction of the female genitalia as the bats. The various types of female genitalia can be so arranged as to recall the different stages of the evolution of the Mullerian ducts in the mammals.

Jones[58] mentioned, "the internal genitalia of the female exhibit the curiously primitive condition of completely separated uteri". While in some form, such as *Epomophorus* and *Cynopterus*, each uterus possesses its own os uteri which opens into a common vagina[58] in *Cynopterus sphinx gangeticus*[3] the cervix extends to about half of the length of the vagina in the form of a vertical septum and the uterine cornua open by separate cervical canals near the distal end of the septum. In *Rousettus aegyptiaca*[51] and *Rousettus leschenaulti*[63] the caudal regions of the uterine cornua are joined externally, but

each cornu opens into the vagina by a separate cervical canal which runs through the entire length of the bulbous cervix. Among the Microchiroptera in *Taphozous longimanus*[39], whereas the construction of the female genitalia in the nonparous female is similar to that in *Rousettus*[51,63], the tip of the bulbous cervix is lost during the first parturition, and, hence, the two cervical canals join just before opening into the vagina, and this condition persists throughout the adult life of the female. Most other Microchiroptera have a bicornuate uterus, but the two uterine cornua join to form a small, or a more extensive, uterine body due to the union of the caudal segments of the uterine cornua. In these cases the uterine body opens into the vagina by a common cervical canal[22,25,79,129,131,132]. The culmination of the process of union of the two sides of the genitalia is achieved in some of the phyllostomid bats[53,107,137] where the uterus is simplex almost similar to that in the primates. The various types of female genitalia in the bats are illustrated in text figure 1(a–g). The female genital anatomy of *Hipposideros fulvus fulvus*[62] is unique and unparalleled in any mammal. Although externally it appears similar to that of any mammal with a bicornuate uterus, an histological examination reveals that the vagina is itself bicornuate, and the uterine part of the genitalia is represented by a short segment at the cranial end of the bifid vagina. Earlier, Stephens[117] had mentioned the occurrence of cornified epithelium in the uterus of *Tadarida brasiliensis cynocephala*. In the absence of detailed description of the histology of the genital organs of this bat it is not clear if the vagina is bifid in this bat also.

B. *Physiological asymmetry*

Bats are remarkable in exhibiting varying degrees of physiological asymmetry of the female reproductive organs. Except for a few vespertilionids[22,50,79,80,125], in which both sides of the genitalia are equally and concurrently functional, in all other species, possessing a bicornuate uterus there is a pronounced physiological dominance of one side of the genitalia over the other.

This unique feature of the reproductive biology of the bats has been reviewed and discussed by Wimsatt[133]. Among the monotocous bats, whereas the right side of the genitalia is invariably functional in rhinolophids[48,85], *Taphozous melanopogon*[39] and molossids[133], the left side is functional in megadermatid bats[25,39]. The left side of the female genitalia is physiologically dominant over the right side in most hipposiderid bats[44,81]. In a few monotocous bats there is a physiological alternation of the two sides

of the genitalia in successive cycles[26,27,89]. In all the cases mentioned above ovulation and pregnancy occur on the same side. However, there are some species such as *Myotis lucifugus lucifugus*[138] in which, whereas

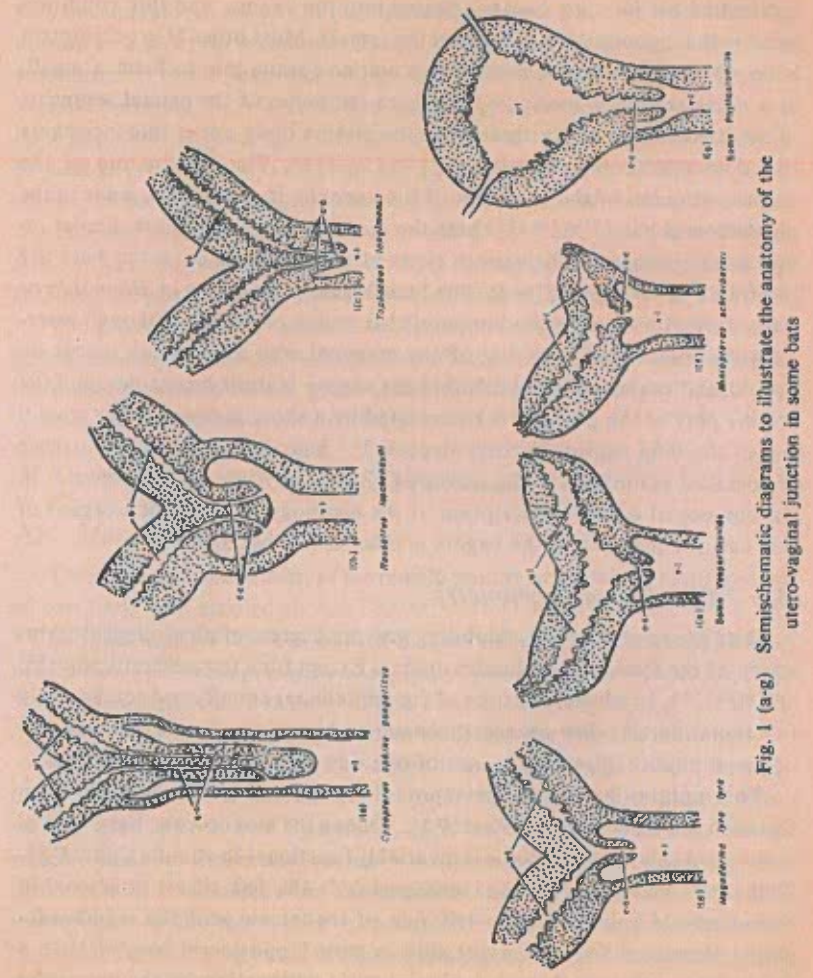


Fig. 1 (a-g) Semischematic diagrams to illustrate the anatomy of the utero-vaginal junction in some bats

ovulation may occur in either of the ovaries with nearly equal frequency, pregnancy is invariably carried in the right uterine cornu. Hence, in this species whenever the ovum is released from the left ovary it migrates to the right cornu for implantation. In *Miniopterus schreibersii fuliginosus*[89] ovulation normally takes place only from the left ovary but the embryo migrates to the right cornu where it implants and undergoes further development. Only in very rare specimens of this species does the right ovary release the ovum and the embryo implants on the ipsilateral side.

III. Preimplantation Development of the Embryo

Our knowledge concerning the preimplantation stages of development of the embryo of bats is confined to only a few species representing eight of the seventeen families[107]. It is remarkable that even among these few species there are wide variations with respect to the stage of development attained by the embryo before it reaches the uterus. The embryo enters the uterus at an early stage of cleavage as in most vespertilionids such as *Myotis lucifugus lucifugus*[128], *Myotis myotis*[115], *Corynorhinus rafinesquei*[97], *Tylonycteris pachypus*[87], *Pipistrellus ceylonicus chrysothrix*[98] and *Scotophilus heathi*[44,72], or as an early or advanced morula as in *Vespertilio murinus*[15], *Pteropus giganteus*[88], *Scotophilus temmincki*[28], *Tadarida brasiliensis*[118,117], *Rhinopoma kinneari*[1], *Rhinolophus rouxi*[4,102], *Hipposideros speoris*[57] and *Miniopterus schreibersii fuliginosus*[10,30]. The embryo develops into the unilaminar blastocyst stage while still in the Fallopian tube in *Desmodus rotundus murinus*[130], *Artibeus jamaicensis*[19], *Glossophaga soricina*[104,105], *Rousettus leschenaulti*[59], *Carollia perspicillata* and *Carollia brevicauda*[8] and *Noctilio albiventor minor*[106]. Probably the period of transit through the Fallopian tube is the major factor which determines the stage of development when the embryo descends into the uterus, apart from factors like the occurrence of delayed implantation or retarded embryonic development due to environmental factors. That the environmental factors play some role in the rate of development of the early embryo has been revealed from the observations of Ramakrishna and Rao[102] and Gopalakrishna and Ramakrishna[100] on *Rhinolophus rouxi* in which the authors showed that the embryonic development proceeds at different rates at different lattitudes. A similar phenomenon has also been reported in *Miniopterus schreibersii* at different Southern lattitudes[16,108].

IV. Blastocyst Uterus Relationship

Typically, the free unilaminar blastocyst (Pl. III, Fig. 9) in bats does not differ from that of other mammals, although minor differences such as the nature of the zona pellucida and the time of hatching of the blastocyst out of the zona may differ in different species. The embryo undergoes implantation when it is in the bilaminar stage of development. The location within the uterus where the blastocyst attaches varies among the different species of bats. Among the species with bicornuate uterus, and in which a single embryo is carried either in only one or in both the cornua of the uterus, the attachment of the embryo is near the cranial end of the uterus as in most megachiropteran species[31,34,82] and in *Taphozous longimanus*[31,46] and *Hipposideros fulvus fulvus*[64] among Microchiroptera, or in the middle of the length of the uterus as in *Megaderma lyra lyra*[31,41] or in the caudal third of the uterus as in *Scotophilus temmincki*[23,31]. In the phyllostomids, which have a simplex uterus, the implantation of the blastocyst occurs intramurally near the opening of the Fallopian tube into the uterus on the fundic side[105,107].

The relationship between the blastocyst and the uterus can be considered under three distinct heads (1) the orientation of the embryonic mass with respect to the morphology of the uterus, (2) the point of first attachment and (3) the depth of implantation or the degree of intimacy of attachment of the blastocyst to the uterine wall. The uterus-blastocyst relationship during implantation has been shown by Mossman[93] to be of considerable systematic and evolutionary significance in mammals. Mossman[93] stated, "The orientation of the (embryonic) disc with reference to the uterus is usually quite constant in all groups known to be closely related." (Parenthesis ours). In a later review on the subject Mossman[95] reiterated this hypothesis and mentioned, "In most major taxonomic groups of eutherian mammals, available data show that at the time of first attachment of the blastocyst to the endometrium the inner cell mass or the embryonic disc has an almost constantly specific directional orientation to the uterus." Regarding the bats Mossman[95] stated, "The literature on the Microchiroptera ('insectivorous' bats) indicates wide variations similar to those in Insectivora. It is particularly confusing since it often indicates lack of consistency even within smaller subgroups, that is, superfamilies and families." He concluded that the Microchiroptera may not violate the general rule, and mentioned,

"instead, we should simply admit that there are insufficient reliable data on which to characterize them".

Table I (*see* at the end of the text) gives the anatomical relationship between the blastocyst and the uterus in representatives of the various families of Chiroptera so far studied. It is evident from the table that some information regarding the early topographical relationship between the blastocyst and the uterus is available for 10 families. The orientation of the embryonic mass is mesometrial in *Cynopterus marginatus*[69], *Cynopterus sphinx gangeticus*[92] and *Rhinopoma kinneari*[38,116]. In *Pteropus giganteus giganteus*[69] the embryonic mass is laterally oriented during early stages but becomes mesometrial later. The embryonic mass is oriented towards the opening of the Fallopian tube in *Rousettus leschenaulti*[59]. It is oriented laterally in *Molossus rufus* and *Molossus obscurus*[110] and *Rhinolophus rouxi*[4]. In *Taphozous longimanus*[28] the embryonic mass is oriented mesometrially during early stages of implantation, but the disc becomes laterally oriented at the neural groove stage of development. Anderson and Wimsatt[2] mentioned that the orientation of the embryonic disc in *Noctilio labialis minor* is "lateral and antimesometrial", and that it is dependant on the orientation of a ridge of the endometrium running the length of each uterine horn. This ridge is lateral and antimesometrial in its location. They further stated, "It should be noted that the orientation is not directly lateral, tending more towards the antimesometrial pole." In *Hipposideros speoris*[57] the embryonic disc is directed towards the side of the uterus which is between the lateral and the antimesometrial sides. It is antimesometrial in all the vespertilionids so far studied[28,65,72,96,99,111,126], *Tadarida brasiliensis cynocephala*[117] and *Desmodus rotundus murinus*[180]. The disc is directed towards the fundic side of the simplex uterus in the phyllostomid bats[58,105,187].

Gopalakrishna and Khaparde[40], who examined several early implanting blastocysts of *Megaderma lyra lyra* at approximately the same stage of development, noticed that the orientation of the embryonic mass is variable in this species being mesometrial in some cases, lateral in some, antimesometrial in some and eccentric in some. A nearly similar situation has been reported also in *Hipposideros fulvus fulvus*[36] and *Rhinolophus rouxi*[100]. In the species, which exhibit either variable orientation of the embryonic mass or in which the orientation of the embryonic disc changes as development advances, the embryonic mass of the blastocyst is perhaps able to migrate

within the blastocyst covering. Such a phenomenon has been reported by Kirby *et al.*[74] and Kirby[78] in the mouse.

The various types of topographical relationships between the blastocyst and the uterus are illustrated in text figure 2. It is evident that the embryonic disc has a tendency for mesometrial orientation in the more primitive families whereas the embryonic disc tends towards antimesometrial orientation in the advanced families. If only this feature is taken into consideration it looks as if this character has not yet stabilized in many families of Chiroptera, and therefore, these families exhibit a variable orientation of the embryonic disc. One other aspect of embryo-uterus relationship needs mentioning here. In all the bats, with the possible exception of *Tadarida brasiliensis cynocephala*[117], the position of the allantoic placental disc bears a specific relationship to the location of the embryonic disc at late neural groove stage of development when the allantoic diverticulum develops. The allantois curves dorsalwards and establishes contact with the chorionic placenta on the dorsal side of the embryonic disc. It is, therefore, reasonable to expect that the orientation of the disc at late neural groove stage is of far greater significance than the orientation of the embryonic mass during early stages of implantation since it is at the late neural groove stage that the allantois develops and establishes contact with the placenta. This feature becomes all the more significant because the location of the placental disc is highly specific even in species in which the embryonic mass has a variable orientation. This appears to lend additional support to the concept that the embryonic mass in some of the species may change its position within the blastocyst covering during early stages of implantation, but the disc acquired its definitive position by the time the embryo reaches the late neural groove stage of development. Hence, at least in the bats, the orientation of the disc at the late neural groove stage is of greater significance than the orientation of the embryonic mass at early attached blastocyst stage of development.

The condition obtaining in *Tadarida brasiliensis cynocephala*[117] appears to be unique and does not follow the pattern noticed in other bats. In this bat Stephens observed, "when implantation is completed, the embryonic disc is located antimesometrially, while the 'pad', which is distinguishable, and which is destined to become the discoidal placenta, is located mesometrially overlying the entrance of the Fallopian tube". In two other molossid bats, *Molossus rufus* and *Molossus obscurus*[110], noted a lateral orientation of the embryonic mass in the early implanted blastocysts. There

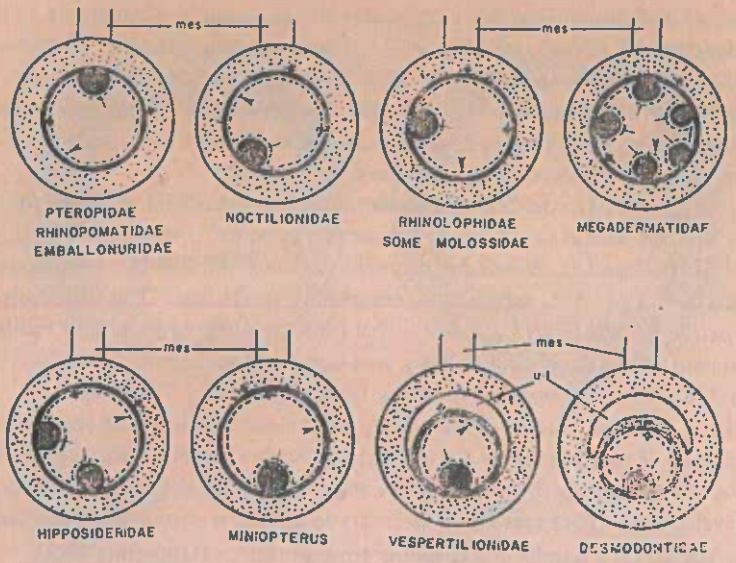


Fig. 2. Schematic drawings to illustrate the nature of implantation in different families of bats. In Megadermatidae and Hipposideridae more than one embryonic mass has been shown to indicate the variability of the orientation of the embryonic mass in these families. Thin arrow points to the embryonic mass, arrowheads to the endodermal layer and thick short arrows to the trophoblastic layer. The dotted region represents the uterine wall

is no description of the orientation of the embryonic disc at neural groove stage in any of the molossid bats mentioned above. Either the disc becomes mesometrially orientated as the embryo reaches the late neural groove stage, or the situation in *Tadarida brasiliensis cynocephala* is unique and needs explanation as to how the allantois grows towards the mesometrial placental pad from the antimesometrially oriented embryonic disc since the presence of the yolk sac in close approximation with the uterine wall has to be displaced to allow the allantois to grow towards, and come into contact with, the preplacental pad in this animal.

The first attachment of the blastocyst to the uterus is antimesometrial in all vespertilionidae[23, 65, 73, 98, 99, 111, 128] and *Desmodus rotundus murinus*[130], lateral in *Pteropus giganteus giganteus*[89] and *Noctilio labialis*

minor[2] and circumferential in *Cynopterus marginatus*[69], *Cynopterus sphinx gangeticus*[92], *Rhinopoma kinneari*[116], *Taphozous longimanus*[28], *Megaderma lyra lyra*[41], *Rhinolophus rouxi*[4] and *Tadarida brasiliensis cynocephala*[117]. In *Rousettus leschenaulti*[50], while the first attachment of the blastocyst is mesometrial, the blastocyst soon expands and establishes contact on its entire surface to become circumferentially attached.

With regard to the depth of implantation it is superficial in all the bats so far studied except in *Pteropus giganteus giganteus*[89], where it is partially interstitial, and in *Desmodus rotundus murinus*[130], and the phyllostomid bats[8,19,53,105,127], where it is completely interstitial. The situation in *Noctilio labialis minor*[2] appears to fall into the category of being "Partially interstitial" although the authors mentioned, "the depth of implantation in *Noctilio labialis minor* resembles both *Myotis* and *Desmodus*"[126,130]. Figure 5, Plate-2 indicates that nearly half of the blastocyst on the embryonic pole is surrounded by the expanding decidua basalis. This feature has also been mentioned by the authors while describing the particular stage of development. They stated, "The embryonic mass is completely surrounded by the decidua basalis and invading trophoblast". (Underline ours).

V. Foetal Membranes

A. *General course of development*

The general course of development of the foetal membranes follows a common pattern in all the bats except for specific variations to be mentioned later. Four distinct stages can be recognized in the morphogenesis of the foetal membranes of all bats—(1) the neural groove stage, (2) the early limb-bud stage, (3) the late limb-bud stage and (4) full term. In some species there are interesting changes in the foetal membranes occurring during the third quarter of pregnancy.

Text figures 3(a-d) and 4(a-e) illustrate the development of the foetal membranes in most vespertilionids[28,24,65,72,98,111,128,130] and in pteropids[59,89,90] respectively. These represent two extreme methods of morphogenesis of foetal membranes among bats, and the various modes of development noticed in the other species are modifications of these extreme types.

As already mentioned, the neural groove stage is a crucial stage of development in bats since the definitive disposition of the foetal membranes

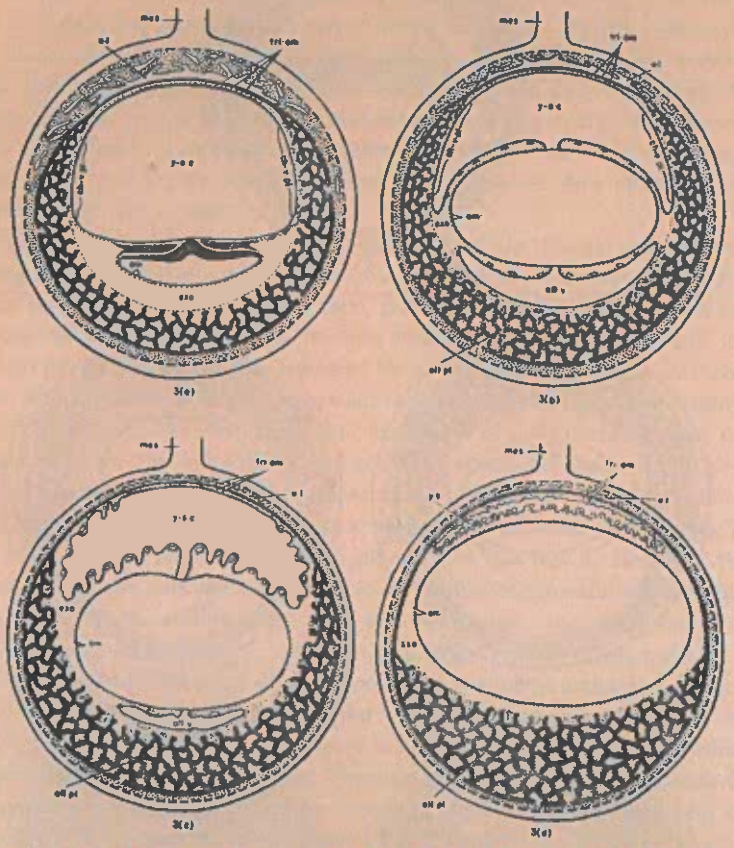


Fig. 3. Semischematic drawings to illustrate the arrangement of the foetal membranes in *Pipistrellus mimus mimus*. (a) Neural groove stage; (b) Early limb-bud stage; (c) Late limb-bud stage; (d) Full term

is established at this stage. It is at this stage that the final position of the chorio-allantoic placenta is established since the final placental disc in all bats, except *Tadarida brasiliensis cynocephala*[117], is formed on the side of the uterus towards which the embryonic plate faces and where a chorionic placenta would have been already established. The yolk-sac placenta, which is at first nonvascular and which later becomes vascularized (chorio-vitelline placenta), develops in relation to the yolk-sac wall which is in contact with

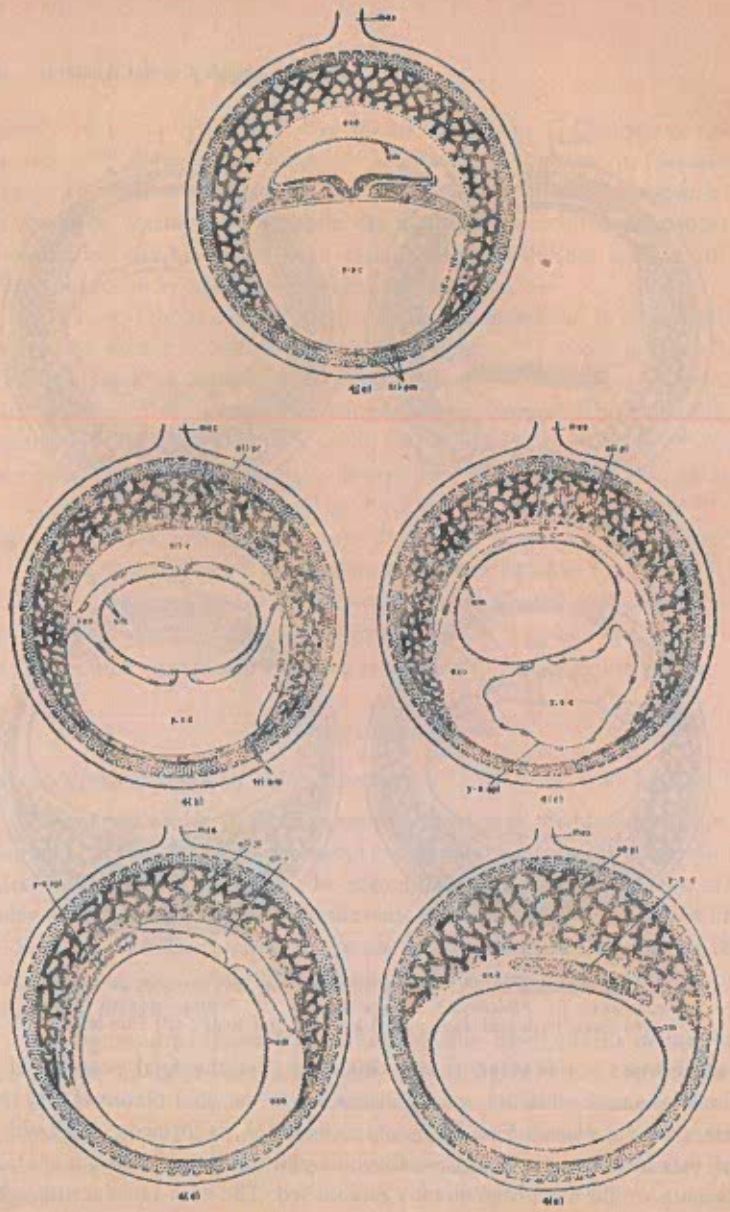


Fig. 4. Semischematic drawings to illustrate the arrangement of the foetal membranes in *Rousettus leschenaulti*. (a) Neural groove stage; (b) Early limb-bud stage; (c) Late limb-bud stage; (d) Mid-pregnancy; (e) Full term

the uterus. Whereas in the vespertilionids the abembryonic part of the yolk-sac wall does not come into contact with the uterus but remains free in the persistent uterine lumen, in other bats the blastocyst wall comes into contact with the uterus on its entire surface resulting in the obliteration of the uterine lumen at the level of attachment of the embryo. Hence, while the yolk-sac placenta is confined to the lateral sides of the uterus in vespertilionids, it is formed in relation to the entire yolk-sac wall including the abembryonic side in other bats.

The early limb-bud stage of development of the foetus presents two changes in the foetal membranes as compared to the previous stage. The allantois, carrying foetal blood vessels, grows towards the dorsal side of the foetus and gets attached to the placenta which was originally opposite the embryonic plate—that is, the region of the chorionic placenta. The allantois spreads fan-wise (in sectional views) on the foetal surface of the placenta. Secondly, due to the formation and expansion of the exocoelom and the enlargement of the amniotic cavity enclosing the foetus the proximal part of the vascular yolk-sac splanchnopleure gets separated from its contact with the uterine wall in the embryonic half of the chorionic sac resulting in the abolition of the chorio-vitelline placenta in this region. However, the distal half of the yolk-sac splanchnopleure retains its placental relation with the uterus. Thus, at this stage both chorio-allantoic and chorio-vitelline placenta occur concurrently, although in different regions of the uterus.

The late limb-bud stage of development of the foetus is characterized by the complete separation of the vascular yolk-sac splanchnopleure from its contact with the uterus so that there is no remnant of the chorio-vitelline placenta from this stage onwards. The placenta is cup-shaped and is entirely supplied by allantoic vessels. After this stage the course of development of the foetal membranes differ in different bats. The chorio-allantoic placenta progressively becomes consolidated in its definitive position. Its final location, shape and structure vary considerably in the different families of bats.

B. *Amnion*

Three main methods of amnion formation have been noticed in Chiroptera : (1) The primitive amniotic cavity arises within the embryonic mass either by the destruction of some of the cells in the center of the embryonic mass or by a radial rearrangement of the cells of the embryonic mass leaving a central cavity. With the expansion of the embryonic mass

into a disc, the roof of the primitive amniotic cavity becomes stretched and gets progressively thin until it ruptures and is lost. Hence, the amniotic cavity is roofed over by the trophoblast for sometime. Later, folds develop from the margins of the embryonic plate and fuse dorsally to form a new roof for the amniotic cavity. This is the ectodermal component of the true amnion which gets an investment of extra-embryonic mesoderm to become the definitive amnion (Fig. 5 a–d). Such a mode of amnion formation has been described in *Vespertilio murinus*[12-14], *Vesperugo noctula*[126], *Molossus rufus* and *Molossus obscurus*[110], *Vesperugo leisleri*[103], *Myotis lucifugus lucifugus*[128], *Megaderma lyra lyra*[41], *Pipistrellus ceylonicus chrysothrix*[98], *Pipistrellus dormeri*[111], *Hipposideros speoris*[57] and *Pipistrellus mimus mimus*[05]; (2) The primitive amniotic cavity is formed as in (1) above, but the roof of the primitive amniotic cavity does not disappear, but persists as the ectodermal component of the definitive amnion. The extension of extraembryonic mesoderm over the roof of the primitive amniotic cavity establishes the definitive amnion (Fig. 6 a–d). Such a method of amnion formation has been noticed in *Glossophaga soricina soricina*[58], *Scotophilus temmincki* (*Scotophilus wroughtoni*)[28], *Pteropus giganteus giganteus*[80], *Desmodus rotundus murinus*[130], *Rousettus leschenaulti*[59], *Scotophilus heathi*[72] and probably *Noctilio labialis minor* (*Dirias albiventor minor*)[2]; (3) Amniogenesis by only fold formation has been reported in *Tadarida brasiliensis cynocephala* by Mossman[93] and Stephens[117] (Fig. 7 a–d).

The mode of development of the amnion in *Miniopterus schreibersii fuliginosus*[10] follows method (2) above up to a certain stage but its further development is unlike that in all other bats, and hence needs special mention. The primitive amniotic cavity in this bat is formed within the embryonic mass as in all other bats. The rapid expansion of the embryonic disc tears away the roof of the primitive amniotic cavity. Since the roof of the primitive amniotic cavity is also the trophoblast layer of the blastocyst overlying the embryonic mass, the loss of the roof of the primitive amniotic cavity results in exposing the embryonic ectoderm to the superficial layer of uterine endometrium, the uterine epithelium having become lost. The amniotic cavity, therefore, becomes confluent with the potential uterine lumen, and is roofed over by the superficial layer of endometrium until the definitive amniotic folds develop from the margins of the embryonic disc, and fuse on the dorsal side to enclose the open space on the dorsal side of the disc. A parallel situation has not been noticed in any other bat except

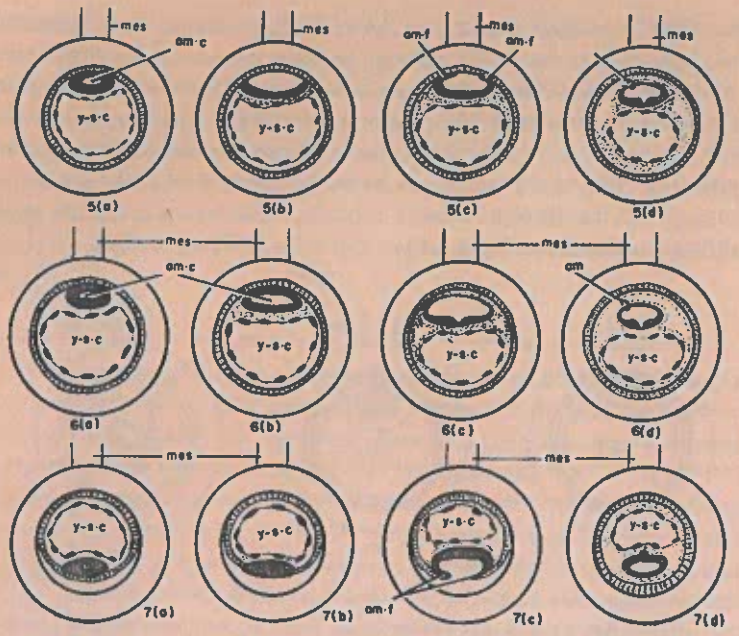


Fig. 5 (a-d). Schematic representation of the development of the amnion in *Megaderma lyra lyra*.

Fig. 6 (a-d). Schematic representation to illustrate the development of the amnion in *Rousettus leschenaulti*.

Fig. 7 (a-d). Schematic representation to illustrate the development of the amnion in *Tadarida brasiliensis cynocephala*.

probably *Eomops*[54], where it is reported that the amniotic cavity is roofed over by the uterine wall for a short period. While describing amniogenesis in *Miniopterus schreibersii* of Europe, Da Costa[11] interpreted that the roof of the amniotic cavity, after the loss of its original roof, was formed by the 'plasmodiblaste'. His Figures 7, 8 and 9 show that the trophoblastic layer, presumed to be overlying the embryonic mass, is also in a state of rupture. His Figures 14 and 16 leave little doubt that Da Costa's interpretation of the tissue roofing the primitive amniotic cavity after the rupture of the original roof was erroneous, and the roof is actually composed of the uterine endometrial tissue and disintegrated cells of the uterine glands.

A fourth, rather unusual, method of amnion formation has been reported to occur in *Cynopterus sphinx gangeticus*[92] and possibly in *Cynopterus*

marginatus[69]. In these animals the "primary" amniotic cavity is formed by cavitation in the center of the embryonic mass. But a second cavity referred to as the "secondary" amniotic cavity is formed on the "dorsal" side of the primary amniotic cavity by the development of folds. Soon the two cavities become confluent due to the breakdown of the roof of the primary amniotic cavity. The roof of the secondary amniotic cavity forms the ectodermal component of the definitive amnion after it is encompassed by the extra-embryonic mesoderm (Fig. 8 a–f).

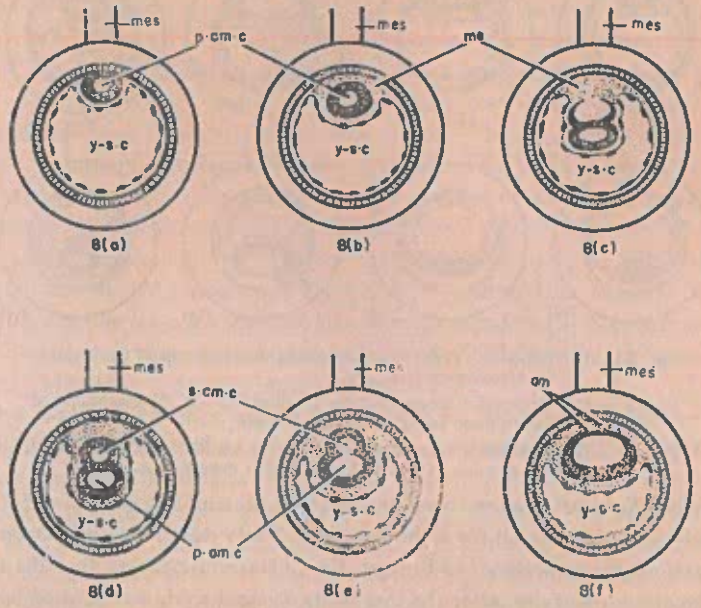


Fig. 8 (a-f). Schematic representation to illustrate the development of the amnion in *Cynopterus sphinx gangeticus*

The mode of formation of the amnion in mammals appears to be dependant on the location of the embryonic disc in relation to the region of the blastocyst which becomes attached to the uterus. Normally, where the attachment is abembryonic, amniogenesis takes place by fold formation, while, if the attachment is embryonic, it occurs by cavitation. Perhaps, this is because, while in the former case the embryonic disc faces the uterine

lumen, and hence there is sufficient space for fold formation, in the latter case there is no space for the development of folds. Consequently, the embryo, while abutting against the uterine wall, has to adapt to the new situation by producing a primitive amniotic cavity, either by the destruction of some of the central cells of the embryonic mass or by a radial rearrangement of the cells of the embryonic mass leaving a cavity in the center. Hence, the mode of amnion formation may be considered as a developmental adaptation to suit the different types of uterus-blastocyst relationships.

C. *Yolk sac*

The development and the final structure of the yolk sac in bats display three main trends : (1) The abembryonic segment of the yolk sac remains permanently bilaminar or trilaminar since the exocoelom does not invade this region of the yolk-sac wall as in all vespertilionids[24,36,49,65,103,129], *Desmodus rotundus murinus*[130], *Noctilio labialis minor*[2], *Artibeus jamaicensis parvipes*[137], *Megaderma lyra lyra*[25,42] and *Thyroptera tricolor*[134]. Whereas the abembryonic yolk-sac wall remains in contact with the uterine wall in *Noctilio labialis minor*, *Artibeus jamaicensis parvipes*, *Desmodus rotundus murinus*, *Megaderma lyra lyra* and *Miniopterus schreibersii fuliginosus*[10,38], it remains hanging in the persistent uterine lumen in the vespertilionid bats. In all these cases the free segment of the vascular yolk-sac splanchnopleure becomes progressively folded and is pushed or "invaginated"[129] towards the persistent bi- or tri-laminar segment of the yolk sac due to the expansion of the amniotic cavity ; (2) The extension of the exocoelom on all the sides of the yolk sac results in the complete separation of the vascular splanchnopleure from the chorion, and the separated yolk-sac splanchnopleure, therefore, lies freely in the exocoelom, undergoes progressive collapse, and is withdrawn from its original position until it is finally pulled towards the chorio-allantoic placenta. Thus, in the final stages a major part of the collapsed yolk-sac splanchnopleure lies adjacent to the placental disc. The progressive hypertrophy of the endodermal, and, in some cases, the mesodermal cells of the collapsed yolk-sac splanchnopleure gives to it a gland-like appearance (Pl. I—Figs. 1 and 2). The original lumen of the yolk sac, however, persists in the form of streak-like spaces thereby accentuating the gland-like appearance since the remnants of the lumen appear like ducts of the gland. Such a method of development of the yolk sac occurs in *Rhinopoma*

kinneari[28,116], *Taphozous longimanus*[28], *Taphozous melanopogon*[28,75], hipposiderid bats[28,37,45,56], and *Rhinolophus rouxi*[5,32] ; (3) The yolk-sac splanchnopleure develops as in (2) above until about mid-gestation after which the splanchnopleure undergoes complete collapse accompanied by the hypertrophy of the endodermal and mesodermal elements. Whereas the close approximation of the endodermal layer has been reported to obliterate the yolk-sac lumen, and the endodermal cells form a nearly continuous layer except for the regions where it is highly folded and, hence, appears to be separated into different blocks in *Tadarida brasiliensis cynocephala*[117,120,121,122], the lumen of the yolk sac persists as small discontinuous spaces surrounded by large endodermal cells in *Tadarida aegyptiaca*[26] (Pl. I—Figs. 3 and 4). In pteropid bats in addition to the enormous hypertrophy of the cells the endodermal lining itself becomes pinched off into numerous acinar groups without any remnant of the lumen in them. Each "acinus" is enveloped by a distinct lining of mesothelial cells which also form a complete outer envelope to the yolk sac. Thus, at full term the yolk sac assumes a solid, vascular gland-like appearance and has the histological appearance of an endocrine gland (Pl. II—Figs. 5–8). Such a type of development of the yolk sac has been consistently noticed in all megachiropteran bats[24,35,61,68,69,90,92,100,112,131].

One interesting feature of the histogenesis of the yolk sac, which seems to be common to all the bats irrespective of the final structure of the yolk sac, appears to have escaped the notice of all earlier workers. The endodermal cells of the yolk-sac wall are squamous and flat until the yolk-sac wall establishes contact with the uterus, after which the endodermal cells becomes slightly enlarged in the region where the yolk-sac wall comes into contact with the uterine wall. A further hypertrophy of the endodermal cells occurs concomitant with the extension of the vitelline vessels. Thus, in bats, in which the vitelline vessels extend to all the areas of the yolk sac, the endodermal cells also hypertrophy on all the sides, whereas in those bats, in which the area vasculosa does not extend to the abembryonic region, the hypertrophy of the endodermal cells takes place only in those regions receiving vascular supply. The separation of the yolk-sac splanchnopleure from the somatopleure triggers the hypertrophy of the mesodermal cells also. These cytological changes in the cellular elements of the yolk sac should reflect changes in the physiology of the yolk sac. But, the present stage of our knowledge does not encourage any speculations regarding the precise

functions of the yolk sac at different stages of pregnancy in different bats. The unique gland-like conversion of the entire yolk sac in molossid bats and all the Megachiroptera has defied a physiological explanation so far. Ultrastructural studies by Stephens and Easterbrook[120–122] on the yolk sac of *Tadarida brasiliensis cynocephala* have shown the presence of peculiar "membranous organelle", which develop directly from the rough endoplasmic reticulum within the endodermal cells of the yolk sac, and the authors indicated that these may be involved in glycogen and lipid metabolism. They have shown that glycogen and lipid accumulates in the endodermal cells of the yolk sac of this animal until late stage of pregnancy, but get depleted just prior to parturition. Histochemical examination of the yolk-sac gland of *Rousettus leschenaulti*[50] has revealed that, whereas the mesodermal cells are intensely PAS-positive even after salivary digestion, the endodermal cells are faintly stained with PAS after salivary digestion. Further, whereas the endodermal cells take a deep stain with Best's carmine procedure, the mesodermal cells are faintly stained. These staining reactions suggest that the endodermal cells are in some way involved in glycogen metabolism. Stephens and Cabral[119] speculated that the mesothelial cells of the yolk sac of *Tadarida brasiliensis cynocephala* are involved in the absorption of proteinaceous material from the exocoelom for the use of the embryo. Enders *et al.*[18], who studied the ultrastructure of the yolk sac of *Myotis lucifugus lucifugus*, indicated that the hypertrophied mesothelial cells performed the function of absorption of proteins from the exocoelomic fluid and that the hypertrophied endodermal cells may be concerned with secretion of substances to be taken into the vitelline blood vessels. Karim *et al.*[68], studying the ultrastructure of the yolk-sac gland of *Rousettus leschenaulti*, noted the presence of an absorption apparatus in the form of elongated microvilli on the exocoelomic surface of the mesothelial cells. The hypertrophied endodermal cells in this animal may be primarily concerned with glycogen metabolism. It is too premature to postulate and specify physiological role(s) to the unique yolk-sac gland of the Megachiropteran bats, and, in the words of Wimsatt[130], a fascinating histophysiological problem awaits solution.

D. *Allantois*

In some bats the allantois has a prominent vesicle, which expands concomitantly with the spreading of the allantoic mesenchyme on the foetal surface of the placenta, and reaches its maximum size by about the limb-bud

stage of development. After this stage, as the placenta becomes progressively consolidated to a discoidal structure, the allantoic vesicle also becomes reduced in size until it is almost completely obliterated or reduced to a small space in the allantoic mesenchyme on the foetal surface of the placenta. The endodermal allantoic duct either persists within the umbilical cord until term or disappears after mid-gestation. Such a history of the allantois, with minor modifications, has been noticed in *Pteropus giganteus giganteus*[69], *Rousettus leschenaulti*[59], *Rhinopoma kinneari*[26,116], *Noctilio labialis minor*[2], *Tadarida brasiliensis cynocephala*[117], *Megaderma lyra lyra*[42], *Rhinolophus rouxi*[5], *Hipposideros bicolor pallidus*[28,45], *Hipposideros speoris*[56], and most vespertilionids[24,66,72,98,111,129]. The allantois is small, nonvesicular and transient in *Desmodus rotundus murinus*[130]. In the phyllostomid bats, *Glossophaga soricina soricina*[53], whereas there is no distinct allantoic vesicle, there is an endodermal allantoic duct, which runs through the length of the umbilical cord and branches several times beneath the placenta. This duct may be occluded to form a solid cord of cells at some places. In another phyllostomid bat, *Artibeus jamaicensis parvipes*[157], the endodermal allantois is a simple tubular duct ending blindly in the mesenchyme below the placenta without any distinct terminal expansion or branching. *Taphozous longimanus*[28] does not have a distinct allantoic vesicle at any stage of development. However, serial sections of the umbilical cord of this species at full term show the presence of distinct allantoic duct[66] throughout the length of the umbilical cord without terminal dilatation into a vesicle.

E. *Placenta*

Although the general plan of development of the placenta follows nearly the same pattern in all bats, there are some differences amongst them with respect to details, particularly with regard to those concerning the position and shape of the placenta and the fate of the maternal endothelium.

1. *Location and Shape*

From a morphological point of view, in most bats the placenta occupies an extensive area of the uterine wall during early stages, but it becomes progressively reduced in extent until it assumes a discoid structure whose location varies among bats (Text figures 3 and 4). The final placental disc is

mesometrial in all Megachiroptera[67,89,90,98], Rhinopomatidae[26,116], Megadermatidae[42], Rhinolophidae[5,52], Hipposideridae[26,37,45,56,61] and Molossidae[117] among the Microchiroptera. The disc is antimesometrial in Phyllostomatidae (fundic)[52,53,104,105,187], Desmodontidae[130] and Vespertilionidae[23,24,66,70,72,98,103,111,129]. In Emballonuridae[26,135] the placenta lies on the mesometrial and the lateral sides, but the mesometrial moiety becomes converted into an haematoma after mid-gestation. In Noctilionidae[2] the placental disc is located between the lateral and antimesometrial sides.

The conversion of the cup-shaped placenta into a discoid structure is brought about either by a consolidation of the placenta towards its definitive position as in most bats, or by the progressive destruction and disappearance of the placental tubules near the margins of the placental cup, and a progressive consolidation of the placenta towards the mesometrial side as in *Hipposideros fulvus fulvus*[61,66]. In all hipposiderids[28,37,45,56] a cleft is developed in the center of the allantoic placental disc, and this becomes progressively deep and wide so that during advanced stages of gestation the placenta appears to be composed of two discs.

The morphological changes in the placenta of *Miniopterus schreibersii fuliginosus*[33] needs mention because it is unmatched in any mammal. The placenta in this bat develops in three stages. The primary chorio-allantoic placenta occurs in the form of a cup-shaped structure occupying the antimesometrial half of the uterus until about the limb-bud stage of development of the foetus, when the secondary chorio-allantoic placenta develops in the form of two caruncle-like structures at the margin of the primary placental disc, one at the cranial and the other at the caudal end of the placental cup. The secondary placental caruncles enlarge progressively and encroach upon the region of the primary placenta until they form two large discs on the antimesometrial side separated by a narrow gap, which is all that remains of the primary placenta. After mid-pregnancy, a tertiary placental disc develops within each of the secondary placental discs. Histologically, the primary placenta consists of a large number of placental tubules, each composed of a loop of maternal blood capillary surrounded by trophoblast. Allantoic mesenchyme and foetal blood capillaries occur between the placental tubules. The primary placental tubules undergo progressive disintegration after mid-pregnancy and occur in the form of a few degenerated empty tubules at full term. The histogenesis of the secondary

placental disc is similar to the chorio-allantoic placenta of other bats and results in the development of a labyrinthine endotheliochorial placenta. Each tertiary placental disc is composed of numerous inter-connected lacunae, the walls of which are made up of a layer of trophoblast. There is no remnant of the maternal endothelium in the tertiary placenta. The tertiary placenta is, therefore, haemochorial.

2. *Development*

In all the bats there are three major kinds of placenta developed during gestation. These are the "Trophoblastic placenta" or "Pre-placenta", the "Yolk-sac placenta" and the "Allantoic placenta". Although these are developed in a chronological sequence there is a certain degree of overlapping of these so that two or more types may be present at the same time but at different regions of the gestation sac during early stages of gestation.

(a) *The trophoblastic placenta or pre-placenta*

The trophoblast is the first embryonic membrane which comes into contact with the uterus. With the advancement of pregnancy there is not only an increase in the area of apposition between the trophoblast and the uterine wall but also a progressively more intimate relationship between the two. After the uterine epithelium is lost and is replaced by the trophoblastic layer of the blastocyst the trophoblast undergoes proliferation and invades the endometrium destroying all but the maternal blood vessels in the endometrium. Thus, at advanced stages of implantation a thick area of the endometrium is invaded by the trophoblast in which maternal capillaries occur. At this stage the trophoblast can be recognized into a basal layer of cellular trophoblast—the cytotrophoblast—and a thick shell of invading syncytiotrophoblast. Until now the foetal component of the placenta is only the trophoblast and hence can be termed the "trophoblastic placenta".

(b) *The yolk-sac placenta*

The part of the original placenta, with which the yolk-sac wall comes into contact, is referred to as the yolk-sac placenta. While the yolk sac lies on the ventral side of the embryonic area, the dorsal side of the embryonic disc is occupied by the amnion, which is in contact with the trophoblastic

placenta, and this may be designated as the "chorio-amniotic placenta", which is a very transitory structure and is abolished by the extension of the exocoelom. The yolk-sac placenta is present on the lateral and the abembryonic sides in *Rousettus leschenaulti*[59] and *Hipposideros fulvus fulvus*[87] and only on the lateral sides in all vespertilionid bats[28, 66, 72, 96, 111, 129]. It is first non-vascular, but soon gets vascularized progressively from the embryonic towards the distal segment thereby establishing the chorio-vitelline placenta. The formation and extension of the exocoelom separates the yolk-sac splanchnopleure from the chorion, and thus abolishes the chorio-vitelline placenta. Such a history of the yolk-sac placenta is common to all the bats with the exception of *Desmodus rotundus murinus*[130] and possibly in *Glossophaga soricina soricina*[58] where a chorio-vitelline placenta is not formed. Wimsatt[130] stated, "A chorio-vitelline placenta never is formed in *Desmodus rotundus*, nor in the phyllostomid bats, for the extra-embryonic mesoderm never extends beyond the margin of the allantoic placental disc, and there is, therefore, no vascular trilaminar omphalopleure, which is, by definition, the fetal component of the chorio-vitelline placenta".

(c) *The allantoic placenta*

In all the bats, except *Tadarida brasiliensis cynocephala*[117], the allantois grows dorsalwards and gets attached to that region of the chorionic placenta, which was for sometime the chorio-amniotic placenta. Hence, the allantoic placenta is formed on the side opposite to the one towards which the yolk sac lies in all the species except *Tadarida brasiliensis cynocephala* which is unique in having the allantoic placenta develop in a region opposite to the location of the embryonic disc. This is of considerable significance because in all the other bats the ultimate location of the placenta depends on the early disposition of the embryo, and the allantois grows dorsalwards from the embryonic disc and attaches to the uterine wall at a region diametrically opposite to the location of the yolk sac which lies on the ventral side of the embryo. However, in *Tadarida brasiliensis cynocephala*[117] the location of the allantoic placenta is on the mesometrial side, where the yolk sac is also present, whereas the embryonic disc is oriented antimesometrially. Apparently, in this animal the allantois grows ventralwards from the foetus and through the exocoelom between the yolk-sac splanchnopleure and the somatopleure—a situation unparalleled in any other bat. Perhaps this aspect of the embryology of this bat needs further investigation.

(d) *Histogenesis of the placenta*

The early histogenesis of the placenta follows nearly the pattern in all the bats whose embryology is known. This has been described in detail by Wimsatt[120-121], Gopalakrishna[39], Gopalakrishna and Moghe[45] and Karim[62]. The trophoblastic layer of the implanting blastocyst destroys the uterine epithelium, undergoes proliferation and progressively invades the endometrial tissue. The trophoblast can now be recognized into a basal layer of cytotrophoblast and a thick mantle of syncytiotrophoblast which destroys all but the maternal vascular channels as it progressively invades the endometrium. As the syncytiotrophoblast penetrates deeper into the endometrium the basal cytotrophoblast sends into the syncytiotrophoblast solid cords of cells (Pl. III—Fig. 10), which become hollowed out from the foetal surface and undergo expansion resulting in the formation of numerous hollow cytotrophoblastic villi (Pl. III—Fig. 11). Consequently, the syncytiotrophoblast becomes progressively reduced into a thin lamina surrounding maternal blood vessels so that, in stained sections, the placenta appears to be formed of tubules—the "placental tubules"[45]—hanging from the uterine wall (Pl. III—Fig. 12). Allantoic mesenchyme and foetal capillaries enter the cytotrophoblastic villi from the foetal side and establish the foetal vascularization of the placenta. By the time the embryo reaches the limb-bud stage of development the chorionic villi undergo branching and anastomosis resulting in the formation of the labyrinthine placenta. Thus, at the limb-bud stage the placenta appears to be composed of a three-dimensional network of tubules hanging from the uterine wall and embedded in allantoic mesenchyme in which foetal blood vessels are present (Pl. IV—Fig. 13). Each placental tubule at this stage in most bats is composed of a central maternal blood capillary with a distinct endothelial lining, an inner layer of syncytiotrophoblast and an outer layer of cytotrophoblast. A prominent eosinophilic cytoplasmic layer incorporating a PAS-positive membrane lies between the endothelial cells and the row of syncytiotrophoblastic nuclei.

The further histogenesis of the placenta follows one of two trends depending upon the fate of the maternal endothelium in the placental tubules. The endothelial lining becomes progressively attenuated until it disappears altogether leading to the establishment of an haemochorial condition in many species such as *Rousettus leschenaulti*[59], phyllostomids[58,127], desmodontids[6,130], some vespertilionids[17,129] and some molossids[38,117,131]. On the other hand in *Pteropus giganteus, Cynopterus sphinx gangeticus*[38,131]

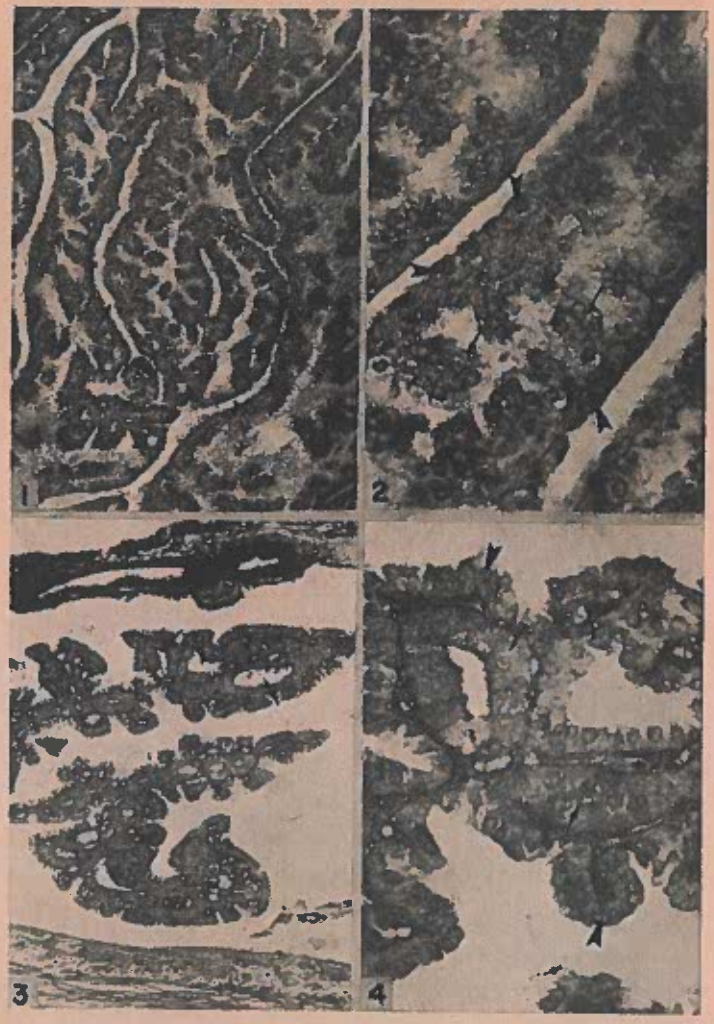


Plate I

Fig. 1. Part of the collapsed yolk-sac splanchnopleure of *Taphozous longimanus*. Arrow points to the remnants of the yolk-sac lumen. × 140

Fig. 2. Part of the collapsed yolk-sac splanchnopleure of *Taphozous longimanus*. Note the hypertrophied endodermal (arrowhead) and mesodermal (arrow) cells. × 560

Fig. 3. Part of the collapsed yolk-sac splanchnopleure of *Tadarida aegyptiaca*. Arrow points to the remnant of the yolk-sac lumen. × 112

Fig. 4. Part of Fig. 3 enlarged to show the hypertrophied endodermal (arrow) and mesodermal (arrowhead) cells. × 560

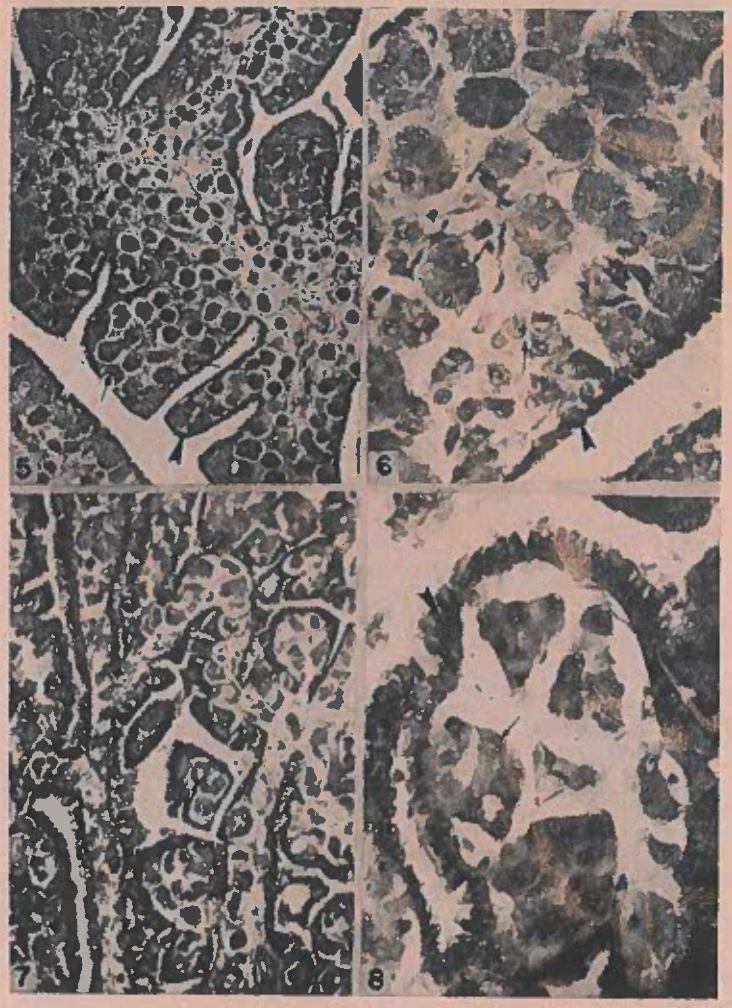


Plate II

Figs. 5 and 6. Yolk-sac gland of *Pteropus giganteus giganteus*. Arrow points to the acinus and arrowhead to the mesodermal layer. 5×88 ; 6×560

Figs. 7 and 8. Yolk-sac gland of *Rousettus leschenaulti*. Arrow points to the acinus and arrowhead to the mesodermal layer. 7×112 ; 8×560

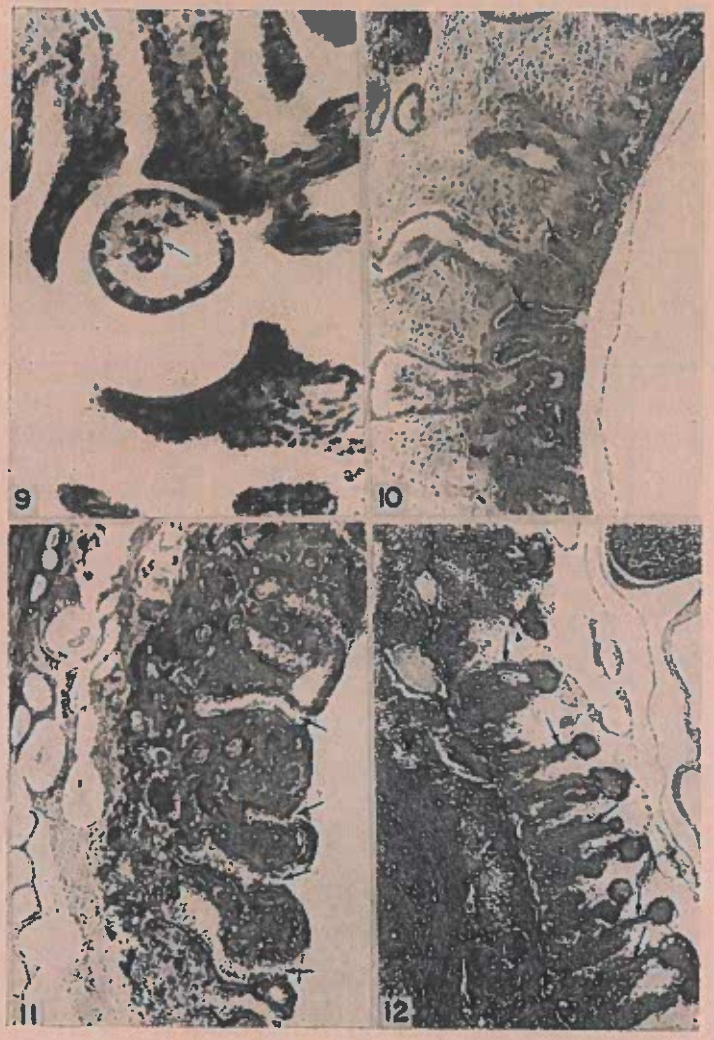


Plate III

Fig. 9. An unilaminar blastocyst of *Rousettus leschenaulti*. Arrow points to the embryonic mass. × 270

Fig. 10. Part of the uterine wall to show the invasion of solid cords of cytotrophoblast into the syncytiotrophoblast layer. Arrow points to the solid cord of cytotrophoblast. × 165

Fig. 11. Part of the uterine wall of *Rousettus leschenaulti* at late neural groove stage. Note the hollowing out of the cytotrophoblastic cords to form hollow villi (arrow). × 106

Fig. 12. Part of the placenta of *Pipistrellus mimus mimus* at late limb-bud stage of development. Note the numerous placental tubules (arrows). × 140

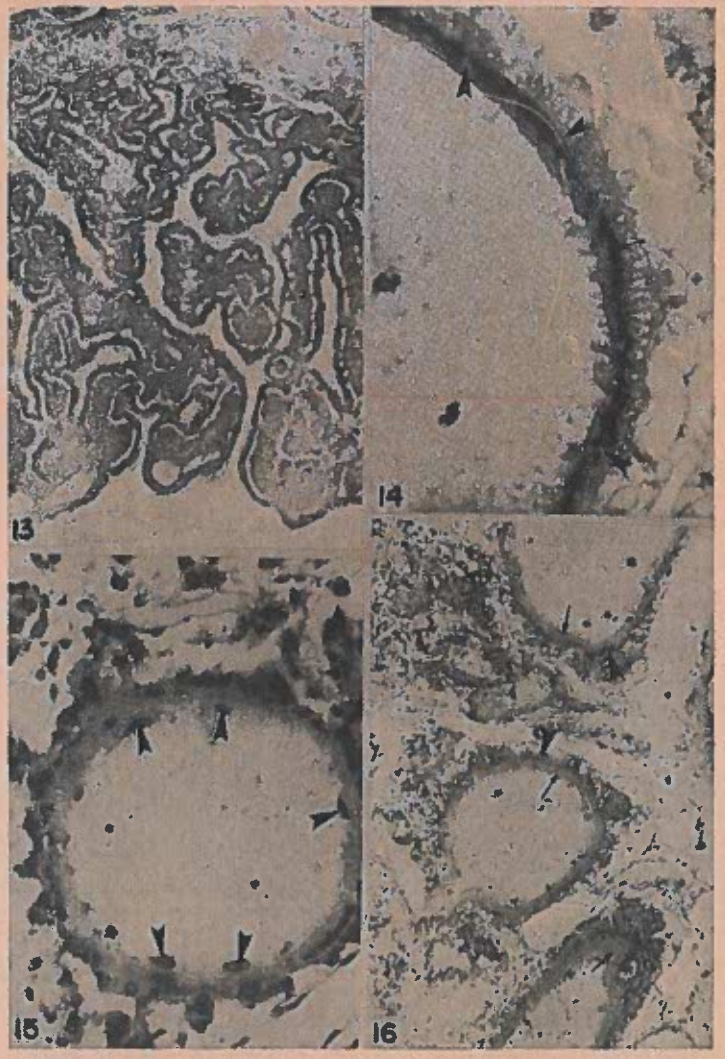


Plate IV

Fig. 13. Part of the placenta of *Rousettus leschenaulti* at mid-gestation. Note the network of placental tubules. × 80

Fig. 14. Part of the placental tubule of *Rousettus leschenaulti* at full term to show the presence of a thick PAS-positive membrane (arrow) surrounding the tubule. Arrowheads point to places where the membrane is discontinuous. × 560

Fig. 15. Section of the placental tubule of *Hipposideros fulvus fulvus* at full term Note the presence of maternal endothelial cells (arrowheads). × 650

Fig. 16. Section of a few placental tubules of *Rousettus leschenaulti* at full term. Note the presence of an eosinophilic lamina (arrow) on the inner surface of the tubules. The arrowheads point to the PAS-positive interstitial membrane. × 110

Rhinopoma kinneari[28,116,131], emballonurid bats[28,71,131], *Megaderma lyra lyra*[28,42], and hipposiderids[26,37,45,56] the endothelial cells of the maternal capillaries remain as a component of the placental barrier throughout gestation so that the placenta is endotheliochorial in these species (Pl. IV—Fig. 15).

The establishment of the definitive placental barrier has been studied in considerable detail in *Rousettus leschenaulti*, *Hipposideros fulvus fulvus*, *Megaderma lyra lyra*, *Pipistrellus ceylonicus chrysothrix* and *Pipistrellus mimus mimus* by the present authors and their coworkers, and these illustrate nearly all the kinds of interhaemal membranes met with among bats. In *Rousettus leschenaulti* and *Pipistrellus mimus mimus*, with the progressive attenuation of the maternal endothelium, there is an increase in the thickness of an eosinophilic cytoplasmic lamina between the layer of nuclei of the syncytiotrophoblast and the endothelium. The topographical position of this eosinophilic cytoplasmic lamina, taken along with the fact that there is a progressive decrease in the size of the endothelial cells, would lead one to the tempting conclusion that this membrane may have been derived from the maternal endothelium. Perhaps, a similar situation might have led some of the earlier authors[9,20] working on bat placenta to the conclusion that the maternal endothelium persists as an enucleate membrane in the placenta of the species examined by them. This conclusion is, however, not tenable in the case of *Rousettus leschenaulti* and *Pipistrellus mimus mimus* because (1) the homogeneous eosinophilic membrane occurs as a thick layer of cytoplasm belonging to the syncytiotrophoblast even from the early stages of development, (2) the thickness of the membrane in most placental tubules is greater than the thickness of the endothelial lining even at the latter's highest hypertrophy, and (3) the endothelial lining can be recognized as a distinct layer retaining its individuality from the homogeneous eosinophilic membrane and lies on the maternal border of the eosinophilic membrane up to the limb-bud stage.

A significant and constant component of the placental tubules of all the bats is the PAS-positive, azan-positive and eosin-negative interstitial membrane (Pl. IV—Fig. 14). In *Rousettus leschenaulti* and *Pipistrellus mimus mimus* this occurs embedded within the homogeneous eosinophilic membrane, which is present on the maternal border of the placental tubule, so that a thin eosinophilic cytoplasmic lamina occurs between the PAS-positive membrane and the lumen of the placental tubule (Pl. IV—Fig. 16). In both

these species the interstitial membrane is discontinuous and network-like so that the cytoplasm of the syncytiotrophoblast is in continuity with the cytoplasmic lamina bordering the lumen of the placental tubules. In *Hipposideros fulvus fulvus* the PAS-positive membrane lies immediately beneath the persisting endothelial lining. Wimsatt[131], who recognized the presence of the interstitial membrane in the placental tubules of many species of bats, adduced compelling evidence to indicate that the material of this membrane may be wholly or in part derived from the syncytiotrophoblast. It is interesting to note that the thickness of the interstitial membrane has some direct relationship to the calibre of the maternal vascular channels, which undergo dilatation and become progressively wider as pregnancy advances, while the endothelial lining becomes progressively attenuated. These facts indicate the possibility that the interstitial membrane may not entirely be the remnant of the basement membrane of the maternal blood capillaries in the placenta. Further, since the only other structure, which is topographically adjacent to the interstitial membrane, is the syncytiotrophoblast, it is evident that the syncytiotrophoblast is in some way responsible for the increase in the thickness of the interstitial membrane. However, these facts do not preclude the possibility that the basement membrane of the original endothelial lining may also get incorporated in the definitive interstitial membrane of the ripe placenta in all these bats.

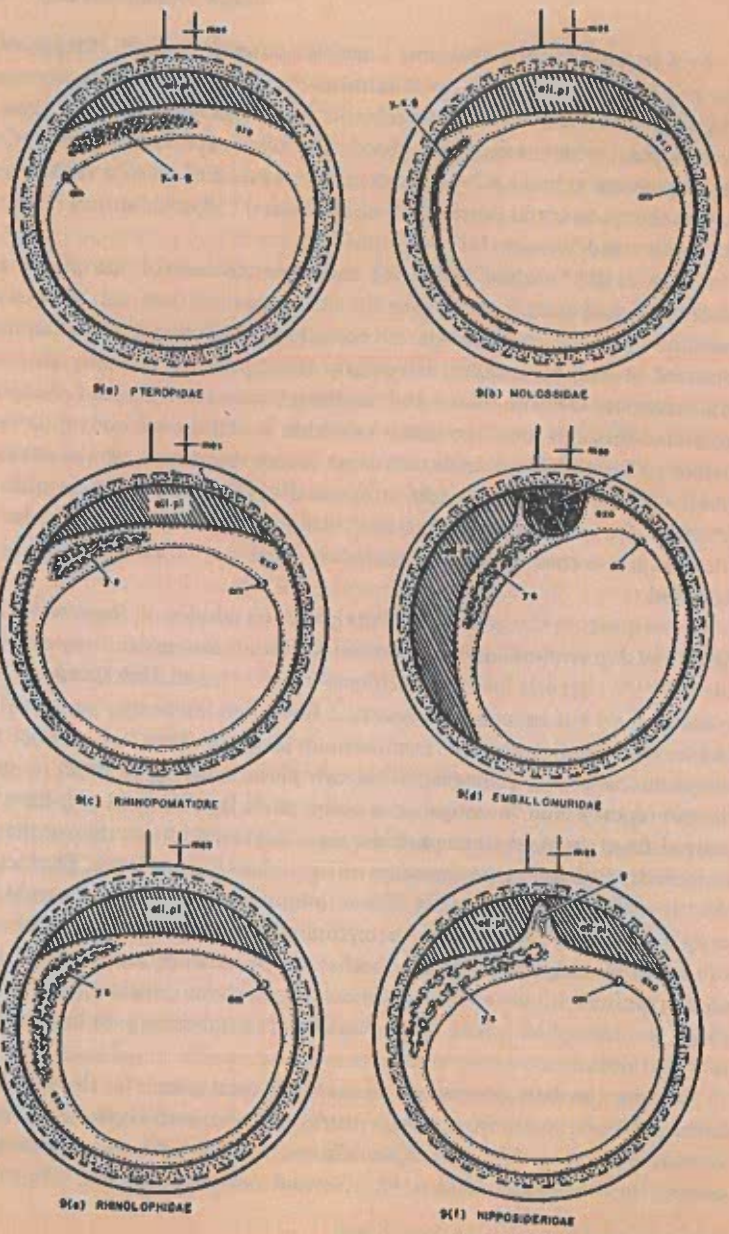
Bjorkman and Wimsatt[6], while studying the fine structure of the placenta of *Desmodus rotundus murinus*, have shown that the interstitial membrane in this animal "is never in direct contact with the lumen of the blood space, but is always covered over with cytoplasm, which is continuous with the syncytiotrophoblast". They also observed that the interstitial membrane may have a dual origin, part of it being the remnant of the basement membrane of the maternal endothelium, and part of it being contributed by the trophoblast. Working on the ultrastructure of the placenta of *Myotis lucifugus lucifugus*, Enders and Wimsatt[17] noted that cytoplasmic processes from the syncytiotrophoblast push through the basement membrane, and spread out underneath the maternal endothelium, thus loosening its attachment to the basement membrane. The endothelial cells thus become separated, and are lost. Hence, the "interstitial membrane", which was present earlier as the basement membrane of the endothelium, gets a covering of cytoplasm from the syncytiotrophoblast, so that it is converted into an "*Intrasyncytial lamina*" "perforated only at irregular intervals by syncytial tongues which spread

out over its upper surface forming a nearly continuous layer". Stephens[118] also noticed the presence of a discontinuous envelope of "homogeneous substance" over which the trophoblastic cytoplasm remains as a layer in direct contact with the maternal blood. The topographical situation of the "homogeneous substance" of Stephens[118] leaves little doubt that it is the same as the "interstitial membrane" of Wimsatt[131], Bjorkman and Wimsatt[6] and Enders and Wimsatt[17].

Karim *et al.*[67], while describing the ultrastructure of the placenta of *Rousettus leschenaulti*, noted that the Schiff-positive lamina, which was a constant feature of the placenta of this animal at all stages of development, consisted of a discontinuous, irregularly thickened, horizontally disposed, homogeneous, electron-dense and acellular material, syncytial cytoplasm protruded through the discontinuities within it and spread out on its upper surface to form a thin cytoplasmic layer facing the lumen of the maternal tubule. While it was extremely attenuated at some places, a continuous cytotrophoblast layer was observed beneath the syncytium in electron micrographs so that the interhaemal membrane may be classified as haemodichorial.

On comparing the structure of the placental tubules of *Rousettus leschenaulti* and *Pipistrellus mimus mimus* with that of *Desmodus rotundus murinus*[6,130,131], *Myotis lucifugus lucifugus*[17,129,131], and *Tadarida brasiliensis cynocephala*[118] it becomes evident that the placental barrier in these bats has a prominent interstitial membrane over which there is a layer of cytoplasm on its maternal border. This cytoplasmic layer, referred to in the present report as the "homogeneous eosinophilic layer", could only have been derived from the syncytiotrophoblast since it is several times thicker than the endothelial lining even at the stage of its highest hypertrophy. The fact that the "interstitial membrane" is discontinuous and network-like would lend support to the hypothesis that the cytoplasm from the syncytiotrophoblast can easily percolate through the meshes of the network towards the lumen of the placental tubule and form the continuous homogeneous layer of cytoplasm bordering the lumen of the tubule thus establishing an haemochorial relationship.

Among the bats possessing an haemochorial placenta, the placenta is haemodichorial in *Myotis lucifugus lucifugus*[17], *Desmodus rotundus murinus*[6], *Rousettus leschenaulti*[68] and *Pipistrellus mimus mimus*[35] and haemomonochorial in *Vesperugo leisleri*[103], *Scotophilus temmincki*[24], *Pipistrellus*



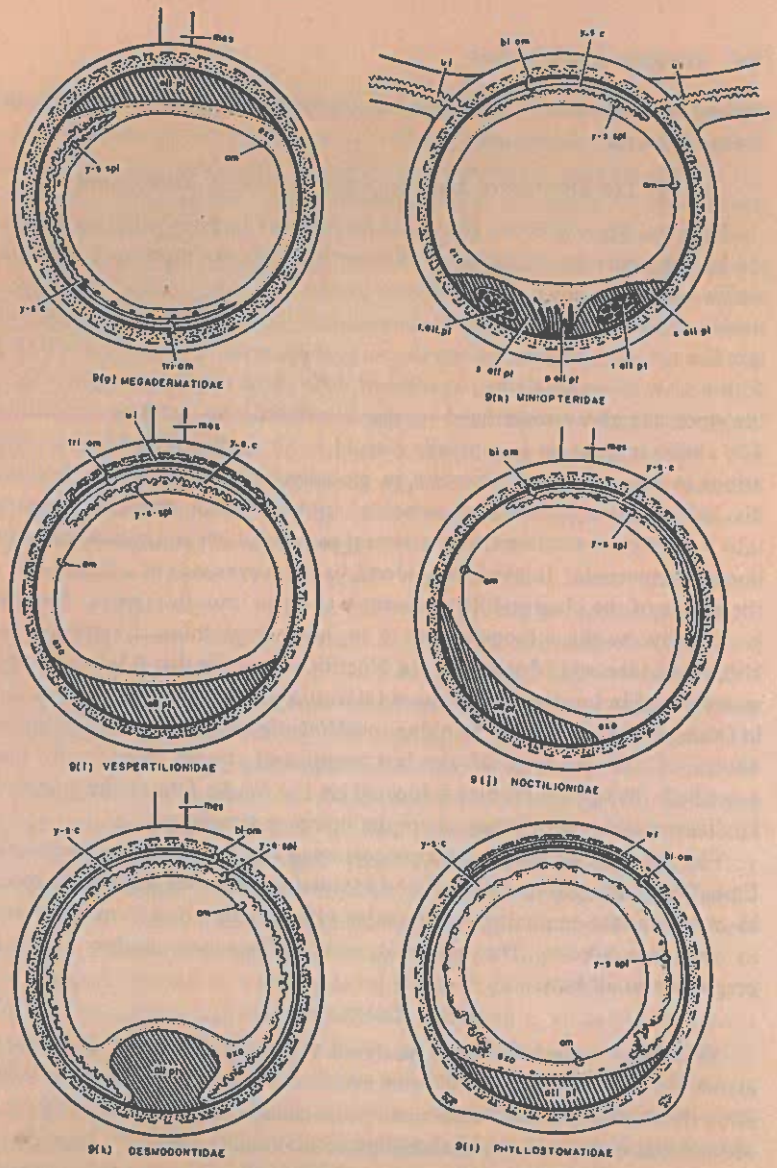


Fig. 9 (a-l). Semischematic representation of the arrangement of the foetal membranes at full term in different families of Chiroptera. The figures have been arranged as per the positions proposed now on the taxonomic sequence of the various families

ceylonicus chrysothrix[98], *Pipistrellus dormeri*[40], *Scotophilus heathi*[101] and *Tadarida brasiliensis cynocephala*[118].

VI. The Definitive Arrangement of Foetal Membranes

The final disposition of the foetal membranes in Pteropidae (as illustrated by *Rousettus leschenaulti*) and Vespertilionidae (as illustrated by *Pipistrellus mimus mimus*) has been already given (Text Fig. 4(e) and 3(d) respectively). However, semischematic diagrams of the foetal membranes of these families are also included in the accompanying series of text figures 9 (a-l) with a view to emphasizing the marked differences among the various families since the above-mentioned species illustrated two extreme conditions. The allantoic placenta is generally discoid in all the families with a few deviations in some families. The disc in emballonurid bats[28,71,185] becomes divided into two parts—a mesometrial moiety, which becomes converted into a sac-like haematoma, and a lateral moiety which is a typical endotheliochorial placenta. In the hipposiderid bats the presence of a deep cleft in the center of the placental disc separates the disc into two parts. The disc is squarely on the mesometrial side in Megadermatidae, Rhinolophidae, Hipposideridae and Molossidae. In Noctilionidae[2] the disc is lateral to antimesometrial in location. The placental disc is on the antimesometrial side in Desmodontidae, Vespertilionidae and *Miniopterus schreibersii*. The unique nature of the placenta of the last mentioned species has already been described. The placental disc is located on the fundic side of the uterus in Phyllostomatidae (in species where the uterus is simplex).

The yolk sac lies freely in the exocoelom in Pteropidae, Rhinopomatidae, Emballonuridae, Rhinolophidae and Molossidae, whereas, the amembryonic bi- or tri-laminar omphalopleure remains without separation from the chorion in all other species. The allantoic vesicle disappears during advanced pregnancy in all bats.

VII. Conclusions

Mossman[93] assembled and analysed the available data on the foetal membranes of mammals and gave an excellent working hypothesis for evaluating developmental and foetal membrane characters for determining inter-ordinal and intra-ordinal relationships of mammals. He[93,95] argued that in the mammals, unlike skin, teeth, limbs and alimentary canal, which are directly under the influence of the environment, and are, therefore, prone to undergo adaptive modifications, the characters of the genital structures

and foetal membranes are relatively free from direct environmental influence and are, therefore, more conservative. Consequently, these characters, according to the author, can be employed as far more reliable criteria for determining phylogenetic and interrelationships of mammals than other morphological characters. Torpin[194] supported this view and mentioned, "It is logical that classification of animal evolution be founded upon embryological progression rather upon morphological similarities, which were ecologically determined." He also mentioned that during the course of evolution, with the progressive change from the simple apposition of the blastocyst to the uterine wall to the one where the blastocyst burries itself in the uterine mucosa, there is an independent, but simultaneous transformation of the bicornuate uterus into a simplex uterus. According to the author, this was first achieved in the phyllostomid bats. Foetal membrane characters have been employed by several authors for determining relationships of different mammalian groups[28,55,75,76,77,78,91,93,188].

On the basis of the study of the genital organs of bats, Jones[58] mentioned that the Magachiroptera form an fairly well-defined group and possess "a rather extraordinary blend of characters which are stigmatised as low with characters which are ranked as high. In many of these characters they overlap with the Microchiroptera". About the Microchiroptera he stated, "It may be suggested that we are dealing with a group which is not truly monophyletic, and that more than one primitive mammalian stock may have produced flying members, the finished individuals, derived from whatever stock, having become extraordinarily similar in their general forms by a process of convergent specialisation to their very peculiar mode of life." He further stated, "The phyllostomids may represent the terminal products of a stock which became 'bats' in the neotropical region, while the rest of the Microchiroptera evolved on similar lines, but may be from a different stock in other parts of the world." He argued that the Phyllostomids are obviously more "Primate-like" and that "this family has taken to an aerial life after a fair proto-Primate apprenticeship had been served, whereas the remaining members of the group may have had a much more lowly origin." However, he concluded that "connecting forms, not hitherto examined in sufficient detail, may be found uniting the two types and pointing to the Phyllostomids as being the highest evolution within a monophyletic group".

Studies during the past about forty years have revealed that the female genital organs of Chiroptera form an excellent example to illustrate the

possible stages of evolution of these structures. In the Megachiroptera the two uterine cornua remain separate throughout their lengths and open independently. The most primitive condition among these is exhibited by pteropids[3] in which the two sides of the female genitalia are completely separated. In *Cynopterus sphinx gangeticus* the cranial half of the vagina is separated into two canals by a prolonged extension of the cervix in the form of a dorso-ventral septum, and each uterine cornu opens by an independent cervical canal, which traverses through the septum-like cervix and opens into the vaginal canal of the respective side at about the middle of the length of the cervix. In *Rousettus leschenaulti*[68] the lumina of the two uterine cornua remain separate and open by independent cervical canals into a common vagina. In all the Microchiroptera, except the phyllostomid bats, the uterus is bicornuate, but there are varying degrees of confluence of the uterine lumina in different species as mentioned earlier. Some of the phyllostomids exhibit the most advanced condition of the union of the cornua in possessing a simplex uterus[7,8,52,58,104,137].

Amongst the developmental and foetal membrane characters, Mossman[93-95] listed four characters as being fairly constant within each of the mammalian orders. These are, (1) the uterus-blastocyst relationship, (2) the condition of the vascular splanchnopleure of the yolk sac, (3) the nature of the foetal placenta and (4) the allantois.

Employing embryological characters Mossman[93] stated with regard to Chiroptera, "Largely because of its mesometrial orientation of the disc I have placed it (Megachiroptera) near the Rodentia quite apart from the bats." (Parenthesis ours). This conclusion was based on the meagre information on the embryology of a few Megachiroptera, and fairly detailed information possibly only on the embryology of a few members of Vespertilionidae among Microchiroptera. Moghe[89] mentioned, "The two groups (Megachiroptera and Microchiroptera) are widely separated from each other in a large number of other characters and probably represent independent offshoots from some primitive insectivore." (Parenthesis ours). But Gopalakrishna[88] and Moghe and Gopalakrishna[91] noted many embryological similarities and transitional embryological characters between the two suborders and concluded that the two groups are closely related to each other.

While assessing the evolutionary significance of the developmental characters of mammals one is confronted with the problem as to which of

the characters are of ordinal value and which are of familial significance. Thus, for example, does the character of the blastocyst-uterus relationship, which is the first feature in mammalian ontogeny reflect a higher order of relationship than the location of the definitive placental disc, which is the last feature to be established in ontogenetic sequence ? One cannot also be sure if the chronological sequence of ontogeny reflects evolutionary trend. Taxonomists have employed mostly morphological and anatomical characters for classification, and in many mammalian groups this does not really reflect their phylogeny or interrelationships. This has been amply demonstrated for several mammalian groups[75,76,77,98]. In the case of Chiroptera the morphological modifications to suit their flying habit are so unique and so different from all other mammals that it is hard to accept any theory which proposes that several divergent mammalian groups developed such closely similar morphological characters due to convergent evolution. Even from the morphological point of view there seems to be little doubt that the group is an homogeneous one although some degree of divergence has undoubtedly occurred with respect to dentition, stomach and such other ecologically determined characters.

Details of the embryology of members of several families have become available during recent years. Tables I and II give the summary of the available information on the various aspects of the embryology of Chiroptera. The position of the embryonic disc at the time of implantation of the blastocyst (Table I) is variable not only among different species in some of the families but in different specimens of the same species. Perhaps, it would be best to consider that this character has not yet become stabilized in all the families. However, the location of the allantoic placental disc is constant for a family and bears a definite relationship to the orientation of the embryonic disc at late neural groove stage when the allantois develops, the final placental disc being located in that part of the uterine wall towards which the embryonic disc faces during the neural groove stage. The only exception to this appears to be Molossidae in which the two have been reported to be at opposite sides of the uterus[117].

If developmental characters, as mentioned by Mossman[93-95], are employed for determining the inter-familial relationships among the Chiroptera, the currently accepted familial hierarchy needs to be slightly modified. One of the features revealed from such an analysis is the close developmental similarity between Pteropidae and Molossidae in three of the four characters

mentioned by Mossman[93], namely, the unique changes undergone by the yolk sac to become ultimately converted into a solid, vascular, gland-like structure, the course of development of the allantois and the mesometrial location of the definitive chorio-allantoic placental disc. There is some divergence of opinion regarding the early orientation of the embryonic disc in the two families[89,90,110,117] and, as already mentioned earlier, this is probably a variable character which has not yet stabilized in the Order. In Rhinopomatidae, Emballonuridae, Rhinolophidae and Hipposideridae the yolk-sac splanchnopleure becomes free and undergoes partial collapse, and its cellular elements hypertrophy giving it a gland-like appearance. However, in these families remnants of the yolk-sac lumen persist as narrow streak-like spaces. In Noctilionidae, Megadermatidae, Phyllostomatidae, Desmodontidae and Vespertilionidae there is a persistent abembryonic bilaminar or trilaminar omphalopleure over which the embryonic segment of the splanchnopleure invaginates.

Another interesting feature, which is revealed in the present analysis concerns the allantoic placental disc, which is mesometrially located in most of the "lower" families, whereas it is antimesometrial in the "higher" families. However, the placental disc is mesometrial in Molossidae, which has been ranked as an advanced family by taxonomists.

In most families the allantoic vesicle, which is large during earlier stages becomes progressively reduced until it is ultimately lost. A vesicular allantois has not been noticed in *Taphozous longimanus*[88], *Desmodus rotundus murinus*[180] and *Glossophaga soricina*[58]. The allantoic duct, however, persists in the umbilical cord in most families.

In all the families the placenta is labyrinthine. The finer morphology of the placenta, however, varies, being endotheliochorial (vasochorial) in some and haemochorial in others.

In a recent review of the phylogenetic analysis of the foetal membranes and reproductive features of Chiroptera, Luckett[78] supported the earlier view of Gopalakrishna[28] and Moghe and Gopalakrishna[91] that the Chiroptera is a monophyletic group, and that the Megachiroptera share a number of developmental characters with several microchiropteran families.

Theorizing on lines similar to those of Mossman[93] it is apparent that there are considerable embryological similarities among Pteropidae, Molossidae, Rhinopomatidae, Emballonuridae, Rhinolophidae and Hipposideridae. Megadermatidae shows greater similarities to these families than to the

others except for the persistence of the bilaminar omphalopleure. The partially interstitial implantation of the blastocyst in Noctilionidae would place this family near Desmodontidae and Phyllostomatidae. Vespertilionidae would find its place between Megadermatidae and Noctilionidae.

In the absence of adequate palaeontological record, and because of extreme morphological specializations, we have to depend largely on embryological characters for determining inter-familial relationships of Chiroptera. This is not to minimize the possibilities of developmental adaptation. However, one can certainly accept that developmental characters, specially in eutherian mammals whose development takes within the confines of the uterus, must be far less prone to undergo drastic modifications than the ecologically determined morphological characters. Allowing that there is some possibility of developmental adaptation, and allowing that ontogenetic chronology may not always strictly reflect evolutionary trends, text figure 10 has been drawn to indicate the possible trends of evolution of the foetal membranes culminating in the extant families of Chiroptera purely on embryological grounds.

Presumably the two major divisions, namely the Megachiroptera and the Microchiroptera, were erected by earlier workers on the basis of the gross size of the animals as the terms imply. The larger Old World fruit bat (*Pteropus*)—the 'flying fox'—is so much larger than all the Old World insectivorous bats that it was justifiable to apply the terms "Megachiroptera" and "Microchiroptera" to the two groups respectively. Later, other morphological and anatomical characters such as, for example, the presence of a claw in the index finger in the Megachiroptera, the nature of dentition and characters of alimentary canal were also noticed as differentiating the two groups. The terminology and the subdivision of the Order into two major Suborders have persisted until now although it is now known that some of the Megachiroptera like *Cynopterus* and *Rousettus* are nearly of the same size as some phyllostomids in the New World and *Taphozous kacchensis* and *Hipposideros lankadiva* of the Old World.

On purely embryological grounds this division into the two Suborders is not justified as revealed from the foregoing discussion since there are numerous significant embryological characters in which the two resemble each other very closely. The question then would pertain to the interrelationships of the various families.

There appears to be little doubt that the group Chiroptera is monophyletic and has arisen from a common insectivore-like ancester from which

several other Orders of mammals have also arisen. The branch which has diversified into Pteropidae, Molossidae, Rhinopomatidae, Emballonuridae, Rhinolophidae, Hipposideridae and Megadermatidae can be presumed to have arisen from this stem after the basic chiropteran characters were established. The rodents probably diverged off from the base of this stem before the chiropteran characters were established. A common ancestral group, which diversified into Phyllostomatidae, Desmodontidae and Noctilionidae, probably arose as another branch from about the same level as, but considerably divergent from, the branch mentioned above. It is reasonable to presume that Noctiliomidae must have arisen from the base of this branch, while Desmodontidae and Phyllostomatidae diversified much later in the evolutionary history of this group. The similarities in the construction of the female genital organs and the developmental characters between the Phyllostomatidae and the primates are so striking that it is tempting to consider that the primates could have diverged off from the base of this branch before the chiropteran characters were established. The Vespertilionidae may be considered as a direct offshoot from the base of the chiropteran stem. These relationships as proposed here are shown in text figure 11.

Most taxonomists include the genus *Miniopterus* in the Family Vespertilionidae in the Sub-family Miniopterinae[114]. However, on comparing the developmental characters of *Miniopterus schreibersii fuliginosus*[10,33] with those of other vespertilionids so far known, it becomes evident that *Miniopterus* differs from all vespertilionids in some major respects, such as :

(1) Whereas the blastocyst in all other vespertilionids attaches to the antimesometrial side of the uterus by the embryonic hemisphere, and the abembryonic part of the blastocyst lies freely in a persistent uterine lumen, the blastocyst of *Miniopterus* establishes contact with the uterine wall on its entire surface circumferentially so that the uterine lumen is obliterated at the level of nidation. In this respect *Miniopterus* resembles some Pteropidae[59,90], Rhinopomatidae[28,116], Emballonuridae[26], Megadermatidae[41], Rhinolophidae[4], Hipposideridae[28,45,56,57,64] and Molossidae[110].

(2) In most other vespertilionids the primitive amniotic cavity arises by cavitation, and this cavity becomes roofed over by the trophoblastic layer after the original roof of the primitive amniotic cavity is lost[11,12,13,14,66,103,126,128]. In a few vespertilionids the roof of the primitive amniotic cavity continues as the ectodermal component of the definitive amnion[23,72]. In

Miniopterus, on the other hand, the loss of the roof of the primitive amniotic cavity exposes the embryonic disc to the potential uterine lumen since there is no trophoblastic layer in this region. Further, there is a considerable delay in the development of the definitive amniotic folds, and during this period the amniotic cavity is roofed over by the uterine endometrial layer itself.

(3) The development of the yolk sac with the abembryonic segment remaining permanently bilaminar is unlike that in all other vespertilionids.

(4) The development and the ultimate structure of the placenta in *Miniopterus* do not find a parallel among any mammal, let alone any bat.

The foregoing analysis of the embryological characters of *Miniopterus* based on the study of the embryology of *Miniopterus schreibersii fuliginosus* calls for a reassessment of the taxonomical status of *Miniopterus*. On purely embryological grounds it appears justifiable to separate *Miniopterus* from Vespertilionidae, and to erect a new Family—Miniopteridae to include this genus with its various species. From this analysis it is evident that this genus must have arisen from the base of the vespertilionid branch before Vespertilionidae acquired their special characters. The position of the newly proposed family, Miniopteridae, has also been shown in text figures 10 and 11.

Mein and Tupinier[88] have made a similar suggestion on the basis of their studies on the dentition of *Miniopterus schreibersii*. They noticed the presence of a supplementary vestigeal tooth between the upper canine and the first premolar. This, according to the authors, justified the creation of a separate family, Miniopteridae, to include this genus.

Surprisingly the schemata shown in text figures 10 and 11, based only on embryological data, correspond very nearly to the taxonomic hierarchy given by Simpson[114] except for the position of Noctilionidae, Megadermatidae and Molossidae, and the creation of the family Miniopteridae. Perhaps, if details of the embryology of the other six families are also known, it may be possible to determine not only the phylogeny of the Order as a whole but also the interrelationships of the various families in this interesting group of mammals.

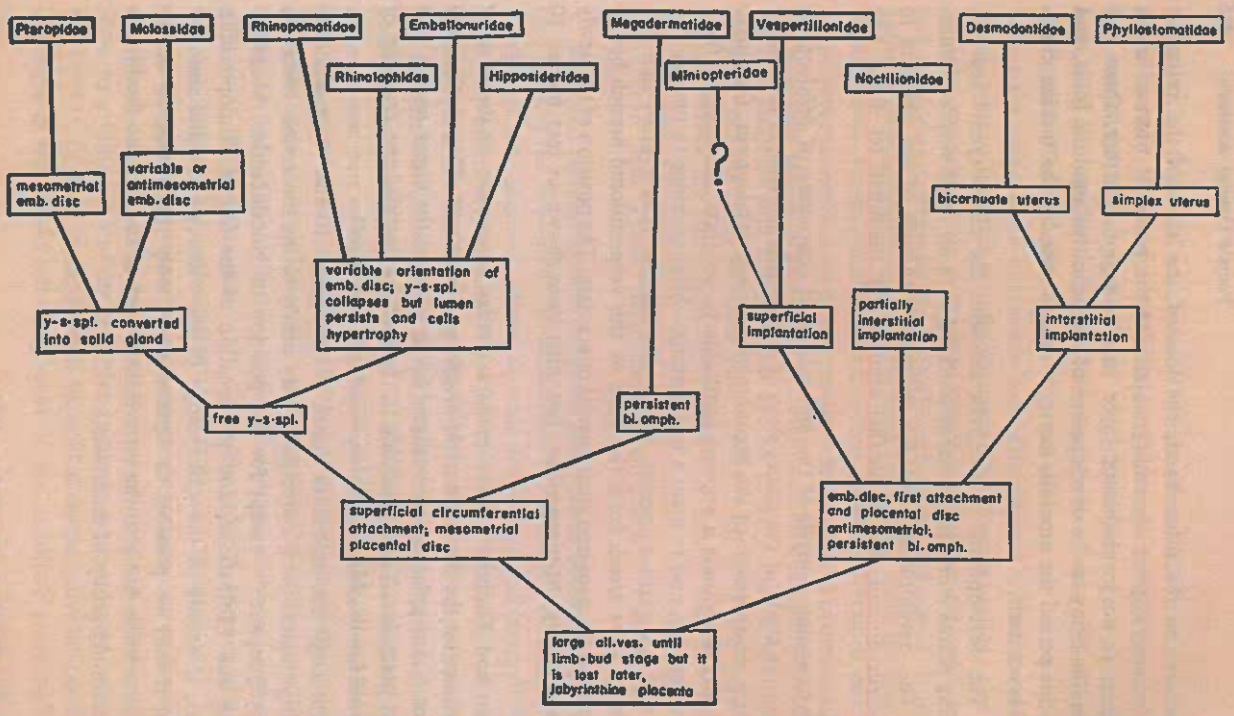


Fig. 10. Schematic representation based on embryological grounds to indicate the trends in the evolution of the different families. The families about which the details of embryology are not known are not included in the schema

National Academy of Sciences, India

Golden Jubilee Commemoration Volume, 1980

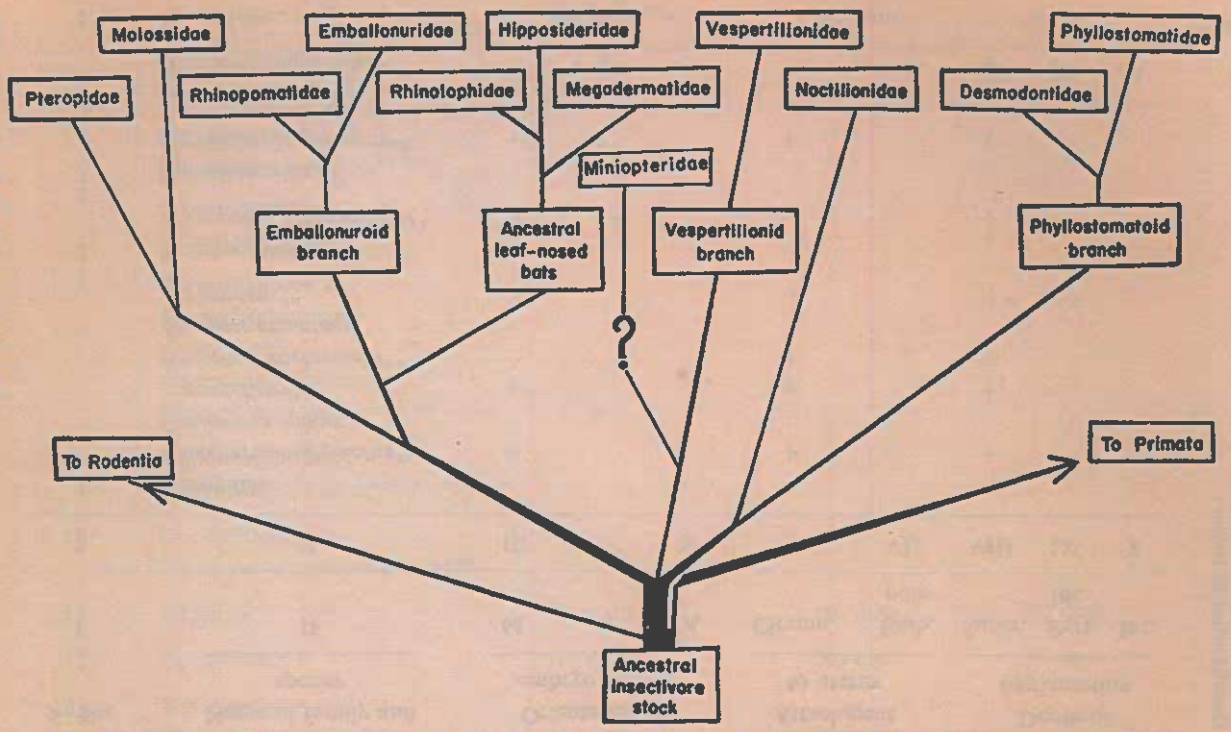


Fig. 11. Schematic representation of the relationships of the various families of Chiroptera based on embryological characters. The question mark indicates that "Miniopteridae" has been proposed now, and its position is open for debate

Table I: BLASTOCYST-UTERUS RELATIONSHIP IN CHIROPTERA

S. No.	Name of family and species	Orientation of embryonic disc			Attachment to uterus		Depth of implantation		
		M	L	A	Circum.	Emb. pole	Super.	Part. Int.	Int.
I	II	III	IV	V	VI	VII	VIII	IX	X
1.	Pteropidae :								
	Cynopterus marginatus[69]	+			+		+		
	Cynopterus sphinx gangeticus[90]	+			+		+		
	Rousettus leschenaulti[59]	+			+		+		
	Pteropus giganteus giganteus[89]	+[1]			+		+[2]	+[2]	
2.	Rhinopomatidae :								
	Rhinopoma kinneari[28,116]	+[1]	+[1]		+		+		
3.	Emballonuridae :								
	Taphozous longimanus[28]	+[3]	+[3]		+		+		
4.	Noctilionidae :								
	Noctilio labialis minor[2]		+[4]	+[4]		+	+[2]	+[2]	
5.	Nycteridae :	No data			No data		No data		

Table I (Contd.)

I	II	III	IV	V	VI	VII	VIII	IX	X
6.	Megadermatidae :								
	Megaderma lyra lyra[40,41]	+[5]	+[5]	+[5]	+		+		
7.	Rhinolophidae :								
	Rhinolophus rouxi[4]	+[1]	+[1]		+		+		
8.	Hipposideridae :								
	Hipposideros bicolor pallidus[28]				+		+		
	Hipposideros fulvus fulvus[36,64]		+[6]	+[6]	+		+		
	Hipposideros speoris[57]		+[4]	+[4]	+		+		
9.	Phyllostomatidae :								
	Glossophaga soricina[53,104,105]			+[7]					+[8]
10.	Desmodontidae :								
	Desmodus rotundus murinus[130]			+		+			+
11.	Natalidae :	No data			No data		No data		
12.	Furipteridae :	No data			No data		No data		
13.	Thyropteridae :								
	Thyroptera tricolor[134]	No data			No data				+
14.	Myzopodidae :	No data			No data		No data		

Table I (Contd.)

I	II	III	IV	V	VI	VII	VIII	IX	X
15.	Vespertilionidae :								
	Vespertilio murinus[96]			+		+	+		
	Vesperugo noctula[126]			+		+	+		
	Vesperugo leisleri[103]			+		+	+		
	Myotis lucifugus lucifugus[128]			+		+	+		
	Scotophilus wroughtoni[93]			+		+	+		
	Pipistrellus pipistrellus[99]			+		+	+		
	Pipistrellus ceylonicus chrysothrix[98]			+		+	+		
	Pipistrellus dormeri[111]			+		+	+		
	Pipistrellus mimus mimus[65]			+		+	+		
	Scotophilus heathi[72]			+		+	+		
16.	Mystacinidae :	No data			No data			No data	
17.	Molossidae :								
	Molossus rufus and *M. obscurus*[110]		+			+	+		
	Tadarida brasiliensis cynocephala[93,117]			+		+	+		

Table I (Concld.)

I	II	III	IV	V	VI	VII	VIII	IX	X
18.	Miniopteridae : (New family suggested now)								
	Miniopterus schreibersii fuliginosus			+	+		+		

M—Mesometrial L—Lateral A—Antimesometrial

1. Changes from lateral to mesometrial side
2. There is a growth of decidua basalis around the embryonic pole of the blastocyst
3. Changes from the mesometrial to lateral side
4. Embryonic disc faces between lateral and antimesometrial sides
5. Variable orientation of the embryonic mass
6. Changes from antimesometrial side to lateral side
7. Fundic implantation
8. Utero-tubal implantation.

Table II : SUMMARY OF FOETAL MEMBRANE CHARACTERS IN CHIROPTERA

Sl. No.	Family	Amniogenesis			Yolk sac	Allantoic vesicle
		A	B	C		
I	II	III	IV	V	VI	VII
1.	Pteropidae	+			Free splanchnopleure converted into a solid gland-like structure	Large during early stages ; obliterated during late pregnancy ; duct persists in the umbilical cord
2.	Rhinopomatidae		No data		Free yolk-sac splanchnopleure ; cellular elements undergo hypertrophy ; lumen persists as streak-like spaces	As in No. 1
3.	Emballonuridae		No data		As in No. 2	Allantoic vesicle not present ; duct present in the umbilical cord
4.	Noctilionidae	+			Persistent trilaminar omphalopleure	As in No. 1
5.	Nycteridae		No data		No data	No data
6.	Megadermatidae		+		As in No. 4	Large during early stages ; becomes progressively reduced and is finally lost
7.	Rhinolophidae		+		As in No. 2	As in No. 1
8.	Hipposideridae		+		As in No. 2	As in No. 1
9.	Phyllostomatidae	+			Persistent bilaminar omphalopleure	Allantoic duct without vesicle
10.	Desmodontidae	+			As in No. 9	Allantois small, non-vesicular, transient
11.	Natalidae		No data		No data	No data
12.	Furipteridae		No data		No data	No data
13.	Thyropteridae	+			As in No. 4	No data
14.	Myzopodidae		No data		No data	No data
15.	Vespertilionidae		+		Persistent bilaminar or trilaminar omphalopleure	As in No. 6
16.	Mystacinidae	-	No data		No data	No data
17.	Molossidae			+	As in No. 1 or small lumina persist	As in No. 1
18.	Miniopteridae (New Family suggested now)		+		As in No. 9	As in No. 15

A : Cavitation ; B : Primary amniotic cavity formed by cavitation, definitive amnion formed by the development of folds ; C : folding ; D : Mesometrial ; E : Lateral ; F : Antimesometrial.

1. Allantoic placenta mesometrial and lateral ; the mesometrial part is converted into an haematoma.
2. The allantoic placenta is located between the lateral and antimesometrial sides.
3. Placental disc is on the fundic side of the simplex uterus.

Table II (Concld.)

Location of placental disc			Finer Morphology of placenta	Authorities
D	E	F		
VIII	IX	X	XI	XII
+			Endotheliochorial* or Haemochorial	Keibel[59] ; Moghe[89,90] ; Wimsatt[131] ; Karim[66] ; Gopalakrishna and Karim[39] ; Karim *et al.*[67]
+			Endotheliochorial	Srivastava[116] ; Gopalakrishna[28] ; Wimsatt[131]
+[1]	+[1]		Endotheliochorial	Gopalakrishna[28] ; Wimsatt and Gopalakrishna[135] ; Karim[66]
	+[2]	+[2]	Endotheliochorial	Anderson and Wimsatt[2]
No data			No data	
+			Endotheliochorial	Gopalakrishna and Khaparde[42] ; Gopalakrishna and Karim[38]
+			Endotheliochorial	Bhiwgade[5]
+			Endotheliochorial	Gopalakrishna[28] ; Gopalakrishna and Karim[37] Karim[66] ; Jeevaji[56]
		+[3]	Haemochorial	Hamlett[53,58] ; Wislocki and Fawcett[137]
		+	Haemochorial	Wimsatt[130] ; Bjorkman and Wimsatt[6]
	No data		No data	
	No data		No data	
		+	Haemochorial	Wimsatt and Enders[134]
	No data		No data	
		+	Haemochorial	Wimsatt[128,129] ; Gopalakrishna[28,24] ; Enders and Wimsatt[17] ; Karim[66]
	No data		No data	
+			Haemochorial	Mossman[93] ; Stephens[117,118] ; Gopalakrishna and Karim[38] ;
		+[4]	Endotheliochorial[5] and Haemochorial	Gopalakrishna and Chari[38] ; Chari[10]

4. Two discs in the cranio-caudal axis on the antimesometrial side.
5. Secondary placental discs endotheliochorial ; tertiary placenta haemochorial ; primarily placenta endotheliochorial in the earlier stages, but undergoes degeneration during final stages of gestation.

* Wislocki (1954) suggested that the term "Vasochorial" is more appropriate than "Endotheliochorial" (Reference : Discussion in "*Gestation*", *Transaction of the First Conference*, page 221).

Abbreviations Used

all.pl	: allantoic placenta
all.v	: allantoic vesicle
am	: amniotic
am.c	: amniotic cavity
am.f	: amniotic fold
bi.om	: bilaminar omphalopleure
c.c	: cervical canal
ch-v-pl	: chorio-vitelline placenta
exo	: exocoelom
g	: cleft in the center of the placental disc
h	: haematoma
me	: mesoderm
mes	: mesometrium
p.all. pl	: primary allantoic placenta
p.am.c	: primary amniotic cavity
s.all.pl	: secondary allantoic placenta
s.am.c	: secondary amniotic cavity
t.all.pl	: tertiary allantoic placenta
tri.om	: trilaminar omphalopleure
u.l	: uterine lumen
v.l	: vaginal lumen
y.s	: yolk sac
y-s.c	: yolk-sac cavity
y-s.g	: yolk-sac gland
y-s spl	: yolk-sac splanchnopleure

REFERENCES

1. Anand Kumar, T. C. (1965) *J. Zool.* (London), **147** : 147.
2. Anderson, J. W., and Wimsatt., W. A. (1963) *Amer. J. Anat.*, **112** : 181.
3. Ashfaque, M. and Tungare, S. M. (1960) *Bull. Zool. Soc., Coll. Sci. Nagpur.*, **3** : 1.
4. Bhiwgade, D. A. (1976) *Proc. Ind. Acad. Sci.*, B **84** : 201.
5. Bhiwgade, D. A. (1977) *Ibid.*, B **86** : 61.
6. Bjorkman, N. H. and Wimsatt, W. A. (1968) *Anat. Rec.*, **162** : 83.
7. Bleier, W. J. (1975) *Anat. Rec.*, **182** : 237.
8. Bonilla, H. De. and Rasweiler IV, J. J. (1974) *Anat. Rec.*, **179** : 385.
9. Branca, A. (1927) *Arch. de Zool. Exper. et Gen.*, **66** : 291.
10. Chari, G. C. (1980) Early development, development of the foetal membranes and placentation in one species of Microchiroptera, *Miniopterus schreibersii fuliginosus*. Ph.D. thesis. Nagpur University.
11. Da Costa, A. C. (1920) *Mem. Soc. Port. de Sci. Nat., Biol. Ser.* **3** : 1.
12. Duval, M. (1894) *Jour. de l'Anat. et de la Physiol.*, **31** : 93-160 and 427.
13. Duval, M. (1895) *Ibid.*, **32** : 105-164 and 420.
14. Duval, M. (1896) *Ibid.*, **33** : 1.
15. Duval, M. (1899) *Etudes sur l'Embryologie des Cheiropteres, Premiere partie.*. Paris.
16. Dwyer, P. D. (1963) *Aust. J. Zool.*, **11**(2) : 219.
17. Enders, A. C. and Wimsatt, W. A. (1968). *Amer. J. Anat.*, **122** : 453.
18. Enders. A. C., Wimsatt, W. A. and King, B. F. (1976) *Ibid*, **146** : 1.
19. Fleming, T. H. (1971) *Science. N.Y.*, **171** : 402.
20. Gerard, P. (1928) *Arch. de Biol.*, **38** : 327.
21. Gohre, R. (1892) *Keimbildung des Kalong (Pteropus edulis). Selenka's studien uber Entwicklungsgeschichte*, Weisbaden, **5** : 209 and 218.
22. Gopalakrishna, A. (1947) *Proc. Ind. Acad. Sci.*, **26** : 219.
23. Gopalakrishna, A. (1949) *Ibid.*, **30** : 226.
24. Gopalakrishna, A. (1950a) *Ibid.*, **31** : 235.
25. Gopalakrishna, A. (1950b) *Proc. Nat. Inst. Sci. India*, **16** : 93.
26. Gopalakrishna, A. (1954) *Curr. Sci.*, **23** : 60.
27. Gopalakrishna, A. (1955) *Proc. Nat. Inst. Sci. India*, **21** : 29.
28. Gopalakrishna, A. (1958) *J. Morph.*, **102** : 157.
29. Gopalakrishna, A. (1964) *Curr. Sci.*, **33** : 558.
30. Gopalakrishna, A. (1969) *Ibid.*, **38** : 388.
31. Gopalakrishna, A. (1971) *J. Zool. Soc. India*, **23** : 55.
32. Gopalakrishna, A. and Bhiwgade, D. A. (1974) *Curr. Sci.*, **43** : 516.
33. Gopalakrishna, A. and Chari, G. C. (1980) 2nd All India Symposium on Life Sciences. Abstr., **154** : 63.
34. Gopalakrishna, A. and Karim, K. B. (1971) *Curr. Sci.*, **40** : 490.
35. Gopalakrishna, A. and Karim, K. B. (1972) *Ibid.*, **41** : 144.
36. Gopalakrishna, A. and Karim, K. B. (1973) *Ibid.*, **42** : 860.

37. Gopalakrishna, A. and Karim, K. B. (1976) *Rev. Roum. Biol. Ser. Zool.*, 20 : 257.
38. Gopalakrishna, A. and Karim, K. B. (1979) *J. Repr. Fert.*, 56 : 417.
39. Gopalakrishna, A., Karim, K. B. and Chari, G. C. (1979) *Curr. Sci.*, 48 : 716.
40. Gopalakrishna, A. and Khaparde, M. S. (1972) *Curr. Sci.*, 41 : 738.
41. Gopalakrishna, A. and Khaparde, M. S. (1978a) *Proc. Ind. Acad. Sci.*, 87 : 91.
42. Gopalakrishna, A. and Khaparde, M. S. (1978b) *Ibid.*, 87 : 179.
43. Gopalakrishna, A. and Madhavan, A. (1977) *J. Bombay Nat. Hist. Soc.*, 74 : 511.
44. Gopalakrishna, A. and Madhavan, A. (1978) *J. Exptl. Biol.*, 16 : 852.
45. Gopalakrishna, A. and Moghe, M. A. (1960) *Zeitschr. f. Anat. u. Entwickl.*, 122 : 137.
46. Gopalakrishna, A. and Murthy, K. V. R. (1960) *Bull. Zool. Soc. Coll. Sci, Nagpur*, 3 : 19.
47. Gopalakrishna, A. and Ramakrishna, P. A. (1977) *Curr. Sci.*, 46 : 767.
48. Gopalakrishna, A. and Rao, K. V. B. (1977) *Jour. Bombay. Nat. Hist. Soc.*, 74 : 213.
49. Gopalakrishna, A. and Sapkal, V. M. (1974) *J. Zool. Soc. India*, 26 : 1.
50. Gopalakrishna, A., Thakur, R. S. and Madhavan, A. (1975) *Ibid.* Dr. B. S. Chauhan Comm. Volume, 225.
51. Grasse, P. P. (1955) *Ordre des Cheiropteres.* in *Traite de Zoologie*, 17 : 1729. Masson et Cie, Paris.
52. Hamlett, G. W. D. (1934) *Anat. Anz.*, 79 : 146.
53. Hamlett, G. W. D. (1935) *Amer. J. Anat.*, 56 : 327.
54. Hamlett, G. W. D. (1947) *Anat. Rec.*, 97 : 340.
55. Hill, J. P. (1932) *Phil. Trans. Roy. Soc. London*, B 221 : 45.
56. Jeevaji, I. H. (1973) Early development, development of the foetal membranes and placentation in the Indian Leaf Nosed bat, *Hipposideros speoris* (Schneider). Unpublished Ph.D. thesis, Nagpur University.
57. Jeevaji, I. H. (1979) *Proc. Ind. Acad. Sci.*, B 88 : 287.
58. Jones, F. W. (1917) *Jour. Anat.*, 51 : 36.
59. Karim, K. B. (1971) Early development, development of foetal membranes and placentation in the Indian fruit bat, *Rousettus leschenaulti* (Desmarest). Unpublished Ph.D. thesis, Nagpur University.
60. Karim, K. B. (1972a) *Proc. Ind. Acad. Sci.*, 76 : 71.
61. Karim, K. B. (1972b) *J. Zool. Soc. India*, 24(2) : 135.
62. Karim, K. B. (1973) *Curr. Sci.*, 42 : 62.
63. Karim, K. B. (1975a) *Ibid.*, 44 : 706.
64. Karim, K. B. (1975b) *Rev. Roum. Biol.*, 19(4) : 251.
65. Karim, K. B. (1975c) *J. Zool. Soc. India*, 27(122) : 119.
66. Karim, K. B. (1976) Embryology of some Indian Chiroptera. D.Sc. thesis, Nagpur University.
67. Karim, K. B., Wimsatt, W. A. and Gopalakrishna, A. (1978) *Anat. Rec.*, 190 : 438.
68. Karim, K. B., Wimsatt, W. A., Enders, A. C. and Gopalakrishna, A. (1979) *Anat. Rec.*, 195 : 493.

69. Keibel, F. (1922) *Arch. f. Mikros-Anat.*, 96 : 528.
70. Kempermann, C. T. (1929) *Zeit. f. Anat. u. Entwickl.*, 91 : 292.
71. Khaparde, M. S. (1975) *Jour. Zool. Soc. India*, 28(1 & 2) : 25.
72. Khan, M. S. (1980) Early development, development of foetal membranes and placentation in one species of Microchiroptera—*Scotophilus heathi* (Horsefield), Ph.D. thesis, Nagpur University.
73. Kirby, D. R. S. (1971) *Biology of the blastocyst.*, University of Chicago Press.
74. Kirby, D. R. S., Potts, D. M. and Wilson, I. B. (1967) *J. Embryol. Exptl. Morph.* 17 : 527.
75. Luckett, W. P. (1969) *J. Rep. Fert. Suppl.*, 6 : 419.
76. Luckett, W. P. (1974a) *Contrib. Primat.*, 3 : 142.
77. Luckett, W. P. (1974b) *In Prosimian Biology*, Duckworth, London, 475.
78. Luckett, W. P. (1978) *Proc. Vth. International Bat Conf. (in press).*
79. Madhavan, A. (1971) *Mammalia*, 35(2) : 283.
80. Madhavan, A. (1979) *Jour. Bombay Nat. Hist. Soc.*, 75(2) : 426.
81. Madhavan, A., Datil, D. R. and Gopalakrishna, A. (1979) *Ibid.*, 75 : 96.
82. Marshall, A. J. (1948) *Proc. Linn. Soc. London*, 159 : 103.
83. Marshall, A. J. (1949) *Ibid.*, 161 : 26.
84. Marshall, A. J. (1953) *J. Endocrinol*, 9 : 42.
85. Matthews, L. H. (1937) *Trans. Zool. Lond.*, 23 : 224.
86. Matthews, L. H. (1942) *Proc. Zool. Soc. London*, 23 : 224.
87. Medway, Lord. (1972) *Zool. J. Linn. Soc.*, 51 : 33.
88. Mein, P. and Tupinier, Y. (1977) *Mammalia*, 41(2) : 207.
89. Moghe, M. A. (1951) *Proc. Zool. Soc. London*, 121 : 703.
90. Moghe, M. A. (1956) *Proc. Nat. Inst. Sci. India*, 22 : 48.
91. Moghe, M. A. and Gopalakrishna, A. (1959) *Bull. Zool. Soc., Coll. Sci. Nagpur*, 2 : 19.
92. Moghe, S. M. (1956) *Proc. Nat. Inst. India*, 22 : 34.
93. Mossman, H. W. (1937) *Contrib. Embryol. Carnegie Inst. Wash.*, 26 : 127.
94. Mossman, H. W. (1953) *J. Mammal*, 34 : 289.
95. Mossman, H. W. (1971) *In* The Biology of the Blastocyst, University of Chicago Press, 52.
96. Nolf, P. (1896) *Arch. de Biol.*, 14 : 561.
97. Pearson, O. P., Koford, M. R. and Pearson, A. K. (1952) *J. Mammal*, 33 : 273.
98. Phansalkar, R. B. (1972) Early development and placentation in the vespertilionid bat, *Pipistrellus ceylonicus chrysothrix*, Ph.D. thesis, Nagpur University.
99. Potts, D. M. and Racey, P. A. (1971) *Micron.*, 2 : 322.
100. Ramakrishna, P. A. (1977) *Curr. Sci.*, 46 : 651.
101. Ramakrishna, P. A. and Madhavan, A. (1977) *Proc. Ind. Acad. Sci.*, 86 B : 117.
102. Ramakrishna P. A. and Rao, K. V. B. (1977) *Curr. Sci.*, 42 : 270.
103. Ramaswami, L. S. (1933) *Half-yrly. J. Mysore Univ.*, 7 : 1.
104. Rasweiler IV, J. J. (1972) *J. Rep. Fert.*, 31 : 249.
105. Rasweiler IV, J. J. (1974) *Amer. J. Anat.*, 139 : 1.

106. Rasweiler IV, J. J. (1977) *Ibid.*, **150** : 269.
107. Rasweiler IV, J. J. (1979) *J. Rep. Fert.*, **56** : 403.
108. Richardson, E. G. (1977) *J. Zool. Soc. London*, **183** : 353.
109. Robin, H. A. (1881) *Ann. Sci. Nat.*, **12** : 1.
110. Sansom, G. S. (1932) *Proc. Zool. Soc. London, Part I*, 113.
111. Sapkal, V. M. (1973) Early development, development of the foetal membranes and placentation in the vespertilionid bat, *Pipistrellus dormeri* (Dobson), Ph.D. Thesis, Nagpur University.
112. Selenka, E. (1892) *Studien uber Entwicklungsgesrhite der Tiere*, **5** : 209.
113. Sherman, H. B. (1937) *J. Mammal*, **18** : 176.
114. Simpson, G. G. (1945) *Bull. Amer. Mus. Nat. Hist.*, **85** : 1.
115. Sluiter, J. W. and Bels, L. (1951) *Proc. K. Ned. Acad. Wet. Sec. C.*, **54** : 585.
116. Srivastava, S. C. (1952) *Proc. Zool. Soc. Bengal*, **5** : 105.
117. Stephens, R. J. (1962) *Amer. J. Anat.*, **111** : 259.
118. Stephens, R. J. (1969) *J. Ultrastruct. Res.*, **28** : 371.
119. Stephens, R. J. and Cabral, L. (1971) *Anat. Rec.*, **169** : 243.
120. Stephens, R. J. and Easterbrook, N. (1968) *J. Ultrastruct. Res.*, **24** : 239.
121. Stephens, R. J. and Easterbrook, N. (1969) *Amer. J. Anat.*, **124** : 47.
122. Stephens, R. J. and Easterbrook, N. (1971) *Anat. Rec.*, **169** : 207.
123. Thakur, R. S. (1972) Reproduction in the Indian vespertilionid bat, *Pipistrellus mimus mimus* (Wroughton) Ph.D. thesis, Nagpur University.
124. Torpin, R. (1976) *Anat. Rec.*, **184** : 457.
125. Uchida, T. (1951) *Sci. Bull. Fac. Agr. Kuyushu University*, **12** : 11.
126. Van der Stricht, O. (1899) *Anat. Anz. Verh. d. Anat. Gessel.*, **13** : 76.
127. Walker, E. P. (1964) *Mammals of the World.*, Baltimore.
128. Wimsatt, W. A. (1944) *Amer. J. Anat.*, **74** : 355.
129. Wimsatt, W. A. (1945) *Ibid.*, **77** : 1.
130. Wimsatt, W. A. (1954) *Acta. Anat.*, **21** : 285.
131. Wimsatt, W. A. (1958) *Ibid.*, **32** : 141.
132. Wimsatt, W. A. (1975) *Biol. Repr.*, **12** : 1.
133. Wimsatt, W. A. (1979) *J. Rep. Fert.*, **56** : 345.
134. Wimsatt, W. A. and Enders, A. C. (1977) *8th Annual Symp. on Bat Res.* 18.
135. Wimsatt, W. A. and Gopalakrishna, A. (1958) *Amer. J. Anat.*, **103** : 35.
136. Wislocki, G. B. (1929) *Carnegie Inst. Washington Publ. No 394. Contrib. Embr. No. 111*, **20** : 51.
137. Wislocki, G. B. and Fawcett, D. W. (1941) *Anat. Rec.*, **81** : 307.

Solid-State Battery —Prospects and Limitations

Suresh Chandra and Rakesh C. Agrawal

Department of Physics, Banaras Hindu University
Varanasi 221 005, India

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

1. Introduction

Galvanic cells such as Leclanché cell, Daniell cell, etc., were the earliest known electrochemical batteries. These cells use aqueous electrolytic solutions. The movement of ions takes place through the aqueous electrolyte medium and the e.m.f. between the electrodes is generated as the ions liberate their charges at the electrodes. These batteries have certain limitations, such as (Chandra—*to be published*) :

(a) limited temperature range of operation (i.e., between 0–100 °C)
(b) electrode material corrosion by the electrolyte solution
(c) miniaturisation is not possible
(d) less rugged.

Hence, the need of discovering some materials which could overcome these limitations of aqueous batteries was strongly felt. Historically, there had been considerable attempts to develop solid electrolyte batteries with long shelf life which could operate in a wide range of temperature. In the earlier attempts normal ionic solids, such as alkali halides, silver halides were used as the electrolyte (*see* Section 3 for details). The room temperature operation was not possible since these solids are low conducting at room temperature. The year 1967 brought a dramatic turn in this field after the announcement of anomalously high-conducting solids, MAg_4I_5 at room temperature by Bradley and Greene[9] and Owens and Argue[77] independently and β-alumina at moderately high temperature by Yao and Kummer[189]. Since then a large number of ionic conductors have been discovered mainly with Ag^+, Cu^+, Li^+, Na^+, K^+, F^- and O^{2-} ion conduction. Owing to the high ionic conduction, these solids are termed as "Superionic Solids" or "Solid Electrolytes"[32,38,64,90,182]. Gross[35] has given an excellent review on the materials which are candidates for battery application. Some of the important solids likely to be good candidates for solid-state battery are listed in Table 1 along with their room temperature conductivity. Fig. 1 gives a comparative plot of σ vs. $1/T$ for some typical superionic solids and aqueous electrolytes.

In general an ideal solid electrolyte for battery applications should possess the properties as given in Table 2. If we compare the Tables 1 and 2, we see that the ideal solid electrolytes are yet to be discovered. However, some silver and copper ion conductors have sufficiently high room temperature conductivities but their decomposition potentials are low. Some

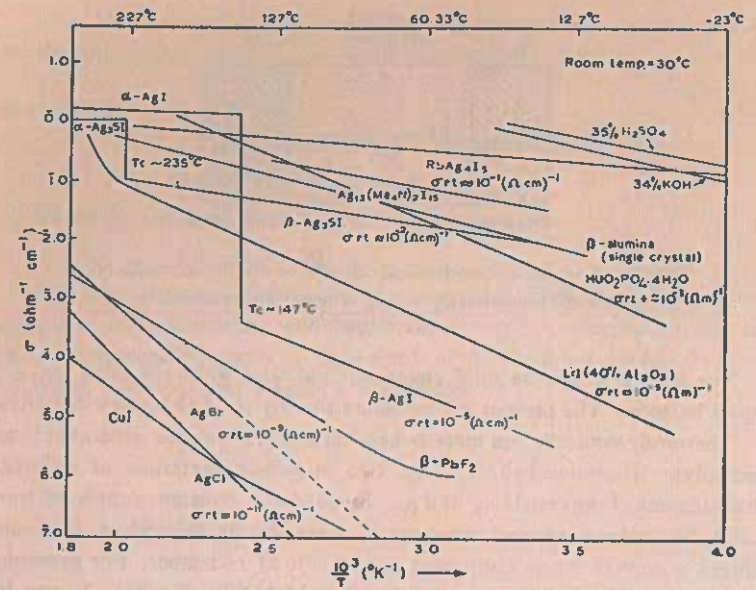


Fig. 1. σ vs. $1/T$ plot of some important solid electrolytes (normal ionic and superionic solids) along with a few aqueous electrolytes. σr_t indicates the electrical conductivity at room temperature

reasonably good alkali metal ion conductors with high decomposition potentials are also available.

In the following paragraphs, we review the recent progress made in the field of solid-state battery.

2. Principle of Solid-State Battery

A typical electrochemical battery is shown in Fig. 2. The electrolyte (also cathode and anode) can either be a solid(S) or a liquid(L) giving rise to the following important configurations[97] :

(i) $S/S/S$
(ii) $S/S/L$
(iii) $L/S/L$
(iv) $S/L/S$
(v) $L/L/L$

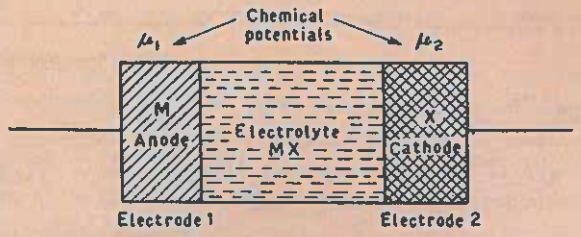


Fig. 2. A typical electrochemical cell with M and X electrodes of chemical potentials μ_1 and μ_2 respectively separated by an electrolyte MX

The first three are the solid electrolyte batteries while the other two are liquid batteries. The present review refers to only solid electrolyte batteries.

Thermodynamically, an electrochemical e.m.f. would be generated if an electrolyte is sandwiched between two electrode materials of different electrochemical potentials μ_2 and μ_1. Further, if a constant supply of ions could be maintained and transported through the electrolyte, it would deliver a current when connected across a load resistance. For example, the reaction involved in cell configuration M/MX/X (*see* Fig. 2) can be given as :

At the anode

$$M - e = M^+ \text{ (dissolved in MX)}$$

At the cathode

$$X + e = X^- \text{ (dissolved in MX)}$$

The M^+ and X^- ions come together in the solid electrolyte by ion transport to form additional MX at either or both the electrolyte/electrode interfaces. Hence, the total cell reaction would be

$$M + X \rightleftharpoons MX$$

The above is a general case for all solid electrolyte batteries. Now, let us consider a particular case of a silver ion conducting solid electrolyte battery. In the cell configuration given in Figure 2, let M electrode be the silver anode, X electrode be the iodine cathode and the electrolyte is some silver ion conductor. The cell reaction would be :

At the anode

$$Ag - e = Ag^+ \text{ (dissolved in the electrolyte)}$$

This Ag^+ ion moves towards the cathode through the electrolyte (Ag^+ are the only mobile ions in these solid electrolytes).

At the cathode

$$I + e = I^- \text{ (remains at the electrolyte/}I_2\text{ interface)}$$

$$\therefore \quad Ag^+ + I^- = AgI$$

Hence the total cell reaction can be written as

$$Ag + \tfrac{1}{2}I_2 \rightarrow AgI$$

This chemical reaction involves an energy (i.e., Gibbs' free energy) which gives the electrochemical cell e.m.f. This energy depends upon the chemical potentials μ_2 and μ_1. The e.m.f. of such systems can be calculated following Kinkkola and Wagner's, expression[48] :

$$E = \frac{1}{|Z|.F}\int_{\mu_1}^{\mu_2} t_{ion}\, d\mu$$

$$\text{or} \quad E = \frac{\mu_2 - \mu_1}{|Z|.F}\, t_{ion}$$

$$\text{or} \quad E = \frac{\Delta G}{|Z|.F}\, t_{ion}$$

where E is the electrochemical e.m.f. developed across the electrolyte sandwiched between the two electrodes of electrochemical potentials μ_2 and μ_1; t_{ion} is the ionic transference number ($t_{ion} = 1$ in most silver ion conducting solid electrolytes, since the sole-charge carriers are Ag^+ ions) ; Z is the valency of the conducting species ($= 1$ for Ag^+) and F is the Faraday's number. If we use elemental iodine as cathode, the value of theoretically calculated voltage is $\sim 0{\cdot}687$ volts, but if the chelated iodine (such as KI_3, RbI_3, $(CH_3)_4NI_5$, etc.) is used, the calculated voltage is $\sim 0{\cdot}66$ volts[75].

The values of electrochemical potentials which can be obtained for iodine cathode with different metal ions are listed in Table 3 (for some other specific cell configurations, *see* Tables 4 and 5).

3. Solid-State Batteries—Earlier Attempts

Perhaps Haber and Tollozzko[57] were the first to develop the following solid-state cells :

$$Pb/PbCl_2/AgCl/Ag \quad \text{and} \quad Cu/CuCl/AgCl/Ag$$

Their batteries did not give any appreciable current due to very high internal resistance. Since then and till early 1960's a large number of attempts were made with solids $PbCl_2$, $PbBr_2$, SrF_2, $SrCl_2$, $BaCl_2$, CaF_2, β- and α-AgI, AgBr, AgCl, CuCl, CuBr, Ag_2HgI_4, Ag_2S and many others like stabilised zirconia[31,69]. The batteries with above solids, though rugged and compact, gave very low current drain $\sim 1\ \mu a/cm^2$. Some of the earlier cells are given in Table 4. The major breakthrough in the solid-state electrochemical cell technology was given by Reuter and Hurdel[85,86,87] and Takahashi and Yamamoto[119,120,121,122] independently. They investigated the double salt of Ag_2S+AgI and found that the conductivities of Ag_3SI, primarily due to Ag^+ ions, is as high as 10^{-2} ohm^{-1} cm^{-1} at room temperature[121]. The situation improved dramatically again in 1967 after the announcement of high conduction in β-alumina[139] at moderately high temperature and at room temperature in MAg_4I_5 where M = Rb, K, NH_4[9,77]. Since then a large number of such solids have been reported (*see* Table 1). This discovery of fast ion-conducting solids renewed the activity into the development of various solid-state batteries and the recent progress is discussed in Section 4.

4. Solid-State Battery—Recent Attempts

Table 5 reports some important solid-state batteries developed recently using the various superionic solids listed in Table 1. These batteries can be divided broadly into two groups on the basis of the energy drain :

(1) Low Energy Density Batteries

(2) High Energy Density Batteries.

Most of the low energy density batteries are strictly all solid-state batteries as their constituents (cathode, electrolyte and anode) are in solid form. Whereas, high energy batteries use either solid or liquid constituents. We will consider only those batteries whose electrolytes are in the solid form. Silver and copper ion conducting batteries come in the first group (energy density $< 0{\cdot}1$ Wh/cc). Because of the low decomposition potentials of Ag and Cu, high energy cannot be drained out from these batteries. Lithium and sodium ion batteries are used both for low energy and high energy density operations. The high or low energy drain depends also upon the temperature of operation. Some anion conductors (listed in Table 1) have also

been employed for various electrochemical batteries, but these solids give reasonable ionic conduction only at very high temperatures. Hence, no successful battery of practical utility could be obtained with these solids. However O^{2-} ion conductors are widely used for fuel cells. For a discussion on fuel cells, reference may be made to Rohr[89], Markin *et al.*[66], Takahashi[112], Nowick and Park[71] (Chandra—*to be published*). Some room temperature proton conductor batteries have also recently been developed[21]. If proper materials could be developed, proton conducting batteries are very promising for future energy applications. Average theoretical power density and average theoretical energy density of some of the important electrochemical batteries are plotted in Fig. 3. Some acid and aqueous electrolyte batteries are also included for comparison. It is clear from the plot that Li/S, Na/S and proton conductor batteries are very attractive for low as well as high energy density operations. In the following paragraphs, we review some of these batteries and their discharge characteristics.

4.1. *Low Energy Density Batteries*

(a) *Silver ion conductor batteries*

This class has a large number of compounds as given in Table 1. These compounds are normally AgI-based solid electrolytes. The first battery reported by Takahashi and Yamamoto[121,122] had *S/S/S* configuration :

$$Ag/Ag_3SI/C+I_2$$

Iodine gives the e.m.f. through chemical reaction :

$$I_2(g)+Ag_3SI\ (s) \rightarrow 3AgI+S$$

Electronic conductor graphite is necessary to accept the charge liberated at the electrolyte/cathode interface. The most widely studied cell is MAg_4I_5 solid electrolyte[75,79]. For better electrode/electrolyte contact and to minimise polarisation, the electrolyte is mixed with the electrode materials. Elemental iodine is highly volatile and has tarnishing action towards the electrolyte and the metal electrodes. Hence, in place of elemental iodine, chelated iodine such as KI_3, RbI_3, etc., is used. A typical discharge curve of the cell[75]

$$Ag/RbAg_4I_5/RbI_3, C$$

is given in Fig. 4 and its significant properties are listed in Table 6. Obviously, these batteries are useful for low energy drain and give a long

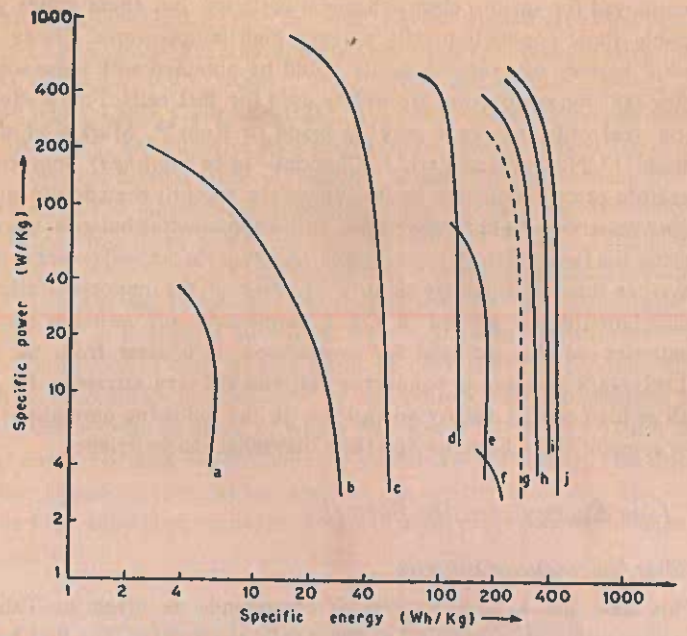


Fig. 3. Average theoretical power density vs. average theoretical energy density for different galvanic cells. (a) Ag-I_2, (b) Pb-acid, (c) Ni-Cd, (d) Ag-Zn, (e) Zn-air, (f) Li-PbI_2, (g) CaH_2/HUP /$TiFeH_2$, (h) Na-S, (i) Li-S, (j) Li-Cl_2, (after Schottens and Van Gool[97]

shelf life. But $RbAg_4I_5$ is thermodynamically unstable and dissociates into low conducting compounds below 27 °C as follows :

$$2RbAg_4I_5 \rightarrow Rb_2AgI_3 + 7AgI$$

The presence of moisture and iodine enhances this reaction. As such, other stable electrolytes like $[(Me \text{ or } Et)_4N]_2Ag_{18}I_{15}$, $PyAg_5I_5$, silver iodide—silver-oxy salt electrolytes, etc., have been used to construct *S/S/S*-type batteries[88, 85, 79].

Iodine gives highest voltage, but its use as cathode causes degradations of the battery as discussed above. Hence many other cathode materials such as Se, Te, $RbTeI_5$[36, 111], etc., are used in place of iodine. These cathodes give low Open-Circuit Voltage (OCV) (*see* Table 5). Some charge

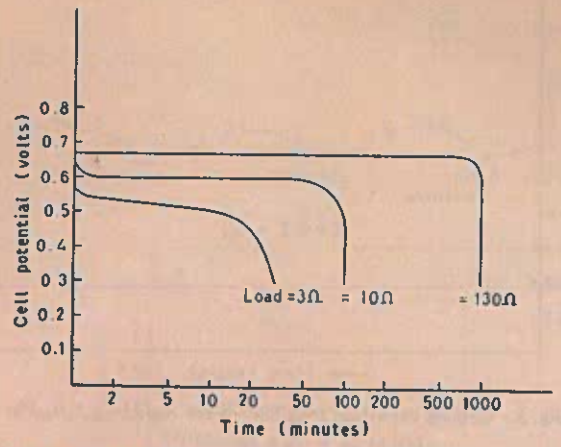


Fig. 4. Typical discharge curves at 25°C of the Ag/$RbAg_4I_5$/RbI_3, C cell (after Owens)[75]

transfer complexes (like I_2–Ph)[81] and non-stoichiometric solid solution electrodes[106] have also been used as cathode from which successful rechargeable batteries have been reported[81]. The rechargeability of a silver cell using I_2-Ph charge transfer complex as cathode with the configuration :

$$Ag/RbAg_4I_5/I_2\text{-}Ph$$

was tested. It was recharged from a state of complete discharge which may be assumed to be :

$$Ag/RbAg_4I_5/AgI,\ Ph$$

The charging efficiency is reported to be as high as 90% when the depth of active silver layer transported does not exceed 0·2 μm[10]. A typical cycling behaviour of this cell is shown in Fig. 5.

Many miniaturised batteries have been reported with the electrolyte in the film form. An excellent review on film electrolyte batteries has been given recently by Kennedy[44]. The electrolyte film growth by vacuum evaporation technique involves many parameters which are difficult to control and results in improper films. We have developed a new technique and successfully obtained films of MAg_4I_5[18, 19, 68], $[(CH_3)_4N]_2Ag_{15}I_{16}$ (Chandra and Agrawal—*to be published*) and $Ag_7I_4PO_4$[16] directly on the

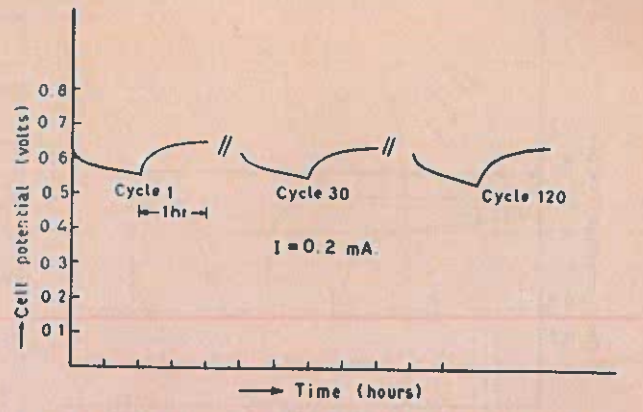


Fig. 5. Cycling behaviour (Shallow) of the $Ag/RbAg_4I_5/I_2$-Ph cell at 25 °C (after Scrosati)[99]

silver substrate and used them for film batteries. The discharge characteristics of the best obtainable films are given in Fig. 6.

The Open-Circuit Voltages (OCV) for some film batteries developed in our laboratory are given in Table 7. It may be noted that in all cases, the OCV is less than the theoretically calculated voltage of ~ 0 68 volt. This is due to the presence of partial electronic conduction in the superionic solid films. The electronic and ionic transference numbers of the electrolyte films as evaluated by us are also given in Table 7.

In conclusion, one may say that the use of silver solid-state batteries is somewhat limited at the moment by its high cost, low voltage and poor rechargeability. Nevertheless, the great amount of work which has been devoted to studying their properties remains important for the development of alternative solid-state power sources.

(b) *Copper ion conductor batteries*

Good copper ion solid electrolytes have been recently discovered by Takahashi and coworkers[116,123,126], Sammels *et al.*[91], Geller *et al.*[88], Lazzari and Scrosati[55]. Amongst the various copper ion solid electrolytes, N, N'dimethyle triethylenediamine dibromide [TED2CH$_3$Br-CuBr (94m/O)] is the one which has received most attention in view of its practical applications[118,126,127]. However, recently discovered $Rb_4Cu_{16}Cl_{18}I_7$ [118] and

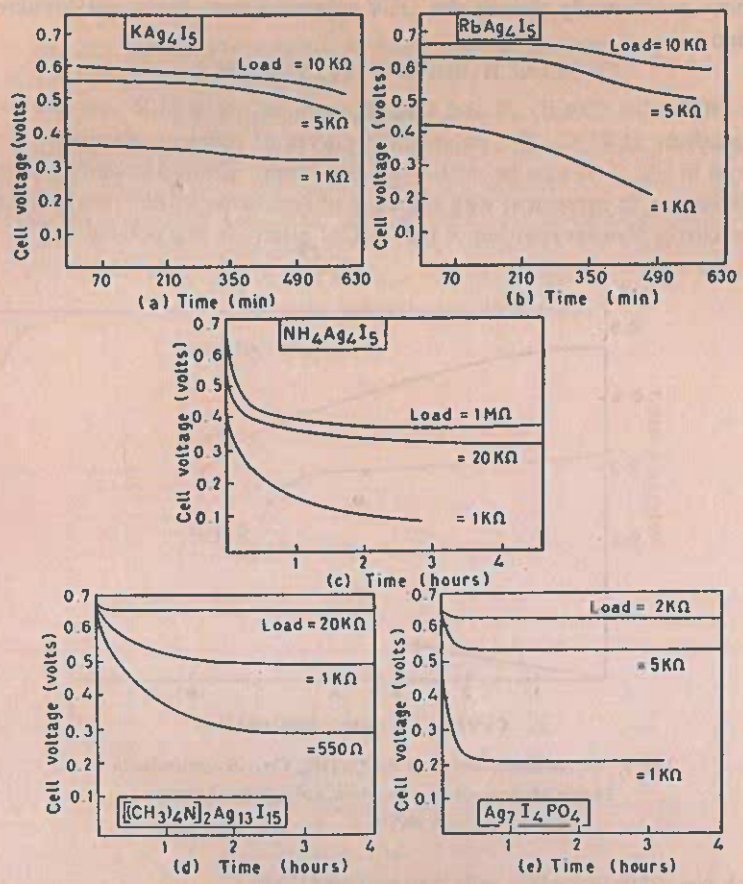


Fig. 6. Typical room temperature discharge curves for cells using same electrocodeposited films

$RbCu_4Cl_3I_2$ [33] electrolytes show highest room temperature conductivity ever reported. The copper ion solid electrolytes are incompatible with the halogen electrodes and have very low decomposition voltage like Ag^+ ion electrolytes. They also show very low value of exchange current for the copper electrodes[8]. Hence various non-stoichiometric and layer-structured electrode materials. Such as $Cu_{1.8}S$, TiS_2 have been developed which

behave satisfactorily though the OCV values are low. In the cell configuration :

$$Cu/TED2CH_3Br\text{-}CuBr/Cu_{1\cdot8}S \text{ (or } TiS_2)$$

the OCV with $Cu/Cu_{1\cdot8}S$ and Cu/TiS_2 couples are 0·34 V and 0·47 V respectively at 25 °C. The polarisation curves of different electrodes are shown in Fig. 7. It may be noted that the copper electrodes show higher polarisation, in agreement with the work of Armstrong *et al.*[8] in which the slow charge transfer reaction of the Cu/Cu^+ electrode was pointed out.

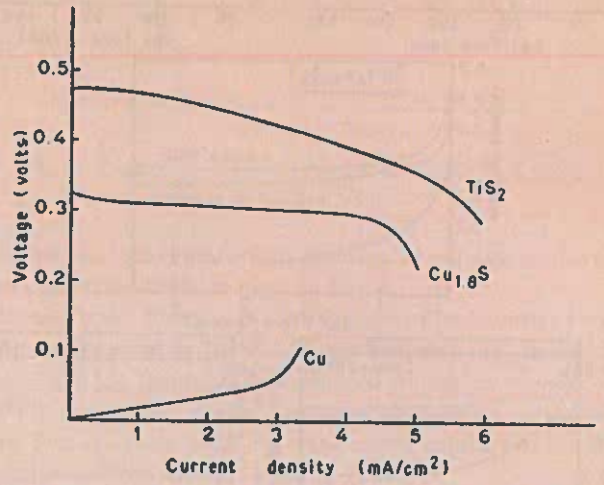


Fig. 7. Polarisation curves of the $CuTiS_2$, $Cu_{1\cdot8}S$ electrodes in $TED2CH_3Br$-CuBr electrolyte at 25 °C (after Lazzari *et al.*[84])

A few other important cells reported are[91,126] :

$$Cu/PhBr\text{-}CuBr/C$$

where PhBr stands for pyridinium bromide, and

$$Cu/HMTCH_3Br\text{-}CuBr/X, Cu_2X$$

where $HMTCH_3Br$ stands for hexamethylene tetramine bromide and X = S, Se or Te. The typical battery characteristics are given earlier in Table 5.

The present status of energy and rechargeability for copper ion batteries are still too poor to open wide ranges of applicabilities.

(c) *Alkali ion conductor batteries*

The first battery was reported by Liang and coworkers[57,59,61] which had the following configuration :

$$Li/LiI\,(Al_2O_3)/PbI_2, PbS$$

The cell reaction and other characteristics are given in Table 5. Typical discharge curves at room temperature for a 136 mAh cell is shown in Fig. 8. These batteries have very long shelf life with very low energy drain. The other commercially available Li^+ battery developed by Schneider and coworkers[93,95,96] use charge-transfer complex $P2VP.nI_2$ (i.e., Poly 2-vinylpyridineiodine) and has the following configuration :

$$Li/LiI/P2VP.nI_2$$

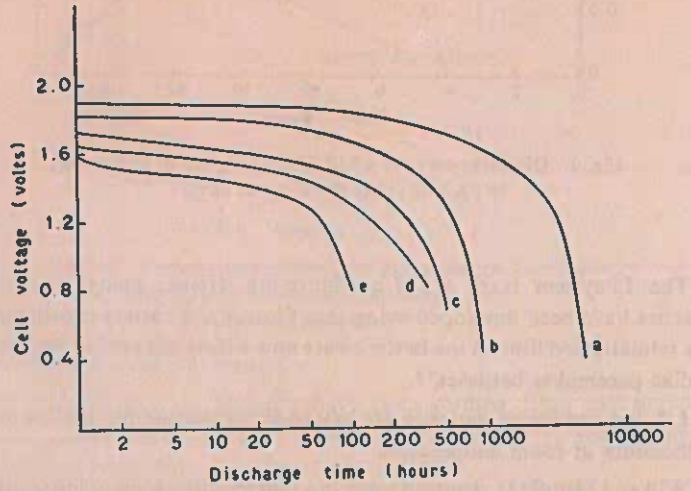


Fig. 8. Typical discharge curves of 136 mAh $Li/LiI\,(Al_2O_3)/PbI_2$, PbS cell under various constant currents and at room temperature.
(a) 18 μA, (b) 36 μA, (c) 4 μA, (d) 72 μA, (e) 90 μA
(after Liang and Barnette[60])

The self discharge losses have been estimated to be less than 10% over a period of 10 years. The cell reaction and characteristics are given in Table 5 along with some other Li^+ ion batteries. Figure 9 gives the discharge curve of this battery at 37 °C under 30 μA load.

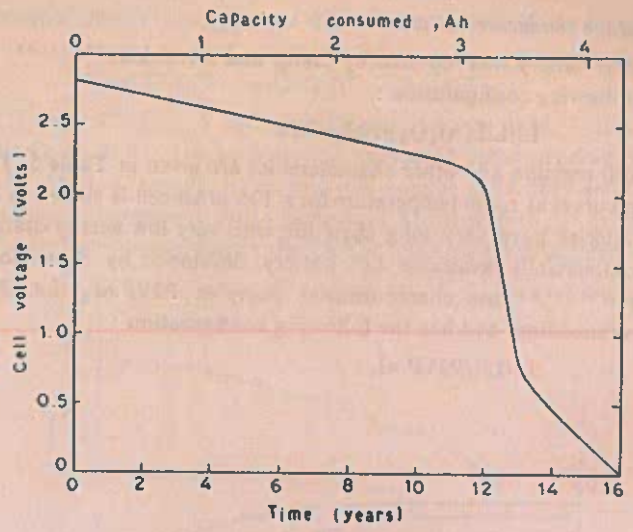


Fig. 9. Discharge curve of a Li/LiI/P2VP.nI_2 cell at 37 °C under 30 μA load (after Schneider *et al.*[94])

The Li-system is of *S/S/S* configuration. Hence, many miniaturised batteries have been developed using thin films of the battery constituents[44]. The miniaturised lithium ion batteries are now widely adapted as implantable cardiac pacemaker batteries[76].

Li^+ ion conductor batteries are extremely promising for the low energy applications at room temperature.

Will and Mitoff[188] reported some low temperature sodium ion conductor batteries with the following configuration :

$$Na\text{-}Hg/Na\text{-}\beta\text{-alumina}/O_2,\ H_2O \text{ or } Br_2 \text{ or } I_2$$

The halogen cathode is dissolved in water or in an organic solvent. The properties and cell characteristic curves of some ambient temperature primary battery with Na-β-alumina as an electrolyte are given in Table 8 and Fig. 10 respectively. The cell voltage depends upon the state of discharge because of the changing sodium activity in the amalgam anode. This battery has a shelf-life of $\sim$5 years and commercially available.

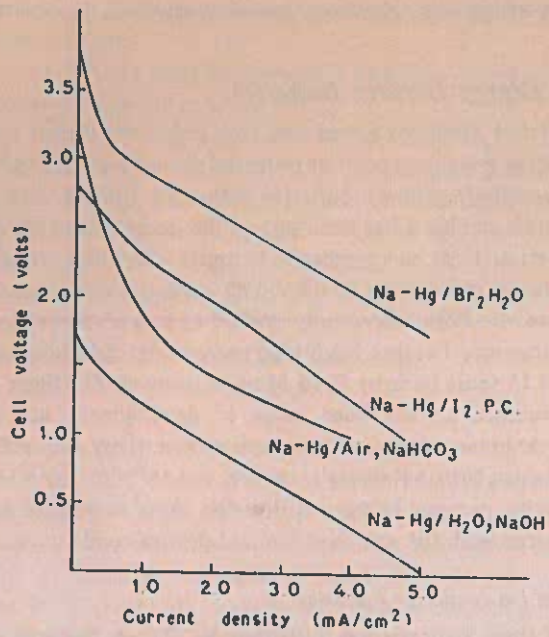


Fig. 10. Current-voltage curves for sodium-mercury amalgam primary cells with various cathodes at 25 °C (after Will and Mitoff[188])

(d) *Fluoride ion conductor batteries*

Very few fluoride ion batteries have been studied. Thin film batteries have been reported by Kennedy[44], Kennedy and Hunter[46], and LaRoy[55]. These batteries have the following configurations :

$$Pb/\beta\text{-}PbF_2/PbF_2,\ CuF_2/Cu$$

and

$$Bi/LaF_3/Au$$

The cell characteristics are given earlier in Table 5. The cathode utilisation is only 30-40% and very low current drain is possible. The cell performance was reported to be poor because a layer of low conducting α-PbF_2 material got formed at the anode during discharge giving an anodic passivation[46]. Schoonman[98] has reported a cell :

$$Pb/\beta\text{-}PbF_2\text{-}AgF/BiO_{0.09}F_{2.82}/Bi$$

with better performance. However, anode passivation is observed in this cell also.

4.2 *High Energy Density Batteries*

The batteries discussed above can have only very special applications, because of either low decomposition potential or very high internal resistance. They are essentially primary batteries with very limited rechargeability. Recently, much emphasis has been put on the development of high energy density batteries. High energy density batteries which are rechargeable are useful as a storage medium for load levelling in electric utility services. These batteries have also been successfully applied as power sources for vehicular (car or rail) traction. The first Na/S high energy density battery was demonstrated about 13 years back by Ford Motor Company[185]. Since then these cells have reached an advanced stage of development at Ford Motor Company, Argonne National Laboratory, Electricity Council Research Center at Chester, British Railway at the UK, etc.[26,30,38,41,45,49,64,66,188]. In the following paragraphs we will describe some important high energy density batteries with Li^+ ion, Na^+ ion and proton conductors.

(a) *Lithium ion conductor batteries*

Most of these batteries use Li/S couple. These batteries have high theoretical energy density (*see* Fig. 3). The first Li/S cell had *L/L/L* configuration employing lithium melt as anode, LiCl-KCl eutectic melt as electrolyte and sulphur melt as cathode[51]. The idea behind using the molten constituents, is to get very high current due to rapid diffusion of the active electrochemical species and also to have an intimate electrode/electrolyte contact. But several problems like sealing, maintaining the molten state, etc., were encountered. Further, liquid lithium attacks the electrolyte and insulating seal materials, liquid sulphur dissolves into the electrolyte and attacks lithium electrode to form Li_2S which leads to the self discharge of the battery. These shortcomings led to the development of solid lithium-aluminium alloy anode and solid iron sulphide (FeS or FeS_2) cathode. The disadvantage of these substitutions are the decrease of the cell voltage, specific energy and power. However, cells with long life could be produced. Several designs of the cell configuration :

Li-Al/LiCl-KCl/FeS or FeS_2

were developed at Argonne National Laboratory[22] and General Motors[11].

For detailed battery discharge characteristics reference may be made to the original papers (*see* also Table 5).

Ohzuku *et al.*[73] have recently reported a battery of S/L/S configuration with graphite compound as cathode :

$Li/LiClO_4$-propylene carbonate $/C_{96}ClO_4$

which has high coulombic efficiency (80-100%) and high positive potential (4·75 V *vs.* Li/Li^+). It was observed that no deterioration occurred up to 250 cycles of charge-discharge. It is, however, reported that the cathode compounds collapse under overcharging conditions.

Liang *et al.*[62] have reported three *S/S/S* battery configurations with an operating temperature of 300 °C :

Li-Si/LiI $(Al_2O_3)/TaS_2$/Ta,

Li-Si/LiI $(Al_2O_3)/TiS_2$/Ti

and

Li-Si/LiI $(Al_2O_3)/TiS_2$, Sb_2S_3, Bi

The battery structure is simplified because of the absence of any liquid phase (solution or fused-salt electrolytes and/or molten electrodes). This also increases package efficiency with respect to the active components. Furthermore, no material compatibility problem is encountered in the system. As a result, these solid-state battery systems have excellent charge retention characteristics and a long projected operating life. A typical charge-discharge test results for the three systems are shown in the Fig. 11. It was reported that Li-Si/TaS_2 and the Li-Si/TiS_2, Sb_2S_3, Bi systems operate at higher efficiencies than the Li-Si/TiS_2 system.

(b) *Sodium ion conductor batteries*

Several systems have been developed so far with Na/S couple using Na-β-alumina as electrolyte. The operating temperature is $\sim$300 °C. The theoretical energy density is $\sim$760 Wh/kg. Basically, the battery consists of molten sodium anode and molten sulphur cathode separated by a β-alumina electrolyte. Several constructions of these cells are possible. Two possible geometries (i.e., Tubular and Flat Plate) are shown in Fig. 12.[103,109]. A tubular cell can be used in the case of high energy capacity whereas a flat cell is advantageous in high power applications. The advantages and disadvantages of both designs have been discussed[48,108,109,110]. The cathode material is a graphite felt impregnated with sulphur since sulphur itself is a non-conductor.

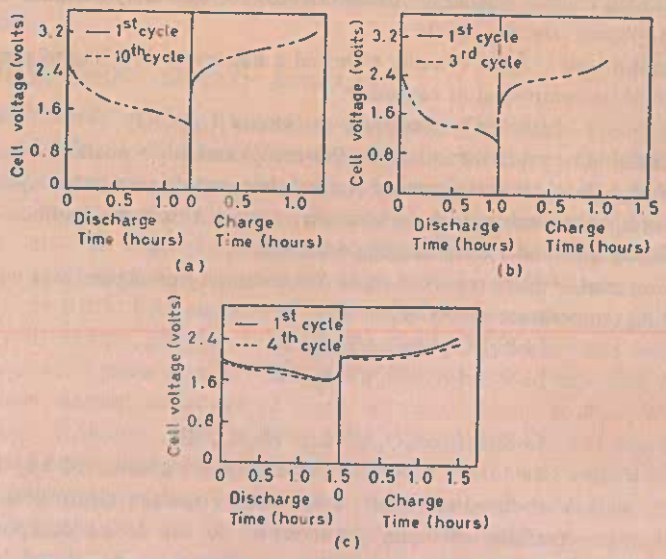


Fig. 11. Discharge-charge cycle tests at temperature 300±10°C and current 0.1A

(a) Li-Si/LiI (Al_2O_3)/TaS_2/Ta, discharge capacity, 125 mAh
(b) Li-Si/LiI (Al_2O_3)/TiS_2/Ti, discharge capacity 100 mAh
(c) Li-Si/LiI (Al_2O_3)/TiS_2, Sb_2S_3Bi, discharge capacity 150 mAh
(After Liang *et al.*[80])

In the *L/S/L* battery configuration of Na/S cell, the cell reaction is (*see* Table 5 also) :

$$2Na + x.S \rightarrow Na_2S_x$$

For x=5 (and some excess sulphur) the cell is considered fully charged. As it discharges, more Na enters and the value of x decreases. Discharge can be carried to as far as Na_2S_3 (actually a mixture of Na_2S_4 and Na_2S_2)[73]. Beyond this point the polysulphide melting point starts to rise rapidly as can be seen from the Na-S phase diagram given in Fig. 13[36]. Operation of the Na/S battery is usually maintained at ~ 300-350°C and further discharge leads to freezing of the cathode material. As a result, the OCV is constant (~2·08 V) in the first stage, but as soon as the sodium content in the cathodic compartment increases and the composition Na_2S_3 is formed, Na_2S_3 compound begins to solidify and the OCV of the cell gradually

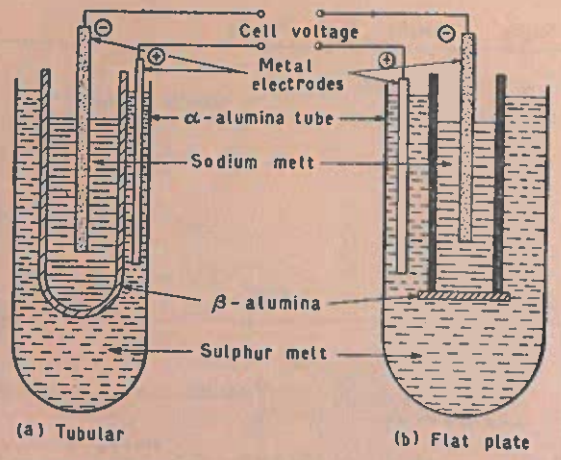


Fig. 12. Two possible designs for β-alumina electrolytes (a) Tubular, (b) Flat plate (after Sudworth *et al.*[110])

decreases to a value ~1·76 V. This can be easily visualised from the Fig. 14. During recharging, the sodium ions move back to Na-anode compartment through the ceramic electrolyte. Successful charge-discharge cycles up to as much as 2000 times have been demonstrated[5]. The current status of Na/S advanced battery development programmes are excellently reviewed by Fischer *et al.*[30] and Asher[5]. Asher[5] has described a new technique for developing Na-β-alumina ceramic cylinders following electrophoretic forming method and their successful application for small (16 Ah) and full (168 Ah) size cells.

Several problems are encountered in the manufacture of these batteries as given below :

(1) *Sealing* : Since both the cathode and anode materials are in the molten state, high temperature leak proof seals capable of withstanding thermal shocks have to be provided[109].
(2) *Corrosion*
(3) *Electrolyte Failure* : It has been found that during the charging cycles cracks develop in the electrolyte due to sodium dentritic growth[43, 88, 129].

Although the major efforts in battery development using β-alumina has been the Na/S battery, some research has been directed towards other

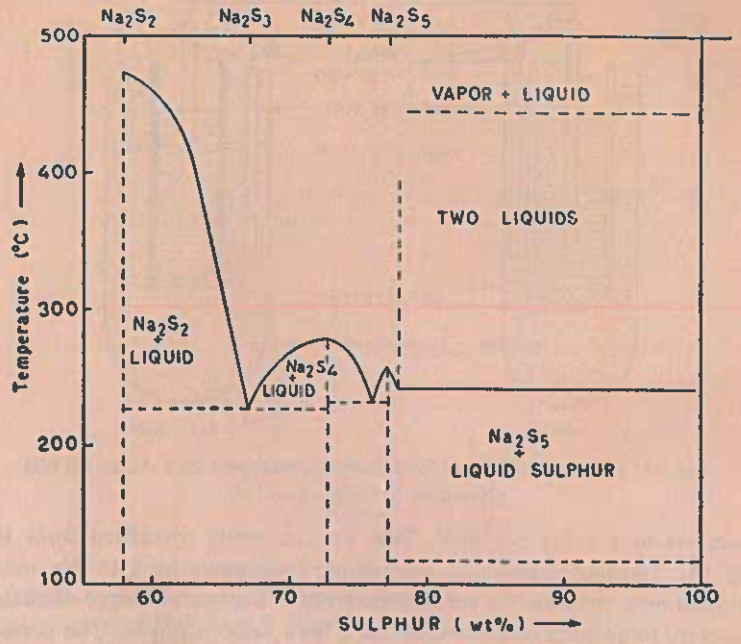


Fig 13. Phase diagram of sodium-sulphur system (after Oei)[72]

systems also. Many other cathode materials apart from sulphur or polysulphides have been used. Some of these couples are listed in Table 5.

(c) *Proton conductor batteries*

Recently, some solids have been reported to allow substantial H^+ ion conduction (*see* Table 1). This has given high hopes for development of solid-state batteries based on hydrogen transport. Recently, Childs *et al.*[21] have reported a successful proton solid-state battery of the following configuration :

$$\underset{(1)}{\text{Metal hydride}} \,/\, \underset{(HUP)}{\text{Electrolyte } (HUO_2PO_4)} \,/\, \underset{(2)}{\text{Metal hydride}}$$

The metal hydrides with different hydrogen activities serve as the cathode and anode of the battery. If the high vapour pressure hydride such as hydrides of Pb, $LaNi_5$ or TiFe and the low vapour pressure hydrides such as alkali earth hydride CaH_2 are used as different electrodes, the OCV

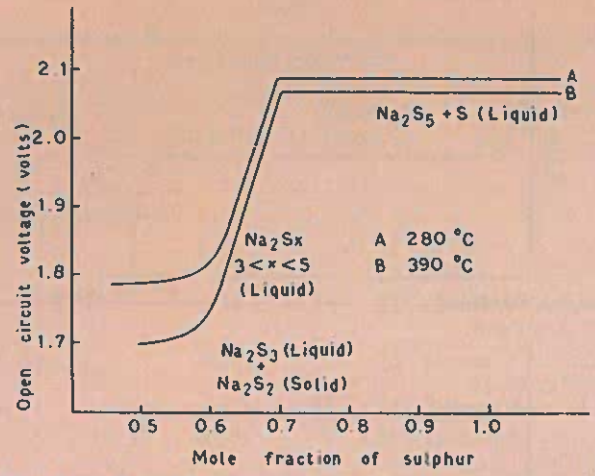


Fig. 14. Open-circuit voltage of a sodium-sulphur battery *vs.* composition of the sodium polysulphide at different temperatures (after Weiner[186])

of the cell would be approximately ~ 0·75–0·8 V. Childs *et al.*[21] have reported that the couple CaH_2/HUP/$TiFeH_2$ is likely to give a very good performance. The theoretical energy density is about 294 Wh/kg.

The discharge characteristics of the above hydrogen cell will depend upon the phase relationships of the hydrides. When the metal and the hydride phases are separated by a significant two-phase region, the voltage of the cell will be constant whilst the two two-phase regions coexist during discharge. A typical case reported by Childs *et al.*[21] is illustrated in Fig. 15 (a), where non-stoichiometric region existing at both extremities of the phase diagram results in the curves at the start and finish of discharge, which are reflected in the corresponding voltage variation shown in Fig. 15 (b).

The study on HUP solid-state battery opened a prospect for several solid electrolyte electrochemical devices using proton conductors. Thin membranes of HUP can also be developed and it is possible to have light weight and low cost high energy density batteries.

5. Concluding Remarks

From the above description of various low and high energy density batteries, it seems that Li^+ ion conductor batteries offer maximum promise

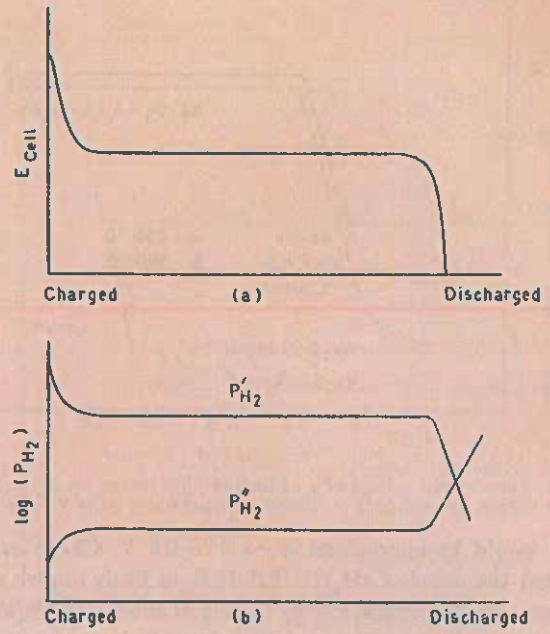


Fig. 15. Changes in the cell voltage (a) and in hydrogen pressure (b) during discharge (after Childs *et al.*[21])

for long life application. Already Li^+ ion batteries in various designs are commercially available and are being used in heart pacemakers. However, good Li^+ ion conductors with high conductivity at ambient temperatures are yet to be discovered. For high energy density applications Na/S system is extremely promising and gives a wide range of applicability particularly in vehicular tractions and storage for load levelling. However, more research and development is needed to bring down its temperature of operation which is normally $\sim$ 300 °C. Though not much work has been done on proton conductor, but it seems that a very good alternative, for different energy demands, lies in proton conducting batteries. Childs *et al.*[21] first attempt in developing such a proton conductor battery should attract more attention for further extensive research.

Table 1

Ionic conductivity of some important solid electrolyte materials for solid-state battery applications

Solid Electrolyte	σ ($ohm^{-1}cm^{-1}$)	*Temperature* °C	*References*
1	2	3	4
(A) *Silver Ion Conductor*			
α-AgI	1·34	>147	Tubandt & Lorenz[181]
Ag_3SI	10^{-2}	25	Reuter & Hardel[87]
$RbAg_4I_5$	0·27	22	Owens & Argue[77], Bradley & Greene[9]
$(CH_3)_4N_2Ag_{13}I_{15}$	0·04	22	Owens[74] & Shahi & Chandra[100]
$Ag_7I_4PO_4$	0·019	25	Takahashi *et al.*[114]
$Ag_6I_4WO_4$	0·047	25	Takahashi *et al.*[115], Shahi & Chandra[107]
$RbAg_4I_4CN$	0·18	25	Mellors & Luzos[67]
Ag-β-alumina	0·006	25	Kennedy[45]
(B) *Copper Ion Conductor*			
(N, N'-dimethyl triethylene diamine dibromide)$_{0·06}$ $(CuBr)_{0·94}$	0 035	20	Takahashi & Yamamoto[123]
$(C_5H_5NH)_2Cu_2Br_7$	0·017	25	Chan *et al.*[12]
$Rb_4Cu_{16}Cl_{13}I_7$	0·35	25	Takahashi[113]
$RbCu_4Cl_3I_2$	0·47	25	Geller *et al.*[33]
(C) *Lithium Ion Conductor*			
LiI (pure)	$5·5\times10^{-7}$	30	Schlaikjer & Liang[92]
LiI (+40% Al_2O_3)	10^{-5}	25	Liang[58]
Li_3N (Li-rich)	10^{-3}	25	Alpen *et al.*[2]
Li-β-alumina	5×10^{-3}	25	Farrington & Roth[27]
$Li_{14}Zn(GeO_4)_4$	$<10^{-5}$	25	Hong[39]
LiX, H_2O, Al_2O_3 ($X=I^-, Br^-, Cl^-, F^-$)	10^{-5}-10^{-2}	25	Owens & Hanson, US Patent 4,007,122
(D) *Sodium Ion Conductors*			
Na-β-alumina	$1·4\times10^{-2}$	25	Kennedy[45]
$NaSbO_3$.1/6NaF	$5·6\times10^{-3}$	300	Goodenhough *et al.*[34]

Table 1 (Concld.)

1	2	3	4
Na-β-Gallate (β-alumina analog)	3×10^{-3}	25	Chandrashekhar & Foster[20]
Na-β''-Gallete	$6{\cdot}3\times10^{-2}$	25	Chandrashekhar & Foster[20]
$Na_3Zr_2Si_2PO_{12}$	2×10^{-4}	25	Goodenhough *et al.*[34], Hooper[41]
$NaYSi_4O_{12}$	5×10^{-4}	25	Shannon *et al.*[102], Hong *et al.*[40]
(E) *Potassium Ion Conductor*			
K-β-alumina	$6{\cdot}5\times10^{-5}$	25	Kennedy[45]
$KMgAlF_2$	5×10^{-5}	300	Goodenhough *et al.*[34]
$K_2O.(6-x)\,Fe_2O_3.0{\cdot}8ZnO$	$1{\cdot}8\times10^{-2}$	300	Takahashi *et al.*[125]
(F) *Anion Conductors*			
(a) *Fluoride Ion Conductor*			
β-PbF_2	$1{\cdot}5\times10^{-5}$	25	Bonne & Schoonman[8]
$Pb_{1-x}Bi_xF_{2+x}$ (x=0·25)	6×10^{-5}	25	Lucat *et al.*[63]
$Pb_{1-x}Th_xF_{2+2x}$ (x=0·125)	2×10^{-4}	25	Trichet & Rouxel[130]
$RbBiF_4$	2×10^{-4}	25	Lucat *et al.*[63]
β-$PbSnF_4$	10^{-3}	25	Rean *et al.*[83]
(b) *Oxygen Ion Conductor*			
ZrO_2 : 12% CaO	$5{\cdot}5\times10^{-2}$	1000	Etsell & Flengas[35]
ThO_2 : 8% Y_2O_3	$4{\cdot}8\times10^{-3}$		
CeO_2 : 7% SrO	$1{\cdot}1\times10^{-1}$		
(G) *Proton Conductor*			
$HUO_2PO_4.4H_2O$ (Hydrogen Urinyl Phosphate)	4×10^{-3}	20	Shilton & Howe[105], Howe & Shilton[42], Childs *et al.*[21]
H^+-exchanged montmorillonite claysmono & bi-layer hydrates	$1{\cdot}4\times10^{-4}$	20	Sheffield & Howe[104]
$H_3PMo_{12}O_{40}.30H_2O$	0·02	20	Nakamura *et al.*[70]
$H_3UO_2(IO_3)_2.4H_2O$	0·007	20	Shilton & Howe[106]
$Sb_2O_5.4H_2O$	3×10^{-4}	20	England *et al.*[24]
H_3O^+-β''-alumina	5×10^{-3}	20	Ferrington & Briant[28]

Table 2

Requirements for an ideal solid electrolyte (after Owens & Skarstad[89])

Property	*Value*
Ionic conductivity	High ($\sim$1 $ohm^{-1}\,cm^{-1}$)
Electronic conductivity	Low ($<10^{-8}$ $ohm^{-1}cm^{-1}$)
Temperature dependance of conductivity	Low
Mobile ion equivalent weight	Low
Decomposition potential	High (3-4 volts)
Thermodynamic stability	Stable with respect to dissociation type reactions
Temperature range of stability	Wide
Specific gravity	Low
Ease of fabrication	Conducive to dense electrolyte elements
Mechanical strength	Withstand stresses of thermal electrical cycling

Table 3

Electrochemical potentials of iodine with different metal ions (after Owens & Skarstad[89])

Metal ions	*Electrochemical potential (volts)*
Li^+	2·80
Na^+	2·95
K^+	3·35
Rb^+	3·37
Ag^+	0·69
Cu^+	0·72

Table 4

Solid-state battery systems investigated till early 1960's (after Liang[59])

System	*Open-circuit Voltage (OCV) (volts)*
$Pb/PbCl_2/AgCl/Ag$	0·46
$Ag/AgI/Ag_2S/S$	0·2
$Ag/AgI/I_2$	0·7
$Ag/AgBr/CuBr_2$	0·74
$Ag/AgI/V_2O_5$	0·46
$Ag/AgCl/Ca(OCl)_2$	0·9-1·0
$Ag/AgCl/KICl_4$	1·04
$Ag/AgBr\text{-}Te/CuBr_2$	0·8

Table 5

General features of some important low and high energy density batteries (after Chandra—*to be published*)

Common Anode	*Common Electrolyte*	*Common Cathode*	*Total cell reaction (Typical)*	*Approximate open-circuit voltage (volts)*	*Order of current density (per cm²)*	*Remarks and References*
1	2	3	4	5	6	7
(A) *Low-Energy-Density Batteries*						
(a) *Silver ion conducting electrolyte*						
Silver or Ag-Hg amalgum or mixture of silver, carbon and the electrolyte	AgI-based solid electrolyte (Ag_3SI, MAg_4I_5, $Ag_7I_4PO_4$, etc.)	(a) I_2+C	$Ag+\frac{1}{2}I_2 \rightarrow AgI$	0·68	mA	Room temperature operation, moderate current drain, low voltage, high cost, extremely limited rechargeability[6, 26, 41, 55, 59, 75, 79, 81, 111, 124]
		(b) Ag_2Te-Te+C	$2Ag+Te \rightarrow Ag_2Te$	0·22		
		(c) Ag_2Se-Se+C	$2Ag+Se \rightarrow Ag_2Se$	0·27		
		(d) I_2-Ph+C	$Ag+\frac{1}{2}I_2$-Ph $\rightarrow AgI+\frac{1}{2}$ Ph	0·62		
		(e) Rb_2TeI_6+C	$4Ag+Rb_2TeI_6 \rightarrow Te+4AgI+2RbI$	0·53		
(b) *Copper ion conducting electrolyte*						
Copper	(a) TED2CH$_3$ Br-CuBr	(i) $Cu_{1.8}S$+C	–	0·34	50 μA	Low cost, operating temperature 25°C, high polarisation of copper electrode, low exchange current, limited rechargeability, poor battery for wide-range applicability[3, 56, 91, 124]
		(ii) TiS_2 C	–	0·47	–	
	(b) PHBr-CuBr		–	–	–	
	(c) HMTCH$_3$ Br-CuBr	X+CuX+C (X=S, Se or Te)	$2Cu+X \rightarrow Cu_2X$	0·45 (for S)		
				0·37 (for Se)	50 μA	
				0·26 (for Te)		

Table 5 (Contd.)

1	2	3	4	5	6	7
(c) *Alkali ion conducting electrolyte*						
Lithium	LiI (pure or Al_2O_3 or oliovalent doped)	(a) I_2, C	$Li+\frac{1}{2}I_2 \rightarrow LiI$	3·0		Long life room temperature operation, very low current density batteries[29, 57, 59, 60, 82, 94, 95, 96, 97]
		(b) AgI, Ag	$Li+AgI \rightarrow LiI+Ag$	2·11		
		(c) CuI, Cu	$Li+CuI \rightarrow LiI+Cu$	2·08	10 μA	
		(d) PbI_2, Pb	$2Li+PbI_2 \rightarrow 2LiI+Pb$	1·9		
		(e) PbS, Pb	$2Li+PbS \rightarrow Li_2S+Pb$	–		
		(f) P2VP.NnI_2	$2Li+P2VP.nI_2 \rightarrow P2VP(n-1)I_2+2LiI$	2·8		
	$4LiI.NH_4I$	$NH_4Ag_4I_5$	$4Li+NH_4Ag_4I_5 \rightarrow NH_4I+4Ag+4LiI$	2·1		
Na-Hg (amalgum)	Na-β-alumina	(a*) O_2, H_2O or Br_2 or I_2 $NaHCO_3$, NaOH	$2Na+Br_2 \rightarrow 2NaBr$	3·8 (for Br_2)		Long life, low energy drain battery with room temperature operation[138]
			$2Na+I_2 \rightarrow 2NaI$		mA	
		(b*) I_2 or Br_2+ propylene carbonate	–	3		
(d) *Fluoride ion conducting electrolyte*						
Pb	PbF_2	PbF_2+CuF_2/Cu	$Pb+CuF_2 \rightarrow PbF_2+Cu$	0·7	10 μA	Low temperature (25°C) operation, poor cell performance due to PbF_2 layer giving anodic passivation. Cell is sensitive to oxygen pressure[44, 46, 53, 96]
Pb	β-PbF_2-AgF	$BiO._{09}F_{2.82}$/Bi	$3Pb+2BiF_3 \rightarrow 3PbF_2+2Bi$	0·33	6·3 μA	
Bi	LaF_3	Au	–	0·5	10^{-9}A	

Table 5 (Contd.)

1	2	3	4	5	6	7
(B) *High-Energy-Density Batteries*						
(a) *Lithium ion conducting electrolyte*						
Li*	LiCl-KCl*	Sulphur*	$2Li+x.S \rightarrow Li_2S_x$	–	amp.	High temperature (>300 °C). The use of liquid electrode posed many problems like fabrication, etc.[51]
Li-Al or Li-Si	LiCl-KCl*	FeS or FeS_2	$2Li+FeS \rightarrow Li_2S+Fe$ $4Li+FeS_2 \rightarrow 2Li_2S+Fe$	2·06	150 mA	High temperature operation (300 °C), suitable for vehicular traction, rechargeable[52, 97, 133]
Li	$LiClO_4$-propylene carbonate*	$C_{96}ClO_4$	$Li+C_{96}ClO_4 \rightleftharpoons LiClO_4+C_{96}$	4·75	—	High energy, rechargeable. But compound collepses when overcharged[73]
Lithium*	Li-β-alumina	Molten Sulphur impregnated graphite or polysulphide*	$2Li+x.S \rightarrow Li_2S_x$	2·3	amp.	High temperature (≥300 °C) operation suitable for vehicular traction rechargeable[58, 59, 61, 62]
Li-Si* (Intermetallic)	Li-β-alumina	(a) TaS_2/Ta (b) TiS_2/Ti (c) TiS_2, Sb_2S_3, Bi	– – –	2·4	15 mA	

Table 5 (Concld.)

1	2	3	4	5	6	7
(b) *Sodium ion conducting electrolyte*						
Sodium*	Na-β-alumina	(a) Sulphur impregnated graphite (or polysulphides C_8CrO_3)*	$2Na+x.S \rightarrow Na_2S_x$	2·0	amp.	Operating temperature (≥300 °C), long life and rechargeable
		(b) $NaAlCl_4+Cl_2+C$*	$2Na+Cl_2 \rightarrow 2NaCl$ (in $NaAlCl_4$)	3·5	50 mA	Moderate temperature operation (200 °C)
		(c) Fe_2O_3 doped-β-alumina $Na_2O.11(Fe_xAl_{2-x}O_3)$ x=0·9–1·2	$2Na+Fe_2O_3 \rightarrow Na_2O+2FeO$	2·1	50 mA	Operation temperature (120 °C). Non-stoichiometric (SSE) are used
Sodium*	MgO-β-alumina	$SbCl_3$ ($NaAlCl_4$)	$2Na+Cl_2 \rightarrow 2NaCl$	2·9	–	Operating temperature (~200 °C)[7, 23, 41, 47, 50, 66, 103, 108, 109, 134, 137]
Al	$AlCl_3$-NaCl-KCl*	Sulphur*	–	1·2	–	Operating temperature (200-300 °C), OCV diminishes fast to 0·6 V due to dedriatic growth[84]
(c) *Proton conducting electrolyte*						
CaH_2	$HUO_2PO_4.4H_2O$	$TiFeH_2$	$M'H_x+M' \rightarrow M'+M'H_x$	0·8	100 mA	Room temperature operation rechargeable[21]

* Starred electrolytes or electrodes are in liquid phase, rest are in solid phase

Table 6

Significant properties of $Ag/RbAg_4I_5/RbI_3$ solid-state battery at 25°C (after Owens[75])

OCV	0·66 volt
Theoretical specific energy	48 Wh/kg
Practical specific energy (at 25 $\mu A/cm^2$ discharge current)	4·8 Wh/kg
Energy density (at 25 $\mu A/cm^2$)	0·015 Wh/cm^3
Shelf life (predicted)	10 years
Shelf life ascertained with $(CH_3)_4NI_5$ chelate	5 years

Table 7

The open-circuit cell voltage, ionic transference number (t_{ion}) and electronic transference number (t_e) for best electrocodeposited films of various superionic solids (after Agrawal[1])

Solid electrolyte film	OCV (E_{meas}) (Volts)	$t_{ion}\left(=\frac{E_{meas}}{E_{calc}}\right)$	t_e $(=1-t_{ion})$	Reference
$Ag_7I_4PO_4$	0·675	0·99	0·01	Chandra *et al.*[16]
$[(CH_3)_4N]_2Ag_{13}I_{15}$	0·67	0·99	0·01	Chandra and Agrawal[14]
$NH_4Ag_4I_5$	0·644	0·95	0·05	Chandra *et al.*[15]
KAg_4I_5	0·61	0·90	0·10	Chandra and Mohabey[17]
$RbAg_4I_5$	0·66	0·97	0·03	Mohabey[68]

Table 8

Cell characteristics of some Na-β-alumina room temperature batteries (after Will and Mitoff[138])

Cell couple	OCV (Volts)	Voltage at 0·1 mA/cm^2 (Volts)	Energy density Wh/kg (Theoretical)	Energy density Wh/kg (Practical)
$Na-Hg/Br_2, H_2O$	3·77	3·5	1026	330–440
$Na-Hg/Br_2$, P.C.*	3·34	3·1	910	330–440
$Na-Hg/I_2, H_2O$	3·25	3·0	579	220–330
$Na-Hg/I_2$, P.C.*	2·82	2·8	502	170–260
$Na-Hg/O_2, NaHCO_3$	2·90	2·8	1553	660
$Na-Hg/H_2O, NaOH$	1·85	1·7	991	440

*P.C.—Propylene Carbonate

REFERENCES

1. Agrawal, R. C. (1979) Thesis, Ravishankar University, Raipur, India.
2. Alpen, U. V., Rabeneu, A. and Talat, G. H. (1977) *App. Phys. Lett.*, 30 : 621.
3. Armstrong, R. D., Dickinson, T. and Taylor, K. (1974a) *J. Electroanal. Chem.*, 57 : 157.
4. Armstrong, R. D., Sherwood, P. M. A. and Wiggins, R. A. (1974b) *Spectrochim. Acta.* 30A : 1213.
5. Asher, J. A. (1979) in "Fast Ion Transport in Solids", (Eds.) P. Vashistha, J. N. Mundy, and G. K. Shenoy, (North Holland), p. 39.
6. Beyeler, H. U. and Briiesch, P. (1977) in "1976 Crystal Growth and Materials", (Eds.) E. Kaidis, and H. J. Scheel, (North Holland).
7. Birk, J. R. (1976) in "Superionic Conductors", (Eds.) G. D. Mahan, and W. L. Roth, (Plenum Press) p. 1.
8. Bonne, R. W. and Schoonman, J. (1977) *J. Electrochem. Soc.*, 124 : 28.
9. Bradley, J. N. and Greene, P. D. (1967) Trans. Farad. Soc., 63 : 424.
10. Butherus, A. D. (1973) 144th Meet. Electrochem. Soc., Abst. No. 10.
11. Cairns, E. J. (1976) in "Critical Materials Problems in Energy Production", (Ed.) C. Stein, (Academic Press, N. Y.), p. 684.
12. Chan, L. Y. Y. and Geller, S. (1978) *J. Solid State, Chem.*, 28 : 85.
13. Chandra, S., "Superionic Solids : Principles and Applications"—*to be published*, (Submitted to North Holland, Amsterdam).
14. Chandra, S. and Agrawal, R. C. : Submitted to J. Phys. D.
15. Chandra, S., Agrawal, R.C. and Pandey, R.K. (1978) *Nat. Acad. Sc. Lett.*, 1 : 112.
16. Chandra, S., Agrawal, R. C. and Pandey, R. K. (1980b) *Phys. Stat. Sol.*, (a)57 : 30.
17. Chandra, S. and Mohabey, V. K. (1976) *Proc. Nue. Phys. Sol. St. Phys. Symp.*, 19 C : 134.
18. Chandra, S. and Mohabey, V. K. (1979) *Phys. Stat. Sol.*, (a) 53 : 63.
19. Chandra, S., Sharma, J. N., Mohabey, V. K. and Agrawal, R. C. (1980a) *J. Phys.*, D. 13 : 493.
20. Chandrashekhar, G. V. and Foster, L. M. (1977) *J. Electrochem. Soc.*, 124 : 329.
21. Childs, P. E., Howe, A. T. and Shilton, M. G. (1978) *J. Power Sources*, 3 : 105.
22. Chilenskas, A. A., Bernstein, G. J., Craae, J. E. A., Hornstra, F., Kolba, V. M., Slawecki, M. A. and Ivins, R. O. (1976) Int. Electric Vehicle Symp. 4th, Dusseldorf.
23. Dell, R. M. and Evans, L. S. (1976) in "Electrode Processes in solid-state Ionics", (Eds.) M. Kleitz and J. Dupuy, (D. Reidel Pub. Co., Holland), p. 411.
24. England, W.A., Cross, M. G., Hamnett, A., Wiseman, P. J. and Goodenough, J. B. (1979) Paper presented at Faraday Div. Electrochem. Group Meet., Oxford, 28-30 March.
25. Etsell, T. H. and Flengas, S. N. (1970) *Chem. Rev.*, 70 : 339.
26. Ervin III, G. (1972) U. S. Patent 3, 701, 685.
27. Farrington, G. C. and Roth, W. L. (1977) *Electrochem. Acta.*, 22 : 767.
28. Farrington, G. C. and Briant, J. L. (1978) *Mat. Res. Bull.*, 13 : 763.
29. Fester, K., Doty, R. L. and Velde, R. U. (1976) 16th A. Am. Soc. Mech. Eng. Symp. Energy. Alternatives, Albuquerque.
30. Fischer, W., Haar, W., Hartmann, B. and Weddingen, G. (1978) J. Pow. Sources, 3 : 299.

31. Foley, R. T. (1969) *J. Electrochem. Soc.*, **116** : 13c.
32. Geller, S. (Ed.) (1977) "Solid Electrolyte", Topics in App. Phys. V. 21, (Springer Verlag, Berlin).
33. Geller, S., Akridge, J. R. and Wilber, S. A. (1979) *Phys. Rev.*, B **19** : May 15.
34. Goodenhough, J. B., Hong, H. U. P. and Kafalas, J. A. (1976) *Mat. Res. Bull.*, **11** : 203.
35. Gross, S., (1976) *Energy Conv.*, **15** : 95.
36. Gupta, N. K. and Tischer, R. P. (1976) *J. Electrochem. Soc.*, **119** : 1033.
37. Haber, F. and Tollozzko, St. (1904) *Z. Anorg. Chem.*, **41** : 407.
38. Hagenmuller, P. and Van Gool, W. (Eds.) (1978) "Solid Electrolytes", Material Science Series (Academic Press).
39. Hong, H. (1978) *Mat. Res. Bull.*, **13** : 117.
40. Hong, H. Y. P., Kafalas, J. A. and Bayard, M. (1978) *Mat. Res. Bull.*, **13** : 757.
41. Hooper, A. (1978) *Comptemp. Phys.*, **19** : 147.
42. Howe, A. T. and Shilton, M. G. (Part I) (1979) *J. Sol. Stat. Chem.*, **28** : (*in press*).
43. Huggins, R. A. (1974) Report of ARPA Mater. Res. Council Meeting.
44. Kennedy, J. J. (1977a) *Thin Solid Films*, **43** : 41.
45. Kennedy, J. J. (1977b) in "Solid Electrolyte", (Ed.) S. Geller, (Springe-Verlag, N.Y.), p. 105.
46. Kennedy, J. H. and Hunter, J. (1976) *J. Electrochem. Soc.*, **123** : 10.
47. Kennedy, J. H. and Sammells, A. F. (1974) *J. Electrochem. Soc.*, **121** : 1.
48. Kinkkola, K. and Wagner, C. (1957) *J. Electrochem. Soc.*, **104** : 308 and 379.
49. Klietz, N. and Dupuy, J. (Eds.) (1976) "Electrode Process in Solid-State Ionics", (D. Reidel Pub. Co., Holland).
50. Kummer, J. T. and Weber, N. (1967) *Proc. Pow. Sou. Conf.*, **21** : 37.
51. Kyle, M. L., Cairns, E. J. and Webster, D. S. (1973) Li/S Batteries for Off-peak Energy Storage, ANL Rep., 7958, March.
52. Lai, S. C. (1976) *J. Electrochem. Soc.*, **123** : 1196.
53. LaRoy, B. C., Lilly, A. C. and Tiller, C. O. (1973) *J. Electrochem. Soc.*, **120** : 1668.
54. Lazzari, M., Razzini, G. and Scrosati, B. (1976) *J. Pow. Soc.*, **1** : 57.
55. Lazzari, J. and Scrosati, B. (1976/77) *J. Pow. Soc.*, **1** : 333.
56. Lazzari, M., Scrosati, B. and Vincent, C. A. (1977) *Electrochemica. Acta.*, **22** : 51.
57. Liang, C. C. (1971) Proc. Intersoc. Energy Conv. Eng. Conf., p. 38.
58. Liang, C. C. (1973) *J. Electrochem. Soc.*, **120** : 1289.
59. Liang, C. C. (1974) *Appl. Sol. Stat. Sc.*, **4** : 95.
60. Liang, C. C. and Barnette, L. H. (1976) *J. Electrochem. Soc.*, **123** : 453.
61. Liang, C. C., Joshi, A. V. and Barnette, L. H. (1976) 27th Am. Proc. Pow. Soc. Conf., p. 141.
62. Liang, C. C., Joshi, A. V. and Hamilton, N. H. (1978) *J. Appl. Electrochem.*, **8** : 445.
63. Lucat, C., Sorbe, P. Portier, P., Reau, J. M., Hagenmuller, P. and Grannee, J. (1977) *Mat. Res. Bull.*, **12** : 145.
64. Mahan, G. D. and Roth, W. L. (Eds.) (1976) "Superionic Conductors", (Plenum Press).
65. Markin, T. L., Bones, R. J. and Dell, R. M. (1976) in "Superionic Conductors", (Eds.) G. D. Mahan, and W. L. Roth, (Plenum Press), p. 15.

66. May, G. J. (1978) *J. Pow. Sou.*, **3** : 1.
67. Mellors, G. W. and Luzos, D. V. (1971) *J. Electrochem. Soc.*, **118** : 846.
68. Mohabey, V. K. (1978) Thesis, Ravishankar University, Raipur, India.
69. Mrgudich, J. N. (1964) in "Encyclopedia of Electrochemistry", (Reinhold, N. Y.).
70. Nakamura, O-, Kodama, T., Ogino, I. and Miyake, Y. (1976) Japan Pat. 76/106, 683.
71. Nowick, A. S. and Park, D. S. (1976) in "Superionic-Conductors", (Eds.) G. D. Mahan and W. L., Roth. (Plenum Press), p. 395.
72. Oei, D. G. (1975) *Inorg. Chem.*, **12** : 435.
73. Ohzuku, T., Takehara, Z. and Yoshizwa, S. (1978) Denki Kagaku, **46** : 438.
74. Owens, B. B. (1970) *J. Electrochem. Soc.*, **117** : 1536.
75. Owens, B. B. (1971) "Solid Electrolyte Batteries", *In* Adv. in Electrochem. and Electrochem. Engg., (Ed.) C. W., Topias, 8, (J. Wiley and Sons, N. Y.), p. 1.
76. Owens, B. B. (1980) in "Application of Solid Electrolytes", (Eds.) T. Takahashi, and A. Kozawa, The Electrochemical Soc. of Japan.
77. Owens, B. B. and Argue, G. R. (1967) *Science*, **157** : 308.
78. Owens, B. B. and Hanson, H. J. : U. S. Patent 4, 007, 122.
79. Owens, B. B., Oxley, J. E. and Sammells, A. F. (1977) in "Solid Electrolyte", *Topics of App. Phys.*, **21**, (Ed.) S. Geller, (Springer-Verlag, Berlin).
80. Owens, B. B. and Skarstad, P. M. (1979) in "Fast Ion Transport in Solids", (Eds.) P. Vashistha, J. N. Mundy and G. K. Shenoy, (North Holland).
81. Pampallona, M., Ricci, A., Scrosati, B. and Vincant, C. A. (1976) *J. App. Electroanal. Chem.*, **29** : 301.
82. Rao, M. L. B. (1969) U. S. Patent No. 3, 455, 742.
83. Rean, J. M., Portier, J., Lenassenn A., Villeneuve, G. and Pouchard, M. (1978) *Mat. Res. Bull.*, **13** : 1415.
84. Re'dey, L., Pornbszky, I. and Molnar, I. (1975) in "Power Sources", (Ed.) D. H. Collins (Oriel Press, England), p. 559.
85. Reuter, B. and Hardel, K. (1961) *Naturwiss*, **48** : 161.
86. Reuter, B. and Hardel, K. (1965) *Z. Anorg. Allg. Chem.*, **340** : 168.
87. Reuter, B. and Hardel, K. (1966) *Ber. Bunsenges Physik Chem.*, **70** : 82.
88. Richmen, R. H. and Tennenhouse, G. J. (1975) *J. Amer. Ceram. Soc.*, **58** : 63.
89. Rohr, F. J. (1978) in "Solid Electrolytes" (Eds.) P. Hagenmuller, and W. Van Gool, (Academic Press, N. Y.), p. 431.
90. Salamon, M. B. (Ed.) (1979) "Physics of Superionic Conductors", *Topics in current Phys.*, **15** (Springer-Verlag, N. Y.).
91. Sammells, A. F., Gougoulas, J. Z. and Owens, B. B. (1975) *J. Electrochem. Soc.*, **112** : 1291.
92. Schlaikjer, C. R. and Liang, C. C. (1973) in "Fast Ion Transport in Solids", (Ed.) W. Van Gool, (North Holland), p. 685.
93. Schneider, A. A. (1976) Int. Symp. Solid. Ionic and Ionic-Electronic Conductors, Rome, Sept.
94. Schneider, A. A., Greatbach, W. and Mead, R. (1975) in "Power Sources 5", (Ed.) D. H. Collins, (Academic Press, London), p. 651.
95. Schneider, A. A. and Moser, J. R. (1972) U. S. Patent No. 3, 674, 562.
96. Schneider, A. A., Moser, J., Webb, T. and Desmond, J. (1970) Proc. 24th Power Sources 5 Atlantic City, May.
97. Scholtens, B. B. and Van Gool, W. (1978) in "Solid Electrolytes", (Eds.) P. Hagenmuller and Van W. Gool, (Academic Press, N. Y.). p. 451 and 463.

98. Schoonman, J. (1976) *J. Electrochem. Soc.*, **123** : 1772.
99. Scrosati, B. (1976) Proc. SAEST, Madras, India.
100. Shahi, K. and Chandra, S. (1975a) *Phys. Stat. Sol.*, (a) **28** : 653.
101. Shahi, K. and Chandra, S. (1975b) *J. Phys. C. (Sol. Stat. Phys.)* **8** : 2255.
102. Shannon, R. D., Chen, H. Y. and Berting, T. (1977) *Mat. Res. Bull.*, **12** : 969.
103. Silverman, H. P., Development Program for Solid Batteries, Interim Rep. EPRI Res. Project, 127-2, 1 July 1974-30 June '75.
104. Sheffield, S. H. and Howe, A. T. (1979) *Mat. Res. Bull.*, **14** : 929.
105. Shilton, M. G. and Howe, A. T. (1977) *Mat-Res. Bull.*, **12** : 701.
106. Shilton, M. G. and Howe, A. T. (1979) in "Fast Ion Transport in Solids" (Eds.) P. Vashistha, J. N. Mundy, and G. K. Shenoy, (North Holland), 727.
107. Skarstad, P. M. and Parker, H. S. (1978) *J. Cryst. Growth.*, xxx.
108. Steele, B. C. H. (1976) in "Superionic Conductors", (Eds.) G. D. Mahan and W. L. Roth, (Plenum Press, N. Y.,) 47.
109. Sudworth, J. L. (1975) IEEE 10th Conf., 616.
110. Sudworth, J. L., Hames, M. D., Storey, M. A., Azim, M. F. and Tilley, A. R. (1973) in "Power Sources' (Ed.) D. H. Collins, 4 (Oriel Press, England), 1.
111. Takahashi, T., (1973) *J. Appl. Electrochem.*, **3** : 79.
112. Takahashi, T. (1976) in "Superionic Conductors" (Eds.) G. D. Mahan, and W. L. Roth, (Plenum Press) 379.
113. Takahashi, T. (1979) in "Fast Ion Transport in Solids" (Eds.) P. Vashistha, J. N. Mundy, and G. K. Shenoy, (North Holland), 521.
114. Takahashi, T., Ikeda, S. and Yamamoto, O. (1972) *J. Electrochem. Soc.*, **119** : 477.
115. Takahashi, T., Ikeda, S. and Yamamoto, O. (1973) *Ibid.*, **120** : 647.
116. Takahashi, T., Kuwabara, K. and Kase, Y. (1959a) *Denki Kagaku*, 43 : 273.
117. Takahashi, T., Kuwabara, K. and Kase, Y. (1975b) Nippon Kagaku No. 8, 1305.
118. Takahashi, T., Walabayashi, N. and Yamamoto, O. (1977a) *J. App. Electrochem.*, **7** : 253.
119. Takahashi, T. and Yamamoto, O. (1964) *Denki Kagaku*, **32** : 610.
120. Takahashi, T. and Yamamoto, O. (1965) *Ibid.*, **33** : 346.
121. Takahashi, T. and Yamamoto, O. (1966a) *Electrochemica Acta.*, **11** : 779.
122. Takahashi, T. and Yamamoto, O. (1966b) *Ibid.*, **11** : 911.
123. Takahashi, T. and Yamamoto, O. (1975) *J. Electrochem. Soc.*, **122** : 83.
124. Takahashi, T. and Yamamoto, O. (1977) *J. App. Electrochem.*, **7** : 37.
125. Takahashi, T., Yamamoto, O., Matsuyama, F. and Noda, Y. (1976) *J. Solid. Stat. Chem.*, **16** : 35.
126. Takahashi, T., Yamamoto, O., and Sawai, A. (1978) *J. App. Electrochem*; **8** : 161.
127. Takahashi, T., Yamamoto, O., and Takahashi, H. (1977 b) *J. Solid Stat. Chem.*, **21** : 37.
128. Takahashi, T., Yamamoto, O., Yamadan, S. and Hayashi, S. (1979) *J. Electrochem. Soc* , **126** : 1654.
129. Tennenhouse, G. J., Ku, R. C., Richman, R. H. and Whalen, T. J. (1975) *Am. Ceram. Soc. Bull.*, **54** : 523.
130. Trichet, L. and Rouxel, J. (1977) *Mat. Res. Bull.*, **12** : 345.
131. Tubandt, C. and Lorenz, E. (1914) *Z. Phys. Chem.*, **87** : 513
132. Van. Gool, W. (Ed) (1973) "Fast Ion Transport in Solids", (North Holland, Amsterdam).
133. Vissers, D. R., Tomczuk, Z. and Stennenberg,, R. K. (1974) *J. Electrochem. Soc.*, **121** : 665.
134. Weber, N. (1976) in "Superionic Conductors", (Eds.) G. D. Mahan, and W. L. Roth, (Plenum Press), p. 37.
135. Weber , N. and Kummer, J. T. (1967) Adv. Energy Conv. Eng. ASME Cont. Florida, p. 913.
136. Weiner, S.A. (1975) in "Research on Electrodes and Electrolyte for the Ford Sodium Sulphur Cell", Ford Motor Co. Annu. Rep., July, Contract NSFC, 805.
137. Werth, J.J. (1974) U.S. Patent No. 3, 874, 667.
138. Will, F.G. and Mitoff, S. P. (1975) *J. Electrochem. Soc* , **122** : 457.
139. Yao, Y. F. Y. and Kummer, J. T. (1967) *J. Inorg. Nue. Chem.*, **29** : 2453.

Radioisotopes in Medicine

R. S. Mani

Head, Radiopharmaceuticals Section
Bhabha Atomic Research Centre
Bombay

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

Introduction

Among the manifold beneficial applications of atomic energy in agriculture, industry, biology, medicine, hydrology and geology, the use of radioisotopes and radiation in medicine and health-care has recorded the most spectacular growth in the last three decades. Radioisotopes are now routinely used for diagnosis and management of several diseases, and also in their treatment. The operation of several nuclear reactors and accelerators in various parts of the world has made available large quantities of useful radioisotopes which are commercially processed and supplied as ready-to-use radiopharmaceutical formulations thus expanding their role in clinical medicine and opening new vistas for their applications. Radioisotopes are now used to investigate and manage a host of vexed clinical problems ranging from dose adjustment of potent drugs and hormones, differentiating medical from surgical jaundice or an abscess from hepatoma, detecting early metastases or tuberculosis of the bone to diagnosing impending rejection in a kidney transplant, measuring the size of myocardial infarct and also in treatment of cancer.

Production of Medical Radioisotopes[1,2]

The radioisotopes required for medical applications have to be of high purity and are produced by bombardment of suitable targets with neutrons in a nuclear reactor or with charged particles in an accelerator. Over 150 different radioisotopes and compounds incorporating them are now used in medical diagnosis and treatment. The half-lives of these radioisotopes range from a few hours to a few years. Generally isotopes with short half-life are used for internal administration to avoid excessive radiation exposure and the recent trend is to use radioisotopes with half-lives of a few minutes (for example N 13 with 10 minutes half-life) to a few hours (for example Tc 99m with a half-life of 6 hours) while isotopes with long half-lives (greater than 1 year) are used as external radiation sources.

The production and purification of radioisotopes is an intricate technology and requires specially equipped laboratories and remote controlled processing cells to avoid radiation exposures to the operator. Irradiation of targets in the reactor or cyclotron has to be carefully planned and various parameters have to be carefully controlled to obtain the highest purity and specific activity. Contamination of one radioisotope with another

product due to impurities in the target element undergoing nuclear reactions and undesirable side reactions taking place concurrently is often a problem which can be avoided by using electromagnetically enriched targets or targets of high purity. The period of irradiation and the energy of the projectile can also be selectively controlled. Special provisions have to be made for reactor irradiation of gaseous targets like Xenon 124 used for production of Iodine 125, a widely used tracer for hormones and other biologically important compounds.

The projectiles from an accelerator have a short range compared to neutrons and in view of their high heat dissipation, the targets have to be provided with cooling arrangements. The most common reactions in the reactor used for isotope production are the (n, γ) reaction (also known as radioactive neutron capture), (n, γ) reaction followed by beta decay and (n, p) reactions. The possible reactions in the accelerator are more complex depending on the type of particles and their energy. For production of isotopes, protons, deuterons and alpha-particles in the energy range 10–75 MeV are commonly used as projectiles.

Some of the important reactor produced radioisotopes of medical interest are iodine 131, iodine 125, chromium 51, molybdenum 99, tin 113, and phosphorus 32. The important applications of accelerator produced radioisotopes are of recent origin and the most widely used are gallium 67, iodine 123, indium 111, and thallium 201. Table 1 gives a listing of radioisotopes both reactor and cyclotron produced, their production methods and half-lives. It may be mentioned that the reactor is a more economical source of radioisotopes than the accelerator because of the large number of available target irradiation positions in a reactor and the possibility of producing many isotopes simultaneously. The bulk of the radioisotopes both in numbers and quantity used in medical applications is produced in the reactor while accelerator produced isotopes are used in specialized applications.

Important Applications in Medicine

The important medical uses of radioisotopes can be broadly classified into five categories :

1. Applications of radiopharmaceuticals for *in vivo* diagnosis and therapy

2. Radioimmunoassays and other *in vitro* assays
3. Therapy with sealed radiation sources
4. Applications for medicare and public health
   (a) Radiation sterilization of medical products
   (b) Radiation hygienisation of sewage
5. Radioisotope applications in medical research.

1. *Radiopharmaceuticals in Diagnosis and therapy*[3-6]

Radiopharmaceuticals are preparations incorporating radioisotopes intended for use as pharmaceuticals by oral or parenteral administration in humans. They have to be non-toxic and acceptable like conventional pharmaceutical preparations. Many radioisotopes are employed in their simple pharmaceutical forms and are known as primary medical isotopes. They include sodium iodide I 131, sodium orthophosphate P 32, sodium fluoride F 18, ferric citrate Fe 59 and sodium chromate Cr 51. More sophisticated radiopharmaceuticals are prepared by incorporating the radioisotopes into complex organic and inorganic compounds with different bio-distribution and pharmacodynamics. The development of these unique types of compounds has really made the radioisotope a versatile tool in clinical medicine. In Table II some important labelled radiopharmaceuticals and methods for their preparation are listed.

During the last two decades a great deal of ingenious effort has been put in for synthesis and purification of a variety of labelled radiopharmaceuticals. An evident method of producing many labelled compounds is the synthesis of the compound starting from a suitable radioactive raw material or intermediate. For example Hg 203 chlor merodrin (3 chloromercury 2 methoxy propylurea), an agent used for kidney and brain imaging is synthesised by addition of Hg 203 as mercury acetate across the double bond in ally urea followed by saturation with sodium chloride for replacement of the acetoxy group by chloride. Where the molecule to be synthesised is more complex and is of biological origin, bio-synthesis using micro-organisms can be used with advantage. Vitamin B_{12} labelled with Co 57 or Co 58 is thus synthesised by fermentation of radiocobalt with nutrients using the micro-organism streptomyces olivaceous. Sometimes it is possible to exchange a non-radioactive atom in a complex molecule for its radioactive

counterpart under suitable experimental conditions. These exchange reactions are of particular value in labelling a variety of iodo-organic compounds with radioiodine. Rose Bengal (tetraiodo tetrachloro fluorescein) and Hippuran (ortho iodo hippuric acid) the two commonly used radiopharmaceuticals for liver and kidney function evaluation respectively are labelled with iodine 131 by heating the compounds in a buffered medium with sodium iodide I 131. Sometimes the molecules to be labelled do not contain an atom which can be replaced by its radioactive counterpart. This situation is encountered in proteins and fatty acids whose constituent atoms are oxygen, carbon, nitrogen, sulphur and hydrogen which do not have gamma-ray emitting isotopes with a reasonably long half-life which is essential for external monitoring of these compounds after administration. In such cases, a foreign atom not originally present in the molecule is introduced into it in trace quantities such that the biological systems are not able to sequester it. Such compounds are known as 'foreign labelled compounds'. Many proteins like human serum albumin, fibrinogen and insulin have been trace-iodinated with radioiodine (I 131 or I 125) and these labelled tracers have been found suitable for study of their metabolism. In producing foreign labelled compounds, the quantity of foreign atom introduced should be very small and hence the process requires carrier free or high specific activity radioisotopes.

In Vivo Applications [7-12]

The earlier applications of radioisotopes were in the study of thyroid physiology, endocrinology and metabolism and haematological studies, since these applications involved the use of radioisotopes in their simple chemical forms and involved simple instrumentation like a well type scintillation counter. The thyroid gland traps iodine for synthesising two important hormones, triiodothyronine and thyroxine which regulate the body's basic metabolic rate. Radioactive Iodine Uptake Test (RAIU) using iodine 131 helps to differentiate the clinical states of hyperthyroidism, euthyroidism and hypothyroidism. The normal gland shows a steady increase in uptake of radioactivity during the 24 hour, reaches a plateau for the 24 to 36 hour period, after which the radioactivity decreases as synthesised hormones are released. In hyperthyroidism there is excessive thyroid hormone production and hence enhanced and sometimes more rapid radioiodine uptake. In hypothyroidism, there is insufficient production of thyroid hormones and the 24 hour radioiodine uptake is below normal. RAIU for assessing thyroid

function is still widely used in our country and offers several advantages over other biochemical tests.

Radioactive tracer studies of blood components and their precursors and metabolites as well as important biological products have made enormous contribution to our knowledge of haematologic physiology and pathology. However, the techniques employed in clinical management of patients are much more limited and can be divided into (a) measurement of blood volumes, (b) cell survival studies, (c) iron turnover studies, and (d) Vitamin B_{12} absorption studies. Blood volume is measured by injecting radioiodine labelled human serum albumin or Cr 31 labelled red blood cells. After a suitable period of mixing, an aliquot is withdrawn from another site. The blood volume is obtained by dividing the total activity of the tracer injected by the concentration of tracer in the aliquot. Blood volume determinations are helpful in indications of shock, burns, congestive, heart failure, guidance of postoperative fluid replacement and differential diagnosis of polycythemia.

Red blood survial studies are carried out with cells labelled with Chromium 51 or diisopropyl fluorophosphate–P32. It is helpful in assessment of hemolytic anemias resulting in the shortening of RBC life span and spleenic sequestration. Ferro kinetic studies are carried out usually with iron 59. The normal plasma iron clearance is exponential with a normal half-life of 90 ± 30 minutes. Variation in normal clearance rates are observed in cases of hemolytic anemia, polycythemia rubra vera and aplastic anemia. Megaloblastic anemia resulting from Vitamin B_{12} deficiency can be evaluated with Co 57 or Co 58 labelled Vitamin B_{12} to diagnose the cause of malabsorption.

Scintigraphy [15-16]

Another important area where radiopharmaceuticals have made a significant contribution to medical diagnosis is radioisotope scintigraphy or organ imaging. The procedure is in many ways similar to X-ray roentgenography, but here the radioactivity is administered to the patient and the radiopharmaceutical which concentrates in the organ of interest serves as the source of radiation. The radiopharmaceutical is biorouted to the organ of interest making use of organ's physiology and the radioactivity distribution is mapped out using a rectilinear scanner or gamma camera. In the rectilinear scanner, the radiation detector, usually a sodium iodide crystal, moves

over the area of interest in a zig-zag manner and quantitatively records the radioactivity. The gamma camera uses a giant sodium iodide crystal detector which can cover the entire organ in a single view and simultaneously records the activity distribution. One important factor which has contributed to the widespread application of radioisotope scintigraphy is the availability of a wide spectrum of radiopharmaceuticals with specificities for various organs (*see* "Radioisotope Generators" below). Using them, now it is possible to get scintigraphic pictures of the brain, the liver, kidney, lungs, skeleton, spleen, thyroid, salivary glands, stomach and pancreas. The scintigrams provide functional information of the organs and are extremely useful in detecting space-occupying lesions like tumours, abscesses, cysts, infarcts and embolism.

As an example, the liver is imaged by intravenous injection of a radioactive colloid such as Au 198 gold colloid which is phagocytised by the recticuloendothelial cells of the liver or a radioactive dye like I 131 Rose Bengal which is excreted by the polygonal cells of the liver. The liver scintigram provides information regarding the size, shape, position of low-lying normal liver, helps to locate lesions such as amaebic abscess, primary and secondary tumours, diffuse liver diseases like hepatitis and cirrhosis, as well as in differential diagnosis of various types of jaundice. The thyroid scan performed after oral administration of iodine 131 is indicated to evaluate a simple palpable nodule which may be hyperfunctioning or non-functioning or functioning normally. Most pathological processes in the brain tissue lead to a breakdown of blood brain barrier and allows diffusion of the radiopharmaceutical into the brain tissue which concentrates in the defective tissue. Brain infarcts, hemorrhage, aneurysm, primary and secondary neoplasms of the brain, trauma, brain abscess, meningitis, and hydrocephalus are some focal diseases of the brain where brain scans are indicated. Radioisotopic imaging is very useful in the study of Cerebro Spinal Fluid (CSF) dynamics particularly in the diagnosis of communicating hydrocephalus. The procedure known as "cisternography" is performed using I 131 human serum albumin or In 111 DTPA complex. It may be mentioned that I 131 human serum albumin meant for cisternography should be highly purified and free of pyrogen contamination to avoid aseptic meningitis. It was in this connection that the rapid and highly sensitive *in vitro* pyrogen test known as "limulus test" was developed. The limulus test is reported to be a thousand times more sensitive than the

conventional rabbit pyrogen test and can detect even minute amounts of pyrogen contamination.

Lung imaging is carried out by either inhaling a radioactive gas Xenon 133 or radioactive aerosols or by injecting intravenously radioisotopically tagged particles of size greater than 15 um which are trapped in the lung capillaries. The major indication for a lung scan is pulmonary embolism, particularly in a patient with a clear chest radiography. The lung scan is also useful in cases of emphysema, asthma, pulmonary fibrosis, pleural effusion and infarction. Kidney scan which demonstrates functional morphology is most frequently indicated where other means of visualising kidneys fail and is used to evaluate kidney size, location, presence of filling defects, tumours, abscesses, cysts, infarctions and identify the functional region of horse-shoe kidney and other anomalies. The spleen can be visualised with heat denatured Cr 51 red blood cells which are trapped by the spleen. Since lesions in spleen are relatively rare, the important indications for spleen scan are estimation of spleen size and the discovery of spleen ruptures from blunt trauma to the abdomen. Other methods to assess spleen size such as physical examination and X-ray are unreliable. Radionuclide bone imaging for detection of cancer metastases is rapidly becoming one of the most important and most widely applied clinical procedures in nuclear medicine. The reasons are two fold ; firstly, the high incidence and great importance of metastatic bone disease and secondly, the fact that radiographic procedures are much less sensitive in detecting them.

Dynamic Studies[17,18]

The above procedures for organ imaging are mostly based on radiopharmaceuticals which are specifically trapped in them. In another set of tests, the passage of the radiopharmaceutical through an organ during a short interval is followed and these are known as dynamic function tests. The I 131 Hippuran renogram for kidney function tests is one of the well-standardised dynamic testing procedures. The renogram is performed after a bolus injection of about 25 u Ci I 131 hippuran intravenously. Two sodium iodide detectors with straight bore collimators are placed over the kidneys and are connected to a strip chart recorder, which records the activity over the kidneys. The renogram indicates the degree of severity of kidney diseases and is most valuable in comparing function either between the two kidneys or in one kidney over a period of time. It is the most sensitive

test available for demonstrating minimal residual function in a badly diseased kidney. It is frequently employed in conjunction with intravenous pyelogram in the work up of hypertension patients in whom unilateral renal artery stenosis is suspected. The renogram is also indicated in patients who are iodide sensitive, in patients with multiple myeloma or pheochromocytomas in whom dehydration, part of the preparation for intravenous pyelography may be dangerous. The renogram also gives lower radiation dose to the fetus as well as mother in cases of pregnancy where kidney tests must be performed. The safety, lack of discomfort to patient, ease of performance and immediate availability of results have made renogram especially valuable in following function of transplanted kidney. The test is now frequently done with a camera with both kidneys in the field of view.

The most recent advance in nuclear medicine imaging is the coupling of a gamma camera to the computer which can store, reproduce and make computations of the data. It can make many calculations of great interest to the cardiologist like cardiac output, ejection fraction, pulmonary blood volume and mean pulmonary transit time. The ECG gated cardiac blood pool imaging using gamma camera and computer provides a convenient tool to assess global and regional cardiac function and can be used to check routinely the benefits of surgical or drug treatment. Myocardial infarct imaging with technetium 99 m pyrophosphate or thallium 201 is now widely used for assessing injury to the myocardium. It is also very useful in diagnosis of ischaemic heart disease when ECG stress test or history is inconclusive.

Radioisotope Generators[19,20]

The introduction of faster imaging devices like gamma camera in nuclear medicine requires that the radioactivity administered should be of the order of several millicuries to get a high photon flux and a good image. To achieve this a variety of short-lived radioisotopes with half-lives of only a few hours have been introduced for organ imaging. To overcome the economic limitations and transporation difficulties of supplying these radioisotopes from the site of production to the medical centres, devices known as "radioisotope generators" have been developed. These generators are separation systems containing a relatively long-lived parent isotope from which a short-lived daughter can be extracted at periodic intervals. These are also known as "radioisotopic cows" or milking units in popular terminology, since the most

common type of generator system is based on column chromatography in which the long-lived parent isotope is firmly absorbed on an absorbent like alumina and the short-lived isotope is obtained by simply passing a small volume of a suitable eluent. Other chemical separation procedures are also used where limitations of specific activity preclude the use of column chromatography. For example, the technetium 99 m generator used in India is based on solvent extraction using methyl ethyl ketone and uses indigenously produced low specific activity Mo 99. Some important radioisotopes used widely in nuclear medicine obtained from radioisotope generators are detailed in Table III.

Technetium 99 m is at present the most widely used short-lived isotope in nuclear medicine. The 140 keV energy of its gamma radiation is ideally suited for use with gamma cameras since it can be easily collimated and efficiently detected. The use of technetium 99 m for organ imaging offers several advantages over conventional radiopharmaceuticals. Its short half-life and absence of beta radiation results in very low patient radiation dose. Hence it can be used even on pregnant women and children who are normally excluded from radiopharmaceutical administration. It is possible to inject multimillicurie quantities to get high quality scan pictures, and also carry out repeat studies without undue radiation dose to the patient. Because of the versatile chemistry of technetium 99 m, it can be incorporated into a variety of complex compounds which can be used for imaging almost all important body organs.

The use of technetium 99 m and its radiopharmaceuticals accounts for more than 75% of all organ imaging procedures in advanced countries. The use of Tc 99 m generator is so widespread that commercial manufacturers excel in offering improved designs with the maximum ease and safety of operation. The presently available generator systems are automated and sterile systems and Tc 99 m in the form of pertechnetate is eluted by connecting an evacuated collection vial or a pressurised eluent container. Tc 99 m generators of a different type based on solvent extraction using methyl ethyl ketone are in use in India. Even though the operation of this generator is somewhat more complicated than the column type generator, it offers several advantages. It can use low specific activity Mo 99 since the solubility of molybdate is much higher than the capacity of absorbents like alumina to adsorb molybdenum. The final product can be obtained in a very small volume suitable for "bolus" injection and is free of aluminium, the most

common chemical impurity from alumina generator. Most of the radionuclidic impurities found in Tc 99 m obtained from chromatographic type generators are not extractable into methyl ethyl ketone. It may be mentioned here that India is the only country having a network of Tc 99 m generators operating on the solvent extraction process thus enabling a self-reliant Tc 99 m applications programme.

Technetium 99 m obtained from the generator is in the form of pertechnetate and as such can be used for brain and thyroid imaging. By tagging Tc 99 m, to a variety of inorganic, organic and biomolecules it is possible to route it to other important organs. For example, Tc 99 m complexed to methylene diphosphonate on intravenous injection concentrates in the skeleton after two hours and can be used for bone imaging. The introduction of Tc 99 m labelled bone seeking radiopharmaceuticals was singularly responsible to make bone scanning a valuable nuclear medicine procedure. In Table IV, a list of Tc 99 m radiopharmaceuticals and their applications are given. In view of the short half-life of Tc 99 m, these compounds have to be necessarily prepared in the hospital prior to administration and the tests for sterility and apyrogenicity can be carried out only on a *post facto* basis after the products are administered. Hence great care has to be exercised in the operation of the generator and in the preparation of the various radiopharmaceuticals at the hospitals from the generator milk; these practices form the core of a new discipline called "Hospital Radiopharmacy". It is the duty of the qualified hospital radiopharmacist to ensure the pharmaceutical and radiological safety of these activities. Ready-to-use sterile kits are now available from commercial sources for various Tc 99 m radiopharmaceuticals. The kits contain the active ingredients in a freeze dried form and are pretested for sterility, apyrogenicity, toxicity and purity. Labelling with Tc 99 m is achieved by mere addition of sterile Tc 99 m pertechnetate and shaking to dissolve the contents and in some cases followed by heating. Use of these kits ensures a high degree of pharmaceutical safety and reproducibility and is of particular help in small hospitals with limited nuclear medicine activities.

Indium 113 m, another short-lived isotope (half-life 1·7 hours) is usually used as a standby for Tc 99 m, particularly, in remote areas. The generator for In 113 m has a longer useful life in view of the long half-life of the parent Sn 113. This system consists of a hydrous zirconium oxide column on which Sn 113 is firmly adsorbed as the chloro complex and In 113 m is eluted with dilute acid. In 133 m obtained as indium chloride from the

generator is a good blood pool agent. It can also be incorporated into colloids, microspheres, DTPA, phytate, phosphonates and a variety of other compounds with localisation in liver, lung, brain, bone, etc. The future may find increasing use of ultra short-lived radioisotopes such as Ir 191m ($t_{\frac{1}{2}}$=4·9 sec), Kr 81m ($t_{\frac{1}{2}}$=13 sec) and Rb 82 ($t_{\frac{1}{2}}$=80 sec) particularly for serial gamma camera dynamic imaging of the cardiovascular system. Some of these procedures may complement or even replace radiographic procedures using contrast media.

Radioisotopes from Accelerators[21]

In recent years, accelerator produced short-lived radioisotopes are increasingly used in organ imaging and with this trend, the cost of these isotopes which is now quite high is expected to come down. They offer certain specific advantages such as short half-life and absence of beta radiation which reduces the patient dose. Most of them are neutron deficient isotopes which decay by emitting two annihilation radiations of 0·5 MeV. Using these two radiations in a coincidence circuit, one can detect smaller tumours such as for example brain tumours, than is possible with other radioisotopes. Cyclotron produced isotopes are carrier free and radiopharmaceuticals with very high specific activities can be produced with them.

Iodine 123, a cyclotron produced radioisotope of short half-life (13 hr) is fast replacing I 131 for thyroid imaging and study of liver and kidney functions. Ga 67 has been found to be very useful for localisation of soft tissues tumours. In 111 DTPA Complex has proved to be an ideal agent for cisternography. Thallium 201 gives an earlier picture of myocardial infarction than other agents.

In addition, the accelerator is the only source of the medically useful radioisotopes of physiological elements namely carbon, nitrogen and oxygen which form 82% by weight of the human body. Carbon 11 ($t_{\frac{1}{2}}$=20·4 mt), Oxygen ($t_{\frac{1}{2}}$=2·05 mt) and Nitrogen 13 ($t_{\frac{1}{2}}$=10 mt) and their labelled compounds hold great potential for clinical use.

Therapy Using Radiopharmaceuticals[22]

It is the goal of radiotherapy to destroy by delivering large amounts of radiation to selected abnormal areas and tissues of the body (usually malignancies) with negligible radiation to neighbouring vital normal tissues.

Table V lists radiopharmaceuticals utilised therapeutically by localisation in target tissues either metabolically or physically by a specific route of administration. Iodine 131 because of its unique avidity for the thyroid gland and strong beta emission with an energy of 606 keV is used widely for treatment of hyperthyroidism, thyrotoxicosis and thyroid carcinoma. It is also used for therapeutic induction of hypothyroidism in patients with cardiac disease and as an adjunct to surgery in the management of thyroid cancer. Phosphorus 32, a pure beta emitting isotope, with a maximum energy of 1·7 MeV, is used in the treatment of haematologic diseases, such as polycythemia vera, leukemia and thrombocytosis. The treatment is based on the fact that radiophosphorus enters the miscible body phosphate pool and is concentrated by rapidly proliferating tissues particularly blood cell precursore in the bone marrow which divide and proliferate rapidly even in healthy persons, but do so more rapidly under the above conditions. Effusions in the pleural and peritoneal cavities due to malignancies of the pleura or peritoneum or tumours in adjacent tissue can be palliatively treated by the intracavitary administration of colloids of hard beta emittes such as colloidal gold Au 198 and chromic phosphate P 32.

However, no radiopharmaceutical known so far with the exception of radioiodine is concentrated and retained in tumours with sufficient selectivity to be therapeutically useful. To our knowledge, the metabolic characteristics of any given tumour are not sufficiently different from those of its mother tissue to allow for adequate target to non-target concentration ratios. Therefore, the hopes for wider therapeutic applications of internally administered radionuclides is rather slim unless a radical change occurs in our understanding of tumour metabolism.

2. *Radioimmunoassays*[23–25]

A very important advance in recent years in the application of radiotracers in medical diagnosis has been the development of a wide spectrum of *in vitro* assay techniques known as radioimmunoassays. In these procedures, there is no need to administer the radioisotope to the patient and consequently even the small radiation exposure involved in most organ imaging and other *in vivo* procedures is eliminated. The impact of the radioimmunoassays on clinical chemistry and diagnosis and basic medical research has been so phenomenal that the originator of this procedure,

Dr Rosalyn Yalow was awarded the Nobel Prize in medicine for 1978 for this discovery.

In radioimmunoassay the extreme detection sensitivity of the radiotracer and the highly specific reaction of an "antigen" with its "antibody" are made use of to attain a high degree of detection sensitivity and specificity in the estimation. To give an example, estimation of the levels of insulin present in the blood serum under various conditions, including fasting and after glucose stimulation is of considerable importance in the management of some forms of diabetes and detection of certain tumours of the pancreas called insulinomas. Radioimmunoassays offer an easy and reliable means of carrying out such an estimation. For this essay, a small measured volume of the patient's blood is withdrawn and the serum is separated out by centrifugation. To this serum are added known quantities of anti-insulin serum and isotopically tagged insulin (tagged with radioiodine I 125). The anti-insulin serum is prepared by immunising guinea pigs with increasing doses of porcine insulin, injected over a period of 4-6 months. The guinea pigs are then bled and the serum rich in insulin antibodies is separated out. The avidity or potency of the anti-insulin can be determined by direct titration with isotopically labelled porcine insulin. Porcine insulin can be tagged or labelled with radioiodine by a simple chemical process involving controlled iodination of the tyrosyl groups of insulin.

For the actual assay, the mixture containing the patient's serum, anti-insulin serum and I 125 tagged insulin prepared as indicated above is incubated for a few hours. The insulin antibody molecules specifically bind the insulin molecules to form a complex and the insulin present in the patient's serum as well as the isotopically labelled insulin molecules added extraneously react with the antibody molecules. Since the total quantity of antibody molecules added is not sufficient to complex all the insulin molecules present in the mixture, some fraction of the insulin remains unbound. The unbound insulin molecules can be readily separated from the antibody-complexed-insulin by a simple procedure such as direct absorption of the former on a specially prepared dextran coated charcoal. The radioactivity associated with the bound and free or unbound insulin fractions is measured accurately and their ratio (denoted as B/F, i.e., Bound to Free) is determined. B/F ratios are also determined, under identical conditions for standard sera samples containing known quantities of pure insulin, and a standard curve relating the observed B/F values with the levels of added standard insulin is constructed.

From this curve, the content of insulin in the patient's serum can be read off knowing the value of B/F for the patient's serum which can be readily determined experimentally.

Radioimmunoassay procedures designed for the estimation of several other important physiologically active ingredients including hormones, steroids, digitalis glycosides, vitamins and exogenous antigens such as Australia Antigen or hepatitis associated antigen are also based on the same principle. In Table VI, substances for which RIAs have been developed are listed. In most cases, the substance to be assayed is "antigenic" so that its antibodies can be produced by injecting graded doses of the substance to animals as described earlier. In the case of steroids, digitalis glycosides, and a few other drugs which are not "antigenic", conjugation with proteins such as bovine serum albumin has to be carried out as the first step ; these conjugates are then injected into animals for production of specific antibodies. Isotopic labelling of a pure sample of the substances to be estimated can be carried out in most instances by selective incorporation of radioactive isotopes of iodine (I 125 or I 131) or by substitution of tritium (H3) for hydrogen.

Application of RIA Procedures

The great potentialities of RIA procedures in the field of medicine can be readily appreciated from the following typical applications :

HPL Estimation

Human Placental Lactogen—HPL—is a hormone which is present in the blood serum of a pregnant woman. The concentration of this hormone rapidly rises during the early weeks of pregnancy if the pregnancy is normal and if no complications are likely to arise at a later stage. In case of complications which may lead to foetal distress or neonatal asphyxia or abortion, the concentration of HPL will be significantly lower than in the case of pregnancy which is normal. It may be mentioned that the majority of such complications do not manifest any other symptoms of abnormality, and radioimmunoassay is the only method now available for the estimation of HPL in the early stages of pregnancy. If such estimations indicate low HPL concentrations, it can be taken as an indication of complications at a later stage of pregnancy. Closer obstetrical management would be required for all such patients and in cases where there is a positive indication of abortion at a later stage, early medical termination of pregnancy would be

warranted. HPL determinations are also of considerable diagnostic value in pregnancies complicated by diabetes or erythroblastosis.

It is now widely recognised that HPL radioimmunoassay would be very valuable screening procedure to detect risk pregnancies and enables better medical management of such cases.

Australia Antigen

A very large number of patients who receive blood transfusions develop serum hepatitis as a result of contamination of the donors blood with Australia Antigen, and the mortality among such patients is high. Even an extremely minute concentration of the order of millionth of a gram of Australia Antigen in the donor's blood can produce serum hepatitis. Conventional methods of detection of Australia Antigen in blood (based on agar gel diffusion, hemagglutination inhibition and counter electrophoresis) suffer from the drawback that the detection limits are too high. Radioimmunoassay offers an easy and reliable method for estimating Australia Antigen at concentration levels which are below the limits of detection of other procedures and can thus serve as a screening procedure in blood collection and transfusion.

RIA of Drugs

The extension of radioimmunoassay procedures to cover certain important drugs like digoxin, digitoxin and phenytoin in recent years has been a significant development in the field of radioisotopic diagnosis. Digoxin and digitoxin (commonly known as digitalis) are very useful but toxic drugs employed for treatment of heart patients. In most cases, patients undergoing digitalis therapy may require these drugs for long periods for maintenance therapy, the dose requirements varying from patient to patient. Unfortunately, the margin between the dose levels required for optimum therapeutic efficacy on the one hand and manifestation of toxicity on the other is relatively narrow for these drugs and digitalis toxicity has been one of the most common of all adverse reactions following the use of these very valuable drugs. No sensitive method is now available for careful control of the concentration levels of these drugs in the blood sera of individual patients. The problem is further compounded since inadequate dosage levels of digitalis can evoke symptoms similar to those evidenced in instances of toxicity due to excess dose administration.

The radioimmunoassay technique developed recently provides a reliable means of estimation of the serum digitalis levels rapidly, specifically and precisely, thus enabling careful control of dose administration for the individual patients. Similarly, a radioimmunoassay procedure has recently been developed for accurate dose adjustment of phenytoin, an anti-covulsant drug used in the treatment of epilepsy. Radioimmunoassay is also useful in the monitoring of broad spectrum antibiotics like gentamycin which are highly potent and toxic and have a narrow therapeutic index. The simplicity can be appreciated considering the complexity and lack of sensitivity of the alternate microbiological assay procedures.

Radioimmunoassay of the thyroid hormones triiodothyronine and thyroxine, thyroid stimulating hormone and thyroid binding globulin are now very widely used in routine clinical diagnosis for differentiating hypo, hyper and euthyroid states. Measurement of angiotensin I by RIA for plasma renin activity helps in differential diagnosis of renal hypertension from other forms of hypertension. The spectrum of substances covered by Radioimmunoassay is constantly expanding and one of the recent inclusions is the RIA of myoglobin, a protein found in patients of myocardial infarcts, and its detection and estimation is of great value in early diagnosis of mycardial infarctions. Closely related to radioimmunoassay are competitive protein binding assay and radio receptor assays. In Competitive Protein Binding assay (CPB), instead of the antibody, a specific binding protein is used whereas in radio receptor assay, a specific biological receptor is used. Serum levels of Vitamin B_{12} are estimated by using intrinsic factor in a CPB and is useful in diagnosis of Vitamin B_{12} deficiency anaemias. Human chorionic gonodotropin is estimated by using leuteinised rat ovaries as the binding agent in place of antibody and the estimation of this hormone is useful in early detection of pregnancies, diagnosis of complications in pregnancies and placental tumours.

3. *Sealed Radiation Sources in Therapy*[26-29]

The use of external radiation sources in the treatment of cancer is one of the most well-known applications of radioisotopes in medicine. Radiotherapy is now recognised as one of the most valuable modes of cancer treatment and is often used as an adjunct to surgical and drug treatment.

Radioisotopes are used in a variety of ways in radiotherapy, and can be classified generally into (a) gamma ray beam units where radioactive material in large amounts is housed inside a suitable shielded unit having an aperture through which a beam of gamma ray emerges ; (b) localised gamma ray application with small sources by introducing them in the body cavities or by surgical implantation as in intracavitary or interstitial applications.

Gamma Ray Beam Units (Teletherapy Units)

A gamma ray beam unit consists essentially of a container adequately thick to house a large quantity of gamma emitting material and yet to reduce the radiation intensity at the walls of the container to permissible level. In one direction, a beam of gamma rays is allowed to emerge from the container and is indicated by a beam-defining system.

There are several essential physical properties that an isotope must have in order to be of use in teletherapy units. The gamma ray emission should be having as high an energy as possible. The isotope should have a practicable long half-life so that the source need not be changed often. The radioactivity of the source should be so high that an adequate gamma ray intensity can be achieved with a source of small dimensions. In practice, a source greater than 2 to 3 cm in diameter would give very large penumbra effects and one many centimeters in length would result in high self-absorption of radiation.

Table VII gives a list of isotopes which have so far been used in teletherapy and their characteristics.

The first gamma ray beam units were produced using radium since this was the only suitable gamma emitter available then. The great disadvantage of this isotope was that it was available only in small quantities of 10 gm or less and was very expensive. In modern teletherapy units, the radioactive source used is Cobalt 60, which can be produced in large quantities by irradiation of cobalt metal in a nuclear reactor. Gamma ray emission per curie of activity is much higher for cobalt 60 than for radium and cobalt 60 is relatively inexpensive and easily produced. Caesium 137 is the next most important isotope available for gamma ray beam units. This has a half-life of 33 years which is very substantial advantage over radiocobalt. Its gamma radiations have energy of 0·66 MeV ; hence compared to cobalt 60, housing for Cs 137 can be smaller. Cs 137 is present in substantial

quantities in the fission products from uranium and can be recovered from wastes of nuclear fuel reprocessing plants. However, the process of separation of Cs 137 from the fission liquids is a difficult one. The great disadvantage of Cs 137 compared to radiocobalt is that gamma emission per curie of material is relatively low and the activity per unit volume that can be obtained from fission liquids is also relatively low since active Cs 137 cannot be separated from inactive caesium isotopes present. The result of this is that in any source of reasonable size for beam therapy, the dose rate is likely to be rather small. Table VII also lists two other possible materials for gamma ray beam units. Of these, iridium 192 has been used in such a unit for therapy, but its short half-life causes great practical difficulties. Thulium 170 gives gamma radiation which is of very low energy and has been tried successfully as a source suitable for diagnostic radiography.

Interstitial and Intracavitary Application

The most common type of small source of gamma radiation for local application in tissues is a needle or tube containing radium or radon or radiocobalt and seeds of radon and grains of gold 198 and iridium 192.

Radioactive needles are thin hollow containers, made of a suitable metal such as platinum having a sharp end and are intended for surgical insertion into tissues. Such treatment with the needles actually buried in tissues is called interstitial therapy. The radioactive tubes are similar to needles but have no sharp points and are meant for intracavitary or surface applications. In this case, platinum can be replaced by a greater thickness of a cheaper metal.

Cobalt 60 Needles and Tubes

For interstitial and intracavitary use, cobalt wire activated in a reactor may be coated with a thin layer of a heavy metal to absorb beta particles or it may be enveloped in a plastic tubing. Thin-walled hollow stainless steel needles into which cobalt 60 wire is sealed can also be used.

Cobalt 60 has several advantages over radium ; (a) there is no associated gas like radon that could escape and be a source of hazard, (b) cobalt wire is more flexible and it is easier to make sources for curved surfaces such as bladder and tongue, (c) inactive ends of the needle can be practically eliminated and a better physical distribution of activity obtained.

The chief disadvantages of cobalt 60 is its half-life (5·3 years) which is relatively short. Needles and tubes made of tantalum 182, gold 198 and iridium 192 have also been used in place of radium and cobalt 60.

Radioactive Seeds and Grains

For special conditions, radioactive seeds and grains made of short-lived radioisotopes are sometimes more convenient. These may be left inside the patient unlike the needles. This is of particular value, for example, when a growth is exposed by surgical operation in bladder or stomach or lung. In such cases, if needles were used, a second operation would be necessary for their removal which can be averted by the use of short-lived isotopes. Radon seeds are made by containing the gas in a very narrow bore gold capillary tube. The seeds are usually 5 mm long and of 1·35 mm dia. Gold 198 grains are nowadays more frequently used in place of radium seeds. These are small pellets of gold made radioactive by neutron irradiation.

Typical Applications of Radiotherapy

Radiotherapy is one of the major modes of treatment of various malignancies. It has been widely used in the treatment of breast cancer in women ; as a mode of radical treatment in person unsuited for surgical treatment ; as a supplementary to surgical removal, to improve on the results of surgery and as palliation in cases where complete cure cannot be hoped for. In case of cancer of the oesophagus, surgical treatment is obviously a most difficult undertaking and for the great majority of the cases radiotherapy is the sole hope ; even if complete cure is not achieved, palliation for many months is generally possible. Radiotherapy is now used with increasing frequency and success in the treatment of cancer of the bladder. It can be used in the treatment of antral neoplasm often in combination with limited surgery. The primary treatment in case of medulloblastoma, a highly malignant tumour arising in the posterior part of the brain called cerebellum is normally surgical. But rapid recurrence always follows and it is essential to treat the whole of this large area with radiation. Skin cancer can be treated with radiation without undesirable cosmetic effects. Cancer of the cervix represents one of the major triumphs of radiotherapy. In the earlier days the patient's sole chance lay in severe major operation. But radiotherapy has completely

altered the picture. It can accomplish much the same as surgery for the operable cases while it has a great deal to offer in later stages where surgery is unsuitable. The mode of treatment is usually by inserting radioactive applicator within the central canal and across the vaginal vault up against the surface of the cervix and fornices. In cancer of the body of the uterus, radiotherapy is used with good results in cases where surgery is not preferred.

Radiation from radioisotopes have also been used in treatment of certain nonmalignant conditions such as rheumatism and menorrhagia. Strontium 90, a pure beta emitter is used widely in ophthalmic applicators for the treatment of superficial lesions of the eye. It is useful for treatment of corneal ulcer, corneal vascularization, and neoplasms on the conjuctiva.

4. *Applications for Medicare and Public Health*[30,31]

(a) *Radiation Sterilization of Medical Products*

Gamma radiation from high intensity radioisotopic sources are finding now increasing use in pharmaceutical and medical products industry. One of the most important processes in the pharmaceutical industry is sterilization i.e., the process of inactivating the micro-organisms. The efficiency of the sterilization process for a micro-organism is conveniently expressed in terms of the inactivation factor. It is the ratio of the initial number of micro-organisms to the surviving number at the end of the sterilization process. The inactivation factor for steam sterilisation at 121 °C for 20 mts is 10^{15} for many of the commonly occurring organisms though it could be as low as 500 for exceptionally heat-resistant strain of B. Stearomophillus. For sterilization with ethylene oxide which is equally effective against both vegetative cells and spores, the inactivation factor is of the order of 10^9. For radiation sterilization with a dose of 2·5 megarads, the inactivation factors for many organisms are several orders of magnitude more, whereas for some others they are comparable to those for steam sterilization.

Sterilization using gamma radiation also offers several other advantages. The sterilization can be carried out in the final package and the materials remain sterile as far as the packaging is not damaged. It is ideally suited for sterilization of thermolabile products made of plastics which cannot be steam sterilized. There is no problem of gas retention as in the case of ethylene oxide gas sterilization. Of practical significance is low radio resistance

of pathogenic micro-organisms such as pseudomonas species, E. Coli and Salmonella species.

The radiation sterilization plant consists of a well-shielded facility with mega curie quantity of cobalt 60 arranged in an array. The container containing the material to be sterilized travels through the radiation field which is of the order of 10-100 rads/sec. till it receives a uniformly distributed sterilization dose of 2·5 megarads. Rigorous dosimetry control is instituted to ensure that each container receives 2·5 megarads dose uniformly. Dosimetry is usually carried out with Perspex dosimeters sent along with the containers and radiation induced unsaturated species is measured spectrophotometrically. Radiation sterilization can also be useful for certain parenteral drugs in solid form as these are difficult to be sterilized by steam or ethylene oxide. It can also be used for sterilization of injectable aqueous solutions of some electrolytes, dusting powders of antibiotics and steroids, ophthalmic preparations of ointments, cosmetic material such as talcs and fatty acid esters and health care items such as sanitary pads.

There is a variety of disposable sterile medical products, devices and equipments made of plastics which can be advantageously sterilized by gamma radiation. These include hypodermic syringes, sutures, catheters, connectors, implants, hydrocephalus valves, arteriovenous shunts, blood transfusion and infusion sets and connectors for artificial kidney units. These are particularly useful for use in remote places as in camp surgery where sterilization facilities are inadequate as well as for emergency surgical procedures. Cellulosic products such as absorbent cotton wool, plastic wound dressing, surgical dressings, plain or impregnated with euflavin are conveniently sterilized by gamma radiation in polyethylene packings. Radiation sterilized vesectomy and tubectomy kits for camp use have several practical advantages. The kits can guarantee sterility and high degree of safety, reduce post-operative complications due to infection and rapid turnover of cases. A number of radiation sterilized medical products are now available commercially in India and other countries and studies to extend the process for several other items are under way.

(b) *Radiation Hygienisation of Sewage*[32]

A recent notable contribution of gamma radiation from radioisotopes to public health is treatment of sewage from cities which contain a high concentration of micro-organisms and parasites causing cholera, tuberculosis,

gastroenteritis, typhoid, polio, hepatitis, dysentry, etc. Presently the sewage after collection is released directly or after some form of treatment into the sea, river or water courses or on the land. Such unsafe disposal methods not only pose a serious threat to the environment and public health but also lead to wastage of the high nutrient value of the sewage. Experiments have shown that properly treated sludge can also be used as supplementary cattle feed apart from its safe usage as manure.

The application of radiation for treating sewage sludge offers a number of technical advantages. Besides totally destroying the pathogens and parasites, radiation also improves the physical properties of the sludge. It has been shown that a dose of 300 kilorads brings about the following changes in the digested sludge :

(a) Reduction of foliforms bacteria by 10 log cycles
(b) Faecal streptococcus reduced by 2 logs
(c) Salmonella species reduced by 8 logs
(d) Viable parasite ova reduced by at least 4 logs
(e) Viruses reduced at least by 1 log
(f) Improvement in the ability of solids to settle and enhancement in their capability to be filtered by a factor of over three
(g) Organics and inorganics degraded and disintegrated by at least one log unit

Radiation treatment is a cold process and is effective in rendering pathogens harmless. In the treatment of sludge both electrons from high voltage accelerators as well as gamma radiation from isotope sources are of potential interest. Economic analysis indicates that cobalt 60 radiation sources would be ideal for plants with capacities below 250 cu m of sludge per day.

5. *Radioisotope Applications in Medical Research*[34]

Starting with the understanding of the mechanism of photosynthesis with C 14 labelled carbon dioxide, the availability of radioisotopically labelled nutrients, metabolites, precursors, enzymes, nucleotides, lipids and other biomolecules have made enormous contributions to our understanding of many complex biochemical processes and their control. Most of these compounds are labelled with Carbon 14, tritium, sulphur 35 and phosphorus 32. As in the case of radiopharmaceutical development, medical research in the initial stages covered mostly study of the accumulation and fate of trace

chemical and micro-biological stability. Specific examples of radioisotopic usage in drug product design is in measuring *in vitro* release and dissolution, in measurement of permeability, penetration and absorption, in bio-availability studies and in evaluation of drug interaction with package components, notably plastic and rubber.

Status in India

Thanks to the vision of Jawaharlal Nehru and Homi Bhabha who ushered in the atomic age in this country, India now ranks high among the six or seven countries in the world with a self-sufficient and comprehensive programme of isotope production and applications. The radiopharmaceuticals produced at the Bhabha Atomic Research Centre have been acclaimed for their quality not only by leading hospitals and medical institutions in this country, but by several advanced centres abroad to whom these products are regularly supplied. Other important achievements have been the development of technology for high intensity radiation sources, radiation sterilization of medical products and teletherapy of cancer. An active programme of research and development is also under way at BARC to make available newer and more advanced techniques and applications for the understanding and diagnosis of disease and alleviation of suffering.

Table I

Some Important Radioisotopes used in Medical Diagnosis and Therapy

	Isotope	*Half-Life*	*Method of Production*
I	**Reactor-Produced Isotopes :**		
1.	Hydrogen 3	12·26 y	Li 6 (n, α) H3
2.	Carbon 14	5760 y	N 14 (n, p) C14
3.	Fluorine 18	110 mt	O 16 (α, np) F18
4.	Sodium 24	15 h	Na 23 (n, γ) Na24
5.	Magnesium 28	21·4 h	Mg 26 (t, p) Mg28
6.	Phosphorus 32	14·3 d	S32 (n, p) P32
7.	Sulphur 35	87·2 d	Cl 35 (n, p) S35
8.	Calcium 45	165 d	Ca 44 (n, γ) Ca45
9.	Potassium 42	12·4 h	K 41 (n, γ) K42
10.	Chromium 51	27·8 d	Cr 50 (n, γ) Cr51
11.	Manganese 54	314 d	Fe 54 (n, p) Mn54
12.	Iron 59	45 d	Fe 58 (n, γ) Fe59
13.	Cobalt 58	71 d	Ni 58 (n, p) Co58
14.	Cobalt 60	5·26 y	Co 59 (n, γ) Co60
15.	Copper 64	12·84 h	Cu 63 (n, γ) Cu64
16.	Zinc 65	245 d	Zn 64 (n, γ) Zn65
17.	Bromine 82	35·4 h	Br 81 (n, γ) Br82
18.	Selenium 75	121 d	Se 74 (n, γ) Se75
19.	Strontium 90	28 y	U (n, f) Sr90
20.	Molybdenum 99	67 h	Mo 98 (n, γ) Mo99
21.	Yttrium 90	64·2 h	Y 89 (n, γ) Y90
22.	Technetium 99 m	6 h	Mo 98 (n, γ) Mo99 $\xrightarrow{\beta^-}$ Tc99 m
23.	Ruthenium 103	40 d	Ru 102 (n, γ) Ru103
24.	Cadmium 115	2·3 d	Cd 114 (n, γ) Cd115
25.	Indium 115 m	4·5 h	Cd 114 (n, γ) Cd115 $\xrightarrow{\beta^-}$ In115 m
26.	Tin 113	118 d	Sn 112 (n, γ) Sn113
27.	Indium 113 m	100 m	Sn 112 (n, γ) Sn113 $\xrightarrow{EC}$ In113 m
28.	Iodine 131	8·04 d	Te 130 (n, γ) Te131 $\xrightarrow{\beta^-}$ I 131
29.	Iodine 125	60 d	Xe 124 (n, γ) Xe125 $\xrightarrow{EC}$ I 125

metals like arsenic, zinc, copper, manganese, etc., in body tissues using the respective radioactive elements. But soon, procedures were developed for the synthesis and purification of a host of biomolecules which rapidly gained acceptance as very useful tools in biomedical research. To cite an example, amino acids are the basic constituents of proteins, compounds essential to all biochemical functions. Investigations into the synthesis, metabolism and function of amino acids and proteins have been greatly assisted by the availability of amino acids labelled with tritium, carbon 14 and sulphur 35. Apart from the common amino acids such as aspartic acid, threonine, serine, glutamic acid, proline, glycine, alanine, valine, methionine, isoleucine, leucine, tyrosine, phenylalanine, histidine, lysine and arginine, a wide variety of less common amino acids such as dihydroxy phenylalanine, taurine and kainic acid which are important in the investigation of neurochemical processes are now available commercially with radioactive labels.

Among the rapidly expanding areas of biomedical research, of particular importance is the field associated with transmission of nerve impulses. Radiolabelled neurochemicals play a key role in this research by providing tools to study the transmission processes. From a medical viewpoint, a knowledge of the mechanisms involved in neurotransmission has wide implications in understanding and control of many physiological conditons, for example, schizophernia, Huntigton's chorea and depressive illness. It is postulated that neurotransmitters exert their effects on target tissues by interacting with specific receptor sites. Receptor site binding can be studied using radiolabelled neurotransmitters themselves like acetyl choline or agonists like nicotine or antogonists like quinuclidinyl benzilate which are tagged with radioactivity. Other compounds of neurological interest which have been radiolablled include catecholamines, α-neurotoxins (such as α-bungaro toxin) and inhibitory transmitter 4-amino butyric acid. The availability of P 32 labelled nucleotides such as (α-P 32) ATP (α-P 32) GTP (α-P 32) CTP and (α-P 32) UTP is of great help in understanding the synthetic and metabolic pathways of DNA and RNA.

The use of radiotracers in biomedical research is so widespread and common that it can be safely presumed that future investigations of any complex biochemical system will not be complete without the use of radiotracer techniques.

In cancer research, radioactive tracers are used widely to investigate the biochemical processes associated with the disease, for instance, the changes in

the synthesis and metabolism of hormones, vitamins, nucleic acid, proteins, lipids, etc.

Radioisotopes in Drug Research[33]

Radioisotopic tracers have contributed to a great extent in drug research and development by providing simple and fast methods for the study of their absorption, distribution, metabolism and excretion. Major drug companies today routinely prepare radiolabelled forms of new drugs in order to learn about their absorption fate and elimination. Initial drug studies are conducted in test animals often in two or three species. A known amount of labelled drug is administered after which the animals are placed in metabolic cages where urine, faeces and exhaled air are collected. The tracers of labelled materials in these samples as well as in selected organs of sacrificed animals are identified and quantified. To obtain more detailed information on the nature of the metabolites, the radioactive sample is typically separated into its various fractions by chromatographic procedures and the various metabolites are then identified and quantified. Most of the drugs are labelled with carbon 14 or tritium (H 3) or sulphur 35 and the radioisotope is located at an appropriate site of the molecule where the radioactive atom will not readily exchange or will not be readily degraded. They can enter the molecule in place of their stable counterparts, namely C 12, H 1 and S 32. For drug metabolic studies, C 14 is preferred to tritium for several reasons. The figure of merit of H 3 versus C 14 is in favour of the latter with conventional counting systems ; the position of the label is more precisely known with C 14 than with H 3. The C 12–C 14 bond is less labile than the C 12–H 3 bond ; the mass ratio of C 14/C 12 leads to less isotopic effect than the mass ratio of H 3/H. In favour of tritium are the price of the labelled molecule and the high specific activity available. Aspirin, barbital, amphetamine, chloroquine, digitoxin, ephedrine, griseofulvin, isoniazid, morphine, nicotine and phenobarbital are typical examples of drugs which have been labelled with C 14. Acriflavine, chloramphenicol, and tetracycline are among drugs which have been incorporated with tritium. Important drugs labelled with S 35 include chloropromazine and penicillin.

Apart from study of drug metabolism and biological and pharmacological properties, these labelled drugs are also used extensively in drug product design and development to identify the physical and chemical interactions between the drug and any proposed excipients and to establish their physical,

Table I (Concld.)

	Isotope	*Half-Life*		*Method of Production*
30.	Xenon 133	5·3	d	Xe 132 (n, γ) Xe133m $\xrightarrow{IT}$ Xe133- U (n, f) Xe 133
31.	Ceasium 131	9·7	d	Ba 130 (n, γ) Ba131 $\xrightarrow{EC}$ Cs 131
32.	Ceasium 137	30	y	U (n, f) Cs 137
33.	Ytterbium 169	31	d	Yb 168 (n, γ) Yb 169
34.	Thulium 171	680	d	Er 170 (n, γ) Er171 $\xrightarrow{\beta^-}$ Tm171
35.	Rhenium 188	17	h	Re 187 (n, γ) Re 188
36.	Iridium 192	74	d	Ir 191 (n, γ) Ir192
37.	Gold 198	2·7	d	Au 197 (n, γ) Au198
38.	Mercury 198	65	h	Hg 196 (n, γ) Hg197
39.	Mercury 203	47	d	Hg 202 (n, γ) Hg203
40.	Thallium 204	3·76	y	Tl 203 (n, γ) Tl 204
41.	Radium 226	1620	y	Naturally occurring radioisotope
42.	Radon 222	3·8	d	Naturally occurring radioisotope
43.	Plutonium 238	87·75	y	Np 237 (n, γ) Np 238 $\xrightarrow{\beta^-}$ Pu 238
II	**Cyclotron-Produced Radioisotopes :**			
44.	Carbon 11	20	m	B 10 (d, n) C11
45.	Oxygen 15	2	m	N 14 (d, n) O15
46.	Nitrogen 13	10	m	C 12 (d, n) N13
47.	Sodium 22	2·6	y	Mg 24 (d, α) Na22
48.	Bromine 77	58	h	As 75 $(\alpha, 2n)$ Br77
49.	Iodine 123	13	h	Sb 121 $(\alpha, 2n)$ I 123
50.	Potassium 43	22	h	Ar 40 (α, p) K43
51.	Iron 52	8·3	h	Cr 50 $(\alpha, 2n)$ Fe52
52.	Cobalt 57	270	d	Ni 60 (p, α) Co57
53.	Zinc 62	9·33	h	Ni 60 $(\alpha, 2n)$ Zn62
54.	Gallium 67	78	h	Zn 66 (d, n) Ga67
55.	Rubidium 81	4·7	h	Br 79 $(\alpha, 2n)$ Rb81
56.	Yttrium 87	80	h	Rb 85 $(\alpha, 2n)$ Y87 m $\xrightarrow{IT}$ Y87
57.	Indium 111	2·8	d	Ag 109 $(\alpha, 2n)$ In111
58.	Lead 203	52	h	Tl 203 (p, n) Pb203
59.	Thallium 201	73	h	Tl 203 $(p, 3n)$ Pb201 $\xrightarrow{\beta^+, EC}$ Tl 201

Table II

Important Labelled Radiopharmaceuticals

Sl. No.	*Labelled Radiopharmaceuticals*	*Method of Preparation*
1.	I 131 Rose Bengal (tetraiodo tetrachloro fluorescein)	Isotopic exchange reactions
2.	I 131 Hippuran (orthoiodo hippuric acid)	Isotopic exchange reactions
3.	I 125 or I 131 triiodothyronine	Isotopic exchange reactions
4.	I 125 or I 131 thyroxine	Isotopic exchange reactions
5.	I 131 human serum albumin	Trace iodination
6.	I 131 fibrinogen	Trace iodination
7.	I 131 insulin	Trace iodination
8.	I 131 triolein	Trace iodination
9.	I 131 oleic acid	Trace iodination
10.	Hg 197 or Hg 203 neohydrin (3 chloromercury methoxy propylurea)	Chemical synthesis
11.	Hg 197 bromomercury hydroxy propane	Chemical synthesis
12.	Co 57 or Co 58 Vitamin B_{12}	Biosynthesis
13.	Se 75 selenomethionine	Biosynthesis

Table III

Radioisotope Generators of Medical Interest

Sl. No.	*Daughter Isotope*	*Half-life*	*Parent Isotope*	*Half-life*
1.	Technetium 99 m	6 hrs	Molybdenum 99	67 hrs
2.	Indium 113 m	1·7 hrs	Tin 113	118 days
3.	Indium 115 m	4·5 hrs	Cadmium 115	2·3 days
4.	Yttrium 90	64 hrs	Strontium 90	28 yrs
5.	Strontium 87 m	2·8 hrs	Yttrium 87	80 hrs
6.	Gallium 68	68 mts	Germanium 68	280 days
7.	Iodine 132	2·3 hrs	Tellurium 132	3·2 days
8.	Barium 137m	2·6 mts	Ceasium 137	30 yrs
9.	Rubidium 82	80 sec	Strontium 82	25 days
10.	Krypton 81m	13 sec	Rubidium 81	4·7 hrs

Table IV

Important Tc 99 m Labelled Radiopharmaceuticals used in Organ Scintigraphy

Sl. No.	*Radiopharmaceuticals*	*Use*
1.	Tc 99 m pertechnetate	Brain and thyroid imaging
2.	Tc 99 m phytate Tc 99 m sulphur colloid	Liver and spleen imaging
3.	Tc 99 m albumin micro-spheres Tc 99 m macroalbumin aggregates	Lung imaging
4.	Tc 99 m glucoheptonate Tc 99 m dimercaptosucci-nate	Kidney imaging
5.	Tc 99 m pyridoxal glutamate Tc 99 m 2, 6 dimethyl ace-tanilidoiminodiacetate	Imaging of hepatobiliary system
6.	Tc 99 m ethane hydroxy diphosphonate Tc 99 m methylene diphosphonate	Skeletal scanning
7.	Tc 99 m pyrophosphate	Bone and myocardial imaging
8.	Tc 99 m red blood cells	Cardiac dynamic studies

Table V

Therapy with Internally Administered Radiopharmaceuticals

Sl. No.	*Radiopharmaceuticals*	*Indication*
1.	I 131 sodium iodide	Hyperthyroidism ; thyroid cancer
2.	I 125 sodium iodide	Hyperthyroidism
3.	P 32 sodium phosphate	Haematologic proliferating dis-orders
4.	Y 90 chelates	Polycythemia rubra vera, leukemia
5.	S 35 sodium sulphate	Thrombocytosis ; chondrosarcoma
6.	Au 198 gold colloid	Malignant pleural and peritoneal effusions
	P 32 chromic phosphate colloid	Lymphomas and lymph node meta-stases from solid tumours.
7.	P 32 or Y 90 labelled ceramic or plastic micro-spheres.	Solid organ tumours.

Table VI

Important Antigens for which Radioimmunoassay has been Developed

Protein hormones	*Drugs*
Human placental lactogen	Digoxin
Human growth hormone	Phenytoin
Insulin, thyroglobulin	
Human chorionic gonodotropin	Morphin
Angiotensin I	LSD
Thyroid stimulating hormone	Gentamycin
Leutinising hormone	Bleomycin
Follicle stimulating hormone	
Adrenal corticotropic hormone	
Non-protein hormones	*Other Antigens*
Triiodothyronine	Australia Antigen
Thyroxine	α-fetoprotein
Cartisol	Myoglobin
Estriol	Choriocarcinogenic Antigens
Progesterone	Calcitonin

Table VII

Radioisotopes for Teletherapy

Sl. No.	*Radioisotope*	*Radiation*	*Energy (MeV)*	*Half-life*
1.	Radium 226	alpha gamma	4·78 0·185 to 0·77	1620 yrs
2.	Cobalt 60	beta gamma	1·48 1·17, 1·33	5·26 yrs
3.	Ceasium 137	beta gamma	1·176 0·662	30 yrs
4.	Iridium 192	beta gamma	0·24 to 0·536 0·296 to 0·612	74 days
5.	Thulium 170	beta gamma	0·97 0·084	127 days

REFERENCES

1. Radioisotope Production and Quality Control. Technical Report Ser. 128 (1971). International Atomic Energy Agency, Vienna.
2. Manual of Radioisotope Production. Tech. Report Ser. 63 (1966). International Atomic Energy Agency, Vienna.
3. Andrews, G. A., Kniselay, R. M. and Wagner, H. N. (1966). Radioactive Pharmaceuticals (Ed.) United States Atomic Energy Commission.
4. Mani, R. S., Desai, C. N. and Iya, V. K. (1971) Synthesis of labelled molecules. *In* Radioisotope Production and Quality Control. Technical Report Ser. 128. International Atomic Energy Agency, Vienna.
5. Narasimhan, D. V. S. and Mani, R. S. (1978) Radiopharmaceuticals. *Ind. J. Pharm. Edu.*, **12** : 80.

6. McAfee, J. C. and Subramanian, G. (1969) Radiopharmaceuticals. *In* Clinical Scintillation Scanning, (Ed.), L. M. Freeman and P. M. Johnson, Harper and Row, N. Y.
7. Bland, W. H. (1971) (Ed.), Nuclear Medicine, McGraw Hill, N.Y.
8. Wagner Jr, H. N. (1968) (Ed.), Principles of Nuclear Medicine, W. G. Saunders, Co.
9. Siemensan, J. K. and Telfer, N. (1976) Nuclear Medicine. *In* Radiopharmacy, (Ed.), M. Tubis and W. Wolf, John Wiley and Sons.
10. McIntyre, P. A. (1972) Radioactive tracers in haematologic disease. *Hosp. Prac.*, 7 : 147.
11. Herbert, V. (1972) Detection of Malabsorption of Vitamin B_{12} due to gastric or intestinal dysfunction, *Sem. Nucl. Med.*, 2 : 220.
12. Belcher, E. H. and Vetter, H. (1971) Radioisotope in medical diagnosis. (Ed.), Appleton-Century-Crofts, N.Y.
13. Herbert, J. C. (1971) Radionuclide Cisternography, *Sem. Nucl. Med.*, 1 : 90.
14. Deland, F. H. and Wagner, H. N. (1973) Atlas of Nuclear Medicine (Ed.), W. B. Saunders Co.
15. Medical Radioisotope Scintigraphy (1968) Proceedings of Symposium held at Selzburg, International Atomic Energy Agency, Vienna.
16. Medical Radioisotope Scintigraphy (1972) Proceedings of Symposium held at Monte Carlo, International Atomic Energy Agency, Vienna.
17. Blaufox, D. (1972) Evaluation of Renal Function and Disease with Radionuclides (Ed)., University Park Press, USA.
18. Lele, R. D. (1978) Cardiovascular Nuclear Medicine. *Bull. Jaslok Hospital and Research Centre*, 3 : 81.
19. Radiopharmaceuticals from generator produced radionuclides. Panel Proceedings Series, International Atomic Energy Agency, Vienna.
20. Mani, R. S. and Narasimhan, D. V. S. (1973) Development of kits for short-lived generator produced radioisotopes. Radiopharmaceuticals and labelled compounds. Proceedings of Symposium held at Copenhagen, International Atomic Energy Agency, Vienna, 1 : 39.
21. Tubis, M. and Wolf, W. (1976) Cyclotron Produced Radiopharmaceuticals. *In* Radiopharmacy. (Ed.), M. Tubis and W. Wolf, John Wiley and Sons.
22. Wasserman, L. R. and Glass, J. L. (1968) Present status of the therapy with radioactive isotopes, *Prog. Atom. Med.*, 2 : 113.
23. Mani, R. S. (1976) Radioimmunoassay—a unique radiometric technique. *Nuclear India*, 5(14) : 11.
24. Chard, T. (1978) An introduction to radioimmunoassay and related techniques. *In* Laboratory techniques in biochemistry and molecular biology, 6 (Ed.), J. S. Work and E. Work, North Holland Pub. Co.
25. Besch, P. K. (1975) Clinical Radioassay Procedures—A Compendium (Ed.) Amer. Assn. Clinical Chemists.
26. Walter, J. and Miller, H. (1959) A short textbook of Radiotherapy (Ed.), J. A. Churchill Ltd., UK.
27. Deelay, J. J. (1976) Principles of Radiation Therapy, Butterworth, U.K.
28. Buschks, F. and Parker, R. G. (1972) Radiation Therapy in Cancer Treatment, Grime and Stratton.
29. Fenner, M. L. (1968) Elementary Clinical Radiotherapy, Butterworth, UK.
30. Manual of Radiation Sterilisation of Medical and Biological Material. Technical Report Ser. 149. International Atomic Energy Agency, Vienna.
31. Gopal, N. G. S. (1977) Recent Advances in Radiation Applications in Pharmaceutical and Medical Product Industry. Proc. Lecture Ser. on Recent Advances in Chemical Industry held at Bombay, Ind. Chem. Manufacturers Assn.
32. Krishnamurthy, K. (1979) Reusing the City's Sewage. *Nuclear India*, 17 : 12.
33. Banker, G. S. (1976) Radionuclides and Radiation in Pharmaceutical Industry. *In* Radiopharmacy. (Ed.), M. Tubis and W. Wolf, John Wiley and Sons.
34. Technical Bulletin periodically published by Radiochemical Centre, Amersham, UK.

A Review on Quarantine for Seed

Paul Neergaard

Danish Government Institute of Seed Pathology for Developing

Ryvangs allé 78, DK-2900 Hellerup, Denmark

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

1. Biological Background

1.1. *International Spread of Diseases by Seed*

The international exchange of seed in trade and agricultural development is more extensive now than ever before, and is ever-increasing. The size of individual consignments of seed may vary from shiploads of grains for consumption or tons of seed intended for sowing down to a few grams of seeds for experimentation—or just some for the pleasure of the owner of a garden. No statistics provide a complete picture of the amounts involved but the total number of seed consignments in the broadest sense moving annually across boundaries runs quite certainly into hundreds of thousand.

A recently published Annotated List of Seed-borne Diseases[46] includes more than one thousand host/pathogen combinations of seed-borne diseases, and the number of seed-borne diseases now known to be caused by bacteria, fungi, viruses and nematodes, is continuously increasing.

The risk of seed transmission of pathogens varies according to host and pathogen. Some crop plants have a wide range of harmful seed-borne diseases, while others seem to be practically free. Crop plants grown extensively and in mass culture appear to be particularly afflicted, partly because such cultivation favours building up of inoculum, such as in cereals, certain fibre crops, and legumes for food and fodder. There are 10–20 serious seed-borne diseases in each of rice, wheat, maize, barley, sorghum, and oats ; in cotton and flax ; and in leguminous food crops such as soybean (about 20 serious seed-borne diseases), Phaseolus bean, and groundnut.

Among the most important categories of seed-borne diseases are cereal smuts, leaf spot, leaf blight and root rot diseases caused by *Drechslera* spp. in cereals, root and stem rot diseases caused by *Fusarium* spp. in many crops, wilts caused by *Fusarium oxysporum* and *F. moniliforme* in many crops, anthracnoses, caused by *Colletotrichum* spp. particularly in leguminous crops, bacterial blights and wilts caused by *Pseudomonas* spp. and *Xanthomonas* spp. and numerous diseases caused by viruses in many different crops, though particularly in leguminous crops.

Many seed-borne pathogens have already been spread by seed throughout the world. Nevertheless their pathogenic potentials, pertinent to resistance of local cultivars, may vary considerably between different parts of the world. This is due to differences from place to place in the populations

of pathogenic races, some races being different in aggressiveness ("horizontal" races), others in their host range of cultivars, being specific pathogenic races ("vertical" races). The rice blast fungus, *Pyricularia oryzae* has several hundred specific pathogenic races, and *Tilletia caries* and *T. foetidae* combined have a total of at least 30 pathogenic races, to mention just two examples. In some instances, introduction of races new to an area has had catastrophic effects.

A considerable number of case histories of international spread of seed-borne diseases from one country or part of the world to another is known. The oldest well-established example is perhaps that of the spread of bunt of wheat, *Tilletia caries* to USA from Australia at about the middle of the last century. The disease appeared for the first time in 1854 in the Sacramento Valley, California, in a field where Australian seed had been used.

Neergaard[41] reviewed 25 examples of seed-borne pathogens that have spread internationally by means of seed. They include 5 species of bacteria, 4 species of rust fungi, 4 species of smut fungi, 3 species of Oomycetes (downy mildew), 7 species of Fungi Imperfecti/Ascomycetes, and 2 viruses.

A few other, recent examples of international spread :

In South Australia Randles and Dube[43] recorded infection of *Ascochyta fabae* in a crop of broad bean or field bean, *Vicia faba* which was isolated from other broad bean crops and was sown in an area where *V. faba* was being grown for the first time. The seed had been imported directly from the United Kingdom.

For the first time *Xanthomonas oryzae*, bacterial leaf blight in rice, has been observed in recent years in the Caribbean region and several countries of South America[81]. Although the disease had been known for some years in America, it was concluded that the pathogen was introduced by the use of infected seed imported from infested areas in other parts of the world. This is indicated by the mode of spreading of the disease and the origin of the affected cultivars.

Echtes Ackerbohnenmosaik-Virus, in field bean or broad bean, *Vicia faba* was detected for the first time in South Australia in plants from seed imported from the United Kingdom[48].

Finally an example illustrative of the danger of introducing new pathogenic races to a geographic area. In Upper Volta rice fields have for ages carried miserable yields, 700 kg/ha, but utilising the technology of the Green Revolution, the damming of the Volta River, and using the new

cultivars, etc., dramatic yield increases were obtained, 14t/ha with two crops per year. Then in 1971 devastating epidemics of blast (*Pyricularia oryzae*) spread on the crops like wildfire, destroying the new cultivars as well as the old ones, reducing yields to the original level of 700 kg/ha, it was a disaster. The situation was now much worse than before the Green Revolution with all of its investments[22]. The introduction of new seeds from breeding centres, such as the International Rice Research Institute, enabled pathogenic races that were new for African rice to come in, to spread and attack the introduced as well as the African rice which was resistant to the African pathogenic races but susceptible to the introduced Asian races. Also in Nigeria domestic resistant rice cultivars suddenly were attacked by the blast, again the work of introduced races to which they were susceptible[40].

1.2. *Epidemiological Types of Seed-borne Pathogens. Categorisation of Quarantine Objects*

Any control of seed-borne disease must be related to its epidemiological type. This certainly applies also to exclusion by quarantine, including selection of appropriate inspection procedures and other precautions. Neergaard[41] outlined the following six epidemiological types of seed-borne pathogens :

(1) Pathogens depending exclusively or almost exclusively on seed transmission
Examples : most cereal smuts, *Drechslera graminea*, barley stripe mosaic.

(2) Pathogens that are soil/plant residue invaders, crop rotation being considered in control
Examples : *Alternaria, Drechslera, Fusarium, Phoma, Septoria.*

(3) Pathogens for which vectors play a dominant role
Examples : vector-borne viruses, *Erwinia stewartii*

(4) Pathogens in which the perfect state provides important air-borne inoculum
Examples : *Sclerotinia sclerotiorum, Mycosphaerella pinodes, Gloeotinia temulenta.*

(5) Pathogens with potential for rapid epidemiological spread, often transmitted from trace amount of seed-borne inoculum.
Examples : downy mildews, rusts, bacteria.

(6) Pathogens often or exclusively present in the host as systemic infection, often carried in a symptomless manner by the host.
Examples : *Fusarium oxysporum, F. moniliforme*, certain other Fungi Imperfecti, wilt producing bacteria, most viruses.

Categories of Quarantine Objects

With special regard to quarantine precautions it is necessary to categorise pathogens considered in this connection. To be considered as true quarantine objects are (1) pathogens not present in the area to be protected by quarantine precautions or present only in limited localities in which it is under effective control (domestic quarantine), (2) pathogens dangerous by their direct pathogenic potential, and (3) pathogens dangerous because of their epidemic potential, as particularly type 5 above.

Based on these considerations the following three categories of quarantine objects have been proposed[40, 41].

Category A

Definition : Dangerous plant pathogens which are not present in the region of introduction and which have a high or considerable epidemic potential. Many pathogens which belong to this category occur in seed in only trace amounts, thus rendering sampling unreliable.

Precautions : Complete prohibition against introduction from areas with infection. Testing based on sampling is not adequate. Particularly valuable material from infested areas may be "filtered" through postentry control by growing seed-bearing plants from introduced seeds under closed quarantine, i.e., in special greenhouses that exclude escape.

Examples : Arabis mosaic virus, barley stripe mosaic virus, Echtes Ackerbohnenvirus, groundnut stunt virus, pea seed-borne mosaic virus (pea), stone fruit viruses, sunblotch virus of avocado, *Corynebacterium flaccumfaciens, Corynebacterium rathayi, Alternaria triticina, Peronospora tabacina, Plasmopara halstedii, Tilletia contraversa.*

Category B

Definition : Dangerous plant pathogens not present in the region of introduction (or present only in more or less restricted areas being under effective control) and which have a moderate epidemic potential.

Precautions : The seed may be tested, based on adequate sampling, if a suitable testing procedure is available. To be suitable for release, the tested samples must be found completely free from infection or contamination.

Examples : *Xanthomonas vesicatoria*, *Colletotrichum lindemuthianum* (pathogenic races), *Phoma lingam*, *Pyricularia oryzae* (pathogenic races ; in rice), *Tilletia caries* and *T. foetida* (races ; wheat), *Ustilago nuda* (races ; barley), *Ustilago tritici* (races ; wheat), *Ditylenchus dipsaci* (races).

Category C

Definition : Other plant pathogens of importance to the planting value of seed. These pathogens are not strict quarantine objects but may be of extreme importance for quality. However, a definite risk pertaining to quarantine is the introduction of new aggressive races into the region under consideration—aggressiveness in relation to the horizontal resistance already established in local cultivars during years of isolation.

Precautions : Testing by adequate procedures of representative samples is advisable. While disease tolerance may be acceptable for ordinary seed trade consignments according to certification standards of the importing country or according to international standards, it is advisable that material from gene centers be handled with more care in an attempt to prevent the introduction of pathogenic races, including both specific races ("vertical pathotypes") and aggressive races ("horizontal pathotypes") from centers of diversity. Seed treatment may be considered.

Examples : Alfalfa mosaic virus, soybean mosaic virus, soybean stunt virus, *Corynebacterium michiganense*, *Pseudomonas lachrymans*, *Pseudomonas phaseolicola*, *Xanthomonas phaseoli*, *Ascochyta rabiei*, *Septoria appiicola*, *Septoria nodorum*.

It must be emphasized that, quite obviously, referring any pathogen to any of the above categories cannot be generally valid. Such categorization of a pathogen depends on whether or not the pathogen (including pathogenic races, if any) is endemic in the region of introduction, and on the effectiveness of the surveys of pathogens present in both the countries and regions involved.

It is worth-mentioning in this connection that EPPO has established the following two lists of pathogens under quarantine[10] :

List A_1 : referring to pests and diseases not yet introduced into the area covered by EPPO. For these quarantine organisms a *zero tolerance* is required for all countries.

In this list are included the following seed-borne pathogens : *Xanthomonas oryzae*, *X. oryzicola* and *Tilletia indica*.

List A_2 : referring to pests and diseases already introduced into some countries belonging to the area covered by EPPO. For these organisms a *zero tolerance* is required.

Seed-borne pathogens included: *Corynebacterium flaccumfaciens*, *C. insidiosum*, *C. michiganense*, *Pseudomonas glycinea*, *P. pisi*, *Xanthomonas phaseoli* f. sp. *fuscans*, *X. phaseoli* f. sp. *phaseoli*, *X. stewartii*, barley stripe mosaic virus, potato spindle tuber "virus", tomato ringspot virus, *Drechslera carbonum*, *D. maydis*, *Didymella ligulicola*, *Diplodia macrospora*, *D. maydis*, *Glomerella gossypii*, *Isariopsis griseola*, *Septoria linicola*, *Tilletia contraversa*, *Aphelenchoides besseyi*.

A *List B* considered will include organisms important to planting value of materials for propagation, but not regarded as quarantine organisms proper.

2. Legislation and Regulation

2.1. *International Plant Protection Convention and International Cooperation in Plant Quarantine*

Phytosanitary regulations in connection with plant protection by quarantine are workable only through permanent international cooperation and well-established mutual confidence between the countries involved.

In international law this cooperation is based on the 1951 International Plant Protection Convention, the so-called "Rome Convention", drawn up under the aegis of FAO[14]. By October 1977, 77 countries had signed or acceded to the Convention. The Convention provides for National Organisation for Plant Protection, for issuance of Phytosanitary Certificates, it defines requirements in relation to imports, defines international cooperation between the contracting governments, and requires the setting up of Regional Plant Protection Organisations. Finally, it provides the text of a Model Phytosanitary Certificate to be adopted by the contracting governments the "Rome Certificate". The Rome Certificate is now in general use and is accepted not only by the contracting countries but also by many countries who have not yet signed the Convention.

The Regional Plant Protection Organisations covering practically all geographic regions of the world are the following :

European and Mediterranean Plant Protection Organisation (EPPO)
Inter-African Phytosanitary Council (IAPSC)
Organismo Internacional Regional de Sanidad Agropecuaria (OIRSA)
Plant Protection Committee for the South-East Asia and Pacific Region (SEAPPC)
Near East Plant Protection Commission (NEPPC)
Comité Interamericano de Protection Agricola (CIPA)
Organismo Bolivariano de Sanidad Agropecuaria (OBSA)
Caribbean Plant Protection Commission (CPPA)
North American Plant Protection Organization (NAPPO)

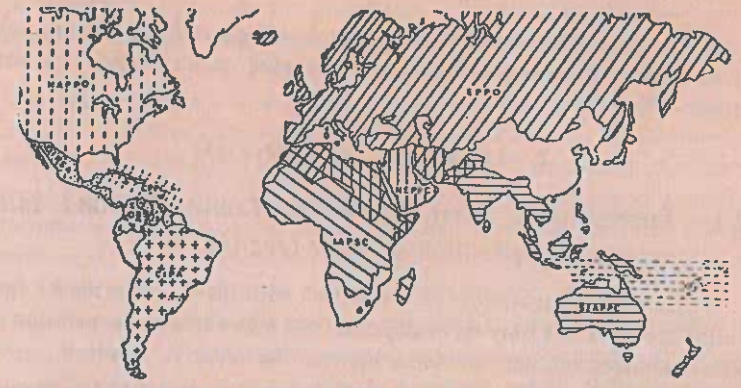


Fig. 1. Regional Intergovernmental Plant Protection Organisations

The primary tasks of the Regional Organisations are coordination within their area of legislation and regulations, including agreement on foreign pests and diseases to be considered as dangerous for introduction into the region in general, agreement on inspection procedures and on interpretation of national and international quarantine rules, such as interpretation of the wording of the FAO Phytosanitary Certificate "substantially free from injurious diseases and pests". Moreover, increasingly the Regional Organisations have taken up extension work to improve inspection practices in the member countries to ensure proper identification of pests and diseases and to ensure distinction between what is essential and what is not in drafting

regulations and in inspection practises. Examples of extension are the Handbook for Phytosanitary Inspectors in Africa[6] and the very informative EPPO Bulletin, especially the data sheets on quarantine organisms, provided with colour plates and detailed descriptions, being issued since 1977 by EPPO[11,12].

Particularly two of the Regional Organisations have taken up the problems of quarantine for seed, EPPO and IAPSC.

EPPO studied the related problems in a series of working parties in which representatives of some member countries of the organisation and of the International Seed Testing Association (ISTA) took part. The first meeting, held in 1954, dealt with seed-borne diseases of temperate crops and included a list of such diseases that were considered as dangerous to Europe, indicating adequate inspection procedures[7], the second working party similarly dealt with seed-borne diseases of tropical and subtropical crops[8], a third working party, in 1965, studied the inspection procedures for seed-borne diseases in general and the controversial problem of interpretation of rules in an attempt to standardise practises in issuance of Phytosanitary Certificates of seeds for export[9]. Recently, in December 1977, a working party, with representatives of ISTA and of the International Federation of Seed Trade (FIS), discussed the problem of obtaining more international uniformity in practices of issuing Phytosanitary Certificates and in assessing the economic importance of seed-borne diseases that may be considered for quarantine, and would be pertinent to International Seed Trade[13]. These meetings resulted in a series of important basic recommendations, contained in the above-mentioned reports on seed-borne diseases to be considered for quarantine from a European point of view, and on principles pertinent to inspection procedures and interpretation of test results. They stimulated and supported the efforts by the Plant Disease Committee of ISTA to build up international standard testing procedures, applicable in testing seed for quarantine, and to outline principles for quarantine policy[28,29]. The main concern and aim of these joint efforts towards providing a sound basis for testing seeds for quarantine was and still is : to *urge* governments to make quarantine regulations precise, factual and relevant, and to *provide* conditions for implementation of uniform procedures for detection of disease organisms in seeds, both for quality and quarantine. This would improve the situation which today is unsatisfactory to trade, to agriculture and to science. If proper rationalisation within this field could be attained, the risks of disease

losses would be reduced, and issuance of misleading certificates and much waste of time on empty formalities would be prevented[39].

In Africa much emphasis is being laid on improving provisions and inspection practices directed against introduction into the continent of dangerous foreign pathogens carried on seed. The Inter-African Phytosanitary Council is active in efforts to set up seed health testing laboratories at all the five regional quarantine centres being developed in Africa. A seed health testing laboratory has for some years been in operation at the Regional Post-Entry Plant Quarantine Station at Moor Plantation, Ibadan, Nigeria[8]. The five Regional Quarantine Stations are at various stages of development under an existing OAU/STRC Joint Project IP 29 : Ibadan in Nigeria (mentioned above), Muguga in Kenya, Ambatobe in Madagascar, Buumbire in Malawi, and Alexandria in Egypt. A joint-African project specific for establishing seed health testing laboratories for plant protection in Africa has been adopted by the Organisation of African Unity (OAU) and has been designated OAU/STRC Joint Project IP 34.

In a recent conference in Kathmandu, Nepal, held by the Plant Protection Committee for the South-East Asia and Pacific Region, problems of seed-borne diseases were taken up. The possibilities for pollen transmission of diseases were emphasised as pertinent to the increasing international exchange of pollen for plant breeding[15].

2.2. *General Regulations on Quarantine for Seed*

In summarising, quarantine regulation may include the following items[41] :

I. *Specifications*

Specification of pathogens and inspection procedures.

II. *Export control*

(Regulations related to the controlling agency of the exporting country).

(1) *Certification of complete freedom from infestation*, in the exporting country, or in a defined growing area of the seed crop concerned. For this purpose, issuance of an international standard Phytosanitary Certificate according to the provisions of the importing country.

(2) *Field inspection of the seed crop* by properly qualified personnel to ascertain freedom of infection (or accordance to requirements of tolerances)

in this crop. Issuance of an international standard Phytosanitary Certificate according to the provisions of the importing country.

(3) *Inspection of the seed lot* by drawing representative samples and testing these according to standard procedures approved by the importing country. Issuance of an international standard Phytosanitary Certificate according to the provisions of the importing country.

(4) *Seed treatment* whenever practicable. This provision should be considered with great reservation. It may be an additional precaution against quarantine object ; seed treatment should not be regarded as a complete safeguard in itself.

III. *Import control*

(Regulations related to the controlling agency of the importing country).

These may appropriately include the following provisions :

(1) *Embargo* : Total prohibition of importation of certain kinds of seed, or restricted, under the conditions of issuance of an *import permit* previously granted by the authorities of the importing country for individual consignments. The importer must request authorisation to import, usually from the Director of Agriculture or the Chief of the Plant Protection Service.

(2) *Inspection of the seed lot* by drawing representative samples and testing these in accordance with standard procedures, such as incubation tests.

(3) *Post-entry growing-on tests* of small seed consignments of special importance, cultivated from seed to mature seed-bearing plants in adequately equipped quarantine greenhouse units. The seed grown under these conditions should be tested and again found free from any infection. It may still be required to be inspected before ultimate release.

2.3. *Specification of Quarantine Objects in Regulations on Quarantine for Seed*

Most seed importing countries have some quarantine provisions on quarantine for seed. Many of these countries require a general plant health certificate for all or nearly all kinds of seed but do not specify any pathogen at all. Consequently it is left entirely to the discretion of the agency of the seed exporting country to decide which seed-borne pathogens should be

considered, and which inspection procedures should be used. As a consequence, the seed importing country has no guarantee that a seed lot, accompanied by a formally duly issued Phytosanitary Certificate, has been correctly inspected for the presence of dangerous parasitic fungi, bacteria, nematodes and viruses; indeed, most often no microbiological test at all has been carried out, before the Certificate is signed, and the seed may have been checked by visual inspection only, if at all! Needless to say that a Certificate, issued under such conditions is worth than useless, it is positively misleading.

There is a growing understanding that to make quarantine regulations workable, the provisions to be observed must be concrete, precise and specific. They must specify both pathogens to be considered, procedures to be used in inspecting the seed, and tolerances, if any, to be given. In most cases no tolerance is acceptable, and it should be stated that the seed consignment should be completely free from the pathogen specified.

The following section is a survey of seed-borne pathogens specified in the quarantine regulations of seed importing countries or otherwise specified by plant quarantine authorities.

The pathogens included in the list below are those specified in quarantine regulations and required to be observed in issuing plant health certificates. Pathogens given in quarantine regulations merely as reasons for embargo are not included.

No specifications on seed health testing methods are apparently given in the quarantine regulations referred to, at least not as far as seen in the abstracts available to the reviewer. Some countries, however, do prescribe *inspection of the seed crop concerned during active growth* for bacterial and virus diseases and certain other diseases, and in general this procedure must be followed even if, as in lettuce mosaic, reliable laboratory tests are available.

Indications of permissible tolerances are given in very few instances: New Zealand has indicated 1·5 per cent bean common mosaic as the maximum infection permissible. For pathogens specified in quarantine regulations without such indications, a tolerance of zero must be assumed. For those pathogens which are not specified in quarantine regulations it is left entirely to the certifying agency of the exporting country to decide which tolerances should be allowed under the conditions of the Rome Certificate, the text of which says that the inspected plant material has been

found "to be substantially free from injurious diseases and pests". The author has already fully discussed some basic principles to be considered in observing tolerances and interpretations of the words "substantially free" [33, 34].

On request lists of seed-borne quarantine objects have kindly been sent to the author in 1966 by the Plant Protection Services of Japan, New Zealand, Poland and USSR. They may not still be valid, but the pathogens concerned are included as examples of specifications that have been given apart from formal regulations.

Some references are made to "Danish Seed Exporter(s)"; they refer to Danish Seed Companies who have informed the Danish Plant Protection Service, Copenhagen, about specific requirements issued by the importing country although the specified pathogen may not have been named in the Plant Quarantine Regulations of the country concerned.

When there is *no reference indicated at countries listed*, the information was given in Neergaard 1970, where references are given, most of them are from Plant Quarantine Announcements in FAO Plant Protection Bulletin, Vol. 7-16, 1958-1968.

The survey below should by no means be regarded as complete or up-to-date. Many countries may in practice follow rules that imply observation of specified pathogens as quarantine objects, and not all countries have communicated to FAO their Quarantine Regulations for inclusion in the FAO Plant Protection Bulletin, and changes and amendments may be made any time. Dr. Robert P. Kahn of Plant Protection and Quarantine Programs, USDA Animal and Plant Health Inspection Service, has surveyed 125 countries for plant quarantine regulations in general, the survey being at present in the press. Out of these 125 countries 33 are included in the following below, and another three countries, Brunei, USSR and Zimbabwe (formerly Rhodesia), making a total of 36 countries.

The nomenclature used here is generally that used by the authorities supplying the information. Author names have been added.

Bacteria

Corynebacterium betae Leyworth and Dowson (*Beta*)
Greece, FAOPPB, 20 : 95.

Corynebacterium flaccumfaciens (Hedges) Dowson (*Phaseolus*)
Australia[82] ; Brunei, FAOPPB, 23 : 194 ; Chile ; East Africa, FAOPPB, 26 : 34 ; Greece, FAOPPB 20 : 95 ; Hungary, FAOPPB, 24 : 62 ("all bean bacteria") ; Israel ; Kenya (Danish seed exporter); Malta, FAOPPB, 25 : 138 ; New Zealand (Pl. Prot. Service, N. Z. 1966 pers. comm.) ; Senegal (also *Pisum*, Lathyrus, Vicia, Vigna, Dolchos lablab) ; Turkey, FAOPPB, 21 : 69.

Corynebacterium insidiosum (McCulloch) Jensen (*Medicago sativa*)
Australia[82] ; Greece, FAOPPB, 20 : 95 ; Hungary, FAOPPB, 24 : 61 ; Jersey, FAOPPB, 25 : 134 ; New Zealand (Pl. Prot. Service, N. Z. 1966) ; Nigeria ; Turkey, FAOPPB, 21 : 68 ; United Kingdom ; Zimbabwe.

Corynebacterium michiganense (E. F. Smith) Jensen (*Lycopersicon esculentum*)
Chile ; Greece, FAOPPB, 20 : 95 ; Israel ; Jersey, FAOPPB, 25 : 134 ; Kenya (Danish seed exporter) ; Malta, FAOPPB 25 : 138 ; Mozambique ; New Zealand (Pl. Prot. Service, N. Z. 1966 ; the seed must have been extracted by hydrochloric acid or treated by some other approved method) ; Nigeria, Poland (Pl. Prot. Service, Poland 1966) ; Romania (also Capsicum) ; Senegal ; South Africa, Republic of ; Turkey, FAOPPB, 21 : 70 ; United Kingdom ; Yugoslavia ; Zimbabwe.

Corynebacterium rathayi (E. F. Smith) Dowson (*Dactylis glomerata*)
Japan (Danish seed exporters).

Erwinia stewartii (E. F. Smith) Dye (*Zea mays*)
Australia[82] ; Brunei, FAOPPB, 23 : 195 ; Chile ; Greece, FAOPPB, 20 : 95 ; Hungary, FAOPPB, 24 : 61 ; Mozambique ; New Zealand (Pl. Prot. Service, N. Z. 1966) ; Nigeria ; Poland (Pl. Prot. Service, Poland 1966) ; Senegal ; Turkey, FAOPPB, 21 : 70 ; USSR ; Yugoslavia ; Zimbabwe.

Pseudomonas glycinea Coerper (*Glycine max*)
Greece, FAOPPB, 20 : 95 ; Hungary, FAOPPB, 24 : 62.

Pseudomonas mangiferae (*Doidge*) (*Mangifera*)
Brunei, FAOPPB, 23 : 195.

Pseudomonas Phaseolicola (Burkh.) Dowson (*Phaseolus*)
Greece, FAOPPB, 20 : 95 ; Hungary, FAOPPB, 24 : 62 ("all bean bacteria') ; New Zealand (Pl. Prot. Service, N. Z. 1966).

Pseudomonas pisi Sackett (*Pisum*)
Chile ; Greece, FAOPPB, 20 : 95 ; Israel, Jersey, FAOPPB, 25 : 134 ; New Zealand (Pl. Prot. Service, N. Z. 1966) ; Senegal ; South Africa, Republic of ; Turkey, FAOPPB, 21 : 69 ; United Kingdom ; Zimbabwe.

Pseudomonas punctulans (Bryan) Nagel (*Pseudomonas tomato*) (*Lycopersicon esculentum*)
New Zealand (Pl. Prot. Service, N. Z. 1966 ; extraction of seed by hydrochloric acid or by some other approved method required).

Pseudomonas tabaci (Wolf and Foster) F. L. Stevens (*Nicotiana*)
Australia[82] ; Senegal.

Pseudomonas tomato (Okabe) Alstatt, see *Pseudomonas punctulans*.

Pseudomonas viridiflava (Burkholder) Clara (*Phaseolus lunatus*)
Australia[82].

Xanthomonas alfalfae (Riker, F. R. Jones and Davis) Dowson (*Medicago*)
Australia[82].

Xanthomonas campestris (Pamm.) Dowson (*Brassica oleracea*)
Senegal (seed treatment required) ; South Africa, Republic of (Danish seed exporter 1966) ; Turkey, FAOPPB, 21 : 70.

Xanthomonas citri (Hasse) Dowson (*Citrus*)
Greece, FAOPPB, 20 : 95 ; Senegal (seed treatment by hydrogen peroxide required).

Xanthomonas malvacearum (E. F. Smith) Dowson (*Gossypium*)
Greece, FAOPPB, 20 : 95.

Xanthomonas oryzae (Uyeda and Ishiama) Dowson (*Oryza sativa*)
Australia[82] ; Chile ; Hungary, FAOPPB, 24 : 61 ; Turkey, FAOPPB, 21 : 70.

Note : EPPO requests *zero tolerance* for all member countries, FAOPPB, 22 : 143.

Xanthomonas aryzicola Fang *et al.* (*Oryza sativa*)
Note : EPPO requests *zero tolerance* for all member countries, FAOPPB 22 : 143.

Xanthomonas phaseoli (E. F. Smith) Dowson (*Phaseolus*)
Chile ; Greece, FAOPPB, 20 : 95 ; Hungary, FAOPPB, 24 : 62 ("all bean bacteria") ; Israel ; Malta, FAOPPB, 25 : 138 ; New Zealand (Pl. Prot. Service, N. Z. 1966) ; Turkey, FAOPPB, 21 : 95.

Xanthomonas phaseoli (E. F. Smith) Dowson var. *fuscans* (Burkh.) Starr and Burkh. (*Phaseolus*)
Australia[82].

Xanthomonas translucens (L. R. Jones, A. G. Johnson and Reddy) Dowson (*Hordeum, Triticum, Oryza*)
Greece, FAOPPB, 20 : 95.

Xanthomonas vesicatoria (Doidge) Dowson (*Capsicum, Lycopersicon esculentum*)
Chile ; Greece, FAOPPB, 20 : 95 ; Israel ; Malta, FAOPPB, 25 : 138 ; Turkey, FAOPPB, 21 : 70.

Xanthomonas vignicola Burkh. (*Dolichos, Vigna sinensis*)
Australia[82].

Fungi

Alternaria brassicae (Berk.) Sacc. (*Brassica oleracea*)
Turkey, FAOPPB, **21** : 70.

Alternaria brassicicola (Schw.) Wilts. (*Brassica oleracea*)
Turkey, FAOPPB, **21** : 70.

Alternaria carthami Chowdhury (*Carthamus tinctorius*)
Australia[32].

Alternaria radicina Meier, Drechsl. and Eddy (*Apium graveolens, Daucus carota*)
Greece, FAOPPB, **20** : 95 ; South Africa, Republic of (Danish seed exporter 1966).

Alternaria raphani Groves and Skolko (*Raphanus sativus*)
South Africa, Republic of (Danish seed exporter 1966).

Alternaria spp. (many hosts)
Japan (Pl. Prot. Service, Japan 1966).

Alternaria tomato (*Lycopersicon*)
Australia[32].

Ascochyta chlorospora Speg., perf. st. *Gnomonia circumscissa* McAlp. (*Prunus amagdalus*)
Italy.

Ascochyta desmazierii (*Gramineae*)
Spain (Danish seed exporters).

Ascochyta graminicola Sacc, var. *brachypodii* (*Gramineae*)
Spain (Danish seed exporters).

Ceratocystis fagacearum (Bretz.) Hunt (*Endoconidiophora fagacearum* Bretz.) (*Fagus* spp., *Nothofagus* spp.)
New Zealand (seed treatment required ; Pl. Prot. Service, N.Z. 1966).

Ceratocystis ulmi (Buism.) Cl. Moreau (*Ulmus*)
Senegal.

Cercospora kikuchii (Mats. and Tomoy.) M. W. Gardner (*Glycine*)
Australia[32].

Cercospora puerariicola Yamam. (*Pueraria*)
Australia[32].

Chalara quercina B. W. Henry (*Castanea molissima, Quercus*)
Czechoslovakia.

Claviceps microcephala (Wallr.) Tul., see *C. purpurea*

Claviceps purpurea (Fr.) Tul. (*Gramineae*), including *C. microcephala*
Japan, slight infestation allowed, degree of infestation must be indicated on the certificate (Pl. Prot. Service, Japan 1966) ; Zimbabwe.

Claviceps sulcata (*Brachiaria*)
Australia[32].

Cochliobolus heterostrophus (Drechsl.) Drechsl., see *Drechslera maydis*

Cochliobolus miyabeanus (Ito and Kuribay.) Drechsl. ex Dast., see *Drechslera oryzae*

Colletotrichum cajani Rangel (*Cajanus cajan*)
Australia[32].

Colletotrichum coffeanum Noack (*Coffea*)
Brunei, FAOPPB, **23** : 195.

Colletotrichum dematium (Pers. ex Fr.) Grove (syn. *C. indicum* Dast.) (*Gossypium, Phaseolus lunatus*)
Australia[32] ; USSR.

Colletotrichum glycines Hori ex Hemmi (*Glycine*)
Australia[32].

Colletotrichum gossypii South, see also perf. st. *Glomerella gossypii* Edg. (Gossypium)
USSR.

Colletotrichum graminicola (Ces.) Wils. (*Gramineae*)
Japan, slight infestation allowed, degree of infestation must be indicated on the certificate (Pl. Prot. Service, Japan 1966) ; Spain (Danish seed exporter).

Colletotrichum higginsianum Sacc. (*Raphanus sativus*)
South Africa, Republic of (Danish seed exporter 1966).

Colletotrichum lagenarium (Pass.) Ell. and Halst. (*Cucumis sativus*)
Greece, FAOPPB, **20** : 95.

Colletotrichum lindemuthianum (Sacc. and Magn.) Bri. and Cav. (*Phaseolus*)
Israel ; Malta, FAOPPB, **25** : 138 ; New Zealand (Pl. Prot. Service, N. Z. 1966).

Colletotrichum lini (Westerdijk) Tochinai (*C. linicola*) (*Linum*)
Czechoslovakia (Danish seed exporter).

Colletotrichum linicola Pethybr. and Laff., see *Colletotrichum lini*

Colletotrichum tabacum Böning (Nicotiana)
Australia[32] ; Senegal ; Zimbabwe.

Diaporthe phaseolorum (Cke. et Ellis) Sacc. var. *batatis* (Harter et Field) Wehm. (*Glycine*)
Australia[32].

Diaporthe phaseolorum (Cke. and Ell.) Sacc. var. *sojae* (Lehman) Wehm. (*Glycine*)
Australia[32].

Didymella lycopersici Kleb., imperf. st. *Diplodina lycopersici* Hollós (*Lycopersicon esculentum*)
Australia[32] ; Greece, FAOPPB, **20** : 95 ; New Zealand, the seed must have been extracted by hydrochloric acid or treated by other approved method, (Pl. Prot. Service, N. Z. 1966).

Diplodia maydis (Berk.) Sacc. (*D. zeae*) (*Zea mays*)
Hungary, FAOPPB, **24** : 61 ; Poland (Pl. Prot. Service, Poland 1966) ; Yugoslavia.

Diplodia sp.
Japan (Pl. Prot. Service, Japan 1966).

Diplodia zeae (Schw.) Lév, see *Diplodia maydis*

Dothidella ulei P. Henn. (*Hevea*)
Malaysia, FAOPPB, **23** : 163 ; (Netherlands) New Guinea ; Nigeria.

Drechslera avenae (Eidam) Scharif (*Avena*)
Greece, FAOPPB, **20** : 95.

Drechslera maydis (Nisik.) Subram. and Jain, perf. st. *Cochliobolus heterostrophus* (*Zea mays*)
Greece, FAOPPB, **20** : 95 ; Hungary, FAOPPB, **24** : 61.

Drechslera nodulosa (Berk. and Curt.) Subram. and Jain (*Triticum*)
Australia[82].

Drechslera Oryzae (Breda de Haan) Subram. and Jain, perf. st. *Cochliobolus miyabeanus* (*Oryza sativa*)
Chile.

Drechslera sigmoidea (Cav.) var. *irregulare* Tull. (*Oryza*)
Australia[82].

Drechslera spp. (cereals)
Zimbabwe.

Drechslera tritici-vulgaris (Nisik.) Ito (*Triticum*)
Australia[82].

Drechslera turcica (Pass.) Subram. and Jain (*Zea mays*)
Hungary, FAOPPB, **24** : 61.

Drechslera victoriae (Meehan and Murphy) Subram. and Jain (*Gramineae*)
Australia[82] ; Turkey, FAOPPB, **21** : 70.

Drechslera vignicola (*Glycine*)
Australia[82].

Endoconidiophora fagacearum Bretz., see *Ceratocystis fagacearum*

Endothia parasitica (Murr.) And. and And. (*Castanea*)
Czechoslovakia ; Hungary, FAOPPB, **24** : 61 ; India[1], Nigeria ; Turkey, FAOPPB, **21** : 70 ; Yugoslavia ; Zimbabwe.

Entyloma oryzae H. and P. Syd. (*Oryza sativa*)
Mexico.

Epicoccum sp.
Japan (Pl. Prot. Service, Japan 1966).

Fusarium oxysporum Schlecht. ex Fr. f. sp. *carthami* Klis. and Houston (*Carthamus tinctorius*)
Australia[82].

Fusarium oxysporum Schlecht. ex Fr. f. sp. *conglutinans* (Wr.) Snyd. and Hansen (*Brassica oleracea*)
Bulgaria (Danish seed exporter) ; Yugoslavia (*Fusarium oxysporum*).

Fusarium oxysporum Schlecht. ex Fr. f. sp. *lini* (Bolley) Snyd. and Hans. (*F. lini* Bolley) (*Linum*)
Chile ; Peru.

Fusarium oxysporum Schlecht. ex Fr.f.sp. *vasinfectum* (Akt.) Snyd. and Hans. (*Gossypium*)
Australia[82] ; USSR.

Gloeotinia temulenta (Prill. and Delacr.) Wilson, Noble and Gray (*Ciboria temulenta*) (*Lolium perenne*)
Israel ; Spain (Danish seed exporters).

Glomerella gossypii Edg., see also imperf. st. *Colletotrichum gossypii* (*Gossypium*)
Greece, FAOPPB, **20** : 95.

Gnonomia circumscissa McAlp., see *Ascochyta chlorospora.*

Helminthosporium, see *Drechslera.*

Hemileia vastatrix B. and Br. (*Coffea*)
Honduras.

Kabatiella zeae (*Zea mays*)
Hungary, FAOPPB, **24** : 62.

Leptosphaeria maculans (Desm.) Ces. and de Not., see *Phoma lingam.*

Leptosphaeria sp.
Japan (Pl. Prot. Service, Japan 1966).

Marasmius perniciosus Stahel (*Theobroma cacao*)
India[1].

Monilia roreri Cif. (*Theobroma cacao*)
India[1].

Mycosphaerella linorum (Wr.). Garcia Rada (*Sphaerella l.*), imperf. st. *Septoria linicola* (*Phlyctaena linorum*) (*Linum*)
Czechoslovakia ; Hungary, FAOPPB, **24** : 62 ; Israel ; Poland (Pl. Prot. Service, Poland 1966) ; USSR.

Mycosphaerella puerariicola Weim. and Luttr. (*Pueraria*)
Australia[82].

Neovossia indica (Mitra) Mundkur, see *Tilletia indica*

Oospora oryzatorum Sacc. (*Oryza sativa*)
Mexico.

Ophiobolus graminis (Sacc.) Sacc. (*Triticum*)
Mexico.

Periconia circinata (Mangin) Sacc. (*Sorghum*)
Australia[82].

Peronospora manshurica (Naum.) Syd. (*Glycine max*)
Australia[32] ; Hungary, FAOPPB, 24 : 62.

Peronospora tabacina Adam (*Nicotiana*)
Brunei, FAOPPB, 23 : 197 ; Greece, FAOPPB, 20 : 95 ; India[1] ; Nigeria ; Poland (Pl. Prot. Service, Poland 1966) ; Senegal ; Turkey ; Zimbabwe.

Phaeoseptoria sp.
Japan (Pl. Prot. Service, Japan 1966).

Phialea temulenta Prill. and Delacr., see *Gloeotinia temulenta*

Phoma betae Frank (*Beta*)
Greece, FAOPPB, 20 : 95.

Phoma glumarum Ell. and Tr. (*Oryza sativa*)
Mexico.

Phoma herbarum Westend. (*Medicago sativa*)
Greece, FAOPPB, 20 : 95.

Phoma lingam (Tode ex Fr.) Desm. (*Plenodomus lingam* (Tode ex Fr.) (Höhnel) (*Brassica oleracea*)
Greece, FAOPPB, 20 : 95 ; Pakistan (Danish seed exporter 1966) ; Romania ; Yugoslavia.

Phoma rostrupii Sacc. (*Daucus carota*)
Bulgaria (Danish seed exporter).

Phyllosticta maydis Arny and Nelson (*Zea mays*)
Hungary, FAOPPB, 24 : 61.

Phyllosticta spp. (many hosts)
Japan (Pl. Prot. Service, Japan 1966).

Physalospora perseae Doidge (*Mangifera*)
Brunei, FAOPPB, 23 : 195.

Physalospora rhodina (Berk. & Curt.) Cke. (*Arachis hypogaea*)
Australia[32].

Physoderma maydis (Miyabe) Miyabe (*P. zeae-maydis*) (*Zea mays*)
Australia[32] ; New Zealand, the seed must have been treated (Pl. Prot. Service, N. Z. 1966).

Physoderma zeae-maydis Shaw, see *Physoderma maydis*

Phytophthora capsici Leonian (*Capsicum* spp.)
Brunei, FAOPPB, 23 : 196 ; Senegal.

Phytophthora phaseoli Thaxter (*Phaseolus lunatus*)
Australia[32].

Plasmopara halstedii (Farl.) Berl. and de Toni (*Helianthus annuus*)
Australia[32] ; Hungary, FAOPPB, 24 : 62 ; Turkey, FAOPPB, 21 : 70.

Pleospora bjoerlingii Byford, see *Phoma betae*

Polyspora lini Lafferty (*Linum*)
Czechoslovakia ; Hungary, FAOPPB, 24 : 62.

Puccinia arachidis Speg. (*Arachis*)
Australia[32] ; India[1] ; Senegal.

Puccinia glumarum (Schm.) Erikss. and E. Henn. (*Hordeum*)
Australia[32].

Pyrenophora avenae Ito and Kuribay., see *Drechslera avenae*

Pyricularia oryzae Cav. (*Oryza sativa*)
Chile ; Greece, FAOPPB, 20 : 95.

Pythium arrhenomanes (*Sorghum*)
Australia[32].

Rhabdocline pseudotsugae Syd. (*Pseudotsuga*)
Yugoslavia.

Sclerospora macrospora Sacc. (*Oryza sativa, Zea mays*)
Hungary, FAOPPB, 24 : 61 ; Mexico.

Sclerospora philippinensis Weston (*Zea*)
Australia[32].

Sclerospora sorghi Weston and Uppal (*Sorghum*)
Australia[32].

Sclerospora spp. (*Gramineae*)
Mozambique.

Sclerotinia sclerotiorum (Lib.) de Bary (many hosts)
Japan, slight infestation allowed, degree of infestation must be indicated on the certificate (Pl. Prot. Service, Japan 1966).

Sclerotinia trifoliorum Erikss. (*Trifolium*)
Japan, slight infestation allowed, degree of infestation must be indicated on the certificate (Pl. Prot. Service, Japan 1966).

Septoria glycines Hemmi (*Glycine*)
Australia[32].

Septoria linicola (Speg.) Garcia Rada, see *Mycosphaerella linorum*

Septoria nodorum Berk. (*Triticum*)
Hungary, FAOPPB, 24 : 62.

Septoria secales Prill. and Delacr. (*Secale*)
Australia[32].

Septoria spp. (many hosts)
Japan (Pl. Prot. Service, Japan 1966).

Sphaceloma arachidis (*Arachis*)
India[1].

Sphaceloma pamparum (*Secale*)
Australia[32].

Sphacelotheca cruenta (Kühn) Potter (*Sorghum*)
Australia[32].

Sphacelotheca sorghi (Link.) Clint. (*Sorghum*)
Japan, slight infestation allowed, degree of infestation must be indicated on the certificate (Pl. Prot. Service, Japan 1966).

Stemphylium radicinum (Meier, Drechsl. and Eddy) Neerg., see *Alternaria radicina*

Tilletia barclayana (Bref.) Sacc. and Syd. (*Brachiaria, Oryza, Panicum, Pennisetum*)
Australia[32].

Tilletia brevifaciens G. W. Fisch., see *T. contraversa.*

Tilletia contraversa Kühn (*T. brevifaciens*) (*Agropyron* sp., *Secale, Triticum*)
Australia[32] ; Czechoslovakia ; Greece, FAOPPB, **20** : 95 ; Japan, slight infestation allowed, degree of infestation must be indicated on the certificate (Pl. Prot. Service, Japan 1966) ; Poland (Pl. Prot. Service, Poland 1966) ; Turkey, FAOPPB, **21** : 70.

Tilletia fusca Ell. and Ev. (*Gramineae*)
Japan (Pl. Prot. Service, Japan 1966).

Tilletia indica Mitra (*Neovossia indica* (Mitra) Mundkur) (*Triticum*)
Hungary, FAOPPB, **24** : 61 ; USSR.

Tilletia pallida G. W. Fisch. (*Agrostis*)
Japan (Pl. Prot. Service, Japan 1966).

Tilletia pancicii Bubak and Ranoj. (*Hordeum*)
Australia[32].

Tilletia secalis (Corda) Kühn (*Secale*)
Australia[32] ; Turkey, FAOPPB, **21** : 70.

Tilletia texana Long (*Hordeum*)
Australia[32].

Tilletia triticoides Savul. (*Triticum*)
Australia[32].

Tolyposporium ehrenbergii (Kühn) Patouillard (*Sorghum*)
Australia[32].

Trachysphaera fructigena Tab. and Bunt. (*Theobroma cacao*)
India[1].

Trichoconiella padwickii (Ganguly) Jain (*Oryza sativa*)
Australia[32] ; Greece, FAOPPB, **20** : 95 ; Japan (Pl. Prot. Service, Japan 1966).

Trichoconis padwickii, see *Trichoconiella padwickii.*

Urocystis cepulae Frost (*Allium*)
Hungary, FAOPPB, **24** : 62 ; Senegal ; Yugoslavia.

Urocystis tritici Koern. (*Tuburcinia tritici*)(*Secale, Triticum*)
Greece, FAOPPB, **20** : 95 ; Israel ; Malta, FAOPPB, **25** : 137 ; Mexico ; Turkey, FAOPPB, **21** : 70.

Urophlyctis alfalfae (Lagerh.) Magn. (*Medicago sativa*)
Nigeria ; Senegal ; Zimbabwe.

Ustilaginoidea virens (Cke.) Tak. (*Brachiaria*)
Australia[32].

Ustilago bullata Berk. (*Bromus*)
Japan, slight infestation allowed, degree of infestation must be indicated on the certificate (Pl. Prot. Service, Japan 1966).

Ustilago crameri Koern. (*Setaria*)
Australia[32].

Ustilago maydis (DC.) Corda (*U. zeae*)(*Zea mays*)
Australia (Morschel 1970) ; New Zealand, the seed must have been treated (Pl. Prot. Service, N. Z. 1966).

Ustilago neglecta Niessl (*Setaria*)
Australia[32].

Ustilago nigra Tapke (*Hordeum*)
Australia[32].

Ustilago spegazzini Hirschhorn var. *agrestis* (Syd.) Fischer and Hirschhorn (*Gramineae*)
Japan (Pl. Prot. Service, Japan 1966).

Ustilago striiformis (Westend.) Niessl (*Gramineae*)
Japan, slight infestation allowed, degree of the infestation must be indicated on the certificate (Pl. Prot. Service, Japan 1966).

Ustilago syntherismae (Schw.) Peck (*Digitaria*)
Australia[32].

Ustilago trebouxii H. and P. Syd.
Japan (Pl. Prot. Service, Japan 1966).

Verticillium alboatrum Reinke and Berth. (*Medicago sativa*)
New Zealand (Pl. Prot. Service, N. Z. 1966) ; USA (the seed must have been treated).

Viruses

Alfalfa dwarf (Medicago virus 3) (*Medicago sativa*)
Nigeria ; Senegal ; Zimbabwe.

Barley stripe mosaic (*Hordeum, Secale, Triticum*)
Australia[32] ; Greece, FAOPPB, **20** : 93 ; Malta, FAOPPB, **25** : 138 ; Turkey, FAOPPB, **21** : 70.

Bean black legginess (*Phaseolus*)
Australia[32].

Bean common mosaic (*Phaseolus*)
Australia ; Bulgaria (Danish seed exporter) ; Greece, FAOPPB, **20** : 93 ; New Zealand (Pl. Prot. Service, N. Z. 1966).

Bud blight (*Glycine*)
Australia[32].

Citrus psorosis (*Citrus*)
Greece, FAOPPB, 20 : 93.

Citrus xylosporosis (*Citrus*)
Greece, FAOPPB, 20 : 93.

Cowpea mosaic (*Vigna sinensis*)
Australia[32].

Cucumber green mottle mosaic (Cucumis virus 2)(*Cucumis*)
Bulgaria (Danish seed exporter).

Exocortis virus
Greece, FAOPPB, 20 : 93.

Groundnut green rosette (*Arachis*)
Australia[32].

Groundnut mosaic (*Arachis*)
Australia[32].

Groundnut veinbanding (*Arachis*)
Australia[32].

Groundnut witches broom (*Arachis*)
Australia[32].

Groundnut stunt (*Arachis*)
Australia[32].

Hoja blanca du Riz (*Oryza sativa*)
Greece, FAOPPB, 20 : 95 ; Mexico, FAOPPB, 21 : 70.

Lettuce mosaic (Lactuca)
Greece, FAOPPB, 20 : 93 ; Israel ; Jersey, FAOPPB, 25 : 134 ; Malta, FAOPPB, 25 : 138 ; United Kingdom ; Yugoslavia (Danish seed exporter).

Maize dwarf mosaic (Zea)
Australia[32].

Medicago virus 3, see Alfalfa dwarf

Prunus necrotic ringspot
Greece, FAOPPB, 20 : 93.

Southern bean mosaic (Phaseolus ; *Vigna sinensis*)
Australia[32].

Soybean mosaic
Australia[32] ; Turkey, FAOPPB, 21 : 70.

Stonefruit viruses (*Prunus*)
Denmark (Order of the min. Agric. 1968) ; Israel ; Malta, FAOPPB, 25 : 138 ; Yugoslavia (specification of viruses given).

Sunflower mosaic (*Helianthus annuus*)
Australia[32] ; New Zealand (Pl. Prot. Service, N. Z. 1966) ; Senegal (virus diseases in general) ; South Africa, Republic of (Danish seed exporter) ; Turkey, FAOPPB, 21 : 70 ; Zimbabwe.

Tobacco mosaic
Greece, FAOPPB, 20 : 93.

Tobacco necrosis
Greece, FAOPPB, 20 : 93.

Viruses (*Hicoria pecan*)
Israel.

Viruses (*Glycine max*)
Hungary, FAOPPB, 24 : 62.

Nematodes

Anguina agrostis (Steinbuch) Filipjev, see *Tylenchus agrostis*

Anguina tritici (Steinbuch) Chitwood (*Triticum*)
Turkey, FAOPPB, 21 : 70.

Aphelenchoides besseyi Christis (*Oryza*)
Nigeria.

Cyst-forming nematodes (Heterodera spp.)
Senegal.

Ditylenchus angustus (Butler) Filipjev (*Oryza sativa*)
Australia[32] ; Nigeria ; USAR (Pl.Prot. Service, USSR 1966)

Ditylenchus destructor Thorne (*Allium*)
Turkey, FAOPPB, 21 : 70.

Ditylenchus dipsaci (Kühn) Filipjev
Ireland (*Allium*) ; United Kingdom, fumigation required (*Medicago sativa*).

Heterodera schachtii A. Schmidt (*Beta vulgaris*)
Romania (Danish seed exporter).

Rhadinaphelenchus cocophilus (Cobb) Goodey (*Cocos*)
Equador, FAOPPB, 18 : 121.

Tylenchus agrostis (*Gramineae*)
Japan (Pl. Prot. Service, Japan 1966).

3. Techniques

The first book on seed health testing methods was published by Doyer (1938). Since 1958, a Handbook on Seed Health Testing is being published *seriatim* by the International Seed Testing Association (ISTA), previously comprising four series : (1) General topics, (2) Specific crops or specific group of pathogens, (3) Working sheets each dealing with specific tests for one organism (edited by Mary Noble), based on cooperative work by the ISTA Plant Disease Committee, and (4) Seed-borne fungi (mostly saprophytes). Descriptions and illustrations of 77 species, mainly on oats and flax.

The new edition of the Handbook, now in preparation, will consist of the following parts :

Section 1.1. Annotated list of seed-borne diseases[46].

Section 1.2. Introduction to methods of seed health testing (being revised by J. de Tempe and J. Binnerts).

Section 2. Working sheets, each dealing with one pathogen on one host (being edited by Johs. Jorgensen).

Other recent surveys on seed health testing methods have been published by Naumova (1972) who previously published three editions of her handbook in Russian (the last one in 1970) and Neergaard[40,41].

A brief summary of the most important types of seed health testing procedures in use is given in Table 1.

Selection of procedures for export inspection of seed for quarantine depends, of course, on the regulation of the importing country. However, few, if any, countries include in their quarantine provisions specification on seed testing procedures, in fact, field inspection of seed crops seems to be the only specific procedure considered in the regulations only of a few countries such as Greece and Yugoslavia, especially as precaution against viruses. This lack of specification means that it is left to the certifying agency of the exporting country which testing procedures should be used. If, however, it is clearly stated on the report or certificate then its worth should be obvious to qualified personnel.

Neergaard[41] has discussed methods applicable for testing seed especially for quarantine with regard to fungi and bacteria. The "generalised tests", such as the blotter test and the agar plate test, reveal a wide range of Fungi Imperfecti. The "specialised tests" are specific and used for detection and identification of individual pathogens, the embryo count procedure for loose smut of wheat and barley, the serological tests for specific viruses and bacteria, the phage-plague tests for specific bacteria and the indicator-plant test to reveal pathogenic races of fungi and bacteria, or special strains of a virus. Selective agar media may be used in testing seed for a special group of fungi, such as for *Fusarium*, thus providing a sensitive procedure suitable for quarantine testing.

Table 1

Important seed health testing methods (Neergaard 1980)

Method	Technique	Application
Blotter method (with modifications: 2, 4-D method and deep-freezing blotter method)	*Fungi* Standard : Germination, moist paper, 7 days, 20°, 12 h NUV/12 h darkness	Many fungi, primarily Fungi Imperfecti on many hosts
Agar method PDA or MA	5-7 days, 20°, 12 h NUV/ 12 h darkness	
Selective media	5-7 days, 20°, 12 h NUV/ 12 h darkness	e.g., *Fusarium*
Brönnimann's method	7 days, 15°, PDA permanent standard light	*Septoria nodorum*
Fluorescence method	4-5 days, 20-25°, oxgall-agar light conditions less important	*Septoria nodorum*
Mangan's method	7 days, 20°, water agar ("holdfast structures")	*Phoma betae*
Hiltner symptom test	germination in medium of crushed bricks 14 days, 10-12°, daylight/darkness	Germination energy (vigour test) seedling diseases, e.g., Fusarium on cereal ; evaluation of the effect of fungicide treatment
Sand symptom test	germination in sand, 2-4 weeks, 10°, daylight/ darkness	
Standard soil symptom test	germination in soil, 2 weeks, 10°, daylight/darkness	
Test-tube symptom test	germination on water agar 10-12 days, 20°, daylight, darkness	Demonstration of transmission of infection from seed to seedling; quarantine control
	Bacteria	
Direct Isolation		Potentially applicable for all bacteria
Indicator method		*Xanthomonas campestris*

Table 1 (Concld.)

Method	Technique	Application
Bacteriophage method	*Bacteria*	Depends on the specificity of the bacteriophage concerned. Used for *Corynebacterium michiganense*, *Xanthomonas vesicatoria*, *Xanthomonas phaseoli*
Serological methods:		
Agglutination method		Potentially applicable for all bacteria (dependent on antiserum-specificity)
A-G-double-diffusion method		
Precipitin method		
Immunofluorescence method		
	Viruses	
Indicator method		Standard technique for identification
Symptom method		Bean common mosaic virus, lettuce mosaic virus
Serological methods:		
Precipitin method (capillary tube)		Arabis mosaic virus, bean yellow mosaic virus, soybean mosaic virus
A-G-double-diffusion method (capillary tube)		Tobacco ringspot virus, tomato ringspot virus
Latex method (capillary tube)		Virus with thread- and rod-shaped particles and isometric virus, Tobacco mosaic virus, Soybean mosaic virus
Enzyme-linked immuno-sorbent assay		Pea seed-borne mosaic virus, prune dwarf virus, soybean mosaic virus, tobacco ringspot virus

4. Quarantine for Seed in Practice

4.1 *Issuance of Rome Certificate*

As it appears from previous discussions in this review, the basis for issuance of the Rome Certificate for seed consignments for export varies exceedingly from country to country. In some seed exporting countries regular seed health testing is carried out using incubation procedures, in others only quick superficial procedures, such as dry seed inspection under a hand lens or stereomicroscope are used. In many other countries, unfortunately, apparently no inspection of the seed is done and certificates are issued upon presentation of seed for export.

In some countries seed accompanied by the Rome Certificate may be checked on entry to the importing country before it is released. In Israel a considerable number of seed consignments accompanied by Certificates issued in European countries or in USA have been intercepted and rejected because of the presence of dangerous pathogens not present in Israel.

Statistical information on numbers of Rome Certificates issued for seed in different countries is very scattered. In 1964–1965 the Plant Protection Service in Copenhagen issued 1603 Rome Certificates for about 12000 seed lots exported to 80 countries[40], in 1976 the Service issued 1786 Certificates for an unreported number of seed lots exported to 106 countries, and in 1977, 1707 Certificates for an unreported number of seed lots exported to 105 countries[44,45]. The United Kingdom and the Netherlands are other European countries with a considerable seed export and charged with similar responsibilities in issuance of seed health certificates.

4.2 *Interceptions*

Neergaard (1979) discussed 30 examples of interceptions of seed-borne pathogens encountered on seed in 6 importing countries, the seeds being imported from more than 20 countries. The intercepted pathogens include 3 species of bacteria, 2 species of rust fungi, 3 species of smut fungi, 16 species of Fungi Imperfecti / Ascomycetes, 2 species of Oomycetes (downy mildews), 2 species of nematodes, and 2 viruses.

For a number of years Dr. E. E. Leppik was in-charge of screening seeds for plant breeding at the Regional Plant Introduction Station, Ames, Iowa. Most of the screening was based on growing-on tests in greenhouse

condition. Based on 4500 plant introductions he found that an average of 2 per cent were infected with foreign pathogens of potential danger to the agriculture of America. The highest percentage was detected in large-seeded legumes, about 5 per cent of the consignments being infected. Introduced pea seed produced 17 per cent infection, *Lathyrus* up to 21 per cent infected plants, sunflower and sesame 10 per cent[28]. The post-entry surveillance of introduced plants and seed, and production of disease-free seed are performed by several regional plant introduction stations in the United States. The seeds are tested, stored and multiplied in greenhouses or in the field[29]. Leppik[30] has pointed out the paramount importance of gene centres as centres of diversity not only of cultivated plants but also of their pathogens.

Recent investigations on plant breeding material of pea and soybean confirm the danger of uncritical use of seed introductions in breeding programmes, and the need for well-organised quarantine in this connection. Zimmer and Ali-Khan[50] recorded, in testing the seed only for pea seed-borne mosaic virus, that about 9% of 1235 breeding lines of field peas, *Pisum arvense*, were infected, 15 out of 100 commercial seed samples had 1–2% infected seed. In USA Hampton and Braverman[21] tested 1835 plant introduction lines of pea serologically for presence of pea seed-borne mosaic virus only and found 23% of the lines to be infected. The incidence in lines introduced in recent years was much higher than in lines from previous years, the increase being attributable principally to lines imported from Northern India. Virus-free stocks of infected lines are being produced. Goodman *et al.*[20] tested serologically for soybean mosaic virus only 897 accessions of soybean germ plasm collections. Out of 497 lines of temperate maturity groups only 83 had no seed transmission when a sample of 200 seeds from each was planted. In the remaining 414 lines infection ranged from 1 to 36%. Of 400 tropical germ plasm lines similarly tested 294 lines were infected, infection ranging from 1 to over 10%.

There is now a growing appreciation of the vital importance of the spread of diseases by international exchange of plant genetic resources. The International Board for Genetic Resources has recently published a book on this issue[28] which deals with methods for detection of seed-borne fungi and bacteria (Neergaard), seed-borne viruses (Bos), principles and methods of quarantine (Kahn), quarantine policy for seed in transfer of genetic resources (Neergaard), post-entry and international quarantine stations (Berg) and phytosanitary regulations and transfer of genetic resources

(Mathys). The book is fundamental for any further progress in quarantine for seed in general and for genetic resources in particular.

In 1974 and 1975 cacao seeds were imported into the Dominican Republic from Trinidad. The seeds developed 70 per cent seedlings infected by "witches broom", *Marasmius perniciosus*. All seed consignments and seedlings were destroyed since the pathogen is not known to be present in the Dominican Republic (Manuel Mendez, see Neergaard[41]).

From India recent examples of interceptions of parasitic fungi in a range of crop seeds have been given by Lambat[26]. Most of the interceptions were based on incubation procedures without which the pathogen would not have been detected.

Encountered by incubation methods : *Fusarium nivale* in wheat, *Diplodia zeae* and *D. macrospora* in maize ; *Colletotrichum glycine* and *Diaporthe phaseolorum* in soybean ; and *Colletotrichum* sp. and Phomopsis sp. in cotton.

By other procedures : *Tilletia caries* and *T. foetida* in wheat, and *T. barclayana* (*T. horrida*) in rice ; *Claviceps purpurea* in *Triticales*, wheat, oats, barley and a number of grasses ; and *Puccinia helianthi* in sunflowers. Although most of these pathogens are already present in India, there is great risk of introducing new pathogenic races attacking resistant Indian cultivars of the hosts concerned. For example, *Tilletia caries* and *T. foetida* have a total of about 30 specific pathogenic races according to international standard identificaton ; many of these races may not be present in the country.

The United States Animal and Plant Health Inspection Service has issued a series of "List of Intercepted Plant Pests" [2,16,17,18,19] ranging from insects, mites, mollusks, nematodes to bacteria and fungi but not including viruses. Total number of tabulated interceptions due to pests and diseases on plant materials during the 4-year period recently published[2] amounts to more than 31000. The seed-borne fungi and nematodes intercepted are listed in Table 2. All of these fungi are detectable by hand lens or by stereomicroscope (usually up to × 60 magnification). This implies that apparently no provision has been made to routine testing by incubation or growing-on in sharp contrast to the procedures and interceptions reported by Leppik also working in the United States.

Table 2

Seed-borne fungi and nematodes intercepted in USA, July 1, 1973 to September 30, 1978

(Animal and Plant Health Inspection Service, United States, Department of Agriculture, 1979 : Anonymous 1979a, 1980)

Pathogen	Host	Country of Origin
Ascochyta sp.	Pisum sativum	Not specified
Cintractia caricis (Pers.) Magn.	Carex sp.	Iran
Claviceps sp.	Digitaria vestita Hyparrhenia cymbaria Pennisetum americanum	Kenya, Uganda, Zaire
Coniothyrium atriplicinum Wint.	Atriplex semibaccata Atriplex sp.	Australia, South Africa
Cytospora sp.	Rosa kochiana	Not specified
Diaporthe eres Nits.	Acer buergeranum	Japan
Diplodia sp.	Brassaia actinophylla Hydriastele sp.	Not specified
Diplodia coccocarpa Sacc.	Cocos nucifera	Mexico
Dothiorella sp.	Brassaia actinophylla	Not specified
Dothiorella caryotae (Syd.)	Caryota sp.	India
Fusarium sp.	Chloris gayana Hyparrhenia lintonii	Not specified
Macrophoma araucariae De Lac	Araucaria columnaris	New Zealand
Macrophoma palmarum (Cke.) Sacc. and M.	Cocos nucifera	Bahamas (?)
Macrophoma sp.	Liriodendron tulipifera	Not specified
Phaeoseptoria oryzae Miyake	Oryza sativa	Not specified
Phleospora pistaciae Petr.	Pistacia sp.	Not specified
Phoma sp.	Pennisetym americanum Pennisetym sp.	Not specified
Phoma eucalyptica	Eucalyptus caesia	Australia
Phoma insidiosa Tassi	Sorghum bicolor Sorghum sp.	Australia, Korea

Table 2 (Contd.)

Pathogen	Host	Country of Origin
Phoma samararum Desm.	Fraxinus excelsior	Italy
Phomopsis sp.	Acer buergeranum Dais cotinifolia	Not specified
Phomopsis mandevilla Costa and Camara	Mandevilla laxa	Australia
Phomopsis palmicola (Wint.) Sacc.	Howea forsterana	Australia
Phyllachora chusqueae P. Henn. and Lind	Chusquea sp.	Costa Rica
Phyllosticta oryzae Hori	Oryza sativa	Not specified
Phyllostictina sp.	Beaucarnea recurvata	Not specified
Poliotelium hyalospora (Saw.) Mains	Acacia confusa	Republic of China
Prospodium pithecoctenii (Paz.) Cumm.	Pithecoctenium echinatum	Mexico
Pucciniastrum areolatum (Fr.) Otth.	Picea abies (cone) Picea omorika (cone) Picea sp. (cone)	Not specified
Pyrenochaeta sp.	Acer sp.	Italy
Septoria sp.	Oryza sativa (rice)	Not specified
Septoria oryzae Catt.	Oryza sativa	Not specified
Sphacelia sp.	Panicum coloratum Panicum maximum Pennisetum americanum Pennisetum polystachyum	Not specified
Tilletia ayresii Berk.	Panicum maximum Setaria sp. Setaria sphacelata	Not specified
Tilletia rugispora Ell. and Gall.	Paspalum sp.	Not specified
Tolyposporium ehrenbergii (Kuehn) Pat.	Sorghum bicolor Sorghum sp.	India, Niger, Sudan, Upper Volta
Tolyposporium penicillariae Bref.	Pennisetum americanum Pennisetum sp.	Africa (Country ?), Chad, India, Niger, Senegal, Sudan, Uganda

Table 2 (Concld.)

Pathogen	Host	Country of Origin
Uredo dodonaeae Koord.	Dodonaea sp.	Unknown
Ustilago bullata Berk.	Bromus brevis Bromus madritensis Bromus parodii Bromus scoparius	Argentina, Soviet Union, Unknown
Anguina tritici	Triticum sativum	Austria
Ditylenchus dipsaci	Vicia faba	United Kingdom

In previous USDA lists of Intercepted Plant Pests[16, 17, 18, 19], some of the above, and other similar fungi, that all are detectable by hand lens, or by stereomicroscope, have been reported, including pycnidial fungi, such as *Phomopsis cocoes*, *P. conorum*, *P. imperialis* and *P. palmicola*.

5. Policy of Quarantine for Seed

As a conclusion we may outline some essential points to be observed in drafting regulations on quarantine for seed. Some of these have been discussed above[8, 9].

(1) Each country must provide *reliable published surveys* on their endemic plant diseases, pinpointing their actual importance.

(2) The world should be divided into *geographical regions with similar climatic conditions*, with similar prevailing types of crops and with a similar situation of prevailing plant diseases. The existing regional plant protection organisations, such as EPPO, provide the first firm basis for such subdivisions according to more decidedly local interests.

(3) Each quarantine region should, by cooperation between the Plant Protection Services of the member countries provide *selected lists of foreign seed-borne pathogens* of potential danger to the region. These lists should be discussed with plant pathologists and the seed trade ; the views of other quarantine regions should be observed.

(4) Regulations should lay emphasis on the *precise identity of the pathogens* to be considered and the *testing methods* to be used.

(5) In drafting lists of seed-borne pathogens for quarantine, provisions should be based on the *epidemiological type* of the pathogens considered (Categories A and B, Neergaard 1977, discussed above).

(6) After thorough investigation of the technical and practical implications, recommendations should be put forward for consideration by member countries of a Regional Organisation.

(7) Quarantine precautions should aim at acting as an efficient *filter*, not as a barrier. *Post-entry control*, use of *eradicating seed treatments*, and *purification* of plant material such as by meristem cultivation should be applied for particularly valuable seed, e.g., as intended for plant breeding.

(8) Inspection techniques should be standardised and the necessary training of personnel should be organised to ensure uniform implementation of inspection.

REFERENCES

1. Anonymous (1973) Import of plants and plant materials into India by air, *Plant Quarantine Information for Customs Officers*, **1** : 35.
2. Anonymous (1979a) List of intercepted plant pests, APHIS, 82-5, *United States Department of Agriculture*, 568 pp.
3. Anonymous (1979b) The first African Workshop on Seed Pathology, *Seed Pathol. News*, **12** : 3.
4. Anonymous (1980) List of intercepted plant pests, APHIS, 82-6, *United States Department of Agriculture*, 178 pp.
5. Anselme, C. and Mathys, G. (1977) Seed health protection on an international level, technical, legislative and commercial aspects, *ISTA Report, 18th International Seed Testing Congress*, Madrid, May 1977, 1-25.
6. Caresche, L., Cotterell, G. S., Peachey, J. E., Rayner, R. W. and Jacques-Felix, H. (1969) Handbook for phytosanitary inspectors in Africa, *OAU/STRC, PMB* 2359, Lagos, Nigeria, 445 pp.
7. EPPO (1954) Danger from seed-borne diseases. A practical approach to the problems of international safeguards, *Report of the Working Party on Seed-borne Diseases*, Paris, July 1954, 31 pp.
8. EPPO (1959) Seed-borne diseases of tropical and subtropical crops, *Report of the Second Working Party on Seed-borne diseases*, Paris, July 1958, 20 pp.
9. EPPO (1966) Health certification of seeds for export, *Report of the Joint EPPO/ ISTA Working Party*, London, July 1965, 39 pp.
10. EPPO (1975) EPPO Recommendations on New Quarantine Measures, EPPO Bulletin, Special Issue, 30 pp.
11. EPPO (1977) Data sheets on quarantine organisms, *EPPO list* A_1, Paris, September 1977.
12. EPPO (1979a) Data sheets on quarantine organisms, Set 2, Paris, February 1979.
13. EPPO (1979b) Joint EPPO/FAO/EEC/ISTA/FIS Panel on phytosanitary regulations referring to seed-borne pathogenic organisms, EPPO Bulletin, **9** : 370.
14. FAO (1951) International Plant Protection Convention, *Food and Agriculture Organisation of the United Nations*, 34 pp.
15. FAO (1978) Eleventh Session of the Plant Protection Committee for the South-East Asia and Pacific region, *FAO*, 1.
16. Girard, D. H. (1969) List of intercepted plant pests 1968, *U. S. Dept. Agric., Agric. Res. Service*, 86 pp.
17. Girard, D. H. (1971) List of intercepted plant pests 1969, *U. S. Dept. Agric., Agric. Res. Service*, 77 pp.
18. Girard, D. H. (1973) List of intercepted plant pests 1971, *U.S. Dept. Agric., Animal and Plant Health Inspection Service*, 81 pp.

19. Girard, D. H. (1974) List of intercepted plant pests 1972, *U. S. Dept. Agric., Animal and Plant Health Inspection Service*, 97 pp.
20. Goodman, R. M., Bowers, G. R. and Paschal, E. H. (1979) *Crop sci.*, **19** : 264.
21. Hampton, R. O. and Braverman, S. W. (1979) *Pl. Dis. Reptr.*, **63** : 95.
22. Harlan, Jack R. (1977) How green can a revolution be. *In* Crop Resources (Davis S. Seigler, Ed.), 105-110, Academic Press Inc., 233 pp.
23. Hewitt, W. B. and Chiarappa, L., Eds. (1977) Plant health and quarantine in international transfer of genetic resources, CRC Press Inc., 347 pp., 1977.
24. ISTA (1966a) International Rules for Seed Testing, 1966, *Proc. Int. Seed Test Ass.*, **31** : 1.
25. Khare, M. N., Mathur, S. B. and Neergaard, Paul, (1977) *Seed Sci. and Technol.*, **5** : 613.
26. Lambat, A. K. (1974) *Ind. Nat. Sci. Aca., Abstracts of Papers*, 2-3.
27. Leppik, E. E. (1968) *FAO Plant Prot. Bull.*, **16** : 57.
28. Leppik, E. E. (1969) List of foreign pests, pathogens and weeds detected on introduced plants, *Plant Introduction Paper*, No. 15, Beltsville, Maryland.
29. Leppik, E. E. (1970a) Detection and interception of foreign pests, pathogens and weeds from introduced seeds, *Plant Introduction Investigation Paper* No. 17, ISTA, 9 pp.
30. Leppik, E. E. (1970b) *Annual Rev. Phytopath.*, **8** : 323.
31. Lozano, J. C. (1977) *Pl. Dis. Reptr.*, **61** : 644.
32. Morschel, J. R. (1970) Some quarantine aspects of seed-borne diseases, *Commonwealth Department of Health*, Canberra, 1970, 44 pp.
33. Neergaard, Paul (1962) *ISTA Proceedings*, **27** : 386.
34. Neergaard, Paul (1962a) *ISTA Proceedings*, **27** : 400.
35. Neergaard, Paul (1965) *Proc. Int. Seed Test. Ass.*, **30** : 99.
36. Neergaard, Paul (1969) Seed-borne diseases, (Inspection for quarantine in Africa), *Handbook for Phytosanitary Inspectors in Africa*, 380-393, Lagos, 1969.
37. Neergaard, Paul (1970) Seed-borne pathogens and pests in quarantine, *FAO, Near-East Plant Protection Commission*, 31 pp., Cairo, 1970.
38. Neergaard, Paul (1972) *Proc. Int. Seed Test. Ass.*, **37** : 117.
39. Neergaard, Paul (1974) *Seed Pathology News*, **6** : 7.
40. Neergaard, Paul (1977a) Seed Pathology, Vol. I and II, The Macmillan Press Ltd., London and Basingstoke, 1187 pp., Rev. ed. 1979, 1191 pp.
41. Neergaard, Paul (1977b) *In* Plant Health and Quarantine in International Transfer of Genetic Resources, W. B. Hewitt and L. Chiarappa, (Eds.) 33, CRC Press, 347 pp.
42. Pathak, V. K., Mathur, S. B. and Neergaard, Paul (1978) *EPPO Bulletin*, **8** : 21.
43. Randles, J. W. and Dube, A. J. (1977) *Aust. Plant Pathol. Soc. Newsletter*, **6** : 37.
44. Regenberg, Paul (1977) 25, Beretning, Statens Plantetilsyn (The Danish Government Plant Protection Service), 98 pp.
45. Regenberg, Paul (1978) 26, Beretning, Statens Plantetilsyn (The Danish Government Plant Protection Service), 84 pp.
46. Richardson, M. J. (1979) An annotated list of seed-borne diseases, 3rd ed., *Commonwealth Mycological Institute*, Kew, 320 pp.
47. Shetty, H. S., Neergaard, Paul and Mathur, S. B. (1977) *Proc. Ind. natn. Sci. Acad.*, **43** : 201.
48. Shetty, H. S., Khanzada, A. K., Mathur, S. B. and Neergaard, Paul (1978) *Seed Sci. Technol.*, **6** : 935.
49. Sheffield, F. M. L. (1968) *Rev. Appl. Mycol.*, **47** : 1.
50. Zimmer, R. C. and Ali-Khan, S. T. (1976) *Canad. Agric.*, **21** : 6.

Thyroid in Reptiles

J. P. Thapliyal, F.N.A., F.N A.Sc.

Head, Deptt. of Zoology and Director, Institute of Life Sciences,
Banaras Hindu University, Varanasi, India

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

Thyroid or structures homologous to it are present in all chordates[2]. There is also evidence that iodinated amino acids are available even in non-chordates[2,62]. What is, however, a matter of speculation is the circumstance that influenced the evolution of discrete thyroid gland, with capacity to synthesize and release the hormones in the body of animals ?

Typically, thyroid gland is made up of follicles. A follicle consists of epithelial cells that surround a central lumen which is filled with a colloid called thyroglobulin.

In most mammals, birds and amphibians thyroid is present as a paired organ. In fishes it appears either as dispersed follicles or as a compact unpaired gland. Among reptiles, in general, it is single in turtles and snakes and paired in lizards.

In homeothermous vertebrates, the most characteristic effects of thyroid hormones are calorigenesis and regulation of body temperature. Besides, it also influences growth, water and electrolyte transport and intermediary metabolism[64]. In poikilotherms, apart from the well-known effect upon the metamorphosis of amphibians which is inhibited following ablation[12,18,24] ; very little is known of the functions of thyroid hormones[1].

Thyroid hormone is synthesized by the gland through a series of well-known steps[64]. The secretory epithelium of the thyroid gland traps iodide from circulating blood, exidizes it to free iodine (I_2) or IO^-, which then binds to tyrosine, forming monoiodotyrosine (MIT) or diiodotyrosine (DIT). The amino acid tyrosine is present as a part of the glycoprotein called thyroglobulin. The thyroglobulin forms the major part of the thyroid colloid and is stored within the follicles. Triiodothyronine (T_3) and tetra-iodothyronine (T_4) also called thyroxine, are then formed by the coupling of one molecule of MIT with one molecule of DIT and of 2 molecules of DIT respectively. T_3 may also be formed by the deiodination of T_4. The chemical structure of the intrathyroidal hormone and of the circulating is the same. While the various steps in the biosynthesis of thyroid hormones (T_3 and T_4) are known yet many details remain unclear[64].

Possibly because of the small size of the molecule, there has been no change in its chemical structure throughout the evolutionary history of the vertebrates. But, the chemical structure of the plasma protein(s) that bind the hormone(s) and carry it from the gland to the effector sites are different in different groups of animals[4].

Information on thyroid metabolism in reptiles is fragmentary, contradictory and based upon work in a few species only[5,7,8,11,18,31,33,44,47,48,50,53,65]. Earlier studies were restricted to thyroid gland only. These showed that T_4, T_3, DIT and MIT were present within the gland of the reptiles. Licht and Rosenberg[38], were the first to publish information on the nature of thyroid hormone in the plasma of *Anolis*. But, Thapliyal and Chandola[50] provided first complete information on iodine metabolism within the thyroid and the blood of a reptile during different seasons of a year. Since then a similar study has been made in snakes also[65].

In general, it seems that in reptiles thyroidal uptake of radioiodine is slow, maximum values, depending upon species, may occur between one to seven days after the administration of the isotope, ^{131}I. This is followed by a long period of retention of radioactive iodine and thyroxine in the thyroid and the blood of the animal[5,50]. However, rate of the incorporation of ^{131}I in hormone and the release of the hormone into blood suggest that reptilian thyroid[50], contrary to the general belief[31], may not be sluggish.

It was also pointed out for the first time by Thapliyal and Chandola[50] that in reptiles ^{131}I uptake and PB ^{131}I values are not reliable parameters of thyroid activity. There is now enough evidence for an annual cyclicity in the activity of the reptilian thyroid[35,50,65]. There is also evidence to show that there is significant variation in thyroidal uptake of ^{131}I and intra-thyroidal thyroxine during a day (24-hour period).

Further, conditions may vary from species to species and either both the hormones T_3 and T_4[31,33,47,66] or only one, the thyroxine[50,65,66] may be present. MIT and DIT have been reported in the blood of snakes[65], but since these are normally present in the blood of a vertebrate, the report needs to be confirmed[65].

The activity of the thyroid in reptiles, as in other vertebrates, is regulated by thyroid stimulating hormone (TSH) from the pituitary[12,22,23]. However, temperature seems to play an important role, since during winter mammalian TSH was unable to stimulate hormone production in *Calotes*[50].

The parietal eye, pineal and the hormones from the gonads are also known to influence thyroid activity in reptiles. Orchidectomy reduced and male hormone administration stimulated thyroid activity in *Calotes*, *Hemidactylus* and *Natrix*. The male hormone was effective both in the gonadectonized and the *intact* animals[7,8,9,53]. On the other hand, ovariectomy increased and oestradial decreased thyroid activity in the

female Indian *Gecko*[9]. This is surprising especially since ovulation and egg laying in female reptiles are associated with increased thyroid activity[35]. Obviously the relationship between thyroid activity and ovulation and egg laying in reptiles need to be reinvestigated with the help of reliable techniques for the estimation of thyroid activity.

An inverse relationship has been suggested between parietal eye and thyroid activity and also between pineal and thyroid activity[35,49]. Recent work shows that in *Calotes*, also pinealectomy stimulates and melatonin depresses thyroid activity[26]. However, neither the mechanism(s) through which the effect is brought about are known, nor is there sufficient information to show that similar relationship exists in other reptiles also.

Until recently, excepting for the suggestion that thyroid may influence the periodicity of skin shedding[5] and behaviour[19,20], there was practically no information on its role in the physiology of reptiles[2,15,22]. However, since the publication of the observations of Maher and Levedhal[36], and Maher[37,38,39] information has accumulated indicating that, as in homeotherms, in reptiles also thyroid may have multiple effects.

Life Span

There is no record of any premature death in birds, mammals, amphibians and fishes following chronic or acute hypothyroidism. In reptiles, removal of thyroid in *Lacerta*, *Calotes* and *Natrix* leads to decrease in body weight, anemia and ultimately to death[16,17,59]. The thyroidectomized animals do not survive more than a few months. However, during winter, absence of thyroid seems to have no deleterious effect. Surprisingly it may even protect the animal from anemia that results from cold exposure[59]. Further, the influence of thyroid on survival seem to vary in the two sexes. Female *Natrix* unlike male, tolerates thyroid ablation without any apparent ill effect. Obviously, female snake has some mechanism(s) that helps it to withstand the ill effects of thyroidectomy better than the male. But, since this has been studied in one species only, it is difficult to make any generalization.

Moulting

Alteration in thyroid activity has profound effect on moulting. The effects differ in different species[23,44,54]. This diversity may be because other endocrine mechanisms may also be involved in the process[23].

Egg Laying and Development

Information on the role of thyroid in yolk deposition, evolution and gestation is scanty[6,35,40,48,51]. Absence of thyroid delays egg laying and leads to death of the embryo. Even when the embryo develops practically to full size it is not able to hatch out of the shell and dies inside the egg[10]. The young ones that are taken out of the eggs also do not survive long. Obviously there is need for more information on this aspect of thyroid function. No one has so far studied the effect of thyroxine or of antithyroid drugs on the development of embryos in reptiles.

Mineral Metabolism

Mineral metabolism in only two species, of reptiles, viz. *Natrix* and *Calotes* has been studied. Thyroidectomy after a transitory rise was followed by prolonged hypocalcemia. Phosphorus showed a reverse trend resulting in low Ca : P ratio in *Natrix*[28,29]. The rate of removal of phosphate from the body through the urine was also accelerated. The turnover in kidney and liver became high and rate of incorporation of phosphate low[28]. It is not clear whether the effects were due directly to the absence of thyroid hormone(s) in the blood, or due to reduced metabolism or even due to alteration in the functioning of some other endocrine gland like ultimobranchial bodies[41]. Obviously available information is too little to help make any generalisation.

Unlike in mammals, birds and amphibions, excepting for one paper by Thapliyal and Kaur[60], there is no information on the role of thyroid in hemopoiesis of reptiles. l-Thyroxine, had no effect in *intact Natrix* maintained at 30°C but was able to suppress the decrease in RBC number, hemoglobin concentration and hematocrit that followed thyroid removal.

At 16°C, surprisingly thyroidectomy appeared to protect the animal from the ill effects of exposure to low temperature, e.g. decrease in RBC, hemoglobin concentration and hematocrit. While at relatively higher ambient temperatures the effects of altered thyroid activity on blood morphology could be explained by suggesting that thyroidectomy decreased gonadal activity which, in turn, reduced the output of erythropoietin from the kidney[30,58], the results reported at 16°C are not so easily explained. The authors have suggested that at low temperature, reptilian thyroid probably secretes some substance(s) that helps to maintain the level of erythropoietin in the body fluid or alters the metabolic processes in such a way as to lead

to a deterioration of the blood picture or perhaps the *intact* snakes were more metabolically active and this, in turn, might lead to greater demand on the erythropoietic system under conditions (low temperature regression) in which pituitary function was reduced[60].

Intermediary Metabolism

The role of thyroid in the intermediary metabolism of reptiles has not been studied in any depth. The effect of altered thyroid activity on protein, fat and carbohydrate metabolism seems to vary with the season of the year and the temperature at which the animals are maintained. There are only two studies available on the role of thyroid in the intermediary metabolism of a reptile[25,58]. In *Natrix* and in *Calotes* plasma and hepatic free and esterified fatty acids, adipose tissue free fatty acid levels and muscle glycogen decreased and liver glycogen increased both at the low and the high ambient temperatures following thyroid removal[59]. l-thyroxine restored these to *intact* levels[25]. While these results indicate the importance of the fatty acids in the physiology of the reptiles at all temperatures, they raise an interesting question regarding the mechanism through which the effects are brought about at low temperatures when thyroid is supposed to be practically completely inactive[50].

At higher ambient temperature thyroidectomy leads to hypo- and T_4 administration to hypercholesterolemia. This suggests that the rate of synthesis of esterified cholesterol and the release of free cholesterol from the liver and the rate of their removal from the blood are affected by thyroid activity in a manner opposite to that of the homeotherms[1,45]. The plasma glucose was not influenced by thyroidectomy but in *Natrix* was elevated by T_4. In *Calotes* it was reduced by both the treatments[59]. There is also evidence that in *Natrix* thyroidectomy reduces the level of protein in the blood[50]. l-thyroxine administration was able to restore the levels of the chemical constituents of the blood, and adipose tissue of the thyroidectomized snakes[25].

Respiration

In general, at relatively higher ambient temperatures severe hypothyroidism depresses and mild hyperthyroidism elevates respiration of the body, brain, liver, heart, skeletal muscle and lung of reptiles, both *in vivo* and *in vitro*[6,9,32,36,39,44,53,68].

In *in vitro* experiments both T_4 and T_3 have been shown to stimulate oxygen uptake of liver and skeletal muscle slices of *Calotes* and *Natrix* during spring and summer only. During the rest of the year (September to March), neither thyroidectomy nor l-thyroxine had any influence. This indicates that the response of tissues to the thyroid hormones depends to a large extent upon the physiological status (time of the year) of the animal[59,61]. It seems that in addition to temperature, some other factors (hormonal ?) play a role in the response of tissues to l-thyroxine[59]. There is enough evidence indicating to the role of gonadal steroids in respiration of reptiles. Male hormone even restores the decrease in oxygen uptake following thyroidectomy[8,9,53,57]. Male hormone effect on respiration, at least in some species, are brought about by way of the thyroid activity[7,8,52,53,57]. While the question of the mechanism(s) of action of thyroid hormones on oxidative metabolism in reptiles has not yet been elucidated, indirect evidence based upon the results obtained with antibiotics and l-thyroxine, suggest that thyroid hormones probably act at the mitochondrial level.

Reproductions

Considerable information exists pointing to a direct relationship between thyroid activity and reproduction in lizards[35]. Recently it has been shown that castration lowers and administration of male hormone elevates thyroid activity in *Calotes* and *Hemidactylus*[8,9]. In *Natrix*, however, castration and male hormone administration were without any effect on thyroidal uptake of ^{131}I [52]. However, since plasma T_4 was not measured it is not clear whether thyroid activity was, in fact, not influenced. Ovariectomy and oestrogen also influence thyroid activity but in a manner opposite to that of the male hormone[8]. Reverse experiments, involving thyroidectomy or l-throxine administration and assay of gonadal steroids in the blood have not been performed so far. There is evidence, that during summer in *Coleonyx variegatus* maintained at 6-hour daily photoperiod and at $29 \pm 0{\cdot}5$ °C temperature, both thyroidectomy and l-thyroxine suppressed spermatogenesis, and l-thyroxine suppressed the ill effects when given to thyroidectomized animals[46]. Recently, Haldar and Thapliyal[27] have studied, the effect of hypo- and hyper-thyroidism on the testes of *Calotes* during all the phases of the annual gonad development cycle. Their study is the first to show that the effect of thyroidectomy on testes are not the same throughout the year. There was no effect of thyroid ablation when

testes were completely inactive and maximum regression occurred in the fully developed gonads. The effect was less in the regressing and the developing gonads. Similarly, the effect of l-thyroxine varied with the physiological status of the gonad. But, in general, l-thyroxine was able to suppress the ill effects of thyroid removal. It thus appears that for proper functioning of the male gonad of a reptile there is an optimum level of thyroid hormone below and above which only gonadal inhibition results[27]. But, whether the amount of thyroid hormone(s) required during different phases of gonad development is the same or varies is yet to be determined. Similarly, the mechanism(s) through which thyroid effects are brought about on the gonad is not known.

Conclusion

There is now sufficient information to show the involvement of thyroid in many metabolic processes of reptiles. It is rather surprising that thyroid hormone(s) may be practically absent during part of the year and essential for survival and reproduction during the rest of the year.

The beneficial influence of absence of thyroid on blood morphology of *Natrix* during winter certainly needs to be explained. But, more than that perhaps more information on many more species is required.

There is inadequate information on the role of thyroid hormone (s) in the intermediary metabolism of reptiles. Obviously, not only more information, but investigation of the metabolic pathways is required to understand the role of thyroid is the physiology of these cold-blooded animals.

It is rather interesting, or could it be a peculiarity of the cold-blooded vertebrates, that both the thyroidal and gonadal hormones influence respiration and that the decrease in respiration following thyroidectomy (or gonadectomy) can be reversed by male hormone (or T_4). Obviously there is need not only to extend these studies to other species of reptiles but also to other poikilothermous vertebrates. There is also need to know what factors determine the sensitivity of tissues to thyroid hormone(s) at a particular time of the day or of the year. Although information on thyroxinogenesis and on the role of thyroid in reptilian physiology has increased substantially during the last decade, a lot more information has to be collected before we will be able to understand the role of thyroid in the physiology of these cold-blooded, first truely terrestrial vertebrates, that gave rise to the homeothermous birds and mammals.

References

1. Barker, S. B. (1964) Physiological activity of thyroid hormones and analogues. *In* "The thyroid gland", (Eds. R. Pitt-Rivers and W. R. Trotter), 199-236, London, Butterworths.
2. Barrington, E. J. W. (1978) Evolutionary aspects of Hormonal structure and function. *In* "Comparative Endocrinology", (Eds. P. J. Gaillard and H. H. Boer), 381-396, Elsevier/North Holland Biomedical Press, Amsterdam.
3. Bellairs, A. D'A. (1960) Reptiles : Life history, evolution and structure, 1-192, Harper and Brothers, New York.
4. Bentley, P. J. (1976) Osmoregulation. *In* "Biology of the Reptilia", 5A, (Eds. C. Gans and W. R. Dawson,) 365-412, Academic Press, New York.
5. Berg, O., Gorbman, A. and Kobayashi, H. (1959) Thyroid hormones in invertebrates and lower vertebrates. *In* "Comparative Endocrinology", (Ed. A. Gorbman), 302-319, Weiley and Sons, New York.
6. Chandola, A., Singh, K. S. and Thapliyal, J. P. (1972) *Ann. Embryol. Morph.*, 6 : 135.
7. Chandola, A., Thapliyal, J. P. and Gupta, S. C. (1973) *J. Endocr.*, 57 : 325.
8. Chandola, A., Kumar, D. S. and Thapliyal, J. P. (1974a) *J. Endocr.*, 61 : 285.
9. Chandola, A., Kumar, D. S. and Thapliyal, J. P. (1974b) *J. Endocr.*, 63 : 191.
10. Chatterjee, S. and Thapliyal, J. P. (1967) *Endokrin.*, 51 : 234.
11. Chiu, K. W., Lynn, W. G. and Leichner, J. P. (1970) *Biol. Bull.*, 139 : 107.
12. Chiu, K. W. and Phillips, J. G. (1971) *J. Endocrinol.*, 49 : 611.
13. Chiu, K. W. and Lynn, W. G. (1972) *J. Endocrinol.*, 32 : 401.
14. Dent, J. N. (1968) Survey of amphibian metamorphosis. *In* "Metamorphosis", a problem in Developmental Biology, (Eds. W. Etkin and L. I. Gilbert), pp. 271-311, New York, Appleton-Century, Crofts.
15. Dodd, J. M. and Matty, A. J. (1964) Comparative aspect of thyroid function. *In* "The thyroid gland", (Eds. R. Pitt-Rivers and W. R. Trotter), 1, pp. 303-356, London, Butterworths.
16. Eggert, B. (1936) *Z. Wiss. Zool.*, 149 : 280.
17. Eggert, B. (1938) "Morphologie und Histophysiologie der normalen Schilddrüse", "Zwanglose Abhandl. inneren Sekretion", 3 : 1-113, J. A. Barth, Leipzig.
18. Etkin, W. (1970) *Mem. Soc. Endocrin.*, 18 : 137.
19. Evans, L. T. (1939) *Anat. Rec.*, 78 (Suppl.) : 113.
20. Evans, L. T. and Clapp, M. L. (1939) *Anat. Rec.* 75 (Suppl.) : 126.
21. Gorbman, A. (1946) Qualitative variation of the hypophyseal thyrotropic hormone in the vertebrates, *Univ. Calif. Publ. Zool.*, 51 : 229.
22. Gorbman, A. (1959) Problems in the Comparative Morphology and Physiology of the vertebrate thyroid gland. *In* "Comparative Endocrinology", (Ed. A. Gorbman), John Wiley and Sons, Inc., London.
23. Gorbman, A. (1963) Thyroid Hormones. *In* "Comparative Endocrinology", (Eds. U. S. Ven Euler and H. Heller), 1, 291-317, Academic Press, London.
24. Gudernatsch, J. F. (1912) *Wilhelm Roux' Arch. Entwicklungsemch., Organ*, 35 : 457.
25. Gupta, S. C., Thapliyal, J. P. and Garg, R. K. (1975) *Gen. Comp. Endocrinol.*, 27 : 223.
26. Haldar, C. and Thapliyal, J. P. (1976) *Indian J. Zootomy XVII*, 2 : 107.
27. Haldar, C. and Thapliyal, J. P. (1980) Thyroid in reproduction of reptiles, Gen. Comp. Endocrinol, (*in press*).
28. Kar, I., Thapliyal, J. P. and Mukherjee, K. L. (1968) Calcium phosphate interaction in the blood of thyroidectomized rat snake, Symp. Comp. End., B.H.U., pp. 35.
29. Kar, I., Thapliyal, J. P. and Mukherjee, K. L. (1969) *Indian J. Zootomy*, 10 : 180.
30. Kaur, R. J. and Thapliyal, J. P. (1975) *Gen. Comp. Endocrinol.*, 27 : 203-208.
31. Kobayashi, H. and Gorbman, A. (1959) *Annot Zool. Jap.*, 32 : 179.
32. Kumar, D. S., Oommen, O. V. and Thapliyal, J. P. (1974) 26th International Congress of Physiological Sciences, Delhi (Abstr.).
33. Light, P. and Rosenberg, L. L. (1969) *Gen. Comp. Endocrinol.*, 13 : 439.

34. Lynn, W. G. (1960) Structure and functions of the thyroid gland in reptiles, *Ann. Midl. Nat.*, 64 : 309-326.
35. Lynn, W. G. (1970) The thyroid. *In* "The Biology of the Reptiles", (Eds. C. Gans and T. S. Parsons), 201-234, Academic Press, New York.
36. Maher, M. J. and Levedhal, B. H. (1959) *J. Exp. Zool.*, 140 : 579.
37. Maher, M. J. (1961) The effect of environmental temperature on metabolic response to thyroxine in the lizard, "*Lacerta muralis*", *Am. Zool.*, 1 : 461 (Abstr.).
38. Maher, M. J. (1964) Metabolic response of isolated lizard tissues to thyroxine administered, "*In vivo*", *Endocr.*, 74 : 994.
39. Maher, M. J. (1965) The role of thyroid gland in oxygen consumption of lizards, *Gen. Comp. Endocr.*, 5 : 320.
40. Miller, M. R. (1955) Cyclic changes in the thyroid and interrenal glands of the Viviparous lizard, "*Xantusia vigilis*", *Anat. Rec.*, 123 : 19.
41. Öhman, U., Granberg, P. O. and Lindell, B. (1978) *Surgery, Gynecology and obstetrics*, 146 : 773.
42. Pandha, S. K. and Thapliyal, J. P. (1966) *Copeia*, 1 : 121.
43. Pandha, S. K. and Thapliyal, J. P. (1967) *Copeia*, 4 : 865.
44. Pandha, S. K., Chandola, A. and Thapliyal, J. P. (1969) *Indian J. Zootomy*, 10 : 180.
45. Pitt-Rivers, R. and Tata, J. R. (1959) The thyroid hormones, Pergamon Press, London, New York, Paris, Loss Angels.
46. Plowman, M. M. and Lynn, W. G. (1973) *Gen. Comp. Endocrinol.*, 20 : 342.
47. Shaham, Y. and Lewitus, Z. (1971) Radioiodine metabolism in the thyroid gland of the lizard, "*Agama stellio*", *Gen. Comp. Endocrinol.*, 17 : 142-148.
48. Sheliabarger, C. J., Gorbman, A., Schatzlein, F. C. and McGill, D. (1956) Some quantitative and qualitative aspects of ^{131}I-metabolism in turtles, *Endocrinology*, 56 : 331-339.
49. Stebbins, R. C. (1960) Effects of pinealectomy in the Western fence lizard, "*Sceloporus occidentalis*", *Copeia*, 4 : 276-283.
50. Thapliyal, J. P. and Chandola, A. (1973) Seasonal variation in thyroid hormonogenesis in the Indian Garden lizard, "*Calotes versicolor*", *J. Endocrinol.*, 56 : 451-462.
51. Thapliyal, J. P., Singh, K. S. and Chandola, A. (1973) Pre-laying stages in the development of Indian Garden lizard, "*Calotes versicolor*", *Annl. Embryol. Morph.*, 6 : 253-259.
52. Thapliyal, J. P., Chandola, A., Kumar, D. S. and Oommen, O. V. (1974) Testosterone versus thyroid hormone in the regulation of oxidative metabolism in two species of reptiles, *J. Endocrinol.*, 63 : 419-420.
53. Thapliyal, J. P., Kumar, D. S. and Garg, R. K. (1974) Effect of castration and male hormone administration on the tissue respiration and thyroid activity of the Chequered water-snake, "*Natrix piscator*", *Gen. Comp. Endocrinol.*, 22 : 308-311.
54. Thapliyal, J. P. and Gupta, S. C. (1974) Effect of thyroidectomy on the Chequered water-snake, "*Natrix piscator*", *J. Endocrinol.*, 60 : 517-524.
55. Thapliyal, J. P. and Kumar, D. S. (1975) "*Calotes versicolor*", Proc. VIII. Conf. Eup. Comp. Endo., 20.
56. Thapliyal, J. P. and Oommen, O. V. (1975) Proc. VIII. Conf. Eup. Comp. Endo., 19.
57. Thapliyal, J. P., Kumar, D. S. and Oommen, O. V. (1975) *Gen. Comp. Endocrinol.*, 26 : 100.
58. Thapliyal, J. P., Gupta, S. C. and Garg, R. K. (1975a) *Endocr. Exp.*, 9 : 43.
59. Thapliyal, J. P. (1979) Thyroid in Reptiles and Birds, Int. Symp. Hormones and Evolution, Japan, 27 (Abstr.).
60. Thapliyal, J. P. and Kaur, R. J. (1976) *Gen. Comp. Endocrinol.*, 30 : 182.
61. Thapliyal, J. P. and Sharan, S. (1979) *Ind. Journ. Exptl. Biol.*, 18 : 236.
62. Thorndyke, M. C. (1978) *Nature*, 271 : 61, London.
63. Turner, J. E. and Tipton, S. R. (1972) *Gen. Comp. Endocrinol.*, 18 : 98.
64. Turner, C. D. and Bagnara, J. T. (1976) General Endocrinology, W. B. Saynders Company, Philadelphia, London, Toronto.
65. Wong, K. L. and Chiu, K. W. (1974) *Gen. Comp. Endocrinol.*, 23 : 63.
66. Wong, K. L., Chiu, K. W. and Wong, C. C. (1974) *Gen. Comp. Endocrinol.*, 23 : 71.

Use of Polygraph as an Aid in Tracing Affinities in Polygonaceae

A. K. Sharma and S. N. Sahay*

Centre for Advanced Studies in Cell and Chromosome Research
Department of Botany, University of Calcutta
35 Ballygunj Circular Road
Calcutta-700 019

*Department of Botany, D.S. College, Katihar, Bihar

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

Background

A significant amount of work on biosystematics has been carried out in recent years on different families of angiosperms. Investigations of populations of the same taxon from different regions of the world have shown the presence of cytological races[8, 7, 11]. Even populations collected from within a comparatively small area have been reported to contain different chromosome numbers[13, 14, 16, 17].

The family Polygonaceae excepting the genus *Rumex* with its different subsections has not been extensively dealt with. This is specially true of the genus *Polygonum* which is widely distributed throughout India. The lacuna in this investigation is mainly due to the difficulty in studying their cytology which forms a major component of biosystematical investigations. The heavy cytoplasmic contents and small chromosomes are the principal limitations for chromosome analysis.

In Bentham and Hooker's system of classification of 1886[3], the family Polygonaceae is the last number of the Curvembryeae under the Monochlamydeae. Engler and Diels[4] included the family in a distinct order Polygonales, lying between the Aristolochiales and Centrospermae. Bessey[1] accepted the family as an advanced taxon of the Caryophyllales. Hallier[6] considered its position under the Caryophyllales, seeking its origin directly from early Ranales on a line parallel to the Papaveraceae. Hutchinson[9] to some extent agreeing with Hallier[6] opined that the Polygonales constitutes a degraded and reduced type of the Caryophyllales, descending from the Ranales. The family is related to Amaranthaceae, Chenopodiaceae and Nyctaginaceae in the plan of floral structure, but differs from them in the presence of ochreate stipules, triangular ovary, solitary erect ovule and S-shaped embryo. The family shows similarity with Urticaceae in stipulate leaves and orthotropous ovules. A unique feature of Polygonaceae, as suggested by Rendle, is the multiplication of the stamens and the relation between the two whorled trimerous and cyclic perianth.

Cytological studies in the genus *Polygonum* with special reference to the interpretation of its taxonomy, have been carried out in detail by Löve and Löve[12]. The chromosome counts of members of this genus have been made by several workers [2, 5, 10, 15, 18, 19, 20, 21].

The present text deals with different species of Polygonaceae, both cultivated and wild, from different populations and localities with special

reference to the genus *Polygonum*. In addition to chromosomal analysis, our attempt has been made to pool other important morphological characters and evidences which are expected to provide useful information about the evolutionary tendencies and affinities of different taxa within the family in general and the genus *Polygonum* in particular.

Methodology of Polygraph Preparation

Along with the cytological study carried out of the various members of Polygonaceae in detail, certain morphological characters have been taken into consideration. Of these, the following parameters were chosen for the preparation of polygraphs.

(A) Chromosome number in diploid complement
(B) Number of chromosome bearing secondary constrictions
(C) Total length of chromosome in haploid complement
(D) Leaf area
(E) Length of internode
(F) Height of the plant

The relationship of the plant varieties or their alliance with each other as regards the above-mentioned characters have been traced through comparison of the polygraphs (which are the diagrammatic representation of these characters for the individual species). It is true that different chromosome numbers have been found in various members of Polygonaceae. However, there is a similarity in the general pattern of chromosome morphology in that, the chromosomes are graded and secondary constrictions vary from two to six in number. But the total amount of chromatin does not show a marked difference. Cytotypes seem to exist in the species of Polygonaceae. A detailed analysis of the karyotype shows that each species is characterised by its own karyotype which differs in minute details from that of others. This is clearly an indication of the role of structural alterations of chromosomes in the evolution of species. The karyotypes differ mostly in the number and morphology of the secondary constrictions.

The differences in morphological characters of the species are quite distinct. There is, no doubt, a relationship between the structural differences in chromosome complement and in phenotypic characters, but the extent to which these differences are correlated to each other are yet to be worked out. It is quite remarkable that widely different populations from the same species

often show a gross similarity with each other species and as such populations be grouped, on the basis of these polygraphs.

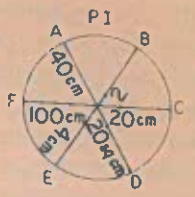
Interpretation of Polygraphs

In the polygraphs, six parameters have been chosen of which three deal with chromosome characteristics and the other three with the morphological nature of the plant. The diploid chromosome number, number of chromosome bearing secondary constrictions as well as total length of haploid complement are the cytological characteristics whereas mean leaf area, length of internode and height of the plant have been chosen on the basis of phenotypic characters (vide Table).

In *Fagopyrum esculentum* (Fig. P II) in absence of any other species studied it is difficult to assess its affinity. The other two species of *Fagopyrum* could not be grown at Calcutta, hence its polygraphs have not been drawn. Of the two populations of *P. limbatum*, polygraphs show remarkable similarity in the characters chosen, the features being almost identical (Fig. P III, P IV).

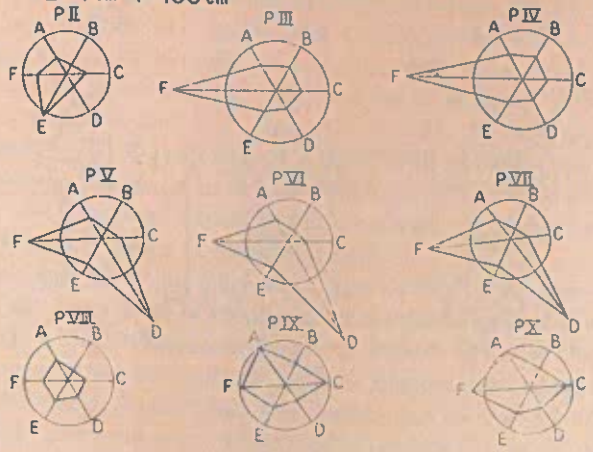
In *P. orientale*, the three populations show polygraphic features quite different from *P. limbatum* and have certain similarities amongst themselves or amongst the different populations of some species. It is remarkable that though the population at Katihar and Calcutta show more or less identical chromosomes features, their polygraphs differ from one another to some extent (Fig. P V–P VII). On the other hand, whereas the population from Darbhanga differs in certain minute karyotypic details from the other two populations, its polygraph is more or less identical with that of Katihar population. It is quite likely that similarities and dissimilarities with respect to certain chromosomes in these three populations are not connected with the particular characters responsible for polygraphic differences.

In *P. barbatum*, four populations have been studied of which three are tetraploids and one diploid. The diploid population from Garia shows markedly diminutive features in phenotypic characters as compared to tetraploid from Calcutta. The Calcutta tetraploid population also shows certain distinctive features from (Fig. P X and P X¹) those of Katihar and Darbhanga, the latter two showing some similarities with each other. In the three populations of tetraploids, the chromosome types are slightly different from each other. The extent to which these differences can be correlated with phenotypic differences from different populations is yet to be investigated.



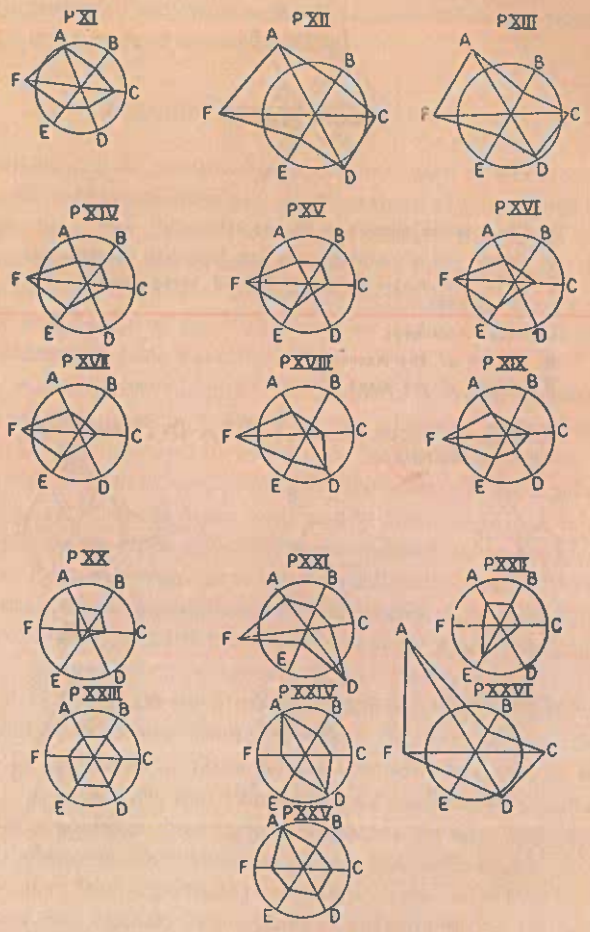
A Chromosome number in diploid set
B Number of chromosome bearing secondary constrictions
C Total chromatin length in haploid chromosome complement
D Mean leaf area
E Length of the internode
F Height of the plant

A = 40 cm B = 2 chromosome C = 20 μ D = 20 square cm
E = 4 cm F = 100 cm



Polygraph I : characters employed in preparing the polygraphs
Polygraphs II to X : belong to the species *Fagopyrum esculentum* ; *Polygonum limbatum*, from Katihar and Darbhanga ;
P. orientale from Darbhanga, Calcutta and Katihar ;
and *P. barbatum* from Garia, Calcutta and Katihar respectively

However, Katihar and Darbhanga populations do not show marked dissimilarity in the polygraphs though they differ in minute karyotypic details. Evidently such karyotypic differences might have involved chromosomes



Polygraphs XI to XXVI : represent the species *Polygonum barbatum* from Darbhanga ; *P. glabrum* from Howrah and Santiniketan ; *P. hydropiper* from Darbhanga, Katihar, Santiniketan, and Garia ; *P. lenigerum* from Katihar ; *P. capitatum* ; *P. chinense* ; *P. runcinatum* ; *P. molle* ; *Rumex vesicarius* ; *R. dentatus* from Calcutta and Katihar and *R. nepalensis* respectively

which do not necessarily contain genes for characters chosen as criteria for the preparation of polygraphs.

The populations of *P. glabrum* show 2n = 60 chromosomes and the two populations have a general similarity in their polygraphs, but dissimilarity in certain aspects (vide polygraphs), the number of chromosomes with secondary constrictions also being four in Santiniketan and six in Howrah populations. The general similarities in the polygraph, as well as the minor difference between the two populations, indicate simultaneously the homogeneity of the species and importance of cryptic chromosomal alteration in evolution. The overall gigantic character in *P. glabrum* as compared to *P. barbatum* and *P. orientale* is an index of the role of high degree of polyploidy in bringing about giganticism. The former two species have tetraploids and diploids whereas *P. glabrum*, studied here, shows hexaploid chromosome number.

All the four populations of *P. hydropiper*, namely Darbhanga, Katihar, Santiniketan and Calcutta, have diploid chromosome constitution (2n = 20) but four chromosomes with secondary constrictions are present in the first two and two constrictions in the last two. The polygraphs of the four populations show a remarkable similarity with slight differences from one another which may be accounted for, on the basis of structural differences of chromosomes. The species, *P. lenigerum* 2n = 22, *P. capitatum* 2n = 22, *P. chinense* 2n = 22, *P. molle* 2n = 32 (Figs. P XVIII to P XXII) are presented in this investigation by one population each. The differences in chromosome number, and in certain aspects of the karyotypes are also reflected in the differential manifestation in their polygraphs.

If the polygraph is to be taken as an index of affinity, as specially evident between the populations of the species, it is remarkable that *P. hydropiper* (Figs. P xiv to P XVII) has strong similarity with *P. capitatum* (Fig. P XIX). The extent to which such affinity indicates phylogenetic relationship is yet to be investigated. Similarly certain populations of *P. orientale* (Fig. P VII) have remarkable similarity with *P. runcinatum* 2n = 22 (Fig. P XXI) though in the latter, there has been an overall reduction in size. The chromosome number of the two species are identical. On the basis of these facts it appears that polygraphs can be utilised to work out the relationship between different species. However, in order to arrive at precise conclusions, it is desirable to have as many important characters in the preparation of polygraphs as possible.

Three species of *Rumex* have been studied, namely *R. vesicarius* 2n = 18, the two populations of *R. dentatus* (Katihar and Calcutta) 2n = 40 and

R. nepalensis 2n = 108 chromosomes (Figs. P xxiii to P xxvi). The three species differ in chromosome number from each other. Remarkably, however, the chromosome number with secondary constrictions are identical in all of the three populations. *R. vesicarius* has largest chromosome size (total chromatin length of the haploid complement 23·38μ, low number 2n = 18) and it is quite distinct from *R. dentatus*.

The polygraphs clearly indicate similarities between two populations of *R. dentatus*. *R. nepalensis* is quite distinct from *R. dentatus* in the nature of polygraph as well as its size which reflects clear giganticism due to polyploidy. The genus *Rumex*, as is well known, represents quite a diversified assemblage which is clearly reflected in the polygraphs as well.

The polygraph analysis of *Fagopyrum*, *Polygonum* and *Rumex* have helped to delineate not only the homogeneity of certain taxa and the heterogeneity wherever in existence, but also the affinities between the different species.

Summary of the External Morphological and Karyological Data on Various Members of Polygonaceae

No. 1 is key diagram

Sl. No.	Name of the species	A	B	C	D	E	F
		Diploid Chromosome Number	Number of Chromosomes bearing Secondary Constriction	Total chromatin length (n) in μ	Leaf area (sq cm)	Length of internode (cm)	Height of plant (cm)
2.	*Fagopyrum esculentum* Moench (Katihar)	16	2	17·35	13·1	9	70
3.	*Polygonum limbatism* Meissn. (Darbhanga)	22	4	19·38	15·5	3·6	200
4.	*P. limbatum* Meissn. (Katihar)	22	4	17·94	16·6	3·9	225
5.	*P. orientale* Linn. (Darbhanga)	22	2	16·23	89·7	5·2	174
6.	*P. orientale* Linn. (Esplanade, Calcutta)	22	2	15·23	103·4	5·5	180
7.	*P. orientale* Linn. (Katihar)	22	2	15·81	88·2	5·8	190
8.	*P. barbatum* Linn. (Garia)	20	2	15·38	17·6	3	50
9.	*P. barbatum* Linn. (Calcutta)	40	4	32·66	17·3	4·5	95
10.	*P. barbatum* Linn. (Katihar)	40	4	35·06	19·3	5	126
11.	*P. barbatum* Willd. (Darbhanga)	40	4	30·46	19·6	4·7	120
12.	*P. glabrum* Willd. (Howrah)	60	6	46·14	47·3	4	190
13.	*P. glabrum* Willd. (Santiniketan)	60	4	45·14	38·5	3·5	150
14.	*P. hydropiper* Linn. (Darbhanga)	20	4	18·51	10·5	5·3	125
15.	*P. hydropiper* Linn. (Katihar)	20	4	15·94	8·5	5·5	132
16.	*P. hydropiper* Linn. (Santiniketan)	20	2	19·12	9·0	5	120
17.	*P. hydropiper* Linn. (Garia)	20	2	14·26	12·0	5·2	123
18.	*P. lenigerum* Roxli. (Katihar)	22	2	14·50	36·3	5	160
19.	*P. capitatum* Ham. (Darjeeling)	22	4	13·94	2·9	2	18
20.	*P. chinense* Linn. (Darjeeling)	24	2	19·66	14·7	4	140
21.	*P. runcinatum* Ham. (Darjeeling)	22	4	17·08	10·3	6·5	36
22.	*P. molle* D. Don. (Darjeeling)	32	4	20·20	54·6	2	150
23.	*Rumex vesicarius* Linn. (Katihar)	18	4	23·38	27·0	5	50
24.	*R. dentatus* Linn. (Calcutta)	40	4	22·75	37·8	3·5	50
25.	*R. dentatus* Linn. (Katihar)	40	4	24·30	25·2	4	65
26.	*R. nepalensis* Spreng (Darjeeling)	108	4	57·15	41·6	5	140

REFERENCES

1. Bessey, C. E. (1914) The Phylogenetic taxonomy of flowering plants, *Ann. Missouri Bot. Gard.*, 2 : 109-164.
2. Castro, D. de, and Fontes, F. C. (1946) Primeiro contacto citologico Com a flora holofila dos Salgados de Sacavem, *Broteria*, 15 : 38-46.
3. Darlington, C. D. and Wylie, A. P. (1955) *Chromosome Atlas of flowering plants*, George Allen and Unnein Ltd., London, pp. 510.
4. Engler, A. and Diels, L. (1936) *Syllabus der Pflanzenfamilien*, Aufl, 11, Berlin.
5. Flovik, K. (1940) Chromosome numbers and Polyploidy within the flora of Spitzbergen, *Hereditas*, 26 : 430-440.
6. Hallier, H. (1911) Provisional Scheme of the natural (Phylogenetic) system of flowering plants, *New Phytol.*, 4 : 151-162.
7. Harada, I. (1956) Cytological Studies in Helobiae. I. Chromosome idiograms and a list of chromosome numbers in seven families, *Cytologia*, 21 : 306-328.
8. Hooker, J. D. (1886) *Flora of British India*, V. L. Reeve and Co., Kent, pp. 110.
9. Hutchinson, J. (1959) *The families of flowering plants*. I. Dicotyledon, Clarendon Press, Oxford.
10. Jaretzky, R. (1928) Histologische und karyologische studien on Polygonaceen, *Jahrb. Wiss. Bot.*, 69 : 35-490.
11. Larsen, K. (1963) Studies in the flora of Thailand. 14. Cytological studies in vascular plants of Thailand, *Dansk. Bot. Ark.*, 20 : 211-275.
12. Love and Love (1956) Chromosome and Taxonomy of Eastern North, *Polygonum*, *Can. J. Bot.*, 34 : 501-521.
13. Mishra, M. P. (1966) Cytology of certain angiospermous species, with special reference to taxonomy and ecology of Hydrocharitaceae, Ph.D. Thesis, Magadh University.
14. Mishra, M. P. (1971) Cytological studies in *Hydrilla Verticillata* Presl., *J. Cytol. Genet.*, 6 : 59-65.
15. Schnack, B. and Fernandez, O. (1946) Numerous Cromosomicos de cuatro especies cultivadas, *Bol. Soc. Argentina Bot.*, 1 : 285-286.
16. Sharma, A. K. (1974) Plant Cytogenetics, *The Cell Nucleus*, 2 : 263-291.
17. Sharma, A. K. and Chatterjee, T. (1967) Cytotaxonomy of Helobiae, with special reference to the mode of evolution, *Cytologia*, 32 : 286-307.
18. Skalinska, M. (1950) Studies in chromosome numbers of Polish angiosperms, *Actasoc. Bot. Polon.*, 20 : 45-68.
19. Sugiura, T. (1936) Studies on the chromosome numbers in higher plants, with special reference to cytokinesis, I, *Cytologia*, 7 : 544-595.
20. Suzuka, O. (1950) Chromosome numbers in pharmaceutical plants, I, *Seiken Ziho* (*Rept. Kiharadust. Biol. Res.*), 4 : 57-58.
21. Wulff, H. D. (1939) Chromosomen studied another schleswigholsteinschen Angiosperman Flora, III, *Ber. Deutsch. Bot. Ges.*, 57 : 84-91.

Scales in Atmospheric Science

M. P. Singh

Indian Institute of Technology, Delhi

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

There is no systematic analytical way to find the solution of a nonlinear system. One of the analytical methods for tackling such problems is the so-called "perturbation technique". The system is reduced to a nondimensional form by referring different physical variables to characteristic scales. The resulting problem involves a nondimensional parameter (parameters) in terms of which the dependent variables are expanded asymptotically. The terms of such a perturbation expansion satisfy, in general, simpler system in comparison to the original system and so quite frequently the solution by perturbation expansion becomes mathematically more amenable.

It is assumed that each term of the perturbation series is smaller in magnitude than its preceding term throughout the region of interest. If this condition breaks down in any part of the region of interest, the asymptotic expansion is no longer valid there. Some of the reasons for the breakdown of the asymptotic solution are :

(i) the highest derivative is multiplied by a small parameter giving rise to what is known as boundary layer, and

(ii) the so-called large integrated effect—the disturbance imparted to a system may be small, but its cumulative effect may become quite large. This generates secular term which leads to the breakdown of the asymptotic solution. For periodic motion, it can be shown that secular terms are produced by resonance. In such a situation, the physics of the problem admits more than one characteristic scale and the different scales are disparate. Let us, for example, consider disparate scales called the "long" and the "short" scales which represent, respectively, the slow and fast changes of the system. The secular terms can be avoided if we introduce both slow and fast variables when the system is made non-dimensional. If there are more than two disparate scales an extension of this technique to cover such cases is fairly straightforward. This method, known as the "method of multiple scales", has been applied to several problems in atmospheric science.

It is recognised that the atmosphere interacts with the underlying surface through a boundary layer. Vertical transport of momentum, heat and moisture take place through the boundary layer. The length and time scales of motion within the boundary layer are small compared to the larger scale motion in the free atmosphere. The time taken for the vorticity in an atmospheric vortex to be reduced to $1/e$ of its initial value by friction in the

boundary is approximately 4 days (spin-down time). The spin-down time for molecular diffusion is about 100 days which is indicative of the different time scales for frictional dissipation and diffusion in the atmospheric boundary layer.

In the Ekman-type boundary layer in rotating fluid[16] inertial oscillations, spin-down response and diffusion give rise to three different time scales, the longest being due to diffusion. An expansion in terms of Ekman number brought out the interesting result that, for a stratified fluid, the structure of the final steady state is resolved only when long-time problem is solved.

The problem of Geostrophic adjustment is gaining importance in numerical integration of primitive equation models[1]. It seeks to determine how a field of pressure and wind which is initially unbalanced, ultimately reaches a state of Geostrophic balance. If f is the Corolis parameter, then the time scale for adjustment in mid-latitudes is $T \sim \frac{1}{f}$. But if weather systems move with a speed U, and L is the characteristic length, then another time scale is $\tau \sim \frac{L}{U}$. The ratio of these scales is the Rossby number : $R_0 = T/\tau = U/fL$.

For mid-latitude $R_0 \sim 0{\cdot}1$ but at low latitudes, R_0 becomes very large. An expansion in R_0 (for mid-latitudes) yields secular terms in the first-order system whose removal provides a physical condition which implies conservation of the important meteorological variable-potential vorticity of the system.

More recently, multiple scales have been used to study resonant interactions between atmospheric waves. Thus, Newell[13] found resonant interactions between Rossby waves, while Stone[15] studied a similar problem with baroclinic waves. Sideband resonance was considered by Shabbar[17] for Rossby waves, and Lindzen[10] examined the vertical propagation of waves near the equator. These investigations have been extended to beta-plane models by Duffy[7] and by Domaracki and Loesch[6].

While considering the equations of motion for weather forecasting, the role of frictional stresses is important enough to merit separate consideration because in the lowest kilometre of the atmosphere these forces are comparable to the force due to pressure gradient, or the Coriolis force. This region of the atmosphere is known as the planetary boundary layer which contains nearly 10 per cent of the mass of the atmosphere. The transition

from laminar to turbulent flow is usually determined by the Reynolds' number R_e. This expresses the ratio between the inertia and viscous forces ($R_e = Ud/\nu$, where ν is the kinematic viscosity and d is the atmospheric depth).

The planetary boundary layer consists of several layers, each of which is characterized by a separate dominant mechanism. The principal layers are :

(i) Layer of skin friction or laminar sub-layer

This is roughly a few mm deep. In this layer the molecular properties influence the structure of turbulence. The properties of this layer need consideration when we wish to determine the characteristics of a film over an extensive water surface, for example.

(ii) Constant stress or surface layer (30 metres)

Above the laminar sub-layer, the molecular properties lose their importance. In the next 30 metre (approx) above the laminar sub-layer, the wind stress changes very slowly (by less than 20%) ; consequently, this is often referred to as the surface of constant stress layer. In the surface layer, the earth's rotation is not important, but thermal stratification is important. The depth of the surface layer varies with the stability of the atmosphere. For a neutral calm atmosphere it is of the order of 30 metres.

(iii) Ekman layer

Above the surface layer, the earth's rotation (Coriolis force) becomes important up to about 1·5 km. This is generally referred to as the Ekman layer.

The dynamics of the Ekman layer is studied, as a first approximation, by assuming balance between (a) the Coriolis force, (b) the pressure gradient force, and (c) the eddy stress terms.

According to quasi-equilibrium assumption and growth of disturbances (Arakawa and Schubert : *Journal of Atmospheric Science*, 31 : 674), the concept of quasi-equilibrium between the destabilising large-scale forcing and stabilising effects of the cumulus ensemble are based on the hypothesis that the time scale of the changes of large-scale forcing is much larger than the time scale over which the cumulus activity adjusts to a given large-scale forcing.

It has been shown recently (J. Shukla : Ph.D. thesis, M.I.T.) that a small variation of surface temperature of Arabian Sea entail large variation of rainfall over the Indian subcontinent (large integrated effect).

Other well-known instances where asymptotic expansion is not uniformly valid relate to (i) passage from inviscid to viscous flows, (ii) transition from heat conduction to convective heat transfer, (iii) passage from Stokes's flow to Oseen flow, (iv) theory of shocks arising in the limit of small viscosity and heat conductivity, (v) shallow water theory, etc.

In hydrostatics the problem of the surface tension of a liquid in a vertical cylinder for large values of the bond number will have a non-uniform solution[14]. In turbulent shear flows viscous stress remains significant in a thin layer close to the wall and the production of turbulent kinetic energy from the mean flow and its diffusion due to the fluctuating motion remain significant in the boundary layer ; however, outside this boundary layer Reynolds stresses alone remain dominant[19]. In the non-gray radiative transfer problem arising in Stellar atmosphere, the governing equations are boundary layer type—the boundary layer region being near to the zero optical depth in the outer layers of the atmosphere where the temperature and the amount of ultraviolet flux emerging from the atmosphere are greatly reduced ; on the other hand in the region outside the layer the temperature is important[4].

Disparate scales occur while studying the Orr-Sommerfeld equation[18], the tidal equations[8], the free surface oscillations in a tank resulting from drainage[11] Rayleigh-Taylor instability[9, 12], gravity waves in water of variable depth[8], the resonant interaction of Rossby wave packets in atmospheric science[13], the propagation of equatorial Tanai and Kelvin waves through shear[10]. Carrier[2] studied the ocean current problem invoking the ideas of boundary layer.

The importance of scales can be very well illustrated by the Linear spring-mass-damping system in which we will discuss two limiting cases that give rise to two different types of problems.

Let us consider a spring-mass-damping system where the mass is subjected to an initial impulse I_0. The governing equation is

$$m\ddot{x} + \beta\dot{x} + kx = 0 \tag{1}$$

with the initial conditions

$$x(0) = 0, \quad \dot{x}(0) = I_0/m \tag{2}$$

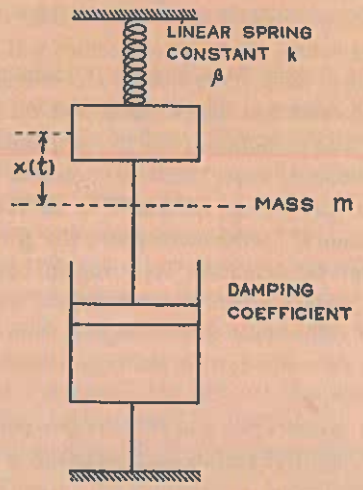
Case A : Slightly damped oscillations (i.e , β is small)

To write the system in non-dimensional form, the characteristic time scale corresponds to the natural frequency of the oscillations, i.e., $\sqrt{\frac{m}{k}}$ and

characteristic length scale is chosen as $l_0/\sqrt{mk}$. This leads to

$$\ddot{x}+2\varepsilon\dot{x}+x=0 \tag{3}$$

$$x(0)=0,\ \dot{x}(0)=1,\ \text{where}\ \varepsilon=\frac{\beta}{2\sqrt{mk}}\ll 1 \tag{4}$$



There is a second time scale (long time scale) m/β corresponding to the decay time. The first one (normal time scale, also called "fast variable") represents fast changes of the system and the long time scale represents slow changes of the system and hence is called "slow variable". Since the oscillations on the short time scale are modulated on the long time scale, the two time scales will have to be considered simultaneously. Such problems are treated by applying the method of multiple scales described below.

Many physical processes are characterised by the slow change of a basically periodic motion through the action of small (ε) imperfection. While the perturbation technique may work for any finite time interval, it can be expected to fail for infinite one ; given ε however small, the change is bigger than ε at sufficiently large values of t, i.e., cumulative effects of small perturbations are large.

It is important to notice that there are two effects of small damping : to decrease the amplitude and to alter the frequency. The time scale for the

former is long $O(\varepsilon^{-1})$, being based on the damping coefficient so that

$$\tau=\varepsilon t \tag{5}$$

is an appropriate variable. The change in frequency involves only a perturbation of time scale, and

$$t^*=t(1+\varepsilon w_1+\varepsilon^2 w_2+\cdots) \tag{6}$$

is a suitable variable. Note that we have assumed, as is adequate for most cases, that the frequency change is time independent ; otherwise $w_1, w_2, \cdots$ are functions of t.

Accordingly, x is now a function of t^* and τ and hence

$$\dot{x}=\frac{\partial x}{\partial t^*}\frac{dt^*}{dt}+\frac{\partial x}{\partial \tau}\frac{d\tau}{dt}=(1+\varepsilon w_1+\cdots)\frac{\partial x}{\partial t^*}+\varepsilon\frac{\partial x}{\partial \tau} \tag{7}$$

Similarly,

$$\ddot{x}=(1+\varepsilon w_1+\cdots)^2\frac{\partial^2 x}{\partial t^{*2}}+2\varepsilon(1+\varepsilon w_1+\cdots)\frac{\partial^2 x}{\partial t^{*2}}+\varepsilon^2\frac{\partial^2 x}{\partial \tau^2} \tag{8}$$

we now expand

$$x(t^*, \tau, \varepsilon)=x_0(t^*, \tau)+\varepsilon x_1(t^*, \tau)+\cdots \tag{9}$$

and substituting in (3) and (4) we obtain

$$\frac{\partial^2 x_0}{\partial t^{*2}}+x_0=0,\ x_0(0, 0)=0,\ \frac{\partial x_0(0, 0)}{\partial t^*}=1 \tag{10}$$

$$\frac{\partial^2 x_1}{\partial t^{*2}}+x_1=-2w_1\frac{\partial^2 x_0}{\partial t^{*2}}-2\frac{\partial^2 x_0}{\partial t^*\partial \tau}-2\frac{\partial x_0}{\partial t^*} \tag{11}$$

$$x_1=0,\ \frac{\partial x_1}{\partial t^*}=-w_1\frac{\partial x_0}{\partial t^*}-\frac{\partial x_0}{\partial \tau}\ \text{at}\ (0, 0),\ \text{etc.}$$

The very first function gives a modulated oscillation :

$$x_0=A_0(\tau)\cos t^*+B_0(\tau)\sin t^*,\ A_0(0)=0,\ B_0(0)=1 \tag{12}$$

But to determine A_0 and B_0 we must go to the next equation

$$\frac{\partial^2 x_1}{\partial t^{*2}}+x_1=2(w_1 A_0-B_0'-B_0)\cos t^*+2(w_1 B_0+A_0'+A_0)\sin t^* \tag{13}$$

secular terms $t^*\cos t^*$, $t^*\sin t^*$ will appear (indicating non-uniformity) unless the parentheses vanish, i.e.,

$$A_0=-e^{-\tau}\sin w_1\tau,\ B_0=e^{-\tau}\cos w_1\tau \tag{14}$$

Then

$$x_0=e^{-\tau}\sin(t^*-w_1\tau) \tag{15}$$

For A_0, B_0 to have the character of amplitude modulator, we need

$$w_1=0 \tag{16}$$

At the next step we have

$$x_1=A_1(\tau)\cos t^*+B_1(\tau)\sin t^*,\ A_1(0)=B_1(0)=0 \tag{17}$$

and, in order to suppress secular terms in the equation for x_2, we need

$$A_1' + A_1 + \left(w_2 + \frac{1}{2}\right) e^{-\tau} = B_1' + B_1 = 0 \tag{18}$$

clearly, $B_1 = 0$ and, to avoid terms $\tau e^{-\tau}$ which would eventually make A_1 large compared to A_0, we need $w_2 = -\frac{1}{2}$ so that $A_1 = 0$ also.

The complete solution is, therefore,

$$x = e^{-\tau} \sin t^* + O(\varepsilon^2) \tag{19}$$

where

$$\tau = \varepsilon t, \; t^* = t\left(1 - \tfrac{1}{2}\varepsilon^2 + O(\varepsilon^3)\right) \tag{20}$$

This is a uniform asymptotic approximation for

$$0 \leqslant t < \infty$$

Case B : *Large damping or small mass* (i.e., $mk \ll \beta^2$)

For small values of m, not only is the equation (1) reduced to first order when $m = 0$, so that only one of the boundary conditions (2) can be satisfied, but also one of the conditions becomes infinite.

The physical picture is of the mass moving rapidly away from equilibrium, but being decelerated even more rapidly by the damping until it has displacement but no velocity. Then it is less abruptly returned by the spring acting to overcome the damping. During the initial stages of the motion (i.e., $t \ll 1$), inertia effect and damping would be important and so the short time scale m/β is effective near $t = 0$. The normal time scale effective after initial stages of the motion is essentially the decay time β/k (in which both spring action and damping are important).

Such problems are best treated by applying Method of Matched Asymptotic Expansions which is described in the following paragraphs.

If we refer the length to I_0/β and time to the normal time scale β/k (i.e., the primary scale), equations (1) and (2) becomes

$$\varepsilon \ddot{x} + \dot{x} + x = 0, \quad x(0) = 0, \quad \dot{x}(0) = \frac{1}{\varepsilon} \tag{21}$$

where

$$\varepsilon = mk/\beta^2$$

clearly, we set

$$x(t, \varepsilon) = x_0(t) + \varepsilon x_1(t) + \cdots \tag{22}$$

(The straightforward expansion (22) is commonly called the "outer expansion" and the variable t is called the "outer variable"). Although one of the boundary conditions $x(0) = 0$ can be satisfied, in this instance neither should be because the outer expansion is not valid in the initial stages of the

motion. (The region of validity of outer expansion is called "outer region"). Thus $\dot{x}_0 + x_0 = 0$, $\dot{x}_1 + x_1 = -\ddot{x}_0$, etc... and so

$$x_0 = A_0 e^{-t}, \; x_1 = -A_0 t e^{-t} + A_1 e^{-t}, \text{ etc...} \tag{23}$$

To write a valid expansion near $t = 0$, we introduce the inner variable

$$\tau = t/\varepsilon \text{ and} \tag{24}$$

$$x(t, \varepsilon) \equiv X(\tau, \varepsilon) \tag{25}$$

It can be checked in a straightforward manner that the inner variable τ given by (24) is referred to the short time scale m/β (i.e., secondary scale) which is dominant during the initial stages of the motion (commonly called the "inner region" or the "boundary layer"). Accordingly, we set up the inner expansion

$$X(\tau, \varepsilon) = X_0(\tau) + X_1(\tau) + \cdots \tag{26}$$

Equations (21) now become

$$X'' + X' + \varepsilon X = 0, \quad X(0) = 0, \quad X'(0) = 1 \tag{27}$$

where

$$X' = \frac{dX}{d\tau}$$

Hence, we find

$$X_0'' + X_0' = 0, \quad X_0(0) = 0, \quad X_0'(0) = 1,$$

$$X_1'' + X_1' = -X_0, \quad X_1(0) = 0, \; X_1'(0) = 0, \text{ etc...}$$

The (unique) solutions are

$$X_0 = 1 - e^{-\tau}, \quad X_1 = 2(1 - e^{-\tau}) - \tau(1 + e^{-\tau}), \cdots$$

In the inner region, i.e., in any fixed finite interval $0 \leqslant \tau \leqslant T$,

$$X = 1 - e^{-\tau} + \varepsilon[2(1 - e^{-\tau}) - \tau(1 + e^{-\tau})] + O(\varepsilon^2) \tag{28}$$

is a uniform asymptotic approximation.

Likewise, we expect that, after A_0 and A_1 have been determined,

$$x = A_0 e^{-t} + \varepsilon[-A_0 t e^{-t} + A_1 e^{-t}] + O(\varepsilon^2) \tag{29}$$

is a uniform asymptotic approximation in the outer region, i.e., $T_0 \leqslant t < \infty$

The solutions (28) and (29) are respectively called "inner" and "outer" solutions.

But how are A_0 and A_1 determined without knowing the exact solution? Here we invoke the asymptotic matching principle : set $\tau = t/\varepsilon$ in the first two terms of (28) and expand to two terms in ε ; set $t = \varepsilon\tau$ in the first two terms of (29) and expand to two terms in ε ; the results would match, in case regions of validity of (28) and (29) have an overlapping domain, i.e.

$$1 + \varepsilon\left(2 - \frac{t}{\varepsilon}\right) \text{ should be equal to } A_0(1 - \varepsilon\tau) + \varepsilon A_1$$

This implies that $A_0=1$, $A_1=2$

It is useful to have an asymptotic expansion which is uniformly valid in $0 \leq t < \infty$. The inner and outer expansions have terms in common, namely those which match. If we add the two expansions and subtract the common terms, the resulting composite expansion will be identical to the inner expansion where it is valid and to the outer where it holds. Since here is, presumably, overlap, the expansion is uniformly valid.

It is customary to subtract the common part from the inner solutions so that the part left decays exponentially on coming out into outer region. It then represents a correction of the outer solution to accommodate the boundary conditions. In the present case the result is

$$x = e^{-t} - e^{-\tau} + \varepsilon\,[(2-t)e^{-t} - (2+\tau)e^{-\tau}] + O(\varepsilon^2) \qquad (30)$$

Even the first term $e^{-t}-e^{-\tau}$ gives a good indication of the motion. There is a rapid rise to a peak at $t = \frac{\varepsilon}{1-\varepsilon}\log\frac{1}{\varepsilon}$ and then slow decay.

With linear equations the region of rapid change can only occur at boundaries ; hence the term boundary layer. The reason is that the solutions are locally exponentials and these grow in one direction if they die out in the other : to be an acceptable inner solution the growth must be blocked by a boundary. With nonlinear equations we may expect free layers and a famous example is the Van der Pol Oscillator for large tube conductance (relaxation oscillations) (*see* Cole 1968, p. 38).

Governing Equations for Atmospheric Motion

We would now apply Scale analysis to the equations of motion in the atmosphere which would enable us to determine the role played by various terms in bringing about the changes in the weather.

The basis of the equations of motion is Newton's second law. By this law the acceleration of a particle of air is related to the sum of the forces acting on it. Newton's law is applicable to an inertial system. This means a system which is fixed relative to the stars. But, meteorological observations are made at points which are fixed on the earth's surface, i.e., a rotating body. Consequently, Newton's law must be modified in such a way

that it applies to a system fixed relative to the rotating earth. Newton's law may be summed up by

$$\frac{d_a \mathbf{v}}{dt} = \sum_{i=1}^{n} \mathbf{F}_i \qquad (31)$$

where the left-hand side is the absolute acceleration in the absolute system, and the right-hand side is the sum of the forces acting on the particle which has unit mass.

If the earth is rotating with an angular velocity $\vec{\Omega}$ the absolute velocity $\mathbf{v}_a$ of a particle may be expressed as the sum of a velocity relative to the earth and a velocity due to rotation.

$$\mathbf{v}_a = \mathbf{v}_r + \vec{\Omega}\times\vec{r} \qquad (32)$$

where $\vec{r}$ is the position vector of the particle.

Hence,

$$\frac{d\vec{r}}{dt} = \mathbf{v} \qquad (33)$$

Consider a vector,

$$\mathbf{A} = A_x\mathbf{i} + A_y\mathbf{j} + A_z\mathbf{k}$$

where i, j, k are unit vectors in a Cartesian system.

From (32) and (33) it is seen

$$\frac{d_a\vec{r}}{dt} = \frac{d_r\vec{r}}{dt} + \vec{\Omega}\times\vec{r} \qquad (34)$$

It can be established (Haltiner and Martin, 1957) that this is true not only for $\vec{r}$, but for any vector.

Hence,

$$\frac{d_a \mathbf{v}_a}{dt} = \frac{d_r \mathbf{v}_a}{dt} + \vec{\Omega}\times\mathbf{v}_a$$

Substitute (32) in above,

$$\frac{d}{dt}(\mathbf{v}_a + \vec{\Omega}\times\vec{r}) + \vec{\Omega}\times(\mathbf{v}_a + \vec{\Omega}\times\vec{r}) = \frac{d_a\mathbf{v}_a}{dt}$$

Simplifying

$$\frac{d_a\mathbf{v}_a}{dt} = \frac{d\mathbf{v}}{dt} + 2\vec{\Omega}\times\mathbf{v} + (\vec{\Omega}\times\vec{\Omega}\times\vec{r})$$

The last term represents a centripetal acceleration due solely to the rotation of the earth. Note,

$$\vec{\Omega}\times\vec{\Omega}\times\vec{r} = -\Omega^2 R \quad (R = r\sin\theta)$$

The second term is the Coriolis acceleration. It is a fictitious force at right angles to v, and does no work. Newton's second law may be thus expressed as

$$\frac{d\mathbf{v}}{dt}+2\vec{\Omega}\times\mathbf{v}-\Omega^2 R=\sum_{i=1}^{n}\mathbf{F}_i$$

The external driving forces on the right-hand side are mainly

(a) gravity
(b) force due to pressure gradients
(c) friction.

Since both gravity and centripetal acceleration ($-\Omega^2 R$) are functions of the distance R from the axis of rotation, it is mathematically convenient to combine them into a single vector $\vec{g}$, whence

$$g\,\mathbf{k}=\vec{g}_a+\Omega^2\vec{R}$$

Introducing the other forces, the complete equation of motion for a rotating frame of reference may be written as

$$\frac{d\mathbf{v}}{dt}+2\vec{\Omega}\times\mathbf{v}=-\alpha\,\nabla p+g\,\mathbf{k}+\mathbf{F} \qquad (35)$$

where

$$\alpha=\frac{1}{\rho}$$

Equation of Continuity

Apart from Newton's second law another fundamental law is the conservation of mass. This is expressed by

$$\frac{d\rho}{dt}+\rho\,\nabla\cdot\mathbf{v}=0 \qquad (36)$$

To express the fact that the total derivative denotes differentiation following the motion, it is sometimes expressed as $\frac{D}{Dt}$, where

$$\frac{D}{Dt}(\ )=\frac{\partial}{\partial t}(\ )+\mathbf{v}.\nabla(\ )$$

Thus,

$$\frac{DF}{Dt}=\frac{\partial F}{\partial t}+\mathbf{v}.\nabla F \qquad (37)$$

Equation of State

A third fundamental law for the atmosphere is the equation of state. For a perfect gas, this is

$$p\alpha=RT \qquad (38)$$

although no perfect gas exists (38) generally suffices for most meteorological purposes. Although the atmosphere is a mixture of several gases, (38) is sufficient if R is a weighted average for all individual gases.

First Law of Thermodynamics

Omitting chemical, magnetic or electrical effects for the present, the first law of thermodynamics is expressed by

$$\delta Q=dU+\delta W$$

where

δQ=heat added per unit mass
dU=change in internal energy per unit mass
δW=work done by a unit mass of the system.

Here the symbols δQ, δW are used to denote the fact that they are not exact differentials. When work is done by a non-viscous gas, the above equation is

$$\delta Q=dU+p\,d\alpha$$

For a perfect gas, this becomes

$$\delta Q=C_v dT+p\,d\alpha$$

An alternative form is obtained by differentiating the equation of state

$$p\alpha=RT$$

whence

$$p\,d\alpha+\alpha\,dp=R\,dT$$

On substitution we find,

$$\delta Q=(C_v+R)\,dT-\alpha\,dp$$

or

$$\delta Q=C_p\,dt-\alpha\,dp$$

In an adiabatic process,

$$C_p\,dT-\alpha\,dp=0$$

or

$$C_p\frac{dT}{T}-R\frac{dp}{p}=0$$

Integrating the above from an initial state (p_0, T_0) to a final state (p, T) gives the following relation for an adiabatic process

$$\frac{T}{T^0}=\left(\frac{p}{p_0}\right)^{\frac{R}{C_p}}$$

This is known as Poisson's relation.

In meteorology the potential temperature (θ) is defined as the temperature attained by the air when it is brought adiabatically, either by expansion or compression to 1000 mb. From Poisson's relation it follows that

$$\theta=T\left(\frac{1000}{p}\right)^{\frac{R}{C_p}}$$

Summary

The above brief description of the physical law governing atmospheric motion suggests that there exists the right number of equations for the unknowns to form a closed system. Given appropriate boundary or initial conditions it is thus possible, in principle, to formulate meaningful problems, which seek to evaluate the future state of the atmosphere.

It would be interesting to see the disparate scales present in studies related to atmospheric sciences. The following tables will give some insight in that direction.

Typical Scales of Atmos. Motion System

	Type of motion	*Horizontal Scale*
I	Molecular mean free path	10^{-5} cm
II	Minute turbulent eddies	1-10 cm
III	Small eddies	10-10^2 cm
IV	Dust devils	10^2-10^3 cm
V	Gusts	10^3-10^4 cm
VI	Tornados	10^4 cm
VII	Thunder clouds	10^5 cm
VIII	Squall lines ; Fronts	10^6-10^7 cm
IX	Hurricanes	10^7 cm
X	Synoptic cyclones	10^8 cm
XI	Planetary wave	10^9 cm
XII	Atmospheric Tides, Mean Zonal Wind	10^9 cm

Characteristic Scales of Field Variables

(Mid latitude Synoptic System)

I	Horizontal speed	U	10^3 cm/sec
II	Vertical speed	w	1 cm/sec
III	Horizontal length	L	1000 km
IV	Depth scale	H	10 km
V	Time scale	L/U	10^5 sec~1 day
VI	Horizontal Pressure fluctuation	p	10^4 dynes/cm^2

Scales for Atmospheric Waves

	Classification	*Wavelength*	*Time Scale*	*Scale of Speed*
1.	Micro-Scale (Planetary Boundary Layer)	Few hundred metres	A few minutes	2-5 m/s
2.	Meso-Scale (Cumulus Thunder lake breeze and sea breeze)	A few km to a few hundred km	A few minutes to a few hours	15-25 m/s
3.	Synoptic Scale (Large Scale meteo. process, Rossby wave, monsoon depression, cyclones)	A few hundred to a few thousand km	A few days	10-20 m/s
4.	Macro-Scale (Ultra long waves, monsoon Troughness, Monsoon Trough)	Tens of thousands of km	Weeks and Seasons	10-15 m/s

Atmospheric Waves

I Longitudinal Waves (Compression Waves)

Sound waves	(Propagate both eastward and westward)
Prominent feature	Adiabatic compression and rarefaction-pressure gradient due to adiabatic compression
Phase speed	$c=\bar{u}\pm\sqrt{\gamma RT}$~330 m/s relative to the medium
Period	<270 sec

II Vertical Transverse Waves

Gravity waves	(Propagate both eastward and westward)
Mechanism	Pressure gradient due to variation in air density above a given horizontal surface—change in pressure due to solely variation of hydrostatic pressure
Phase speed	$c=\bar{u}\pm\sqrt{gH\left(1-\frac{P_2}{P_1}\right)}$ $H=1$ km to 10 km $\frac{P_2}{P_1}=0{\cdot}90$ to $0{\cdot}99$
Speed scale	10 m/s to 100 m/s essentially independent of wavelength and amplitude
Period	270 sec to 3 hours

III Horizontal Transverse Wave

Rossby wave	(Westward propagating w.r.t. medium)
Characteristic features	Speed depends upon wavelength, play important role for meteorology Associated with large-scale fluctuations in the atmospheric flow pattern which characteristically travel much slower than the high speed sound and gravity waves The relevant horizontal speed is at least 100 times (2 order) greater than the corresponding vertical speed
Phase speed	$c=\bar{u}-\frac{\beta}{\kappa^2}$ $\left(\beta\text{—Coriolis parameter},\ \kappa=\frac{2\pi}{\lambda},\ \lambda\text{—wavelength}\right)$
Period	>12 hours

(Period of atmospheric tides $=24/m$ hours, $m=1, 2, \cdots$, important for geomagnetic variations)

Because speed of Rossby waves depends on wavelength we consider Rossby wave in terms of wave numbers referred to the latitudinal circle at which study is made.

I Planetary waves (Ultra long waves)

Wave numbers	1 – 4
Wavelength	Tens of thousands of km, $=2\times\frac{2}{\gamma}\times 6370$ e.g., ($\lambda=2\pi$ a for wave No. 1 40,000 km $\lambda=2\pi$ a for wave No. 4 $\lambda\sim$ 10,000 km $\lambda=2\pi$ a for wave No. 8 $\lambda\sim$ 5,000 km) (Zonal flow : wave No. zero) Flow pattern in monsoon regions, monsoon trough corresponds to wave numbers 2 – 4

II Synoptic waves : Wave numbers : 5 – 8

Corresponds to normal monsoon disturbances such as monsoon depression, mid-tropospherical cyclones

Simplification of Governing Equation of motion by Scale Analysis

Order of magnitude (Large-Scale flow)

Consider a typical cyclone at about 45° N

Horizontal wind speed	10 m/s
Vertical speed	5 cm/s
Radius R of the storm	1000 km
Difference in pressure from the centre of cyclone to undisturbed surroundings	10 mb

$\frac{du}{dt},\ \frac{dv}{dt}\sim\frac{V^2}{R}$ (centripetal acceleration)

$$\frac{du}{dt}-\frac{uv\tan\phi}{a}+\frac{uw}{a}=-\frac{1}{\rho}\frac{\partial p}{\partial x}+2\Omega(v\sin\phi-w\cos\phi)+\text{i.f, CGS}$$

$$\frac{10^6}{10^8}\quad\frac{10^6}{6\times10^6}\quad\frac{5\times10^3}{6\times10^6}\quad\frac{10^3\cdot10^4}{10^8}\quad10^{-4}\ [10^3,5]\quad\frac{?\text{ cm}}{5^2}$$

$$10^{-2}\quad 1{\cdot}6\times10^{-3}\quad 10^{-5}\quad 10^{-1}\quad 10^{-1}\,5\times10^{-4}\quad\frac{?\text{ cm}}{5^2}$$

(Earth's radius $\sim$ 6400 km)

1 mb $=10^3$ dynes/cm^2
 $\simeq 10^3\times g$ gm wt/cm^2
 $\simeq 10^6$ gm wt/cm^2

Hence, to an accuracy of about 10%, we retain only first term on the left, and first and second on the right.

For the vertical equation, we estimate that vertical velocity might change sign in 12 hours and that 500 mb surface is at about 6 km.

$$\frac{dw}{dt} - \frac{u^2+v^2}{a} = -\frac{1}{\rho}\frac{dp}{dz} - g + 2\Omega u\cos\phi + \mathbf{k}\cdot\mathbf{f}_r$$

$\frac{10}{4\cdot3\times10^4}$	$\frac{2\times10^6}{6\times10^6}$	$\frac{10^3\times5\times10^5}{6\times10^5}$	10^3	$10^{-4}\cdot10^3$	$\frac{?\,cm}{s^2}$
$2\cdot5\times10^{-4}$	3×10^{-3}	10^3	10^3	10^{-1}	$\frac{?\,cm}{s^2}$

We need to retain only the first two terms on the right-hydrostatic approximation.

Thus, approximation equation for large-scale flows are

$$\frac{du}{dt} = -\frac{1}{\rho}\frac{\partial p}{\partial x} + fv$$

$$\frac{dv}{dt} = -\frac{1}{\rho}\frac{\partial p}{\partial y} - fu$$

$$\frac{\partial p}{\partial z} = -g$$

where

$$f = 2\Omega\sin\phi$$

Since frictional effects are neglected. We may also ignore heat conduction term and so we have

$$C_v\frac{dT}{dt} + p\,\nabla\cdot\mathbf{v} = -\nabla\cdot\mathbf{R}$$

$$\frac{d\rho}{dt} + \rho\,\nabla\cdot\mathbf{v} = 0$$

$$p = \rho\,RT$$

Notice, we no longer have an equation for $\frac{dw}{dt}$.

The first law may be combined with the equation of Stone to give

$$\frac{\partial p}{\partial t} + \mathbf{v}_H\cdot\nabla p - wg\rho - \frac{C_p}{C_v}\frac{p}{\rho}\frac{d\rho}{dt} = \frac{p}{C_v}\frac{ds}{dt}$$

using continuity equation

$$\frac{\partial p}{\partial t} + \mathbf{v}_H\cdot\nabla p - wg\rho + \frac{C_p}{C_v}p\left[\nabla_H\cdot\mathbf{v} + \frac{\partial w}{\partial z}\right] = \frac{p}{C_v}\frac{ds}{dt}$$

From hydrostatic condition

$$p = \int_z^\infty g\rho\,dz \qquad (p=0 \text{ at the top of atmosphere})$$

gives

$$\frac{\partial p}{\partial t} = \int_z^\infty g\frac{\partial\rho}{\partial t}\,dz = -\int_z^\infty g(\nabla\cdot\rho\mathbf{v})\,dz$$

$$= -\int_z^\infty g\nabla_H\cdot(\rho\mathbf{v})\,dz + g\rho w$$

(w vanishes at top of atmosphere)

Using this in the above equation, we obtain

$$\frac{\partial w}{\partial t} = -\nabla_H\cdot\mathbf{v} + \frac{C_v}{C_p}\cdot\frac{1}{p}\left[\int_z^\infty g\nabla_H\cdot\rho\mathbf{v}\,dz' - \mathbf{v}_H\cdot\nabla p + \frac{p}{C_v}\frac{ds}{dt}\right]$$

$$\therefore\quad w = \int_0^z\left[-\Delta_H\cdot\mathbf{v} + \frac{C_v}{C_p}\frac{1}{p}\left\{\int_{z'}^\infty g\nabla_H\cdot\rho\mathbf{v}\,dz'' - \mathbf{v}_H\cdot\nabla p + \frac{p}{C_v}\frac{ds}{dt}\right\}\right]dz'$$

where $w=0$ at $z=0$

Thus, vertical velocity is dependent on horizontal motion and the pressure field and is precisely the amount of vertical motion required to maintain hydrostatic equilibrium. This equation was presented by Richardson in his attempt to design a numerical prediction model long before computers were introduced.

Note :

1. Earth's rotation : linear term in u and v
2. Friction is present but it acts on much smaller scale than that of the motion
3. Effect of water vapour and its changes of phase have been ignored.

Most of the theory of the dynamics of atmospheric motion has been described by above equations. In actual practice, sphericity of earth is included by 2 further assumptions : (*a*) Cyclic continuity (periodicity) x, thus modelling the closed nature of atmospheric flow along a latitude cycle ; (*b*) We replace the variable Coriolis force with a linear approximation

$$f = f_0 + \beta y \qquad \left(\beta = \left.\frac{\partial f}{\partial y}\right|_{y=0}\right)$$

$$= \frac{\partial}{a\partial\phi}(2\Omega\sin\phi)_{y=0}$$

$$= \frac{2\Omega}{a}\cos\phi = \beta$$

Scale Analysis and the Geostrophic Approximation

If we retain the dominant terms in the last analysis, we have

$$fv=\frac{1}{\rho}\frac{\partial p}{\partial n},\quad fu=-\frac{1}{\rho}\frac{\partial p}{\partial y}$$

$$\therefore\quad \mathbf{v}_g=\frac{1}{\rho f}\mathbf{k}\times\nabla_H p,\ \text{i.e.,}\quad -\frac{1}{\rho}\nabla_H p=f\mathbf{k}\times\mathbf{v}_g$$

(Geostrophic wind $\mathbf{v}_g$)

(Greek : geo : earth, strophe : turning)

This can be used to replace pressure gradient by an equivalent velocity. Hence

$$\frac{d\mathbf{v}_H}{dt}=f\mathbf{k}\times\mathbf{v}_g-f\mathbf{k}\times\mathbf{v}+\mathbf{f}_{rH}$$

Hence, the actual wind will be geostrophic when $\frac{d\mathbf{v}_H}{dt}-\mathbf{f}_{rH}=0$ and so when parcels are unaccelerated and friction force vanishes, wind will be geostrophic. This situation is rarely satisfied in the atmosphere, but they are often approximately satisfied a km or more above the earth's surface and so the flow above is often geostrophic.

Equation of Motion

$$\frac{\partial\mathbf{v}_H}{\partial t}\times\mathbf{v}\cdot\mathbf{v}_H=f\mathbf{k}\times(\mathbf{v}_g-\mathbf{v}_H)+\gamma\nabla^2\mathbf{v}_H+\gamma\nabla_H(\nabla\cdot\mathbf{v})$$

where $$\mathbf{v}=\mathbf{v}_H+\mathbf{k}w$$

Characteristic scales $\mathbf{v}_H\sim V$, $w\sim W$ horizontal wavelength $\sim L$, vertical wavelength $\sim D$, period $\sim T$

$$\therefore\quad \mathbf{w}^*=\frac{\mathbf{v}_H}{V},\ x^*=\frac{x}{L},\ w^*=\frac{w}{W},\ y^*=\frac{y}{L},$$

$$\mathbf{u}_g{}^*=\frac{\mathbf{v}_g}{V},\ z^*=\frac{z}{D},\ t^*=\frac{t}{T}$$

Introducing the parameter

Rossby number $R_0=\dfrac{\text{Inertia force}}{\text{Coriolis force}}=\dfrac{V}{fL}$

and

Reynolds number $R_e=\dfrac{VL}{\gamma}=\left(\dfrac{\text{Inertia force}}{\text{Viscous force}}\right)$

Hence, we have

$$\frac{1}{fT}\frac{\partial u^*}{\partial z^*}+R_0\mathbf{u}^*\cdot(\nabla_H\mathbf{u}^*)+\frac{W}{V}\frac{L}{D}R_0\ \mathbf{w}^*\frac{\partial\mathbf{u}^*}{\partial z^*}=$$

$$\mathbf{k}\times(\mathbf{u}_g{}^*-\mathbf{u}^*)+\frac{R_0}{R_e}(\nabla_H^2\mathbf{u}^*)+\frac{L^2}{D^2}\frac{R_0}{R_e}\frac{\partial^2\mathbf{u}^*}{\partial z^{*2}}$$

$$+\frac{R_0}{R_e}\left(\nabla_H\nabla_H\cdot\mathbf{u}^*+\frac{W}{V}\frac{L}{D}\frac{R_0}{R_e}\nabla_H\frac{\partial w^*}{\partial z^*}\right)$$

For large-scale flow in middle atmosphere

$$R_0=\frac{V}{fL}=\frac{10\ \text{m/s}}{10^{-4}/\text{s}\ 10^6\,\text{m}}=0{\cdot}1$$

$$\frac{WL}{VD}=\frac{(0{\cdot}01\ \text{to}\ 0{\cdot}1)\,\text{m/s}\ 10^6\,\text{m}}{10\ \text{m/s}\times5\times10^3\,\text{m}}\sim0{\cdot}2\ \text{to}\ 2$$

$$R_e=\frac{VL}{\gamma}=\frac{10\ \text{m/s}\ 10^6\,\text{m}}{0{\cdot}15\times10^{-4}\,\text{m}^2/\text{s}}\sim6{\cdot}6\times10^{11}$$

$$\frac{D^2R_e}{L^2}\sim\left(\frac{5}{1000}\right)^2 6{\cdot}6\times10^{11}\sim1{\cdot}6\times10^7$$

$$|\mathbf{u}_g{}^*|\sim|\mathbf{u}^*|\sim1$$

Thus for large-scale flow in the atmosphere we may drop friction terms because $R_e\gg1$. Further $R_0\ll1$, hence, we have

$$\frac{1}{T}\frac{\partial u^*}{\partial t^*}=\mathbf{k}\times(u_g{}^*-\mathbf{u}^*)$$

and thus for time scales such that $T\gg\frac{1}{f}$ acceleration terms can be ignored.

Hence, criterion for geostrophic balance is that $R_0\ll1$, $R_e\gg1$ and the flow is slowly varying. This is obviously not applicable in tropics where Coriolis parameter becomes very small and vanishes at the equator, and second is small-scale motion such as atmospheric turbulence, which is not affected significantly by the rotation of the earth.

REFERENCES

1. Blumen, W. (1972) *Rev. Geophys. and Space Physics*, 10(2) : 485-528.
2. Carrier, G. F. (1953) *Adv. in Appl. Mech.*, 3 : 1-19.
3. Carrier, G. F. (1966) *J. Fluid Mech.*, 24 : 641-659.
4. Carrier, G. F. and Avrett, E. H. (1961) *Astrophys. Journal*, 134(2) : 468-81.

5. Cole, J. D. (1968) Perturbation methods in Applied Maths, Blaishdell Publ. Co.
6. Domaracki, A. and Loesch, A. (1977) *J. Atmos. Sci.*, 34 : 486-498.
7. Duffy, D. G. (1974) *J. Atmos. Sci.*, 31 : 1218-1231.
8. Jacobs, S. J. (1967) *J. Fluid Mech.*, 30 : 417-438.
9. Kiang, R. L. (1969) *Phys. Fluids*, 12 : 1333-1339.
10. Lindzen, R. S. (1971) *J. Atmos. Sci.*, 28 Part 1 : 609-622.
11. Murray, J. D. (1968a) *Appl. Sci. Res.*, 19 : 234-249.
12. Nayfeh, A. H. (1973) Perturbation methods, John Wiley and Sons.
13. Newell, A. C. (1969) *J. Fluid Mech.*, 35 : 255-271.
14. Concus, P. (1968) *J. Fluid Mech.*, 34(3) : 481-495.
15. Stone, P. H. (1969) *J. Atmos. Sci.*, 26 : 376-389.
16. St. Maurice, J. P. (1975) *J. Fluid Mech.*, 68, Part 3 : 417-445.
17. Shabbar, M. (1971) *J. Atmos. Sci.*, 28 : 345-349.
18. Tam, K. K. (1968) *J. Fluid Mech.*, 34 : 145-158.
19. Yajnik, K. S. (1970) *J. Fluid. Mech.*, 42 : 411-427.

The Living Mutagens

G. K. Manna

National Fellow
Department of Zoology, University of Kalyani
Kalyani, West Bengal

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

Mutation often called "sport" is the building block of evolution, a change that may be inherited in a Mendelian manner though "somatic mutation" obeys different principles. Since the discovery of the molecular structure of the DNA[171], our attention has been directed more to physiochemical approach of mutation[176]. Conventionally, mutations are classified as (i) gene mutation, molecular mutation or point mutation and (ii) chromosome mutation[135]. White[173] opined that besides evolutionary polyploidy, mutations are of three types : (i) point mutation, not visible even in the highest magnified view of chromosome, is relative to rearrangement or substitution at the level of the individual nucleotide pairs in the DNA molecule ; (ii) gene duplication of the type studied in *Chironomous* species and (iii) chromosomal rearrangements involving many genetic loci visible under microscope. According to him as the satellite DNAs may differ in closely related species, e.g., of *Chironomous*[78,80], large-scale deletion and duplication of this material seemed to play a separate role in addition to the two conventional forms mentioned before. The conservativeness of heredity is counteracted by mutation and the genetic control of a character is known when the gene mutates. In nature the rate of mutation is extremely low, one in 2000 or so.

Since the pioneering discovery of the artificial production of mutation by X-rays in *Drosophila*[134] and in barley[161] and by chemicals in plants and animals[10,148] mutations were distinguished as spontaneous and artificially induced ones deploying some mutagen. The present advancement of knowledge on mutagenesis leaves hardly any scope to think that mutations may occur spontaneously because no environment is free from mutagens of one kind or the other [103,113]. Mutagens are now put under three main categories, viz., (a) physical, (b) chemical and (c) living ones[99,100] though a fourth category as "genetic mutagen" should also be recognised. The progress in the field of radiation mutagenesis is very extensive and has been reviewed time to time[13,29,47,67,76,86,114,155,160,174]. The same is true of chemical mutagenesis which covers a still wider area[9,20,40,45,48,81,82,98,104,151,158,169,174]. In contrast, progress in the field of genetic mutagens is limited. It was shown that certain genes might initiate mutation in another, e.g., the "dotted" gene, *Dt* located in short arm of chromosome 9 induced the gene *a* located in chromosome 3 to mutate to *A* in maize. The gene *A* is necessary for anthrocyanin production[157] has been discovered. Another genetic mutagen, known as *Ac-Ds* system, *Ac* stands for activator

action and *Ds* for producing chromosome breaks. In the presence of *Ac*, in the same cell, the *Ds* locus would produce chromosome breaks studied genetically in marked stock *G* in maize. In *Drosophila*, the presence of some gene influencing the effect of other is known. Anyhow, the genetic mutagens differ from other mutagens as the latter come from extrinsic source.

The mutagenic effect of living organisms was first shown by Hampar and Ellison[61] who reported the chromosomal aberrations induced by an animal virus in mammalian cells. This study started studies on the hazardous effects of living mutagens, the importance of which is yet to be realised. The remarks of Halkka[60] significantly remarked that the true agents responsible for spontaneous mutations are obscure and the commonness of viral and other intracellular infection in animals and plants necessitates serious consideration of microbial mutagenesis. The study on the effects of living mutagens on mammalian models is of special significance since advocates of the study of the hazardous effects of environmental mutagens have paid little attention to these[108].

Since microbes like viruses, rickettsia, mycoplasma, bacteria, pathogenic fungi, etc., have been found to induce chromosome mutation in host cells of the infected or experimentally treated animals, they are included under the category of living mutagen[100,113]. According to Cruickshank[36], microorganisms are put into two classes as (1) Microtatobiotes which include two orders, Virales and Rickettsiales, and (2) Schizomycetales bacteria comprising five orders, namely (a) Actinomycetales, (b) Spirochaetales, (c) Mycoplasmatales, (d) Pseudomonatales and (e) Eubacteriales.

I. Virus-Induced Chromosome Aberrations

Since the discovery of the virus-induced chromosome aberrations[62] investigations have been mainly carried out on the changes in the somatic chromosomes in mammalian *in vivo* and *in vitro* system[5,15,113] because of its relevance to carcinogenesis and teratogenesis[23,31,137]. The study on germinal effect has been limited to insects[25]. The discovery of chromosomal aberrations by the attenuated virus in live-vaccine drew attention to undesirable mutational risks[88,128,129,130]. The virus-induced dominant lethal testing in mammals is yet to be undertaken while the study on the point mutation induced by them on *Drosophila* is not free from controversy[18,25,145].

Viruses (Latin)=poison) are simplest and smallest among microbes, measuring from 0·1 to 0·3 micron in diameter and parasitic to animals, plants and bacteria. Characteristically they lack biosynthetic and energy yielding enzymes and possess only one kind of nucleic acid, DNA or RNA and not both. According to Zinsser Microbiology[71], animal viruses are grouped into (A) DNA containing viruses comprising six major families, viz., (i) Poxviridae (ii) Harpesviridae (iii) Adenoviridae (iv) Papovaviridae (v) Iridoviridae and (vi) Parvoviridae of which the chromosome breaking effects of Irido and Parvo viruses have not been tested, and (B) RNA containing viruses divided into ten major families, viz., (i) Picornadiridae (ii) Togaviridae (iii) Reoviridae (iv) Orthomyxoviridae (v) Paramyxoviridae (vi) Rhabdoviridae (vii) Retraviridae (viii) Bunyaviridae (ix) Arenaviridae and (x) Coronaviridae of which viruses belonging to last three families have not been tested for chromosome breaking effects.

(a) *Chromosome Breaking Effects of DNA Viruses*

Among DNA containing viruses Vaccinia and Variola of Poxviridae produce all types of aberrations, including gross effect in blood culture of patients as well as cultured cells infected with Vaccinia[5,156], while treatment of attenuated small-pox virus in life-vaccine induced breaks and diffused type of chromosomal alterations[128].

Among Herpesviridae, chromosome aberrations induced by Herpes simplex virus were reported in the infective cells of Chinese hamster by Hamper and Ellison[61,62] who reported that the treatment did not cause cytopathic or mitotic inhibition but in the successive cell generations, there were significant chromosome aberration in the form of chromatid breaks and diffused alterations. The effect of this virus has been further tested on chromosomes of *in vitro* and *in vivo*, bone marrow cells of mice[14], kidney cells of monkey[21], embryonic rats[69], Chinese hamster[126], and on human culture cells[12,131,144] and Hela cells and HEp-2[152]. The aberrations comprised breaks of various kinds and diffused alterations, leading to pulverization, pycnosis, etc., but there has been reported of polyploid change. Some workers[69,103] found that the Herpes virus-induced breaks were somewhat localised and nonrandom in occurrence. O'Neill and Rapp[144] found that G_2 phase of DNA cycle of the host cell was mostly affected by the treatment with Herpes simplex virus. Huang[70] observed frequent pulverisation

effects caused by Herpes simplex and Cytomegalo virus. The effect of chicken-pox virus, Varicella studied only on human chromosomes of patients and cultured cells[5,55,150] showed predominently break type aberrations while Harnden[68] failed to observe these aberrations. With Herpes zoaster virus, break type aberrations and pulverization effects occurred in the cultured cells of human lung[10] but Harnden[68] failed to see such effects.

In the family Adenoviridae, several types of adenoviruses have been tested for their chromosome breaking activity. In a comparative study of the chromosome breaking effects in human and hamster cells by adenovirus type 2, 4, 7, 12 and 18 and Simian adenovirus type 7 and Simian virus 15, Cooper *et al.*[54,55] found that all of these induced more or less similar, localised chromosome aberrations and gross effects but others found that adeno type 7, 12 and 18 induced, in addition, the numerical changes in chromosomes[43,78]. Nichols' observations with adeno type 2 were different but not with adeno type 12. Stich and Yohn[164] studied the mutagenic capacity of adeno type 2, 4, 12 and SA 7 and SV 15 and found that the effect was of diffused nature and comprised breaks in most cases. The effects of adeno type 12 has also been studied on chromosomes of embryonic rat[69] and on chromosomes of Chinese and Armenian hamster[28,78]. Adenoviruses cause respiratory trouble, pharyngitis and also tumourous changes in cells[165,178]. Aula *et al.*[8] observed various types of chromosomal aberrations by adeno virus. Some workers claimed that adeno type 12 induced-breaks were localised and nonrandomly distributed[69,180].

The Papovaviridae includes tumour forming DNA viruses. Prunieras *et al.*, showed both structural and numerical changes in chromosome complement of mammalian cells infected with Papilloma shope virus. Some workers[44,170] have observed structural changes in chromosomes in neoplastic transformations of hamster embryonic cells by Polyoma virus. The chromosome breaking effects of SV 40 (Simian virus) belonging to this family has been studied in human diploid cells[52,152,154] and in bovine lung fibroblast[148]. Wolman *et al.*[175] recorded localised breaks, pulverization and heteroploid changes in chromosome of human fibroblast culture infected with SV 40. Yerganian[177] recorded similar chromosome aberrations in Armenian hamster cells infected with this SV 40 while Lehman and Defendi[87] noted early chromosome damages in Chinese hamster induced by the same virus. The breaks were nonrandomly distributed between and within the chromosome in both human and hamster cells[172,177].

Thus the nature of chromosome aberration induced by the four families of DNA containing viruses are basically similar and comprise breaks and other structural changes besides gross effects in the whole chromosome complement and heteroploid alterations. The oncogenic Papovaviruses and certain adenoviruses (type 12) showed aberrations comparable to those of chemically induced ones, specially those obtained by the treatment of base analog of DNA and antimetabolites[88,89,99].

(b) *Chromosome Breaking Effects of RNA Viruses*

Among RNA containing viruses in the family Picornaviridae, Poliomyelitis viruses type 1, 2 and 3 occur in man and monkey and Columbia SK occurs in man and these have been tested for their chromosome breaking effects by different workers. Bartsch *et al.*[16] observed breaks and diffused aberrations in chromosomes after infection with Poliovirus type 1. Mice immunised poliovaccine type 2 contained chromosome aberration in the form of breaks, constrictions, heteroploidy and physiological effects of diffused nature, affecting the entire chromosome[130]. However, Aula and Nichols[7] failed to observe specific forms of chromosome aberrations in cells infected with polio type 3. Nachtigal *et al.*[136] found both breaks, constrictions and heteroploidy in cell culture in normal and in chronic infection with Columbia SK virus.

In the family Togaviridae the chromosome breaking effects by arboviruses and rubella viruses have been studied by different workers. The chromosome breaking effect of mosquito-borne yellow fever virus belonging to genus Flavirus was shown in the bone marrow cells of mice due to some hepatocephalotrophic infection[14], while in man the cytogenetic effect due to infection and vaccination against yellow fever was studied by Harnden[68] who found both gross effects and breaks, constrictions, etc., in certain individual chromosome. The tick-borne enchephalitis virus of arbo group induced heteroploid changes while the effect on individual chromosome structure was negligible in human diploid cells strain carrying persistent infection of the virus[90]. Several workers studied chromosome aberrations in human cells induced by rubella viruses on cultured embryonic tissue[29], in other *in vivo* and *in vitro* cells[81], in leucocyte culture of patients[84] and in congenital rubella patients[142]. The effects consisted mainly of breaks, gaps, etc., involving structural damages in certain parts while diffuse effect and heteroploidy

were not very obvious. In some cases different workers found different types of aberrations[22,63,128].

In the family Reoviridae, chromosome breaks and heteroploidy was produced by reoviruses in normal and chronically infected cell cultures[136]. Keast[77] observed that heteroploid chromosome number in lymphoma (2731/L) resulted from infection of reovirus type 3 in neonatal mice.

In the family Orthomyxoviridae, the effect of human influenza virus did have marked chromosome breaking effect while influenza viruses A(PR-8) were introduced in the serial passages of human embryonic lung culture[51], in human diploid cell strain and in BHK-21 cells[167].

In the family Paramyxoviridae, parainfluenza virus type 1 and 3, Newcastle disease, mumps and measles viruses parasitic to man were tested for their chromosome breaking effect by different workers. Cantell *et al.*[27] reported predominant occurrence of gross type of chromosome damages in Hela cells induced by the treatment of mumps and parainfluenza virus 1 and 3. Renda[150] found various types structural damages in human chromosome infected with mumps and similar effects by the infection of measles virus were reported *in vivo* and *in vitro* human cells by Aula[5], while pulverization effects, in addition to structural changes, were reported by others[140,141]. Several workers found the pulverization effects induced by parainfluenza virus (Sendai) in human chromosomes[6,75,153,162] and in chromosomes of kidney cells of rabbit[125] while Homma *et al.*[69] noted genomic changes induced by this virus. The infection of Newcastle disease virus of chick was found to induce chromosome pulverization in treated Hela cells[27] and also pulverization and heteroploidy in human cell line[166].

Michajlova and Gorsunova[129] reported diffused type of chromosome aberrations to be caused by rabbies virus (family Rhabdoviridae) while Marquez-Monter *et al.*[122] did not notice any localised break in chromosomes as a result of this virus.

The family Retraviridae includes oncogenic RNA containing viruses, causing tumours in chick, mice, etc.[71]. The chromosome breaking effect of several strains of Rous sarcoma virus have been tested by different workers. According to Kato[73,74] the administration of Schmidt-Ruppin strain of RSV into Chinese hamster induced all kinds of chromosome aberrations, including heteroploid changes while Ahlstrom *et al.*[2] observed only heteroploid changes. Mark[121] also reported similar chromosomal changes induced by this strain in mice while Boue *et al.*[24] obtained both structural

and numerical changes in human cell culture transformation and Nichols *et al.*[140] found localised break-type aberrations. Kato[78] studied effect of Mill-Hill strain of RSV and found both localised break and pulverization in Chinese hamster chromosome which were similar to those of SR strain mentioned above. On the other hand, certain workers[147,148] did not find such chromosome aberration in Bryan strain treated bovine lung fibroblasts. Similarly nonconfirmatory result was reported in the transformation of monkey embryonic tissue culture treated with Carr strain of RSV[1]. The same was true for Engelbreth-Holm strain of RSV with bovine lung fibroblast culture[146]. The chromosome damaging effects induced by several strains of murine-leukemia viruses were studied by different workers. Sofuni *et al.*[159] in their study on the Friend virus strain of MLV recorded all types of chromosomal damages including genome changes while Tsuchida and Rich[168] reported earlier only pulverization and heteroploid changes of chromosome. The latter authors also found similar chromosomal effects with the treatment of Rausher strain of MLV. Joneja and Stich[72] observed chromosome breakages and heteroploidy but not pulverization by the treatment of Gross virus MLV while Boiron *et al.*[21] did not find structural changes in chromosome in cultured cells of monkey treated with Moloney virus, but the transformation of cells took place.

(c) *Chromosome Breaking Effect of Miscellaneous Groups*

In regard to several viruses, it has not been settled if they contain DNA or RNA. Joklik and Willet[71] put them into a miscellaneous group. Hepatitis viruses come under this category. Clinically recognised types A and B contain RNA and DNA respectively but the chromosome breaking properties of hepatitis viruses have been tested by different workers. Kerkis *et al.*[76] and Aya and Makino[11] found in the leukocyte culture of patients suffering from acute hepatitis, chromosome aberrations in the form of structural changes, pulverization and genome modifications. El-Alfi *et al.*[46] also made similar observation in human leukocyte culture. Barinsky and Dementiev[14] found various types of break-type aberrations in the bone marrow chromosomes of mice infected with hepatoenchephalotrophic viral infection. Chromosome breaks and mitotic suppression were noticed in infectious hepatitis by other workers as well[124,127]. Yoshida *et al.*[179] reported chromosome aberrations of various kinds in cultured blood and kidney cells of dog, infected with canine hepatitis virus.

The types of aberrations induced by different viruses in the chromosomes of man and laboratory mammals, mainly in the cultured cells reviewed above, possess some common features and some special features. Aberrations induced by viruses are predominantly of chromatid type and manifested as breaks, constrictions, gaps, etc., localizable at some region of the affected chromosomes much in the same way as in the case of certain chemical mutagens, including some nucleic acid base components, analogues and antimetabolites of DNA and RNA[33,69,82,99,107,108,139,140]. Extreme regionwise localization of breaking point in chromosomes have been shown to be induced by viruses and chemical mutagens in various mammalian chromosomes including man[33,69,82,101,102]. However, it has been claimed that the breaks induced by viruses undergo delayed restitution. Various suggestions have been forwarded to explain virus-induced breakages in chromosomes. Of these the hypothesis of Zur Hausen[180] regarding the direct enzymatic effect of some virus protein has limited applicability, while the role of untimely liberation of lysosomal hydrolytic enzymes in host cells by the infection of viruses proposed by Allison[8] appears to be more attractive. However, the latter also does not have universal applicability as Bartsch *et al.*[17] observed that the breaks occurred in higher frequency in the presence of a polio dependent lysosomal enzymes as well as in the absence of additional lysosomal enzymes in their study of poliomylitis virus-induced breaks. A more attractive hypothesis proposed was that virus-induced inhibition of cellular synthesis, specially DNA synthesis[17,33,140], this hypothesis has wider applicability to correlate the data of aberration induced by chemical and viral agents. Moreover, the virus-induced impairment of cellular DNA, RNA and protein synthesis would have wider acceptability because it can explain virus-induced gross effects like pulverization, inhibition of mitosis and even numerical changes.

Besides the common occurrence of structural changes in chromosomes, genome changes leading to heteroploidy, sometimes polyploidy, induced by almost all groups of viruses (though more by heterogenous oncogenic groups)[15] were related mainly with mitotic disturbance. Endoreduplication occurred sometimes, but C-mitotic configuration and gross effect. Genome changes are regarded to have adaptive value in genome selection applicable for virus-induced ones[54] in cancer[92,95,97]. Structural and genome changes caused by infection or otherwise reduce chances of survival of aberrated

cells, but variations due to chromosomal changes may affect adaptive selection in changed environment[97].

In contrast to the structural and numerical changes commonly induced by viruses, other effects, described here as gross effects, diffused effect, etc., are produced in which the entire chromosome is involved. Of these, pulverization type evokes special interest as some workers characterized its occurrence by viral agents only, while Levan and Tjio[89] recorded it as chemical effect. Pulverization of chromosome has been reported in human and other mammalian materials induced by measles virus, mixovirus, mumps, cytomegalo virus, etc. Stenman and Saksela[162] opined that it is a precocious and incomplete contraction of chromatin functionally yet to mature and is possibly caused by a triggering effect of the mitotic nucleus fused into the same cytoplasm. However, this type of effect along with other gross effects like stickiness, pycnosis, despiralisation, etc., were seen to be produced not only by viruses but also by bacteria[100, 109], rickettsia[58], mycoplasma[49] and by toxic chemicals[99] and physical agents. Since the effect involved surface structure and coiling of chromosome, some workers[4] suggested that it is due to the breaking of the DNA-histone binding. Since pulverization occurs along with cytopathic effects, this process of degeneration has been attributed to proteolysis, followed by depolymerisation of DNA in which the role of lysosome seems implicit[88].

Research interest in the study of virus-induced cytogenetical effect seems to have dwindled after a short period of vigorous studies between 1960 and 1970. However, considerable more attention is now being paid to the study of virus-induced transformation of cells. The question is relatively controversial and complicated[15]. Recently, Nichols[188], using virus-induced transformation as a model, opined that the genetic effect of the transforming virus in the infected cells would help to probe into its role in the genetic alterations of virus-induced oncogenesis. He advocated that the discovery of the virus-induced sister chromatid exchange and gene mutation would provide a broad spectrum of genetic alterations caused by viruses. This diversion in the study of the virus-induced cellular transformation is perhaps be due to the fact that oncogenic viruses have shown chromosomal alterations after the lapse of considerable time and even by the time cell-transformation has occurred. Thus Bartsch[15] opined that "carcinogenesis by way of somatic mutation is unlikely". On the other hand, Aula *et al.*[6] and Manna[100] suggested that the mutation theory proposed by Bovery and viral theory

proposed by Rous in the etiology of cancer could be rolled into one since we know now that virus can induce mutation to cancer in host cells so also the cell itself in case of human cancer of nonviral origin. It may, however, be pointed out that chromosome aberrations induced by all types of viruses could not have the same bearing in inducing cancer and controversy exists as to whether chromosome aberrations is the cause or the effect of cancer[97]. Anyhow the etiology of cancer still remains unsolved.

The second outcome of the study of virus-induced chromosome aberrations is their implication in teratogenesis. Rubella and Cytomegalovirus infection has been found to alter the foetal "anlage", possibly by somatic mutation which led to malformation in the offspring, but the possibility of damage to the "anlage" due to cytopathic effect cannot be ruled out. The toxic product produced by virus infection may affect the metabolism of host cell. However, Bartsch[15] is in favour of considering the teratogenic effect to be due to somatic mutations caused by viruses, although while several workers have observed chromosome aberrations to be induced by Rubella but not by Cytomegalovirus. Mutational potentiality of viruses to the host cells would be strengthened from our study on bacteria-induced chromosome aberrations (vide *infra*).

It bas already been pointed out that virus-induced point mutations have been reported only in *Drosophila*[85] but not in mammalian models. Studies on the latter should be worthwhile because hereditary effects have already been documented in case of bacterium *Pseudomonas aerugenosa* treated mice by Manna and Dey (vide *infra*) and have important implication in public health.

II. Rickettsia-Induced Chromosome Aberrations

In the class Microtatobiotes, besides viruses, species belonging to the order Rickettsiales[36] have rarely been tested for their chromosome breaking activity in host cells. Rickettsial infection causes typhus which often influenced the course of history[71]. Charles the First, abandoned his plan to march on London while the French Army occupied Prague due to mass deaths of defenders due to epidemic typhus[71]. The first microscopic study of Rickettsia was made in the beginning of the 20th century by Dr. Howard Ricketts who discovered the agent of Rocky mountain spotted

fever and cultivated the causative organism *Rickettsia rickettsii* on laboratory animals. Three genera *Rickettsia, Rochalimae* and *Coxiella* under Rickettsiales are parasitic on man. Of these the cytogenetic effect of only *Rickettsia* has been investigated. Rickettsia are pleomorphic, rod-shaped, coccoid organisms, measuring $0{\cdot}3-0{\cdot}8\times0{\cdot}6$ mμ. In ultrastructure and chemical composition, they resemble gram positive bacteria. Unlike viruses, they possess both DNA and RNA and their respiratory and synthetic capabilities are independent in the cell. They are placed between viruses and bacteria.

Rickettsia grylli-induced chromosome aberration have been studied by Halkka and his collaborators in the male meiotic cells of the gryllid *Acheta* sp.[58,60]. True break-type aberrations were rare and the effect was manifested in the form of stickiness, clumping, pulverization of chromosomes, etc. They were more or less of the same nature as induced by Mycoplasma (vide *infra*).

III. Bacteria-Induced Chromosome Aberrations

According to Cruickshank[36] bacteria, class Schizomycetes comprise five orders, viz., (1) Actinomycetales, (2) Spirochaetales, (3) Mycoplasmatales, (4) Pseudomonadales and (5) Eubacteriales. So far as the present author is aware there is no report on chromosome aberrations induced by Actinomycetes and Spirochaetes though antibiotics derived from Actinomyces have been found to induce chromosome aberrations in both plants and animals[58,104,174]. Actinomycetales are known as higher bacteria, having a mycellium and special cell for reproduction while the remaining orders are called lower bacteria.

(a) *Effect of Mycoplasma*

Mycoplasmatales or PPLO (Pleuro-Pneumonia-Like Organism), among lower bacteria, are widespread in nature and form an indigenous microbial flora of the orthopharynx and genito-urinary tract of birds and mammals, including man and the type species *M. mycoides* was isolated from cattle. Mycoplasma are unicellular prokaryotic organisms, lacking a cell wall but bounded by a triple-layered membrane. The size ranges between 0·5 and 1 μ while the shape is changeable. Except for the absence of cell wall, its fine structure is similar to that of bacteria[71]. It is the causative agent for human atypical pneumonia.

The progress in the study of Mycoplasma induced chromosome aberration like Rickettsia is very limited. It was studied in the male germ cells of two insects *Nepa* and *Acheta*[59]. The meiotic anaphase chromosomes showed changes somewhat comparable to those caused by radiation. The aberrations consisted of laggard, acentric fragments, stickiness translocations, inversions, aneuploidy, polyploidy, cytoplasmic budding, without chromosomes, etc. The effect on goniomitosis in rat-kangaroo cells was studied by Lastikka *et al.*[86] and on somatic chromosomes of man by several workers. Paton *et al.*[146] studied chromosomal changes in the cultured human cells while Fogh and Fogh[49] found in modified line of human amnion cells irreversible pulverization. Kindsin[83] studied chromosomal aberrations induced by T-strain Mycoplasma.

According to Halkka and Heinonen[59], two properties of Mycoplasma infection, viz. arginine degradation and cleavage of thymidine to thymine could be responsible for chromosome damage because arginine is an important components of histone and its degradation might affect the histone of chromosome. Anyhow, the type of chromosome aberration induced by mycoplasma has basic similarities with those of viruses and rickettsia.

(b) *Effect of Other Lower Bacteria*

According to Joklik and Willet[71] and Cruickshank[36], lower bacteria constitute the most important members of bacteria. These microorganism never have mycelium, and are characterised by having naked, nonmembrane-bound, nuclear DNA without associated basic protein and a semirigid cell wall of unique composition. The cell lacks mitochondria and other cell organelles. The single layered cell wall with adherents, known as capsules, forms the cell envelope. The cell size ranges between 0·4 and 1·5 μ in diameter and in many forms, the relatively rigid cell wall gives them the characteristic shape on which their classification is based, e.g., Cocci- spherical, Bacilli- rod-shaped, Vibrios- curved rods, Spirilla- spirally twisted rods, and Spirochaetales- spirally twisted flexous filaments. The biological classification is based on Linnaean scheme and will be followed here, while genetic classification based on nucleotide base composition has also been advocated[71]. Quite a few species of bacteria, like *E. coli, Salmonella, Pneumococcus*, etc., form important materials for genetical researches and the progress of

microbial genetics within the short span of time has been very startling. Presently, we would be concerned with their role as mutagens.

Progress in the study of the lower bacteria-induced chromosome aberrations, besides mycoplasma, has almost wholely been due to Manna and his collaborators, using mice as the main experimental animals[30,37,41,56,57,64,65,98,99,100,104,105,109,110,111,113,114,115,116,117,118,119,120]. The only other observations are by Makino *et al.*[98] in leucocyte culture of patients suffering from scarlet fever, the likely causative agent being *Streptococcus pyogenes*. The results obtained by using different human pathogenic bacteria, some toxin and vaccine are briefly reviewed here.

(i) *Pseudomonas-Induced Chromosome Aberrations*

In the order Pseudomonadales, among two families Pseudomonadaceae and Spirillaceae, the chromosome breaking property of one species each has been studied by us. *Pseudomonas aeruginosa*, the pathogenic species of man out of 140 species belonging to this genus, is frequently present in the intestinal flora and also on skin of man and animal. It is a gram-negative, non-sporing bacillus, measuring 0·5 to 1·5 μ. It moves by a single flagellum and the only species that produces the chloroform soluble pigment pyocianin[30]. A number of enzymes and toxins, in addition to slime endotoxins, produce pathogenic effects on animals. It has been claimed now that its infection might lead to the induction of cancer [52]. Though chromosome breaking effects have been shown by us[41,111], the cytotoxic effects of leukocydenes of this pathogen, tested on human polymorpholeukocytes, were shown within a few minutes. Culture of *P. aeruginosa* at the log phase, isolated bacteria suspended in saline, log culture filtrate after removal of all bacteria and the heat-killed bacterial suspension were tested for chromosome breaking affects on bone-marrow cells of mice against suitable controls. The treatment proved fatal in some cases while others had full life though some treated parent mice and their progeny developed physical abnormalities like kingking and shedding of tail, loss of hind legs, etc., in later part of life. The chromosome aberration data of the adult, and parent and their progeny, the result of dominant lethal testing and the effect on sperm count have breifly been reviewed.

The aberrations induced by the different samples of *P. aeruginosa* in the bone-marrow cells of treated mice consisted of subchromatid, chromatid

and rarely isochromatid breaks, fragments of unknown origin and translocations taken as break types, while constrictions and gaps occurred in some individual chromosomes and there were metaphases with gross effect consisting of stickiness, pycnosis, polyploidy, pulverization, etc. The aberration frequency at 14 intervals between 10 min and 192 hr among 4200 metaphases examined was 16·4% individual type (378 break and 314 physiological), and 17·6% gross type in the log culture series ; 21·6% individual type (428 break and 480 physiological) and 8·8% gross type in saline suspended series ; 24·4% individual type (386 break and 640 physiological) and 3·1% gross type in culture filtrate series and 12·3% individual type (265 break and 251 physiological) and 2·5% gross type in heat-killed bacterial suspension injected series. As compared to nutrient broth, saline and distilled water injected controls, the frequency of aberrations was appreciably high in all the treated series leaving not the least doubt that the *Pseudomonas* samples induced very high frequency of aberrations. The individual type aberrations were predominantly chromatid type.

Chromatid breaks were nonrandomly distributed between and within chromosomes. Out of 927 total breaks of all treated series 164 were in Gr I, 387 in Gr II, 285 in Gr III, 86 in Gr IV and 5 in Gr V against the expected number of 120, 213, 464, 111 and 19 respectively, calculated as per relative percentage length ($p < 0{\cdot}001$) and 93, 139, 464, 185 and 46 respectively as per proportional number of chromosomes ($p < 0{\cdot}001$) in different groups[99,106]. Chromosome belonging to Gr I and Gr II were more vulnerable while those of remaining groups were quite resistant to the treatment. Further, the distribution of breaks along the length of chromosomes in 3 equal segments of proximal, middle and distal was also nonrandom because the number was 16, 69 and 79 respectively. Out of 164 breaks in Gr I ; 40, 211 and 136 respectively among 387 breaks in Gr II ; 30, 134 and 121 respectively among 285 breaks in Gr III ; 23, 42 and 21 among 86 breaks in Gr IV ; 3, 1 and 1 among 5 in Gr V and 112, 457 and 358 respectively among the grand total of 927 chromatid breaks. With random distribution each region should have had 1/3 of the total number. Thus it was evident that the proximal region was fairly resistant while the mid-region was most vulnerable.

The most interesting finding in *Pseudomonas* treated mice was the continuation of more or less the same type of aberrations in the F_1 and F_2 generations of the treated parents. As controls, a male and two females in

each set were injected with fresh nutrient broth and allowed mating. One male parent sacrificed on day 120 had in 300 metaphases of bone-marrow one gross type effect (0·3%), one F_1 male examined on 16th day after birth had in 300 metaphases only one constriction (0·3%) while a 30-day old F_2 male in this control line had no aberration in 300 metaphases. In the treated line, the parents were injected with log culture of *Pseudomonas* (1756×10^7/ml) at the rate of 1 ml per 100 gm body weight. The bone-marrow of the male parent sacrificed on 95th day had among 405 metaphases, 129 individual type aberrations and 7 plates with gross effect (33·5%) while one 1-day old F_1 litter had in 137 metaphases, 5 chromatid aberrations and a sticky metaphase (4·3%). In the 2nd set of treated parents, the male sacrificed on 120th day had, in 300 plates, 76 individual aberrations and 8 plates with gross effects (28%) while one 7-day old F_1 litter of this parent had in 90 metaphases, 22 individual and 4 gross type aberrations (28·8%) ; another 11-day old litter had in 80 metaphases only 8 individual type aberrations (10%) and a third 16-day old F_1 litter had in 50 metaphases 9 individval and 2 gross type aberrations (21·1%). Peculiarly enough, one 30-day old female litter of F_2 produced by mating of F_1 in the treated line had in 245 metaphases, 89 individual and 6 gross type (21·1%) aberrations.

The injection of double dose (2 ml/100 gm body weight) of log culture of *Pseudomonas* into different sets of parents had, in the male parent of 1st set on 60th day, out of 300 metaphases 37 individual and 9 gross type aberrations (15·3%) and another 23-day old F_1 litter had in 300 metaphases, 51 individual and 9 gross type aberrations (20%). One 21-day old litter of F_1 of the second treated pair had in 300 metaphases 49 individual and 5 gross type aberrations (18%). The male of the 3rd treated set had on day 120 out of 300 metaphases 60 individual and 7 gross type aberrations (22·3%). One 7-day old F_1 litter of this pair had in 73 plates only 31 individual type aberrations (34·2%) while another 11-day old F_1 litter had in 120 plates, 16 individual and 3 gross type aberrations (15·8%). The 3rd 16-day old F_1 litter of this parent had in 73 plates only 31 individual type aberrations (42·4%). It was surprising that one 30-day old male F_2 litter obtained from F_1 in this double dose treated line had in 170 metaphases only 34 individual type aberrations (20%). The control line of the double dose series had 1·3% in 120-day old male and 0·3% in a 16-day old F_1 litter and 0·3% in a 30-day old F_2 litter. Thus treatment of *Pseudomonas*

definitely induced chromosome mutations which perpetuated in F_1 and F_2 generations. The occurrence of the same type of aberrations in two successive generations after treatment would suggest, pending further verification, that the treatment-induced certain loci to mutate which very likely control the aberrations. This resembles the action of genetic mutagens like the effect of *Ac-Ds* genes in maize on *C*-locus[9,1].

That *Pseudomonas*-induced chromosomal mutations were heritable was also supported from the dominant lethal testing. Male mice treated at different time of spermatogenesis with log culture of *Pseudomonas* were introduced to virgin females at 9 intervals between 48 hour and 6 week after the treatment and the pregnant females were sacrificed on the 15th day of gestation. Out of 97 total implants in the treated series, 44 were dead and 53 were alive while in the corresponding series, out of 155 implants only 6 were dead and 149 were alive. The difference in the number of dead implants between the treated and control lots was statistically significant at 0·05 level. The present data indicates that the bacterial sample induced both somatic and germinal mutations.

The study of the sperm count of the treated and control male at 9 intervals between 48 hour and 6 week yielded in the epididymal sperm suspension preparation the number 485 in treated males against 861 in controls, indicating that the treatment caused loss of considerable number of sperms.

(ii) *Vibrio Cholerae-Induced Chromosome Aberrations*

Among two genera *Vibrio* and *Spirilum* in the family Spirillaceae, chromosome-breaking effects of *V. cholerae* have been studied by us. *V. cholerae* or the comma bacillus is the causative organism of Asiatic cholera which sometimes occurs in epidemic form. It measures 0·5 to 1·5μ, with the ends slightly pointed.

Mice injected with 1 ml dose (37×10^6 cells/ml) had in 6 intervals between 15 min. and 72 hr., among 1700 metaphases 43 break-type and 26 physiological and 71 gross-type aberrations (8·2%) in the log culture series ; out of 2000 metaphases 51 break-type, 30 physiological-type and 104 gross-type aberrations (9·6%) in the saline-suspended series ; among 1500 metaphases 37 break-type, 24 physiological-type and 50 gross-type aberrations (7·4%) in cultured filtrate series, indicating that the effect was quite appreciable. Qualitatively the effect was more like *Pseudomonas*-induced

chromosome aberrations. The treatment of *V. cholerae* was fatal to the specimens as they hardly survived far beyond 72 hours. The analysis of the chromatid breaks revealed that it was nonrandomly distributed between and within the chromosomes[87, 88].

Effects of Species of Bacteria under Eubacteriales

Among lower bacteria the Order Eubacteriales is formed of 9 families. Of these the mutagenic property of no member of the families Achromobacteriaceae, Brucellaceae, Bacteroidaceae, Corybacteriaceae and Bacillaceae has so far been studied. The effect of only some species of the Enterobacteriaceae, Micrococcaceae, Lactobacillaceae, Neisseriaceae and Bacillaceae or their toxins have been studied by us. In the family Enterobacteriaceae, chromosome breakages induced in mice by *Klebsiella pneumoniae*, *Salmonella typhi*, *Salmonella paratyphi* have been studied.

(iii) *Klebsiella Pneumoniae-Induced Chromosome Aberrations*

This gram negative species is responsible for less than 1% of bacterial pneumonia but the fatality is very high. Treated mice also did not survive much longer than 48 hours. The chromosome aberrations were more or less of the same type as those caused by *V. cholerae* but the frequency of gross effect was low. The frequency of aberrations determined in 5 intervals between 1 hr and 48 hr had in 1500 metaphases, 34 break-type and 1 constriction (2·3%) in log culture treated series (84×10^9 cells/ml), 39 breaks and 4 physiological type (2·8%) in saline-suspended bacteria injected series, 13 breaks and 4 physiological type in culture filtrate injected series. (1·1%) and 12 breaks and 4 physiological type (1.06%) in heat-killed bacteria suspension injected series[56, 118].

The effects of the treatment of *K. pneumoniae* culture on mitotic index were interesting as the bacteria seemed to have accelaratory action rather than inhibitory action as is the case with other bacterial speceis (vide *infra*). The average frequency was 3·36% against 2·14% in control.

(iv) *Salmonella Typhi-Induced Chromosome Aberrations*

Salmonella typhi is a gram-negative, non-sporing motile bacillus with numerous long peritrichous flagellae which measures between 1·4 by 0·5 μ and does not possess a capsule. It is the causative agent of typhoid fever. The administration of this bacterium to mice proved fatal within a few

hours so that the study of chromosome aberrations was rather restricted in scope [89, 109].

The aberrations consisted of chromatid-type breaks, constrictions fragmentation, etc., as well as pulverization, pycnosis, etc. The frequency of aberrations seen in 300 metaphases, fixed at 6 hour was 4% individual type and 59·3% gross type in log culture, 1% individual and 79·3% gross type in isolated bacteria in saline suspension and 1·3% individual and 63% gross type in culture filtrate treated series. Thus the effect was not very severe so far as the individual type was concerned but the gross effect was very high. The injection of lower dose (0·5ml) induced 4·7% individual and 11·4% gross type in log culture, 5·3% individual and 18·7% gross type in suspension and 4·4% individual and 14·2% gross-type aberrations in culture filtrate treated series. Thus the frequency was higher in higher does of 3 treated series.

(v) *Salmonella Paratyphi A-Induced Chromosome Aberrations*

Log culture of *S. paratyphi* A was relatively less toxic to mice. It induced chromosome aberrations more or less similar to those of *S. typhi*[87, 110]. Out of 1800 metaphases of 6 intervals between 1 hr and 48 hr, the frequency of individual types was 2·1% and that of gross type 26·8% while the treatment of heat-killed suspension in 11 intervals between 30 and 168 hours yielded out of 3300 metaphases 1·06% individual types and 2·1% gross-type aberrations which was a little higher than in the control series.

(vi) *Staphylococcus Aureus-Induced Chromosome Aberrations*

Of the four genera of Micrococcaceae *Staphylococcus aureus* has been tested for its chromosome breaking activity[80, 105, 112, 113]. *S. aureus* are gram-positive spherical cocci arranged in irregular clusters. There are strain differences in virulence. Some are strongly toxic.

The aberrations induced by the treatment of *S. aureus* in bone-marrow chromosomes mainly consist of chromatid-type breaks, constrictions, gaps, fragments, etc., while the gross effects were limited. Mice treated with log culture and fixed at 20 intervals between 5 min. and 312 hr. yielded 6000 metaphases, 155 breaks and 91 physiological types (4·1%). The treatment with isolated bacteria in saline yielded at 29 intervals between 5 min. and 312 hr., among 9000 metaphases, 417 break-type and 241 physiological-type (7·6%) aberrations ; culture filtrate induced in 22 intervals, among 6600

metaphases, 213 break-type and 152 physiological-type (5·5%) and the heat killed sample yielded in 15 intervals between 5 min. and 312 hr., among 4500 metaphases, 83 breaks and 79 physiological-type aberrations (3·6%) indicating that the *S. aureus* samples were mutagenic to bone-marrow chromosomes of mice.

As in other bacteria treated series chromatid breaks were nonrandomly distributed between and within the chromosomes because out of 425 breaks analysed 27 were in Gr I, 106 in Gr II, 277 in Gr III, 14 in Gr IV and 1 in Gr V against the expected number of 42·5, 85·0, 212·5, 63·7 and 22·2 respectively ($P < 0·001$). Further, breaks were also nonrandomly distributed in different regions because out of the 425 breaks, 29 were in the proximal, 284 in the middle and 112 in the distal region against the expected number of 142 breaks in each region ($P < 0·001$). However the proximal half of the chromosome was less vulnerable than the distal half.

With a view to finding out the chromosome-breaking effect of samples of *S. aureus* in some other mammals, Syrian hamster, albino laboratory rat, wild rat, dog and monkey were injected the same dose of log culture and at 7 hr, the frequency was 6·3%, 5·3%, 34·3%, 5% and 33·4% respectively. The treatment of saline-suspended bacteria and the culture filtrate also yielded, in 300 metaphases each, at 7 hr 6% and 6% aberrations in Syrian hamster, 2·6% and 4·3% in albino rat, 4·3% and 6% in wild rat, 4% and 8% in dog and 33·4% and 32·4% in monkey respectively, showing that the mutagenic substance of *S. aureus* was also effective in other mammals and the effect was not host specific.

The effect of the log culture on the mitotic frequency in bone-marrow cells revealed the average frequency of 2·04% in normal, 2·14% in control and 2·01% in treated mice while in Syrian hamster, the frequency was 1% in the treated specimens. Though the inhibitory effect was not seen in mice, it was quite evident in hamsters[64].

With a view to finding out the dose-dependant effect of the aberration frequency, log culture (513×10^7 cells/ml) of *S. aureus* was injected into Syrian hamsters with different doses of which 1 ml dose yielded 2·5% aberrations, 2 ml dose yielded 5·08% aberrations and 3 ml dose 6·9% aberrations. The treatment of saline suspended bacterial samples also yielded 1·9% aberrations in 1 ml dose, 3·8% in 2 ml and 4·5% in 3 ml doses while the injection of culture filtrate yielded 3·5% aberrations in 1 ml dose, 7·5% in 2 ml and 8·1% at 3 ml dose. Thus in all the three treated

series, there was disproportionate increase in the frequency with the higher doses and the rise in frequency was relatively less between 2 ml and 3 ml doses. Fractionation experiment with 2 ml dose given in 1 ml doses at an interval of 1 hr, yielded more or less the same frequency of aberrations as that with 2 ml dose given at a time. It was 5·4% and 5·08% in log culture, 3·1% and 3·8% in saline suspension and 12·0% and 13·3% in culture filtrate respectively (Manna and Hazra—*unpublished*).

The dose-dependant effects tested on bone-marrow chromosome of mice also showed disproportionate increase in frequency with higher doses since in log culture series, it was 10·3% with 1 ml, 12% with 2 ml and 13·3% with 3 ml ; in saline suspension series it was 9·3% with 1 ml, 6·6% with 2 ml and 11·3% with 3 ml and in culture filtrate series it was 9·3% with 1 ml, 10·3% with 2 ml and 12% with 3 ml.

(vii) *Neisseria Meningitidis-Induced Chromosome Aberrations*

Among the two genera of this family, the chromosome-breaking effects of *Neisseria meningitidis* have been studied by us[50,115,117]. Morphologically this species is an oval diplococcus with opposite surfaces concave, and measuring about 0·8 μ in diameter. The natural habitat is nasopharynx but it is the causative agent for meningitis when spread into meninges.

The injection of this bacterium to mice is fatal a little beyond 8 hours. The aberration types induced by its treatment were similar to those of other bacteria and comprised mainly chromatid type breaks, gaps, constrictions, fragments and a few translocations and metaphases with gross effects. In bone-marrow cells of mice, the treatment of log culture yielded in 4 intervals between 30 min. and 8 hr. among 1200 metaphases, 68 break-type and 68 physiological-type (11·3%) aberrations, isolated bacteria in saline suspension yielded, in 1200 metaphases, 76 break-type and 70 physiological-type aberrations (12·1%) and the culture filtrate yielded out of 1200 metaphases, 29 break-type and 52 physiological-type aberrations (6·7%) while the treatment of heat killed bacterial suspension in 6 intervals between 30 min. and 24 hr. yielded, among 1800 metaphases, 25 break-type and 3 physiological-type aberrations (1·5%).

The treatment of the log culture in mice had retarded effect on the mitotic frequency because the average frequency in fixation interval was 1·4% against 2·09% in control. The maximum inhibition was 1·25% at 12 hours.

There was dose-dependant effect as the frequency of aberration in 0·5 ml and 1 ml doses respectively, was 7·6% and 11·1% in log culture series, 7·5% and 11·6% in saline-suspended series and 5·2% and 6·7% in culture filtrate series.

Effect of Species of Bacteria under Lactobacillaceae

Among 4 genera belonging to this family, the chromosome-breaking effects of two species belonging to *Streptococcus* have been studied by us[56,57,116,119].

(viii & ix) *Streptococcus Pyogenes and Streptococcus Faecalis-Induced Chromosome Aberrations*

Streptococcus are gram-positive spherical or oval cells arranged in pairs. It is the principal etiologic agent in tonsilitis and scarlet fever. *S. faecalis* leads a commensal existence in the human and animal intestine but may cause urinary tract infection, and sometimes causes subacute bacterial endocarditis. Mice treated with these two bacterial species did not survive much longer than 12 hours.

The aberration types induced by the treatment of the two species comprised chromatid-type breaks, gaps, constrictions, fragments and a few translocations. Some metaphases showed pulverization and pycnotic effects, the frequency of which was not determined. The frequency of aberrations in *S. pyogenes* in 5 intervals between 10 min. and 12 hr. and *S. faecalis* in 4 intervals between 30 min. and 12 hr. was 5·5% and 6·5% in log culture, 13·4% and 53% in saline suspended, 7·8% and 4·4% in culture filtrate and 1·06% and 19% respectively in heat-killed bacteria treated series, indicating slightly higher effect in *S. pyogenes* treated series.

In both species, chromatid breaks were nonrandomly distributed between and within the chromosome. In *S. pyogenes* out of 217 breaks, 24 were in Gr I, 58 in Gr II, 130 in Gr III, 5 in Gr IV and nil in Gr V against the expected number of 21·7, 43·4, 108·5, 32·5 and 10·8 respectively ($P<0.05$). Out of the 217 breaks, 22 were in the proximal, 108 in the middle and 87 in the distal region against the expected number of about 72 breaks in each region ($P<0.01$). In *S. faecalis* out of 74 chromatid breaks, 8 were in Gr I, 14 in Gr II, 51 in Gr III, 1 in Gr IV and nil in Gr V against the expected number of 7·4, 14·8, 37, 11·1 and 3·7 breaks respectively. The nonrandom

distribution was not significant ($P<0.05$) but the sample was small. However, the analysis of length-wise distribution of chromatid breaks revealed the phenomenon to be significantly nonrandom ($P<0.05$) because out of 74 breaks, 5 were in the proximal, 45 in the middle and 24 in the distal region against the expected number of about 25 breaks in each region. As with other bacteria, the distal half was more vulnerable while the proximal half was fairly resistant.

Treatment with log cultures of *S. pyogenes* and *S. faecalis* retarded the mitotic frequency in the bone-marrow cells of mice but it was more pronounced in the case of the former species. The average frequency was 1·03% and the peak effect was 0·9% at 10 min in *S. pyogenes* and 1·74% with the maximum retardation of 1·45% at 6 hours in *S. faecalis*.

IV. Bacterial Toxin-Induced Chromosome Aberrations

The chromosome-breaking effects of some bacterial toxins in the root tip cells of *Vicia faba* was reported quite early[188]; we have studied the same on bone-marrow chromosomes of mice. The dose used was arbitrary.

(a) Cholera toxin : Toxin of *Vibrio cholerae* diluted as mg/ml saline was intraperitoneally injected into mice at the rate of 1 ml/100 gm body weight and their bone-marrow cells, fixed at 12 intervals between 15 min. and 192 hr. yielded in 3600 metaphases, 50 break-type aberrations like subchromatid and chromatid break fragments and translocations and 5 gaps and constrictions, making 1·6% individual-type aberrations while there were 365 metaphases with gross effects (10·1%) of which 304 comprised pulverization. As compared to controls, the gross-type effect was very severe, (Manna and Das, S. K. —*unpublished*).

(b) Gas-gangrene toxin : Toxin of *Clostridium perfringens* (Bacillaceae) was diluted as 10 mg in 1% peptone solution and this stock solution was diluted further and injected into mice at the rate of 1 ml/100 gm body weight. The chromosome-breaking effect (Das and Manna,—*unpublished*) in bone-marrow cells in stock solution injected series was, in 600 metaphases, 3 break type and 3 gaps and constrictions (2%) while 356 metaphases showed gross effect (59·3%) a good number of which consisted of polyploidy and pulverization. Injection of 1 : 9 dilution of the stock solution induced, among 700 metaphases, only 2 fragments (0·3%) but 239 metaphases with gross effect (34·1%). The injection of 1 : 30 dilution yielded, in 800

metaphases, 15 break type and 8 gaps and constrictions, making a frequency of 2·9% while the number of metaphases with gross effect was 229 (28·6%). Injection of 1 : 50 dilution yielded, among 800 metaphases, 17 break type and 7 gaps and constrictions (3%) while the number of metaphases with gross effect was 187 (23·3%). Thus the injection of higher doses of toxin induced disproportionately higher frequency of aberrations because the total frequency was 26·3% at 1 : 50 dilution, 31·5% in 1 : 30 dilution, 34·4% in 1 : 9 dilution and 60·3% in stock solution injected series. The data of toxin-induced chromosome aberrations revealed that the gross-type effect was strikingly high and among this pulverization effect was the highest. On the other hand, chromosome aberration data induced by most of the bacteria indicated that they were less cytopathic than that of their toxins since the frequency of individual types of aberration was considerably high while gross effect was relatively low. Therefore, the aberrations induced by the different bacteria might not solely be due to the toxin liberated by bacteria. The real mutagenic factor is yet to be found as the frequencies were considerably high in log culture, saline suspension and culture filtrate series while it was as low as in the control in the heat-killed treated series.

V. Vaccine and Formalin Killed Bacteria Induced Chromosome Aberrations

Mice were injected with human dose equivalent of TABC (*S. typhi*, *S. paratyphi* A, B and *V. cholerae*) vaccine and their bone marrow, fixed at 6 intervals between 15 and 48 hr., yielded in 1800 metaphases 20 individual type breaks and 10 gaps and constrictions (1·6%) and 25 gross-type aberrations (1·4%). The data indicated that the vaccine contained some mutagenic property which was also concluded from chromosome aberrations in persons vaccinated against viral diseases[68,125,129,130].

VI. Pathogenic Fungus Induced Chromosome Aberrations

Since we have been reviewing the chromosome-breaking effects of living mutagens, the effect of pathogenic fungus on chromosome of mice come within our scope (Manna and Patel—*unpublished*). So far as the present author is aware, chromosome-breaking effect of pathogenic fungus has not been reported by earlier workers but this effects and the antimitotic property of some toxins on human cultured cells have been reported[48].

Reiss[151] also reported the chromosome-breaking effect of aflatoxin B in onion root tip cells. Aflatoxins, the toxic metabolites of a few fungi, e.g., *Aspergillus flavous*, *A. paraciticus*, *Penicillum puberulum* have been recognised as 8 compounds (aflatoxin B_1, B_2, B_2a, G_1, G_2, G_2a, M_1 and M_2). The biological

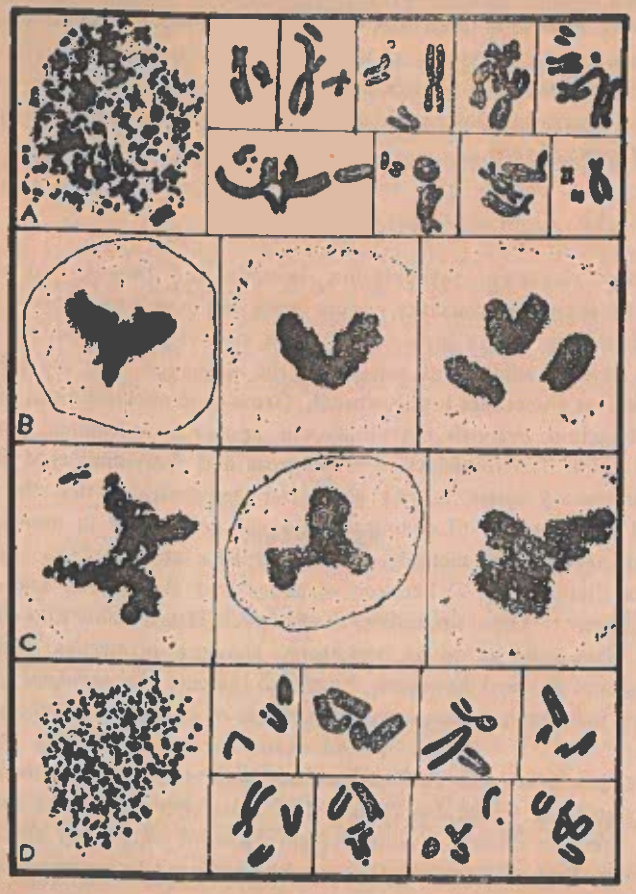


Fig. 1. *A Row* : Some representative type of virus-induced chromosome aberrations in man redrawn from Makino and Aya (1968) ; *B Row* : Virus-induced polykaryomitosis in man redrawn from Heneen *et al.* (1970) ; *C Row* : Polykaryomitosis in human cancerous cervix uteri redrawn from Manna (1962) ; *D Row* : bacteria-induced chromosome aberrations in mice drawn from Manna and collaborators.

effects of aflatoxins on cell culture and their role in carcinogenesis have been studied by some workers[48].

Intraperitoneal injection of the cultured spores of *Aspergillus fumigatus* into mice (Manna and Patel—*unpublished*) yielded in their bone-marrow cells, fixed at different intervals, the same types of chromosome aberrations as found in bacteria-induced ones viz., chromatid breaks, constrictions, gaps, fragments, translocations, etc., as individual type while some metaphases had gross effects in the form of stickiness, pycnosis, pulverization, etc. The frequency of aberrations increased considerably within 12 hours and it dropped with the lapse of time.

Concluding Remarks

Viruses, rickettsia, mycoplasma, bacteria and fungus, and in some cases their toxins and vaccines, induce some common form of chromosome aberrations (Fig. 1) like gaps, constrictions and fragments, chromatid type breaks while isochromatid ones and the exchange type are relatively infrequent or sometimes totally absent. Gross-type aberrations in the form of pulverization, pycnosis, polyploidy, etc., appear more frequently in toxin-treated series. The frequency of both gross and individual type differ in different treated series. It has also been demonstrated that the breaks induced by a number of chemical mutagens are similar in nature. Even the polykaryomitotic metaphase and anaphase abnormalities appear to be quite similar (Fig. 1) between virus-induced aberrations and human cervix cancer[96]. These similarities in the effects clearly show that microbes can be considered as living mutagens, showing properties similar to chemical and physical mutagens. They also indicate the existence of some common factor responsible for the induction of chromosome aberrations. Manna[99,100,101,102,104], has advocated a hypothesis that the odd mutagen produces some physicochemical stress which possibly breaks the chromosomes at vulnerable points. The gross effects could be produced by the effect on surface structure of chromosome which may also lead to true break, heteroploidy, etc. Certain chemicals and living mutagens have been found to induce localised break in the heterochromatic region as well as some other region of chromosomes in different materials[8,11,88,69,73,82,99,101,102,140,168] which led Manna[101] to postulate that there are some inherent weak regions which are liable to break oftener than other regions

due to physicochemical stress rather than due to specific interaction between the segment and the mutagen. Cohen and Shaw[88] also opined that the specific localised breaks are not so much due to specific action of certain agents but rather due to the specific reaction of certain chromosomal segments which Manna[101,102], envisaged as inherent weaker region in chromosome.

The hereditary effects in the progeny of *Pseudomonas* treated mice shown by us and in *Drosophila* induced by some virus[25,145], no doubt establish the claim that microorganisms may act as mutagen both to somatic and germinal chromosomes much as the chemical mutagens. Virus-induced point mutation has been suggested[18] while we also suspect some such events as the same type of chromosome aberration was seen to perpetuate in two generation. Thus, a systematic study on the hazardous hereditary effects of mutagens need be done if we really want to have a comprehensive knowledge of environmental mutagens[103]. Studies on the effects of living mutagens on mammalian models are all the more important because we are being frequently infected by these mutagen during our life time.

REFERENCES

1. Adzhighitov, F. I., Dzhemiev, Z. A. and Kosichenko, L. P. (1967) *Vop. Virus*, **12** : 439.
2. Ahlstrom, C. G., Kato, R. and Levan, A. (1964) *Science*, **44** : 1232.
3. Allison, A. C. (1967) *Sci. Amer.*, **217** : 62.
4. Arrighi, F. E. and Hsu, T. C. (1965) *Exp. Cell Research*, **39** : 305.
5. Aula, P. (1965) *Suomal. Tiedeakat.* Sr. A4, **89** : 1-75.
6. Aula, P. and Saksela, E. (1966) *Hereditas*, **55** : 362.
7. Aula, P. and Nichols, W. W. (1968) *Exp. Cell Res.*, **51** : 595.
8. Aula, P., Nichols, W. W. and Levan, A. (1968) *Ann. N.Y. Acad. Sci.*, **155** : 737.
9. Auerbach, C. (1967) *Science*, **158** : 1141.
10. Auerbach, C. and Robson, J. M. (1946) *Proc. Roy. Soc. Edin. B.*, **62** : 271.
11. Aya, T. and Makino, S. (1966) *Proc. Jap. Acad.*, **42** : 648.
12. Aya, T., Makino, S. and Yamada, T. (1967) *Proc. Jap. Acad.*, **43** : 239.
13. Bacq, Z. M. and Alexander, P. (1961) *Fundamentals of radiobiology*, 2nd Ed., Pergamon Press, Oxford.
14. Barinsky, L. F. and Dementiev, I. V. (1968) *Acta. Virol.*, **12** : 464.

15. Bartsch, H. D. (1970) Virus induced chromosomal alterations in mammals and man, Art. 26, pp. 420-432, *In Chemical mutagenesis in mammals and man*, (Eds.) F. Vogel and G. Rohrborn, Springer-Verlag, New York.
16. Bartsch, H. D., Habermehl, K. O. and Diefenthal, W. (1967) *Exp. Cell Res.*, **48** : 671.
17. Bartsch, H. D., Habermehl, K. O. and Diefenthal, W. (1969) *Arch. Ges. Virusforsch.*, **27** : 115.
18. Baumiller, R. C. (1967) *Nature*, **214** : 806.
19. Benyesh-Melnick, M., Stich, H. F., Rapp, F. and Hsu, T. C. (1964) *Proc. Soc. Exp. Biol.* (N.Y.), **117** : 546-549.
20. Biesele, J. (1958) *Mitotic poisons and cancer problem*, Amsterdam Elsevier Publ. Co.
21. Boiron, M., Tanzer, J., Thomas, M. and Hampe, A. (1966) *Nature*, **209** : 737.
22. Boue, J. G., Boue, A., Moorhead, P. S. and Plotkin, S. A. (1964) *C. R. Acad. Sci.* (Paris), Sr. D, **259** : 687.
23. Boue, A., Plotkin, S. A. and Boue, J. G. (1965) *Virusforsch.*, **16** : 443.
24. Boue, J. G., Boue, A., Montagnier, L. and Vigier, P. (1968) *C. R. Acad. Sci.* (Paris), Sr. D, **266** : 178.
25. Burdette, W. J. and Yoon, J. S. (1967) *Science*, **155** : 340.
26. Burdette, W. J. and Yoon, J. S. (1968) *Am. Assoc. Cancer Res.*, **9** : 38.
27. Canteli, K., Saksela, E. and Aula, P. (1966) *Ann. Med. Exp. Fenn.*, **44** : 255.
28. Cassingena, R. and Tournier, P. (1965) *Ann. Inst. Pasteur*, **108** : 277-297.
29. Catcheside, D. G. (1948) *Adv. in Genet.*, **2** : 271.
30. Chakraborti, S. (1974) Effects of *S. aureus* on the dividing bone-marrow cells of *Mus musculus*, Ph.D. Thesis, University of Kalyani.
31. Chang, T. H., Moorhead, P. S., Boue, J. G., Plotkin, S. A. and Hoskins, J. M. (1966) *Proc. Soc. Exp. Biol.* (N.Y.), **122** : 236.
32. Chumakov, M. P., Mustafina, A. N., Chumakova, Y. A., Karmysheva, N. M., Shestopalova and Reinfold, V. N. (1964) *Acta. Virol.*, **8** : 217.
33. Cohen, M. M. and Shaw, M. W. (1965) Specific effects of viruses and antimetabolites on mammalian chromosomes, pp. 50-66, *In* G. Yerganian, *The Chromosome, Boston Publ. of Tissue Cult. Assoc. Inc.*
34. Cooper, H. L. and Black, P. H. (1963) *J. Nat. Cancer Inst.*, **30** : 1015.
35. Cooper, J. E. K., Yohn, D. S. and Stich, H. F. (1968) *Exp. Cell Res.*, **53** : 225.
36. Cruichshank, R. (1970) *Medical microbiology*, 11th Ed., The English Book Society.
37. Das, S. K. (1976) Effects of some pathogenic bacterial culture, toxin and vaccine on bone-marrow chromosomes of mice, Ph.D. Thesis, University of Kalyani.
38. Das, S. K. and Manna, G. K. (1974) *Proc. 43rd Ann. Sess. Nat. Acad. Sci. India*, Jodhpur, Abs., p. 20.
39. Das, S. K. and Manna, G. K. (1975) *Proc. 44th Ann. Sess. Nat. Acad. Sci. India*, Bhagalpur, Abs., p. 5.
40. Darlington, C. D. (1953) *Symposium on Chromosome breakage*, Heredity Suppl., Oliver and Boyd.
41. Dey, M. and Manna, G. K. (1978) *Proc. 3rd All India Congr. Cytol. Genet.*, Hissar, (*In press*).

42. Dipaolo, J. A., Rabson, A. S. and Malmgren, R. A. (1968) *J. Nat. Cancer Inst.*, **40** : 757.
43. Dolimpo, D. A. (1968) *Proc. Soc. Exp. Biol.* (N.Y.), **127** : 559.
44. Dulbecco, R. (1965) Interaction of viruses with genetic material of the host cells. *In Reproduction, molecular subcellular and cellular*, (Ed.) M. Locke, Acad. Press, N.Y.
45. Eigsti, O. J. and Dustin, P. (1955) *Ames*, Iowa State College Press.
46. El-Alfi, O. S., Smith, P. M. and Biesele, J. J. (1965) *Hereditas*, **52** : 285.
47. Evans, H. J. (1962) *Intern. Rev. Cytol.*, **13** : 221.
48. Fishbein, L., Flamm, W. G. and Falk, H. L. (1970) *Chemical mutagens—environmental effects on biological system*, Acad. Press, N.Y.
49. Fogh, J. and Fogh, H. (1967) *Proc. Soc. Exp. Biol.* (N.Y.), **126** : 67.
50. Freese, F. (1963) Molecular mechanism of mutations, pp. 207-269. *In Molecular Genetics*, (Ed.) J. H. Taylor, Pt. I, Acad. Press, N.Y.
51. Gavrilov, V. I., Yershow, F. I., Blyumkin, V. N., Kvorov, I. I., Levina, D. S. and Zmiyeva, R. G. (1965) Peculiarities of morphogenesis of the CA-SV 40-63-I cell line, *Vop. Virus.*, **10** : 323.
52. Gaydos, J., Michael, Lee, Carrick, J. R. and Richards, B. (1976) *Proc. Soc. Exp. Biol. Med.*, **149** : 908.
53. Gebhart, E. (1970) The treatment of human chromosomes *in vitro* : results, pp. 367-382, *In Chemical mutagenesis in mammals and man*, (Eds.) F. Vogel and G. Rohrborn, Springer-Verlag, N.Y.
54. Genest, P. and Daniel, P. (1966) *Proc. Soc. Exp. Biol.* (N.Y.), **123** : 722.
55. Grover, S., Solanki, B. K., Khandelwal, M. K. and Jaiswal, R. B. (1968) *Indian Pediat.*, **5** : 326.
56. Gupta, S. (1976) Effects of four pathogenic bacteria on the chromosomes and mitotic index in bone-marrow cells of mice and its alterations by antibiotics, Ph.D. Thesis, University of Kalyani.
57. Gupta, S. and Manna, G. K. (1975) *Proc. 61st Ind. Sci. Congr.*, Pt. 3 : 28.
58. Halkka, O. (1967) *Hereditas*, **58** : 248.
59. Halkka, O. and Heinonen, L. (1968) *Hereditas*, **58** : 253.
60. Halkka, O., Meynadier, G., Vago, C. and Brummer-Korvenkontio, M. (1970) *Hereditas*, **64** : 126.
61. Hampar, B. and Ellison, S. A. (1961) *Nature*, **192** : 145.
62. Hampar, B. and Ellison, S. A. (1963) *Proc. Nat. Acad. Sci.*, **49** : 474.
63. Harnden, D. G. (1964) *Amer. J. Human Genet.*, **16** : 204.
64. Hazra, M. K. (1979) The bacterium *Staphylococcus* induced mammalian chromosome aberrations and their alterations by the antibiotic, penicillin, Ph.D. Thesis, University of Kalyani.
65. Hazra, M. K. and Manna, G. K. (1978) Bone-marrow chromosome aberration in four mammalian species injected with *Staphylococcus aureus.*, *Proc. 3rd All India Congr. Cytol. Genet.*, Hissar (*In press*).
66. Heneen, W. K., Nichols, W. W., Levan, A. and Norbby, E. (1970) *Hereditas*, **64** : 53.

67. Hollaender, A. (1954) *Radiation Biology*, McGraw-Hill Publ.
68. Homma, M., Ohira, M. and Ishida, N. (1968) *Virology*, 34 : 60.
69. Huang, C. C. (1967) *Chromosoma*, 23 : 162.
70. Huang, C. C. (1971) *Mamm. Chrom. Newsletter*, 12 : 7.
71. Joklik, W. K. and Willet, H. P. (1976) *Zinseer Microbiology*, 16th Ed., Prentice-Hall International, Inc., Englewood Cliffs.
72. Joneja, M. G. and Stich, H. F. (1965) *Exp. Cell Res.*, 40 : 148.
73. Kato, R. (1967) *Hereditas*, 58 : 220.
74. Kato, R. (1968) *Hereditas*, 59 : 63.
75. Kato, H. and Sandberg, A. A. (1968) *J. Nat. Cancer Inst.*, 41 : 1117-1123.
76. Kaufmann, B. P. (1954) Chromosome aberration induced in animal cell by ionizing radiations. *In Radiation Biology*, (Ed.) A. Hollaender, pp. 627-711, McGraw-Hill, N.Y.
77. Keast, D. (1966) *Exp. cell Res.*, 44 : 295-302.
78. Kerkis, Ju. Ja., Sablina, O. V., Radzabli, S. I. and Bocarov, E. F. (1967) *Genetika* (Leningrad), 5 : 85-94.
79. Keyl, H. G. (1965) *Experientia*, 21 : 191-193.
80. Keyl, H. G. (1966) Increase of DNA in chromosomes, *Chromosomes To-day*, 1 : 99-101, (Eds.) C. D. Darlington and K. R. Lewis, Oliver Boyd.
81. Kihlman, B. A. (1961) *Adv. Genet.*, 10 : 1-59.
82. Kihlman, B. A. (1966) *Actions of chemicals on dividing cells*, Prentice-Hall, Englewood Cliffs.
83. Kindsin, R. B. (1976) *J. Med. Genet.*, 8 : 181-187.
84. Kuroki, Y., Makino, S., Aya, T. and Nayagama, T. (1966) *Jap. J. Human Genet.*, 11 : 17-23.
85. Lastikka, L., Virsu, M. L., Halkka, O., Erriksson, K. and Estola, T. (1976) *Med. Biol.*, 54 : 146-149.
86. Lea, D. E. (1962) *Actions of radiations of living cells*, 2nd Ed., University Press, Cambridge.
87. Lehman, J. M. and Defendi, V. (1970) Early chromosome damage in Chinese hamster cells by SV 40 virus, *Mamm. Chrom. Newsletter*, 11 : 69.
88. Leuchtenberger, C. (1950) *Chromosoma*, 3 : 449-473.
89. Levan, A. and Tjio, J. H. (1948) *Hereditas*, 34 : 453-484.
90. Libikova, H. and Motajova, J. (1967) *Acta. Virol.*, 11 : 380.
91. McClintock, B. (1955) *Brookhaven Symposium Biol.*, 8 : 58-74.
92. Makino, S. (1956) *Int. Rev. Cytol.*, 6 : 25.
93. Makino, S., Yamada, K. and Kajii, T. (1965) *Chromosoma*, 16 : 372.
94. Makino, S. and Aya, T. (1968) *Cytologia*, 33 : 370.
95. Manna, G. K. (1656) *Proc. Intern. Genet. Symp., Cytologia Suppl.*, 182-187.
96. Manna, G. K. (1962) *Proc. Zool. Soc.*, Calcutta, 15 : 1-12.
97. Manna, G. K. (1967) *Proc. 3rd All India Cancer Congr., Indian J. Cancer*, 4 : 64-80.
98. Manna, G. K. (1969) Some aspects of chromosome cytology, *Presidential address, Sec. Zoology, Proc. 56th Ind. Sci. Congr.*, 2 : 185-214.

99. Manna, G. K. (1971) Bone-marrow chromosome aberrations in mice induced by physical, chemical and living mutagens, *Proc. 1st All India Congr. Cytol. Genet., J. Cytol. Genet., Congr. Vol.* : 144-150.
100. Manna, G. K. (1973) Chromosome aberrations induced by living mutagens, *Presidential Address, Biol. Sci., 43rd Ann. Sess. Nat. Acad. Sciences*, India, 1-15.
101. Manna, G. K. (1975) Induced breakage in chromosomes, mechanical and/or chemical, Symp. on *Structural and functional aspect of chromosomes*, BARC, Bombay, pp. 194-204.
102. Manna, G. K. (1978a) Nonrandom aberrations and their implication on the possible existence of weaker regions in chromosomes, *Proc. 3rd All India Congr. Cytol. Genet.*, Hissar (*In press*).
103. Manna, G. K. (1978b) Mutational hazards in changing environment, *Proc. 47th Ann. Sess. Nat. Acad. Sci.*, India, Bhopal (*In press*).
104. Manna, G. K. (1980) Antibiotics in different roles in the induced chromosome aberrations. *Intern. Seminar on Management of Environment*, BARC, Bombay, Feb. 11-15, (*In press*).
105. Manna, G. K. and Bhowmick, S. (1969) *Proc. 56th Indian Sci. Congr.*, Pt. 3 : 454.
106. Manna, G. K. and Das, R. K. (1972) *Nucleus*, 15 : 180-186.
107. Manna, G. K., Das, R. K. and Das, P. K. (1976) *Nucleus*, 19 : 40-46.
108. Manna, G. K., Das, R. K. and Khuda-Bukhsh, A. R. (1977) *Nucleus*, 20 : 252-257.
109. Manna, G. K. and Das, S. K. (1974) Chromosome aberrations in mice injected with the bacterium, *Salmonella typhi*, its culture and filtrate, *Nucleus*, 17 : 71-75.
110. Manna, G. K. and Das, S. K. (1975) *Proc. 62nd Ind. Sci. Congr.*, Pt. 3 : 104.
111. Manna, G. K. and Dey, M. (1979) *Proc. 47th Ann. Sess. Nat. Acad. Sci.*, Nagpur, 27-29 Nov., Abstr.
112. Manna, G. K. and Chakraborti, S. (1970) *Nucleus*, 12 : 167-176.
113. Manna, G. K. and Chakraborti, S. (1975) Proc. Sym. on *Structure and functional aspects of chromosomes*, BARC, Bombay, March 25 and 26, pp. 194-204.
114. Manna, G. K. and Mazumder, S. C. (1967) *Nucleus* 10 : 128-139.
115. Manna, G. K. and Gupta, S. (1975a) *Proc. 62nd Indian Sci. Congr.*, Pt. 3 : 44.
116. Manna, G. K. and Gupta, S. (1975b) *Proc. 44th Ann. Sess. Nat. Acad. Sci.*, Abs. p. 7.
117. Manna, G. K. and Gupta, S. (1976) *Proc. 63rd Ind. Sci. Congr.*, Pt. 3 : 150.
118. Manna, G. K. and Gupta, S. (1977) *CIS* (Japan), 23 : 30-31.
119. Manna, G. K. and Gupta, S. (1978a) *CIS* (Japan), 24 : 8-10.
120. Manna, G. K. and Gupta, S. (1978b) Meningococcus associated chromosome aberrations with their alternation by penicillin and mitotic inhibition in mice, *Nucleus*, (*In press*).

121. Mark, J. (1967) *Hereditas*, 57 : 23-82.

122. Marquez-Monter, H., Higuera-Ballesteros, F. J., Alfaro-Kofman, S., Monte-Mayor, E. G. and Gonizalez-Angulo, A. (1967) Chromosomes of human rabies, *Lancet*, II : 102.

123. Matsaniotis, N., Maounis, F., Kiossoglou, K. A. and Anagnostakis, D. E. (1968a) Chromosomes after small-pox vaccination, *Lancet*, I : 978.

124. Matsaniotis, N., Maounis, F., Kiossoglou, K. A. and Anagnostakis, D. E. (1968b) Chromosome in infectious hepatitis, *Delt. Pedia. Klin. Panep. Athin.*, 15 : 12-26.

125. Mauler, R. and Hennessen, W. (1965) *Arch. Ges. Virusforsch.*, 16 : 175-181.

126. Mazzone, H. M. and Yerganian, G. (1963) *Exp. Cell Res.*, 30 : 591-592.

127. Mella, B. and Lang, D. J. (1967) *Science*, 155 : 80-81.

128. Mellman, W. J., Plotkin, S. A., Moorhead, P. S. and Hartnett, E. M. (1965) *Amer. J. Dis. Child.*, 110 : 473-476.

129. Michajlova, G. R. and Gosunova, L. P. (1969a) *Vop. Virus*, 14 : 326-331.

130. Michajlova, G. R. and Gorsunova, L. P. (1969b) *Genetica*, 5 : 48-56.

131. Miles, C. P. and O'Neill, F. J. (1969) *Amer. Assoc. Cancer Res.*, 10 : 230.

132. Moorhead, P. S. and Saksela, E. (1965) *Hereditas*, 52 : 271-284.

133. Montezuma-De-Carvalho, J. (1955) *Biol. Soc. Broter.*, 29 : 145-183.

134. Muller, H. J. (1927) *Science*, 66 : 84-87.

135. Muller, H. J. (1956) *Brookhaven Symp. Biol.*, 8 : 126-147.

136. Nachtigal, M., Aderca, I., Matache, A. and IF Timovici, M. (1966) Chronic infections with RNA viruses in cell cultures, III Chromosomal changes in normal and chronically infected cell cultures, *Rev. Roum. Inframicrobiol.*, 3 : 43-52.

137. Nichols, W. W. (1966) *Hereditas*, 55 : 1-27.

138. Nichols, W. W. (1978) Genetic effects of virus transformations, *Proc. XIV Inter. Congr. Genet.*, Moscow, Plenary Sessions and Symposium Abstract, p. 83.

139. Nichols, W. W. and Heneen, W. K. (1965) *Hereditas*, 52 : 402-410.

140. Nichols, W. W., Levan, A. and Khilman, B. A. (1964) Chromosome breakages associated with viruses and DNA inhibitors, pp. 255-271, *In Cytogenetics of cells in culture*, (Ed.) R. J. C. Harris, Acad. Press, N.Y.

141. Norrby, E., Levan, A. and Nichols, W. W. (1965) *Exp. Cell Res.*, 41 : 483-491.

142. Nusbacher, J., Hirschhorn, K. and Cooper, L. Z. (1967) *New Engl. J. Med.*, 276 : 1409-1413.

143. Oehlkers, F. (1943) *Z. Induktive Abstammungs-u, Vererbungslehre*, 81 : 313-341.

144. O'Neill, F. J. and Rapp, F. (1971) *Mamm. Chrm. Newsletter*, 12 : 28.

145. Parnes, V. A., Rapoport, I. A. and Levina, D. M. (1968) *Dokl. Akad. Nauk, SSSR*, 182 : 953-956.

146. Paton, G. R., Jacobs, J. P. and Perkins (1965) *Nature*, 207 : 43-45.

147. Ponten, J. (1963) *J. Nat. Cancer Inst.*, 30 : 897-921.

148. Ponten, J. and Lither, F. (1966) *Int. J. Cancer*, I : 589-598.

149. Prunieras, M. and Jacquemont, C. (1965) *Ann. Inst. Pasteur*, 109 : 185-203.

150. Renda, Y. (1968) *Turk. J. Pediat.*. 10 : 13-22.

151. Rosen, G. Von (1964) *Hereditas*, 51 : 190-294.

151a. Reiss, J. (1971) *Experientia*, 27 : 971.

152. Sablina, O. V. and Bocarov, E. F. (1968) *Genetica*, 4 : 128-134.

153. Salkinder, M. G. and Gear, J. H. S. (1964) Virus and chromosomes interaction, Meeting of South African Society of Pathology.

154. Saksela, E. (1964) *Hereditas*, 52 : 250-251.

155. Sax, K. (1940) *Genetics*, 25 : 41-68.

156. Schuler, D., Hervei, S., Gacs, G. M., Kirchner, M. and Szathmary, J. (1968) *Humangenetik*, 6 : 55-60.

157. Singleton, W. R. (1965) *Elementary genetics*, Affiliated East-West Press, New Delhi.

158. Sharma, A. K. and Sharma, A. (1960) *Inter. Rev. Cytol.*, 10 : 101-136.

159. Sofuni, T., Makino, S. and Kobayashi, H. (1967) *Proc. Jap. Acad.*, 43 : 389-394.

160. Sparrow, A. (1951) *Ann. N.Y. Acad. Sci.*, 51 : 1508-1540.

161. Stadler, L. J. (1928) *Science*, 68 : 186-187.

162. Stenman, S. and Saksela, E. (1969) *Hereditas*, 62 : 323-338.

163. Stich, H. F., Hsu, T. C. and Rapp, F. (1964) *Virology*, 22 : 439-455.

164. Stich, H. F. and Yohn D. S. (1967) *Nature*, 216 : 1292-1294.

165. Stoltz, D. B., Stich, H. F. and Yohn, D. S. (1967) *Cancer Res.*, 27 : 587-598.

166. Szanto, J., Motajova, J. and Pristasova, S. (1967) *Acta. Virol.*, 11 : 363-371.

167. Tolstova, T. I., Kulikova, K. S., Bogdanova, N. S., Pershin, G. N. and Mironova, L. (1966) *Acta. Virol.*, 10 : 315-321.

168. Tsuchida, R. and Rich, M. A. (1964) *J. Nat. Cancer Inst.*, 33 : 33-49.

169. Vogel, F. and Rohrborn, G. (1970) *Chemical mutagenesis in mammals and man*, Springer-Verlag, New York.

170. Vogt, M. and Dulbecco, R. (1963) *Proc. Nat. Acad. Sci., USA.*, 49 : 171-179.

171. Watson, J. D. and Crick, F. H. C. (1953) The structure of DNA, *Cold Spring Harbour Symp. Quant. Biol.*, 18 : 123-131.

172. Weinstein, D. and Moorhead, P. S. (1965) *J. Cell Comp. Physiol.*, 65 : 85-92.

173. White, M. J. D. (1973) *Animal cytology and evolution*, 3rd Ed., Cambridge Univ. Press.

174. Wilson, G. B. (1960) *Intern. Rev. Cytol.*, 9 : 293-304.

174a. Wolff, S. (1963) *Radiation-induced chromosome aberration*, Columbia Univ. Press.

175. Wolman, S. R., Hirschhorn, K. and Todaro, G. J. (1964) *Cytogenetics*, 3 : 45-61.

176. Woodward, D. O. and Woodward, V. W. (1978) *Concepts of molecular genetics*, TMH Ed., Tata McGraw-Hill Publ. Ltd., New Delhi.

177. Yerganian, G. (1966) Relative growth and oncogenic potential of normal, malignant and virus-transformed euploid cells of dwarf species hamster, pp. 112-122, *In Recent results in cancer research*, 6, (Ed.) W. H. Kirsten, Springer-Verlag, N.Y.

178. Yerganian, G., Cho, S. S., Ho, T. and Nell, M. N. (1965) Euploidy and chromosome alteration in normal, malignant and tumour virus transformed cells of the Chinese and Armenian hamsters, pp. 349-360, *In Symposium mutational process*, Prague.

179. Yoshida, M. C., Makino, S. and Kinjo, T. (1967) *Proc. Jap. Acad.*, 43 : 662-667.

180. Zur Hausen, H. (1967) *J. Virol.*, 1 : 1174-1185.

181. Zur Hausen, H. and Lanz, E. (1966) *Z. Med. Mikrobiol. Immunol.*, 152 : 66-72.

Perspectives in the Development of Radiation Chemical Concepts and Mechanisms

K. N. Rao and P. N. Moorthy

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

I. Introduction

Radiation chemistry is the study of chemical changes brought about by the action of radiations on matter. In such a broad definition the word "radiation" is all inclusive : both electromagnetic (γ, X, UV and visible) and corpuscular (fast charged particles such as the α and β "rays", protons, deuterons and charged atomic species farther up in the periodic system ; neutrons ; mesons, etc.). However, photochemistry or the study of the chemical effects brought about by UV and visible radiation on matter considerably predates what we now call radiation chemistry. As the energy content of UV and visible photons is of the order of excitation energies of molecules and is hardly ever greater than their ionization potentials, the initial act of such photons is molecular excitation and photochemistry is therefore concerned mainly with the chemistry of these excited molecules. On the other hand, all the rest of the radiations mentioned above are endowed with an energy content (per photon or particle) much greater than the ionization potentials of atoms and molecules and therefore these can produce both excited and ionic species. The field that has now come to be known as radiation chemistry comprises the study of the chemical action of these high energy radiations and its mechanistic definition could be given as the chemistry of concurrently produced ionization and excitation. It must be noted, however, that concurrent production of ionized and excited species is not restricted to high energy radiations. Such a phenomenon also occurs during electrical discharges in gases at low pressures and the chemical changes occurring in such systems (known as plasma chemistry) are rather similar to radiation chemical changes in gases.

In a sense, radiation chemistry is a truly multidisciplinary science. In fact, Haissinsky[57], pondering over the question "which scientific disciplines are historically and genetically linked to radiation chemistry " comes to the conclusion that "in short, physical and nuclear chemistry, radiobiology, photochemistry, chemical kinetics and electrochemistry are radiation chemistry's main sources. When of age, radiation chemistry may, in turn, contribute to its sources and other disciplines".

From its humble but significant beginnings with the observation of evolution of gaseous hydrogen and oxygen from aqueous solutions of radium bromide by Runge and Bodlaender (quoted by Giesel[52]) at the beginning of the present century, radiation chemistry in general and that

of water and aqueous solutions in particular has made enormous progress. This is quite evident from the growth of publications in this field (Fig. 1). Besides the basic curiosity to understand minute details as to *how much* of *what species* are generated and reacting at *what periods of time* from the initial physical act of ionisation and excitation of molecules in the system, many other factors have contributed to the enormous growth of this field. Notable among these is the proliferation of nuclear technology in the post-Second World War years. The presence of high intensities of high energy radiations within a nuclear reactor necessitated a knowledge of the chemical (as well as physical) damage occurring in its functional components such as the coolants and moderator (generally water, graphite, CO_2) and structural components such as lubricants in machine parts. Because of the fact that in such nuclear installations as well as in areas such as therapeutic applications of radioactive substances human beings are exposed to varying levels of radiation doses, study of the biological effects of radiation both at the cellular and molecular level became important. The availability of large radiation sources as by-products of the nuclear industry attracted attention to possible industrial applications of the former in areas such as synthesis of new materials like polymers, pasteurization of foods, sterilization of medical products and so on.

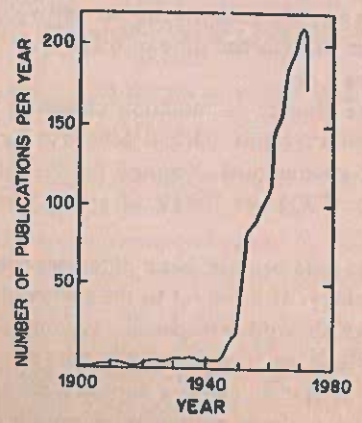


Fig. 1. Growth of publications in the field of radiation chemistry of water and aqueous solutions (after Draganic and Draganic[28])

It is the purpose of this article to review the various facets in the evolution of radiation chemistry with an emphasis on the elucidation of basic concepts and mechanisms. From the discussion presented it would be obvious that despite its resemblances to and overlaps with other fields, radiation chemistry has a unique identity of its own.

The types of systems whose radiolytic behaviour has been studied to-date is innumerable. These include gases, liquids, solids, solutions and even liquid crystals—to name only a few. But by far the one that has been most extensively studied and best understood is the aqueous system (water and aqueous solutions). This has been partly due to the fact that water is available in plenty in nature and can be easily purified, but the main reason for the large attention towards this system has been the importance of water as a coolant and moderator in nuclear reactors and as a major constituent of all living organisms. The discussion in the present article will, therefore, pertain largely to the aqueous system.

II. Definitions and Concepts

In order to facilitate understanding the material presented in this article the jargon peculiar to radiation chemistry is explained in this section.

Absorbed dose is the amount of radiation energy absorbed in the system and is expressed in units of electron volts per unit volume or mass (eV/dm^3 or eV/g). A quantum of absorbed dose of $6{\cdot}242\times10^{15}$ eV(= 100 ergs) per gram is referred to as a rad.

G-value : It is the practice in radiation chemistry to quantify changes (or yields) in terms of a *G*-value which is defined as the number of molecules of a compound *M* decomposed (denoted by $G(-M)$) or of a product *P* formed (denoted by $G(P)$) per 100 eV of energy absorbed in the *system as a whole.*

It is important to note here the basic difference between photochemistry and radiation chemistry with respect to the energy absorption mechanism. In the former, energy absorption is specifiic, i.e., only those incident photons will be absorbed whose energies fall within the spectral absorption profile of a component in the system, and if a photon is absorbed, all its energy is initially deposited on a single molecular or atomic species in the system. Therefore it is meaningful to speak of a quantum yield, i.e., the extent of chemical (or physical) change per quantum absorbed. On the other hand, in

radiation chemistry energy absorption is by the system as a whole via collisions with the extra nuclear electrons of all the component atoms of the system, and the energy of one photon or particle is initially deposited on the extra-nuclear electrons of more than one species in the system. It is, therefore, more meaningful to speak of the chemical change occurring per a fixed amount of energy absorbed and this fixed amount has been adopted as 100 eV. The concept of such a *G*-value was first introduced into radiation chemistry by Burton[38].

The experimental determination of the *G*-value requires the measurement of two quantities—the energy absorbed and the extent of chemical change that has occurred as a result. The former is known as dosimetry. Absolute dosimetry can be carried out by measuring the temperature rise in a material having good thermal conductivity and in which *all* the absorbed energy is converted to heat (e.g., graphite or metal), but is too cumbersome for routine use. Therefore many secondary dosimeters have been developed whose responses have been calibrated by the absolute method, and the more popular among these secondary dosimeters are chemical systems wherein the extent of chemical change observed can be taken as a direct measure of the absorbed dose. Depending on the magnitude of the dose one can choose a suitable chemical dosimetric system, but by far the most convenient and reliable one is that developed by Fricke (and hence designated after him) in the early days of radiation chemistry[49]. This consists of an air-saturated millimolar solution of ferrous (or ferrous ammonium) sulphate in 0·8 NH_2SO_4, containing millimolar NaCl (to counteract the effect of organic impurities[50]. The concentration of ferric formed is measured spectrophotometrically at 304 nm, knowing $\epsilon_{3\,O_4}^{Fe^{+3}}(t^\circ C)=2193+(t-25)\,0{\cdot}0007$ and the absorbed dose is calculated knowing $G(Fe^{+3})$ for different types of radiations (*see* Appendix I). For the quantitative determination of the chemical change in the system of interest, any appropriate method can be selected from among a host of physical, physicochemical, chemical or biological techniques.

Initial G-value : In practice the extent of measured change may or may not increase linearly with increasing dose. The latter situation arises either as a result of saturation of some pre-existing condition in the system (for example, depletion of some minor component) or competitive participation of a product in the overall reaction scheme. Under such a situation the quantity which is more meaningful is the initial *G*-value and this can often

be evaluated by a linear back extrapolation of a plot of measured G-values versus dose to zero "dose" (*see* Fig. 2).

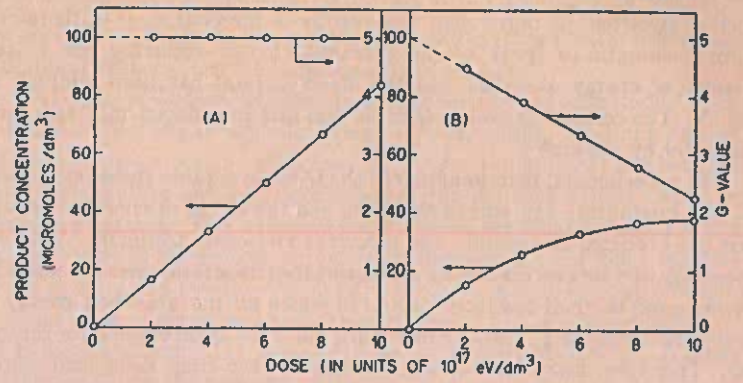


Fig. 2. Variation of product concentration with dose—A ; linear, constant G-value ; B : G-value decreases with increasing dose

Direct and Indirect Effects : Because of the fact that in radiation chemistry, energy absorption is via collisions with the extranuclear electrons of all the component atoms of the system, one can partition the energy absorbed in the system (E) in terms of the electron fractions contributed by the different component molecules (f_i)

$$E_i/E = f_i = (\bar{Z}/\bar{A})_i / \sum_i (\bar{Z}/\bar{A})_i \qquad (1)$$

where $\bar{Z}$ and $\bar{A}$ are the mean atomic numbers and mean atomic weights of the component molecules of the system[165]. Hence, in a multicomponent homogeneous system if the concentrations of all components (solutes) except one (the solvent) are much smaller than that of the solvent, then energy absorption is almost exclusively by the major component and any changes in the solutes are referred to as arising from an *indirect* effect. In the early days of radiation chemistry it was the study of these indirect effects in aqueous solutions under different conditions that contributed greatly to unravelling the basic mechanisms of water radiolysis. The changes in the major component which are obviously a consequence of radiation absorption in that component are referred to as *direct effects*. Thus radiolytic decomposition of water molecules in dilute aqueous solutions is a direct effect, whereas oxidation of ferrous ions to ferric in dilute acidic

solutions of ferrous salts is an indirect effect. Depending on the chemical composition of the system both direct and indirect effects may contribute to the formation of a particular product. Thus in irradiated water containing dissolved oxygen, hydrogen peroxide is formed by both direct effect (via radiolysis of H_2O molecules) and indirect effect (via chemical reactions in which the dissolved oxygen participates). As suggested by Daniels[85] one can partition the measured yield into the two contributions in terms of the electron fractions (f)

$$G(P)/f_m = G(P)_s f_s/f_m + G(P)_m \qquad (2)$$

where the subscripts s and M denote the solute and the medium (i.e., all components except the solute). A plot of $G(P)/f_m$ versus f_s/f_m would be linear with a slope equal to the direct effect contribution from the solute and (extrapolated) intercept equal to the indirect effect contribution from the medium. In practice, the plot may deviate appreciably from linearity at the low concentration end (i.e., low f_s/f_m) because of insufficient solute to "sense" all the changes which occurred in the medium.

Time scale : In developing a complete picture of the evolution of radiation-induced chemical changes, it is convenient to divide the entire time scale of events into three domains or stages : physical, physicochemical and chemical. It must be stressed at the very outset that the interfaces between these domains are not very sharp nor is the time span so rigid. As a rough working guide, the time scale of events in liquid water and aqueous systems at 300 K is given in Appendix II. As the names would imply, the physical and chemical stages are dominated respectively by well-known physical processes and chemical reactions which are sufficiently well understood. On the other hand, the processes occurring in the physicochemical stage which forms a link between the other two, are even today not completely clear and are the subject of much experimental and theoretical research.

III. Physical Stage

The physical stage of radiolysis comprises the physical interaction of the radiation with the medium through which it passes culminating in the formation of reactive zones which contain ions and excited molecules. In the case of high energy electromagnetic radiation (X- and γ-rays) the primary interactions with the medium are through photoelectric effect, Compton

scattering and pair production processes. These are shown schematically in Fig. 3. The relative importance of these processes depends both on the photon energy and the mean atomic number of the medium. In radiation chemistry γ-rays from Co^{60} having energies of 1·17 and 1·33 M eV are the

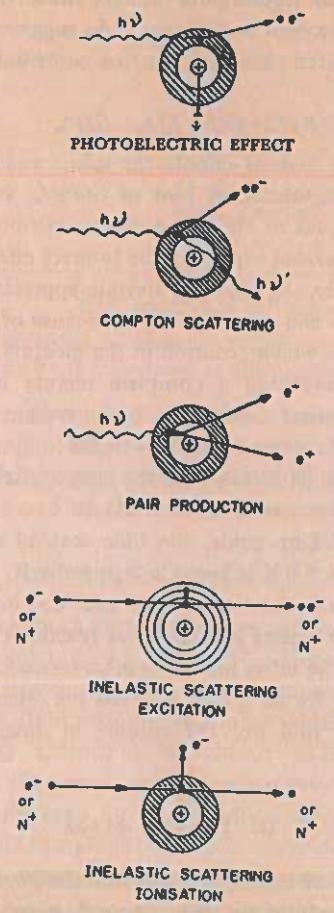


Fig. 3. Schematic of photoelectric, Compton and pair production processes for photons and inelastic scattering for fast-charged particles

most commonly used electromagnetic radiations, and for these the only important process in water (and also other low atomic number media such as organic compounds, biological systems, etc.) is Compton scattering (Fig. 4). In all the three processes, however, the net effect on the medium is ionization of the atom or molecule thereby generating a fast electron. The latter, during its travel in a medium of macroscopic dimensions, can bring about interactions in its path in a way similar to other charged particles such as α- and β-rays, protons, etc.

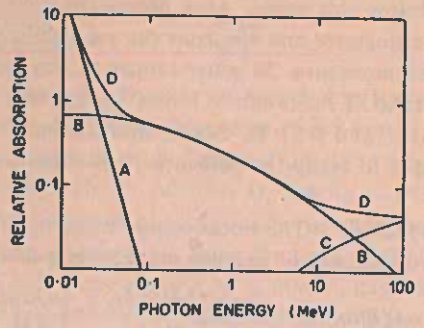


Fig. 4. Relative contribution of (A) photoelectric, (B) Compton and (C) pair production processes in water as a function of photon energy. (D) is the total absorption due to all the three processes (after Spinks and Woods[165])

The dominant mode of interaction of charged particles with matter leading to energy deposition in the latter is mainly inelastic scattering via Coulombic interaction with the extranuclear electrons (Fig. 3). Depending on the energy transferred to the target atom or molecule the result of an inelastic collision can be either excitation or ionization. Such events occur all along the path of the incident particle until its energy is insufficient to cause these effects. In the case of ionization the ejected electron (referred to as a secondary electron) can repeat the above process of ionisation and/or excitation until its energy falls below the lowest excitation potential of the molecules of the medium and such an electron is referred to as a subexcitation electron. As will be seen later the participation of subexcitation electrons in the physicochemical stage of radiolysis has often been invoked although there is no positive experimental proof of their definite role in radiation chemistry. Platzman[182] has also called attention to the fact that

polyatomic molecules can have superexcited states having energy in excess of the ionization potential. However, the net relative contribution of excitations (both normal and superexcitation) and ionizations in radiation chemistry cannot be computed with any degree of certainty. This is partly because the optical approximation[180, 182] which is the starting point of such computations is not sufficiently satisfactory at electron energies below ~100 eV whereas the secondary and succeeding generation of electrons formed in ionizing collisions have a wide spectrum of energies both above and below this value. Also, during the physicochemical stage excited states can autoionize and electrons can recombine with positive ions to generate excited molecules. In water subject to irradiation by electrons the value for the ratio of excitation to ionization has been computed to be 0·42 by Platzman[182] and 0·97 by Santar and Bednar[185]. There are no experimental findings to verify the authenticity of either of these.

Linear Energy Transfer : The linear energy transfer (LET), i.e., the rate of energy loss along the path of charged particles in a medium is given by

$$-(dE/dx)_{\text{coll}} = \frac{2\pi e^4 Z_p^2 N Z M_p}{E_p m_0} \ln\left(\frac{4E_p m_0}{I M_p}\right) \qquad (3)^*$$

where Z_p, M_p and E_p are respectively the charge, mass and kinetic energy of the impinging particle, m_0 the rest mass of the electron, $\bar{Z}$ the mean atomic number of the medium and I a constant for a given medium (~ 66 eV for water). From this expression it is obvious that for a given medium and a given particle the LET increases as the particles slow down and for different particles of the same energy the LET in a given medium is directly proportional to the mass of the particle and to the square of its charge. Thus fast electrons (and hence also X- and γ-rays whose effect is essentially those of fast electrons) have much lower LET than protons and other more massive particles of comparable energies (*see* Appendix I). As the LET along the track of a given particle increases as it slows down a more practically useful quantity is the mean or average LET defined as

$$-(dE/dx)_{\text{av}} = \frac{1}{R}\int_0^R -(dE/dx)_{\text{coll}}\, dx \qquad (4)$$

* This equation has to be modified somewhat at relativistic values of E_p.

where R is the range of the particle in the medium. A comparison of the values of initial and average LET for different radiations is given in Table 1.

Table 1

LETs of different high energy radiations

Radiation	*Initial LET* (eV/Å)	*Average LET* (eV/Å)
^{60}Co γ ; 1–2 MeV electrons	0·02	0·023
200–250 keV X-rays	0·10	0·17
^{210}Po α-rays (5·3 MeV)	8·8	4·3

Formation of Spurs and Tracks : As each energy loss event (via inelastic collision) is associated with the formation of an ion or excited molecule, the density of ions and excited molecules along the path or track of high energy radiation through the medium is directly related to the LET and hence the great importance of this quantity in radiation chemistry. When the energy loss event at any point along the track of the primary particle is an ionization the secondary electron so produced, depending on its kinetic energy can either generate its own independent track (and hence is called a δ-ray) or lose its entire energy in collisions over a path spanning a few molecules from its point of origin. Thus, from Fig. 5, it can be seen that

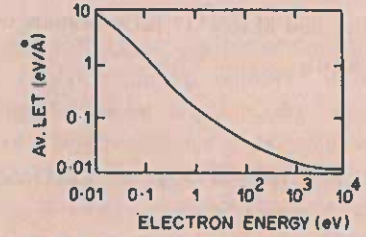


Fig. 5. Average LET for electrons in water (after Spinks and Woods[165])

when the secondary electron has an energy of 100 eV or less it can travel a distance not exceeding ~10 Å in water and within this range most of its energy is dissipated. If we assume an average of 20 eV per event (ionization or excitation) we see that such a secondary electron is capable of producing about 5 events in a region of radius not exceeding 10 Å. However, such groups of ionization and excitation events, called "spurs (or clusters)" are rather sparsely spaced along the path of the primary particle, e.g., at

distances of ~ 1000 Å in the case of ^{60}Co γ-rays, or 1–2 MeV electrons, progressively decreasing with increasing LET. The picture which would emerge from such considerations is that of a string of beads, the latter coming closer and closer together as we approach the end of the string. Further, for particles of higher LET the beads or the spurs are more closely spaced than for the electrons produced by the ^{60}Co γ-rays, and in fact overlap each other when the LET is more than about 10 eV/Å, producing a continuous column of ions and excited molecules with the primary particle trajectory as the axis. A rather crude and qualitative picture of how the track structure varies with LET is shown in Fig. 6. The actual geometric structure and composition of the track can be more complicated than pictured above. In addition to the spherical entities (isolated spurs, representing energy loss less than ~ 100 eV) and a cylindrical short track of closely spaced spurs (representing energy loss between 500–5000 eV over a continuous path). Mozumder and Magee[119] have introduced one more track entity, the blob (essentially a large spur representing energy loss of 100–500 eV) and computed relative contributions on these entities. The energy partition between these entities as a function of electron energy in electron tracks is shown in Fig. 7. As compared to the original estimate of Samuel and Magee[166] arriving at spurs of ~ 20 Å diameter spaced about 10^4 Å apart in 1 MeV electron and ^{60}Co γ-ray tracks in water the recent estimates of Mozumder and Magee[119] favours spurs of ~ 40 Å diameter spaced about 3000 Å apart.

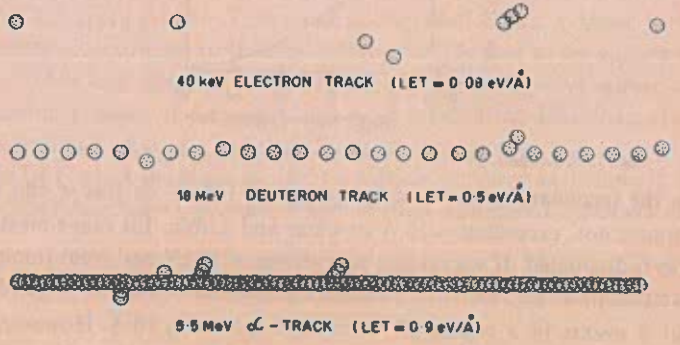


Fig. 6. Dependence of track structure on LET

A balance sheet of electronic energy distribution at the end of ~ 10^{-15} sec after the passage of high energy radiations in water has been computed by Mozumder and Magee (Table 2).

Table 2

Distribution of Electronic Energy in Water at the End of Electronic Phase

Type of electronic energy	*G-value*	*Average energy (eV) per molecule or particle*	*Contribution (eV)*
Ionization	3·33	12·6	41·96
Molecular excitation	0·54	9·0	4·86
Superexcitation	0 92	15·0	13·80
Excitation of positive ions	3·33	8·0	26·64
Kinetic energy of subexcitation electrons	3·33	3·83*	12·74*
		Total ···	100,00

This estimate is for the gas phase and as the authors have emphasised "energy balance is a difficult task at best and as the time after initial excitation increases, the uncertainties increase and eventually the analysis becomes meaningless".

Diffusion of Species : The track entities or the reactive zones in a system subjected to high energy radiation represent isolated regions of high concentrations of reactive species. Thus, a spherical spur of 40 Å diameter having an average of 5 reactive species corresponds to a concentration of ~1·5 mole dm^{-3} whereas if these species were uniformly distributed in the medium the average concentration would only be ~$4{\cdot}4 \times 10^{-7}$ mole dm^{-3} (assuming inter spur and inter track separation of 3000 Å). As a result of diffusion the species formed initially at high local concentrations within the track entities tend to attain homogeneous distribution with passage of time. Therefore the relative extent of various concentration dependent chemical processes among the species within the reactive zones or between these species and other molecules in the system is time dependent. This aspect of radiation chemistry whose importance is being increasingly recognised in

* By default.

recent years, is difficult both to tackle theoretically and assess experimentally and as will be seen later has been responsible for much of the confusion regarding the origins and yields of some of the species. Again the diffusion aspect introduces a phase dependence : whereas the tracks are nonexistent in the gas phase over any significant period after energy deposition, they are virtually frozen in a rigid solid wherein they can often be observed by more or less direct methods. In fact, the latter forms the basis of one of the methods for charged particle detection : the tracks formed in a plastic can be developed later by etching and directly observed (Fig. 7).

Fig. 7. Chemically etched particle tracks in a solid plastic material

Dose Rate Effect : The picture of non-overlapping tracks is valid only at low dose rates. With increasing dose rate the tracks will be spaced progressively closer until at very high does rates they can practically overlap each other producing ultimate effects not too different from those that arise from overlapping of spurs at high LET.

IV. Homogeneous Kinetic Chemical Stage

The discussion of the physicochemical stage will be deferred until the next section except to mention here that this stage terminates and gives way to the chemical stage once the active species formed in the reactive zones have reached a homogeneous distribution as a result of diffusion. Obviously chemical reactions would occur also during this intervening period and the yield (G-value) of the different active species that have survived reactions till the beginning of the chemical stage is referred to as the primary yield, symbolised either by a g or G followed by the formula of the species concerned as a subscript, e.g., g_{H_2} or G_{H_2}. It may be recalled that the corresponding experimentally observed net yield is witten as $G(H_2)$. With the recent developments in ultrafast kinetic techniques such as picosecond pulse radiolysis it has been possible to measure the time dependence of the yields of the active species during the diffusion and physicochemical stage and from this it is often possible with the help of proper diffusion kinetic models to extrapolate and obtain the initial spur yields of these species.

1. *Limiting Yields and Competition Kinetics*

The reactions which occur in the chemical stage are governed by the known laws of homogenous chemical kinetics. One can treat the primary species now present in the system simply as any other reactants and the time dependence of their concentration can be computed from the known primary yield, dose rate and chemical reaction rate constants. For example, if in a given system, H atom is one of the primary species and its *only* reaction is

$$H + RH \rightarrow H_2 + R \tag{5}$$

one can write

$$d[H]/dt = G_H I/100\, A - k_5[H][RH] \tag{6}$$

and

$$d[H_2]/dt = G_{H_2} I/100\, A\, k_5[H][RH] \tag{7}$$

where A is the Avogadro number, I the dose rate in eV dm^{-3} sec^{-1} and the square brackets denote the concentrations of the enclosed species in moles dm^{-3}. For a situation like this wherein a species is formed and consumed simultaneously and hence does not accumulate in the system the usual steady state principle of chemical kinetics can be applied so that $d[H]/dt = 0$ and hence from (6) and (7) we get

$$G(H_2) = G_{H_2} + G_H \tag{8}$$

If, in addition to (5) the species undergoes other reactions such as

$$H + X \rightarrow P(\text{not } H_2) \tag{9}$$

Equation (6) now takes the form :

$$d[H]/dt = (G_H I/100\,A) - \{k_5[RH] + k_9[X]\}[H] \tag{10}$$

From (7) and (10) and applying the steady state principle we get

$$G(H_2) = G_{H_2} + G_H \frac{k_5[RH]}{k_5[RH] + k_9[X]} \tag{11}$$

It can be seen that $G(H_2)$ increases with [RH] until at high values of the latter it attains the limiting value $G(H_2)_{lim} = G_{H_2} + G_H$ which is the same as given by equation (8) derived on the assumption that reaction (5) is the only pathway for reaction of the species.

Equation (11) can be rearranged to the form

$$\frac{1}{G(H_2)} = \left\{\frac{1}{G_{H_2}} + \frac{1}{G_H}\right\} + \frac{1}{G_H}\frac{k_9}{k_5}\frac{[X]}{[RH]} \tag{12}$$

from which it follows that a plot of $1/G(H_2)$ versus the concentration ratio [X]/[RH] (known as the competition kinetic plot) should be linear with slope equal to $(1/G_H)(k_9/k_5)$ and intercept equal to $1/G_{H_2} + 1/G_H$. In the above treatment, reactions between the primary species themselves such as

$$H + H \rightarrow H_2 \tag{13}$$

have been neglected. This is justified *in the homogeneous chemical* stage because, as seen before, the homogeneous or bulk concentration of these species at the beginning of this stage are very small ($\sim 10^{-7}$ mole dm^{-3}) as compared to the concentration of reactants such as RH.

Equations similar to (8) and (12) have formed the heart of all experimental radiation chemistry in the early phase of its development from the late fifties through the sixties. For, it can be seen that if we choose experimental conditions in such a way that k_9/k_5 is known, then from the competition plot one can evaluate both G_H and G_{H_2}. Alternatively, from the limiting yield of $G(H_2)$ one can calculate either G_{H_2} or G_H knowing the other from an independent reaction system, and knowing the primary yields one can evaluate the rate constant ratio (and hence the absolute rate constant for one of the reactions knowing that for the other). Thus, in water, the primary yields have been evaluated accurately from the measured

product yields in suitably chosen systems wherein equations of the type discussed above can be formulated on the basis of authentic reaction schemes involving the primary species and added solutes. Spinks and Woods[166] have given a compilation of suitable systems for this purpose.

Similarly over the past two decades, the rate constants of a large number of reactions of a wide variety of molecules with the primary species in water have been determined by the competition kinetic method and compilations of these values along with those determined by the use of the pulse radiolysis technique to be described later are available[10,12,41].

It is important to note here that in all these methods making use of steady-state kinetic conditions for the evaluation of primary *G*-values and rate constant ratios the experimentally measured product yields used in the calculations must be true initial yields.

2. *Primary Species*

Although the early observations of Runge and Bodlander[52], Cameron and Ramsay[24] and Debierne[39] had established beyond doubt that water is decomposed by radium emanations to give a mixture of H_2 and O_2, it was Kernbaum[63] who performed quantitative measurements of the gases evolved and showed that the quantity of O_2 produced is more than what is expected from the stoichiometry observed in electrolysis

$$2H_2O \rightarrow 2H_2 + O_2 \tag{14}$$

He established as well the presence of H_2O_2 as a product and showed that H_2 and H_2O_2 are produced in the stoichiometry expected for overall decomposition according to

$$2H_2O \rightarrow H_2 + H_2O_2 \tag{15}$$

To relate this observed chemistry to the then known ionizing behaviour of these emanations, Debierne[40] hypothesized that the ionized molecules in water give rise to H and OH radicals which then react either to reform water or the product molecules H_2 and H_2O_2

$$H + OH \rightarrow H_2O \tag{16}$$

$$H + H \rightarrow H_2 \tag{17}$$

$$OH + OH \rightarrow H_2O_2 \tag{18}$$

Oxygen is then only a minor chemical decomposition product of H_2O_2. Risse[149] also assumed that water is decomposed into H and OH radicals under the action of radiation.

This contains the essence of the modern theory of water radiolysis but it remained dormant until it was revitalised by the theories of Weiss[178] and Frank and Burton[5].

Activated water and free radical theories : In the intervening period, Fricke and Morse[49], working on the action of X-rays on water and aqueous solutions, observed that these radiations, unlike the radium emanations affect pure water little or not at all. But when solutes or impurities were present decomposition was observed and the solutes were oxidized or reduced, the extent of changes so induced in the solutes being independent of their concentrations over a wide range. This led Fricke to the activated water hypothesis : the X-rays somehow activate the water molecules and the activated water molecules revert back to water unless there are solutes or impurities to react with them.

Weiss[178], in addition to putting forward a theory to explain the primary formation of H and OH radicals, also invoked the Frank-Rabinowitch cage effect[48] to explain why these primary species readily undergo the reverse reaction (16) in pure water. However, he appears to have ignored Kernbaum's observation[88] of H_2O_2 as a major product in water subjected to radium emanations and proposed the product forming reactions to be (17) and (20)

$$2OH \rightarrow H_2O + O; \quad O + O \rightarrow O_2 \qquad (20)$$

The solutes present were considered to be acceptors for the H and OH radicals and hence able to compete with the reverse reaction.

The involvement of free radicals in the reactions brought about by γ-rays in aqueous solution was demonstrated more directly by Dainton[88]. It was known at that time that the polymerization of acrylonitrile in aqueous solution can be initiated by any source of free radicals to give an insoluble precipitate of polyacrylonitrile and this could be accomplished by the use of γ-rays also.

Both the free radical and activated water theories, however, failed to explain the difference between the observed effects of α-radiations and X-rays on pure water : a net decomposition being observed only in the former case. Besides, as shown by Allen[6] the free radical theory in conjunction with homogeneous kinetics predicts a dose rate effect on the variation of rate of product formation with solute concentration. Such an effect has not been experimentally observed.

Molecular products as primary species : The explanation for the different behaviour of α- and X-rays came from the work of Allen, Hochanadel and Davis[4] on water decomposition in the Oak Ridge nuclear reactor. These experiments permitted investigations under different LETs and revealed that in water containing dissolved H_2 and H_2O_2 as solutes these are rapidly consumed according to the overall reaction

$$H_2O_2 + H_2 \rightarrow 2H_2O \qquad (19)$$

Further, they showed that the extent of this reaction can be quantitatively accounted for on the basis of the chain reaction sequence (reactions 20 and 21) involving the H and OH radicals primarily produced in water

$$OH + H_2 \rightarrow H + H_2O \qquad (20)$$
$$H + H_2O_2 \rightarrow OH + H_2O \qquad (21)$$

As the overall reaction (19) was slower at higher peroxide concentrations they also postulated the following chain terminating steps involving H_2O_2.

$$OH + H_2O_2 \rightarrow H_2O + HO_2 \qquad (22)$$
$$HO_2 + HO_2 \rightarrow H_2O_2 + O_2 \qquad (23)$$
$$H + HO_2 \rightarrow H_2O_2 \qquad (24)$$

In order to maintain the chain terminating character, reaction (25) is not expected to occur.

$$HO_2 + H_2 \rightarrow H_2O_2 + H \qquad (25)$$

The authors concluded that the radiations decompose water to give *both* radicals (H and OH) and complete molecules (H_2 and H_2O_2) as the primary species. To explain the observed difference between the effect of X- and γ-rays on the one hand (low LET radiations) and α-particles on the other (high LET radiations) it was logical to suppose that the primary yields of radicals are higher than those of the molecular products in the case of low LET radiations and the opposite situation holds in the case of high LET radiations. As a result there are more than enough radicals to decompose all the molecular products in the former case whereas in the latter molecular products survive this decomposition. The results of various workers on the primary yields at different LETs as compiled together by Allen[7] are shown in Fig. 8. These, more than amply, demonstrate the truth of the above supposition. It may be noted that HO_2 is now recognised as one of the primary species. The origin of this species along with those of others will be discussed in Section V.

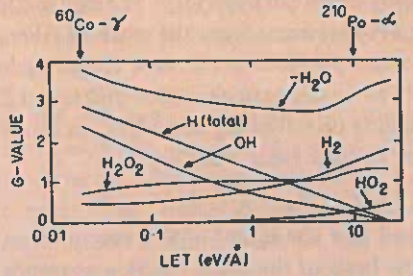


Fig. 8. Variation of the water decomposition and primary yields with LET at neutral pH (after Allen[7])

Two 'kinds' of hydrogen atoms : In addition to the primary $G_{H_2O_2}$, hydrogen peroxide can also be formed in water containing dissolved oxygen via reactions (26) and (23)

$$H + O_2 \rightarrow HO_2 \qquad (26)$$

However, Hochanadel[73] and Barr and Allen[16] observed that when water initially containing H_2, H_2O_2 and O_2 was irradiated the rapid chain decomposition of H_2O_2 according to reaction (19) occurred only after irradiation to doses corresponding to complete depletion of oxygen. This seemed to indicate that the primarily produced hydrogen atom reacted equally fast with both O_2 and H_2O_2 but the one produced via reaction (20), and designated as H', preferentially reacted with O_2 as compared to H_2O_2, so that as long as O_2 was present, *this hydrogen atom* did not destroy H_2O_2 but produced more of it via reactions (26) and (23).

Studying the pH effects on the $G(H_2)$ and G (acetone) in deaerated aqueous solutions of isopropanol subjected to γ-radiation, Allen and Scholes[8] also concluded that two kinds of H atoms are primarily formed, one of them, designated H_α, capable of dehydrogenating isopropanol to give H_2 and the other capable of reacting with acetone and probably also with H_2O_2. From their results they estimated the yields of the two kinds to be 0·6 and 2·3 respectively (at neutral pH) and also suggested that the former are produced by the dissociation of excited water molecules. Their results also suggested that the species formed with $G = 2·3$ are convertible to the other form at low pH.

Hydrated Electron : Although in the gas phase only one kind of H atom is known, it was surmised that the hydrogen atom in aqueous solutions can

exist in conjugate acid-base forms in equilibrium just as many other proton containing chemical species do. Two such conjugate pairs are possible for the hydrogen atom

$$H_2^+ \rightleftharpoons H + H^+ \qquad (27)$$

$$H \rightleftharpoons e^- + H^+ \qquad (28)$$

The two forms observed in irradiated water must be one of these. Allen[8] believed it to be the H, e^- pair, and more definite indication to this effect came from the experiments of Hayon and Weiss[65] and Hayon and Allen[66]. In deaerated solutions of chloroacetic acid (maintained at pH = 1·0) Hayon and Weiss[65] observed that with increasing concentration of the solute $G(H_2)$ decreased and $G(Cl^-)$ increased, the two variations following almost a mirror-image relationship (Fig. 9). This seemed to indicate that two processes, one leading to the formation of H_2 and the other to Cl^- were mutually competing, and the authors suggested the two competing pathways to be

(a) $$e^- + H_3O^+ \rightarrow H + H_2O \qquad (29)$$

followed by, $$H + ClCH_2CO_2H \rightarrow H_2 + ClCHCO_2H \qquad (30)$$

and (b) $$e^- + ClCH_2CO_2H \rightarrow Cl^- + CH_2CO_2H \qquad (31)$$

thus identifying the H' or H species to be a true hydrogen atom and the other kind of "hydrogen atom" to be an electron.

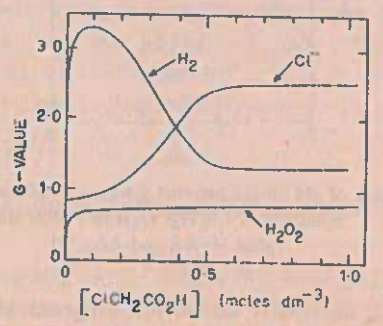


Fig. 9. Effect of chloroacetic acid concentration on product yields in irradiated aqueous solutions (pH = 1·0) (after Hayon and Weiss[65])

In the experiments carried out by Hayon and Allen[66], the chloroacetic acid concentration was kept constant at 0·1 mole dm^{-3} and the pH varied

from ~0 to 8. Again a mirror-image relationship was observed between $G(H_2)$ and $G(Cl^-)$ (Fig. 10). These results more convincingly demonstrated that the real competition is between H_3O^+ ions and $ClCH_2CO_2H$ for the electron, i.e., reactions (29) and (31). The radiation produced electron in water has been variously designated : hydrated electron (symbol e^-_{aq}), polaron (symbol, $(H_2O)^-$), but the former has gained almost universal acceptance. The fact that this species bears a unit negative charge was conclusively established[82] by a study of the effect of ionic strength on the rate constant ratios k_{33}/k_{32}, k_{34}/k_{32} and k_{35}/k_{32}

$$e^-_{aq} + H_2O_2 \rightarrow OH + OH^- \quad (32)$$

$$e^-_{aq} + O_2 \rightarrow O_2^- \quad (33)$$

$$e^-_{aq} + H_3O^+ \rightarrow H \quad (34)$$

$$e^-_{aq} + NO_2 \rightarrow NO_2^{2-} \quad (35)$$

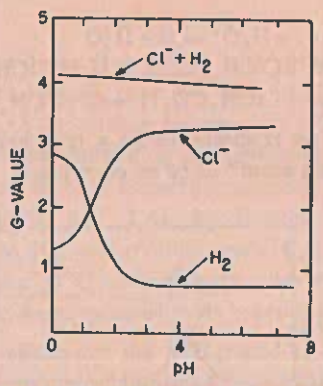


Fig. 10. Effect of pH on the product yields in irradiated aqueous solutions of $ClCH_2CO_2H$ (0·1 mole dm^{-3}) (after Hayon and Allen[66])

As shown in Fig. 11, the results satisfactorily agreed with the expectations based on the Bronsted model for kinetic salt effect on ionic reactions in solution. The absolute rate constant for reaction (36)

$$e^-_{aq} + Fe(CN)_6^{3-} \rightarrow Fe(CN)_6^{4-} \quad (36)$$

has also been found[55] to exhibit a kinetic salt effect expected for reactions between two like-charged ions.

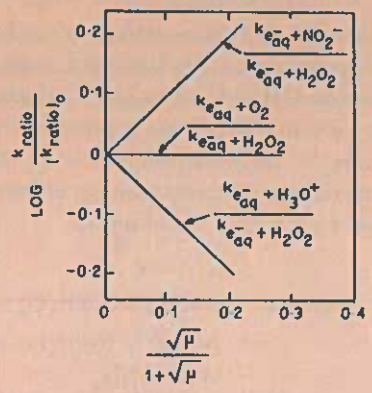


Fig. 11. Primary salt effect on hydrated electron reactions (after Czapski and Schwarz[82])

Since the above pioneering reaction systems which led to the discovery of e^-_{aq} many other solutes which react preferentially with either e^-_{aq} or H have been found and Table 3 lists such solutes along with the corresponding bimolecular rate constants. These rate constants have been evaluated both by competition kinetic methods as well as by the absolute pulse radiolysis method to be described later. Study of such preferential reactions in irradiated aqueous systems have also led to precise evaluation of G_H and $G_{e^-_{aq}}$ and these are listed in Appendix III along with the G-values for the other primary species in water.

The yields of the primary species are interrelated by the material balance equation

$$G(-H_2O) = 2G_{H_2} + G_{e^-_{aq}} + G_H = G_{OH} + 2G_{H_2O_2} + 3G_{HO_2} \quad (37)$$

Besides, charge balance would require that there should be a positively charged species, and this has now been recognised to be the H_3O^+ ion [15,25,157]. From the known value of $G_{e^-_{aq}} = 2{\cdot}63$ for ^{60}Co γ-rays in water at pH 3 – 13, we would expect $G_{H_3O^+}$ also to be 2·63. However, pulse radiolytic experiments to evaluate this quantity employing pH indicators

and conductivity techniques gave much higher values : 2·9[167], 3·25[26] and 3·38[15]. Any value in excess of 2·63 would again for charge balance require that there should be yet another entity bearing a negative charge and this has been recognised to be OH^- [157] with $G=0·7$. If we accept this number, a value of $G_{H_3O^+} \simeq 3·3$ is nearer to what is expected from charge balance. Recent redetermination[128] has confirmed this value. We can thus quantitatively represent the radiolytic decomposition of water (at the beginning of the chemical stage) by low LET radiations as

$$4·08H_2O \longrightarrow 0·45H_2 + 0·68H_2O_2 + 2·63e^-_{aq} + 0·70H^- + 3·33H_3O^+ + 0·55H + 2·720H + 0·026HO_2 \quad (38)$$

where the symbol $\longrightarrow$ originally introduced by Burton[33] is the accepted way for denoting radiation chemical changes and has a significance similar to $\xrightarrow{h\nu}$ in photochemistry. The origin or mechanism of formation of the various primary species enumerated above will be taken up in the discussion of the diffusion kinetic and physicochemical stage.

Conjugate Acid-base Forms of Primary Species : Among the various primary species summarized in equation (38), the ones that are commonly involved in reactions occurring in the homogeneous kinetic chemical stage are : e^-_{aq}, H, OH, HO_2 and H_2O_2, and are generally referred to as reactive primary species. As it is the common practice to employ well-buffered solutions in aqueous radiation chemistry the primarily produced H_3O^+ and OH^-, whose bulk concentration in experiments performed at low to moderate dose rates is sub-micromolar, do not enter into any reactions other than with the buffer. These two species are therefore generally ignored in steady-state aqueous radiation chemistry. The conjugate base forms of the other reactive species are *not primarily produced by radiation* but may be subsequently formed and react in these secondary forms depending on the pH of the medium. The acid-base dissociation constants (pK_a) of the primary species which would govern the relative contribution of the two forms at any given pH are given in Table 4.

Table 3

Comparison of bimolecular rate constants for e^-_{aq} and H atom reactions with various solutes (numbers in parentheses are pH values)*

Solute(s)	$k_{e^-_{aq}+S}$ ($dm^3\ mole^{-1}\ s^{-1}$)	k_{H+S}
H_3O^+	$2·36\times10^{10}$ (4–5)	$2·6\times10^8$
H_2O_2	$1·23\times10^{10}$ (7)	9×10^7 (2)
O_2	$1·9\times10^{10}$ (7)	$1·9\times10^{10}$ (2)
CO_2	$7·7\times10^9$ (7)	$<10^6$ (2)
N_2O	$9·4\times10^9$ (7)	$1·25\times10^4$
NO_3^-	$1·1\times10^{10}$ (7)	9×10^6 (7)
Cu^{2+}	$3·4\times10^{10}$ (1·5–4·5)	$4·2\times10^7$ (1·5)
Fe^{2+}	$3·5\times10^8$	$1·6\times10^7$ (0·4)
Fe^{3+}	$\sim5\times10^{10}$	5×10^7
Cd^{2+}	$5·3\times10^{10}$	$<10^5$
Acetone	$5·9\times10^9$ (7)	$3·5\times10^5$ (1)
Allyl alcohol	$<10^6$	$2·3\times10^9$
Chloroacetate ion	$1·2\times10^9$ (7–11)	$2·6\times10^5$ (7–11)
Chloroacetic acid	$6·9\times10^9$ (1–1·5)	8×10^4 (1–1·5)
Ethanol	$\leq 4\times10^8$	$1·3\times10^7$ (2)
HCO_2^-	$2·4\times10^4$ (11)	$2·6\times10^8$ (1)
Methanol	$\leq 4\times10^8$	$1·56\times10^6$ (2)
2-Propanol	very low	5×10^7 (2)
Phenol	$1·8\times10^7$ (6·3–6·8)	$4·2\times10^9$ (7)
Pyridine	3×10^9 (5·5–7·3)	$6·4\times10^7$ (7)
Tetranitromethane	5×10^{10} (5·5–7)	$2·6\times10^9$ (7)

*Taken from Ref. 4 and 8.

Table 4

pK_a values of the primary species

Species and Equilibria	pK_a
$H \rightleftharpoons e^-_{aq} + H^+$	9·7
$OH \rightleftharpoons O^- + H^+$	11·9
$HO_2 \rightleftharpoons O_2^- + H^+$	4·8

The evidence for an acid form of H and the equilibrium : $H_2^+ \rightleftharpoons H + H^+$ is rather meagre. The H_2^+ species has often been invoked[31,148,150] to explain the experimentally observed oxidising property of H atoms at very low pH both in radiation chemistry and in the interaction of H atoms produced by a microwave discharge :

$$H_2^+ + Fe^{+2} \rightarrow Fe^{+3} + H_2 \quad (39)$$

$$H_2^+ + I^- \rightarrow \tfrac{1}{2}I_2 + H_2 \quad (40)$$

However, other schemes not involving the H_2^+ species have also been suggested[70,71] to explain these reactions and it is debatable whether H_2^+ at all exists.

The discovery of e_{aq}^-, the conjugate base form of hydrogen atoms as the major reducing primary species in water led to the question as to the form of the species hitherto considered as OH. Hummel and Allen[74] determined the rate constant ratio for the reaction of this species with Br^- and ethanol as competing solutes in oxygenated neutral solutions and found this to be independent of the ionic strength, thus establishing that the species is uncharged and must be really OH. However, in alkaline solutions the rate constant for the reaction of this species with a number of solutes such as $Fe(CN)_6^{4-}$, CNS^- and CO_3^{2-} have been observed to decrease very markedly indicating the reacting species to be in a different form (O^-) and the pK_a for the $OH \rightleftharpoons O^- + H^+$ equilibrium has been evaluated to be 11·9[1,133,172]. In Table 5 are listed the rate constants of some of the reactions of OH and O^- to illustrate the marked differences in their reactivities towards certain solutes.

Table 5

Comparison of rate constants of reaction of OH and O^-
(numbers in parenthesis refer to pH)

Solute(s)	k_{OH+S} ($dm^3 mole^{-1} s^{-1}$)	k_{O^-+S}
CNS^-	$2{\cdot}8 \times 10^{10}$ (2, neutral)	1×10^9 (13)
CO_3^{2-}	$4{\cdot}2 \times 10^8$ (10·8)	$< 3 \times 10^6$ (13·5)
$Fe(CN)_6^{4-}$	$1{\cdot}1 \times 10^{10}$ (3 – 10)	$< 7 \times 10^7$ (13·5)
Benzoate	6×10^9 (6 – 9·4)	$< 6 \times 10^6$ (13)
Ethanol	$1{\cdot}6 \times 10^9$ (neutral)	$8{\cdot}4 \times 10^6$ (13)

Although the participation of H_2O^+ the conjugate acid form of OH as the reactive species has at times been suggested to explain certain radiation induced hydroxylation reactions[174] and pH dependence of the rate constant for reaction of OH with halide ions[156] the assumptions behind these explanations are not very plausible. In the absence of more convincing data, H_2O^+ is to be regarded as having no role to play *in the chemical stage* of radiolysis in fluid aqueous media.

O_2^-, the conjugate base form of the hydroperoxyl radical HO_2 is also not a primary species, but as the pK_a for the equilibrium $HO_2 \rightleftharpoons O_2^- + H^+$ is 4·8[186] it is the reacting form at neutral pH produced either via the above equilibrium or directly by reaction of e_{aq}^- with O_2 in oxygenated solutions.

$$e_{aq}^- + O_2 \rightarrow O_2^- \quad (33)$$

As in the case of H and OH, there are no convincing evidences for a conjugate acid form $H_2O_2^+$ of the hydroperoxyl radical.

3. *Net Chemical Changes*

Aqueous systems : Ever since the early days of radiation chemistry, study of radiation chemical changes in aqueous solutions has been one of the important areas of research. Although, often quantitative measures of these changes coupled with plausible reaction schemes have led to the evaluation of the yields of the primary species the reverse procedure is also commonly adopted : from a knowledge of the primary yields in water and the measured product yields in the system a plausible reaction scheme is developed to explain the observed chemical changes. Besides, the selectivity in the reactions of primary species towards certain solutes (e.g., e_{aq}^- with N_2O and SF_6, H with formate and allyl alcohol, OH with formate and isopropanol, e_{aq}^-, and H with O_2, etc.) are made use of to identify the different reaction pathways in the unknown system and attribute them to reaction of the solute with the appropriate primary species. In such studies it is also possible to evaluate the rate constant ratios for reaction of the appropriate primary species with the investigated solute and one of the above reference solutes. In what follows, the work carried out at BARC will be described in some detail to illustrate these aspects of radiation chemical research.

(i) *Vitamins* : In their study of the radiolysis of B-group vitamins in neutral aqueous solution Kishore, Moorthy and Rao[86] observed that

whereas in an aerated 10^{-4} mole dm^{-3} solution these compounds undergo very rapid decomposition such was not the case when the solutions were saturated with oxygen or N_2O and contained 0·1 mole dm^{-3} glucose (Gl). As a typical example the case of riboflavin (RF) is illustrated in Fig. 12. The expected reactions in this system are :

$$RF + e^-_{aq} \rightarrow \text{products} \quad (41)$$
$$RF + H \rightarrow \text{products} \quad (42)$$
$$RF + OH \rightarrow \text{products} \quad (43)$$
$$N_2O + e^-_{aq} + H_2O \rightarrow OH + OH^- + N_2 \quad (44)$$
$$O_2 + e^-_{aq} \rightarrow O_2^- \quad (33)$$
$$O_2 + H \rightarrow HO_2 \quad (26)$$
$$Gl + H \rightarrow R \rightarrow \text{products} \quad (45)$$
$$Gl + OH \rightarrow R' \rightarrow \text{products} \quad (46)$$

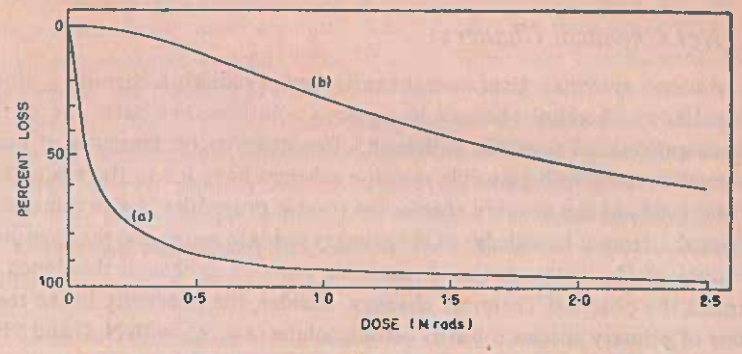


Fig. 12. Radiolytic decomposition of 10^{-4} mole dm^{-3} riboflavin solution (pH 6·8) : (a) in aerated solution (b) oxygen-saturated solution containing 0·1 mole dm^{-3} glucose (after Kishore, Moorthy and Rao[88])

Glucose is unreactive towards hydrated electrons[12]. At low doses when the conditions are such that $k_{44}[N_2O] \gg k_{41}[RF]$, $k_{33}[O_2] \gg k_{42}[RF]$, $k_{45}[Gl] \gg k_{42}[RF]$ and $k_{46}[Gl] \gg k_{43}[RF]$, the primary species react exclusively with N_2O, O_2 and glucose. If the product radicals from reactions 33, 26, 45 and 46 are unreactive towards the vitamin as would appear to be the case from a consideration of redox potentials[88] and

assuming that reaction by electron transfer is the predominant mode of interaction of these radicals with the vitamins, virtually no decomposition of the vitamin is expected. With increasing dose, however, decomposition is observed although at a much slower rate than in the absence of the additives which indicates progressive depletion of N_2O and O_2 and hence the above inequalities become progressively unfulfilled. Also the glucose radicals and O_2^- may react to some extent with the vitamin by pathways other than electron transfer. The observed results are in qualitative accord with this picture. Effect of this type wherein the presence of one or more additives reduces the rate of radiolytic decomposition of a given solute is known as protection and there are abundant examples of such a behaviour in the literature. The opposite effect wherein the radiolytic decomposition of a solute is enhanced in presence of an additive is known as sensitization. Thus, for example the radiolytic degradation of glucose in neutral aqueous solution would almost be doubled in presence of N_2O as compared to that in a deaerated solution because N_2O converts e^-_{aq} which is unreactive with glucose to OH radicals which are highly reactive.

In oxygen-saturated 10^{-4} mole dm^{-3} solution of thiamine (Thm^+) Kishore, Moorthy and Rao[89] observed $G(-Thm^+) = 1{\cdot}2 \simeq \frac{1}{2}G_{OH}$ and $G(H_2O_2) = 2{\cdot}6 \simeq G_{H_2O_2} + \frac{1}{2}(G_{e^-_{aq}} + G_H)$. These results are in agreement with the scheme

$$O_2 + e^-_{aq} \rightarrow O_2^- \quad (33)$$
$$O_2 + H \rightarrow HO_2 \rightleftharpoons O_2^- + H^+ \quad (26)$$
$$H^+ + HO_2 + O_2^- \rightarrow H_2O_2 + O_2 \quad (47)$$
$$Thm^+ + OH \rightarrow Thm^+(OH) \xrightarrow{O_2} \tfrac{1}{2}\,Thm^+ + \tfrac{1}{2}P^+ + O_2 \quad (48)$$

Additionally, when 10^{-3} mole dm^{-3} formate was also present, $G(-Thm^+)$ was found to be zero, and $G(H_2O_2) = 3{\cdot}8\ G_{H_2O_2} + \frac{1}{2}(G_{e^-_{aq}} + G_H + G_{OH})$. This conforms to the scheme : reactions (33), (26), (47) and

$$HCO_2^- + OH \rightarrow H_2O + CO_2^- \quad (49)$$
$$CO_2^- + Thm^+ \rightarrow Thm + CO_2 \quad (50)$$
$$Thm + O_2 \rightarrow O_2^- + Thm^+ \quad (51)$$

(ii) *Amino acids* : Mukherji and Mittal[120] have carried out a detailed study of the products formed in irradiated aqueous solutions of phenyl

alanine and inferred the site of attack of the different primary species on this polyfunctional molecule as shown in Fig. 13.

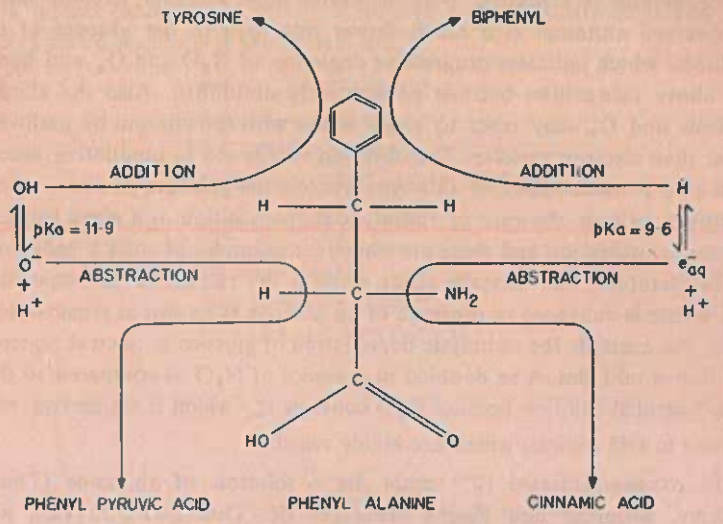


Fig. 13. Site of attack of primary species on phenylalanine (after Mukherjee and Mittal[120])

Lal[94,95] has carried out estimation of the yields of all major products in irradiated aqueous solutions of cysteine (CySH) and reduced glutathione (GuSH), and obtained good material balance between $G(-\text{solute})$ and the weighted sum of the yields of the amino fragments.

$G(-\text{CySH})$ was shown to be related to the yields of primary species and products by equation (52).

$$\begin{aligned} G(-\text{CySH}) = {} & G(\text{CyS}-\text{SCy}) + G(\text{H}_2\text{O}_2) - G_{\text{H}_2\text{O}_2} \\ & + \tfrac{1}{2}G_{\text{OH}} + 2G(\text{CyH}) + 3G(\text{CyOH}) \\ & + G(\text{CySO}_2\text{H}) - \tfrac{1}{2}G_{\text{H}} - \tfrac{1}{2}G_{e^-_{aq}} \\ & + 2G(\text{CySCy}) + G_{\text{H}}/(1 + k_{58}/k_{52}) \end{aligned} \qquad (52)$$

In aqueous solution of a more complex species, viz., the enzyme papain[96], the $-\text{SH}$ group on the 25th amino acid fragment cystein was identified to be the site of attack by OH radicals and H_2O_2.

(iii) *Alkali halides* : Shankar, Rama Rao and Shastri[161] have studied the radiolysis of aerated aqueous solutions containing methyl iodide and I^-. The products whose yields have been measured are I_2, H_2O_2 and CH_3O_2H. They have suggested the following scheme of reactions :

$$CH_3I + e^-_{aq} \rightarrow CH_3 + I^- \qquad (53)$$
$$CH_3I + H \rightarrow CH_3 + HI \qquad (54)$$
$$O_2 + e^-_{aq} \rightarrow O_2^- \qquad (33)$$
$$O_2 + H \rightarrow HO_2 \rightleftharpoons O_2^- + H^+ \qquad (26)$$
$$O_2 + CH_3 \rightarrow CH_3O_2 \qquad (55)$$
$$I^- + OH \rightarrow OH^- + I \qquad (56)$$
$$I^- + I \rightarrow I_2^- \qquad (57)$$
$$I_2^- + I_2^- \rightarrow I_3^- + I^- \qquad (58)$$
$$I_2^- + CH_3O_2 + H^+ \rightarrow CH_3O_2H + I_2 \qquad (59)$$
$$O_2^- + CH_3O_2 + H^+ \rightarrow CH_3O_2H + O_2 \qquad (60)$$
$$O_2^- + I_2^- \rightarrow 2I^- + O_2 \qquad (61)$$
$$H_2O_2 + I_2^- \rightarrow HO_2 + 2I^- + H^+ \qquad (62)$$

Based on this scheme, the equation relating the observed product yields to the primary yields has been derived as

$$\begin{aligned} & G_{OH} + G(CH_3O_2H) + 2G(H_2O_2) - 2G_{H_2O_2} - 2G(I_2) = A \\ & = \frac{G_{e^-_{aq}}}{1 + \dfrac{k_{58}[CH_3I]}{k_{38}[O_2]}} + \frac{G_H}{1 + \dfrac{k_{54}[CH_3I]}{k_{26}[O_2]}} \end{aligned} \qquad (63)$$

A plot of $1/A$ versus $[CH_3I]/[O_2]$ was found to be linear*. From the intercept of this plot $G_{e^-_{aq}} + G_H$ was evaluated as 3·2. This together with the slope gave a value of 0·4 for the above rate constant ratios.

(iv) *Uranium (IV)* : Bhattacharya and Saini[18] have studied the radiolytic oxidation of U(IV) in strongly acid media. In deaerated 0·8 mole dm^{-3} HCl solutions they observed $G(\text{U(VI)}) = 2{\cdot}1$ which is close to what one would expect on the basis of oxidation of U(IV) by Cl_2^- radicals and H_2O_2 according to

$$OH + 2Cl^- \rightarrow Cl_2^- \qquad (64)$$
$$U^{4+} + Cl_2^- + 2H_2O \rightarrow UO_2^+ + 4H^+ + 2Cl^- \qquad (65)$$

* which indicated that $k_{58}/k_{38} \approx k_{54}/k_{26}$

$$2UO_2^+ + 4H^+ \rightarrow UO_2^{2+} + U^{4+} + 2H_2O \quad (66)$$

$$U^{4+} + H_2O_2 \rightarrow UO_2^{2+} + 2H^+ \quad (67)$$

The derived stoichiometry is

$$G(U(VI))_{vac} = G_{H_2O_2} + 0{\cdot}5G_{OH} = 2{\cdot}1 \quad (68)$$

On the other hand, in aerated solutions the HO_2 radicals formed via reaction (26) also contribute to U(IV) oxidation according to

$$U^{4+} + HO_2 + H_2O \rightarrow UO_2^{2+} + 2H^+ + OH \quad (69)$$

followed by reactions (64)–(66). Under this condition U(VI) yield is given by

$$G(U(VI))_{air} = 1{\cdot}5G_H + G_{H_2O_2} + 0{\cdot}5G_{OH} = 7{\cdot}8 \quad (70)$$

which is in agreement with the experimentally observed limiting value at high U(IV) concentration in 0·8 mole dm^{-3} HCl medium.

(v) *Oscillatory reactions* : Rama Rao and Prasad[138] have observed a radiation-induced dormancy in the oscillatory behaviour of the Belousov-Zhabotenskii system (*see* Fig. 14). The reactions leading to oscillations of the Co (IV)/Ce (III) ratio (monitored by the potential of a Pt electrode) in this system are

$$BrO_3^- + Br^- + 2H^+ \rightarrow HBrO_2 + HOBr \quad (71)$$

$$HBrO_2 + Br^- + H^+ \rightarrow 2HOBr \quad (72)$$

$$BrO_3^- + HBrO_2 + H^+ \rightarrow 2BrO_2 + H_2O \quad (73)$$

$$BrO_2^+ + Ce(III) + H^+ \rightarrow HBrO_2 + Ce(IV) \quad (74)$$

$$2HBrO_2 \rightarrow BrO_3^- + HOBr + H^+ \quad (75)$$

$$HOBr + H_2C(CO_2H)_2 \rightarrow BrCH(CO_2H)_2 + H_2O \quad (76)$$

$$4Ce(IV) + 2BrCH(CO_2H)_2 + 3H_2O \rightarrow 6H^+ + 4Ce(III) + 2Br^- + OHCH(CO_2H)_2 + 2CO_2 + HCO_2H \quad (77)$$

The control intermediate is Br^- and radiolysis amounts to a constant input of this ion by reactions of primary H atoms with bromomalonic acid

$$H + BrCH(CO_2H)_2 \rightarrow H^+ + Br^- + CH(CO_2H)_2 \quad (78)$$

This would lead to complete conversion of Ce(III) to Ce(IV) via reactions (71)–(74). Once radiolysis is terminated the constant input of Br^- ceases and the system returns to its oscillatory state.

Organic Compounds : In sharp contrast to water, the radiation chemistry of organic compounds is much more complex because of the presence of

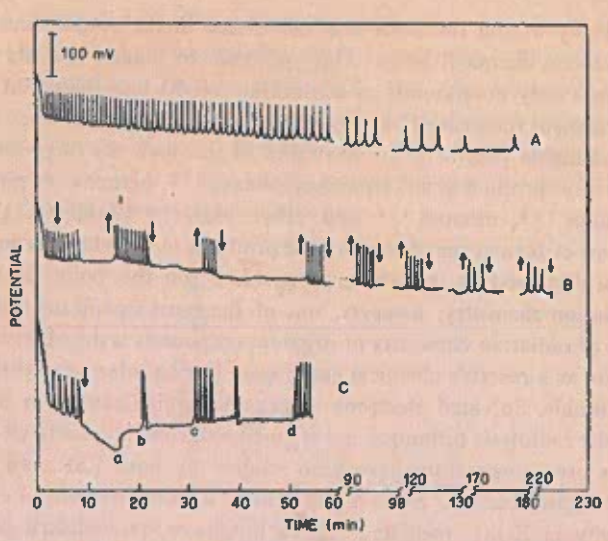


Fig. 14. Radiation effect on the oscillatory behaviour of the B-Z system.
A. Initial composition : $BrO_3^- = 0{\cdot}048$ mole dm^{-3}, malonic acid 0·275 mole dm^{-3}, Ce(IV) = 0·002 mole dm^{-3}, $H_2SO_4 = 1{\cdot}0$ mole dm^{-3}. Unirradiated.
B. Composition : As in A. Irradiated at $8{\cdot}4 \times 10^{18}$ eV dm^{-3} s^{-1}. Arrows indicate the onset (↓) and cessation (↑) of irradiation.
C. Excitability of a dormant oscillator. Initial composition : $BrO_3^- = 0{\cdot}018$ mole dm^{-3}, malonic acid = 0·275 mole dm^{-3}, Ce(IV) = 0·002 mole dm^{-3}, $H_2SO_4 = 1$ mole dm^{-3}. Continuously irradiated. Additional bromate added *in situ* at points a = 2 dm^3 of 0·18 mole dm^{-3} to 30 dm^3 mixture, b, c, d = 1 dm^3 each time (after Rao and Prasad[69])

diverse types of atoms and functional groups. Nevertheless, in a general way and in analogy with H_2O one can represent the situation existing at the beginning of the homogeneous kinetic chemical stage in terms of equation (79)

$$MHX \longrightarrow \begin{cases} H, X, MH, HX, H_2, m, MHX^-, MHX^+ \\ e^-_{solv}, MH^+, X^- \end{cases} \quad (79)$$

In the representation of the parent molecule as MHX it is implied that it contains a moity M linked to one or more H atoms and one or more singly linked other atoms or functional groups. Among the primary products, m denotes a small molecular fragment. Even without going into detailed mechanisms for the formation of these primary species one can appreciate

the complexity of the reactions that can ensue in the subsequent homogeneous kinetic chemical stage. The spectrum of final products would comprise not only compounds of molecular weight less than that of the parent (fragment molecules) but also higher molecular weight ones such as dimers and higher polymers. As examples of this diversity of products are cited below the product yields from cyclohexane [167], benzene [69], piperidine and pyridine [122], ethanol [121] and ethyl iodide[58] (Table 6). Detailed mechanisms of formation of some of the products in the above compounds have been discussed in the original papers. From the point of view of basic radiation chemistry, however, one of the most significant findings in the realm of radiation chemistry of organic compounds is the observation of the electron as a reactive chemical entity even in nonpolar or slightly polar organic liquids. Solvated electrons in organic liquids have been observed by the pulse radiolysis technique and trapped electrons in irradiated organic glasses at low temperature have been studied by both ESR and optical absorption spectroscopy. More significantly, by measurement of electrical conductivity in X-ray irradiated organic liquids of low dielectric constant, such as hydrocarbons, ethers, etc., at very low electric fields, Allen and Hammel[9] have demonstrated that electrons and positive ions escape recombination in significant yields. These free ion yields can also be determined from chemical scavenging experiments. Ward and Hamill[170] and Mittal[108] have also demonstrated the formation of carbonium ions (R^+) in irradiated organic liquids by observing the products formed from subsequent chemical reactions of these species.

Energy transfer : As can be seen from the data given in Table 6 it is generally true that aromatic compounds are more stable than corresponding aliphatic or alicyclic compounds. This stability is not only evident in the low product yields in these aromatic compounds but their presence has been found to confer a stabilising influence on the otherwise more labile compounds. Thus, as shown in Fig. 15, $G(H_2)$ in mixtures of pyridine or benzene with cyclohexane or piperidine[122] is much lower than what is expected on the basis of the mixture law equation (2). Both energy transfer and charge transfer mechanisms have been proposed to explain this effect, e.g., in the case of benzene cyclohexane mixtures

$$C_6H_{12}^* + C_6H_6 \rightarrow C_6H_{12} + C_6H_6^* \quad (80)$$

$$C_6H_{12}^+ + C_6H_6 \rightarrow C_6H_{12} + C_6H_6^+ \quad (81)$$

It being assumed that the excited benzene molecule or benzene cation is more stable towards product formation as compared to the corresponding cyclohexane species. This is reasonable from the observed inherently greater radiations stability of benzene as compared to cyclohexane. However, there are no clear-cut experimental evidences to decide in favour of either one or the other of the above mechanisms for the protective effect and it is possible that both are operative.

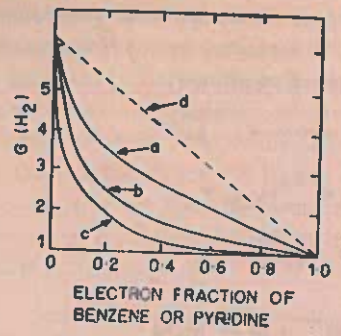


Fig. 15. Protective effect of aromatics. (a) Benzenepiperidine, (b) pyridine-piperidine, (c) pyridine-cyclohexane, (d) approximate mixture law behaviour (after Rao and Ramanan[122])

Table 6

Product yields in the *H*-radiolysis of some organic compounds in the liquid state

Compound	*Products and G-values*
Cyclohexane	H_2 6·6 ; CH_4 0·22 ; C_2H_2 0·21 ; C_2H_4 0·98 ; C_2H_6 0·18 ; C_3H_6 0·14 ; C_3H_8 0·24 ; Butenes 0·05 ; Butanes 0·25 ; Methylcyclopentane 0·10 ; Cyclohexene 0·5 ; C_7 to C_{11} 0·87 ; Bicyclohexyl 0·55
Benzene	H_2 0·039 ; CH_4 0·019
Piperidine	H_2 3·69 ; C_2H_4 0·15 ; NH_3 0·06 ; Polymer 4·5
Pyridine	H_2 0·025 ; C_2H_2 0·016 ; C_2H_4 0·005 ; Bipyridyl 0·96 ; Polymer 3·86
Ethanol	H_2 5·0 ; H_2O 0·5 ; CO 0·06 ; CH_4 0·6 ; CH_2O 0·13 ; C_2H_2 0·02 ; C_2H_4 0·14 ; C_2H_6 0·24 ; CH_3CHO 2·8 ; $(CH_3CHOH)_2$ 2·2 ; $CH_3CH(OH)CH_2OH$ 0·13 ; other C_4 glycols 0·06
Ethyliodide	H_2 0·23 ; HI 0·33 ; I_2 2·12 ; CH_4 0·01 ; C_2H_2 0·09 ; C_2H_2 2·20 ; C_2H_6 1·92 ; C_4H_{10} 0·33

Polymerisation : As mentioned earlier, the formation of free radicals in irradiated aqueous systems was demonstrated by Dainton[88] via the polymerisation of acrylonitrile. However, there are other monomers which are known to undergo polymerisation by either cationic or anionic initiation mechanisms also. Since irradiation can produce both ionic and free radical species in any given medium (*see* e.g., equation 79) one can expect radiation-induced polymerisation to occur by one or more of these mechanisms depending on the nature of the compound and other additives present in the system. In a general way one can represent radiation-induced polymerisation in terms of equations

$$M \longrightarrow M^{\#} \tag{82}$$

$$M + M^{\#} \xrightarrow{k_{p_1}} MM^{\#} \tag{83-1}$$

$$MM^{\#} + M \xrightarrow{k_{p_2}} M_2M^{\#} \tag{83-2}$$

etc.

$$M_{n-1}M^{\#} + M \xrightarrow{k_{p_n}} M_nM^{\#} \tag{83-n}$$

$$M_{n_1}M^{\#} + M_{n_2}M^{\#} \xrightarrow{k_{bt}} P_1 + P_2 \tag{84}$$

or

$$M_nM^{\#} + X \xrightarrow{k_{ut}} P + X^{\#} \tag{85}$$

In this general chain reaction scheme $M^{\#}$ is the initiating active species which can be either a free radical, cation, anion or ion pair, P the polymeric species, X a chain terminating agent ($X^{\#}$ being incapable of propagation) and $k_{p_1} = k_{p_2} = \cdots k_{p_n} = k_p$. Because of the occurrence of such a chain reaction the *G*-value for the conversion of monomer to polymer can be very high although that for the initiating species is small (generally < 10). Depending on whether chain termination is bimolecular (reaction 84) or unimolecular (reaction 85) one can derive the following equations for the rate of polymerisation

$$R_p = \frac{k_p}{k_{bt}^{\frac{1}{2}}} G(\#)^{\frac{1}{2}} I^{\frac{1}{2}} [M]/10A^{\frac{1}{2}} \tag{86}$$

or

$$R_p = \frac{k_p}{k_{ut}} G(\#) I [M]/100A \tag{87}$$

Accordingly, it is seen that the rate is proportional to I^x, the dose rate exponent x being either 0·5 or 1 depending on the chain termination mechanism. Values of x between 0·5 and 1 could suggest a complicated mechanism with both types of termination occurring. The above scheme would predict a variation of the type shown in Fig. 16, curve A. However, in practice the observed variation may depart considerably from this (Curves B and C). For example, if additives or impurities are present which could scavenge the initiating species to form nonpropagating species, an induction period will be observed, no polymerisation occurring until the impurity is all used up. Alternatively if the impurity could be regenerated in the system during the course of irradiation, the limiting conversion may never reach 100 per cent. The type of behaviour represented by curve B is usually exhibited in the polymerisation of liquid systems wherein with increasing extent of conversion the material becomes highly viscous and as a result the termination step is considerably slowed down thereby indirectly enhancing the overall polymerisation rate.

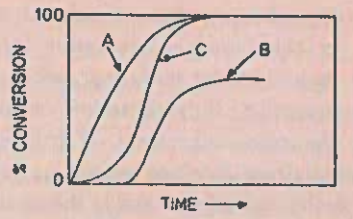


Fig. 16. Kinetics of radiation-induced polymerisation.
(Explanation discussed in the text)

For the two different mechanisms of chain termination, the net activation energy for the rate of polymerisation is given by

$$E_a = E_i + E_p - \tfrac{1}{2} E_{bt} \tag{88}$$

and

$$E_a = E_i + E_p - E_{ut} \tag{89}$$

where E_i, E_p and E_t are the activation energies respectively for the initiation, propagation and termination steps. As the initiation step in radiation-induced polymerisation, viz., the formation of the primary species (ions and free radicals) is temperature independent, E_i is zero. Depending on the relative magnitudes of E_p and E_{bt} or E_{ut} the observed overall activation energy can be positive, zero or negative.

In any given system whether polymerisation takes place by a free radical or ionic mechanism can be generally established by the use of specific scavengers.

Over the years the radiation-induced polymerisation of a large number of compounds have been investigated under a variety of conditions and there are available excellent books and monographs on this subject[26, 145]. One of the aspects that has attracted attention in recent years is the polymerisation of trioxane to give polyoxymethylene in the solid state. In a detailed investigation of this system Rao and Rao[141] have established that the polymerisation proceeds by a cationic chain mechanism and that traces of impurities such as water and formic acid unless scrupulously removed can lead to long induction periods and limiting conversions ($\sim 50\%$). Water for example, can scavenge the initiating cation and later get regenerated

$$RH^+ + H_2O \rightarrow R + H_3O^+ \quad (90)$$

$$H_3O^+ + e^- \rightarrow H + H_2O \quad (91)$$

They have also investigated the radiation induced postpolymerisation of this compound[142]. In this case, the irradiation is carried out at a low enough temperature such that the initiating species are trapped in the matrix and hence unreactive. Polymerisation occurs on annealing at a higher temperature, the rate and extent of conversion to the polymer depending on the irradiation dose and annealing temperature. A typical behaviour observed in the case of trioxane is shown in Fig. 17.

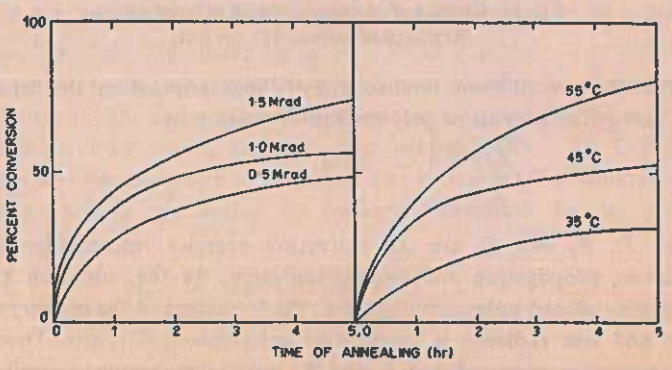


Fig. 17. Radiation-induced postpolymerisation of trioxane in the solid state (after Rao and Rao[142])

Irradiation of Polymers : This generally leads to two important effects : degradation to lower molecular weight fragments and crosslinking to give a more rigid three-dimensional network polymer. The relative extent to which these effects occur depend on various factors such as the chemical nature of the monomeric units, crystallinity and stereoregularity of the polymer, and conditions of irradiation such as the temperature and ambient gas. These radiation-induced effects in the polymer substantially affect its physical properties and have been exploited for several practical applications. Both degradation and crosslinking are the result of production of active sites such as ions and free radicals on the polymer backbone by the radiation. These active sites, once produced, can also be made use of to initiate polymerisation of another monomer so that a new polymeric chain is grafted on to the backbone polymer. The resulting graft copolymer will have the properties of both the constituent units and thus it is possible to tailor-make materials of desired properties. Rao *et al.*[142, 146, 147] have carried out detailed investigations of radiation-induced grafting of monomers with polar-end groups such as acrylic and methacrylic acids, acrylamide and acrylonitrile on synthetic fibres of nylon, polypropylene and polyester to impart on to the latter the practically useful and important properties of antistaticity, good dye uptake and moisture regain.

Copolymerisation without grafting can also be brought about by irradiation of mixtures of two or more monomers. In a porous material like wood it is possible to fill the voids with a monomer and polymerise it by radiation to give wood plastic composites which have considerably superior mechanical and aesthetic properties (*see* e.g., Rao *et al.*[139, 140]).

Other chain reactions : Chain reactions with *G*-values for conversion are not restricted to polymerisation reactions. Oxidation, sulfoxidation and sulfochlorination of organic hydrocarbons under irradiation can also proceed by a radiation-induced free radical chain mechanism. Thus, e.g., Moorthy and Rao[115] have observed high *G*-values for the conversion of a number of paraffinic and alicyclic hydrocarbons to the corresponding sulfonyl chlorides and have proposed the sequence of reaction steps which involve Cl as the chain carrier

$$RH \xrightarrow{\sim\sim} R + H \quad (92)$$

$$R + SO_2 \rightarrow RSO_2 \quad (93)$$

$$RSO_2 + Cl_2 \rightarrow RSO_2Cl + Cl \quad (94)$$

$$Cl + RH \rightarrow R + HCl \quad (95)$$

$$Cl + Cl \rightarrow Cl_2 \quad (96)$$

The practical importance of this reaction lies in the fact that the long chain ($C_{11} - C_{16}$) linear alkyl sulfonates which can be obtained by alkali hydrolysis of the corresponding sulfonyl chlorides have detergent properties and are biodegradable.

Biological Systems : The principal chemical components of most living systems are sugars, proteins, nucleic acids and lipids, in addition to water which is often a major constituent. The constituent units of the organic components, viz. the amino acids, pyrimidine and purine bases and the monomeric sugars are all known to be reactive with e^-_{aq}, H and OH radicals to varying extents. Reactions of the active species from water with the above biomolecules as well as direct effect of radiation on them leads to the formation of reactive sites on the macro-biomolecules causing scission and crosslinking of the latter. Some of the radical reactions can also be considerably affected by the presence of oxygen due to formation of peroxy-radicals and peroxides. Extensive researches have been carried out on these aspects at the molecular, cellular and organic levels and accounts of these can be found in references[2,14,159].

4. *Study of Transient Species*

Pulse radiolysis : The various primary radical species which were postulated to explain the kinetics of radiation-induced chemical changes had, because of their short lifetime, eluded direct observation. The experimental techniques then familiar to kineticists for the study of such transient or shortlived species were : stopped flow, rotating sector and flash photolysis methods[60]. An adaptation of the last-named technique to study radiation-induced transient species was rendered feasible with the advent of charged particle accelerators which could give intense bursts or pulses of electrons of short duration and it was thus that the technique of pulse radiolysis came to be evolved[81,102,104,152]. The first successful attempts to use this technique to directly observe and monitor the decay of the hydrated electron were reported by Boag and Hart[20], Hart and Boag[61] and Keene[82].

The principle of the pulse radiolysis method is briefly as follows : If a system is subjected to a pulse of radiation of duration τ which is just greater than the duration of the diffusion kinetic and physicochemical stage, the

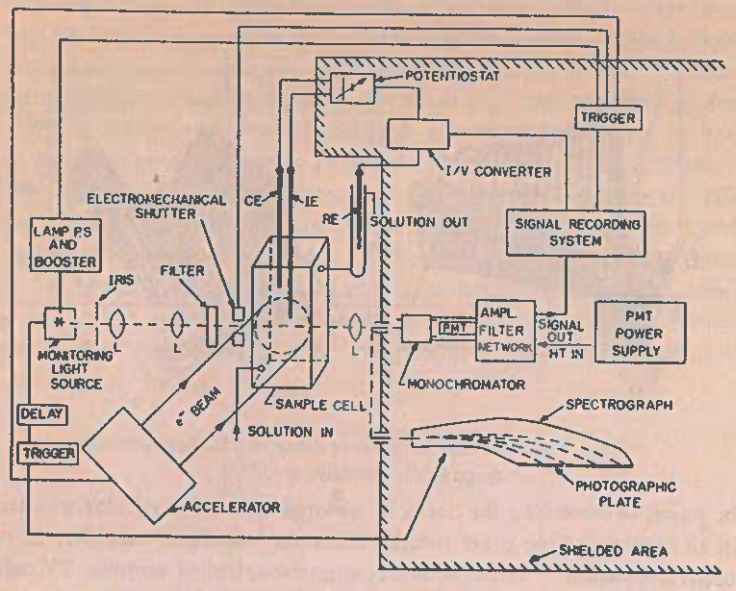


Fig. 18. Block diagram of experimental setup for pulse radiolysis

system will contain, at the end of the pulse, all the primary species formed from the major component of the system at concentrations which can be computed from their respective initial yields and the dose delivered to the system by the pulse. If we choose some specific observable quantitative property of a particular species we can follow the time profile of change of this property in the post-pulse period and relate it to the reaction kinetics of the species via appropriate rate equations. The time involved in the measurement of the property should be much less than the lifetime of the species to which it pertains or in other words the detector employed for monitoring the property should have a response faster than the rate of decay of the species. This limits the possibilities to optical absorption (employing photographic plate or photomultiplier or silicon photodiode) current due to electroreduction or oxidation at a suitable electrode (generally a hanging mercury drop) and electrical conductivity (in the case of charged species). The first of these is the most popular and widely used and can be adapted to either photographing the absorption spectrum at different periods after

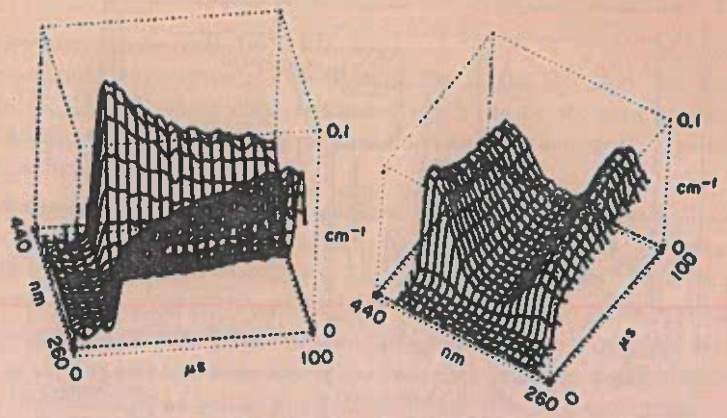


Fig. 19. Three-dimensional plots of decay of transient absorption spectra (after Schmidt *et al.*[158])

the pulse, or observing the decay of absorbance at a particular wavelength on an oscilloscope or other suitable transient waveform recorder. A more recent innovation[158] is the use of a computer-controlled scanning TV camera to give a three-dimensional picture exhibiting the entire absorption spectrum as well as its decay profile (Fig. 19). Starting from microsecond pulses in the early sixties, the time resolution has been reduced to the nanosecond region in the seventies[75] and later to a few tens of picoseconds employing the fine structured micropulses contained in a nanosecond pulse[22]. A lower limit on the time resolution for the optical detection scheme is the time of flight of the analysing light beam over the cell-path length and is about 3 p. sec. for a 0·1 cm path. With picosecond pulses it is possible to probe the events occurring in the physicochemical stage of radiolysis. Pluse radiolysis work in the nano-micro second time regime has helped to evaluate the physicochemical properties and yields of primary species and the rate constants for reactions of these among themselves and with a large number of compounds. The more significant among these are briefly described below.

Hydrated electron

Undoubtedly the most important contribution of pulse radiolysis to radiation chemistry is the direct observation of the hydrated electron via its optical absorption spectrum. This spectrum along with those of other

species which were subsequently mapped by the same technique are shown in Fig. 20. The evidence in favour of assigning this spectrum to the hydrated electron was based both on its shape (which was similar to that of solvated electrons formed in metal ammonia solutions) but more particularly from the strong inhibitory effect of additives known to be efficient hydrated electron scavengers from earlier radiation chemical work. The fact that the species bears a unit negative charge was again established from the effect of ionic strength on the rate constant of its reaction with ferricyanide[184]. Its molar extinction coefficient at 578 nm was determined by Rabani *et al.*[185] making use of its rapid and quantitative reaction with tetranitromethane to give nitroform anion whose extinction coefficient is known. Avery *et al.*[18] observed the EPR spectrum of e^-_{aq} ($g=2{\cdot}0002$, $\Delta H_{ms}=0{\cdot}5$ gauss) in pulse-irradiated water.

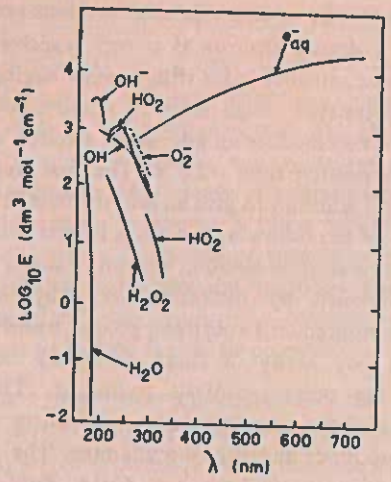


Fig. 20. Absorption spectra of transient species in irradiated water. The spectra of stable species are also given for comparison (after Matheson and Dorfman[103])

Various other physical, physicochemical and radiation chemical properties of the hydrated electron have been accurately evaluated[162]. Many of these have been evaluated by the use of the pulse radiolysis technique. The kinetic data and the diffusion constant have been used in many

diffusion kinetic model calculations to predict the time history of evolution of the primary yields in the physicochemical and diffusion kinetic stages.

Rate constants for reactions of e^-_{aq} : Because of the high molar extinction coefficient of e^-_{aq} in the visible spectral region experimental determination of the rate constants for the reactions of the hydrated electron is relatively easy. In the tens of nanosecond and longer time regime, all intra- and inter-spur reactions are over and only homogeneous kinetics prevail. One can choose experimental conditions such that only the desired reaction predominates and is pseudo first order and all others involving the hydrated electron are suppressed. The rate constant for the desired reaction can then be directly computed from the semilog plot of the decay of absorbance of e^-_{aq} with time. Even when the above conditions cannot be realized, the rate constants can be evaluated by a curve fitting procedure using a computer[62] if the rate constants for the undesirable side reactions are known.

Chemically the hydrated electron is a very reactive entity and reacts with almost diffusion-controlled rates with species having one or more of the following characteristics : high electron affinity, good electrophilicity, a vacant orbital to accommodate an additional electron and redox potential (for oxidation) more positive than $-2{\cdot}6$ V. The first step in the reaction is invariably an electron addition to give an intermediate. This electron adduct then undergoes one of the following reactions to give relatively longer lived species : monomolecular dissociation, proton transfer from the solvent, protonation accompanied by hydration or dehydration, association, association with the nonreduced substrate, charge transfer to other solutes, disproportionation, etc. Many of these secondary reactions have been studied by using the pulse radiolysis technique. Thus Moorthy and Hayon[117] have studied the transformations of electron adducts of various heteroaromatic compounds and B-group vitamins. The unfamiliar valence states of a number of inorganic ions, e.g., Cd^{+1}, Zn^{+1}, Ag^0, Ag_2^{+1}, Tl_2^{+1}, etc., have also been identified in pulse radiolysis experiments[17]. The reactions of e^-_{aq} with itself, H and H_3O^+ exhibit significant isotope effects, and this has led to the conclusion that these involve the intermediate species : e^-_2aq, H^- and H_3O^{63}. The activation energies of many reactions involving e^-_{aq} are very low ($\sim$3-4 kcal mole^{-1}). It is possible that an electron tunnelling mechanism is involved in many hydrated electron reactions.

Rate constants for H and OH reactions : The direct measurement of the rate constants for reactions of H atoms and OH radicals by the pulse radiolysis method is rather difficult because they absorb in the UV where interfering absorptions are likely, and the extinction coefficients at the absorption maxima are rather low (Fig. 20). Wherever convenient spectral regions and high enough extinction coefficients permit, these measurements have been attempted by following the kinetics of growth of the product transient or by competition methods involving observations of either the kinetics of growth of a product species from a reference solute or from steady state concentrations of the product transient from the reference solute. Unless the mechanisms of the concerned reactions are sufficiently well known to establish that the observed transient is the first or the only step in the reaction of H or OH with the solute concerned, the rate constant values so determined may be erroneous.

The kinetics and mechanisms of reactions of the radical products from reactions of e^-_{aq}, H and OH with a large number of compounds have also been studied using the pulse radiolysis technique. Among these, particular mention may be made of oxygenated species O_2^-, HO_2, O_3^-, HO_3, etc.[19], because of their possible formation and reactions in many radiation chemical system studied in the presence of ambient air. More importantly, Rao and Hayon[144] have developed a pulse radiolytic method for the determination of the redox potentials of various radical species by redox titration of these radicals with electron acceptors and donors of known redox potential. Henglein[68] has employed a combination of polarography with pulse radiolysis to deduce the redox potentials of transient radicals.

Matrix isolation : Even before the advent of pulse radiolysis, attempts had been made to demonstrate the formation and reactions of H atoms, OH radicals and hydrated electrons in irradiated aqueous systems by another technique well known to free radical chemists : matrix isolation and EPR which makes use of the paramagnetic property of free radical species and their relatively high stability in a rigid matrix at low temperatures. Thus Livingston, Zeldes and Taylor[99] were the first to observe the theoretically expected doublet EPR spectrum of H atoms in γ-irradiated frozen samples of mineral acids at 77 K. Piette *et al.*[127] observed this doublet and another, attributable to OH radicals, in the EPR spectrum of pure ice irradiated and monitored at 4 K. On warming, the former disappeared at $\sim$20 K and the

latter at ~110 K. This early work was followed by more detailed investigations on the radiation chemistry of frozen aqueous systems by several groups of researchers. Kevan, Moorthy and Weiss[84] were the first to establish that in acidic ices the H atoms are formed by reaction of the radiation produced mobile electron $(e_m^-)^*$ with the acid anions according to

$$e_m^- + HSO_4^- \rightarrow H + SO_4^= \quad (97)$$

These H atoms as well as the ones formed directly by radiation action on the water molecules in the solvation shell of oxyanions were found to be stabilized at the site of formation at 77 K. Moorthy and Weiss[111] studied reactions of mobile electrons with Cd^{+2}, Zn^{+2}, Hg^{+2}, Mg^{+2} and $CrO_4^=$ in γ-irradiated frozen aqueous systems at 77 K. The EPR spectra of the product species are shown in Fig. 21. It is noteworthy that unlike the hydrated electron in water the mobile electron in ice does react with Mg^{+2}. Also, whereas the Hg(I) state in aqueous solutions is the dimeric diamagnetic form Hg_2^{+2}, the one formed by mobile electron reaction in ice is the monomeric Hg^{+1} and hence is paramagnetic. In addition to the mobile electron, Moorthy and Weiss[106] were the first to establish the radiation produced hole (h_m^+) to be a chemically significant radiation produced species in frozen aqueous systems capable of reacting with appropriate solutes, e.g.,

$$h_m^+ + HSO_4^- \rightarrow H^+ + SO_4^- \quad (98)$$

Further, they also demonstrated reversible interconversion between the SO_4^- and similar species formed in reaction (98) and the trapped hole, e.g.,

$$SO_4^- \xrightarrow{h\nu} SO_4^- + h_{tr}^+ \quad (99)$$

In subsequent years Moorthy, Rao and Shankar[116] found that in γ-irradiated frozen aqueous azide, NH_2 radicals are formed at 77 K by reaction of H atoms with the solute

$$N_3^- + H \rightarrow NH^- + N_2 \quad (100)$$

$$NH^- + H_2O \rightarrow NH_2 + OH^- \quad (101)$$

In acidic solutions of alcohols and other organic compounds both mobile electrons and holes were found to react with the solute to give a greatly enhanced yield of the organic radical[114].

* Originally this was thought to be the equivalent of the hydrated electron but now it is considered to be a different species and possibly a precursor of the latter.

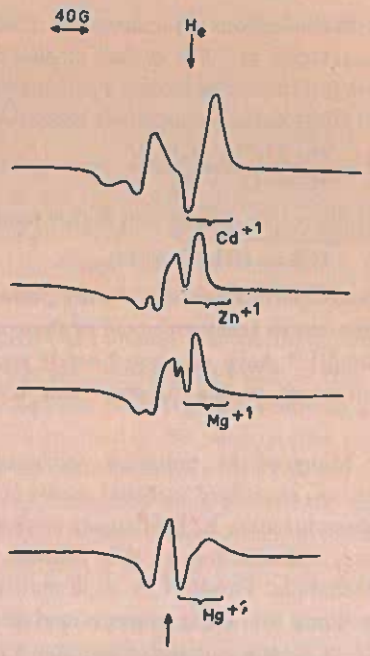


Fig. 21. EPR spectra of Cd^{+1}, Zn^{+1}, Hg^{+1}, Mg^{+1} and Cr(V) species in frozen aqueous media at 77 K (after Moorthy and Weiss[109])

In γ-irradiated alkaline glasses trapped electrons were observed both by EPR[110] and optical absorption[84] and mobile holes found to react with OH^- to give O^-

$$e_m^- + \text{trap} \rightarrow e_{tr}^- \quad (102)$$

$$h_m^+ + OH^- \rightarrow O^- + H^+ \quad (103)$$

whereas trapping and reaction with solutes was competitive at 77 K, the trapped species were found to undergo various transformations and reactions with dissolved solutes on photobleaching and warming to higher temperatures. Thus, with O_2 as a solute in the alkaline matrices, Moorthy *et al.*[118] have found evidence for the formation of O_2^- and $O_3^=$ species :

$$e_{tr}^- \xrightarrow[\Delta]{h\nu_{or}} e_m^- \left\{ \begin{array}{l} \xrightarrow{O_2} O_3^- \\ \xrightarrow{O^-} \\ \longrightarrow O^= \end{array} \right\} \rightarrow O_3^= \quad (104)$$

One of the indirect conclusions from the study of the radiation chemistry of frozen aqueous systems at 77 K is that in pure ice the H and OH are formed by a process not involving mobile electrons and holes and undergo reactions (105) and (106) in the temperature range ~20 K :

$$H_m + OH_m \rightarrow H_2O \quad (105)$$

$$H_m + H_m \rightarrow H_2 \quad (106)$$

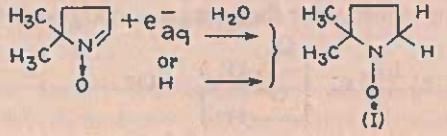
whereas reaction (107) occurs only at still higher temperatures (~110 K)

$$OH_m + OH_m \rightarrow H_2O_2 \quad (107)$$

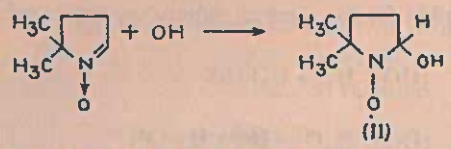
The reactions of mobile electrons and holes are observed only when appropriate solutes are present to react with *both* of these species.

Mittal and Hamill[106] have also conclusively established that electron trapping and reactions of mobile electrons and holes occur in organic glasses.

Spin trapping : Many of the radiation produced radicals are highly reactive and hence too shortlived in liquid media at ambient temperatures to permit their observation by EPR in steady state radiolytic experiments. As discussed before, one solution to this problem was afforded by the matrix isolation technique. However, a solid matrix is not comparable to the corresponding liquid in many respects and the quantitative radical yields evaluated from matrix isolation experiments may not hold good for the liquid. Even the nature and properties of some of the species may be different in the two matrices. Thus for example the hydrated electron in water and the trapped electron in aqueous glasses have very different optical absorption maxima. In order to permit direct EPR observation in the liquid, the technique of spin trapping has been recently developed[101, 153, 164, 171]. This is based on the fact that nitroso compounds and nitrones react with free radicals to give nitroxides which are relatively more stable and exhibit well-defined narrow line EPR spectra with hyperfine constants characteristic of the original radical. Thus with 5, 5' dimethyl pyroline-N-oxide as the spin trapping agent for the radicals from water, the nitroxides formed are[154] :



(108)



(109)

(I) was found to exhibit hyperfine interaction with one N nucleus and two equivalent protons whereas (II) exhibited interaction with one N and H each. In presence of N_2O the yield of (I) was greatly (but not completely) reduced and that of (II) almost doubled which is in agreement with the above assignment since $G_{e^-_{aq}} \simeq G_{OH} \gg G_H$ and N_2O is known to scavenge e^-_{aq} quantitatively to give OH radicals. The technique has since been used to identify and quantitatively measure the yields of free radicals formed in a number of organic liquids. However, the yields so evaluated must be interpreted with caution in view of the uncertainties involved in measurements of absolute radical concentrations by EPR. Also, as will be seen in Section V the measured yields may be in error if the spin traps are highly reactive towards the primary species and/or their precursors.

V. Diffusion Kinetic and Physicochemical Stage

It is in this stage of radiolysis that we look for a rational mechanism for the formation of the different primary species discussed in the previous Section. In this connection some of the important questions that have arisen over the years are : how are electrons solvated, what is the relative contributions of excitations, ionizations and subexcitation electrons, whether molecular primary products are formed by a direct process involving radiation action on a single molecule of the medium, do parent positive ions have long enough life times to react with solutes in competition with the solvent, etc. In order to answer these and similar questions one or more of the following approaches have often been adopted in the past : speculation based on the known theories of physics and chemistry, experimentation aimed at directly probing the processes occurring in the time regime characteristic of this stage and theoretical model calculations to best fit the experimental data.

1. *Early Models*

Weiss[178] was the first to propose a mechanism for the formation of H and OH as primary species in irradiated water based on the then

prevalent explanation for the electron affinity spectra of anions in aqueous solution

$$(HO)^- H^+ \rightarrow HO + H \quad (110)$$

or

$$(HO)^- H_2O \rightarrow HO + H + OH^- \quad (111)$$

the radiation being assumed to play the role of facilitating the electron transfer in the direction indicated by the curved arrow.

Electron escape model : Unlike the explanation given by Weiss subsequent ones have specifically taken into account the ionizing behaviour of the high energy radiations. Thus Lea[96], Gray[56] and Platzman[129] calculated that the secondary electron released in the ionization act would travel on the average $\sim$ 50 Å from the parent positive ion before thermalization. At this distance the electrostatic attraction between the electron and the parent positive ion in a polar medium like water would be too small to force recombination of the two and as a result the electron would be solvated (to give e^-_{aq}) while the sibling H_2O^+ ion would undergo rapid ion molecule reaction to give OH and H_3O^+.

$$H_2O \longrightarrow H_2O^+ + e^-_{fast} \quad (112)$$

$$e^-_{fast} \xrightarrow{H_2O} e^-_{thermal} \xrightarrow{H_2O} e^-_{aq} \quad (113)$$

$$H_2O^+ + H_2O \rightarrow OH + H_3O^+ \quad (114)$$

Weiss[175] held the view that the charges of the positive ion and electron are delocalized over the water dipoles in the vicinity as happens in an ionic solid, to give positive and negative polarons symbolized as $(H_2O)^+$ and $(H_2O)^-$ and explained the formation of molecular H_2 and H_2O_2 via reactions

$$2(H_2O)^- \rightarrow H_2 + 2OH^- \quad (115)$$

$$2(H_2O)^+ \rightarrow H_2O_2 + 2H^+ \quad (116)$$

occurring in competition with

$$(H_2O)^+ + (H_2O)^- \rightarrow 2H_2O \quad (117)$$

The occurrence of a reaction such as (115) has subsequently been directly demonstrated in pulse radiolysis experiments. Although evidences for reactions of mobile holes in frozen aqueous systems[109] and of dry holes in fluid aqueous systems[46] have been obtained, there is no convincing proof for the

participation of a hydrated positive hole $(H_2O)^+$ in aqueous media and it is possible that hydration of the hole does not successfully compete with its ion molecule reaction with water to give OH radicals.

Electron capture model : An entirely different viewpoint was put forward by Samuel and Magee[151] according to which the thermalization distance was much smaller so that electrostatic attraction between the electron and H_3O^+ could be strong enough to allow recombination of the two to give eventually H, OH pair.

$$e^- + H_2O^+ \rightarrow H + OH \quad (118)$$

Molecular product formation was accounted for via reactions (17) and (18)

$$H + H \rightarrow H_2 \quad (17)$$

$$OH + OH \rightarrow H_2O_2 \quad (18)$$

occurring during the subsequent diffusion of these species out of the spur. Although in its original form this model does not predict formation of the hydrated electron, the subsequent discovery of this species in the early sixties could be accommodated into the parent ion recapture scheme by assuming that the thermalized electron returns to H_3O^+ to give H_2O which dissociates according to

$$H_2O \rightarrow H_3O^+ + e^-_{aq} \quad (119)$$

The principal drawback of the Samuel-Magee theory both in the original and modified forms lies in the fact that it predicts similar spatial distribution, and hence similar concentrations, for H and OH in the spur (Fig. 22a) so that, assuming $k_{17} \simeq k_{18}$ and $D_H \simeq D_{OH}$, one would expect nearly equal yields of molecular H_2 and H_2O_2. This is contrary to the experimental finding $G_{H_2O_2} > G_{H_2}$. As we now know $k_{17} > k_{18}$ and $D_H > D_{OH}$, and hence we should expect, if at all, $G_{H_2} > G_{H_2O_2}$ on the basis of the Samuel-Magee model. On the other hand, the Lea-Gray-Platzman model predicts electron thermalization and solvation away from the site of reaction (114) and hence the spatial distribution of e^-_{aq} is expected to be broader than that of OH (Fig. 22b). As will be seen later, diffusion kinetic calculations which quantitatively take this aspect into account and the known or plausible values of diffusion coefficients and rate constants, predict molecular product yields in better agreement with the experimental data. Although the Lea-Gray-Platzman theory thus seems to be favoured in the

case of water, in general, both parent ion recapture followed by radical formation, and electron escape followed by solvation processes may be concurrently occurring, the relative extents of the two being governed to some extent by the physical properties of the matrix such as the dielectric constant, velocities of charge migration and dipole relaxation time.

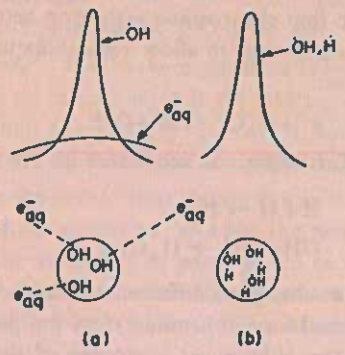


Fig. 22. Expected spatial distribution of radicals in the spur : (a) Lea-Gray-Platzman theory, (b) Samuel-Magee theory

2. *Diffusion Kinetic Models*

In order to quantitatively test the validity of any theory for the formation of the primary species in irradiated systems one requires to compute their yields as a function of experimental variables such as LET, dose rate, solute concentration and time of observation and compare these with the corresponding experimentally measured quantities. Such computations are based on the consideration that the primary species (either all or some) are initially formed in reactive zones (spurs, blobs and short tracks) in high concentrations and as these reactive zones are separated by regions (or the bulk) where such species are not present to start with the species diffuse out from the reactive zones into the bulk. The diffusion continues until homogeneity is attained and during the *diffusion* the species can react among themselves as well as with solutes if present according to known *kinetic* laws —hence the name : diffusion kinetic model. The various inputs required in such diffusion kinetic model calculations are : the initial yields of the species in the reaction zones, their spatial distribution and diffusion constants, concentrations of solutes and rate constants of all relevant reactions

(including both formation and loss of the species) and interzonal (reflecting LET) and intertrack (reflecting dose rate) separations. Making use of Fick's laws of diffusion and chemical rate equations one can write the variation of the concentration of any species in the form

$$\frac{\partial C_i}{\partial t} = D_i \nabla^2 C_i - \sum_j k_{ij} C_i C_j + \sum_{j,\, h' \neq i} k_{jk'} C_j C_{k'} - \sum_j k_i C_{ij} C_i \qquad (120)$$

Here ∇ is the Laplacian operator, D, C and k are respectively the diffusion coefficients, concentrations and the bimolecular rate constants, and the subscripts i, j, k and s refer to the ith, jth, kth species and solute respectively. The required solution is the time variation of C_i till diffusion is complete and can, in principle, be obtained by a numerical integration of the above set of partial differential equations. The initial spatial distribution of the radicals is invariably assuumed to be Gaussian, and further, as a useful simplification (called the prescribed diffusion approach which the calculations of Kuppermann[93] have shown to be quantitatively valid) assumed to remain so all through the diffusion stage. Samuel and Magee[154] were the first to perform such calculations, but assumed for simplicity only one radical (R) and no solute so that the only reaction to be considered was

$$R + R \rightarrow \text{products} \qquad (121)$$

In the years that followed, calculations involving various degrees of complexity have been performed such as : one radical and one solute (Ganguly and Magee[51], Fricke[50], Flanders and Fricke[47]), two radicals and one solute (Dyne and Kennedy[45]) and many radical – many solute (Kuppermann[91], Schwarz[160] and Girija and Gopinathan[54]). In the last model, all the species ("radicals" which term for the purposes of these calculations does not necessarily exclude the molecular species H_2 and H_2O_2) conceivably present a few picoseconds after the passage of the ionizing particle are taken into account and typical solutes capable of reaction with each of these species (except for H_3O^+, OH^-, H_2 and H_2O_2) are considered. The initial G-values of the various species taken into account in the last three many radical calculations[54,91,160] are summarized in Table 8 along with their diffusion constants and initial spur radii. The rate constants used are tabulated in the original references. As compared to the older models the additional improvement introduced by Girija and Gopinathan is the assumption that the spur radii themselves have a Gaussian distribution.

Table 8

Initial Parameters of Some of the Diffusion Kinetic Model Calculations

Species	Kuppermann[a]			Schwarz[b]			Girija and Gopinathan[c]		
	G°	D $\times 10^{-5}$ cm^2 s^{-1}	γ_0 (Å)	G°	D $\times 10^{-5}$ cm^2 s^{-1}	γ_0 (Å)	G°	D $\times 10^{-5}$ cm^2 s^{-1}	γ_0 (Å)
e^-_{aq}	2·08	4·5	18·75	4·78	4·5	23	5·31	4·9	
H	0·18	8·0	6·25	0·62	7·0	7·5	0·69	7·0	
H_3O^+	2·08	10·0	6·25	4·78	9·0	7·5	5·31	9·0	
OH	2·10	2·0	6·25	5·70	2·8	7·5	6·0	2·8	(See Text)
O	0·08	2·0	6·25	0	—	—	—	—	
OH^-	0	2·0	—	0	5·0	—	—	5·0	
H_2	0	—	—	0·15	—	—	—	—	
H_2O_2	0	1·4	—	0	2·2	—	—	2·2	

Initial time : (a) and (b) $10^{-11} - 10^{-10}$ sec, (c) 10^{-12} sec

As such, all values of initial spur radii from 0 – 20 Å are taken into account with the restriction $r^\circ_{e^-_{aq}}/r^\circ_{OH,H_3O^+} \simeq 3{\cdot}0$. Their calculations predict for $G_{e^-_{aq}}$ and G_{OH} at 200 psec values of 4·25 and 5·0 as compared to the experimentally measured values[78,79] of 4·48 and 5·9 respectively. Similarly, the computed variation of $G_{e^-_{aq}}$ with time and of $G(H_2)$ and $G(H_2O_2)$ with the concentration of hydrated electron and OH radical scavengers agree with the experimental data very closely (Figs. 23–25).

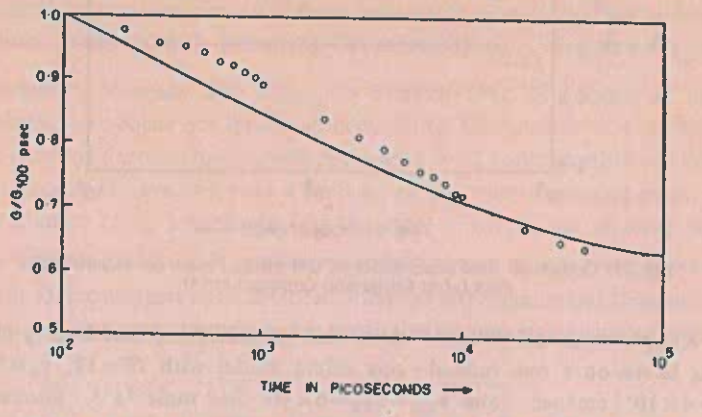


Fig. 23. Computed time dependence of e^-_{aq} yield. Points are experimental data (after Girija and Gopinathan[54])

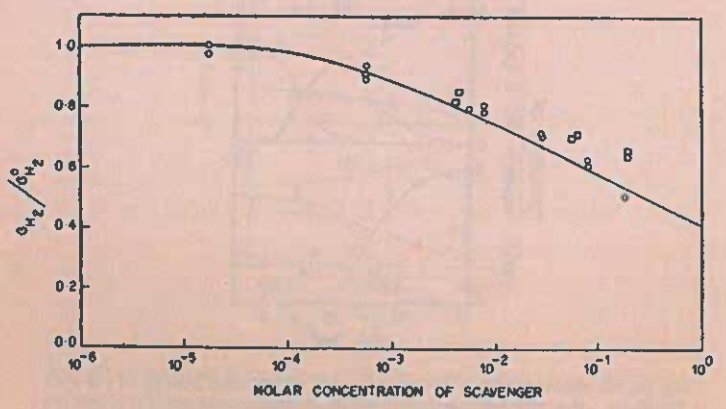


Fig. 24. Variation of G_{H_2} with the concentration of e^-_{aq} scavenger. Points are experimental data (after Girija and Gopinathan[54])

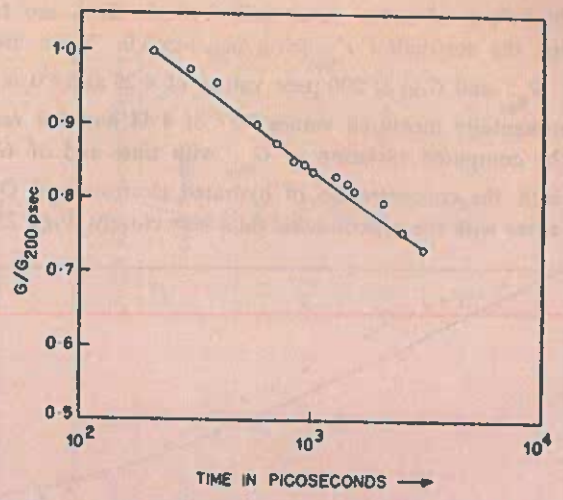


Fig. 25. Computed time dependence of OH yield. Points are experimental data (after Girija and Gopinathan[84])

Kuppermann and Belford[90] have also computed the time history of the spur based on a one radical – one solute model with $G^{\circ}=12$, $\gamma_0=10$ Å, $D=4\times10^{-5}$ cm² sec⁻¹ and $k_{RR}=k_{RS}=6\times10^{9}$ dm³ mole⁻¹s⁻¹. The results (Fig. 26) reveal that the radical – radical reactions are complete in about

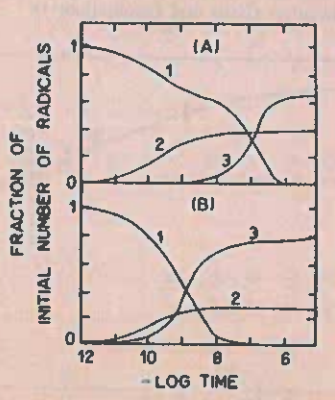


Fig. 26. Computed diffusion kinetic curves of the time history of the spur. Ordinate : Fraction of initial number of radicals—that survive (curve 1), that are built into molecular produce (curve 2) and that react with the scavenger (curve 3) for (A) 0·001 mole dm⁻³ and (B) 0·1 mole dm⁻³ solute concentrations (after Kuppermann and Belford[90])

1 nanosec. Hence, in order to intercept these reactions one requires solute concentration > 0·1 mole dm⁻³ and subnanosec observation times. In the 1–10 picosec region even 0·1 mole dm⁻³ concentration of solute is insufficient to intercept the radical-radical reactions.

The many radical diffusion kinetic calculations also predict that with increasing hydrated electron reactivity ($k\times C_S$) there should be an increase in $G_{e^-_{aq}}$, G_{OH} and $G_{H_2O_2}$ and a decrease in G_H and G_{H_2} (Fig. 27) due principally to the interception of intra-spur reactions involving the hydrated electron. Similarly, with increasing OH reactivity $G_{e^-_{aq}}$, G_{OH} and G_{H_2} are predicted to increase and $G_{H_2O_2}$ to decrease (Fig. 28). Some of these predictions are borne out by experiment. Thus Chouraquie and Sutton[88] have reported decrease in G_{H_2} with increasing NO_3^- concentration and Peled and Czapski[126] have observed a decrease in G_{H_2} with increasing reactivity for a number of e^-_{aq} scavengers. In the case of $G_{H_2O_2}$ an increase with increasing reactivity for e^-_{aq} scavengers and a decrease with increasing reactivity for OH scavengers have also been observed by Draganic and Draganic[44]. The conclusion can, therefore, be made that the major portion of atomic

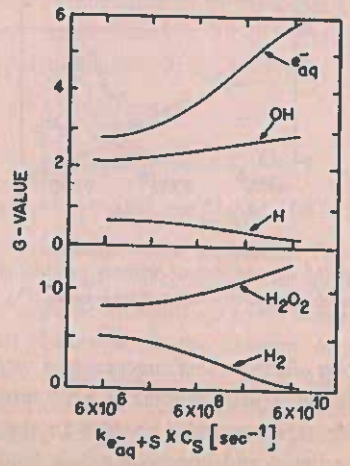


Fig. 27. Computed dependence of primary product yields on hydrated electron reactivity (after Kuppermann[91])

and molecular hydrogen and molecular H_2O_2 yields and HO_2 originate from the intra-spur reactions

$$e^-_{aq} + H_3O^+ \rightarrow H + H_2O$$
$$e^-_{aq} + e^-_{aq} \rightarrow H_2 + 2OH^-$$
$$OH + OH \rightarrow H_2O_2$$
$$H + H_2O_2 \rightarrow H_2O + OH$$

These and water reformation reactions

$$e^-_{aq} + OH \rightarrow OH^-$$
$$H + OH \rightarrow H_2O$$

and the reactions of these species with solutes are all competing pathways during the diffusion stage.

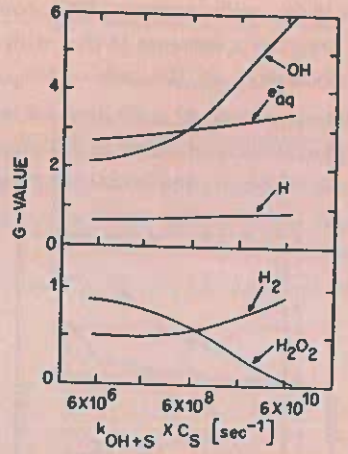


Fig. 28. Computed dependence of primary product yields on OH radical reactivity (after Kuppermann[91])

Role of subexcitation electrons and superexcited molecules : In spite of the success of the diffusion kinetic models in accounting for the variation of product yields with experimental variables in the right direction, the agreement between predicted and observed values tends to be progressively poorer with increasing solute concentration. In order to explain such discrepancies alternative and additional sources of primary species have

been proposed from time to time. Some of these are briefly discussed here. Thus Platzman[181] has suggested subexcitation electrons to be possible precursors of molecular H_2 :

$$e^-_{sc} + H_2O \rightarrow H^- + OH$$
$$H^- + H_2O \rightarrow H_2 + OH^-$$

while Hamill[59] has proposed the scheme :

$$e^-_{th} + H_2O^+ \rightarrow H_2O^* \rightarrow H\,(hot) + OH$$

or

$$e^-_{th} + H_3O^+ \rightarrow H_3O^* \rightarrow H\,(hot) + H_2O$$
$$H\,(hot) + H_2O \rightarrow H_2 + OH$$

Similarly Anbar[11] has suggested that superexcited H_2O^{**} may give rise to molecular H_2O_2 :

$$e^-_{fast} + H_2O \rightarrow H_2O^{**} + e^-$$
$$H_2O^{**} \rightarrow OH^+ + H + e^-$$
$$OH^+ + H_2O \rightarrow H_3O^+ + O$$
$$O + H_2O \rightarrow H_2O_2$$

Normal excited water molecule has also been often invoked[67] as the source of molecular products. Anbar[11] has discussed these alternative pathways in some detail and there do not seem to be available positive experimental proofs for any of these. More recently Singh, Chase and Hunt[164] have also proposed additional pathways for e^-_{aq} formation :

$$H_2O^* \xrightarrow{H_2O} H_2O^+ + e^-_{aq}$$
$$H_2O + e^-_{se} \rightarrow H_2O^{*V} \xrightarrow{H_2O} H^+_{aq} + OH^-_{aq}$$
$$H_2O^{*V} + OH^-_{aq} \rightarrow H_2O + OH + e^-_{aq}$$

and presented experimental data which apparently (but not unquestionably) support the proposed schemes. Participation of excited molecules in the radiation chemistry of organic compounds have, however, been more convincingly demonstrated. In water, in the absence of positive experimental evidences to support any or all of the above additional pathways, no attempt has been made so far to include their effects in diffusion kinetic model calculations. Introduction of such refinements in the calculations would undoubtedly further increase their complexity because of the fact that additional reactions of solutes with subexcitation electrons, hot hydrogen atoms and excited H_2O need to be taken into account.

As pointed out earlier electron escape as considered by Lea, Gray and Platzman and parent ion recapture as proposed by Samuel and Magee are two extremes and the real situation may be a mix of the two. The recent many radical models all assume initial distributions as appropriate to the former viewpoint. Inclusion of a contribution to initial distribution as appropriate to the Samuel-Magee case and correlating spur sizes with the number of radicals therein may lead to better agreement between calculated and experimental trends. Probably one of the major drawbacks of diffusion kinetic model calculations is the assumption of validity of macroscopic parameters such as chemical rate and diffusion constants for a small aggregate of species in a restricted region.

3. *Electron Solvation*

The experimentally observed upper limit for rate constants of hydrated electron reactions are close to those expected for diffusion-controlled processes of hydrated ions in solution and hence the inference that the species is a hydrated or localized electron rather than a free thermal electron which, if that were the case, could react much faster. The term hydrated (or solvated in general) implies that the electron is surrounded by a sheath of oriented solvent dipoles as shown in Fig. 29. This picture has derived

Fig. 29. Picture of a solvated electron as an electron surrounded by oriented solvent dipoles

considerable support from the classical studies on metal-polar solvent systems (e.g., Na in liq. NH_3) wherein one of the species formed has been inferred to be a solvated electron. As pointed out by Moorthy and Shankar[112] the dynamics of thermal electrons and dipole relaxation are not compatible with such a simplified picture. The dynamics of electron ejection from a metal in a polar solvent is rather different from the radiation chemical case, and the question, therefore, arises as to whether in an irradiated medium the electron really gets solvated in the above sense and if so, how a thermalized electron

passes over to this state, or in other words does the solvated electron in these systems have a structure different from that implied in the above simple picture. Much research has been carried out to arrive at the answer, which however has not satisfactorily emerged yet although many links have been established. The experimental approach has been to probe the "solvation" process by working in picosecond time scales in fluids (wherein dipole relaxation timer are of the order of 10^{-11} sec.) or to delay this and diffusion limited processes as is possible in rigid frozen matrices.

Models for electron solvation : The classical model of the hydrated electron as proposed by Jortner[80] and Weiss[175] is based on the "polaron" model for excess electrons in ionic solids developed by Davydov[86], Diegen[87] and Pekar[125] and treats the medium as a dielectric continuum. By comparison of observed dependence of the absorption maximum of the hydrated electron in both water and ice on temperature with the trends predicted on the basis of the polaron model, Moorthy and Shankar[112] concluded that if this model is correct then it is the electronic and not dipole orientational polarisation which is to be taken into account. Whereas the polaron model takes into account only the dielectric properties of the medium and ignores the microscopic structure altogether, the structural model proposed by Natoari and Watanabe[138] does just the opposite. According to this model defective sites (with the central water molecule of the normal tetrahedral structure missing) are potential traps for solvating a mobile thermalized electron as shown in Fig. 30.

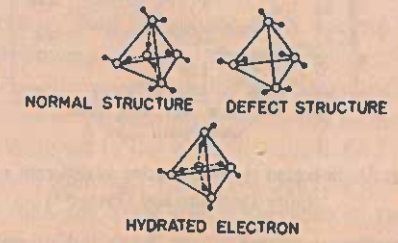


Fig. 30. Structural model of the hydrated electron

Presolvation processes : Studies on trapped electrons in various aqueous and organic glasses[84, 85] revealed a number of clues to the mechanism of electron trapping. These were : the broad, structureless optical absorption is due to electronic transitions to the conduction band from bound states of

varying depths; the λ_m of the species trapped at ~ 4 K is red shifted as compared to the ones trapped at ~ 77 K, the shift being greater in polar glasses as compared to hydrocarbon glasses; the transition from the "red" species to the "blue" species is readily observed on warming to higher temperatures. As dipole relaxation time is longer at lower temperatures, the above results can be interpreted to mean that the "red" species represent the earlier stage of solvation as compared to the "blue" species and are the precursors of the latter. Magnetic resonance experiments ont he trapped species revealed that the electron is surrounded by protons as its immediate neighbours in what might be called a first solvation sheath, the distance from the electron to the nearest proton estimated as ~ 2 Å in aqueous glasses[87] and 3 Å in MTHF glasses[86]. Similar results emerged from pulse radiolysis experiments in fluid media. Thus it was observed that the hydrated electron reaction rate constants were different at high solute concentrations (i.e., at short times) than in dilute solutions[77]. Hamill[58] obtained good experimental evidence for reactions of the precursor of the hydrated electron and designated the former as dry electron. Investigation of a large number of solvents revealed that the solvated electron spectra in these fall into definite groups depending on solvent types[42] (*see* Fig. 31). In mixed

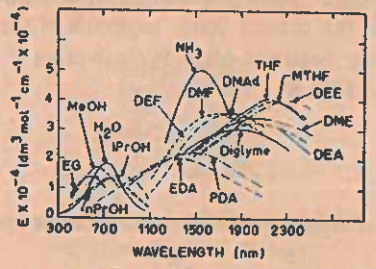


Fig. 31. Solvated electron spectra in different solvents (after Dorfman and Gavlas[42])

solvents the absorption maxima either depended on solvent composition or were exclusively characteristic of one component irrespective of composition in which case the dominating component did not necessarily have the higher dielectric constant. It was thus apparent that solvent structure rather than dielectric constant is the more important factor in electron solvation. The spectra observed at very short times ($\sim$ a few ps) after the

pulse were appreciably red shifted as compared to the nanosecond spectra, the shift being more in polar solvents. The observed time constant for the decay of the "red" species and the growth of the "blue" species correlated with the dielectric relaxation time corresponding to single molecular rotation[27]. All these observations, both in fluids and glasses strongly suggested the existence of a precursor of the solvated electron and the picture which has currently emerged can be summarised as follows[87]. A thermolized mobile electron in a medium is a delocalized excess electron which can be described by the wavefunction $\psi = i^{ikr}$ (r = distance from the origin, k, a constant and $i = \sqrt{-1}$), and whose probability density is constant in the medium. Depending on the microscopic structural properties of the medium there can exist preformed (i.e., even before the introduction of the thermalized electron) regions or traps at which the electron energy can be less than the energy corresponding to the bottom of the conduction band appropriate to the medium. The wavefunction of an excess electron at this site is given by $\psi = e^{-\lambda r}$ (where λ is a real constant and hence the probability density falls off with distance, or in other words the electron gets localized at this site. The localized electron can then induce structural changes at the site such as dipole orientation and thereby get solvated fully. In this picture, as the dipole orientation takes place around a localized and static electron, there is no conflict between electron dynamics and dipole relaxation, and the electron solvation time corresponds to the dipole relaxation time. In water, for example, hydrated electron can be considered to be fully in ~ 10 psec.

The various electron states proposed from time to time can now be rationalised in the following manner: "mobile" or "dry" electrons are delocalized electrons, "damp", "presolvated", "partially solvated" or "quasifree" electrons are localized at pre-existing traps in the medium and "solvated" electrons are really solvated in the sense in which this term is applied to ions in solution. In a frozen matrix both the localized and solvated electrons are trapped species (like any other species) because of very small diffusion coefficients which preclude their chemical reactions unless detrapped to the delocalized state by photo or thermal stimulation. As revealed from the work of Moorthy and Weiss in the frozen aqueous system[109] and Hamill in fluid aqueous solutions[58] the delocalized state starts off as a geminate pair with the parent ion or at later times with a secondary positive ion (formed in ion molecule reactions of the parent). As discussed by Moorthy[118]

such a geminate pair may be considered an exciton. In fluid hydrocarbons there are experimental evidences to indicate that the geminate pair may evolve into a solvated ion pair before diffusion out of the spur, whereas in a frozen glass, depending on the relative magnitudes of the time required for dipole relaxation and charge migration the localization and solvation processes may occur well after the geminate pairs have migrated out from the spurs. Recent experimental findings[105] on electron reactions in frozen media strongly point to a tunneling mechanism for these.

Role of excited molecules : We have earlier defined radiation chemistry as the chemistry of concurrently produced ionized and excited molecules. Singh[168] has reviewed in detail the role of excited molecules in radiolysis. Although in nonpolar organic (particularly aromatic) systems there is strong evidence for excited state formation based on the results of radiation induced luminescence[91], radiation induced geometrical isomerization of olefins[80,72,168] and direct observation by pulse radiolysis[76] such is not the case in polar media, particularly water. Even in the former the contribution of direct excitation (as in the optical case) is relatively small[100,169] and the possible routes involve ionization followed by electron (or negative charge) recapture.

$$MX \longrightarrow \begin{cases} MX^+ + e^- \\ MX^{+*} + e^- \end{cases}$$

$$\left.\begin{matrix} MX^+ + e^- \\ MX^{+*} + e^- \end{matrix}\right\} \rightarrow MX^*$$

$$e^- + S \rightarrow S^-$$

$$MX^+ + S^- \rightarrow MX^* + S$$

Obviously, such recapture processes (geminate or otherwise) will be more probable in nonpolar media and the availability of stable (i.e., nondissociative) excited levels as for example, in aromatic molecules would favour channelization of the energy of recombination into excited molecule formation. In water the first excited singlet state is dissociative and part of the H and OH radicals may originate in this manner as implied in the Samuel-Magee model. In condensed media any radicals originating from excited molecule dissociation must be "hot" in order to be able to survive recombination within the solvent cage. Although the existence of a triplet excited state of water has been experimentally demonstrated[20,92,97,107] there is no unambiguous evidence for its involvement in aqueous radiation

chemistry. Considering the fact that in radiation chemistry excited molecules may originate by both direct excitation and charge recombination a satisfactory and unambiguous assessment of the relative contributions of excitation and ionization processes in any given system either experimentally or theoretically is yet to be achieved.

VI. Epilogue

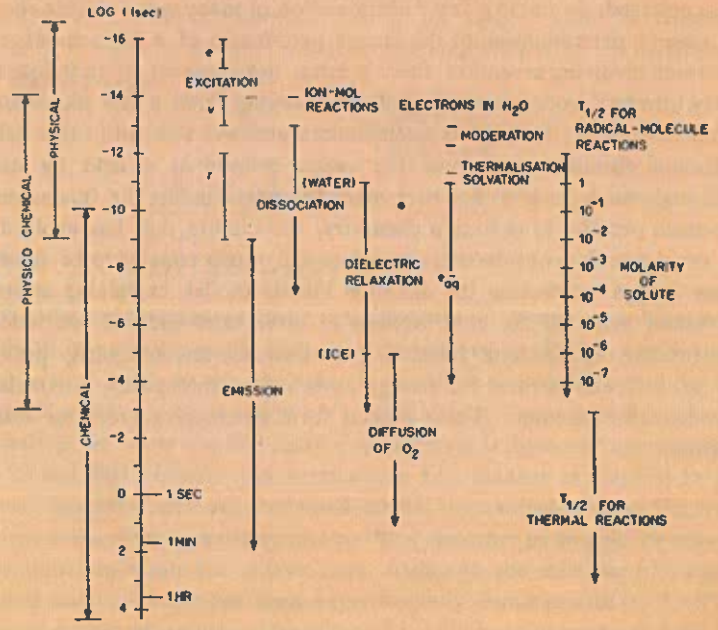
From what has been said so far it is clear that commencing from the interaction of radiation with the medium a large number of processes occur culminating in the formation of stable products. These are ionization (and to a lesser extent excitation) of matrix molecules, charge recombination, ion-molecule reactions, electron solvation, formation of free radicals, and diffusion and reactions of all these species among themselves, with the matrix and with solutes. Radiation chemical studies have been largely aimed at understanding these processes and in so doing many new contributions have emerged. To name a few : identification of many new unstable chemical species, determination of the kinetic parameters of a large number of reactions involving several of these species, development of techniques to study ultrafast processes (with half-lives ranging from a few picoseconds to microseconds). Despite its resemblances and overlaps with other fields, radiation chemistry has, over the years, evolved as a field by itself. Although much progress has been made in understanding the fundamental processes peculiar to radiation chemistry, the picture that has evolved so far cannot yet be considered as complete and much remains to be done in areas such as perfecting the diffusion kinetic model, explaining electron solvation, assessing the contributions of direct and indirect excitations, superexcited species and subexcitation electrons, and ultimately provide the proper balance sheet for energy absorbed in the system, particularly in a condensed medium. These aspects form challenging areas for future research.

APPENDIX I*

Dependence of $G(Fe^{+3})$ in the Fricke dosimeter on the LET of the radiation

Radiation	Initial LET (eV/Å)	$G(Fe^{+3})$
^{60}Co-rays (av. 1·25 MeV) } 1–2 MeV electrons }	0·02	15·5
250 KeV X-rays	0·10	14·3
10 KeV X-rays	0·20	13·5
^{3}H-β	~ 0·5	12·8
2 MeV protons	1·7	8·0
^{210}P$_0$-α (5·3 MeV)	8·8	5·1
^{10}B(n, α) ^{7}Li recoil nuclei	10–17	4·2
^{235}U fission fragments	700	3·0

APPENDIX II TIME SCALE OF EVENTS IN RADIATION CHEMISTRY



*Adapted from References 7 and 8.

APPENDIX III

Primary yields in water (adapted from Reference 8)

Medium	Radiation	Av. LET eV/Å	pH	G_{-H_2O}	G_{H_2}	$G_{H_2O_2}$	$G_{e^-_{aq}}$	G_H	G_{OH}	G_{HO_2}
$H_2O(v)$	X- or γ-rays, electrons	0·023	—	8·2	0·5	0		7·2	8·2	
$H_2O(l)$	γ-rays, e^- (0·1–20 MeV)	0·023	0·46	4·45	0·40	0·78	0	3·65	2·90	0·008
$H_2O(l)$	"	0·023	3–13	4·08	0·45	0·68	2·63	0·55	2·72	0·026
$D_2O(l)$	"	0·023	1·3	4·32	0·32	0·66		3·67	3·0	
$D_2O(l)$	"	0·023	4–13	4·12	0·36	0·66	2·96	0·43	2·84	
$H_2O(l)$	$^{3}H-\beta$	0·47	1	3·97	0·53	0·97	0	2·91	2·0	
$H_2O(l)$	32 MeV He^{2+}		7	3·01	0·96	1·00	0·72	0·42	0·91	0·05
$H_2O(l)$	12 MeV He^{2+}		7	2·84	1·11	1·08	0·42	0·27	0·54	0·07
$H_2O(l)$	^{210}P$_0$–α	4·3	0·46	3·62	1·57	1·45	0	0·60	0·50	0·11
$H_2O(l)$	^{10}B(n, α) ^{7}Li recoil nuclei		0·46	3·55	1·65	1·55	0	0·25	0·45	
$H_2O(l)$	(Extrapolated)	∞	0·46	2·9	1·45	1·45	0	0	0	

REFERENCES

1. Adams, G. E., Boag, J. W., Currant, J. and Micheal, B. D. (1965) *In* Pulse Radiolysis, (Eds. M. Ebert, J. P. Keene, A. J. Swallow and J. H. Baxendale), Academic Press : 117.
2. Alexander, P. and Lett, J. T. (1967) *In* Comprehensive Biochemistry, (Eds. M. Florkin and E. H. Stots), Elsevier, Amsterdam, **27** : 267.
3. Allan, J. T. and Scholes, G. (1960) *Nature*, **187** : 218.
4. Allen, A. O., Hochanadel, C. J., Ghormley, J. A. and Davis, T. W. (1952) *J. Phys. Chem.*, **56** : 575.
5. Allen, A. O. (1961) *In* Actions Chimiques et Biologiques des Radiations, (Ed. M. Haissinsky), Academic Press, London, V : 14.
6. Allen, A. O. (1961) *In* Actions Chemique et Biologiques des Radiations, (Ed. M. Haissinsky), Academic Press, London, V : 13.
7. Allen, A. O. (1961) *In* The Radiation Chemistry of Water and Aqueous Solutions, D. Van Nostrand Co. Inc., New York, 58.
8. Allen, A. O. (1961) *In* Actions Chemique et Biologiques des Radiations, (Ed. M. Haissinsky), Academic Press, London, V : 25.
9. Allen, A. O. and Hummel, A. (1966) *J. Chem. Phys.*, **44** : 3426.
   Allen, A. O. and Hummel, A. (1966) *J. Chem. Phys.*, **44** : 3431.
10. Anbar, M. and Neta, P. (1967) *Int. J. Applied Radiat. and Isotopes*, **18** : 493.
11. Anbar, M. (1968) *In* Fundamental Processes in Radiation Chemistry, (Ed. P. Ausloos), Wiley-Interscience, New York : 672.
12. Anbar, M., Bambach, P. and Ross, A. (1972) *NSRDS-NBS*, **43**, National Bureau of Standards, USA.
13. Avery, E. C., Remko, J. R. and Smaller, B. (1968) *J. Chem. Phys.*, **49** : 951.
14. Bacq, Z. M. and Alexander, P. (1961) Fundamentals of Radiation Biology, Pergamon Press, Oxford.
15. Barker, G. S., Fowles, P., Sammon, D. S. and Stringer, B. (1970) *Trans. Faraday Soc.*, **66** : 1498.
16. Barr, N. F. and Allen, A. O. (1959) *J. Phys. Chem.*, **63** : 928.
17. Baxendale, J. H., Filden, E. M. and Keene, J. P. (1963) *Proc. Chem. Soc.* : 242.
18. Bhattacharya, P. K. and Saini, R. D. (1979) *Radiat. Phys. Chem.*, **13** : 52.
19. Bielksi, B. H. J. and Gebiski, J. M. (1970) *In* Advances in Radiation Chemistry, (Eds. M. Burton and J. L. Magee), **2**, Wiley-Interscience, New York : 133.
20. Boag, J. W. and Hart, E. J. (1963) *Nature*, **197** : 45.
21. Brocklehurst, B. (1970) *Radiat. Res. Rev.*, **2** : 149.
22. Bronskill, M. J., Taylor, W. B., Wolff, R. K. and Hunt, J. W. (1970) *Rev. Sci. Instrum.*, **41** : 333.
23. Burton, M. (1947) *J. Phys. Colloid Chem.*, **51** : 611.
24. Comeron, A. T. and Ramsay, W. (1908) *J. Chem. Soc.*, **92** : 966.
    Comeron, A. T. and Ramsay, W. (1908) *Ibid.*, **92** : 992.
25. Cerecik, B. and Kongshaug, M. (1969) *J. Phys. Chem.*, **73** : 2056.

26. Chapiro, A. (1962) Radiation Chemistry of Polymeric Systems, Interscience, New York.
27. Chase, W. J. and Hunt, J. W. (1975) *J. Phys. Chem.*, **79** : 2835.
28. Chouraquie, M. and Sutton, J. (1966) *Trans. Farad. Soc.*, **62** : 2111.
29. Compton, R. N., Huebuer, R. H., Reinhardt, P. W. and Christophoru, L. G. (1968) *J. Chem. Phys.*, **48** : 901.
30. Chundall, R. B. and Tippett, W. (1968) *Adv. Chem. Ser.*, **82** : 387.
31. Czapski, G. and Stein, G. (1959) *J. Phys. Chem.*, **63** : 850.
32. Czapski, G. and Schwarz, H. A. (1962) *J. Phys. Chem.*, **66** : 471.
33. Dainton, F. S. (1948) *J. Phys. Chem.*, **52** : 490.
34. Dainton, F. S. and Gopinathan, C. (1969) *Trans. Farad. Soc.*, **65** : 151.
35. Daniels, M. (1969) *J. Phys. Chem.*, **73** : 371.
36. Davydov, A. S. (1948) *Zhur. Exptl. Teoret Fiz.*, **18** : 913.
37. Deigen, M. F. (1954) *Ibid.*, **26** : 300.
38. Dewhurst, H. A. (1952) *Trans. Farad. Soc.*, **48** : 905.
39. Dibierne, A. (1909) *C.R. Acad. Sci. Paris*, **148** : 703.
40. Dibierne, A. (1974) *Ann. Phys. Paris*, **2** : 97.
41. Dorfman, L. M. and Adams, G. (1972) *NSRDS-NBS*, **45**, National Bureau of Standards, USA.
42. Dorfman, L. M. and Gavlas, J. F. (1974) *In* Radiation Research. Proc. V Intern. Cong. Rad. Res. Seattle, USA, (Eds. O.F. Nigard, H. I. Adler and W. K. Sinclair): 326.
43. Draganic, I. G. and Draganic, Z. D. (1971) The Radiation Chemistry of Water, Academic Press, New York.
44. Draganic, Z. D. and Draganic, I. G. (1969) *J. Phys. Chem.*, **73** : 2571.
45. Dyne, P. J. and Kennedy, J. M. (1964) *Can. J. Chem.*, **36** : 1518.
    Dyne, P. J. and Kennedy, J. M. (1964) *Ibid.*, **38** : 61.
46. Fischer, M. M. and Hamill, W. H. (1973) *J. Phys. Chem.*, **77** : 171.
47. Flanders, D. A. and Frocke, H. (1958) *Ibid.*, **28** : 1126.
48. Frank, J. and Rabinowitch, E. (1933) *Trans. Faraday Soc.*, **29** : 120.
49. Fricke, H. and Morse, S. (1929) *Phil. Mag.*, **7** : 129.
50. Fricke, H. (1955) *Ann. Acad. Sci. N.Y.*, **59** : 567.
51. Ganguly, A. K. and Magee, J. L. (1956) *J. Chem. Phys.*, **25** : 129.
52. Giesel, F. (1902) *Ber. Deut. Chem. Ges.*, **35** : 3608.
    Giesel, F. (1903) *Ibid.*, **36** : 342.
53. Gillis, H. A., Williams, R. R. and Hamill, W. H. (1961) *J. Am. Chem. Soc.*, **83** : 17.
54. Girija, G. and Gopinathan, C., *Radiat. Phys. Chem. (In press)*
55. Gordon, S., Hart, E. J., Matheson, M. S., Rabani, J. and Thomas, J. K. (1963) *J. Amer. Chem. Soc.*, **85** : 1375.
56. Gray, L. H. (1951) *J. Chem. Phys.*, **48** : 172.
57. Haissinsky, M. (1968) *In* Radiation Chemistry of Aqueous Systems, (Ed. G. Stein), The Weizmann Science Press of Israel, Jerusalem, : 1.
58. Hamill, W. H. (1969) *J. Phys. Chem.*, **73** : 1341.

59. Hamill, W. H. Quoted by Anbar in Ref. 11.
60. Hammes, G. (Ed.) (1974) Investigation of Rates and Mechanisms of Reactions, Pt. 2, Wiley-Interscience, New York.
61. Hart, E. J. and Boag, J. W. (1962) *J. Amer. Chem. Soc.*, **84** : 4090.
62. Hart, E. J., Gordon, S. and Fielden, E. M. (1968) *J. Phys. Chem.*, **72** : 577.
63. Hart, E. J. and Anbar, M. (1970) The Hydrated Electron, Wiley-Interscience, New York.
64. Hase, H. (1979) Radiation Research, (1979). Proc. VI Intern. Cong. Rad. Res., Tokyo, Japan, (S. Okada, M. Imamura, T. Terashima and H. Yamaguchi) : 309.
65. Hayon, E. and Weiss, J. (1958) Proc. Int. Conf. Peaceful uses of Atomic Energy, 2nd Geneva, 1958, **29** : 1517.
66. Hayon, E. and Allen, A. O. (1961) *J. Phys. Chem.*, **65** : 2181.
67. Hayon, E. (1962) *Nature*, **194** : 737.
68. Henglien, A. (1976) *In* Pulse, Radiolysis and Polarography, in Electroanalytical Chemistry, (Ed. A. J. Bard), **9**, Marcel Dekker : 163.
69. Henry, V. P., Maxwell, C. R., White, W. C. and Peterson, D. C. (1952) *J. Phys. Chem.*, **56** : 153.
70. Hentz, R. R., Farhataziz, D. M. and Burton, M. (1967) *J. Chem. Phys.*, **46** : 2995.
71. Hentz, R. R. and Johnson, C. G. (1969) *Ibid.*, **51** : 1236.
72. Hentz, R. R. and Lehman, H. P. (1969) *J. Phys. Chem.*, **73** : 4283.
73. Hochanadel, C. J. (1952) *Ibid.*, **56** : 587.
74. Hummel, A. and Allen, A. O. (1962) *Radiat. Res.*, **17** : 302.
75. Hunt, J. W. and Thomas, J. K. (1967) *Ibid.*, **32** : 149.
76. Hunt, J. W. and Thomas, J. K. (1967) *J. Chem. Phys.*, **46** : 2954.
77. Hunt, J. W., Lam, K. Y. and Chase, W. J. (1974) *In* Radiation Research 1974, Proc. V Intern. Cong. Rad. Res., Seattle, USA (Eds. O. S. Nygaard, H. I. Adler and W. K. Sinclair) : 345.
78. Jonah, C. D., Matheson, M. S., Miller, R. J. and Hart, E. J. (1976) *J. Phys. Chem.*, **80** : 1267.
79. Jonah, C. D. and Miller, J. R. (1977) *Ibid.*, **81** : 1974.
80. Jortner, J. (1957) *J. Chem. Phys.*, **27** : 823.
Jortner, J. (1959) *Ibid.*, **30** : 839.
81. Keene, J. P. (1960) *Nature*, **188** : 843.
82. Keene, J. P. (1963) *Ibid.*, **197** : 47.
83. Kernbaum, M. (1909) *C.R. Acad. Sci., Paris*, **149** : 116.
84. Kevan, L., Moorthy, P. N. and Weiss, J. J. (1963) *Nature*, **199** : 689.
Kevan, L., Moorthy, P. N. and Weiss, J. J. (1964) *J. Amer. Chem. Soc.*, **86** : 771.
85. Kevan, L. (1974) *In* Advances in Radiation Chemistry (Eds. M. Burton and J. L. Magee), Wiley-Interscience, New York, **4** : 181.
86. Kevan, L., Browman, M. K., Narayana, P. A., Boekmann, R. R., Yudanov, V. F. and Tsvetkov, Y. D. (1975) *J. Chem. Phys.*, **63** : 409.
87. Kevan, L. (1978) *J. Phys. Chem.*, **82** : 1144.

88. Kishore, K., Moorthy, P. N. and Rao, K. N. (1976) *Radiat. Effects*, **27** : 167.
Kishore, K., Moorthy, P. N. and Rao, K. N. (1976) *Ibid.*, **29** : 165.
Kishore, K., Moorthy, P. N. and Rao, K. N. (1978) *Ibid.*, **38** : 97.
89. Kishore, K., Moorthy, P. N. and Rao, K. N. *Radiat. Phys. Chem.* (*In press*).
90. Kuppermann, A. and Belford, G. G. (1962) *J. Chem. Phys.*, **36** : 1427.
91. Kuppermann, A. (1967) *In* Radiation Research 1966, Proc. III Intern. Cong. Rad. Res., Evian, France, 1966 (Ed. G. Silini), North Holland, Amsterdam : 212.
92. Kuppermann, A. (1970) Proceedings of International Symposium on Radiation Chemistry and Biology, 161, USAEC.
93. Kuppermann, A. (1974) *In* Physical Mechanisms in Radiation Biology (Eds. R. D. Cooper and R. W. Wood), USAEC 155.
94. Lal, M. (1970) *Radiat. Effects*, **2** : 225.
Lal, M. (1974) *Ibid.*, **22** : 49.
Lal, M. (1974) *Ibid.*, **22** : 237.
95. Lal, M. (1976) *Can. J. Chem.*, **54** : 1092.
96. Lea, D. E. (1946) *In* Actions of Radiations on Living Cells, Cambridge Univ. Press, Cambridge.
97. Lewis, D. and Hamill, W. H. (1969) *J. Chem. Phys.*, **51** : 456.
98. Lin, W. S., Gaucher, G. M., Armstrong, D. A. and Lal, M. (1976) *Can. J. Chem.*, **54** : 242.
99. Livingston, R., Zeldes, H. and Taylor, E. H. (1955) *Disc. Faraday Soc.*, **19** : 166.
100. Makarov, V. I. and Binderskii, V. A. (1970) *High Energy Chem.*, **4** : 1.
101. Mao, S. W. and Kevan, L. (1974) *Che. Phys. Lett.*, **24** : 505.
102. Matheson, M. S. and Dorfman, L. M. (1960) *J. Chem. Phys.*, **32** : 1870.
103. Matheson, M. S. and Dorfman, L. M. (1969) *In* Pulse Radiolysis, The MIT Press, Cambridge, 68.
104. McCarthy, R. L. and Maclachlan, A. (1960) *Trans. Faraday Soc.*, **56** : 1187.
105. Miller, J. R. and Butz, J. V. (1979) *In* Radiation Research, 1979, Proc. VI Intern. Cong. Rad. Res., Tokyo, Japan, (Eds. S. Okada, M. Imamura, T. Terashima and H. Yamaguchi) : 301.
106. Mittal, J. P. and Hamill, W. H. (1967) *J. Amer. Chem. Soc.*, **89** : 5749.
107. Mittal, J. P. (1977) *Nature, P. S. Letters*, **230** : 160.
108. Mittal, J. P. (1971) *Radiat. Effects*, **7** : 279.
109. Moorthy, P. N. and Weiss, J. J. (1965) *Adv. Chem. Sers.*, **50** : 180.
Moorthy, P. N. and Weiss, J. J. (1964) *Nature*, **204** : 776.
Moorthy, P. N. and Weiss, J. J. (1965) *J. Chem. Phys.*, **42** : 3127.
110. Moorthy, P. N. and Weiss, J. J. (1964) *Phil. Mag.*, **10** : 659.
111. Moorthy, P. N. and Weiss, J. J. (1964) *Nature*, **201** : 1317.
112. Moorthy, P. N. and Shankar, J. (1969) *Radiat. Effects*, **2** : 91.
113. Moorthy, P. N. (1969) DAE Chemistry Symposium, Chandigarh, India, Proceedings, II : 186.
114. Moorthy, P. N., Gopinathan, C. and Rao, K. N. (1970) *Radiat. Effects*, **2** : 175.
115. Moorthy, P. N. and Rao, K. N. (1970) DAE Chemistry Symposium, Madras, India, Proceedings, III : 23.

116. Moorthy, P. N., Rao, K. N. and Shankar, J. (1971) *Nature (P.S.)*, **234** : 17.
117. Moorthy, P. N. and Hayon, E. (1974) *J. Phys. Chem.*, **78** : 2615.
Moorthy, P. N. and Hayon, E. (1975) *Ibid.*, **79** : 1059.
Moorthy, P. N. and Hayon, E. (1975) *J. Amer. Chem. Soc.*, **97** : 2048.
Moorthy, P. N. and Hayon, E. (1975) *Ibid.*, **97** : 3345.
Moorthy, P. N. and Hayon, E. (1976) *J. Org. Chem.*, **41** : 1067.
Moorthy, P. N. and Hayon, E. (1977) *Ibid.*, **42** : 879.
118. Moorthy, P. N., Rao, K. N., Mishra, K. P. and Singh, B. B. (1974) *Radiat. Effects*, **22** : 243.
119. Mozumder, A. and Magee, J. L. (1975) Physical Chemistry, An Advanced Treatise, VII, Reactions in Condensed Phases, (Ed. H. Eyring), Academic Press, New York : 699.
120. Mukherjee, T. and Mittal, J. P. (*Personal Communication*).
121. Myron, J. J. J. and Freeman, G. R. (1965) *Can. J. Chem.*, **43** : 381.
122. Narayana Rao, K. and Ramanan, G. (1968) *Ind. J. Chem.*, **6** : 444.
123. Natoari, M. and Watanabe, T. (1966) *J. Phys. Soc. (Japan)*, **21** : 1573.
124. O'Donnell, J. H. and Sangster, D. F. (1970) Principles of Radiation Chemistry, Edward Ernold, London.
125. Pekar, S. I. (1946) *J. Phys. USSR*, **10** : 341.
126. Peled, E. and Czapski, G. (1970) *J. Phys. Chem.*, **74** : 2903.
127. Piett, L. H., Rempel, R. L., Weaver, M. W. and Siegel, S. (1959) *J. Chem. Phys.*, **30** : 1623.
128. Pikaev, A. K., Kabakechi, S. A. and Zansokhova, A. A. (1977) Faraday Discussions of the Chemical Society, **63** : 112.
129. Platzman, R. L. (1953) *In* Basic Mechanisms in Radiobiology, National Res. Council Publ. : 305, Washington D.C.
130. Platzman, R. L. (1962) *Radiat. Res.*, **17** : 419.
131. Platzman, R. L. (1962) Proc. II Intern. Cong. Rad. Res., Harrogate : 128.
132. Platzman, R. L. (1967) Radiation Research (1966) (Ed. G. Silini), Proc. III Intern. Cong. Radiat. Res.), North Holland Publ., Amsterdam.
133. Rabani, J. and Matheson, M. S. (1964) *J. Amr. Chem. Soc.*, **86** : 3175.
134. Rabani, J. and Matheson, M. S. (1965) *In* Pulse Radiolysis by Dotfman, L. M. and Matheson, M. S. in Progress in Reaction Kinetics, (Ed. P. Porter), Pergamon Press, Oxford : 237.
135. Rabani, J., Mulac, W. A. and Matheson, M. S. (1965) *J. Phys. Chem.*, **69** : 53.
136. Rabani, J. and Nielsen, S. O. (1969) *J. Phys. Chem.*, **73** : 3736.
137. Rabani, J., Gratzel, M., Chaudhuri, S. A., Beck, G. and Hanglein, A. (1971) *J. Phys. Chem.*, **75** : 1759.
138. Rama Rao, K. V. S. and Prasad, D. (1979) Discussion meeting on Kinetics of Physico-Chemical Oscillations, Deutsche Bunseugeselleschaft fur Physikalische Chemie, Aachen.
139. Rao, K. N., Gopinathan, C., Moorthy, P. N., Rao, M. H. and Kumar, V. (1967) *BARC*-331.

140. Rao, K. N., Moorthy, P. N., Rao, M. H., Kumar, V. and Gopinathan, C. (1968) *BARC*-369.
141. Rao, M. H. and Rao, K. N. (1970) *J. Polymer. Sci.*, **B8** : 643.
Rao, M. H. and Rao, K. N. (1971) *Radiat. Effects*, **10** : 87.
142. Rao, M. H. and Rao, K. N. *Radiat. Effects (In press)*.
143. Rao, K. N., Rao, M. H., Moorthy, P. N. and Charlesby, A. (1972) *J. Polymer*, Science Letters Edition, **10** : 893.
144. Rao, P. S. and Hayon, E. (1974) *J. Amer. Chem. Soc.*, **96** : 1287.
Rao, P. S. and Hayon, E. (1975) *Ibid.*, **97** : 2986.
145. Rao, K. N. (1979) DAE Symposium on Industrial Polymers and Radiation, Anand, India, Proceedings.
146. Rao, M. H. and Rao, K. N. (1979) *Polymer Bulletin*, **1** : 727.
147. Rao, K. N., Rao, M. H., Lokhande, H. T., Mody, N. R. and Jog, A. G. (1979) *J. Appl. Polymer. Science*, **23** : 2133.
148. Rigg, T., Stein, G. and Weiss, J. (1952) *Proc. Roy. Soc.*, A **211** : 375.
149. Risse, O. (1930) *Ergeben-Physiologie*, **30** : 242.
150. Rothschild, W. G. and Allen, A. O. (1958) *Radiat. Res.*, **8** : 101.
151. Samuel, A. H. and Magee, J. L. (1953) *J. Chem. Phys.*, **21** : 1080.
152. Sangster, M. (1954) *Disc. Faraday Soc.*, **17** : 112.
153. Sargent, F. P., Gardy, E. M. and Fall, H. R. (1974) *Chem. Phys. Lett.*, **24** : 120.
154. Sargent, F. P. and Gardy, E. M. (1974) *Can. J. Chem.*, **52** : 3645.
Sargent, F. P. and Gardy, E. M. (1976) *Ibid.*, **54** : 275.
Sargent, F. P. and Gardy, E. M. (1976) *J. Phys. Chem.*, **80** : 854.
155. Santar, I. and Bendar, J. (1967) *Collect. Czech. Chem. Commun.*, **32** : 953.
Santar, I. and Bendar, J. (1968) *Ibid*, **33** : 1.
Santar, I. and Bendar, J. (1969) *Ibid.*, **34** : 1.
156. Sarrah, D. (1965) *Kernenezgie*, **8** : 220.
157. Schmidt, K. H. and Ander, S. M. (1969) *J. Phys. Chem.*, **73** : 2846.
158. Schmidt, K. H., Gordon, S. and Mulac, W. A. (1976) *Rev. Sci. Instrum.*, **47** : 356.
159. Scholes, G. (1963) *In* Progress in Biophysics, **13**, Pergamon Press, New York : 59.
160. Schwarz, H. A. (1969) *J. Phys. Chem.*, **73** : 1928.
161. Shankar, J., Rama Rao, K. V. S. and Shastri, L. V. (1969) *J. Phys. Chem.*, **73** : 52.
162. Shankar, J. (1971) *J. Ind. Chem. Soc.*, **48** : 97.
163. Singh, A. (1972) *Radiat. Res. Rev.*, **4** : 1.
164. Singh, A., Chase, W. J. and Hunt, J. W. (1977) *Faraday Disc. Chem. Soc.*, **63** : 28.
165. Spinks, J. W. T. and Woods, R. J. (1976) An Introduction to Radiation Chemistry, Wiley, New York.
166. Spinks, J. W. T. and Woods, R. J. (1976) An Introduction to Radiation Chemistry, Wiley, New York.
167. Spinks, J. W. T. and Woods, R. J. (1976) *In* An Introduction to Radiation Chemistry, Wiley, New York.
168. Van Ingen, J. W. F. and Cramer, W. A. (1970) *J. Phys. Chem.*, **74** : 1134.
169. Voltz, R. (1968) *Radiat. Res. Rev.*, **1** : 301.

170. Ward, J. A. and Hamill, W. H. (1967) *J. Amer. Chem. Soc.*, **89** : 5116.
171. Wargon, J. A. and Williams, F. (1972) *Ibid.*, **94** : 9717.
172. Weeks, J. L. and Rabani, J. (1966) *J. Phys. Chem.*, **70** : 2100.
173. Weiss, J. (1944) *Nature*, **153** : 748.
174. Weiss, J. (1964) *Radiat. Res. Suppl.*, **4** : 141.
175. Weiss, J. J. (1960) *Nature* (*London*), **186** : 751.
Weiss, J. J. (1963) *Ibid*, **199** : 589.
Weiss, J. J. (1964) *Rad. Res. Suppl.*, **4** : 141.

Biotechnological Applications of Plant Tissue and Cell Cultures : Problems and Prospects

M. S. Chadha and M. R. Heble

Bio-Organic Division, Bhabha Atomic Research Centre
Trombay, Bombay 400 085

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

I. Introduction

Progressive depletion of the natural resources as against rising global population has caused concern towards our future requirements of food, fibre and energy. The availability of forest products, edible and non-edible oils and raw materials for pharmaceuticals is also under continuous stress. Consequently, the biotechnological aspects of plant sciences have gained considerable importance. Plant tissue culture technology has developed into a potential area of plant sciences for obtaining improved plants of economic importance[64]. It has also provided potential for cultivation of plant cells in culture with the ultimate hope of getting plant products of pharmaceutical importance[87].

The possibilities of utilizing plant tissue and cell cultures for medicinal compounds was visualised in the early 40's[12]. However, the efforts made during the initial stages were empirical in approach and attention was essentially focussed towards establishing a rapidly growing tissue culture system and ascertaining its synthetic ability for desired constituents. The developments in this area of research during the past two decades have shown that the biosynthesis of plant products is under the control of several regulatory mechanisms. It is expected that with a clear understanding of these aspects, it should be possible to exploit the tissue culture systems to advantage like the microbial systems. While the microbial systems have been successfully used for the production of antibiotics and vitamins, using the plant tissue culture systems, it may be possible to produce a wide spectrum of physiologically active compounds like alkaloids, steroids, phenolics, vitamins and antibiotics[79]. Furthermore, cell cultures could also be used for the purposes of specific biotransformations of organic molecules[22].

It is now possible to establish tissue cultures from any part of a plant either as static cultures or as liquid suspensions and such cultures have been maintained for several years[88]. Significant efforts have also been made in growing cultured tissues on large-scale using batch culture, chemostat and fermentors[84]. Mass cultivation of plant cells, therefore, may not pose serious problems, although the growth conditions may have to be ascertained for each tissue.

The concept of morphogenetic and biosynthetic totipotency has been demonstrated in several plant tissue culture systems[41,18,46] and it is generally accepted that plant cells in culture possess all the genetic

information of the parent plant. However, during the initiation of tissue cultures from an organised explant, several biochemical, physiological and cytological changes take place. On subculture, these changes may persist or lead to further changes and as a result, the behaviour of tissue culture systems is often unpredictable. Since these events may have strong influence on the biosynthetic potential of the tissues, it is essential to understand why such changes take place and how these changes could be controlled or regulated.

In spite of impressive advances in the area of synthetic organic chemistry, our dependence on plant products continues. A list of a few commonly used important plant products in the pharmaceutical industry is given in Table 1. The prospects of obtaining drugs for cancer chemotherapy are also actively explored and a number of plant products are under trial[55]. Successful cultivation of plant cells and organs for indefinite period under controlled conditions offers unlimited scope for the production of pharmaceuticals and other products through tissue cultures.

Plant tissue, organ and cell cultures have been established from a large number of medicinal plants. Some of these cultures synthesize secondary products in significant yields. Of particular interest are *Dioscorea deltoidea* callus cultures (diosgenin)[80], *Dioscorea composita* shoot culture (diosgenin)[35], cell culture (diosgenin)[58], *Panax ginseng* callus cultures (ginsengoside)[23], *Catharanthus roseus* cell cultures (ajmalicine and serpentine)[97], *Atropa belladonna* root cultures (atropine)[61], *Nicotiana tabacum* callus cultures (nicotine)[54], *Nicotiana tabacum* cell culture (ubiquinone)[49] and *Mucuna sloanei* callus cultures (L-dopa)[71]. Several interesting reactions are also performed by plant cell cultures of which some are common to plants and microbes and others specific to plant cells. The biotransformation potentiality of plant cells is yet another interesting aspect which could be exploited by industry. The major reactions encountered are stereospecific reductions, hydroxylations, oxidation, decarboxylation and glycosylation[22].

Although the potentialities of plant cells and tissue cultures have been demonstrated in several important medicinal plants, yet they have not been commercially exploited like in the microbial systems. This is largely due to the fact that we have far more to understand about the complexities of cellular metabolism particularly those related to the regulatory mechanisms of secondary products biosynthesis. These are problems of fundamental interest which have a direct bearing on the successful exploitation of plant tissue

cultures. In this review an attempt has been made to outline the present status of work concerning the biosynthesis and biotransformation of secondary products in plant tissue and cell cultures and to discuss the problems and prospects involved in translating this knowledge into a viable technology.

II. The Role of Differentiation and Development on Product Biosynthesis

The events leading to the synthesis of secondary products in higher plants are very diverse and it is rather difficult to generalise these. However, certain common features can be identified in regard to the formation of products such as alkaloids, steroids, terpenes and essential oils. The alkaloids have received maximum attention because of their importance in pharmaceutics and considerable literature is available with respect to the factors controlling alkaloid production in tissue cultures and in whole plants. The concentration of alkaloids is generally high in growing fruits, vascular tissues, epidermis and cells just behind shoot and root, laticiferrous cells and latex vessels. These observation suggest that alkaloid production may be closely associated with the process of differentiation. In this context, tissue cultures derived from certain Solanaceous plants have been studied extensively.

The formation of nicotine in tissue and organ cultures of *Nicotiana* have been reported. The general levels of nicotine produced by cells and tissue cultures have been reported to range from 0·1 to 1·0 μg/mg dry weight as compared to 29 μg/mg dry weight from the roots of the intact plant[54]. Dawson reported that nicotine is produced in the roots from where it is translocated to the leaves. Furuya *et al.*[25], have demonstrated that nicotine synthesis in tobacco could be regulated by the exogenous hormonal application without recourse to organogenesis. Peters *et al.*[70], have reported that nicotine can induce rooting in tissue and cell cultures of *Nicotiana rustica*. Tabata *et al.*[89], have demonstrated a close relationship between organization and nicotine production in tissue cultures of *N. tabacum*. By obtaining single cell clones from *N. tabacum* cell cultures, they have isolated cell strains, which contained as high as 3·4% dry weight of nicotine. The original callus cultures contained less than 0·05% nicotine[67]. These results indicate that nicotine synthesis can take place in the absence of organogenesis and

strongly suggest that more than one factor regulate the biosynthesis of nicotine in cultured tissues.

The synthesis of tropane alkaloids in tissue cultures have been the subject of considerable exploration. Bhandari *et al.*[5] observed that in *Atropa belladonna*, the alkaloid synthesis is dependent upon the emergence of root organization. The total alkaloidal content of the root cultures was 0·048%. Similar observations have been made in tissue cultures of *Scopolea parviflora*[90]. In these cultures, the authors could observe an increase in the alkaloid content when the regenerated roots were allowed to grow further into thickened roots in liquid medium. These observations indicated that the alkaloid synthesis is influenced not only by the process of root differentiation but also further development of the roots. However, the total alkaloid content of these cultures represented only 10% of the alkaloid content in the original roots or rhyzomes of the parent plant. Mitra[61] established continuous cultures of *A. belladonna* roots which yielded as high as 0·41 to 0·45% of atropine. In tissue cultures of *Hyoscyamus niger*, the initiation of root organogenesis caused only a marginal increase from 0·05 to 0·075% in the alkaloid content[15]. These observations suggest that besides rooting there could be other important factors which control the synthesis and accumulation of tropane alkaloids.

Further studies of these aspects have been carried out in our laboratory with tissue cultures of *A. belladonna*.[18] The roots of this plant are supposed to be the active site of alkaloid synthesis. Tissue cultures established from the leaf segments of haploid and diploid plants of *A. belladonna* exhibited organogenetic potential during the first eight serial subcultures (Fig. I). Subsequently, the diploid cultures lost their organ-forming capacity but the haploid cultures continued to differentiate shoot buds. The alkaloid profile of the unorganised cultures and differentiating tissues was similar. The total alkaloid content of these tissues was very low as compared to the parent plant (Table 2). The shoot buds isolated from the differentiated cultures also contained low levels of alkaloids like the unorganised cultures. However, the relative proportion of atropine was more in the shoot buds than in the callus tissues. On the basis of these results, it was concluded that neither the origin of the cultures nor their ability to form unorganised structures influenced the synthesis of belladonna alkaloids. Nevertheless, the process of shoot bud formation does affect the alkaloid profile as they show a tendency towards acquiring the metabolic status of the whole plant (Fig. II).

The synthesis of alkaloids in the callus and the regenerated shoot buds (devoid of roots) showed that alkaloid synthesis is not exclusively associated with the initiation and development of roots. The observations further supported the hypothesis that several factors govern the synthesis and accumulation of tropane alkaloids.

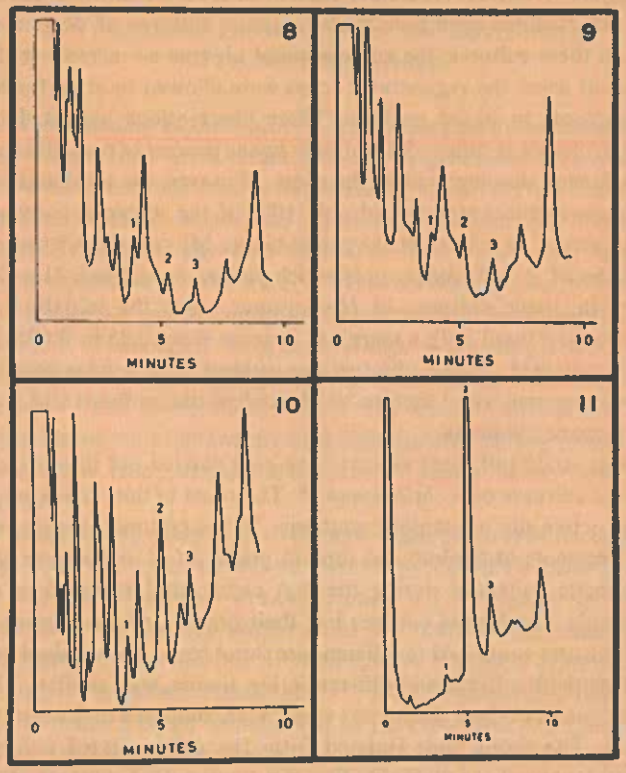


Fig. II, 8-11. Gas-liquid chromatography on 3% SE-30 of the tropane alkaloids isolated from ; 8 : diploid callus cultures grown on W and B medium ; 9 : haploid callus cultures grown on MS medium ; 10 : 30-day-old regenerated shoot buds from haploid callus grown on MS medium ; 11 : regenerated plant at flowering stage. Peak 1, tropine ; peak 2, atropine ; peak 3, scopolamine. From Eapen *et al.*, 1978.

There are reports of certain other alkaloids whose synthesis is associated with the development of plant. Ikuta *et al.*[50] examined callus cultures and

redifferentiated plantlets of certain Papaveraceae species. All the tissue cultures contained a similar profile of alkaloids, comprising benzophenanthridine, protopine and aporphine-type alkaloids. The shoot buds regenerated from *P. bracteatum* and *P. somniferum* yielded similar alkaloid composition as the callus cultures. The regenerated plantlets, on the other hand, showed an alkaloidal pattern that resembled closely that of the original plant.

Our studies with tissue cultures of *Tylophora indica* (a plant containing phenanthroindolizidine alkaloids) have shown that the alkaloid synthesis is associated with the development of the regenerated plant to maturity[4]. In this plant the alkaloids are distributed more or less uniformly in all parts (Table 3) indicating that specialized cells may not be involved in the alkaloid synthesis. Tissue cultures initiated from different parts of the plant contained only trace amounts of alkaloids, their pattern being similar. Exogenous treatment with either growth hormones, inorganic nutrients or alkaloid precursors had profound influence on growth and organogenesis but failed to bring about any changes in the alkaloid synthesis of the cultured tissues (Fig. III). While the alkaloid profile of the regenerated shoot buds and roots resembled that of the callus tissues, the regenerated plantlets at (10 leaf stage) showed a resemblance to that of the mature plant in alkaloid composition. The regenerated mature plants and the normal plants contained similar alkaloid profile and alkaloid concentration. The results clearly demonstrate that in tissue cultures of *T. indica*, the alkaloid synthesis is not influenced by the exogenous chemical factors or by the processes of root or shoot differentiation, even-though the callus cells retained the genetic information necessary for alkaloid biosynthesis. However, recapitulation of alkaloid synthesis by the regenerated plantlet and the mature plant indicate that the alkaloid synthesis is dependent upon the growth and development of the plant. The processes of growth and development of the plant are related to the degree of cellular differentiation and the presence of highly differentiated and quiescent cells. The degree of cellular differentiation, therefore, appears to be an important factor for alkaloid biosynthesis. In this context, studies with *in vitro* cultures of isolated organs, having the biosynthetic capacity of the whole plant, might prove more rewarding. Such organ cultures have been successfully maintained[55].

There are a number of instances where the alkaloid synthesis occurs

without the processes of organ differentiation. Interesting examples of the formation of potential antitumor alkaloids in cell cultures are *Catharanthus roseus*[95], *Cephalotaxus horringtonia*[18] and *Camptotheca acuminata*.

The relation between steroid formation and organogenesis has also been studied in certain tissue cultures. Our studies with tissue cultures of *Solanum xanthocarpum* showed that the active synthesis of diosgenin occurs in the unorganised cultures[87]. While the formation of steroids is considerably influenced by the growth hormones used in the medium, shoot and root organogenesis did not cause appreciable change in the steroid synthesis[88]. Hirotani and Furuya[45] reported the absence of cardiac glycosides in callus cultures and regenerated shoot buds of *Digitalis purpurea*. However, the synthesis of the cardenolides was restored in the regenerated shoots.

Tissue cultures of *Bupleurum falcatum* is an excellent example where product biosynthesis is directly related to the root organogenesis[98]. The roots of the plant, known to be antipyretic and antiinflammatory, contain saikosaponins a, b, c and d. The callus cultures of the plant failed to synthesize the saikosaponins. When roots were initiated from the tissues in suspension, the synthesis of saikosaponins, in concentrations as high as in the roots of the original plant, was observed. Our studies with tissue cultures of *Hemidesmus indicus* showed that only stem callus exhibited the potential to form organised structures (Fig. IV)[91]. The synthesis of 16-dehydropregnenolone was observed only in these cultures.

It is, therefore, evident that differentiation in the form of organogenesis is not the only important parameter which figures in relation to secondary metabolism. More important is the process of cellular differentiation that occurs during the course of growth and development of the tissues and the plant. There is strong evidence to indicate the importance of cyto-differentiation on product biosynthesis which would be discussed in the following sections.

III. Regulation of Product Biosynthesis by Growth Hormones and Precursors

The important prerequisites for the biosynthesis of secondary products by plant cells in culture are : (1) The genetic information needed for the biosynthetic events, (2) The expression of the information through appropriate biochemical steps and (3) an abundant supply of the primary

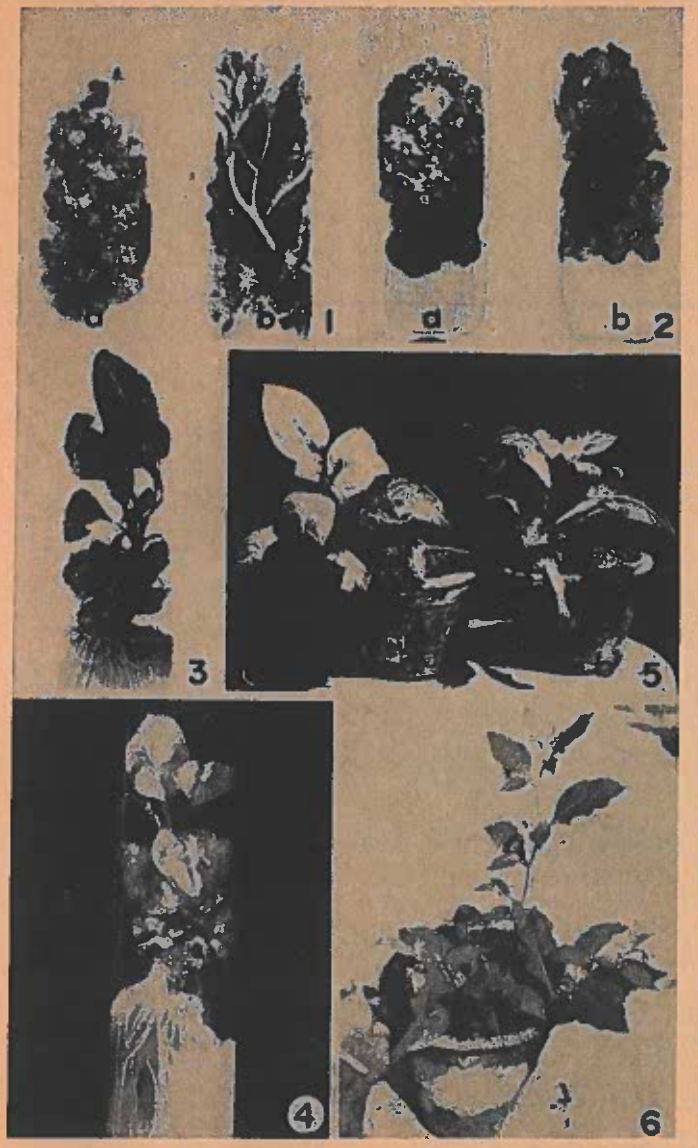


Fig. I—1-6. (1) Four-week-old haploid callus cultures on W & B+NAA (2mg/l)+Kn (0·5 mg/l) : (a) Callus showing unorganized growth ; (b) Callus with well-differentiated shoots. (2) Four-week-old diploid callus cultures : (a) Unorganized callus growth on W & B+NAA (2 mg/l)+Kn (0·5 mg/l) ; (b) Shoot bud differentiation on MS+BA (1 mg/l). (3) and (4) Four-week-old plantlets from haploid (3) and diploid (4) calli with well-developed roots on modified White's medium+NAA (1 mg/l). (5) Regenerated plantlets from haploid (a) and diploid (b) calli, about 6 weeks after transfer to soil. (6) *In vitro* obtained regenerants from haploid callus bearing flowers (arrow-marked) about 8 weeks after transfer to field

Abbreviations : BA—6-benzyladenine ; 2, 4-D—2, 4-dichlor ophenoxyacetic acid ; IAA—indoleacetic acid ; Kn—kinetin ; NAA—naphthaleneacetic acid

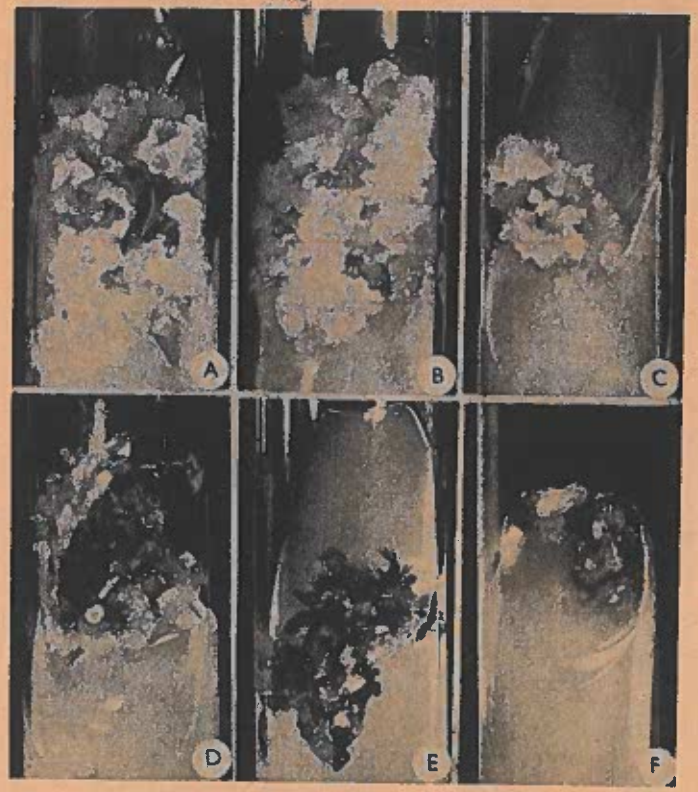


Fig. III—A-F. Effect of auxins on different organs of T. indica. 2, 4-D : A, Leaf ; B, Stem ; C, Root. NAA : D, Leaf ; E, Stem ; F, Root (from Benjamin *et al.*, 1978).

Abbreviations : BA—6-benzyladenine ; 2, 4-D—2, 4-dichlor ophenoxyacetic acid ; IAA—indoleacetic acid ; Kn—kinetin ; NAA—naphthaleneacetic acid

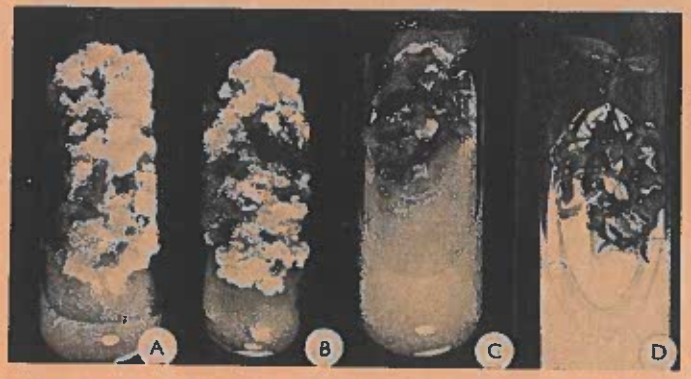


Fig. IV—A-D. Cultured tissues of *Hemidesmus indicus* : A—shoot tip callus, B—root callus, C—stem callus, D—stem callus with regenerated plantlet (from Heble and Chadha 1978)

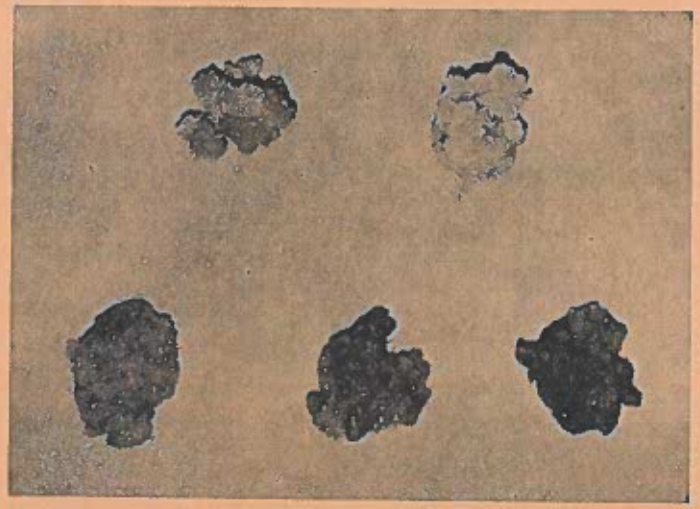


Fig. V. Cell strains isolated from stem callus of *Plumbago zeylanica* (from Heble *et al.*, 1974)

precursors leading to the formation of the end products. A number of tissue cultures are known to retain the genetic information necessary for product biosynthesis[59] and such cultures would be more suitable for studies concerning the influence of growth hormones and precursors.

Plant growth hormones are known to bring about morphogenetic and biochemical changes in cells and tissues. The biochemical changes that occur due to the hormonal stimuli may have profound influence on the metabolic events leading to the synthesis of secondary products. While the influence of growth hormones on morphogenesis has been well recognised for sometime[75], their effect on product biosynthesis was recognised only in the late 60's. Our studies with tissue cultures of *Solanum xanthocarpum* showed that the tissues synthesised the glycoalkaloid solasonine[86] and spirostanol diosgenin[87] thereby indicating that the cultures have retained the genetic information necessary for product biosynthesis. Growth hormones, such as 2, 4-D, IAA, IBA, GA and Kn were found to influence the formation of diosgenin and solasodine in the cultured tissues[86]. Of the different hormones tested, IAA and IBA, augmented the formation of diosgenin as compared to 2, 4-D and GA-Kn. The steroidal alkaloid solasodine formation was, however, suppressed by IAA and IBA (Table 4). On the other hand, in tissue culture of *Dioscorea deltoidea*. 2, 4-D was found to be more favourable for diosgenin synthesis than IAA or IBA[87].

The influence of auxins on nicotine synthesis has been extensively investigated. In tissue cultures of *Nicotiana tabacum*, 2, 4-D was found to be inhibitory to nicotine synthesis[20], but IAA and kinetin were found to favour the formation of nicotine[89]. In cell cultures of *Catharanthus roseus*, IAA was found to be superior for both cell growth and alkaloid synthesis as compared to 2, 4-D, NAA, BA and GA[97].

In cell cultures of *Morinda citrifolia*, anthraquinone formation occurred in the presence of IAA but not in the presence of 2, 4-D[96]. However, in tissue cultures of *Cassia tora*, 2, 4-D favoured the synthesis of anthraquinones to the extent that cell cultures synthesised more anthraquinones than the plant itself[91]. Our studies with tissue cultures of *Plumbago zeylanica* have shown that 2, 4-D favoured the synthesis of plumbagin and anthocyanins as compared to IAA, NAA or BA (Table 5)[80].

It is evident from these observations that at this stage no generalisation is possible with respect to the influence of growth hormones on product biosynthesis. The same auxins could promote the synthesis of a product

in one tissue and retard the synthesis of the same product in another as observed in case of diosgenin and anthraquinones. This is possibly due to the fact that the mode of action of growth hormones is diverse and their effect could be different in each system studied.

It is likely that the growth hormones could act at the enzymatic level involved in the biosynthetic pathways. Mizusaki *et al.*[68] reported that auxin inhibits the activity of putrescine-N-methyl transferase, catalyzing the N-methylation of putrescine which is a key intermediate in the biosynthesis of nicotine. The hormonal control of phenylalanine ammonia-lyase (PAL) activity has been demonstrated in several tissue cultures[48,66]. Auxins may also function through selection of cell lines with special biochemical potential by selectively favouring the growth of a particular cell type over the other. In tissue cultures of *P. zeylanica* pigmented strains which synthesized higher levels of plumbagin and anthocyanin grew well only with 2, 4-D in the medium[31]. Other auxins like NAA and IAA inhibited the growth of this cell strain.

Much of the work done in the area of hormonal influence on product biosynthesis is rather empirical. The present-day knowledge is confined to the effect of the hormones in terms of product formation only. More critical studies are needed to understand the hormonal regulations of product biosynthesis particularly in view of the fact that the hormones can exercise their influence at the nuclear, enzymatic, cellular, tissue and organ levels and bring about chromosomal aberrations, cytodifferentiation and organ differentiation. These events may have a direct effect on the biosynthetic potential of the cultured tissues.

Cultured cells which possess the capacity for product biosynthesis should also have the capacity to synthesise the primary and intermediate precursors. However, it is observed that in a number of instances an exogenous supply of precursors to the cultures results in increased product biosynthesis. Interesting examples of this type are the production of tropane alkaloids[47], indole alkaloids[97] and diosgenin[5]. On the other hand, a large number of tissue cultures do not respond to exogenous precursor addition.

Working with cell cultures of *Catharanthus roseus*, Zenk and coworkers[97] screened a large number of precursors of vinca alkaloids and observed that L-tryptophan was effective in augmenting the alkaloid synthesis by nearly three fold and in increasing the cell yield. Whereas D-tryptophan and tryptamine were detrimental to cell growth as well as alkaloid synthesis.

However, Doller *et al.*[17], observed that in cell cultures of *C. roseus*, L-tryptophan did not influence the alkaloid synthesis. It is, therefore, essential to screen a large number of primary and direct precursors for each strain of tissue to determine whether a precursor is suitable for a particular tissue or not.

Our studies with tissue cultures of *S. xanthocarpum* showed that the addition of cholesterol (an intermediate) did not influence the formation of diosgenin or solasodine. Addition of the precursors of plumbagin to the medium failed to influence the product biosynthesis in the pigmented strain of *P. zeylanica* (Table 6)[80]. In tissue cultures of *Tylophora indica*, the primary precursors of tylophora alkaloids, phenylalanine, tyrosine and ornithine had no influence on alkaloid formation although they were found to be beneficial for the growth of the tissues (Table 7)[4].

It is also important to ascertain (1) the acceptability of precursors (2) the optimal levels at which they function and (3) the exact phase of growth at which they are effective. The last aspect has not been studied in detail. Studies with tissue cultures of *D. deltoidea* have shown that the synthesis of diosgenin is more active during the late stationary phase of growth[84]. In suspension cultures of *Perilla ocymoides* the accumulation of caffeic acid and the enzyme PAL activity were observed during the lag phase and early exponential phase[48]. In this context, the use of primary precursors like acetate, simple amino acids, and sugars may be more beneficial, since they are involved in more than one metabolic event. In tea suspension cultures addition of ammonium salts to the medium increased the caffeine content by nearly four times[68].

IV. Heterogeneity of Cell Population and Selection of Biochemically Potent Cell Strains

Plant tissue and cell cultures comprise a heterogenous cell population. The heterogeneity could be cytological and/or morphological and/or biochemical. While the cytological and morphological heterogeneity have been well studied in several cultured tissues[85], the biochemical heterogeneity has received comparatively less attention. Blakely and Steward[7] reported the presence of pigmented and non-pigmented cells in tissue cultures of *Haplopappus gracilis*. They were able to isolate colonies which differed in their growth characteristics, pigmentation and the amino acid compositions. Working with tissue cultures of *Daucus carota*, Alfermann and Reinhard[1]

isolated anthocyanin pigmented and non-pigmented strains. Street and coworkers[11] isolated several colonies from cell cultures of *Atropa belladonna*, showing distinct differences in their growth characteristics and chlorophyll spectrum.

Successful isolation of biochemically potent cell strains, in regard to secondary products, was, reported from our laboratory[41]. Tissue cultures, initiated from stem segments of the plant *Plumbago zeylanica* comprised cells differring in pigmentation. By careful selection during serial subcultures, it was possible to establish three strains showing distinct variations in their growth characteristics, morphogenetic responses, pigmentation pattern (Fig. V) and the potential for the synthesis of plumbagin (5-hydroxy-2-methyl-1, 4-naphthoquinone) the principal antimicrobial constituent of the plant. The three strains namely PZ_1, PZ_2 and PZ_3, grew well on modified Murashige and Skoog's medium and maintained their morphogenetic and biochemical characteristics. Strain PZ_1 was pale yellow, compact, slow growing and turned chlorophyllous when grown on NAA+BA medium. This strain contained β-carotene and low levels of plumbagin. Strain PZ_2 was nodular, dark, initially slow growing and amenable to root organogenesis. After 10 to 12 serial subcultures the tissues became friable and lost the organogenetic potential. The tissue synthesized low levels of plumbagin. Strain PZ_3 was highly pigmented, friable and fast growing. This tissue synthesised high levels of plumbagin, the amounts being comparable to those found in the leaves of the mature plant (Table 8). The observations clearly demonstrated that in a given condition, not all cells in culture possess the synthetic capacity for a particular secondary product. It is also possible to select and grow high yielding biochemically potent cell lines from a heterogeneous mass of tissue.

Subsequently (1976), by using the technique of single cell plating, Tabata and Hiraoka[88] successfully isolated a number of colonies from cell cultures of *Nicotiana rustica*. The clones thus isolated showed wide variations in their growth characteristics and their ability to synthesize nicotine. One of the clones gave rise to a relatively stable strain capable of producing high levels of nicotine (0·29% dry weight). The authors emphasised that selection of such stable cell strains with high-yielding capacity would be of great importance in the potential exploitation of plant tissue cultures. High nicotine yielding cell lines have also been isolated from cell cultures of *N. tabacum*[67].

Using the technique of radioimmunoassay in clonal selection, Zenk and coworkers[97] were able to examine variability in cell cultures of *Catharanthus roseus* and *Solanum laciniatum* and select high-yielding cell strains. The alkaloids, serpentine and ajmalicine content of some selected cell lines is reported to be as high as 1·3% dry wt which is more than that observed in the parent plant.

Although selection of high-yielding cell strains has been accomplished in a number of plant tissue cultures, several uncertainties concerning the application of this method remain unsolved. One of the major considerations is the stability of the selected cell line. Plant tissue and cell cultures exhibit high genetic instability. The instability starts right at the initiation of the culture, possibly due to the exogenous hormonal application and the loss of the organised system. The instability continues during serial passages due to several factors which are discussed in the following section. Instability of cell strains over prolonged subculture has been reported in *Catharanthus* cell cultures. Some selected strains lost the capacity for alkaloid synthesis and in one case the synthetic capacity was restored[14]. Tabata and Hiraoka[88] isolated a high nicotine-yielding clone by selection from cell cultures of *N. rustica*. This clone spontaneously formed variant clones, of which one synthesised only trace amounts of nicotine and the other produced more than the original clone. More attention should, therefore, be paid to devise methods for stabilizing cell strains with high synthetic capacity. The use of growth hormones for selecting and maintaining the desired strains by altering their growth characteristics may prove advantageous. Continuous selection of the strains, although painstakingly tedious seems to be one of the important approaches to maintain high-yielding characteristics of the strains. A number of cultures are known to exhibit stability in regard to the synthesis of secondary products. Studies of the factors responsible for their stability may prove rewarding in understanding the behaviour of different cell strains.

V. Nuclear Cytology and Product Biosynthesis

Plant tissues and cells, in culture, display considerable nuclear irregularities since they are free from the stabilizing environment of the intact organism. Although numerous studies have been carried out in regard to the nuclear variability of cultured tissues, the impact of these events on the biosynthetic potential has not been critically examined. The influence of

ploidy status on the formation of secondary metabolites was recognised in the early 40's. Thus Rawson[76] reported increased tropane alkaloid production in autotetraploids of *Atropa* and *Hyoscyamus*. Similar observations have also been made in *Solanum khasianum*, a steroidal alkaloid yielding plant (Table 9)[61,82]. Soloman and Crane[78] have suggested that heredity in alkaloidal phenotypes of *A. belladonna* was under the influence of polygenic system. Our studies with tissue cultures derived from the leaves of haploid *A. belladonna* indicated that the ploidy status of the regenerated plants not only influences the total alkaloid production (Table 10) but also has a bearing on selective control over the biosynthetic pathway (Fig. VI)[19]. The results lend further support to the hypothesis that the ploidy status has a strong influence on the metabolic events of the plant.

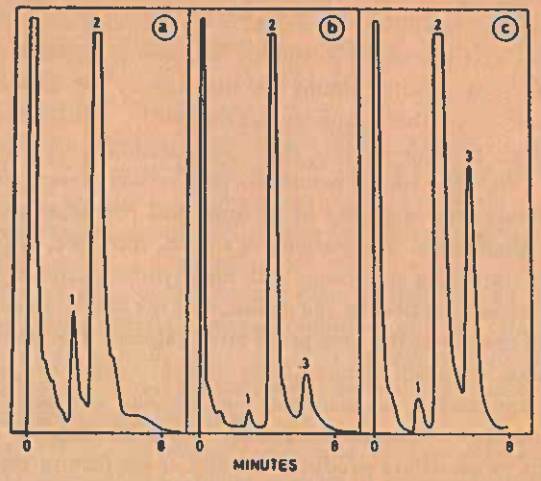


Fig. VI. a-c. Gas-liquid chromatography of basic extracts isolated from diploid (a), triploid (b), and tetraploid (c) regenerants of Atropa belladonna using 3% SE-30. 1 = tropine, 2 = atropine and 3 = scopolamine. (From Eapen *et al.*, 1978)

The situation in callus and cell cultures, however, is different and complicated. During the process of *in vitro* proliferation of an organised tissue, essentially two types of chromosomal changes may take place, which would result in nuclear irregularities. These are : (1) endoreduplication followed by mitosis and (2) chromosomal breakage followed by mitosis. Unlike in

the intact organism, in cultured tissues there exists a rigid process of elimination of the unwanted variants. As a result, under given condition of growth, the most actively proliferating cell types predominate over the others causing a loss of the stem line or the cell lineage. These aspects have been reviewed recently by Sunderland[86] and D'Amoto[10]. Since the process of biosynthesis is essentially under genetic control, it would be worthwhile to ascertain the extent to which the chromosomal irregularities affect the metabolic status of the cultured tissues.

While the degree of polyploidization in certain plants, may contribute significantly towards the biosynthesis of end products, in cultured tissue there are no studies in this regard. On the other hand, chromosomal mutations caused due to chromosome breakage in cultured cells may result in structural changes and loss of genetic material. Such events could obviously give rise to tissue cultures which either fail to synthesize the characteristic plant products or produce compounds unknown to the parent plant. Although several plant tissue cultures are known to exhibit such metabolic irregularities as they produce altogether new products not found in the parent plant[25,39,42], we have no knowledge of their chromosomal status. It is, therefore, essential to select cell lines which have the stem lineage and exhibit cytological stability during prolonged subcultures.

Several tissue culture systems exhibit cytological stability during serial subcultures. Sacristan[77], working with haploid and diploid plants of *Crepis capillaris*, observed that 34% of the haploid cultures and 79% of the diploid cultures maintained the original ploidy level over a period of more than a year. Reinert and Kuster[74] observed that tissue cultures of *C. capillaris* did not undergo changes in the chromosome numbers after one year in culture. *C. capillaris* (2n = 6) is a non-polysomatic species and therefore the explants taken from any part of the plant consist entirely of cells of one ploidy level. Tissue cultures derived from such plants would be ideal for biosynthetic studies.

Yet another type of stability is the one observed in cultures derived from highly polyploid species or a species containing large number of small chromosomes. In such a population the degree of change is comparatively slow and as a result they tend to maintain some equilibrium. Our studies with cultured tissues of *Atropa belladonna*, a species containing large number of small chromosomes (2n = 74) showed that the tissues maintained chromo-

somal stability during serial subcultures[20(a)]. The biosynthetic potential of these cultures, however, was rather poor.

Although several tissue cultures are known to exhibit stability in their biosynthetic characteristics over prolonged period of subculture, their cytological status has not been studied[21,52]. Tissue cultures of *Morinda citrifolia* maintained over a period of seven years exhibited stability in regard to the production of anthraquinones[26]. The nuclear DNA composition of these cultures did not show major variations when compared with the root tips of the differentiated plants[25].

There are also instances where mutant cell lines with high selective advantage and distinct characteristics emerge from a heterogeneous cell population. These cells, under controlled nutritional conditions may attain stability to certain extent. In tissue cultures of Haplopappus a mutant strain producing anthocyanin was obtained from a pale yellow callus[61]. The mutant strain exhibited biosynthetic stability over many years of subculture despite the fact that 90% of the cells were of irregular karyotype. Spontaneous emergence of distinct strain from a heterogeneous cell population has been demonstrated in a few cultures[40], but very little insight is available into their cytological status. The stability of such mutants during prolonged subculture is always an open question.

While examples quoted above refer to a class where stability is natural there are a few instances where stability is achieved through physicochemical factors. One such example is the haploid tissue cultures of *N. tabacum*. Gupta and Carlson[29] were able to select haploid tobacco cells using increasing levels of parafluorophenylalanine. Further investigations by Dix and Street[16] revealed that the effect of this compound varied significantly with different tissue cultures. Nevertheless, the fact that selection through the use of chemicals is possible gives incentive to explore the utility of other compounds.

In establishing a stable cell line which carries the lineage of the stem line, several factors are to be taken into consideration. The nature of the explant is an important factor which contributes towards the ploidy status of the cultured tissues[68,69]. Subculture of the tissues at regular intervals is known to be beneficial in maintaining the stability of the cell lines. Since growth hormones used in the maintenance of the cultures are known to exert profound influence on the ploidy status[62], minimal use of these compounds may prove beneficial towards stability. Another approach could be

the continuous selection of stable and high-yielding strains through clonal selection. It is well known that even these strains selected through cloning would give rise to further substrains[58]. Semiorganised cultures or organ cultures may be more rewarding in obtaining stable lines. In these cultures, although the tissues are taken out of the environment of the whole plant the cells are still protected by the organised structures. Several organ cultures have been established and studied for their biosynthetic potential[55,56], but very little is known about their cytological and biosynthetic stabilities. Critical studies of these aspects could prove more rewarding.

Even though the cytological studies of the cells and tissues in culture constitute a formidable effort, but in view of their implications in the processes of biosynthesis, such studies warrant serious consideration.

VI. Biotransformations

The ability of plant tissue and cell cultures to transform organic molecules was recognised as early as 1964, by Stohs and Staba[62], who demonstrated the transformation of cardiac glycosides. They were, however, unable to isolate the transformed products. Subsequent reports by Graves and Smith[28] and Furuya *et al.*[24], revealed a number of biotransformations of steroids. The usual reactions involved in these transformations were reduction of double bonds, reduction of keto to hydroxyl groups, esterification and glycosylation. Our studies with tissue cultures of *H. antidysenterica* have shown that cholesterol C^{14} is biotransformed to campesterol and sitosterol which involves successive methylation of the sterol side chain. However, since 1974, a great deal of progress has been made in the field of biotransformations by plant cell cultures. Of particular interest are the transformations of cardiac glycosides by cell cultures of *Digitalis lanata* and *Digitalis purpurea* studied by Furuya *et al.*[22], and Reinhard *et al*[3]. Two important reactions, viz., glycosylation and specific hydroxylation of cardenolides have been demonstrated.

The major source of cardeotonic glycosides is the plant *D. lanata*, which contains two important glycosides, digoxin and digitoxin. Since digoxin is a more favoured drug used in cardiac therapy, there is considerable accumulation of digitoxin. Significant work has been done in Reinhard's laboratory at Tubingen, with regard to the transformation of digitoxin by cell cultures of *D. lanata*. This group has standardised the techniques involved in the biotransformation of β-methyldigitoxin to its 12β-hydroxy derivative which

is an important drug used in cardiac therapy[44]. The transformations have been successfully carried out in 20 liter capacity airlift fermentors.

Other important biotransformations reported in plant cell cultures are stereospecific reductions and isomerisations. Itokawa *et al.*[51], reported trans⇌cis isomerization like geraniol⇌nerol. Stereospecific reduction of (–)-codeinone to (–)-codeine was observed in cell cultures of *Papaver somniferum*, *Digitalis* and *Nicotiana*[32]. The transformation of D-tryptophan to L-tryptophan by the racemase enzyme in tobacco tissue cultures, reported by Miura and Mills[90] is of particular interest since this observation illustrates the potential use of cell cultures for the conversion of racemic compounds to optically active forms. In addition to these, biotransformations involving glycosylation esterification, methylation and oxidation have also been reported[23].

Further search for rare and complex reactions which are specific to plant cells would be more rewarding. The progress achieved in this area of plant cell cultures indicates that the biotransformation potential of plant cell cultures would be an important aspect of practical value.

VII. Conclusions

Studies of growth and metabolism of plant cells, during the past three decades, have resulted in several important scientific contributions. Several high-yielding tissue cultures have been established in the past few years, though a number of problems remain to be solved for their practical utilization. This is mainly due to the fact that different aspects of plant tissue culture research are carried out by scientists with high degree of specialisation. A deep insight into the processes of secondary metabolism in plant cells, on the other hand, requires knowledge of several functions like growth, differentiation, development, cytology, etc. A clear understanding of these processes is, therefore, possible only through an integrated approach by plant tissue culture scientists belonging to different disciplines.

A dialogue between the drug industry and the basic scientists for the identification of priority areas and selection of problems which are otherwise impossible or very difficult to solve by the application of conventional techniques is very important.

The plant kingdom has provided a wide range of therapeutics some of which occur in very minute quantities along with a large number of related compounds. Examples of this type are alkaloids like reserpine, codeine,

scopolamine, quinidine, vincristine and vinblastine. Tissue culture techniques could be exploited to advantage with these plants. For this purpose induction of genetic variability through the techniques of anther cultures may be worth trying. By this method it could be possible to synthesize a large population of variants and homozygous recombinants[89]. Extensive screening of these derivatives could possibly yield strains with desired biosynthetic capacity. Tissue and cell cultures of such strains would be of particular interest.

Maintenance of the selected strains for longer period is an important aspect. In this context organ cultures of the selected strains which would be relatively more stable, may prove advantageous since they are less amenable to nuclear irregularities caused by exogenous factors.

The overall progress made during the past few years reflects an optimistic view towards the utilization of plant tissue and cell cultures by the industry. Nevertheless, there are several unsolved problems which need critical examination. The plant tissue culture research has also made significant contributions in several areas of plant sciences. To quote examples from our own laboratory, these include clonal propagation[73], cultures of isolated protoplasts and somatic hybridization[72], production of haploids and their implications in plant breeding[27] and hormonal interactions in the processes of growth and differentiation[3]. Besides these, there are other new and novel studies like the uptake of microorganisms, host-parasite interactions, production of virus and micoplasma free and resistant strains and genetic engineering[94].

It is, therefore, altogether possible that from an area of scientific enquiry which was basic and fundamental to start with, there may result in not too distant a future, plausible technological applications. Already, the benefits of the application of plant tissue culture in practical agriculture, horticulture and forestry are reportedly many and similar benefits in biotechnology are perhaps on the threshold of realisation.

Table 1

Therapeutics Obtained from Higher Plants

Type	Components	Plant	
		Family	Genus and Species
Tropane alkaloids	Atropine	Solanaceae	*Atropa belladonna*
	Scopolamine		*Datura stramonium*
			Hyoscyamus niger
Quinoline alkaloids	Quinine	Rubiaceae	*Cinchona succirubra*
	Quinidine		
Isoquinoline alkaloids	Morphine	Papaveraceae	*Papaver somniferum*
	Codeine		
	Berberine		*Argemone mexicana*
	Emetine	Rubiaceae	*Cephaelis ipacacuanha*
Indole alkaloids	Reserpine	Apocynaceae	*Rauvolfia serpentina*
	Serpentine		*Catharanthus roseus*
	Vincristine		
	Vinblastine		
Steroid alkaloids	Conessine	Apocynaceae	*Holarrhena antidysenterica*
	Solasodine	Solanaceae	*Solanum khasianum*
Alkaloid amines	Ephedrine	Ephedraceae	*Ephedra sinica*
Steroids	Diosgenin	Dioscoreaceae	*Dioscorea deltoidea*
			D. floribunda
			D. composita
	Digitoxin	Scrophulariaceae	*Digitalis purpurea*
	Digoxin		*D. lanata*
	Scillaren A and B	Liliaceae	*Urginea maritima*
Quinones	Sennoside	Fabaceae	*Casia aungustifolia*
	Aloe-emodin	Liliaceae	*Aloe barbadensis*

Table 2

Alkaloid Content in Cultured Tissues and Regenerated Mature Plant of *Atropa Belladonna*

Material	Growth medium	Passage No.	Morphogenetic response	Total alkaloid µg/g dry wt.
Diploid				
Callus	W & B	5+6+7	Unorganized growth	8
Callus	MS	7+8*	Unorganized growth	11
Callus	MS	5	Shoot bud formation	12
Regenerated shoot buds	MS	5	—	20
Haploid				
Callus	W & B	5+6+7	Shoot bud formation	8
Callus	MS	7+8	Shoot bud formation	11
Regenerated shoot buds	MS	7+8	—	18
Regenerated flowering plant	—	—	—	1160

*Cultures exhibited loss of organogenetic potential

Table 3

Alkaloid Content in Different Organs of Normal and Regenerated Flowering Plants of *Tylophora Indica*

Material	Normal plant Alk % dry wt.	Regenerated plant Alk % dry wt.
Leaves	0·29	0·33
Stem	0·25	0·21
Root	0·41	0·27

Table 4

Steroidal Content of Callus Tissues of *Solanum Xanthocarpum* Treated with Different Hormones

Treatment	*Dry wt. of callus tissues (g)*	*β-Sitosterol (mg)*	*Diosgenin (mg)*	*Solasodine presence (+) absence (−)*
2, 4-D (1 ppm) (control)	35	17·8	1·3	+
IAA (2 ppm)	25	6·0	2·0	−
IBA (2 ppm)	35	10·2	3·6	−
Kn (2 ppm) + GA (1 ppm)	35	12·8	1·6	+

Table 5

Influence of Growth Hormones on Plumbagin and Anthocyanin Synthesis in Pigmented Strain of *Plumbago Zeylanica*

Growth hormones (ppm)	*Growth value*	*Relative anthocyanin content*	*Plumbagin μg/g fresh wt.*
2, 4-D (2) + Kn (0·2)	5·0	0·27	20·0
IAA (2) + Kn (0·2)	2·7	0·07	4·9
NAA (2) + Kn (0·2)	2·2	0·10	3·9
2, 4-D (0·2) + BA (2)	2·6	0·16	8·4

Table 6

Influence of Precursors on Plumbagin and Anthocyanin Synthesis in Pigmented Strain of *Plumbago Zeylanica*

Precursor (ppm)	*Growth value*	*Relative anthocyanin content*	*Plumbagin μg/g fresh wt.*
Control	5·1	0·28	21·0
Shikimic acid (200)	5·7	0·29	19·8
Tyrosine (200)	3·5	0·21	21·2
Phenylalanine (200)	4·5	0·31	14·6

Table 7

Effect of Precursors on Growth and Alkaloid Formation in Leaf Callus of *Tylophora Indica*

Precursors mg/l	*Growth value fresh wt.*	*Growth value dry wt.*	*Dragendorff +ve components*
1 Control	9·2 ±0·354	12·9 ±0·398	+*
2 Phenylalanine (100)	11·5 ±0·743	12·7 ±1·282	+
3 Tyrosine (100)	8·34±0 536	9·0 ±0·624	+
4 Ornithine (100)	9·8 ±0·949	9·38±0·810	+
5 Phenylalanine (50) + Tyrosine (50) + Ornithine (50)	9·16±0·687	13·6 ±0·932	+

* Indicates trace amounts of Dragendorff-positive components as observed on TLC.

Table 8

Influence of Plant Hormones on the Growth and Biosynthesis of Plumbagin in Variant Cell Strains Derived from *Plumbago Zeylanica*

(Growth period = 30 days.)

Cell line	*Growth hormone*	*Response of the tissue*	*Relative anthocyanin content (absorbance)*	*Plumbagin mg/100 g (fresh wt.)*	*Growth value (fresh wt.)*
Pz_1	A	Compact, pale yellow	—	0·17	4·4
	B	Compact, pale yellow	—	0·11	3·2
	C	Compact, yellow	—	0·26	2·2
	D	Compact, intense green	—	0·40	2·7
Pz_2	A	Dark, nodular	—	0·37	1·4
	B	Rooting (90%)	—	0·84	1·4
	C	Rooting (90%)	—	0·14	2·1
	D	Dark, nodular	—	0·42	2·4
		Isolated roots in culture	—	0·97	—
Pz_3	A	Friable, intensely red	0·27	2·00	5·0
	A[a]	Loss of pigments	0·05	3·30	3·9
	B	Loss of pigments	0·07	0·49	3·8
	C	Rooting (5%), loss of pigments	0·10	0·39	3·7
	D	Loss of pigments	0·19	0·82	2·2

[a] Growth in dark

(A) 2, 4-D (2·0)+Kn (0·2) ; (B) IAA (2·0)+Kn (0·2) ; (C) NAA (2·0)+Kn (0·2) ; (D) NAA (0·2)+BA (2·0).

Table 9

Alkaloid Content in Diploids and Tetraploids of *Solanum Khasianum*

Plant type	*Average glycoalkaloid content % dry wt. (by wt.)*		*Average solasodine content % of dry wt. (by wt.)*		*Average solasodine content % dry wt. (by colorimetry)*	
	1974	*1975*	*1974*	*1975*	*1974*	*1975*
Control (2n)	4·38	—	1·78	—	1·43	—
Curved spine mutant (2n)	4·75	—	2·02	—	1·47	—
Control tetraploid (4n)	8·09	6·58	2·8	3·04	1·89	1·97
Curved spine mutant tetraploid (4n)	7·72	6·54	2·5	3·02	1·85	1·88

Table 10

Alkaloid Content in Tissue Cultures and Regenerated Plants of *A. Belladonna*

Material	*Total alkaloid mg dry wt.*	*Relative concentration %*		
		Tropine	*Atropine*	*Scopolamine*
Leaf callus	0·8	21·1	29·5	49·4
2n plants	116·0	7·35	92·66	Traces
3n plants	228·0	3·27	92·08	4·63
4n plants	268·0	1·69	80·54	17·78

REFERENCES

1. Alfermann, A. W. and Reinhard, E. (1971) *Experientia*, **27** : 353.
2. Alfermann, A.W., Boy, H. M., Doller, P. C., Hagedorn, W., Heins, M., Wahl, J. and Reinhard, E. (1977) *In* "Plant Tissue Culture and its Biotechnological Application", (Eds. W. Barz, E. Reinhard and M. H. Zenk), Springer-Verlag, Berlin, Heidelberg, New York : 125.
3. Bapat, V. A. and Rao, P. S. (1979) *Ann. Bot.*, **44** : 629
4. Benjamin, B. D., Heble, M. R. and Chadha, M. S. (1979) *Z. Pflanzenphysiol*, **92** : 77.
5. Bhandary, R., Collin, H. A., Thomas, E. and Street, H. E. (1969) *Ann. Bot.*, **33** : 647.
6. Bhat, B. and Heble, M. R. (1978) *Environmental and Experimental Bot.*, **18** : 127.
7. Blakely, L. M. and Steward, F. C. (1964) *Am. J. Bot.*, **51** : 809.
8. Collins, G. B., Leg, P. D. and Kasperbauer, M. J. (1974) *Crop. Sci.*, **14** : 77
9. Corduan, G. (1978) *In* "Production of Natural Compounds by Cell Culture Methods", (Eds. A. W. Alfermann and E. Reinhard), BPT-Report 1/78, Munchen : 295.
10. D'Amato, F. (1978) *In* "Frontiers of Plant Tissue Culture", (Ed. T. A. Thorpe), International Association for Plant Tissue Culture, Calgory, Alberta, Canada : 287.
11. Davey, M. R., Fowler, M. W. and Street, H. E. (1971) *Phytochem.*, **10** : 2559.
12. Dawson, R. F. (1942) *Am. J. Bot.*, **48** : 813.
13. Delfel, N. E. and Rothfus, J. A. (1977) *Phytochem.*, **16** : 1595.
14. Deus, B. (1978) *In* "Production of Natural Compounds by Cell Culture Methods", (Eds. A. W. Alfermann and E. Reinhard), BPT-Report, 1/78 Germany, 118.
15. Dhoot, G. K. and Henshaw, G. G. (1977) *Ann. Bot.*, **41** : 943.
16. Dix, P. J. and Street, H. E. (1975), *Plant Sci. Letters*, **5** : 231.
17. Doller, G., Alfermann, A. W. and Reinhard, E. (1976) *Planta Med.*, **30** : 14.
18. Eapen, S., Rangan, T. S., Chadha, M. S. and Heble, M. R. (1978) *Plant Sci. Letters*, **13** : 83.
19. Eapen, S., Rangan, T. S., Chadha, M. S. and Heble, M. R. (1978) *Can. J. Bot.*, **56** : 278.
20. Eapen, S., Rangan, T. S., Chadha, M. S. and Heble, M. R. (1980) Proc. Symp. Plant Tissue Culture, Genetic Manipulation and Somatic Hybridization of Plant Cells, DAE, Bombay, India (*in press*).
21. Furuya, T. and Kojima, H. (1968) *Kitasato Arch. of Exp. Med.*, **21** : 47.
22. Furuya, T. (1978) *In* "Frontiers of Plant Tissue Culture", (Ed. T. A. Thorpe), International Association for Plant Tissue Culture, Calgary, Alberta, Canada : 191.
23. Furuya, T. and Ishi, T. (1972) German Patent 2 143 946.
24. Furuya, T., Hirotoni, M. and Shinohara, T. (1970) *Chem. Pharm. Bull.*, **18** : 1080.
25. Furuya, T., Ikuta, A. and Syono, K. (1972) *Phytochem.*, **11** : 3041.
26. Furuya, T., Kojima, H. and Syono, K. (1971) *Phytochem.*, **10** : 1529.
27. George, L. and Narayanaswamy, S. (1973) *Protoplasma*, **78** : 467.
28. Graves J. and Smith, W. K. (1967) *Nature*, **214** : 1248.
29. Gupta, S. and Carlson, P. S. (1972) *Nature*, **239** : 86.
30. Heble, M. R. (1975) Ph.D. Thesis, University of Bombay, India.

31. Heble, M. R. (1977) *In* "Cultivation and Utilization of Medicinal and Aromatic Plants", (Eds. C. K. Atal and B. M. Kapur), Regional Research Laboratory, Jammu-Tawi, India : 510.
32. Heble, M. R. and Bhat, B. (1978) *Indian J. Exp. Biol.*, **14** : 77.
33. Heble, M. R. and Chadha, M. S. (1978) *Z. Pflanzenphysiol.*, **89** : 401.
34. Heble, M. R. and Staba, E. J., *Planta Med. (in press)*.
35. Heble, M. R. and Staba, E. J., *Planta Med. (in press)*.
36. Heble, M. R., Narayanaswamy, S. and Chadha, M. S. (1968) *Naturwissenschaften*, **55** : 350.
37. Heble, M. R., Narayanaswamy, S. and Chadha, M. S. (1968) *Science*, **161** : 1145.
38. Heble, M. R., Narayanaswamy, S. and Chadha, M. S. (1971) *Phytochem.*, **10** : 2393.
39. Heble, M. R., Narayanaswamy, S. and Chadha, M. S. (1971) *Z. Naturforsch.*, **26** : 1382.
40. Heble, M. R., Narayanaswamy, S. and Chadha, M. S. (1973), Proc. Symp. Control Mechanisms in Cellular Processes, DAE, Bombay, India : 557.
41. Heble, M. R., Narayanaswamy, S. and Chadha, M. S. (1974) *Plant Sci. Letters.*, **2** : 405.
42. Heble, M. R., Narayanaswamy, S. and Chadha, M. S. (1976) *Phytochem.*, **15** : 681.
43. Heble, M. R., Narayanaswamy, S. and Chadha, M. S. (1976) *Phytochem.*, **15** : 1911.
44. Heins, M., Wahl, J., Lerch, L., Kaiser, F. and Reinhard, E. (1978) *Planta Med.*, **33** : 57.
45. Hirotani, M. and Furuya, T. (1977) *Phytochem.*, **16** : 610.
46. Hiraoka, N. and Tabata, M. (1974) *Phytochem.*, **13** : 1671.
47. Hiraoka, N., Tabata, M. and Konoshima, M. (1973) *Phytochem.*, **12** : 795.
48. Ibrahim, R. K. and Edgar, D. (1976) *Phytochem.*, **15** : 129.
49. Ikeda, T., Matsumoto, T. and Noguchi, M. (1976) *Phytochemistry*, **15** : 568.
50. Ikuta, A., Syono, K. and Furuya, T. (1974) *Phytochem.*, **13** : 2175.
51. Itokawa, H., Takeya, K. and Akasu, M. (1976) *Chem. Pharm. Bull.*, (Tokyo), **24** : 1681.
52. Kaul, B. and Staba, E. J. (1967) *Planta Med.*, **15** : 145.
53. Khanna, Pushpa (1977) *In* "Cultivation and Utilization of Medicinal and Aromatic Plants", (Eds. C. K. Atal and B. M. Kapur), Regional Research Laboratory, Jammu-Tawi, India : 495.
54. Krikorian, A. D. and Steward, F. C. (1969) *In* "Plant Physiology", (Ed. F. C. Steward), Academic Press, New York, London : 227.
55. Lewis, W. H. and Memory, P. F. Elvin-Lewis (1977), Medical Botany, John Wiley and Sons, New York, London : 133.
56. Lui, J. H. C. and Staba, E. J. (1979) *Phytochem.*, **18** : 1913.
57. Marshall, J. G. and Staba, E. J. (1976) *Phytochem.*, **15** : 53.
58. Mehta, A. R. and Staba, E. J. (1970) *J. Pharm. Sci.*, **59** : 864.
59. Misawa, M. (1977) *In* "Plant Tissue Culture and its Biotechnological Application", (Eds. W. Barz, E. Reinhard and M. H. Zenk), Springer-Verlag, Berlin, Heidelberg, New York : 17.
60. Misawa, M., Sakato, K., Tanaka, H., Hayashi, M. and Samejima, H. (1974) *In* "Tissue Culture and Plant Science", (Ed. H. E. Street), Academic Press, London : 405.
61. Mitra, G. C. (1972) *Indian J. Exp. Biol.*, **10** : 217.
62. Miura, H. and Mill, S. E. (1971) *Plant Physiol.*, **47** : 483.

63. Mizusaki, S., Tanabe, Y., Noguchi, M. and Tamaki, E. (1971) *Phytochem.*, **10** : 1347.
64. Murashige, T. (1978) *In* "Frontiers of Plant Tissue Culture", (Ed. T. A. Thorpe), International Association for Plant Tissue Culture, Calgary, Alberta, Canada : 15.
65. Murashige, T. and Nakano, R. (1967) *Am. J. Bot.*, **54** : 963.
66. Nash, D. T. and Davies, M. E. (1972) *J. Exp. Bot.*, **23** : 75.
67. Ogino, T., Hiraoka, N. and Tabata, M. (1978) *Phytochem.*, **17** : 1907.
68. Ogutuga, D. B. A. and Northcote, D. H. (1970) *J. Exp. Bot.*, **21** : 258.
69. Partanen, C. R. (1965) *In* "Proc. Intern. Conf. Plant Tissue Culture", (Eds. P. R. White and A. R. Grove), McCutchan Publ. Corp., Berkely, California : 463.
70. Peters, J. E., Wu, P. H. L., Sharp, W. R. and Paddock, E. F. (1974) *Physiol. Plant*, **31** : 97.
71. Rai, P. P. (1978) *In* "Frontiers of Plant Tissue Culture", (Ed. T. A. Thorpe), International Association of Plant Tissue Culture, Calgary, Alberta, Canada : 1422.
72. Rajasekhar, E. W. (1973) *Nature*, **246** : 223.
73. Rao. P. S. and Bapat, V. A. (1978) *Can. J. Bot.*, **56** : 1153.
74. Reinert, J. and Kuster, H. J. (1966) *Z. Pflanzenphysiol.*, **54** : 213.
75. Reinert, J., Bajaj, Y. P. S. and Zbell, B. (1977) *In* "Plant Tissue and Cell Culture, (Ed. H. E. Street), Blackwell Scientific Publications, Oxford, London, Edinburgh, Melbourne : 389.
76. Rowson, J. M. (1945) *J. Pharm. Pharmacol.*, **18** : 185.
77. Sacristan, M. D. (1971) *Chromosoma*, **33** : 273.
78. Solomon, M. J and Crane, F. A. (1970) *J. Pharm. Sci.*, **59** : 1670.
79. Staba, E. J., Plant Tissue Culture as a Source of Biochemicals (*in press*).
80. Staba, E. J. and Kaul, B. (1971) US Patent 3 628 287.
81. Stickland, R. G. and Sunderland, N. (1972) *Ann. Bot.*, **36** : 671.
82. Stohs, S. J. and Staba, E. J. (1964) *J. Pharm. Sci.*, **54** : 56.
83. Street, H. E. (1969) *In* "Plant Physiology", (Ed. F. C. Steward), Academic Press, New York and London, 3.
84. Street, H. E. (1977) *In* "Plant Tissue and Cell Culture" (Ed. H. E. Street), Blackwell Scientific Publications, Oxford, London, Edinburgh, Melbourne, 61.
85. Street, H. E. (1977) *In* "Plant Tissue and Cell Culture", (Ed. H. E. Street), Blackwell Scientific Publications, Oxford, London, Edinburgh, Melbourne : 207.
86. Sunderland, N. (1977) *In* "Plant Tissue and Cell Culture", (Ed. H. E. Street), Blackwell Scientific Publications, Oxford, London, Edinburgh, Melbourne : 177.
87. Tabata, M. (1977) *In* "Plant Tissue Culture and its Biotechnological Application", (Eds. W. Barz, E. Reinhard and M. H. Zenk), Springer-Verlag, Berlin, Heidelberg, New York, 3.
88. Tabata, M. and Hiraoka, N. (1976) *Physiol. Plant*, **38** : 19.
89. Tabata, M., Marumoto, Y. and Konoshima, M. (1971) *Phytochem.*, **10** : 723.
90. Tabata, M., Yamamoto, H., Hiraoka, N. and Konoshima M. (1972) *Phytochem.*, **11** : 949.
91. Tabata, M., Hiraoka, N., Iknoue, M., Sano, Y. and Konoshima, M. (1975) *Lloydia.*, **38** : 131.
92. Torrey, J. G. (1967) *Physiol. Plant*, **20** : 265.
93. Uomori, A. and Tomita, Y. (1974) *Japan. J. Pharmacog.*, **28** : 152.
94. Vasil, I. K. (1976) *In* "Advances in Agronomy", Academic Press, New York, San Francisco, London : 119.
95. Zenk, M. H. (1978) *In* "Frontiers of Plant Tissue Culture", (Ed. T. A. Thorpe), International Association for Plant Tissue Culture, Calgary, Alberta, Canada : 1.
96. Zenk, M. H., El-Shagi, H. and Schutle, U. (1975) *Planta Medica* (Supplement), 75.
97. Zenk, M. H., El-Shagi, H., Arens, H., Stockigt, J., Weiler, E. W. and Deus, B. (1977) *In* "Plant Tissue Culture and its Biotechnological Application", (Eds. W. Barz, E. Reinhard and M. H. Zenk), Springer-Verlag, Berlin, Heidelberg, New York : 27.

Reproductive Endocrinology of Decapod Crustacea

R. Nagabhushanam, G. K. Kulkarni and R. Sarojini

Department of Zoology, Marathwada University, Aurangabad-431 004

National Academy of Sciences, India : Golden Jubilee Commemoration Volume, 1980

1. General Introduction

Of major significane to investigators in the field of crustacean endocrinology, which has its beginning in the latter half of the 1920's, is the fact that many of the crustacean hormones are neurosecretory products. The early efforts in this area were essentially directed toward identification of the sources of the hormones and their functions at the organismic level. Normally, the term "neurohumors" (Chemical intermediaries of nerve transmission which are essentially of low molecular weight, water soluble compounds, charged at physiological pH and delivered from "sender" to "deliver" cell : Wurtman[197]) and neurohormones (which usually possess a long chemical reach both in time and space and are released through vascular channels, exceptions : neurosecretomotor cell). In the present review the literature concerning the physiological role of neurohormones alone has been considered. Neurohormones are produced by neurosecretory cells, which are modified neurons and are transported by axoplasmic flow either to discharge peripherally[117] or carried to storage sites in modified axonal terminals prior to release (e.g., the medulla terminalis X-organ-sinus gland system in crustaceans). Groups of such modified terminals are also known as "neurohaemal organs". This concept of neurosecretion; largely developed by Scharrer and Scharrer[158,159], has attracted a large number of investigators during the last few decades.

In decapod crustaceans, central nervous system is mainly comprised of eyestalks, brain, thoracic ganglion and abdominal ganglia, which are generally six in number (but are not present in brachyurans) (Fig. 1). The neurohormones produced from these neuroendocrine centres are known to regulate a number of diverse physiological phenomena in crustaceans. Among all the neuroendocrine centres extensive investigation has been done on the structural and functional aspects of eyestalk profile and reported that eyestalks contain a variety of hormones or factors apparently governing diverse functions such as growth, moulting, metabolic rates, heart rate, carbohydrate and protein metabolism, water balance, pigment migration and sexual activity[64], and the active principles appear to be elaborated by groups of neurosecretory cells present in the three parts, viz., medulla externa, medulla interna and medulla terminalis of optic lobe peduncle, contained within the eyestalks. The neurosecretory cells in the central nervous system of crustaceans can be grouped into various types based on

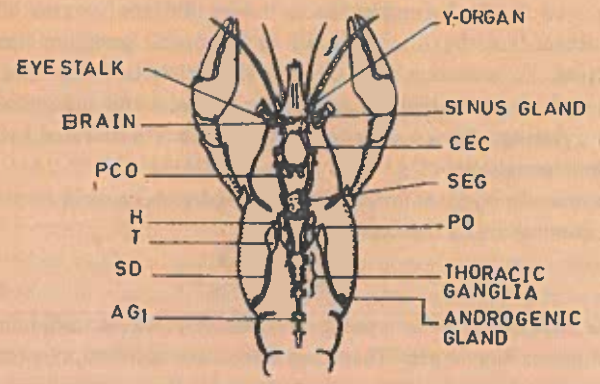


Fig. 1. The endocrine system of a generalized male decapod crustacean. Principal endocrine areas are indicated on the figure with complete labels. The endocrine system includes neurohemal organ such as sinus gland and non-neural epithelial glands such as the *Y*-organ and androgenic gland (from Tombes, 1970)

Abbreviations :

AG_1—First abdominal ganglion, CEC—Circumesophageal connective, H—Heart, PCO—Post-commisural organ, PO—Pericardial organ, SD—Sperm duct, SEG—Subesophageal ganglion, T—Testis

differences in size, shape, locality, cytological and cytophysiological details and tinctorial behaviour[1,59,61,116,134,167]. In the neurosecretory system of the individual eyestalk, the morphology of the *X*-organ sinus gland complex deserves special mention. It is believed to secret a hormone controlling both reproduction and moulting. In adult females of several crustaceans, eyestalk ablation resulted in ovarian growth and precocious yolk deposition in the ovary[11,16,87,107,124,127,164]. Whereas destalking in non-breeding adult males has been claimed to induce precocious spermatogenesis, enlargement of the vasa deferentia and hypertrophy and hyperplasia in the androgenic glands[54,57,58,74,103,104,139,145,146,152,154,189]. It is now well known that eyestalks from males and females are reciprocally effective in restoring the normal maintenance of the gonads in eyestalkless decapods, so there is no reason to believe that the ovary-inhibiting and testis-inhibiting principles differ from one another. Therefore, the convenient term *inhibitory gonadotropin* will commonly be used to indicate the principle(s) coming under this category in both the sexes, and

plausibly also in the hermaphrodites, during different phases of gonadal activity. Regarding the role of brain and thoracic ganglion hormone in reproduction, it was reported that implantation of brain and thoracic ganglion into young individuals resulted in a precocious maturation of the gonads suggesting a gonad stimulating substance is elaborated by the brain and thoracic ganglion[2,12,21,73,74,89,90,103,104,127,138,141,147,190]. The term *stimulatory gonadotropin*, therefore, has been used to denote the factor(s) coming under this category.

2. Sex Hormones

In Crustacea, just as among vertebrates, sexual differentiation is mediated by sex-hormones. These hormones are secreted, respectively, by the androgenic glands and the ovary; the testis has no endocrine function.

(a) *Androgenic Glands*

The androgenic glands, whose presence and function were discovered in *Orchestia gammarella*[40] are paired structures attached to the subterminal region of the vas deferens (for references[46]) in coxopodite muscles of fifth walking leg only in males. In some other pericarida, excluding amphipods, the androgenic gland shows an equally variable position along the genital tract[24]. The presence of androgenic glands is now well established in nearly all groups of the Malacostraca[43,45,46]. Although, located outside the testes, the androgenic glands belong to the reproductive system, their primordia being connected to the end of the rudimentary deferent canal. Both rudiments are present in all individuals but develop only in males[110]. Based on the arrangement of cells, four types of glands have been distinguished by Charniaux-Cotton *et al.*[47] in Malacostraca : (1) Simple cellular strands (Cords) e.g., *Carcinus*, *Cancer* (2) Compound strands, e.g., *Orchestia* (3) Lobules, e.g., *Maja*, *Eriocheir*, *Gecarcinus* and (4) Syncytial spherules, e.g., *Squilla* (for figures *see* review by Aidyodi and Adiyodi[1]. Generally the nuclei of the component cells are oval ranging from 2-10 μ in diameter depending on the species and endowed with dense chromatin. The fine structure of the androgenic gland in decapods has been studied by King[98] in *Pachygrapsus*, and by Meusy[120] in *Carcinus*. It is now well known that androgenic glands show seasonal activity in some of the species studied, being smaller in winter, the period which coincides with reproductive quiescence, but large in summer[88].

The information regarding the chemical nature and the mode of secretion of Androgenic Gland Hormone (AGH) is not yet clear. Earlier studies depending upon the ultra-structural and histochemical profile of the gland emphasized that AGH is of proteinacious or polypeptide in nature[1,36,37] and furthermore it was stated that the hormone is available only in aqueous extracts and not in lipid-solvent extracts of glands. Recently, Katakura *et al.*[97] obtained from the androgenic glands of the isopod, *Armadillidium vulgare*, a proteinic fraction of molecular weight 15000. But according to Sarojini[157] the AGH is lipidic because of its behaviour during the isolation. Berreur-Bonnenfant *et al.*[28] have purified a lipid substance from androgenic glands of *Carcinus maenas* and reported its two biological activities : (1) inhibition of vitellogenesis when injected in female *Orchestia gammarella* and (2) it is responsible for the accumulation of carotinoid pigments in the antennae when administered in *Talitrus saltator*. But more recently, Ferezou *et al.*[63] have again obtained a purified lipid fraction from the androgenic glands of the male crab, *Carcinus maenas* and supported the view of Berreur-Bonnenfant and others, reporting the lipid fraction of androgenic gland inhibits the incorporation of 3H-Leucine in ovary subcultures of *O. gammarella* and produce pink colouration of the antennae, a secondary sexual characteristic of the male *T. saltator*. Masculinization, i.e., induction of spermatogenesis, has never been obtained using this preparation. Therefore, it is postulated that the water soluble fraction might be responsible for masculinization. But our study on a marine female prawn, *Parapenaeopsis stylifera* has indicated that when these prawns were injected repeatedly with aqueous extract of androgenic glands, there was a total inhibition of ovarian growth, but no signs of masculinization of ovary or morphological changes in the secondary sexual characters were encountered[88]. Therefore, it is suggested that neither lipidic nor proteinic (water soluble) fractions are responsible for the masculinization, but possibility of a *third* fraction cannot be ruled out and needs further investigation.

The androgenic glands are not neurosecretory. The mode of secretion of AGH remains enigmatic. Charniaux-Cotton[43,45] assumes to be holocrine on the basis of her observation of "degenerate areas", having picnotic nuclei and small amounts of deeply-stained cytoplasm, in some parts of the gland. Degenerate areas have also been reported in some decapods by Charniaux-Cotton *et al.*[47] but earlier King[98] could not find them in *Pachygrapsus*. The AGH are not only responsible for the development of secondary sex

characters but also that of primary ones like the differentiation of testis, sperm duct primordia, maintenance of germinative zone, and normal spermatogenesis[25,46]. AGH is believed to act selectively on the growth of male secondary sex characters only during critical periods separating stages of their allometric growth, but the maintenance of these characters apparently requires a continued optimum titer of AGH in hemolymph[1].

(b) *Ovaries*

There is some evidence to suggest that the ovaries of crustaceans are the source of a substance or substances which may control hormonally the differentiation of secondary sexual characters and vitellogenin synthesis in amphipods and isopods[43]. But surprisingly, Katakura[96] has reported in an isopod, *Armadellidium vulgare* that the development of female sexual characters is not under the direct control of ovarian hormone. The existence of a Female Hormone (FH) by performing the extirpation and transplantation experiments, a classical practice in endocrinological studies, of ovaries in an amphipod, *Orchestia* has recently confirmed by Junera *et al.*[91] and Meusy *et al.*[121], who postulated that vitellogenin synthesis in females require an ovarian factor called "*vitellogenin stimulating ovarian hormone*" (VSOH) (Fig. 2). Earlier, Croisille *et al.*[50] reported that the experimentally-induced sex reversal in *Orchestia gammarella* indicated that the synthesis of the female specific protein (Vitellogenin protein) is under the control of ovarian hormone. It is believed that the hormone secreted by the mature ovary probably comes from the germinative zone or from young follicles[41]. But, it is not clearly shown so far, which ovarian cells, mesodermal or germinal, are responsible for VSOH secretion and the chemical nature of this active material. It would be of interest to know if VSOH is identical to the hormone which controls secondary sexual characteristics (ovigerous hairs) whose existence was established in *O. gammarella*[41].

It will be interesting to know whether the permanent female characteristics, such as oostegites, and the temporary ones such as ovigerous hairs connected with the incubation of eggs, are under the influence of different female endocrine glands as they are in an isopod, *Asellus aquaticus*, wherein the sternal glands appear to act on permanent, and the ovarian follicular cells on temporary characters[1,19]. The development of several temporary sex characters, which usually appear as the ovary matures and which are

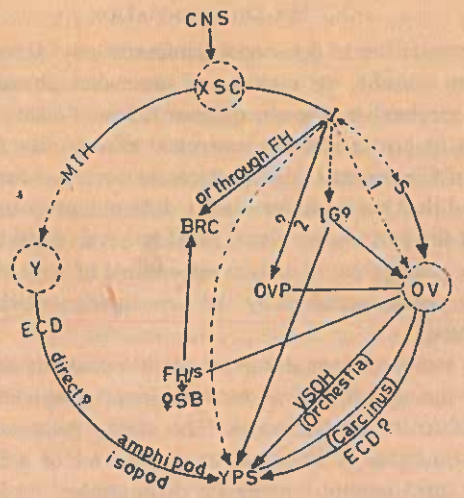


Fig. 2. Schematic diagram illustrating endocrine relations in the reproduction of decapod crustaceans. Unbroken arrows indicate stimulation and the broken ones, inhibition (from Adiyodi, 1978)

Abbreviations :

BRC—Brooding characters
CNS—Central nervous system
ECD—Ecdysteroids
FH/S—Female hormone/s
I —Inhibitory gonadotropin
IG —Intermediate gland
MIH—Moult inlubiting hormone
OV —Ovary
OVP—Oviposition
S —Stimulatory gonadotropin
♀SB —Female sexual behaviour
VSOH—Vitellogenin—stimulating ovarian hormone
XSC —X-organ sinus gland complex
Y —Y-organ
YPS —Vitellogenic protein synthesis
1 —Inhibition and phasing of vitellogenesis
2 —Stimulation during the breeding season

plausibly female hormone dependent, is however, either retarded or ceased on eyestalk ablation in *Carcinus maenas* and *Pachygrapsus marmoratus*[46,49]. If FH is strictly ovarian in origin in the Decapoda, as appears to be the case in amphipods and isopods, it is unlikely that FH plays any role in stimulating oogenesis, in contrast to AGH which promotes spermatogenesis. It is interesting that in crustaceans the testes are not known to be a source of any endocrine substance.

3. Sex Differentiation

Sexual differentiation in decapod crustaceans may increase progressively with subsequent moults, or may occur somewhat abruptly at a critical moult[174]. The mechanism whereby differentiation of male and female reproductive systems in crustaceans is controlled was poorly known until the classical work of Charniaux-Cotton, which she reviewed comprehensively in 1960. She found that the basis for sexual differentiation in males was the development of the androgenic glands, which never developed in females. Nowadays, the conception of hormonal control of sex differentiation in crustaceans is generally accepted by the investigators working on invertebrate endocrinology.

Most of the malacostracan decapods are bisexual and sex is determined genetically, but the morphological and functional expression of this genetic sex is largely influenced by hormones. The young malacostracan acquires its particular form either at hatching or at the end of a larval period in either case, the sexes cannot usually be distinguished and this asexuality may persist for several moults. External differences between the sexes may be due to the presence of special sex organs, such as the male's penis or petasma (in marine prawns) or the female's oostegites or thelycum (in marine prawns), but also to variations in the morphology of structures common to both[80]. In undifferentiated young crustacea, the anlagens of the vasa deferentia are present in both sexes. Attached near the end of each vas deferens in higher crustaceans and to the cephalic end of each testis in isopods, is a gland known, because of its function, as the androgenic gland, which lies in the first and second thoracic segments in isopods and between the muscles of the coxopodites of the last pair of walking legs in higher crustaceans. In a genetic female, the androgenic glands fail to develop, but in the male, the glands enlarge to form solid strands of cells, folded several times. It will be of interest to know that how the androgenic glands are involved in the control of sexual differentiation.

(a) *Male Differentiation*

The androgenic glands are known to be the exclusive source of a hormone, responsible for the differentiation and development of all male primary (testis) and secondary (mating behaviour for example) sexual characters. Transplantation of testis or of a portion of the vas deferens without androgenic gland into a female is ineffective, so far there is no

evidence that testes secrete any hormone at all, but grafting an androgenic gland masculinizes the ovary and the secondary sexual characters[64, 80, 96]. The ovaries become transformed into testes, the primary germ cells give rise to spermatocytes, spermatids and spermatozoa. The primordium of the vas deferens, present in the female grows and genital papillae appear, and the secondary sexual characteristics gradually aquire male form. If the female is ovariectomized (this operation could be possible only in amphipods and isopods so far), before implanting androgenic gland, external masculinization still takes place, and this proved that the androgenic gland acts directly on the external characters. The masculinized females have a male sexual behaviour[46]. On the contrary, removal of the androgenic glands from an adult male resulted in degeneration of the testes. In *Orchestia*, when the androgenic glands are removed from a young male, the primary germ cells in the testis produce not spermatocytes but oocytes. Therefore, it is suggested that a hormone from the androgenic glands is necessary for the normal development of the testes, in its absence, the primary germ cells develop spontaneously into oocytes[80].

(b) *Female Differentiation*

In genetic females, the androgenic glands do not develop. But when androgenic glands from a young male are transplanted into female, the ovaries rapidly transform into testes, the primary germ cells giving rise to spermatocytes, which develop fully through spermatids to spermatozoa. The rudimentary vasa deferentia in the female also grow and develop. The androgenic gland hormone can act in the female as it does in the male and moreover, the hormone is not species-specific[80].

It is ambiguous so far, whether ovarian hormone has any role to play in sex differentiation or not and need further investigation.

The genetic determination of the primary sex characters like testis and ovary is expressed through the androgenic gland. But it is of profound interest to know that: (1) how the secondary sexual characteristics develop, and (2) are they determined genetically, or are they also controlled hormonally. The male secondary sex characters are induced by the direct action of androgenic hormones upon the tissues. When the androgenic glands are removed from a male *Orchestia*, and at the same time the appendages are amputated, new appendages regenerate without the typical

male form, as expected. But the new appendages also do not have acquired the female form either they are undifferentiated rather than being male or female. Therefore, it is speculated that in the female, where the androgenic glands are undeveloped, the sex characters must be determined by something other than the mere absence of androgenic hormone[80]. The androgenic hormone not only induces male differentiation, but inhibits ovarian controlled female characters[46]. After amputation of a thoracic leg in normal or castrated female *Orchestia*, regeneration of the leg with an oostegite occurs, but in masculinized females the oostegite does not regenerate (Charniaux-Cotton, 1961); it is interesting to note that in certain intersexes of *Orchestia*, genetic maturation can transform these inhibitory properties of the androgenic gland into inductive properties[45].

Females of *Orchestia* possess permanent secondary sex characters, the oostegites which form a brood pouch for the young, as well as temporary characters, ovigerous hairs which appear on the oostegite during a moult preceding the egg laying, replacing the short juvenile hairs of immature forms, or of females in sexual repose. Oostegites appear in males whose androgenic glands have been removed and implanted with ovarian tissue from a young or a maturing female[41]. This permanent female secondary sex character is thus induced by an ovarian hormone. When an ovary is implanted into an ovariectomized female, yolk deposition in the oocytes begins after the first post-operative moult, and ovigerous hairs are developed after the second post-operative moult. Therefore, it is suggested that both, the temporary and permanent, secondary sexual characteristics are not determined genetically but are controlled by an ovarian hormone[109]. Our knowledge of the hormonal control of the development of female secondary sex characters in higher decapods is so fragmentary that very little can be said regarding the possible relationship between these characters and stimulatory gonadotropin, though the antagonism between inhibitory and stimulatory gonadotropins may perhaps suggest a brood inhibiting role for stimulatory gonadotropin. The problem of the involvement of stimulatory gonadotropins in sexual differentiation in decapods is likely to prove interesting in view of the fact that parasitic castration in crabs may involve destruction of the thoracic ganglion by the root-system of the parasites[1].

4. Regulation of Gonadal Activity

(a) *Relation between Moult Inhibiting Hormone (MIH) and Inhibitory Gonadotropin*

Reproductive activity and growth in decapod Crustacea are two antagonistic phenomena, as they are in the polychaete worms. Ovigerous shrimps and crabs will not moult until after the eggs have hatched. The problem whether MIH and inhibitory gonadotropin actually represent a single hormone (Panouse, 1947) "growth-inhibiting hormone" affecting two diverse, but integrated functions known as somatic (regenerative) growth and reproductive (genital) growth, or are alternatively distinct hormones has been debated[144]. Female reproduction, comprising various phenomena like egg maturation, ovulation (release of eggs from the follicles), spawning (oviposition), incubation of eggs within the pleopods of the mother and lastly the release of the larvae or young ones, is completed during the intermoult in higher decapods. Nevertheless, Cheung[48] and Adiyodi[4] have observed an interesting and a very rare phenomenon in nature that the moulting crabs were bearing growing ovaries. Spawning follows ecdysis in some of the Natantia which are also remarkable for their short intermoult and relatively very long premoult. In the Caridean prawn, *Palaemon serratus* each spawning is preceded by a moult[148]. These instances of apparently synchronous occurrence of moulting and reproduction, together with the observation that, besides the maturation of the ovary, there occurs a shortening of the intermoult, with an exception of *P. serratus*[60], suggest the possibility that the hormonal mechanism involved in moulting and reproductive activities are the same, or if they are different, they act synergistically.

Moulting cycle of two types have been distinguished among decapod crustaceans : (1) *Diecdysis*, in which the individual moults throughout the year (e.g , *Palaemon, Macrobrachium, Caridina, Lysmata*) or several times during a season (e.g., *Pachygrapsus*) and (2) *Anecdysis*, in which the animals have a long intermoult period and they moult on a seasonal or annual basis (e.g , *Cambarus, Uca, Paratelphusa*). The crustaceans attained maximum size have no need for further growth and may cease moulting permanently, which can be termed as *permanent* or *terminal anecdysis* (e.g., *Carcinus, Callinectes, Maja*). When *Carcinus* is in stage of terminal anecdysis the Y-organs, though reduced, are functional and moulting may be induced

by eyestalk ablation. What restrains the Y-organs from functioning in terminal anecdysis stage may, therefore, be a heightened activity of MIH circulating in hemolymph[85]. If it is the same hormone that inhibits both the gonad maturation and moulting, it is not clear how reproduction in the terminal anecdysis crustacean can be explained[1]. Drach[60] has reported a difference in temperature coefficients for moult initiation and stimulation of ovarian development but later on Passano[144] has argued for a inhibitory gonadotropin which is distinct from MIH. The antagonism between reproductive and proecdysial events has been shown by a large number of crustacean species (*see* Sect. I of review by Adiyodi and Adiyodi, 1970a for details). Adiyodi[4] has reported that the reproduction and moulting are under different hormonal controls in a freshwater crab, *Paratelphusa hydrodromous*, since the normal relationships between the moulting stages and ovarian growth could be changed experimentally and further stated that reproduction in *Paratelphusa* occurs normally in stage C, but in starved animals it persists into stage D. The physiological and histological evidence suggests a high titer of MIH and a comparatively low titer of inhibitory gonadotropin in the hemolymph during intermoults in reproducing individuals[3]. The high concentration of MIH is essential for reproduction, which it facilitates indirectly by restraining the Y-organ. Post-moult stage A and B in normal females of *Paratelphusa* and also stage C in non-breeding adults are rich in MIH and inhibitory gonadotropin[4]. This, together with the observation that in starved females of *Paratelphusa* ovarian growth continues even in premoult, indicates that MIH and inhibitory gonadotropin could be synergistically active and inactive under both normal and altered conditions.

Naturally, in young juvenile crustaceans the emphasis is on somatic growth and the gonads hardly being capable of any positive growth even under the stimulation of eyestalk ablation. Reproductive growth starts usually with puberty ; in some crustaceans gonadal growth can be induced even in prepuberty. Several decapods continue both to grow and reproduce in the adult state. Demeusy[56] has admirably elucidated the relation between somatic and reproductive growths in *Carcinus maenas*. Oocyte growth in decapods appears to be generally accomplished in two phases ; a primary phase by the close of which oocytes have grown to a size that permits yolk deposition, and secondary phase during which yolk appears. In species such as *Carcinus maenas*[56] and *Paratelphusa*[5], the second phase is completed

in two stages, the first and second vitellogenesis. In normal young adult females an autotomized limb undergoes basal-limb growth during the period of first vitellogenesis, but further growth of the regenerate occurs only during premoult, i.e., after reproduction has been completed. Though, regeneration and reproduction may take place simultaneously in eyestalkless young adult females of *Carcinus* there is some indication that the balance is slightly in favour of genital growth during the second vitellogenesis. However, eyestalk ablation at the end of the basal-limb growth allows premoult growth and second vitellogenesis, which are antagonistic events, to occur simultaneously[1]. Therefore, it can be believed that this process may be related to the abrupt elimination of the sources of MIH and inhibitory gonadotropin at a critical period, when both the gonad and regenerate have reached a stage allowing further growth and differentiation under withdrawal of the inhibitory principles.

(b) *Inhibitory Gonadotropin*

(i) *Source*

In many decapod crustaceans, the gonad maturation, both ovary and testis, has been shown to be profoundly influenced by the inhibitory gonadotropin produced from eyestalks (Figure 3). In the early half of the twentieth century, Panouse[142] in a shrimp, *Leander*, Brown and Jones in a crayfish, *Cambarus*[31] and a fiddler crab, *Uca*[32] and Takewaki and Yamamoto[171, 172] in a shrimp, *Paratya* have evidenced, for the first time, the existence of an ovary inhibiting hormone in the eyestalks of these crustaceans. Later on the above studies were extensively elaborated in several other crustaceans, such as *Pandalus gracilis* (*Kessleri*)[15], *Carcinus maenas*[56], *Scylla serrata*[80, 87, 151], *Crangon crangon*[26], *Penaeus japonicus*[107], *Penaeus orientalis*[16], *Homarus americanus*[123], *Uca pugilator*[196], *Penaeus monodon*[77, 150] and *Parapenaeopsis hardwickii*[127]. Demeusy[54] has reported the existence of a testis inhibiting hormone in the eyestalks of *Carcinus maenas*, but Carlisle[34] found no effect of eyestalk ablation on the testis weight and development of male genital ducts in *Lysmata seticaudata*. But later on Demeusy's, results were strengthened by Otsu[140] in *Potamon dehaani*, Gomez[78] in *Paratelphusa hydrodromous*, Rangnekar *et al.*[152] in *Scylla serrata*, Ryckaert *et al.*[154] in *Palaemon serratus*, Shirwadkar[164] in *Caridina rajadhari* and Kulkarni *et al.*[103] in *Parapenaecopsis hardwickii*.

Rangnekar *et al.*[152] in *S. serrata*, *payen et al.*[146] in *Rhithropanopeus harrissi* and *Callinectes sapidus* and Touir[189] in *L. seticaudata* have emphasized that in these crustaceans the eyestalk extirpation led to an increase in the activity of androgenic glands as evidenced by their hypertrophy, essentially due to hyperplasia and the increased spermatogenetic activities seemed to be a consequence of the hyperactivity of androgenic glands. The precocious maturation of the gonads after eyestalk ablation is prevented either completely as found in majority of the decapods, or incompletely as in the crab, *Scylla serrata*[151], if sinus gland alone is implanted or an extract of complete eyestalk is administered. This obviously suggested that in the decapod crustaceans, the gonad maturation and function appear to be regulated by the inhibitory gonadotropin produced in the X-organ-sinus gland complex in eyestalks, and is stored in the sinus gland. This structure is formed of swollen nerve terminals whose cell bodies are located in different regions within the central nervous system. Though cyclic histological changes have been described in the sinus gland in some crustaceans (Gabe, 1966 for referenees), there is very less evidence in the literature to suggest that the inhibitory gonadotropin is secreted by the sinus gland itself. Inhibitory gonadotropin like MIH is belived to be initially elaborated in the neurosecretory cells of the ganglionic X-organ and transported intra-axonally along the X-organ-sinus-gland tract to the endings in the sinus gland for storage and release[144]. Recently, Shivers[165] and Andrew and Shivers[18] have reported in detail regarding the mechanism of release of neurosecretion from the sinus glands in crayfish, *Orconectes virilis*.

In brachyuran crustaceans the main neurosecretory cells in the ganglionic part of the X-organ, despite the various names assigned to them, may perhaps all correspond to the β-cells described by Enami[61] and Matsumoto[117]. These are generally polygonal and unipolar with uni- or bi-nucleolated nucleus. When the cells are active the cytoplasm contains granules positive to Aldehyde Fuchsin (AF) and Chromealum-Haematoxylin-Phloxine (CHP). The secretory products appear to be transported along the axons and these constitute the X-organ-sinus-gland tract which stains a deep lilac with AF in sections, whereas the other neurosecretory pathways in the eyestalk, brain and thoracic ganglia take up only a pink colour. The distinctive staining shown by the β-cells of the X-organ has been considered as evidence for the different nature of their secretory products, and in fact, these cells have been considered to be the sources of both MIH and

inhibitory gonadotropin[117]. Histomorphological variations in the activity of the X-organ-sinus-gland complex, correlated with seasonal activity of the gonad, have been described by Aoto and Nishida[15] in the shrimp, *Pandalus gracilis* (*Kessleri*) and by Hanaoka and Otsu[79] in the freshwater crab, *potamon dehaani*. It is possible that under the influence of different incoming signals the β-cells in the X-organ of *Potamon* can produce inhibitory gonadotropin as well as MIH which are perhaps chemically related molecules. The β-cells in the X-organ may well be the source of inhibitory gonadotropin in *Potamon* in particular[117], and in decapods in general, but the cytological evidence for this is unfortunately very fragmentary so far. Further, investigations are needed to show whether the inhibitory gonadotropin can be elaborated by all or by only few specific types of cells present in the X-organ. Adiyodi's observations[8] on *Paratelphusa* suggested the possibility of the involvement of eyestalk e-cells in the production of inhibitory gonadotropin. Surprisingly enough, recently, Demassieux and Balesdent[52, 53] have reported the cyclic variation in histological profile of β-type neurosecretory cells of cerebral and ventral cord ganglion in an isopod, *Asellus aquaticus*, in relation to reproduction and suggested that β-cells secretion has an inhibitory effect on vitellogenesis. Kurup[105] in an anomuran, *Petrolisthes linctipes* has reported that the Z-organs supposedly play the role of the X-organs in moulting and have also been found to secrete a principle controlling oocyte maturation. Each Z-organ is composed of a number of bunches of elliptical or elongated "neurosecretory" cells aggregated in the form of follicles.

(ii) *Variations in Secretion*

Otsu[140] is of the opinion that inhibitory gonadotropin is present not only in adults, but in immature stages too ; since in eyestalk of immature crabs, *Potamon dehaani* of both sexes have found to contain the hormone. When the vitellogenesis is already in full swing, eyestalk extirpation does not perceptibly accelerate ovarian growth suggesting that during this period the synthesis of inhibitory gonadotropin and/or its release from the eyestalk into the general circulation, may be at low ebb. From a survey of the literature on the reproductive and also moulting phenomena in various decapod crustaceans it appears that the synthesis and/or release of inhibitory gonadotropin, like some of the other eyestalk hormones, may be cyclic in both sexes. Few earlier experiments involving eyestalk ablation

have revealed that in some phases of the female reproductive cycle this operation fails to accelerate vitellogenesis[4, 115, 195]. Therefore, it is speculated that some of the activities, such as synthesis and release of inhibitory gonadotropins and response of gonads to these hormones appear to be cyclic in crustaceans. Many investigators (*see* review by Adiyodi and Adiyodi[1]) have indicated the gonad maturation and, therefore, possibly secretion of inhibitory gonadotropin, like the secretion of MIH[10], may be under the influence of such factors as the season, photoperiod, thermal fluxes, nutrition, salinity and other features in the external environment acting through sensory receptors. Of the two phases in the growth of the oocytes, inhibitory gonadotropin appears to inhibit only the secondary phase and probably has no effect on the primary growth[46], the latter being under the control of Moult Hormone (MH). But Adiyodi and Adiyodi[1] have evidenced that in *Paratelphusa*, the inhibitory gonadotropin principally inhibits the first vitellogenesis ; eyestalk ablation greatly accelerates the normally long first vitellogenesis in these crabs, but its effect on the second vitellogenesis is far from impressive. Klek-Kawinska and Bomirski[99] have reported on the variation in the titre and activity of Ovary Inhibiting Hormone (OIH) particularly, in the eyestalks of female shrimps, *Crangon crangon* during the course of the annual reproductive cycle. In the early part of the breeding season, after the first oviposition, OIH activity was apparently absent. During the latter part of the breeding season, after the second oviposition, it is present, but at a much lower level than after the breeding season when its activity was found highest. Recently, we have reported, more concisely, on the role of OIH from eyestalks of marine female prawns, *Parapenaeopsis hardwickii*, during ovarian developmental cycle[101]. The OIH activity was found to be very high in the immature (Stage I) and post-spawned phases of ovarian development. During the phase of full vitellogenesis (Stage IV), the inhibitory activity was negligible. After spawning (Stage V) it was present, but at a lower level than during inactive and post-spawned phases. Therefore, it is suggested that the ovary inhibiting hormones, in both *C. crangon* and *P. hardwickii* appear to be responsible for the phase of ovarian quiescence and to control the rate of vitellogenesis. Unfortunately, no study has been done so far in this regard in male prawns.

(C) *Stimulatory Gonadotropin Source and Function*

More valid support for the presence of a stimulatory gonadotropin, however, comes from the analysis of the hormonal situation in those decapods which respond by precocious spermatogenesis or vitellogenesis to eyestalk extirpation.

(a) *Males*

While investigating the role of central nervous system in reproduction of crustaceans, Gomez[78] found that the implantation of the brain and the thoracic ganglion into males of *Paratelphusa hydrodromous* resulted in a precocious maturation of the testis and an eventual hypertrophy of the vasa deferentia (Fig. 3). This observation together with that of Otsu[140] that the

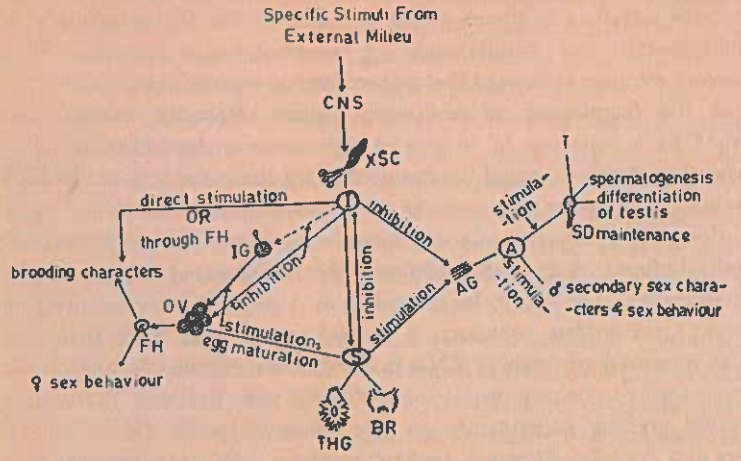


Fig. 3 : Diagram showing the possible relationships involved in crustacean reproduction between the inhibitory and stimulatory gonadotropic principles and the sex hormones. (From Adiyodi and Adiyodi, 1970)

Abbreviations

A—Androgenic hormone
AG—Androgenic gland
BR—Brain
CNS—Central nervous system
FH—Female hormone
I—Inhibitory gonadotropin
IG—Intermediate gland
O—Ovary
S—Stimulatory gonadotropin
SD—Sperm duct
T—Testis
ThG—Thoracic ganglion
XSC—X-organ sinus gland complex

thoracic ganglion of adult males of *Sesarma* is effective in accelerating ovarian development in young females of *Potamon*, suggested that a stimulatory gonadotropin, which is effective in both sexes and in different genera, is present in the thoracic ganglion and perhaps also in the brain of male crabs. Berreur-Bonnenfant[21] has reported in *Orchestia gammarella* that a cervical factor of the male brain is necessary for the proliferation of testicular germinative zone. But Rangnekar *et al.*[152] have not found any effect of injection of extracts of thoracic ganglion and brain on the spermatogenetic activity of *Scylla serrata*. Adiyodi and Adiyodi[2] have found in *P. hydrodromous* that the B-cells of implanted brains undergo a secretory cycle and suggested that gonadal acceleration resulting from brain implantation may be related to some gonad stimulating factor released from the brain neurosecretory cells. Recently, Touir[190] in *Leander serratus* and *Crangon crangon* reported that the median protocerebrum of these crustaceans secretes a hormone which is necessary for the stimulation of spermatogenesis and maintenance of spermatogonia. Juchault[89] in *Sphaeroma serratum* evidenced that a neurosecretory centre of protocerebrum control the functioning of androgenic gland. Recently, Amato and Payen[12] have reported in a crayfish, *Pontastacus leptodactylus leptodactylus* that the mesodermal tissue surrounding the primary gonia inhibits the initiation of spermatogenesis and its disappearance led to the transformation of the primary gonia into secondary gonia and this can be experimentally induced by eyestalk ablation. The initiation of the meiosis of spermatogenesis appeared to be controlled by a male factor originating in the brain ; its complete achievement can take place only when there has been an increased synthesis of RNA in the secondary gonia ; the androgenic hormone seems to be responsible for mediating this. But after performing the cross grafting experiments on two oniscoid species (*Armadillidium vulgare* and *Porcellio dilatatus*), Juchault and Legrand[90] drew the conclusion that an androgenic gland introduced into a heterospecific receiver hypertrophies, but remain functional. A simultaneous introduction of protocerebrum of the same species on the androgenic gland prevents this hypertrophy and slows down the synthesis of the male hormone. The grafting of ventral nerve chain also inhibits the synthesis but does not prevent hypertrophy. Therefore, it was suggested that two types of specific neurohormones are present, one controlling the growth of the androgenic gland, and the other, the synthesis of the male hormone. The former comes from the

protocerebrum and the latter from the protocerebrum and the ventral nerve chain and speculated that a third nonspecific neurohormone may control the release of the androgenic hormone. We found in the prawn, *Parapenaeopsis hardwickii* that injection of brain and thoracic ganglion extracts induced hypertrophy of the androgenic gland, essentially due to hyperplasia, and consequence of which was viewed in the acceleration of spermatogenesis, both in eyestalkless and normal individuals[108]. Therefore, it was suggested that the stimulatory gonadotropins were elaborated in the brain and thoracic ganglion neurosecretory complex. So far, we are not knowing the specific cell types synthesizing these hormones either in brain or in thoracic ganglion.

(b) *Female*

When studying the role of central nervous system in reproduction of crustaceans, Otsu[136, 140, 141] found that the repeated implantation or injection of thoracic ganglion into the sexually inactive female crab, *Potamon dehaani* resulted in considerable ovarian growth. Gomez[73] and Gomez and Nayar[74] have also evidenced that besides the thoracic ganglion, the brain also secretes an ovarian growth accelerating hormone in the crab *Paratelphusa hydrodromous*, but other parts of the nervous system appear to be ineffective. By these procedures gonadal development could be accelerated to nearly the same extent as with eyestalk removal, and it is therefore possible that stimulatory gonadotropin must be originating in the central nervous system. It is debatable whether the gonad accelerating effect found in crabs implanted with thoracic ganglion and brain is caused by neuroendocrine factors contained therein, or is only one of the general metabolic or pharmacological effects exerted by central nervous system extracts. The former appears to be more likely in view of the histomorphological characteristics of these organs. Perryman[147] has reported three endocrine factors influencing the ovarian growth in malacostracans ; two from the eyestalks and one from the supraoesophageal ganglion and stated that these factors, in common with several other crustacean hormones, are probably neurosecretory. The brain and the thoracic ganglion in decapods are endowed with various types of neurosecretory cells assembled into well defined groups[8, 59, 105, 116, 117, 118].

Matsumoto[117] described three cell types (A, B and E) in the brain and five cell types (A, A', B, B' and C) in the thoracic ganglion of the crabs,

Eriocheir japonicus, *Potamon dehaani*, *Sesarma intermedia*, *Neptunus tritubercularus* and *Chionoecetes opilio*. The A, A′ and E cell appeared to be maximally active in these crabs when the ovaries were generally in their early stages of maturation and ovulation was coincided with active secretion in A cells but A′ and E cells contain no stainable secretory granules. A cell secretion was in the form of granules positive to CHP and positive or negative to AF. Matsumoto[118] concluded that the secretory products of A and E-cells in the brain and A and A′ cells in the thoracic ganglion of crabs may contain principles closely related to ovarian development, and that the product of the A-cells may further be concerned with ovulation. Kurup[105] observed that both the neurosecretory cell types in the brain, namely the elliptical A-cells and the small almost round B-cells, have found to function with the principle of the eyestalk Z-organs in the maturation of the eggs, but the details supporting this statement is lacking. Adiyodi and Adiyodi[2] reported that in *Paratelphusa*, implanted brain, though naturally isolated, functions apparently normally for a time. In such tissues retrieved 2 weeks after implantation, the various neurosecretory cells, their axonal tract, and the glial cells responsible for gliosecretion still preserved their normal cytological appearance ; the B-cell groups even showed an axonal discharge of their CHP-positive granular secretion. Perryman[147] has examined the sections of the brain of a crayfish, *Procambarus simulans* and correlated the stages of ovarian development with the varying amounts of neurosecretory material in the various cell types and she postulated that cell type three produced granules responsible to hasten ovulation. Larger amounts of granular material were present when the ovaries composed a greater percentage of the body weight. This was the stage at which the greatest amount of yolk was deposited in the ovaries. Diwan and Nagabhushanam[59] described four types of neurosecretory cells (A, B, C and D) in the brain and thoracic ganglion of a freshwater crab, *Barytelphusa cunicularis*. They reported that the number of A-cells showing secretory material in the perikarya was more from March through June, the period in which the gonads of the crabs begin to mature. The secretion of the A-cells was remarkably high in April and May and during the breeding season. It has been found that the spawning season of *B. cunicularis* starts in June and extends up to September. Majority of the crabs have spent gonads from October to February. During post-spawning period most of the cells of A-type did not contain secretory material. This correspondence

between the secretion of the A-cells and gonadal activity is indicative of the role of brain and thoracic ganglion neurohormone in gametogenesis. However, the variation in the secretory activity of B-cells was not in correlative correspondence with gametogenesis. We have reported in a marine prawn, *Parapenaeopsis hardwickii* that injections of brain and thoracic ganglion extracts into both eyestalkless and normal individuals led to the acceleration of ovarian maturation[127]. Similar observations were reported earlier by Nagabhushanam and Diwan[124] in *B. cunicularis* and suggested that supraoesophageal and thoracic ganglia contain a gonad stimulating factor.

Passano[144] was of the opinion that the Y-organ which usually initiates the moulting activities appears to be inhibited in the ovigerous state by neurosecretory cells of the thoracic ganglion, based on cues received from sensory elements in the pleopods which hold the eggs in place. The role of stimulatory gonadotropin, therefore, appears to be dual in that it accelerates oocyte maturation and inhibits Y-organ activity ; the latter is accomplished either directly or indirectly by raising the level of MIH and/or lowering the level of inhibitory gonadotropin (cf. Adiyodi and Adiyodi[1]). More detail and critical investigation is warranted to confirm the claim of Otsu[140] that stimulatory gonadotropin is produced even during the sexually quiescent periods in adult females, and that the suppression of the ovary during the nonbreeding season is due to inhibitory gonadotropin and not owing to the lack of stimulatory factor. The observations of Matsumoto[117] and Diwan and Nagabhushanam[59], may be considered here, which argue for a seasonal activity of the thoracic ganglion. The question why stimulatory gonadotropin should be synthesized and stored in two separate ganglionic centres is yet not answered, nor it is known whether the factor or factors produced by either of them solely triggers the secretion of the other. It would be of great interest to know how extirpation of thoracic ganglion affect reproduction and the secretion of stimulatory gonadotropin by the brain in normal and eyestalkless females during the breeding and nonbreeding seasons.

(D) *Functional Significance of Y-organ Hormones*

The Y-organs which are paired bodies were shown by Gabe[67] to be characteristic of Malacostracan Crustaceans and declared from histological study with the light microscope that these glands have a role in moulting

and growth. The Y-organs are not neurosecretory structures. The Y-organ is innervated from the suboesophageal ganglion, but there is no evidence either at the light microscopic or ultrastructural level for an intraaxonal passage of the neurosecretory material along the ecdysial nerves' to the gland. Conceivably the Y-organ is regulated by neuroendocrine signals transmitted through vascular channels[1]. They are presumed to be the sources of the Moulting Hormones (MH) which are a group of steroids, cumulatively known as ecdysones, effective in diverse classes of arthropods[160]. Knowles[100] published some preliminary observations of the ultrastructures of the Y-organ which indicated that electron-dense inclusions pass out of the nucleus into the surrounding cytoplasm. The ultrastructural machinery of the Y-organ is suggestive of an organ engaged in the formation of a steroid from a precursor like cholesterol, an observation in keeping with the known chemical structure of both crustacean and insect MH. In Crustaceans, MH may comprise different molecules like crustecdysone (20-hydroxyecdysone) and the deoxycrustecdysone (*see* Highnam and Hill[80]). MH is primarily responsible for initiating premoult, but it also influences other functions like the hardening of the cuticle, colour change, regeneration and gonadal maturation[1].

It is now well established that the gonads need to reach a particular stage of differentiation before they respond to the waning titer of the circulating inhibitory gonadotropin. MH has been shown to be essential for the initial stages of gonad maturation, namely oogonial[17, 56] and spermatogonial mitoses[18]. Charniaux-Cotton and Kleinholz[46] are of the opinion that spermiogenesis and vitellogenesis, which do not involve mitosis, can proceed in the gonad even in the absence of MH. It is reported that when reproduction generally occurs in decapods, the Y-organ does not store MH[94] and that, histologically, the gland is almost inactive[118]. In *Maja*, the Y-organ regresses simultaneously with the pubertal moult[85]. Whereas *Carcinus* devoid of Y-organs may ovulate and carry embroys[17, 144]. These observations indicate that MH is not essential for reproduction, but only a high titre of MIH. It was thought that this effect may be related to the removal of a specific vitellogenesis-inhibitory factor associated with the Y-organ. The possibility of that the Y-organ secretes a separate positive gonadotropic principle in prepubertal stages was thought of remote but recently at least in one crustacean, *Scylla serrata*, Farooqui and Nagabhushanam[69] have studied changes in the histological profile of

Y-organ during different phases of ovarian maturation. They found an increased synthesis of stainable material in the Y-organ cells in the immature crabs as revealed by their increased nuclear diameters and simultaneously, there was an enhancement in the transport of the stainable secretory material as evidenced by the less cytoplasm in the perikarya. However, Y-organ cells of crabs bearing maturing and mature oocytes were not active as judged by their small nuclei and less cytoplasm. They have postulated that the Y-organ is necessary for the ovarian development in *S. serrata*. Earlier, Vernet *et al.*[191] emphasized that reproduction hastens in the Y-organ less mature *Carcinus maenas*, but Y-organ ablation causes degeneration of gonads in immature crabs. Meusy *et al.*[121] have evidenced that the cauterization of Y-organs in *Orchestia gammarella* had followed by a depression of vitellogenin synthesis. Further, critical investigation is yet required to delineate the exact role of Y-organ hormone in reproduction of crustaceans.

(E) *Sex Reversal and Hermaphroditism*

Following the discovery by Charniaux-Cotton[40] of the androgenic gland in the amphipod, *Orchestia gammarella*, experiments on sex reversal by way of the extirpation of this gland from the males or its implantation into the female have been made on various forms of Malacostraca[96]. The location of the secretory cells as well as the histological structure of the androgenic gland is different in each form so far observed but the results of its extirpation from male or the implantation into female indicated unequivocally that the androgenic gland secretes a substance(s) which produces the male differentiation of the reproductive system including both the gonads and the secondary sexual features. Several decapod crustaceans, like the prawns, *Lysmata* and *Pandalus* and the two isopod groups Cymothoidae and Cryptoniscidae are functional protandric hermaphrodites. The androgenic glands are present during the male phase and disappear before the female phase is initiated (Charniaux-Cotton and Kleinholz[64] for references). Vitellogenesis does not occur in the ovarian region of the gonad during the male phase, even after eyestalk ablation[15], an expected result since androgenic hormone inhibits vitellogenesis. Disappearance of the androgenic glands suffices to explain sex reversal[42] for it allows complete transformation of the gonad into ovary and the appearance of external female characters ; vitellogenesis, no longer inhibited by the androgenic hormone, can

initiate if the eyestalks have ceased producing the OIH. It is of interest to know whether the disappearance of the androgenic glands is controlled hormonally ? Eyestalk ablation during the male phase in *Lysmata* and *Pandalus* is effectless[37] and the androgenic glands are fully active even at the second post-operative moult and claimed that a hormone from the eyestalk inhibits sex reversal and that this hormone is the OIH ; destalking of *Lysmata* and *Pandalus* close to their sex-reversal stage results in a higher rate of reversal in the operated than in the control. This difference in rate, obviously suggest that eyestalk ablation hastens the already initiated degeneration of the androgenic gland by accelerating their secretion[46]. Therefore, it looks probable that the disappearance of the androgenic glands may not be hormonally controlled. Females no longer possess inhibitor of the androgenic glands, since a gland implanted into female *Lysmata*[44] and *Armadillidium vulgare*[95] survives and masculinizes the host. It is probable that androgenic glands disappear when regeneration no longer succeeds in counterbalancing their degeneration[43]. Protandric hermaphroditism or at least the tendencies to sex reversal may be wide spread in crustaceans. This is borne out by recent finding on inter-sexuality and sex reversal in a mole crab, *Emerita* (where such phenomena are hitherto not reported) and the possibility of sexual inversions in non-hermaphroditic prawns by transplantation of androgenic gland in the opposite sex[170]. Nagamine and Knight[185,186] have, for the first time, succeeded in inducing the feminization by androgenic gland ablation (Andrectomy) in males and masculinization of females by implanting androgenic glands, in a non-hermaphroditic decapod, *Macrobrachium rosenbergii*. They suggested an inverse relationship between degree of feminization and age at andrectomy in males and degree of masculinization and age at implantation in females. It is hypothesized on the basis of above results that in Decapoda determination of the genetic sex is correlated with length of larval development, with those with extended larval development ; e.g., *M. rosenbergii*, being determined at a slower rate. Induction of sexual transformation by andrectomy or androgenic gland implantation is thus comparatively easy.

Yaldwyn[198] has thoroughly reviewed the incidences of protandric hermaphroditism in decapod crustaceans belonging to the families Pandalidae, Hippolytidae and Campylonotidae. Of the twelve species described, information on endocrine mechanisms controlling reproduction

is available only in four, three belonging to the family Pandalidae, viz. *Pandalus borealis*, *P. gracilis* and *P. platyceros* and one belonging to the family Hippolytidae like *Lysmata seticaudata*[1]. Recently, Thampy and John[188] have reported an incidence of protandric hermaphroditism in a fish parasite isopod, *Irona far* belonging to the family Cymathoidae, and studied the mode of hormone secretion in the androgenic glands. Brockenbrough and Hoffman[60] studied the effect of eyestalk ablation and injection of β-ecdysone on RNA synthesis in the androgenic glands of the protandric shrimp, *Pandalus platyceros*. Injection of eyestalk extract into the hermaphrodite crustaceans at about the time of sex reversal inhibits the growth of the ovary and suppresses the rate of sex transformation, contradictorily eyestalk ablation accelerates ovarian maturation and sex reversal[33]. With the discovery of the androgenic gland, it has become obvious that the factor suppressed the ovarian growth in the male phase of protandric hermaphrodites is not the inhibitory gonadotropin, as presumed by Carlisle[33] earlier, but is androgenic hormone. This hormone promotes testicular development, but inhibits vitellogenesis in the ovarian part of the gonad even after eyestalk removal[46]. The androgenic gland disappears at the end of the male phase in *L. seticaudata*[42], *Pandalus borealis*[36, 37] and *P. platyceros*[92] ushering in sex reversal. Sex reversal is delayed if the androgenic gland is implanted into males about to change their sex[48]. In hermaphrodite crustacea the ovary, released from the inhibition of AGH, may undergo development during the female phase at periods of low titer of inhibitory gonadotropin in the blood. Implantation of the androgenic gland causes masculinization of external sex characters during the female phase in protandric hermaphrodites[1, 145]. Maintenance of both the male gonads and the androgenic gland is known to be under the inhibitory control of inhibitory gonadotropin in crabs ; it is perhaps not illogical to consider that the same situation probably holds good in the protandric hermaphrodites, the titer of inhibitory gonadotropin being low during seasons of peak reproductive activity in the male phase and high during non-breeding months[1]. Such a view was supported earlier by Hoffman[82] in *Pandalus platyceros*. A threshold of inhibitory gonadotropin appeared to be necessary for maintenance of the normal androgenic gland, as it undergoes what appears to be non-functional hypertrophy in the absence of the eyestalks ; sex reversal results after eyestalk extirpation and can be suppressed by replacement therapy. But in *Lysmata* it is suggested

that the maintenance of androgenic gland is not solely dependent upon the inhibitory gonadotropin, moreover, it is controlled by some other hormone, plausibly stimulatory gonadotropin. Atrophy of the androgenic gland at the end of the male phase may be related to a disturbance in the balance between inhibitory and stimulatory gonadotropins[1].

5. Mobilization of Organic Reserves During Reproduction and Hormonal Control

It is now well established that the important organic materials, which are used as energy reserves in many crustaceans are glycogen and lipid[71,88,187] which are stored mainly in the midgut gland. Accumulation of these organic reserves in the storage depots is characteristic of the intermoult stage, i.e., C_4 in decapods. The reserves so accumulated are diverted either for the building up of a new exoskeleton through moulting or for reproduction depending upon the prevalent shift in hormonal milieu. These organic substances may be mobilized from the midgut gland to the gonads during their maturation, particularly glycogen, which can be utilized in the formation of gonadal products[101,102,103,104,127]. Fluctuations in the biochemical composition are also reported in freshwater and marine crustaceans, particularly on brachyuarans[5,59] and anomurans[192,193], in relation to reproductive activities. Many authors have described the annual reproductive and nutritional cycles of marine prawns[148,149,160,164]. Rice and Armitage[153] reported that in crayfish, *Orconectes nais*, exposed to short and normal photoperiods, lipid reserves are apparently shifted to the ovaries during ovarian development. We have found in *Parapenaeopsis hardwickii* that there was a continuous decline in the levels of protein, glycogen and fat in the midgut gland during ovarian development and a similar increase in their concentration in the ovary[101]. Lawrence *et al.*[108] while studying the biochemical changes associated with maturation of the shrimp, *Penaeus setiferus*, have determined the midgut gland and gonad indices, % dry weight of carbohydrate, lipid and protein in midgut gland and gonad and carbohydrate, lipid and protein indices for midgut gland and gonad in maturing female and male prawns and suggested the importance of these organic substances in gamete formation and gonad maturation. Similar results were recorded by Shyamasundari and Erribabu[166] in a female crab *Menippe rumphii*. The midgut gland which is an important site of intermediary metabolism in crustaceans, may well be the major source of vitellogenic precursors,

though apparently not of the major vitellogenic protein in *Paratelphusa*[8]. Supportingly, we have analysed electrophoretically the protein patterns in gill, midgut gland and abdominal muscles in four species of prawns (*Metapenaeus affinis* ; *M. monoceros* ; *Parapenaeopsis hardwickii* and *P. stylifera*) in relation to sex and found the weakest bands in the midgut gland of all prawns[104]. Free sugars (glucose, galactose and sucrose) in the midgut gland of *Paratelphusa* undergo quantitative and qualitative cyclic fluctuations related to the ovarian cycle[1]. The changes in the level of different biochemical parameters are reflected in the hemolymph[66]. A doubling of the hemolymph glucose was observed by Dean and Vernberg[61] in *Callinectes sapidus* and by Telford[187] in *Cancer borealis* as the crabs become ovigerous. We have found a direct proportionality between hemolymph glucose level and ovary growth in *Parapenaeopsis hardwickii*[128]. There was an incessant hike in hemolymph glucose level as ovarian development proceeds from Stage I to IV but a drastic downfall was noticed in Stage V, i.e., in the spawned prawns. The sugars present in the midgut gland and hemolymph are also found in some abundance in the ovary during early stages of vitellogenesis, but disappear progressively as the proteins in the ovary become conjugated in the course of yolk formation[1]. This suggests that the sugars may be utilized mainly in the construction of yolk proteins by the ovary and perhaps also as fuel during that process. Ganion and Kessel[70] have evidenced the intracellar synthesis, transport and packaging of proteinaceous yolk in oocytes of *Orconectes immunis* during the ovarian maturation and vitellogenesis, moreover, recently, Zerbib[200] reported that in the crayfishes, *Astacus astacus* and *Astacus leptodactylus*, the yolk protein material is derived from intra- and extra-oocytic sources. In *Paratelphusa* the sex-related and other vitellogenic plasma proteins may act as blood carriers of lipids and carbohydrates, presumably from the midgut gland to the ovary[5,7], the ovarian proteins becoming conjugated to the carbohydrate and lipid moieties apparently only within the ovary. Wallace *et al.*[195] have demonstrated the chemical nature of the principle protein component of mature ovaries and also of the eggs of six species of decapods as a lipoprotein, whereas chemical nature of the yolk protein of mature oocytes in *Parapenaeopsis hardwickii* was found to be lipovitellin[102].

There is some evidence to show that the accumulation of organic reserves in the midgut gland and their subsequent liberation may be regulated by eyestalk hormones since eyestalk ablation results in a depletion

of such reserves [8,6,168] and degeneration of the midgut gland[122,109] in decapods. Ablation experiments have shown in *Paratelphusa hydrodromous* that eyestalk hormone is necessary for the normal maintenance of the midgut gland. Eyestalk extirpation resulted in the vacuolation of metal cells, also the cytoplasmic RNA content of the iron cells was visibly reduced and the distinction between iron and copper cells had almost disappeared[14]. In the crayfish, *Procambarus clarkii*[92] and the prawn, *Caridina weberi*[92] a neuroendocrine substance from the eyestalk appears to stimulate the synthesis of RNA in the midgut gland, and this RNA is probably responsible for amylase synthesis in the same organ. Earlier, Thornborough[186] stated in the prawn, *Palaemonetes vulgaris* that the eyestalk extract repressed the activity of midgut gland RNA. He suggested that the eyestalk factor(s) may operate as a repressor of the structural gene or it may act as an activator of the regulator gene or as a corepressor of its products (Fig. 4).

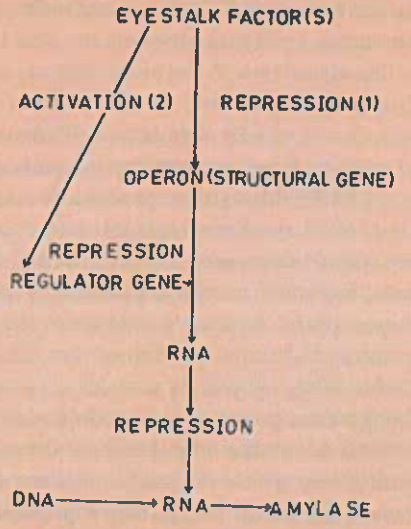


Fig. 4. Diagrammatic summary of the possible mechanism of the action of eyestalk hormonal factor controlling amylase secretion (from Jyoti and Nagabhushanam, 1977)

In anomurans, the Z-organs, whose secretory activity is correlated with moulting and oocyte maturation[105] contain a factor controlling protein synthesis in the midgut gland[106]. In *Paratelphusa*, mobilization of yolk

protein is known to be arrested by the inhibitory gonadotropin as evidenced by the precocious decrease in the quantity of yolk proteins in the plasma following the eyestalk ablation[5]. In view of the antagonism between inhibitory and stimulatory gonadotropins, the latter may be considered as probably stimulating the synthesis and utilization of the yolk proteins under low titres of inhibitory gonadotropin[1], but direct experimental evidence, such as implantation therapy remains yet to be furnished. Decapod eyestalk is known to contain a hyperglycaemic (diabetogenic) hormone which increases hemolymph sugar levels when administered into normal as well as eyestalkless crustaceans[64,114,188]. We found interesting results in *Parapenaeopsis hardwickii*. There was a notable enhancement in the hemolymph glucose level in the cauterized eyestalkless and normal prawns injected with eyestalk extract, but on contrary, a considerable decrease in the level of hemolymph glucose in normal prawns was found after injecting little higher doses of brain and thoracic ganglion separately. This suggested that eyestalks of this prawn also contains a hyperglycaemic hormone but it is expected that brain and thoracic ganglion neurosecretory machinery might be synthesizing a factor responsible for hypoglycaemia. Of course further investigation is needed in this direction before coming to any firm conclusion. Moreover, we observed an obvious change in hemolymph glucose level due to the endocrine hormones injected from other prawns of the related family. Eyestalk extracts of *Parapenaeopsis stylifera* when injected into cauterized eyestalkless, *P. hardwickii* produced a remarkable increase, whereas brain and thoracic ganglion caused considerable decrease in hemolymph glucose level. The extracts of the same endocrine centres from *Metapenaeus monoceros* did not induce any significant alteration in quantity of hemolymph glucose of *P. hardwickii*. Therefore, it is indicated, for the very first time in prawns, that there is an intrageneric neuroendocrine hormonal relation and may be suggested that they influence the physiological phenomena of one another[129]. Also a transglucosylase inhibitor which controls glycogen synthesis[188], and whose activity probably varies with the stages in the moulting cycle[119,194] has been obtained from eyestalk extracts. We have reported in *Parapenaeopsis hardwickii* that eyestalk hormones have inhibitory influence over the activity of the enzyme phosphorylase of midgut gland and uridine-diphosphate glucose-glycogen trans-glucosylase (UDPG-GT) of abdominal muscle, and accelerate the phosphorylase activity in abdominal muscle.

Furthermore, it is also suggested that in this prawn, eyestalk hormones have profound influence on carbohydrate metabolism and the hormone activity alters according to moulting stages[180]. The cyclic changes observed in the levels of carbohydrates in the midgut gland and hemolymph in relation to gonadal development suggest the possibility that the carbohydrate metabolism related to yolk formation may perhaps be under the control not only of the specific reproductive hormones, but also of some of the metabolic hormones not usually implicated in reproduction. The problem may be ultimately solved only with the isolation, purification and characterization of the hormones from various neuroendocrine centres, a beginning in this direction is already initiated[1].

6. Steroid Hormones in Crustacea

The occurrence of steroid hormones in Crustacea is not new. The well-known and characterized hormone, ecdysone, secreted by Y-organ, is steroidal in nature and is able to accelerate moulting in Crustacea. Therefore, it is called Crustecdysone. This steroid has now been isolated and identified. It apparently has the same configuration as β-ecdysone (20-hydroxyecdysone)[78]. But in addition, a second ecdysone has been isolated from the marine crayfish, *Jasus lalendei*, in which the hydroxyl group on C atom 3 is replaced by hydrogen. The compound is therefore called deoxycrustecdysone[69]. At present it is not known whether the individual hormones have specific effects upon different aspects of metabolism and development, or whether they are in fact precursors or breakdown products of a single steroid hormone and warrant further investigation to confirm the exact phenomenon offing in. Blanchet *et al.*[28] have reported that the ecdysteroid titer in the whole body extract of *Orchestia gammarellus* fluctuate during the course of the moult cycle. Ovaries, at the end of vitellogenesis, contain ecdysone and low polarity substances. Ecdysteroids are also present in eggs after laying. Variations in ecdysteroid levels influence oocyte growth, vitellogenin synthesis, however, is independent of hormonal fluctuations.

During the last one-and-half decade, efforts have been made to determine the occurrence of vertebrate steroid hormones in gonads and their role in the reproduction of Crustaceans. Many invertebrates are able to synthesize these compounds[112]. There are a few studies on the occurrence of steroids in the reproductive tissues of some invertebrates like molluscs[39,75,76],

echinoderms[28,161,162,163], insects[111] and lobsters[72,118]. As far as the knowledge on the steroidogenesis in crustaceans is concerned, the information is very fragmentary. There are also some reports on the effect of vertebrate steroid hormones on oogenesis of prawns[102,182]; shrimp[27] and crabs[156]. Teshima[175,176,177,178] has evidenced the bioconversion of dietary brassicasterol and ergosterol to cholesterol and its utilization and metabolism during steroidogenesis in some marine crustaceans like *Artemia salina* and *Penaeus japonicus*. Teshima and Kanazawa[179] have shown the biosynthesis of sterols in the lobster, *Panulirus japonicus*, the prawn, *Penaeus japonica* and the crab, *Portunus trituberculatus*. Further, it is reported by the same authors[180] that the ovaries of crab, *P. trituberculatus*, convert progesterone-4-^{14}C to 17-hydroxyprogesterone, testosterone, and deoxycorticosterone, and suggested the presence of steroid 17-α-hydroxylase, steroid 21-hydroxylase and steroid $C_{17} - C_{21}$ lyase in the ovaries of the crab, while studying the steroidogenesis in the testis of Crustaceans. Teshima and Kanazawa (1971c) found in the crab, *P. trituberculatus* that after incubation of the sliced testes with progesterone-4-^{14}C it was observed that 17α-hydroxyprogesterone and testosterone were synthesized by the testes and suggested that the testes of the crab, *P. trituberculatus* contain the enzyme system for 17α-hydroxylation and side-chain cleavage of progesterone. Cholesterol has been reported to be the basic substance or precursor for the synthesis of steroid hormones in many invertebrates[39,75]. Kanazawa and Teshima[98] have reported on the conversion of cholesterol to steroid hormones in the spiny lobster, *Panulirus japonicus*. Previously, Gilgan and Idler[72] demonstrated that the testes and androgenic glands of the lobster, *Homarus americanus* can convert androstenedione to testosterone and Tcholakian and Eik-Nes[178] found that the androgenic glands of the blue crab, *Callinectes sapidus* can convert progesterone to 11-deoxycorticosterone, but these steroids have not yet been identified in the blood. Earlier, Sarojini[157] suggested that when the androgenic glands of *Ocypoda platyotarsis* are extracted for sterols, the non-saponifiable lipid fraction mimics the effects of androgenic gland transplants when injected into female crabs. Moreover, injected testosterone will convert ovarian to testis tissue, although it will not by itself induce the secondary male characters. The evidences so far suggest that the crustacean androgenic gland could produce a sex steroid, perhaps in combination with a protein hormone to fulfil the complete range of androgenic gland activities. Bomirski and Klek-Kawinska[27] have documented that a human

chorionic gonadotropin (HCG) has stimulated the vitellogenesis in *Crangon crangon*. Existence of an estrogen-like compound has been observed by Jeng *et al.*[85] in the ovary of the shrimp, *Parapenaeus fissurus*. Recently we have reported the trophic influence of progesteron on the ovarian maturation in *Parapenaeopsis hardwickii* and suggested one significant difference of its action in crustaceans than in vertebrates. In vertebrates this hormone has no effect on germ cells until they are close to maturity[20,29] but unlike in vertebrates, it has stimulated proliferation and development of germinative cells in this prawn at an immature stage to form oogonia which later rapidly grew into oocytes[102]. But interestingly, when the same prawn was injected with vertebrate male steroid hormones, testosterone propionate and testosterone acetate, there was a total arrest of ovarian maturation but as in *Ocypoda platytarsis*[156] there was no masculinization of the ovary[126]. Similar arrest of ovarian growth was also obtained by Joshi *et al.*[88] in *Parapenaeopsis stylifera* after injecting the aqueous extract of androgenic gland and suggested that the aqueous fraction (mostly proteinaceous) is responsible only to retard the maturation and lipoidal factor, as reported by other, may be responsible to induce the masculinization of ovary and modification of secondary sexual characters. The exogenously injected Testosterone propionate and Testosterone acetate have been reported to have stimulatory influence on the androgenic gland activity, which hypertrophies possibly due to hyperplasia of the component cells, and consequently the testicular growth in *Parapenaeopsis hardwickii*[181]. The possible mechanism of action is suggested that these hormones, like in vertebrates along with other steroid hormones, can penetrate the plasma membranes of androgenic gland cells and affect organized sites deeply within them. The changes in nuclear structure may suggest that the penetrated hormones might trigger the nuclear activity and consequent rapid synthesis of secretory granules in the cytoplasm. This increased activity of androgenic gland cell synthesize and release the hormone faster which can initiate spermatogenesis. The hormones may be evoking the cells of germinative portion to undergo changes to form spermatogonia which ultimately grow into mature spermatocytes, spermatids and spermatozoa.

The techniques adopted to get artificially the rapid gonad maturation[16] and spawning[155] is that of eyestalk ablation which is no more dependable. The occurrence of steroid hormones in the gonads of crustaceans is already

evidenced and such steroid hormones can be used in the reproductive processes. Recently we, perhaps for the first time in literature, have succeeded in artificially spawning the prawn, *Parapenaeopsis stylifera* using a steroid preparation, 17-hydroxy progesterone[195] at low temperature (20°C) at which this prawn does not spawn in nature[185]. Therefore, it was suggested that a steroid mediator is essential to the crustaceans for spawning which is generally synthesized at high temperature, but for the prawns to spawn at low temperature, the administration of exogenous steroid mediator like 17-hydroxy progesterone is needed, because synthesis of such steroid mediator is inhibited at low temperature. Similar effect of another steroid mediator like 17-hydroxy-20 dihydro-progesterone on spawning in carps has been reported and the possible endogenous action of it as suggested above, was strengthened by Jalbert *et al.*[84].

ACKNOWLEDGEMENTS

Financial support extended by ICAR and UGC, New Delhi, through a project on Reproductive Biology of Edible Prawns are gratefully acknowledged. We are also indebted to Dr. S. N. Dwivedi, Director, Central Institute of Fisheries Education (CIFE), Versova, Bombay for generously providing laboratory facilities to carry out the investigation on marine prawns.

REFERENCES

1. Adiyodi, K. G. and Adiyodi, R. G. (1970a) *Biol. Rev.*, 45 : 121.
2. Adiyodi, R. G. and Adiyodi, K. G. (1970b) *J. Ker. Acad. Biol.*, 3(1) : 48.
3. Adiyodi, R. G. (1967) Endocrine physiology of moulting and regeneration in the crab, *Paratelphusa hydrodromous*, Ph.D. Thesis, Kerala University, India.
4. Adiyodi, R. G. (1968a) *Physiol. Zool.*, 41 : 204.
5. Adiyodi, R. G. (1968b) *Indian J. Exp. Biol.*, 6 : 144.
6. Adiyodi, R. G. (1969a) *Experientia*, 25 : 43.
7. Adiyodi, R. G. (1969b) *Indian J. Exp. Biol.*, 7 : 13.
8. Adiyodi, R. G. (1969c) *Ibid.*, 135.
9. Adiyodi, R. G. (1978) *Comp. Endocr.*, (Eds. P. J. Gaillard and H. H. Boer) : 25-28.
10. Aiken, D. E. (1969) *Science*, N.Y., 164 : 149.
11. Alikunhi, H. H., Poernomon, A., Adisukresne, S., Budiono, M. and Busman, S. (1975) *Bull. Shrimp. Cult. Res. Ctr.*, 1(1) : 1.
12. Amato, G. D. and Payen, G. G. (1978) *Gen. Comp. Endocr.*, 36 : 487.
13. Andrew, R. D. and Shivers, R. R. (1976) *J. Morphol.*, 150(1) : 253.

14. Anilkumar, G. and Adiyodi, K. G. (1976) V Internatl. Congr. Histochem. Cytochem., (Abstract) Buckharest.
15. Aoto, T. and Nishida, H. (1956) *J. Fac. Sci. Hokkaido Univ. (Ser. VI)*, **12** : 412.
16. Arnstein, D. R. and Beard, J. W. (1975) *Aquaculture*, **5** : 411.
17. Arvy, L., Echalier, G. and Gabe, M. (1954) *C.R. Acad. Sci. Paris*, **239** : 1853.
18. Arvy, L., Echalier, G. and Gabe, M. (1956) *Ann. Sci. Nat. Zool., Serie 11*, **18** : 263.
19. Balesdent, M. L. (1965) *Bull. Acad. Soc. Lorraines Sci.*, **5** : 1.
20. Ben-Or, S. (1970) *In* "Gonadotrophins and ovarian development", (Eds. W. R. Bhutt, A. C. Crooke and M. Ryle) : 271.
21. Berreur-Bonnenfant, J. (1967) *Arch. Zool. Exp. Gen.*, **108** : 521.
22. Berreur-Bonnenfant, J., Meusy, J. J., Ferezou, J. P., Devys, M., Quesneau-Thierry, A. and Barbier, M. (1973) *C.R. Acad. Sci. Paris*, **277** : 971.
23. Blanchet, M. F., Porcheron, P. and Dray, F. (1979) *Internatl. J. Invert. Reprod.*, **1**(2) : 133.
24. Bonnenfant, J. (1961) *Compt. Rend. Acad. Sci.*, **252** : 1518.
25. Bonnenfant, J. (1964) *Annls Endocr.*, **25** (5 suppl.) : 14.
26. Bomirski, A. and Klek-Kawinska, E. (1974) *Mar. Biol.*, **24** : 329.
27. Bomirski, A. and Klek-Kawinska, E. (1976) *Gen. Comp. Endocr.*, **30** : 239.
28. Boticelli, C. R., Hisaw, F. L. and Wotiz, H. H. (1961) *Proc. Soc. Exp. Biol. Med.*, **606** : 887.
29. Brandau, H. (1970) *In* "Gonadotrophins and ovarian development", (Eds. W. R. Bhutt, A. C. Crooke and M. Ryle), 311.
30. Brockenbrough, F. N. and Hoffman, D. L. (1974) *Gen. Comp. Endocr.*, **22** : 439.
31. Brown, F. A. and Jones, G. M. (1947) *Anat. Rec.*, **99** : 657.
32. Brown, F. A. and Jones, G. M. (1949) *Biol. Bull. mar. biol. Lab. Woods Hole*, **91** : 228.
33. Carlisle, D. B. (1953) *Pubbl. Staz. Zool. Napoli*, **24** : 355.
34. Carlisle, D. B. (1954) *Ibid.*, 79.
35. Carlisle, D. B. (1957) *J. mar. biol. Assoc. UK*, **36** : 291.
36. Carlisle, D. B. (1959a) *Ibid.*, **38** : 381.
37. Carlisle, D. B. (1959b) *Ibid.*, 481.
38. Carpenter, M. B. and DeRoos, R. (1970) *Gen. Comp. Endocr.*, **15** : 143.
39. Carreau, S. and Drosdowsky, M. (1977) *Gen. Comp. Endocr.*, **33**(4) : 554.
40. Charniaux-Cotton, H. (1954) *C.R. Acad. Sci.*, **238** : 953.
41. Charniaux-Cotton, H. (1957) *Ibid.*, **245** : 1655.
42. Charniaux-Cotton, H. (1958) *C.R. Hebd. Seanc. Acad. Sci. Paris*, **246** : 2817.
43. Charniaux-Cotton, H. (1960) *Ibid.*, **250** : 4046.
44. Charniaux-Cotton, H. (1961) *Bull. Soc. Zool. France*, **86** : 484.
45. Charniaux-Cotton, H. (1962) *Gen. Comp. Endocr. (Suppl.)* **1** : 241.
46. Charniaux-Cotton, H. and Kleinholz, L. H. (1964) *In* "The Hormones", (Eds. C. G. Pincus, K. V. Thimann and E. B. Astwood), Chapt. 3 : 135.
47. Charniaux-Cotton, H., Zebbib, C. and Meusy, J. J. (1966) *Crustaceana*, **10** : 113.
48. Cheung, T. S. (1966) *Biol. Bull. mar. biol. Lab. Woods Hole*, **130** : 59.
49. Cornubert, G., Demeusy, N. and Veillet, A. (1952) *C.R. Hebd. Seanc. Acad. Sci. Paris*, **234** : 1405.

50. Croisille, Y., Junera, H., Meusy, J. J. and Charniaux-Cotton, H. (1974) *Amer. Zoologist*, **14** : 1219.
51. Dean, A. M. and Vernberg, F. J. (1965) *Comp. Biochem. Physiol.*, **14** : 29.
52. Demassieux, C. and Balesdent, M. L. (1977) *C.R. Acad. Sci. Ser. D.*, **284**(3) : 207.
53. Demassieux, C. and Balesdent, M. L. (1978) *Arch. Biol. (Bruxelles)*, **89** : 89.
54. Demeusy, N. (1953) *C.R. Acad. Sci. Ser. D.*, **239** : 780.
55. Demeusy, N. (1962) *Cah. Biol. Mar.*, **3** : 37.
56. Demeusy, N. (1965) *Archs. Zool. Exp. Gen.*, **106** : 625.
57. Demeusy, N. (1967a) *C. R. Hebd. Seanc. Acad. Sci. Paris*, **265** : 568.
58. Demeusy, N. (1967b) *Ibid.*, 628.
59. Diwan, A. D. and Nagabhushanam, R. (1975) *Riv. Di. Biol.*, **28** : 79.
60. Drach, P. (1955) *C. R. Seanc. Soc. Biol.*, **149** : 2079.
61. Enami, M. (1951) *Biol. Bull. mar. biol. Lab. Woods Hole*, **101** : 241.
62. Farooqui, U. M. and Nagabhushanam, R. (1980) Symp. Crustacean Endocr., (Abst.) 27-29 March, Aurangabad.
63. Ferezou, J. P., Berreur-Bonnenfant, J., Tekitec, A., Rojas, M., Barbier, M., Suchy, M., Wipf, H. K. and Meusy, J. J. (1977) *In* "Marine Natural Products Chemistry", (Eds. D. J. Faulkner and W. H. Fenical), N.Y. : 361.
64. Fingerman, M. (1970) *Scientia* : 1.
65. Fingerman, M., Dominiczak, T., Miyawaki, M., Oguro, C. and Yamamoto, Y. (1967) *Physiol. Zool.*, **40** : 23.
66. Florkin, M. (1960) *In* "The Physiology of Crustacea", (Ed. T. H. Waterman), N.Y., **1** : 141.
67. Gabe, M. (1953) *C.R. Acad. Sci.*, **237** : 1111.
68. Gabe, M. (1966) *Neurosecretion*, Oxford Pergamon Press.
69. Galbraith, M. N., Horn, D. H. S., Middleton, E. J. and Hackney, R. J. (1968) *Chem. Commun.*, **2** : 83.
70. Ganion, L. R. and Kessel, R. G. (1972) *J. Cell Biol.*, **52** : 420.
71. Gilbert, L. I. and O'Connor, J. D. (1970) *In* "Chemical Zoology", (Eds. M. Florkin and B. T. Scheer), **5** : 229.
72. Gilgan, M. W. and Idler, D. R. (1967) *Gen. Comp. Endocr.*, **9** : 319.
73. Gomez, R. (1965) *Naturwissenschaften*, **9** : 216.
74. Gomez, R. and Nayar, K. K. (1965) *Zool. Jb. (Abst. Physiol.)*, **71** : 694.
75. Gottfried, H. and Dorfman, R. I. (1970) *Gen. Comp. Endocr.*, **15** : 120.
76. Gottfried, H., Dorfman, R. I., Forschielli, E. and Wall, P. E. (1967) *Gen. Comp. Endocr.*, **9** : 454 (Abstract only).
77. Halder, D. D. (1978) *Aquaculture*, **15** : 171.
78. Hampshire, F. and Horn, D. H. S. (1966) *Chem. Commun.* : 37.
79. Hanaoka, K. I. and Otsu, T. (1957) *J. Fac. Sci. Hokkaido Univ.*, **13** : 379.
80. Highnam, K. C. and Hill, L. (1969) *The Comparative Endocrinology of the Invertebrates*, Edward Arnold Ltd., London : 169.
81. Hisano, S. (1974) *J. Fac. Sci. Hokkaido Univ.*, **19**(3) : 503.
82. Hoffman, D. L. (1968) *Nature Lond.*, **218** : 170.
83. Hohnke, L. and Scheer, B. T. (1970) *In* "Chemical Zoology", (Eds. M. Florkin and B. T. Scheer), N.Y., V. : 147.
84. Jalbert, B., Breton, B., Bruzuska, E., Fastier, A. and Wieniawski, S. (1977) *Aquaculture*, **10** : 353.

85. Jeng, S. S., Chia-Mowan, W. and Chang, C. F. (1978) *Gen. Comp. Endocr.*, 36(2) : 211.
86. John, S. and Sivadas, P. (1978) *Mahasagar*, 11 (1 and 2) : 62.
87. John, S. and Sivadas, P. (1979) *Ibid.*, 12(1) : 11.
88. Joshi, P. K., Nagabhushanam, R. and Kulkarni G. K. (1980) *Indian J. mar. Sci.*, (*in press*).
89. Juchault, P. (1977) *C.R. Acad. Sci. Serie D.*, 285(2) : 179.
90. Juchault, P. and Legrand, J. J. (1978) *Gen. Comp. Endocr.*, 36 : 175.
91. Junera, H., Zerbib, C., Martin, M. and Meusy, J. J. (1977) *Ibid.*, 31 : 457.
92. Jyoti, M. and Nagabhushanam, R. (1977) *Comp. Physiol. Ecol.*, 2(2) : 70.
93. Kanazawa, A. and Teshima, S. I. (1971) *Bull. Jap. Soc. Sci. Fish.*, 37 : 891.
94. Karlson, P. and Skinner, D. M. (1960) *Nature Lond.*, 185 : 543.
95. Katakura, Y. (1960) *Annot. Zool. Jap.*, 33(4) : 241.
96. Katakura, Y. (1967) *Gunma Symp. Endocr.*, 4 : 49.
97. Katakura, Y., Fujimaki, Y. and Unno, K. (1975) *Annot. Zool. Jap.*, 48(4).
98. King, D. S. (1964) *Gen. Comp. Endocr.*, 4 : 533.
99. Klek-Kawinska, E. and Bomirski, A. (1975) *Ibid.*, 25 : 9.
100. Knowles, F. (1965) *Arch. Anat. Microscop.*, 54 : 343.
101. Kulkarni, G. K. and Nagabhushanam, R. (1979) *Aquaculture*, 18 : 373.
102. Kulkarni, G. K., Nagabhushanam, R. and Joshi, P. K. (1979) *Indian J. Exp. Biol.*, 17 : 986.
103. Kulkarni, G. K., Nagabhushanam, R. and Joshi, P. K. (1980a) *Abst. Ist India Symp. Crust. Endocr., Aurangabad*, 2 March 26-28.
104. Kulkarni, G. K., Nagabhushanam, R. and Joshi, P. K. (1980b) *Hydrobiologia*, 69 (1 and 2) : 25.
105. Kurup, N. G. (1964) *Gen. Comp. Endocr.*, 4 : 99.
106. Kurup, N. G. and Scheer, B. T. (1966) *Comp. Biochem. Physiol.*, 18 : 971.
107. Laubier, A. B. (1975) *C.R. Acad. Sci. Serie D.*, 281 : 2013.
108. Lawrence, A. L., Chamberlin, G. W., Hutchins, D. L. and Middleditch, B. S. (1979) ISIR Symp. Davis (Abstract), 27-29 August.
109. Legrand, J. J. (1955) *C.R. Acad. Sci. Paris*, 241 : 1083.
110. Legrand, J. J. and Juchault, P. (1960) *C.R. Soc. Biol.*, 154 : 676.
111. Lehoux, J. G., Sandor, J., Lanthier, A. and Lusis, O. (1968) *Gen. Comp. Endocr.*, 11 : 481.
112. Lehaux, J. G. and Sandor, T. (1970) *Steroids.*, 16 : 141.
113. Lisk, R. P. (1961) *Can. J. Biochem.*, 39 : 659.
114. Lockwood, A. P. M. (1968) *Aspects of the Physiology of Crustacea*, London, Oliver and Boyd. Publ.
115. Lowe, M. E. (1961) *Tulane Stud. Zool.*, 8 : 157.
116. Madhyastha, M. N. and Rangnekar, P. V. (1973).
117. Matsumoto, K. (1958) *Biol. J. Okayama Univ.*, 4 : 103.
118. Matsumoto, K. (1962) *Gen. Comp. Endocr.*, 2 : 4.
119. McWhinnie, M. A. and Chua, A. S. (1964) *Ibid.*, 4 : 624.
120. Meusy, J. J. (1965) *C.R. Hebd. Seanc. Acad. Sci. Paris*, 260 : 5901.
121. Meusy, J. J., Blanchet, M. F. and Junera, H. (1977) *Gen. Comp. Endocr.*, 33 : 35.

122. Miyawaki, M. and Tanoue, S. (1962) *Kumamato J. Sci.* (*Ser. B, Sect. 2, Biol.*), 6 : 1.
123. Mauviot, J. C. and Castell, J. D. (1976) *J. Fish. Res. Board Can.*, 33(1) : 1922.
124. Nagabhushanam, R. and Diwan, A. D. (1974) *Nat. Sci. J.*, 13(6) : 59.
125. Nagabhushanam, R., Joshi, P. K. and Kulkarni, G. K. (1980) *Indian J. mar. Sci.*, (*in press*).
126. Nagabhushanam, R. and Kulkarni, G. K.(1979) All India Symp. Develop. Biol., Gwalior (Abst.) 13.
127. Nagabhushanam, R. and Kulkarni, G. K. (1980a) Proc. Symp. Coastal Aquacult. Cochin (*in press*).
128. Nagabhushanam, R. and Kulkarni, G. K. (1980b) *Hydrobiologia* (*in press*).
129. Nagabhushanam, R. and Kulkarni, G. K. (1980c) *Ibid.* (*in press*).
130. Nagabhushanam, R. and Kulkarni, G. K. (1980d) *Ibid.* (*in press*).
131. Nagabhushanam, R. and Kulkarni, G. K. (1980e) *Aquaculture* (*in press*).
132. Nagabhushanam, R. and Kulkarni, K. M. (1977) *Monitor Zool.* (*Ital.*), 11 : 57.
133. Nagabhushanam, R. and Shirwadkar, S. K. (1976) *Nat. Sci. J.*, 15(8) : 205.
134. Nagabhushanam, R. and Vasantha, N. (1972) *Broteria*, XLI(3-4) : 177.
135. Nagamine, C. and Knight, A. (1979a) ISIR Symp. on Invert. Reprodn. Davis, (Abst.) 27-30 Aug.
136. Nagamine, C. and Knight, A. (1979b) *Ibid.* (Abst.).
137. O'Connor, J. D. and Gilbert, L. I. (1968) *Amer. Zoologists*, 8 : 529.
138. Otsu, T. (1960) *Annot. Zool. Japon*, 33 : 90.
139. Otsu, T. (1961) *Bull. Yamagata Univ.* (*Ser. Nat. Sci.*), 5 : 379.
140. Otsu, T. (1963) *Embryologica*, 8 : 1.
141. Otsu, T. (1964) *Gunma Symp. on Endocr.*, 1 : 25.
142. Panouse, J. B. (1943) *C.R. Hebd. Seanc. Acad. Sci. Paris*, 217 : 553.
143. Panouse, J. B. (1947) *Bull. Biol. Fr. Belg.* (*Suppl.*), 33 : 160.
144. Passano, L. M. (1960) *In* "The Physiology of Crustacea", (Ed. T. H. Waterman), N.Y., I : 473.
145. Payen, G. (1968) *C.R. hebd. Seanc. Acad. Sci. Paris*, 266 : 1056.
146. Payen, G., Costlow, J. D. and Charniaux-Cotton, H. (1971) *Gen. Comp. Endocr.*, 17 : 526.
147. Perryman, E. K. (1969) *Trans. Amer. Micro. Soc.*, 88(4) : 514.
148. Pillay, K. K. and Nair, N. B. (1970) *Curr. Sci.*, 40 : 161.
149. Pillay, K. K. and Nair, N. B. (1973) *Mar. Biol.*, 18 : 167.
150. Primavera, J. H. (1978) *Aquaculture*, 15 : 355.
151. Rangnekar, P. V. and Deshmukh, R. D. (1968) *J. Anim. Morphol. Physiol.*, 15 : 116.
152. Rangnekar, P. V., Madhyastha, M. N. and Latey, A. N. (1971) *Ibid.*, 18 : 17.
153. Rice, P. R. and Armitage, K. B. (1974) *Comp. Biochem. Physiol.*, 47A : 243.
154. Rychaert, M., Vial, N. and Ceccal di, H. J. (1974) *C.R. Seanc. Soc. Biol.*, 168 : 1364.
155. Santiago, A. C. (1977) *Aquaculture*, 11 : 185.
156. Sarojini, S. (1963) *Curr. Sci.*, 32(9) : 411.
157. Sarojini, S. (1964) *Ibid.*, 33(2) : 55.
158. Scharrer, E. and Scharrer, B. (1954a) *In* "Handbuch de Mikroskopischen Anatomie des Menschen", W. Bergmann et. Berlin : 953-1066.

159. Scharrer, E. and Scharrer, B. (1954b) *In* "Recent progress in hormone research", **10** : 183.
160. Schneidermann, H. A. and Krishnakumaran, A. (1968) *Biol. Bull. mar. biol. Lab. Woods Hole*, **135** : 435 (Abst. only).
161. Schoenmakers, H. J. N., Lambert, I. G. D. and Voogt, P. A. (1976) *Gen. Comp. Endocr. (Abstr.)*, **29**(2) : 43.
162. Schoenmakers, H. J. N., Colenbrander, P. H. J. M. and Peute, J. (1977) *Cell Tiss. Res.*, **182** : 275.
163. Schoenmakers, H. J. N., Dieleman, S. J., Van Bohemen, Ch. G. and Voogt, P. A. (1978) *Gen. Endocr.*, (Eds. P. J. Gaillard and H. H. Boer), Amsterdam : 33.
164. Shirwadkar, S. K. (1977) *Physiological ecology of freshwater prawn, Caridina rajadhari.*, Ph.D. Thesis, Aurangabad.
165. Shivers, R. R. (1976) *J. Morphol.*, **150**(1) : 227.
166. Shyamasundari, K. and Erribabu, D. (1979) ISIR Symp. on Invert. Reprod. Davis, (Abstract) 27 August.
167. Shyamasundari, K., Govardhan, L. and Hanumantharao, K. (1978) *Boll. Zool. Ital. (in press)*.
168. Sinha, R. C. and Mooswi, H. K. (1978) *Comp. Physiol. Ecol.*, **3**(1) : 49.
169. Subrahmanyam, C. B, (1963) *Curr. Sci.*, **32** : 165.
170. Subramoniam, T. (1979) Souvenir of Crust. Reprod. Biol. : 7.
171. Takewaki, K. and Yamamoto, Y. (1950a) *Annot. Zool. Jap.*, **23** : 187.
172. Takewaki, K. and Yamamoto, Y. (1950b) *Zool. Mag. (Tokyo)*, **59** : 5.
173. Tcholakian, R. K. and Eik-Nes, K. B. (1969) *Gen. Comp. Endocr.*, **12** : 171.
174. Teissier, G. (1960) *In* "The Physiology of Crustacea", (Ed. T. H. Waterman), A.P. N.Y., **1** : 537
175. Teshima, S. (1971a) *Comp. Biochem. Physiol.*, **37** : 603.
176. Teshima, S. (1971b) *Bull. Jap. Soc. Sci. Fish.*, **37** : 720.
177. Teshima, S. (1972) *Ibid.*, **38** : 1305.
178. Teshima, S. (1973) *Mem. Fac. Fish. Kagoshima Univ. Jap.*, **22** : 15.
179. Teshima, S. and Kanazawa, A. (1971a) *Comp. Biochem. Physiol.*, **50B** : 485.
180. Teshima, S. and Kanazawa, A. (1971b) *Gen. Comp. Endocr.*, **17** : 152.
181. Teshima, S. and Kanazawa, A. (1971c) *Bull. Jap. Soc. Sci. Fish.*, **37**(6) : 524.
182. Teshima, S. and Kanazawa, A. (1973) *Mem. Soc. Fish. Kagoshima Univ. Jap.*, **22**(1) : 15.
183. Thampy, D. M. and John, P. A. (1974) *Inter. J. Parasitol.*, **4** : 535.
184. Thomas, M. M. (1974) *Indian J. Fish.*, **2**(1) : 152.
185. Thomas, M. M., George, K. V. and Kathirvel, M. (1974) *Indian J. Fish.*, **21** : 266.
186. Thornborough, J. R. (1968) *Comp. Biochem. Physiol.*, **24** : 625.
187. Telford, M. (1968) *Biol. Bull. mar. biol. Lap. Woods Hole*, **135** : 574.
188. Telford, M. (1975) *Comp. Biochem. Physiol.*, **51B** : 69.
189. Touir, A. (1973) *C.R. Acad. Sci. Paris Serie D*, **277** : 2541.
190. Touir, A. (1977) *Ibid.*, **285** : 539.
191. Vernet, G., Bressac, C. and Trilles, J. P. (1978) *Arch. Zool. exp. gen.*, **199** : 201.
192. Vijayaraghavan, S., Wafer, M. V. M. and Royan, J. P. (1976) *Mahasagar*, **8**(3/4) : 165.
193. Wallace, R. A., Walker, S. L. and Hauschka, P. V. (1967) *Biochemistry*, N.Y., **6** : 1582.
194. Wang, D. H. and Scheer, B. T. (1963) *Comp. Biochem. Physiol.*, **9** : 263.
195. Weitzman, M. C. (1964) *Am. Zoologists*, **4** : 329.
196. Webb, H. M. (1977) *Biol. Bull. mar. biol. Lab. Woods Hole*, **153**(3) : 630.
197. Wurtman, R. J. (1972) *Symp. Am. Soc. Pharmacol. exp. Therapeuttcs*, 1769.
198. Yaldwyn, J. C. (1966) *Nature Lond.*, **209** : 1366.
199. Yamamoto, Y. (1960) *Zool. Mag.*, **69** : 195.
200. Zerbib, C. (1979) *Internat. J. Invert. Reprod.*, **1**(5) : 289.

THE NATIONAL ACADEMY OF SCIENCES, INDIA

(Council for 1980)

President

Dr. Hari Narain, M.Sc., D.Phil., Ph.D. (Sydney), F.N.A.Sc., Vice-Chancellor, Banaras Hindu University, Varanasi.

Two Senior-most Past Presidents according to Seniority by Rotation

Prof. N. R. Dhar, D.Sc., Dr.Es.Sc., F.R.I.C., F.N.A , F.A.M., F.A.A., F.N.A.Sc., I.E.S. (Retd.), Corr. Member F.A.Sc., Director, Sheila Dhar Institute of Soil Science, University of Allahabad, Allahabad.

Prof. B. D. Nag Chaudhuri, M.Sc. (Alld.), Ph.D. (Calif.), F.N.A., F.N.A.Sc., Saha Institute of Nuclear Physics, 92, Acharya Prafulla Chandra Road, Calcutta-700 009.

Vice-Presidents

Prof. H. D. Srivastava, D.Sc., F.N.A., F.N.A.Sc., Emeritus Scientist, I.C.A.R., Zoology Department, Allahabad University, Allahabad.

Prof. R. S. Kushwaha, M.Sc., Ph.D., D.Phil., F.N.A., F.N.A.Sc., Prof. and Head, Department of Mathematics, University of Jodhpur, Jodhpur.

Treasurer

Prof. U. S. Srivastava, M.Sc., M.Ed., D.Phil., D.I.C., F.N.A., F.N.A.Sc., Professor and Head, Deptt. of Zoology, University of Allahabad, Allahabad.

Foreign Secretary

Prof. H. L. Nigam, M.Sc., D.Phil. (Alld.), Ph.D. (London), F.N.A., F.N.A.Sc., Professor of Chemistry, University of Allahabad, Allahabad.

General Secretaries

Prof. Krishnaji, M.Sc., F.N.A.Sc., M.I.E.T.E., Professor of Physics and National Fellow, Allahabad University, Allahabad.

Prof. P. T. Narasimhan, M.Sc., Ph.D., F.N.A , F.N.A.Sc., Senior Professor of Chemistry, Indian Institute of Technology, Kanpur.

Members

Prof. G. K. Manna, M.Sc., D.Phil., D.Sc., F.N.A., F.N.A.Sc., Professor of Zoology and National Fellow, Kalyani University Kalyani, Distt. Nadia, West Bengal.

(Council for 1980)—*Contd.*

Prof. K. S. Bilgrami, M.Sc., D.Phil., D.Sc., F.N.A.Sc., Prof. and Head, Deptt. of Botany, Bhagalpur University, Bhagalpur, Bihar.

Prof. R. N. Tandon, M.Sc., Ph.D., D.I.C., F.A.Sc., F.N.A., F.N.A.Sc., Retd. Prof. and Head of the Department of Botany, Allahabad University, C/o Sri G. N. Tandon, 4/3 Multistory Building, Sahajahan Road, New Delhi.

Prof. B. Ramchandra Rao, D.Sc., D.Sc. (*Honoris Causa*), F.N.A., F.N.A.Sc., Vice-Chairman, University Grants Commission, Bahadurshah Zafar Marg, New Delhi-110 002.

Prof. M. D. L. Srivastava, D.Sc , F.N.A.Sc., Retired Prof. and Head of the Department of Zoology, University of Allahabad, 4–D, Beli Road, Allahabad.

Dr. V. P. Sharma, M.Sc., D.Phil. (Alld.), F.N.A.Sc., F.E.S.I., F.I.S.C.D., Dy. Director, Malaria Research Centre, 22, Sham Nath Marg, Delhi-110 054.

Dr. S. P. Ray Chaudhuri, M.Sc., Ph.D., D.Sc., F.R.I.C., F.N.A., F.N.A.Sc., C-52, Inderpuri, New Delhi-110012.

Prof. V. K. Thakare, M.Sc., Ph.D., F.N.A.Sc., Prof. and Head, Deptt. of Zoology, Nagpur University, Nagpur.

Prof. H. C. Khare, M.Sc., Ph.D. (McGill), F.N.A.Sc., Professor of Mathematics, University of Allahabad, Allahabad.

Prof. S. Rameshan, M.Sc., A.I.I.Sc., D.Sc., F.N.A.Sc., Dy. Director and Head of Material Science, Dn., National Aeronautical Lab., Bangalore.

Dr. Nityanand, M.Sc., Ph.D. (Bom.), Ph.D. (Cantab.), F.N.A., F.N.A.Sc., Director, Central Drug Research Institute, Lucknow.

Prof. Suresh Chandra, M.Sc., D.Phil., Grad. Brit. I.R.E., F.N.A.Sc., Professor of Physics, Banaras Hindu University, Varanasi-5

Prof. M. K. Singhal, M.A., Ph.D., LL.B., F.N.A.Sc., Professor and Head, Department of Mathematics, Meerut University, Meerut.

Prof. S. Ghosh, D.Sc., F.R.I.C., F.N.A., F.N.A.Sc., Retired Head, Chemistry Department, Allahabad University, 21–B, Batuk Krishna Banerji Road, Allahabad.

We regret we are unable to accommodate the following five well-written articles for want of time and space :

1. Some Mathematical Problems in Biosciences
   by J. N. Kapur
2. Malaria : Current Situation and Area
   by V. P. Sharma
3. Elastic Collisions of Intermediate Energy Electrons with Atoms and Molecules
   by S. P. Khare and Deo Raj
4. Kachchh Mesozoic : Some Problems and Recent Contributions
   by S. K. Agrawal
5. Nitrogen Problem in the Production of Food and its Assimilation
   by N. R. Dhar

EDITOR

on behalf of the Printer