SOVIET PHYSICS USPEKHI VOLUME 12, NUMBER 4 JANUARY-FEBRUARY 1970

577.3

PROBLEMS OF BIOPHYSICS

S. E. BRESLER

Institute of High-Molecular-Weight Compounds, USSR Academy of Sciences, Leningrad

Usp. Fiz. Nauk 98, 653-708 (August, 1969)

1. INTRODUCTION

In living organisms we find unusually complex systems that perform extremely refined and subtle functions. We can classify the phenomena of the living world as being chemical or physical, just as in non-living nature. The chemical processes of living nature are studied by biochemistry. This science became distinct and developed much earlier than biophysics. The reason for this is that molecular-atomistic concepts became the basis of biochemistry long before they did in biophysics. Hence biochemistry is a more organized, systematic science, and shows more completeness.

Biophysics is supposed to study the physical phenomena of living nature. However, it has done this phenomenologically for a long time, without investigating the atomic-molecular nature of living matter. Hence, at present it is still more a collection of unsolved problems than a perfected science. Only in the last 15 years has a breakthrough taken place and a powerful modern weapon been forged: a system of experimental methods and theoretical outlooks. In essence, this is a new scientific method that starts with studying the structures and interactions of the molecules that living beings consist of.

It is precisely the atomic-molecular approach that has made it possible to solve in principle one of the most fundamental problems, that of understanding the mechanism of heredity and variation of organisms. This branch of biophysics comprises the content of what is now commonly called molecular biology. However, besides the solution of concrete problems, it was even more important that a new scientific method was created, which can and must be applied to other problems that are even more complex and difficult. Thus biophysics is going through a very interesting period now. There is a multitude of unsolved problems, and there are powerful methods based on current advances in physics and chemistry. Finally, there is a great number of very capable enthusiasts solving these problems. The advances of this science are multiplying not daily, but hourly.

Turning to the concrete material of biophysics, we can classify it in various ways. Undoubtedly, the number one problem is the molecular mechanism of genetic coding, which was mentioned above, or the mechanism of self-reproduction in living nature. In one of its essential aspects involving single-celled organisms, this problem has already been solved, As for the self-reproduction of a complex organism and its development from a single sexual cell by differentiation, there is still almost nothing known about it. Finally, when we say "self-reproduction of organisms," we are speaking of the synthesis of the materials that the cell consists of, and about which the nucleus contains information. However, biological molecules exist not in the free form, but as a rule, are incorporated into specialized, morphologically highly varied, insoluble structures, or membranes.

All the organelles constituting the cell structure are made of membranes: the outer membrane, the nucleus, the mitochondria, ribosomes, etc. How this structure arises and how the information about it is carried on to posterity—all this, which is usually called the problem of morphogenesis, should be considered the second very important unsolved problem of biophysics.

Very important biological functions, e.g., that of motion, are assigned to the membranes. Spontaneous motion, the direct transformation of chemical energy into mechanical work, is seen in all organisms at all stages of evolution from viruses to man and in all types of cells of a complex organism, starting with specialized organs, i.e., muscles, and ending with the contraction of mitochondria and ribosomes that occurs in any cell. We already know a great deal on the nature of biological movement. However, even this vigorously growing branch of biophysics is far from perfection.

Another fundamentally important function of membranes is active transport of various substances against a concentration gradient, i.e., in the direction opposite to ordinary diffusion. It is precisely the transport of substances through membranes, both through the outer membrane of the cell and within the cytoplasm, that regulates the supply of chemical substances to the places where chemical reactions are occurring in the cell. Hence, active transport coordinates chemical processes in space and time, and is one of the most important regulators of the rates of metabolic reactions. Some very important and widespread processes, in particular, are the so-called sodium and potassium pumps, or the active transport of potassium ions into cells and of sodium into the outer medium. The active transport of ions gives rise to electromotive forces on the membranes, which resemble galvanic cells or charged storage batteries. Under certain circumstances, the cells of the nervous system (neutrons) undergo membrane breakdown, and are short-circuited. Then an electric current flows through the membrane. The relation of transmission of nervous excitation to electric current has been known since the time of Galvani and Volta.

The brief electric signals arising in the neurons are currently considered to be the basis of nervous activity. These signals are just how all nervous activity of man and animals is coded.

Naturally, the mechanism of action of membranes, their function as chemical or electrochemical (in the case of ions) pumps, and the mechanism of generation of electric currents in nerve cells—all these are fundamental, not yet really solved problems of modern

biophysics. I shall treat these problems briefly in this review. A great deal of experimental material has been collected on them. However, as I've said above, biophysics lacked for a long time the atomic-molecular approach, and this restricted its potentialities extremely. I should emphasize that the atomic-molecular methods, ideas, and way of thinking became established in physics in general only at the beginning of the 20th Century. This was a half-century after these ideas had triumphed in chemistry, where any other approach was no longer possible after Dalton. This is explained simply by the fact that it proved to be an extremely complex matter to invent experimental methods to study atoms and molecules as physical objects. Only in the 20th Century (with the introduction into practice of x-ray and electron diffraction, molecular spectroscopy, and radio spectroscopy) did it become possible to study the atomic-molecular world by direct methods. The situation was the same in biophysics as well.

Biological molecules (proteins, nucleic acids, lipids, and carbohydrates) are complicated, as compared with those on which modern molecular physics grew up. It has proved possible in principle to study the structure of these molecules by using the same methods. However, it was necessary to develop an extremely refined technique in order to do this. This has been carried out during the last 15 years, and it has undoubtedly been the fruit of the scientific-technical revolution that has marked the middle of the 20th Century. For example, structural analysis of proteins would be simply impossible, not only without the modern microfocus x-ray tubes, but also without fast computers. Hence, the past 15 years have added an unusual amount to our knowledge of the structure, thermal motion, and molecular interaction of biological molecules. And this is what has made possible the atomic-molecular treatment of biological phenomena. This method has shown its great power with the example of the genetic code.

The time has now come to attack other biological problems.

In beginning this review, I shall discuss the fundamental data on the structure of biological molecules and on the problem of the genetic code. Then I shall discuss some biophysical problems that have been worked out less well or have been hardly touched on at all: the mechanism of development and differentiation in the formation of a complex organism, the mechanism of morphogenesis, the structure and function of membranes, electric phenomena in membranes and their meaning in nervous activity.

I shall omit a number of problems of biophysics, such as the nature of biomechanical phenomena, the mechanism of sight and hearing, the action of ionizing radiation on living matter, etc.

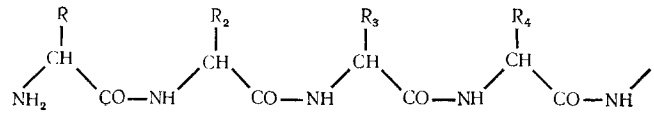
The wealth of material and brevity of presentation forces me to describe the results of a vast amount of experimental work without going into the methods of obtaining various data. It is clear to us how faulty such a method of presentation is. However, if I were to do otherwise, we would have not an article, but a book. I have tried to give a list of the literature in which interested readers can find more detailed information on the problems mentioned in the article.

II. BIOLOGICAL MOLECULES

1. Proteins[1]

The proteins occupy a central position in material exchange (metabolism). The proteins are enzymes, or catalysts for chemical reactions. They accelerate to a great extent the chemical reactions that metabolism consists of. Here each individual enzyme acclerates one strictly-defined reaction, and acts on strictly-defined substances, or substrates. Of course, the reactions accelerated by enzymes must be thermodynamically possible. However, in the absence of enzymes, these reactions do not proceed even when this condition is satisfied. The enzymes make the substrates react with one another in the range of moderate temperatures from 10° to 40°C in which life takes place. Enzymes are far more powerful and selective catalysts than those known in chemical technology. The basis of each enzyme is a protein molecule, which is often combined with a relatively small molecule of a so-called coenzyme. The vitamins or their derivatives often play the latter role. The structures and properties of the coenzymes have been well studied. However, as for the proteins, they are what give the enzymes their distinctiveness (since one particular coenzyme can be involved in tens of different enzymes), and their structures have become fully known in the last 10–15 years.

Proteins are polymers. They are long molecular chains. The smallest protein, insulin, consists of 51 links, and the longest protein chains apparently contain several thousand links. The links that protein chains consist of are amino acids. There are 20 of the so-called "magic" amino acids, which are universal, and which occur in almost all proteins. We shall not give their chemical structures, which can easily be found in the books. The general formula of a protein is



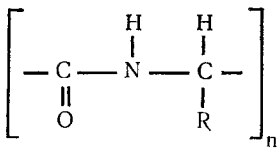
The position of each amino-acid link in a protein chain is fixed with mathematical exactitude. Replacement of one link with a new one is the result of a genetic mutation, and it can lead to a catastrophe in the structure of the protein and the function that it fulfills. Thus, Pauling and Ingram have shown that the replacement of a single link in the protein of the red blood cells, hemoglobin, gives rise to a very severe hereditary disease, sickel-cell anemia. One can compare a protein chain to a long line of text. Just as the letters in the text succeed one another in an absolutely fixed way, and not one can be dropped or replaced without distorting the text, a protein must be assembled from a fixed sequence of amino acids. Each substitution of a link is equivalent to a misprint. Of course, some amino acids are repeated often, like some letters in a text, and others more seldom. However, the essential point is that this is a fully ordered system that allows no fluctuations.

This type of protein structure, or the so-called primary structure, has been fully proved and cannot be disputed. There are excellent analytical methods

that permit one to determine the structural formula of a protein chain, i.e., the order of succession of the different amino acids along the chain. This has been carried out already for some tens of proteins.

The final proof that one knows the structure of any molecules in chemistry is considered to be its complete synthesis from simple substances. The total synthesis of a protein (insulin) by organic chemical methods has been performed within the last three years. Furthermore, the methods of synthetic chemistry have been developed so well that Merrifield, one of the outstanding chemists of our time, has invented an automatic machine that synthesizes any protein chain sequentially, step by step, according to an assigned program. Such a machine must be able to synthesize any protein, rather than any one particular object. The enzyme ribonuclease, a chain of 124 amino-acid residues, has already been synthesized on it. The synthetic enzyme proved to be active.

The proteins were the first example of polymers whose macromolecules have a secondary structure. What does this mean? It turns out that the links of a given chain interact strongly, owing to formation of hydrogen bonds. A protein chain is a sequence of links having the general formula



The peptide groups —CO—NH— form such strong hydrogen bonds that the latter do not break down, even in aqueous solution. Only such substances as trifluoroacetic or dichloroacetic acid or urea can completely disrupt the intramolecular hydrogen bonds in a protein macromolecule. Pauling predicted that the structures of minimum energy for a protein chain can be either the so-called α helix, in which each link is bound by two hydrogen bonds to its fourth nearest neighboring links along the chain, or the pleated β structure, in which the chain is folded in parallel pleats, and again each link forms two hydrogen bonds with the juxtaposed links in neighboring folds. And these are the two types of regular secondary structure in proteins. In essence, this is a way to pack a polymer chain in space, when it must minimize the energy of the attractive forces. We can call this phenomenon intramolecular crystallization.

The types of secondary structure predicted by Pauling have been confirmed experimentally in all details by x-ray structure analysis, both on model polymers (polypeptides consisting of identical links, i.e., any repeated amino acid) and on actual proteins. Furthermore, it has been shown that formation of a secondary structure, or respectively, its breakdown, is a cooperative process. They are actually analogous to crystallization and melting, although they occur within the confines of an individual macromolecule (proteins have molecular weights from 600 to hundreds of thousands, the average weight of a link being close to 110).

Another important detail becomes clear if we compare proteins with simple model polypeptides. The polypeptides have the following alternatives: either every macromolecule is entirely in the form of an α helix, and it then behaves like a rigid rod, or it is completely folded into a pleated β structure, forming a rather symmetric object like an ellipsoid, or, finally, it remains as an amorphous clump permeated with the solvent, i.e., water. Each of the secondary structures is selected by nature, and this depends on the relation between the energy of the attractive forces and the thermal motion of the links of the chain. In the final analysis, everything depends on the specific features of the side groups R that define the nature of the amino acids. The situation differs in proteins. Each macromolecule contains all 20 amino acids or most of them. Hence, all three types of secondary structure, α, β, and amorphous, can exist within a single protein macromolecule. A protein macromolecule is a system of small elements built according to the three possible types of secondary structure. Furthermore, all of these elements of varying secondary structure are packed in space in such a way as to produce a compact object, or protein globule. A protein macromolecule contains relatively little empty space filled with solvent molecules, usually about 20—30% of its volume. What determines such a topologically complex, so-called tertiary structure of a protein molecule? Naturally, it is determined by the balance of intramolecular forces and thermal motion.

Why shouldn't proteins form a single strictly regular structural element, as in polypeptides, e.g., a single α helical region? There are a number of reasons for this. First, as we see from Fig. 1, where the primary structure of a typical protein is shown in a diagram, the protein chain is cross-linked by chemical bridges at many places. If we use the terminology of technical chemistry, we can say that protein chains are vulcanized by sulfur bridges, as rubber is. Evidently, the α structure must be interrupted near the bridges. In addition, there are two types of links (proline and hydroxyproline), near which the α and β structures prove to be geometrically impossible.

Consequently, there are a number of varied factors that favor breakdown of long-range order, or amorphization of the protein chain. However, these include one most general factor. Undoubtedly, a protein globule is an equilibrium figure of all types of molecular forces acting among the links of the chain. The fact that the tertiary structure of proteins is the thermodynamic-

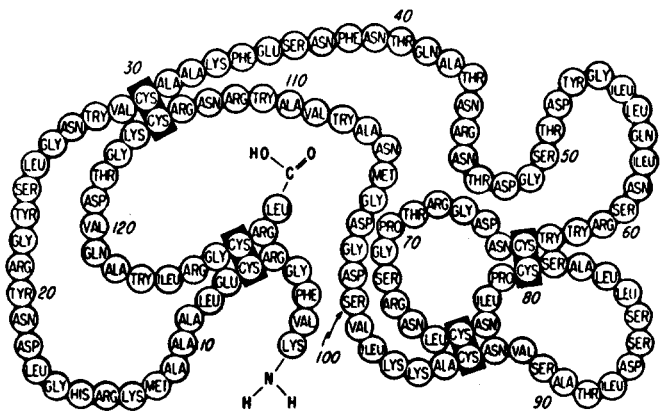
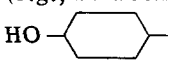


FIG. 1. Primary structure of lysozyme.

equilibrium structure has been recently proved by direct experiments. That is, they showed that one can disorganize this structure reversibly. The macromolecule is transformed into a random coil, and it then returns spontaneously to its initial state when the denaturing factors are removed.

Among all the types of molecular forces acting within a protein globule, there is one that favors its coiling into a compact object. This is the van der Waals, or dispersion forces between the hydrocarbon-type side groups. In a protein chain, up to 50% or more of the side-chain radicals are purely hydrocarbon in nature. Hence, the corresponding regions of the protein chain become hydrophobic. In an aqueous medium, these hydrocarbon radicals try to make contact among themselves, as if to merge into a hydrocarbon drop, for exactly the same reasons why kerosene forms in water round drops of minimum surface. The dispersion forces, or as they are unfortunately called, hydrophobic interactions, tend to gather the protein macromolecule into a compact ball. However, this is hindered by the secondary structure, which calls for rigid helices. Hence we get a compromise solution: the protein consists of ordered and disordered fragments mixed together, as if compressed into a compact, relatively symmetrical structure.

Of course, the "hydrophobic forces" are not the only ones that determine the tertiary structure of a protein. There are also highly characteristic hydrogen bonds between certain specifically adapted side groups (e.g., between the phenol group of tyrosine



and the carboxyl group of glutamic or aspartic acid). These hydrogen bonds are detected by their spectroscopic manifestations, since the phenol ring has a characteristic absorption band.

The Coulomb forces between the ionized side groups play a substantial role in proteins, and the number of these in an average protein molecule is as much as several tens. Finally, entire α-helical regions within the macromolecule possess substantial dipole moments of the order of tens and even hundreds of Debye units (depending on the dimensions of the α-helical regions). The dipole interactions of the helical regions also make an appreciable contribution to the energy of the attractive forces. Different helical fragments tend to arrange themselves antiparallel to one another. The distinctive feature of proteins consists in the fact that all the types of molecular forces make contributions of the same order to the energy of the protein globule. This is why proteins are so strikingly changeable. As the composition or the ionic forces of the medium vary, the protein macromolecules change: they swell or contract, elongate or become more symmetrical. For example, the blood protein hemoglobin seems to breathe as it combines with oxygen. Combination with each oxygen molecule is accompanied by less of one proton from a histidine side group in the protein. This gives rise to a variation of charge and change of dimensions of the macromolecule.

Enzymes change their dimensions when they combine with a substrate. People consider that this fact is important in effecting catalysis. In one way or another, all the internal contradictions in a protein between the secondary and tertiary structures, and between attractive forces (hydrophobic forces and hydrogen bonds) and repulsive forces (Coulomb forces between ionized groups) lead to some state of minimum energy. This is what characterizes the equilibrium structure of a protein macromolecule.

Some interesting attempts have been made recently to calculate the equilibrium secondary and tertiary structure of a protein, starting with the known primary structure, i.e., from the known sequence of amino-acid residues along the chain. In principle, the point is to minimize the total energy of all types of attractive and repulsive forces. The attempt to perform this calculation with modern computers has not yet succeeded. However, the very posing of this problem is instructive. We should assume that it can be solved, although this will require surmounting of considerable difficulties.

The study of the secondary and tertiary structures of a protein is currently an empirical matter. Here x-ray structure analysis gives the fullest and most all-encompassing information. The analysis of protein crystals is a striking success of the x-ray method.* These crystals contain several thousand atoms per unit cell. The analysis gives the coordinates of each atom. That is, it permits one completely to reproduce the primary, secondary, and tertiary structures without any additional data. The amount of x-ray structural information that one must get experimentally and work up is enormous. Ordinarily, each analysis requires exact measurement of the coordinates and intensities of many thousands of reflections. This is followed by mathematical treatment of this information by Fourier-series procedures. In practice, this can be done only with the fastest computers. In essence, currently all the work of measuring the x-ray patterns and entering the information into a computer is performed automatically, using a unified complex of instruments.

Having no opportunity to spend time on this technique, I shall give only an example of the space models of a pair of proteins that have been fully studied by the x-ray method: myoglobin (Fig. 2) and lysozyme (Fig. 3). These models are intentionally coarse, for they correspond to a resolution of 6 Å, at which the individual atoms are not visible, but only the overall topology of packing of the chain. Besides, the side groups that fill almost all the gaps are omitted in the models. All of this has been done for pictorial reasons. We see the amazingly complex tertiary structure of the proteins. As I have stated, it consists of ordered and amorphous regions. We can see quite well the segments of Pauling α helices as elements of this structure. In the enzyme lysozyme, one can identify the site in the structure of the protein where the substrate is adsorbed. This is the so-called active center. Naturally, it is small in comparison with the dimensions of the entire globule.

With all its automation, the x-ray structural method is still difficult, and demands great expenditures of effort and time. Hence, various physical methods are useful. Although they are rather qualitative, they nevertheless permit one to follow changes in the

*Cf. the article by W. L. Bragg, Scientific American 219, 518 (1969). (1969).

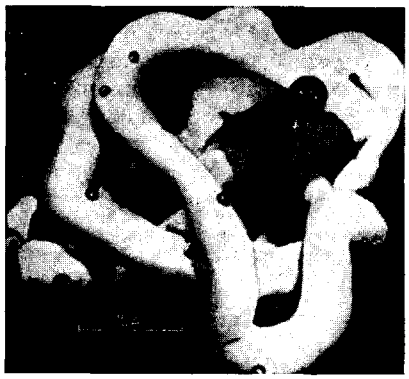


FIG. 2. A model of myoglobin at 6 Å resolution.

FIG. 3. A model of lysozyme. The arrow indicates the active center of the enzyme.

secondary and tertiary structures of a protein. These include:

1) the spectrophotometric methods (the peptide bond has an absorption band near 220 nm; the intensity of this band decreases when regular α and β structures are formed; this is the so-called hypochromic effect);

2) spectropolarimetry;

3) variation of the circular dichroism near an absorption band, or the so-called Cotton effect.

All these phenomena permit one to get a certain amount of information on the structure. However, the lack of an exact theory rules out using them to characterize quantitatively the degree of regularity of a protein or to distinguish elements of α from those of β structures. The optical methods are still highly qualitative, and not always unequivocal. However, they are useful as a palliative because of their speed and lack of difficulty.

Another group of experimental methods, the hydrodynamic methods, gives an overall picture of the dimensions of a protein globule, and thus of its tertiary structure. Those most used are to measure the viscosity of protein solutions and the sedimentation constant (the rate of sedimentation due to centrifugal acceleration) of a protein in the ultracentrifuge. All these methods, however limited, permit one to make advances in the field of protein physical chemistry. However, only x-ray structure analysis gives decisive answers permitting no alternatives.

In closing this section, I can state that the structure of proteins now holds no unsolved secrets. The primary structure is already known for tens of proteins, and the secondary and tertiary structure for nine. There are precedents for the total synthesis of proteins by organic chemical methods, and here the synthetic protein showed the same functional activity as the natural one.

The problem now is quantitative growth of information and increase in the number of objects from all classes of proteins. This is an important and useful task, but it no longer has the novelty and complete unexpectedness that distinguished this field even ten years ago. However, one fundamental question remains practically completely untouched. Why do proteins have the structures that we find them to have? On the one hand, this problem involves the function fulfilled by the proteins. In other words, we can ask: how does the structure of a given actual protein serve to fulfill this useful function? What features in the structure are obligatory and unchangeable? On the other hand, each protein is a product of evolution, being the result of accumulation of random, inheritable changes that have become established by natural selection. Therefore, not only the present function of a protein is written in the structure of a protein, but also its history, i.e., the functions that it fulfilled at some time in more primitive organisms. We must not look for total functionality in the structures of proteins. As we know, evolution is a statistical process, full of chance. An understanding of the relation of the structure of a protein to its different functions and the mutual dependence of structure and function are the almost universal problem linking most of the problems of biophysics mentioned in the introduction.

2. Nucleic acids

This is the second very important class of biological molecules playing a principal role in the process of self-duplication of a cell or an entire organism. The nucleic acids are cybernetic substances, or the material substrate of heredity. They contain fixed in them the structures of all the proteins of the cell and organism, and all its constituent substances. The nucleic acids are polymer chains in which four types of links, the so-called nucleotides, take part in various combinations. The primary-valence chain of the nucleic acids contains alternately molecules of sugar and of phosphoric acid linked to the former by two ester bonds. Purine and pyrimidine bases are attached to the side of the primary-valence chain: adenine, guanine, cytosine, and thymine (or uracil). Two types of sugar are found: ribose in ribonucleic acid (RNA) and deoxyribose in deoxyribonucleic acid (DNA). It has been established that DNA is a fundamental constituent of cell nuclei, and that the amount of DNA in all the nuclei of the cells of a given organism is absolutely standard. We should also distinguish primary, secondary, and tertiary structures in the nucleic acids. The primary structure is the structural formula of the substance. For the nucleic acids, this involves knowing the sequence of the four types of nucleotides along the chain. Various ingenious methods have been devel-

oped for studying the primary structure, but the problem here has proved to be considerably more difficult than in studying proteins, and the results are still meager. At present one can study exhaustively only sequences of nucleotides that are not too long, not exceeding 100 units. The greatest advances are the elucidation of the structural formulas of the transfer RNA chains, which consist of 75—80 nucleotide links.

As for DNA, we find here unusually long chains of as many as millions of links. The sequence of the four types of nucleotides along such a long chain cannot be studied by the existing methods. Hence, we must still be content with little. One can find out what the 2—3—4 terminal links of the chain are, or one can split the chain into short oligonucleotides (dimers and trimers), and then find the statistical distribution of neighbors in these short fragments.

As a whole the problem of determining the primary structure of nucleic acids has proved to be more difficult than for proteins. Here the fact that the nucleic acids are made of four types of monomers rather than twenty, like the proteins, has not turned out in our favor. This is because the fragments after partial hydrolysis prove to be too highly standardized. The apparent simplicity is a hindrance, rather than an advantage.

As for the secondary structure of the nuclei acids, I should take this up in more detail, since its discovery (Watson and Crick, 1953) was one of the major stages in the founding of molecular biology. They analyzed the x-ray diffraction patterns of DNA, in conjunction with much other data (here an x-ray structural analysis is by itself inadequate, since the nucleic acids give imperfect diffraction patterns having small numbers of reflections, owing to a low degree of order). Thus they showed that DNA always forms double helices (with very rare exceptions). Here the two strands of the double helix are strictly complentary to one another (Fig. 4). Namely, opposite adenine in the first strand we always find thymine, and always cytosine opposite guanine. This mutual complementarity of the nucleotides stems from the fact that adenine forms two hydrogen bonds with thymine (or uracil), while guanine forms three hydrogen bonds with cytosine. There are no other possibilities for compact packing in the long DNA chains. If we take simply the side groups, i.e., free purine and pyrimidine bases, rather than polymeric DNA chains, they can form complexes with one another. This is manifested it their ability to cocrystallize in stoichiometric ratios. However, here the variety of possible complexes is much greater, as has been shown by x-ray structural analysis. This is not remarkable, since the small molecules can be rotated to various angles with respect to one another in the unit cell. Conversely, no variety is observed in the polymers, and the complementarity rules formulated by Watson and Crick are strictly obeyed.

An attempt has been made in recent years to estimate the contribution of various molecular forces to the interaction energy of the two chains of DNA in the double helix.[2] It gave the impression that the van der Waals dispersion forces between the purine and pyrimidine groups, which are packed like a pile of coins inside the double helix, is the dominating factor, and contributes most. However, the hydrogen bonds create the specificity of bonding, namely, A—T and G—C. That is, they are the basis of the complementarity principle. Unfortunately, the quantum-chemical methods are crude, and give varied results in the hands of different authors. Hence, we don't yet have good calculations physically justifying the quantity 14—15 kcal/mole, which is the energy barrier protecting the complementary pairing from "errors."

However, a fact is still a fact. The secondary structure of DNA resembles an intramolecular crystallization even more than with the proteins, and its breakdown by thermal motion or at extreme pH values resembles the melting of a crystal. It occurs under critical conditions, as is characteristic of cooperative phenomena. However, the most important feature in the structure of DNA and in the complementarity principle found in it is that the two chains mutually determine each other. Each of the two chains of DNA completely determines the structure of the complementary chain. This fact is the basis of a new fundamental principle of nature. The point is that the function of DNA is to preserve and transfer information on the structures of all the proteins and nucleic acids and on their supermolecular organization in the cell. Evidently, the first requirement imposed on DNA is its capacity for literal self-reproduction (autoreplication). As we know, when a cell divides, the two nuclei of the daughter cells acquire the full information contained in the nucleus of the mother cell, and this means a complete copy of the maternal DNA. However, DNA is a polymer in which the four types of nucleotides alternate in the most arbitrary way. How can a polymer chain of the four types of monomer links be synthesized in such a way as to reproduce a given sequence of monomers with absolute accuracy?

The probability of error, i.e., incorporating a wrong link, is a stupendously small quantity. We can estimate it to be 10^{-10}—10^{-11}. The methods of synthetic chemistry known to us are unsuitable for attaining such accuracy. Hence, the special concept of template synthesis has been introduced, i.e., a process in which an existing polymer chain serves as a template or stencil for a second newly-formed chain, and imposes its own structure on it, as though transferring the information stored in it. The concept of a template was only a logical abstraction until the physical structure of DNA had been discovered. Watson and Crick transformed it into a real physical object, and showed what molecular forces it was based on. Let us imagine that the double helix of DNA begins to unwind at one of its ends. Then monomers can be adsorbed from the solution on the liberated DNA strand. They will be adsorbed with required obedience to the complementarity principle. Then, when they have been linked into a chain, they will consequently form two Watson-Crick double helices. In this synthesis, each of the two chains has served as a template for assembling the complementary chain. This mechanism completely explains the process of autoreplication of DNA.

As one usually says, this process must be semiconservative, since each new daughter molecule of DNA contains one chain from the maternal DNA and one newly synthesized chain. This concrete prediction can

be tested by simple experiments using an isotopic label to permit distinguishing the old DNA chains from the new. The experiments of Meselson, Stahl, and Vinograd on bacteria and those of Taylor on higher plants have completely demonstrated that the mechanism of DNA synthesis in living nature is a mechanism of semiconservative copying. In other words, the DNA molecule is a template for synthesizing similar DNA molecules. One may ask what molecular forces ensure such a selectivity in choosing the nucleotide links, such that the probability of error is no greater than 10^{-10}. We can estimate from this number the energy difference when a right or wrong (non-complementary) link is incorporated into the chain being synthesized. Since according to Boltzmann the probability of fluctuation $W = e^{-U/kT}$, where U is the sought energy difference, then if we set $W = 10^{-10}$, we find that $U = 15$ kcal/mole. This is a large quantity, and according to all the data, it is ultimately determined by the adenine-thymine and guanine-cytosine hydrogen bonds.

All that I've said makes it clear what an enormous role was played by the discovery of the secondary structure of DNA. The situation differs somewhat with the secondary structure of RNA, since RNA is synthesized in the form of a single chain. RNA is not found in cells in the free state, but is incorporated into special protein-polynucleotide particles, the ribosomes. They contain three types of RNA molecules, having molecular weights of 30,000, 600,000, and 1,000,000. Among viruses containing RNA, the latter is also incorporated into the structure of the virus corpuscle. Only the small transfer RNA (tRNA) can be found in the free dissolved state.

The tRNA chain is folded back on itself by formation of adenine-uracil and guanine-cytosine hydrogen bonds, and apparently forms a clover-leaf shape. It contains several (apparently four) ordered fragments (lengths of double helix), with amorphous connections between them. All the data indicate that such a tertiary structure of tRNA is important in its function. Destroying it in various ways (e.g., by heat, with subsequent fixation of the denatured states with multivalent ions) leads to loss of function.

Why is the shape of the tRNA molecule of such substantial importance? In all probability, the enzymes acting on tRNA as a substrate are sensitive to changes in its tertiary structure. A phenomenon takes place between the enzyme and the substrate that we can call "molecular recognition." Enzymes are extremely specific. They adsorb only their own substrates. Undoubtedly, the specificity is based on ordinary molecular forces, but this phenomenon is cooperative. The different parts of the substrate molecule must simultaneously match the complementary groups of the enzyme molecule. Evidently, the tertiary structures of both the enzyme and the substrate are pertinent in such a recognition. The same optical methods are widely applied to study changes in the tertiary and secondary structures of the nucleic acids as for the proteins. Here one can measure either the hypochromic effect in the absorption band of the nucleic acids (i.e., near 260 nm), or the optical activity (which in the nucleic acids is almost entirely due to the helical secondary structure), or the optical rotatory dispersion and circular dichroism near the absorption band (the Cotton effect), or the flow birefringence (the Maxwell effect). All these methods are in general use. In addition to them, the method of x-ray diffraction of solutions of nucleic acids has been rather successfully applied.

Electron microscopy gives valuable information for DNA. It permits one to measure the dimensions of the DNA double helix, and gives rather precise values for the molecular weight of DNA.

Incidentally, the problem of the molecular weight of DNA hasn't been cleared up yet. For viruses containing DNA (e.g., bacteriophages), one can sometimes isolate entire DNA molecules from the virus particle (Fig. 5). One gets a molecularly-homogeneous DNA preparation with a molecular weight of the order of $(100-300) \times 10^6$. However, when one prepares DNA from bacteria and from cells of higher organisms, one cannot avoid mechanical breakage of the long macromolecules. Consequently one gets DNA particles of random molecular weights. The suggestion has been repeatedly made that the entire chromosome of a bacterium contains one molecular chain of DNA. Then its molecular weight would have to be $(2-3) \times 10^9$. However, no one has been able to prove this yet. In one way or another, when one makes DNA preparations from various organisms, one gets usually a mixture of fragments of mean molecular weight $(10-20) \times 10^6$.

3. Lipids[3]

The third group of very important biological molecules is the lipids. They are the simplest substances in structure. They have low molecular weights and their structures are completely known. However, they

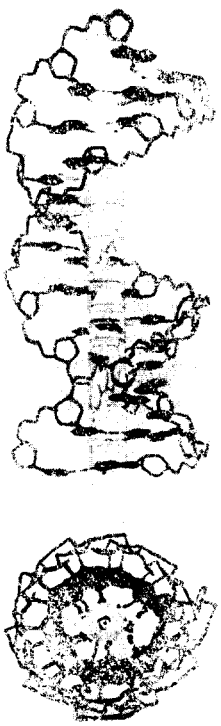


FIG. 4. The Watson-Crick helix (wire model).

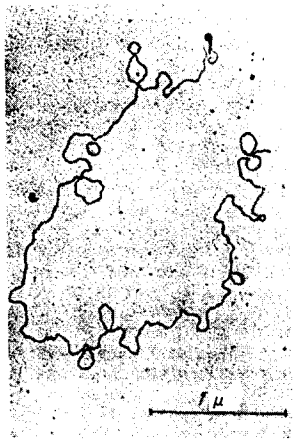


FIG. 5. Electron micrograph of DNA. The instant when the phage T1 ejects its DNA is shown.

are very complex in their functions, and it is not yet clear exactly what they are supposed to do nor why nature has given them such distinctive structures. The lipids are surface-active substances that structurally resemble the detergents that are so widely used in technology and everyday life. The lipids are a necessary component of biological membranes. The lipids are precisely what makes the membranes difficultly permeable for ions, sugars, amino acids, and other substances readily soluble in water. Their permeability to these substances ceases to be a question of solubility and passive diffusion, but becomes the chemical process of the so-called active transport. If we remove the lipids from a membrane, it becomes "porous," and water-soluble substances leak through it without hindrance. How the lipids are incorporated into the membranes is not exactly known at present. I shall return later to these problems.

The functions of lipids are not limited to membranes. They form complexes with many proteins. Apparently some proteins are carriers for lipids (the latter are in themselves insoluble in water, but the appropriate proteins solubilize them). An example is the serum albumin of blood. It can even absorb benzene. In other proteins, lipids serve as regulators of enzymatic activity. (Cytochrome oxidase completely loses its activity if one removes the lipids from it, and regains activity if the lipids are replaced.) We should recall that certain hormones, e.g., those of the adrenal cortex, also belong to the lipids. Apparently they also form complexes with certain enzymes or other active proteins, and thus affect their function.

What is the mechanism of action of lipids on proteins based on? Why can a particular lipid affect the activity of various proteins? According to all the data, lipids change the conformation (i.e., the geometry) of the protein chain, and deform the protein globule. They are conformational regulators. I shall return below to this function that they have. Here I shall only emphasize that the structures of lipid molecules are adapted to adsorbing on proteins and affecting their conformations. The lipids can be classified into several groups. These are:

a) Simple lipids (fatty acids and neutral fats or triglycerides of the fatty acids). Quite probably, these substances play a subordinate role, being intermediate products for the synthesis of the more complex lipids or sources of energy for oxidative processes in the cell.

b) Steroids. This class of compounds includes cholesterol, the bile acids, the sex hormones, and vitamin D.

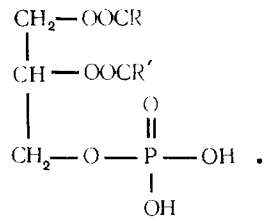
c) Hydrocarbons. This class of compounds includes vitamin A and other carotenoids, and also vitamins E and K.

d) Complex lipids (phospholipids and glycolipids; the latter contain carbohydrate in addition to fatty acids and amino alcohols).

e) Waxes: the esters of higher monofunctional alcohols and higher fatty acids.

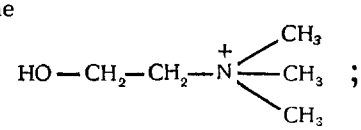
Apparently it is the complex lipids that fulfill the specific biological functions in membranes and in certain enzymes. The phospholipids are especially abundant in the brain cells and in the nervous system as a whole. This involves the self-sufficient importance of membrane processes in the nervous system.

I shall breifly take up the chemical structure of the phospholipids. A rather considerable number of phospholipids are known. Since we do not understand clearly what their functions are, and correspondingly we can't formulate any rational requirements on their structures, it would hardly be expedient to list all these substances. I shall refer those interested to the specialized literature.[3] I shall give only several typical examples. Many phospholipids consist of a so-called phosphatidic acid combined by an ester link with an amino alcohol. The formula of a phosphatidic acid is



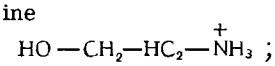
The radicals R and R′ are fatty-acid residues, R′ most often being an unsaturated fatty acid and R saturated. Hence, a phosphatidic acid amounts to glycerol in which two of the hydroxyls are esterified with fatty acids and the third hydroxyl with phosphoric acid. Evidently, a phosphatidic acid can form another ester bond through its phosphoric acid residue. Such bonds are formed in the lipid series with one of the following alcohols:

a) choline



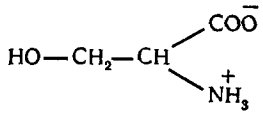
The corresponding class of lipids is called the lecithins. They are the most widespread phospholipids in nature.

b) ethanolamine



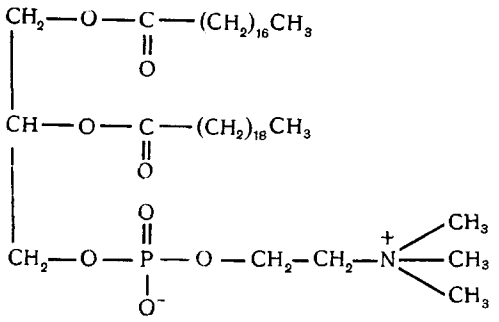
The corresponding class of lipids is called the cephalins.

c) serine



The phosphatidylserines, like the cephalins, were initially isolated from the brain, where they occur in considerable quantity.

We see that the structures of the discussed phospholipids have the following inherent features. They have two long aliphatic chains R and R′ that give the molecule hydrophobic properties. At the same time they have a positively-charged group arising from the amino alcohol and an anionic group from the doubly-esterified phosphoric-acid residue. Let us give the complete structural formula of a lecithin:



Evidently, the end of the molecule containing the charged groups is extremely hydrophilic. That is, it is distinguished by high affinity for water. Apparently, the phospholipids form complexes with proteins with a structure such that the long hydrocarbon chains are absorbed in hydrophobic (water-repellent) drops within the protein globule, while the charged groups stick out and tend to approach the charged groups of the protein. These general features of molecular architectonics suffice to permit the lipids to fulfill their functions. Thus, for example, various lecithins activate the enzyme cytochrome oxidase with equal success after it has been inactivated by removing the lipid component. One doesn't observe a very rigid specificity here. There is apparently also not a strict specificity in the experiments on modeling and reconstruction of membranes, as will be discussed below. Hence, the phospholipids are relatively simple surface-active substances without which biological structures (membranes) cannot exist and certain enzymes extractable from the membranes cannot function. However, the picture of the molecular structure of membranes, the way in which proteins and lipids are incorporated into this structure, and the mechanism of action of the lipids on the proteins are all as yet unsolved problems. For this reason we should consider this very important class of biologically active substances to be less understood than the incomparably more complex proteins and nucleic acids.

The next class of compounds that is universally distributed throughout living nature is the sugars and polymeric carbohydrates: cellulose, starch, glycogen, and some of their derivatives. The role and significance of these substances have been elucidated to a considerable degree, and have been described in detail in the biochemical literature.

III. SYNTHESIS OF BIOPOLYMERS. THE GENETIC CODE

The very important biopolymers, nucleic acids and proteins, are linear sequences of the different nucleotides (of total number 4) or amino acids (20 in number). The alternation of the different monomer units in the sequence is fixed with absolute accuracy. The replacement of one link of the chain by any other wrong link is a mutation, which is pregnant with catastrophic consequences for the structure of the protein and its function. We recall the above-mentioned example of hemoglobin. Replacement of a single link out of the 287 by a wrong one produces protein molecules with changed properties (Pauling, Ingram). The pathological hemoglobin becomes less soluble and crystallizes within the erythrocytes. The latter then break down and clog the blood capillaries. This gives rise to a severe, fatal disease, sickle-cell anemia. Therefore, organisms have a mechanism for synthesizing linear sequences of nucleotides and amino acids that acts with unprecedent accuracy. Ordinary copolymerization in organic synthesis always gives a statistical scatter in the structure of a copolymer. Fluctuations of the order of tens of percent are a quite general matter. The probability of error in the copying of nucleic acids does not exceed 10^{-10}–10^{-11}. This is made possible by synthesis of the polymers on a template or pattern. As I've stated, each double chain of DNA is a template for synthesizing two complementary chains. The principle of template synthesis is embodied in the very structure of DNA.

Error-free operation of the assembly machine is maintained by the Watson-Crick principle of complementarity of bases. Here we understand in general outline the nature of the molecular forces that provide exact correspondence between the template and the chain being synthesized. Synthesis of DNA on DNA is brought about by the enzyme DNA polymerase. The enzyme does not in itself determine the sequence of nucleotides. With equal success, it copies any template with any suitable arrangement of nucleotide links. In the last few months, Kornberg has succeeded in using this reaction carried out in solution, i.e., in vitro, to show that one can exactly copy the DNA of a virus (the bacteriophage φX 174), and infect cells with the synthetic DNA just as effectively as with natural DNA.

DNA is the polymer that contains information on the structure of all the proteins of the cell. However, DNA doesn't directly participate in synthesis of protein. The template for protein synthesis is the so-called informational or matrix RNA (mRNA), synthesized on DNA as a template. Hence, RNA is continually being synthesized on DNA in the cell. This process is called transcription. It is carried out by the enzyme RNA polymerase, which attaches itself to certain points on the DNA chains contained in the chromosome of the cell. The RNA polymerase moves along the DNA chain, and makes a copy of one of the two Watson-Crick chains. The copy is built according to the same principle of complementarity of bases (adenine versus uracil, and guanine versus cytosine). The structure of the RNA is complementary but not identical with that of

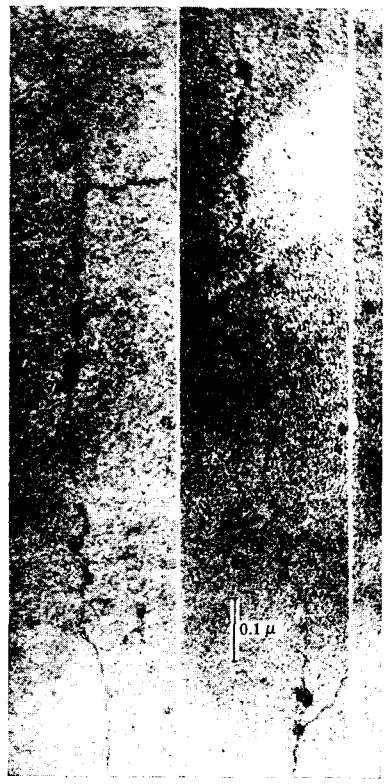


FIG. 6. Electron micrograph of several molecules of RNA polymerase attached to a DNA filament.

the DNA. First, the sugar differs, ribose being substituted for deoxyribose; second, one of the bases is modified: uracil serves instead of thymine. However, these variations do not interfere with the main feature, which is the formation of hydrogen bonds between the bases and the Watson-Crick complementarity principle, which serves as the main regulator in assembling the chain to be synthesized from the four monomers.

An excellent electron micrograph by Hall (Fig. 6) shows a DNA template and molecules of RNA polymerase strung on it like beads with definite intervals between them. They all begin their motion from a starting point, the so-called promoter. When one of them has moved far enough along, another attaches to the promoter, then a third one, etc. Single-stranded molecules of template RNA are the result. It is of interest that some promoters initiate copying of one of the DNA chains (one of these chains is jokingly called the "Crick" chain, and the other the "Watson" chain), while other promoters cause the enzyme to copy the other chain. Hence, matrix RNA can sometimes result from transcription from the "Crick" chain, and at other times from the "Watson" chain. However, transcription from both chains simultaneously with the same promoter never occurs.

As we shall see below, the promoters are accompanied by operators, which are the points on the chromosome to which repressors are attached. The latter prevent transcription and stop synthesis of certain proteins. This is the basis of the mechanism of regulation of protein synthesis. Without it, the cell would produce proteins in excess, without regard to its actual needs. The structures of the operators have some curious details, as yet not elucidated. However, on the whole, the transcription process is clear enough,

Then the synthesis of protein on the matrix RNA follows. Here there is a difficulty in principle. There are no structural analogies nor complementarity between the amino acids and the nucleotides. In addition, there are four nucleotides, but twenty amino acids. Hence, if the linear sequence of nucleotides determines the linear sequence of amino acids, a special code must exist, which establishes a rigid correspondence between several consecutive nucleotides and one amino-acid link. The relations here are about the same as between the ordinary alphabet and the Morse code, which is much poorer in variety of symbols, so that we must represent in it each letter by a sequence of several dashes and dots. The code translating the language of nucleic acids into the language of proteins is called the genetic code, and the process of protein synthesis itself is called translation. It is easy to understand that the most economical code must be a triplet code in order to get from four nucleotides to twenty amino acids (a two-letter code does not suffice, since the number of possible combinations is $4^2 = 16$, i.e., less than 20; a three-letter code gives $4^3 = 64$, i.e., about a threefold excess of combinations). Exact and irreproachable experiments have shown that the genetic code is actually a triplet code. A triplet of adjacent nucleotides in the template, coding for a single amino acid, is called a codon.

The number of triplet codons exceeds the number of amino acids. This leads to the so-called degeneracy of the code. Almost every amino acid corresponds to several alternative codons. The genetic code has been fully worked out now (Table I), and it has shown to be universal for all of living nature from bacteria to man. The whole epic of decipherment of the genetic code took less than six years, and is an amazing intellectual accomplishment. The principal method of decipherment was discovered by Nirenberg. This is a so-called in vitro experiment on simple artificial models. It was shown that simple synthetic polynucleotides of a given structure can serve as templates for synthesis of simple polypeptides. Since the chemist Khorana had developed to perfection the synthesis of the polynucleotide models, it was possible to discover in full the entire correspondence between amino acids and triplet codons. This is the genetic code shown in Table I.

Of course, the question arose of how exactly do model experiments (i.e., in vitro) correspond to what happens in the living cell (i.e., in vivo). Ingenious methods were devised to test the code in vivo. They are all much more indirect than the in vitro experiments, but there have been so many of them, and they are so varied that, on the whole, they are extraordinarily convincing. To my great regret, I can spend no time here on the experimental foundations of molecular biology. I shall refer the reader to the appropriate books and articles.[1]

The table of the genetic code contains three triplets that are not codons for any amino acids (two of these triplets are denoted in the table by the words "amber" and "ochre"). When a point mutation occurs in a cell, one nucleotide is replaced by another. Here one triplet

Table I. The genetic code.

First nucleotide	Second nucleotide U	C	A	G	Third nucleotide
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Ochre	Non-sense (?)	A
	Leu	Ser	Amber	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met*	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val*	Ala	Glu	Gly	G

*At the 5'-end of the chain, the codon serves to initiate growth of the protein chain, and codes for formylmethionine.

code is changed, and one amino-acid link in the protein chain is replaced by a wrong one. As it were, there is a misprint in the text. However, if a point mutation results in one of the three stated meaningless, or nonsense codons, the protein chain is simply interrupted at this point, and grows no further. One gets a shorter or longer fragment of the protein. This has been shown by direct experiment. It is currently assumed that the nonsense codons are punctuation marks signaling the end of the protein chain. This is a quite plausible hypothesis. However, it has not yet been proved by in vivo experiments. The terms "amber" and "ochre" mean the following. Special mutants have been found in bacteria that can correct the nonsense codons, and read them as being meaningful. For example, let us imagine a bacteriophage whose DNA contains a mutation having a nonsense codon. Such a phage will not grow, since one of its proteins fails to be produced. However, if the bacterium itself contains a mutation that "corrects" the nonsense codon of the phage and makes it meaningful, then such a cell will produce viable phage. In other words, this type of cells can correct an entire class of point mutations. This phenomenon is called suppression, while the mutations giving rise to the two types of nonsense codons that can be corrected by suppressors are usually called "amber" and "ochre" mutations (at the whim of the researchers that discovered them). We shall now see what the true physical meaning of the suppressors is,

Let us take up the problem of how the "assembly machine" for protein synthesis in the cell works.[4] Figure 7 shows the essence of our knowledge about this process (the system was suggested by Lipman). The basic workshop in which any of the proteins of the cell is synthesized is the ribosome. It consists of two particles (one subunit is somewhat larger, having a sedimentation constant of 50 S, and the other is smaller, with a constant of 30 S). The ribosomes consist of a special type of RNA and of several tens of different proteins. The exact structure of the ribosomes and the way that they function have not yet been elucidated. The chain of matrix RNA (having a molecular weight of the order of several million) passes through the ribosome. The connection between the codon and the corresponding amino acid is not made directly, but by way

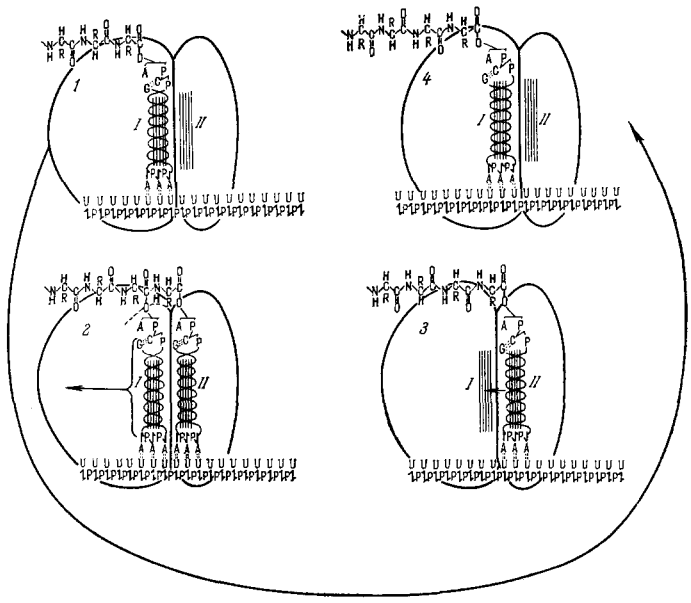


FIG. 7. Diagram of the four-step cycle of protein synthesis.

of an intermediate molecule, or adapter. The adapters are small polymers, or special transfer RNA's (tRNA) having molecular weights of the order of 25,000 (78—80 links per chain). In all, there is a somewhat greater number of different adapters (40—50 in a cell) than of amino acids. One of the ends of the adapter is chemically combined to some particular amino acid. The tRNA is linked to its own amino acid by a special enzyme, aminoacyl-tRNA synthetase. The number of these enzymes in the cell is approximately equal to the number of amino acids. The enzyme recognizes its own amino acid and one of its own tRNA's (the number of tRNA's for a given amino acid is usually more than unity, owing to the degenerate code), and it combines them through an ester link. On the other hand, the tRNA contains a triplet anticodon that is complementary to the codon on the template. Hence, the tRNA bearing the amino acid finds its place on the template, owing to the very same Watson-Crick complementarity principle. The codon-anticodon bond proves to be rather strong, so that error-free assembly of the protein chain will occur in accord with the structure of the template and the genetic code. This is how translation, or protein synthesis, takes place.

The suppressor mutations, which I mentioned above, are errors in the structure of tRNA. It has been shown that all of the tRNA's are synthesized by transcription from certain regions of the chromosome. Hence, the structure of the tRNA's can also contain mutational errors, just like the mRNA'S. If the error occurs in the anticodon, and in such a way that the mutated anticodon of any concrete tRNA comes to resemble a nonsense codon, we get a suppressor strain. The error in the tRNA anticodon can correct an entire class of mutational errors in mRNA. This suggested the idea of a very important experiment. One can now determine fully the sequence of nucleotides in transfer RNA's. Somewhat in the chain of 75 links is the anticodon. We don't know exactly where, but we can guess, since we know the structure of the codons.

However, once a suppressor strain has one particular tRNA with a mutated anticodon, then if we isolate it and study its structural formula, we can exactly determine by comparison with the normal non-mutated tRNA where the anticodon is and what its structure is. This has been done in the last few months by Smith and his associates, and it gave exactly the expected result. Thus, a codon-anticodon interaction by the Watson-Crick complementarity rules has been confirmed by direct analysis.

I shall not take up the experimental methods here. Nevertheless, I can't help but mention a stupendous achievement: the isolation of one individual tRNA from the many tens of them, and the complete determination of its structure. As I have stated, the transfer RNA's are small polymers. To separate them from one another is a problem quite comparable in difficulty with that of separating the isotopes of uranium. It suffices to say that this separation is performed with a countercurrent extraction cascade (extraction is the process of partition between two liquid phases). Cascades of 1500—2000 stages are used for complete purification of an individual tRNA! Evidently the recent experience of mankind with isotope separation hasn't played its final role in solving rapidly this problem of molecular biology.

Preparation of individual tRNA's requires working on an industrial scale, although thus far only scientific laboratories use them. This is one of the difficulties that shows how complex and expensive modern experimentation is.

To return to the mechanism of protein synthesis in the ribosomes (see Fig. 7), we can formulate its fundamental stages. The ribosome contains two positions, or sites. One of these is occupied by a molecule of tRNA to which all of the incomplete protein molecule is chemically attached (we are considering an intermediate stage in the synthesis, rather than the beginning or the end). The intermediate compound in protein synthesis is called a peptidyl-tRNA. We have isolated and studied it in our laboratory. The second site in the ribosome is occupied by a tRNA bearing the next amino acid. Evidently, the particular tRNA adsorbed on the ribosome is the one whose anticodon is complementary to the codon.

We shall call one of the sites on the ribosome the acceptor site, namely, the one containing the tRNA bearing the amino acid. We shall call the second site that holds the tRNA bearing the peptide the donor site. One of the enzymes acts to separate the end of the peptide chain from the tRNA, and attach it to the amino group of the next amino acid. This forms the next peptide bond. Thus a liberated tRNA is left in the donor position, and it is desorbed and goes into solution. Then the tRNA bearing the peptide (peptidyl-tRNA) must shift from the acceptor to the donor site. We can convince ourselves that this shift is really necessary, since it can be inhibited by some poisons (e.g., by the antibiotic fusidic acid), and then protein synthesis stops. When the peptidyl-tRNA has shifted to the donor site, then tRNA with a molecule of the next amino acid is adsorbed again on the vacated acceptor site. In order that the particular tRNA next in line should be adsorbed, the mRNA must shift with respect to the ribosome by three nucleotide links, i.e., by one codon. Perhaps both motions occur simultaneously, i.e., the shift of the peptidyl-tRNA from the acceptor to the donor site and the shift of the mRNA by the length of the next codon. This is not known yet. It is not clear how this ribbon-feeding mechanism works, but undoubtedly, the ribosomal proteins generate the necessary forces, and this motion is based on conformational changes in the protein molecules, as in the contraction of the muscle proteins.

We see that synthesis of proteins is performed by a real machine. It also needs special devices to initiate synthesis of the protein chain and to complete the chain. I have mentioned above the possible role of the nonsense codons as chemical signals for chain termination.

At present we do not know yet many of the details of protein synthesis, since the structure and function of the ribosomes have been studied little thus far. In all probability, we should expect many refinements. However, the major, most fundamental aspects of the phenomenon have been elucidated. The roles of mRNA and of tRNA, and of codons and anticodons are evident. We understand the fundamental sequence of events taking place in the ribosome. At the current tempo of development of molecular biology, refinement of the details may take another 3—5 years, and no more. However, as I've stated, beyond this fundamentally solved problem, others already arise that are even more tempting and difficult.

IV. REGULATION OF SYNTHESIS OF PROTEINS. DIFFERENTIATION

To proceed to problems that are much less clear, I shall first take up the problem of regulation. The problem of regulation of synthesis of proteins arises especially distinctly in one-called organisms: bacteria, yeasts, and algae. All these cells must constantly adapt themselves to the changing outer medium. The pattern differs in a compiex many-celled organism: each type of cells, or as we usually say, each tissue, exists under constant external conditions. Only the organism as a whole adapts itself to changes in the medium. The adaptive mechanisms at the level of individual cells are far from being as substantial as in one-celled organisms.

Why in general does the problem of regulation arise? An autonomous bacterial cell contains in its chromosome the genetic information for about 1000—1500 different proteins. However, the number of these proteins actually being synthesized does not exceed 10—20%. The bulk of the potentialities is not realized here, and this is advantageous to the cell, since its apparatus for synthesizing proteins works under great stress. As I've stated, the proteins are synthesized sequentially in the ribosomes. A cell of the bacterium Escherichia coli contains 5000—6000 ribosomes on the average. The time for synthesis of an individual protein chain is of the order of 5—10 seconds. This means that something of the order of 10^6 protein molecules can be synthesized in the time between divisions of the

bacterial cell (20—30 minutes). The cell contains in all several thousand molecules of each enzyme. This is not so very many, if we consider the huge amount of work that these enzymes must do in a short time.

If we force a bacterial cell to synthesize many additional proteins, then its condition deteriorates. A good example is given by the transfer of bacteria from a rich to a poor medium. Let us assume E. coli growing in a medium supplied with all the amino acids. This frees the cell from having to synthesize hundreds of enzymes that are superfluous for it, but necessary for fabricating the twenty necessary amino acids. When the bacterial culture is transferred to a medium containing no amino acids, then the cells must produce all the lacking enzymes. Here growth completely stops for several hours, and then gradually sets in again, although it remains much slower than in the rich medium (the period for doubling is one hour instead of twenty minutes). This means that the cell is handling its problems with more difficulty and more slowly. However, it is clearly evident from this that the bacterial cell does not synthesize proteins that it does not need, and thus it budgets its resources and energy for synthesis of those needed. This implies the existence of automatic regulation of enzyme synthesis and feedback from the products of enzyme activity to the synthesis of the enzymes themselves.

The fundamental mechanism of this automatic regulation in bacteria is already known, mainly through the studies of two scientists, J. Monod and F. Jacob.[5] All of the regulation occurs at the level of transcription (formation of mRNA from DNA). This was shown by direct experiments in which they measured the amount of a certain particular mRNA corresponding to one small group of genes. They found that the amount of it varies sharply during the process of regulation.

The fundamental tool that performs the process of regulation is a special protein, the repressor. It is synthesized in the usual way by its own mRNA. The function of the repressor is to suppress a group of associated genes and to prevent reading of the information from this group, i.e., to inhibit the process of transcription. Since suppression involves a group of genes, regulation occurs quite economically. Let us give an example.

In order to synthesize the amino acid histidine, the cell needs ten enzymes (according to the number of stages of biochemical synthesis). The structures of these enzymes are imprinted in ten adjacent genes, forming a so-called operon. One mRNA chain is produced in transcription from the entire group of genes, i.e., from the entire operon. Synthesis of mRNA begins from a particular starting point of the operon, or the so-called operator. This point is where the repressor is attached, and it prevents the enzyme RNA-polymerase from transcribing the entire operon.

Many details of this complex picture have currently been confirmed by experiment. In two cases, repressors have been isolated and purified. Here they could even measure the binding constant of the repressor to the operator. This binding proved to be strong (the dissociation constant proved to be of the order of 10^{-10}—10^{-11} moles/liter, instead of the usually-observed enzyme-substrate binding constants of the order of 10^{-4}—10^{-6} moles/liter). One can easily calculate that 10—20 molecules of repressor per cell are quite sufficient to keep the operator always saturated with repressors.

How then does feedback occur and is the process of transcription of mRNA started? The point is that the repressor must combine with the end product of synthesis, i.e., histidine. Only in this state is the repressor active and reacts with the operator. When histidine is lacking in the medium, the repressor does not contain histidine (the so-called effector), and it does not fulfill its retarding function. Hence transcription can occur at this moment. The corresponding mRNA is produced, and the necessary enzymes are synthesized. Histidine is produced by these enzymes. When the amount of it has become enough, the repressor is saturated with histidine, becomes activated, and interrupts further transcription. This means that the automatic mechanism has reacted to oppose increasing concentration of the amino acid. This example is typical of the so-called anabolic (i.e., synthetic) enzymes.

Regulation occurs in a slightly different way for a second group, the so-called catabolic enzymes. This group of enzymes does not synthesize anything, but conversely, participates in reactions of cleavage and oxidation of organic substances. The essential factors for these enzymes are not the reaction products, but the original substance that serves as a source of energy. Let me give an example.

Let us assume that some bacteria use glucose as an energy source. At a certain instant of time, the glucose in the medium is replaced by another sugar, e.g., lactose. In order to ferment lactose, new enzymes are needed. There have been practically none of them in the cell (the amount of these enzymes has amounted to several molecules). However, the chromosome contains information on these enzymes, but the corresponding operon is completely suppressed by a repressor, and transcription of mRNA does not occur. At the instant that lactose appears in the medium, it combines with the repressor (lactose is the effector) and inactivates it. Here the process is exactly opposite to the case of histidine. The effector exerts a neutralizing influence on the repressor. This phenomenon is called enzyme induction, and in this example lactose acts as the inductor. The final result is the same as in the previous case. The operator is released from the repressor, transcription of mRNA from DNA begins, and the necessary enzymes are finally synthesized. The amount of these enzymes in the cell after induction increases several thousandfold. This is the range of variation of concentration of proteins owing to regulation.

This is the overall picture. Individual variants can occur in it. Thus, is the synthesis of some enzymes (e.g., those used in the metabolism of arabinose), regulation does not occur negatively, using a repressor, but positively, using an inductor protein. However, this changes nothing in principle.

A number of interesting details of regulation are yet far from being elucidated. Thus, the total number of repressors (and inductors) in a cell can be estimated to be 200—300 when the total number of possible pro-

teins is 1000—1500. Since the repressors evidently are synthesized without any regulation, or as is usually said, constitutively, then the amount of each of them must be very small for reasons of economy. We have seen that we can estimate the amount of each repressor in the cell to be 10—20 molecules (and this means that the total number of them is 3000—5000 molecules). How does the cell contrive to synthesize so little of each of these proteins, when the transcription of their matrix RNA is not suppressed in any way? This situation still remains unclear.

One might ask way it has been so important to study the process of regulation in bacteria. Naturally, we are interested both from the theoretical and practical standpoint primarily in the higher organisms, many-celled plants and animals. However, molecular biology has developed mainly on the example of one-celled organisms, and has used them as an excellent model.

A many-celled organism also shows regulation of the rate of synthesis of proteins. However, it responds to signals that enter from the external medium to the organism as a whole, rather than any particular cell. The material agents that bring about regulation are the hormones, which are special substances produced by gland cells. They enter the blood or go directly into organs by way of the nervous system in response to external stimuli. The hormones have many regulatory functions. They affect the permeability of membranes and the activity of certain key enzymes. Apparently, one of the most important functions of hormones is to affect the synthesis of protein in cells. Possibly they do this by derepression of the transcription process from certain genes. This phenomenon can be observed in the whole organism or in individual surviving organs. However, the observations are especially convincing and important that involve cell and tissue cultures, those very important model systems for modern biological experimentation. In tissue cultures, the cells of a given tissue grow continuously in the form of a single layer on the inner surface of a vessel containing a rather complex nutritive medium. With a culture of hepatoma (liver tumor) cells, Tomkins and his associates were able to show that introduction into the medium of certain ketosteroid hormones of simple structure stimulates synthesis of the enzyme tyrosine transaminase by a factor of 10—20.

Now I shall proceed to formulate one of the most important unsolved problems of biology, the mechanism of development and differentiation. A whole many-celled organism arises from a single fertilized sexual cell, or zygote. The original cell is converted by successive divisions into thousands or billions of cells. While the first divisions are in their major features a copying of the original zygote, differentiation occurs abruptly at a quite definite instant. That is, the cells belonging to different regions of the embryo become somewhat specialized, and differ from one another in their structure and metabolism. In the first stages of development, the instants of differentiation are approximately determined by the number of prior mitoses. Actually, each cell acquires through its cell membrane information on its immediate environment. Depending on this information, sudden shifts occur in the synthesis of proteins, and this leads to the first stages of differentiation. Later, when there are many cells in the embryo, the times and directions of differentiation are determined by the action of certain types of cells on others by way of special chemical inductors. Thus, embryonic development (ontogenesis) proceeds by a series of abrupt changes until the complex organism has been produced.[6]

It is extremely important that the cell nuclei are identical in all cells, beginning with the zygote and ending with all the tissues of the complex organism. More exactly, the nuclei are identical in the amount and nature of the DNA. That is, they contain the very same store of genetic information as the nucleus of the first sexual cell. We might ask what this store of DNA is like, if one compares a cell of a higher organism with a bacterium. It turns out that the amount of DNA in the nucleus of a complex organism is three orders of magnitude (1000 or more times) larger than that in a bacterial cell. We have seen that the amount of DNA in a bacterial cell is enough to code for the order of 1000 different proteins. As I've pointed out, the cell actually uses about one tenth of this information. It has recently been possible to measure even how much of its DNA a bacterial cell is using in transcribing mRNA. It turned out that this fraction is close to 10%. A cell of a differentiated higher organism occurs under simpler conditions than a bacterial cell. It is specialized, and gets many necessary substances from the other tissues of the organism. Hence, the number of enzymes actually being produced in each type of cells is smaller than in the universal cell of a bacterium. Then, what happens with the whole abundance of genetic information that occurs in the nucleus of each concrete cell of a complex organism? This information is locked up, as we might suppose, if we compare the store of DNA in the cells of the bacterium and the complex organism, and repressed to the extent of 99.9%. The corresponding operons are completely inactivated, and not subject to transcription. No matrix RNA is formed from almost the entire length of the chromosome, and the corresponding proteins are not synthesized. Only a very small fraction of the genetic material (of the order of 1/1000) keeps its activity in each concrete type of cells. This makes it possible to generate many hundreds of different types of cells during the development of the embryo.

The store of information suffices for this. The main question is how protein synthesis is regulated here, and how newer and newer regions of the chromsomes are opened up successively during differentiation, and how this process regulates itself in time and space. Undoubtedly, the same concepts are applicable here as with the bacteria (gene repression and induction), and differentiation is based on "chemical signals" that the cell receives from its immediate environment. The experiments conducted by Steward on many ordinary plants are highly noteworthy in this regard.

Steward took up the complex differentiation of the organism of a plant, e.g., a carrot. By special methods (chiefly mechanical), he was able to disperse a particle of a tissue of the plant (e.g., the root of the carrot) into individual cells. He got what is called a cell culture. Evidently, he was dealing here with differentiated, specialized cells of a complex organism. These cells

were living, respiring, and metabolizing in a specially chosen sterile liquid medium. When certain growth hormones were added to the medium, the cells began to divide and form small cell colonies. In this state, the experimenter deposited them on a solid nutritive medium (agar) containing all the ingredients. As a result, the cells continued to divide and differentiate, and passed through all the stages of embryonic development, and whole plants grew out of them.

This experiment is extremely instructive. Isolated cells of a carrot root are put under conditions where they cannot affect one another. The "chemical signals" produced by the cells are diluted in the enormous volume of the medium. It turns out that the cells then dedifferentiate and lose their specialization, and their nuclei regain all their potentialities. In other words, a tissue cell of an individual formed organ acquires the nature of a zygote. When we transfer a small colony to agar, we facilitate differentiation, since the "chemical signals" produced by the cells remain in situ. Here a cell passes through its entire embryonic cycle, differentiation occurs, and we get a small plant from each cell.

One of the important features of ontogenesis of many-celled organisms is limited growth. Every organ grows to definite dimensions. This arises from the mutual influence of the cells through the chemical signals that they produce. As we know, formation of malignant tumors is manifested in the fact that the cells cease to obey the regulating action of their surroundings. The tumor cells grow and divide independently of the surrounding tissues—nothing limits them. In addition, they partially or completely dedifferentiate. That is, they lose their character as successors of the cells of the organs and tissues from which the tumor arose. However, the amount and nature of the DNA in the nuclei remain the same in tumors as in cells of the healthy organism. (In any case, no difference has yet been found.)

One of the most remarkable advances in biological experimentation is that one can find models for these phenomena in a relatively simple system. These systems are the above-mentioned tissue cultures. In special, empirically chosen sterile media, prolonged growth occurs in tissues taken from almost any organs of man or animals. It takes the form of single-layer structures (one cell thick). One can elicit malignant transformation of single cells in tissue culture, either by action of certain special oncogenic viruses (e.g., the Rous sarcoma virus) or by special poisons (carcinogenic hydrocarbons), or finally, by x-ray irradiation. The malignization of individual cells become visible, since volume growth of the culture begins at this site. The culture ceases to be single-layer in thickness, and an excrescence is formed. Transplanting such an excrescence into an animal gives a malignant tumor. Hence, malignant transformation is intrinsic even to such a simplified model as a tissue culture.

One of the most striking properties of tissue cultures is their spontaneous transformation after a period of prolonged cultivation. This transformation takes place randomly, with great fluctuations in time. If one cultivates in parallel several tens of identical tissue cultures, then they all undergo malignant transformation after a long enough time. However, this happens at different times in different specimens, in accord with statistical laws. Apparently, when one or a small number of transformed cells have arisen in a culture, they subsequently grow vigorously and exhibit selection. That is, they suppress the vital activity of the normal cells. In a certain simplified way, the malignization of a tissue culture is a model for what happens in the diseased organism.

Let us take up now the molecular mechanism of differentiation, or respectively, dedifferentiation. Evidently there are a number of problems here. The first question is: What is the material nature of the chemical signals for differentiation, how are they produced, and what in particular do they act on? (They ultimately act on a region of the chromosome, but their action may be mediated.) The second question is: How do the events of differentiation evolve in time, and why? The third question is: How does differentiation evolve in space, and why? That is, how does it lead to coordinated growth of different tissues in different spatial regions of the embryo?

There are some indirect data on all these questions, based on suppressing the differentiation process by poisons, on mechanical and radiation injury to embryos, on experiments on transplantation of small regions of certain embryos into others, and on obtaining various mutants showing defective differentiation, etc. There are as yet no direct data on the higher organisms.

I shall give some ideas based mainly on logic. We can estimate the number of different types of cells, beginning with the first stages of embyonic development and ending with the complex organism, to be of the order of a thousand. If we consider that each type of cells has its own region of the chromosomes which is repressed and derepressed as a whole in a single event, we arrive at a thousand repressors, or inductors. Hence, 99.9% of the DNA in each concrete type of cells is repressed. Again we come to an informational problem. There are chemical signals that must recognize certain points on the chromosome and interact with them completely selectively. This molecular recognition must take place for 10^3 cases. What is such a molecular recognition based on? Since one of the partners is evidently DNA, we must suppose that the second partner, i.e., the repressor or inductor, must contain a polynucleotide sequence, even when it is a protein. In fact, the Watson-Crick complementary interaction of bases is the only example that we know of molecular forces that make possible error-free recognition. However, a triplet code sufficed for recognition of the 20 amino acids. Recognition of a thousand different points on the chromosome requires a sequence of five nucleotides ($4^5 = 1024$). (As experiment has shown, these points are scattered chaotically along the chromosome.) The inductors or repressors are apparently proteins, but it is difficult to imagine how they can function without containing a pentanucleotide in their structures. This conclusion is purely speculative, but some data have now appeared to confirm it.

J. Bonner recently discovered the following important fact. If one extracts the chromatin (i.e., the nuclear nucleoprotein) from the nuclei of the cells of

any differentiated tissue, then one can transcribe it, using a separately prepared RNA polymerase, i.e., use it as a template for the enzymatic synthesis of RNA. Then one can measure what fraction of the DNA is complementary to the RNA obtained in this way. It turns out that about 5% of entire DNA has been transcribed under these experimental conditions. This is much more than one would expect for the synthesis of mRNA in cells. However, an in vitro experiment cannot exhibit as rigid a control as one done in vivo. The essential point is that most of the chromatin is repressed, and is not transcribed even in vitro. Chromatin contains DNA covered with special basic polypeptides, the histones, which form an envelope around it that prevents transcription. If one dissociates the histones with a solution of high ionic concentration (2 molar NaCl), then all the DNA begins to be transcribed. It would seem that addition of histones to DNA should give the opposite effect, i.e., cover a greater fraction of the chromosome, leaving a small part of it in a state accessible to transcription. It was found that this actually happens, but the histones combine with the DNA and inactivate it without specificity and without any order. That is, the RNA that is transcribed is not at all the form characteristic of the given type of cells. It was also found that there is a special form of chromosomal RNA that must be present simultaneously with the histones, in order that a particular region of the chromosome should not be inactivated, but should be subject to transcription. If one reconstitutes chromatin without removing this RNA, then the ability is regained for synthesis of the same completely specific RNA as before dissociation. This special form of RNA found in the chromsomes of cells of differentiated organisms is probably the substance that determines the repression and derepression of genes upon differentiation. Everyone is waiting impatiently for confirmation and further development of this work.

In conclusion, I shall proceed to the bacteria, and shall tell about the recent experiments of Jacob, Gros, and their associates, where the problem of differentiation was posed in an extremely simplified system. The problem is that of synthesis of a virus, the λ bacteriophage, which is parasitic on E. coli cells. This virus is gradually assembled from several types of proteins and a DNA molecule. It is a certain sort of differentiated organism. It consists of several organelles: head, tail, and fibers. In addition, when a cell is infected with this virus, certain additional enzymes are produced in the cell itself. They are necessary for synthesis of virus components, but don't enter into its final construction. According to the time of formation, one distinguishes the early proteins (they are the mentioned enzymes) and the late proteins, or structural proteins of the phage itself. Detailed study has already made it possible to show the existence of at least four operons in the chromosome of the λ phage. Each operon produces its own mRNA upon transcription. The types of matrices all differ in molecular weight, and hence one can distinguish them well in the ultracentrifuge. Only one of the operons (the "early protein" operon) is functioning at the initial moment of development of the bacteriophage (the first few minutes after infection). All the other operons are not being transcribed. One of the proteins (a quite definite one) that is synthesized by the cell from the first mRNA is the inductor that derepresses the synthesis of mRNA from the next operon. Similarly, one particular protein synthesized on the second matrix is the inductor for synthesis of the third matrix by transcription from the third operon, etc. These facts have been irreproachably established by genetic and concurrent biophysical experimentation. There is no reason to doubt them. We can hope that in the near future, the inductors of differentiation for this elementary case will be isolated and studied. The future will show how far we can carry the analogy with a complex organism and its development.

There is one logically evident detail in the problem of differentiation that demands special explanation. The problem is that the cells are copied literally in ordinary division, but differentiation leads to different types of cells. While the nuclear material in them is identical, the repression-derepression pattern in them evidently differs. This then gives rise to morphological, physiological, and biochemical differences.

One might ask how such a non-equivalent division can occur. The answer to this question is given in a very simple form by the experiments of Sonneborn on sex differentiation in paramecia. The paramecia are one-celled protozoa. At a particular moment of their development, the cells divide into two cells of differing sexual type, which differ physiologically from one another. Such a non-equivalent division occurs because the newly formed nuclei separate, and end up in different regions of the cytoplasm. The structure of the cytoplasm is genetically determined. That is, the information for it is contained in the nucleus. Here the cytoplasm is anisotropic or inhomogeneous: its properties differ in opposite regions of the cell. When the nuclei separate at the moment of division into different regions of the cytoplasm, they fall under the influence of their environments and of the surrounding proteins, which differ in composition. Hence, the different parts of the chromosome prove to be repressed in different ways in the two nuclei, owing to interaction of the nuclei with the adjacent regions of the cytoplasm. This gives to resultant cells of different structures and differentiated functions.

Many analogous examples are known in the ontogenesis of animals. Embryologists have introduced the special concept of a biological field, which denotes the existence of spatial gradients of various determining factors. As a result of this, the structure of the cell in itself has potentialities of non-equivalent division and divergent development, i.e., transformation into different types of cells.

As a whole, the situation in this field of science is such that biophysical studies have just begun. I have found it useful to point out the most important problems. How to approach them experimentally is a matter for the intuition of the investigators.

V. MORPHOGENESIS

How in nature do formed elements arise, in which the biological molecules (proteins, nucleic acids, lipids, etc.) are combined in a strictly defined order?

How are membranes formed, how is information on their structure transmitted to succeeding generations, and how is this information realized upon cell division? Until recent years, one couldn't give even conjectural answers to these questions. Only now have they become somewhat clarified. The first important result in this field was obtained in a very simple system, tobacco mosaic virus. This virus consists of one RNA chain and about 2000 identical protein globules, which are combined to form a protein shell in which the helical RNA filament is embedded.

Hence, the virus consists of one type of RNA and one type of protein. Nevertheless, formation of tobacco mosaic virus in a very simple example of morphogenesis. One of the most sensational discoveries of past years was the discovery of the fact (by Fraenkel-Conrat) that one can disassemble tobacco mosaic virus into individual protein subunits by a simple shift in pH to the acid side. Here the RNA is released from the virus, and can be separated from the protein. Then one can mix solutions of the RNA and the protein of the virus. On bringing back to neutral pH, one observes total reconstruction of the virus particles to their former dimensions (as measured in the electron microscope). The most amazing fact is that the reconstructed virus has the same infectivity as the natural virus. Thus, morphogenesis arises by self-assembly. All that is needed is the presence of both necessary ingredients, RNA and protein, and a favorable medium.

Why does this self-assembly occur? What forces participate in it? Now I must take up the concept of the quaternary structure of proteins. I've discussed above the primary, secondary, and tertiary structure of proteins. They resulted in formation of protein globules. However, the subject isn't exhausted with these three stages of structure. Entire protein globules can combine with one another in definite stoichiometric complexes. Most enzymes consist of several identical subunits combined into such a quaternary structure. The quaternary structure can be disrupted in various ways, the complexes dissociate, and the enzymes then usually lose their activity. What then holds the subunits together? Usually it is forces that arise between functional groups on the surface of the protein globules. In particular, they can be Coulombic forces between positively and negatively charged groups. The essential point is that the surface of the globule has a discrete, small number of such groups. This leads to formation of complexes of varying coordination numbers. In the special case in which the coordination number is two, the protein globules can associate like beads into a chain. This type of polymerization of protein subunits to form long chains is observed rather often in living nature. For example, this is how the muscle proteins myosin and actin are constructed. It is easy to imagine how polymerization of subunits with a coordination number of two might occur to form a helically-wound chain, rather than a straight one. We see such a case of quaternary structure in tobacco mosaic virus.

The protein shell (Fig. 8), which is hollow inside, is formed by a helically packed chain of protein subunits. Experiment has shown that such a polymerization of protein globules occurs even in the absence of RNA. However, then the length of the cylindrical particles produced is arbitrary. It can be small, but it can even exceed the length of the natural virus particles (2200 Å). In the presence of the viral RNA, the helical RNA chain is interwoven inside the helically-packed chain of protein subunits. As a result, one gets virus cylinders of strictly defined length, namely, the length sufficient for complete covering of the nucleic acid chain. What forces make the RNA chain pack itself inside the protein shell? Apparently, hydrogen bonds are implicated again here, along with Coulombic interactions. In this case, they occur between the side groups of the RNA and the side groups of the protein. This is not yet known exactly, but this phenomenon is not surprising.

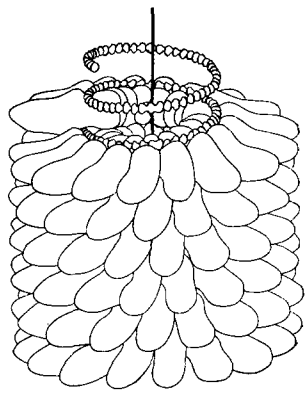


FIG. 8. A model of tobacco mosaic virus. The protein subunits are partially removed in order to show the packing of the RNA.

Formation of virus particles is the simplest case of morphogenesis. A much more complex and interesting case, which has been the object of further studies, is the morphogenesis of bacteriophages. Here we are dealing with several tens of different proteins combined into a complex differentiated structure. A phage possesses a number of organelles: the head (i.e., the reservoir containing a DNA molecule that is its chromosome); the tail, which attaches to the bacterial cell wall, and through which the phage injects its DNA into the bacterium; and finally, the fibers, which cover the end of the tail. The bacteriophage is the simplest example of a differentiated biological object, however small, that possesses organelles differing in structure and function.

How does morphogenesis occur in a bacteriophage? Two studies of extreme importance have succeeded in elucidating this. One of these (R. Edgar[7]) showed that the DNA of the phage bears the information for all the types of proteins of the phage, plus the information for some types of proteins that aren't incorporated into the phage, but are necessary specifically for morphogenesis. Just what are these extra proteins? It turned out that they are enzymes that fulfill the following function: they specifically link together the individual simple organelles of the phage, the head, tail, and fibers, which are produced by self-assembly from definite protein subunits. The function of these enzymes in assembly of the phage is quite clear. In order to understand it, we can appeal to simpler, well-studied examples.

I have said that the quaternary structure of proteins

arises by interaction of complementary groups on the surfaces of protein globules, e.g., positively and negatively charged, or donor and acceptor (in case of hydrogen-bond formation). However, these groups can be masked or demasked by special enzymes that detach the masking fragment from the protein molecule. It turns out that nature often uses this clever method. For example, the digestive enzymes, trypsin and chymotrypsin, are secreted by the pancreas in the form of the inactive precursors trypsinogen and chymotrypsinogen. The groups needed to form the active enzymes are masked in these precursers. These precursors encounter in the duodenum specific enzymes that activate them by converting them into trypsin and chymotrypsin (I. P. Pavlov).

It has now been precisely determined what happens here. Small fragments are split out of the chains of trypsinogen and chymotrypsinogen. Then, not only the primary, but also the secondary and tertiary structures of these proteins are changed. Consequently, an active site is "assembled" in the protein, since all the needed groups prove to be concentrated in one place in the necessary conformation. Such is the process of activation of trypsinogen and chymotrypsinogen, which is now well studied.

The second example, which resembles even more the assembly of a bacteriophage, is the process of blood coagulation. This is based on the protein fibrinogen, an elongated ellipsoidal structure of molecular weight 600,000. This protein is always present in blood, and is inert, since the two active groups existing on its ends are masked. At the necessary moment, when hemorrhage occurs, the enzyme thrombin acts on fibrinogen and splits short polypeptide chains from its ends, thus demasking its active groups. One says that the fibrinogen is converted into fibrin monomer. This is followed by polymerization of the fibrin monomer into fibrin fibers of macroscopic dimensions. These fibers block the injury in the wall of the blood vessel and stop the hemorrhage. Here we have a process of enzymatic demasking of active groups on the surface of a protein. The result is polymerization. Something similar occurs in the assembly of a phage from its structural elements. This also requires specific enzymes specialized for connecting certain parts of the phage.

Edgar made his discovery in a very interesting way. He obtained a multitude of phage mutations that had some particular genetic injury or defect. Some phages could form heads but not tails, and others vice versa. The mutations all belonged to the so-called amber type. This means that they showed up when the phage was grown on ordinary strains of bacteria. However, normal phages were formed on suppressor strains, i.e., those that correct the errors. Then Edgar demonstrated a remarkable thing: complementation, i.e., mutual complementarity of defective mutants. For example, if one grows phages that don't form heads and those that don't form tails, breaks open the cells, and mixes these incomplete preparations simply in the form of suspensions, then complete phages are directly formed from them in vitro whenever the preparations are complementary to one another. Suppressor strains of bacteria could be infected with the phages thus obtained. Hence, self-assembly of viruses from their individual structural parts has been observed.

By proceeding in this way, Edgar studied all his phage mutants and classified them. Here it turned out that there are mutants that possessed separately all the organelles of the phage, but didn't attach them together. The enzymes needed for this were absent. These enzymes could be added upon extracting them from cells infected with a phage having any differing defects. Complementation of the two types of mutants was found. Finally, he showed by extremely sensitive immunological experiments that the attachment proteins are not themselves incorporated into the phage particle. That is, they are enzymes that facilitate assembly, but remain outside the virus. This is the logic of the facts that led to the above-described picture of the self-assembly of a phage.

One might ask whether self-assembly of a phage occurs at once, or is a chain of processes extended in time in some way. We can now give a highly probable answer to this question. We have seen that, according to the latest data of Jacob and Gros, the proteins of the bacteriophage are not all synthesized simultaneously, but sequentially, in groups. Here one of the proteins of the first operon starts transcription of mRNA from the second operon. That is, it is an inductor for mRNA synthesis. Similarly, one of the proteins of the second operon starts synthesis of mRNA from the third operon, etc. Such a regulation of transcription brings the different groups of proteins into action sequentially. Hence it is natural to assume that assembly of the structural elements of the phage occurs in time in the form of a regular sequence of events.

One might ask whether one can extrapolate these first results to morphogenesis in cells: formation of the structural elements of a cell and assembly of entire cells. Such direct data are lacking as yet. However, the experiments with bacteriophages give a rational approach to this problem. These experiments say that the supermolecular structure is completely determined by the proteins and other components. They are formed by self-assembly with the participation of special enzymes.

The time-coordinated induction of synthesis of definite groups of proteins gives the self-assembly of structures a sequential, organized nature. The cells of higher organisms contain certain multiply-repeated structural elements. The ribosomes, which I've discussed above, are the simplest of these elements. The mitochondria, in which an entire arsenal of respiratory enzymes is concentrated, are a second example. Finally, cell membranes in the broadest sense, and in particular the outer membranes of cells, apparently consist of a mosaic of periodically repeating standard structural units (D. Green). In green plants, the plastids have such a periodic structure. In recent years it has been possible to show that all these structural elements (plastids, mitochondria, and membranes) contain DNA in small amount (1—4% of the nuclear DNA), and that this DNA plays a genetic role. That is, it contains information for certain enzymes. The experiments of Sager are especially important. Mutants of green cells were found in which the DNA of the plastids was specifically injured. Less convincingly,

such mutants have also been found for mitochondria. Phenomena of cytoplasmic heredity supplement the fundamental nuclear heredity. Many investigators are now inclined to think that the plastids, mitochondria, and subunits of membranes are capable of autoreproduction, i.e., self-copying. If we understand by this the replication of the cytoplasmic DNA, transcription of its operons, immediately followed by synthesis of the corresponding structural proteins, and finally, the self-assembly of the mitochondria, plastids, or subunits of membranes, then we shall continue the same ideas as were noted in morphogenesis by the experiments of Fraenkel-Conrat and Edgar. Of course, the more complex the structure is, the harder it is to imagine how morphogenesis is carried out by self-assembly, and the greater number of stages this process must have. However, apparently we have here fully surmountable quantitative difficulties. Biologists have repeatedly proposed introducing a higher-order template or pattern then the mere DNA chain, in order to explain the autoreplication of structures. However, this hypothesis always lacked concrete biophysical and biochemical mechanisms. It was an abstraction devoid of concrete content. Hence, the hypothesis of "supertemplates" afforded nothing toward setting up new experiments. And this is still the only thing that we demand of biological hypotheses.

VI. STRUCTURE AND FUNCTIONS OF MEMBRANES

Biological membranes are special two-dimensional structures 50—100 Å thick. They form the partitions, both between cells and between the different organelles within each cell. Their significance is quite understandable. They isolate the chemical reactions occurring within the cell from one another, and thus organize them in space and time. Their role is the same as that of the reaction vessels in chemical technology. Thus, for example, as we have seen, protein synthesis occurs in the ribosomes, while oxidative reactions occur in the mitochondria. The nucleus also is an organ separated by a membrane from the cytoplasm.

A remarkable function of membranes is selective permeability and active transport of various substances. The latter is the movement of molecules against a concentration gradient, i.e., against the direction of thermal diffusion. Evidently, active transport reduces the entropy of the substance being transported, and hence it requires continual expenditure of energy. In essence, active transport is a chemical process involving the same energy sources as the cell uses.

One can give many examples of active transport. Thus, nutrition of cells by transport of nutrients from the outer medium into the protoplasm always occurs by active transport. Secretion by glandular cells is another obvious example. A case of active transport that is very important for living nature and highly universal is the so-called potassium-sodium pump. The outer membrane of many cells absorbs potassium from the outer medium, and expels sodium. Red blood cells, muscle cells, and nerve cells behave in this way. As a result, the potassium concentration in a nerve cell is 30 times as high as in the outer medium, while the sodium concentration in 10—15 times lower. This situation is continually maintained as long as the cells live. If the cells are poisoned or their metabolic processes are inhibited, the concentrations of ions level out. Operation of the pump continually uses up a source of chemical energy, adenosine triphosphate (ATP), which is cleaved in the membrane by a special enzyme.

One might ask what the structure of membranes is, and how it is involved in their function. It is beyond our ability to answer these questions yet. Membranes are an exceedingly difficult object to study. They are insoluble, and hence it is difficult to observe chemical reactions in them. One cannot even prepare them suitably for electron microscopy, and hence the amount of reliable structural data on membranes is quite small. However, they evidently consist of special proteins and lipids. The lipids, which are insoluble in water, create the barrier that is impermeable to water-soluble metabolites, and without which active transport would be impossible. Any removal of the lipids from membranes renders them freely permeable to water and soluble substances.

One recently has a new approach developed in the study of membranes: to separate the membranes by chemical methods into subunits that can be brought into solution when the lipids have been removed.[8] These experiments apparently imply that membranes consist of repeating, regular spatial structures, or groups of protein macromolecules. In this sense, membranes are reminiscent of crystals. When the subunits of membranes are brought into solution, the enzymes contained in them retain their activity, and the membranes can be reconstituted from these protein subunits by adding to them the lipid components.

Interestingly, total removal of the lipids completely destroys the activity of some of the membrane proteins (e.g., cytochrome oxidase). Adding phospholipid to a solution of such an enzyme restores its activity. This is a sign that the lipid interacts strongly with the protein, and affects the conformation of the protein macromolecule. Apparently, it is precisely the protein composition of membranes that unambiguously determines their lipid composition, since lipids strongly interact with the side groups of the proteins, and are structurally complementary to the proteins to a certain degree.

Very interesting results have been obtained in studying the protein subunits of membranes. As I've said above, they are complexes of enzymes that are functionally interrelated, and catalyze a chain of successive reactions. A good example of this sort is fatty-acid synthetase, which was isolated by Lynen from cell membranes. One can see in the electron microscope that it is a cluster of many proteins, or enzymes that carry out the multistage synthesis of fats. Evidently the combination of a group of enzymes into a single spatial unit is not fortuitous. Perhaps the problem of transport of the reacting molecules from one catalyst to the next is solved especially simply under these conditions, just as a machine part moves on a conveyer belt or automatic assembly line, and is worked on successively by many machines and instruments. Therefore, membranes are not only tools of active transport, but are also carriers of complicated

enzyme complexes and sites of integration of multistage chemical reactions.

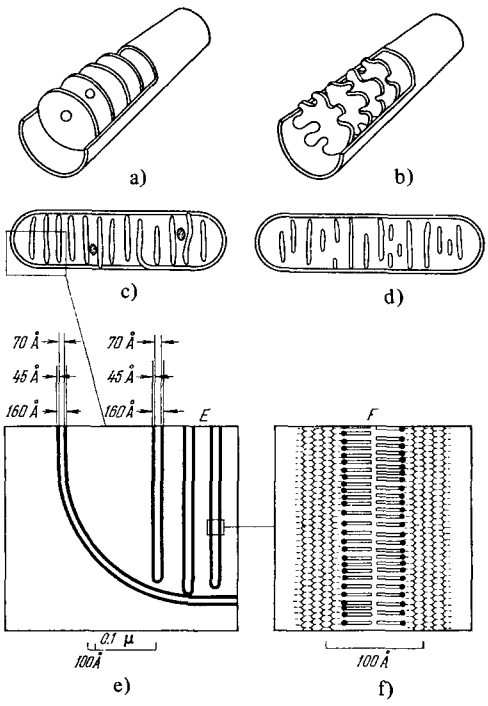


FIG. 9. Reconstruction of the structures of various mitochondria, based on electron micrographs of ultrathin sections. Details are shown at increasing magnifications of the microscope.

Unfortunately, we cannot now determine at all reliably the internal structure of a membrane. When they form morphologically-complex folds and interweavings, as, e.g., in mitochondria (Fig. 9), their functional meaning is also not yet clear. We also do not know yet even one example of an active-transport reaction that has been studied thoroughly. Indeed we still know little even about the most well-studied potassium-sodium pump. We have not identified the compound in which potassium and sodium are transported through the membrane. Neither have we elucidated how this transport is coupled with expenditure of chemical energy.

Let us take up another extremely important function of membranes, their excitability. This is the function that is decisive in the operation of the nervous system, and in particular, the brain. We have noted that lipids are characteristic of the substance of membranes. It is not fortuitous that lipids constitute such a great percentage of the mass of brain tissue (from 30 to 50% of the dry weight in different regions of the brain). Excitability means the ability of membranes to lose the property of impermeability to water-soluble substances when acted on by an electric field or special chemical agents. Here the structural peculiarities of the proteins and lipids are undoubtedly significant. As I've said, the insulating properties of a membrane involve the continuity of the lipid film. Lipids, and in particular the phospholipids, have a rather large dipole moment. When acted on by large potential gradients arising in the membrane (of the order of 2×10^5 V/cm), the lipid molecules must orient, and the corresponding electrical energy makes a substantial contribution to the free energy of the lipid film. Hence, this is important for its stability. As we know, a large potential jump near a phase boundary surface facilitates adsorption and orientation of dipole molecules on this surface. Hence, the importance of the electric field in a membrane for its stability is understandable in general outline. When an external electric potential depolarizes the membrane, the field in it declines, the lipid film loses its stability, and breaks appear in it. Then ions, sugars, and other water-soluble substances can be transported through it. Another mechanism of loss of stability involves action of mediators on the membrane proteins. Proteins often show conformational transitions when acted on by specific substances of low molecular weight. For example, x-ray structural analysis of the enzyme carboxypeptidase has shown that the protein is deformed and individual groups are shifted by the enormous distance of 14 Å upon formation of the enzyme-substrate complex. This indicates strong conformational changes in the protein macromolecule. If the membrane proteins undergo conformational transitions when they combine with mediators, it is easy to imagine processes that will alter the combination of protein with lipid, and thus affect the stability of the lipid film. Conformational transitions are highly cooperative. The identical subunits from which membranes are built behave like molecules in a crystal structure. As it were, they all simultaneously transform from one allotropic modification to another. Kittel et al.[9] have worked out the statistical mechanics of such transitions for membranes. However, the concrete mechanisms remain unknown. Excitable membranes generate electric signals, and respond to them. The activity of nerve cells, which will be discussed below, is fully based on this.

Lack of knowledge of the physical structure of membranes has brought into existence an entire new field of study: attempts to find physicochemical models of biological membranes, using films of fats, soaps, and lipids. All of these models lack proteins having enzymatic activity, and show no active transport. Hence they are very far from reality. And nevertheless, curiously, one can reproduce with these models some of the features of excitable membranes.[10] This makes clearer the types of molecular forces participating in formation of biological membranes. One can easily draw out thin films of water-insoluble lipids bounded on both sides by an aqueous phase. Such films of the Langmuir type are bimolecular lamellae. Their interiors are hydrophobic, i.e., water-repellent, and consist of hydrocarbon chains. The outer surfaces are covered with ionogenic groups that are wetted by water. These films prove to be non-conducting for ions, and their electric capacitance per unit area is of the same order of magnitude as that of biological membranes ($\sim 1\ \mu F/cm^2$). This means that their thickness is of the order of 50 Å. Furthermore, by introducing certain substances into the lipid films, one can furnish them with selective permeability to certain ions, e.g., the sodium and potassium cations. Then we obtain semipermeable membranes that behave like sodium or potassium electrodes. If the aqueous medium on the two sides of the membrane contains differing concen-

trations of any potassium salt, then potassium begins to diffuse from the one region to the other, whereas an equivalent number of anions cannot pass in the same direction. Consequently a potential jump arises at the membrane, and the latter becomes charged. This potential jump will depend on the ratio of potassium ion concentrations on the two sides of the membrane, according to the well-known Nernst formula.

One of the substances that gives a lipid model the properties of a semipermeable partition is the antibiotic alamethicin. Its chemical structure is peculiar. It is a so-called macrocylic polypeptide consisting of 19 amino-acid residues. Inside this doughnut is a carboxyl group that forms a salt with a potassium or sodium ion. Apparently, in the film five or six of these rings form a stack within which a channel filled with potassium ions remains. This permits specifically the cations to pass along the channel in a chain, while such a movement is ruled out for anions. This gives rise to an electrode potential on such a semipermeable partition. If in addition we apply to the film an external potential difference, then at critical values of the latter, the conductivity for cations will increase by a factor of 10^4. That is, breakthrough of the membrane will occur. It is assumed that it is exactly under these conditions that six-membered channel complexes are formed within the non-conducting film. All of this remotely resembles the properties of an excitable membrane, although conductivity (i.e., drop in the potential difference across the film) arises in the latter at the moment of depolarization, as will be shown below.

Figure 10 shows a hypothetical structure of the outer membrane of a nerve cell. The outer surface is drawn as being covered with gangliosides, which are lipids containing a molecule of the so-called sialic acid. According to all the data, this outer membrane of the neuron specifically interacts with calcium and potassium ions alternately. It reacts with the former in a state of rest, and with the latter in a state of excitation. In other words, the outer membrane releases calcium ions at the moment of excitation. We shall return to this fact below.

As for muscle cells, their excitable membrane, as distinguished from the nerve cell, transmits to the muscle protein a signal to contract. The nature of this signal has now been elucidated. It involves penetration into the muscle of calcium ions. Without the latter, the muscle protein actomyosin cannot interact with the energy source, ATP. This interaction requires a 10^{-5}—10^{-6} molar concentration of calcium ions. However, in fact, the calcium concentration within the cells

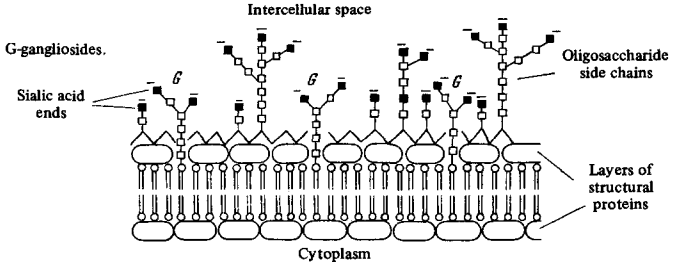


FIG. 10. Diagram (hypothetical) of the membrane of a neuron.

is maintained by active transport at a level of 10^{-7} molar, as compared with 10^{-4} molar in the outer medium. When the cell membrane is excited, the barrier hindering passage of calcium ions is broken. Hence a medium is created in the cells that favors reaction of the muscle protein with ATP.

Thus, membranes show four types of functional activity: 1) spatial separation of chemical processes (compartmentalization), 2) active transport, 3) integration of multistage biochemical syntheses, and 4) excitability.

VII. ELECTRICAL SIGNALS AND CODING IN THE NERVOUS SYSTEM

The central nervous system of man and animals continually receives varied information from the outer world via the sensory organs. This information is transmitted by the nerve conductors to the brain, where it is reworked and transformed into commands. The latter are transmitted by other nerve conductors to the motor organs, glands, and blood vessels, or in a word, to practically any functionally active structures. A fraction of the primary information, as well as its products of reworking, remains in the brain for a prolonged time in the form of memory. The cells of the nervous system, or neutrons (according to various estimates, man has from 10 to 100 billion of them) are occupied with all these very complex processes.

Physiologists studied the higher nervous activity of man and animals initially by using concepts developed by psychology. Then these were replaced by physiological concepts: unconditioned and conditioned reflexes, excitation and inhibition. It has always been clear that these physiological phenomena are based on material processes, physical and chemical, occurring in the matter of the nervous system. However, people began to study these material processes systematically only at the beginning of the 20th Century, and our understanding of them has now reached such a level that we can now formulate the fundamental scientific problems. The fundamental physical process that has been found in all nerve tissues is transmission of electrical signals. Rapidly-varying and substantial currents (of the order of tens of microamperes) flow in individual regions of the nervous system, at levels as small as individual neurons. One can record a portion of these currents, even by using external electrodes, e.g., as applied to the skull, i.e., by tapping off a tiny fraction of them to the outside. This is exactly how one studies a certain overall electrical process in the brain, by an electroencephalogram. However, one can also record potentials occurring in small regions of the nervous system, down to the single neuron.

Modern technique makes it possible to introduce and "implant" a thin electrode (several microns thick or less) into a single particular neuron in an animal. One can tap the potentials of a single neuron even in man during neurosurgical operations. The impulse activity of a group of neurons can be recorded in man under the so-called chronic conditions. The electroencephalogram of man gives a highly summated effect, resulting from the combined action of millions of neurons. One can easily convince one's self that the elec-

troencephalogram reflects all the possible information incident on a person, e.g., the effect of stimuli (light, sound, etc.). However, the question arises of whether the pattern of electric potentials in the brain actually corresponds to psychic reality and is a coded reflection of it. There is as yet no final proof of this, since the code functioning in the nervous system has not been deciphered. The conclusion that the pattern of potentials actually reflects the functioning of the nervous system is mainly based on observations of small assemblies of neurons in certain relatively simple animals (mollusks and crustacea). It is a great advantage to neurobiology that these animals have few nerve cells, and the latter are large in dimensions. On the other hand, these small assemblies apparently possess certain features of electrical activity that we find in highly developed organisms. On the other hand, one can actually hope to analyze the electrical behavior of such an assembly during a surveyable time period. It is with these primitive objects that people have been able to relate the detailed pattern of electrical signals in the neurons to the action of the stimuli eliciting these signals. We shall take up this point again below.

The fundamental observation that convinces us at least that transmission of an electric signal along a nerve suffices to make the acting mechanism (e.g., a muscle) function is the old experiment of Galvani (in 1791). The isolated foot of a frog jerks when an electrical impulse is applied to it from the outside. We know how many disputes this experiment aroused in the 18th Century. Its importance in physiology is eternal.

Let us examine now how action currents arise. Their source is a membrane, the outer envelope of the neurons. A potential difference arises across the membrane, owing to operation of the potassium-sodium pump. Every neuron resembles a charged storage battery: its source of electromotive force is the excess of potassium ions within the cell. In a state of rest, the outer envelope of the cell retains only a small permeability to potassium ions, and is impermeable to sodium ions and anions. However, the surface membrane of neurons of a particular type differs from the outer membrane of other cells. It is an excitable membrane. When a nerve impulse is being transmitted, a state of conductivity suddenly and abruptly arises in certain regions of the membrane. An ion current (potassium and sodium) arises at this point in the membrane, and it tends to level out the potentials and concentrations inside and outside the neuron.

Figure 11 schematically depicts a neuron. We see that it is a large cell with a highly developed surface covered with excrescences, the dendrites. On one side, the neuron transforms into a long, thin filament, the axon. The axon is the transmission line for the nerve impulse. It can attain a meter in length. The axon of one neuron joins with the dendrites of the adjacent one, and that one with the next, etc. Other forms of connection of neurons can also occur. Networks of neurons are formed in this way. The sites of contact have a special structure (Fig. 12). They are called synapses. This is where the initial breakdown of the insulating action of the membrane occurs. That is, this is where the current pulse is generated. The current pulse is transmitted from one synapse to another, where a

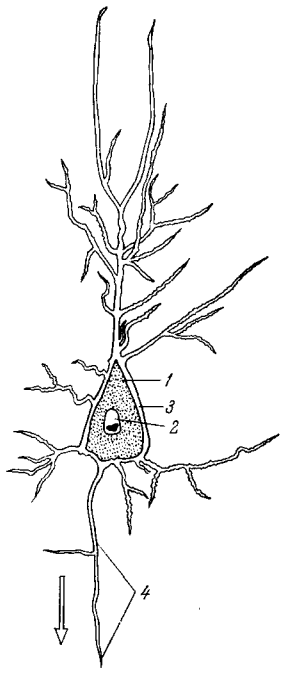


FIG. 11. A typical configuration of a neuron. The axon emerges at the bottom in the direction of the arrow. 1 – body of the cell, 2 – nucleus, 3 – envelope or membrane, 4 – axon. The arrow indicates the direction of propagation of impulses.

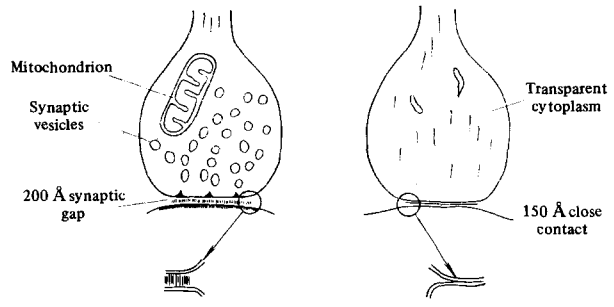


FIG. 12. Diagrams of synapses having a chemical (left) or an electrical (right) mechanism of transmitting impulses.

state of conductivity also arises, etc. Hence the neuronal net is not simply a passive conductor for an electric signal. Rather, it is a radio relay line, in which the signal is continuously amplified at the expense of the electric energy stored by the potassium-sodium pump. This is a very important circumstance, since nerve signals would decay very rapidly without a special amplification mechanism, owing to the good conductivity of the cytoplasm.

In all of the discussion below, I shall have to eschew many complications that have no fundamental significance for the problem under discussion, the coding of information in the nervous system. In particular, I shall not treat the second type of cells in the nervous system, the glial cells, and the role of the myelin sheaths that they produce around the axons.

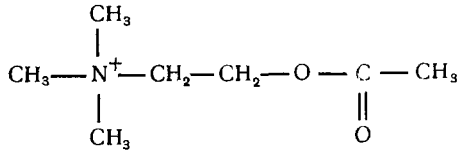
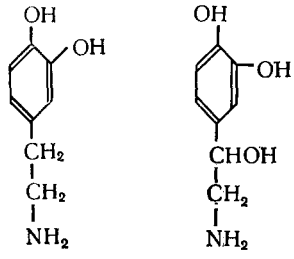
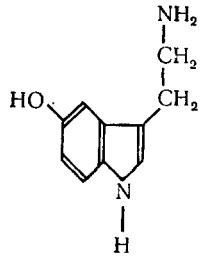
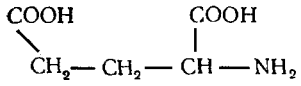
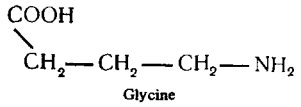
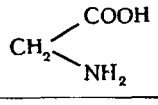
Neurons are not restricted to transmitting electric signals along their axons. This is only their simplest function. There are neurons of differing structures, and the number of synapses per individual neuron can vary (depending on its functional role) from a few to tens of thousands. The more complex neurons are analyzers: they add up many electric signals coming

from the neuronal net via the synapses. Here the summation is non-linear. Depending on the position of a particular synapse on the dendrites, the signal excited in it is attenuated and combined with other signals that come from other synapses and which are also attenuated in various ways. One signal arises from summation of all the attenuated signals, and it can prove to be lower or higher than the threshold of excitation of the axon membrane. If the totaled signal from all the synapses proves to be above the excitation threshold, then an excitation wave travels along the axon, and an output signal will be given at the appropriate synapse at its end to act on the next neuron.

Therefore, the outer membrane of a neuron is not homogeneous, but contains structurally different regions that perform different functions. First, there are the receivers, or input synapses situated in the dendrites. They produce the primary signals. Second, there are the regions of branching of the dendrites. The signals are partially damped in them (they play the role of attenuators) and are mixed or summed. Third, there is the axon membrane, which undergoes threshold excitation when the sum of all the partial signals that reach it is high enough. Further on, the resulting signal is propagated without decay along the axon. And its opposite end contains regions of membrane of a fourth type, the transmitters, or output synapses. In addition to these four types of functionally different membranes, we also find in neurons a fifth type, which serves as a generator of self-oscillations, independent of external sources of excitation. Later on, we shall encounter again the existence of such undecaying electrical oscillations in the central nervous system, or the so-called rhythms.

We shall now discuss what the reason is for the distinctive breakthrough of the membrane in a synapse. It was shown more than 40 years ago that there are special substances, or mediators of nerve conduction, that act directly on the synapses. Table II gives the most firmly established mediators. Apparently, these substances act on the proteins of the membrane by combining with them and causing conformational transitions. Consequently, the membrane in the synapse region becomes conductive for the water-soluble ions

Table II. Mediators of nerve conduction.

Acetylcholine 	— mediator of excitation
Noradrenaline and dopamine 	То же
5-Hydroxytryptamine (serotonin) 	» »
Glutamic acid  γ-Aminobutyric acid	— mediator of excitation and inhibition
 Glycine 	— mediator of inhibition » »

of potassium and sodium. Within a few milliseconds, the membrane regains its impermeability.

The structure of various synapses has been studied in recent years by electron microscopy. Figure 12 shows a synapse between a muscle cell and a motor neuron (a neuron transmitting a command for the muscle to contract). The diagram is an illustration generalizing the electron-microscopic data. The neuron ends in a sack. We can see mitochondria in the cytoplasm, as well as vesicles about 500 Å in diameter containing acetylcholine (this has been shown by direct methods). The neuron is terminated at the bottom by the presynaptic membrane, beyond which are the synaptic gap and the postsynaptic membrane, which belongs to the other neuron. When an electrical signal arrives at the presynaptic membrane, it releases a certain portion of the mediator, apparently that contained in about one vesicle (of the order of 2000 molecules of acetylcholine). Then the postsynaptic membrane is excited, and an electrical impulse arises in it (the postsynaptic impulse). Right after this, the enzyme cholinesterase immediately cleaves the liberated acetylcholine and stops the excitation. Whenever the synapse links two neurons, the postsynaptic impulse travels a short distance while being attenuated. It is added to other similar impulses, and when it arrives at the excitable membrane of the axon, it can start propagation of a non-decaying signal. When the latter arrives at the output synapse, it again liberates a mediator , and thus the excitation wave runs through the neuronal net from synapse to synapse. In fact, in the great majority of cases, the situation is not limited to one single signal. Thus, a motor neuron liberates at the nerve-muscle synapse a whole series of vesicles of acetylcholine, resulting in a series of electrical impulses. This then leads to muscular contraction. Single impulses do not suffice for this. Further, it has been possible in recent years to make observations on such a synapse in a state of rest and find random single signals in it. Apparently, this is the noise that arises from the spontaneous release of individual vesicles of mediator at the membrane. The amplitude of this noise is of the same sort as that of the useful signals in the transmission of excitation.

The detailed mechanism of operation of synapses having chemical transmission has not yet been elucidated. An essential point is that synapses having a purely electrical mechanism of transmission have been discovered as well. They differ in structure, in particular, in not containing vesicles full of mediator. The synapses having different chemical mediators also differ somewhat in their microstructure. It has been shown that one can trace chains of neurons in the nervous system that operate with a particular type of mediator, beginning with peripheral cells and ending with the cerebral cortex. Apparently the reason why the nervous system has such "cables" having different chemical mediators is in order to make certain functionally distinct regions of the nervous system somewhat autonomous, and to decrease the possibility of crosstalk in certain regions. On the one hand, all neurons are linked into nets. As will be shown below, this is necessary for the functioning of the nervous system. On the other hand, according to all data, the brain is a highly differentiated organ consisting of different types of cells, and the existence of many mediators, rather than one, is a manifestation of differentiation.

The central nervous system of animals and man is the object of action of various pharmacological agents: narcotics, hallucinogens, stimulants, tranquilizers, etc. In many cases one can detect a similarity of molecular structure of these agents with that of the mediators of nerve impulses. One might ask what are the advantages of the chemical mechanism of conduction in synapses as compared with the electrical mechanism. It is not so easy to answer this question. However, one fact is quite evident. The chemical mechanism ensures unipolar conduction, i.e., propagation of a signal in the required direction alone.

Let us examine now in somewhat more detail the process of formation of a single electrical signal. The potassium-sodium pump makes the sodium concentration within the neuron about 15 times as small than in the outer medium (which is a 0.15 M NaCl solution, while it is 0.01 M inside). The concentration of potassium ions is 30 times greater than in the outer medium (it is of the order of 0.03 M within the neuron, but 0.001 M outside). In the normal state (state of rest), the membrane is impermeable to both sodium and to anions, but is freely permeable to potassium. Only the potassium ions are potential-producing. They diffuse into the outer medium and produce a double layer at the membrane, the negative layer being within the neuron. The potential jump at the polarized membrane (rest potential) amounts to 60—70 mV. When it has been established, further diffusion of potassium ions ceases. Thus, the potassium battery is charged. If the membrane behaved like a semipermeable partition permeable to potassium ions and impermeable to anions, the electrochemical potential across it would be defined by the Nernst formula:

$$RT \ln\left(\frac{C_1}{C_2}\right) = eV_p,$$

where C_1 and C_2 are the potassium-ion concentrations inside and outside the cell, V_e is the equilibrium potential difference, and e is the charge of an electron.

If we calculate the expected potential difference in our case, it should amount to 90 mV. In fact, however, the potassium pump is operating in the resting membrane, and it increases the concentration of potassium within the cell by expending the free energy $\Delta\Phi$ of metabolic reactions. Hence, the true potential difference V will be less than the equilibrium value, and will be determined by the equation

$$RT \ln\left(\frac{C_1}{C_2}\right) = eV + \Delta\Phi.$$

As I have said, the rest potential is 60—70 mV. When a mediator acts on the membrane at a synapse, its insulating action ceases. A local flux of sodium ions into the neuron arises (since sodium is the predominant cation in the outer medium), and one of potassium into the outer medium. The conductivity for anions remains small. Here the membrane begins to be depolarized at the site of "breakthrough", and a damped depolarization wave runs along it.

When the damped electrical signal arrives at the excitable membrane of the axon, then conduction can

FIG. 13. A typical oscillogram of biocurrents. Time markers above (one division = 10 msec).

arise in it, starting at the very beginning point. The axon membrane becomes depolarized, starting at this point. The cation deficit within the axon will be compensated mainly by sodium ions entering from the outside, since the sodium concentration in the outer medium is high (0.15 M), while the potassium concentration is relatively small everywhere. The anions practically do not participate in current transport.

The depolarization current of the membrane propagates from the site of breakthrough, that is, from the beginning point of the axon. When the negative potential jump at some point of the membrane becomes low enough (10–15 mV instead of the initial 60–70 mV), its resistance sharply decreases, and it becomes the site of a secondary breakthrough. Consequently a steady-state depolarization wave is formed, and it propagates at a definite velocity along the axon. This wave has some additional peculiarities. Thus, the sodium-ion flux from the outer into the inner medium does not cease when the potential difference becomes zero. This is quite natural, since the sodium concentration in the outer medium is 15 times larger. Hence, diffusion of sodium continues, and is not balanced by diffusion of anions. Hence, the membrane is recharged from −60 to +50 mV. When the action potential has been generated, i.e., the membrane has been recharged from −60 to +50 mV, it again regains its initial properties: it becomes impermeable to sodium and only slightly permeable to potassium. This happens long before the equilibrium sodium potential has been reached. Consequently, the local action potential exists only for about 1 msec (Fig. 13), and it has the form of a sharp burst (spike). Its amplitude attains 110 mV. An essential point is that this signal propagates along the membrane (in other words, along the axon) without decaying, until it reaches the next synapse. We see that the excitable membrane actively particpates in the process of producing and maintaining the electrical signal. The proteins and lipids comprising it change their structures. Most probably, they undergo conformational transitions. This is what we find if we measure the optical properties of an axon at the instant of propagation of an excitation wave. Synchronously with the electrical signal, a wave of variation of light scattering and birefringence propagates along the axon.[13] The changes in the structure of the membrane are the reason for the change in its electric parameters. Let us examine how a passive semipermeable membrane polarized in an electrolyte solution behaves. A potential jump of the same order of magnitude is produced on it. If we puncture it at any point, it begins to depolarize, and the potential gradually falls to zero. We can write the equation for the current flowing along the membrane:

$$\frac{\partial i}{\partial x}=\frac{\partial q}{\partial t}=C\frac{\partial v}{\partial t},$$

where i is the current at the given point, v is the potential at this point, q is the charge at this point (per unit cable length), and C is the capacitance of the membrane per unit cable length. Further,

$$i=\frac{1}{R}\frac{\partial v}{\partial x},$$

where R is the overall resistance of the outer and inner media per unit cable length. Hence,

$$\frac{\partial v}{\partial t}=D\frac{\partial^2 v}{\partial x^2},$$

where D = 1/CR.

The solution of the differential equation for an initial perturbation expressed by a δ-function will be:

$$v=v_0\left(1-\frac{1}{\sqrt{4\pi Dt}}e^{-\frac{x^2}{4Dt}}\right).$$

We see that the solution is a strongly-damped depolarization region propagated along the x axis.

This is the situation with the original signal arising at a synapse. Upon proceeding 1-2 mm along the dendrite, the amplitude of the signal is substantially reduced. The path length traveled by the signal is the main factor determining the action of the "attenuator." The situation differs in the excitable membrane of an axon. As we have seen, the electrical signal propagates along it without damping, since the axon membrane does not behave passively, but becomes conductive at the site of passage of an impulse. We can estimate the rate of propagation of the signal as follows.[11] The potential along the axon will obey the ordinary telegraph equation, which can be written for any cable having distributed capacitance and resistance as:

$$\frac{\partial v}{\partial t}=\frac{1}{RC}\frac{\partial^2 v}{\partial x^2}-\frac{\mathcal{J}}{C}, \qquad (1)$$

where C is the capacitance per unit length of axon, R is the resistance (external and internal) per unit axon length, and $\mathcal{J}$ is the current flowing through the membrane per unit axon length. If we write the solution of the telegraph equation in the form of a stationary wave, then we can introduce the new variable ξ instead of x:

$$\xi=x-Vt,$$

where V is the rate of propagation of the signal. Then the equation is transformed into the form

$$\frac{d^2 v}{d\xi^2}+VRC\frac{dv}{d\xi}-R\mathcal{J}(\xi)=0. \qquad (2)$$

We can easily solve the problem if we assume a simple form of electrical signal $\mathcal{J}=\mathcal{J}(\xi)$.

Let us consider the propagation of a signal made of two rectangular waves (Fig. 14). We have retained the most characteristic features of the phenomenon, i.e., alternating discharges of the membrane, followed by recharging to the initial potential. For simplicity, we shall replace the actual time-variation curve of the current with a discontinuous function. If the membrane were to behave according to the all-or-none principle, i.e., were excited in a stepwise fashion, this pattern would be correct. In fact, excitation occurs sharply, but not abruptly. However, this is only slightly reflected

in the qualitative pattern. We get the solution shown in Fig. 14b for the potential v.

The analytic expression for the region $\xi \geq 0$ will be:

$$v(\xi) = \frac{1}{V^2RC}\left[\mathcal{J}_1 + \mathcal{J}_2 e^{-V^2RC(\tau_1+\tau_2)} - (\mathcal{J}_1 + \mathcal{J}_2)\, e^{-V^2RC\tau_2}\right] e^{-VRC\xi}.$$

We can easily write expressions for the potential in the other regions as well, but we shall not need them.[11]

In order to determine the parameter V, which is the rate of propagation of the signal, we must introduce an additional condition: at the instant that the membrane current starts, i.e., when $\xi = 0$, the potential v is equal to the excitation threshold v_1: $v(0) = v_1$. When we substitute this into Eq. (3), we get:

$$\frac{1}{V^2RC}\left[\mathcal{J}_1 + \mathcal{J}_2 e^{-V^2RC(\tau_1+\tau_2)} - (\mathcal{J}_1 + \mathcal{J}_2)\, e^{-V^2RC\tau_2}\right] = v_1.$$

This transcendental equation for V has two solutions, only one of which is stable. We can get a quite simple solution if we neglect the exponential terms. It is valid when

$$\left(1 + \frac{\mathcal{J}_2}{\mathcal{J}_1}\right) e^{-\frac{\mathcal{J}_1\tau_1}{v_1C}} \ll 1.$$

This inequality implies that the maximum potential in the impulse is considerably higher than the excitation threshold. This condition is obeyed in practice. Then

$$V = \sqrt{\frac{\mathcal{J}_1}{v_1RC^2}}.$$

The rate of propagation of the signal depends only on the current $\mathcal{J}_1$ of the discharge at the instant of breakthrough, and is practically independent of the reverse charging current.

If we take from the study of A. Hodgkin and A. Huxley[12] the values $v_1 = 15$ mV, $\rho = 35$ ohm-cm, and the radius $a = 0.02$ cm, then the resistance of the inner medium of the axon will be $\rho/\pi a^2 = 3 \times 10^4$ ohm/cm. The capacitance of the membrane is 1 μf/cm^2, or $C = 2\pi a \times 10^{-6} = 1.2 \times 10^{-7}$ F/cm. Finally, the current $\mathcal{J}_1 = 10$ μA/cm. Hence we get the rate of propagation of the signal

$$V = 30 \text{ m/sec},$$

which corresponds in order of magnitude with the experimental values. In practice, the velocity V varies over a broad range in different neurons. In man it usually amounts to tens of meters/sec. It is as much as 100 meters/sec in thick axons, and diminishes to a few meters/sec in the thinnest fibers. The rate directly depends on the diameter of the axon: the thicker the axon, the greater the velocity (owing to the larger discharge current $\mathcal{J}_1$).

Let us take up briefly some experiments that have verified the principles described here of formation and behavior of a single electrical impulse. The success of the experimenters lay in finding a good object for study. This object was the giant axon of the squid. Its diameter is as much as 1 mm, whereas the axons of most nerve cells of man are only about 10 μ in diameter. The length of the giant axon attains tens of centimeters. Apparently, evolution has rewarded the squid with a giant axon just so that a nerve signal can pass along it very quickly. The squid moves on the reaction principle by ejecting water. All of its muscles must contract simultaneously. In order to achieve this, the

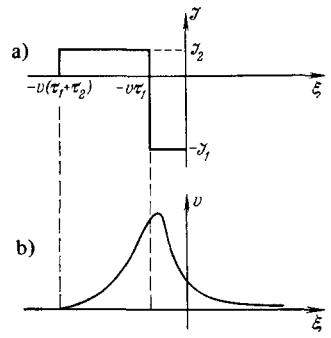


FIG. 14

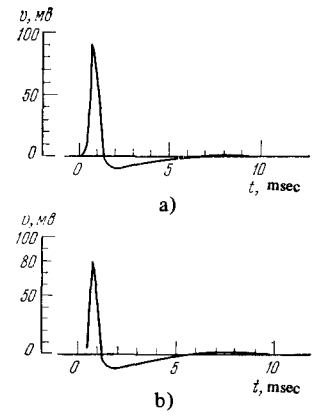


FIG. 15

FIG. 14. Idealized curve of a single impulse used for calculation. a) Current through the membrane; b) variation in potential along the membrane.

FIG. 15. Curve of a single impulse: a) experimental; b) calculated by the Hodgkin-Huxley theory.

signal to contract must reach all the muscles with a very small phase shift. For experimenters, the giant axon is a bonanza. It is easy to make a preparation of it, and to insert insulated microelectrodes several microns thick at various points in it. One can squeeze all of the inner medium, or axoplasm, out of it, and perfuse the axon with any salt solution through a capillary cannula. Finally, one can immerse the giant axon in any outer medium (only taking care that the osmotic pressure is balanced). When the electrolyte concentrations have been decreased, one can at will restore the osmotic pressure with uncharged sucrose. Experiment has shown that the axon keeps its fundamental properties under all these manipulations, and can act as a mechanism for transmitting an electrical impulses.

By using experiments on the giant axon, A. Hodgkin, A. Huxley, B. Katz, and other neurobiologists have discovered the mechanism of production and transmission of a unit signal. The first thing found was that the rest potential in the membrane is determined only by the difference in potassium concentration, and can change sign when there is more potassium in the outer medium. Hence, it has been shown experimentally that neither sodium ions nor any anions penetrate the resting membrane, while potassium passes through. Further, an electrical impulse was applied at a certain point on the axon, using special "stimulating" electrodes. A current was applied through these electrodes that reduced the potential difference at this point to the critical 15 mV. Then a signal arose from this point and was propagated, and could be recorded on an oscillograph at any distance from the point of excitation. It was shown further that the recharging of the membrane at the moment of excitation (the so-called "overshoot") involves the existence of a gradient of sodium-ion concentration. Introduction of sodium inside the axon decreased the recharging potential, while replacing the existing sodium (at constant potassium concentration) with sugar increased it. Finally, it was shown precisely that the rate of propagation of the excitation potential

is determined by the electric conductivity of the solution inside the axon. Replacing the sodium ions with sugar decreased the rate of propagation. All this agrees exactly with theory.

The elementary theory of the rate of propagation of a nerve impulse that I have discussed above describes the phenomenon qualitatively correctly. However, it is unsuitable for quantitative treatment. If one must study theoretically the true signal shape, then one should use a non-linear differential equation, since the electric resistance R of the membrane must be considered to depend on the potential. Hence, one must solve this equation by numerical methods. Hodgkin and Huxley made a systematic study of the flux of potassium and sodium ions through the membrane when a d. c. potential difference is suddenly applied to the membrane. Figure 16 gives the curves for variation of the flux of potassium and sodium ions as functions of the time and the potential difference.[13] We see that the process of establishment of the ion flux is a relaxation process for potassium. In addition, the dependence of the current on the potential is very sharp, but not discontinuous. The situation is still more complex for sodium ions. The sodium flux increases for 1 msec, then falls to zero according to a relaxation law in 3–5 msec. The reason for this phenomenon is not understood, but I have already mentioned it when I pointed out that the sodium flux through the membrane in itself ceases long before the overshoot potential reaches the equilibrium potential difference. Hodgkin and Huxley expressed the curves of Fig. 16 by empirical formulas, and used the latter to calculate the shape and rate of propagation of the electrical signal along the axon. They solved the differential equation for signal propagation numerically, and the results proved to agree quite satisfactorily with experiment (see Fig. 15), both in the shape of the signal and in the rate of propagation. Thus the phenomenological pattern of generation and propagation of a single electrical signal was perfected. Attempts are being made currently to improve further and to simplify the phenomenological theory. However, the cognitive importance of these studies is not great. As for the physical nature of the phenomena occurring within the excitable membrane, they remain unclear in many ways. In recent years, the studies of I. Tasaki and his associates have succeeded in making some progress on this problem as well. The Japanese researchers have completely replaced the biologically active sodium and potassium ions inside and outside the giant axon with a variety of others: lithium, rubidium, cesium, and ammonium. When these non-physiological ions were perfused inside the axon, while the outer medium contained the divalent ions of calcium, barium, or strontium, the axon membrane retained the property of excitability and the ability to generate electrical signals although they observed no active transport of these foreign ions through the membrane. Consequently, the function of excitability is independent of active transport. According to Tasaki's ideas, excitation of a membrane occurs in several stages. Initially, in a state of rest, the outside of the membrane contains divalent cations, apparently calcium, combined with a polymeric polyanion. When an external potential difference is applied, the calcium is replaced by univalent ions, and this is the ultimate reason for the conformational transitions in the subunits of the membrane. Consequently the membrane becomes conductive for potassium and sodium. When the impulse ceases, there is a reverse exchange of the bound potassium ions by calcium, resulting in a reverse conformational transition to the non-conductive state.

Data have been obtained in the laboratory of Nachmansohn that confirm this idea and show that calcium is adsorbed in the synaptic membrane by the same molecules that can combine with the mediator acetylcholine. When acetylcholine is applied to the membrane, calcium ions are released, and this is apparently followed by functioning of the same trigger mechanism that Tasaki studied (see[14]).

In addition to the neurons that carry out excitation processes in the nervous system, there are neurons that carry out inhibitory processes. The excitatory neurons and the inhibitory neurons form a mosaic, and are linked by synapses. An electric signal is also generated and propagated along the axon in the inhibitory neurons, but it is opposite in sign to the signal in the excitatory neurons. The electric current in these impulses flows from the inside to the outside, rather than vice versa. This behavior of the inhibitory neurons is explained by differing properties of their membranes.

How can an ion flux flow from the inside to the outside when the inner part of the cell is negatively

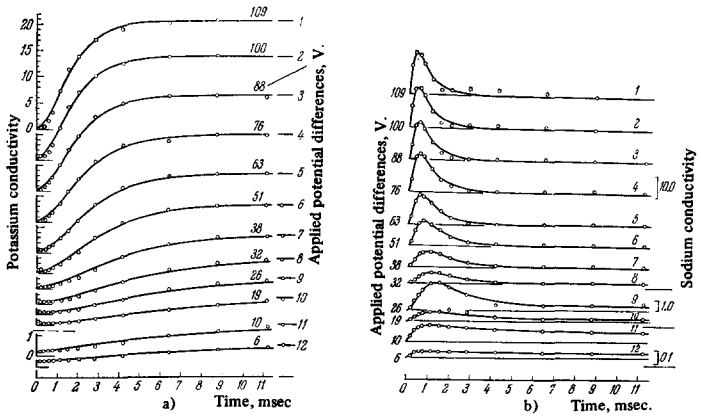


FIG. 16. Characteristic curves for conductivity of an axon membrane upon instantaneous application of potential difference (in mV). Conductivity for potassium – a); for sodium – b). 1–12 are the serial numbers of the curves.

charged with respect to the outer medium?

It turned out (this was elucidated by Eccles with the motor neurons of mammals, which are relatively large cells in the spinal cord) that the membrane of the inhibitory neurons becomes quite permeable to potassium ions upon breakthrough, but not at all for sodium ions. Hence a flux of potassium ions from the inside to the outside takes place. Consequently the negative potential within the inhibitory neuron rises even higher.

What is the limit of this increase? If the potassium concentration inside the cell is 30 times greater than outside, then the equilibrium potential jump at the membrane is 90 mV. Since the rest potential does not exceed 60 mV, a possibility still remains of appreciably increasing the negative charge inside the neuron. And in fact, the amplitude of the inhibitory signal in the neuron is 80 mV.

I shall discuss now the problem that is next in importance and complexity: the behavior of a neuron when not isolated, but incorporated into the system of connections natural to it. Very simple assemblies of neurons are just what we need in order to correlate the pattern of currents with the qualitative and quantitative pattern of external stimuli and with the behavior pattern of the animal, e.g., coordinated contraction of different muscles. As I've stated, nature has "met us halfway" here as well. There are a number of animals with a nervous system of elementary construction (without a brain, but with neural nodes, or ganglia). The gastropod mollusk Aplysia (the sea hare) serves as one of the favorite objects of neurophysiologists. The first thing that distinguishes the electrical behavior of a neuron under natural conditions of stimulation through synapses is the fact that it does not give individual impulses as a rule, but entire packets or volleys of signals periodically following one another (see Fig. 13). The intensity of stimulation (e.g., by light) is reflected not in the amplitude, but in the frequency of repetition of signals. The amplitude of a single signal is entirely determined by the physical parameters of the membrane: the capacity, the electric conductivity of the axoplasm, and the critical potential. The frequency of the impulses is determined by the number of vesicles of mediator arriving at the synaptic membrane. Consequently, the neuronal net is characterized by frequency modulation.

In their structures and functions, neurons are classified as "detonators" and "integrators." The former only transmit a nerve impulse along the net. Hence they have a relatively small number of very efficiently working synapses. They carry out fast, efficient transmission. Other neurons carry out integrating activity or analysis of information. They have numerous synapses, and a rather low level of efficiency of transmission. Hence, the individual postsynaptic impulses that arrive at the axon membrane do not suffice to excite the membrane. However, the action of nearby synapses is added together, and in an algebraic fashion, since the inhibitory synapses contribute with a negative sign. In all, when such an integrator neuron is excited, this is the result of a summation or analysis made on the information.

It was possible to observe in the very simple neuronal net of the mollusk Aplysia the elementary process of production of a conditional reflex, or learning. To do this, a particular neuron was chosen, with two nerve trunks leading to it. Stimulating one of the trunks by an electrical impulse elicited a strong response reaction (a strong excitation wave) in the form of a volley of impulses. Stimulation through the second trunk elicited a weak action (several single impulses). Then both stimulating signals were applied, one after the other with a small interval. Then a temporary relation arose between them. It developed gradually, and persisted then for tens of minutes.

When a temporary relation existed, it sufficed to apply to the neuron the weak or poorly-effective impulse to get a reaction equivalent to the action of the strong stimulus. Interestingly, development of the temporary relation required a regular, rapid alternation of the two signals in time. When the two impulses, the strong and the weak one, were applied at random intervals of time, then no conditioned reflex was developed. These results obtained on small assemblies of neurons permit us to rely on these assemblies as objects for studying the mechanism of formation of electrical connections.

One can observe very simple mechanisms of interneuronal interaction with the small assemblies. The treatment of visual information has been studied with the example of the simple camera eye of the mollusk Hermissenda. The eye of this animal consists of a lens, a special black diaphragm isolating a narrow cone of light, and a retina containing only five receptor cells. A microelectrode could be inserted into each of these cells. It turned out that some cells showed rapidly repeated electrical impulses when acted on by light, while their neighbors, conversely, showed inhibition (the so-called lateral inhibition). That is, they all form a single electrical network. The signal frequency in the excitable cells varies regularly as one changes the position of the light source. Shifting the light source from left to right increases the frequency of discharges in one of the two excitable neurons. Therefore, certain very simple aspects of formation of visual perception could be studied in principle with this elementary example.

Such examples of studying small neuronal nets can be multiplied. Apparently, this furnishes possibilities for relatively rapid advance in our knowledge of the nervous system. We shall now try to consider the nervous system in all its complexity. Evidently, the central point is the informational problem, i.e., the problem of the code. It is far from simple to formulate it. We have already seen that the activity of nervous cells is manifested in generating short electrical impulses. The amplitude of these impulses is fundamentally standard, while the frequency increases within certain limits with increasing intensity of stimulation (e.g., with increasing illumination in the eye, intensity of sound in the ear receptor, etc.). The amplitude changes observable in electroencephalograms can be ascribed to the fact they add up the potentials of a large and variable number of neurons.

One may ask where the possibilities for coding of information lie. How in general should we understand this coding? In order to answer these questions, we should take up the question of what type of information

arises in the brain.[14] Evidently it arises from the temporary interrelation of different sensory signals (i.e., from the sense organs). A very simple example is the Pavlovian conditioned reflex. An animal (or man) receives simultaneously from the outer world two different signals (e.g., food and a bell). Each of these signals passes through the nervous signal by its own pathway. One arises in the visual or smell receptor, and the other in the sound receptor. The two signals are linked by the fact of simultaneity. If this combination is repeated many times, then the two signals are associated in the brain (this is the formation of a temporary connection), and saliva is secreted when the bell rings, since the bell means food. If we wish, this is also a very simple example of associative memory or a very simple example of learning. One can try to represent any complex psychological events as associations of signals belonging to different receptors, i.e., belonging to regions characterizing a phenomenon from the standpoint of sight, hearing, smell, taste, or touch, etc. These associations occur between sensations, but undoubtedly they have a material nature. That is, they are actual physical associations between neurons and the currents that flow in them.

And here we come to the coding in the nervous system. If we consider the vast number of neurons in the brain, and assume that these cells are preassigned to different receptors and different effector mechanisms (muscles, glands, etc.), then we can illustrate the brain as a vast panel of flashing signal lights. Certain regions of the tableau indicate that the animal sees something, and others that it hears. A certain combination of lights in the first region (the visual receptor region) will denote a certain visible configuration of an object, the object of study. When combinations of lights in different regions of the tableau flash simultaneously, temporary connections can arise between them. Then it will suffice for only the lights belonging to the auditory receptor to flash under an external action (the bell). This will automatically elicit a flash in the combination of lights belonging to everything connected with food, and will start the effector mechanisms (salivary glands, etc.). This is the biophysical mechanism of a very simple conditioned reflex, or a very simple association.

One might ask how such connections arise between the simultaneously flashing lights. Evidently, a commutation system must exist that will permit electrical connection of different neurons when necessary, even those widely separated, when they are excited simultaneously. And there is another point: there must exist an automatic mechanism that tests this simultaneity. That is, it continually tests which connections between events should be fixed in memory, and which should be rejected as being fortuitous.

Let us return to our illustration with the tableau of lights. How could we begin to establish on it the regular combinations of flashing lights? We would examine the entire tableau from beginning to end, i.e., scan it once, twice, and three times. It is just this type of experiment that would give us a picture of how the flashing lamps are combined together, and what the statistical weight of these combinations are. Obviously, the brain must perform both of these operations. It must scan the entire neuronal net at high enough a frequency, as if examining all the random and nonrandom combinations of excited neurons that arise here, and it must make an internal electrical connection between those that are continually excited at the same time under an external influence. Consequently, the nervous code must be based, on the one hand, on the frequency of electrical signals in the neurons as a measure of the intensity of the signal, and on the other hand, on the position, or coordinate of the excited neurons. The coding mechanism consists in the connections, combinations, and associations of simultaneously excited neurons.

This entire picture is a logical consequence of the above-described experiments on primitive assemblies of neurons. We can imagine such a picture of integration of elements of the nervous system, although we still know very little on the detailed mechanisms of either the scanning of the neural net, or of the electrical connections effected in it, which materialize, as it were, the short-term memory. Nevertheless, there are already some experimental data confirming it.

I have said that the pattern of potentials taken from the brain, either through external electrodes (i.e., from rather large regions), or by implanting electrodes in individual cell populations, is rather complex. Psychological activity must amount to aperiodic signals (or volleys of signals), since the sensations and the results of their treatment are continually varying.

In order to visualize the results of measuring biopotentials in the brain, physiologists have invented the tableau method, which is an almost exact model of the picture that I have described. Electroencephalograms are taken simultaneously from several (and sometimes from some tens of) regions of the brain. Each electrode or pair of electrodes has its own channel in the electronecephaloscope or its own electron-beam tube in the toposcope. A change in the potential in the region of any of the electrodes is accompanied by a change in the luminosity of the corresponding tube of the toposcope. The entire picture is photographed on movie film. Here one can see the topography of potential variations in the resting brain (hence the name of toposcope for the instrument), together with its changes when acted on by external stimuli, or whenever the brain is busy with work demanding imagination of logical effort (e.g., solving a mathematical problem). The English have proposed the term "pattern" for the dynamic pictures depicting the biopotentials. No corresponding Russian word has yet been adopted, and our physiologists use the Anglicism "pattern." Evidently, the resolving power of the experimental methods is still small, and each of the recorded potentials is an overall effect from a large region of the brain. Thus, there is still a long way to go to the microscopic pattern of excitation and inhibition of individual neurons.

Let us return to the problem of coding. We have seen that the amplitude and shape of an individual electrical signal (a "spike") is determined by the properties of the membrane. That is, it is relatively standard. However, the frequency of repetition of signals, as determined by the sequence of mediator particles arriving at the synaptic membrane, bears information. The simplest case, in which a signal is generally propa-

gated, occurs in the neurons involving the touch receptors. Here the simplest "yes" or "no" information is transmitted. The essential instant of time is when the "yes" signal appears.

Packets of evenly-spaced signals are prodcued in the visual receptor, as in many others, and their frequency codes the intensity of illumination (or in general, the intensity of the stimulus). There are systems of neurons in which the duration or width of the packet of impulses bears a certain amount of information. Apparently, this language is adopted in the neurons controlling the high-voltage electrical discharges in certain marine fishes. We can imagine even other refined examples of frequency modulation that could be used. The biopotential patterns give examples of this sort, but it hasn't yet been determined whether these variations have a coded meaning or are noise.

Finally, it has been shown without doubt that self-oscillations arise in certain neurons at characteristic frequencies not related to external stimuli. If one superposes many electroencephalograms, averages them, and expands them in a Fourier series, then one can sometimes distinguish several periodic electric oscillations, or the so-called α-, β-, δ-, and θ-rhythms.

As certain physiologists have proposed (e.g., W. G. Walter[15]), the periodic rhythms reflect the scanning mechanism in the brain. By certain ingenious comparisons, one can even pin down the fact that the α-rhythm of man, which has a period of the order of 0.1 sec, is most closely involved with visual stimuli, while the θ-rhythm having a period of 0.2 sec sometimes proves to be associated with emotions. These rhythms vary in a very interesting way during childhood growth, in sleep, under action of narcotics, or with certain afflictions of the brain (tumors, hemorrhages, etc.).

One of the curious experimental methods consists in stimulating sensory organs of man with a rhythmic stimulus at the same frequency as a certain rhythm found upon analyzing the biopotentials. Thus, for example, one can stimulate the eye with periodic flashes of light using an electronic stroboscope, which can easily be synchronized with the α-rhythm of the brain that has been isolated from the noise. Such a stimulation, synchronized with a rhythm and primitive in content (simple changes of light and darkness), can elicit many interesting phenomena, both psychological and physiological, and pathological as well.

Some people experience something like hallucinations when acted on by such a stimulus. At first, it is limited to visual sensations. The person sees complex moving periodic patterns and figures. Then this is combined with non-visual sensations. Finally, a small fraction of the subjects develop pathological phenomena in the form of epileptic fits. If one simultaneously takes an electroencephalogram from many points of the brain and examines their pattern on the toposcope, intensive changes are first shown in a limited region of the cortex involving the visual receptor. They apparently involve states of excitation. Then the excitation state spreads to neighboring regions of the brain, and propagates more and more widely. This peculiar resonance phenomenon leads, as it were, to a breakdown of the boundaries of the local regions of excitation that are directly related to the given sense organ. A curious feature of the α-rhythm consists in the fact that most people show it when their eyes are closed, so that a direct action of the sense organs is ruled out. If this is actually scanning, then it belongs precisely to the stage of searching for a signal in the appropriate regions of the nervous system. Rhythmic oscillations are a sign of continual searching for connections, while a cessation in it indicates the end or completion of the searching process.

We shall not take up here the other periodic rhythms, for we are not interested in physiological details, but in biophysical mechanisms.

To sum up, I can say that the following points are important for drawing an overall picture of information transfer in the brain:

1) knowledge of the relation of the topography of the neurons with the content of the signals that they transmit;

2) knowing how to record experimentally the formation of electrical connections between distant neurons;

3) understanding the material mechanism that establishes these connections;

4) understanding what additional information is carried by the variations in time intervals between electrical impulses and what the noise is here;

5) elucidating the mechanism of scanning or searching for connections.

At present the work is just beginning on identifying the functions of individual regions of the brain. This is greatly aided by data obtained by implanting thin gold electrodes in the brain for diagnosis and therapy. For a long time, physiologists have restricted themselves to functional study of the cortex alone. Now people have started studying the subcortical structures, including those responsible for emotions and many psychological phenomena. The next few decades will show how far we can succeed in advancing along this line.

The problem of the material mechanism that establishes the permanent connections between neurons (the so-called plasticity of synapses) is now almost a total mystery. In plain language, this is the mechanism of memory. There are two fundamental types of memory: short-term and long-term. In long-term memory the information is stored for years, rather than seconds. Perhaps the only thing that we can state now, based on rather crude experiments with specific inhibitors, is that long-term memory may involve changes in the synapses, and requires synthesis of specified proteins in the brain cells, and hence also the synthesis of matrix RNA.

The problems of neurobiology, including the nature of memory, stand now on the order of the day. The conviction has repeatedly been stated that our brain and the mechanism of consciousness are unknowable. These ideas have been periodically revived even in the modern epoch. In fact, we can expect that the next few years will attract biophysics toward studying the brain on the molecular level, and many of the puzzles discussed above will prove to be solved even during the lifetime of our generation.

In conclusion, I would like to thank my many associates who have read this article in manuscript, in par-

ticular N. P. Bekhtereva, S. A. Neĭfakh, and P. G. Svetlov, for their kind criticism and for much advice and remarks.

[1]Reviews of the structure of biopolymers and of the problem of molecular biology and neurobiology are given in the books: J. D. Watson, Molecular biology of the gene, Benjamin, New York, 1965 (Russ. Transl., M., "Mir", 1968); V. M. Ingram, The Biosynthesis of Macromolecules, Benjamin, New York, 1965 (Russ. Transl., M., "Mir", 1966); G. S. Stent, Molecular Biology of Bacterial Viruses, Freeman, 1963 (Russ. Transl., M., "Mir", 1966); S. E. Bresler, Vvedenie v v molekulyarnuyu biologiyu (Introduction to Molecular Biology), M., "Nauka", 1966; Collected Volume, Molekuly i kletki (Molecules and Cells), Ed. G. M. Frank, Nos. 1–3, M., "Mir", 1968.

[2]B. Pullman, Quantum Biochemistry, Interscience, 1963 (Russ. Transl., M., "Mir", 1965); Electronic Biochemistry Academic, 1964 (Russ. Transl., M., "Nauka", 1966).

[3]D. J. Hanahan, Lipide Chemistry, Wiley, New York, 1960.

[4]A. S. Spirin and L. P. Gavrilova, Ribosomy (Ribosomes), M., "Nauka", 1968.

[5]F. Jacob, Science 152, 1470 (1966); Genetika 6, 3 (1965).

[6]J. F. Bonner, Molecular Biology of Development, Oxford, 1965 (Russ. Transl., M., "Mir", 1967).

[7]R. S. Edgar and I. Lielausis, J. Mol. Biol. 32, 263 (1968).

[8]D. E. Green and R. Goldberger, Molecular Insights into the Living Process, Academic, 1966 (Russ. Transl., M., "Mir", 1968); Mekhanizmy integratsii kletochnogo obmena (Mechanisms of Integration of Cellular Exchange), Ed. S. A. Neĭfakh, M., "Nauka", 1967.

[9]J.-P. Changeux, J. Thiéry, Y. Tung, and C. Kittel, Proc. Nat. Acad. Sci. 57, 335 (1967).

[10]P. Mueller and D. O. Rudin, Nature 217, 713 (1968).

[11]V. S. Markin and Yu. A. Chizmadzhev, Biofizika 12, 900 (1967).

[12]A. L. Hodgkin and A. F. Huxley, J. Physiology 117, 500 (1952); A. L. Hodgkin, The Conduction of the Nervous Impulse, Liverpool Univ. Press, 1964 (Russ. Transl., M., "Mir", 1965).

[13]L. B. Cohen, R. D. Keynes, and B. Hille, Nature 218, 438 (1968); I. Tasaki, A. Watanabe, R. Sandlin, and L. Carnay, Proc. Nat. Acad. Sci. 61, 883 (1968).

[14]Symposium, Frontiers in Molecular Neurobiology, Chmn, F. O. Schmitt, Proc. Nat. Acad. Sci. 60, 1055 (1968).

[15]W. G. Walter, The Living Brain, Norton, New York, 1953 (Russ. Transl., M., "Mir", 1966); D. S. Vorontsov, Élektrofiziologiya (Electrophysiology), M., "Nauka", 1967.

Translated by M. V. King