Free-Energy Calculations of the Interactions of Helical Poly(L-Proline) with Water*

(polypeptides/solvent interactions/conformational energies/conformational entropies)

S. KRIMM AND C. M. VENKATACHALAM

Harrison M. Randall Laboratory of Physics, and Biophysics Research Division, Institute of Science and Technology, University of Michigan, Ann Arbor, Mich. 48104

Communicated by J. L. Oncley, July 26, 1971

ABSTRACT Conformational energies, consisting of nonbonded and electrostatic interactions, have been calculated for poly(L-proline) I and II helices, of six-residues length, with a water molecule hydrogen-bonded to each imide group. For polypyrroline I, the Traub-Shmueli prolyl-ring geometry was used, whereas for polypyrroline II, calculations were done for the 2 ring geometries of Leung and Marsh and, in some cases, for the several ring geometries of Ramachandran, *et al.* The chain torsion angles $\omega(\text{N}-\text{C}')$ and $\psi(\text{C}^\alpha-\text{C}')$ were varied, as were the two angles that specify the orientations of the water molecules. By summation over the latter, it was possible to calculate a free energy at each ω, ψ that incorporated the conformational entropy of the water molecules. Such ω, ψ maps reveal that binding of water can cause changes in the equilibrium conformation of polypyrroline, as well as in the energetics of the chain with respect to ω and ψ . This has important consequences on the statistics of the chain. Considerations of prolyl-ring self-energies, as well as the conformational entropies of the chains, show that polypyrroline II is more stable in water than is polypyrroline I, and determine which ring geometry is favored for polypyrroline in water.

In recent years, a large effort has been devoted to the calculation of stable conformations of polypeptides and proteins (1-3). The goal of this work has been to predict the spatial organization of atoms within polypeptide chains, on the basis of knowledge of the primary structure and of the nature of the interactions between the atoms.

In most of the theoretical work to date, the polypeptide chain has been considered essentially as an isolated molecule in a vacuum. This approximation may be adequate for structures such as the α -helix, in which peptide groups interact with each other internally in the chain via hydrogen bonds. It is undoubtedly far less satisfactory for more open chain structures, in which the peptide groups interact with molecules of solvent or with other small molecules present in solution. While it has been generally recognized that interactions with solvent molecules can influence the structure of a polypeptide chain, no attempt has as yet been made to take into account explicitly the possible effects of these interactions. Yet, some of the most significant properties of synthetic polypeptide systems, and certainly of proteins, are studied in solution. It is, therefore, less than adequate to ignore the contributions of such interacting small molecules to the determination of the properties of a polypeptide chain.

In this series of papers, we propose to examine explicitly the ways in which the interactions of solvent and other small molecules influence the conformations and statistical properties of polypeptide chains. We will be concerned first with the hydrogen-bonded interaction of water with the peptide group, and we will therefore concentrate our attention on structures in which such interactions are feasible. In later papers, we will discuss the interactions of other solvating molecules.

For the general polypeptide chain, at least two water molecules can be hydrogen-bonded to each peptide group, one via the O—H bond to the oxygen atom of the $\text{>C}'=\text{O}$ group and the other via the O atom to the H atom of the $\text{>N}-\text{H}$ group. These water molecules will have many degrees of freedom in their orientation with respect to the peptide group. Postponing study of the general polypeptide chain with two hydrogen-bonded water molecules to a subsequent paper, we deal here with the simpler case of the poly(L-proline) $[(\text{Pro})_n]$ chain, in which only hydrogen bonding to a $\text{>C}'=\text{O}$ group is possible. This calculation is, however, relevant to determining the role that water plays in stabilizing the $(\text{Pro})_n$ structure, and to establishing the preferred conformations in solution and their statistical properties. We treat here only helical conformations; nonhelical structures will be considered in a later paper.

In this treatment, we are neglecting the interaction of the polymer chain with the solvent continuum. Our model of the chain in effect enlarges the monomer unit by associating with it the specifically bound small molecule.

CALCULATIONS

Much is known about the structure of $(\text{Pro})_n$ (4). X-ray diffraction studies in the solid state show that two helical conformations are possible: form I, a right-handed helix of *cis* residues having ten units in three turns, and form II, a left-handed helix of *trans* residues with three units per turn. Both types of conformations are also thought to occur in solution under suitable conditions, although only $(\text{Pro})_n\text{II}$ is stable in water. Infrared studies of $(\text{Pro})_n\text{II}$ in aqueous solution (5) show conclusively that water molecules are hydrogen bonded to the $\text{>C}'=\text{O}$ groups of the chain.

The model we have used for our calculations consists of a helix of six proline residues, to each $\text{>C}'=\text{O}$ of which is attached one molecule of H_2O . (This is based on an energy calculation that shows that steric interactions with the

Abbreviations: $(\text{Pro})_n$, poly(L-proline); TS, Traub and Shmueli; LM, Leung and Marsh.

* This paper is the first in a series: "Theoretical Studies of the Interactions of Polypeptides with Small Molecules."

chain completely forbid the hydrogen bonding of two water molecules to a single $>C'=O$ group.) While we recognize that this water molecule is exchangeable with molecules in the medium, it is likely that a water molecule is hydrogen bonded to a $>C'=O$ group for a large enough fraction of the time so that its presence predominates in determining the conformation of the chain.

The orientation of the water molecule with respect to the peptide group was specified as follows. The H^1-O^w bond of the water molecule was placed in the plane of the imide group (see Fig. 1). The $O \cdots O^w$ distance was set equal to 2.60 Å. (The effect of varying this distance has also been studied.) Two more parameters completely specify the position and orientation of the water molecule: α , the supplement of the $C'OH^1$ angle, and θ , the $C^{\alpha}C'O^wH^2$ dihedral angle at $\alpha = 0^\circ$ ($\theta = 0^\circ$ for H^2 *cis* to C^{α}). The $O-H$ bond length was taken as 1.0 Å, and the bond angle at the O atom as 105° .

The prolyl-ring geometry was obtained from previous studies. For $(Pro)_nI$, all bond lengths, bond angles, and torsion angles (except ψ) were taken from Traub and Shmueli (TS) (6). For $(Pro)_nII$, calculations were done for the 2 ring geometries of Leung and Marsh (7), corresponding to the C_6'' position (which we call LM1) and the C_6' position (which we call LM2) of the C^{γ} atom. In addition, in order to evaluate the influence of ring geometry on the energy, some calculations were made using the 80 different ring conformations recently given by Ramachandran, *et al.* (8). The geometry of the prolyl-ring is specified in Table 3 of ref. 8. The bond angles $\tau(C^{\alpha}C'O)$ and $\tau(C^{\alpha}C'N)$ were both fixed at 119° , in agreement with experimental data (7). In all calculations the $\tau(HCH)$ angles in the prolyl ring were taken as 110° . Subsequent calculations in which this angle was allowed to vary with the CCC angle according to a relation obtained using the Complete Neglect of Differential Overlap (CNDO) method (E. Ronish and S. Krimm, to be published), did not give significantly different results.

Two contributions to the energy were incorporated in the calculation. The first was a nonbonded potential of the Lennard-Jones 6-12 form, namely, $V_{nb} = a_{ij}/r^{12}_{ij} - b_{ij}/r^6_{ij}$. The values of the potential constants a_{ij} and b_{ij} for the various interactions were taken from Scott and Scheraga (9). The second energy contribution was an electrostatic interaction computed by using the monopole approximation, namely, $V_{es} = e_i e_j / \epsilon r_{ij}$. The dielectric constant ϵ was taken as 4.0, and the partial charges, e_i , on the atoms of the $(Pro)_n$ chain were those of Hopfinger (10). The partial charges on the water molecule were determined so as to reproduce its known dipole moment, and were found to be $e_O = -0.66$ and $e_H = +0.33$, in units of the electronic charge. With respect to the torsion potentials, the energy contribution due to the ψ rotation (11) was neglected, but energy changes due to the ω rotation (11) were included on the basis of 0.61 kcal/mol for a 10° rotation (2). The hydrogen-bond energy has not been included, since it is found to be almost constant with α in the range of relevance to $(Pro)_n$ (12).

The conformational energies of the $(Pro)_n$ -water system are functions of ω , ψ , α , and θ , namely, $E(\omega, \psi, \alpha, \theta)$. Using computer programs discussed elsewhere (13), we computed $E(\omega, \psi, \alpha, \theta)$ by independently varying ω , ψ , α , and θ . The angle ω was varied in steps of 5° from -20° to $+20^\circ$ for

FIG. 1. Geometry of the imide group with a hydrogen-bonded water molecule.

$(Pro)_nI$ and from 160° to 200° for $(Pro)_nII$; the angle ψ was varied in steps of 5° from 100° to 210° ; the angle α was varied in steps of 15° from -60° to $+60^\circ$; and the angle θ was varied in steps of 30° from 0° to 360° . The nonbonded plus electrostatic energies were computed for each of these conformations.

The energies obtained by the above procedure are functions of ω , ψ , α , and θ . We wish, however, to know the conformations of the chain that prevail in aqueous solution, i.e., the energies as a function of ω and ψ . It would, of course, be possible to specify at a given ω and ψ the structure that is of lowest energy in α and θ . We would thereby, however, be missing an important factor, namely the conformational entropy of the water molecules. This is particularly important since the available α , θ space varies with the ω and ψ .

We have, therefore, incorporated the conformational entropy of the hydrogen-bonded water molecules. We have assumed that α and θ represent completely the freedom of the water molecule. This is quite reasonable, since the energy of the hydrogen bond will increase comparatively rapidly with variation in the $O \cdots O^w$ distance. (The hydrogen-bond energy is the major factor in this energy change, the steric energy change being minimal. Thus, changing $O \cdots O^w$ to 2.70 Å increases the steric energy by about 0.2 kcal/mol.) The conformational entropy of the water molecule is then given by $RT \ln \nu$. Here ν is the number of orientational "isoenergetic" positions of the water molecules for a given polymer conformation, i.e., the number of sets of (α, θ) corresponding to equal or comparable potential energy. More accurately, the entropy S can be calculated from (14)

$$S(\omega, \psi) = R \ln Z + \langle E \rangle / T, \quad (1)$$

where the partition function Z and the average potential energy $\langle E \rangle$ are obtained from the following summations over the various water configurations:

$$Z = \sum_{\alpha} \sum_{\theta} \exp[-E(\omega, \psi, \alpha, \theta) / RT] \quad (2)$$

$$\langle E \rangle = Z^{-1} \sum_{\alpha} \sum_{\theta} E(\omega, \psi, \alpha, \theta) \exp[-E(\omega, \psi, \alpha, \theta) / RT]. \quad (3)$$

Using these formulas, we then associate a free energy, $F(\omega, \psi)$, with each polymer conformation, given by

$$F(\omega, \psi) = E_{\min}(\omega, \psi) - TS(\omega, \psi), \quad (4)$$

where $E_{\min}(\omega, \psi)$ is the minimum potential energy for a given ω and ψ .

FIG. 2. Map of conformational free-energy of poly(L-proline) I as a function of ω and ψ . (A) Without water, (B) with water. Contours in kcal/mol of residues.

RESULTS AND DISCUSSION

The conformational free energies of helical (Pro)_nI, both with and without water, are given in Fig. 2 as a function of ω and ψ . The global minimum occurs at $\omega = 0^\circ$, $\psi = 155^\circ$ in the absence of water and at $\omega = 5^\circ$, $\psi = 155^\circ$ in the presence of water. We see, therefore, that interaction of solvent molecules with a polypeptide chain can result in a change in the equilibrium conformation of the polypeptide. An interesting result of this calculation, not revealed in previous calculations (15, 16) in which only $\omega = 0^\circ$ was used, is that the (Pro)_nI helix conformation can as easily be distorted from equilibrium by an ω rotation as by a ψ rotation. Since the possibility of ω rotations is relevant in determining such chain properties as end-to-end length, we feel that it is important to include the energetics of the ω rotation in conformational energy calculations.

In the case of (Pro)_nII, we have made the calculations for LM1- and LM2-ring conformations. Since other ring geometries may be possible (8), however, we have attempted to evaluate our use of LM1 and LM2 by comparing the energies of helices having these two ring geometries with helices in which the 80 lowest-energy ring geometries (8) can occur. To simplify this calculation, only the case of $\omega = 0^\circ$ and $\alpha = 0^\circ$ was examined, although for the case of attached water molecules the energy was minimized with respect to θ . The variation of $E_{\min}(\psi)$ with ψ is shown in Fig. 3. In Fig. 3A, the one ring geometry out of the 80 that gives a helix of lowest energy at a particular ψ is indicated by its symbol from Table 3 of ref. 8. The A geometries are of the LM1 type and the B geometries are of the LM2 type. The $E_{\min}(\psi)$ in this case include the relative self-energies of the different ring structures (8), and it is significant to note that the ring geometry changes as the helix geometry is altered by changes in ψ under the conditions stated above. In Fig. 3B and 3C, we give $E_{\min}(\psi)$ as a function of ψ for LM1 and LM2 geometries, respectively. The latter calculations do not include prolyl-ring self-energies, so that it is not feasible to compare directly the magnitudes of $E_{\min}(\psi)$ for LM1 and LM2. However, our preliminary calculations of

FIG. 3. Potential energy (minimum with respect to θ) of poly(L-proline)II as a function of ψ for $\omega = \alpha = 0^\circ$, with (—, and lower row of symbols) and without (---, and upper row of symbols) water. (A) For Ramachandran, et al. (8) 80 ring geometries, (B) for Leung-Marsh-1 ring geometry, (C) for Leung-Marsh-2 ring geometry.

prolyl-ring self-energies show that the self-energy of LM1 is about 1.5 kcal/mol higher than that of LM2. If we add this amount to the curves of Fig. 3B, then we see that the combination of the respective curves of Fig. 3B and C is in good agreement with that calculated in Fig. 3A from the Ramachandran, et al. (8) ring geometries.

The results of Fig. 3 indicate that by calculating energy maps with only LM1 and LM2 ring geometries, we have, with reasonable satisfaction, spanned the limiting cases of the conformations of the prolyl ring.

The conformational free energies of helical (Pro)_nII for LM1 and LM2 geometries are given in Figs. 4 and 5, respectively. In each case, prolyl-ring self-energies are not included; for each ring geometry the energies are shown with and without water.

Some general observations may be made with respect to the energy plots in Figs. 4 and 5. First, whereas the addition of water produces a slight change in chain conformation in the case of LM1 (from $\omega = 180^\circ$, $\psi = 175^\circ$ to $\omega = 180^\circ$, $\psi = 180^\circ$), no significant change occurs for LM2 (the minimum remains at $\omega = 180^\circ$, $\psi = 160^\circ$). Second, although the addition of water diminishes the ω , ψ space accessible to the chain (consider for example the energy contours that are 2 kcal/mol above the minimum), this reduction is somewhat larger for LM1 than for LM2. Third, while addition of water does not significantly change the shape of the energy contour

FIG. 4. Map of conformational free energy of poly(L-proline)-II, LM1 ring geometry, as a function of ω and ψ . (A) Without water, (B) with water. Contours in kcal/mol of residues.

near the minimum for LM1, there is an important shape change for LM2: the trough at the minimum shifts from being essentially parallel to the ψ -axis to being significantly inclined to this axis. Thus, incorporation of solvent molecules that can bind to a chain leads to important alterations in the energy surface that governs the conformation of the chain.

The above results also permit us for the first time to establish in a detailed way the energetic basis for the stable conformation of (Pro)_n in aqueous solution. Our preliminary self-energy calculations for the ring show that the TS ring geometry is about 5.3 kcal/mol higher in energy than the LM2 geometry. Thus, the global minimum for TS (Pro)_nI with water has an energy of -2.0 kcal/mol whereas that for LM2 (Pro)_nII has an energy of -7.4 kcal/mol. In addition, it can be seen by comparing Fig. 2B with Fig. 5B that the conformational entropy of the chain for LM2 (Pro)_nII is much higher than that of TS (Pro)_nI, thus further favoring the former structure. (The energies of a (Pro)_nI chain without water, but with LM1 or LM2 ring geometries, are about 1 kcal/mol higher than energies of the corresponding (Pro)_nII chains. Since addition of water lowers the energy of (Pro)_nII chains more than it does (Pro)_nI, and since the entropic factor is still dominating, we expect that (Pro)_nI with other than TS ring geometry will still be disfavored with respect to (Pro)_nII in water.) Similar considerations readily show that LM2 (Pro)_nII is favored over LM1 (Pro)_nII in water: the energy difference at the minimum is about 1.6 kcal/mol, when ring self-energies are included, and the chain-conformational entropy adds over 1 kcal/mol. Thus, LM2 (Pro)_nII is the preferred conformation of (Pro)_n in aqueous solution.

In summary, the bonding of water molecules to the (Pro)_n chain is seen to have important consequences with respect to the conformation and energetics of the chain and, therefore,

FIG. 5. Map of conformational free-energy of poly(L-proline)-II, LM2 ring geometry, as a function of ω and ψ . (A) Without water, (B) with water. Contours in kcal/mol of residues.

to its statistics. This should be true for other solvating molecules and other polypeptide chains, and the detailed consequences of such interactions will be examined in future papers in this series.

This research was supported in part by National Science Foundation grant GB-15682.

1. Liquori, A. M., *Ciba Foundation Symposium on the Principles of Biomolecular Organization* (1966), p. 40.
2. Ramachandran, G. N., and Sasisekharan, V., *Advan. Protein Chem.*, **23**, 283 (1968).
3. Scheraga, H. A., *Advan. Phys. Org. Chem.*, **6**, 103 (1968).
4. Mandelkern, L., in *Poly- α -Amino Acids*, ed. G. D. Fasman (Marcel Dekker, Inc., New York, 1967), chap. 13.
5. Swenson, C. A., and R. Formanek, *J. Phys. Chem.*, **71**, 4073 (1967).
6. Traub, W., and U. Shmueli, *Nature*, **198**, 1165 (1963).
7. Leung, Y. C., and R. E. Marsh, *Acta Crystallogr.*, **11**, 17 (1958).
8. Ramachandran, G. N., A. V. Lakshminarayanan, R. Balasubramanian, and G. Tegoni, *Biochim. Biophys. Acta*, **221**, 165 (1970).
9. Scott, R. A., and H. A. Scheraga, *J. Chem. Phys.*, **45**, 2091 (1966).
10. Hopfinger, A. J., Ph.D. thesis, Case Western Reserve University, 1970.
11. IUPAC-IUB Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains, *J. Mol. Biol.*, **52**, 1 (1970).
12. Rao, C. N. R., A. Goel, K. G. Rao, and A. S. N. Murthy, *J. Phys. Chem.*, **75**, 1744 (1971).
13. Ramachandran, G. N., C. M. Venkatachalam, and S. Krimm, *Biophys. J.*, **6**, 849 (1966).
14. Flory, P. J., in *Statistical Mechanics of Chain Molecules* (Interscience Publishers, New York, 1969), p. 274.
15. De Santis, P., E. Giglio, A. M. Liquori, and A. Ripamonti, *Nature*, **206**, 456 (1965).
16. Hopfinger, A. J., and A. G. Walton, *J. Macromol. Sci.-Phys.*, **B3**, 171 (1969).