Electrostatic Effects on Polyproline I–II Transitions*

G. Holzwarth and K. Backman†

ABSTRACT: The electrostatic energy of an array of peptide dipoles in the polyproline I and polyproline II geometries has been calculated. The energy per peptide is markedly dependent upon chain length, *i.e.*, large unfavorable end effects occur. For dielectric constant unity, each end of a form I helix contributes +7 kcal/mole, whereas each end of form II contributes only +2 kcal/mole. This may explain the observed preference of short polyproline chains for the form II geometry. The corresponding excess electrostatic energies of a junction between form I and form II, and between form II and form I, are +4 and +7 kcal per mole for dielectric constant unity. The observed cooperative nature of the polyproline I–II transition thus originates in part in the electrostatic interactions of the peptide groups.

Poly-L-proline is known to undergo a reversible cooperative conformational transition from a right-handed helix, designated form I, to a left-handed helix of different geometry, designated form II, in certain solvent systems (Katchalski *et al.*, 1963; Mandelkern, 1967). In form I, the peptide groups adopt a *cis* conformation whereas in form II the more usual *trans* peptide group is found. A detailed analysis of the factors determining the conformation of this synthetic polymer is not only of intrinsic physical–chemical interest but may also increase our understanding of the molecular basis of conformation in biological macromolecules. This is especially true because low molecular weight polyproline, like some proteins, can exist in only two conformational states. Earlier studies have concluded that the observed secondary structure of polyproline, as well as the nature of the cooperativity, originate from repulsive steric interactions and attractive van der Waals forces (De Santis *et al.*, 1965; Schimmel and Flory, 1967; Rifkind and Applequist, 1968). Intramolecular hydrogen bonds cannot occur because of the lack of an amide hydrogen; electrostatic interactions have not been examined in detail.

The central features of the present paper are (a) a calculation of the intramolecular electrostatic interaction of the highly polar peptide groups of polyproline I and polyproline II as a function of the length of the polymer chain and (b) a calculation of the electrostatic energies of the two types of molecules in which one half is form I and the other half form II. These four different structures are shown diagrammatically in Figure 1. The importance of dipole–dipole electrostatic interactions to polypeptide conformation has been recognized only recently (Arridge and Cannon, 1964; Arridge, 1965; Brant and Flory, 1965a,b; Gibson and Scheraga, 1966; Ptitsyn, 1967) and the calculations are in part similar to those reported earlier by Arridge and Cannon for the α helix and the β form (Arridge and Cannon, 1964; Arridge, 1965).

The method of calculation is given in the paragraph

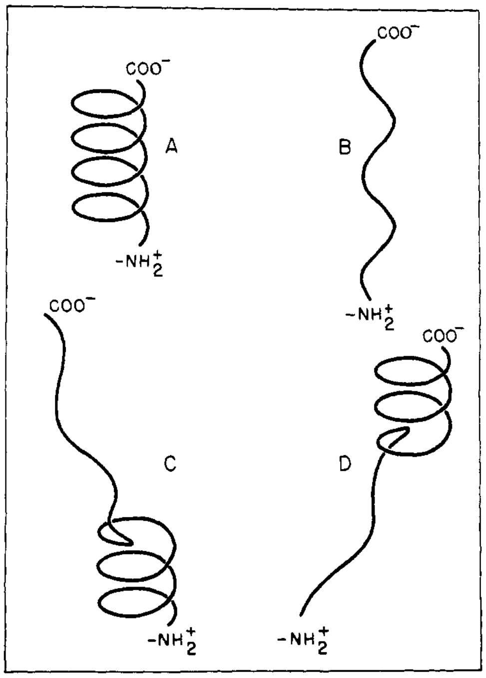

FIGURE 1: Diagram of the four structures for which calculations were performed. A represents form I; B represents form II; C represents a molecule with the amino-half form I and the carboxylic-half form II; D represents a molecule whose amino-half is in form II and whose carboxylic-half is form I. C contains a I–II junction, whereas D contains a II–I junction.

---

* From the Department of Biophysics and the Department of Chemistry, The University of Chicago, Chicago, Illinois 60637. *Received October 11, 1968.* Supported by U. S. Public Health Service Grant No. NB 07286 from the National Institutes of Health.

† National Science Foundation undergraduate research participant.

below. Also included is an analysis of the dependence of the molecular electrostatic energy upon the number of peptide units in a chain. This analysis permits one to determine excess free energies associated with the chain ends of form I and form II. An excess free energy of the I–II and II–I junctions can be similarly obtained.

Method of Calculation. The electrostatic dipole–dipole energy $W(N)$ of an array of N polyproline residues can be calculated with the residues in each of the four geometries shown in Figure 1, using the standard dipole–dipole approximation. Thus the mutual potential energy of two dipoles $\mathbf{m}_i$ and $\mathbf{m}_j$ at residues i and j is given by

$$W_{ij} = \frac{\mathbf{m}_i \cdot \mathbf{m}_j}{\epsilon r_{ij}^3} - \frac{3(\mathbf{m}_i \cdot \mathbf{r}_{ij})(\mathbf{m}_j \cdot \mathbf{r}_{ij})}{\epsilon r_{ij}^5} \qquad (1)$$

Here $\mathbf{r}_{ij}$ is the vector from residue i to residue j, r_{ij} is equal to $|\mathbf{r}_{ij}|$, and ϵ is an effective dielectric constant characterizing the space around the two dipoles. For an array of N dipoles, one has

$$W(N) = \sum_{i,j>i}^{N} W_{ij} \qquad (2)$$

In order to evaluate $W(N)$ one must establish the location in space of the N, C, and O atoms in each peptide unit. The dipole locations and directions are then readily found.

The geometries of form I and form II are taken from analyses of X-ray data (Traub and Shmueli, 1963; Sasisekharan, 1959). These data give directly the position of the N, C, and O atoms of each residue for a helix of arbitrary length. The positions in space of the atoms at a I–II and II–I junction can then be calculated. For these geometries the generating matrix methods are useful. These methods have been extensively applied to the analysis of polymer conformation (see, for example, Birshtein and Ptitsyn, 1966). In brief, each bond in a polypeptide structure is assigned a length, L_i, and a 3×3 transformation matrix, T_i. The transformation matrix T_i is usually characterized by the two angles θ_i and Φ_i. θ_i defines the bond angle between bond i and bond $i - 1$, whereas Φ_i defines the rotation of the bond i about an axis defined by bond $i - 1$. Thus, for example, the backbone of form I is fully characterized by four bond lengths (C_α–C′, C′–O, C′–N, and N–C_α) and four transformation matrices. Form II is similarly characterized by four different bond lengths and four different matrices. The values for these eight bond lengths and eight matrices are readily obtained from the data on the structure of form I and form II. Since the proline peptide residue can exist in only the *cis* or the *trans* form, with little variation sterically permitted in bond lengths and angles (De Santis *et al.*, 1965; Schimmel and Flory, 1967), the position of the atoms at a I–II and II–I junction region can be calculated by using appropriate combinations of bond lengths and matrices for the *cis* and *trans* peptide units. The junction is set at the midpoint of the molecule, *e.g.*, a 8-mer with a I–II junction would be of the form *cis-cis-cis-cis-trans-trans-trans-trans*, starting from the amino end, as shown in Figure 1.

The location and direction of each dipole moment are next established. The dipole moment of the proline residue is assumed to originate solely in the polar peptide group. The direction of the dipole moment is set identical with that measured in formamide (Kurland and Wilson, 1957), *i.e.*, the dipole is in the NCO plane and at an angle of 39.6° to the CN bond. This procedure seems to us to be preferable to the adoption of bond moment summations in view of the known resonance stabilization of the amide group, the observed invariance of amide dipole moments with substitution (Meighan and Cole, 1964), and the failure of bond moment summations to reproduce the observed magnitude of the dipole moment. It has in any case been verified that a ±20° variation of the dipole angle in the plane of the peptide group does not significantly affect our conclusions for form I and form II. Following Arridge and Cannon, we further locate the dipole center in the amide plane at the intersection of a line through the O nucleus and a second line passing through the amide C nucleus and perpendicular to the OC bond. The magnitude of the dipole moment is taken as 3.8 D, an average obtained from closely agreeing data for alkyl amides (Meighan and Cole, 1964). Once the locations and directions of the peptide dipoles are established, the electrostatic energy of molecules of various lengths is readily calculated using eq 1 and 2. All the calculations were carried out on an IBM 7094 computer.

Results

The electrostatic energy per residue for polymers of various lengths in form I and form II shows a strong dependence upon the number of peptide units, N, and the geometry, as shown in Figure 2. Similar data for the α helix (Arridge and Cannon, 1964) are included in the figure for comparison. For degrees of polymerization less than 18, the electrostatic energy of form I is less favorable than that of form II. The difference in molecular electrostatic energy reaches a maximum of 5.5 kcal/mole for $N = 6$. Only for $N > 18$ is form I more highly stabilized by electrostatic interactions than form II.

In order to calculate an electrostatic "end effect," the *total* electrostatic energy $W(N)$ of a molecule in form I (or form II) is assumed to be of the form, $W(N) = NW_{\text{I}} + 2E_{\text{I}}$, that is, each form I residue contributes an energy, W_{I}, and each end contributes an additional energy, E_{I}. For form II, W_{II} and E_{II} are similarly defined. E_{I} is readily evaluated by plotting $W(N)$ against N and then extrapolating the linear data for $N > 10$ to $N = 0$. The values of E_{I} and E_{II} found are +7 and +2 kcal per mole for dielectric constant unity. The difference between the end effects in the two forms is expected to manifest itself primarily in a tendency for short chains to adopt the form II geometry rather than form I.

The electrostatic energy of molecules with I–II and II–I junctions are substantially more positive than the pure form I and pure form II helices of equal length, as shown in Figure 3. It is found that a molecule with

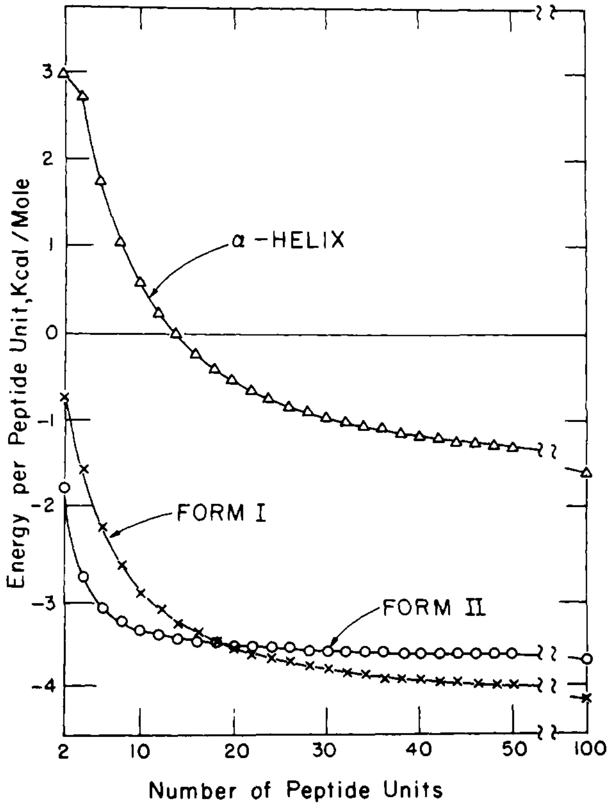

FIGURE 2: The intramolecular electrostatic dipole-dipole energy per residue for the α helix, polyproline I, and polyproline II as a function of chain length. The dielectric constant has been set equal to unity.

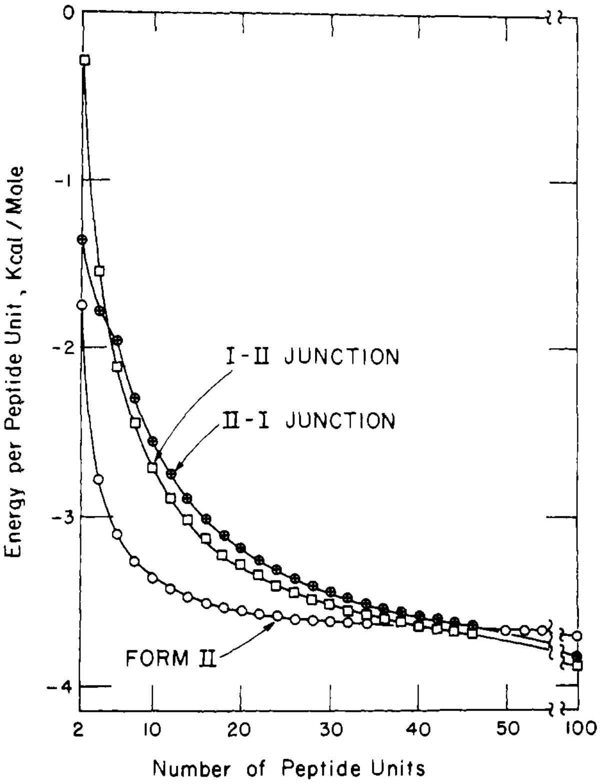

FIGURE 3: The electrostatic dipole–dipole energy per residue for polyproline II, for a polymer with a single I–II junction at the middle, and for a polymer with a single II–I junction at its midpoint. The dielectric constant is taken as unity. The data for form II are included to provide a reference value for the electrostatic energy of a molecule without a junction.

a I–II junction is slightly more stable than the molecule with a II–I junction. It is illuminating to reduce the data of Figure 3 to yield effective junction excess free energies. For example, the I–II geometry can be assigned a junction excess free energy, $J_{\text{I-II}}$, if the total energy $W(N)$ of a I–II N-mer is assumed to be of the form, $W(N) = N(W_{\text{I}} + W_{\text{II}})/2 + E_{\text{I}} + E_{\text{II}} + J_{\text{I-II}}$. A plot of $W(N)$ *vs.* N is linear for $N > 10$, permitting accurate extrapolation. $J_{\text{I-II}}$ is found to equal +4 kcal/mole for $\epsilon = 1$. A similar analysis yields $J_{\text{II-I}}$, the excess energy of a II–I junction. It is found that $J_{\text{II-I}}$ equals +7 kcal/mole for dielectric constant unity. These results are summarized in Table I. The 12-kcal/mole excess free energy for ends of an α helix, which was calculated in a similar way, is also included in the table.

TABLE I: Excess Electrostatic Free Energies Evaluated for Dielectric Constant Unity.

	kcal/mole
Form I end	+7
Form II end	+2
I–II junction	+4
II–I junction	+7
α-Helix end	+12

The values for the calculated electrostatic energies presented above are subject to at least four sources of adjustment. The first and most important is the effect of solvent in the region between and surrounding any two dipoles. This may reduce the absolute magnitudes of the calculated energies by the factor $1/\epsilon$, as given in eq 1, where ϵ is the dielectric constant of the medium. Many of the experiments on polyproline have been done in media with $5 < \epsilon < 20$. One should note, however, that the bulk dielectric constant will probably *overestimate* the solvent effect, since its use presupposes that solvent molecules can freely enter the region between the dipoles and can orient themselves so as to generate a minimum total system energy. For the short-range interactions present in, for example, the II–I junction, steric factors will prevent the solvent molecules from entering the region between near-neighbor dipoles.

The second adjustment factor is the polarizability of the solute itself. This correction has been examined for the α helix by Arridge (1965) and found to stabilize that structure by a factor of 2. No correction for this factor has been made in the present work but it is unlikely that its inclusion would alter the qualitative conclusions from the calculations.

A third possible adjustment factor is the geometry adopted for the I–II and II–I junctions. We have assumed that the junction bond lengths and bond angles are the same as those of the simple helices, whereas in fact some accommodation may occur. Model studies (Rifkind and Applequist, 1967) suggest that, especially for the II–I junction, steric overlap requires some variation in bond angles. However, the long-range nature of the Coulombic interaction makes unlikely a major change in electrostatic energy with the minor adjustment in bond lengths and bond angles required to achieve steric compatibility.

A fourth possible correction would be the use of a point monopole interaction to replace the dipole–dipole interaction. In the absence of any data on the locations of the monopoles, there seems little point to introducing this refinement.

Discussion

The molecular basis of the configuration of high molecular weight polyproline has been the subject of several previous investigations (De Santis *et al.*, 1965; Schimmel and Flory, 1967; Rifkind and Applequist, 1968). These investigations attribute the secondary structure of polyproline and the cooperativity of the I–II transition to steric and van der Waals effects. Indeed, the importance of steric factors in limiting the configurational freedom of the proline residue to either the *cis* (form I) or *trans* (form II) geometry is quantitative and well founded. In contrast, however, the molecular nature of the energetic barrier to intrachain I–II and II–I junctions has not been so convincingly established in earlier work. Molecular model studies (Schimmel and Flory, 1967; Rifkind and Applequist, 1968) suggest a steric contribution, but the quantitative evaluation of steric interactions at such junctions has not been reported to date.

End Effects. Experimental data are available to test the predictions of the electrostatic calculations. For example, optical studies on proline oligomers with degrees of polymerization between two and eight (Okabayashi *et al.*, 1968) show that molecules with four to eight peptide groups are in form II geometry in acetonitrile, water, and methanol, whereas the dimer and trimer are not entirely in this conformation. High molecular weight material, in which end effects are unimportant, adopts form II in water and presumably in acetonitrile, but adopts form I in methanol (Steinberg *et al.*, 1958). The failure of the dimer and trimer to adopt a helical geometry may in part be due to unfavorable electrostatic end effects. Steric restrictions must also be diminished in these short chains. More important is the observation that chains with four to eight proline residues, when dissolved in methanol, are in form II rather than form I. This observation supports the contention that form II has more favorable electrostatic end effects than does form I. However, evidence that the electrostatic end effect is not universally dominant is provided in a study of proline oligomers by Yaron and Berger (1961). These authors showed that the 6-mer and the 12-mer can exist both in form I and in form II in suitable (unspecified) solvents.

Additional evidence supporting the electrostatic end effect is provided in a study of the I–II transition for degrees of polymerization 34 and 215 (Engel, 1966). Under solvent conditions which, for high molecular weight, lead to equal concentrations of form I and form II, the I–II equilibrium is shifted to 10% form I:90% form II for the low molecular weight. The diminished stability of short form I helices has been previously ascribed to loss of favorable van der Waals interactions at the chain ends (Schimmel and Flory, 1967; Rifkind and Applequist, 1968). However, the direction of the shift is consistent with an electrostatic origin of the end effects as well.

A more detailed comparison between molecular hypothesis and experiment requires that the I–II transition be described thermodynamically. Appropriate theories are those developed by Schwarz (1967, 1968) and Applequist (1968). These theories are an extension of the Zimm–Bragg (1959) and Lifson–Roig (1961) theories, which use an Ising model for the polymer chain. In these theories the conformational transition is in part described by parameters σ_i which are the Boltzmann factors $e^{-\Delta F/RT}$ corresponding to the excess free energies of the ends and internal junctions. A small value of σ at a junction thus corresponds to a highly cooperative transition. Using this theory, the data of Engel can be analyzed and the difference between the excess free energies of the ends of form I and form II helices estimated (compare Figure 2 of Engel (1966) with Figure 3 of Schwarz (1968)). Such an analysis yields $E_{\text{I}} - E_{\text{II}} = 1.4$ kcal/mole ($\sigma_{\text{B}}/\sigma_{\text{A}} = 0.1$). The difference in calculated electrostatic end effects is 5.5 kcal/mole for $\epsilon = 1$. Since the measurements were done in a solvent of mean dielectric constant 16, the electrostatic end effects are probably reduced to 0.3–0.6 kcal/mole. Thus although peptide dipole–dipole interactions are probably not the sole origin of the end effects in polyproline, they make a significant contribution of the observed sign.

Junction Excess Free Energy. A comparison between theory and experiment for the junction excess free energy is also possible but is complicated by a lack of agreement on the experimental value. Values ranging from +3 to +7 kcal per mole ($10^{-2} > \sigma > 10^{-5}$) have been suggested on the basis of pressure and kinetic studies (Rifkind and Applequist, 1968; Engel, 1966). The electrostatic effects can be calculated from the average of J_{I-II} and J_{II-I}; both of these must contribute because a molecule cannot adopt several I–II junctions without possessing II–I junctions as well. The electrostatic junction free energy is thus $5.5/\epsilon$ kcal/mole. For ϵ between 3 and 15, the calculated junction free energy is +0.4 to +2 kcal per mole. Thus electrostatic effects alone can make a large contribution to the observed positive cooperative effect in polyproline.

Comments on the α-Helix–Random Coil Transition. The analysis presented above, relating the electrostatic energy calculations to the parameters of the statistical thermodynamic theory of the I–II transition, can also be applied to the equilibrium between the α helix and disordered polypeptide chains. The importance of electrostatic effects in this case has been demonstrated by Arridge and Cannon (1964) and by Brant and Flory (1965a, b); we wish merely to point out that the Arridge–Cannon calculations can be used to estimate the electrostatic contributions to the nucleation parameter σ or v in the Zimm–Bragg or Lifson–Roig theory, respectively. In these theories the parameter σ or v gauges the difficulty of nucleating and terminating an α-helical region. The random coil is assigned a nucleation parameter equal to unity. The calculated excess electrostatic free energy, corresponding to σ in the Zimm–Bragg theory, is $+24/\epsilon$ kcal/mole, since this theory includes the effects of both ends in a single parameter. For $3 < \epsilon < 10$, one finds $8 > \Delta F > 2.4$ kcal/mole, corresponding to $10^{-6} < \sigma < 10^{-2}$.

The calculated contributions of electrostatic end effects to σ can be compared with experimental findings. For example, poly-γ-benzyl-L-glutamate in dichloroethylene–dichloroacetic acid mixtures ($\epsilon < 10$) shows σ equal to 2×10^{-4} (Zimm *et al.*, 1959), poly-L-glutamic acid in water shows σ equal to 2×10^{-3}, and poly-L-glutamic acid in water–dioxane mixtures of intermediate dielectric constant shows intermediate values of σ (Ptitsyn, 1967). The order of magnitude of σ has been previously ascribed to the entropy of nucleation of a helical turn (Zimm and Bragg, 1959). The dependence upon dielectric constant has been qualitatively assigned to a solvent-dependent constriction of configuration space of the random coil by dipole–dipole interactions (Ptitsyn, 1967). Our calculation suggests that the observed data and their dependence upon the dielectric constant also reflect long-range electrostatic dipole–dipole end effects within the helical regions. Further experimental support for the existence of a large, energetically unfavorable electrostatic end effect in the α helix is provided by studies on the conformation of L-glutamate oligomers with degrees of polymerization between 2 and 11 (Goodman and Rosen, 1964). These studies show that chains with fewer than seven residues are disordered in solvents in which longer molecules become α helical. The large value of the dipole moment of the peptide groups in proteins, and the compactness of most proteins suggest that such interactions may have a profound influence upon protein conformation as well.

Acknowledgments

We wish to thank E. P. Geiduschek for stimulating discussions and R. Mandel for confirming the conformation of polyproline in methanol.

Added in Proof

The discussion on electrostatic effects in the α helix, given in the last two paragraphs above, is similar to findings presented in a recent paper by Brant (1968), which we overlooked.

References

Applequist, J. (1968), *Biopolymers 6*, 117.

Arridge, R. G. C. (1965), *Proc. Phys. Soc. 85*, 1157.

Arridge, R. G. C., and Cannon, C. G. (1964), *Proc. Roy. Soc.* (*London*) *A278*, 91.

Birshtein, T. M., and Ptitsyn, O. B. (1966), Conformations of Macromolecules Vol. 22, New York, N. Y., Interscience.

Brant, D. A. (1968), *Macromolecules 1*, 291.

Brant, D. A., and Flory, P. J. (1965a), *J. Am. Chem. Soc. 87*, 663.

Brant, D. A., and Flory, P. J. (1965b), *J. Am. Chem. Soc. 87*, 2791.

De Santis, P., Giglio, E., Liquori, A. M., and Ripamonti, A. (1965), *Nature 206*, 456.

Engel, J. (1966), *Biopolymers 4*, 945.

Gibson, K. D., and Scheraga, H. A. (1966), *Biopolymers 4*, 709.

Goodman, M., and Rosen, I. (1964), *Biopolymers 2*, 537.

Katchalski, E., Berger, A., and Kurtz, J. (1963), *in* Aspects of Protein Structure, Ramachandran, G. N., Ed., New York, N. Y., Academic, p 205.

Kurland, R. J., and Wilson, E. B., Jr. (1957), *J. Chem. Phys. 27*, 585.

Lifson, S., and Roig, A. (1961), *J. Chem. Phys. 34*, 1963.

Mandelkern, L. (1967), *in* Poly-α-amino Acids, Fasman, G. D., Ed., New York, N. Y., Marcel Dekker, Chapter 13.

Meighan, R. M., and Cole, R. H. (1964), *J. Phys. Chem. 68*, 503.

Okabayashi, H., Isemura, T., and Sakakibara, S. (1968), *Biopolymers 6*, 323.

Ptitsyn, O. B. (1967), *in* Conformation of Biopolymers, Ramachandran, G. N., Ed., New York, N. Y., Academic, p 381.

Rifkind, J. M., and Applequist, J. (1968), *J. Am. Chem. Soc. 90*, 3650.

Sasisekharan, V. (1959), *Acta Crystal. 12*, 897.

Schimmel, P. R., and Flory, P. J. (1967), *Proc. Natl. Acad. Sci. U. S. 58*, 52.

Schwarz, G. (1967), *Biopolymers 5*, 321.

Schwarz, G. (1968), *Biopolymers 6*, 873.

Steinberg, I. Z., Berger, A., and Katchalski, E. (1958), *Biochim. Biophys. Acta 28*, 647.

Traub, W., and Shmueli, V. (1963), *Nature 198*, 1165.

Yaron, A., and Berger, A. (1961), *Bull. Res. Council Israel 10A*, 46.

Zimm, B. H., and Bragg, J. K. (1959), *J. Chem. Phys. 31*, 526.

Zimm, B. H., Doty, P., and Iso, K. (1959), *Proc. Natl. Acad. Sci. U. S. 45*, 1601.