SYMPOSIUM: INTERDISCIPLINARY ASPECTS OF LIPID METABOLISM

conducted by The American Oil Chemists' Society

at its 57th Annual Meeting, Los Angeles, California

April 24-27, 1966

GEORGE ROUSER, Program Chairman

The Use of Molecular Models in Evaluating Protein and Peptide Conformations

DONALD T. WARNER, The Upjohn Company, Kalamazoo, Michigan[1]

Abstract

A new theoretical conformation of polypeptides and proteins designated as the "hexagonal conformation," which was developed through the use of molecular models, is discussed. The term "hexagonal" applies to the arrangement of the peptide chain O-atoms since the placement of the polymer backbone used in this conformation results in a planar hexagonal assignment of these atoms. For a cyclic hexapeptide, one hexagon of O-atoms results; for a cyclic decapeptide, two hexagons result; and a protein subunit (e.g., cytochrome C) can form an extensive honeycomb network of hexagons. The coincidence of this network with a similar "second neighbor" oxygen network in water is discussed in relation to the importance of water in biological processes. Several models of complete peptide sequences assembled in the hexagonal conformation are evaluated in detail. It is suggested that in some instances the "reactivity" of these peptides may reside in an ability to disturb resonance patterns of the surrounding water in a precise manner at specific locations. Finally, the foreseeable problems involved in a uniform application of the "hexagonal concept" are discussed along with the extension of the hexagonal conformation of proteins to cell membrane structures with their lipid bilayers.

Introduction

THE USE OF MOLECULAR MODELS as a tool in scientific research has been much neglected. Molecular models are often considered to be too expensive for the little definitive information that can be gained from their use. However, in the science of "molecular biology" the use of models is a growing necessity as more and more complex structures are elucidated. Given these complex structures, it is now virtually impossible to develop an understanding of their role in biological reactions without a firmer appreciation of the possible molecular organization of these components at the atomic level. Generally, this appreciation can best be gained through the use of molecular models assembled from carefully designed atom units.

Our interest in the use of molecular models as a scientific tool began with a program for the synthesis of peptides initiated several years ago at The Upjohn Company. Prior to attacking the synthetic aspects of this peptide program, a study was made of some known peptide sequences with molecular models to see if any common structural features were present in these highly reactive biological entities. Several cyclic hexapeptides and the slightly more complex cyclic decapeptides were selected for detailed study. Cyclic peptide structures pose many difficulties when one

[1] These studies were carried out in part while the author was an Institute Fellow at the City of Hope Medical Center, Duarte, California, during a sabbatical leave from The Upjohn Company.

INDEX

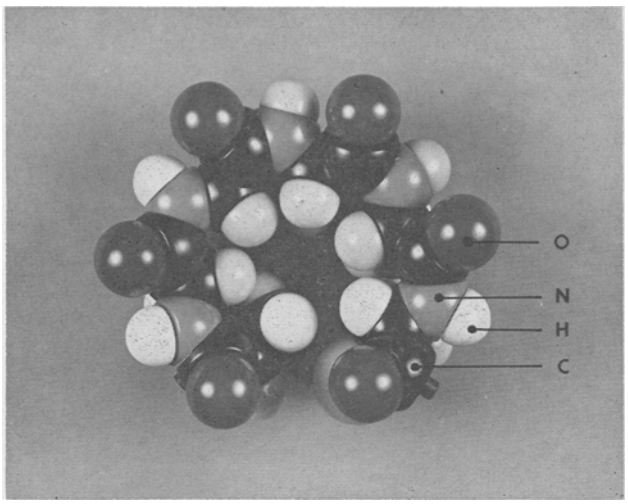


FIG. 1. Hexapeptide in ring structure stabilized by cystine. C,H,N,O indicate the various atoms in the ring. (Included with permission of *Nature* magazine.)

attempts to apply either the α-helix (1) or the pleated sheet conformations (2) to the limiting confines of their known amino acid sequences. In attempting to assemble models of these small cyclic polypeptides, a new theoretical concept of protein and peptide structure or conformation has emerged. I would like to recall briefly the steps leading up to this new concept, and then apply it to a wide variety of peptide and protein molecules.

Types of Models

Our first studies of the peptide sequences were made with space-filling molecular models. In these model sets individual atoms are designed to show the space that is occupied by that atom as measured by its van der Waals radius or some appropriate fraction of that radius. The choice of proper dimensions is still a point of difference among the several manufacturers of space-filling models. Some use the full van der Waals radii for the various atoms in general agreement with the values of Pauling (Ref. 2, p. 257 ff.) while others, particularly with regard to H-atoms, use the somewhat smaller "interference" radius (3). A new set of models (C-P-K Atomic Models, Ealing Corporation, Cambridge, Mass.) is now available and represents the latest consensus about the ideal dimensions and characteristics of space-filling atoms. On the matter of the van der Waals radius of hydrogen, this set uses the minimum value suggested by Pauling (1.0 A) and allows for some further compression by making the hydrogen shell of soft plastic. In addition to the space-filling models, we have made some use of the stereo models which show bond angles and other features such as planarity of benzene rings and boat and chair forms of cyclohexane. These models usually consist of some sort of rod and tube arrangement. For example, in a research model available under the name "Dreiding stereomodels" (W. Buchi Company, Flawil, Switzerland), the C-atoms are made by welding together a rod and tube so that the four stems (two hollow, two solid) form correct tetrahedral angles and each stem projects toward the corner of a tetrahedron. The rod-tube welding junction designates the position of the carbon nucleus. When two atoms are joined together by inserting the rod of one into the tube of the other, a simple indentation coupling holds the centers of the two nuclei the correct distance apart.

Early Models Studied and The Protein-Water Concept

Disulfide Stabilized Structures

The disulfide stabilized ring structures found in oxytocin or vasopressin were studied first (4). These rings are stabilized by a covalent disulfide bond joining Cys-1 and Cys-6 to yield a cyclic structure containing six peptide residues. For simplicity of illustration this first model was made with four alanine residues and two cysteine units. It is virtually impossible to apply the conformational concepts of the undistorted α-helix or the pleated sheet to this ring system. One of the possible conformations for this ring is achieved by placing all the amide linkages in one plane so that the peptide carbonyl O-atoms occupy the corners of a regular hexagon (Fig. 1). In this arrangement all of the atoms of the continuous peptide backbone lie exposed on one surface, rather arbitrarily designated as the "hydrophilic" surface. The peptide O-atoms are thus all coplanar and project slightly *above* the backbone plane while the bulk of the side chains project *below* the backbone (when viewed edge-on from the same point of reference) and constitute another surface having considerable "hydrophobic" character. This term is used somewhat loosely since certain side chains in a complex protein may also be polar or partially "hydrophilic" in character, but for simplicity the terms "hydrophilic" and "hydrophobic" will be used in discussing these structures to designate the peptide bond face and side chain face, respectively, of the various models. A more detailed consideration of this is given in the earlier article (4).

Cyclic Decapeptides

Advancing from the cyclic ring with one hexagon of O-atoms, the cyclic decapeptides represent an additional step in complexity. If the hexagonal pattern of peptide O-atoms is a significant common structural feature, we are immediately confronted with the problem of maintaining this structural feature in this series. One possible answer was found in a consideration of the benzene-naphthalene analogy. In benzene we have six C-atoms yielding one ring and in napthalene ten C-atoms yielding two fused rings. It seemed useful to apply the same sort of approach to the peptide problem, with six peptide O-atoms yielding one complete hexagon and ten peptide O-atoms yielding two "fused" hexagons having two central O-atoms in com-

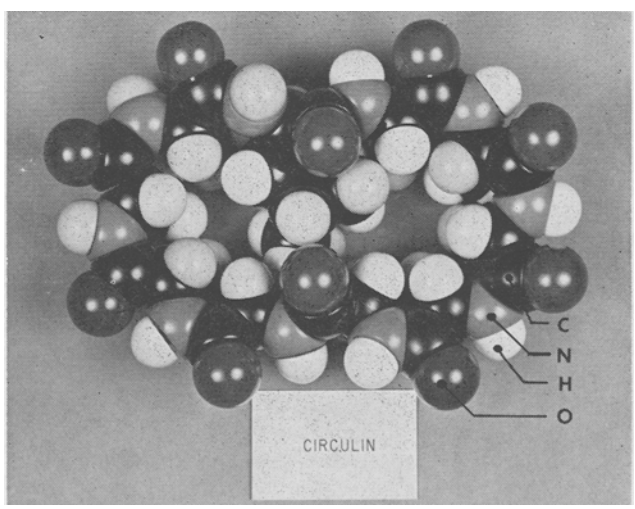


FIG. 2. Cyclic decapeptide ring. (Included with permission of *Nature* magazine.)

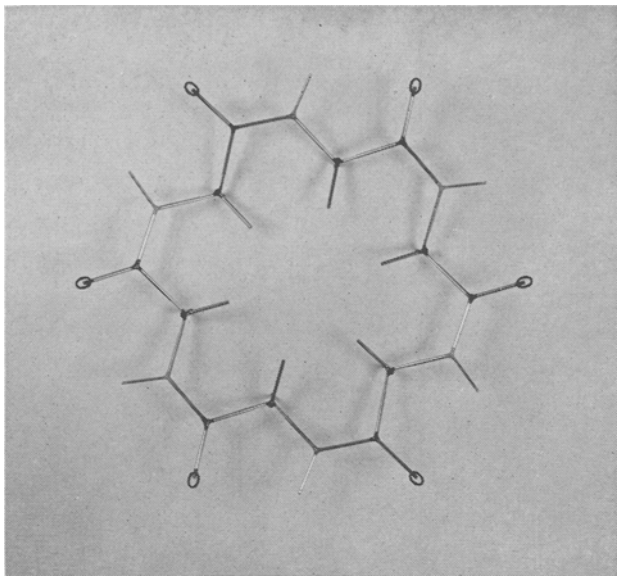


FIG. 3. Cyclic hexapeptide. 0 indicates positions of peptide oxygens.

mon. The peptide backbone of a simplified cyclic decapeptide arranged in this way is illustrated in Fig. 2. The fused hexagonal outlines are readily apparent here and again all the O-atoms are coplanar.

"Second Neighbor" O-atoms of the Ice Lattice

While examining a molecular model of the ice lattice made with stereomodels, it was observed that groups of "second neighbor" O-atoms in the ice lattice are also in a hexagonal array and, like the hexagonal peptide O-atoms, their layers are also rigorously coplanar (5). The possible significance of this similarity in general shape and planarity was further enhanced by the surprising similarity of the dimensions of this "second neighbor" oxygen pattern in ice and the hexagonal peptide oxygen pattern. Although the "second neighbor" distance in ice would be about 4.51 A, this value may be about 4.74 A in liquid water at 25C (6). This agrees reasonably well with the value of about 4.8 A calculated for water at 37C based on a "first neighbor" oxygen distance in water of 2.94 A and the tetrahedral angle between bonds (4,7). Vandenheuvel (8) by using an H–O–H angle of 104° 45′ and a "first neighbor" distance of 2.77 A calculates a "second neighbor" distance in water of 4.63 A. The important point is that any one of these "second neighbor" distances is reasonably close to the distance between adjacent peptide O-atoms in a peptide chain, since the latter value is about 4.8 A. Therefore, the planar hexagonal oxygen pattern of our proposed conformations could make precise collinear hydrogen bonds with a similar hexagonal planar array of "second neighbor" water O-atoms lying adjacent to it. The closeness of this correspondence is shown for a cyclic hexapeptide and the same hexapeptide superimposed on a segment of the water lattice, in Fig. 3 and 4, respectively. Both units were made with Dreiding stereomodels calibrated to the same scale for the two systems. The two oxygen patterns are exactly superimposed on each other as the photograph (Fig. 4) clearly shows. It is suggested that such a lamination of matching layers could impart additional stability to the peptide layer and also serve to orient the water in an ice-like lattice around the macromolecule. In addition, the overlying water layer would provide an excellent proton transfer pathway available to the entire amino acid sequence for proton interaction. Furthermore, it is well to be cognizant of the "two-sided" character of a water layer in the ice lattice structure. An examination of a model of the ice lattice will show that each "second neighbor" layer has all of its available bonding positions pointing in one direction so that in our proposed conformation all the bonds would be available for interaction with a peptide oxygen network. However, this "second neighbor" layer also lies in a "first neighbor" relationship with an adjacent "second neighbor" layer on a different plane which has *all its available bonding positions pointing in the opposite direction.* These bonding positions could simultaneously be contacting another protein layer, so that in essence the ice-like lattice, having this unique bifunctional "second neighbor" attachment possibility, could serve as a cement between two "hydrophilic" peptide faces.

Open Chain Peptides and Proteins

These possible water-peptide interactions through their respective hexagonal oxygen lattices prompted the examination of open chain peptides to see if the hexagonal concept could be extended to chains having no covalent cyclizing members. After some initial success with the ACTH N-terminal decapeptide (5), the study was further extended to the B chain of insulin (Fig. 5). This model shows the hexagonal pattern persisting throughout the chain of 30 amino acids, yielding 9 contiguous hexagons. A model of the entire sequence of cytochrome C (104 amino acids) was also prepared, but the photograph is too complex for a meaningful presentation. (Later this model will be presented by an alternate device.) It is sufficient to note that the honeycomb hexagonal pattern is maintained throughout the sequence, yielding a total of 41 hexagons. The continuation of this possible water-protein relationship throughout an entire peptide chain emphasizes what I believe to be a fundamental principle of this conformational proposal. Although there have been other suggestions concerning the mode of interaction of protein with water (9,10) and these alternate proposals also have considerable validity, there is, however, only one feature of every protein

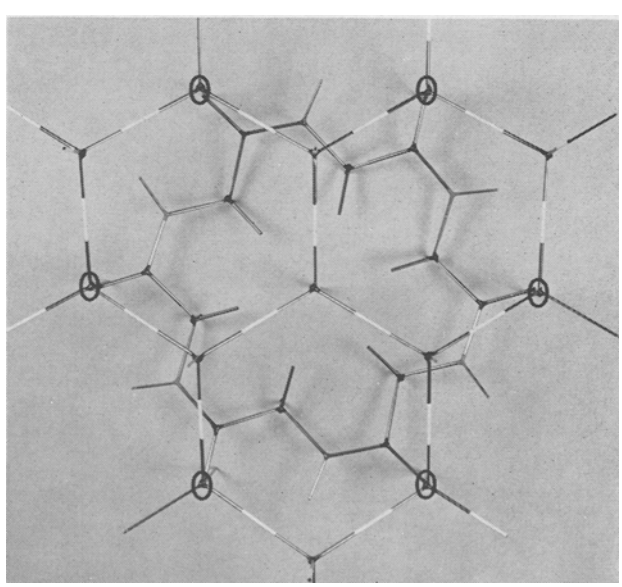


FIG. 4. Cyclic Hexapeptide superimposed on water lattice. 0 indicates junction points between peptide lattice and water lattice.

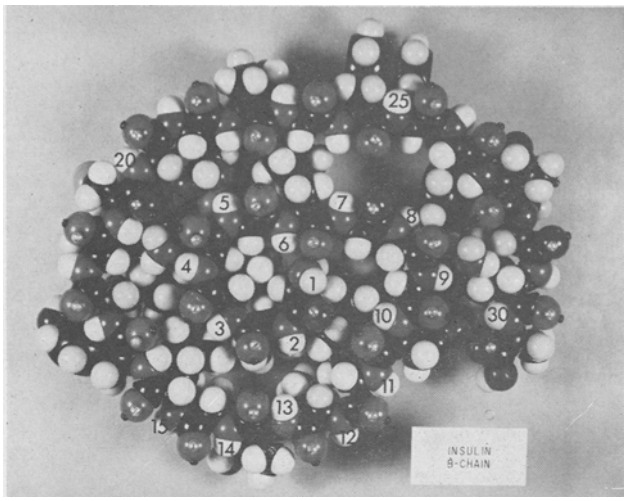


FIG. 5. Insulin B chain (*cis-trans model*). Hydrogen atoms of most of the peptide –N–H positions are numbered as indicated, screw heads show oxygen positions. (Reproduced from Warner, D. T., "A New Approach to the Structure of Polypeptides and Proteins in Biological Systems such as the Membranes of Cells," in Mechanisms of Hormone Action, P. Karlson, ed., Georg Thieme Verlag, Stuttgart, 1965.)

or peptide (regardless of side chain composition) that allows an ordered contact with water molecules arranged in an ice-like lattice. This feature is the uniformly repeating carbonyl O-atom of the peptide backbone. The uniformity is maintained irrespective of the presence of proline units, polar or nonpolar side chains or aromatic residues in the amino acid sequence. Since the studies with models have now emphasized the feasibility of arranging the peptide backbone in a conformation that favors the interaction of every backbone O-atom with a matching O-atom of the water lattice, this suggested hexagonal protein conformation presents the most meaningful general theory yet proposed for the mechanics of the water-protein interaction. Later it will be shown that such models also make provision for the possible interaction of each peptide –N–H group with the water lattice, so that mechanistically every feature of the hexagonal conformation lends itself to a favorable interaction with a bonded water layer.

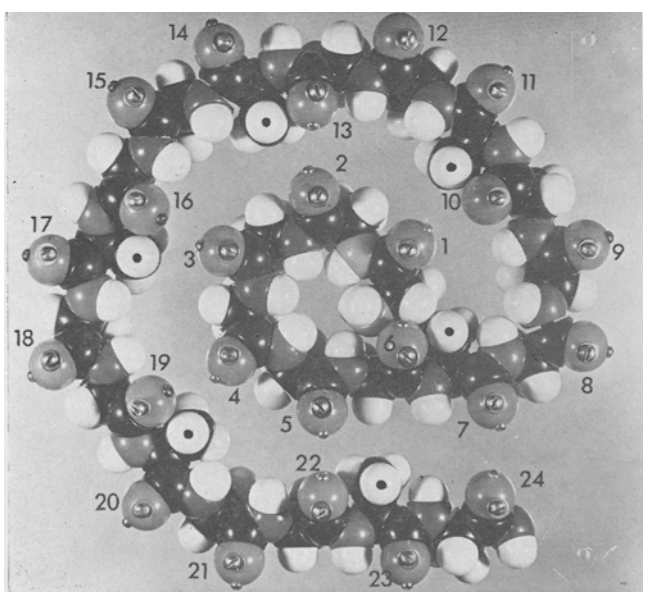


FIG. 6. Poly-L-Alanine (all-trans model) showing hydrophilic surface. Peptide oxygen positions are indicated by number.

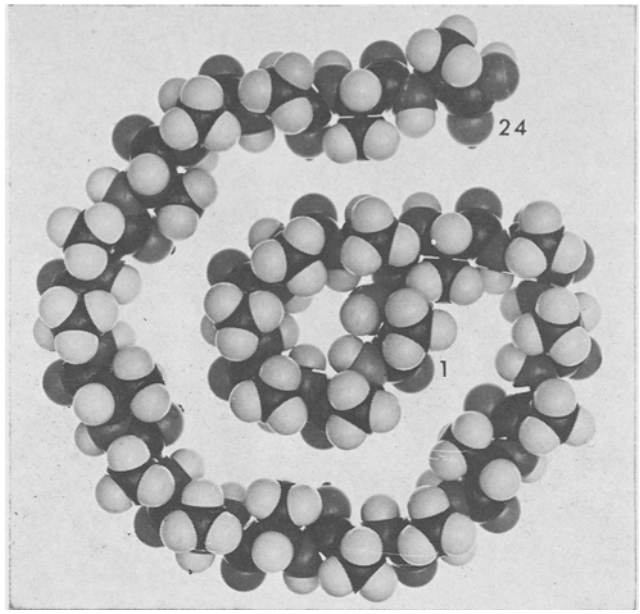


FIG. 7. Poly-L-Alanine (all-trans model) showing hydrophobic surface. 1 = N-terminus 24 = C-terminus.

We will now consider the variables which are allowable within this hexagonal framework without markedly altering the uniform honeycomb pattern of peptide O-atoms on which the water may arrange itself in an ice-like structure. The first consideration here should be the possibility of *cis* and/or *trans* amide junctures at the peptide bonds. In the early presentations of the small ring compounds all of the peptide bonds were in the *cis* form (see e.g., Fig. 3). In discussing these models in a previous publication (4), it was suggested that the *cis* form seemed to permit closer hydrophobic interaction between groups on adjacent α-carbon atoms. However, Badger and Rubalcava (11) have presented evidence suggesting that for amides and substituted amides, the *trans* form is perhaps stabilized in comparison with the *cis* form by more than 2 kcal/mole. Dickerson (12) has also commented that the use of *trans* amide junctures might be more favorable for models of the general type that I have proposed. Therefore, it seemed worthwhile to take advantage of a feature that could contribute additional stability at various places throughout the conformation and examine the implications of the change with molecular models. Consequently, instead of using a combination of *cis* and *trans* amide linkages for the larger peptides (see e.g., B chain of insulin in Fig. 5), several models were also constructed of known peptide sequences using all-*trans* amide bonds.

To illustrate the basic features of the *trans* backbone arrangement, a model of poly-L-alanine is shown in Fig. 6 (hydrophilic side) and Fig. 7 (hydrophobic side). Some general comparisons can be made between this model and the *cis-trans* model of Fig. 5. First of all, the direction of the chain coiling with all-*trans* bonds is counterclockwise (Fig. 6) instead of the clockwise coiling of the *cis-trans* form. Secondly, the β-carbon atoms of adjacent amino acid units are somewhat farther apart for the all-*trans* form (see Fig. 7), thus providing more space to accommodate larger van der Waals radii for the hydrogen atoms and carbon atoms if that seems desirable. This model was made with the Catalin set having a hydrogen radius of 0.95 A. The new C–P–K models have a hydrogen

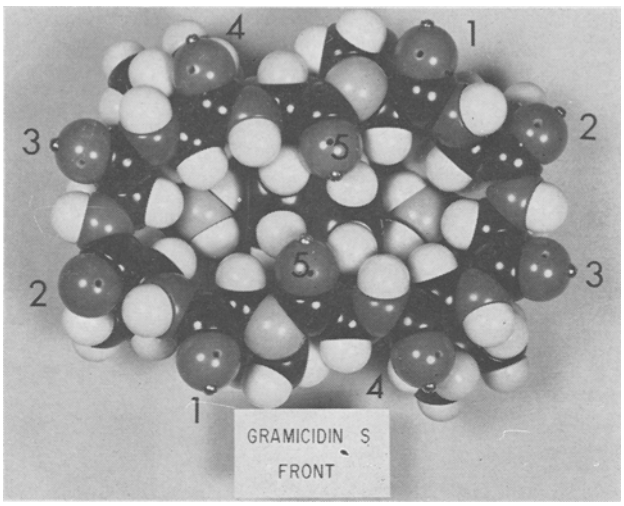


FIG. 8. Gramicidin S (*cis-trans* model) showing hydrophilic surface. $(1\text{–}2\text{–}3\text{–}4\text{–}5)_2$ = (*L*-Pro-*L*-Val-*L*-Orn-*L*-Leu-*D*-Phe)$_2$.

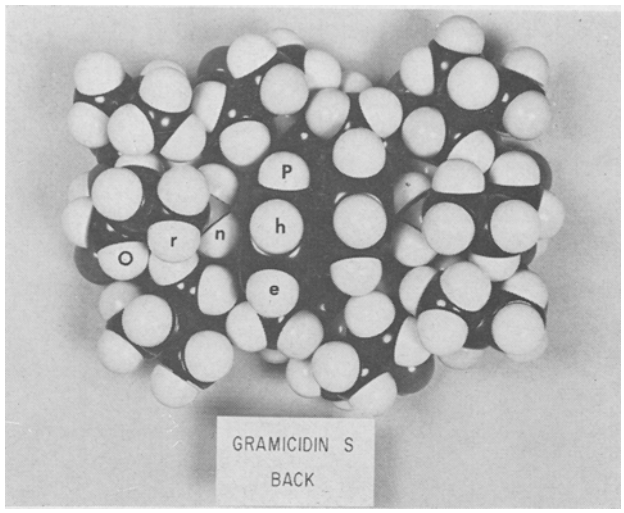


FIG. 9. Gramicidin S (*cis-trans* model) showing hydrophobic surface. Phe = phenylalanine, Oren = ornithine.

radius of 1.00 A, which is not significantly larger. Both larger and smaller van der Waals radii for hydrogen have been proposed as previously mentioned (2,3). Thirdly, considering only the first hexagon at the amino end of the chain (Fig. 6) in this open chain peptide, it will be noted that the H-atoms forming the ring around the hexagonal center are furnished by the –N–H groups of the peptide bonds instead of the H-atoms of the α-carbon atoms in Fig. 5. For such a peptide hexagon in contact with a water layer, there would be available a water molecule lying within the ring of H-atoms, and this water could be directly bonded to at least some of the –N–H groups. Since the bonded H-atoms could resonate along the bonding leg between oxygen and nitrogen, by moving in toward the water oxygen any possible crowding due to van der Waals contacts in the central zone could be relieved. With the Catalin models this contact is minimal even with all the H-atoms "on nitrogen." A fourth significant variation between all-*trans* and *cis-trans* shows up in the positioning of certain side chains. In either model some of the β-carbon atoms (or side chains) instead of pointing directly below the main plane of the backbone below the α-carbon will point to the side of the α-carbon in the plane of the backbone. In general these particular side chains in the all-*trans* model point *inward* toward the central hexagon and at the edge of the model there are no protruding side chains. On the other hand, with the *cis-trans* model this series of side chains point *outward* away from the central hexagon and at the outer edge of the model (see Fig. 5) protruding side chains appear. This has an interesting consequence in the case of cytochrome C with regard to the histidine group at position 18 in the sequence (13). In the *cis-trans* model with the haem group superimposed in the required position the imidazole ring of His-18 lies to one side contacting one of the substituents of the haem ring. In the all-*trans* model the same imidazole ring lies over the haem ring. Since the imidazole is usually pictured over this haem in contact with the iron (14) this could be taken to mean an argument for the all *trans* form, although it seems to me that there is no reason why the imidazole group, making contact with the edge of the haem ring in the *cis-trans* form, could not function equally well for electron transfer through the resonance qualities of the haem ring. In any event the two models of cytochrome C furnish two different presentations of the His-18 imidazole group to the haem ring, vividly illustrating one possible difference achieved by changing from the *cis-trans* to the all *trans* amide juncture in the case of this protein sequence.

The all *trans* model still might not satisfy the exact requirements of the Dickerson critique (12) with regard to the complete planarity of the amide link. This planarity is difficult to evaluate exactly from models, and I prefer to approach it from the standpoint of the theoretical principles involved. Theoretically, the most favorable hydrogen bonding of the peptide layer with an overlying water layer would occur where the dihedral angle formed in the bond members –H–N–C=O between the –N–C=O and the –H–N–C– planes is approximately 142° (for a tetrahedral angle in the water lattice) or else 151° (for the H–O–H angle of 104° 45′). If we now calculate the strain energy involved in twisting the amide bond from planarity to these angles, using the Dickerson formula (12), the value for the dihedral angle of 151° (29° from planarity) is only about 4.6 kcal/bond and for 142° (38° from planarity) is about 8 kcal/bond. Either result is somewhat lower than the average value of 10 kcal/bond used by Dickerson in evaluating the hexagonal conformation (12). In fact, if an "amino-type" nitrogen is used instead of a planar

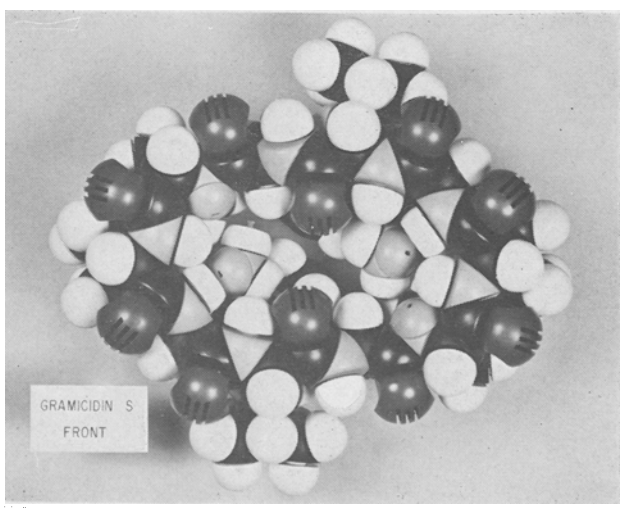


FIG. 10. Gramicidin S (all-*trans* model) showing hydrophilic surface. Grooved atoms = carbonyl oxygens.

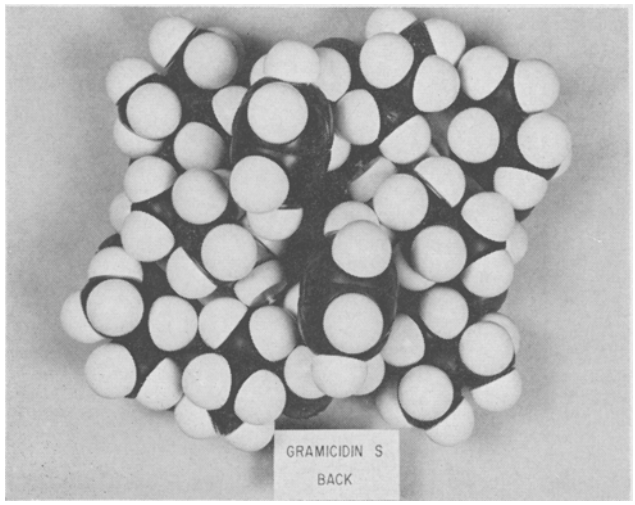


FIG. 11. Gramicidin S (all-*trans* model) showing hydrophobic surface.

nitrogen still another situation results and this also should be kept in mind. The main point that has emerged from the study of the all-*trans* models is the demonstration that here, too, the carbonyl oxygens can adapt themselves to the planar hexagonal water pattern without any interruption of the interaction pattern and with minimal distortion of bond angles. Consequently, if the all-*trans* configuration offers certain advantages from the standpoint of thermodynamic stability and more favorable van der Waals spacing (albeit with a possible decrease in hydrophobic fitting), then there seems to be no good reason why the use of all-*trans* peptide linkages cannot be utilized in the hexagonal theory of protein conformation.

This conclusion from the model studies is no assurance that the all-*trans* peptide juncture does exist in aqueous solution or biological systems. For certain situations, *cis* bonds seem to fit some of the data equally well. Grant (15) has recently repeated some studies of the molecules glycine, diglycine, and triglycine in aqueous solution and determined their charge separation by dielectric constant measurements. He concluded that the peptide chain is curved in space. If his charge separation distances are compared with the expected values for a peptide chain laid out with the Dreiding stereomodels on a hexagonal network (see e.g., Fig. 3), then the correspondence of distances is quite good if the peptide bonds between

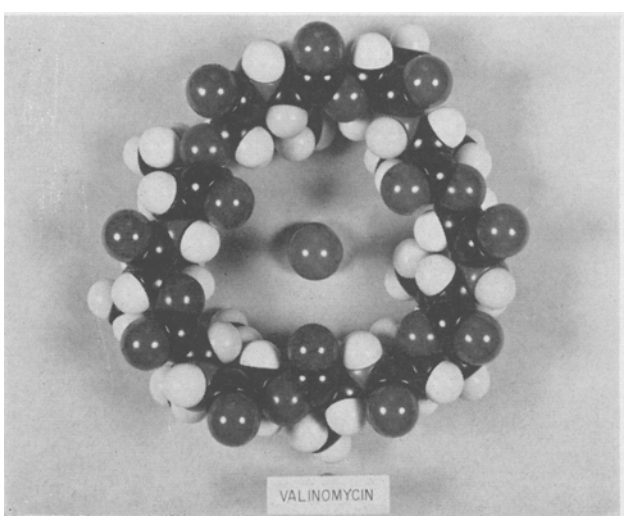
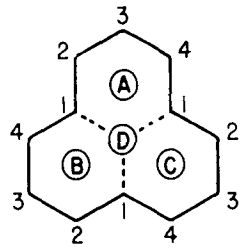


FIG. 12. Valinomycin showing hydrophilic surface.



$(1\text{-}2\text{-}3\text{-}4)_3$=

$(\underline{D}\text{-Val-}\underline{L}\text{-Lactic-}\underline{L}\text{-Val-}\underline{D}\text{-Hydroxyvaline})_3$

FIG. 13. Hexagonal pattern of valinomycin showing chain sequence.

glycine units are in the *cis* form and much less precise for the hexagonal conformation using *trans* peptide bonds. As Grant points out, a model of the extended chain (β-conformation) using *trans* bonds gives no useful correlation whatsoever with the measured values. Beacham et al. (16) have also made a similar study of a series of peptides. Here, too, correlation with a *cis* amide bond is good up to the tripeptide stage, but we have not been able to extend the hexagonal conformation correlation to the tetrapeptide and higher peptide stages. Beacham et al. suggest the possibility that the theoretical relationship between dielectric constant and dipole moment may need to be refined so that the correlations must be taken with reservations. Other interesting studies on glycine, diglycine, triglycine and tetraglycine have been done by Goto and Isemura (17) in aqueous solution. Although their results are not directly indicative or suggestive of the conformation in solution, these workers have shown that the peptide bonds are definitely hydrated in solution. They conclude that the peptide bond is a "hydrophilic" group.

Side Chain Interactions in Polypeptides

Thus far we have shown how the "hexagonal" approach to protein conformation was developed into a general theoretical concept with the use of models. The main subjects of discussion have involved the peptide O-atoms, their possible interaction with water, and the consequences of *cis-trans* interchanges of amide linkages. We will now evaluate the side chain interactions resulting in the hexagonal conformation of some known primary sequences. We have preferred to study models of biologically active peptides or proteins of determined structure, hoping that we are thereby assured that the primary order of amino acid residues in the chain progression is significant for some particular biological function, and approaching the conformation problem with the faith that some specific tertiary arrangement of these residues endows the primary sequence with the framework that imparts to it this elusive quality of biological activity. It is immediately apparent from models that for each specific arrangement of the peptide backbone (e.g., α-helix, pleated sheet or hexagonal conformation) the same amino acid sequence will present entirely different side chain reactions. Thus, for a sequence of 30 amino acids in the α-helical conformation, residue 1 will be relatively close to 4, 4 to 7, 7 to 10, 10 to 13, etc. For the extended chain structure used in the pleated sheet, very few side chains from the *same strand* can effectively interact with each other unless relatively large groupings are present such as lysine, glutamic acid or arginine, so the interactions that can be envisioned are going to be almost entirely

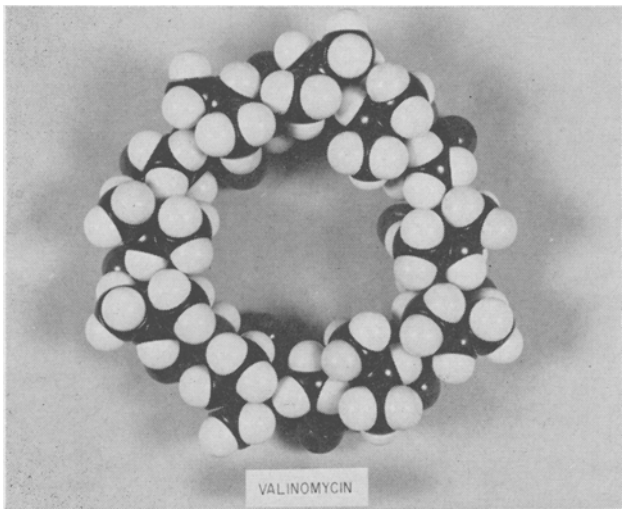


FIG. 14. Valinomycin showing hydrophobic surface.

between side chains on immediately adjacent units since any other units will be too far apart unless the chain folds back on itself. The hexagonal concept, however, allows amino acids which are far apart in the primary sequence to be very close to each other on the hexagonal spiral. A glance at the model of the insulin B chain (Fig. 5) readily shows that residue 1 is close to residue 10, and residue 9 is close to residue 30. Since other workers have extensively studied many models in the α-helical or pleated sheet conformation, the discussion in this section will be mainly concerned with the side chain positioning and interaction in the hexagonal concept of protein conformation. Furthermore, it seemed advisable to study three or four models quite comprehensively rather than to present a whole battery of models superficially.

Gramicidin S

A very interesting array of side chain interactions resulted from the hexagonal conformation with our first selected example, gramicidin S. This cyclic decapeptide is formed by the combination of two identical pentapeptides in the following sequence:

$$\ulcorner\text{(D-Valine-L-Lactic-L-Valine-D-Hydroxyvaline)}_2\urcorner$$

Two D-phenylalanines are present in this peptide, and two ornithines provide the only polar side chains. The present model (Fig. 8) differs in several respects from the model presented earlier (4). First, in assembling larger models with the Catalin set the entire structure, both side chains and backbone, is now threaded on thin nickel wire to avoid separation of the rubber pegs and consequently the atoms can be held tightly together. Second, this model differs conformationally in having the –L–Leu–D–Phe amide junction in the *trans* form, a possibility previously suggested (4) but not previously illustrated in print. The model contains four *trans* and six *cis* amide linkages. The main chain is laid out so that the carbonyl O-atoms form two fused hexagons similar to Fig. 2. For the L-amino acids in the sequence *cis* amide bonds bring the α-C-atoms toward the center of the respective hexagons and the H-atoms on these α-C-atoms form two compact rings in the centers. For the *trans* junction at the –L–Leu–D–Phe position the opposite D-configuration projects the α-H-atom of the D-Phe outward and the benzyl side chain now has the $-CH_2-$ portion just below the α-carbon position. The benzene

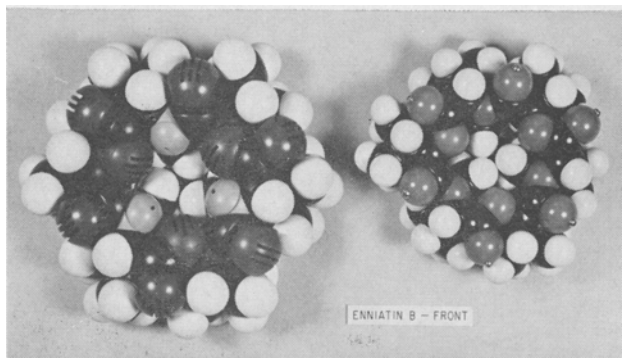


FIG. 15. Enniatin B showing hydrophilic surface: *large model*, C-P-K set; *small model*, Catalin set.

ring standing on edge projects into the central space adjacent to the two central (or common) peptide O-atoms of the fused hexagons. The position of entry is such that the two benzene rings, coming into the center from opposite directions lie exactly side by side. It is seen that the ring of H-atoms of the respective hexagons are thus completed by four H-atoms from the benzene rings, two on each side. The planes of the two benzene rings are perpendicular to the long axis of the molecule, and the resulting benzene "sandwich" has an ornithine amino group on either side, with the amino ends emerging into contact with the hydrophilic surface through the central "holes" in the two hexagons. In a simple backbone model such as poly-L-alanine (Fig. 6 and 7), the hexagonal units have open centers, but in most biological sequences these centers are usually occupied by certain polar side chain groups, as in this gramicidin S model.

The back side of this *cis-trans* model (Fig. 9) clearly shows the two benzene rings lying side by side. The compact arrangement of the side chains is quite obvious, and the general shape and size are in quite good agreement with the 11×16 A rectangular dimensions described by Hodgkin's X-ray data (18).

To complete the study of gramicidin S, a model having all *trans* amide bonds was constructed with the C–P–K models. Fig. 10 shows the peptide surface. Except for the two peptide –N–H groups of the two D-Phe units (which point outward) all of the other peptide –N–H groups and the $-\ddot{N}-$ of the two proline units form a ring around the hexagonal centers. The $-NH_2$ groups of the two ornithine side chains appear in the center of this ring, and here each of them can be bonded to the ornithine peptide –N–H group through a hydrogen bond while one of the amino H-atoms can hydrogen bond with the electron pair of the proline peptide nitrogen. The benzene rings of

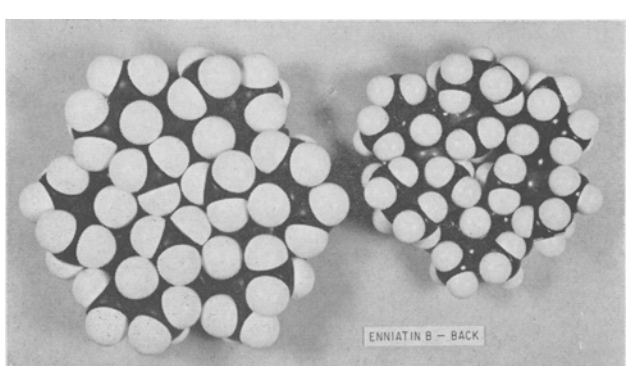


FIG. 16. Enniatin B showing hydrophobic surface: *large model*, C-P-K set; *small model*, Catalin set.

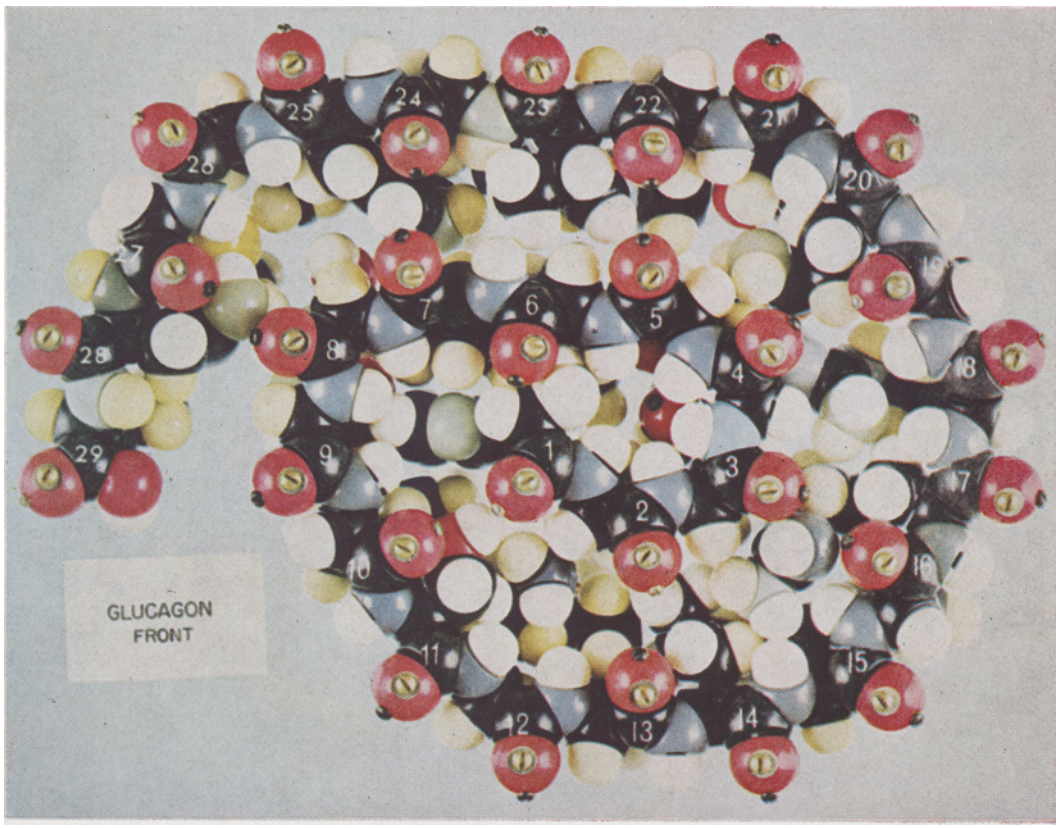


FIG. 17. Glucagon showing hydrophilic surface: *a*) Sequence numbers are placed on the respective carbonyl carbon atoms (sequence given in Table II; *b*) red = oxygen, blue = nitrogen, black = carbon, white = hydrogen, screw heads on peptide oxygens.

the D-Phe units again point toward the central open area of the molecule, but are not seen clearly from this surface in the "side-by-side" position characteristic of the *cis-trans* model (Fig. 8). Again the –N–H–CH(R)–C=O direction within the chain is counterclockwise and the carbonyl O-atoms common to the two hexagons are here supplied by the L-leucine residues instead of the D-phenylalanine units as in Fig. 8. The back side of the all *trans* structure of gramicidin S is illustrated in Fig. 11. Here the position of the two benzene rings of the D-phenylalanine side chains is easier to see and they are not "side-by-side." The general paraffinic nature of this side chain surface is again apparent, but there are no particularly outstanding features.

Depsipeptides

Recently another type of cyclic antibiotic has been completely characterized by a total synthesis. This type of structure is not strictly a "polypeptide" but is composed of a mixture of amide and ester bonds formed from *a*-amino and *a*-hydroxy acids of either D- or L-configuration. Such compounds are commonly known as depsipeptides. One of the compounds in the class is valinomycin, a cyclododecadepsipeptide having the formula:

┌L-Val-L-Orn-L-Leu-D-Phe-L-Pro-L-Val-L-Orn-L-Leu-D-Phe-L-Pro┐

The total synthesis of this compound and several similar products has been described by Shemyakin et al. (19). In the formulation of a possible conformation for valinomycin, it is again difficult to apply the *a*-helical concept, not only by reason of the cyclic structure but also because of the paucity of –N–H groups for the required –C=O- - -H–N– bonds. Similar considerations apply to the formulation of the pleated sheet or *β*-configuration. It is, however, possible to arrange the carbonyl O-atoms on a hexagonal lattice as illustrated in Fig. 12. This arrangement leaves the model with an open center which would be a common point for the completion of each of the three hexagons. In the photograph of the model an O-atom has been inserted there to illustrate the completed hexagonal pattern indicated by the dotted lines radiating from point *D* in Fig. 13. Fig. 13 also shows the chain sequence as noted in the legend. This may be an esthetically satisfying picture, but any suggested conformation should also furnish us with some clue as to why valinomycin is a reactive compound. "Reactive" amino, hydroxyl, or carboxyl groups, or even aromatic rings are not present in valinomycin. The back side of the model shows a completely hydrophobic environment (Fig. 14). However, if we examine the front side, or peptide surface, of Fig. 12 and consider its interaction with a water layer in contact with it, the following points may be noted. Each of the hexagons (or partial hexagons since the common central point is not furnished by the chain) will make contact with water through its carbonyl O-atoms. Such a water lattice in the ice form will present a centrally located water O-atom to the molecule at each of the hexagonal centers lettered *A*, *B*, and *C*, respectively (Fig. 13). To this water molecule our proposed structure will present the following groups: –C–H, –C–H, –Ö–, and –N–H. These groups are available for hydrogen bonding contact with the water but in various ways. Thus, the water molecule can donate a hydrogen to the –Ö– of the ester, accept a hydrogen from the –N–H of the amide and find its resonance possibilities blocked by the –C–H groups, which have available H-atoms although presumably not strong hydrogen bonding partners. In the binding of the water at that

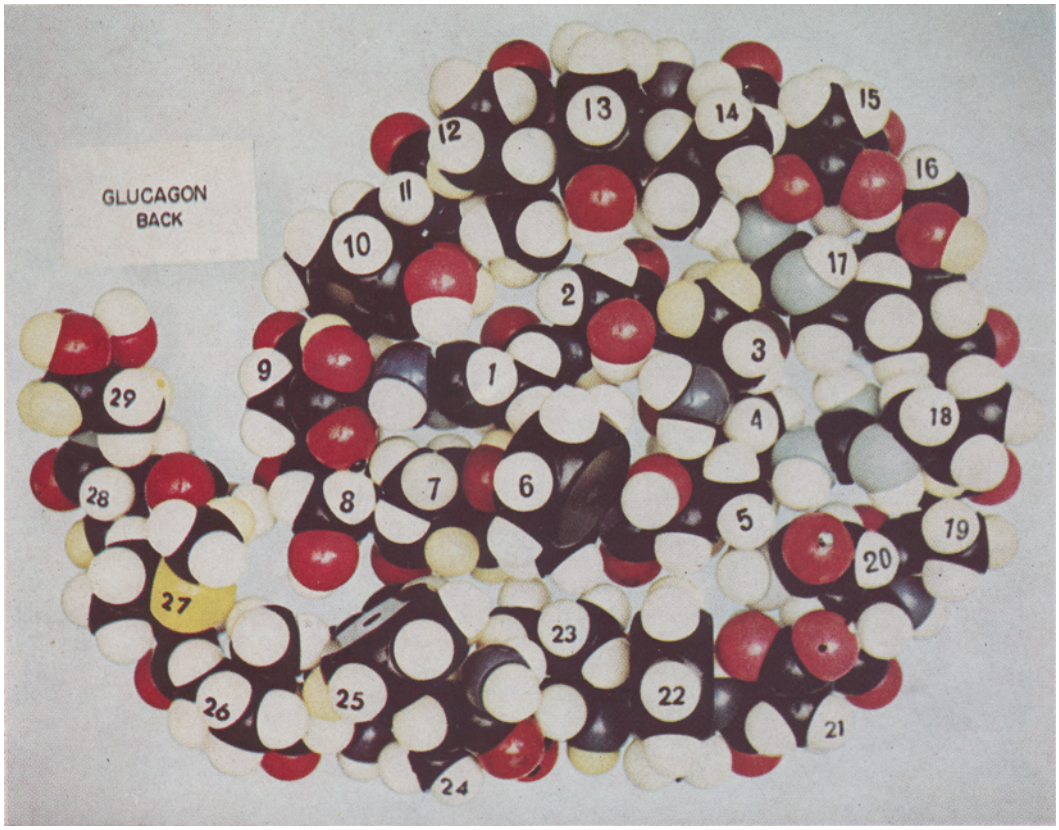
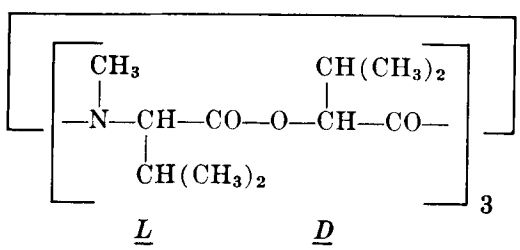


FIG. 18. Glucagon showing hydrophobic surface. *a*) Sequence numbers are located on some part of the proper side chain (sequence given in Table II); *b*) red = oxygen, blue = nitrogen, black = carbon, white = hydrogen.

site, the ester O-atom will have a tendency to draw an H-atom from the water lattice to fulfill its bonding possibilities, and consequently that particular water molecule may have one of its two available H-atoms more or less permanently oriented away from contact with the surrounding water molecules, thus producing a local disturbance in the water resonance. The same thing could apply to the water molecules at each of the positions *A*, *B*, and *C*. Thus, the whole structure acting in this manner could produce quite a concerted change on the resonance of the immediately adjacent water structure. The general effect would be to leave certain of these water molecules with a (δ-) charge, perhaps capable of attracting hydrogens from adjacent protein layers or ions such as potassium ions from the media. Valinomycin is known to have an effect on potassium transport (20). Thus, a highly coordinated system of so-called "neutral" groups may have just as profound an effect on the resonance of the water lattice as that exerted by a more "reactive" group such as an amino group.

It is also noteworthy that the cyclohexadepsipeptides, which would have only one hexagon of carbonyl O-atoms, in some instances also form a reactive group of compounds (21). This is especially so when the hydroxy and amino acid residues alternate regularly. The two most active compounds contain three N-methyl amino acids and three hydroxy acids. When these are joined in the ester-amide linkages to form the cyclic depsipeptide, there are no –N–H groups available for the –C=O- - -H–N– bonds proposed for either the α-helix or pleated sheet. Consequently, there must be many other forces which influence the conformation of molecules and which are capable of producing biologically active arrangements. A proposed "hexagonal" model of a cyclohexadepsipeptide is given in Fig. 15 showing the hydrophilic carbonyl oxygen surface constructed with both the C–P–K models (larger size) and the Catalin models. All amide and ester linkages are in the *trans* form. The sequence shown is that of enniatin B (21), which has the formula:



Both models show clearly the possibility of arranging the carbonyl O-atoms in the hexagonal pattern. In the C–P–K models these O-atoms comprise the outer ring of six O-atoms (identified in the model by the slots used for making hydrogen bonds). The inner ring of O-atoms and N-atoms is indicated by the three ether O-atoms (with slots) alternating with the three N-atoms. In the C–P–K model these N-atoms were represented with amino nitrogens which allow the methyl group on the nitrogen to lie below each N-atom. In this way each N-methyl group can project one of its H-atoms into the very center of the inner ring. If planar N-atoms are used with the C–P–K set these central three H-atoms require slight compression. This is not true when the model is assembled with the Catalin set, for in that model the planar or amide N-atoms were employed without difficulty, perhaps due to slightly smaller radii for the H-atoms in that set. The back side of this molecule as constructed

TABLE I
Hexagonal Units of the Peptide Surface of the Glucagon Model (Figure 13)

Number of Hexagon	Carbonyl Numbers from Figure 13	Possible Side Chain in Hexagon Center
1	1,2,3,4,5,6	Glu(NH_2)-3
2	1,6,7,8,9,10	His-1
3	1,2,10,11,12,13	Lys-12
4	2,3,13,14,15,16	Arg-17
5	3,4,16,17,18,19	Arg-18
6	4,5,19,20,21,22	Glu(NH_2)-20
7	5,6,7,22,23,24	Val-23
8	7,8,24,25,26,27	Ser-8
9	8,9,27,28,29	Thr-29

with the two sets is shown in Fig. 16. The very precise fitting of the nonpolar groups scarcely needs additional comment, and together with the uniform arrangement of carbonyl O-atoms in Fig. 15 represents a very provocative proposal for the conformation of enniatin B.

Glucagon

It has already been suggested that the hexagonal conformation can be extended to open chain peptide sequences in which no cyclic ring is present. One such model has already been presented for the insulin B chain (Fig. 5) to show the persistence of the hexagonal lattice for a combination of *cis* and *trans* amide bonds. No detailed discussion of the resulting side chain interactions was presented. We now would like to take a molecular model of a known compound, glucagon, in the all-*trans* form and document the various side chain interactions that are observable from a sequence in the hexagonal conformation. In order to emphasize again the basic arrangement of the peptide backbone in the all-*trans* form, let us refer back briefly to the 24 unit poly-L-alanine model in Fig. 6 and 7. On the front side (Fig. 6), it will be noted that except for the central hexagon (with 5 –N–H groups and one amino pointing inward) and the second hexagon (with 3 –N–H groups pointing inward) the remaining 5 hexagons and all subsequent hexagons in a longer chain would have only 2 –N–H groups projecting toward their respective hexagonal centers. One methyl group, marked by a black dot on one of the H-atoms, also points inward on hexagons 2–7. The back side of the model (Fig. 7) shows that there is adequate space between the methyl groups on the adjacent α-carbon atoms of the backbone to accommodate larger van der Waals radii for the H-atoms if that seems desirable (Catalin models have a hydrogen radius = 0.95 A). With the small methyl side chains of alanine the backbone spiral is a fairly open structure. However, if we examine the model of a known naturally occurring sequence, it is quite obvious that many side chain groups are considerably larger than methyl, and most of the space between the backbone atoms is filled in to give a more compact assembly. This is well illustrated by the model of the all-*trans* conformation of the glucagon structure (22) shown in Fig. 17 (hydrophilic surface) and Fig. 18 (hydrophobic surface). In Fig. 17 the numbers appearing on the carbonyl carbons of the backbone identify the 29 amino acids in this sequence (See Table II for sequence). This all-*trans* model illustrates the uniform manner in which the hexagonal pattern of carbonyl oxygens is maintained throughout the sequence. The carbonyl positions comprising the nine hexagons and the central group in each hexagon are included in Table I. The side chain surface is given in Fig. 18. A list of some of the possible side chain interactions is given in Table II. It should not be supposed that the side chains are so firmly positioned that these suggested contacts are the only possible ones. The list is given to emphasize the wide and integrated side chain interplay that is possible with the hexagonal conformation. There are several well-directed hydrogen bonds, two and possibly three electrostatic interactions and a few hydrophobic bonds. In some instances the hydrogen bond network interconnects at least three polar groups into a single unit. In other instances the interacting groups may be quite far apart in the primary sequence (e.g., His-1 with Asp-9). There are other possible bondings which were not cited in the table, such as the bonding of the Ser-11 hydroxyl to the main chain carbonyl oxygen of His-1. These bonds could also be equally effective as stabilizing features or activating influences, for the main chain peptide bonds must also qualify as "reactive" groups. In this sense, too, the hexagonal concept of protein conformation offers additional thoughts about the efficacy of proline units or –N–CH_3 amino acids in the role of "reactive" groups. If the peptide chain is in intimate uniform contact with a water lattice, a proline residue in that chain poses the likelihood of a specific alteration in the resonance pattern of the particular water molecule in contact with it. This water molecule may contact and hydrogen bond to the proline nitrogen, but when it does, the H-atom for the bond must always come from the water molecule since the N-atom has only its electron pair to offer. By contrast the –N–H of an H–N–C=O bond can also furnish an H-atom to its nearest water molecule, so it will be apparent that different resonance situations will result in the two instances. Consequently, in the theoretical aspects of the hexagonal concept each proline unit presents itself as a distinctive reactive function, of key importance to biological

TABLE II
Possible Side Chain Interactions in Glucagon (Figure 14)

Sequence Number	Amino Acid	Possible Interaction
1	His-1	Electrostatic with Asp-9
2	Ser-2	Hydrogen bond with amide —NH_2 of Glu(NH_2)-3
3	Glu(NH_2)-3	See Ser-2. Also —C=O at center of Hexagon 1 (Figure 13)
4	Gly-4	Small size facilitates Glu(NH_2)-3 positioning
5	Thr-5	Bonds with —NH_2 of Glu(NH_2)-3
6	Phe-6	No particular interaction in the position shown
7	Thr-7	Hydroxyl group forms hydrogen bond with α-NH_2 of His-1
8	Ser-8	Center of Hexagon 8 (Figure 13). Can also be rotated to contact Asp(NH_2)-28 in hydrogen bond
9	Asp-9	See His-1, above
10	Tyr-10	Phenolic —OH is near the imidazole nitrogen of His-1
11	Ser-11	Hydrogen bond to ε-NH_2 of Lys-12
12	Lys-12	a) See Ser-11, above b) Hydrophobic with Tyr-13
13	Tyr-13	Benzene ring held between side chains of Lys-12 and Leu-14
14	Leu-14	See Try-13, above
15	Asp-15	a) Electrostatic with Arg-17 b) Hydrogen bond with Ser-16
16	Ser-16	See Asp-15, above
17	Arg-17	See Asp-15, above
18	Arg-18	Hydrogen bond with Glu(NH_2)-20, which in turn may form electrostatic bridge to Asp-21
19	Ala-19	Methyl group serves as possible rotation barrier to Arg-18
20	Glu(NH_2)-20	Possible bonding bridge between Arg-18 and Asp-21 (See Arg-18, above)
21	Asp-21	See Glu(NH_2)-20, above
22	Phe-22	Hydrophobic bond with Val-23
23	Val-23	See Phe-22, above
24	Glu(NH_2)-24	No definite interaction. Might be located to point into the center of Hexagon 7 (Figure 13)
25	Try-25	Hydrophobic with Leu-26
26	Leu-26	See Try-25, above
27	Met-27	Hydrogen bond with amide —NH_2 of Asp(NH_2)-28
28	Asp(NH_2)-28	See Met-27, above and also possible Ser-8, above
29	Thr-29	No special interaction in position shown in the model. Possible center of Hexagon 9 (Figure 13).

potency, rather than the weak or absent link in an otherwise uninterrupted helical spiral.

Cytochrome C

Although a complete model of cytochrome C has been assembled with Catalin models and photographed, an 8 × 10 in. print of this model is so small that even the definite hexagonal arrangement of carbonyl oxygens is difficult to follow at that magnification. The side chain details are so hard to label clearly that for the present the publication of such models is best served by resorting to another method. This method employs hexagonally ruled paper to represent the carbonyl oxygen network, and the result is a very useful "model" indeed. The paper was carefully ruled to the scale of 1 cm = 4.8 A so that the dimensions of a protein subunit in the hexagonal conformation could be measured directly. This device was first used for the representation of the tobacco mosaic virus protein (23). A hexagonal paper model of cytochrome C (Fig. 19) outlines the overall shape of the subunit of 104 amino acid residues plus an N-terminal acetyl group furnishing an additional carbonyl oxygen for the network. As in the case of the tobacco mosaic virus unit (23), this terminal acetyl oxygen has been numbered zero to avoid changing the accepted chain numbering (13). From a knowledge of the primary sequence as well as our assembled model of cytochrome C, it is possible to use selected symbols to represent desired side chain groups quite precisely in the paper honeycomb pattern. This is done for certain groups in contact with the haem ring as judged from the all-*trans* model. Since the molecular model of cytochrome C was used as a guide for labeling the paper model, the placement of these groups is quite accurate. The cytochrome C model construction also lends more confidence to the tobacco mosaic virus case, for it seems unlikely that the TMV sequence (158 amino acids) will present any additional structural difficulties not already encountered in the cytochrome C study.

The following points about the cytochrome C—haem interaction were observed in our model. First of all, the two cysteines at positions 14 and 17 in the cytochrome C sequence are appropriately placed so that they contact the vinyl side chains of the haem molecule at the desired positions, forming the necessary thioether bonds (24). The thioether bonds can also be formed with this portion of the cytochrome C sequence arranged in the α-helix conformation (25), a point that I have also verified. The all-*trans* hexagonal model here brings the imidazole group of the His-18 directly into contact with the central portion of the haem. This arrangement should be checked with the more precise C–P–K models. One other major point about the haem interaction deserves comment. In the cytochrome C sequence, position "0" (our numbering) is an acetyl group and position 1 is a glycine. Both of the groups are sterically very small and consequently they provide an open space on the back side of the model, approximately in the area outlined with the dotted line. If the haem group is now positioned over the model so that the thioethers can form with the vinyl side chains and the ring projects toward the central hexagon (positions 0,1,2,3,4,5), then the two propionic acid side chains on the haem ring fit very nicely into that open area. Therefore, although the haem group is attached to the back side of the cytochrome C model (or "hydrophobic" surface), the open space provided by the small groups at positions 0 and 1 in the sequence allows the haem to make contact

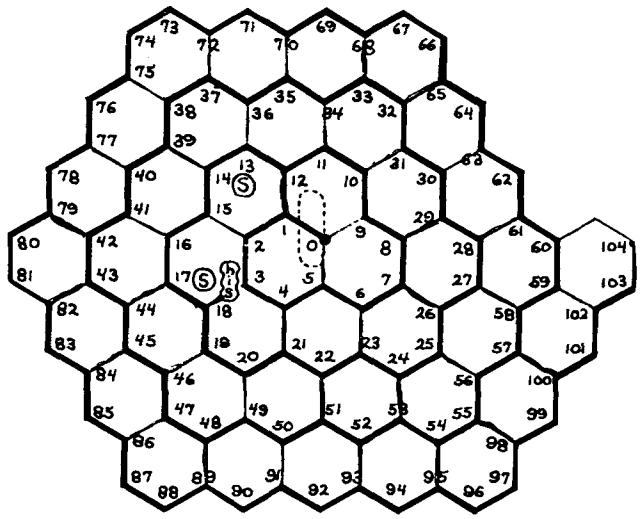


FIG. 19. Cytochrome C subunit on hexagonal paper model. *a*) Numbers indicate primary sequence positions (Ref. 13) *b*) *S* = cysteine, *his* = histidine.

with the front side (or "hydrophilic" surface) of the subunit by having these propionic acid side chains project through this opening, bringing the carboxyl groups into the aqueous layer which the hexagonal conformation presumes to be present at the peptide surface. Although the spacing may be coincidence, it furnishes a plausible reason for having an N-acetyl group followed by an equally small glycine group at the N-terminus of the molecule.

The possible binding of one of the coordinate positions of haem iron to a carboxyl group is not ruled out by the cytochrome C model. The aspartic acid at position three in the sequence can be so placed that it lies exactly under the center of the haem ring. When so placed it is also possible for it to be neutralized by the lysine at position 13 in the sequence so that it could form an ionic bond with the iron. In this placement Asp-3 is also available to form a hydrogen bond with the His-18, thus enabling it to still transfer electrons to the histidine unit. This possible Asp-3 interaction is mentioned because carboxyl attachment of the iron has not been ruled out unequivocally. Williams (26) states that Co, Fe, and Ni are likely to be bound to mixed oxy anion-nitrogen coordinating ligands. Heller and Smith (27) retain the possibility that the sixth ligand of heme iron in cytochrome C may be either an O-atom (glutamyl or tyrosyl) or the indole of tryptophan.

The Aggregation of Protein and Peptide Molecules

Thus far the discussion has mostly concerned models of single molecules of antibiotics, peptide chains, or protein subunits. It is well known that many of these single molecules are capable of dimerizations, trimerizations or aggregations to even higher stages of complexity. The forces operating in these various stages of polymerization are usually thought to be hydrophobic, hydrogen bonding or electrostatic forces. In the case of proteins, the idea that these polymers are built up from monomeric entities designated as "protein subunits" is now quite firmly established. A subunit consists of a single polypeptide chain having a precise and characteristic amino acid sequence. Thus, the 104 amino acid residues of cytochrome C constitute a protein subunit. "Proteins" of multimillion molecular weight are probably always comprised of many subunits, joined together at least in part by the weaker secondary forces mentioned above.

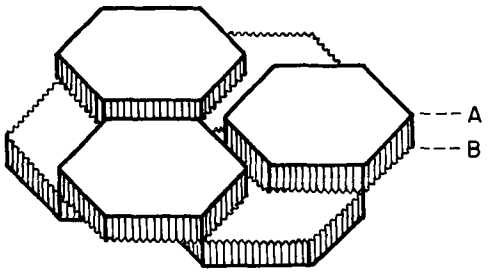


FIG. 20. Schematic drawing of hexamer of subunits. *A* (straight line) hydrophilic surface; *B* (ridged line) hydrophobic surface. (Included with permission of the Journal of Theoretical Biology.)

For each conformational concept of the individual subunit, a different idea related to subunit aggregation will arise, since different groups are exposed for interchain contact. Thus, for the α-helix, aggregates are pictured as three strand ropes which further combine in multistrand ropes or supercoils. For the β-configuration, we have the parallel and antiparallel pleated sheets, with these sheets in turn stacked in layers. I will present some thoughts about the hexagonal concept of subunit conformation with regard to the modes of its potential aggregation properties. These ideas were first employed in attempting to fit the tobacco mosaic virus (TMV) protein subunit into the structural features and known dimensions of the virus rod (23). Beginning with an individual protein subunit, such as that given for cytochrome C (Fig. 19) in its hexagonal conformation, we see that the subunit is essentially a flat-sided plate with hydrophilic (peptide bond) and hydrophobic (side chain) surfaces. Here the hydrophilic surface contains all the peptide O-atoms in what seems to be an ideal hexagonal array for contacting a water layer, which is then pictured as adhering to this surface through hydrogen bonds. The term "hydrophobic" is in some ways an oversimplification used for convenience only, especially for a protein subunit where some polar or "hydrophilic" groups will almost invariably be present in this surface, though not necessarily in a regular array. However, given a protein subunit of this fundamental type, the problem of aggregation was approached from the standpoint of defining mutually attracting surfaces. On the basis of arguments pointed out in some detail for the TMV case (23), the choice was made to have the hydrophobic surface of one subunit interact with the hydrophobic surface of another subunit through the postulated media of hydrophobic bonds, hydrogen bonds between $-CO-NH_2$, $-CO_2H$ and other polar groups such as hydroxyls. Another type of bond, although not present in the TMV example, for bonding together such hydrophobic surfaces is the covalent –S–S– bond of cystine units. This aspect will be referred to later. The general effect of these combined bonding patterns is the formation of a sort of subunit sandwich sealed together by interaction of side chains so that the hydrophobic regions are largely buried, while the hydrophilic peptide groups would be exposed completely on the two outward facing surfaces of the sandwich. This kind of sandwich is like a lipid bilayer of the Danielli membrane model (28) where the bond between the two lipid layers is presumed to be hydrophobic and the outer surfaces are comprised of polar hydrophilic lipid groups. A schematic representation is illustrated in Fig. 20. This is based on the TMV case, in which three upper subunits are pictured in bonding with three lower subunits. It will be noted that the subunits are not directly over each other, but are staggered. This arrangement was selected largely because it fits a hexagonal cross section as previously explained (23). The alternate directly superimposed subunit array is certainly not excluded.

If "back-side" interactions have any validity, the preferred positioning might depend upon specific group attractions and shapes acting to produce favorable energetics at some particular overlap position. This combination of six subunits might be termed a "hexamer," and with TMV protein its molecular weight would be close to 100,000, a value similar to the molecular weight of the A-protein aggregate of Schramm (29). In this aggregation concept based on hexagonal units, the hexamer can be used as a fundamental building block. The equatorial assembly of such hexamers around a central one, employing interpenetration and new hydrophobic overlaps at the hexagonal corners, would serve to form thin sheets of protein with two water-coated surfaces. The axial assembly of such hexamers, in which the hexamers would be piled above each other in the form of stacked discs, with the water layers serving as both cement and "spacer" between adjacent discs, leads to a rod-like assembly.

It was suggested for the TMV model system (23) that these axially stacked hexamers were equivalent to rod formation, for an examination of the basic hexamer assembly had shown that it had practically the same cross-sectional shape and dimension as the virus rod itself. This axial assembly of the hexamers may be a temperature sensitive process. An attempt to analyze the "second neighbor" oxygen distance in water (23) seemed to indicate that at about 25–30C the hexagonal "second neighbor" water oxygen pattern would be dimensionally similar to the hexagonal peptide oxygen pattern. Consequently, it might be anticipated that water would be the most compatible cement at about 25–30C. Lauffer (30) has studied the polymerization of TMV protein in the aqueous system as a function of temperature, using turbidity as a measure. Maximum turbidity is attained at about 30C. The axial aggregation of subunits is therefore equated in this instance with a favorable correspondence of water-peptide patterns at certain temperature ranges, leading to hydrophilic-hydrophilic interaction and rod-like aggregates via a water-cementing action (23). These views have been further elaborated in a later reference (31). Other interesting phenomena involving disaggregation-reaggregation of protein as a function of temperature at a specific pH value have been noted for flagellin and phycocyanin. In the case of flagellin (32) Vegotsky et al. have noted that at pH 4.08 precipitation of the protein was observed to be more rapid at 26C than at 4C, and filaments of different length but uniform width have been observed. In the studies with phycocyanin, Scott and Berns (33) have proposed a hexamer as the aggregation species which is important in vivo, and the factors of pH and temperature which influence its formation have been discussed.

The Hexagonal Concept and Biological Membranes

One of the intriguing possible applications of the hexagonal protein conformation and the suggested protein subunit interaction has been postulated by Hechter for the molecular organization of cell membranes (34). In this proposal water-coated lipid bilayers are visualized as interacting with water-coated protein bilayers of "interlocked discs" (34), with water

serving as the interlayer cement. An important consequence of the hexagonal protein conformation as devised in this model is the allowed formation of membrane pores. These pores are structurally arranged at stated intervals in the protein layer as a direct result of selecting this specific conformation for the protein subunit. Inherent in the proposal is the postulate that when the protein subunit is a different size (i.e., shorter or longer primary sequence) the holes will be closer together or further apart, respectively. Thus, the hexagonal concept of protein conformation brings to the realm of molecular biology a new theoretical approach to membrane structure and function, especially with relation to the protein components, although a few studies of molecular models of phospholipids and triglycerides in the water lattice (31,34) have also served to verify the possibilities of orderly lipid-water interactions.

In the lipid area, the number of known compounds which can still be examined with molecular models for significant interactions with the model of the water lattice is legion. For example, it is intriguing to find that the quinoid "head" of coenzyme Q_{10} as well as the polyunsaturated "tail" are both capable of interesting compatibilities with the water structure. The unsaturated double bonds (all-*trans*) in the fully extended "tail" are uniformly placed at exactly the same distance apart as a row of "second neighbor" O-atoms in the water lattice; also the extended "tail" of coenzyme Q_{10} has about the same length (ca. 45A) as the thickness of a lipid bilayer. One is therefore tempted to speculate that the lipid bilayer "holes" of the Hechter model (34) might be conveniently filled with extended molecules of coenzyme Q_{10}. These strategically placed units could serve for electron transfer at selected points across the lipid bilayer, perhaps via the medium of an iron hydroxide core, hydrogen bonded to the uniformly spaced double bonds of the coenzyme "tail." Presumably more than one coenzyme Q_{10} might be inserted in such a hole, either in parallel or antiparallel arrangement. The quinoid "head," placed perpendicular to the isoprenoid "tail" could serve as sort of a cap over the hole as well as a contact point between the iron core and the outer environment.

It should not be implied that all of the problems that arise in attempting to apply the hexagonal protein conformation to known situations are minor ones. Our proposal for the structure of actinomycin D has been challenged (35) and perhaps rightly so. Although all of the objections may not be valid (36), perhaps still others can be raised (37). While the "model builders" encounter numerous difficulties, the crystallographers who test the validity of these models also have theirs. The old problem of structure in "crystal versus solution" (to which we might add "solvent used") may be with us for a long time. Even the simplest examples in this area deserve considerable reflection. For example, L-hydroxyproline in the crystal form has an angle of buckle (referring to the "envelope conformation" of the ring) of about 20° from the work of Donohue and Trueblood (38), but the work of Abraham et al. (39) suggests that in solution this angle of buckle is about 50–60°. With the aid of molecular models of L-hydroxyproline (Dreiding stereomodels), it is interesting to observe than an angle of buckle of 50–60° brings the ring N-atom and hydroxyl O-atom closer together than in the crystal where the angle of buckle is 20°. In the solution situation the N–O distance is quite exactly equivalent to the "first neighbor" oxygen distance in the water lattice, but in the crystal form the N–O distance exceeds this value appreciably. Equally disconcerting is the report by Karle and Karle (40) that crystalline cyclohexylglycine hemihydrate exists in four distinctly different conformational isomers. One is tempted to ask whether in more complex protein crystals similar conformational isomers might not also complicate the interpretation. Perhaps the same hazard applies to any claims about "single forms" in solution states, and consequently one is not even certain under what experimental conditions to work in order to secure proof for a proposed conformation.

New information on the primary structure of proteins and peptides is rapidly becoming available so that new sequences are continually in hand for molecular modeling. In some instances postulated structures are being synthesized as a final "proof of structure," although even here it is always possible that rearrangements can take place during synthetic steps. Sometimes two structures are suggested for even fairly simple peptides. For example, circulin B was first described as a cyclic decapeptide by Koffler and Kobayashi (41) but recently analytical and synthetic procedures are claimed by Vogler (42) to support a cycloheptapeptide sequence with a peptide side chain. In the total concept of protein conformation these discrepancies may be regarded as minor differences, yet they keep us alert to the possibility that similar small variations in protein sequences could drastically alter the determined primary structure.

Subunit Overlap in Protein Aggregation

At the present time one of the most challenging areas for attempting to exploit the hexagonal conformation lies in the field of the disulfide stabilized proteins. As we pointed out in the case of the TMV hexamer (Fig. 18) if it is postulated that "hydrophobic" faces are joined together by general hydrophobic and hydrogen bonded forces (but *not* covalent –S–S–bonds), one can approach their position of overlap in the study model in a somewhat arbitrary manner. However, when established –S–S– bridges between specific cysteine units are limited to definite positions in the chain sequence, then it becomes necessary to line up the overlap position of the upper and lower hexagonal subunits so that all –S–S– bonds are in the correct register. The simplest known instance involving –S–S– bridges between two chains is the beef insulin molecule, also the first protein whose primary sequence was determined (43). In a cooperative effort, Dr. Oscar Hechter and I have been studying models of beef insulin, and a preliminary proposal for the conformation of this sequence, following the general hexagonal concept, has been presented with reservations (44). In essence, insulin is treated like two separate protein subunits, chain A (21 amino acids) and chain B (30 amino acids), tied together by disulfide bridges emanating from their respective hydrophobic surfaces. Although other bonds between other side chains are also involved in hydrophobic interaction, the *position of the overlap* is specifically determined by the *covalent disulfide bridges.* Only overlap positions permitting the formation of the disulfide bridges between the A and B chain can be tolerated. Although our model did fulfill this requirement reasonably well, there were other difficulties, especially in the conformation of the A chain. Therefore, we are continuing to examine slight changes in its arrangement. A new approach using all-*trans*

amide bonds in both chains looks very promising at present, but we would like to make it with the new C–P–K model set, only recently available to us, before reaching a conclusion.

Ribonuclease

From the present knowledge of protein sequences, the disulfide-linked structure of the next degree of complexity would probably be ribonuclease (45). Its structure continues to provoke our interest, and concerted efforts to solve it with molecular models and hexagonal papers in terms of the hexagonal conformation have been made, thus far without complete success but with some hopeful glimmerings of insight. Even in the realm of primary structure, ribonuclease has been a formidable challenge as reflected in the later revision (45) of an earlier published structure by Hirs, Moore, and Stein (46). This molecule has been receiving diligent study by Harker and his group (47) using X-ray diffraction methods. The conclusions to be drawn at this stage have been presented very guardedly, but one is led to wonder if the statement (47) "There appear to be more 'cross points' in the continuous regions than can be put into correspondence with the four disulfide cross links in the ribonuclease molecule" may not presage the possibility of additional covalent linkages, perhaps involving bonds like ϵ-linked lysines at points such as Lys-61 or Lys-65 in the ribonuclease chain. We have made some models of the sequence around the small disulfide ring using one or more ϵ-linked lysines. Lacking chemical evidence for their validity, the models are not shown here, but this consideration would be in line with our previous attempt (23) to look at ribonuclease as two "half subunits" each composed of about 60–65 amino acid residues so that half of the sequence would be in hydrophobic contact with the other half of the sequence. The position of overlap would here be determined by the disulfide bonds.

Chymotrypsinogen A

Another primary structure containing several disulfide bonds is now available in the recently determined sequence of chymotrypsinogen A (48). No comprehensive attempt has been made to evaluate this structure in terms of the hexagonal conformation. However, it seems to me to be an item of more than passing interest that with a total of 246 amino acids in the chain, the joining of Cys-1 to Cys-122 may involve "setting apart" approximately the first half of the sequence from the last half of the molecule. This first segment also has another –S–S– bridge within it (Cys-42 to Cys-58). Although I have stressed –S–S bonds in our proposed models for their possible value in binding an upper "subunit" to a lower "subunit," there is of course the alternate possibility that they will serve to bind adjacent strands in the same subunit. This is the sense in which this bond is utilized in the intrachain link of the A chain of insulin and also in the case of oxytocin, so that some of the –S–S bridges in chymotrypsinogen A may be "in-plane" bonds while others would be positioned as "between-plane" bonds. However, aside from this modest effort, mostly with hexagonal papers, our main effort with the model sets is for the present still being directed toward the insulin sequence. Until this relatively simple case is resolved in terms of the hexagonal protein concept (assuming it can be), one hesitates to go on to more profound considerations. The intensive synthetic efforts in the direction of a total fabrication of the insulin molecule by Katsoyannis (49) and many others (50) should soon give us added assurance that the primary structure is indeed correct.

Recent articles by Kartha, Bello and Harker (51) and Avey et al. (52) have disclosed additional X-ray data on ribonuclease without any additional comment on the excess "cross points" in the continuous regions of the peptide chain. With regard to insulin structure, the latest report of Katsoyannis et al. (53) has furnished convincing evidence for the identity of natural and synthetic insulins.

ACKNOWLEDGMENTS

Assistance in the preparation of photographs and drawings provided by Ilan Hadar (City of Hope Medical Center) and E. Beals and R. Simonds of The Upjohn Company.

REFERENCES

1. Pauling, L., and R. B. Corey, Proc. Nat. Acad. Sci. U.S. *37*, 205, 235, 241, 251, 256, 261, 272, 282, (1951); *38*, 86 (1952).
2. Pauling, L., "The Nature of the Chemical Bond," 3rd Edition, Cornell University Press, Ithaca, New York, 1960, p. 499.
3. a) Forbes, W. F., and W. A. Mueller, Can. J. Chem. *34*, 1542–1552 (1956). b) Braude, E. A., and F. Sondheimer, J. Chem. Soc., — 3754–3766 (1955).
4. Warner, D. T., Nature *190*, 120–128 (1961).
5. Warner, D. T., J. Theoret. Biol. *1*, 514–528 (1961).
6. Berendsen, H. J. C., J. Chem. Phys. *36*, 3297–3305 (1962).
7. Warner, D. T., Nature *196*, 1055–1058 (1962).
8. Vandenheuvel, F. A., JAOCS *42*, 481–492 (1965).
9. Klotz, I. M., Federation Proc. *24*, 524–533 (1965).
10. Berendsen, H. J. C., and C. Migchelsen, Ann. N. Y. Acad. Sci. *125*, 365–379 (1965).
11. Badger, R. M., and H. Rubalcava, Proc. Nat. Acad. Sci. U.S. *40*, 12–17 (1954).
12. Dickerson, R. E., Nature *208*, 139–143 (1965).
13. Margoliash, E., J. Biol. Chem. *237*, 2161–2174 (1962).
14. Margoliash, E., N. Frohwirt, and E. Wiener, Biochem. J. *71*, 559–570 (1959).
15. Grant, E. H., Ann. N. Y. Acad. Sci. *125*, 418–427 (1965).
16. Beacham, J., V. T. Ivanov, G. W. Kenner and R. C. Sheppard, Chem. Commun. No. 16, 386–387 (1965).
17. Goto, S., and T. Isemura, Bull. Chem. Soc. Japan *37*, 1697–1701 (1964).
18. Hodgkin, D. C., G. M. J Schmidt and B. M. Oughton, Biochem. J. *65*, 744–750, 752–756 (1957).
19. Shemyakin, M. M., N. A. Aldanova, E. I. Vinogradova and M. Yu Feigina, Tetrahedron Letters —1963, 1921–1925.
20. Moore, C. and B. C. Pressman, Biochem. Biophys. Res. Commun. *15*, 562–567 (1964).
21. Shemyakin, M. M., Yu A. Ovchinnikov, V. T. Ivanov, A. A. Kiryushkin, G. L. Zhdanov and I. D. Ryabova, Experientia *19*, 566–568 (1963).
22. Bromer, W. W., L. G. Sinn, A. Staub, and O. K. Behrens, J. Am. Chem. Soc. *78*, 3858–3860 (1956).
23. Warner, D. T., J. Theoret. Biol. *6*, 118–136 (1964).
24. Paleus, S., A. Ehrenberg, and H. Tuppy, Acta Chem. Scand. *9*, 365–374 (1955).
25. Ehrenberg, A., and H. Theorell, Acta Chem. Scand. *9*, 1193–1205 (1955).
26. Williams, R. J. P., "The Enzymes," (eds: H. Lardy, P. D. Boyer and K. Myrback), 2nd Ed., Vol. *1*, Academic Press, New York, 1959. p. 412 ff.
27. Heller, J., and E. L. Smith, Proc. Nat. Acad. Sci. U.S. *54*, 1621 (1965).
28. Danielli, J. F. and H. Davson, J. Cell. Comp. Physiol. *5*, 495–508 (1935).
29. Schramm, G., Z. Naturforsch. *2b*, 112–121, 249–257 (1947).
30. Lauffer, M. A., "The Molecular Basis of Neoplasia," University of Texas Press, Austin, 1962, p. 192, Figure 5; p. 184, Table I.
31. Warner, D. T., Ann. N. Y. Acad. Sci. *125*, (Art. 2), 605–624 (1965).
32. Vegotsky, A., F. Lim, J. F. Foster and H. Koffler, Arch. Biochem. Biophys. *111*, 296–307 (1965).
33. Scott, E., and D. S. Berns, Biochem. *4*, 2597–2605 (1965).
34. Hechter, O., Federation Proc. *24*, S-91 to S-102 (1965).
35. Bachmann, H. G., and W. Muller, Nature *201*, 261–264 (1964).
36. Perutz, M. F., Nature *201*, 814 (1964).
37. Palmer, A. T., R. A. Palmer and R. E. Dickerson, Nature *202*, 1052–1057 (1964).
38. Donohue, J., and K. N. Trueblood, Acta. Cryst. *5*, 419–431 (1952).
39. Abraham, R. J., K. A. McLauchlan, S. Dalby, G. W. Kenner and R. C. Sheppard, Nature *192*, 1150–1153 (1961).
40. Karle, I. L., and J. Karle, Acta Cryst. *16*, 969–975 (1963).
41. Koffler, H., and T. Kobayashi, Federation Proc. *19*, 342 (1960).
42. Vogler, K., Angew. Chem. *77*, 974–975 (1965).
43. Sanger, F., Bull. Soc. Chem. Biol. *37*, 23–35 (1955).
44. Hechter, O., in "Mechanisms of Hormone Action" (ed. P. Karlson), Georg Thieme Verlag, Stuttgart, 1965, and Academic Press, New York, 1965, p. 80–81.
45. Smyth, D. G., W. H. Stein and S. Moore, J. Biol. Chem. *238*, 227–234 (1963).
46. Hirs, C. H. W., S. Moore and W. H. Stein, J. Biol. Chem. *235*, 633–647 (1960).
47. Harker, D., J. Bello, D. R. Harris, H. H. Mills, R. Parthasarathy and F. E. DeJarnette, in "Structure and Activity of Enzymes," (eds. T. W. Goodwin, J. I. Harris and B. S. Hartley), Academic Press, New York, 1964, p. 15–17.
48. a) Keil, B., S. Prusik and F. Sorm, Biochem. Biophys. Acta *78*, 559–561. (1961). b) Brown, J. R., and B. S. Hartley, Biochem. J. *89*, 59P–60P (1963). c) Hartley, B. S., see Reference (47), p. 52.
49. Katsoyannis, P. G., Metab. Clin. Exptl. *13*. 1059–1081 (1964).
50. Zahn, H., O. Brinkhoff, J. Meienhofer, E. F. Pfeiffer, H. Ditschuneit and Ch. Gloxhuber, Z. Naturforsch. *20b*, 666–670 (1965).
51. Kartha, G., J. Bello and D. Harker, Nature *213*, 862–865 (1967).
52. Avey, H. P., M. O. Boles, C. H. Carlisle, S. A. Evans, S. J. Morris, R. A. Palmer and B. A. Woolhouse, Nature *213*, 557–562 (1967).
53. Katsoyannis, P. G., A. C. Trakatellis, C. Zalut, S. Johnson, A. Tometsko, G. Schwartz and J. Ginos, Biochem. *6*, 2656–2668 (1967).