The Native Protein

Dorothy Wrinch
Department of Physics, Smith College, Northampton, Massachusetts

RESULTS OBTAINED IN PHYSICAL chemistry and chemical crystallography yield a rather definite picture of the native protein which fits at many points with that emerging from the various fields of the biology of the proteins, *i.e.* enzymology, immunology, pharmacology, virus studies, and genetics. If we attempt to sum up the general impression derived from a study of various independent sources of information on the proteins, it can be broken down into the following parts:

(1) Many proteins in solution are particles, not individual molecules (*21*). It is useful to use the term *proteon* to connote a native protein unit incapable of division into subunits also having the native protein character. Every native protein may then be regarded as a proteon or a system of proteons, alike or different. That a protein particle is not a mere conglomerate of proteons but an orderly aggregate, *i.e.* a molecular colony, follows from the finding that in many cases where ultracentrifugal studies show that a protein is uniform in size and shape under well-defined external conditions, the protein in question is capable of dissociation into subunits when the environmental conditions are modified (*21*). Much information has now been assembled concerning the particle status of many proteins, ranging from the very large tobacco mosaic nucleoprotein down to horse hemoglobin. The reversible dissociation of protein particles does not imply any change in the chemical status of the proteons. It is therefore to be sharply distinguished from denaturation.

(2) Interproteon associations within particles depend upon the R-substituents of the amino acid backbones (N—C_α—C) which make up the proteon skeletons. Structure chemistry provides three types of association: direct polar, mediated polar, and less polar. Horse hemoglobin, which on dilution dissociates into two subunits, provides an illustration of the first type; the Helix hemocyanin particle whose dissociations in aqueous solution are affected by Ca^{++} and Mg^{++}, an illustration of the first and second types (*21*). That the less polar R-substituents of aliphatic and aromatic types, as well as the more polar types, emerge from protein surfaces is indicated by many data, e.g. the higher solubility of insulin in propylene glycol than in water and the affinity of human serum albumin for the paraffin chains of fatty acid molecules (*13*). Thus, associations of the third type are also a possibility, and the fact that studies in aqueous solution give two subunits per particle (e.g. for horse hemoglobin) does not preclude the possibility that these subunits may themselves be particles in which the proteons are associated by less polar R-groups. Such a situation is increasingly indicated by crystallographic studies (*6*).

(3) The proteon skeleton is an *intrinsically unstable structure*. This is seen, for example, in the fact that each protein has its own pH stability range. Its stability depends upon the energy provided by the association of favorably placed polar R-substituents which can be direct or mediated by foreign molecules or ions. The minimum solubility of a protein in the neighborhood of its isoelectric point and the general instability of proteins on dilution (*4*) are further indications of this situation. The fact that the proteon, apart from its R-groups, is unstable means that there is, in the skeletons of native proteins, a labile bond. That the protein skeleton is rigid, we know from both crystallography (*6*) and physical chemistry (*10*). Evidently it is this labile bond which is responsible for the rigidity. That the labile bond joins skeletal atoms and not atoms belonging to R-substituents is inherent in the viewpoint which emerges from chemistry as a whole. Just as the sterol skeleton embodies the sterol character, the protein skeleton embodies the protein character, irrespective of the particular amino acid compositions of individual proteins. To reject this viewpoint is to exclude the proteins from the fabric of present-day chemical ideas. The protein molecule may be viewed as a triumph of atomic engineering in which there is a "compression" member comprising the skeletal atoms and a "tension" member comprising the R-substituents with hydrogen bridges and hydrogen bridge circuits as "ties." The relation between the skeleton and the associational complex of its substituents is thus essentially symbiotic. Only with the stabilization energy provided by the associations of its substituents can the skeleton continue to exist; only so long as the skeleton remains intact are the sites or roots of the substituents held in spatial positions which permit the formation of associations. In the associational complex of substituents resides the physicochemical individuality of a proteon; upon it depends its biological specificity.

The Protein Skeleton

The major unsolved problem in the structure of proteins is the nature of the protein skeletons. It is to be expected that these skeletons are entities as definite as, for example, the carbon skeleton of the sterols or the carbon-

Based upon the Stieglitz Memorial Lecture, American Chemical Society, Chicago, January 1947.

nitrogen skeleton of hexamethylene tetramine. The information required, before fundamental progress in understanding protein structure can be made, is a blueprint showing the positions in space of every N, every C_α and every C in the skeleton or skeletons. It is necessary, therefore, to discover how the (N—C_α—C) units are interlocked to form a rigid skeleton and, incidentally, the number or numbers of amino acid residues capable of building such integrated structures.

One important issue, generally the first to be discussed in an analytic study of an unknown structure, is symmetry. The fact that all the (N—C_α—C) units in proteins have the C_β atom attached to the C_α in the same (levo-) position narrows the possibilities regarding the symmetries of protein skeletons, since it eliminates all types involving a center of symmetry or a mirror plane. Twelve possibilities comprise: two cubic, O and T; two hexagonal, D_6 and C_6; two rhombohedral, D_3 and C_3; three tetragonal, D_4, C_4, and S_4; one orthorhombic, D_2; one monoclinic, C_2; and one triclinic, C_1. Each will require investigation in turn. So far only cubic symmetry has been considered (*25*). So many facts regarding proteins fall into place if it is assumed that protein skeletons have cubic symmetry that this assumption and its far-reaching implications merit detailed investigation and study. The hypothesis at once explains how it can happen that a number of proteins, unlike organic molecules in general, form crystals of high symmetry (*6*, *7*). It explains, also, the variety of space groups exhibited by protein crystals (*6*, *7*). The body of knowledge comprised in the atomic structure of minerals (*2*) shows exactly the lines upon which such a situation can arise, provided a cubic skeleton is characteristic of all proteons. Proteons form particles, often of immense molecular weights, which, to the limit of sensitivity of the ultracentrifugal technique, are uniform in shape (*21*). The cubic hypothesis explains how this can happen. Molecules built on cubic skeletons can form particles by the association (as it were) of faces, edges, or vertices, after the manner of the silicates (*2*), thereby building aggregates which are of definite morphology, even though they may be large. There seem to be many indications that only a cubic skeleton can build particles of definite morphology having the variety of shapes exhibited by protein particles. The variety of optical characters (*12*) is also explained. Since the cubic skeleton is necessarily isotropic, the optical character of a protein crystal depends upon the nature and arrangement of its particular R-substituents.

In this connection, the idea presents itself that a local situation on the surface of a protein molecule or within a protein particle may be, to all intents and purposes, a replica of a local situation in a mineral (*26*). Metals ions are known to be associated with many proteins, e.g. manganese with arginase, zinc with carbonic anhydrase, copper with phenolases, iron with catalase, peroxidase, cytochromes, and lactic dehydrogenase. With mineral precedents in mind, we can picture a metallic ion, for which a tetrahedral or an octahedral environment of oxygens is appropriate, clamping together protein molecules, each carrying glutamyl or aspartyl substituents. We can also picture hydrogen bridge circuits, of the types found in ice (*1*) and in various organic crystals, such as pentaerythritol (*5*) and acetamide, (*18*) engaging substituents of protein molecules. Such situations involve symmetry elements also. A cubic skeleton permits, though it does not require, such associational complexes of its substituents. It would also permit the integration of individual hydrogen bridge circuits by water molecules into a multiply-connected network covering the whole molecule. Proton jumps or transfers, of the type already postulated in ice, would then allow the entry and exit of protons at widely separated points. The suggestion may be made that the essence of a protein capable of enzymatic activity may reside in such a network over its surface. Long-range forces for proteins, which have been postulated for some time, provide, it would seem, the only explanation of biological phenomena such as those concerned with the pairing of chromosomes and the mitotic cycle in general (*22*). They have recently been postulated, even more definitively, on the basis of revolutionary results concerning protein interactions at distances of the order of 100 A. (*14*, *15*). That an enveloping network of multiply-connected hydrogen bridge circuits may prove to be the seat of such forces is worthy of the most careful consideration.

A Theory of Biosynthesis

Whether our concern is with the possibility of long-range forces or with the nature of enzymes, protein function may be regarded as the outward sign of an inner essence. It is therefore to be expected that the most fundamental of all protein functions, the production of replicas, may throw light upon the nature of protein structure. (It is of interest in this connection that the function of replication is gradually being extended to more and more proteins, *20*.) A theory of biosynthesis is therefore presented for consideration. It is based upon the belief that replication is akin to crystal growth and so belongs essentially to the domain of crystallography, the richest repository of ideas for structure studies in general and indispensable in any attempt to understand the proteins.

The replication of proteins falls into two parts: the replication of protein molecules and the assembling of protein molecules into protein particles. The former is the crucial happening in protein synthesis.

For molecular replication, three basic assumptions are required:

(1) We assume that the inherent capacity of certain species of protein to produce replicas depends on the presence on their surfaces of *active patches* (to use Warburg's term), *i.e.* local specific constellations of R-sub-

stituents. Each of these, we must presume, can function as a mold or templet, permitting the laying down on itself of a complementary constellation. Under what circumstances, then, will such constellations provide appropriate material for new, hitherto-non-existing proteons which are exact replicas of those already existing? Evidently, in the first place, this occurs only if a proteon is in some sense, a surface structure—a stiff fabric of some kind forming a 'space-enclosing' shell. This requirement accords well with the fact that all proteins, no matter how large they may be, produce monolayers of the same thickness, 7–11 A. Secondly, it is necessary that the constellations complementary to those comprising an existing proteon be equivalent to the constellations themselves. Thus, we might take a molecule of the same shape as, for instance, hexamethylene tetramine, but larger and analyze its surface into various patches. Suppose now that each patch of the surface is complementary to the antipodal patch. This would mean that the patch synthesized on, and complementary to, one part would actually be the antipodal patch. In such a case, then, the synthesis by individual molecules of complementary patches would *ipso facto* produce the patches capable, when correctly assembled, of making new molecules which are exact replicas. Thus, the first deduction to be made regarding protein molecules is that they are surface structures, comprising a set of individual patches which form a self-complementary set.

(2) In order to have a mechanism whereby these isolated constellations on different molecules may, on occasion, be integrated so as to interlock, in the same spatial pattern as in the original molecules, something has to be postulated regarding the capacity of the original molecules to form orderly assemblies. In such a case, we can picture the incipient crystallization of the existing molecules permitting complementary patches to interlock, step by step, in a variety of sequences, into precisely the pattern in which the patches are integrated in the existing molecules. We see then that the global self-complementariness of the surface of a protein molecule belonging to a species capable of replication can be more precisely defined. It must be "crystallographically satisfactory"; the incipient crystallization of existing molecules then so sets the individual complementary patches that interlocking into new and identical molecules results. In the example cited, the antipodal patches form pairs of complementary patches. In the case of the phosphotungstic acid anion, patches of the surface are self-complementary. There are, in fact, a well-defined number of ways in which a crystallographically satisfactory global self-complementariness can be achieved. One of these ways must be characteristic of the synthesizing proteins; which, we do not yet know. It is worth emphasizing that whichever alternative is adopted by proteins, one apparent difficulty disappears. It has been stated that there is some objection to postulating the formation of a complementary pattern, since a protein produces not a complementary pattern but a replica of the original (*3*). The difficulty evanesces when it is seen that the molecule results from the integration of individual complementary patches, which may be, but need not be, replicas of the patches on which they are synthesized, and that it is only the appropriate *global self-complementariness* of the molecules which ensures replication.

(3) The energy questions involved in protein synthesis point in the direction of certain ancillary molecules being present during such operations, e.g. molecules containing energy-rich phosphate bonds. It should be noted in this connection that among the many crystalline *nucleo* proteins already well known, bushy stunt virus nucleoprotein achieves even a *cubic* space group (*7*). That nucleic acid plays an essential role in replication has long been assumed (*20, 24*).

Protein Intermediates

This theory of biosynthesis has a number of implications, some of them far reaching. It is sufficient here to refer to one of them—the existence of protein intermediates. These may be visualized as the various individual active patches upon which complementary constellations are formed. There are indications that such structures may already have been found in no less than a dozen cases, e.g. in the lactalbumin precursor (*19*), in the digestion products of beef pseudoglobulin (*11*), β-lactoglobulin (*8*), and casein (*23*), and in hydrolytic products of silk (*9*). These products are all of low molecular weight corresponding to amino acid residue numbers in the range 4–12, clearly indicating that the smallest—perhaps the only—proteon skeleton may comprise relatively few residues. It is to be emphasized that the products in the last case have biological activity, connoting a definite structure. Biological activity has been found in protein monolayers in a variety of cases which include ovalbumin, horse serum albumin and globulin (*17*), and insulin and metakentrin (*16*). In the case of the globulin, it proved necessary to compress the film to 15 dynes or so (*17*). A structure capable of biological activity only under such circumstances may well contain a labile bond. All these facts point to one conclusion, which is exactly that reached in the discussion of biosynthesis. The proteon, it would seem, is an integration of parts of definite structure and individuality. The immense difficulty of the study of the *structure* of proteins (in contrast to *composition* of proteins) by way of organic chemistry needs no emphasis. But a new approach opens when we think in terms of protein intermediates and aim at a more limited objective. The application of the general ideas which presented themselves in the consideration of the symmetry of the protein skeletons to the possible structure of such protein intermediates yields various specific suggestions regarding the particular amino acid residues which may be expected in conjunction (*26*). A systematic study of the

amino acid compositions to be expected of active patches can now be undertaken. The complete characterization of protein intermediates and their resynthesis would mark a crucial stage in protein studies.

References

1. Bernal, J. D., and Fowler, R. H. *J. chem. Phys.*, 1933, **1**, 515.
2. Bragg, W. L. *Atomic pattern of minerals*. Ithaca, N. Y.: Cornell Univ. Press, 1937.
3. Burnet, F. F. *The production of antibodies*. (Univ. Melbourne Monogr., 1941.)
4. Cohn, E. J., and Edsall, J. T. *Amer. chem. Soc. Monogr.*, 1943, p. 571.
5. Cox, E. G., Llewellyn, F. J., and Goodwin, T. H. *J. chem. Soc.*, 1937, 882.
6. Crowfoot, D. *Chem. Rev.*, 1941, **28**, 215.
7. Fankuchen, I. *Proc. N. Y. Acad. Sci.*, 1941, **41**, 157.
8. Haugaard, G., and Roberts, R. M. *J. Amer. chem. Soc.*, 1942, **64**, 2664.
9. Landsteiner, K. *J. exp. Med.*, 1942, **75**, 269.
10. Oncley, J. L. In Cohn, E. J., and Edsall, J. T. *Amer. chem. Soc. Monogr*, 1943, p. 568.
11. Petermann, M. *J. phys Chem.*, 1942, **46**, 183.
12. Reichert, E. T., and Brown, A. *The crystallography of hemoglobins*. Washington, D.C.: Carnegie Institution of Washington, 1909.
13. Rice, R. G., Ballou, G. A., Boyer, P. D., Luck, J. M., and Lum, F. G. *J. biol. Chem.*, 1945, 158, 608.
14. Rothen, A. *Science*, 1945, **102**, 446.
15. Rothen, A. *J. biol. Chem.*, 1946, **163**, 345; 1947, **167**, 299; 1947, **168**, 12.
16. Rothen, A., Chow, B. F., Greep, R. O., and Van Dyke, H. B. *Cold Spr. Harb. Sympos.*, 1941, **9**, 272.
17. Rothen, A., and Landsteiner, K. *Science*, 1939, **90**, 65.
18. Senti, A., and Harker, D. *J. Amer. chem. Soc.*, 1940, **62**, 2008.
19. Sjögren, B., and Svedberg, T. *J. Amer. chem. Soc.*, 1930, **52**, 3650.
20. Spiegelman, S., and Kamen, M. D. *Science*, 1946, **104**, 581.
21. Svedberg, T., and Pederson, K. O. *The ultracentrifuge*. London: Oxford Univ. Press, 1940.
22. Szent-Györgyi, A. *Science*, 1941, **93**, 609; *Nature, Lond.*, 1941, **148**, 157; Muller, H. *Cold Spr. Harb. Sympos.*, 1941, **9**, 290.
23. Winnick, T. *J. biol. Chem.*, 1944, **152**, 465.
24. Wrinch, D. *Protoplasma*, 1936, **25**, 550.
25. Wrinch, D. *Biol. Bull*, 1945, **89**, 192; *Amer. Mineral.*, 1946, **31**, 513.
26. Wrinch, D. *Austr. J Sci.*, 1946, **8**, 103.

Tidal Illumination Diagrams

H. R. Seiwell

Woods Hole Oceanographic Institution[1]

THE TIDAL ILLUMINATION DIAGRAM IS one of the various graphic devices designed during the recent war to represent significant hydrographic conditions on military targets. It was introduced by the Oceanographic Section of Headquarters, Army Air Forces, in 1942 to represent relationships between stage of the tide and state of illumination in target areas of Northwest Africa. Later, its use was extended to all theaters, and preparation on a mass scale was taken over by the U. S. Coast and Geodetic Survey, where facilities are available for computation and graphic presentation of these data for many parts of the world. Recent comments have shown the diagrams to be of practical use in this postwar period.

The tidal illumination diagram is so arranged that for any hour of the day combinations of solar illumination and stage of the tide are readily noted. Heights of high and low tides are referred to the sound datum of nautical charts of the specific locations. In so far as possible, tidal data from the U. S. Coast and Geodetic Survey tide tables are used. The occurrences of high and low tides are shown as a series of connected points enabling the user to visualize the changing sequence throughout the month. Days of the month are represented by vertical lines covering the period from noon through midnight to noon of the next day.

Three types of twilight are indicated to bring out gradation in light from sundown to darkness and from darkness to sunrise. Civil twilight begins with sunset and ends when the sun is 6° below the horizon. During this period, objects may be distinguished readily, and a newspaper may be read without the aid of artificial light. At its end,

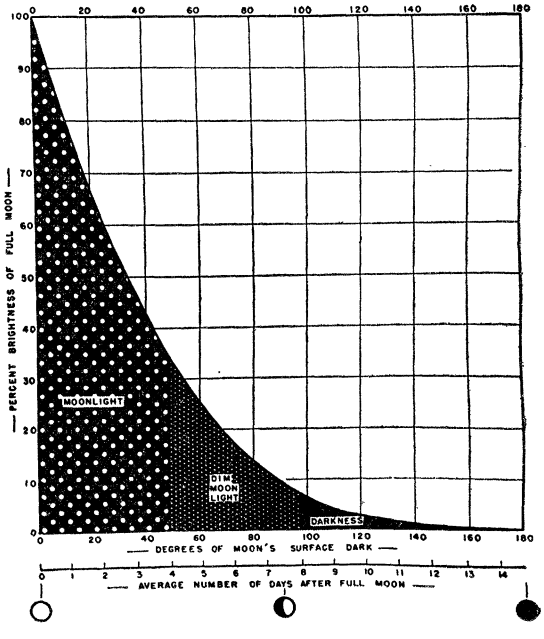

FIG. 1. Relative brightness of moon with change of phase.

sky brightness is approximately 20 times that of the full moon at zenith (Fig. 1). In the following period of nautical twilight, the brighter stars are visible, the horizon is

[1] Contribution No. 382.