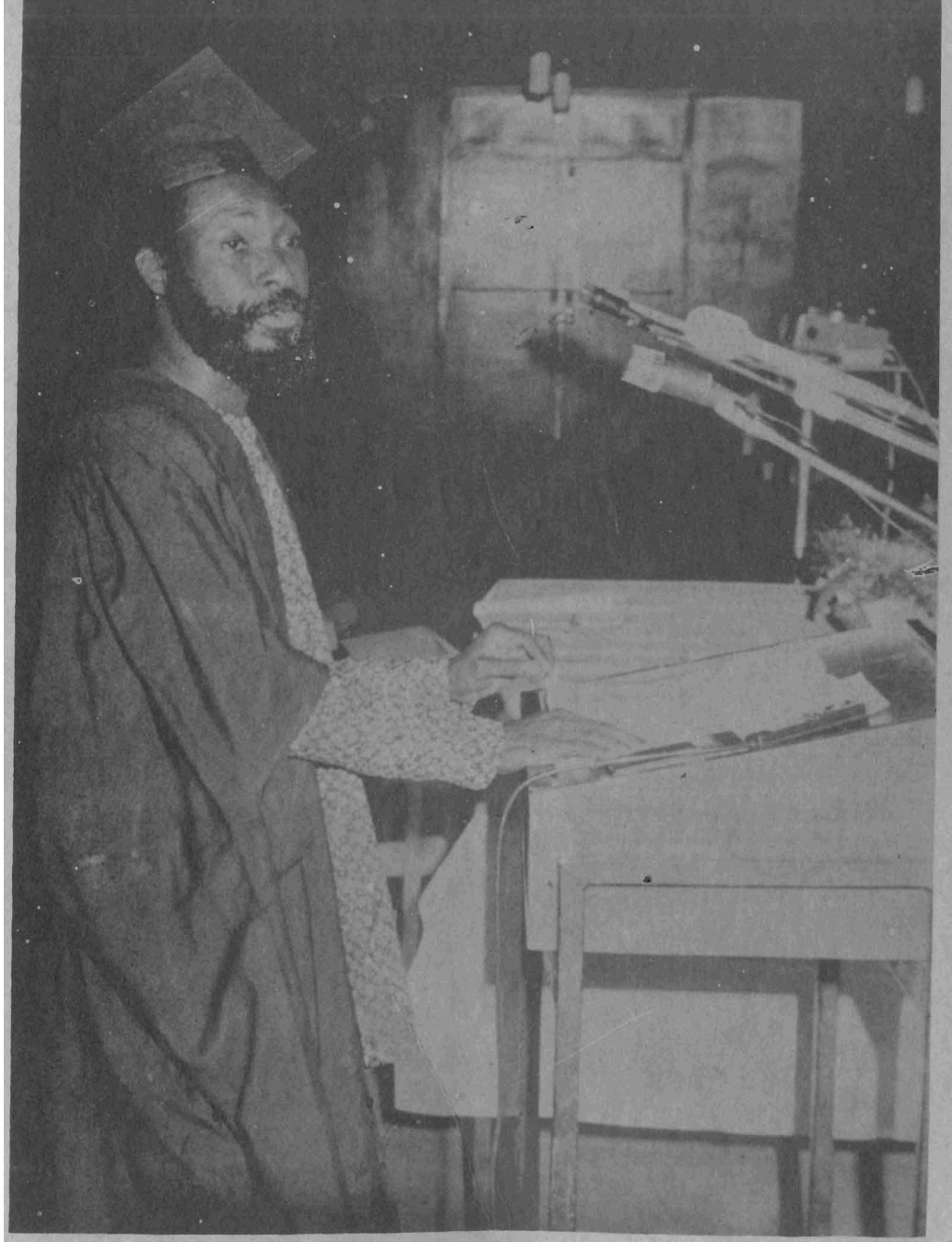
UNIVERSITY OF IFẸ · NIGERIA

Inaugural Lecture Series 52

EVOLUTION OF COMPLEXITY

by Kayode Adetugbo

UNIVERSITY OF IFE PRESS



PROFESSOR H. K. ADETUGBO

EVOLUTION OF COMPLEXITY

by

Kayode Adetugbo

Professor of Chemical Pathology, Haematology and Immunology

An Inaugural Lecture Delivered at the University of Ife on Thursday, 19th May, 1981.

Inaugural Lecture Series 52

UNIVERSITY OF IFE PRESS · ILE-IFE · NIGERIA.

Printed by Bosude Printers Limited, SW8/910, Bateye Street, Oke-Ado, Ibadan.

EVOLUTION OF COMPLEXITY

I

Indeed, evolution has had a fine career. Ever since Charles Darwin's *Origin of Species* in 1859, the theory has become the centrepiece of all biological explanation. This, in spite of Scopes trials, monkey trials, various creationist surges, organised religion and organised crime. As Theodosius Dobzhansky (1900-1975) put it: "Nothing in biology makes sense except in the light of evolution." Evolution today is a Kuhnian paradigm that is in the puzzle-solving phase of Normal Science.

Evolution has had a fine career, indeed. It has even gone into computers; it has become mathematical and molecular. Its first obstacle was overcome rather dramatically. That was in the debate at the 1860 meeting of the BA between Thomas Huxley, an eminent scientist, and Samuel Wilberforce, the Bishop of Oxford. The bishop had demanded that Huxley reveal on which side he claimed descent from the apes. Huxley, in turn, remarked that an ape would be a preferable grandfather to an intelligent man who used his skills to bring ridicule to science. The debate ended in favour of Huxley, and the bishop thereafter devoted his skills to ecclesiastical matters.

The significance of the theory of evolution is larger, much larger, than the drama of the electrified Huxley-Wilberforce debate. The significance is still very much with us and for us today. It is this: Darwin had completed the Copernican revolution by drawing out for biology the ultimate conclusions of the notion of nature as a lawful system of matter in motion. The diversity of organisms, the emergence of novel and highly complex forms, even the origin of man himself would henceforth be explained by an

orderly process of change governed by natural laws. These changes are changes in Mendelian factors, in genetic material, and are hereditable. Molecular biology helps place the Darwinian revolution in permanence by elucidating the mechanisms involved in evolution and by reducing teleology to causality.

In certain quarters,[1] however, the scientific nature of the theory of evolution has been questioned. Popperian epistemologists argue that evolution is not a scientific theory, and that it is in the nature of belief. This derives from the Popperian notion of the scientific method[2] which claims that scientific theories are always tentative; that scientific statements should be falsifiable, at least in principle, because they never can be shown to be true; and that such statements are objective only in so far as they are "intersubjectively testable." By this notion, it would seem that there is no (specified) condition or observation that could possibly constitute a falsification of the theory of evolution. It would also seem that the theory has no predictive value.

Now, there are certain difficulties with this view of science and, in particular, with its treatment of the theory of evolution. Obviously, there are many different views on the nature of science. There are also many different *systems* of science. Most of the current (Western) views on the nature of science ignore this latter point; they simply operate as if nothing exists outside of Western science. That Western science is but one of the systems of science has been forcefully put by a remarkable Islamic scholar, Ziauddin Sardar, in his recent book, *The Future of Muslim Civilisation*.[3] He argues that contemporary science is Western science, and only *a* science, and not *the* science. (True, this lecture operates only within the confines and limitations of Western science. But this is without prejudice to the validity of other systems of science. Western science is indeed only one

of the many systems of science. And it does not exclusively have the answer, nor does it have all the answers).

To return to Popperian hypothetico-deductivism: The decision to retain a scientific theory or to reject it is often subjective. This decision often depends on a concensus among scientists. This is implied in a published statement of Francis Crick's about the nature of molecular biology, my field of training and of present practice: "Molecular biology is what molecular biologists (choose to) do." The Popperian view of science does not take enough cognisance of what scientists really do, or of the aims of science. Simply put, the aim of science is the creation of knowledge. This creative enterprise needs guidance, hence the usefulness of theories. The knowledge thereby created may be immediately useful or not so. More importantly and decisively, however, knowledge, scientific knowledge, may be right, or it may be woefully wrong.

The views[4] on the nature of science expressed by Lakatos and by Fayerabend seem more real and more tolerant. Science is really not substantially different from other human activities -- the aim of science is the creation of knowledge, the establishment of relationships between observation and reality, the linkage between theories and the real world. If ongoing research programs continue to stimulate the process of the creation of knowledge, then, they are an acceptable part of the scientific corpus.

True, we cannot prove scientific theories; nor can we disprove them. After Einstein, it has become well accepted that scientific knowledge is not and cannot be proven truth. Science does not grow by an accretion of eternal truths. A proposition is scientific if it aims at expressing a causal connection. To quote Einstein:

> The ground aim of science is to cover the greatest number of empirical facts by the logical deductions from the smallest possible number of hypotheses

The synthetic theory of evolution is therefore science *par excellence*

Now, let us return to the status of evolutionary theory within the scientific enterprise. There are two kinds of questions in the study of evolution. One concerns history--the study of phylogeny, the description of the course of events that has led to the present state of the biological world. Systematics, paleontology, biogeography, comparative anatomy, comparative embryology, comparative biochemistry, all describe and inform this. There is a second kind of question, however. It concerns the elucidation of the mechanisms and processes that bring about evolutionary change. The questions deal with causal rather than historical relationships. These studies proceed by the formulation and empirical testing of hypotheses, the same (hypothetico-deductive) methodology characteristic of the physical sciences and other disciplines concerned with causal processes.

Molecular biology extends our understanding of the processes of life and its phenomena to the molecular level, and in terms of the co-ordinative interactions of small and large molecules. Molecular biology has also shown us that in all cases, the physical and chemical laws applicable to small molecules also apply, when we extend these to the study of the molecules of life - large molecules like proteins and nucleic acids, including DNA, the primary genetic material. Molecular biology applied to evolution is molecular evolution. Molecular evolution is a predictive science and employs the methodology of the physical sciences. It also allows the elucidation of causal mechanisms in evolution.

II

We have approached the study of molecular evolution from the perspective of the ontogeny of B lymphocytes. B lymphocytes are a class of lymphocytes, a type of white blood cells, that is responsible for the production of antibodies. Each of us has one million million lymphocytes: this cell type constitutes over 1% of the total number of cells in the body. Slightly less than half the lymphocyte population are B cells whose main concern is the elaboration of one kind of molecule - - the immunoglobulin (Ig) or antibody molecule. The other half or so are T lymphocytes. As we all know antibodies are body proteins which are made specifically in response to antigens. When parasites, bacteria, or other foreign substances enter the tissues, the vertebrate host produces antibodies which circulate dissolved in body fluids and which help to eliminate the infection or foreign organism. (Other types of cells involved in imune responses include other white blood cells -- T lymphocyte subsets, macrophages, eosinophils, basophils, neutrophils. We shall not deal with these other cell types in this lecture).

Antibodies and B lymphocytes have become much more interesting than their routine protective and other housekeeping functions. As it has emerged, the life and reproduction of the B lymphocyte is a microcosm of biological evolution, and is thereby relevant to advances in genetics.

Immunoglobulin molecules are made up of four chains: Two identical light chains of about 220 amino acid residues and two identical heavy chains. Each heavy chain is twice as large as a light chain, and sometimes larger.[5] Each chain is divided functionally into two regions. The variable (V) region, which is the aminoterminal 100 or so residues, mediates antigen recognition and binding. The remainder of the chain is the constant (C) region. The variable region can be further sub-divided into four segments which vary very little from one antibody molecule to another. These are the framework regions. The remainder of the variable region is in three segments, called the hypervariable regions. These

hypervariable regions, as their name implies, are very much variable from antibody to antibody. In the folded antibody protein molecule, these hypervariable segments come together to form the walls of the antigen binding pocket, rather like the active site pockets of enzymes.

The division between variable and constant segments can be illustrated by considering a few words in the English language:

SPEEDING
SPENDING
BLEEDING
PLEADING
BLENDING

When we write these words immediately below each other, we notice several things. Each word can be broken into a variable region and a constant region. The last 4 letters are always DING - - this is the constant region of our set of words. The first four letters constitute the variable region. The C and V regions are of equal length, as in immnoglobulin (Ig) light chains. A closer look shows that the V region has a portion that is invariant. At position 3 each word possesses the letter E. Finally, note that in two of the other positions (1 and 4) there are three alternative letters, but at position 2, there are only two alternative letters. This is a fair approximation of the way in which evolution has designed antibody chains. The mode of diversity of these variable regions has profitably consumed biologists (not to speak of mice) in the last fifteen years or so.[6]

Apart from being secreted into body fluids, antibody molecules are also found as surface membrane structures on B cells. In these circumstances, they are cell surface receptors for antigen molecules. The interaction between antigen

and antibody, whether it be in body fluids or in cell surfaces, occurs as a result of a stereochemical fit. It is rather like a lock and key mechanism, the right antibody being able to "fit" the right antigen.

In the biological world, there is virtually a limitless array of antigens. This might imply a correspondingly limitless number of molecular types or species of antibodies, if the interaction between antibody and antigen was as specific as I have just implied.

* * *

When I started out in molecular biology ten years ago, I was involved with the study of the mode of diversity of antibody molecules. At the time, the four chain structure of immunoglobulins had been elucidated by Rodney Porter; Gerry Edelman was coming out with the details of the first complete amino acid sequence of an imunoglobulin molecule. (Both these workers[7] were to share the 1972 Nobel Prize for Medicine and Physiology). Roberto Poljak was still trying to construct x-ray maps of immunoglobulin fragments, so we did not have a coherent idea about the shape of the molecule in space[8] It was known by then, of course, that several structural gene clusters control the expression of the various classes and types of immunoglobulins. But the most important discovery at that time, as far as the evolution of antibody proteins was concerned, was that encoded in the phrase "two genes one polypeptide chain". Dreyer and Bennet had deduced from the fragmentary amino acid sequence data on light chains that the V and C segments are separately encoded in the genome.[9] This proposal was confirmed ten years later by the first sequence studies on mouse light chain DNA. The existence and function of hypervariable regions[10] were being contemplated by Kabat and his coworkers in 1971.

Now, to locate and situate ourselves within the biology scene in 1971: Fred Sanger had published his RNA sequencing technique two years earlier. Walter Fiers and Sanger were yet to publish the complete sequence of the coat protein genes of the RNA phages MS2 and R17 - these came in 1972. Of course, there was no chemical or enzymatic method to obtain DNA sequences. To be sure, nucleic acid hybridisations, a method of counting the numbers of copies of genes kinetically, had been popularised by Britten, Kohne, Davidsohn and others. But in 1971, there was a lot of interest in refinement of the kinetic equations to enable gene counting become routinely possible. Also, Boris Ephrussi in Paris and Harry Harris in Oxford had popularised somatic cell fusion studies; these were being used increasingly to locate specific genes encoding defined proteins and enzymes to individual chromosomes. But the hybridoma was to wait another 5 years to be discovered in our own Cambridge laboratories. The central dogma had just been, if you like, reversed slightly with David Baltimore, Sol Spiegelman, Mizutani and Howard Temin independently showing that RNA tumor viruses contain an enzyme which uses the viral RNA as a template for the synthesis of DNA. This discovery, which had a considerable bandwagon effect, has profound implications for the whole of molecular biology as well as for the mechanism of cancer induction by RNA viruses. Mammalian messenger RNAs were about to become common property in laboratories: Labrie had just employed conventional electrophoresis and centrifugation to purify rabbit globin mRNA. John Gourdon had convinced most of us by his elegant nuclear transplantation experiments that every animal cell, differentiated or otherwise, had the same genetic complement or potential. In 1971, discussions on moral dilemmas in biology focused mainly on death and dying, thenatology and euthanasia, and not as yet on genetic engineering.

At the same time, in 1971, the distinction between T and B lymphocytes had been firmly established. While the T cell and its products seemed like black magic, the B cell and its product, the antibody molecule, were enigmatic and provocative. As a problem in biology, the B cell had arrived. Just a few years after Francis Crick had made his pronouncement that molecular biology had solved just about all the fundamental problems in biology that it was capable of solving, there was a new rush of molecular biologists, mathematical biologists, biophysicists, chemists and other physical scientists to biology - all aiming to unravel the mode of diversity of the B lymphocyte and of the antibody.

* * *

The life and reproduction of B lymphocytes has turned out to be a good system for the study of those hereditary changes that are the basis of all biological evolution. First, an accident of nature has produced a disease state in which the plasma cell, the end cell of the differentiation of B lymphocytes, becomes a cancer and produces uncontrollably a lot of monoclonal antibodies. These tumors, myelomas, arise spontaneously in man and are easily induced in mice[11] The antibody proteins that the tumors produce can be purified and studied.

When these tumor cells of mice are put in cell culture, we found that 16-20% of their synthetic machinery is devoted exclusively to the production of immunoglobulin. This is remarkable when one realizes that over 10,000 different kinds of proteins are made in a cell. Secondly, immunoglobulin production is not known to confer any advantages to the cell. The cell does not require antibody for growth or reproduction. So, when we look at immunoglobulin mutations, we can more or less be certain of the non-interference of selection pressures.

Using an elaborate but very sensitive method, we were able to isolate, from screening 7,000 or so individual B cells derived from a single tumor clone, five different cells whose antibody molecules showed small differences in net charge from those of the parent. The differences shown in each case were not the same as in any other. Because we were dealing with mouse cancer cells in tissue culture, we were able to propagate each individual variant. We were also able to reinject these into the same strain of mice to form tumors, and to study these tumors and their protein products.[12] From detailed structural studies of protein and gene, we showed that the mutants were derived from the same 'wild type' parent.

I should here stress two points and with force. First, we had not used any physical or chemical agents to induce the changes that we had found. We were merely studying spontaneously arising variations. The whole object is to use the antibody system as a model in molecular evolution. Secondly, we had started out with cells, all derived from a single clone. Now all the members of a clone derive from a single cell. A clone is the most genetically uniform population of cells known. Within limits, the members of a clone are genetically identical; the limits being the occurrence in certain members of mutations, hereditable changes or errors in the base sequence of (and hence in the structure of and information contained in) DNA, such as the ones we are now talking about. I should anticipate myself a little and state that these mutations, rare and accidental processes, errors, if you like, are the basis of life in its evolution. They equally are the basis of the cycles of life (through ageing and its final consequences) and of biological accidents like cancers. Our study of spontaneously arising mutations in the absence of any known selection pressures could be said to be a study of an aspect of evolution right there in the test tube. Such mutation studies had been

carried out with profit in micro-organisms before, but the methods for molecular analysis of mutations in single cells in higher forms had not been available.

The mutations of immunoglobulin genes which we were able to detect and to analyse their structures both at the levels of protein sequence and nucleic acid structure showed a non-randomness. What I mean is this: we found that there are favoured sites for acceptable (and therefore detectable) mutations. Besides, some of these spontaneously arising mutations are analogous to processes that have been observed in nature. We think these are not mere coincidences.

For example, let us look at the structure of the mutant we christened IF4.[13] It has an altered base sequence in its gene, an A to G alteration in its structural cistron or mRNA. This corresponds to an altered protein amino acid sequence an aspartate for asparagine at position 415. The mutation occurs at a position that is exactly structurally homologous to those in human immunoglobulin light chains that bear the stamp of evolution. In human immunoglobulin lambda (λ) light chains, there is a lysine/arginine interchange at position 190 of the amino acid sequence. Each one of us has light chains some of which carry lysine and others arginine at that position, products of different segments of DNA (genes) that have diverged relatively recently in evolutionary time. Position 190 of the λ chain is an isotypic marker nicknamed *Oz*. Position 191 of human kappa (k) light chains contains (part of) a genetic marker named *Inv*. For the *Inv* marker, there are three alternatives, and their inheritance follows simple rules of Mendelian genetics. The point is this. Both *Oz* and *Inv* markers have arisen as a result of changes in DNA base sequence during evolution.

In IF4 we have the result of a *contemporary mutation in a test tube*, at a position exactly homologous to the markers *Inv* and *Oz*. Moreover, the mutation in IF4,

and the *Oz* and *Inv* interchanges, are all conservative. By this, I do not mean that they belong to the "postcursors" of the NPC, the Action Group, NCNC, NEPU or of the Borno Youth Movement, or that they are Reagannites or Thatcherites. What it means is this: The processes involve base changes in DNA that result in replacement of a similar amino acid, so that there are no drastic implications for the structure of the protein gene product. These observations tell us that there are favoured sites for "acceptable" gene mutations, evolutionary and contemporary, both in germ line and in somatic genes.

There is yet another interesting deduction, this time from the structure of our mutant named IF2. This has to do with carcinogenesis, or the origin and the mechanism of the formation of cancers. The structure of IF2[14] is strikingly homologous to the structure of heavy chain disease (HCD) proteins. HCD is another accident of nature, a B cell tumor that produces Ig heavy chain fragments. The abberant heavy chains produced by many of these tumours have an internal deletion in their amino acid sequence. When these deletions occur, and whatever the extent, these altered chains always regain normal sequence at position 215 (MOPC 21[12] numbering) of the wild type sequence. Our IF2 has an internal deletion of the whole of the Cγ1 pseudo-subunit of its normal prototype and regains wild type sequence at position Valine-215. We showed that it represents a mutation that is accompanied by a deletion of the Cγ1 exon. The implication of the structure and mechanism of IF2 for the mutation origin of cancers is inescapable. Here, we have a normal protein sequence "change" to an abnormal sequence, the type associated with a defined human cancer, by mutation, right there in the test tube. (incidentally, the structure of IF2 allowed us to precisely locate the end of the V region. This has since been confirmed by x-ray crystallography and DNA sequence studies)

* * *

Immunoglobulin gene expression has turned out to be quite complex, indeed. The cell uses quite a few tricks to ensure that only the right immunoglobulin genes are expressed by the right cells and at the right time. How this happens has puzzled biologists, until very recently. We had early in the game ascribed it to epigenetic changes, the kind of shuffling that is not well understood but which occurs in the genome in development. More is now known; it has turned out to be an apt lesson on DNA mechanisms in development.

To digress a little. There are two kinds of cells in our body, differing by the absolute number of chromosomes and by function. The germ cells, or sex cells, sperm and eggs, are responsible solely for (genetic) reproduction and continuity. The other specialised body cells, somatic cells, (for example, the specialised cells of liver, brain, muscle, heart, teeth, blood, lymphocytes, etc.) are more differentiated, and ultimately derive from the union of germ cells of different polarity. Germ cells have half the number of chromosomes as somatic cells: the number in a somatic cell is the sum for both germ cells giving rise to it. This is roughly the situation in animal cells. As for what is called differentiation, we have always taught our students that it does not involve differential partitioning of genes to different nuclei, but that it represents differential expression (or derepression) of genes; that differentiation does not involve the loss of genes. The conclusive experiments[15] on this line were carried out by John Gourdon in Christchurch College in the late sixties. He was able to transplant the nucleus of a differentiated cell (from brain or intestine) of a frog into an ennucleated frog egg. From this, develop-

ment through tadpole to adult frog resulted. So, the differentiated and highly specialised nucleus of the intestine or brain, has in it, afterall, all the genetic information of the whole body, although it may make use of only a few of these at its own leisure. Likewise, our differentiated liver cells do not make immunoglobulin, although they have the genes for Ig structures.

We now know why immunoglobulin genes are expressed only in certain cells, e.g. B cells, but not normally in others, like lung cells and liver cells, and why certain Igs are expressed in a B cell at one time and another Ig in the same cell at a later time. It has to do with how immunoglobulin DNA is arranged within the genome, and with what rearrangements take place within this DNA. Let us look more closely at the segment(s) of DNA that specify immunoglobulins.

Each heavy chain is now known to be coded for by four separate genes: a C gene coding for the constant region; a J gene coding for the fourth framework region; a D (D for diversity) gene coding for the third hypervariable region; and a V gene coding for the first three framework regions including hypervariable regions one and two[16]. This is the germ line arrangement of immunoglobulin genes. There is a limited number of V genes at some distance from the cluster of C genes in the germ line configuration. Of course, germ cells do not express Ig. When the Ig genes of B cells are studied, a different arrangement is found: There has been a series of rearrangements compared with germ line DNA

During B cell maturation, one of the V genes is brought in proximity with a C gene, mostly by excision (and deletion) of intervening DNA sequences. This translocation, as it has turned out, provides some explanation for some of the diversity observed in Ig sequences. The V region is itself constructed from three genes - the V gene, the D gene,

and one of four J genes adjacent to the C gene on the germline DNA.

The random combination of V, D and J segments together with the codon variation at the joins of these segments accounts for a lot of V region diversity. Point mutations at hypervariable regions one and two account for further diversity during the differentiation of the B lymphocyte. Our own studies and those of Martin Weigert and his colleagues had suggested that point mutations are important in the diversification of antibody combining sites. Besides. DNA sequence studies are now able to confirm some differences in hypervariable region sequences between germline DNA and differentiated somatic DNA. These differences turn out to be similar to the kinds of exchanges and replacements seen in the combining site sequences of antibodies, and are of the same nature as the spontaneously arising point mutations which we described.

Some of the impetus for the recent explosive activity in the molecular biology of eukaryotic genes has been the intellectual controversy around the somatic mutation model. We had posited that some immnoglobulin diversity arises from mutations in somatic cells, and are not represented in the information contained in germ line DNA. To many biologists, that a somatic copy or gene product could be different from the original germ line master is heresy. To have *somatic* as opposed to germ line diversification is just not cricket. Biologists are a conservative lot. That is precisely why Popperian falsifiability in its most naive form is hard to dislodge. In any case, many lines of approach - - using computers to compare Ig protein amino acid sequences; nucleic acid hybridisations to " count" Ig gene copies; (these are just two of the most sophisticated ones) - - had failed to convincingly disprove (or prove) somatic mutation. The situation was such that the same set of data was interpreted differently by somatic mutationists and germ liners,

each to confirm his own world view. So much for objectivity in science. However, things are now much clearer. Our evolution-in-the-test-tube experiments show that somatic mutations do occur in immunoglobulin genes, and suggest that these mutations contribute to antibody diversity. Besides, these mutations are stable. DNA sequence studies confirm this.

I should stress another point. The workings of our somatic mutation model do not controvert or contravene Weinstein's principle of the inviolability of germplasm. It is not a facet of what some have described as a Lamarckist rebirth, nor is it a contribution to the saucy factions in Sociobiology. The available knowledge cannot support the transfer of somatically acquired traits to the germplasm as a routine mode of inheritance. The germplasm is not and cannot be involved because of certain tricks the cell uses.

* * *

During the life of a B lymphocyte, the first immunoglobulin secreted is of the IgM class, expressing μ chains. As the immune response progresses, the antibody producing cell now starts expressing another immunoglobulin class. This it does, retaining the same V region, retranslocating the same V region on to a different C gene. This is known as the class switch. It is now known that the class switch or retranslocation event is a DNA strategy. Our own structural analysis of the immnoglobulin and Ig gene of the mutant IF2 indicated to us in 1975 or so that translocation and retranslocation occur at the level of DNA. That conclusion provoked controversy and a lot of discussion at the time: others had their pet theories. What is now also firmly established is that the heavy chain switch involves the deletion of segments of DNA coding for C

regions. The deletion of the constant region genes[17] including $C\mu$ occurs in a way which is consistent with a C_H gene order μ-γ3-γ1-γ2b-γ2a- etc and which is consistent with a model in which the genes preceding the expressed gene are deleted. So, the differentiation of immunoglobulin genes involves deletions of segments of DNA resulting in rearrangements of the DNA to configuration that are different from the germ line arrangement. This reminds us of the process of "activation" of some enzymes and hormones, which involves the excision of segments of the primary protein sequences.

The loss of gene segments in the ontogeny of B lymphocytes is quite an interesting phenomenon. In 1975, after comparing the protein amino acid sequence I had derived for the mouse γ1 chain with those available for human and other sequences, we came up with a model of gene expansion-contraction[18] in the evolution of constant regions. Gene expansion, or more correctly, the expansion of gene pools, (or of isologous genes) is easy enough to grasp. It represents the series of gene duplications that have occured in evolution which have given us, for example, apart from myoglobin and the cytochromes, the different haemoglobin classes that have all risen from a putative common ancestor[19] This process of gene duplications has been postulated to have given rise to immunoglobulin domains, all deriving from a common ancestor. We did not think that told the whole story. There are immunoglobulin classes in mouse, for example, whose structural (sequence) homologues are not found in man. Also, we have found that our own African land tortoise shows a remarkable restriction of immunoglobulin types. If immunoglobulin sub-classes evolved 70 million years ago by a series of duplications, it would then seem that one likely reason for these differences has to do with a process whereby genes have got lost

(or deleted) through evolution. This we call the process of gene contraction.

* * *

One of the numerous problems posed by B Cells is how they contrive to express the immunoglobulins on only one of the two homologous chromosomes of a diploid cell. Diploid cells, like the body cells of you and me, have chromosomes in pairs, one from the father, the other from the mother. These two chromosomes coexpress, except in the case of immunoglobulin gene expression. In no case has a myeloma, for example, been shown to express more than one functional chain of each type. In our own mutant immunoglobulin from MOPC 21 IF4 cells, whose structure we looked at both at the protein amino acid sequence level and at the nucleic acid sequence level, we found no evidence for the wild type component in the mutant.[20] This led us to characterise Ig genes as " effectively monosomic". This means that these genes behave as if only one copy - the father's or the mother's is present in each cell. (Another reason we thought Ig genes were effectively monosomic was indirect. Our own studies show that the mutation rate of immunoglobulin genes is of the same order of magnitude as those for genes in bacteria, which are monosomic, and is about the square root of the rate for two copied eukaryotic genes).

The reason for this behaviour is now clear. In some myelomas, the Ig genes on the " excluded" chromosome remain in the germ line configuration, without the necessary rearrangement of the DNA. In others, they do not; the cell simply uses one of several methods at abortive rearrangement.[21] In this kind of rearrangement, an error is deliberately introduced so as to make the the gene segment " useless". The situation is analogous to that of the sex chromosomes in a woman: although there are two X chromosomes, only one functions in each cell; the other is inactivated.

This much is clear: That in cells which express immunoglobulin genes, there is *effective* rearrangement in only one of the two chromosomes. In cells that do not produce antibody, the Ig genes remain in the germ line configuration and are not " activated". It also follows from this that the events we described as the " class switch" occur on only one chromosome.

One last problem for mention. We said earlier that Ig molecules also serve as cell membrane structures, in which context they are receptors for antigen. It was always thought that secreted antibody and membrane surface Ig were *identical.* Well, in 1978, we started building some models, using computer, on how different Ig structures would fold into a spatial configuration. One problem became apparent quickly: that the interaction between the carboxyterminal end of secreted Ig (whose structure we knew) and the cell membrane that was supposed to anchor it would not be stable. This is because the membrane anchor is hydrophobic (water-not-loving) while the carboxy end of secreted Ig is hydrophilic (water loving). Unable to come up with a suitable enough mode of coexistence between our Ig and the plasma membrane, we reluctantly came to the conclusion that secreted Ig and membrane Ig are of different forms. This, I must emphasize, was based purely on theoretical model building. We then went into the laboratory to confirm our suspicion. Using two lines of evidence - products of hybrid cells we had constructed for the occasion and CNBr fragmentation pattern of membrane Ig - we showed that there was a difference between the secreted and membrane forms of Ig. By late 1979, we were able to locate the differences to within the last 50 or so residues of the molecule.

These observations have now been confirmed. Besides it has recently been shown[22] that the same stretch of DNA is used for these two different molecules. How does the cell do it? It simply splices the same piece of DNA in two different fashions: one for secreted Ig; the other for surface Ig. That way, you have two *different* proteins from the same "gene". Benzer's "one gene, one polypeptide chain" has been a good reflection of the situation as we know it, until now. It is now more useful to talk of transcription units and replication units. With what we now know, with what strategies the cell has used in the evolution of complexity, these terms are more useful and less ambiguous than "genes".

To summarise: In recent years, research into the molecular mechanisms of the life of B cells has led to decisive impacts on eukaryotic cell biology. It has led to the clarification of the role of gene mutations in carcinogenesis; to the discovery of new evolutionary mechanisms; to new methods of studying cells of higher forms; and to the definition of a new concept of the gene, among other discoveries.

III

Bishop Wilberforce might have posed a slightly different question to Huxley: That, since mitochondria and ribosomes travel with ova, might Huxley not agree that his descent from the apes was more likely to be from the mother's side? Today, however, a much different question suggests itself. Sequences; DNA; cloning; hybridoma; splicing; etc. What is the point of all these? Especially as the problems that are with us daily are malaria, malnutrition; hypertension; diabetes; sickle cell. Is a knowledge of the way(s) immunoglobulin genes are spliced about to contribute to our "Health for All by the Year 2000", or is it about to bring closer a cure for the common cold? Is the study of

the organisation and strategies of DNA useful, or is it merely eristic?

True, it is difficult to get emotional about how four bases, A, T, G, C, are strung together. However, unlike in the physical sciences, advances in biological knowledge can still powerfully affect progress in applied areas like medicine (including pharmacy) and agriculture. To be able to manipulate molecules and cells to our own benefit, we have to understand the strategies of these molecules and cells. I should now give some illustrative examples to show that millions of people need this knowledge, now and in future generations.

1. It has now become fashionable to regard cancers as *mutations*.[23] This however has not always been so. One station on this road was our own analysis of the structure of the mouse immunoglobulin mutant protein IF2. Its structure was so strikingly similar to the products of the human cancer heavy chain disease (HCD) protein. The lessons drawn from this remarkable homology are inescapable. It now becomes possible to rationalise chemotherapy and other treatment modalities for cancer.

2. Diabetes in all its forms occurs in about 10% of the human population. The best characterised variant requires insulin. Usually, hog insulin is used as replacement with its own deleterious side effects. There just is not enough human insulin from necropsy pancreases to go around. In the last two years, the human insulin gene has been isolated, sequenced and cloned.[24]. Human insulin is now available in commercial quantities from bacteria into which the human insulin gene had been inserted. *E.coli*, that enteric pathogen, may yet turn out to be man's best friend!

The use of gene transfers is not limited to problems of human medicine. Currently, there is a lot of interest in exploiting gene transfers in microorganisms and plants impor-

tant to agriculture, and in the search for new food and energy sources.

3. In the last few years, there has been some sensation about the possible anti-tumor activity of interferon. It can only be used in pure form, and it has hitherto been available in hopelessly small quantities. Recently it has become possible to obtain active interferon through bio-technology. The method uses a combination of cloned production by engineered bacteria and purification using antibodies obtained by the hybridoma technique.

Again, cloned antibody production is being used in other areas apart from the purification of difficult molecules. It has made available (and cheap) important, well-defined antisera used in the clinical diagnostic laboratory.

4. To come nearer home, malaria, even today, is responsible for the death of one million black African children annually. Apart from this, we have immense human debility and economic waste from its morbidity. In spite of numerous programmes aimed at its eradication (aerial sprays, chemoprophylaxis, drainage systems etc.) malaria rages on. In some areas, India and Thailand, for example, the strains of the parasite that have emerged post "eradication" are more resistant to drugs than before, making malaria a more dangerous problem. It would seem like we need more than eradication programmes.[25]

A molecular biologist's look shows that many of the parasite's coat proteins are capable of generating an immune response. Indeed, one of these protein spots could be the key to the immunoprophylaxis of malaria. A strategy could be along this line: isolate the transcription unit corresponding to a protective" antigen protein; package this into *E.coli,* which should oblige with synthesis of the antigen; purify enough of the antigen to go round for vaccination.

It becomes even more compelling to apply this new knowledge about the gene and these new techniques to genetic diseases. One very important genetic disease is sickle cell anaemia. In this disease, globin, the protein portion of haemoglobin, has an altered structure. All there is to sickle haemoglobin is this slight alteration in its structure which derives from an altered globin gene. Now, the isolation of human chromosome fragments that encode globins has become commonoplace. The following programme sounds reasonable: culture bone marrow cells, the cells that are responsible for making haemoglobin; transfer "good" globin genes into these; repopulate the bone marrow with these engineered cells. That way, the synthesis of normal haemoglobin is set in competition with that of the abnormal variant. This should place patients in various degrees of functional heterozygosity.

* * *

There is no doubt that the modern excursion into the molecular mechanisms of life has in its store large spinoffs for parasitic and infectious diseases, cancer, genetic diseases, as well as for birth defects, ageing degenerative diseases and mental illness. However, the use of our new knowledge and technology poses very important questions for society, questions which are so important that they cannot, must not, be entrusted to specialists and politicians. These are questions which should be discussed openly and in every community. The questions and moral dilemmas could be the subject of a separate lecture; these can however be sketched thus : What is the genetic burden on the society, now and in the future? What are the social costs? The cost-effectiveness? Who pays? Who should benefit? Who decides who benefits? Who decides who decides?

In whatsoever way these questions may be resolved, it is certain that the human future will be quite different

from the human present, and this, mainly because of man's spirited explorations along biology's frontiers. Our future evolution will probably be mostly cultural and artificial-genetic. We should have to assume some godlike prerogatives as we become self-annointed trustees of our own evolution.

IV

The aim of science is the acquisition of knowledge. The goal of science is, according to physicist Jean Perrin, " to explain the complicated visible by some simple invisible". The goal of modern biology is the reduction of teleology to causation. We have seen how apparently complex processes of the B lymphocyte can be understood at the molecular (genetic) level, in simple terms of the various DNA mechanism: mutation; deletion; splicing; rearrangement; joining. Of course, we all know that genes are made of DNA, whose molecular configuration is known in detail. It is known how four fairly small chemical bases (A, T, G, C,) are strung on two paired strands of DNA and how the arrangement makes for accurate copying during replication. Many life's processes are now explainable in exquisite detail, starting from the gene.

As an example, to explain sickle cell anaemia, we may start from the gene. In DNA, the second letter of the three-letter code word that is used for the sixth position of one of the protein chains of haemoglobin is altered. Sickle haemoglobin thereby carries the amino acid valine in that position in its protein, rather than glutamic acid. This is a nonconservative replacement which alters the configuration of the haemoglobin molecule, especially when it is not carrying oxygen. These configuration changes are known at atomic resolution.[26] If there are enough sickle haemoglobin molecules that are not combined with oxygen, these molecules tend to stick together. They form a sickle polymer, which

is no longer soluble in the cell. The polymerisation distorts the cell into an elogated sickle form. These sickle cells become rigid and may obstruct blood flow in capillaries. This in turn starves the tissues of oxygen, allowing for more haemoglobin that is not combined with oxygen, and so on, and so on. All the problems go back to a structural change in the molecule, directed by a change in a structural gene.

Such levels of explanation can be made for many biological processes, from antigen-antibody interaction to enzymic action; from muscle contraction to hormonal action: from receptor signaling to pain; etc. etc. In all these, simple laws of physics and of chemistry are obeyed. In the analysis of the cell, which is the atomic unit of the organism, no laws outside of those of physics and chemistry are required. The basic sequences and structures of life's processes follow from those of dead nature without the intervention of any special powers or acts.

Let us look at our DNA again. DNA sequences, any DNA sequences, are capable of replication, of copying themselves. This is the sole purpose of DNA, to make copies of itself, so as to survive within the genome. This it fulfils very well, within certain limits of tolerance.

DNA molecules have little else to do but to joust with each other to get ahead. Much DNA has no phenotypic expression. Much of this DNA, some of which we have talked about as intervening between and within the coding sequences, has no known function in the genome; they may be just no more than efficient self-replicators. This DNA makes little contribution to the organismal phenotype. Much of this DNA may therefore be considered as junk; selfish junk. It is " selfish" DNA. Selfish, because while it uses the cell's energy resources to replicate, it may contribute nothing to the life of the cell. Such DNA is truly parasitic[27] besides, many of these species are capable of interchromosomal movement.

The DNA of higher organisms consist of a minority of sequences with highly specific functions plus a majority with little or no specificity. The latter category, which includes transposable genetic elements and middle-repetitive sequences, apparently conveys little or no selective advantage on to the organism. This mandatory replication of selfish DNA has an analogue in an axiom of high energy physics: what (state transition) is not forbidden is mandatory.

We can now look at selection at two different levels - at the level of the gene, and at the level of gene product. At the level of the gene (DNA), those sequences that are efficient replicators will survive in genomes. Most of these apparently have no specific function; so no design or purpose is obvious in this kind of selection. At the level of gene product, we have roughly the same kind of neutral process, neutral from the viewpoing of the *functioning* of the cell or organism. What is selected for are those molecule that are able to survive within the cell, not necessarily those that are the most efficient functionally[28]. Those molecules that can exist in contrapuntal harmony with their own micro-environment are not necessarily those with the "best" functions. Selection, therefore, has to do with the (differential) survival of molecules within cells; in that sense, it is molecular. Evolutionary changes, the hereditable changes in DNA base sequence, are themselves neutral and occur in a stochastic manner. Therefore, the important distinction about life is *not* that it is complex or that it is different in its governance from inanimate matter. The distinction is that it is quite improbable, quite accidental and therefore unique. It is indeed quite improbable that so many different molecules, different structures, have evolved in the cell and act together in such harmony in a process we call life.

All evolutionary changes result from hereditable changes in DNA molecules. We have seen how some of these have

come about. Ironically, Queen Victoria, the supreme head of Bishop Wilberforce's church, was carrying such a mutation in her genes. She faithfully passed on to her descendants the mutant gene for the antihaemophiliac factor. There is another more profound irony, however.

Life is a process. The study of the cell has to be the study of a changing structure. Life is not an accurate copying, as is neatly carried out in the geometrical scaffolding of a dead crystal. Life is an evolutionary process, which moves forward because there are accidents or errors in the DNA copy. Once in a while, one of these errors is successful enough to be incorporated as another step in the progression. That is evolution. Essentially the same process occurs in the cyclical nature of life, the time-dependent progression of living forms. We live, grow old and die from disease or "old age". Ageing is an accumulation of these errors, mutations, the same kind of errors that are responsible for the progression of evolution, the same quantum effects that are essential to explain the uniqueness of a living form. Ageing, he correlate of the cyclical (as opposed to the evolutionary and open) process of life is an accumulation of mutations, the accumulation of these errors ultimately leading to a cellular catastrophe - cell death. And this is the irony: The errors which destroy the individual are also the origin of species. Therefore the nature of life is expressed in its perpetual evolution - it is the succession (and the success) of its errors.

In physics, it is commonplace to think of any material body as an arrangement of a large number of atoms, repeated in some regular manner, and to explain its properties and behaviour by going back to its basic units. We can build the same model for biology. Is the organism a sum total of its chemical processes? Can life be explained by its component processes and phenomena? Or is life, represented by the living cell, a product of a manifest design? Is there a place

for vitalism in biological explanation?

Vitalism is a persistent belief which holds that the laws of inanimate physics, or, rather, the *kinds* of laws operative in inanimate physics, will not suffice to explain the phenomena of life. Vitalists find two types of arguments against our own materialist, mechanistic approach.

First, they hold that the cell or organism functions in a way that physics cannot explain. This implies the existence of another kind of laws[29], for example Walter Elsasser's "biotonic laws". Bilingbroke, early in the eighteenth century, similarly ascribes this function of higher co-ordination to God. Also, vitalists note the apparent direction of evolution in time, that evolution is usually towards complexity and not in the reverse direction. Does this not violate the statistical provisions of the Second Law of Thermodynamics? they will ask.

The design of a watch is the classical illustration of God's design in man.[30], William Paley, Henry St. John and Viscount Bolingbroke in *Evidences of Christianity* used it to claim that man is a more ingenious machine than is a watch and so must have been created by a more ingenious creator. Michael Polanyi gives this argument a new look by saying this: that just as the design of a watch points to and is only understood in its purpose, so the *design* of the machinery of life points to and is only understood as a higher level of explanation by its purpose. The mechanism, then, must fit into and serve some overall plan outside of itself. The watch, for example, is designed to tell the time.

This is a curious argument. It is intended to show that living forms are not mere machines. In order to do this, living forms are compared with a typical machine, namely, a watch. It is then concluded that man is more purposeful than the mechanism that drives the watch. But also, the watch is more purposeful than its own mechanism. In short, even the watch is not merely a machine:

Man, therefore, is not a machine because he (like the purposeful watch) is also a machine that is not a machine! I do not intend to belabour this point.

One of the more frustrating experiences I have had here in Ife as a teacher of medical biology has been to get my students to even *consider* that evolution may not have been designed to fulfil a designer's purpose of bigger and better things anthropocentric. And that there is no evidence that "fitness" is ever maximised; this incidentally constitutes serious disagreement between us and the optimalist schools of sociobiology. To most students around, life is a purpose to be fulfilled, and its molecules are under the direction of an all-wise evolutor, who probably understands the Second Law of Thermodynamics and the Laws of Chance much better than we do. This variant of prime movement is paramount in everyday thinking, and is irrational.

The direction of evolution, which can be traced for three thousand million years, is consistent, and gives it the appearance of a planned program. It may appear to conform to a master plan, some plan larger than the laws of physics. In a history of three billion years, evolution has not run backwards. Is it possible to have such a mechanism that is not planned? Or, how can disorder on the small scale be consonant with order on the large scale, in time and in space?

We can explain this problem by looking at *level of order.*[31] First, in physics, in the evolution of chemical elements, a process that does not require the intervention of selective forces, the chemical elements are built up in different stars, step by step: hydrogen to helium to carbon, and on to higher elements. The encounter of two hydrogen nuclei makes helium. Each helium nucleus, which is stable, is now used as raw material to build up higher elements. The wildly improbable encounter of three helium nuclei builds carbon. The more complex stratum is built on simpler ones.

The more complex stratum is built on simpler ones.

Similarly, the levels of order in a cell may be enumerated as atoms, bases and amino acids, nucleic acids and proteins, genes, etc. With genes, there is another hierarchy, another set of levels of order. Each level is made up of stable structures. And the cell itself is stable as a *topological* structure in space and time. Cells build up tissues, organs, organisms.

Natural selection establishes the stability of each level of order. Each level of order is built upon the next lower level. If a back mutation occurs, it may be inappropriate if it does not fit into the level of stability which the system has already reached. For this reason, it does not reverse the direction of evolution.

One last question that some vitalists may hold on to is this: Does the interaction of genes on the same or on different chromosomes require any kind of master law? It is true that there are what we call "supergenes": genes that control the expression of other genes. There are other master genes which control groups of other genes by making them all more mutable or more stable. Recently, the interaction between certain genes, for examples between the *cro* and *lambda* repressor genes[32], has been explained in detail at the molecular level. These interactions follow the simple rules of physics and chemistry. There is no reason to suspect that gene action will need other kinds of laws to explain. Indeed, no special powers or acts intervene, no laws outside of those of inanimate physics operate in life's processes. There is no place for vitalism in the analysis of the cell.

Scientific study of the cell and of cellular processes will no doubt continue to unravel the mysteries of life. It is however pertinent to ask if there are any limits to such explanation. For example, can molecular studies of the nervous system ever resolve the mind-matter paradox?

Can there be a molecular mechanism for consciousness,[33] a very private experience? Is the brain capable of providing an explanation for itself? We should leave this as a question. We should also stress that this question cannot be a loophole for vitalism. There may be limits to human understanding, limits imposed by the level of evolution of human structures, including the processes of the brain. We should also conjecture that should our own brains solve the mind-brain problem, then the barriers between the social sciences and the physical sciences, and indeed between all the branches of knowledge, would have withered away![34]

Ile-Ife, May, 1981

REFERENCES AND NOTES

1. Barbour, I.G. (1970) *Myths, Models and Paradigms* (SCM Press, London).
2. Halstead, B. (1980) *New Scientist* 87, 215-217; Popper, K. (1968). *The Logic of Scientific Discovery* (Hutchinson Press, London).
3. Sardar, Z. (1979) *The Future of Muslim Civilisation* (Croom Helm, London). It however can be argued that science is a corpus of knowledge that has no geographical or national boundaries, i.e. that there is no such thing as "Western" science. This view is not entirely correct. Whereas there *ought not* to be an entity "Western" science as opposed to Science, there exists a certain corpus characterised by its own peculiar myths and models, methodologies, explanations and cult practices, and distinguishable from, say, Chinese science. For example, the scientific nature of acupuncture was established *by concensus* in Chinese science centuries ago; until very recently, Western science was not capable of accomodating acupuncture.
4. Lakatos, I. (1970) in *Criticism and the Growth of Knowledge*, Lakatos, I and Musgrave, A., eds. (Cambridge University Press London), pp. 91-196; Fayerabend, P. (1970) *ibid*, pp. 197-230; Fayerabend, P. (1975), *Against Method*, (NLB, London). Sir Karl Popper's views on evolution are based on mistaken ideas, and he is apparently unaware of population–genetic theories that involve natural selection. Also see: Popper K. (1978) *Dialectica* 32, 344; Ruse M (1977) *Phil. Sci.* 44, 638; Wasserman, G.D. (1978) *Br. J. Phil. Sci.* 29, 223.
5. See Gally, J. (1973), in *The Antigens* (Sela, M. ed.) Vol.1, pp. 161-298 (Academic Press, New York); Hood, *L.* (1973) *Proc. Stadler Genetics symp.* 5, 73.
6. Many volumes have been devoted entirely or almost entirely to the origins of lymphocyte diversity. Two readable examples will suffice: *Cold Spring Harbor Symposia on Quantitative Biology* (1977), vol.41; Cunningham, A.J., (ed.), *The Generation of Antibody Diversity*. (Academic Press, London).
7. Porter, R.R. (1973) *Science* 180, 713; Edelman, G.M.(1973) *Science* 180. 830

8. Poljak, R.J. *et al.*, (1974), *Proc. Natn. Acad. Sci. U.S.A.* 71, 3440-3444; Poljak R.J. (1975). *Nature* 256, 373-376. The three dimensional structure of immunoglobulins is now known at atomic resolution. How the molecule is folded can be related to immunoglobulin genetics and evolution.
9. Dreyer W.J. and Bennet, J,C. (1965) *Proc. Nath. Acad. SCi. U.S.A.* 54 864-869; Hood, L. and Talmage, D.W. (1970) *Science* 168, 325-334; Hozumi, N. and Tonegawa, S. (1976) *Proc. Nath. Acad. Sci. U.S.A.* 73, 3628-3632
10. Kabat, E.A. and Wu, T.T. (1971) *Ann.N.Y.Acad. Sci.* 190, 382-391; Wu, T.T. and Kabat, E.A. (1970) *J. Exp. Med.* 132,211-251.
11. Potter, M. (1972) *Physiol. Revs.* 52, 631-719; Also see the UICC Technical Report (No.13, 1974), *Multiple Myeloma and related immunoglobulin - producing neoplasms.* In man in a neoplastic condition called multiple myeloma, proliferating plasma cells produced structurally homogeneous immunoglobulin molecules The study of these tumours, their genes, gene products, receptors and associated signals has, more than study of any other eukaryotic system, increased our knowledge of eukaryotic molecular biology, especially in the last fifteen years or so.
12. Adetugbo, K., Milstein, C., & Secher, D. (1977). *Nature* 265, 299-304; Adetugbo, K., & Milstein, C. (1977) *J. Molec. Biol.* 115, 75-90; ——.(1978), *J.Molec. Biol.* 121, 239-254; Adetugbo, K., (1978) *J.Biol. Chem.* 253, 6076-6080; Milstein *et al* (1977), *Cold Spring Harbor Symp. Quant. Biol* 41, 793-803.
13. Adetugbo, K. & Milstein, C. (1977) *J. Molec. Biol.* 115,75-90.
14. Adetugbo, K. (1978). *J. Biol. Chem.* 253, 6076-6080; Milstein, C. *et al.* (1974). *Progr. Immunol.* (2)1, 157-168. However, whereas both the mutant plasmacytoma IF2 and its wild type parent MOPC 21 secrete the four chain immunoglobulin protomer as their tumor cell product, the human heavy chain disease tumors altered heavy chains without light chains.
15. For an interesting and eminently readable account fo these experiments, see Gourdon, J.B. (1974). *The control of gene ex-*

pression in animal development (Oxford University Press). Chapters 1 and 2 are especially relevant.

16. Tonegawa, S. *et al.* (1978). *Proc. Nat. Acad. Sci. U.S.A.* 74, 3518-3522; Molgaard, H.V. (1980) *Nature* 286, 657; Dunnick, W. (1979) *Nature* 281 253; Maki, R. *et al.* (1980) *Proc. Nat. Acad. Sci. U.S.A.* 77, 2142; Sakano, H. *et al.* (1979) *Nature* 286, 676; The description of immunoglobulin molecules in the text is that for heavy chains. Immunoglobulin light chains are encoded by *three* separate genes – a C gene coding for the constant region; a J gene coding for the fourth framework region and part of the third hypervariable region; and a V gene coding for the first three framework regions including hypervariable regions one and two, and part of hypervariable region three. There is no D gene.

17. Rabbitts, T.H. *et al.* (1980) *Nature* 283, 351, 356; Sakano *et al.* (1980) *Nature* 286, 676-683; Davis, M. M. *et al.* (1980) *Nature* 283, 733-739; Yaoita, Y. and Honjo, T. (1980) *Nature* 286, 850-853. In the germ line arrangement of heavy chain DNA, before recombinational events, for example, there is, first, some 50 or so genes, each with a short leader or "precursor" sequence, coding for part of the variable region (V genes). Then, there is a cluster of D gene segments, which code for most of the third hypervariable regions. Some distance away from the D segments there are four J segments which complete the V region coding sequence. Each of these segments and the different coding sequences (or "genes") constituting each segment is separated from others by intervening sequences or introns. The J segments lie about 8 kilobases from the Cμ gene, which lies at the start of a cluster containing all the C region genes. Again, the C region sequences are interrupted by non-coding sequences. This interruption gives rise to a series of exons corresponding to the domains (pseudosubunits) and the huige region. Thus, two types of recombinational events are necessary to obtain a mature immunoglobulin – a VDJ recombination with Cμ gene to give a complete μ chain transcription unit, and a heavy chain switch which involves the deletion of C μ and all other C segments that are necessary to place the VDJ segment next to the appropriate C gene. Following transcription, the introns are removed by splicing leading to a continous coding sequence in the mRNA.

18. Adetugbo, K. *et al.* (1975) *Eur J. Biochem.* 56, 503-519; Adetugbo, K. (1978) *J. Biol. Chem.* 253, 6068-6075.

19. Archer, R. (1974). *Angew. Chem. (Internat. Edit.)* 13, 186-197.

20. Adetugbo K. & Milstein C. (1977) *J. Mol. Biol.* 115, 75-90; Adetugbo, K. *et al.* (1977) *Nature (London)* 265, 299-304.

21. Coleclough, C. *et al.* (1980) *Proc. Natn. Acad. Sci. U.S.A.* 77, 1422; Perry R. P. *et al.* (1980) *Proc. Natn. Acad. Sci. USA*, 77, 1937; Seidman, J. & Leder, P. (1980) *Nature* 287, 603.

22. Singer, P.A. *et al.* (1980) *Nature* 285, 294-300.

23. McCann, J. & Ames, B.N. (1976) *Proc. Natn. Acad. Sci. U.S.A.* 73, 950-954. Cairns, J. (1981) *Nature* (London 289, 353-357

24. Bell, G.I. *et al.* (1980) *Nature* 284, 26-32.

25. The WHO recently had a rethink: epidemiology should now become only part of its "fight" for health in Third World countries; science, especially modern biology, is now allowed a role in the strategy against tropical diseases. WHO's imperialist tactics remain however: the crucial modern biology and other technical aspects of tropical diseases may only be tackled in nontropical metropolitan countries. But this is in the nature of things. Molecular biology or no, just like the sociological (or "eradication") programs before it, meaningful progress cannot be achieved in the fight against the tropical parasitic diseases of tropical peoples until the defeat of imperialism, which defeat will make the WHO irrelevant.

26. Perutz, M.F. Lehman, H. (1968) *Nature* (London) 219, 902-909. Sickle cell haemoglobinopathy is a major public health problem in Nigeria. The main approach currently has been the search for chemicals which can modify the configuration of the S-haemoglobin molecule, making it less "sticky". There have not yet been many remarkable advances in this area (see Benesch, R. and Benesch, R. (1980) *Nature* 289, 637). The "genetic engineering" option remains speculative.

The effects of a successful genetic engineering approach on the population has been questioned. The S gene is supposed to have survived because of "balanced polymorphism" – As heterozygotes surviving malaria much better than normal individuals. Today, decidedly most of the S gene is carried in heterozygotes, who are also responsible for just about all its perpetuation. A strategy that allows a normal life for SS homozygotes and enables them normal reproductive activity has the same effects on society as the introduction of insulin replacement therapy for diabetes. The effect on gene frequence is small.

27. Dawkins, R. (1976) *The selfish Gene* (Oxford University Press); Doolittle, W.F. and Sapienza, C. (1981) *Nature*, 284, 601-603; Orgel, L.E. and Crick, F.H.C. (1981) *Nature* 284, 604-607; Orgel, L.E. Crick, F.H.C. and Sapienza, C. (1981) Nature 288, 645-646; Dover, G. and Doolittle, W.F. (1981) Nature 288, 646-647. There has been a lot of debate about the concept of "selfish", "parasitic", "ignorant" or "incidental" DNA, and about its relevance for modes of genome evolution. Some things are however clear:

(1) Such DNA arises when a DNA "spreads" by forming more copies of itself within the genome; such amplification and dispersion may be sequence-independent or sequence dependent;

(2) It makes not highly specific contribution to the organismal phenotype. The self-perpetuation and self-selection of these DNA sequences lead to certain types of repetitive DNA. The relative contributions of such accumulated DNAs to the evolution of eukaryotic genomes is however still in dispute.

The bulk of eukaryotic DNA has no known (phenotypic or nucleotypic) function. The usual assumption that they are regulatory is unsatisfactory. DNAs with no immediate phenotypic or nucleotypic benefit are of no immediate selective advantage.

28. Compare the ideas adumbrated in Adetugbo *et. al.* (1977) *Nature* 265, 299-304 with those in Jacque Monod's 1972 book *Chance and Necessity*.

29. Elsasser, W.M. (1958) *The Physical Foundation of Biology* (New York).

30. Polanyi, M. (1967) *Chem. Eng. News*, 45, 54-66; *Science* 160, 1308-1312.

31. (a) see notes under Ref. 33 below.
(b) the arguments on strata of stability or levels of order have been outlined by Bohm, D. (1967) in *Towards a Theoretical Biology* (Waddington, C.H., ed; Edinburgh) pp. 18-40.

32. Plashne, M. *et al.* (1980) *Cell* 19, 1-8; Maurer, R. *et al.* (1980), *J. Molec. Biol.* 139, 147.

33 Erwin Schrodinger, in his 1945 book, *What is Life?* (Cambridge University Press), took the position that physics and chemistry can account for all the processes that take place in a living organism. The peculiar quality of living matter, namely, that it creates order from disorder, in apparent violation of the Second Law which asserts that, in the Universe as a whole, order decays into disorder, he explains by the fact that life feeds on the gigantic decay processes that occur in the sun.

Niehls Bohr in his address "Light and Life" in 1932 (published in *Nature* 131, 421, 457 (1933) thought that it would be wise to at least keep in mind the possibility that in the study of life, phenomena might not be wholly accountable in terms of conventional physical concepts; that atomistic features in the functions of biological organisms may not be sufficient for a comprehensive explanation of biological phenomena.

The view expressed here is not identical to Bohr' "irrational element". Rather, it has to do with the ability to integrate very large numbers of information, and fast enough, to make meaning.

34. I thank many friends, colleagues and students for their criticism of the early draft of this address. Of course, we collectively share responsibility for any errors of omission or of commission or of expression contained therein.