

## INFORMATION IN BIOLOGY

AN international colloquium on information in biology was recently held at Royaumont, near Paris. The aim of the organizing committee was to bring together workers interested in various fields of biology in which problems of information are faced—genetics, biochemistry, immunology, embryology, neurophysiology. Certain sessions were devoted to reports covering a broad field, followed by a general discussion; others were seminars with brief reports and original contributions on a limited subject. The discussions used to continue around the luncheon table, in the cloister or in the gardens of the Abbey. The value of such a meeting resides as much in these conversations, in improvised discussions and remarks, as in the prepared reports and contributions themselves.

M. Delbrück was chairman of the seminar on genetic information which occupied the first half of the colloquium. It is only one year since Nirenberg reported the specific stimulation by synthetic polynucleotides of amino-acid incorporation into protein material and an almost complete amino-acid code has already been proposed. Nirenberg and Lengyel both reported experiments which led to assigning one or more nucleotide triplets to each amino-acid; there is good agreement between most results of the two laboratories. The genetic code, however, is not yet completely solved; many questions and some difficulties remain. M. Bretscher directed the attention to two sources of error: in many of the experiments from which the present tentative code was derived the composition of the synthetic polynucleotide was assumed to be the same as that of the mixture of nucleoside diphosphates used to make them. According to M. Bretscher and M. Grunberg-Manago, this assumption is not warranted, and in many cases therefore the calculated frequencies of triplets on which the code rests might be wrong. The assumption that phenylalanine is coded by *UUU* exclusively is also uncertain and the triplet assignments rest on this assumption. Nirenberg reported that most of the polynucleotide added to the *E. coli* preparation is destroyed within a few minutes; a small part only, which is bound to the ribosomes, remains and acts as the template. It is not certain whether all the polynucleotides have equal chances of being bound. The presence of *U* in all the proposed coding units is not what one would expect, for no known nucleic acid is as rich in *U* or *T* as would be a template corresponding to the proposed code. According to the Cambridge-Paris group, poly-*A-C*, free of *U*, does stimulate the incorporation of three amino-acids, indicating the existence of codons (a word proposed by Crick for coding sequences) which do not contain any uracil. The proper interaction of polynucleotides with the ribosomes may depend on the solubility or the secondary structure, and *U*-containing polynucleotides might have an advantage over the others in this respect when assayed in the Nirenberg system.

The present code is but a first approximation: after these discussions, one would even doubt its significance, if it were not for the striking agreement between the code as it stands and many of the replacements of amino-acids observed in mutants. That Nirenberg's system can indeed make defined proteins when a natural RNA is provided to it, is shown by experiments from Fraenkel-Conrat's laboratory. Isolated tobacco mosaic virus RNA added to Nirenberg's system stimulates the incorporation of amino-acids into protein material which can be precipitated with an antiviral serum. A fraction of this material closely resembles the coat protein of tobacco mosaic virus; by hydrolysis, it gives apparently the same peptides. This fraction, however, is not perfect coat

protein, for it does not aggregate into rods as does the real tobacco mosaic virus protein, possibly because the acetylation of the  $\text{NH}_2$  end of the polypeptide fails to occur in *E. coli* extracts.

Incidentally, tobacco mosaic virus RNA causes the synthesis of other polypeptides besides that which resembles the coat protein, for the RNA stimulates the incorporation into protein material of two amino-acids which are not found in the coat protein. It is indeed most probable that tobacco mosaic virus RNA, like phage DNA, also codes for enzymes required for virus multiplication, and not only for the protein sheath of the virus. The previously established fact that many tobacco mosaic virus mutants have no amino-acid replacements in their coat protein already indicates that only part of the RNA molecule codes for the virus protein and that the rest of the RNA nevertheless is important for the process of virus infection.

Wittmann reported on further investigations concerning the amino-acid replacements obtained in nitrite mutants of tobacco mosaic virus. A given amino-acid is replaced by a very restricted choice of other amino-acids only, indicating that the number of meaningful substitutions within a codon is small. Since the replacement of an amino-acid does not lead to the replacement of its neighbours, the coding units are largely independent—the code is probably non-overlapping: but it is degenerate, for part of the amino-acids are coded by more than one triplet. If one assumes that de-amination results in the replacement of *C* by *U* and of *A* by *G* (*H* pairs like *G* and is ultimately replaced by *G*), the 64 possible triplets can be ranged in 8 groups of filiation; within each group, each triplet is related to three others by a one-step mutation obtainable by de-amination.

It is striking that all the transitions so far observed fall into three of the eight possible filiation groups. Unless some restriction to the observable transitions is imposed by factors irrelevant to the code, Wittmann's results suggest that the degeneracy of the code might obey certain rules. An interesting proposal for a logically degenerate code was presented by Woese on theoretical grounds. It is assumed that the three bases in each codon do not carry the same amount of information. In one of the positions, the four bases are recognized as different, but in the other two positions, only the nature of the group ( $\text{NH}_2$  or  $\text{OH}$ ), either in position 2 or in position 6 of the heterocyclic ring, is read. This is a degenerate code since in the member which codes according to its position 6, *U* is equivalent to *G* and *C* to *A*, and in the member which codes for ring positions 2, *C* is equivalent to *U*. Such a code would account for the wide differences in DNA composition of different micro-organisms without much change in protein composition, and also for the high *U* content of the codons derived from Nirenberg's system, since in one of the three members of the triplet *U* would have the same meaning as *G* and in another member *U* would be equivalent to *C*. Woese's code fits rather well with the coding assignments of the *E. coli* system, and with existing replacement data. In an excellent discussion, F. Crick analysed the present state of the coding problem, concluding that the genetic code is probably universal, partly degenerate, that there is little, if any, overlapping of the codons, and that the codons may indeed be made of three nucleotides.

The information coded in the template is read by transfer RNA. This is clearly proved by works done by F. Chapeville in F. Lipmann's group: cysteine bound to its specific transfer RNA was chemically changed into

either alanine or cysteine acid and the resulting product assayed for its transfer properties in a *E. coli* system programmed by poly-*UG* which normally stimulates cysteine incorporation. The modified amino-acids when bound to cystein transfer RNA behaved like cysteine, showing that the positioning of an amino-acid on the template does not depend on the amino-acid itself, and that it is determined exclusively by the properties of the transfer RNA to which it is bound. By fractionating transfer RNA's from *E. coli*, Sueoka showed the existence of more than one transfer RNA for certain amino-acids, for example, two each for Leu, Ser, Try, Asp, Glu, three for Ile. Moreover, results from S. Benzer's laboratory show that the two separable transfer RNA's specific for Leu must carry different reading sequences since one of them transfers Leu to *UC* polymers only, the other one to *UG*, not to *UC*; therefore, the two transfer RNA's which read leucine are not interchangeable. The concept of degeneracy will have to be reconsidered in the light of such experiments: the code is degenerate in this sense that one amino-acid can be coded by more than one codon, but if these different codons are nevertheless recognized as different by the reading system, the code is not degenerate in the same sense as Woese's code. Plurality, rather than degeneracy, is indicated by these results. It will be interesting to know whether the different leucine residues of the same protein are coded by the same word or whether certain positions are coded in one language and others in the other.

Certain mutants can be restored to an almost wild phenotype by mutations occurring far from the structural gene in which the original mutation occurred. Genetic studies on these 'suppressor' mutations provide extremely interesting data on the reading mechanism. An example was presented by A. Garen, who studied mutants of the alkaline phosphatase cistron, which do not produce any cross-reacting material, and therefore possibly do not produce any protein in place of the phosphatase in the locus of which mutations occurred. It is assumed that the mutation in the phosphatase locus changes a codon into a nonsense triplet, with the result of blocking the synthesis of the phosphatase polypeptide. The synthesis of active enzyme with normal immunological properties can be restored by a 'suppressor' mutation. This second mutation probably changes the structure of a transfer RNA in such a way that this RNA now reads the nonsense triplet as if it was a good codon. The function of the normal transfer RNA which was changed by the mutation cannot be lost, otherwise the mutation would be lethal; on the other hand, the change in reading specificity should not cause too much damage to the other proteins; there are several escapes to these difficulties. A misreading, anyway, is better than no reading at all, and this might justify the existence of several codons for the same amino-acid: the more there are of meaningful codons, the smaller are the chances for a mutation to produce a nonsense triplet (cf. Benzer).

As an introduction to the next section of the discussion, F. Jacob briefly outlined the problem of transfer of information from the gene to the machine which makes proteins; he summarized his theory of information transfer and regulation and emphasized the support it receives from current works on short-lived RNA's and on the Nirenberg system. Jacob then raised a few questions which introduced the next papers: Do rapidly labelled RNA's carry genetic information? What are the biochemical and physical properties of messenger RNA? At what level does regulation operate?

Rapidly labelled RNA now appears very heterogeneous in size. It is quite reasonable to assume that large messengers carry the message for long polypeptides. By combining chromatography and density gradient centrifugation, R. Monier was able to resolve rapidly labelled RNA into several components. Certain fractions have a sedimentation constant as high as 30S; this corresponds

to rather big RNA molecules, larger than those of ribosomal RNA, and large enough to code for several proteins. One wonders whether messages for several proteins can be delivered as a unit to the ribosomal machine, whether some messenger RNA's contain the information of a complete operon. F. Gros showed that the activity of various RNA fractions when assayed in Nirenberg's system depends essentially on their content in rapidly labelled RNA. Free ribosomal RNA is inert; it is not even attached to ribosomes. Another question is whether each molecule of messenger RNA is able to direct the synthesis of only one copy of a polypeptide or whether it is used over and over again. The differential effects of inhibitors on the synthesis of proteins and of messenger RNA in bacteria indicate that each molecule of messenger RNA can function several times, perhaps 10 or 20 times before losing its activity.

Evidence that the different genes of  $T_2$  bacteriophage send their messages at different times after infection was presented by two independent groups (R. Hessin and S. Spiegelman). Short pulses of tracer were given at various times after infection, the labelled complementary RNA was isolated by hybridization with phage DNA; the RNA's which are produced at various times after infection do not compete for the DNA, indicating that they correspond to different regions of the DNA molecule.

Another investigation by S. Spiegelman concerns the fate of the infecting RNA of a small RNA bacteriophage. Phage RNA binds to the 80S particles and stays there for a long time without being degraded. The association between this RNA and the ribosomes is very stable even at low magnesium concentration. This is in sharp contrast with what happens when the same virus RNA and ribosomes are confronted *in vitro*; association depends on high magnesium concentration, a large part of the RNA is destroyed during incubation and lowering the magnesium concentration results both in the separation of virus RNA from ribosomes and in the dissociation of the ribosomes themselves. Spiegelman also reported that this phage RNA causes the labelling of two proteins in the *E. coli* system. One of them is the coat protein; the function of the other one is not known.

The series of lectures on information in protein synthesis was closed by a report by M. Beljanski. In a soluble preparation from *A. faecalis* amino-acids bind to RNA's isolated from ribosomes; the binding depends on the presence of nucleoside triphosphates; analogues of purines or pyrimidines interfere with this fixation. Under certain conditions, rather large free peptides are formed.

M. Cohn very clearly summarized basic facts of immunological response, the intricate problems they raise and the current theories, of which he considered only the genetic models. The specificity is assumed to reside in the genetic material, the antigen regulating the production of a specific antibody or selecting the antibody producing system among the many existing potentialities. M. Cohn then summarized present knowledge on the structure of  $\gamma$ -globulin as deduced from immunological, chemical and genetic studies. The structural basis of antibody activity was discussed by R. Porter; his approach may answer the fundamental question as to whether the specificity of antibodies rests on differences in amino-acid sequence. Rabbit  $\gamma$ -globulins appear to be made of two identical A chains, flanked by two smaller B chains. Papain splits the antibody into three fragments: one of them, which was crystallized, is made of a rather large piece of the double A chain; it does not carry specific combining sites. The other two are probably identical to each other, each containing one of the two combining sites; they are formed of a B chain associated with the remnant of an A chain. The B chain was isolated; by itself it does not possess binding specificity. The fragment of A associated with B could not be studied easily because of a tendency to aggregate at neutral pH but such activity as survived the separation was always in the A chain. Comparable

studies by G. Edelman and B. Benacerraf on guinea pig antibodies indicate differences in electrophoretic mobility between the *B* chains of different antibodies; it would seem that *B* may take part in the specificity although it does not bind antigen when separated from *A*. The combining site appears to be in the *A* chain remnant, but it may be formed by the association of *B* with part of *A*.

The amount of work presented and discussed at Royumont was so great that when the programme was half way through, it was realized that there was not enough time left for the second half. It was agreed that the discussion on information in embryology would be replaced by two brief talks meant to state the problems and present perspectives. Thus C. H. Waddington considered essential problems of embryonic differentiation in terms of the mechanisms which control the synthesis of specific proteins. Experiments on nuclear transplantation and other less-direct evidence indicate that differentiation does not necessarily depend on permanent genetic changes; differentiation must involve switching on or off whole batteries of genes. Obvious questions occur: how are genes co-ordinated? How rigid is this co-ordination? What is the nature of 'genotropic substances', that is, of cytoplasmic substances which must control the activity of the genes? For if the cytoplasm in bacterial systems appears ancillary in regard of the all-mighty genome, embryologists view the cells and embryos as a much more involved system in which the cytoplasm exerts a very strong control on the genome.

Reasons why embryologists give a great importance to the cytoplasm were also illustrated by J. Brachet, who summarized fundamental facts about the early determinism of morphogenesis during cleavage. During this period of development, DNA replicates whereas protein and RNA synthesis are non-existent. Ribosomes isolated from eggs at this stage nevertheless incorporate amino-acids *in vitro*; the inactivity of ribosomes *in vivo* at that stage thus raises a clear problem of regulation. Hybrids between different species offer a means of studying the specific interactions between a genome and a foreign cytoplasm. Thus hybrids between frog and toad show a normal cleavage but are blocked precisely at the stage when specific protein synthesis normally begins. RNA metabolism in many nuclei of such hybrids is disturbed.

After this glimpse over challenging problems facing the molecular biologist of to-morrow, the symposium turned to considerations on enzyme structure and function, which were introduced by a brief report by R. Wurmser. The information content of an enzyme can be looked at from different points of view. Considering only the specificity of enzymes toward substrates, one obtains a small amount of information; but if one considers the actual support of enzyme activity, the protein itself, and calculates the amount of information from the negentropy of such a structure as estimated from the possible permutations of the amino-acids, figures of the order of several thousand bits are obtained. The reports and discussions which followed showed indeed that many structural features of the protein are important for enzymatic activity.

B. Keil raised the question as to whether the arrangements of amino-acids in protein chains obey certain rules. Common themes in amino-acid vicinity are obviously connected with function: thus the presence of aspartic acid close to the reactive serine is a common feature of esterases. Analogies of sequences between enzymes of similar function can also be recognized in the rest of the molecule, although this requires much more faith or enthusiasm. The notion of an active centre comprised of a few chemical groups is too narrow and altogether inadequate; the catalytic properties of an enzyme indeed depend on the rest of the protein as well—they can be altered by subtle changes occurring in other parts of the protein. Activation of trypsinogen and of chymotrypsinogen are good examples. Here the active

centre is created by a transition of the secondary and tertiary structures made possible by the rupture of one peptide bond. P. Desnuelle gave another example: lipases which are active on insoluble substrates only, as if adsorption at the interface would cause a structural change necessary for enzymatic activity to appear. J. Yon discussed factors from which the activity of trypsin may depend. A fragment of the trypsin molecule containing less than one-third of the polypeptide chain can be isolated by pepsin treatment. It is enzymatically active, but its specificity is not as restricted as that of intact trypsin; it can hydrolyse substrates of chymotrypsin, beside those of trypsin. On the other hand, the specificity of trypsin becomes narrower in the presence of copper ions: activity on lysyl linkages is increased, whereas hydrolysis of arginyl linkages is inhibited.

Another example was given by F. Labeyrie. Mild treatments may change the properties of a yeast lactic dehydrogenase: the enzyme which has been crystallized does not have the same affinity for a competitive inhibitor, D-lactate, as the native enzyme. A very striking type of enzyme modification was discussed by G. Tomkins. A glutamic dehydrogenase with a molecular weight in the range of one million can dissociate under various conditions into four molecules which have lost the ability of oxidizing glutamic acid, but which now possess all the properties of an alanine dehydrogenase. The type of chemical reaction catalysed is the same; but the substrate specificity of the enzyme depends on its state of aggregation. Extracts of HeLa cells contain alanine dehydrogenase, but no glutamic dehydrogenase. Addition of ADP causes the aggregation of alanine dehydrogenase into glutamic dehydrogenase. ADP in this case is not a metabolite, it acts physically, causing association; it is a kind of 'configurational co-factor'. Certain steroid hormones, on the contrary, split glutamic dehydrogenase into alanine dehydrogenase. These and similar processes suggest new mechanisms for the regulation of enzyme activity.

Delicate interactions between proteins or regions of proteins are also illustrated by complex enzymes. C. Gunsalus discussed the stereospecificity of dehydrogenases involved in the metabolism of camphor, the co-operation of two enzymes in a process of lactonization and the induction specificity of these enzymes. E. Racker described the case of glyceraldehyde dehydrogenase, a double-headed enzyme which catalyses two successive reactions; the oxidation of the aldehyde by DPN with the formation of the acylenzyme, and the phosphorolysis of the acylenzyme with formation of free acylphosphate. Chymotrypsin splits the dehydrogenase releasing a smaller protein which still catalyses the first reaction, but has lost the phosphorolysis activity. Erythrose diphosphate inhibits the first reaction in the native enzyme; but it has no effect on the split enzyme; binding of this inhibitor depends on a part of the protein which does not contain the active centre. Negative feed-back, that is, the process by which the end-product of a metabolic chain inhibits the enzyme which catalyses the first step in that particular chain probably rests on effects of a comparable type. G. Cohen gave a few examples of feed-back regulatory processes, with emphasis on branched metabolic chains. The kinetics of inhibition of the controlled enzymes is not quite of the competitive type; the structure of the inhibitor is indeed very different from that of the substrate of the enzyme, hence the suggestion to name this type of process 'allosteric inhibition'. The sensitivity of enzymes to allosteric inhibition is often a very labile feature of the enzyme protein; it can be abolished without impairment of the catalytic activity of the enzyme. Beside the catalytic site, there must be an allosteric site specific for the controlling metabolite.

On the basis of these facts, J. Monod proposed a generalization according to which induction and repression of enzyme synthesis as well as feed-back inhibition of enzyme activity operate by allosteric action. A regulated

enzyme has two sites: the enzymatic site and the allosteric site. Fixation of the regulating metabolite on the allosteric site slightly changes the shape of the protein and thus abolishes the activity of the enzyme. In the case of repression and induction of protein synthesis, the repressor produced under the control of the regulator gene is now assumed to be a protein (the evidence in favour of its not being a protein was not very good, and in the case of a lysogenic system there is now evidence that the repressor is probably a protein). The repressor-protein is believed to possess two sites: one of them recognizes the operator and interacts with it, thus locking the operon; the other site can bind the inducer metabolite. When the inducer metabolite is bound, the repressor protein is slightly changed, it loses affinity for the operator and the operon unlocks; a repressor metabolite, on the contrary, increases the affinity of the repressor protein for the operator. This new hypothesis affords a simple model of action of the inducing or repressing metabolite and thus removes one of the greatest difficulties of the previous models of enzyme induction. It also introduces more unity since the regulator gene is now regarded as a structural gene which specifies the repressor protein. The last day of the colloquium was devoted to a completely different field, namely, information in neurophysiology. A. Fessard gave a fascinating course on the mechanisms of brain, the nature of the messages, how they are processed and stored. He first outlined the several levels of organization: wiring diagram of neurones, neurone populations, cellular organization, synapses, membranes, macromolecular structures. He described the activity of the nerve fibre as revealed by the action potentials, and discussed how modulation of the time sequence of successive signals carry information. Depolarization, hyperpolarization and other processes occurring in the dendrite, and the role of the trigger zone in transmitting the messages were described. Then came the spatial aspect of impulse

reception in the brain, the interconnexions between neurones, the filtration of signals and the increase of contrast. Great emphasis was laid on the fact that the signals arriving to the brain are not restricted to a small primary area, that they go to many other parts where signals from other types of excitations also arrive. The code in which a certain excitation is translated is assumed to be in terms of patterns.

Finally, the problem of memory was outlined. Fessard described striking experiments on short-term and long-term memory, insisted on the absence of localization of the stored information, and indicated present interpretations. When hearing this lecture after those on enzyme structure and regulation one was led inescapably to think of memory as a regulation process, and to imagine a model in terms of allosteric effects on some synaptic structure. The conclusion of Fessard's lecture was that many of the elaborate properties of the nervous system can be explained by random connexion between neurones, regulated by external actions. A similar conclusion applies to the evolution of the highly organized structure of proteins and nucleic acids. The high information content of the genes, the 'antihazard' which is so obvious in the structure of a protein like haemoglobin or a double-headed enzyme can have their origin in known random events corrected by the purely deterministic challenge of the external medium. Teleological features of living matter are thus already accounted for in terms of causality.

The participants in this symposium are grateful to their French colleagues, especially Prof. E. Aubel, secretary of the organizing committee, and to the Délégation générale à la Recherche scientifique et technique which provided financial support. The Abbey of Royaumont, the gardens, the musical evenings, and supper by candlelight are delightful memories for those who were privileged to take part in this most successful colloquium.

H. CHANTRENNE

## THE INTERNATIONAL ATOMIC ENERGY AGENCY

AN outline of the programme for 1963 of the International Atomic Energy Agency is given in the October issue of the Agency's *Bulletin* (4, No. 4, 11; 1962). A total budget of 9,562,100 dollars has been proposed, representing an increase of approximately 7.5 per cent on the budget for 1962. Estimated expenditure under the heading of the "Regular Budget", financed out of assessed contributions by member states of the Agency, amounts to 7,337,500 dollars. The remainder of the total budget is assessed under the heading of the "Operational Budget" and is financed out of voluntary contributions by member States and other resources, such as funds made available under the United Nations Expanded Programme of Technical Assistance.

The Agency's technical assistance programme is to be extended. During 1962 requests for assistance involving the services of 40 experts, and equipment worth nearly 230,000 dollars were approved. This was in addition to some 36 experts already serving during 1962 under previously sanctioned projects. During 1963 the need will arise for the services of some 90 field experts, in addition to those whose assignments continue from previous years. Requests for equipment are expected to exceed the value of 250,000 dollars.

The Agency's training programme includes the award of fellowships and research grants. The number of applications for fellowships has increased steadily, but there has not been a corresponding increase in the funds available for the purpose. Hence during 1963 more extensive use is to be made of cost-free fellowships made available by member States, but even then it is anticipated

that not all deserving requests will be able to be met. For use in training on the spot the Agency provides two mobile laboratories. Approximately 750 trainees in the Far East, Latin America and Europe have used these laboratories in the past years, and it is expected that one of the laboratories will move to Africa late in 1962 or early in 1963.

A number of conferences and symposia are planned for 1963. A panel of experts is to meet to discuss the Agency's report on the results of an investigation into the cost of nuclear power in an integrated system and its comparison with the cost of conventional power production, and another group of experts will discuss the physics of heavy-water lattices. Symposia are to be held on materials for reactor control, including their metallurgical and physical properties and their design and fabrication; on new nuclear materials technology, with special reference to the use of such materials in the fabrication of non-metallic fuel elements; on the use of radioisotopes and radiation sources for the control of plant pests; on the use of isotopes in hydrology; on the biological effects of neutron irradiation and on the effects of radiation on the endocrine system; and on radiological health and safety in the mining and milling of nuclear materials; and the Agency's Scientific Advisory Committee has recommended that a seminar be held on some aspect of nuclear, solid-state or plasma physics.

Research contracts in tropical medicine are to be awarded with the view of promoting the application of established radioisotope techniques to the study of diseases affecting large groups of people in the less-