Bull. Inst. Chem. Res., Kyoto Univ., Vol. 66, No. 4, 1989

Review

Computational Studies of the Structure and Assembly of Triple-Stranded Models of Collagen

George NÉMETHY and Harold A. SCHERAGA*

Received August 16, 1988

Conformational energy computations on the structure of collagen-like poly(tripeptide)s at various levels of structural complexity, carried out mostly in the laboratory of the authors, are reviewed. Poly-(Gly-Pro-Pro) was used as the main model for collagen, although other sequences have also been investigated. An exhaustive study of regular single- and triple-stranded conformations has been possible because of the small number of independent variables. The proposed role of β-bend formation in single strands in post-translational prolyl and lysyl hydroxylation has been confirmed. The computed triple-stranded structure with the most favorable energy agrees closely with an observed single crystal structure and with proposed models derived from fiber diffraction studies on collagen. This structure is favored even in the presence of hydration. Computations on the packing of triple helices have indicated that specific residue-residue interactions are crucial in stabilizing the observed parallel packing of triple helices in microfibrils. The computations have predicted that hydrogen bonds involving hydroxyl groups of Hyp residues stabilize the aggregation of triple helices. This series of studies confirms the working principle that complicated protein structures can be investigated by considering successively structural elements of increasing levels of complexity.

KEY WORDS: Collagen structure/ Conformational energy computations/ Protein conformation/

I. INTRODUCTION

Theoretical analysis of protein conformations results in an improved understanding of the interactions that determine protein folding, and it can lead to predictions concerning the structure and stability of proteins.[1,2] The scope of theoretical studies has expanded over the last 25 years from oligopeptides to entire protein molecules and to protein-protein association. Several significant steps of progress in our laboratory have been achieved in collaboration with Professor Tatsuo Ooi, ranging from computations on the stability of α-helices[3,4] to the analysis of the flexibility of a protein molecule[5] and to a novel treatment of the thermodynamics of hydration of proteins.[6] It gives us great pleasure to dedicate this review to Professor Tatsuo Ooi on the occasion of his retirement.

We summarize here the main results of the ongoing investigation in our laboratory of one particular protein, collagen. The conformational energy computations have been carried out at various levels of structural complexity. The work on collagen serves as a good illustrative example of the potential and scope of conforma-

* To whom requests for reprints should be addressed.
Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853–1301, U.S.A.

tional energy computations.

The structure of collagen can be described as a hierarchical order of self-assembly of repeating structures at various levels. The smallest such structural unit occurs at the level of the covalent structure, corresponding to the Gly-X-Y tripeptide. The regular repetition of the conformation of this unit along a polypeptide chain constitutes the regular helical structure of the individual polypeptide strands. Three helical strands assemble into the triple-helical collagen molecule. The molecules associate laterally and longitudinally to form microfibrils. These, in turn, form fibrils which continue to aggregate and constitute various forms of connective tissue. The reader is referred to recent reviews for detailed descriptions of the structure of collagen[7-9] and, in particular, of the energetics and thermodynamics of self-assembly[10].

The sequence of our conformational energy computations, summarized here, has been designed to follow this hierarchical structure. It proceeded from the conformational analysis of oligopeptides to that of higher-level structures. So far, it has reached the level of pairs of packed triple helices, serving as a model for microfibril formation. Starting from the known conformational properties of individual amino acid residues[11,12] (blocked at both termini to represent a residue in the interior of a polypeptide chain), we have analyzed first the conformational preferences of Gly-X-Y tripeptides and of their oligomers.[13-16] Based on these results, the preferred regular helical structures for single-chain $(Gly\text{-}X\text{-}Y)_n$ polypeptides were deduced. These, in turn, formed the basis for determining the energetically most favored ways of assembling three regular chains, according to various symmetry criteria.[13-16] The conformational flexibility of various amino acid side chains in positions X and Y[17] and the rigidity of the triple helix as a function of its sequence[18] have also been analyzed.

Following this, the energetically most favorable ways of packing two triple helices have been computed,[19,20] and the important role of the hydroxyprolyl residue in stabilizing the packing has been pointed out.[21] The effects of hydration on the stability of the structures have also been considered.[22]

Short non-triple-helical sequences (the telopeptides) that occur at both ends of the molecule will not be considered in this review, even though they also play an important role in collagen self-assembly.[7-9]

Poly(Gly-Pro-Pro) is the simplest polypeptide sequence resembling collagen. Because of the frequent occurrence of Pro and Hyp in positions X and Y, respectively, any model for the conformation of the collagen molecule and for higher levels of assembly must be consistent with the conformational constraints inherent in the Gly-Pro-Pro or Gly-Pro-Hyp sequence. Low-energy conformations for these sequences allow the substitution of any other amino acid in positions X and Y, corresponding to the natural collagen sequence, because any conformation that is permitted for Pro (or Hyp) is accessible for any other residue as well, but the converse is not necessarily true. For this reason, poly(Gly-Pro-Pro) is used frequently as a model in experimental and theoretical studies of the conformational properties of collagen.[10,13] Every stage of our computations summarized here was first carried

out on the (Gly-Pro-Pro) tripeptide or on poly(Gly-Pro-Pro), and these models were then applied to other sequences as well.

Throughout this work, atomic coordinates and energies were computed with the original[23] (ECEPP) or revised[24] (ECEPP/2) version of the "Empirical Conformational Energy Program for Peptides." A recently developed hydration shell model[25,26] was used to compute the free energy of hydration. Interactions between triple helices were computed by a general method developed for the analysis of packing of helices.[27]

II. SINGLE-STRANDED GLY-X-Y POLYPEPTIDES

A. Regular Conformations of Poly(Gly-Pro-Pro) and Poly(Gly-Pro-Hyp)

Miller and Scheraga have carried out an exhaustive search for all regular low-energy conformations of the $(Gly\text{-}Pro\text{-}Pro)_4$ molecule.[13] Only 80 conformations were found within an energy range of 15.0 kcal/mol of the Gly-X-Y repeat unit. 26 of them fell within an energy range of 3 kcal/mol of repeat unit. The exhaustive computation was made feasible by the following special features of the problem. The allowed conformational space of the Gly-Pro-Pro tripeptide is strongly limited. Because of the constraint of the ring, there exist only five minimum-energy backbone conformations for Pro (or Hyp) residues, even though cis-trans isomerization of the preceding peptide bond can occur, in contrast to other residues for which the cis form has a high energy.[11,12] Furthermore, the conformation of the residue preceding Pro is also limited, because of steric constraints.[28] Therefore, a complete search of conformational space was possible for di- and tripeptides.

A randomly coiled poly(Gly-Pro-Pro) chain still would possess an enormous number of conformations. On the other hand, for the purposes of analyzing the assembly of the chains into triple-stranded molecules (Sec. III), only regular chain conformations had to be considered, in which all Gly-X-Y units have the same conformation.[13] Therefore, all possible conformations of the Gly-Pro-Pro tripeptides were computed and inserted into a regular chain structure. The latter step allowed the elimination of a few tripeptide conformations that could not be extended into a regular polymer structure, because of steric hindrance between successive tripeptides.

The results showed that the poly(Gly-Pro-Pro) molecule has considerable flexibility, in spite of the severe constraints caused by the Pro residues, and that numerous low-energy single-chain conformations that contain cis peptide bonds exist.[13]

Recently, the free energy contribution of hydration was incorporated into the computation,[22] using a newly developed hydration shell model.[25,26] This contribution does not result in an alteration of the dihedral angles of the minimum-energy conformations, but it reduces the range of the free energy differences between various conformations by a factor of two. Thus, the presence of hydration enhances the flexibility of the ensemble of single chains.

The dihedral angles and relative energies of the single-strand conformations of $(Gly\text{-}Pro\text{-}Pro)_4$ and of $(Gly\text{-}Pro\text{-}Hyp)_4$ are identical, both in the absence and presence of hydration.[14,22] The free energy of hydration for a given conformation of a Gly-

Pro-Hyp peptide is about 4.5 kcal/mol of GXY unit lower than for the corresponding (Gly-Pro-Pro) conformation, as a result of the favorable hydration of the 4-Hyp residue, but this does not affect the conformational distribution.

B. Regular Conformations with Other Sequences

Similar computations have been carried out for other sequences that can serve as general models of collagen, viz. for $(Gly\text{-}Pro\text{-}Ala)_4$ and $(Gly\text{-}Ala\text{-}Pro)_4$, in which a non-imino acid is substituted in positions Y or X, respectively.[15,16] The poly(tripeptide) conformations were generated by combining all minimum-energy conformations of the constituent dipeptides.[28] The number of regular low-energy single-chain structures is roughly doubled, as a result of the increased flexibility of Ala, as compared with Pro.

C. Post-translational Hydroxylation and β-bend Formation

Our computations have demonstrated that there is a very high probability of occurrence of a β-bend at the Pro-Gly position in the Pro-Pro-Gly-Pro tetrapeptide and that the β-bend can also be accommodated in single-stranded poly(Gly-Pro-Pro), while the probability of β-bend formation is very low at the Gly-Pro position.[29] This result supports the hypothesis[30] that a β-bend at Pro-Gly is required for post-translational hydroxylation of Pro during the biosynthesis of collagen, before assembly of the chains into the triple-helical structure. It also explains the observed specificity for hydroxylation of Pro in position Y, to the exclusion of position X. Furthermore, similar computations have shown that a hydrogen bond between the OH of Hyp and the C=O of the preceding Gly residue stabilizes the β-bend conformation as well as a partially extended conformation that may be a kinetic intermediate in the formation of the triple helix.[31]

It has also been suggested[30] that a β-bend may play a role in the preferential post-translational hydroxylation of Lys residues in position Y. Computations on Lys-containing tetrapeptides have indicated that they exist as an ensemble of many conformations, but that a β-bend is one of the favored conformations, and the type of β-bend formed depends on the position (X or Y) of the Lys residue.[32]

III. STRUCTURE AND STABILITY OF THE TRIPLE HELIX

A. Three-stranded Poly(Gly-Pro-Pro) Structures

Three equivalent regular poly(Gly-X-Y) chains can be assembled into a three-chain structure in various ways, corresponding to various symmetry arrangements.[13,14] These include coiled coils (forming a triple helix) with screw symmetry and parallel-chain complexes with either screw or rotational symmetry. In principle, coiled coils with rotational symmetry would also be possible, but they are sterically less satisfactory.

The energy has been computed for all possible ways of arranging three identical $(Gly\text{-}Pro\text{-}Pro)_4$ chains, starting from the computed low energy conformations (Sec. II. A) and minimizing the energy of the three-chain structure.[13] The structures considered included the various symmetry arrangements mentioned. Coiled-coil

structures generally had much lower energies than parallel-chain arrangements. The energetically most favorable structure is a coiled-coil triple helix with all-trans peptide bonds. Its helical parameters are close to the models derived for collagens from fiber X-ray diffraction measurements.[33-36] The atomic coordinates of the computed structure agree with those obtained subsequently from a single-crystal X-ray study[37] of (Pro-Pro-Gly)$_{10}$ to within an r.m.s. deviation of 0.3 Å. The total energy of this structure is at least 4.2 kcal/mol of Gly-Pro-Pro unit lower than that of any other triple-stranded arrangement, i.e. it is strongly stabilized over other, hypothetical structures. The inclusion of hydration reduces the free energy difference between various structures, but the collagen-like triple helix is still more stable than any other structure, by at least 2.0 kcal/mol of Gly-Pro-Pro unit.[22] Other calculations, using ECEPP, confirmed that the structure with triple-helical parameters closest to collagen is energetically most stable.[36,38] There is a slight improvement in the fitting of the computed structure to the single-crystal structure if flexibility (i.e. puckering) of the prolyl ring is allowed.

B. Other Three-stranded Poly(tripeptide) Structures

Identical conclusions have been reached for poly(Gly-Pro-Hyp)[14] and for poly (Gly-Pro-Ala)[15] as for poly(Gly-Pro-Pro). Essentially the same collagen-like triple helix has the lowest energy for all three poly(tripeptide)s, with a somewhat larger value of the angular repeat for the poly(Gly-Pro-Ala) structure. For poly(Gly-Ala-Pro), on the other hand, several packing arrangements have comparable energies, including triple-stranded coiled coils and parallel-strand structures.[16] This agrees with film diffraction data for the latter polypeptide which showed that films with both kinds of symmetries can occur, depending on the solvent used to prepare the film.[16,39] The result indicates that the balance between several structures with similar energies can be shifted by interactions with the solvent.

Tumanyan et al. have carried out computations on all four of these poly(tripeptide)s, taking the flexibility of the prolyl ring into account.[40,41] Their conclusions with regard to the structure of the most favorable triple helical structures agree with those described above. With the use of flexible proline, the computed helical parameters for poly(Gly-Pro-Pro) and poly(Gly-Pro-Ala) come even closer to each other. Poly(Gly-Ala-Pro) and poly(Gly-Ala-Ala) were found to be similar to each other and somewhat more flexible than the other structures, in agreement with the result for poly(Gly-Ala-Pro), cited above.

The presence of 4-Hyp instead of Pro in position Y enhances the thermal stability of the triple helix both in collagen and in synthetic model poly(tripeptide)s.[42-44] The stabilization cannot be attributed to direct interactions, such as hydrogen bond formation, between the hydroxyl group of Hyp and backbone atoms of the triple helix, as demonstrated by both conformational energy computations[14] and the stereochemical consideration of models.[45,46] The mere presence of water around the triple helix, expressed in terms of a hydration-shell model, does not provide an explanation of the stabilization.[22] On the basis of the geometry of the triple helix, several models have been proposed that demonstrate that a water molecule, *specifically*

hydrogen-bonded to the hydroxyl group, could form a bridge to backbone carbonyl groups.[45-47] We plan to investigate the energetics of such localized interactions with water.

On the other hand, Hyp in position X decreases the thermal stability of the triple helix.[48] No explanation is available for this effect. The possible role of hydration of poly(Gly-Hyp-Pro) is also to be studied.

C. The Effect of Local Substitutions

Most of the conformational energy studies cited have been carried out on repeating poly(tripeptide)s. Substitution of individual residues at particular positions in a poly(Gly-Pro-Pro) or poly(Gly-Pro-Ala) triple helix could conceivably have a destabilizing effect because of changes in the residue geometry (bond lengths and bond angles), steric effects, or altered non-covalent interactions. Acutally, however, it was found that changes in the helix geometry and in intra- and interchain energies are very small when such substitutions are made.[18] As a consequence, it is possible to use the poly(tripeptide) models for the study of the general physical properties of triple-helical collagen. Computations with flexible prolyl ring geometry corroborate this conclusion.[41]

D. Conformational Preferences of Side Chains

We have shown that the placement of a residue (into position X or Y) and the nature of neighboring residues in the same and neighboring strands may have a profound effect on the conformational freedom of side chains in the triple helix.[17] The presence of a bulky neighbor (especially an imino acid) can reduce the number of allowed conformations of the side chain and alter the relative energies of the remaining ones. Intra- and interchain interactions are strongly sequence-dependent. Thus, side chains in position X interact sterically with an imino acid in position Y of the same strand, so that the presence of Pro (or Hyp) in position Y restricts their conformational freedom. Their interactions with the other strands are negligible, however. The opposite is seen for side chains in position Y. Their conformation is not influenced by the nature of residues in position X in their own strand, but they are strongly constrained by the presence of Pro in the neighboring strand. Although every amino acid residue can be accommodated in any position in the triple helix, large differences exist in the conformational freedom of various side chains, depending on their placement and on the nature of their neighbors. Observed preferences[7] of some amino acids for positions X or Y can be correlated with these differences, for example the preference of Leu and Phe for position X or Thr for position Y.

These preferences were also seen in model studies[46,49] and also in stereochemical (hard-sphere) computations,[50,51]. A quantum-mechanical study of collagen-like tripeptides reached similar conclusions about the constraining role of Pro in position Y and about the preference of Phe and Ile for position X, from the comparison of side-chain conformational freedom in various single-stranded tripeptide sequences.[52,53]

E. The Entropy of Local Destabilization of the Triple Helix

There is recent evidence that sequence-dependent transient local destabilization of the triple helix may occur at particular locations along the natural collagen sequence.[10] Most of this evidence has come from studies of proteolytic cleavage.[54-57] Collagenase cleaves collagen at a Gly-Ile or Gly-Leu peptide bond at a specific location, in a region of 12 amino acids not containing Pro and Hyp, but not at other Gly-Ile peptides along the sequence. Several other proteolytic enzymes also act specifically in the same region.[54] It has also been suggested that the triple helix in two other regions of similar length, not containing Pro and Hyp, may be relatively less stable, and that their flexibility plays a role in fibrillogenesis in vitro.[58]

Existing theories for the triple helix-coil transition either deal with short triple helices[59,60] in which the unfolding can be assumed to occur in a zipper-like fashion from either end of the helix, without internal loop formation, or apply only in the limit of an infinitely long chain.[61] The latter study has indicated that the entropy of formation of loops may make a significant contribution to the heat capacity of collagen, suggesting that loops are not to be neglected in collagen.

In a preliminary analysis of the formation of small internal loops, we have carried out a computer simulation study to estimate the effect of ring closure and of excluded volume on the entropy of loop formation in a triple helix.[62] For chain lengths of about 30 residues, the probability of internal loop formation is only about 10 times lower than the probability of unwinding from the end of the triple helix. The possibility of internal loop formation must, therefore, be considered in detailed analyses of the stability of collagen.

IV. THE PACKING OF TRIPLE HELICES

In fibrils of type I collagen, the triple-stranded molecules associate into a parallel array.[7-9] The arrangement in the axial ordering is well established.[63] Neighboring molecules are staggered longitudinally by an axial translation of D=670 Å, which corresponds to 234 or 235 residues or 1/4.4 of the length of each triple helix. The precise mode of lateral organization is less well understood[8] and several arrangements have been proposed. According to one of the most favored proposals, bundles of five triple helices form microfibrils, which in turn aggregate to fibrils.[64] In another view, a quasi-crystalline, near-hexagonal array of slightly tilted molecules has been proposed.[65,66] In both models, the rod-like triple helices pack fairly tightly against their lateral neighbors.

As the initial step in the analysis of the energetics of the aggregation of collagen, we have carried out conformational energy computations to establish the most favorable ways of packing two collagen-like poly(tripeptide) triple helices.[20,21] Later, the computations will be extended to higher aggregates.

The relative position and orientation of two rigid bodies (e.g. two triple helices) can be defined in terms of six variables.[27,67] Three of them describe the position in terms of the components of the vector connecting the midpoints (or appropriate reference points) of the two helices. The orientation can be specified by three

Euler angles of rotation. In the present context, two of these angles (α and γ) represent rotations of the two helices about their own axes, while the third one (β) is the angle of relative orientation of the helix axes. Among these, the Euler angle β (or the equivalent, but coordinate-independent orientation angle Ω_0 derived from it) is the most significant parameter for the description of the packing.[67] It ranges from 0° for parallel helices to $\pm$180° for antiparallel helices. Because of the screw symmetry of collagen-like poly(tripeptide) triple helices, several combinations of the other five variables (with the same Ω_0) can correspond to essentially the same set of interactions between the two helices.

A. The Packing of Poly(tripeptide) Triple Helices

We have analyzed the modes of packing of two $(Gly\text{-}Pro\text{-}Pro)_n$ triple helices of various lengths, with n=3, 4, nd 5.[20] For very short helices (n=3), many different packing arrangements with similar energies can be found, and Ω_0 can take a wide range of values. On the other hand, when the helices are longer (n=5), a very strong energetic preference is seen for a family of near-parallel packing arrangements, with $\Omega_0=-10°$, to the exclusion of other packings, including the antiparallel ones. This result agrees with the observed parallel orientation in fibrils of type I collagen.[7,66] Fiber X-ray diffraction data on collagen indicate that the molecules are tilted at a small angle, about 5°, relative to the axis of the microfibril.[66] This tilting is consistent with the computed angle of orientation.

The packing of several triple helices would also require them to be near-parallel or near-antiparallel, because of geometrical constraints on the close packing of rods (as seen, e.g. from the theory of crystallization of polymer rods[68]), but these constraints would not be able to differentiate between the two orientations. The observed preference for near-parallel packings is, therefore, an energetic consequence of specific residue-residue interactions between the triple helices and in particular of the presence of many imino acid residues in position Y.

The latter conclusion has been corroborated by computations on the packing of two poly(Gly-Pro-Ala) triple helices.[20] In addition to near-parallel arrangements that are similar to those computed for poly (Gly-Pro-Pro), some low-energy near-antiparallel arrangements have been found. Therefore, it appears that preference for parallel packing requires the frequent occurrence of imino acids.

B. Stabilization of the Packing by Hyp

Repetition of the computations[21] with $(Gly\text{-}Pro\text{-}Hyp)_5$ indicated that the same near-parallel packing arrangement is preferred as for $(Gly\text{-}Pro\text{-}Pro)_5$. The stability of this arrangement is enhanced, however, because the OH groups of Hyp in position Y can be accommodated in the space between the two triple helices without any steric hindrance and they can form weak hydrogen bonds with a carbonyl oxygen of the neighboring triple helix. The interaction energy between the two triple helices becomes more favorable by 1.9 kcal/mol per Hyp residue. The possibility of such a hydrogen bond has been proposed on the basis of model construction[69] and studies of the proteolytic susceptibility of collagen fibers.[70] Our computation

represents the first demonstration that Hyp can make a direct energy contribution to the stabilization of collagen microfibrils.

This prediction is an important new finding because it sheds light on the role of hydroxyproline in collagen assembly. As discussed above, the observed thermal stabilization of the isolated triple helix by Hyp in position X is not due to direct interactions of OH groups of Hyp with atoms of the triple helix, and hence it must be mediated by interactions with water. In contrast, the contribution of Hyp to the stabilization of the assembly of triple helices arises from direct interactions involving Hyp. The presence of Hyp in position Y is known to stabilize collagen fibrils, as seen from the elevation of the shrinkage temperature of tendons.[44] The computation reported here provides an explanation of these experimental results.

In the case of $(Gly\text{-}Hyp\text{-}Pro)_5$, the packing energy is weaker because of unfavorable steric interactions.[21] Therefore, the calculations predict that the substitution of Hyp for Pro in position X should decrease the stability of collagen aggregates.

V. CONCLUSIONS

Conformational energy computations have elucidated the energetic reasons for many observed features of collagen structure at various levels. They have led to new predictions, such as the role of Hyp in stabilizing fibril structure. The analysis of the packing of triple helices represents the first application of conformational energy computations to the energetics of a supramolecular assembly in fibrous proteins, and it shows the way for the further analysis of collagen assembly.

In addition, the work reported here is an example that confirms and extends a working principle that has been proposed earlier,[67] viz. that it is possible to account for many features of packing arrangements and other higher-level assemblies of polypeptide structures in terms of the conformational properties of the component structures without requiring the inclusion of all long-range interactions at all stages. The principle can be extended to account for both intra- and intermolecular interactions.

ACKNOWLEDGMENTS

This work was supported by research grants from the National Institute on Aging (AG-00322) and from the National Institute of General Medical Sciences (GM-14312) of the National Institutes of Health, and from the National Science Foundation (DMB84–01811).

REFERENCES

(1) H.A. Scheraga, *Carlsberg Res. Commun.* **49**, 1 (1984).
(2) H.A. Scheraga, *Israel J. Chem.* **27**, 144 (1986).
(3) T. Ooi, R.A. Scott, G. Vanderkooi, and H.A. Scheraga, *J. Chem. Phys.* **46**, 4410 (1967).
(4) H.A. Scheraga, R.A. Scott, G. Vanderkooi, S.J. Leach, K.D. Gibson, T. Ooi and G. Némethy, in "Conformation of Biopolymers", ed. by G.N. Ramachandran, Academic Press, London and New York, 1967, p. 43.

(5) T. Ooi, K. Nishikawa, M. Oobatake and H.A. Scheraga, *Biochim. Biophys. Acta* **536**, 390 (1978).
(6) T. Ooi, M. Oobatake, G. Némethy, and H.A. Scheraga, *Proc. Natl. Acad. Sci., U.S.A.* **84**, 3086 (1987).
(7) P. Bornstein and W. Traub, in "The Proteins", ed. by H. Neurath and R.L. Hill, 3rd ed., Vol. IV, Academic Press, New York, 1979, p. 411.
(8) K.A. Piez, A.H. Reddi, eds., "Extracellular Matrix Biochemistry," Elsevier, New York, 1984.
(9) M.E. Nimni, ed., "Collagen," Vol. I-III, CRC Press, Boca Raton, 1988.
(10) G. Némethy, in "Collagen," ed. by M.E. Nimni, Vol. I, CRC Press, Boca Raton, 1988, p. 79.
(11) S.S. Zimmerman, M.S. Pottle, G. Némethy, and H.A. Scheraga, *Macromolecules* **10**, 1 (1977).
(12) M. Vásquez, G. Némethy, and H.A. Scheraga, *Macromolecules* **16**, 1043 (1983).
(13) M.H. Miller and H.A. Scheraga, *J. Polymer Sci. (Polymer Symposia)*, **54**, 171 (1976).
(14) M.H. Miller, G. Némethy and H.A. Scheraga, *Macromolecules* **13**, 470 (1980).
(15) M.H. Miller, G. Némethy, and H.A. Scheraga, *Macromolecules* **13**, 910 (1980).
(16) G. Némethy, M.H. Miller, and H.A. Scheraga, *Macromolecules* **13**, 914 (1980).
(17) G. Némethy and H.A. Scheraga, *Biopolymers* **21**, 1535 (1982).
(18) G. Némethy, *Biochimie* **63**, 125 (1981).
(19) G. Némethy, *Biopolymers* **22**, 33 (1983).
(20) G. Némethy and H.A. Scheraga, *Biopolymers* **23**, 2781 (1984).
(21) G. Némethy and H.A. Scheraga, *Biochemistry* **25**, 3184 (1986).
(22) G. Némethy and H.A. Scheraga, *Biopolymers*, in press.
(23) F.A. Momany, R.F. McGuire, A.W. Burgess, and H.A. Scheraga, *J. Phys. Chem.* **79**, 2361 (1975).
(24) G. Némethy, M.S. Pottle, and H.A. Scheraga, *J. Phys. Chem.* **87**, 1883 (1983).
(25) Y.K. Kang, G. Némethy and H.A. Scheraga, *J. Phys. Chem.* **91**, 4105, 4109, 4118 (1987).
(26) Y.K. Kang, K.D. Gibson, G. Némethy, and H.A. Scheraga, *J. Phys. Chem.*, **92**, 4739 (1988).
(27) K.C. Chou, G. Némethy, and H.A. Scheraga, *J. Am. Chem. Soc.* **106**, 3161 (1984).
(28) S.S. Zimmerman and H.A. Scheraga, *Biopolymers* **16**, 811 (1977).
(29) E. Lee, G. Némethy, H.A. Scheraga, and V.S. Ananthanarayanan, *Biopolymers* **23**, 1193 (1984).
(30) S.K. Brahmachari and V.S. Ananthanarayanan, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5119 (1979).
(31) M. Bansal and V.S. Ananthanarayanan, *Biopolymers* **27**, 299 (1988).
(32) K. Takahashi, Y.K. Kang, G. Némethy, H.A. Scheraga, and V.S. Ananthanarayanan, *Biopolymers* **26**, 1781 (1987).
(33) G.N. Ramachandran and G. Kartha, *Nature* **174**, 269 (1954).
(34) A. Rich and F.H.C. Crick, *Nature* **176**, 915 (1955).
(35) A. Yonath and W. Traub, *J. Mol. Biol.* **43**, 461 (1969).
(36) R.D.B. Fraser, T.P. MacRae, and E. Suzuki, *J. Mol. Biol.* **129**, 463 (1979).
(37) K. Okuyama, N. Tanaka, T. Ashida, and M. Kakudo, *Bull. Chem. Soc. Japan* **49**, 1805 (1976).
(38) K. Okuyama, H. Eto, I. Nakaya, and M. Takayanagi, *Bull. Chem. Soc. Japan* **53**, 913 (1980).
(39) B.B. Doyle, W. Traub, G.P. Lorenzi, and E.R. Blout, *Biochemistry* **10**, 3052 (1971).
(40) V.G. Tumanyan and N.G. Esipova, *Biopolymers* **21**, 475 (1982).
(41) V.G. Tumanyan, R.A. Abagyan, and N.G. Esipova, *Biopolymers* **23**, 1499 (1984).
(42) J. Rosenbloom, M. Harsch, and S. Jimenez, *Arch. Biochem. Biophys.* **158**, 478 (1973).
(43) R.A. Berg and D.J. Prockop, *Biochem. Biophys. Res. Commun.* **52**, 115 (1973).
(44) T.V. Burjanadze, *Biopolymers* **21**, 1489 (1982).
(45) G.N. Ramachandran, M. Bansal, and R.S. Bhatnagar, *Biochim. Biophys. Acta* **322**, 166 (1973).
(46) W. Traub, *Israel J. Chem.* **12**, 435 (1974).
(47) E. Suzuki, R.D.B. Fraser, and T.P. MacRae, *Int. J. Biol. Macromol.* **2**, 54 (1980).
(48) K. Inouye, Y. Kobayashi, Y. Kyogoku, Y. Kishida, S. Sakakibara, and D.J. Prockop, *Arch. Biochem. Biophys.* **219**, 198 (1982).
(49) G. Salem and W. Traub, *FEBS Letters* **51**, 94 (1975).
(50) M. Bansal, *Int. J. Peptide Protein Res.* **9**, 224 (1977).
(51) M. Bansal and G.N. Ramachandran, *Int. J. Peptide Protein Res.* **11**, 73 (1978).
(52) D. Cabrol, H. Broch, and D. Vasilescu, *Biopolymers* **20**, 269 (1981).
(53) A.M. Tamburro, V. Guantieri, D. Cabrol, H. Broch, and D. Vasilescu, *Int. J. Peptide Protein Res.* **24**, 627 (1984).
(54) E.J. Miller, J.E. Finch, Jr., E. Chung, and W.T. Butler, *Arch. Biochem. Biophys.* **173**, 631 (1976).

(55) E.D. Harris, Jr. and E.C. Cartwright, Mammalian collagenases, in *Res. Monographs in Cell and Tissue Physiology*, ed. by J.T. Dingle, Vol. 2, Elsevier/North Holland Biomed. Press, Amsterdam, 1977, p. 249.
(56) L. Ryhänen, E.J. Zaragoza, and J. Uitto, *Arch. Biochem. Biophys.* **223**, 562 (1983).
(57) M.F. French, K.A. Mookhtiar, and H.E. Van Wart, *Biochemistry* **26**, 681 (1987).
(58) D.L. Helseth, Jr. and A. Veis, *J. Biol. Chem.* **256**, 7118 (1981).
(59) M. Schwarz, Jr. and D. Poland, *Biopolymers* **13**, 687 (1974).
(60) B.R. Shaw and M.J. Schurr, *Biopolymers* **14**, 1951 (1975).
(61) F.Y. Chen and H.W. Huang, *Macromolecules* **14**, 332 (1981).
(62) W.L. Mattice, G. Némethy and H.A. Scheraga, *Macromolecules*, **21**, 2811 (1988).
(63) A.J. Hodge and J.A. Petruska, in "Aspects of Protein Structure", ed. by G.N. Ramachandran, Academic Press, New York, 1963, p. 289.
(64) J.W. Smith, *Nature* **219**, 157 (1968).
(65) R.D.B. Fraser and T.P. MacRae, *J. Mol. Biol.* **127**, 129 (1979).
(66) D.J.S. Hulmes and A. Miller, *Nature* **282**, 878 (1979).
(67) H.A. Scheraga, K.-C. Chou, and G. Némethy, in "Conformation in Biology," ed. by R. Srinivasan and R.H. Sarma, Adenine Press, New York, 1983, p. 1.
(68) D.G. Wallace and A. Thompson, *Biopolymers* **22**, 1793 (1983).
(69) M. Bansal, C. Ramakrishnan, and G.N. Ramachandran, *Proc. Ind. Acad. Sci.* **A82**, 152 (1975).
(70) D. Fujimoto, *Biochim. Biophys. Acta* **154**, 183 (1968).