PRESIDENTIAL ADDRESS AND SYMPOSIA / DISCOURS PRÉSIDENTIEL ET COLLOQUES

W. H. Cook, President, 1962–63

TRANSACTIONS OF THE ROYAL SOCIETY OF CANADA
VOLUME I: SERIES IV: JUNE, 1963

DISCOURS PRÉSIDENTIEL 1963 PRESIDENTIAL ADDRESS 1963

Toward the Margin of Life
Vers la lisière de la vie

W. H. COOK, F.R.S.C.

LE Président de la Société royale du Canada doit prononcer, à chaque année, un discours présidentiel. C'est une tâche rendue très difficile par la variété et l'étendue des disciplines représentées dans cet auditoire distingué.

Le choix du sujet est laissé au Président, et j'ai cru bon de vous parler des réalisations des autres, pour élargir le sujet et le rendre plus intéressant. Plusieurs d'entre vous ne sont probablement pas tout à fait au courant des recherches qui se poursuivent aux frontières des sciences biologiques. J'ai cherché à tenir compte de cette situation.

For some this presentation may serve as an introduction to tomorrow's symposium dealing with biology at the molecular level. May I now request the forbearance of the specialists who have carried the study of biology to its present margin.

Historically, physical science started with a study of matter *en masse* and gradually moved to units of progressively smaller dimensions until it reached the particles in the atomic nucleus. The first world-wide scream of the newborn atomic age broke through the barriers of secrecy in 1945, and society has been trying to control this infant ever since. About the same time life science entered the wonderland of diminishing dimensions and hastened down past the size of Mr. Carroll's Alice to the level of Mr. Maxwell's demon, who played handball with hot molecules.

This new branch of biological science, called molecular biology, is conducting studies on the borderline between the animate and the inanimate. It has grown up quietly, does not threaten human existence, requires a comparatively modest level of financial nutrition, and nature is its only security officer. In consequence, molecular biology has made little impression on society, but it has been well recognized in scientific circles. During recent years the number of individuals who have been awarded Nobel Prizes for their work on molecular biology matches or outnumbers those awarded to individuals working in any other comparable branch of science. In 1962 alone the Nobel Prizes in both biology and chemistry were given to five individuals for their contributions to this subject.

Conceptual Background

The difference between living and non-living things seems simple when we compare a horse and a car. The obvious differences shrink to the level of

NRC 8001

definitions when considering certain viruses that often behave as molecules and even crystallize *in vitro* and yet are highly infective *in vivo*. All species of living organisms must be capable of growth and reproduction in order to survive. Certain inanimate systems, such as crystals, are also capable of growth and reproduction, but there is a fundamental difference between the two systems: crystal growth releases energy, while living systems absorb energy during growth. This is not a contradiction of thermodynamics. Living organisms accumulate energy by burning fuel and not by creating a mysterious order out of their environment or by gulping particles of negative entropy. However, living things do require a complex organization to maintain such a physically less-probable state.

A single-celled organism obviously has all the structures and organization required to maintain growth and reproduction. The simplest cell, however, is still highly complex, since it contains and controls a whole hierarchy of sub-cellular particles and a multiplicity of soluble components. Molecular biology is attempting to discover the physical structure, chemical composition, and biological function of these sub-cellular components, with the object of explaining the physical bases of life. This involves experiments on the simplest living organisms and the development of physical systems that perform some of the basic functions of living things. Physicists can define a "critical mass." Is there a "critical organization" that represents life?

Science starts with observations but observations alone are not enough. It also requires a belief in a universal order in nature that man can discover, and the application of an inductive process to explain the pattern. As each branch of knowledge grows, its concepts gradually acquire a coherent relation until they are fused in a quantitative theory or exact law. This may have been appreciated by Alfred, Lord Tennyson, when he wrote,

> Eye, to which all order festers, all things
> here are out of joint,
> Science moves, but slowly slowly, creeping
> on from point to point

The concept that the highest court of appeal in natural phenomena was not human authority, but valid observation and experiment, was accepted in physical science by Newton's time in the seventeenth century, but did not receive similar acceptance in life science until Darwin's time two centuries later. This lack of a conceptual scheme probably retarded progress toward a fundamental understanding of living processes as much as the complexity of the problems or the lack of adequate tools. During the intervening period most biologists limited themselves to observations, systematic classification, and descriptions of purpose. Legs were for walking, eyes for seeing, and living organisms contained "animal spirits" or "vital forces." Any thought that certain physical structures, processes and reactions must be common to all life was vulgar and an affront to human dignity.

While new concepts introduced in the mid-nineteenth century stimulated

progress in biological science, some of the older views appeared as philosophies that lay outside the realm of science. Henri Bergson still proclaimed the existence of an *élan vital* at the turn of this century. His concept of *élan vital* may be accepted, respected, or rejected but, since its existence cannot be proved or disproved by observation and experiment, it lies beyond scientific studies of biology. At the experimental level biological science has not yet revealed definite categorical discontinuities in the controversial region, and the difficulty of defining a living organism is evidence of this fact.

Science proceeds by describing the complex phenomena of nature in terms of simpler events. All explanations are subject to the dictum, known as Occam's razor, that theoretical entities must not be multiplied beyond necessity. Some molecular biologists who are immune to "vital forces" have shown some reluctance to shave their interpretations with Occam's razor. Whitehead's statement, "seek simplicity and distrust it," is, however, particularly applicable to present-day knowledge of biology at the molecular level. Relative to the complexity of any system that absorbs energy, grows, and reproduces itself, our facts are relatively few and our interpretations are over-simplified. The scientific method is capable of exploring the physical basis of life much farther than we have at present, but ultimately it may be the sheer complexity of organization in the living cell that will hinder further progress. Let us now make a historical descent toward the molecular level.

The Cell

If we are to explain the nature of life, we must find the factors that are common to such diverse forms as bacteria, worms, trees, and elephants. One of the first unifying principles to be discovered was that all living things were composed of cells. Robert Hooke in the seventeenth century named the structures he observed in plant materials "cells." A few years later Anton van Leeuwenhoek was the first to see living single-celled organisms in pond water. Their contemporaries and others throughout the eighteenth century saw and made accurate drawings of many plant and some animal cells, but this was the observational period of biology, and these investigators either failed to realize themselves, or failed to impress others with, the general significance of their observations.

It was not until 1839 that Schwann made the simple generalization that all organisms contained cells, and he is commonly regarded as the originator of the cell theory. This simple statement appears to have been accepted but phrases such as "free cell formation" indicated the prevalent belief that they might be formed from non-cellular material. Spontaneous generation was not disproved conclusively by Pasteur until twenty years later.

It was Virchow, pathologist and politician, who formulated the cell theory in precise form, *Omnis cellula e cellula*—all cells from cells—and stated "that throughout the whole range of living organisms there rules an eternal law of continuous reproduction." Virchow published this concept

at least a year before Darwin's *Origin of Species* and before Pasteur started his final studies that proved all life came from life.

True, Harvey, of blood circulation fame, had foreseen in 1651 in his dictum *Omne vivum ex ovo* that "almost all animals—including man, are produced from eggs," and he also declared against spontaneous generation. These beliefs, while correct, were sheer insight, as the subject was inaccessible to him with the available microscopes, and his proofs were therefore inadequate for such generalizations. In fact, the mammalian ovum was not discovered until 1827.

Sub-Cellular Structures

As optical equipment improved, an increasing number of sub-cellular structures was discovered and several of these were common to all except the most specialized cells. We shall consider only two of these, the nucleus that carries the hereditary factors and the ribosomes that synthesize proteins.

Robert Brown, who gave his name to Brownian movement, appears to have been the first to call the largest sub-cellular body the nucleus. Later in the nineteenth century it was observed that just before cell division the substance of the nucleus broke up into a number of thread-like bodies, called chromosomes. During cell division these chromosomes were observed to pass through a complicated series of geometrical patterns, representing the phases so dear to the cytogeneticist, before they reunited to form two nuclei, one for each of the two new daughter cells.

Until the turn of the present century, these remained as observations with nothing to link them to the hereditary mechanism. Mendel's monumental experiments of the 1860's were neglected for thirty-five years, perhaps because the scientific atmosphere was not receptive to his findings. His experiments with the garden pea showed that a discontinuous variation occurred between generations and that a number of these could be described by simple mathematical rules. Mendel explained this by assuming that each character in the body or somatic cells of the plant was determined by two hereditary factors that were present in equal numbers, and segregated in each succeeding generation in accordance with the laws of probability.

The Chromosome Theory

In the early years of this century Sutton, De Vries, and Boveri, working independently, evolved the chromosome theory that was to link Mendel's findings to the observed behaviour of chromosomes. Germ cells contained only one of each kind of chromosome but, as one set is derived from each parent, they occur in pairs in the somatic cells. Now, if we assume that Mendel's hereditary factors—or genes—are carried on the chromosomes, it is clear that paired genes will occur in *somatic* cells, and that segregation can occur when reduction division occurs in the *germ* cells. This theory was

thoroughly tested and developed by Morgan and proved to be another unifying principle common to all living things.

The chromosome was then visualized as a linear string of genes and this was confirmed by recombination or cross-over experiments. However, this experimental gene of the geneticist was really a small piece of chromosome that acted as a unit, rather than a minimum fundamental unit of heredity. Modern evidence indicates that the early geneticist's experimental gene, as described above, does contain a number of sub-units or sub-genes. Considering the small amount of nuclear material in a fertilized egg that is divided into chromosomes, and then into genes, and then into sub-genes, it is clear that the ultimate units must approach molecular dimensions.

Mutations occur in nature or they may be produced experimentally by various means including irradiation and chemical mutagenic agents. Those that are not fatal cause some change in the organism that is reproduced in subsequent generations. Major changes are usually associated with alterations in the number or character of the chromosomes. Less conspicuous changes that are usually detected biochemically are called "gene mutations." The important point here is that all mutations, whether natural or induced, are unpredictable changes in the genetic material that occur *in vivo*.

The first evidence that the genetic code could be altered in a more predictable manner by actual incorporation of *in vitro* material came when Avery and his colleagues (1) discovered a substance that transformed bacteria. An extract from the encapsulated (smooth) strain of *Pneumococcus*, when added to the growth media of the unencapsulated (rough) strain, transformed some of the latter to the smooth type. This induced ability to synthesize a capsule was transmitted to the descendants of the transformed strain. Purification of these extracts showed that the active transforming substance was the chromosomal material. Some thirty distinct characters have now been introduced in bacteria by this method. These extracts evidently perform two functions normally attributed to genes, namely, a change in the genetic character of the organism, and the initiation of its own duplication.

This suggests that the gene may be an inanimate material, representing a set of specifications that the organism can implement in terms of cell metabolism, and reproduce for the next generation. Evidently a few sentences of the specification from a closely related species can be substituted, implemented, and reproduced. This apparently inanimate nature of the gene or sub-gene suggests investigation at the molecular level.

The Molecular Level

This is not the place to deal with the detailed chemistry of the cellular materials but a few facts about the three components responsible for reproduction and protein synthesis must be mentioned. The nuclear material or chromosomes that carry the hereditary factors are made up largely of

desoxyribonucleic acid, called DNA for simplicity. The smallest cellular particles, or ribosomes, lying outside the nucleus, perform protein synthesis. They contain ribonucleic acid, or RNA, and this form of nucleic acid can be distinguished both biologically and chemically from DNA. Finally, there are the catalytic proteins or enzymes that control cell metabolism.

The nucleic acids and proteins have molecular weights ranging from several thousand to several millions; and are therefore giant molecules by usual chemical standards. These macromolecules are made up of smaller units, somewhat like a chain made up of a great many links, but there are relatively few different types of links. Thus, DNA is made up of only four kinds called nucleotides (thymine, adenine, guanine, cytosine, combined to a pentose sugar and phosphate group) which we shall call T, A, G, and C, but there are several thousand of each of these in one DNA molecule. The second type of nucleic acid, RNA, has a similar but not identical composition.

Under the temperature and other conditions prevailing in cells the essential reactions cannot take place without enzymes. These enzymes are proteins that break down the nutrients, release energy, and synthesize such essential components as the nucleotides themselves. Each enzyme can promote only one reaction and is therefore highly specific.

Proteins consist of a chain of several hundred amino acids of which there are twenty different types, but this is not the place to learn their twenty-letter alphabet. In fact, protein structure is much more complicated than a sequence of letters since the primary chain of amino acids is coiled into a secondary structure and then folded into a tertiary structure. These complicated structures have been described in some recent work (2, 3).

Returning to the biological aspects, Beadle and Tatum (4) produced mutants of *Neurospora* that had lost the power to produce certain vitamins and amino acids. More refined experiments showed that this was caused by the inability of the mutant organism to produce an essential enzyme. In fact, several enzymes were often required to produce a single amino acid. Evidently the gene controls enzyme production, and these enzymes in turn determine hereditary characters. Since the loss of a single enzyme is one of the smallest hereditary changes that can be demonstrated, this led to the hypothesis that one gene produces one enzyme.

Clearly, an initial understanding of living processes requires an answer to two basic questions. First, how does DNA reproduce an exact copy of itself to satisfy the law of reproduction? Second, how does the DNA assemble amino acids to form the enzymes that control cell metabolism and growth?

Reproduction of genetic code

The fidelity with which DNA reproduces the genetic code suggests a template mechanism. This means that the DNA molecule must behave as a mold or pattern capable of casting a replica of itself. If this is correct, its physical structure must be of paramount importance, but the chemical

composition alone is of little value for indicating the precise structure of these giant molecules in which the sequence of the letters TAGC is not known. A major advance was made in 1953 when Watson and Crick (5) proposed a structure for DNA that was consistent with the available chemical and recently acquired X-ray data.

This structure is a double-stranded helix with cross links. For simplification, we shall mentally unwind this two-stranded rope and view it as a ladder. The two uprights are formed by the pentose sugar and phosphate groups of the nucleotides. The rungs are formed by the nitrogen bases of two nucleotides strongly attached to opposite uprights, but attached to each other by comparatively weak chemical bonds. An important point is that the nucleotides will fit only in unique pairs, T is always opposite A and G always opposite C. This complementary structure suggests a self-duplicating mechanism since, if the two uprights separate at the weak bonds in the middle of the rungs, each upright could serve as a template for the formation of its complement. While it has been established that the nucleotides only combine in TA and GC pairs, it has not been established that the template is formed by unwinding the strands.

Proof that DNA acts as a template was provided by Kornberg and his associates (6, 7). They isolated an enzyme called polymerase which was capable of joining the nucleotides to form DNA *in vitro* but only in the presence of a small amount of extracted DNA. Not only was this primer DNA necessary to promote synthesis, but it always duplicated its own structure. This reaction would proceed essentially irreversibly until it had produced as much as twenty times the amount added. Clearly, Kornberg's *in vitro* polymerase system has much in common with a virus which is largely DNA and requires suitable conditions in the host cell to reproduce itself.

Now one might ask if a sequence of four symbols, represented by TAGC, is likely to carry all the information required. The Morse code can carry any message with three symbols—dots, dashes, and spaces. At present it is believed that three nucleotides constitute a coding unit, and with four different nucleotides, this provides 64 different sets of triplet sequences or coding units. At molecular dimensions the length of the DNA chain required to carry this coding unit is exceedingly small. The amount of DNA in a bacterium would provide a double helix long enough to carry the information in about 5,000 printed pages, and each cell of a higher animal could carry about ten times as much.

In summary, DNA evidently reproduces itself by a template mechanism, and at the molecular level this template appears to have adequate capacity to carry the required information.

Enzyme synthesis

Now we must consider our second question. How does the DNA control the assembly of amino acids to form the enzymes that control cell meta-

bolism? Any mechanism proposed must recognize that this involves the translation of the four-symbol code of DNA into the 20-symbol code representing the number of different amino acids in proteins. A 20-symbol code is comparable with the number of letters in our alphabet, and the total number of amino acids in a protein is comparable with the number of letters on a printed page. It is therefore evident that there could be about as many proteins as there are pages of literature from Homer to eternity. Clearly, the DNA code must be able to specify and control the specific sequences to provide a meaningful specification.

The second type of nucleic acid found in cells, the RNA, is evidently responsible for translating the DNA code in terms of protein synthesis. RNA differs from DNA chemically in having an extra hydroxyl group attached to the sugar, and the nucleotide T is replaced by another nucleotide U. Like DNA it has four symbols, and it is believed that a triplet code is required for each amino acid, although more than one triplet sequence may act as the code for one amino acid. RNA occurs in the cell in at least three forms: as messenger RNA; as RNA in the ribosomes, the small particles that synthesize proteins; and finally, as soluble RNA.

Available evidence indicates that DNA directs the synthesis of messenger RNA by a template mechanism and, when this message is attached to the ribosome, it becomes the seat of protein synthesis. In some bacteria this messenger RNA is unstable, having a half life of only a few minutes, and although it may be more stable in other species, this destruction of the message after use may be part of the control mechanism.

Now we come to the translation step. It is believed that a different form of soluble RNA interacts specifically with each amino acid, and that each form has a nucleotide triplet that is complementary to a triplet on the messenger RNA on the ribosome. In this way the activated amino acids are brought into the required sequence. The presence of messenger RNA, soluble RNA, and the synthesis of proteins on the ribosomes are established facts, but the mechanism is still at the level of a working hypothesis and subject to modification.

Our knowledge of this phase is due largely to the work of Ochoa and his colleagues (8, 9). They were the first to isolate an enzyme that would polymerize reversibly one or more nucleotides into a synthetic RNA. By using these synthetic polynucleotides as messenger RNA in cell-free systems containing ribosomes and the other components, it has been possible to direct the synthesis of proteins and make a start on deciphering the nucleotide code. In a recent experiment (10) a synthetic nucleic acid containing only one nucleotide was used as messenger RNA. The product obtained was polyphenolalanine, a kind of protein containing only one amino acid. Neither this monotonic message nor its monotonic echo would be of much use to a cell but scientifically this type of experiment provides a key for decoding messenger RNA.

An extension of these experiments has suggested the code for assembling

most of the amino acids but final proof has not been obtained. If these speculations are correct, they suggest that the code letters may be universal. If so, we have found another unifying principle, namely, that all living things use the same alphabet to write different specifications.

The linear chain of amino acids produced on the ribosome must be coiled and folded into a specific secondary and tertiary structure before it becomes active as an enzyme. Given the proper sequence of amino acids, it is believed that the properties of this unit may be as important as a template mechanism in generating the folding required to produce an active enzyme. This is comparable with the properties of salts that crystallize in a unique form. White and Anfinsen (11) have obtained evidence that supports this hypothesis with the enzyme ribonuclease.

Conclusion

There is more of this story, more about template generation, property generation, and condition generation, of the fundamental units that control reproduction, metabolism, and differentiation of living cells; and much more that we do not yet understand. Like the atomic nucleus, the gene is dissolving into cistrons, mutons, and other particles, as our knowledge advances. Molecular biology is a young science but it has already established a firm foundation of fact plus the inevitable scaffolding that supports its working hypotheses. This construction engages physicists, chemists, and biologists, a fact that reflects a growing interest in the nature of life processes and demonstrates the fundamental unity of all science.

With the information currently available, can we define the minimum unit that will meet Virchow's "law of reproduction"? It has been shown that the infective part of several viruses is DNA—a macromolecule. Clearly, a virus can reproduce its genetic code and control the metabolism of the host cell, but its dependence on the host indicates that its own metabolism is inadequate to provide the building bricks necessary for its independent existence. Since a virus cannot maintain its reproduction without a host cell, is it a living organism or just an abnormal part of the host cell produced by the introduction of the wrong set of specifications? If a virus is a living organism, what about the polymerase system? They both reproduce DNA, one using materials provided by the host cell, the other using extracts provided by man in a test tube. It is a matter of definition rather than discontinuity. The single-celled organism appears to be the smallest unit that can maintain growth and reproduction without assistance from another living system. Proteins and DNA are just macromolecules when removed from the organizational context we call life.

This rough chart of the mechanism of physical inheritance shows that these macromolecules carry the basic specification for each species, which is reproduced for each succeeding generation. Changes in certain words, sentences, or paragraphs of that specification can be tolerated if they are

meaningful, and these alterations are reflected in the differences among individuals within species. Copies with serious faults are usually fatal and are therefore eliminated. This most important code of nature evidently has its symbols and these are reproduced at least in part by template at the molecular level.

This physical inheritance has its parallel in the evolution of human culture. Would Beethoven's symphonies, Shakespeare's plays, or Renoir's paintings have come into being without a cultural inheritance and a contemporary society? Certainly current scientific advances rest on the foundations of earlier discoveries and the intellectual interaction of an increasing number of participants. This cultural heritage was made possible by man-made symbols, codes and communications that are often produced by templates. Our cultural records may not require as frequent reproduction but the message still has to be transmitted to each successive generation. The transmission of this message is our responsibility. May it carry the code that promotes understanding and the growth of wisdom.

References

1. O. T. Avery, C. M. McLeod, and M. McCarty. J. Exp. Med. *79*, 137 (1944).
2. J. C. Kendrew, H. C. Watson, B. E. Strandberg, R. E. Dickerson, D. C. Phillips, and V. C. Shore. Nature, *190*, 663 (1961).
3. M. F. Perutz. Proteins and Nucleic Acids. Elsevier Pub. Co., Amsterdam, 1962.
4. G. W. Beadle and E. L. Tatum. Am. J. Bot. *32*, 678 (1945).
5. J. D. Watson and F. H. C. Crick. Nature, *171*, 737 and 964 (1953).
6. A. Kornberg, I. R. Lehman, M. J. Bessman, and E. S. Simms. Biochim. Biophys. Acta, *21*, 197 (1956).
7. I. R. Lehman, M. J. Bessman, E. S. Simms, and A. Kornberg. J. Biol. Chem. *233*, 163 (1958).
8. M. Gronberg-Manago and S. Ochoa. J. Am. Chem. Soc. *77*, 3165 (1955).
9. S. Ochoa. Fed. Proc. *22*, 62 (1963).
10. M. W. Nirenberg and J. H. Matthaei. Proc. Nat. Acad. Sci. *47*, 1558 (1961).
11. F. H. White, Jr., and C. B. Anfinsen. Ann. N.Y. Acad. Sci. *81*, 515 (1959).