IV. PROTEIN INTERACTIONS

AGGREGATION OF GLOBULAR PROTEINS

BY LINUS PAULING

The Gates and Crellin Laboratories of Chemistry *, California Institute of Technology, Pasadena 4, California

Received 24th October, 1952

Many globular proteins are known to polymerize. It is pointed out that a monomer protein molecule containing a single polypeptide chain is asymmetric, and that a single pair of complementary combining regions on such a molecule would permit it to attach itself to equivalent molecules in such a way as to form a helical structure, giving rise to fibrils. If the pitch of the helix is zero or nearly zero the helix degenerates into a circle, and a small polymer (dimer, trimer) is formed. A reasonable amount (about 5°) of freedom of angular motion in the bond between adjacent molecules would permit a helical fibril of large diameter to change its length by as much as a factor of two, and would also permit tubular aggregates with walls one, two, or three molecules thick to be formed. It is suggested that some of the observed properties of globular proteins can be accounted for in this way.

Protein molecules are sticky, and ordinarily clump together into aggregates. Through the assumption of a certain configuration the molecules of a protein may have a surface of sufficiently diminished stickiness to permit a moderately concentrated solution of the protein to be stable; the special configuration is that of the native soluble globular protein. The inherent stickiness of the protein becomes evident, however, on denaturation; a denaturing agent or condition can change the configuration of the protein molecule from the native configuration, and the molecules can then attach themselves to one another, forming an essentially insoluble coagulum of denatured protein.¹

Aggregation of globular protein molecules would be expected to occur if a force of interaction could operate between two molecules with energy corresponding to a bond of about 20 kcal/mole or more. A single covalent bond formed between two protein molecules would hold them together; examples of bonds of this sort that form under special circumstances are the sulphur-sulphur bond between half-cystine groups, and the sulphur-mercury-sulphur bond in the complex formed by a mercuric ion and two serum albumin molecules. Usually, however, aggregates of globular protein molecules are formed through the interaction of a combining region on the surface of one molecule and a complementary combining region on the surface of a second molecule. The interaction of the two molecules is the sum of a number of small interactions—van der Waals' forces, hydrogen bonds, and the attraction of positively charged and negatively charged groups.²

It has been observed that association or dissociation of protein molecules may result from rather small changes in environment or structure. Haemocyanin may be prepared in solution with molecular weights (particle weights) ranging from 250,000 to 5,000,000 or more. Horse haemoglobin has molecular weight 68,000 in salt solution, and 34,000 in solution containing urea. A change in

* Contribution No. 1747.

pH causes insulin to be transformed from small molecules, with molecular weight 12,000 or 36,000 (perhaps 48,000), to very large fibrils. Sick-cell-anemia oxyhaemoglobin is stable in solution; removal of the oxygen molecules permits aggregation to occur, with the formation, in solutions containing 10 % or more of the protein, of nematic liquid crystals,³ or of true crystals of the haemoglobin. The action of thrombin on fibrinogen, which is stable in solution, causes its conversion into fibrin (perhaps by splitting off a short peptide group), and the fibrin molecules undergo aggregation to form a clot. The enzymatic splitting-off of a short peptide chain from ovalbumin converts it into plakalbumin, with different solubility and different crystal form.

A protein molecule may involve one or more polypeptide chains, coiled about in a reasonably precise way. If a molecule can be separated into two or more identical parts, each consisting of one or more polypeptide chains folded in a definite way, we describe the original molecule as a polymer containing the smaller molecules. If, on the other hand, the large molecule can be separated into its parts only with the destruction of the configuration of the individual polypeptide chains (that is, if the several chains that constitute the original molecule are so coiled together as to require their essentially complete uncoiling in order to disentangle them from one another) we describe the process as involving the destruction of the original molecule, and do not call the complex a polymer.

Let us consider a protein whose molecules consist of single polypeptide chains, with an amino terminal group and a carboxylic terminal group. Such a molecule is necessarily asymmetric. We consider a solution containing asymmetric protein molecules, subjected slowly to conditions that change their composition or structure. These conditions might involve the removal of a side chain, oxidation or reduction of a side chain, a small change in configuration, then larger and larger changes in composition and configuration, and ultimately complete unfolding of the polypeptide chain—complete denaturation. At some stage in this process molecules which originally had no power of combining with one another would necessarily develop complementary regions, A and A^{-1} , with sufficient energy of interaction to correspond to a bond. We assume that the changes occur slowly enough for each molecule to pass through the same sequence of structure and composition; accordingly at this stage each molecule would have a group A and a group A^{-1} , and would be capable of forming two bonds. Because of the assumed lack of symmetry of the molecules, the groups A and A^{-1} would not be exactly on opposite sides, but would be in general positions in the molecule (fig. 1). The operation of conversion of a molecule into an equivalent molecule bonded to it by the AA^{-1} bond would be the general operation for an asymmetric object, the rotatory-translation. This corresponds to a screw axis, and the operation if repeated leads to a helical aggregate (fig. 2).

The precision of definition and the rigidity of the helical aggregate are determined by the properties of the bond between the molecules. If a covalent bond is formed between side chains the link between the molecules may be very flexible, and the configuration of the aggregate will be correspondingly poorly defined. If, however, the bond is due to the co-operation of weak forces operating over a moderately large region A and complementary region A^{-1} the bond will be rather rigid. The study of interactions between antigens and antibodies² has indicated that the combining regions may be about 10 Å in diameter, and that the amount of leeway in juxtaposition of atoms may amount to about 0.5 Å. This amount of leeway corresponds to a flexibility of the bond permitting bending by 5° in any direction from the average bond direction.

If the pitch of the helix is less than the width of the molecule in the direction of the axis of the helix, steric hindrance will prevent the formation of a fibril (fig. 3). The process of aggregation resulting from the complementary groups A and A^{-1} would then result in the formation of finite polymers. Such a polymer may be asymmetric, with the number of bonds one less than the number of molecules,

as shown in fig. 4. However, if the pitch of the helix is nearly zero and the rotational operation is nearly equal to 360° divided by an integer, the flexibility of the bonds might permit the formation of a ring, involving one more bond, the number of bonds then being equal to the number of molecules. This situation is illustrated in fig. 5. In general, it will involve some strain in each of the bonds; if the strain energy for the bonds is less than the energy of the additional bond that is formed in the ring, the ring polymer will be stable relative to the incomplete

FIG. 1.—An asymmetric protein molecule with two complementary combining groups, A and A^{-1} . The arrows indicate roughly the normals to the approximate planes of the combining groups.

FIG. 2.—A helix formed by aggregation of asymmetric protein molecules with a pair of combining groups.

ring. (Consideration must also be taken of the difference in entropy; the incomplete ring, with greater freedom of motion, will have larger entropy than the ring polymer.) It is not unlikely that the molecules of insulin with molecular weight 36,000 that are present in rhombohedral crystals of insulin have the structure illustrated diagrammatically in fig. 5.

FIG. 3.—Steric interference with completion of a helix, when the pitch of the helix is less than the diameter of the molecule in the direction of the axis of the helix.

FIG. 4.—An aggregate of three protein molecules containing two bonds.

The ring polymers that are formed in the way just described have an n -fold axis of symmetry, with n equal to the number of molecules in the polymer. If the pitch of the helix is approximately equal to the thickness of the molecule in the direction of the axis of the helix, additional stabilization of the infinite polymer will be introduced through the interaction of each molecule with the molecules in the adjacent turns of the helix, as illustrated in fig. 6. The flexibility of the bonds might permit this compact tubular structure to be assumed even when the pitch for unstrained bonds is only approximately equal to the molecular diameter. The sheath surrounding the axial filament in the sperm tail may be a tube of this sort.

In general the helical polymer would not involve a rational number of molecules per turn, and would accordingly not be well suited to the formation of crystals. These helical molecules might, however, easily form liquid crystals of the nematic type; moreover, the interaction between a helical polymer molecule and adjacent molecules might deform the helixes slightly, into configurations with a rational number of molecules per turn, resulting in the formation of crystals with well-defined configuration.

FIG. 5.—An aggregate of three molecules containing three bonds.

FIG. 6.—A helical polymer with pitch approximately equal to the thickness of the molecule in the direction of the axis of the helix, introducing additional stabilization through the van der Waals interaction of the molecules in adjacent turns.

If the pitch of the helix is approximately equal to an integral multiple of the diameter of a molecule in the direction of the axis of the helix a tubular fibril might be formed involving two or more interpenetrating helixes, as illustrated in fig. 7. Moreover, if the diameter of the helix is moderately large, in comparison with the molecular diameter, the flexibility of the bonds AA' should permit helixes

FIG. 7.—A cylindrical polymer formed from two identical helixes.

to be formed with diameters differing by one molecular diameter; tubular molecules could then be built, from identical monomers, with walls two or three monomers thick. A tubular molecule of this sort is illustrated in fig. 8.

A tubular polymer with diameter much larger than the diameter of the individual molecules could vary its pitch over a large range without bending the bonds through more than $\pm 5^\circ$. For example, in fig. 9 there is illustrated a helical aggregate of approximately spherical molecules, with $8\frac{1}{2}$ molecules per turn of the helix. The angle of declination of the helix from the horizontal is 7° . The attraction between molecules of one turn and those of adjacent turns of the helix might stabilize the configuration in which adjacent turns are in contact even though the bond angles AA^{-1} were strained by 5° . The configuration corresponding to

unstrained bonds, with angle of declination 12° , is illustrated in the right part of the figure. This corresponds to a pitch 70 % greater than that for the more compact configuration. Assuming that the forces of interaction between molecules in adjacent turns of the helix are short in range, we see that the helical molecule would have two stable (metastable) configurations. The configuration of large pitch would be stabilized by the unstrained bonds AA^{-1} . On decreasing the pitch the potential energy of the molecule would increase, because of the strain of the bonds AA^{-1} , until the force of attraction between molecules in adjacent turns began to operate, when there would occur a decrease in potential energy.

FIG. 8.—A tubular polymer with walls two or three molecules thick, formed from helical aggregates with diameters differing by the diameter of the monomer.

The potential energy as a function of the pitch of the helix is shown in fig. 10, as calculated for Hooke's law energy of distortion of the intermolecular bonds, and inverse sixth power energy of van der Waals' attraction and inverse twelfth power energy of van der Waals' repulsion for the forces between molecules in adjacent turns.

The relative positions of the potential minima for the extended and contracted configurations could be varied, of course, by changing the distribution of charge

FIG. 9.—A helical polymer with variation in pitch of the helix.

on the molecules, as was first suggested by Meyer in his discussion of molecular contraction.⁴ Thus the contracted configuration might be stable under conditions such that the molecules had small electric charge. Increase in the magnitude of the electric charge of the molecules would result in electrostatic repulsion that would stabilize the extended configuration. The process of conversion from the contracted to the extended form would, of course, be a gradual one, beginning at one end of the molecule and advancing toward the other end. No significant activation energy would be involved, such as might be suggested by the curve in fig. 10, which represents the energy of the molecule as a function of configuration when the pitch is held constant throughout the molecule.

It can be predicted that compound helical structures would be formed under certain circumstances. For example, a helix of the sort shown in fig. 9, with variable pitch, could be combined with a similar helix in the way shown in fig. 11.

Here the two helices differ in diameter by an amount somewhat less than the diameter of the monomer molecule. In a segment of short length of the helical aggregate the molecules of each polymer are in contact with those in adjacent turns of the same polymer and also with those in two turns of the other polymer, the distance between adjacent turns on this side of the large tube being equal to the molecular diameter in the direction of the axis of the tube. On the other side of the tube the distance between turns is 70 % greater than the diameter of

FIG. 10.—A potential energy function composed of a Hooke's law term representing distortion of the helix from the value of the pitch for maximum stability (approximately equal to the indicated minimum for separated turns) and inverse sixth power energy of van der Waals' attraction and inverse twelfth power energy of van der Waals' repulsion between molecules in adjacent turns.

the molecule in the direction of the axis of the tube, and each molecule is in contact with molecules in two turns of the other polymer. The helices are accordingly deformed in such a way as to correspond to a twist of the tube. The curve formed by the centre of the tube is itself a helix. The azimuthal directions in which the

FIG. 11.—A compound helical structure formed by two helices differing in diameter by an amount somewhat less than the diameter of the monomer molecule; the axis of the two helical molecules itself describes a helical curve.

contracted and expanded configurations are assumed might be determined by the packing between the molecules of the two helices—because of the difference in diameters of the two helices, there would be a change in relationship that could be described as a change in phase of one helix relative to the other.

It is interesting to speculate that a possible mechanism of the motion of bacterial flagella, sperm tails, and similar biological structures may be formulated in terms of compound helical structures of this nature. The compound helix of fig. 11 can be described as a cylindrical structure bent into helical form. If a structure of this sort were to be attached at one end to a cell which was able to

vary the position of the molecules in the last turn of one polymer relative to the molecules of the last turn of the other polymer by a distance equal to half the molecular diameter the direction toward the contracted side of the flagellum at its base would rotate about the structure. This effect would be transmitted along the length of the structure, through the motion of one polymer relative to the other, and the entire helical cylinder would change its configuration in the way corresponding to the transmission of spiral waves from the attached end to the free end of the cylindrical structure. This is just the sort of motion that is observed for some biological structures, in particular bacterial flagella and sperm tails. The scale of dimensions is satisfactory. For example, monomer molecules 40 Å in diameter would produce, when arranged as shown in fig. 11, compound helical tubes about 150 Å in diameter.

More complex structures might also be constructed; for example, three compound helical tubes of the sort shown in fig. 11 might be twisted around one another to give a three-strand flagellum, resembling those that have been observed in electron micrographs.

The foregoing considerations represent the application of a fundamental principle of great significance, as was pointed out some years ago by Crane¹—the principle that a helical structure results from the continued application of a general identity operation to an asymmetric element, and accordingly that the helix is to be considered as the simplest infinite aggregate of units.

¹ Wu, *Chinese J. Physiol.*, 1931, **5**, 321. Mirsky and Pauling, *Proc. Nat. Acad. Sci.*, 1936, **22**, 439.

² Pauling, *J. Amer. Chem. Soc.*, 1940, **62**, 2643. Pauling, Campbell and Pressman, *Physiol. Rev.*, 1943, **23**, 203.

³ Harris, *Proc. Soc. Expt. Biol. Med.*, 1950, **75**, 197.

⁴ Meyer, *Biochem. Z.*, 1929, **214**, 253.

⁵ Crane, *Sci. Monthly*, 1950, **70**, 376.