Conformational Characteristics of L-Proline Oligomers

Alan E. Tonelli

Contribution from the Bell Telephone Laboratories, Inc.,
Murray Hill, New Jersey 07974. *Received April 1, 1970*

Abstract: The conformational characteristics of L-proline oligomers are studied by calculating the intramolecular energies of all conformers distinguishable from each other by the cis or trans character of each of their imide bonds. Intrinsic torsional and nonbonded van der Waals potentials and dipolar electrostatic interactions which depend on the cis or trans character of the imide groups and the conformation about the C^α–C bonds are considered. Appreciable amounts of most of the four and eight conformers of the di- and tri-L-proline oligomer derivatives (*tert*-butoxy N-terminal blocking group), respectively, are predicted, in agreement with recently reported experimental observations. Neglect of all interactions beyond those between first and second neighbor proline residues enables the intramolecular energies of the tetra- and penta-L-proline derivatives to be evaluated from the energies calculated for the dimer and trimer. Beginning with the trimer, all nonterminal residues are predicted to be exclusively trans, while the N- and C-terminal residues in all the oligomers are found to be mixtures of cis and trans. However, according to the present calculations the reported onset of the form II helical conformation abruptly at the pentamer has no intramolecular origin.

A proton nuclear magnetic resonance (nmr) study of the L-proline oligomer derivatives depicted in Figure 1 has recently been conducted by Deber, Bovey, Carver, and Blout[1] at 220 MHz. Following an analysis of the intensities and the number of benzylic, *tert*-butyl, and α-proton resonances in $CDCl_3$ at 25°, they concluded that all possible conformers, where conformers are differentiated by the rotational states (cis or trans) of the imide bonds, are present to an appreciable extent in L-proline oligomers with less than five residues. With the exception of the N-terminus, all imide bonds in L-proline oligomers with five or more residues were found to be exclusively trans. This implies that in chloroform L-proline oligomers with five or more residues assume a form II left-handed, helical conformation, while those oligomers with fewer than five residues assume most of the conformations consistent with cis–trans isomerization of the imide bonds.

In the present study the intramolecular conformational energies of the L-proline oligomers dimer through pentamer (see Figure 1) are calculated as a function of the conformation about the C^α–C bonds and as a function of the cis or trans character of each imide bond. It was hoped that from these calculations it could be determined if the mixture of cis and trans imide groups found[1] in L-proline oligomers with less than five residues has an intramolecular origin. In addition, the influence of intramolecular interactions on the abrupt onset of the form II helical conformation on passing from four to five residues in these L-proline oligomers is commented upon.

Description of the Calculations

The conformational freedom of poly-L-proline and its oligomers is restricted compared to other polypeptides by the rigid geometry[2,3] of the pyrrolidine ring (see Figure 1). Rotation about the proline backbone is confined to the C^α–C bonds. Additional conformational freedom may be realized if the imide bonds in poly-L-proline are able to adopt either the cis or trans conformation. Thus, the intramolecular conformational energy of the poly-L-proline chain is a function of the angles of rotation ψ and ω (both ψ and ω are taken[4] as 0° in the planar, trans conformation and adopt positive values for right-handed rotations); see Figure 1.

The intramolecular conformation energies of L-proline oligomers may be estimated[5,6] through use of semiempirical potential energy functions[7] for the intrinsic torsional potentials (the inherent torsional potential about the imide bond[8] is neglected and equal inherent torsional energies are assigned to the cis and trans conformers of this bond), for the nonbonded van der Waals interactions, and for the electrostatic interactions between dipolar imide groups. (The lack of an amide hydrogen precludes consideration of intramolecular hydrogen bonding.)

$$\begin{aligned} V(\psi_1, \psi_2, \ldots, \psi_n, \omega_1, \omega_2, \ldots, \omega_n) = {} & V_{\psi_1}(\psi_1) + V_{\psi_2}(\psi_2) + \ldots + V_{\psi_n}(\psi_n) + {} \\ & \sum_{j,k}[V_{r,jk}(\psi_1, \psi_2, \ldots, \psi_n, \omega_1, \omega_2, \ldots, \omega_n) + {} \\ & V_{l,jk}(\psi_1, \psi_2, \ldots,, \psi_n, \omega_1, \omega_2, \ldots, \omega_n) + {} \\ & V_{c,jk}(\psi_1, \psi_2, \ldots, \psi_n, \omega_1, \omega_2, \ldots, \omega_n)] \end{aligned} \quad (1)$$

where

$$V_\psi(\psi) = V_\psi{}^0/2[1 + \cos 3(\psi - \pi/3)] \quad (2)$$

$$V_{r,jk}(\psi_1, \psi_2, \ldots, \psi_n, \omega_1, \omega_2, \ldots, \omega_n) = a_{jk}/r_{jk}{}^{12} \quad (3)$$

$$V_{l,jk}(\psi_1, \psi_2, \ldots, \psi_n, \omega_1, \omega_2, \ldots, \omega_n) = -c_{jk}/r_{jk}{}^{6} \quad (4)$$

and

$$V_{c,jk}(\psi_1, \psi_2, \ldots, \psi_n, \omega_1, \omega_2, \ldots, \omega_n) = 332\delta_j\delta_k/\epsilon r_{jk} \quad (5)$$

---

[1] C. M. Deber, F. A. Bovey, J. P. Carver, and E. R. Blout, *J. Amer Chem. Soc.*, **92**, 6191 (1970).

[2] V. Sasisekharan, *Acta Crystallogr.*, **12**, 897 (1959).

[3] W. Traub and U. Shmueli, *Nature (London)*, **198**, 1165 (1963).

[4] J. T. Edsall, P. J. Flory, J. C. Kendrew, A. M. Liquori, G. Nemethy, G. N. Ramachandran, and H. A. Scheraga, *J. Biol. Chem.*, **241**, 1004 (1966); *Biopolymers*, **4**, 127 (1966); *J. Mol. Biol.*, **15**, 399 (1966).

[5] P. R. Schimmel and P. J. Flory, *Proc. Nat. Acad. Sci. U. S.*, **58**, 52 (1967).

[6] G. Holzwarth and R. Chandrasekaran, *Macromolecules*, **2**, 245 (1969).

[7] D. A. Brant, W. G. Miller, and P. J. Flory, *J. Mol. Biol.*, **23**, 47 (1967).

[8] J. F. Yan, F. A. Momany, R. Hoffmann, and H. A. Scheraga, *J. Phys. Chem.*, **74**, 420 (1970).

Table I. The Energetically Allowed Conformations of tt and cc Di-L-proline and of ttt and ccc Tri-L-proline

Energetically allowed^a ranges of ψ_1, ψ_2, and ψ_3				
	tt ($\omega_1 = \omega_2 = 0°$)		cc ($\omega_1 = \omega_2 = 180°$)	
$\psi_1 = -30$ $-60 \leq \psi_2 \leq 30$ $120 \leq \psi_2 \leq 180$	$\psi_1 = 0$ $-60 \leq \psi_2 \leq 30$ $120 \leq \psi_2 \leq 210$	$\psi_1 = 30$ $-30 \leq \psi_2 \leq 30$ $120 \leq \psi_2 \leq 180$	$\psi_1 = 0$ $-60 \leq \psi_2 \leq 0$ $120 \leq \psi_2 \leq 180$	$\psi_1 = 30$ $-30 \leq \psi_2 \leq 0$ $150 \leq \psi_2 \leq 210$
		ttt ($\omega_1 = \omega_2 = \omega_3 = 0°$)		
$\psi_1 = 30$ $-30 \leq \psi_2 \leq 0$ $-30 \leq \psi_3 \leq 30$ $120 \leq \psi_3 \leq 180$		$\psi_1 = 0 \quad (-30)$ $\psi_2 = -30\ (0)$ $-30 \leq \psi_3 \leq 30$ $120 \leq \psi_3 \leq 180$		$\psi_1 = -30$ $\psi_2 = 30$ $-30 \leq \psi_3 \leq 0$ $\psi_3 = 150$
$\psi_1 = 30$ $\psi_2 = 30$ $-30 \leq \psi_3 \leq 0$	$\psi_1 = 0$ $\psi_2 = 30$ $-30 \leq \psi_3 \leq 0$ $120 \leq \psi_3 \leq 180$		$\psi_1 = -30$ $\psi_2 = -30$ $-30 \leq \psi_3 \leq 30$ $120 \leq \psi_3 \leq 180$	$\psi_1 = 0$ $\psi_2 = 0$ $-60 \leq \psi_3 \leq 30$ $120 \leq \psi_3 \leq 210$
		ccc ($\omega_1 = \omega_2 = \omega_3 = 180°$)		
	$\psi_1 = 30$ $\psi_2 = 0$ $-60 \leq \psi_3 \leq 30$ $120 \leq \psi_3 \leq 210$		$\psi_1 = 30$ $\psi_2 = 30$ $-30 \leq \psi_3 \leq 30$ $150 \leq \psi_3 \leq 210$	

[a] The allowed ranges in ψ_1, ψ_2, and ψ_3 correspond to intramolecular conformational energies within 5 kcal/mol of the conformation of minimum energy of each conformer. The minimum energy conformations for tt and cc di-L-proline are $\psi_1 = \psi_2 = 0°$ and $\psi_1 = 0°$, $\psi_2 = 30°$, respectively, while the conformations of minimum energy for ttt and ccc tri-L-proline are $\psi_1 = \psi_2 = \psi_3 = 0°$ and $\psi_1 = 30°$, $\psi_2 = 0°$ and $\psi_3 = -30$ or 150°, respectively.

The partial charges, δ_j and δ_k, the nonbonded van der Waals interaction constants, a_{jk} and c_{jk}, the torsional barrier height, V_ψ^0, about C$^\alpha$–C, and the dielectric constant, $\epsilon = 3.5$, necessary to the evaluation of the intramolecular conformational energy of L-proline oligomers are taken from Brant, Tonelli, and Flory,[9] except for the partial charges on the N, C, and O atoms of the imide group, which are assigned the values used by Holzwarth and Chandrasekaran.[6] The nonbonded interactions involving N are assigned the constants a_{Nk} and c_{Nk} calculated according to previously described[9] methods.

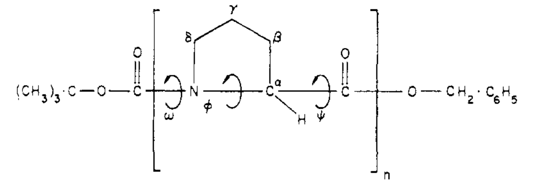

Figure 1. A schematic representation of the structure of the L-proline oligomers studied by Deber, *et al.*,[1] and in the present investigation.

The geometries of the *trans*- ($\omega = 0°$) and *cis*- ($\omega = 180°$) L-proline residues as reported by Sasisekharan[2] and by Traub and Shmueli,[3] respectively, are adopted here. The *tert*-butyl group at the N terminus and the benzyl group at the C terminus (see Figure 1) are replaced by methyl groups for the purposes of the present calculations.

The intramolecular conformational energies of the di- and tri-L-proline oligomers are determined by calculating, according to eq 1–5, the interactions between every pair of atoms or groups whose distance of separation,[10] r_{jk}, depends on one or more of the rotational angles ψ_1, ψ_2, ψ_3, ω_1, ω_2, and ω_3 (see Figure 1). The ψ rotational angles are varied in increments of 30° between $\psi = 0$ and 360°, while each ω is assigned the value 0 or 180° reflecting[4] the trans (t) or cis (c) conformation of the imide bonds. The C-terminal ester group is assumed to be exclusively in the trans conformation. The intramolecular conformational energies of the 16 and 32 conformers of the tetra- and penta-L-proline oligomers, respectively, are evaluated by neglecting all interactions beyond those between first and second neighbor proline residues; *i.e.*, only those interactions accounted for in the dimer and trimer are considered. On this basis the energy of the cctct conformer of the pentamer, for example, is given by $V_{\mathrm{cctct}} = V_{\mathrm{cct}} + V_{\mathrm{ctc}} + V_{\mathrm{tct}} - V_{\mathrm{ct}} - V_{\mathrm{tc}}$.

From the conformational energies of each of the conformers of di-L-proline (tt, tc, cc, and ct), of tri-L-proline (ttt, ttc, tct, tcc, ccc, cct, ctc, and cct), and of the tetramer and pentamer, the fraction of each conformer present, or its probability of occurrence, can be readily evaluated. For example, the fraction f_{tt} of di-L-proline in the all-trans conformation ($\omega_1 = \omega_2 = 0°$) is given by

$$f_{\mathrm{tt}} = SW_{\mathrm{tt}}/(SW_{\mathrm{tt}} + SW_{\mathrm{tc}} + SW_{\mathrm{cc}} + SW_{\mathrm{ct}}) \tag{6}$$

where

$$\begin{aligned}
SW_{\mathrm{tt}} &= \sum_{\psi_1,\psi_2} \exp[-V(\psi_1, \psi_2, \omega_1 = 0°, \omega_2 = 0°)/RT] \\
SW_{\mathrm{tc}} &= \sum_{\psi_1,\psi_2} \exp[-V(\psi_1, \psi_2, \omega_1 = 0°, \omega_2 = 180°)/RT] \\
SW_{\mathrm{cc}} &= \sum_{\psi_1,\psi_2} \exp[-V(\psi_1, \psi_2, \omega_1 = 180°, \omega_2 = 180°)/RT] \\
SW_{\mathrm{ct}} &= \sum_{\psi_1,\psi_2} \exp[-V(\psi_1, \psi_2, \omega_1 = 180°, \omega_2 = 0°)/RT]
\end{aligned} \tag{7}$$

Results and Conclusions

The energetically allowed ranges in ψ_1, ψ_2, and/or ψ_3 in the tt and cc di-L-proline and in the ttt and ccc tri-L-proline derivatives are presented in Table I. The allowed ranges of ψ_1, ψ_2, and ψ_3 correspond to intramolecular conformational energies within 5 kcal/mol

---

[9] D. A. Brant, A. E. Tonelli, and P. J. Flory, *Macromolecules*, **2**, 228 (1969).

[10] D. A. Brant and P. J. Flory, *J. Amer. Chem. Soc.*, **87**, 2791 (1965).

of the conformation of minimum energy for each conformer. Schimmel and Flory[5] and DeSantis, *et al.*,[11] found the calculated allowed ψ domain in poly-L-proline II (all imide bonds trans) to extend from ψ = 275 to 370° (10°). DeSantis, *et al.*,[11] varied the ψ angles in concert and considered only the repulsive nonbonded van der Waals interactions. Schimmel and Flory[5] considered all nonbonded van der Waals interactions (repulsive and attractive) dependent upon only a single ψ angle. Inclusion of the electrostatic interactions and the coarse choice of 30° ψ increments in the present study results in a narrowing of the allowed ψ_2 domain in the ttt tri-L-proline oligomer (see Table I). In addition, inclusion of the intrinsic torsional potential about the C^{α}–C bond (see eq 2) displaces the allowed domain toward larger values of ψ ($\psi_2 \approx$ 330–30°).

In view of the rather narrow energetically allowed ψ domains it would have been preferable to have performed the energy estimates and to have calculated the statistical weights (eq 7) using ψ increments smaller than the 30° value employed here. However, reducing the ψ increments to 10° would require a 27-fold increase in the number of conformations considered for each of the eight trimer conformers, making the computer calculation impractical. Partial confirmation of the validity of using 30° ψ increments is afforded by the observations that nearly all of the allowed ψ domains are rather flat exhibiting no sharp minima, and the breadths of the allowed domains are very similar for each of the conformers of the di- and tri-L-proline oligomers. On this basis it is hoped that the use of 30° increments has not resulted in the neglect of a significant fraction of the allowed ψ conformations.

Inspection of Courtauld space-filling molecular models[12] of the cc and ccc dimer and trimer conformers, respectively, in their allowed conformations (see Table I) show that severe steric interactions involving the *tert*-butyl or benzyl blocking groups are absent in these two conformers. Since the cc and ccc conformers are the most sterically constrained dimer and trimer conformers, it is believed that replacement of the *tert*-butyl and benzyl groups by methyl groups in the present energy calculations has only a minor effect on the calculated results.

The calculated fractions of each conformer of the di- and tri-L-proline oligomers present at 25° are compared in Table II with the corresponding conformer fractions obtained by Deber, *et al.*,[1] from the nmr analysis of the same L-proline oligomers in chloroform also at 25°. It is clear that the experimentally observed presence of substantial amounts of most of the conformers of the di- and tri-L-proline oligomers receives confirmation from the calculated intramolecular conformational energies reported here. Although lacking in quantitative agreement, the calculated relative amounts of the four and eight conformers of di- and tri-L-proline, respectively, generally agree qualitatively with the nmr findings of Deber, *et al.*[1]

Of the 16 conformers of tetra-L-proline, 8 are unmistakably present in the nmr spectrum. Ambiguity concerning the residue adjacent to the N terminus precludes a detailed accounting of the conformer fractions. Calculations show that four conformers (tttt, cttt, tttc, and cttc) should be present in amounts exceeding 1%. At the pentamer only the ttttt and cttttt conformers are detected experimentally, while the ttttt, ttttc, ctttt, and ctttc conformers are calculated to be present.

Table II. The Fractions of Conformers Present in Di- and Tri-L-proline Oligomers

Conformer	Fraction of conformer present: Exptl[a]	Fraction of conformer present: Calcd
tt	0.40	0.49
ct	0.40	0.38
tc	0.20	0.08
cc	0.0	0.05
ttt	0.30	0.48
ctt	0.30	0.25
ttc	0.10	0.19
ctc	0.10	0.07
tct	0.16	<0.01
cct	0.0	<0.01
tcc	0.04	<0.01
ccc	0.0	<0.01

[a] Values taken from ref 1.

Rifkind and Applequist[13] have reported that severe steric overlaps occur at the junction between an all-trans sequence followed by an all-cis sequence when proceeding from the N to C terminal. Their observations are based on the inspection of space-filling molecular models. These severe overlaps at the junction are absent when an all-trans sequence follows an all-cis sequence. These observations receive confirmation from the results presented in Table II (compare f_{ct} and f_{tc} and f_{ctt} and f_{ttc}). Both the experimental[1] and calculated fractions of conformers present in the di- and tri-L-proline oligomer derivatives are generally lower for those conformers where a cis residue succeeds a trans residue, as compared to the conformers where a trans residue succeeds a cis residue.

The preference for a trans sequence to succeed a cis sequence is also indicated in a recent nmr study[14] of the cis–trans isomerization of poly-L-proline. Torchia and Bovey[14] found that when poly-L-proline I (all imide bonds cis) is allowed to mutarotate in aqueous solution to form II (all imide bonds trans) the imide bond isomerization always begins from the C terminal. This leads to a junction only between a cis sequence followed by a trans sequence when proceeding from the N to C terminal.

The probabilities of finding a given residue in the trans conformation are presented in Table III for the dimer, trimer, tetramer, and pentamer. The calculated probabilities show that beginning with the trimer all nonterminal residues must be exclusively trans, while the trans conformation is favored, though not exclusively, for the terminal residues. On the other hand, the nmr data indicate that the trans conformation becomes exclusively preferred for all residues except the N terminus beginning with the pentamer. Thus, in chloroform the form II helical conformation, into which only the N terminus is not fully included, appears abruptly at the pentamer.

---

(11) P. DeSantis, E. Giglio, A. M. Liquori, and A. Ripamonti, *Nature* (*London*), **206**, 456 (1965).

(12) Obtained from Ealing Corp., Cambridge, Mass.

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

(14) D. A. Torchia and F. A. Bovey, in preparation.

Table III. Residue Conformation Probabilities for the Di-, Tri-, Tetra-, and Penta-L-proline Oligomers

Oligomer		Probability of finding residue in trans conformation				
		1 (N-terminus)	2	3	4	5
Dimer	Calcd	0.57	0.87			
	Exptl	0.60	0.80			
Trimer	Calcd	0.67	0.99	0.73		
	Exptl	0.60	0.80	0.76		
Tetramer	Calcd	0.67	0.99	0.99	0.73	
	Exptl	0.60	0.80[a]	0.80	0.80	
Pentamer	Calcd	0.67	0.99	0.99	0.99	0.73
	Exptl	0.60	0.99[a]	0.99[a]	0.99	0.99

[a] Assumed due to experimental[1] ambiguity.

The observed sudden onset at the pentamer of the form II helical conformation in poly-L-proline is not predicted by the present calculation of intramolcular interactions in L-proline oligomers. Only interactions between the fourth and first residues in the tetramer and between the fifth and the second and first residues in the pentamer have not been considered in the intramolecular conformational energy calculations performed here. Since the first and fourth residues and the first or second and fifth residues are separated by considerable distances, due to the extended conformations of the intervening residues (see Table I), their interactions might be expected to be rather weak (see eq 1–5). However, Rifkind and Applequist[13] noted that the hydrogen atoms on C^{α} and C^{β} are located at the van der Waals contact distance with the oxygen atom of the third nearest neighbor residue in the all-cis poly-L-proline I helix. This long-range attractive interaction, which is not accounted for in the calculations including second nearest neighbor residue interactions described here, may stabilize the all-cis helix, or, for that matter, the all-cis conformers of the L-proline oligomers higher than the trimer.

Presumably inclusion of third nearest neighbor interactions in the tetramer and pentamer would increase the calculated fraction of their all-cis conformers. This stabilization would work against the observed onset of the all-trans conformer at the pentamer.

Based on intramolecular interactions alone, the addition of a fifth L-prolyl residue to the tetramer will not result in the exclusive adoption of the trans conformation by the imide bonds in each of the residues except the N terminus, as is observed.[1] Thus, the present intramolecular approach to the study of the conformational characteristics of L-proline oligomers has not taken into account all significant contributions. The most obvious omission, but one that has yet to be successfully accounted for, is the failure to consider the interactions of the solvent molecules with the L-proline oligomers. The role played by solvent in the cis–trans isomerization of poly-L-proline is well established.[15,16] Although the present intramolecular calculations predict a preference for the trans imide bond conformation in the L-proline oligomers studied, the observed[1] abrupt onset of the form II helix at the pentamer in chloroform and the stability[13] of the form I helix in other solvents[15,16] result predominantly from solvent interactions.

Acknowledgments. The author is grateful to Drs. F. A. Bovey, D. A. Torchia, and C. M. Deber, and to Professors J. P. Carver and E. R. Blout for making available the results of their investigation prior to publication. In addition, several helpful discussions with Professor D. A. Brant and Dr. F. A. Bovey and the constructive criticism of the referees are also appreciated.

(15) F. Gornick, L. Mandelkern, A. F. Diori, and D. E. Roberts, *J. Amer. Chem. Soc.*, **86**, 2549 (1964).

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