

PROTEIN STRUCTURE AND BIOLOGICAL ACTIVITY

PARIS SYMPOSIUM

IN 1951, a Commission on Proteins was created within the Section on Biological Chemistry of the International Union of Pure and Applied Chemistry. Its aim was to consider the possible need for, and provision of, a few standard proteins, rigorously tested by both chemical and physical methods, and available internationally to anyone who might wish to check his own methods against a standard. The first protein to be distributed in this way was a large batch of crystalline beef insulin*, prepared in Great Britain by the British Insulin Manufacturers.

The description 'standard' is used advisedly, for one is hesitant to apply any longer within the field of protein chemistry a term with the finality of 'homogeneous'. In recent years, the most respectable proteins have been resolved by several kinds of simple and sensitive techniques into two or more components, each perhaps with the same biological activity, but differing slightly in some chemical or physical feature. In some cases, it is not known how such 'microheterogeneity' arises; but just as Kossel in 1901 squarely faced the great diversity and complexity of proteins as a class—attributes which had taken fifty years to establish—so now the biochemist must anticipate the diversities which may arise in any one protein, either genetically as a mutation, as the result of species differences, as slight modifications in the course of preparation, or perhaps, as Pauling suggested, as an occasional mistake in the biosynthetic mechanism.

The Commission decided at its meeting in Zurich in 1955 that a symposium on protein structure and its relation to biological activity would greatly assist the Commission in evaluating the relevance of its earlier objectives. The executive committee of the International Union of Pure and Applied Chemistry, through Dr. Murray Luck, gave generous financial help, which enabled the organizers, Dr. Stanford Moore (United States), Prof. A. Neuberger (Great Britain) and Prof. J. Roche (France), to invite some twenty distinguished speakers and as many others to a symposium held in Paris during July 25–29, following the nineteenth Conference of the International Union of Pure and Applied Chemistry itself. Both papers and discussions will eventually appear as a monograph†, and the present account (prepared in the main without reference to manuscripts) cannot possibly do justice to the uniformly high standard of the communications, reflected by the enthusiasm of the audience in the congenial atmosphere of the Collège de France.

Following the usage of Linderström-Lang, it is convenient to consider the structural concepts of proteins in terms of the primary structure (amino-acid sequence), secondary structure (*intra*-molecular configuration of the polypeptide backbone—CH.CO.NH—) and the tertiary structure (*inter*-molecular bonding and topography of the side-chains). Although our knowledge of secondary and

tertiary structure is only fragmentary, it became evident during the meeting that a bold and imaginative approach by several workers had yielded fascinating results. It was also quite evident that the revolutionary progress in defining primary structure arises first from the application of many kinds of chromatographic techniques, and second from the introduction and expansion of end-group methods. It is not yet possible to assess in any historical sense the debt we thus owe to the pioneers in these fields, Martin and Syngé on one hand, and Sanger on the other.

Primary structure. The complete elucidation of the amino-acid sequence in insulin has led to more ambitious attempts to determine that of much longer chains. Through the work of S. Moore and W. H. Stein, and of C. B. Anfinsen and their co-workers, one may anticipate very soon the complete sequence in ribonuclease, a chain of 124 residues, in which the positions of the intrachain disulphide bridges have already been allocated. Steady progress was also reported by Cl. Fromageot and his colleagues in the case of hen lysozyme (128 residues), a difficult protein by virtue of its high content of tryptophan. The structure of pig and ox adrenocorticotrophic hormone, described by C. H. Li (39 residues), has a fascinating interest in view of the occurrence of amino-acid sequences common also to the α - and β -melanophore-stimulating hormones, the structures of which have been determined by C. H. Li and by J. I. Harris and their collaborators. Features which augment one type of biological activity at the expense of another are already apparent. The free amino-group in the terminal serine of corticotropin, which is correspondingly substituted by an unknown group in α -melanophore-stimulating hormone and by an additional peptide sequence in β -melanophore-stimulating hormone, appears to be necessary for its activity; if obliterated by oxidation, its corticotrophic activity is lost, although the slight melanophore-stimulating hormone activity persists.

In spite of the extension of sequence studies to these and other proteins and peptides, it is worth noting that the insulins are still the only proteins containing more than one polypeptide chain the complete structure of which is fully known.

Secondary and tertiary structure. The classical denaturing agents such as heat, acid, alkali and urea may induce reversible or irreversible changes of configuration which can now be followed by a variety of methods. K. Linderström-Lang described the use of deuterium to study the exchangeability of the imide hydrogen atoms of the polypeptide backbone, and discussed the results obtained with poly-DL-alanine which is known to exist in the α -helical form of Pauling and Corey. The rate of exchange as a function of hydrogen-ion and hydroxyl concentrations has suggested that H-bonds can be broken singly only at the ends of the helix, whereas in other positions at least three must dissociate simultaneously. In both insulin and ribonuclease, the observations suggest that exchangeability is greatly restricted in

* *Nature*, 179, 567 (1957).

† To be published by Methuen and Co. (Great Britain) and by John Wiley and Sons (U.S.A.).

some parts of the molecule, but is instantaneous in presence of denaturing agents, or after opening out the molecule by oxidation of the restraining intra- and inter-molecular disulphide bridges.

Changes in optical rotation, shifts in ultra-violet absorption, dilatometry and measurement of intrinsic viscosity may also reflect the extent of configurational change, and even where this is considerable, as in ribonuclease, biological activity can be retained. Indeed, the examples of this kind of phenomenon are now quite large, and have led two participants at the conference, Emil Smith and C. B. Anfinsen, to inquire how much of the protein molecule is simply 'junk'. Even covalent linkages may be extensively split as in the case of papain, discussed by Smith, where 50 out of 180 residues can be broken off by aminopeptidase without any loss in activity against five different substrates. B. G. Malmström's enolase—an enzyme occurring in several electrophoretically distinct forms—is also unaffected after the action of carboxypeptidase or of aminopeptidase, the latter splitting some eighty amino-acids from a particle of molecular weight 68,000. It is not very difficult in reality to envisage that an active centre can survive such changes, although we know too little about tertiary structure to predict whether it will or will not. Conversely, some slight change occurring near or within the active centre may abolish activity completely. In papain, only one —SH group out of six is chemically reactive, and this, if substituted, destroys activity. In ribonuclease, a covalent split four amino-acids from the C-terminal end inactivates, as does any modification of tertiary structure which results in the liberation of H-bonded phenolic groups.

The 'junk' hypothesis, while no doubt valid with respect to some kinds of biological activity, need have no general validity. Antibody formation, for example, is evoked by more than one configurational site of the antigen molecule. R. R. Porter described the isolation of a homogeneous fragment from an enzyme digest of serum albumin, less than one-fifth the size of the native protein. This, when studied immunologically, contains some of the sites involved in the binding of antibody (serum albumin itself being used as antigen), but others have been shed, and the behaviour of the fragment is still dependent upon the intactness of its tertiary structure. J. Reynaud has likewise shown that trypsin-digested tetanus toxin contains at least four products antigenically distinguishable in their combination with antibody.

In the broadest sense, the factors affecting the stability of a protein cannot readily be generalized. C. Tanford pointed out that effects due to charge are not symmetrical about the isoelectric point, and that H-bonding involving the ionic groups must also play a part. He inclines to the view that the 'unmasking' of titratable groups, first observed by the writer in the titration of edestin and studied in detail by Steinhardt in the case of haemoglobin, is due primarily to a change in the apparent pK of ionizable groups, induced by a change of charge density due to unfolding. The expansion of a protein in this way may be an entirely reversible process, and the molecule, if treated sympathetically, may regain its configuration and solubility at the isoelectric point. The rate of return, however, is often slow, and may compete with the tendency of the expanded molecule to become aggregated and precipitated, that is, 'denatured'.

Ultimately, one hopes that the biochemist's approach to structure will be complementary to that of the X-ray crystallographer. J. C. Kendrew and M. F. Perutz discussed the value of the isomorphous replacement method, in which heavy-metal compounds or metal atoms have been introduced into myoglobin and haemoglobin at a few selective sites. The two-dimensional Fourier projections at 6 Å resolution are virtually uninterpretable, but the three-dimensional projection at 6 Å of myoglobin begins to show rivulets of electron-dense material consistent with the gross morphology of polypeptide chains. Mrs. D. C. Hodgkin remarked that even a resolution of about 2 Å would be insufficient to resolve individual atoms, though it would suffice to disclose the general topography of some of the side-chains.

Chemical variation and biological activity. By studying the variations in primary structure of a given type of protein, either within a species or between species, important information can be gathered about the essential or dispensable parts of the polypeptide chain. H. Tuppy described how, in the case of cytochromes (where no microheterogeneity occurs), the peptide sequence of seven amino-acids around the half-cystine residues, which are linked covalently to the porphyrin group, is the same in ox, pig, horse and salmon. In the chick, serine replaces alanine, and in the silkworm, arginine replaces lysine; but in all, the span of two amino-acids between the cysteine residues allows a model to be built on the plan of an α -helix, in which the imidazole residue is able to bond with the metal of the protohaem.

Many abnormal haemoglobins have hitherto been characterized in relation to a specific disease and by physical methods. Those responsible for sickle cell- and Hb-C-anæmia arise by a mutation in the locus for normal adult haemoglobin. V. M. Ingram has compared the three forms of chromatography ('fingerprinting') of the tryptic digests, and finds a difference in only one peptide fragment. The sequence —glu.glu.lys— in adult is replaced by —val.glu.lys— in sickle cell anæmia and by —lys.glu.lys— in haemoglobin-C anæmia. That such a slight change can so profoundly modify the physical properties of haemoglobin is a problem far beyond our understanding at present.

The unique phosphorylation of serine in or near the active centres of certain enzymes by diisopropyl phosphofluoridate allows a comparison of sequence at a specific site in very different proteins. J. A. Cohen described how the sequence —gly.asp.ser.gly.gly— of chymotrypsin is replaced by —gly.glu.ser.ala.gly— in liver alicesterase. The substitution of similar side-chains, that of glu for asp and ala for gly, might imply that these are essential to the configuration of the active centre.

Activation of zymogens. The proteolytic enzymes are often secreted as inactive precursors which are converted to a functional form by a limited proteolysis. The schools of H. Neurath and of P. Desnuelle have greatly clarified the detailed events of these activations; but whether the active centre is created in the process, or simply revealed by unmasking, is still a matter for speculation. In trypsinogen, the liberation of one acidic hexapeptide from the N-terminal end produces the active enzyme. Desnuelle reported that the active split of this bond is accelerated by calcium ions, which at the same time

suppress a more extensive proteolysis. Chymotrypsinogen can give rise to several chymotrypsins: a tryptic split at 1 in the sequence —leu.²ser.arg.¹ileu.val— gives an active form (π), and a second split at 2 by chymotrypsin itself gives the δ -form. Further splitting by chymotrypsin at 3 and 4 in the sequence —tyr.⁴thre.asp.³ala— produces the class-

NH₂

ical α -form of Kunitz. Desnuelle has shown that it is possible to start in reverse and in complete absence of trypsin to cleave by the action of chymotrypsin the links at 2, 3 and 4 above to give modifications ('neochymotrypsinogens') which are still enzymically inactive.

Other activations involve quite extensive proteolysis. Recent work of Neurath's group shows that Anson's crystalline carboxypeptidase is in reality only one-third the molecular weight of the zymogen, the rest being lost as dialysable fragments. Miss G. E. Perlmann reported that, by the action of urea, pepsin itself can lose dialysable components, leaving a portion with increased specific activity. This finding again is relevant to the 'junk' hypothesis.

Virus structure. Advances with respect to the structure of tobacco mosaic virus were discussed by H. Fraenkel-Conrat, G. Schramm and Miss R. E. Franklin. An unexpected discovery of recent years is that the ribonucleic acid alone, in absence of protein, is infective; another, that the sub-units of

molecular weight 18,000 can be reconstituted to form the giant virus particle either in the absence of, or in the presence of, the nucleic acid component. This finding should greatly assist the elucidation of the topographical relation of nucleic acid and protein. Fraenkel-Conrat reported a feature of more general importance in protein chemistry—that the N-terminal residue (serine) of the sub-unit is substituted by an acetyl group. The indirect evidence for cyclic chains in proteins, resting as it does on the apparent absence of C- and N-terminal groups, becomes, by this and other examples (cf. α -melanophore-stimulating hormone below), of less and less significance.

Methods. The new methods which have been used so successfully for the separation of proteins, large and small peptides, and amino-acids were discussed by A. Tiselius, J. Porath, J. Borman, S. Moore, F. Turba, F. Sorm and B. Keil. In addition, L. C. Craig described a new and very simple microdiffusion apparatus which can be used not only for fractionation of dialysable peptides, but is also capable of giving information on the heterogeneity of the solute.

In retrospect, the Paris conference very clearly emphasized the great rewards that have accrued from the application of new and novel methods, and justifies the belief that any discussion of methodology should not be confined to a special symposium, but related over a very wide field to the major advances which have been made in our knowledge of the protein molecule.

KENNETH BAILEY

NEWS and VIEWS

Photograph of the Track of the Artificial Earth Satellite

THE Russian artificial Earth satellite is accompanied in its orbit by the remains of the last stage of the rocket from which it was launched. The rocket is at present travelling ahead of the satellite, at a distance which is increasing with time as the relatively greater resistance encountered by the rocket from the Earth's atmosphere causes it to fall to a lower altitude than the satellite. A photograph of the rocket, which is larger and more easily seen than the satellite itself, was taken at Hobart, Tasmania (42° 54' S., 147° 20' E.), at approximately 0938 G.M.T. on October 7, by Mr. T. McMahon, head of the Photographic Section in the University of Tasmania, working with members of the Department of Physics. Prof. A. L. McAulay, of the University of Tasmania, writes: "It shows the track of the rocket (Fig. 1) travelling approximately south-south-west to north-north-east against a background of stars of the constellation Aquarius".

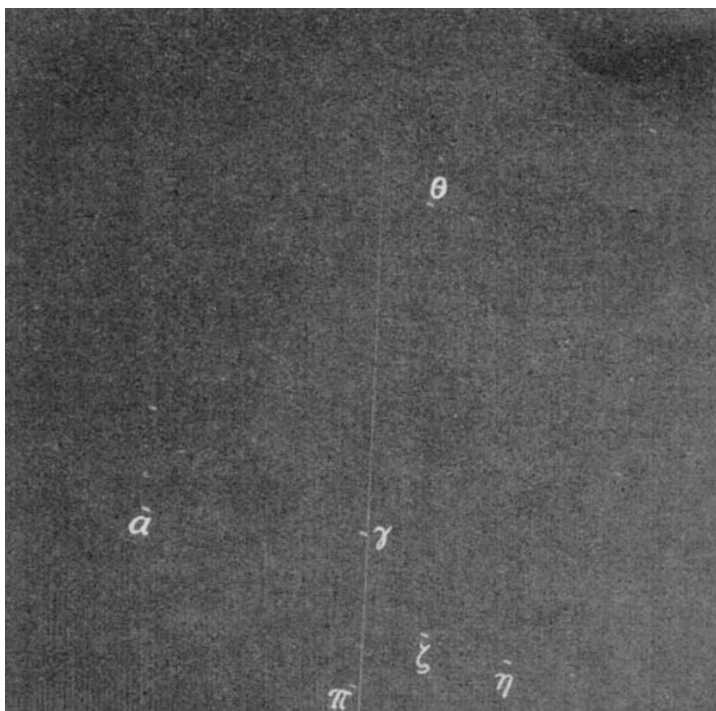


Fig. 1