1990

Proline side chain effects on theoretical [pi-pi] * absorption and circular dichroic spectra of proline-containing peptides

Kathryn Ambler Thomasson
Iowa State University

Order Number 9035121

Proline side chain effects on theoretical $\pi^{-}\pi^{*}$ absorption and circular dichroic spectra of proline-containing peptides

Thomasson, Kathryn Ambler, Ph.D.

Iowa State University, 1990

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

Proline side chain effects on theoretical $\pi-\pi^*$ absorption and circular dichroic spectra of proline-containing peptides

by

Kathryn Ambler Thomasson

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department: Chemistry
Major: Physical Chemistry

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1990

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

A. Why Circular Dichroism?

Circular dichroism (CD), the absorption difference between left and right circularly polarized light, is a form of optical activity that depends upon molecular geometry (1). For chromophores having a plane of symmetry (e.g., the peptide group NC'O), optical activity is a result of the interacting chromophores and perturbing groups (1). Thus, optical activity measurements have been widely used to monitor conformational states in biopolymers (1). A theoretical method applied to peptide structures in the past is the dipole interaction model (2, 3, 4-9). The major feature of the model is the inclusion of the polarizabilities of all parts of the peptide (including side chains) in the interaction problem (4, 5).

The dipole interaction model, used to predict $\pi-\pi^*$ absorption and CD spectra of peptides, was the first model to successfully predict the CD spectra for α-helices $(Gly)_n$, $(L\text{-}Ala)_n$, and $(D\text{-}Ala)_n$. The calculated spectra were missing a negative peak near 180 nm which had been predicted by previous models, but was missing from experiment (8). The same parameters developed for these α-helices also produced predicted $\pi-\pi^*$ absorption and CD spectra that were approximately correct for poly(L-proline) I and II (2). Moreover, the predictions for poly(L-proline) II were better than those predicted by previous models (2).

The dipole interaction model has also been used to successfully predict $\pi-\pi^*$ absorption and CD spectra for polypeptide β-structures (10) and collagen (11). Furthermore, both studies indicated the importance of the

presence of larger amino acid side chains (10, 11). Variations of side chain conformations were studied for poly(L-α-aminobutyric acid), and some backbone regions had predicted CD that were very sensitive to the side chain structure (12). The theoretical π-π* CD spectra of cyclic dipeptides were also side chain dependent (especially for proline), and only certain allowed side chain conformations gave spectra that were in reasonable agreement with experiment (3). Thus, a thorough examination of proline side chain conformations in conjunction with the dipole interaction predictions for π-π* absorption and circular dichroism may provide insight into the conformations of various proline-containing polypeptides (i.e., poly(L-proline) I and II, cyclo(Pro-Gly)$_3$, and cyclo(Gly-Pro-Gly)$_2$).

B. Why the Proline Side Chain?

Proline, one of the 20 naturally occurring amino acids, is of unique interest because it can have a profound influence on the conformations of peptide segments (13). The proline residue has restricted conformational freedom because the peptide nitrogen, α carbon, and three side chain carbons form a five membered ring. As a result, proline is seldom located in α-helical or β-sheet structures of proteins because it destabilizes these structures; on the other hand, proline is often found in bend or loop structures where the polypeptide chain reverses direction (14). In spite of restricting the backbone of the peptide, the proline ring still has some flexibility through puckering of the five-membered ring which can take on many conformations (15). Such flexibility has often been referred to as pseudorotational mobility (16). A fast and easy method of characterizing the wide range of possible proline ring structures will be useful.

Poly(L-proline) is a helical synthetic peptide whose two backbones have been well characterized in the literature. The dipole interaction model has provided excellent predictions for the CD spectra of both forms using the X-ray structures (2). Only the X-ray conformation of the proline rings were included in the earlier study (2). This makes this molecule a prime candidate for exploring different proline ring conformations on the two backbones and the effect the different ring conformations have on the predicted CD spectra.

C. Why Cyclic Peptides?

Cyclic peptides are compounds of interest for two major reasons: they are of biological relevance, and they have few discrete conformational states as opposed to the large multitude of states available to linear peptides (1). Many cyclic peptides are naturally occurring and of biological significance. There are many examples of such molecules. Cyclic peptide antibiotic ionophores include valinomycin (17, 18), enniatin, beauvericin, monamycin, and antamanide (18). Other cyclic peptide antibiotics include serratamolide, tuberactinomycin, viomycin, ilamycin, polymixin, bacitracin, tyrocidine, vernamycin, patricin, stendomycin, telomycin, actinomycin (18), and gramicidin S (17, 18, 19, 20).

Some naturally occurring cyclic peptides are toxins. Examples include: amanitin, phallotoxin (18), tentoxin (21, 22), rosetoxin B (23), and cyclochlorotine (24). Other natural cyclic peptides are hormones; e.g., oxytocin, vasopresin, and tociamide (18). Still others contain iron; e.g., ferrichrysin (25) and asperchrome A (26).

Synthetic cyclic peptides have also been used in the past as models for natural peptides and protein sequences, to demonstrate β- or γ-turns, or to function as ionophores (15, 27). Examples include

BOC-Cys-Gly-Pro-Cys-NHMe

to mimic the active site of thioredoxin (28) and cyclo(Pro-Phe-D-Trp-Lys-Thr-Phe) to mimic somatostatin (29).

D. Why Cyclohexapeptides?

Of all the synthetic cyclic peptides studied in the literature cyclohexapeptides are among the most extensively studied (27). Even with the restrictions of cyclization, however, multiple conformers separated by low energy barriers exist for cyclohexapeptides making structural characterization difficult in some cases (27). For example, the simplest cyclohexapeptide, cyclo$(Gly)_6$, has been shown by energy minimizations to have 24 possible symmetric structures representing local energy minima (30) and at least 81 possible asymmetric structures (31). Addition of proline to cyclohexapeptides reduces the number of conformations considerably (1); moreover, proline-containing cyclohexapeptides having high sequential symmetry have certain preferred conformations (32). Furthermore, there is a bonus to studying proline-containing cyclohexapeptides; they can achieve cis or trans peptide bonds (1). Although trans peptide bonds are more common, cis peptide bonds are found in proteins and naturally occurring cyclic peptides like antamanide (1). Synthetic cyclic peptides having both cis and trans peptide bonds make excellent models for these kinds of molecules.

Examples of cyclohexapeptides containing proline which have had their backbones characterized in the literature and which will be pursued here are cyclo(Pro-Gly)$_3$ and cyclo(Gly-Pro-Gly)$_2$. These peptides are of interest for the above mentioned reasons and the following. First, cyclo(Pro-Gly)$_3$ has only 13 C_3 symmetric conformers (33-36) and 18 asymmetric ones (33, 36-38). Second, cyclo(Pro-Gly)$_3$ is capable of binding and transporting cations (32) making it an excellent model for ionophores like antamanide. Third, cyclo(Gly-Pro-Gly)$_2$ has only 11 C_2 symmetric structures (1, 39-41) and 1 asymmetric one (42). Fourth, cyclo(Gly-Pro-Gly)$_2$ has been shown to exhibit β-turns (1, 39-42) and γ-turns (1). The spectra of both the molecules will be studied using the dipole interaction model.

II. METHODS

A. Structure Generation

Schematic portraits of the structures generated are given in Figure 1. For a given set of structural parameters, a peptide or proline ring is generated by placing N (or a point representing N) and the atoms C^{α} and C' in a coordinate plane and adding the remaining atoms successively from given bond lengths, bond angles, and torsion angles, following the method of Ramachandran and Sasisekharan (43). This method states that given the location of three atoms in a chain with bond angle τ, the torsion angle χ, and distance between the third and fourth atom in a chain, the location of this fourth atom can be determined by the following equation:

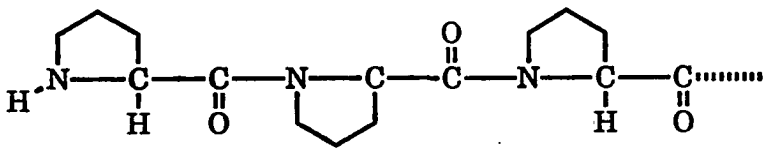
$$\underset{\sim}{r}_4 = \underset{\sim}{r}_3 + [M_u^{\chi}][M^{\pi-\tau}{}_n]\underset{\sim}{v} \qquad (1)$$

where $\underset{\sim}{r}_4$ and $\underset{\sim}{r}_3$ are the position vectors of atoms 4 and 3, $[M_u^{\chi}]$ is the rotation matrix for the torsion angle, $\underset{\sim}{u}$ is the unit vector pointing from atom 2 to atom 3, and $[M_n^{\pi-\tau}]$ is the rotation matrix for the bond angle τ, and $\underset{\sim}{n}$ is the unit vector of $\underset{\sim}{p}\times\underset{\sim}{q}$ (see Figure 2). $\underset{\sim}{v}$ is given by

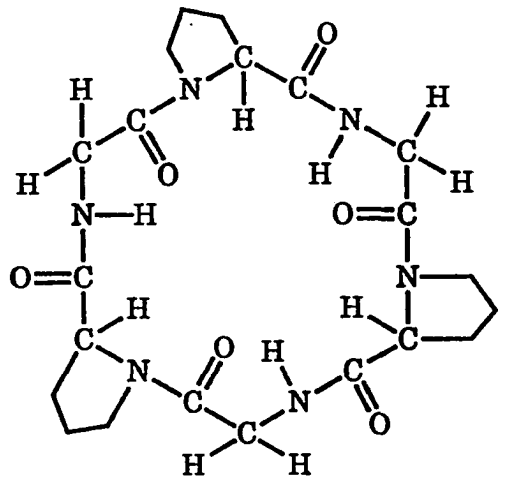
$$\underset{\sim}{v} = l\underset{\sim}{u} \qquad (2)$$

where l is the distance between atoms 3 and 4.

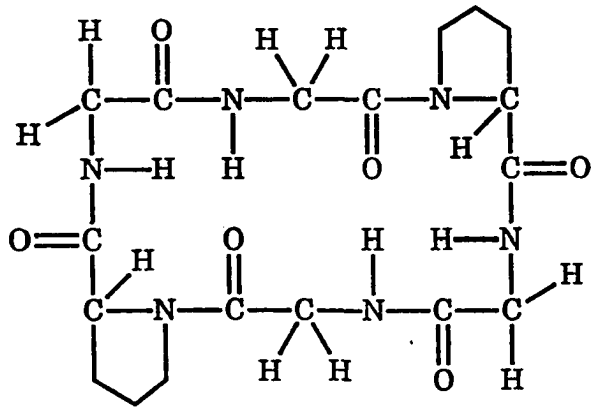
For example, the proline ring atoms are added in the order of C^{β}, C^{γ}, C^{δ} using N, C^{α}, C^{β} as the first three atoms of the chain. C^{β} is located tetrahedrally as follows (44). Derivations with figures are provided in Appendix 1.



poly (L-Proline)



cyclo(Pro-Gly)$_3$



cyclo (Gly-Pro-Gly)$_2$

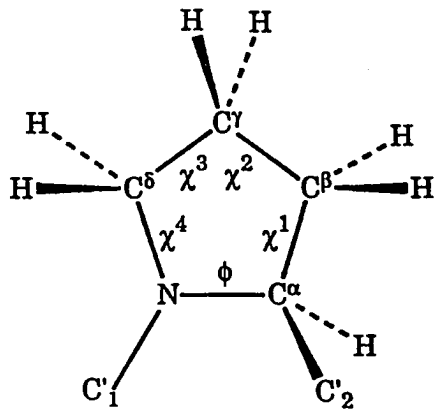
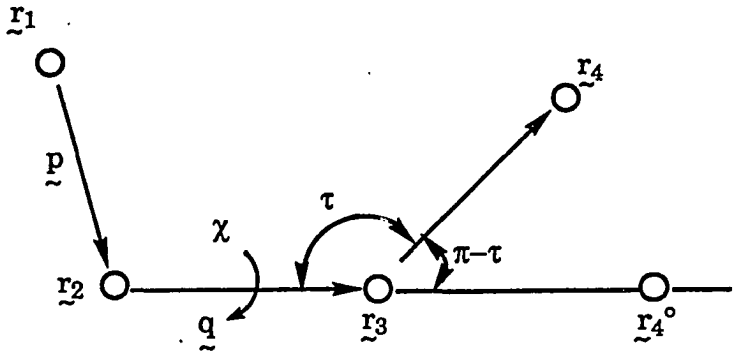


Figure 1. Schematic diagram of the proline-containing structures studied



$$\underset{\sim}{p} = \underset{\sim}{r}_2 - \underset{\sim}{r}_1 \qquad \underset{\sim}{q} = \underset{\sim}{r}_3 - \underset{\sim}{r}_2$$

$$\underset{\sim}{n} = \frac{(\underset{\sim}{p} \times \underset{\sim}{q})}{|\underset{\sim}{p} \times \underset{\sim}{q}|}$$

$$\underset{\sim}{u} = \frac{\underset{\sim}{q}}{|\underset{\sim}{q}|}$$

Figure 2. Four atoms in a chain: Basis for sequential generation of structures

$$C_x^\beta = 1.54 \cos \tau_2$$

$$C_y^\beta = 1.54 \, (\cos \tau_1^R - \cos \tau_1 \cos \tau_2)/(-\sin \tau_1) \qquad (3)$$

$$C_z^\beta = (1.54^2 - C_x^{\beta 2} - C_y^{\beta 2})^{1/2}$$

where C_x^β, C_y^β, and C_z^β are the x, y, and z coordinates of C^β; 1.54 and 1.47 Å are the C^α-C^β and N-C^α bond lengths, respectively. τ_1 = $\langle NC^\alpha C'$; τ_2 = $\langle C'C^\alpha C^\beta$; τ_1^R = $\langle NC^\alpha C^\beta$. The hydrogens on C^β, C^γ, and C^δ are placed by arranging a local two-fold symmetry of the bonds attached to the carbon, giving approximately tetrahedral local symmetry (Equations 4 and 5) (44) (see Appendix A for derivations).

$$\cos \theta = \cos \left(\frac{\alpha}{2}\right) \sin (\tau_2^R/2)/(1 - \cos^2 \left(\frac{\alpha}{2}\right) \cos (\tau_2^R/2))^{1/2} \qquad (4)$$

where θ is the angle between the $C^\alpha C^\beta C^\gamma$ plane and the $C^\alpha C^\beta H$ plane and α is the $\langle HCH$.

$$\sin \zeta = \sin (\tau_2^R/2)/(\cos^2 \theta + \sin^2 \theta \sin^2 (\tau_2^R/2))^{1/2} \qquad (5)$$

where ζ is the $C^\alpha C^\beta H$ angle. So that

$$\underset{\sim}{H}_1^\beta = \underset{\sim}{C}^\beta + [M^{\chi^1 + \pi - \theta}][M^{\pi - \zeta}] 1 \underset{\sim}{u}$$

when 1=CH, and $\underset{\sim}{u}$ is the unit vector between C^α and C^β (44). The bond angles $\langle HCH$, $\langle CCH$, and τ_2 are fixed at 109.5°; and $\langle C'NC^\alpha$ (τ_6) is fixed at 123°. Fixed bond lengths (in Å) are: C-H 1.095 (46), C^α-C' 1.53, C'-N 1.32, N-C^α 1.47 (43), and ring C-C 1.54 (2).

The glycine side chain is trivial compared to proline. First, one hydrogen is located just as C^β except that the C-H bond length is used in

place of the C^{α}-C^{β} length. Again the second hydrogen is located by reflecting the first through the $NC^{\alpha}C'$ plane.

An entire peptide backbone chain is located using the above method beginning with a single residue. C^{α}, C' and a hypothetical point where N would be for the preceding residue, are located in the same plane. The first atom in any chain is always C^{α}, so that the hypothetical N is not present in the final structure unless the final molecule is a closed cyclic peptide. Bond angles (in degrees) and bond lengths (in Å) are fixed as in proline with the following additions: $\langle C^{\alpha}C'O$ (τ_3) 121, $\langle C^{\alpha}C'N$ (τ_4) 114, $\langle C'NH$ (τ_5) 123, $\langle C^{\alpha}NH$ 114, C'-O 1.24, N-H 1.00 (43). The final step is generation of the peptide chain by means of translations and rotations determined by the peptide unit dimensions and the backbone torsions, using an extension of the method of Ramachandran and Sasisekharan (43) which permits ω to take arbitrary values. Once a structure is generated using backbone dihedral angles from the literature and the molecule is cyclic, a geometric optimization procedure is applied to close the ring more precisely so that the final N is as close as possible to the location of hypothetical one.

B. Backbone Ring Closure

The geometric method of ring closure regards certain structural parameters as variables to be adjusted by nonlinear least squares optimization to achieve ring closure. These variables include all torsion angles ϕ and ψ, and one bond angle τ_1. τ_1 was treated as the same for all residues. The objective of backbone ring closure is to close the ring with an optimum fit of these variables to the literature ϕ and ψ angles. The

torsion angles ϕ, ψ, and ω are defined by the atoms $C'NC^{\alpha}C'$, $NC^{\alpha}C'N$, and $C^{\alpha}N'NC^{\alpha}$ respectively. When no τ_1 was listed by the source, the value 109.5° was targeted. The optimum was defined by a minimum in the sum of squares of residuals. Table 1 lists the variables and residuals used for various symmetries of cyclohexapeptides. There were different variable sets for different symmetries because backbone closure variables are symmetry dependent for cyclohexapeptides.

Go and Scheraga have also developed a geometric ring closure method (45) which utilizes local conformational deformations. Their method uses a set of algebraic equations to solve for six unknown dihedral angles where the user must choose n-6 independent dihedral angles arbitrarily (n is the total number of dihedral angles in the cyclic peptide) (45). Not all choices of n-6 dihedral angles will give solutions to the set of algebraic equations (45). Both methods require a structure that is near closure to obtain good solutions. Both methods use the Pauling-Corey peptide geometry (46), but Go and Scheraga's method does not directly address the value of τ_1 which can have a significant effect on ring closure.

Asymmetric peptides (C_1) have the most variables. ψ_1 is not included among the variables for these forms because it only sets the direction of the chain. When symmetry occurs, the number of variables is reduced because the peptide repeats. For C_2 forms: $\phi_1 = \phi_4$; $\phi_2 = \phi_5$; and $\phi_3 = \phi_6$. For C_3 forms: $\phi_1 = \phi_3 = \phi_5$; $\phi_2 = \phi_4 = \phi_6$. ψ and ω have the same relationships. ω angles are not used as variables because during optical calculations they are assumed to be planar so that the literature inputs (which are generally planar) are left as constants. The bond angles and

Table 1. Variables and residuals for backbone ring closure

Variables			
C_1	C_2	C_3	Residuals[a]
ϕ_1	ϕ_1	ϕ_1	$\phi_1^L - \phi_1^c$
ϕ_2	ϕ_2	ϕ_2	$\phi_2^L - \phi_2^c$
ϕ_3	ϕ_3	ψ_1	$\phi_3^L - \phi_3^c$
ϕ_4	ψ_1	ψ_2	$\phi_4^L - \phi_4^c$
ϕ_5	ψ_2	τ_1	$\phi_5^L - \phi_5^c$
ϕ_6	ψ_3		$w_1\ (\phi_6^L - \phi_6^c)$
ψ_2	τ_1		$w_2\ (\psi_1^L - \psi_1^c)$
ψ_3			$\psi_2^L - \psi_2^c$
ψ_4			$\psi_3^L - \psi_3^c$
ψ_5			$\psi_4^L - \psi_4^c$
ψ_6			$\psi_5^L - \psi_5^c$
τ_1			$\psi_6^L - \psi_6^c$
			$w_3\ (\omega_6^L - \omega_6^c)$
			$\tau_1^L - \tau_1^c$
			$w_4\ (123.0 - \tau_6^c)$
			$w_5\ (1.470 - NC^{\alpha}{}_1{}^c)$

Symmetry	w_1	w_2	w_3	w_4	w_5
C_1	1	1	10	10	1000
C_2^1	10; 100	10; 100	10; 100	10; 100	1000; 10,000
C_3^2	10	10	10	10	1000

[a]The subscripts on ϕ, ψ, and τ refer to the residue number. The superscript L refers to literature value. The superscript c refers to the calculated value. All the residuals are in degrees except for the NC_1^{α} residual which is in Å.

the bond length of NC^{α} are chosen among the residuals because they occur around the N_6-C_1^{α} bond where ring closure occurs.

The weighting factors vary with symmetry as well. The asymmetric forms only have substantial weights on angles about the $N-C^{\alpha}$ bond because no symmetry needs to be maintained. C_2 and C_3 symmetric forms need ϕ_6 and ψ_1 weighted more in order to preserve the symmetry. The NC^{α} bond length has the largest weight to bring this residual to the same order of magnitude as the angle residuals. Once backbone closure was achieved, the proline ring was closed using the resulting ϕ and τ_1 values.

C. Proline Ring Closure

The method of proline ring closure is very similar to that of the backbone ring except that it is not necessary to begin with a set of parameters that have been previously optimized in the literature. The proline ring is closed by a technique known as bond optimized ring closure (BORC) based on an arbitrary choice of two independent variables ϕ and χ^2 (47). The essential feature of BORC is that an optimum fit, rather than an exact fit, is found. The term "bond optimized" is intended to distinguish BORC from various energy-optimization methods that are widely used in structural calculations (and which formed the beginnings of the published backbone rings). There exists a related method of optimizing rings whereby the iterative procedure begins with a planar polygonal structure instead of a chain generated stepwise (48). Both methods find an optimum fit to target values for the bond parameters by nonlinear least squares procedures.

When closing a proline ring only part of the structure is necessary (Figure 1). The atoms included are those in the proline ring and the two attached carbonyl carbons C_1' and C_2'. The proline ring is generated stepwise fron N, C^{α}, and C^{β} as described in Section A. The structural parameters to be adjusted to achieve ring closure become the $N\text{-}C^{\delta}$ bond length and all ring bond angles and certain torsion angles. The objective of BORC is to achieve ring closure with an optimum fit of these parameters to certain target values when a minimum set of independent torsion angles are specified.

It was found that, with few exceptions, a unique closed ring structure could be obtained when two torsion angles about ring bonds were fixed arbitrarily within certain limits. ϕ was chosen because it is fixed by the conformation of the peptide chain; χ^2 was chosen because it does not involve any backbone bonds and corresponds to the usage of Anteunis and Sleeckx in their extensive review of proline-containing linear peptides (13).

For a given ϕ,χ^2 pair, the adjustable parameters χ^1, $\langle NC^{\alpha}C^{\beta}$ (τ_1^R), $\langle C^{\alpha}C^{\beta}C^{\gamma}$ (τ_2^R), and $\langle C^{\beta}C^{\gamma}C^{\delta}$ (τ_3^R) are varied from a set of initial trial values to achieve optimum ring closure. For proline, the optimum is defined by a minimum in the sum of the squares of fractional deviations of the following quantities from their target values: $N\text{-}C^{\delta}$, τ_1^R, τ_2^R, τ_3^R, $\langle C^{\gamma}C^{\delta}N$ (τ_4^R), $\langle C^{\delta}NC^{\alpha}$ (τ_5^R), and $\langle C_1'NC^{\delta}$ (τ_6^R). The target values are the mode values in the survey of published proline structures described in the next chapter on proline, except that a slightly different target value of 125° was taken for τ_6^R in order to obtain a sum of 360° for the target values of

bond angles about N. The initial trial values of adjustable bond angles were also set to the target values. The initial value of χ^1 and χ^3 were taken as χ^2 and $-\chi^2$ in separate trials, as the optimum ring closure parameters sometimes depended on the sign of the initial values.

The above described use of the torsion angles alone to define the proline ring conformation differs from the practice of Anteunis and Sleeckx (13) and other workers (49) who use a pseudo-rotational phase angle P and a maximum pucker angle χ_M, which are related to the ring torsion angles by (13)

$$\tan(P) = (\chi^1 - \chi^3 + \chi^4 - \chi^5)/(3.077\chi^2)$$

$$\chi_M = \chi^2/\cos(P)$$

$$\chi^j = \chi_M \cos[P + 144(j-2)],\ j = 2, 3, 4, 5, 6$$

where χ^5 is the torsion angle defined by $C^{\delta}NC^{\alpha}C^{\beta}$, and $\chi^6 = \chi^1$. This treatment has in common with ours the definition of the ring conformation by two variables. The correspondence is a bit loose because we use the backbone torsion angle ϕ in place of χ^5, and the difference between the absolute value of $\chi^5 - \phi$ is found to be in the range of 65 ± 15 (13), depending on the bond geometry about N and C^{α}.

D. Optical Calculations

The theoretical method of calculating π-π^* absorption and CD spectra is the dipole interaction model (4, 5). It has been successfully applied to cyclic peptides and helices in the past (2, 3, 6-9). The major feature of the model is the inclusion of the polarizabilities of all parts of the molecule in the interaction problem. Dispersive parameters (4, 5) are required only for the π-π^* transition of the NC'O group, and all other

atoms are treated as nondispersive polarizabilities (8, 9, 50). The molar absorption coefficient, ε, is expressed in terms of the normal mode dipole strength, D_k, and CD, $\Delta\varepsilon$, is expressed in terms of the rotational strength, R_k (8, 9).

$$\varepsilon = (8\pi^2 \bar{\nu}^2 N_A \Gamma/6909n) \sum_{k=1}^{q} [D_R/(\bar{\nu}_k^2 - \bar{\nu}^2)^2 + \Gamma^2\bar{\nu}^2]$$

$$\Delta\varepsilon = (32\pi^3\bar{\nu}^3 N_A \Gamma/6909n) \sum_{k=1}^{q} [R_k/(\bar{\nu}_k^2 - \nu^2)^2 + \Gamma^2\bar{\nu}^2]$$

where $\bar{\nu}$ is the frequency of the light wave in wavenumbers; N_A is Avogadro's number; Γ is the half peak bandwidth; q is the number of dispersive oscillators; n is the number of residues, and $\bar{\nu}_k$ is the normal mode wavenumber (8, 9). The normal mode dipole and rotational strengths are defined as $D_k = \mu^k \cdot \mu^k$ and $R_k = \mu^k \cdot m^k$ respectively. μ^k is the normal mode electric dipole moment, and m^k is the normal mode magnetic dipole moment (8,9).

In this model, the NC'O group is treated as a single unit, located at the center of the N-C' bond, whose polarizability consists of a complex dispersive contribution from the π-π^* transition and a nondispersive, anisotropic contribution from all other electronic transitions; all other atoms are assigned nondispersive, isotropic polarizabilities. The values for all transition and polarizability parameters are the semiempirical values used in previous polypeptide calculations (7). In particular, the intrinsic wavelength of the π-π^* transition (170.3 nm) is left unchanged, so that any shifts in the spectrum will arise solely from the interactions inherent in the model (2).

All computations were carried out on a NAS 9160 computer. Nonlinear least squares optimization was done using IMSL subroutine ZXSSQ (51). Standard deviations of the adjustable parameters for proline were estimated as described elsewhere (52).

III. PROLINE: COMPARISON OF LITERATURE AND BORC RINGS

A. Introduction

The proline ring is of interest to theoretical studies using the dipole interaction model because previous studies have indicated that CD spectra are sensitive to the proline ring structure (2, 3). The method, BORC, described in the previous chapter, provides a means to carry out systematic studies of these optical properties. In this chapter, the BORC method is tested by, and the method's success is gauged by a comparison of predicted ring conformations with 191 experimentally observed conformations covering a wide range within the broad limits permitted by ring closure. This chapter represents a more detailed expansion of the material in another paper (47). The program itself is listed in Appendix B.

There is a related study of the conformation of proline by DeTar and Luthra (53), who generated proline rings by energy minimization of the molecule actylproline methyl ester. They compared the proline conformations in 40 observed X-ray structures with their calculated energies, obtaining results whose main features are similar to those obtained by our geometric method of optimizing the structures and a larger sample of known structures.

In Appendix B a copy of the FORTRAN code for BORC is listed. It follows the method described in Chapter II and tested in this chapter. The steps are well documented with comment statements. The main program is the basis of BORC. It calls ZXSSQ to optimize a series of proline rings. FUNC is the subroutine that calculates the residuals, F, for ZXSSQ. PROCRD is the subroutine that produces the cartesian coordinates for the proline

fragment. Subroutine ENERGY calculates the proline fragment's energy; the parameters of Set 1 described in Section C of this chapter are used in this listing. COOUT and TRIOUT print the cartesian coordinates and a triangular matrix, respectively. Other subroutines called but not listed are either IMSL routines or subroutines developed by Dr. Applequist for short repeated calculations (44).

B. Survey of Proline Structures

Table 2 contains the structural parameters obtained from published X-ray crystallographic data of 244 proline rings found in small natural or synthetic peptides (34,35,37,42,54-169). As described in the previous chapter, the relevant parameters for this study are the N-C^{δ} distance and the ring bond and torsion angles. The ring C-C distances are not surveyed because they tend to be systematically underestimated in X-ray diffraction studies (2). The more accurate alkane C-C distance (1.54 Å) (46) is used to avoid problems with optical calculations such as those encountered in a previous study (2). The distribution of values for each parameter is unimodal for the bond angles and length (Figure 3), with the modes, means, and standard deviations given in Table 3. The torsion angles χ^1 show bimodal distributions (Figure 4), reflecting the two conformational regions for the ring referred to as "up" and "down" (170). Some of the X-ray structures included two locations for C^{γ}, indicating a mixture of discrete conformations; in these cases, each location was treated as a separate structure for the purpose of this survey. The sum of the angles about N is included in Table 3 to show that the bond geometry rarely deviates significantly from the planar form. The mean of the sum differs slightly

Table 2. Structural parameters for proline from X-ray crystals of proline-containing peptides

ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_6	τ_6^R	τ_5^R	N-C$^\delta$	Ref.
-65.0	-32.4	40.7	-32.5	12.5	102.4	103.5	103.1	103.1	120.6	124.6	111.8	1.476	54
-59	-3.7	5.9	-5.6	2.9	103.0	108.0	112.9	104.1	118.7	129.5	111.7	1.46	55
-62	-17	28.2	-28.5	18	106.1	102.9	106.2	104.6	120.3	128.0	111.7	1.47	55
-89.1	36.4	-35.8	20.4	2.9	103.5	104.9	103.9	101.7	127.2	121.0	112.8	1.478	56
57.3	10.9	-13.2	9.7	-2.3	103.5	107.5	101.3	104.2	120.4	127.6	112.1	1.469	57
82.6	-29.5	35.5	-27.0	9.1	102.4	105.2	103.5	104.1	119.7	127.1	112.7	1.477	57
-61.2	-15.5	29.8	-32.0	23.0	109.1	105.4	106.0	103.8	117.5	132.0	110.4	1.467	58
-67.2	19.0	-30.0	28.2	-16.7	103.7	105.4	106.3	102.7	120.3	126.3	112.8	1.473	59
-68.6	-5.6	22.5	-30.3	27.3	104.8	106.0	106.3	105.6	118.9	132.6	108.4	1.472	60
-69.6	26.5	-32.2	24.9	-7.9	104.2	104.3	106.4	103.7	117.5	130.9	111.5	1.458	61
-88.2	32.9	-37.4	26.7	-6.1	101.8	104.1	103.4	104.2	119.8	126.9	112.7	1.470	62
-62.3	-18.5	31.4	-31.6	20.5	103.7	104.7	105.3	103.6	119.7	128.3	112.0	1.458	63
-63.2[a]	-7.7[a]	13.7[a]	-14.5[a]	9.8[a]	103.4	107.9	110.7	104.3	121.3	126.8	111.8	1.473	64

[a]Calculated by the method of Anteums and Sleeckx (13). All angles are in degrees; the NC$^\delta$ bond length is in Å. Torsions are defined in Figure 1. Bond angles are defined as follows: τ_1^R = $\langle NC^\alpha C^\beta$; τ_2^R = $\langle C^\alpha C^\beta C^\gamma$; τ_3^R = $\langle C^\beta C^\gamma C^\delta$; τ_4^R = $\langle C^\gamma C^\delta N$; τ_5^R = $\langle C^\delta NC^\alpha$; τ_6^R = $\langle C'NC^\delta$; τ_6 = $\langle C'NC^\alpha$.

Table 2. Continued

ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_6	τ_6^R	τ_5^R	N-C$^\delta$	Ref.
-67.1[a]	24.0[a]	-33.0[a]	29.4[a]	-14.6[a]	102.3	105.3	104.8	104.2	120.8	126.4	112.8	1.475	64
-58.0	-37.7[b]	46.2[b]	-44.0[b]	26.3[b]	106.0	103.0	117.0	102.0	118.0	130.0	111.0	1.47	60
70.8[a]	-20.3[a]	34.1[a]	-34.9[a]	22.3[a]	103.1	104.8	103.9	102.7	124.0	122.8	112.8	1.483	65
61.5[a]	-28.0[a]	39.1[a]	-35.3[a]	18.0[a]	102.8	103.5	104.3	101.7	119.8	127.3	112.6	1.478	65
-78.3[a]	32.2[a]	-35.8[a]	25.6[a]	-5.7[a]	102.8	103.3	104.8	103.8	126.5	120.5	112.9	1.486	65
-70	-10	21	-23	17	102.5	108.1	110.4	103.6	126.0	120.5	111.9	1.471	66
84	-17	20	-14	3	103.6	106.8	108.0	104.6	117.5	130.6	111.7	1.465	66
-67	22	-30	25	-11	103.3	104.9	108.0	101.5	123.4	122.5	114.1	1.464	67
-70	24	-31	25	-10	103.2	104.3	107.8	108.1	125.5	121.7	112.8	1.459	68
-72	31	-35	24	-4	102.0	104.1	105.8	102.5	124.9	121.2	113.6	1.45	69
--	21	-32	31	-20	102	108	102	104	--	--	113	1.48	70
-51	-31	37	-30	12	103	104	101	107	120	130	111	1.47	70
-63	-7	14	-14	9	103	108	111	104	121	127	112	1.47	70
-67	24	-33	29	-14	104	105	105	102	121	126	113	1.45	70
70	-32	40	-32	13	102.3	103.3	103.8	102.1	119.5	127.7	112.7	1.466	71
-70	-10	25	-20	24	103.3	106.5	105.0	102.8	123.8	123.0	113.0	1.470	71
70	-28	40	-36	17	103.8	101.0	106.8	100.7	119.3	128.8	112.0	1.443	71
-72	28	-36	29	-10	104.1	101.7	106.6	104.7	124.4	125.0	110.5	1.452	71

[b]Calculated from the published coordinates.

Table 2. Continued

ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_6	τ_6^R	τ_5^R	N-C$^\delta$	Ref.
-61	5	1	-7	11	103.5	105.2	114.5	101.2	123.6	120.8	114.5	1.469	72
-64.8	-17.9	32.2	-33.4	22.8	103.8	104.7	105.4	101.4	124.9	122.1	113.0	1.455	73
-53.3	-26.5	34.2	-27.3	10.0	101.8	104.3	106.7	103.2	119.1	127.8	113.0	1.481	73
-60.0	-19.0	34.5	-36.4	26.0	103.3	104.5	103.1	102.5	123.9	123.3	112.6	1.464	73
-59	-28	36	-30	12	103.5	102.9	105.4	102.4	120.5	125.9	113.0	1.465	74
-55.9[b]	29.4[b]	-39.8[b]	33.6[b]	-15.2[b]	103.3[b]	102.8[b]	104.2[b]	102.8[b]	118.9[b]	127.8[b]	111.5[b]	1.462	75
-80.1	32.0	-38.8[b]	30.1	10.3	103.0	103.3	103.4	103.3	121.1	126.2	112.2	1.484	76
-102.6[a]	12.3[a]	-28.8[a]	34.2[a]	-26.7[a]	98.8	115.2	95.1	107.8	--	--	111.6	1.50	77
-40.4[b]	-13.5[b]	30.8[b]	-34.9[b]	28.0[b]	103.7	103.0	106.6	98.3	114.8	125.4	116.2	1.43	77
--	-16.3[b]	0.9[b]	14.8[b]	-25.8[b]	104.5	107.9	106.8	105.4	--	--	108.4	1.510	78
--	15.8[b]	-.7[b]	-14.6[b]	25.1[b]	104.2	108.2	107.9	104.6	--	--	108.6	1.502	78
-67.3	-7.4	22.0	-26.9	23.1[b]	104.2	107.3	106.9	103.3	121.8	127.0	111.6	1.488	79
-63.7	23.5[a]	-29.8[a]	24.7[a]	-10.2[a]	103.6	104.7	107.4	101.7	123.9	121.8	114.2	1.456	80
-75.3	36.8	-39.4	26.3	-3.1	101.8	103.6	103.5	103.3	126.0	121.7	111.8	1.474	81
-66.6	27.0	-31.1	22.7	-5.3	102.6	105.3	106.9	102.7	124.1	122.0	113.3	1.456	82
-95.1	38.4	-35.7	18.9	5.9	103.3	102.2	106.3	103.3	121.1	128.6	109.5	1.480	82
-71.8	28.5	-33.1	24.2	-5.8	103.7	103.6	106.4	103.4	120.1	127.5	112.3	1.474	83
-59	25	-27	19	-3	105.4	102.0	108.9	102.4	123.5	120.8	113.8	1.465	84
-66	17	-20	14	-2	103.5	104.6	111.0	104.2	118.4	128.3	113.1	1.471	84
-73	29	-38	31	-13	104.1	103.5	104.0	102.9	121.0	127.2	111.8	1.482	84

Table 2. Continued

ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_6	τ_6^R	τ_5^R	N-C$^\delta$	Ref.
--	-30.2	30.7	-18.0	-1.9	103.4	104.4	109.1	106.3	--	--	107.0	1.504	85
-70.8	34.7	-34.3	20.3	1.9	101.9	103.8	103.8	104.9	124.3	121.6	112.5	1.461	85
-54.3	16.0	-20.8	16.5	-5.5	103.0	105.9	111.0	103.1	122.7	122.9	113.2	1.462	86
-66.7	-3.6	11.7	-14.9	12.6	102.5	108.6	111.3	102.9	119.4	127.8	112.8	1.458	87
-69.4	24.3	-22.3	10.6	5.8	102.4	106.8	110.0	102.3	121.4	126.0	112.7	1.478	87
-55	-39	46	-35	10	98.1	101.2	101.1	105.9	120.7	125.9	112.2	1.485	88
-55	28	-41	37	-20	106.9	102.2	103.2	99.6	120.7	125.9	112.2	1.485	88
-62	27	-38	34	-18	103.0	103.5	102.7	101.8	119.1	126.2	114.4	1.490	89
-68.2[a]	26.1[b]	-50.0[b]	36.7[b]	-56.0[b]	103.7	105.9	104.2	103.6	120.6	126.1	113.3	1.45	90
-68.2[a]	-3.4[b]	-11.0[b]	13.7[b]	-39.1[b]	103.7	107.8	107.1	103.2	120.6	126.1	113.3	1.45	90
-66.7	32.0	-37.6	28.6	-7.8	105.1	101.8	105.9	101.2	120.6	127.2	112.2	1.479	91
-66.7	-30.3	45.0	-43.2	25.8	99.6	103.8	98.5	103.9	120.6	127.2	127.2	1.479	91
-75.6	31.7	-34.9	24.2	-4.0	103.2	103.5	105.0	103.5	120.6	126.8	112.6	1.472	92
-70.5	29.4	-35.9	28.0	-9.7	103.4	104.0	104.5	103.2	120.8	126.6	112.5	1.462	92
--	-28	39	-36	19	105.1	103.5	103.4	103.7	--	--	108.5	1.50	93
--	14.0[b]	-27.2[b]	29.4[b]	-22.1[b]	104.5	107.6	103.9	105.5	--	--	109.4	1.48	94
--	-22.0[b]	44.0[b]	-46.9[b]	29.3[b]	108	97	109	96	--	--	108	1.53	95
-67.7	-9.9	14.9	-13.5	7.1	103.3	107.1	109.7	102.5	123.6	120.8	115.3	1.463	96
-69.9	27.8	-31.5	22.6	-3.9	102.9	102.1	107.0	103.3	118.4	126.4	114.8	1.471	96
-61.1	21.6	-28.4	23.1	-9.5	106.0	105.7	108.0	100.9	122.1	125.8	111.9	1.502	96

Table 2. Continued

ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_6	τ_6^R	τ_5^R	N-C$^\delta$	Ref.
-57.8	-13.2	16.6	-12.3	2.7	102.6	106.3	114.1	101.5	118.7	128.0	113.2	1.506	96
-76.3[a]	27.4[a]	-36.2[a]	31.1[a]	-14.2[a]	103.4	104.7	104.2	102.8	121.4	125.6	112.2	1.476	97
--	-35.0	23.0	-3.8	-18.4	106.8	100.8	109.2	106.0	--	--	104.6	1.516	98
-66.9	20.6a	-28.0[a]	24.7[a]	-11.9[a]	102.9	105.7	108.9	100.6	118.2	127.0	114.7	1.468[b]	99
59.1	-27.6[a]	37.9[a]	-33.8[a]	16.8[a]	103.4	103.7	104.0	103.5	117.6	131.0	111.2	1.478	100
-58.6	-24.1[a]	35.9[a]	-33.9[a]	19.0[a]	103.3	104.5	104.0	103.2	120.1	127.9	111.9	1.473	100
-53.2	-26.2	32.8	-26.2	9.4	104.3	103.5	105.9	104.1	119.3	128.5	111.9	1.48	101
-57.8	-26.2	40.5	-39.7	24.8	102.8	104.1	100.5	103.5	116.3	132.5	111.1	1.452	102
-65.0	-17.6	28.2	-27.2	16.2	104.0	107.0	107.5	102.5	119.0	129.7	110.9	1.499	103
-69.3[a]	21.7[a]	-30.8[a]	28.1[a]	-14.7	103.4	104.6	105.9	104.0	118.6	128.4	112.8	1.464	104
-65	19.0[b]	-21.6[b]	15.7[b]	-2.1[b]	107	103	115	101	121	126.9[b]	112	1.51	105
-71	22.4[b]	-31.5[b]	27.6[b]	-13.0[b]	105	103	107	102	124	125.2[b]	112	1.52	105
-70	28.8[a]	-33.1[a]	24.8[a]	-6.9[a]	102	104	105	102	120	126.3[a]	114	1.49	105
-64	26.4[a]	-34.9[a]	30.0[a]	-13.9[a]	105	103	105	101	120	126.4[a]	113	1.49	105
--	-39.1[b]	38.5[b]	-23.1[b]	-1.6[b]	103.7	102.3	105.0	105.3	--	--	107.3	1.533	106
--	35.6[b]	-41.0[b]	33.6[b]	-12.4[b]	106.4	101.0	102.4	106.7	--	--	106.9	1.48	107
--	-18.3	32.0	-33.5	23.1	105.0	106.3	103.0	104.9	--	--	109.1	1.496	108
--	20.7[b]	-32.6[b]	31.3[b]	-20.1[b]	102.2	108.2	102.3	103.8	--	--	112.5	1.468	109
-50.3[b]	-31.3[b]	38.6[b]	-30.7[b]	12.0[b]	102.7	103.9	101.4	106.6	119.9	129.6	110.5	1.471	109
-83	20.7[b]	-19.0[b]	9.1[b]	4.8[b]	105.8[b]	104.5[b]	111.6[b]	100.9[b]	121.9[b]	125.2[b]	112.8[b]	1.510[b]	110

Table 2. Continued

ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_6	τ_6^R	τ_5^R	N-C$^{\delta}$	Ref.
--	-18.2	32.0	-33.0	21.8	105.7	104.1	105.4	104.2	--	--	109.4	1.494	111
-71.3	34.1	-40.2	29.8	-7.4	105.0	100.7	107.1	99.8	118.7	129.4	111.7	1.480	112
-74.4	31.0	-38.2	31.0	-11.7	103.5	103.5	103.3	104.2	118.4	130.6	111.0	1.463	112
-53.3	-30.0	35.7	-26.6	6.5	102.2	101.9	106.9	102.4	127.5b	125.7[b]	106.6[b]	1.470[b]	112
-62.6	30.6	-37.2	28.7	-9.4	102.5	103.2	104.5	103.6	123.6	122.1	112.8	1.467	113
-72.6	32.4	-35.8	25.5	-4.7	102.5	103.7	104.0	105.0	124.0	123.3	112.2	1.46	113
-81.4	24.1	-29.4	21.4	-5.2	101.7	105.7	109.9	102.1	124.4	121.9	113.3	1.48	114
-76.2	22.1	-28.3	21.5	-6.4	102.3	103.9	109.5	102.3	125.0	120.8	114.2	1.48	114
-60.9	-22.2	32.9	-29.5	15.6	101.4	105.4	105.9	102.8	124.3	121.6	114.1	1.463	115
-64.9	24.7	-37.9	35.8	-21.6	103.3	104.5	102.6	102.5	120.1	127.5	112.2	1.475	115
-69.5	28.7	-34.9	27.0	-8.9	103.3	104.2	105.0	103.0	120.5	126.5	112.8	1.475	116
-57.7	-16.9a	20.9[a]	-16.9[a]	6.5[a]	102	108	109	104	122	124	114	1.46	117
-65.0[a]	-24.0[a]	26.6[a]	-19.0[a]	4.2[a]	103	103	108	105	119.5[a]	126	114	1.46	118
-72.2[a]	33.0[a]	-37.9[a]	28.3[a]	-7.9[a]	103	105	108	98	125	121	114	1.55	118
-71.6	-22.4	33.7	-32.1	18.2	104.6	103.8	104.7	104.0	116.2	130.6	111.0	1.490	119
-77.0	28.9	-40.6	36.9	-20.2	105.1	101.5	100.8	105.5	118.8	130.7	109.9	1.456	119
-62.0	-24.5	37.1	-34.4	19.4	102.6	105.9	104.1	120.4	117.7	131.0	111.1	1.485	120
-53.1	-25.4	34.2	-28.2	12.3	101.8	104.3	105.4	103.5	117.1	128.9	113.9	1.472	121
-56.5	-29.9	37.6	-29.9	11.1	102.8	103.2	104.4	103.4	118.7	127.9	112.6	1.474	121
-81.0	28.8	-39.5	33.9	-17.1	103.4	104.0	102.1	104.4	126.5	122.0	110.6	1.49	122

Table 2. Continued

ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_6	τ_6^R	τ_5^R	N-C$^\delta$	Ref.
-68.5	17.9	-22.9	16.9	-5.0	103.0	105.3	110.8	103.2	124.1	122.8	113.1	1.49	123
-69.8	14.0	-18.1	14.5	-4.9	103.0	106.1	110.9	103.9	124.2	122.6	113.1	1.52	123
-71.7	28.6[a]	-32.8[a]	24.5[a]	-6.9[a]	102.3	105.2	106.5	102.9	125.1	122.0	112.9	1.468	124
-38	-34	36	-23	1	102.9	102.9	105.4	103.3	123.4	124.8	111.7	1.468	125
-37	-31	35	-24	5	103.2	104.3	104.1	104.4	123.1	124.9	112.0	1.460	125
-60	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	126
-54	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	126
-58	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	126
-60	28.1[b]	37.8[b]	-30.3[b]	14.4[b]	100	107	102	102	119	125	115	1.44	127
-56	-21.7[b]	32.0[b]	-26.9[b]	14.1[b]	103.9	107.0	106.6	103.7	119.4	128.5	111.7	1.45	128
-72	-18.6[b]	26.4[b]	-22.7[b]	10.3[b]	102.0	106.2	109.2	101.8	118.7	125.3	114.5	1.45	128
69	10.0[b]	-26.6[b]	28.9[b]	-24.6[b]	104	106	107	99	121	125	114	1.54	129
84	-29.0[b]	36.9[b]	-28.4[b]	8.8[b]	103	106	111	98	115	129	112	1.55	129
72	-27.0[b]	36.8[b]	-32.4[b]	14.6[b]	106	100	106	106	120	132	108	1.43	129
81	-26.8[b]	33.4[b]	-35.8[b]	19.7[b]	103	105	102	103	120	126	111	1.46	129
-53	NC	NC	NC	NC	103.4	104.8	104.3	104.4	120.0	128.2	111.5	1.47	130
-46.4	-34.1	39.7	-29.4	7.6	102.6	102.3	104.4	102.2	120.8	126.3	112.9	1.489	131
-58.7	-28.1	39.6	-34.8	18.4	101.8	105.4	102.3	101.4	118.5	127.7	113.5	1.460	131
-73.3	30.1	-39.1	32.0	-13.3	102.5	103.6	103.8	102.7	127.3	120.2	112.5	1.477	132
-60.1	-25.8	36.5	-32.3	16.7	103.4	104.4	104.1	102.6	120.4	126.4	112.4	1.479	132

Table 2. Continued

ϕ	χ^1	χ^2	χ^3	χ^4	τ^R_1	τ^R_2	τ^R_3	τ^R_4	τ_6	τ^R_6	τ^R_5	N-C[δ]	Ref.
-84.3	34.8	-36.1	23.1	-1.1	102.7	103.7	104.4	104.0	119.1	129.3	111.5	1.481	132
-66.6	13.0	-20.3	18.3	-9.5	102.5	107.8	110.7	103.2	126.5	121.3	112.0	1.474	132
83.0	24.1	-30.0	26.2	-11.4	104.7[b]	104.9	107.7	105.2	124.4[b]	128.3	110.6	1.473	133
-89	9	0	-9	15	103.2	102.2	116.7	103.9	127.6	119.6	111.7	1.468	134
-80	18	-23	17	-4	102.8	104.4	113.0	101.5	120.6	125.2	113.7	1.457	134
-69	4.7[b]	-6.5[b]	5.5[b]	-2.1[b]	104.8	106.8[b]	114.3[b]	109.7[b]	120.8[b]	125.0[b]	113.9[b]	1.507[b]	135
-68	-11.0[b]	24.6[b]	-28.0[b]	22.2[b]	102.7	106.9[b]	106.2[b]	102.2[b]	125.1	118.7[b]	113.9[b]	1.464[b]	135
-70	-5.3[b]	8.0[b]	-6.8[b]	3.1[b]	102.6	106.7[b]	113.5[b]	101.7[b]	118.3	126.6[b]	114.8[b]	1.532[b]	135
-71	-15.5[b]	31.3[b]	-33.0[b]	23.7[b]	104.0	104.5[b]	107.4[b]	100.0[b]	126.4	120.7[b]	113.0[b]	113.0[b]	135
-82	23	-20	9	6	102.2	107.5	110.1	103.3	119.7	128.1	111.6	1.464	136
-50	-32	41	-34	15	102.6	103.4	102.3	102.2	118.9	128.1	113.0	1.482	136
-65	-6	17	-21	17	103.5	106.0	111.1	102.3	122.2	125.2	112.5	1.474	137
-41.5	-31.5	36.0	-25.1	4.5	103.2	103.1	107.2	101.5	123.5	124.0	112.4	1.482	138
-52	-27	39	-36	19	104.3	102.5	104.5	101.3	121.5	126.6	111.8	1.499	139
-86	28	-23	9	9	103.1	106.5	106.4	105.4	121.4	127.6	111.0	1.473	139
-75	33	-35	22	0	103.1	103.0	106.6	102.1	122.7	124.8	112.4	1.487	140
-66	22	-18	6	10	102.4	105.7	110.0	104.1	118.5	126.3	112.5	1.460	140
-64	25	-35	31	-16	104.3	104.2	104.1	103.0	118.3	128.9	112.3	1.490	141
-82	28	-26	14	5	103.1	105.7	107.1	106.0	122.6	127.3	110.0	1.471	141
-61	-26	33	-27	11	102.2	104.1	105.5	104.6	120.0	126.4	113.1	1.459	142

Table 2. Continued

ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_6	τ_6^R	τ_5^R	N-C$^\delta$	Ref.
-90	31	-37	28	-9	102.1	104.0	104.5	102.9	127.1	117.7	113.2	1.518	142
-64	-20	29	-26	12	104.0	105.8	108.4	102.3	118.9	128.9	111.5	1.512	142
-78	28	-37	29	-11	101.0	104.9	104.5	101.5	127.4	116.7	115.6	1.487	142
-50	-33	40	-32	10	103.5	101.7	105.4	101.2	121.1	126.6	112.2	1.497	142
-94	30	-39	32	-13	103.7	103.1	106.1	99.6	124.2	119.5	112.7	1.508	142
-65	-18	28	-26	13	102.6	104.1	109.6	102.2	118.4	127.4	114.2	1.451	142
-68	10	-18	18	-10	101.2	109.7	100.8	102.9	128.8	117.8	112.9	1.513	142
-68	-24	35	-32	20	101.2	109.7	100.8	102.9	128.8	117.8	112.9	1.513	142
-51	15	-29	30	-21	103.7	101.9	108.6	102.5	117.0	128.8	114.2	1.452	37
-79	39	-44	30	-6	97.2	106.1	100.2	105.2	119.8	127.7	112.5	1.478	37
-68	17	-10	0	12	98.2	113.4	106.7	101.2	118.9	123.9	116.8	1.482	37
-65	12.0^b	-11.3^b	14.6^b	-6.6^b	105.5	104.8^b	112.6^b	99.8^b	119.2	126.0^b	114.7^b	1.505^b	143
-81	21.7^b	-19.8^b	8.5^b	5.7^b	99.6	110.6^b	106.2^b	103.7^b	123.8	121.0^b	115.2^b	1.457^b	143
-74	24.9^b	-33.5^b	28.3^b	-12.7^b	102.7	103.8^b	105.5^b	102.8^b	118.7	126.9^b	114.2^b	1.489^b	143
-69	-21.3^b	28.3^b	-24.0^b	9.2^b	105.5	101.2^b	105.3^b	104.4^b	129.1	120.0^b	110.6^b	1.487^b	143
-64	-15	26	-26.5^b	15.2^b	106.3	100.5	113.3	100.2	122.9	124.1	112.9	1.49	144
-80	25	-33	26.8^b	-9.8^b	103.9	102.8	109.3	101.0	129.4	116.8	113.0	1.52	144
-62	11	-11	5.9^b	3.0^b	105.9	98.7	115.8	100.4	121.4	120.3	117.8	1.48	144
-92	29	-31	22.8^b	-4.5^b	105.5	101.8	103.8	109.1	125.3	123.8	109.4	1.41	144
-100	33.7^b	-31.4^b	16.6^b	5.0^b	102.6	104.5	104.9	105.1	129.2	118.3	111.1	1.492	145

Table 2. Continued

ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_6	τ_6^R	τ_5^R	N-C$^\delta$	Ref.
-110	13.0[b]	8.0[b]	-26.1[b]	35.6[b]	103.4	105.7	107.4	103.2	127.4	121.6	108.3	1.476	145
-95	27.8[b]	-21.8[b]	7.0[b]	11.6[b]	103.2	104.7	108.0	104.6	128.7	119.3	111.6	1.483	145
-102	31	25	9	12	102	105	107	105	128	120	113	1.48	146
-83	32	-35	23	-2	102.5	103.6	105.5	103.0	127.5	119.3	113.1	1.479	147
-70	31	-36	26	-7	102.4	103.7	104.5	104.2	127.1	119.9	112.7	1.469	147
-71	19	-13	1	12	102.4	106.6	111.3	104.0	127.2	120.5	111.8	1.467	147
-63	-13	29	-33	26	103.6	104.9	106.9	101.9	126.5	121.6	111.7	1.460	147
-56	11	-23	25	-19	NA	NA	NA	NA	NA	NA	NA	NA	34
-78	34	-41	31	-10	NA	NA	NA	NA	NA	NA	NA	NA	34
-77	13	-15	-10	0	NA	NA	NA	NA	NA	NA	NA	NA	34
-63	9	-18	20	-13	NA	NA	NA	NA	NA	NA	NA	NA	34
-57	-9	7	-2	5	NA	NA	NA	NA	NA	NA	NA	NA	34
-64	18	-28	25	-13	NA	NA	NA	NA	NA	NA	NA	NA	34
-63	-21	32	-29	15	NA	NA	NA	NA	NA	NA	NA	NA	34
-59	-36	40	-30	5	NA	NA	NA	NA	NA	NA	NA	NA	34
-55	-29	38	-32	15	NA	NA	NA	NA	NA	NA	NA	NA	34
-54	-28.5	34.8	-27.0	9.1	103.6	104.4	105.2	104.0	120.5	127.5	111.7	1.482	148
-70	-17.6	26.9	-25.1	14.3	104.0	107.8	106.6	102.5	120.9	126.6	111.9	1.498	148
-70	17.7	-31.0	30.6	-18.9	104.0	104.6	108.8	100.9	120.9	126.6	111.9	1.489	148
-66	-26.5[b]	38.4[b]	-35.1[b]	18.9[b]	104.0	102.7	103.8	102.3	119.9	126.9	112.2	1.464	42

Table 2. Continued

ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_6	τ_6^R	τ_5^R	N-C[δ]	Ref.
-53	-30.6[b]	42.4[b]	-36.8[b]	18.2[b]	103.0	102.4	102.7	101.6	121.1	126.5	112.4	1.484	42
-66	NA	NA	NA	NA	103.8	105.5	102.6	102.0	115.0	NA	110.3	1.463	149
-65.9	5.0	10.3	-12.4	25.7	NA	NA	NA	NA	NA	NA	NA	NA	150
-63.7	-2.6	9.0	-11.5	9.8	NA	NA	NA	NA	NA	NA	NA	NA	150
-65.6	14.7	-25.1	24.8	-14.7	NA	NA	NA	NA	NA	NA	NA	NA	150
-58.9	22.8	-36.7	35.2	-22.1	103.8	104.9	103.4	101.9[b]	118.3	129.7	112.0	1.477	151
-60	-31.0[b]	38.0[b]	-30.4[b]	11.7[b]	101	106	101	106	117	129.5	111	1.43	152
-91.0[b]	31.9[b]	-39.1[b]	30.9[b]	-10.8[b]	104.6	100.4	104.4	102.5	125.5	119.4	112.3	1.52	153
-71.1[b]	-28.3[b]	33.4[b]	-24.8[b]	6.9[b]	101.8	107.1	105.3	104.3	124.8	121.5	111.1	1.50	153
-91.4[b]	33.8[b]	-41.4[b]	31.6[b]	-11.6[b]	102.4	102.5	101.6	104.2	124.1	121.7	112.4	1.48	153
-65.3[b]	-30.6[b]	37.1[b]	-30.8[b]	11.1[b]	103.1	104.2	103.0	104.6	125.7	120.7	111.3	1.45	153
-89.9	33.0	-38.3	20.3	0.9	102.3	104.0	104.6	104.6	126.4	120.8	112.4	1.470	154
-90.3	36.6	-35.0	22.2	-.8	102.0	104.2	104.5	104.2	126.3	121.1	112.3	1.479	154
-64	-21.9[b]	50.6[b]	-52.8[b]	22.8[b]	101	108	108	102	120	123	115	1.47	35
-68	-27	39	-34	17	102	105	102	104	119	128	112	1.46	35
-37.3	-32.7	35.6	-24.0	3.2	102.9	103.0	104.8	103.5	122.4	124.2	112.8	1.462	155
-41.5	-31.5	36.0	-25.1	4.5	102.8	103.1	105.2	103.3	122.4	124.1	113.0	1.468	155
--	-10.2[b]	33.2[b]	-41.9[b]	34.9[b]	107	103	107	100	--	--	106	1.55	156
--	8.4[b]	19.0[b]	-39.4[b]	45.6[b]	104	108	102	102	--	--	103	1.52	156
-71	33	-38	28	-7	103	103	104	104	117	130	112	1.47	157

Table 2. Continued

ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_6	τ_6^R	τ_5^R	N-C$^\delta$	Ref.
--	-13.4[b]	32.4[b]	-52.2[b]	31.7[b]	106.8	103.4	103.7	103.3	--	--	105.2	1.49	158
-65.0	-17.7	32.5	-33.9	23.5	103.5	104.8	105.2	102.4	124.2	123.5	112.2	1.462	159
-52.6	-26.2	33.1	-26.1	9.3	102.4	104.8	106.1	103.8	119.6	127.8	112.5	1.467	159
-77.8[b]	-17.4[b]	22.0[b]	-26.0[b]	18.1[b]	105	109	104	105	121.2[b]	123	115	1.49	160
-74.6[b]	18.5[b]	-13.7[b]	3.5[b]	8.9[b]	105.1[b]	105.8[b]	108.0[b]	104.8[b]	121.2[b]	125.9[b]	112.8[b]	1.460[b]	161
-67.4[b]	-6.1[b]	19.1[b]	-24.3[b]	21.6[b]	104.9[b]	106.8[b]	106.3[b]	102.8[b]	121.1[b]	125.7[b]	113.2[b]	1.462[b]	161
-83.0[b]	23.6[b]	-18.2[b]	4.4[b]	12.3[b]	105	101	115	100	126	121	113	1.48	162
-97.2	31.3	-25.2	8.9	11.9	102.4	104.9	106.8	105.1	129.1	119.4	111.0	1.487	163
-94.8	32.6	-28.5	13.0	8.2	102.6	104.6	105.7	105.9	129.6	119.5	110.5	1.483	163
-94.9	29.0	-17.6	-.9	20.5	102.5	105.2	107.4	104.8	129.1	120.5	110.4	1.477	163
-97.6	31.3	-25.7	9.7	10.9	102.2	104.8	106.6	105.0	128.3	119.5	111.4	1.475	163
-95.1	29.6	-18.0	-.9	20.9	102.3	105.3	106.7	105.4	129.0	120.8	110.1	1.474	163
-106.0	34.6	-29.5	12.6	10.0	101.9	104.5	105.4	105.3	128.0	119.2	110.9	1.477	163
-67	-8	13	-13	8	102.8	108	109	107	118.1	130	112	1.52	164
-98	30	-30	17	3	99.3	105	108	101	124.1	118	116	1.50	164
-65	-9.8[b]	19.4[b]	-20.6[b]	15.0[b]	102.5	107.7	108.1	103.2	117.6	128.1	114.2	1.490	165
-96	33.5[b]	-36.4[b]	24.6[b]	-3.0[b]	99.6	107.1	104.5	102.4	125.1	119.7	113.5	1.487	165
-70	29	-36	28	-10	103	104	105	103	121	126	113	1.46	166
-80	33	-40	31	-10	101.6	103.5	103.1	102.9	121.0	125.8	113.1	1.468	167
-91.0[b]	38.1[b]	-34.3[b]	15.8[b]	9.8[b]	100.1[b]	102.7[b]	105.7[b]	102.3[b]	119.0[b]	127.3[b]	113.7[b]	1.472[b]	168

Table 2. Continued

ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_6	τ_6^R	τ_5^R	N-C $^\delta$	Ref.
-83.8[b]	24.0[b]	-25.1[b]	15.7[b]	-.4[b]	102.1[b]	107.0[b]	105.1[b]	104.7[b]	118.8[b]	126.2[b]	114.7	1.471[b]	168
-104	17	-6	-8	20	100	108	109	106	126	122	112	1.43	169
-45	-28	37	-31	13	103	104	105	104	130	119	111	1.33	169
-52	-32	39	-29	10	102	104	104	103	128	118	114	1.44	169
104	-34	18	5	-28	102	104	107	103	133	118	109	1.49	169
-47	-25	41	-39	25	104	102	103	100	127	119	113	1.34	169
-50	-25	31	-23	7	101	107	106	106	128	120	112	1.47	169
108	-38	32	-14	-11	103	102	107	103	129	120	109	1.48	169

Figure 3. Distributions for proline ring bond angles and NC^{δ} bond length

Each of the bond angles is in degrees. The NC^{δ} bond length is in Å. To correlate the angles with data in Table 2, the angles are named as follows: $\tau_1^R = \angle NC^{\alpha}C^{\beta}$; $\tau_2^R = C^{\alpha}C^{\beta}C^{\gamma}$; $\tau_3^R = C^{\beta}C^{\gamma}C^{\delta}$; $\tau_4^R = C^{\gamma}C^{\delta}N$; $\tau_5^R = C^{\delta}NC^{\alpha}$; $\tau_6^R = C'NC^{\delta}$; and $\tau_6 = C'NC^{\alpha}$.

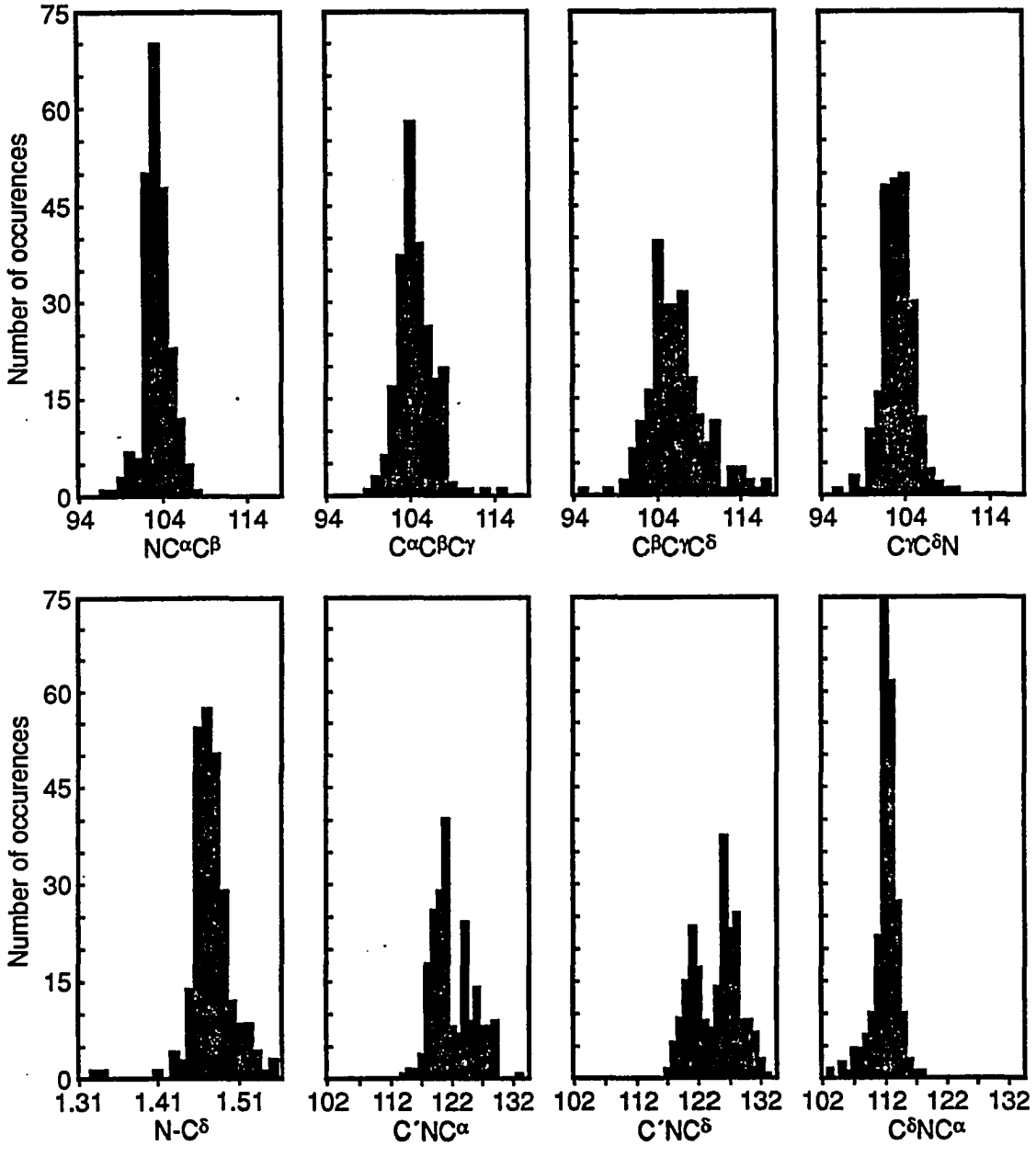

Number of occurences
75
60
45
30
15
0
94
104
114
NCαCβ
CαCβCγ
CβCγCδ
CγCδN
1.31
1.41
1.51
N-Cδ
102
112
122
132
C'NCα
C'NCδ
CδNCα

Table 3. Statistics of X-ray structures for proline-containing compounds

Parameter[a]	Mode	Mean	Std. Dev.	Sample Size
ϕ(L-pro)	-68	-69.0	13.6	211
ϕ(D-pro)	-[b]	79.1	15.9	10
χ^1	-28,29	3.1	24.8	240
χ^2	-36,39	- 0.7	31.2	240
χ^3	-32,31	- 1.9	26.3	240
χ^4	-10,12	3.7	14.8	240
$\angle NC^\alpha C^\beta$	103	103.2	1.6	229
$\angle C^\alpha C^\beta C^\gamma$	104	104.7	2.2	229
$\angle C^\beta C^\gamma C^\delta$	104	106.2	3.3	229
$\angle C^\gamma C^\delta N$	104	103.3	2.2	229
$\angle C'NC^\alpha$	121	122.1	3.5	212
$\angle C'NC^\delta$	126	124.4	8.7	211
$\angle C\ NC^\alpha$	112	112.1	2.0	229
$N-C^\delta$	1.47	1.48	0.02	229
Sum[c]	360	359.5	0.9	211

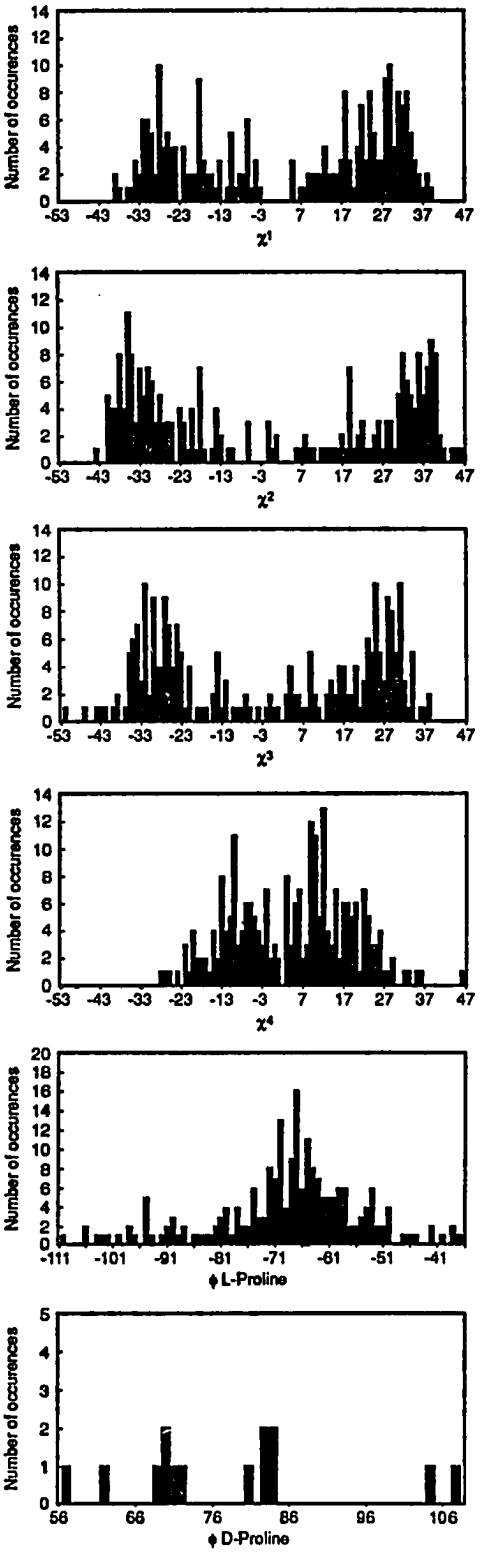
[a] Angles in degrees; bond length in Å.

[b] No distinct mode.

[c] $\angle C'NC^\alpha + \angle C'NC^\delta + \angle C^\delta NC^\alpha$.

Figure 4. Distributions for proline ring torsion angles

All torsion angles are in degrees. The torsion angles are defined as follows: χ^1, $NC^{\alpha}C^{\beta}C^{\gamma}$; χ^2, $C^{\alpha}C^{\beta}C^{\gamma}C^{\delta}$; χ^3, $C^{\beta}C^{\gamma}C^{\delta}N$; χ^4, $C^{\gamma}C^{\delta}NC^{\alpha}$; ϕ, $C'_1NC^{\alpha}C'_2$.


Number of occurences
χ¹
χ²
χ³
χ⁴
ϕ L-Proline
ϕ D-Proline

from the sum of the means of these angles due to cumulative round-off errors.

The success of BORC can be gauged by a comparison of predicted ring conformations with those experimental conformations from the survey. Only those experimental rings complete with all torsions ϕ and χ^i are used as test cases. In addition, the energies of the predicted ring conformations are calculated, using various sets of semi-empirical energy parameters from the literature, to determine whether the calculated energies are consistent with the distribution of observed conformations.

C. Semi-Empirical Energy Functions

The energy, U_{tot}, of the proline structure in Figure 1 is calculated as the sum

$$U_{tot} = U_b + U_a + U_e + U_n + U_t \tag{1}$$

Where U_b is the bond stretching energy for the N-C^{δ} bond length $r_{NC}{}^{\delta}$;

$$U_b = k^b_{NC}{}^{\delta}(r^0_{NC}{}^{\delta} - r_{NC}{}^{\delta})^2 \tag{2}$$

U_a is the angle bending energy for all variable bond angles, θ_i.

$$U_a = \sum_i k^{\theta}_i (\theta^0_i - \theta_i)^2 \tag{3}$$

U_e is the electrostatic energy interaction among fixed atomic charges, q_i.

$$U_e = \frac{1}{2} \sum_{ij} \left(\frac{332\, q_i\, q_j}{D r_{ij}} \right) \tag{4}$$

U_n is the Lennard-Jones nonbonded interaction energy.

$$U_n = \frac{1}{2} \sum_{ij} \left\{ \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right\} \tag{5}$$

U_t is the torsion energy associated with torsion angles χ^i.

$$U_t = \frac{1}{2} \sum_i k_i^{\chi} (1 + \cos n_i \chi^i) \qquad (6)$$

In Equations 2 and 3 $k_{NC}^{b}\delta$ and k_i^{θ} are constants and $r_{NC}^{0}\delta$ and θ_i^0 are equilibrium values of the variables. In Equations 4 and 5 r_{ij} is the distance between atoms i and j; D is the dielectric constant; A_{ij} and B_{ij} are the repulsive and attractive constants respectively that are characteristic of each atom pair; the summations are over all atom pairs separated by three or more bonds, and the factor 332 gives an energy in kcal/mol when q_i is in atomic units and r_{ij} is in Å. In Equation 6, k_i^{χ} is the torsion energy amplitude, and n_i is the symmetry number for the rotational barrier (taken as 3 for all bonds included here); the summation is over the five ring torsion angles.

A number of different sets of parameters have been used by various authors for calculation of energies of peptides by the above equations (171-182). Three sets were selected arbitrarily to calculate the energy of the proline fragment in order to learn which features of the energy maps are consistent among the sets (171-175). The sets and some of their parameters are described as follows. The remaining parameters are given in Tables 4-6.

1) Set 1

The parameters are those of Karplus and Lifson (171) and Lifson and Warshell (172) who give $k_{NC}^{b}\delta$ = 522 kcal mol^{-1} $Å^{-2}$; $r_{NC}^{0}\delta$ = 1.458 Å; D = 1, and k_i^{χ} = 5.672 kcal/mol for C-C bonds and -3.00 kcal/mol for C-N bonds. The zero point of U_t is shifted from that of equation 6 by subtracting $0.5k_i^{\chi}$ from each term.

Table 4. Angle bending parameters

Angle	Set I		Set II	
	k^{θ} (kcal mol^{-1} rad^{-2})	$\theta°$ (deg)	k^{θ} (kcal mol^{-1} rad^{-2})	$\theta°$ (deg)
$NC^{\alpha}C^{\beta}$	42.0	109.47	80.0	109.7
$NC^{\alpha}H$	60.2	109.47	80.0	109.7
$NC^{\delta}C^{\gamma}$	42.0	109.47	80.0	111.2
$NC^{\delta}H$	60.2	109.47	80.0	109.7
$C^{\alpha}NC^{\delta}$	109.0	120.0	50.0	118.0
$C^{\alpha}C^{\beta}H$	54.0	109.47	63.0	112.4
$C^{\alpha}C^{\beta}C^{\gamma}$	43.2	109.47	63.0	112.4
$C^{\beta}C^{\alpha}H$	54.0	109.47	63.0	111.5
$C^{\beta}C^{\gamma}C^{\delta}$	43.2	109.47	63.0	112.4
$C^{\beta}C^{\gamma}H$	54.0	109.47	63.0	112.4
$C^{\gamma}C^{\beta}H$	54.0	109.47	63.0	112.4
$C^{\gamma}C^{\delta}H$	54.0	109.47	63.0	112.4
$C^{\delta}C^{\gamma}H$	54.0	109.47	63.0	112.4
$C^{\delta}NC'$	109.0	120.0	50.0	121.9

Table 5. Fixed atomic charges (a.u.)

Atom	Set I	Set II	Set III
C'	0.449	0.526	0.455
N_α	-0.305	-0.257	-0.285
C_α	0.144	0.112	0.050
H_β	0.000	0.000	0.040
C_β	0.000	0.001	-0.024
H_γ	0.000	0.000	0.015
C_γ	0.000	0.036	-0.050
H_δ	0.000	0.000	0.025
C_δ	0.000	0.084	0.100
H	0.000	0.000	0.010

Table 6. Nonbonded potential parameters

Pair	Set I		Set II		Set III	
	Aij x 10^{-4} (kcal mol^{-1} A^{-12})	Bij (kcal mol^{-1} A^{-6})	Aij x 10^{-4} (kcal mol^{-1} A^{-12})	Bij (kcal mol^{-1} A^{-6})	Aij x 10^{-4} (kcal mol^{-1} A^{-12})	Bij (kcal mol^{-1} A^{-6})
N...C^{β}	56.9	522	84.1	683	21.6	366
N...C^{γ}	56.9	522	84.1	683	21.6	366
N...H	3.98	65.3	6.76	145	2.7	125
C'...H	2.98	49.0	8.55	152	3.8	128
C'...C'	42.6	392	79.0	616	28.6	370
C'...C^{γ}	42.6	392	100.0	695	28.6	370
C'...C^{β}	42.6	392	100.0	695	28.6	370
C'...C^{δ}	42.6	392	100.0	695	28.6	370
C^{α}...C^{δ}	42.6	392	87.0	601	28.6	370
C^{α}...H	2.98	49.0	7.4	132	3.8	128
C^{α}...C^{γ}	42.6	392	87.0	601	28.6	370
C^{β}...C^{δ}	42.6	392	127.0	782	28.6	370
C^{β}...H	2.98	49.0	11.3	174	3.8	128
C^{γ}...H	2.98	49.0	11.3	174	3.8	128
C^{δ}...H	2.98	49.0	11.3	174	3.8	128
H...H	0.184	5.75	0.711	32.9	0.446	46.7

2) Set II

The parameters are those of Weiner et al. (173), which are incorporated into the AMBER program (not used here). For this set $k^{b}_{NC}\delta$ = 377 kcal mol^{-1} $Å^{-2}$; $r^{0}_{NC}\delta$ = 1.449 Å; D is set numerically equal to r_{ij} in Å, and k^{χ}_{i} = 4.00 kcal/mol for C-C bonds and 0.0 kcal/mol for C-N bonds.

3) Set III

The parameters are those of Scott and Scheraga (174) and Momany et al. (175). These authors do not include the U_a, U_b, or U_t terms for proline, and this practice was followed for this set. They adopt D in the range of 2-8; D = 4 was chosen for use with their parameters in this study because this value was frequently chosen by others (174, 178, 181).

D. Results

For 191 of the proline rings in our survey, BORC was used to generate a closed ring with the values of ϕ,χ^2 for the observed structure. The values χ^1, χ^3, and χ^4 were then compared for the observed and calculated structures in each case. Table 7 shows this comparison of the structures surveyed. Usually the BORC solution was unique regardless of the choice of initial values for the adjustable parameters. In 24 cases two solutions were found, depending on whether the initial value of χ^1 was positive or negative. Where two solutions were found, both are given in Table 7. In such cases, the observed structures are bracketed by the two solutions; in some cases one solution is favored by observation; this is usually the solution resulting from initial value of $\chi^1 = -\chi^2$. The double solutions occurred only in the ranges of $-71° \leq \phi \leq -57°$ and $-28° \leq \chi^2 \leq 17°$.

Table 7. Calculated and observed torsion angles in proline (degrees)

		χ^1		χ^3		χ^4		
ϕ	χ^2	exptl.	calcd.[a,b]	exptl.	calcd.[b]	exptl.	calcd.[b]	Ref.
-110	8.0	13.0	12.0(2)	-26.1	-25.2	35.6	35.0	145
-106.0	-29.5	34.6	35.2(2)	12.6	-29.5	10.0	12.3	163
-104	-6.0	17	21.8(3)	-8	-12.4	20	28.1	169
-100	-31.4	33.7	34.6(2)	16.6	15.9	5.0	6.5	145
-98	-30	30	33.7(2)	17	14.5	3	7.4	164
-97.6	-25.7	31.3	32.2(2)	9.7	9.1	10.9	12.2	163
-97.2	-25.2	31.3	32.0(2)	8.9	8.5	11.9	12.7	163
-96	-36.4	33.5	35.4(2)	24.6	23.1	-3.0	-.7	165
-95.1	-35.7	38.4	35.0(2)	18.9	22.5	5.9	-.3	82
-95.1	-18.0	29.6	29.0(3)	-.9	-.2	20.9	20.0	163
-94.9	-17.6	29.0	28.8(3)	-.9	-.7	20.5	20.3	163
-94.8	-28.5	32.6	32.8(2)	13.0	13.0	8.2	8.4	163
-94	-39	30	36.0(2)	32	26.8	-13	-4.3	142
-92	-31	29	33.0(2)	23	16.8	-5	4.5	144
-91.4	-41.4	33.8	36.4(4)	31.6	30.2	-11.6	-7.6	153
-91.0	-39.1	31.9	35.3(3)	30.9	27.7	-10.8	-5.6	153
-91.0	-34.3	38.1	33.6(2)	15.8	21.6	9.8	-.2	168
-90.3	-35.0	33.6	33.7(1)	22.2	22.7	-.8	-1.3	154
-90	-37	31	34.2(2)	28	25.3	-9	-3.8	142
-89.9	-33.3	33.0	33.2(2)	20.3	20.4	.9	0.8	154
-89.1	-35.8	36.4	33.7(2)	20.4	24.0	2.9	-2.7	56
-89	0	9	17.5(4)	-9	-17.8	15	31.0	134
-88.2	-37.4	32.9	34.0(1)	26.7	26.2	-6.1	-4.9	62
-86	-23	28	31.4(3)	9	5.4	9	15.6	139

[a] Figures in parentheses are estimated standard deviations.

[b] Where two values are given, the first corresponds to the initial value $\chi^1 = \chi^2$ and the second corresponds to initial $\chi^1 = -\chi^2$.

Table 7. Continued

		χ^1		χ^3		χ^4		
ϕ	χ^2	exptl.	calcd.[a,b]	exptl.	calcd.[b]	exptl.	calcd.[b]	Ref.
-84.3	-36.1	34.8	33.0(2)	23.1	25.1	-1.1	-4.5	132
-83.8	-28.1	24.0	32.4(3)	15.7	12.7	-.4	8.4	168
-83	-35	32	32.6(2)	23	23.8	-2	-3.3	147
-82	-26	28	32.0(3)	14	9.8	5	11.3	141
-82	-20	23	30.3(4)	9	1.8	6	18.6	136
-81.4	-29.4	24.1	32.3(3)	21.4	15.0	-5.2	5.9	114
-81.0	-39.6	28.8	34.0(4)	33.9	29.5	-17.1	-8.4	122
-81	-19.8	21.7	30.0(4)	8.5	1.7	5.7	18.5	143
-80	-40	33	34.3(4)	31	30.0	-10	-8.9	167
-80	-33	25	32.1(2)	27	21.0	-10	-0.7	144
-80	-23	18	31.1(4)	17	5.9	-4	14.7	134
-79	-44	39	36.9(7)	30	33.8	-6	-11.1	37
-78.3	-35.8	32.2	32.2(2)	25.6	25.3	-5.7	-5.2	65
-78	-41	34	34.8(5)	31	31.1	-10	-9.6	34
-78	-37	28	32.6(2)	29	26.9	-11	-6.6	142
-77.8	22.0	-17.4	-5.2(4)	-26.0	-30.4	18.1	28.9	160
-77.0	-40.6	28.9	34.5(5)	36.9	30.8	-20.2	-9.5	119
-76.3	-36.2	27.4	32.1(2)	31.1	26.1	-14.2	-6.1	97
-76.2	-28.3	22.1	31.2(3)	21.5	14.4	-6.4	5.8	114
-75.6	-34.9	31.7	31.6(2)	24.2	24.6	-4.0	-4.9	92
-75.3	-39.4	36.8	33.7(4)	26.3	29.7	-3.1	-9.0	81
-75	-35	33	31.5(2)	22	24.9	0	-5.3	140
-74.6	-13.7	18.5	26.0(4)	3.5	-4.0	8.9	21.8	161
-74.4	-38.2	31.0	32.8(3)	31.0	28.6	-11.7	-8.4	112
-74	-33.5	24.9	30.8(2)	28.3	23.2	-12.7	-3.9	143
-73.3	-39.1	30.1	33.2(4)	32.0	29.7	-13.3	-9.2	132
-73	-38	29	32.5(3)	31	28.7	-13	-8.6	84
-72.6	-35.8	32.4	31.2(2)	25.5	26.5	-4.7	-7.2	113
-72.2	-37.9	33.0	32.3(3)	28.3	28.7	-7.9	-8.8	118

Table 7. Continued

ϕ	χ^2	χ^1 exptl.	χ^1 calcd.[a,b]	χ^3 exptl.	χ^3 calcd.[b]	χ^4 exptl.	χ^4 calcd.[b]	Ref.
-72	-36	28	31.1(2)	29	26.9	-10	-7.6	71
-72	-35	31	30.6(2)	24	25.8	-4	-6.8	69
-72	26.4	-18.6	-11.6(4)	-22.7	-30.9	10.3	25.2	128
-71.8	-33.1	28.5	29.8(3)	24.2	23.5	-5.8	-4.9	83
-71.7	-32.8	28.6	29.7(3)	24.5	23.2	-6.9	-4.6	124
-71.6	33.7	-22.4	-22.0(3)	-32.1	-32.4	18.2	19.7	119
-71.3	-40.2	34.1	33.7(4)	29.8	31.0	-7.4	-10.3	112
-71.1	33.4	-28.3	-21.8(3)	-24.8	-32.1	6.9	19.5	153
-71	-38	33	32.1(3)	28	29.1	-7	-9.3	157
-71	-31.5	22.4	29.1(4)	27.6	21.7	-13.0	-3.5	105
-71	-13	19	3.9(5), 24.5(5)	1	17.2, -3.5	-4	-15.7, 20.1	147
-71	31.3	-15.5	-19.2(4)	-33.0	-31.3	23.7	20.5	135
-70.8	-34.3	34.7	29.9(3)	20.3	25.4	1.9	-6.9	85
-70.5	-35.9	29.4	30.6(2)	28.0	27.2	-9.7	-8.3	92
-70	-36	31	30.5(2)	26	27.5	-7	-8.7	147
-70	-36	29	30.5(2)	28	27.5	-10	-8.7	166
-70	-33.1	28.8	29.0(3)	24.8	24.4	-6.9	-6.5	105
-70	-31	24	28.4(4)	25	21.6	-10	-3.9	68
-70	8.0	-5.3	10.0(4)	-6.8	-23.6	3.1	31.2	135
-70	21	-10	-3.7(4)	-23	-30.3	17	29.8	66
-70	25	-10	-9.5(5)	-29	30.9	24	26.6	71
-70	26.9	-17.6	-12.6(5)	-25.1	-30.8	14.3	24.4	148
-69.9	-31.5	27.8	28.4(4)	22.6	22.4	-3.9	-4.7	96
-69.8	-18.1	14.0	6.2(8), 26.1(5)	14.5	23.2, 3.2	-4.9	-20.5, 14.2	123
-69.6	-32.2	26.5	28.4(4)	24.9	23.5	-7.9	-5.9	61
-69.5	-34.9	28.7	29.7(2)	27.0	26.6	-8.9	-8.3	116

Table 7. Continued

		χ^1		χ^3		χ^4		
ϕ	χ^2	exptl.	calcd.[a,b]	exptl.	calcd.[b]	exptl.	calcd.[b]	Ref.
-69.4	-22.3	24.3	10.2(7), 27.1(6)	10.6	25.9, 9.0	5.8	-20.7, 8.7	87
-69.3	-30.8	21.7	27.8(4)	28.1	21.9	-14.7	-4.5	104
-69	-6.5	4.7	20.2(7), -3.5(7)	5.5	-9.7, 14.1	-2.3	23.8, -17.3	135
-69	28.3	-21.3	-15.1(4)	-24.0	-30.6	9.2	22.4	143
-68.6	22.5	-5.6	-5.8(5)	-30.3	-30.6	27.3	28.7	60
-68.5	-22.9	17.9	10.2(6), 26.6(7)	16.9	26.9, 10.4	-5.0	-21.8, 6.8	123
-68	-18	10	4.5(6), 25.0(7)	18	24.7, 4.1	-6	-23.3, 12.5	142
-68	-10	17	-2.4(6), 21.7(7)	0	18.6, -5.6	12	-21.3, 20.4	37
-68	24.5	-11.0	-8.8(5)	-28.0	-30.8	22.2	27.0	135
-68	35	-24	-24.9(3)	-32	-31.6	20	16.8	142
-68	39	-27	-29.1(4)	-34	-33.8	17	16.4	35
-67.7	14.9	-9.9	3.0(5)	-13.5	-27.2	7.1	31.0	96
-67.4	19.0	-6.1	-1.5(4)	-24.1	-29.3	21.6	30.2	161
-67.2	-30.0	19.0	25.7(5)	28.2	22.8	-16.7	-6.9	59
-67.1	-33.0	24.0	27.3(3)	29.4	25.9	-14.6	-9.2	64
-67	-33	24	27.3(3)	29	25.9	-14	-9.2	70
-67	-30	22	25.4(5)	25	23.0	-11	-7.3	67
-67	13	-8	4.7(5)	-13	-25.9	8	30.7	164
-67.3	22.0	-7.4	-5.2(5)	-26.9	-30.4	23.1	28.9	79
-66.9	-28.0	20.6	17.7(6), 25.0(7)	24.7	27.4, 20.2	-11.9	-17.3, -4.6	99
-66.7	-37.6	32.0	30.6(3)	28.6	30.0	-7.8	-11.3	91
-66.7	11.7	-3.6	5.9(5)	-14.9	-24.9	12.6	30.4	87
-66.7	45.0	-30.3	-35.2(6)	-43.2	-37.3	25.8	16.1	91

Table 7. Continued

		χ^1		χ^3		χ^4		
ϕ	χ^2	exptl.	calcd.[a,b]	exptl.	calcd.[b]	exptl.	calcd.[b]	Ref.
-66.6	-31.1	27.0	25.6(5)	22.7	24.6	-5.3	-8.9	82
-66.6	-20.3	13.0	5.9(6), 24.7(7)	18.3	27.0, 8.2	-9.5	-24.8, 7.9	132
-66	-20	17	5.2(6), 24.1(7)	14	27.2, 8.2	-2	-25.4, 7.5	84
-66	-18	22	3.1(6), 23.7(7)	6	26.1, 5.4	10	-25.6, 10.1	140
-66	38.4	-26.5	-29.3(3)	-35.1	-32.6	18.9	15.1	42
-65.9	10.3	5.0	6.9(5)	-12.4	-23.7	25.7	29.7	150
-65.6	-25.1	14.7	11.4(7), 23.4(7)	24.8	29.2, 17.1	-14.7	-23.4, -2.4	150
-65.3	37.1	-30.6	-28.2(3)	-30.2	-31.6	11.1	14.6	153
-65	-21.6	19.0	6.6(5), 23.2(8)	15.7	28.5, 11.7	-2.1	-25.9, 3.2	105
-65	-17.3	12.0	1.9(6), 22.7(7)	14.6	26.2, 5.3	-6.6	-26.6, 9.6	143
-65	17	-6	0.2(5)	-21	-27.8	17	29.7	137
-65	19.4	-9.8	-2.5(5)	-20.6	-29.0	15.9	29.3	165
-65.0	26.6	-24.0	-12.4(5)	-19.0	-30.5	4.2	24.2	118
-65	28	-18	-15.1(5)	-26	-30.1	13	21.9	142
-65.0	28.8	-17.6	-16.7(5)	-27.2	-29.7	16.2	20.4	103
-65.0	40.7	-32.4	-31.8(4)	-32.5	-33.8	12.5	14.6	54
-65.0	32.5	-17.7	-23.2(4)	-33.9	-29.2	23.5	15.5	159
-64.9	-37.9	24.7	30.3(3)	35.8	30.8	-21.6	-12.4	115
-64.8	32.2	-17.9	-22.9(4)	-33.4	-29.0	22.8	15.5	73
-64	-35	25	27.4(3)	31	29.0	-16	-12.4	141
-64	-34.9	26.4	27.3(3)	30.0	29.0	-13.9	-12.4	105
-64	-28	18	16.0(6)	25	29.2	-13	-20.4	34
-64	26	15	-11.5(5)	-26.5	-30.5	15.2	24.7	144

Table 7. Continued

		χ^1		χ^3		χ^4		
ϕ	χ^2	exptl.	calcd.[a,b]	exptl.	calcd.[b]	exptl.	calcd.[b]	Ref.
-64	29	-20	-17.4(5)	-26	-29.4	12	19.6	142
-63.7	-29.8	23.5	20.5(6)	24.7	27.5	-10.2	-15.5	80
-63.7	9.0	-2.6	7.2(7), -17.5(7)	-11.5	-21.9, 3.0	9.8	28.0, -14.8	150
-63.2	13.7	-7.7	2.9(7), -19.5(6)	-14.5	-25.2, -2.7	9.8	28.6, -10.2	64
-63	-18	9	1.7(7), 20.6(7)	20	27.5, 8.5	-13	-28.2, 4.9	34
-63	14	-7	2.5(7), -19.8(6)	-14	-25.3, -2.8	9	28.6, -10.3	70
-63	29	-13	-17.8(5)	-33	-29.0	26	19.0	147
-63	32	-21	-23.6(4)	-29	-28.0	15	13.9	34
-62.6	-37.2	30.6	28.9(3)	28.7	31.1	-9.4	-13.6	113
-62.3	31.4	-18.5	-23.2(4)	-31.6	-27.4	20.5	13.5	63
-62	-38	27	29.4(3)	34	31.9	-18	-14.1	89
-62	-11	11	-5.5(6), 18.2(8)	5.9	23.4, -.4	3.0	-28.6, 12.6	144
-62.0	28.2	-17.0	-16.4(6)	-28.5	-29.4	18.0	20.4	55
-62.0	37.1	-24.5	-29.6(3)	-34.4	-30.3	19.4	12.3	120
-61.2	29.8	-15.5	-21.1(6)	-32.0	-26.9	23.0	14.4	58
-61.1	-28.4	21.6	16.1(5)	23.1	29.8	-9.5	-20.9	96
-61	1	5	11.6(7), -15.2(6)	-7	-13.3, 13.7	11	21.8, 24.7	72
-61	33	-26	-26.1(4)	-27	-27.1	11	11.3	142
-60.9	32.9	-22.2	-26.1(4)	-29.5	-27.0	15.6	11.2	115
-60.1	36.5	-25.8	-29.7(3)	-32.3	-29.1	16.7	11.0	132
-60.0	34.5	-19.0	-28.1(3)	-36.4	-27.6	26.0	10.4	73
-60.0	37.8	-28.1	30.9(3)	-30.3	-30.1	14.4	11.2	127
-60.0	38.0	-31.0	-31.0(3)	-30.4	-30.2	11.7	11.2	152

Table 7. Continued

		χ^1		χ^3		χ^4		
ϕ	χ^2	exptl.	calcd.[a,b]	exptl.	calcd.[b]	exptl.	calcd.[b]	Ref.
-59	-27	25	13.0(5)	19	30.6	-3	-23.8	84
-59.0	5.9	-3.7	7.1(8), -19.4(5)	-5.6	-16.7, 9.9	2.9	22.4, -23.5	55
-59	36	-28	-29.8(3)	-30	-28.3	12	10.0	74
-59	40	-36	-33.0(4)	-30	-31.5	5	11.3	34
-58.9	-36.7	22.8	27.1(3)	35.2	32.2	-22.1	-16.0	151
-58.7	39.6	-28.1	-32.7(4)	-34.8	-31.1	18.4	11.0	131
-58.6	35.9	-24.1	-29.8(3)	-33.9	-28.0	19.0	9.7	100
-57.8	16.6	-13.2	-2.3(8), -25.1(5)	-12.3	-24.6, -1.7	2.7	24.5, -15.1	96
-57.8	40.5	-26.2	-33.6(4)	-39.7	-31.6	24.8	11.0	102
-57.7	20.9	-16.9	-6.8(7), -26.7(6)	-16.9	-27.1, -7.0	6.5	24.2, -10.5	117
-57	7	-9	4.8(7), -21.0(6)	-2	-16.2, 9.8	5	20.4, -24.5	34
-56.5	37.6	-29.9	-31.8(3)	-29.9	-28.8	11.1	9.3	121
-56	-23	11	6.4(4)	25	30.7	-19	-28.5	34
-56.0	32.0	-21.7	-29.1(4)	-26.9	-22.6	14.2	4.4	128
-55.9	-39.8	29.4	29.0(4)	33.6	35.2	-15.2	-18.0	75
-55	-41	28	29.8(4)	37	36.3	-20	-18.6	88
-55	38	-29	-32.4(3)	-32	-28.8	15	8.8	34
-55	46	-39	-38.2(8)	-35	-35.6	10	12.3	88
-54.3	-20.8	16.0	-3.4(4)	16.5	30.3	-5.5	-30.0	86
-54.0	34.8	-28.5	-30.9(2)	-27.0	-25.2	9.1	6.0	148
-53.3	34.2	-26.5	-30.9(2)	-27.3	-24.2	10.0	4.9	73
-53.3	35.7	-30.0	-31.5(2)	-26.6	-26.0	6.5	6.5	112
-53.2	32.8	-26.2	-30.7(3)	-26.2	-22.2	9.4	2.9	101
-53.1	34.2	-25.4	-31.0(2)	-28.2	-24.1	12.3	4.8	121
-53.0	42.4	-30.6	-35.7(6)	-36.8	-32.5	18.2	10.6	42

Table 7. Continued

ϕ	χ^2	χ^1		χ^3		χ^4		Ref.
		exptl.	calcd.[a,b]	exptl.	calcd.[b]	exptl.	calcd.[b]	
-52.6	33.1	-26.2	-30.9(3)	-26.1	-22.4	9.3	2.9	159
-52	39	-27	-33.4(4)	-36	-29.3	19	8.7	139
-52	39	-32	-33.4(4)	-29	-29.3	10	8.7	169
-51	-29	15	15.2(4)	30	31.6	-21	-23.5	37
-51	37	-31	-32.5(2)	-30	-27.1	12	6.9	70
-50.3	38.6	-31.3	-33.4(3)	-30.7	-28.8	12.0	8.2	109
-50	31	-25	-31.7(3)	-23	-18.2	7	-2.1	169
-50	40	-33	-34.2(4)	-32	-30.1	10	9.0	142
-50	41	-32	-34.8(5)	-34	-31.1	15	9.6	136
-47	41	-25	-34.9(5)	-39	-31.0	25	9.5	169
-46.4	39.7	-34.1	-34.2(4)	-29	-29.7	8	8.6	131
-45	37	-28	-33.0(2)	-31	-26.6	13	6.0	169
-41.5	36.0	-31.5	-33.1(2)	-25.1	-24.9	4.5	4.1	138
-41.5	36.0	-31.5	-33.1(2)	-25.1	-24.9	4.5	4.1	155
-40.4	30.8	-13.5	-32.6(2)	-34.9	-16.9	28.0	-4.1	77
-38	36	-34	-33.7(2)	-23	-24.3	1	3.1	125
-37.3	35.6	-32.7	-33.7(2)	-24.0	-23.6	3.2	2.3	155
-37	35	-31	-33.6(2)	-24	-22.7	5	1.5	125

The estimated standard deviations in χ^1 are measures of the uncertainty of the optimum found. These are usually on the order of a few degrees, and the optimum χ^1 usually agrees with the observed value to within this range of uncertainty. The uncertainty tends to be largest in the region where double solutions are found because the optimum is less well defined in these cases.

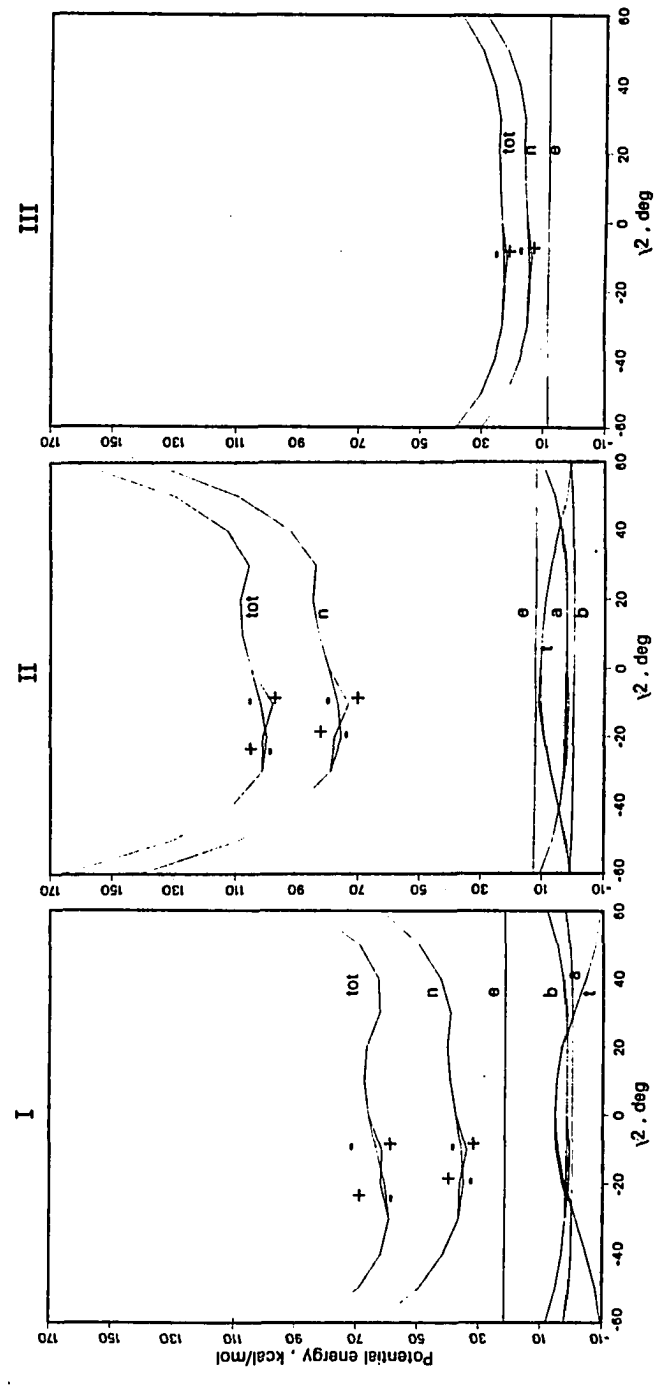
Root mean square deviations between experimental and calculated torsion angles for the 191 structures are 4.8° for χ^1, 4.7° for χ^3, and 8.3° for χ^4. These figures were obtained by taking the best solution in cases where double solutions were found, or in 17 cases where neither solution was clearly superior, the average of the two was taken. The fit to the experimental structures is satisfactory, considering uncertainties in experimental data and the restrictions in the model.

Another measure of success of BORC is the root mean square deviaion of the $N\text{-}C^{\delta}$ bond length and six bond angles of a closed ring from the target values. The sum of squares of fractional deviations was in the range of 5×10^{-5} to 3×10^{-3} for the sample of 191 rings surveyed, corresponding to a root mean square deviation of 0.3% to 2% in each parameter.

Figure 5 illustrates the significance of the various terms in the semi-empirical energy for the case of $\phi = -70°$ (near the mode value for observed structures) using BORC to obtain structures over a range of χ^2. While the three sets give rather different results, they agree in showing a broad energy well in the range of χ^2 shown, the shape of the well being dominated by the Lennard-Jones nonbonded potential. All sets show a shallow double minimum in energy which is enhanced in sets I and II by the

Figure 5. Potential energies for L-proline at $\phi = -70°$

The energies for L-proline at $\phi = -70°$ were calculated using parameter sets I, II, and III. Energies are in kcal/mole. Curves are labelled with the subscripts of the correspondings in the equation for U_{tot}. The "+" and "-" branches correspond, respectively, to initial trial values $\chi_1 = +\chi_2$ and $\chi_1 = -\chi_2$.


I
II
III
Potential energy , kcal/mol
χ2 , deg
tot
n
e
a
b
t
+
-

presence of the torsion energy. Sets I and II show small effects of angle bending and $N-C^{\delta}$ bond stretching near the extremes of the well. It is noteworthy that the bond angles, which are within about 2° of those in Table 3, deviated significantly from the minimum energy values (Table 4), yet the variation in angle strain energy is small over the range of these structures. The variation in electrostatic energy is negligible for all three sets. Where double solutions are obtained, there is a slight bifurcation of the energy curves, most notably in the nonbonded energy. The three sets agree in showing a difference on the order of 10° to 20° in χ^2 at the energy minimum for the two solutions, although the energy difference is so small (< 4 kcal/mol) that the shift is not necessarily significant.

Since, in most cases, a unique ring structure is found when two torsion angles are specified, the energy can be regarded as a function of only two independent torsion angles. Thus, U_{tot} can be represented as a function of various arbitrarily selected pairs of torsion angles (Figures 6, 7, 8 and 9). To limit the selection, ϕ is taken as one member of each pair. The contours are generated from 156 structures produced by BORC spanning the ranges of ϕ and χ^2 in Figures 6-9. In the region where two solutions occur, the solution resulting from initial $\chi^1 = -\chi^2$ was used for these contours. The energy surfaces in which the alternative solution was used differ only in minor respects from those shown here.

The filled circles in Figure 6 represent the torsion angles from experimental structures in the survey. The distributions of the points, as well as the energy surfaces, demonstrate the broad conformational range

Figure 6. Contour map of energy as a function of ϕ and χ^1

Energies were contoured as a function of ϕ and χ^1 for parameter sets I, II, and III. Numbers show U_{tot} in kcal/mole with respect to an arbitrary zero point, which differs for each parameter set. Contour interval is 5 kcal/mole. Filled circles show data from literature structures.

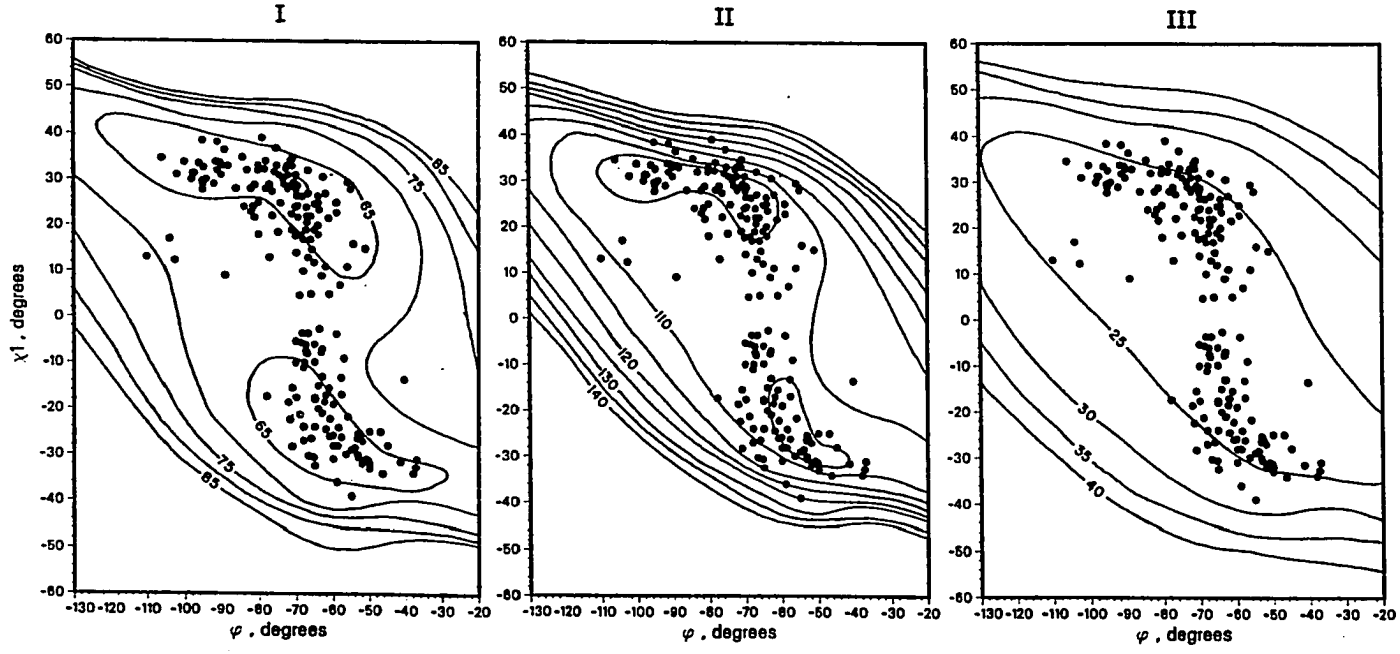

I
II
III
χ1 , degrees
φ , degrees
φ , degrees
φ , degrees
65
75
65
65
75
85
110
120
130
140
25
30
35
40

Figure 7. Contour map of energy as a function of ϕ and χ^2

Energies were contoured as a function of ϕ and χ^2 for parameter sets I, II, and III. Numbers show U_{tot} in kcal/mole with respect to an arbitrary zero point, which differs for each parameter set. Contour interval is 5 kcal/mole. Filled circles show data from literature structures.

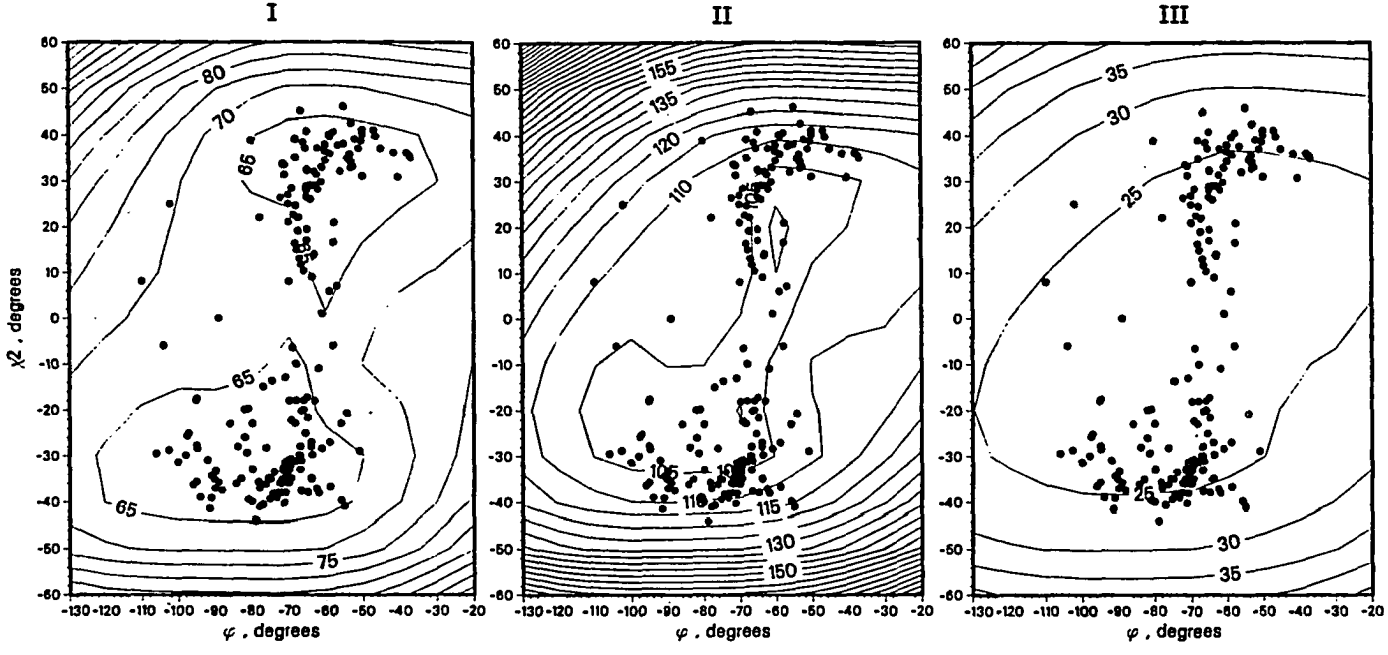

I
II
III
χ2 , degrees
φ , degrees
φ , degrees
φ , degrees
80
70
65
65
65
75
155
135
120
110
105
115
130
150
35
30
25
30
35

Figure 8. Contour map of energy as a function of ϕ and χ^3

Energies were contoured as a function of ϕ and χ^3 for parameter sets I, II, and III. Numbers show U_{tot} in kcal/mole with respect to an arbitrary zero point, which differs for each parameter set. Contour interval is 5 kcal/mole. Filled circles show data from literature structures.

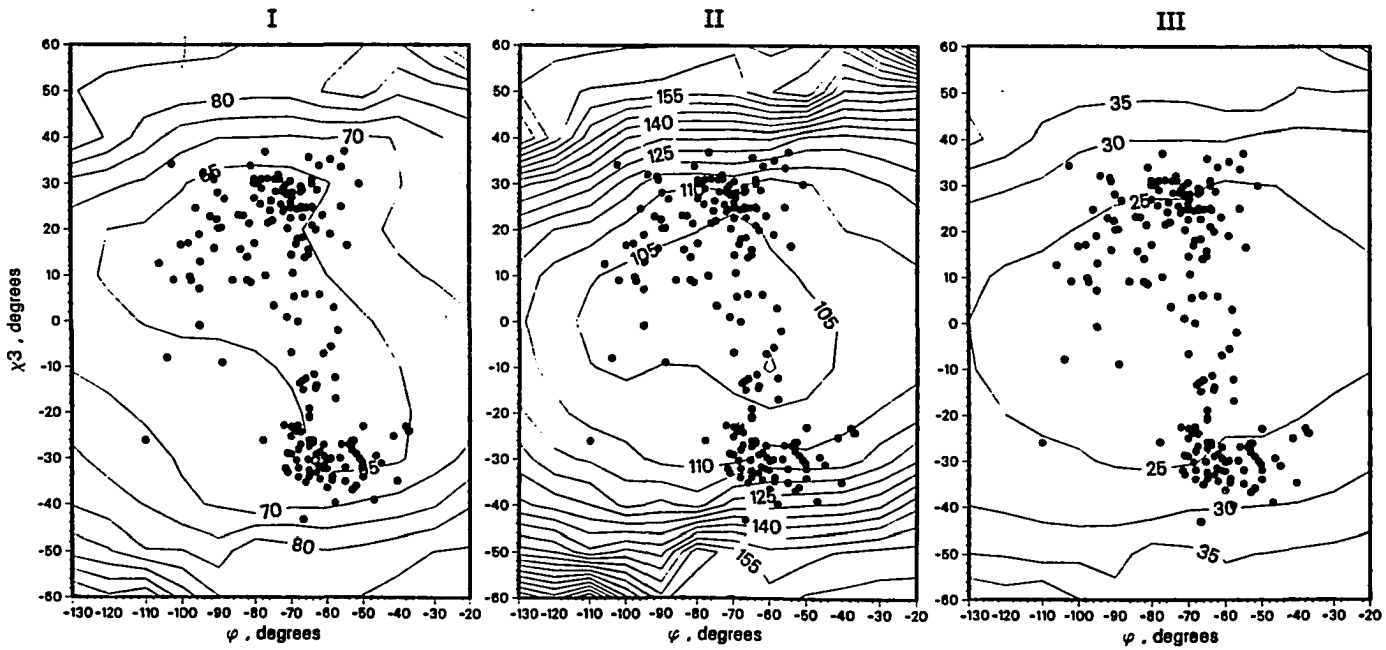
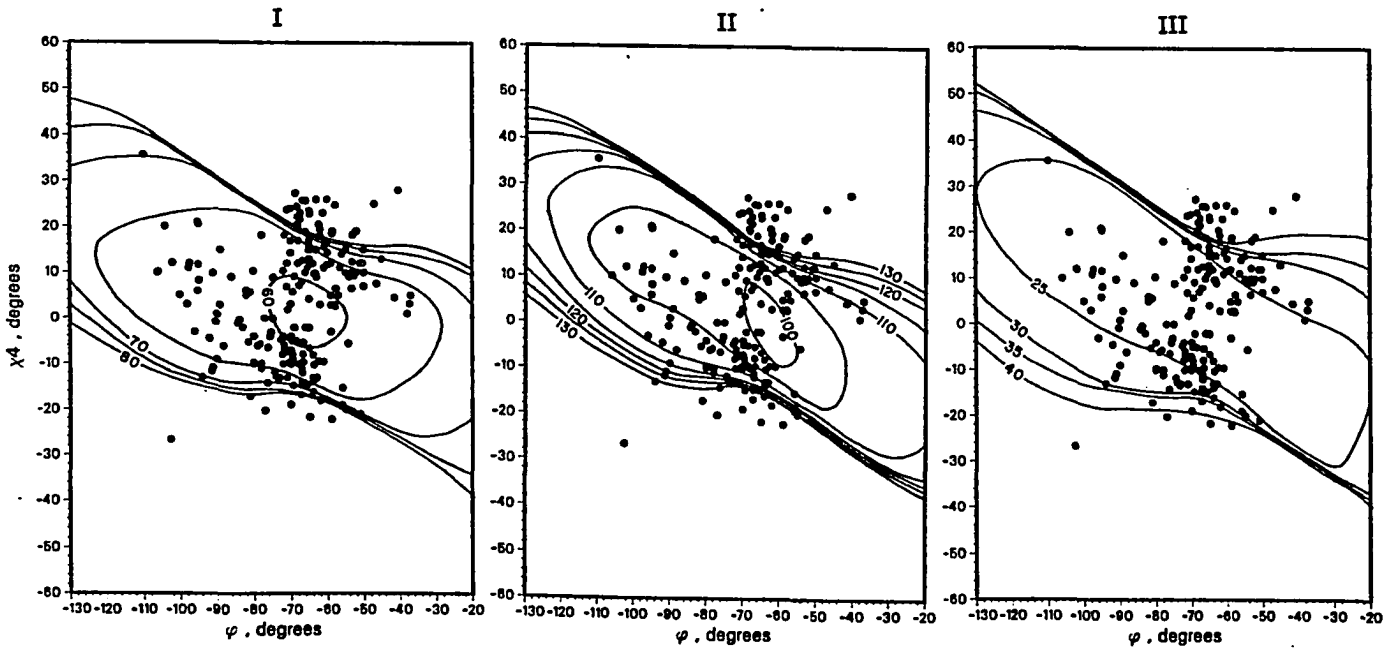

I
II
III
χ3 , degrees
φ , degrees
φ , degrees
φ , degrees
80
70
65
55
70
80
155
140
125
110
105
105
110
125
140
155
35
30
25
25
30
35

Figure 9. Contour map of energy as a function of ϕ and χ^4

Energies were contoured as a function of ϕ and χ^4 for parameter sets I, II, and III. Numbers show U_{tot} in kcal/mole with respect to an arbitrary zero point, which differs for each parameter set. Contour interval is 5 kcal/mole. Filled circles show data from literature structures.


I
II
III
χ4 , degrees
φ , degrees
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-130 -120 -110 -100 -90 -80 -70 -60 -50 -40 -30 -20
90
70
80
130
120
110
100
25
30
35
40

available to the proline ring. This matter is discussed further in the next section.

It may be noted that BORC is equivalent to an energy minimization in which only bond stretching and angle bending terms of Equation 1 are included, using a weighting scheme corresponding to an arbitrary set of force constants. It was found that the BORC solutions were rather insensitive to the weighting factors, so this distinction does not appear to be signficant. It would be of interest, however, to learn whether the inclusion of the remaining energy terms would alter the solutions. This was investigated briefly by substituting the full energy function, U_{tot}, for the sum of squares of residuals that is minimized in the BORC procedure and using the same minimization program. Energy parameter set I was chosen, and ring closure solutions were sought for five structures over a range of ϕ and χ^2 for which experimental structures were available. In two cases the method failed to find any solution, apparently due to the greater complexity of the function. In the remaining three cases, structures similar to those obtained by BORC were found, although some torsion angles differed by as much as 15°. Energies of structures obtained by the two methods agreed to within 0.8 kcal/mol. BORC gave better agreement with experiment in each case. While these results are not conclusive, they do not suggest that better results would be obtained by the more time consuming energy minimization method.

E. Discussion

The main finding of this study is that BORC achieves a satisfactory ring closure which is unique in most cases when only two torsion angles are

fixed independently. The extent of bond angle strain required to achieve closure resembles that found in X-ray crystal structures, and the torsion angles which are dependent variables fall within a few degrees of those in the corresponding experimental structures.

The method led to the conclusion that a broad range of ring conformations is possible with little variation in angle strain, a finding that is consistent with the broad range of experimental structures. The calculated energies offer further insight into the factors influencing conformation.

First, the highest concentrations of observed structures occur in the regions of minimum calculated energy. This is especially true for parameter set I, for which the shapes of the energy wells conform to the shapes of the observed distributions. For all parameters, the majority of the observed structures fall within 5 kcal/mol of the energy minimum. This result is expected if the conformation of the proline ring in the experimental compound is relatively unaffected by interactions with the rest of the molecule, aside from the tendency for ϕ to be fixed by constraints on the peptide backbone. Furthermore, this result may support the proline ring's tendency to disrupt structures such as α-helices and β-pleated sheets and to occur more often in turns of proteins and peptides; in other words, the proline ring affects the backbone structure, not the other way around.

Second, the bimodal distribution of the torsion angles χ^i is evident in the energy contours. Parameter sets I and II produce double energy minima which are nearly coincident with the two modes in the χ^1,ϕ and χ^2,ϕ maps.

The low barrier between the two minima is attributable to the torsion energy term in these sets.

Third, while a given structure generated by BORC has the same energy on any of the maps of Figures 6-9, the experimental structures do not always correspond to the same energy on all maps, since they do not coincide exactly with BORC generated structures. This is particularly evident in the χ^4,ϕ contours, where many structures are seen at energies of 20 kcal/mol or more above the minimum, while such deviations from the minimum are rare in the other ϕ,χ^i contours. This finding may be related to the fact that χ^4 is the least accurately predicted torsion angle in the ring.

Our preliminary effort to carry out energy minimizations that would parallel the BORC method did not suggest that additional computational effort in the energy minimizations would result in more realistic structures. This is consistent with our findings and that of DeTar and Luthra (53) that the ring conformation is rather insensitive to the details of the energy function. The energy map obtained by DeTar and Luthra (53) via energy minimization was comparable to those in Figure 7 in that the two energy minima found were at $\chi^2 = \pm 36°$.

Because of its relative simplicity and speed, BORC lends itself well to a systematic computational study of properties which depend on proline ring conformation. The following chapters report studies on absorption and CD spectra of proline-containing peptides for which BORC was used.

IV. POLY(L-PROLINE)

A. Introduction

Poly(L-proline) is a helical synthetic peptide that has been well characterized in the literature. It exists in two backbone forms, I and II. Form I is a right-handed helix with a cis conformation of the N-C' bond (torsion angle $\omega = 0°$) (162). Form II is a left-handed helix with a trans conformation of the N-C' bond ($\omega = 180°$) (110). The absorption and CD spectra of both forms have been measured in the vicinity of the π-π^* transition near 200 nm and are distinct for the two (183, 184). Moreover, the dipole interaction model has been used with X-ray conformations to correctly predict the π-π^* absorption and CD spectra of both forms; it was the first model to successfully achieve these predictions because it included the nonchromophoric atoms (2). Although the proline ring was included in the previous study, the ring was not closed in a manner consistent with known bond geometry (2). Since small changes in the side-chain structure affect the predicted absorption and CD spectra, a further study of BORC proline rings with various values for χ^2 could provide more information that would help to determine the structure in solution.

Those proline rings generated by BORC for arbitrary choices of χ^2 are referred to as BORC rings. Those rings referred to as nonoptimized merely use the torsions and bond angles published in the X-ray structures with the expanded C-C bond lengths of 1.54 Å. The C-C bonds were expanded because in the X-ray literature these distances are often short and interfere with optical calculations (2).

Since intermolecular interactions present in the crystal lattice are absent in solution, the proline ring conformation may be different from that in the solid state. For example, nuclear magnetic resonance studies of poly(L-proline) II in solution have indicated that individual proline conformations of large puckering are in rapid equilibrium (185). These studies suggest χ^2 values around $\pm$ 35° for ϕ = -60 degrees (185) instead of the X-ray values of χ^2 at -19° (110). The present study is intended to provide further information that would help resolve the uncertainty of the structure in solution.

B. Structure Generation

The poly(L-proline) helices were generated using the Ramachandran-Sasisekharan technique (43) described in the methods chapter. The parameters of the poly(L-proline) backbones are derived from the X-ray structures given in Table 8 (110, 162). The nonoptimized and BORC rings were generated similarly using parameters listed in Table 9.

The $(Pro)_{10}$ I helix generated has 3.34 residues/turn and an axial translation of 1.88 Å/residue compared to the literature values of 3.33 residues/turn and 1.90 Å/residue (162). The $(Pro)_{10}$ II helix generated with X-ray parameters has 3.02 residues/turn and 3.13 Å/residue which compare with the literature values of 3.00 residues/turn and 3.12 Å/residue (110). The poly(L-proline) II helix generated with NMR backbone parameters has only 2.86 residues/turn and an axial translation of 2.89 Å/turn which both fall short of the X-ray data in the fiber state (Table 8).

π-π^* absorption and CD spectra were calculated for both forms of poly(L-proline) with 10 residues using a half-peak bandwidth, Γ = 6000

Table 8. Backbone structural parameters for poly(L-proline) I and II

		$(Pro)_n$I (162)	$(Pro)_n$II (110)	$(Pro)_n$II (185)
Bond Lengths (Å)	C^{α}-C'	1.52	1.53	-
	C'-O	1.24	1.24	-
	N-C'	1.32	1.32	-
	N-C^{α}	1.48	1.48	-
Bond Angles (deg)	C'C^{α}N	114.8	110.0	-
	$C^{\beta}C^{\alpha}$C'	110.5	112.4	-
	C^{α}C'O	119.0	121.0	-
	C^{α}C'N	119.0	114.0	-
	C'NC^{α}	126.2	120.8	-
Torsion Angles (deg)	ϕ	-83.0	-77.9	~-60
	ψ	157.3	147.0	~120
	ω	0.0	180.0	180.0

Table 9. Proline rings of poly(L-proline): BORC versus nonoptimized

Parameter	Poly(L-proline) I Nonptimized	Poly(L-proline) I BORC	Poly(L-proline) II Nonoptimized	Poly(L-proline) II BORC
ϕ	-83.0	-83.0	-77.9	-77.9
χ^1	23.6	29.9	20.7	29.6
χ^2	-18.2	-18.2	-19.0	-19.0
χ^3	4.4	-0.7	9.1	0.7
χ^4	12.8[a] (12.3)	20.8	4.7[a] (4.8)	19.3
τ_1^R	105.0	102.8	105.8	102.6
τ_2^R	100.7	105.1	104.5	105.2
τ_3^R	116.0	106.2	111.6	106.2
τ_4^R	96.5[a] (100.0)	105.4	101.6[a] (100.9)	105.1
τ_5^R	115.4[a] (113)	110.2	112.1[a] (112.8)	111.1
τ_6^R	118.4[a] (121)	122.9	125.9[a] (125.2)	124.5
τ_6	126.2	126.2	120.8	120.8
NC^δ	1.531[a] (1.48)	1.476	1.551[a] (1.51)	1.465
$C^\alpha C^\beta$	1.54[b] (1.52)	1.54	1.54[b] (1.51)	1.54
$C^\beta C^\gamma$	1.54[b] (1.48)	1.54	1.54[b] (1.50)	1.54
$C^\gamma C^\delta$	1.54[b] (1.49)	1.54	1.54[b] (1.53)	1.54

[a]Values calculated from the ring generated with X-ray parameters using C-C bond length of 1.54Å; values in parentheses are the actual X-ray values.

[b]1.54Å was used for the C-C bond length in optical calculations because it prevented too close contacts and better resembled typical hydrocarbon C-C bonds (2).

cm^{-1}. This chain length was chosen for comparison with experimental spectra of high molecular weight $(Pro)_n$.

C. Results

BORC and nonoptimized rings with the same ϕ,χ^2 are compared in Table 9. The BORC rings were similar to the nonoptimized rings with the following substantial differences. χ^1, χ^3, and χ^4 were on the average 9° different; ring angles were on the average 4° different, and the NC^δ bond length deviated 0.064 Å for $(Pro)_{10}$ I and 0.095 Å for $(Pro)_{10}$ II. The largest torsion deviations were found in χ^4 which probably result from the accumulation of other changes in the bond angles and bond length. The other major deviations resulted directly from geometric optimization of the proline ring (both the NC^δ bond and the $\langle C^\beta C^\gamma C^\delta$ angle were decreased to more commonly found values).

The π-π^* spectra calculated for poly(L-proline) I and II for the nonoptimized structures and the BORC structures at the χ^2 X-ray value are given in Figure 10. Calculated spectra for BORC side chain conformations other than those at the χ^2 X-ray value are included in Figures 11 and 12. These figures indicate the sensitivity of the predictions by dipole interaction model to changes in the side chain conformations, and for both molecules some ring structures give a reasonably good overall fit. Lorentzian band shapes used for mathematical convenience account in part for discrepancies between predictions and experiment, especially near the endpoints of the predicted spectra. Torchia's backbone parameters irrespective of the side chain conformations gave structures that had

Figure 10. $\pi-\pi^*$ absorption and CD spectra predicted for nonoptimized and BORC structures of poly(L-proline) I and II

Units for the scales are as follows: nm for λ, $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled experiment for the experimental spectra, BORC for the spectra predicted using BORC proline rings, and nonoptimized for the spectra predicted using the x-ray crystallographic torsion and bond angles. The scale listed for the lower left is the same for all other spectra. Poly(L-proline)I spectra are on the left. Poly(L-proline)II spectra are on the right. The bandwidths in both cases are 6000 cm^{-1}.

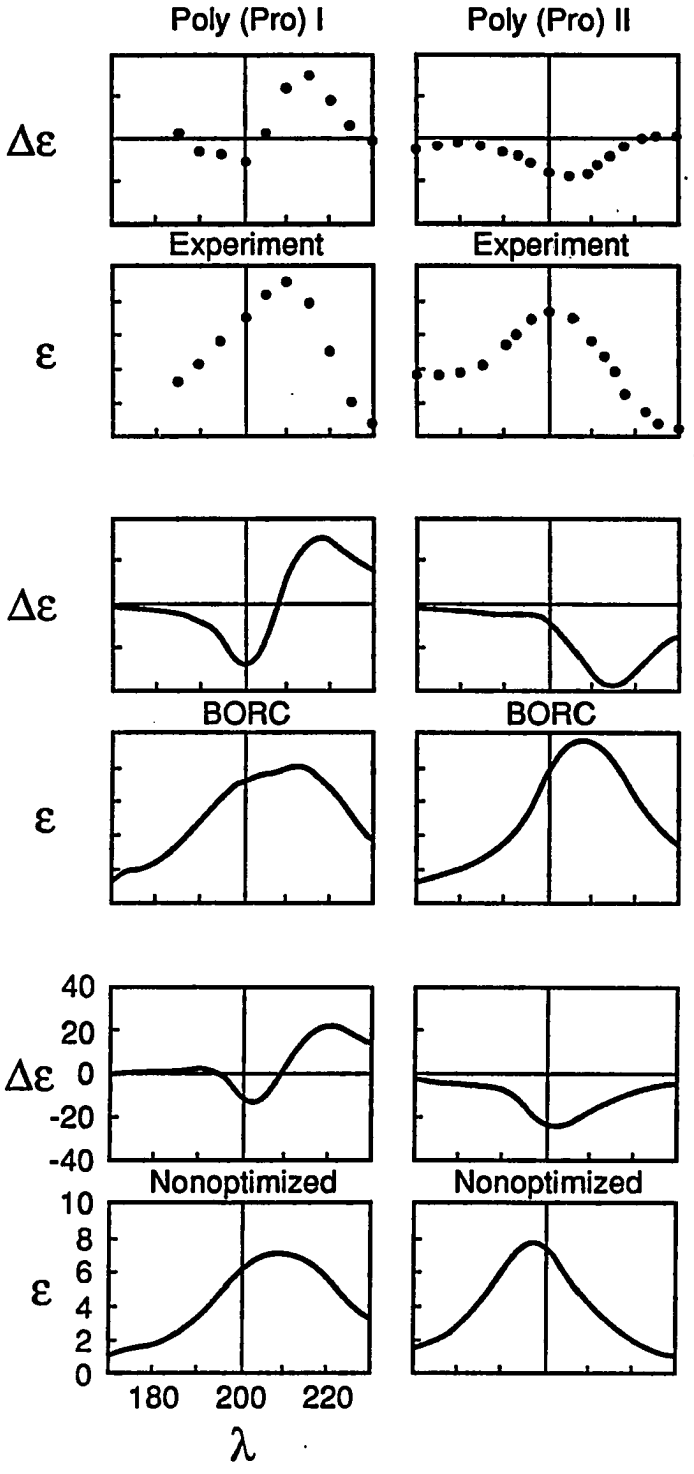

Poly (Pro) I
Poly (Pro) II
Δε
Experiment
Experiment
ε
Δε
BORC
BORC
ε
40
20
Δε 0
-20
-40
Nonoptimized
Nonoptimized
10
8
6
ε 4
2
0
180
200
220
λ

Figure 11. $\pi-\pi^*$ absorption and CD spectra for poly(L-proline) I

Units for the scales are as follows: nm for λ, $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled according to χ^2 value. All spectra are to the scales listed for the lower left pair. The bandwidth is 6000 cm^{-1}.

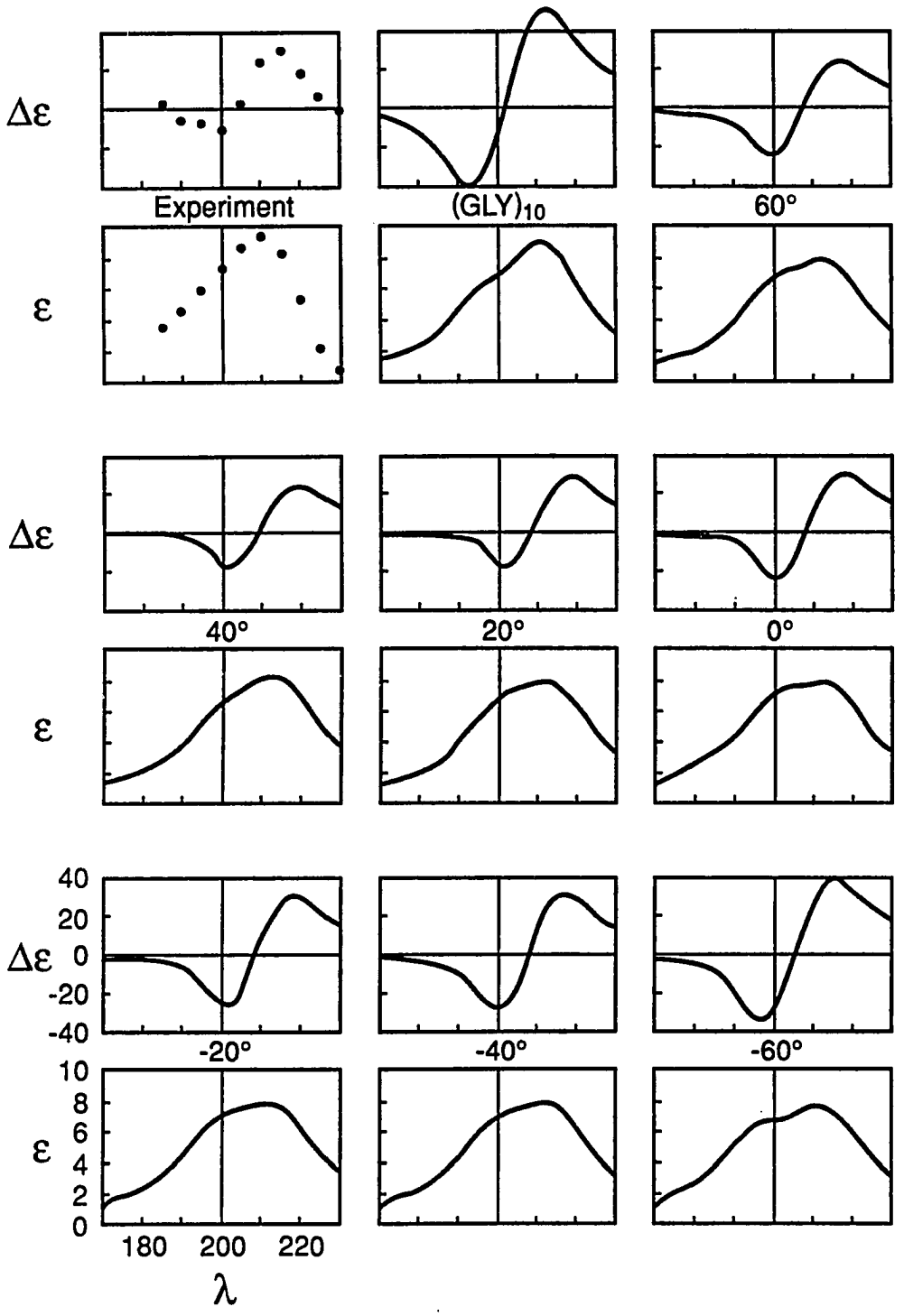
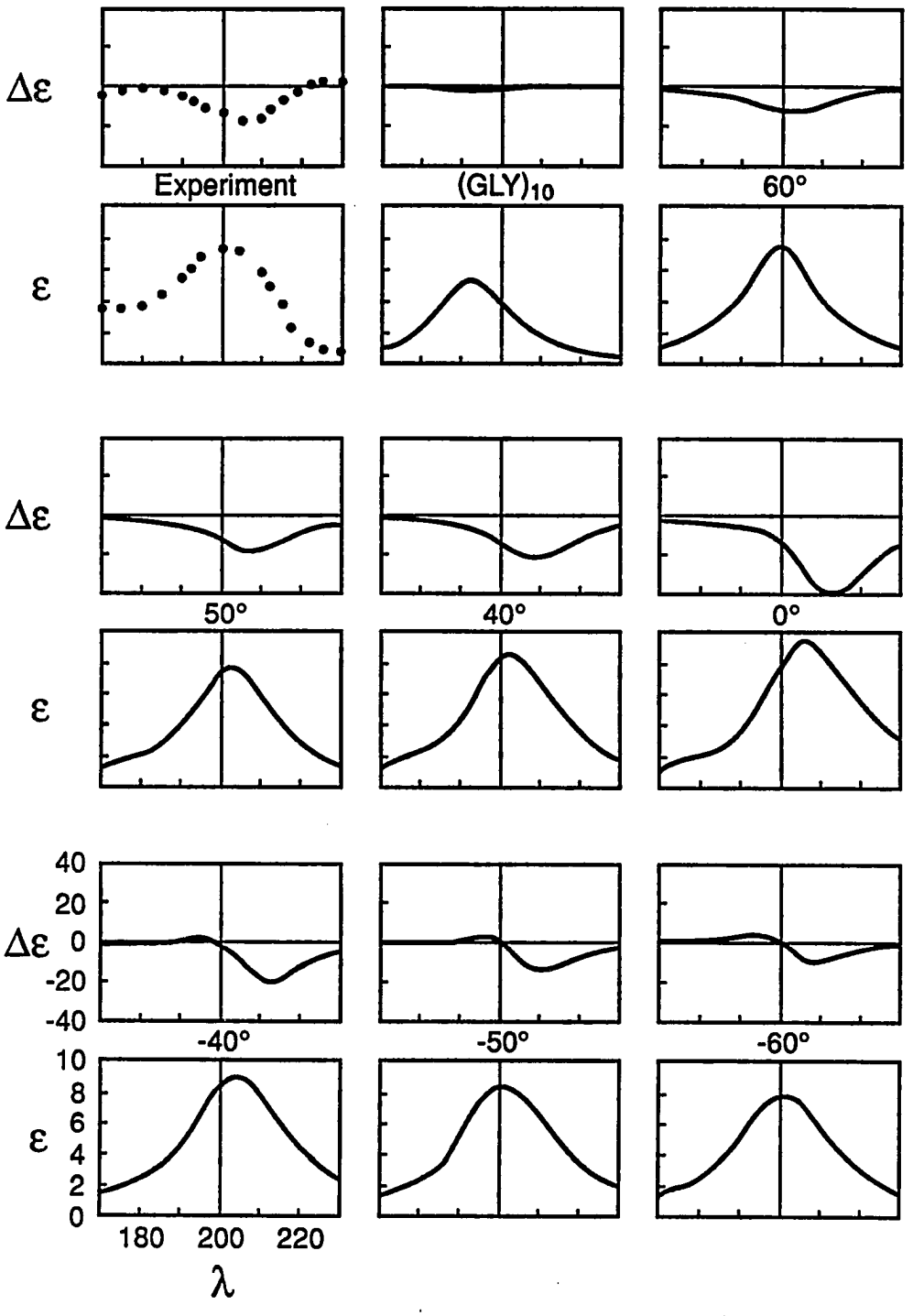

Δε
Experiment
(GLY)10
60°
ε
Δε
40°
20°
0°
ε
40
20
Δε 0
-20
-40
-20°
-40°
-60°
10
8
ε 6
4
2
0
180
200
220
λ

Figure 12. $\pi-\pi^*$ absorption and CD spectra for poly(L-proline) II

Units for the scales are as follows: nm for λ, $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled according to χ^2 value. All spectra are to the scales listed for the lower left pair. The bandwidth is 6000 cm^{-1}.


Δε
Experiment
(GLY)10
60°
ε
Δε
50°
40°
0°
ε
40
20
Δε 0
-20
-40
-40°
-50°
-60°
10
8
ε 6
4
2
0
180
200
220
λ

predicted CD with a positive band near 200 nm which is contrary to experiment (Table 10).

The oscillator strengths, rotational strengths, and wavelengths of the $\pi-\pi^*$ modes are listed for each structure in Tables 11 and 12. Of the 10 modes found for the 10-mer, the four modes above 190 nm dominate the oscillator and rotational strengths for both forms (2). These tables include comparisions of experimental absorption and CD of $(Pro)_n$ solutions for the $\pi-\pi^*$ bands (186-188) whose electric dipoles are parallel (subscript 11) and perpendicular (subscript $\perp$) to the helix axis.

The relative magnitudes of the oscillator strengths of the resolved bands are of interest in connection with Moffitt's theory for the spectra of helical molecules assuming noninteracting chromophores (2, 189). Table 13 shows the resulting values of $f_\perp/f_{11}$ from Moffitt's theory do not agree with the experimental data as pointed out earlier (2). The ratio calculated by the dipole interaction model is in agreement for the nonoptimized structures of both forms and for $\chi^2 = 20°$ for poly(L-proline) I and $\chi^2 = -40°$ for poly(L-proline) II. These χ^2 choices are in partial agreement with the CD spectral matches at $\chi^2 = \pm 20°$ for poly(L-proline) I and $\pm 50°$ for poly(L-proline) II. Some side chain conformations of poly(L-proline) I produce predictions no better than Moffitt's, but no conformations overshoot experiment as much as Moffitt's predictions for poly(L-proline) II.

Table 10. π-π^* Data for two forms of poly(L-proline)II

	Experiment (183)	Backbone by X-ray[a] (110)	Backbone by NMR[a] (185)
A. Absorption			
λ^b_{max} (nm)	200	204	200
ε^c_{max} ($Lmol^{-1}cm^{-1}$)	7100	8010	9203
B. Circular Dichroism			
λ^b_{max} (mn)	180	150	204
$\Delta\varepsilon^c_{max}$ ($Lmol^{-1}cm^{-1}$)	-1.3	-1.0	18.2
λ^b_{min} (nm)	205	208	214
$\Delta\varepsilon^c_{min}$ ($Lmol^{-1}cm^{-1}$)	-18.0	-22.0	-10.0

[a] $\chi^2 = 40°$ for both cases; backbone parameters are given in Table 8.

[b] λ is the wavelength where the maximum (max) or minimim (min) occur in the spectra.

[c] ε is absorption and $\Delta\varepsilon$ CD; when these are maxima they are denoted max; when these are minima they are denoted min.

Table 11. Wavelengths, splittings, oscillator strengths and rotational strengths of dominant π-π* modes in poly(L-proline)I

	exptl.	$(Gly)_{10}$	60	40	20	χ^2 (deg) 0	-20	-40	-60	-18.2[a]	-18.2[b]
λ_{11} (nm)	208.7[c] 211[d]	211.3	213.8	214.9	215.2	215.3	214.6	214.1	213.7	214.8	216.1
$\lambda_{\perp}$ (nm)	200.5[c] 199[d]	193.3	197.4	199.6	199.7	198.9	198.4	198.0	195.4	198.5	201.5
$\bar{\nu}_{\perp}-\bar{\nu}_{11}$ (cm^{-1})	1960[c] 2860[d]	4407	3886	3567	3607	3830	3805	3798	4382	2823	3353
f_{11}	0.122[c]	0.317	0.250	0.243	0.232	0.224	0.228	0.240	0.240	0.227	0.187
$f_{\perp}$	0.091[c]	0.116	0.145	0.158	0.173	0.172	0.159	0.156	0.163	0.160	0.144
$\bar{f}$ (all modes)	0.33[e]	0.455	0.418	0.428	0.429	0.431	0.431	0.431	0.436	0.431	0.395
R_{11} $\frac{DBM}{residue}$	0.78[c] 0.306[f]	3.50	2.17	2.30	2.51	2.71	2.78	2.76	3.12	2.78	2.06
$R_{\perp}$ $\frac{DBM}{residue}$	-.55[c] -.196[f]	-2.96	-1.98	-1.79	-1.98	-2.72	-3.18	-2.92	-3.07	-3.18	-2.43

[a] BORC generated ring.

[b] Improperly closed ring.

[c] Reference 187.

[d] Reference 186.

[e] Reference 2.

[f] Reference 188.

Table 12. Wavelengths, splittings, oscillator strengths, and rotational strengths of dominant $\pi \rightarrow \pi^*$ modes on poly(L-proline)II

	exptl.	$(Gly)_{10}$	60	50	40
λ_{11} (nm)	208.6[c] 209[d]	192.6	203.4	206.3	208.1
$\lambda_{\perp}$ (nm)	201.3[c] 191[d]	191.7	197.3	200.1	201.8
$\bar{\nu}_{\perp} - \bar{\nu}_{11}$ (cm^{-1})	1740[c] 4510[d]	1196	1520	1502	1500
f_{11}	0.097[c]	0.087	0.084	0.098	0.115
$f_{\perp}$	0.106[c]	0.110	0.186	0.202	0.070
$\bar{f}$ (all modes)	0.29[e]	0.240	0.310	0.340	0.360
R_{11} (DBM/residue)	-0.347[f]	-0.22	-0.90	-1.94	-1.34
$R_{\perp}$ (DBM/residue)	0[f]	0.03	-0.30	-0.13	0.16

[a]BORC generated ring.

[b]Improperly closed ring.

[c]Reference 187.

[d]Reference 186.

[e]Reference 2.

[f]Reference 188.

χ^2 (deg)							
20	0	-20	-40	-50	-60	-19[a]	-19[b]
210.6	212.4	213.3	209.3	207.6	205.3	212.4	201.0
203.6	203.8	202.7	199.9	198.2	195.7	202.8	193.8
1632	1987	2231	2247	2284	2389	2229	1848
0.160	0.203	0.212	0.176	0.175	0.176	0.212	0.152
0.204	0.195	0.194	0.181	0.138	0.114	0.194	0.166
0.403	0.439	0.449	0.413	0.397	0.377	0.448	0.356
-1.96	-2.51	-2.51	-1.68	-1.31	-0.96	-2.53	-1.62
0.86	0.86	0.73	-0.86	-3.03	2.88	0.74	0.24

Table 13. Oscillator strength ratios $f_{\perp}/f_{11}$ for (Pro)n

	I	II
Experiment[a]	0.75	1.09
Moffitt Theory	0.54	7.60
Present Theory:		
$(Gly)_{10}$	0.37	1.26
χ^2 = 60	0.58	2.21
(deg) 50	-	2.06
40	0.65	0.61
20	0.74	1.28
0	0.70	0.96
-20	0.55	0.92
-40	0.65	1.03
-50	-	0.79
-60	0.68	0.64
closed modified x-ray structure[b]	0.70	0.92
modified x-ray structure[b]	0.77	1.09

[a]Reference Jenness et al. Biopolymers, 1976.

[b]Modified x-ray structures are those using χ^2 = to the value in the x-ray structure.

D. Discussion

1) Effect of the BORC treatment

Optimizing the ring by BORC has no direct affect on the helix backbone, but it does affect both intensity and band location for abosrption and CD spectra. For poly(L-proline) I, the BORC ring compared to the nonoptimized ring has a greater effect on the intensity of the $\pi-\pi^*$ spectra than on the location of the bands (Figure 10). The increased intensities are reflected in the slightly larger values of the oscillator and rotational strengths for both parallel and perpendicular modes (Table 11). The intensity of the spectra for the BORC ring better approaches experiment.

Optimizing the ring by BORC has a greater effect on the spectra of poly(L-proline) II. Just as poly(L-proline) I spectra were more intense with BORC, so are poly(L-proline) II spectra more intense (Figure 11). Again the oscillator and rotational strengths are stronger for the BORC form (Table 12). These changes, and their effects on the predicted spectra imply that the nonoptimized ring structure itself cannot accurately represent those ring structures found in solution. It should be noted that the oscillator and rotational strengths of the nonoptimized rings are slightly different from those of the previous calculation (2) because the earlier study contained a slight error in locating the β, γ, and δ hydrogens (44).

Based on the ring parameters alone, the BORC rings are more realistic because the ring closure took the increased C-C distances into account when the bond and torsion angles and the $N-C^{\delta}$ bond were adjusted so that they are more consistent with known proline geometry. Based on the predicted

spectra, the location of the band is dependent on the intensity of the puckering, and the BORC rings of intense puckering coincide with experiment (Figure 12). These spectra suggest a higher degree of puckering than the previous calculation.

2) Visual comparison of π-π^* spectra

a) Poly(L-proline) I The most intense absorption and CD spectra produced for poly(L-proline) I occur when the proline residue is replaced by glycine. Although the absorption band resembles experiment in location and intensity, the CD is much too strong so that the proline side chain is needed to reduce the CD intensity into the experimental region (Figure 11). The absorption spectrum is not as sensitive to the side chain conformation as the CD spectrum, but in both cases the main quantity changing with χ^2 is the intensity of the bands. CD intensities most resembling experiment occur for $\chi^2 < 0°$ in the band around 214 nm with the best agreement occurring between $\chi^2 = -20°$ and $-40°$. This does not rule out interconversion with the equivalent positive χ^2s because for $\chi^2 = 20°$ the CD intensity of the 200 nm band most resembles experiment. Extreme puckering ($\chi^2 = \pm 60°$) is not indicated because the predicted spectra are either too weak or too intense.

b) Poly(L-proline) II The spectra of poly(L-proline) II are even more sensitive to side chain conformations than poly(L-proline) I (Figure 12). When the proline residue is replaced by glycine, a weak and slightly blue-shifted absorption band appears and $\Delta\varepsilon$ is about 2 $Lmol^{-1}cm^{-1}$. Once the proline side chain is added, strong CD and absorption spectra appear. In all cases, the intensity of the absorption band increases and slightly

red shifts systematically as χ^2 approaches 0° with the weakest spectra being at $\chi^2 = \pm 60°$. Those CD and absorption most resembling experiment occur around $\chi^2 = \pm 50°$ with the positive value slightly more favorable because the band is more red-shifted for the negative χ^2 values.

3) Side chain effects on π-π^* dominant modes

For poly(L-proline) I the parallel modes predicted by the dipole interaction model are generally slightly red-shifted relative to experiment (Table 11). The perpendicular modes resemble experiment more especially when $\chi^2 = 40°$ through -40°. The red-shift in parallel modes increases the predicted splitting, $\bar{\nu}_{\perp} - \bar{\nu}_{11}$ in wavenumbers, but still those predictions agree with experiment (186) best when $\chi^2 = \pm 40°, \pm 20°$ for poly(L-proline) I. Parallel mode predictions better coincide with experiment (161) for poly(L-proline) II when intensely puckered $\chi^2 = \pm 40°, \pm 50°$ (Table 12). Perpendicular predictions also coincide with experiment (187) the same way.

As in the previous study (2), the oscillator strengths per residue calculated for the resolved bands, f_{11} and $f_{\perp}$, for all poly(L-Proline) I forms and most poly(L-proline) II forms are substantially larger than the experimental values obtained from film spectra. This discrepancy may be related to difficulties in interpreting film spectra (187). Poly(L-proline) II forms showed the best agreement with experiment (187) when $\chi^2 = \pm 50°$. The proline side chain does red shift the parallel and perpendicular modes in both cases to better coincide with experiment.

Moreover, as in the previous study (2), the calculated rotational strengths of the two bands, R_{11} and $R_{\perp}$, although correct in sign, are substantially larger than experiment. The experimental values, however,

depend strongly on assumptions about bandwidth and shape which information is lost in the overlapping of bands of opposite sign (2). Visual comparison of CD spectra (Figures 11 and 12) is a more valid test of the theoretical predictions than the resolved rotational strengths.

E. Conclusions

The principal finding of this study is that approximately correct $\pi-\pi^*$ absorption and CD spectra of poly(L-Proline) I and II are predicted by the dipole interaction model only when the proline side chain is included and intensely puckered. The region of puckering predicted is χ^2 = +20° to +40° and -20° to -40° for poly(L-Proline) I and χ^2 = +40° to +50° and -40° to -50° for poly(L-Proline) II. The proline structures given by nonoptimized structures are not puckered as strongly as solution predictions indicate and are not closed by evenly distributing angle strain throughout the entire ring as the BORC rings. The present evidence implies that the sensitivity of the dipole interaction model to side chain structure can be used to produce information about the side chain conformations in solution. This study supports Torchia's (185) finding that poly(L-Proline) II side chains are represented by a combination of interconverting intensely puckered structures. Such independent support provides additional credibility to the dipole interaction model.

V. CYCLO(PRO-GLY)$_3$ COMPLEXES WITH METAL IONS

A. Introduction

Many naturally occurring cyclic peptides bind ions and participate in ion transport across membranes. Examples of such naturally occurring peptides include enniatin (18, 34), beauvericin (18), valinomycin (17, 18, 34, 35), momomycin (18), ferrichrysin (18, 25, 26), ferrichrome (18), asperchrome (26), and antamanide (1, 18, 35). Cyclo(prolyl-glycyl)$_3$ (referred to as cyclo(Pro-Gly)$_3$ hereafter) is a synthetic peptide that mimics ion binding and transport proprties of natural peptides (33-35, 190, 191). It also makes a good model for the binding of substrate to enzymes becuse it binds larger moieties such as amino acid salts (192). The cations which cyclo(Pro-Gly)$_3$ binds include Mg^{2+} (33, 34, 190), Ca^{2+} (33-35, 126, 190), Na^+ (126, 191), K^+, Li^+, Rb^+, Cs^+ (191), and RNH_3^+ and RNH_2^+ (e.g., Val-OMe and Pro-OBz, respectively) (192).

X-ray crystallographic studies have been done for several of the cation complexes: Na^+ (126), Ca^{2+} (35), and Mg^{2+} (34). Moreover, information on solution backbone conformations has been obtained by NMR, energy minimizations, and CD studies which have included theoretical CD calculations (1, 33, 38). These studies have shown that the cation complexes are C_3 symmetric; furthermore, the CD spectra were somewhat different depending on the cation (33).

The previous solution studies have placed the cation complexes in two conformational classes, one for a Mg^{2+} complex and one for other cations (33). Both classes are C_3 symmetric, have all six peptide bonds trans, and have proline ψ angles near trans. The main differences between the two are

stoichiometries (33). The ratio of two Mg^{2+} to one cyclo(Pro-Gly)$_3$ has been observed in solution where the Mg^{2+} cations are coordinated to the proline carbonyls on one side of the ring and to the glycine carbonyls on the other side forming a Mg^{2+} sandwich (33). Other cations including Mg^{2+} can form 1:1 complexes with the cation coordinated to only the glycine carbonyls (33).

The crystal structures show similarities to these proposed solution forms. The Na^+ and Ca^{2+} mixed cation complex forms a continuous sandwich (Na^+:cyclo(Pro-Gly)$_3$:Ca^{2+}:cyclo(Pro-Gly)$_3$:....) where the cyclohexapeptide is C_3 symmetric with all peptide bonds trans and the glycyl carbonyls and proline carbonyls pointing to opposite sides of the ring (just as the proposed solution forms did) (126). The Ca^{2+} crystal structure was a Ca^{2+} sandwich with two different forms of the peptide on either side of the cation; both were also C_3 symmetric with all trans peptide bonds and one had the glycine and proline carbonyls pointing to opposite sides of the rings as in the solution conformations (34). The Mg^{2+} crystal complex had a 1:1 ratio with only approximate C_3 symmetry; again all peptide bonds were trans, and the proline and glycine carbonyls pointed to opposite sides of the peptide ring (34). Thus, these crystal structures support cation complex structures proposed for the solution conformations. Furthermore, molecular mechanics calculations for the binding of Ca^{2+} to cyclo(Pro-Gly)$_3$ also produce structures with these same features (36).

The theoretical CD calculations by Madison et al. (33) were somewhat different from the dipole interaction model. First, their calculations did not include the proline side chain (33), and the dipole interaction model

can cope with side chains. Second, their calculations attributed the bands above 200 nm to the n-π* transition and those below 200 nm to the π-π* transition (33). The reliability of this approach was never well established and was subject to considerable doubt because it failed to predict correct CD spectra for the α-helix and poly(L-proline) II helix; the dipole interaction model, on the other hand, has successfully predicted both the π-π* CD and absorption spectra of these helices (2, 6-11, 13). Moreover, the earlier CD calculations involved adjustment of the wavelengths of the isolated π-π* and n-π* transitions for different conformations (33). The dipole interaction model uses a single set of established parameters for the π-π* transition regardless of the conformation.

Since the dipole interaction model has predicted reasonable CD for helices, cyclic dipeptides, and β-structures (2,3,6-11, and Chapter IV), and since it includes the proline side chain, application of this model to the structure of cyclo(Pro-Gly)$_3$ cation complexes may provide further insight to the proposed solution backbones and the puckering of the proline ring for this molecule.

B. Structure Generation

Cyclo(Pro-Gly)$_3$ forms were generated using the Ramachandran-Sasisekharan technique (43) described in the methods chapter. The parameters of the cyclo(Pro-Gly)$_3$ cation complex backbones are derived from the published minimum energy structures (33) or the X-ray structures (147, 155). Backbone rings were closed as described in the methods chapter. They needed to be closed in order to keep the end of the chain in

the proper geometry with as few modifications to the literature ϕ and ψ as possible. The literature parameters needed to be revised to remove the differences between the Ramachandran-Sasisekharan parameters and those in the literature. Often in the litrature, no $\angle NC^{\alpha}C'$ angle was provided, and/or the other bond lengths or angles differed from Ramachandran-Sasisekharan values. Furthermore, each residue used in the current study has the same bond angles and bond lengths regardless of whether the residue is proline or glycine. Crystal structures often distinguish the two and even have variations among residues of the same kind. After backbone ring closure, BORC was applied to the proline rings for the resulting values of ϕ and τ_1 to produce a wide range of proline rings for optical studies.

$\pi-\pi^*$ absorption and CD spectra were calculated for both the Mg^{2+} and the Ca^{2+} complexes although the cations themselves were excluded from the optical calculation. Half-peak bandwidths varied from $\Gamma = 4000$ cm^{-1} to 6000 cm^{-1} depending on the form. The backbone forms have been coded C_3 for C_3 symmetry and arbitrarily numbered after that as follows: $C_3 2$ is the solution Mg^{2+} structure; $C_3 3$ is the solution Ca^{2+} structure (both these backbones were proposed by Madison et al. and designated S_2^* and S_1^*, respectively by them (33)); $C_3 5$ and $C_3 6$ are the pseudo C_3 symmetric X-ray structures for the Mg^{2+} complex; $C_3 8$ and $C_3 9$ are the Ca^{2+} crystal structures. Both solution structures were determined by energy minimizations and NMR (33). The Ca^{2+} crystal structure is a true sandwich compound with two different forms of the peptide bound; the Mg^{2+} complex, however, exhibits 1:1 stoichiometry for the peptide and cation, but the unit cell contains two different peptide molecules (34). Bandwidths used

for these forms are as follows: $C_3$2, 4000 cm^{-1}; $C_3$3, 6000 cm^{-1}; $C_3$5, 4000 cm^{-1}; $C_3$6, 4000 cm^{-1}; $C_3$8, 6000 cm^{-1}; and $C_3$9, 4000 cm^{-1}. The choice between the two bandwidths was made according to which one gave the better agreement with experiment.

C. Results

The backbone parameters of the closed rings used for the cation complexes of cyclo(Pro-Gly)$_3$ are listed in Table 14. Backbone ring closure produced changes in the ϕ and ψ angles that were at most 4°. Table 15 contains the proline ring parameters determined by BORC. The completely C_3 symmetric backbones $C_3$2, $C_3$3, $C_3$8, and $C_3$9 all had the same proline ring parameters used for all three rings in a given backbone becuase they all had the same ϕ,χ^2 pair. The pseudo C_3 symmetric peptides, $C_3$5 and $C_3$6, were treated as asymmetric so that the three proline rings could be different. These two structures were called psuedo symmetric because they only approximated C_3 symmetry. Their torsion angles would be C_3 symmetric to within 10° in most cases and within 20° in all cases.

Hori and coworkers calculated energy minima for cyclo(Pro-Gly)$_3$ Ca^{2+} complexes which began at the X-ray structures $C_3$8 and $C_3$9 (36) (Table 16). Although their resulting ϕ, ψ, and ω angles differed from the original X-ray structures by as much as 40°, the gross overall structures (C_3 symmetry, near trans peptide bonds, and the direction of the carbonyls with respect to the peptide ring) were maintained (36). The absorption and CD spectra were calculated for these structures, and the results are listed in Table 17.

Table 14. Parameters for backbones of cation complexes of cyclo(Pro-Gly)$_3$

Cation	Complex Code[a]	Ref	ϕ_p^1	ϕ_g^2	ϕ_p^3	ϕ_g^4	ϕ_p^5	ϕ_g^6	ψ_p^1	ψ_g^2
Mg^{2+}	$C_3$2		-68.2	68.9	-68.2	68.9	-68.2	68.9	150.2	-151.1
		33	-68	69	-68	69	-68	69	150	-150
Mg^{2+}	$C_3$5		-60.9	84.7	-63.3	78.0	-56.8	79.2	142.9	-171.7
		34	-64	83	-63	79	-57	77	142	-171
Mg^{2+}	$C_3$6		-55.6	83.7	-61.5	91.8	-60.9	82.4	142.1	173.6
		34	-55	86	-63	92	-59	82	144	172
Ca^{2+}	$C_3$3		-42.1	83.7	-42.1	83.7	-42.1	83.7	130.0	172.2
		33	-42	84	-42	84	-42	84	130	172
Ca^{2+}	$C_3$8		-64.0	84.1	-64.0	84.1	-64.0	84.1	144	-178.1
		35	-64	85	-64	85	-64	85	144	-180
Ca^{2+}	$C_3$9		-68.0	-83.7	-68.0	-83.7	-68.0	-83.7	-25.0	-157.5
		35	-68	-83	-68	-83	-68	-83	-25	-157

[a]The complex codes are a method of cataloguing the various backbones of cyclo(Pro-Gly)$_3$. $C_3$2 and $C_3$3 are the Mg^{2+} and Ca^{2+} complexes in solution. $C_3$5 through $C_3$9 are crystal backbones. All angles are in degrees. Subscripts p and g stand for proline and glycine, respectively. Superscripts refer to the residue number. The $\langle NC^{\alpha}C'$ is the same for all six residues. The referenced backbone parameters are from the literature; those parameters immediately above them are the optimized backbone parameters. NA means the parameter was not listed in the reference.

ψ_p^3	ψ_g^4	ψ_p^5	ψ_g^6	ω_p^1	ω_g^2	ω_p^3	ω_g^4	ω_p^5	ω_g^6	$\langle NC^{\alpha}C'$
150.2	-151.1	150.2	-151.1	180.0	180.0	180.0	180.0	180.0	180.0	109.5
150	-150	150	-150	180	180	180	180	180	180	NA
138.8	-173.1	151.3	175.8	-174.0	-173.0	179.0	-179.0	-172.0	-177.7	106.4
140	-173	149	172	-174	-173	179	-179	-172	-176	NA
138.7	-176.1	140.8	173.0	-171.0	-175.0	-178.0	-179.0	-178.0	178.0	105.6
141	-175	139	174	-171	-175	-178	-179	-178	177	NA
130.0	172.2	130.0	172.2	180.0	180.0	180.0	180.0	180.0	180.0	109.2
130	172	130	172	180	180	180	180	180	180	NA
144	-178.1	144	-178.1	-175.0	-177.0	-175.0	-177.0	-175.0	-177.0	105.5
144	-180	144	-180	-175	-177	-175	-177	-175	-177	110p, 111g
125.0	-157.5	-25.0	-157.5	177.0	-170.0	177.0	-170.0	177.0	-170.0	110.6
-25	-157	-25	-157	177	-170	177	-170	177	-170	113p, 111g

Table 15. Parameters for the proline rings of cyclo(Pro-Gly)$_3$ cation complexes

Code	ϕ	χ^1	χ^2	χ^3	χ^4	τ^R_1	τ^R_2	τ^R_3	τ^R_4	τ^R_5	τ^R_6	NC^δ
$C_3$2	-68.2	48.1	-60.0	47.0	-17.9	99.1	97.4	96.6	98.9	110.6	124.2	1.50
		32.9	-40.0	31.5	-11.4	103.1	102.9	102.5	103.5	112.0	124.3	1.48
All 3		25.8	-20.0(-)	6.6	10.3	104.2	105.5	106.0	105.9	111.7	122.8	1.47
Prolines		6.6	-20.0(+)	25.8	-22.9	105.4	106.5	105.0	104.5	112.4	122.9	1.46
		15.7	0.0	-15.8	27.3	104.5	106.5	106.2	104.9	110.5	122.6	1.46
		-2.6	20.0	-29.8	30.0	104.8	106.3	104.7	103.7	111.4	124.4	1.46
		-30.0	40.0	-34.5	16.5	103.6	103.2	102.1	103.1	111.9	123.6	1.47
		-47.4	60.0	-47.9	19.5	99.4	97.5	96.3	98.8	110.4	123.2	1.50
$C_3$3	-42.1	42.4	-60.0	53.5	-28.7	101.1	97.8	95.2	97.8	109.1	120.1	1.51
		24.6	-40.0	40.1	-26.2	104.3	103.3	101.3	102.0	111.0	122.8	1.47
All 3		2.6	-20.0	29.7	-30.0	104.6	106.4	104.8	103.0	111.7	125.2	1.46
Prolines		-17.8	0.0	18.1	-31.4	103.2	106.4	106.2	103.7	110.7	125.5	1.45
		-30.4	20.0	-1.6	-18.9	102.4	105.1	106.1	105.0	111.3	125.0	1.46
		-34.7	40.0	-29.7	8.2	102.7	102.8	102.7	103.6	112.1	124.7	1.48
		-48.4	60.0	-46.4	17.1	98.6	97.5	96.8	98.7	111.2	125.7	1.51

Table 15. Continued

Code	ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_5^R	τ_6^R	NC^δ
$C_3 5$	-60.9	47.0	-60.0	48.4	-20.2	99.6	97.5	96.3	98.8	110.3	123.3	1.51
Pro(1)		30.9	-40.0	33.6	-15.1	103.5	103.0	102.2	103.3	111.9	123.8	1.48
$C_3 6$		15.2	-28.0[a]	30.0	-21.7	104.8	105.6	104.1	103.7	112.4	124.5	1.46
Pro(5)		3.2	-20.0	29.8	-28.9	105.0	106.4	104.7	103.5	111.5	123.7	1.46
		-14.5	0.0	14.6	-25.3	104.6	106.8	106.2	105.2	110.9	121.7	1.46
		-23.4	20.0(-)	-9.0	-6.2	104.5	105.9	106.0	106.0	112.2	122.5	1.47
		-4.3	20.0(+)	-28.1	27.0	105.3	106.3	104.8	103.9	111.5	123.4	1.46
		-32.4	40.0[b]	-32.0	12.2	103.2	103.0	102.4	103.5	112.1	124.2	1.48
		-48.4	60.0	-46.7	17.5	99.0	97.4	96.6	98.9	110.8	124.4	1.51
$C_3 5$	-63.3	47.3	-60.0	48.0	-19.7	99.5	97.4	96.3	98.8	110.3	123.3	1.51
(Pro(3)		31.2	-40.0	33.3	-14.5	103.5	103.0	102.3	103.3	111.9	123.9	1.48
		3.5	-20.0	28.9	-28.4	105.0	106.4	104.7	103.6	111.5	123.6	1.46

[a] X-ray value of χ^2 for $C_3 5$ used only for $C_3 5$.

[b] X-ray value of χ^2 for $C_3 6$ and part of the search for $C_3 5$.

All angles are in degrees; the NC^δ lengths are in Å.

Table 15. Continued

Code	ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_5^R	τ_6^R	NC[δ]
		1.3	-18.0[c]	27.9	-28.9	103.7	106.3	106.3	105.1	111.3	123.3	1.46
		-14.0	0.0	14.1	-24.3	104.8	106.9	106.2	105.4	111.0	121.6	1.46
		-21.7	20.0(-)	-10.6	-3.3	104.8	106.0	106.0	106.0	112.5	122.7	1.47
		-4.0	20.0(+)	-28.4	27.6	105.2	106.3	104.8	103.8	111.5	123.5	1.46
		-32.1	40.0	-32.3	12.8	103.3	103.0	102.4	103.4	112.0	124.0	1.48
		-48.2	60.0	-46.9	17.8	99.0	97.4	96.6	98.9	110.7	124.2	1.51
$C_3$5	-56.8	46.0	-60.0	49.6	-22.2	100.0	97.6	96.0	98.6	110.0	122.2	1.51
Pro(5)		29.1	-40.0	35.5	-18.3	103.8	103.2	102.0	103.0	111.7	123.5	1.48
		2.4	-20.0	30.0	-30.4	104.7	106.3	104.7	103.1	111.5	124.6	1.46
		-16.8	0.0	16.9	-29.2	104.0	106.6	106.1	104.5	110.4	122.8	1.46
		-21.5	7.0[c](-)	10.3	-25.4	103.7	106.3	106.3	105.1	110.5	122.5	1.46
		3.2	7.0[c](+)	-14.6	17.6	106.0	107.0	106.4	105.8	112.0	121.9	1.46
		-27.7	20.0(-)	-4.6	-13.8	103.7	105.4	106.0	105.7	111.4	122.8	1.47
		-7.4	20.0(+)	-25.0	21.6	105.6	106.4	105.2	104.7	112.0	122.9	1.46
		-33.7	40.0	-30.7	10.0	102.9	102.8	102.3	103.6	112.1	124.5	1.48
		-48.5	60.0	-46.4	17.0	98.8	97.4	96.8	98.8	111.0	125.2	1.51

[c]X-ray value of χ^2 for that given code.

Table 15. Continued

Code	ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_5^R	τ_6^R	NC[δ]
$C_3$6	-55.6	45.9	-60.0	49.6	-22.3	100.1	97.5	96.0	98.6	110.0	122.2	1.51
Pro(1)		29.1	-40.0	35.4	-18.2	103.8	103.1	102.0	103.0	111.7	123.6	1.48
		2.4	-20.0	30.0	-30.4	104.7	106.3	104.7	103.1	111.5	124.6	1.46
		-16.7	0.0	16.9	-29.2	104.0	106.6	106.0	104.6	110.4	122.8	1.46
		-27.7	20.0(-)	-4.5	-13.8	103.6	105.4	106.0	105.7	111.4	122.7	1.47
		-7.2	20.0(+)	-25.3	22.0	105.6	106.3	105.1	104.6	112.0	123.0	1.46
		-32.2	38.0[c]	-29.0	9.1	103.2	103.3	103.0	103.9	112.2	124.4	1.48
		-33.6	40.0	-30.7	10.0	102.9	102.9	102.6	103.6	112.1	124.5	1.48
		-48.6	60.0	-46.2	16.8	98.7	97.4	96.8	98.8	111.0	125.2	1.51
$C_3$6	-61.5	46.9	-60.0	48.5	-20.4	99.7	97.5	96.2	98.8	110.2	122.9	1.51
Pro(3)		30.5	-40.0	34.0	-15.7	103.6	103.1	102.2	103.2	111.9	123.8	1.48
		2.9	-20.0	29.4	-29.4	104.9	106.3	104.7	103.4	111.4	123.9	1.46
		-15.1	0.0	15.2	-26.3	104.5	106.8	106.1	105.1	110.7	121.9	1.46
		-24.6	20.0(-)	-7.7	-8.4	104.3	105.8	106.0	106.0	112.0	122.5	1.47
		-4.8	20.0(+)	-27.6	26.1	105.3	106.3	104.9	104.0	111.6	123.3	1.46
		-26.0	32.0[c]	-25.6	9.7	104.2	104.6	104.1	104.5	112.5	124.0	1.47
		-32.7	40.0	-31.7	11.7	103.2	102.9	102.5	103.5	112.0	124.2	1.48
		-48.4	60.0	-46.6	17.4	98.9	97.4	96.6	98.9	110.8	124.5	1.51

Table 15. Continued

Code	ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_5^R	τ_6^R	NC^δ
$C_3$8	-63.4	47.6	-60.0	47.7	-19.0	99.4	97.4	96.4	98.8	110.4	123.6	1.51
		32.0	-40.0	32.5	-13.1	103.3	103.0	102.4	103.4	112.0	124.1	1.48
All 3		14.1	-27.0[c]	29.5	-21.9	105.0	105.6	104.3	103.8	112.4	124.2	1.46
Prolines		19.8	-10.0(-)	-3.6	17.1	104.8	106.3	106.6	106.0	111.3	122.0	1.47
		-5.0	-10.0(+)	21.3	-25.9	105.2	107.0	105.7	104.8	111.2	121.9	1.46
		14.2	0.0	-14.3	24.6	105.0	106.0	106.3	105.3	110.8	122.1	1.46
		-16.7	10.0(-)	0.5	-11.8	105.1	106.8	106.5	106.3	112.1	121.6	1.47
		6.7	10.0(+)	-23.0	28.9	105.0	106.6	105.7	104.3	110.8	122.8	1.46
		-3.2	20.0	-29.2	28.9	105.0	106.3	104.7	103.5	111.4	123.9	1.46
		-31.3	40.0	-33.2	14.2	103.4	103.1	102.3	103.3	112.0	123.9	1.48
		-48.0	60.0	-47.2	18.3	99.1	97.4	96.5	98.9	110.6	123.9	1.51
$C_3$9	-68.0	40.7	-50.0	39.4	-14.6	101.4	100.4	99.8	101.6	111.4	124.4	1.49
		33.0	-40.0	31.4	-11.2	103.1	102.9	102.5	103.5	112.1	124.3	1.48
All 3		27.2	-30.0	21.2	-4.3	104.1	104.7	104.7	105.0	112.5	123.9	1.47
Prolines		15.8	0.0	-16.0	27.5	104.4	106.5	106.2	104.8	110.5	122.7	1.46
		-28.9	39.0[c]	-34.0	16.7	103.8	103.4	102.3	103.2	112.0	123.7	1.48
		-39.1	50.0	-41.2	17.7	101.9	100.6	99.5	101.4	111.1	123.3	1.49
		-47.3	60.0	-47.9	19.6	99.4	97.6	96.3	98.8	110.4	123.1	1.51

Table 16. Parameters for the cation complexes of cyclo(Pro-Gly)$_3$ derived from the work of Hori et al. (36)[a]

Backbones

Complex Code[b]	Closed/REF	ϕ_p	ψ_p	ω_p	ϕ_g	ψ_g	ω_g	$\angle NC^{\alpha}C'$
$C_3$10	closed	-62.0	-176.0	168.0	54.6	-153.3	172.0	107.4
	REF	-62	-176	168	54	-154	172	NA
$C_3$11	closed	-51.0	-37.0	175.0	-77.0	-165.6	-177.0	114.6
	REF	-51	-37	175	-78	-166	-177	NA
$C_3$12	closed	-42.0	143.0	-176.0	62.9	174.9	-171.0	110.0
	REF	-42	143	-176	63	175	-171	NA
$C_3$13	closed	-53.5	-33.6	175.0	-74.4	-167.4	180.0	113.8
	REF	-52	-35	175	-76	-168	180	NA

Proline Rings

Complex Code	BORC/REF	χ^1	χ^2	χ^3	χ^4
$C_3$10	BORC	33.4	-43.0	35.8	-15.6
	REF	30	-43	39	-22
$C_3$11	BORC	-38.1	46.0	-35.6	12.3
	REF	-34	46	-40	20
$C_3$12	BORC	25.6	-41.0	40.7	-26.0
	REF	23	-41	43	-31
$C_3$13	BORC	-38.0	46.0	-35.7	12.4
	REF	-34	46	-40	19

[a]All angles are in degrees. p = proline and g = glycine. Those parameters list after REF are taken directly from reference 36. The closed and BORC parameters are the optimized values used in optical calculations.

[b]The complex codes refer to the four forms minimized from the Ca^{2+} crystal structure by Hori et al. (181); in this reference they were referred to as CHARMM1 WHB(A), CHARMM1 WHB(B), CHARM1 NHB(A), and CHARMM1 NHB(B), respectively.

Table 17. Absorption and CD peaks for further structures of cyclo(Pro-Gly)$_3$

FORM[a]	ϕ	χ^2	λ	ε_{max}	λ_1	$\Delta\varepsilon_1$	λ_2	$\Delta\varepsilon_2$
$C_3$10	-62.0	-43.0	196	5852	186	2.8	208	-2.4
$C_3$11	-51.0	46.0	196	5549	190	14.0	204	-16.5
$C_3$12	-42.0	-41.0	198	6864	--	--	204	9.2
$C_3$13	-53.5	46.0	194	5677	188	13.8	204	-17.0
Exp Ca^{2+}	--	--	--	--	193*	-2.9[b]	205	2.9

[a] ϕ and χ^2 are in degrees; λ, λ_1 and λ_2 are in nm. ε and $\Delta\varepsilon$ are in L mol^{-1}cm^{-1}. The bandwidth for all calculations is 6000 cm^{-1}.

[b] This was the last experimental point available.

π-π^* spectra calculated for the cation complexes of cyclo(Pro-Gly)$_3$ are given in Figures 13-19. Figure 13 shows the spectra predicted without the proline side chain; i.e., the backbone parameters were used to produce a cyclo(Gly)$_6$ structure. The experimental CD spectra were obtained from reference 33. No experimental absorption spectra were available (193). The other figures, 14-19, include the proline side chain. These figures reiterate the sensitivity of the dipole interaction model to changes in the side chain conformation. Some ring structures give a reasonably good overall fit to the π-π^* spectra for these complexes.

D. Discussion

1) Mg^{2+} complex of cyclo(Pro-Gly)$_3$

Predicted π-π^* absorption and CD spectra for the Mg^{2+} complexes of cyclo(Pro-Gly)$_3$ can be seen in Figures 13-16. Both absorption and CD are sensitive to the proline ring structure with the CD showing the greatest variations. The most dramatic variations are seen in $C_3 2$. When the proline ring is replaced with a glycine H, the predicted CD is extremely weak even though absorption is strong (Figure 13). Once the proline ring is added, strong CD appear (Figure 14); the absorption spectra broaden and indicate a shoulder around 195 nm to the intense band around 202-205 nm except at the extreme values of χ^2 where only the 195 nm band is present. The CD spectra are extremely dependent on the choice of χ^2. For $\chi^2 \geq 0°$, the CD spectra produced show only negative bands (these bands do not resemble experiment at all). For $\chi^2 < 0°$, a positive band appears. For $\chi^2 = -10°$ and $-20°$ a bifurcation results from BORC. Only those BORC solutions

Figure 13. π-π^* absorption and CD spectra for cyclo(Pro-Gly)$_3$ cation complexes treated as cyclo(Gly)$_6$

Units for the scales are as follows: nm for λ, $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled by the backbone code. Those for the Mg^{2+} complexes are listed on the left ($C_3$2, $C_3$5, and $C_3$6). The Ca^{2+} complexes are listed on the right ($C_3$3, $C_3$8, and $C_3$9). The scale listed for $C_3$2 is the same for all other forms. Bandwidths for each form are as follows: $C_3$2, $C_3$5, $C_3$6, and $C_3$9 have Γ = 4000 cm^{-1}; $C_3$3 and $C_3$8 have a bandwidth of 6000 cm^{-1}.

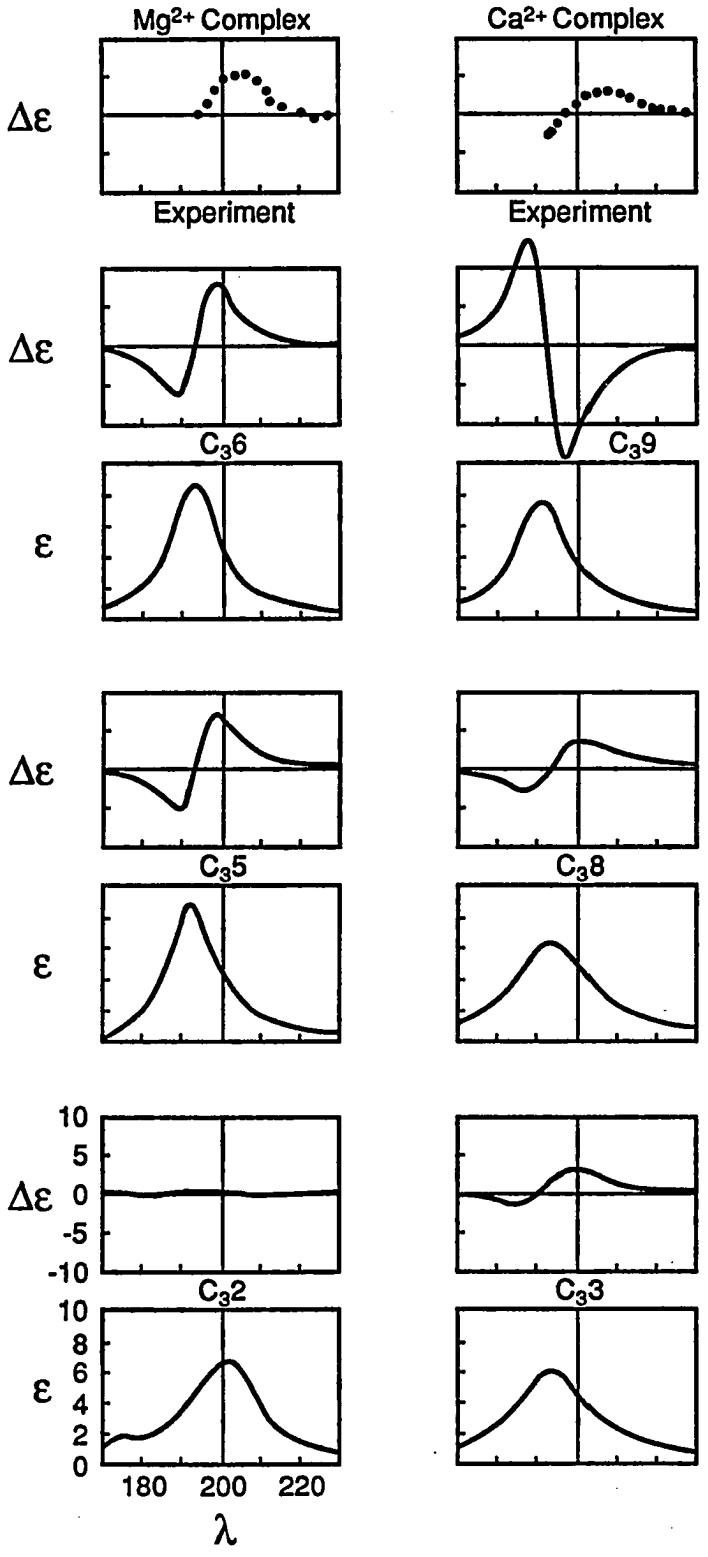

Mg2+ Complex
Ca2+ Complex
Δε
Experiment
Experiment
Δε
C36
C39
ε
Δε
C35
C38
ε
10
5
Δε 0
-5
-10
C32
C33
10
8
6
ε 4
2
0
180
200
220
λ

Figure 14. π-π^* absorption and CD spectra for cyclo(Pro-Gly)$_3$ Mg^{2+} complex backbone C_32

Units for the scales are as follows: nm for λ, $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled according to χ^2 value. For the bifurcation at $\chi^2 = -20$, (-) means the solution for $\chi_0^1 = -\chi^2$, and (+) means the solution for $\chi_0^1 = +\chi^2$. All spectra are to the scale listed for the lower left pair. The bandwidth used for all calculated spectra was 4000 cm^{-1}.

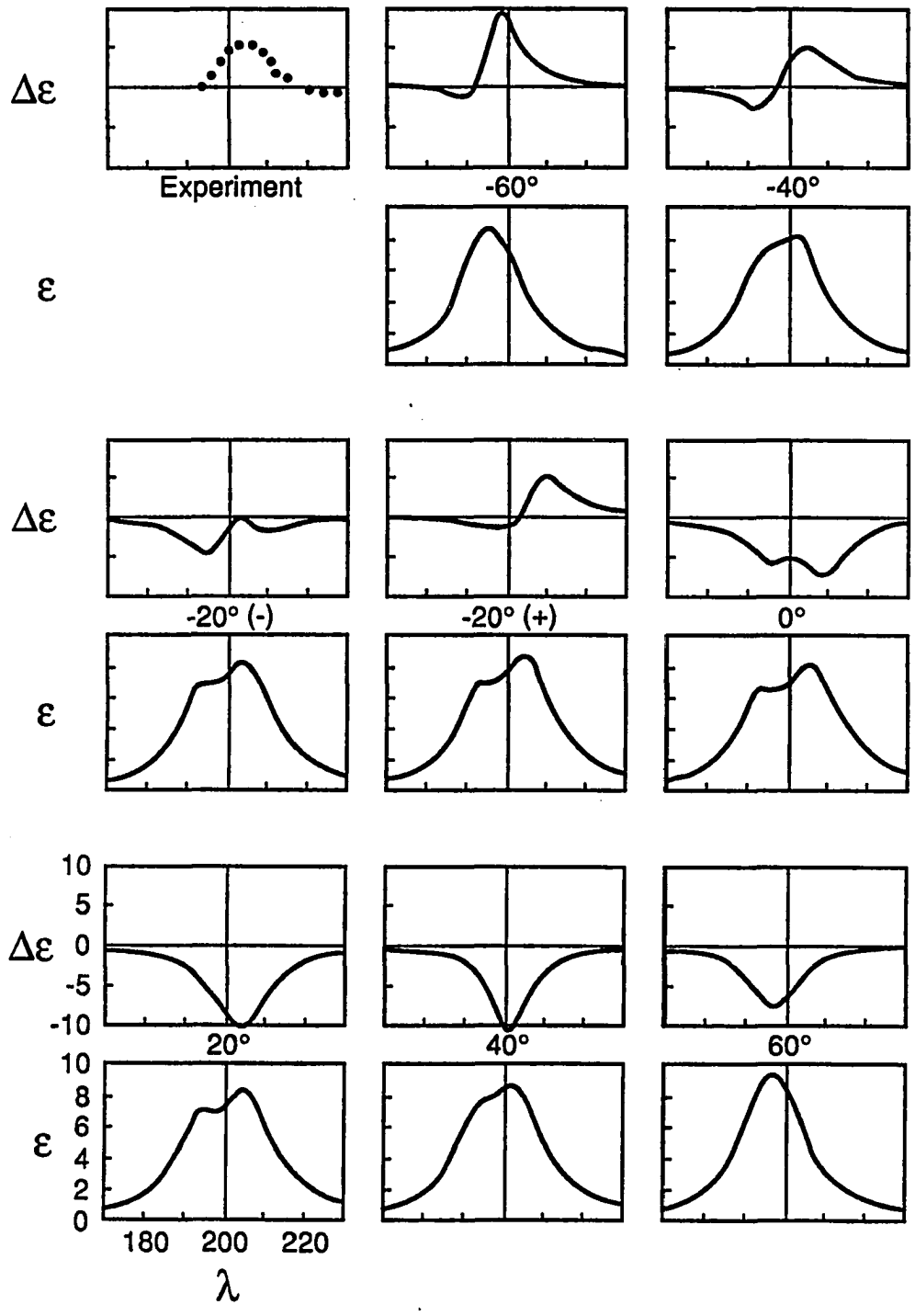

Δε
Experiment
-60°
-40°
ε
Δε
-20° (-)
-20° (+)
0°
ε
10
5
Δε 0
-5
-10
20°
40°
60°
10
8
ε 6
4
2
0
180
200
220
λ

Figure 15. $\pi-\pi^*$ absorption and CD spectra for cyclo(Pro-Gly)$_3$ Mg^{2+} complex backbone $C_3 5$

Units for the scales are as follows: nm for λ, $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled according to χ^2 value. The bifurcations at 20° and 7° are designated (+) and (−) for solutions $\chi_0^1 = +\chi^2$ and $\chi_0' = -\chi^2$ respectively. The x-ray values of χ^2 are -20°, -18°, and 7°. All spectra are to the scale listed for the lower left pair. The bandwidth is 4000 cm^{-1}.

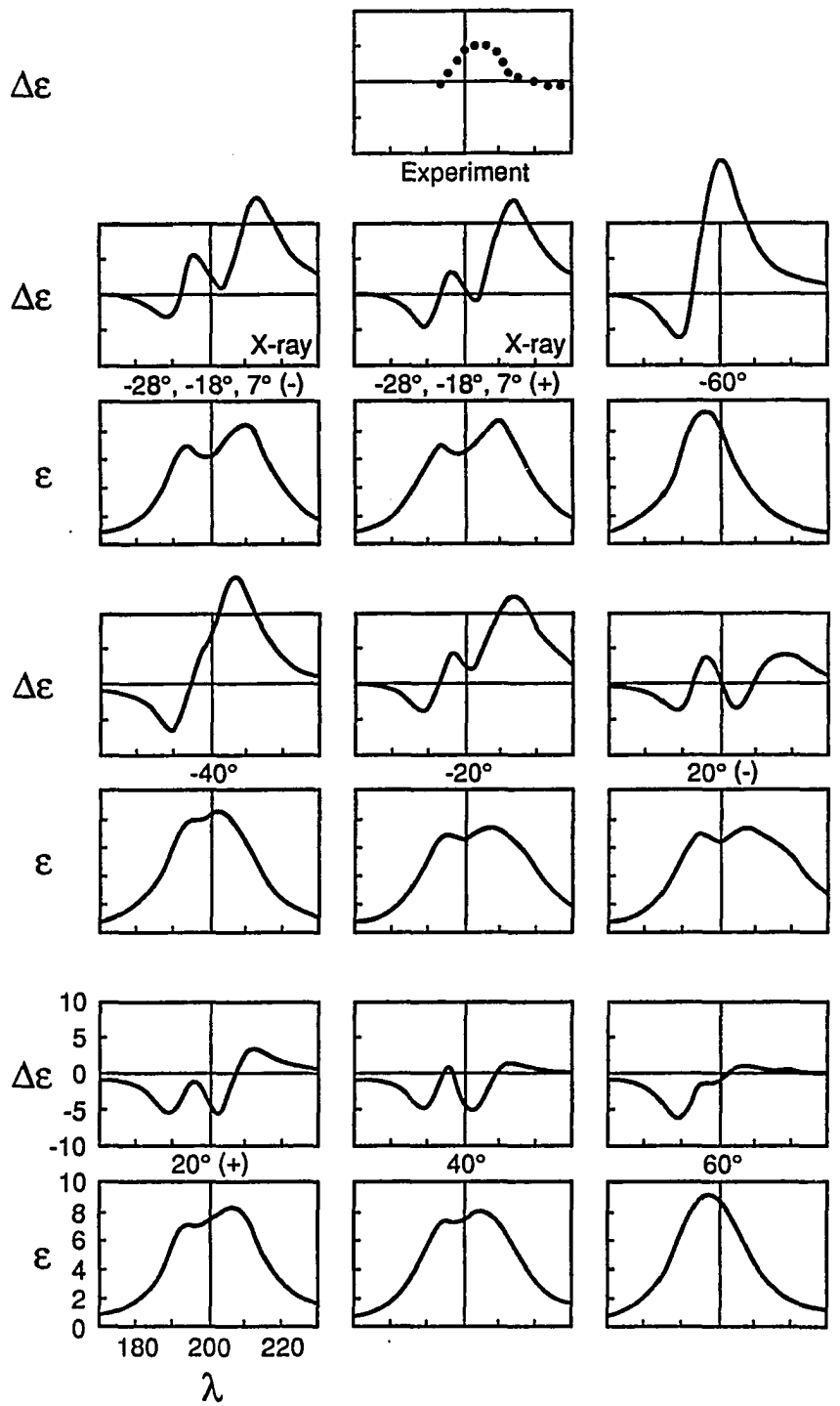

Δε
Experiment
Δε
X-ray
X-ray
-28°, -18°, 7° (-)
-28°, -18°, 7° (+)
-60°
ε
Δε
-40°
-20°
20° (-)
ε
Δε
10
5
0
-5
-10
20° (+)
40°
60°
ε
10
8
6
4
2
0
180
200
220
λ

Figure 16. π-π^* absorption and CD spectra for cyclo(Pro-Gly)$_3$ Mg^{2+} complex backbone $C_3 6$

Units for the scales are as follows: nm for λ, $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled according to χ^2 value. The bifurcated solution for $\chi^2 = -20°$ is designated by (-) for $\chi_0^1 = -\chi^2$ and (+) for $\chi_0^1 = +\chi^2$. The x-ray values of χ^2 are 38°, 32°, and 40°. All spectra are to the scale listed for the lower left pair. The bandwidth is 4000 cm^{-1}.

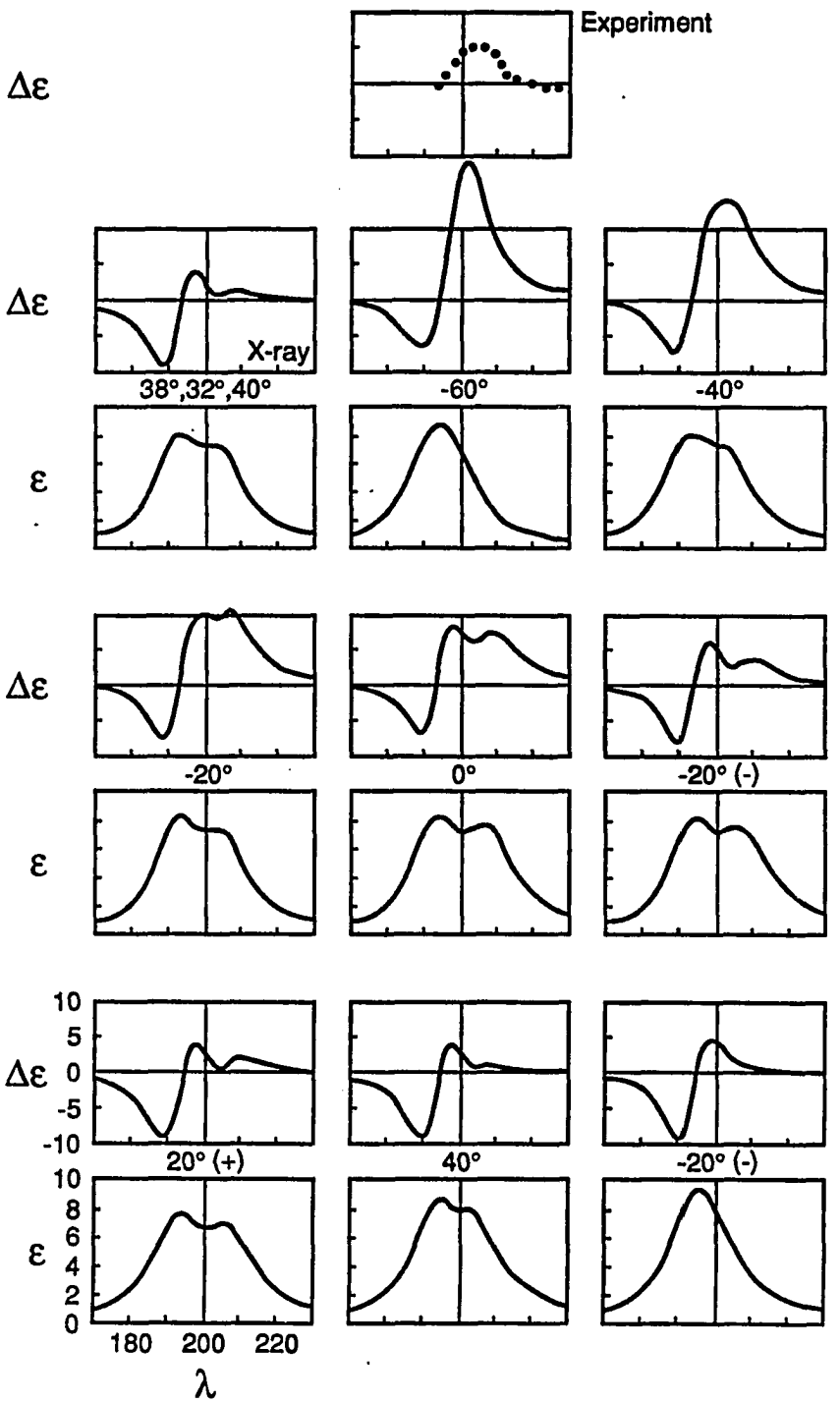

Experiment
Δε
X-ray
38°,32°,40°
-60°
-40°
ε
-20°
0°
-20° (-)
20° (+)
40°
10
5
0
-5
-10
8
6
4
2
180
200
220
λ

Figure 17. π-π^* absorption and CD spectra for cyclo(Pro-Gly)$_3$ Ca^{2+} complex backbone C_3^3

Units for the scales are as follows: nm for λ, $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled according to X^2 value. All spectra are to the scale listed for the lower left pair. The bandwidth is 6000 cm^{-1}.

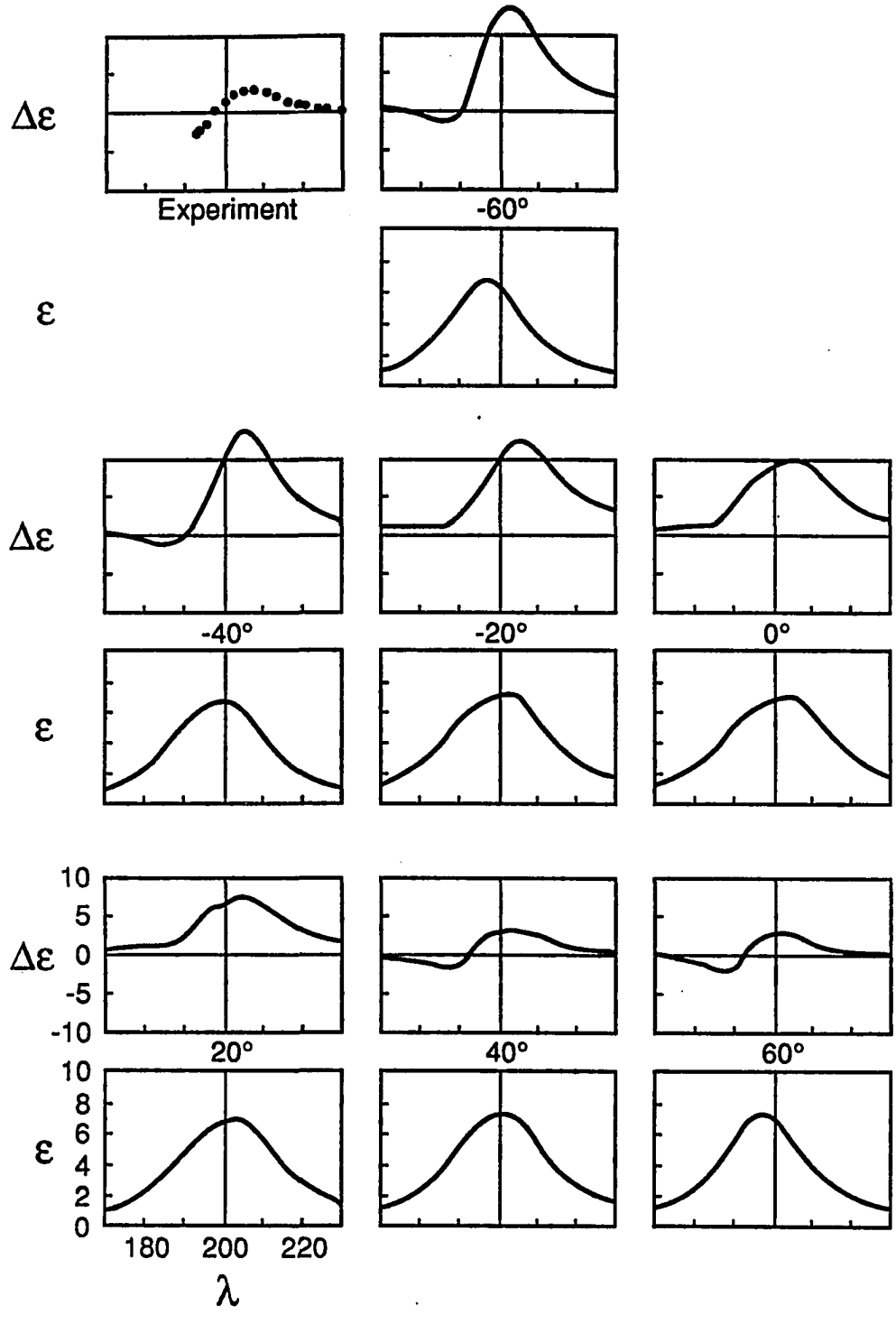

Δε
Experiment
-60°
ε
Δε
-40°
-20°
0°
ε
10
5
Δε 0
-5
-10
20°
40°
60°
10
8
ε 6
4
2
0
180
200
220
λ

Figure 18. $\pi-\pi^*$ absorption and CD spectra for cyclo(Pro-Gly)$_3$ Ca^{2+} complex backbone $C_3 8$

Units for the scales are as follows: nm for λ, $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled according to χ^2 value. The X-ray value of χ^2 are -27°. All spectra are to the scale listed for the lower left pair. The bandwidth is 6000 cm^{-1}.

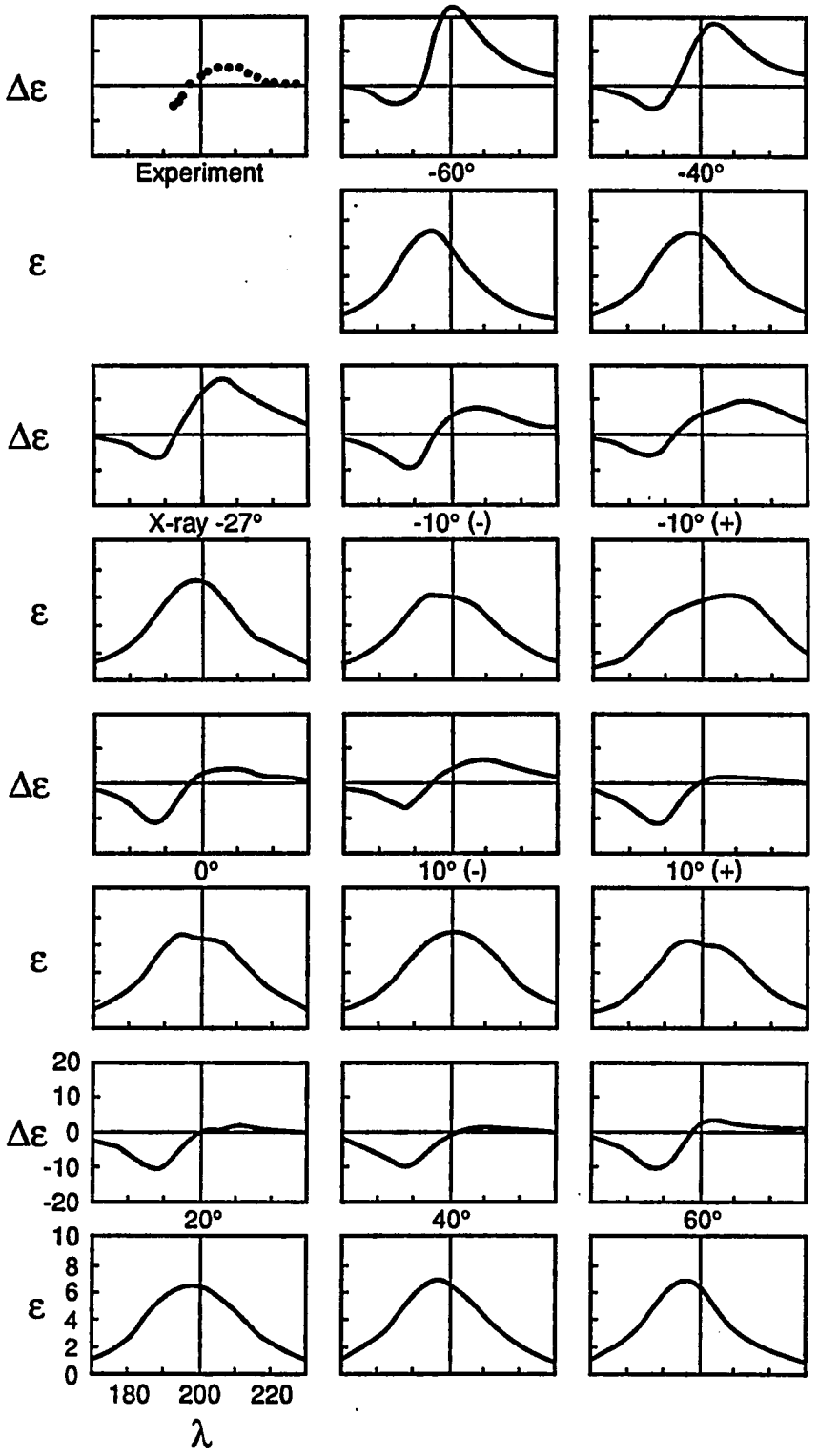
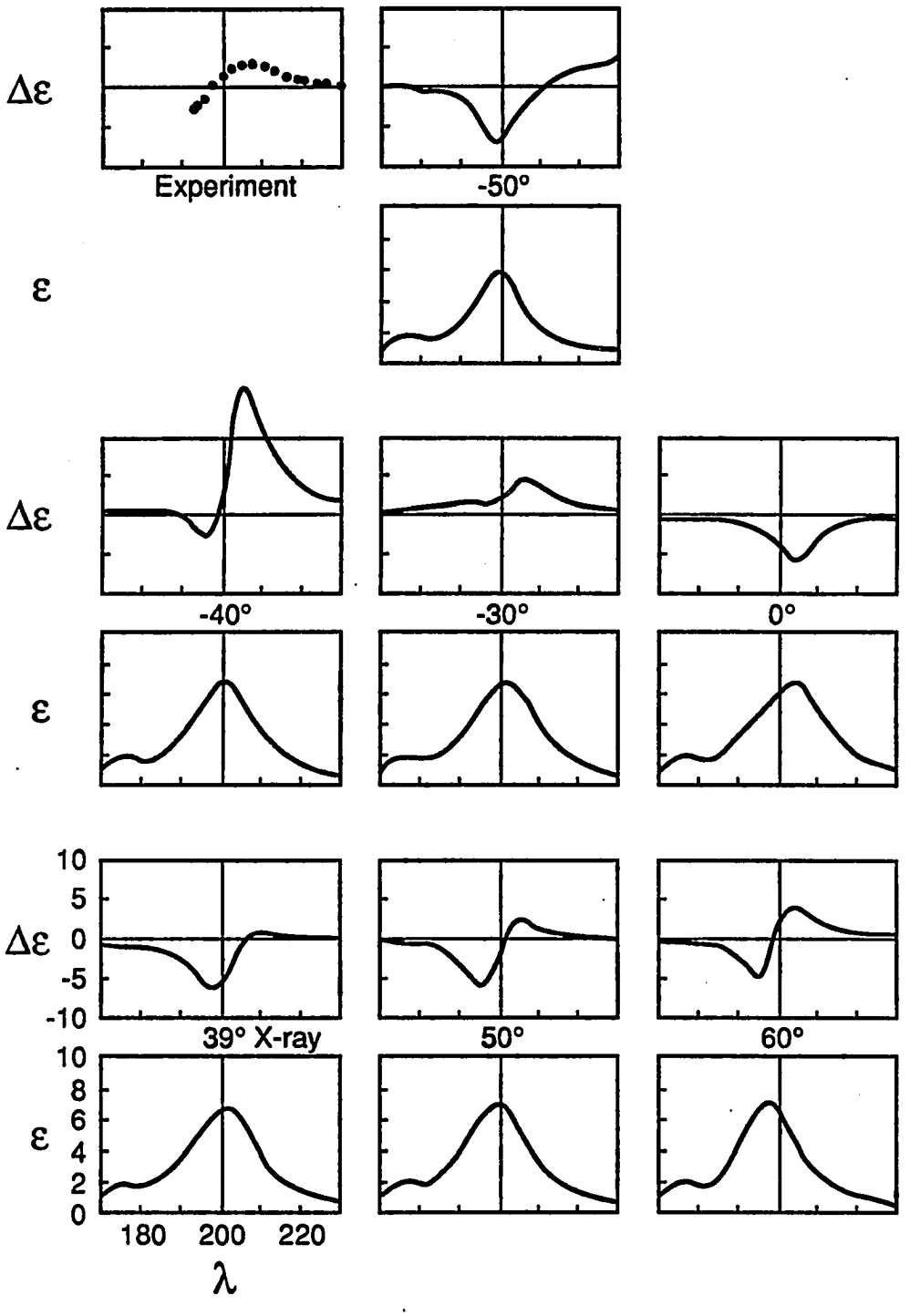

Δε
ε
Experiment
-60°
-40°
X-ray -27°
-10° (-)
-10° (+)
0°
10° (-)
10° (+)
20°
40°
60°
20
10
0
-10
-20
10
8
6
4
2
0
180
200
220
λ

Figure 19. π-π^* absorption and CD spectra for cyclo(Pro-Gly)$_3$ Ca^{2+} complex backbone C_3^9

Units for the scales are as follows: nm for λ, $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled according to χ^2 value. The x-ray value of χ^2 is 39. All spectra are to the scale listed for the lower left pair. The bandwidth is 4000 cm^{-1}.


Δε
Experiment
-50°
ε
Δε
-40°
-30°
0°
ε
10
5
Δε 0
-5
-10
39° X-ray
50°
60°
10
8
ε 6
4
2
0
180
200
220
λ

obtained from $\chi_0^1 = +\chi^2$ produced the positive band. When $\chi_0^1 = -\chi^2$ two CD bands were distinguishable, but they were both negative. For $\chi^2 = -30°$ through -60°, the intensity increases and the positive band is blue shifted so that it most resembles experiment at $\chi^2 = -40°$.

The X-ray structures of the Mg^{2+} complex of cyclo(Pro-Gly)$_3$ ($C_3$5 and $C_3$6) did not produce CD spectra that resembled experiment as well as the solution form ($C_3$2) (Figures 13, 15, and 16). The absorption spectra of $C_3$5 and $C_3$6 resemble those calculated for the solution form, but the CD spectra do not. The cyclo(Gly)$_6$ approximations for both forms produced CD spectra that are a little more intense and blue than experiment. Once the proline side chain is added more bands appear. The spectra that come closest to experiment in the number of bands, sign, and location are again for $\chi^2 < 0°$ especially $-60° \leq \chi^2 < -30°$. This supports predictions for the $C_3$2 form, but the CD intensities predicted for the X-ray structures are very strong. These X-ray structures do not represent solution conformations well. Most of the X-ray forms produce CD spectra with more bands than experiment and generally the longest band is red of experiment. The X-ray structure backbones, besides being asymmetric, tend to have greater torsion angles than solution and the ω torsion angles are not planar. Perhaps the loss of symmetry in the solid is partly responsible for the changes in predicted CD spectra.

2) Ca^{2+} complex of cyclo(Pro-Gly)$_3$

The Ca^{2+} forms of cyclo(Pro-Gly)$_3$ include $C_3$3, $C_3$8, and $C_3$9. The solution form, $C_3$3, does not show as dramatic a change in the CD as the Mg^{2+} complex when the cyclo(Gly)$_6$ approximation is applied. For $C_3$3 the

resulting CD of the cyclo$(Gly)_6$ approximation is merely blue of experiment (Figure 13). The CD is again increased when the proline side chain is added so that most of the predicted CD are more intense than experiment (Figure 17). It is not until χ^2 reaches 40° that the predicted CD show a reduced intensity that resembles experiment. One of the X-ray structures, $C_3 8$, shows a very similar spectral trend to those of $C_3 3$ except that they are weaker in the positive band and stronger in the negative one (Figure 18). Those most resembling experiment occur near χ^2 of $\pm$ 10°. The other part of the X-ray structure, $C_3 9$, also produces good CD predictions for some values of χ^2 (Figure 19). Predicted CD for $C_3 9$ is the most sensitive to side chain structures of the three forms. First, the cyclo$(Gly)_6$ spectrum shows no positive band around 206 nm like experiment. $C_3 9$ supports χ^2 values of 50° to 60° by best resembling experiment in that region. The negative region of χ^2 is not representative because of strong intensitives that become long wavelength normal modes by $\chi^2 = -50°$ and singularities by $\chi^2 = -60°$. Both $C_3 3$ and $C_3 9$ Ca^{2+} forms support positive values of χ^2. This is reinforced by the χ^2 value of 39°, the X-ray structure value. $C_3 8$ supports small values of χ^2. The backbones of $C_3 3$ and $C_3 8$ are similar, but that of $C_3 9$ is substantially different. Because all three did give CD resembling experiment for at least one value of χ^2, they may all be considered representative, but the backbone of $C_3 9$ is the least representative because only extreme puckering greater than what is normally found for proline produced CD resembling experiment for $C_3 9$.

E. Conclusions

The principal finding of this study is that approximately correct π-π^* CD spectra of the Mg^{2+} and the Ca^{2+} complexes of cyclo(Pro-Gly)$_3$ are predicted by the dipole interaction model when the proline side chain is included on certain predetermined backbone structures. The experimental spectra for the two complexes, although very similar, have been considered distinguishable (33); these differences have been attributed to the small differences in backbones in the past (33). The Mg^{2+} complex solution form ($C_3$2) has ϕ_p (the proline ϕ) nearer to the mode value in ϕ for proline (-68°) than does the Ca^{2+} complex solution form ($C_3$3) (see Chapter III); ψ values are the same for the two complexes within 30° (Table 18), and all ω values are 180° (trans). The largest differences between $C_3$2 and $C_3$3 occur between ϕ_p and ψ_g (glycine ψ). Only ϕ_p will affect the proline ring structures. The CD spectra are more sensitive to the proline side chain conformation for the Mg^{2+} $C_3$2 form. ϕ is nearer the mode value in ϕ_p than for the Ca^{2+} $C_3$3 form.

Just as the proposed solution conformations have gross similarities to each other, so the crystal structures of the complexes have similarities among themselves. One crystal backbone even produced CD that resembled experiment. First, the two Mg^{2+} crystal backbones ($C_3$5 and $C_3$6) resemble each other in ϕ and ψ when rounded to the nearest 20° (Table 18). The similarities in ϕ_p account for the similar predicted proline ring conformations for these two forms. The loss of symmetry for $C_3$5 and $C_3$6 makes the resulting CD hard to interpret because of the appearance of multiple bands. For χ^2 = -40° to -60°, however, only two bands appear and

Table 18. Backbone comparisons for cation complexes of cyclo(Pro-Gly)$_3$[a]

Cation	Form	ϕ_p	ψ_p	ω_p	ϕ_g	ψ_g	ω_g	Good similarity to CD
Mg^{2+}	$C_3$2	-70	150	180	70	-150	180	yes
	$C_3$5	-60	140_2	-170_2	80	-170_2	-170	no
			150	180		180	180_2	
	$C_3$6	-60	140	-170	80_2	170_2	180	no
				180_2	90	180		
Ca^{2+}	$C_3$3	-40	130	180	80	170	180	yes
	$C_3$8	-60	140	180	80	180	180	yes
	$C_3$9	-70	-25	180	-80	-160	-170	no

[a]Each backbone torsion angle has been rounded to the nearest 10°. When a value is subscripted by 2 this means that value occurs twice (this occurs only in the two asymmetric backbones $C_3$5 and $C_3$6). The subscripts p and g represent proline and glycine, respectively.

resemble experiment in sign and location (the positive band is considerably more intense than experiment). Because predicted spectra only resemble experiment in a gross sense, these two backbones ($C_3$5 and $C_3$6) may not entirely be representative of cation complexes in solution. The Ca^{2+} X-ray structure had two forms that appeared considerably different from each other (Table 18). One of the backbones, $C_3$8, had ϕ and ψ similar to that of $C_3$5 and $C_3$6, but it had two distinct differences: (1) $C_3$8 was truly C_3 symmetric, and (2) all the peptide bonds were approximately trans. The CD spectra predicted for $C_3$8 resembled those observed for both experimental complexes (Figures 14 and 18). Thus, of the crystal structures, the $C_3$8 backbone may be considered a representation of the solution conformation of cyclo(Pro-Gly)$_3$ cation complexes.

The other part of the Ca^{2+} X-ray structure, $C_3$9, has a backbone that is radically different from all other complex backbones. Although it is C_3 symmetric and its ϕ_p resembles that for $C_3$2, ϕ_g and all ψ angles are different (all the carbonyls point in the same direction); the $C_3$9 backbone also includes nonplanar peptide bonds (ω = -170°). When proline is intensely puckered ($\chi^2 \leq -50°$) long wavelength normal modes appear. Long wavelength normal modes usually indicate unreasonably close contacts within the molecule. Therefore, the $C_3$9 backbone may only be considered representative of solution conformations for cation complexes of cyclo(Pro-Gly)$_3$ if $\chi^2 >> 0°$.

The three backbones that produce CD spectra most resembling experiment of either complex are $C_3$2, $C_3$3, and $C_3$8; the predicted spectra could not distinguish between the two experimental spectra. All three of these

backbones have near planar peptide bonds and differ the most in ϕ_p and ψ_g (Table 18). For $C_3$2 the predicted CD spectrum that most looks like experiment is $\chi^2 = -40°$. For $C_3$3 the one CD spectrum that most looks like experiment is $\chi^2 = 40°$, but those CD for $\chi^2 < 0°$ differed only by being more intense so that they may be considered possibly solution representations. For $C_3$8 the CD spectrum that most resembles experiment is $\chi^2 = -10°$. As in $C_3$3 $\chi^2 << 0°$ produced similar but more intense CD. These choices of χ^2 imply that the proline ring has a tendency to be puckered preferably in one direction for these backbones. This result is also seen in the molecular mechanics calculations of Hori et al. (36).

According to Hori's and coworkers' (36) energy minimizations, proline ring puckering was backbone dependent. For those backbones derived from the $C_3$8 form (which had $\chi^2 = -27°$ (35)), χ^2 was always negative and in the vicinity of -40°; those backbones derived from the $C_3$9 form ($\chi^2 = 39°$ (35)) had χ^2 positive (46°) (36). Although these puckering ranges agree with the puckering ranges favored by the CD spectra for the two backbones ($C_3$8 and $C_3$9) done in this study ($\chi^2 = -40°$ and 50° respectively), the backbones of Hori et al. (36), when closed with addition of a BORC ring, give calculated CD spectra which are usually opposite in sign or lacking the short wavelength negative band entirely (Table 17). These backbones do not appear to be representative of the solution conformation of the Ca^{2+} complex of cyclo(Pro-Gly)$_3$. Thus, the only agreement between solution calculations and these energy minimized structures is in the direction and intensity of the proline ring puckering.

Finally, good π-π^* CD predictions occur for cation complexes of cyclo(Pro-Gly)$_3$ when: (1) the backbone is C_3 symmetric, (2) ω is approximately trans (ω = 180°), and (3) proline side chain is included in the calculation. The predominant backbone types found in solution appear to be those which have $\chi^2 < 0°$, but the predicted CD cannot rule out the possible presence of the $C_3 9$ type as long as $\chi^2 \gg 0°$.

VI. CYCLO(PRO-GLY)$_3$ UNCOMPLEXED

A. Introduction

Since cyclo(Pro-Gly)$_3$ is a synthetic peptide that mimics ion binding and transport properties of natural peptides, it changes conformation from the bound to unbound forms (33-36, 190, 194). The crystal structure has been determined for the uncomplexed form (37). Moreover, information on solution backbone conformations has been obtained by NMR, energy minimizations, Raman spectroscopy, and CD studies which have included theoretical CD calculations (1, 33, 38, 194).

The solution studies indicate that the uncomplexed conformation is dependent on solvent; it is C_3 symmetric in nonpolar solvents and asymmetric in polar solvents (33, 194). NMR results indicated that in nonpolar solvents, two C_3 symmetric forms were possible; one with three-cis peptide bonds and one that was all trans with three γ-turns (33). Madison et al. (33) used theoretical CD calculations to rule out the three-cis structure. Elimination of this structure is strongly supported by energy minimizations because the three-cis structure has a much higher potential than the all trans conformer (3.3 (38) and -6.3 (33) kcal/mole respectively). Raman spectra for uncomplexed cyclo(Pro-Gly)$_3$ in $CHCl_3$ was also all trans (194). Further energy minimization of the all trans conformer led Hori et al. (36) to conclude that this conformer was correct for nonpolar solvents.

Asymmetric structures of cyclo(Pro-Gly)$_3$ occur in polar solvents (33, 194). Even the crystal structure, which was crystalized from a polar solvent, was asymmetric (37). Madison et al. (33) used CD theory to

propose one asymmetric conformation to be representative of the uncomplexed molecule in polar solvents; they called this conformation "A". Several asymmetric backbones had been generated via energy minimization; theoretical CD spectra for these structures were in the region of experiment (38). Madison et al. (33) eliminated most as possibilities based on the potential energies and the quality of the theoretical CD predictions.

Since the dipole interaction model has predicted reasonable CD for helices, cyclic peptides, and β-structures (2, 3, 6-11, Chapter IV, and Chapter V), and since it includes the proline side chain, application of this model to the uncomplexed structures of cyclo(Pro-Gly)$_3$ may provide further insight to the proposed solution backbones for this molecule.

B. Structure Generation

Cyclo(Pro-Gly)$_3$ forms were generated using the Ramachandran-Sasisekharan technique (43) described in the methods chapter. The parameters of the cyclo(Pro-Gly)$_3$ uncomplexed backbones are derived from the published minimum energy structures (33, 36, 38) or the X-ray structure (37). Backbone rings were closed as described in the methods chapter. They needed to be closed in order to keep the end of the chain in the proper geometry with as few modifications to the literature ϕ and ψ as possible. The literature parameters needed to be revised to remove the differences between the Ramachandran-Sasisekharan parameters and those in the literature. Often in the literature no $\angle NC^{\alpha}C'$ angle was provided, and/or the other bond lengths or angles differed from Ramachandran-Sasisekharan values. Moreover, in some cases closure was needed because

the rounded data in the literature were not good enough to give accurate closure by the stepwise method of chain generation. Furthermore, each residue used in the current study has the same bond angles and bond lengths regardless of whether the residue is proline or glycine. Crystal structures often distinguish the two and even have variations among residues of the same kind. After backbone ring closure, BORC was applied to the proline rings for the resulting backbone rings for optical studies.

π—π^* absorption and CD spectra were calculated for both the C_3 symmetric structure and multiple asymmetric structures. Half-peak bandwidths varied from Γ = 4000 cm^{-1} to 6000 cm^{-1} depending on the form. As in the previous chapter, the C_3 symmetric backbone form was coded C_3 for C_3 symmetry; whereas the asymmetric forms were all coded A and arbitrarily numbered. $C_3$1 is the solution structure in nonpolar solvents derived from conformation "S" of Madison et al. (33); $C_3$14 and $C_3$15 are the energy minimized structures of Hori et al. (36) (named C3-1 and C3-2, respectively by them). The asymmetric backbones for polar solvents are coded as follows: A1 through A15 for the energy minimized structures numbered 1-15 in Madison's work (38); A16 for the conformation "A" of Madison et al. (33); A17 for the X-ray crystal backbone (37); and A18 for the minimization 1 done by Hori et al. (36). Bandwidths used for these forms are as follows: $C_3$1, 4000 cm^{-1}; $C_3$14 and $C_3$15, 6000 cm^{-1}; A1 through A18, 6000 cm^{-1}. The choice between the two bandwidths was made according to which one gave the better agreement with experiment.

C. Results

The backbone parameters of the closed rings used for the uncomplexed forms of cyclo(Pro-Gly)$_3$ are listed in Table 19. Backbone ring closure produced changes in the ϕ and ψ angles that were at most 14° (although most changes were less than 6°). Those backbones containing the largest changes were A8 (14° change in ψ_g^2); A12 (8° change in ϕ_p^3); and A16 (9° change in ϕ_g^4). Table 20 contains the proline ring parameters determined by BORC. The C_3 symmetric backbones $C_3$1, $C_3$14, and $C_3$15 all had the same proline ring parameters used for all three rings in a given backbone because they all had the same ϕ,χ^2 pair. Some of the asymmetric forms were also similarly treated because all three prolines had ϕ values within 3° of each other; these include A1, A2, A3, A6, A7, A9, A10, A14, and A15. Asymmetric forms A8, A16, and A17 had different ϕ values but the same χ^2 values so that the three proline rings could be slightly different. For the structures derived from Hori's and coworkers' energy minimizations ($C_3$14, $C_3$15, and A18) (36), only the χ^2 values they found upon minimization were used. Note that Hori et al. (36) did derive two other energy minimized structures, but these deviated severely from the original structures (for $C_3$1 changes of 79° in ϕ_g and 86° in ψ_g and for A16 changes of 61° in ψ_p^1, 153° in ψ_p^3, and 115° in ψ_g^6) so that closure by our method could only be achieved with severe deformations (i.e., values << 100°) to the $\angle NC^{\alpha}C'$.

π-π^* spectra calculated for the uncomplexed forms of cyclo(Pro-Gly)$_3$ are given in Figures 20-25 and Table 21. Figures 20 and 21 show the absorption and CD spectra respectively, predicted without the proline side chain, i.e., the backbone parameters were used to produce a cyclo(Gly)$_6$

Table 19. Parameters for backbones of uncomplexed forms of cyclo(Pro-Gly)$_3$

Code[a]	Ref.	1 φp	2 φg	3 φp	4 φg	5 φp	6 φg	1 ψp	2 ψg	3 ψp
A1		-98.4	64.6	-97.5	159.6	-98.4	115.2	155.0	-144.6	93.8
	38	-98	65	-98	160	-98	115	155	-145	95
A2		-98.7	-59.8	-97.6	78.7	-98.0	-167.4	-85.1	-100.2	154.8
	38	-98	-60	-98	79	-98	-167	-85	-100	155
A3		-97.9	145.7	-97.2	87.2	-99.3	169.2	111.7	-113.9	147.1
	38	-98	145	-98	87	-98	169	110	-115	148
A4		-96.3	100.7	-99.8	45.3	-95.4	162.9	170.8	81.7	-179.0
	38	-98	100	-98	44	-98	163	170	80	179
A5		-97.3	161.3	-97.7	110.6	-98.6	-41.6	171.3	-59.4	160.8
	38	-98	160	-98	110	-98	-42	170	-60	160
A6		-66.7	134.8	-68.8	149.6	-67.5	141.3	96.3	-176.4	81.8
	38	-68	135	-68	151	-68	138	95	-175	84
A7		-67.9	-62.1	-68.1	143.6	-68.1	160.0	-60.2	-173.7	93.3
	38	-68	-63	-68	145	-68	158	-60	-173	95
A8		-62.6	102.8	-59.4	-140.9	-66.6	-136.4	96.0	161.0	51.7
	38	-68	105	-68	-140	-68	-134	95	175	55
A9		-67.8	95.1	-67.3	102.3	-69.7	121.7	151.7	-164.4	136.3
	38	-68	95	-68	102	-68	122	150	-165	138

[a]The complex codes are a method of cataloguing the various backbones of cyclo(Pro-Gly)$_3$. A1 through A16 are proposed solution structures in polar solvents. A17[3] is the crystal structure. $C_3$1, $C_3$14, and $C_3$15 are proposed solution structures for nonpolar solvents. All angles are in degrees. Subscripts p and g stand for proline and glycine, respectively. Superscripts refer to the residue number. The ⟨NCαC' is the same for all six residues. The referenced backbone parameters are from the literature; those parameters immediately above them are the optimized backbone parameters. NA means the parameter was not listed in the reference. TPE is total potential energy calculated for the backbones by the given references; they are in kcal/mole; the energies for reference 36 used a different zero point so that only energies among the same reference may be directly compared.

4 ψ g	5 ψ p	6 ψ g	1 ω p	2 ω g	3 ω p	4 ω g	5 ω p	6 ω g	<NCαC'	TPE
-142.8	104.2	-106.1	180.0	180.0	180.0	180.0	180.0	180.0	110.1	
-143	104	-105	180	180	180	180	180	180	NA	-7.1
-152.8	76.2	-116.4	180.0	180.0	180.0	180.0	180.0	180.0	110.2	
-153	77	-116	180	180	180	180	180	180	NA	-1.1
-95.7	-74.9	157.0	110.0	180.0	180.0	180.0	180.0	0.0	110.0	
-95	-74	154	180	180	180	180	180	0	NA	0.2
-102.2	128.4	-61.5	180.0	0.0	180.0	180.0	180.0	0.1	109.6	
-103	129	-68	180	0	180	180	180	0	NA	11.2
73.0	159.5	119.5	180.0	0.0	180.0	0.0	180.0	0.0	109.9	
72	159	118	180	0	180	0	180	0	NA	35.6
-170.2	92.0	-167.0	180.0	180.0	180.0	180.0	180.0	180.0	109.8	
-172	89	-167	180	180	180	180	180	180	NA	-6.5
-164.5	83.6	-152.4	180.0	180.0	180.0	180.0	180.0	180.0	109.7	
-165	82	-152	180	180	180	180	180	180	NA	-3.8
-118.0	-67.6	-80.8	180	180	180	180	180	-179.9	110.6	
-124	-63	-80	180	180	180	180	180	180	NA	1.0
-120.7	-62.1	155.6	180.0	180.0	180.0	180.0	180.0	0.0	109.5	
-120	-60	152	180	180	180	180	180	0	NA	-.3

Table 19. Continued

Code[a]	Ref.	1 ϕ_p	2 ϕ_g	3 ϕ_p	4 ϕ_g	5 ϕ_p	6 ϕ_g	1 ψ_p	2 ψ_g	3 ψ_p
A10		-70.2	-84.3	-69.0	99.5	-68.4	-22.1	151.2	167.8	148.8
	38	-68	-85	-68	98	-68	-21	150	165	146
A11		-66.6	105.9	-67.9	-91.8	-66.8	-142.3	151.4	97.6	-72.2
	38	-68	105	-68	-94	-68	-142	150	95	-73
A12		-64.2	152.4	-59.5	-177.7	-67.1	86.8	-68.0	113.7	169.6
	38	-68	145	-68	-178	-68	89	-75	115	172
A13		-68.6	-154.4	-68.5	-41.4	-67.7	-167.5	-74.7	136.9	168.8
	38	-68	-155	-68	-41	-68	-168	-75	135	169
A14		-41.6	55.0	-41.9	79.8	-41.4	142.8	100.0	-135.4	127.1
	38	-42	53	-42	80	-42	142	100	-135	127
A15		-41.4	85.0	-41.3	64.0	-41.0	-51.7	99.0	-105.3	99.9
	38	-42	85	-42	65	-42	-51	100	-105	99
A16		-66.6	94.9	-61.6	54	-64.9	135.0	152.1	-147.8	164.5
	33	16	-68	95	-68	45	-68	136	150	-155 159
A17		-47.8	92.6	-81.0	94.5	-68.2	-98.1	144.4	-126.0	-4.1
	37	-51	95	-79	92	-68	-102	151	-128	-5
A18		-53.0	105.1	-64.9	94.5	-66.6	-63.5	130.0	-112.0	-26.8
	36	-53	107	-61	93	-72	-69	130	-111	-25
$C_3$1		-81.2	169.6	-81.2	169.6	-81.2	169.6	71.1	-162.9	71.1
	33	-80	170	-80	170	-80	170	70	-162	70
$C_3$14		-69.1	151.9	-69.1	151.9	-69.1	151.9	82.9	-168.7	82.9
	36	-69	154	-69	154	-69	154	83	-167	83
$C_3$15		-69.1	153.7	-69.1	153.7	-69.1	153.7	79.0	-166.2	79.0
	36	-69	155	-69	155	-69	155	79	-165	79

4 ψ g	5 ψ p	6 ψ g	1 ω p	2 ω g	3 ω p	4 ω g	5 ω p	6 ω g	<NCαC'	TPE
-148.2	103.5	125.9	180.0	0.0	180.0	180.0	180.0	0.0	109.3	
-147	105	125	180	0	180	180	180	0	NA	1.7
-87.4	75.9	-69.6	180.0	0.0	180.0	180.0	180.0	0.0	109.4	
-88	77	-70	180	0	180	180	180	0	NA	6.9
-76.8	141.4	97.8	180.0	0.0	180.0	0.0	180.0	0.0	100.3	
-81	142	101	180	0	180	0	180	0	NA	2.9
159.0	-44.3	144.7	180.0	0.0	180.0	0.0	180.0	0.0	113.0	
159	-45	144	180	0	180	0	180	0	NA	10.5
144.4	89.8	-153.2	180.0	180.0	180.0	180.0	180.0	180.0	109.7	
144	90	-154	180	180	180	180	180	180	NA	-5.7
-156.3	-59.7	163.9	180.0	180.0	180.0	180.0	180.0	180.0	109.9	
-157	-60	163	180	180	180	180	180	180	NA	-4.6
-111.8	-78.0	161.2	180.0	180.0	180.0	180.0	180.0	0.0	111.5	
-116	-73	155	180	180	180	180	180	0	NA	0.4
170.1	-11.0	-161.6	179.0	-176.0	-171.0	-8.0	175.0	-177.0	114.4	
164	-12	-163	179	-176	-171	-8	175	-177	112.7	
149.3	-14.6	154.1	-171.0	-175.0	-171.0	-14.0	171.0	-178.0	117.8	
143	-19	151	-171	-175	-171	-14	171	-178	NA	-89.3
-162.9	71.1	-162.9	180.0	180.0	180.0	180.0	180.0	180.0	108.9	
-162	70	-162	180	180	180	180	180	180	NA	-6.3
-168.7	82.9	-168.7	175.0	176.0	175.0	176.0	175.0	176.0	109.0	
-167	83	-167	175	176	175	176	175	176	NA	-89.1
-166.2	79.0	-166.2	175.0	177.0	175.0	177.0	175.0	177.0	110.0	
-165	79	-165	175	177	175	177	175	177	NA	-88.3

Table 20. Proline ring parameters for cyclo(Pro-Gly)$_3$ uncomplexed

Code	ϕ	χ^1	χ^2	χ^3	χ^4	τ^R_1	τ^R_2	τ^R_3	τ^R_4	τ^R_5	τ^R_6	NC^δ
A1	-98	48.8	-60.0	46.2	-16.7	98.8	97.3	96.7	98.9	110.8	124.8	1.51
A2		37.6	-40.0	26.8	-3.2	102.2	102.4	102.8	104.0	111.6	124.0	1.48
A3		29.9	-20.0	2.1	18.0	102.5	105.2	106.1	105.0	111.5	125.2	1.46
All 3		17.0	0.0	-17.3	30.0	103.3	106.6	106.3	103.8	111.1	125.7	1.45
Prolines		-.6	20.0	-31.8	33.5	104.5	106.2	104.4	102.6	110.9	124.3	1.46
		-19.7	40.0	-45.0	34.7	104.2	103.6	100.3	100.6	110.0	121.5	1.47
		-38.8	60.0	-56.7	35.0	100.9	98.7	94.0	96.6	108.9	117.7	1.51
A6	-68	48.2	-60.0	46.9	-17.8	99.1	97.4	96.6	98.9	110.7	124.3	1.51
A7		40.8	-50.0[a]	39.4	-14.6	101.4	100.4	99.8	101.6	111.4	124.4	1.49
A9		36.2	-44.0[a]	34.4	-12.2	102.4	102.0	101.7	102.8	111.8	124.6	1.48
All 3		33.1	-40.0	31.3	-11.0	103.1	102.9	102.5	103.5	112.1	124.3	1.48
Prolines		26.3	-20.0(-)	6.0	11.3	104.1	105.5	106.1	105.9	111.6	122.8	1.47
A17		6.2	-20.0(+)	26.2	-23.6	105.4	106.4	105.0	104.4	111.9	122.9	1.46
Pro(5)		16.0	0.0	-16.2	27.8	104.4	106.5	106.2	104.8	110.4	122.8	1.46
		-2.5	20.0	-29.9	30.0	104.8	106.3	104.7	103.2	111.5	124.5	1.46
		-29.7	40.0	-34.8	17.1	103.6	103.2	102.1	103.1	111.9	123.6	1.48
		-47.3	60.0	-48.0	19.7	99.4	97.6	96.2	98.8	110.4	123.1	1.51
A8	-63	47.4	-60.0	47.9	-19.5	99.5	97.4	96.4	98.8	110.4	123.4	1.51
Pro(1)		39.8	-50.0	40.5	-16.5	101.8	100.4	99.6	101.5	111.2	123.8	1.49
		31.5	-40.0	33.0	-13.9	103.4	103.0	102.3	103.4	112.0	124.0	1.48
		20.0	-20.0(-)	12.3	0.4	105.1	106.0	106.0	105.9	112.6	123.1	1.47
		3.9	-20.0(+)	28.5	-27.8	105.1	106.4	104.8	103.7	111.5	123.4	1.46
		-13.3	0.0	13.4	-23.2	104.9	107.0	106.2	105.5	111.2	121.4	1.46
		-3.6	20.0(-)	-28.8	28.2	105.1	106.3	104.8	103.7	111.4	123.7	1.46
		-31.8	40.0	-32.6	13.4	103.3	103.0	102.3	103.4	112.0	124.0	1.48
		-48.2	60.0	-46.9	17.9	99.1	97.4	96.5	98.9	110.7	124.1	1.51

[a] X-ray value of χ^2 for the A17. All angles are in degrees; the NC^δ lengths are in Å.

Table 20. Continued

Code	ϕ	χ^1	χ^2	χ^3	χ^4	τ^R_1	τ^R_2	τ^R_3	τ^R_4	τ^R_5	τ^R_6	NC$^\delta$
A8	-59	46.8	-60.0	48.6	-20.6	99.7	97.5	96.2	98.8	110.2	122.9	1.51
Pro(3)		39.0	-50.0	41.3	-17.8	102.0	100.5	99.5	101.4	111.0	123.4	1.49
		30.5	-40.0	34.0	-15.6	103.6	103.1	102.2	103.2	111.9	123.8	1.48
		3.0	-20.0(-)	29.4	-29.3	104.9	106.4	104.7	103.4	111.4	123.9	1.46
		-15.1	0.0	15.2	-26.3	104.5	106.8	106.1	105.1	110.7	121.8	1.46
		-24.6	20.0(-)	-7.7	-8.4	104.3	105.8	106.0	106.0	112.0	122.4	1.47
		-4.7	20.0(+)	-27.7	26.3	105.3	106.3	104.9	104.0	111.6	123.3	1.46
		-32.7	40.0	-31.7	11.7	103.2	102.9	102.5	103.5	112.0	124.2	1.48
		-48.5	60.0	-46.5	17.3	98.9	97.4	96.6	98.9	110.8	124.6	1.51
A8	-67	48.0	-60.0	47.2	-18.3	99.2	97.4	96.5	98.9	110.6	124.0	1.51
Pro(5)		40.5	-50.0	39.7	-15.1	101.5	100.4	99.8	101.6	111.3	124.2	1.49
A16		32.6	-40.0	31.8	-11.9	103.2	102.9	102.5	103.5	112.0	124.2	1.48
Pro(1)		25.5	-30.0	23.0	-7.3	104.3	104.8	104.6	104.9	112.6	124.0	1.47
A10		24.9	-20.0(-)	7.4	8.9	104.4	105.6	106.1	106.0	111.8	122.7	1.47
A11 3		5.8	-20.0(+)	26.6	-24.3	105.4	106.4	104.9	104.3	111.8	122.9	1.46
Prolines		15.3	0.0	-15.4	26.5	104.6	106.5	106.2	105.0	110.6	122.4	1.46
		-2.7	20.0	-29.7	29.8	104.9	106.3	104.7	103.3	111.4	124.2	1.46
		-18.5	30.0	-29.9	19.5	104.7	105.2	103.9	103.8	112.5	124.4	1.47
		-30.5	40.0	-34.0	15.8	103.6	103.1	102.2	103.2	111.9	123.7	1.48
		-47.6	60.0	-47.6	19.1	99.3	97.5	96.3	98.9	110.5	123.4	1.51
A14	-42	42.4	-60.0	53.5	-28.7	101.1	97.7	95.2	97.8	109.1	120.1	1.51
A11 3		24.6	-40.0	40.1	-26.2	104.3	103.3	101.3	102.1	111.0	122.8	1.47
Prolines		2.6	-20.0	29.7	-30.0	104.6	106.4	104.8	103.0	111.7	125.2	1.46
		-17.8	0.0	18.1	-31.4	103.2	106.4	106.2	103.7	110.7	125.5	1.45
		-30.4	20.0	-1.6	-18.9	102.4	105.1	106.1	105.0	111.3	125.0	1.46
		-34.7	40.0	-29.7	8.2	102.7	102.8	102.7	103.6	112.1	124.7	1.48
		-48.4	60.0	-46.4	17.1	98.6	97.5	96.8	98.7	111.2	125.8	1.51

Table 20. Continued

Code	ϕ	χ^1	χ^2	χ^3	χ^4	τ^R_1	τ^R_2	τ^R_3	τ^R_4	τ^R_5	τ^R_6	NC$^\delta$
A15	-41	42.2	-60.0	53.6	-29.0	101.2	97.8	95.1	97.7	104.1	120.0	1.51
All 3		24.4	-40.0	40.3	-26.5	104.3	103.3	101.3	102.0	111.0	122.7	1.47
Prolines		2.6	-20.0	29.7	-30.0	104.6	106.4	104.8	103.0	111.7	125.2	1.46
		-17.8	0.0	18.0	-31.4	103.2	106.4	106.2	103.7	110.7	125.5	1.45
		-30.4	20.0	-1.6	-18.9	102.4	105.1	106.1	105.0	111.3	125.0	1.46
		-34.7	40.0	-29.6	8.1	102.7	102.7	102.7	103.6	112.1	124.6	1.48
		-48.4	60.0	-46.4	17.1	98.7	97.5	96.8	98.7	111.2	125.6	1.51
A16	-62	47.3	-60.0	48.0	-19.7	99.5	97.4	96.3	98.8	110.3	123.3	1.51
Pro(3)		39.6	-50.0	40.6	-16.7	101.8	100.4	99.6	101.5	111.1	123.7	1.49
		31.4	-40.0	33.1	-14.2	103.5	103.0	102.3	103.3	111.9	123.9	1.48
		20.3	-30.0	28.1	-16.3	104.7	105.2	104.1	104.1	112.6	124.3	1.47
		3.7	-20.0	28.7	-28.0	105.1	106.4	104.7	103.7	111.5	123.5	1.46
		-13.6	0.0	13.7	-23.7	104.8	106.9	106.2	105.4	111.1	121.4	1.46
		-3.7	20.0	-28.7	28.0	105.2	106.3	104.8	103.7	111.5	123.6	1.46
		-20.4	30.0	-28.0	16.1	104.7	105.2	104.2	104.2	112.6	124.3	1.47
		-32.0	40.0	-32.5	13.1	103.3	103.0	102.4	103.4	112.0	124.0	1.48
		-48.2	60.0	-46.8	17.8	99.0	97.4	96.6	98.8	110.7	124.2	1.51
A16	-65	47.8	-60.0	47.4	-18.6	99.3	97.4	96.5	98.9	110.5	123.8	1.51
Pro(5)		40.3	-50.0	39.9	-15.5	101.6	100.4	99.8	101.6	111.2	124.1	1.49
		32.3	-40.0	32.1	-12.5	103.3	102.9	102.4	103.4	112.0	124.1	1.48
		23.9	-30.0	24.5	-10.0	104.5	105.0	104.4	104.7	112.7	124.1	1.47
		24.0	-20.0(-)	8.3	7.3	104.6	105.7	106.1	106.0	112.0	122.7	1.47
		5.2	-20.0(+)	27.2	-25.5	105.3	106.4	104.9	104.1	111.7	123.0	1.46
		14.8	0.0	-14.9	25.6	104.8	106.6	106.3	105.2	110.6	122.3	1.46
		-2.9	20.0	-29.5	29.4	104.9	106.3	104.7	103.4	111.4	124.1	1.46
		-18.9	30.0	-29.5	18.8	104.7	105.2	104.0	103.9	112.5	124.4	1.47
		-30.9	40.0	-33.6	15.0	103.5	103.1	102.2	103.3	111.9	123.8	1.48
		-47.8	60.0	-47.4	18.7	99.2	97.5	96.4	98.9	110.5	123.6	1.51

Table 20. Continued

Code	ϕ	χ^1	χ^2	χ^3	χ^4	τ^R_1	τ^R_2	τ^R_3	τ^R_4	τ^R_5	τ^R_6	NC[δ]
A17	-48	43.0	-60.0	52.7	-27.6	100.8	97.8	95.3	98.0	109.4	120.4	1.51
Pro(1)		25.1	-40.0	39.6	-25.3	104.2	103.3	101.4	102.2	111.2	122.8	1.47
		2.6	-20.0	29.7	-29.9	104.6	106.4	104.8	103.0	111.8	125.2	1.46
		-8.7	-10.0[a]	25.0	-32.7	103.9	106.6	105.6	103.2	111.1	125.5	1.45
		-18.0	0.0	18.2	-31.6	103.2	106.4	106.2	103.7	110.6	125.2	1.45
		-30.4	20.0	-1.6	-18.9	102.5	105.1	106.0	105.0	111.2	124.7	1.46
		-34.5	40.0	-29.9	8.5	102.7	102.8	102.7	103.6	112.2	124.7	1.48
		-48.4	60.0	-46.3	17.0	98.6	97.5	96.8	98.7	111.2	125.8	1.51
A17	-81	48.6	-60.0	46.2	-16.8	98.6	97.4	96.8	98.7	111.2	125.7	1.51
Pro(3)		34.5	-40.0	29.8	-8.5	102.7	102.8	102.7	103.6	112.2	124.8	1.48
$C_3$1		32.1	-29.0[a]	14.4	6.4	102.8	104.4	105.2	105.1	112.0	124.8	1.47
All 3		30.4	-20.0	1.7	18.8	102.5	105.1	106.1	105.1	111.2	124.7	1.46
Prolines		18.0	0.0	-18.2	31.6	103.2	106.4	106.2	103.7	110.6	125.2	1.45
		-2.6	20.0	-29.7	30.0	104.6	106.4	104.8	103.0	111.7	125.2	1.46
		-24.6	40.0	-40.0	26.1	104.1	103.5	101.3	102.0	111.2	122.7	1.47
		-43.7	60.0	-51.8	26.4	100.3	98.1	95.2	98.2	109.8	120.5	1.51
$C_3$14 All Proline	-69	-31.7	42.0	-36.0	17.0	103.3	102.7	101.7	102.9	111.6	123.6	1.48
$C_3$15 All Proline	-69	34.5	-42.0	33.0	-11.8	102.7	102.5	102.1	103.2	111.9	124.5	1.48
A18 Pro(1)	-46	25.8	-41.0	40.4	-25.6	104.1	103.1	101.2	102.0	111.0	122.7	1.47
A18 Pro(2)	-65	36.0	-44.0	34.7	-12.6	102.4	102.0	101.6	102.8	111.8	124.5	1.48
A18 Pro(3)	-67	35.5	-43.0	33.6	-11.8	102.5	102.2	101.9	103.8	111.9	124.6	1.48

Table 21. π-π^* predicted maxima and minima for cyclo(Pro-Gly)$_3$ uncomplexed[a]

Form	χ^2	λ_{max}	ε_{max}	λ_1	$\Delta\varepsilon_1$	λ_2	$\Delta\varepsilon_2$	λ_3	$\Delta\varepsilon_3$
A1	-60	192	5231	186	-4.1	200	9.5	214	-3.0
	-40	192	5207	188	-5.8	202	8.4	216	-6.0
	-20	194	5130	188	-7.4	202	4.0	216	-9.3
	0	194	5086	188	-8.6	204	1.4	216	-10.5
	20	194	5121	188	-9.0	204	0.9	216	-9.5
	40	194	5210	188	-8.9	202	1.2	216	-6.5
	60	192	5224	188	-7.9	202	1.9	214	-3.2
A3	-50	194	4700	186	-2.4	200	7.0	216	-3.7
	-40	200	4773	186	-2.5	202	8.9	220	-1.8
	-20	196	4759	186	-2.9	202	6.8	218	-6.1
	0	194	4705	188	-3.4	202	4.8	218	-6.7
	20	196	4722	188	-3.2	204	4.6	218	-5.7
	40	196	4799	188	-2.4	202	4.6	218	-3.6
	60	194	4786	184	-1.2	200	4.5	216	-1.8
A6	-60	190	5565	190	-5.5	206	10.6	-	-
	-40	190	5789	190	-7.3	210	13.5	-	-
	-20(-)	190	5793	190	-8.9	212	11.1	-	-
	-20(+)	190	5790	190	-10.7	216	17.5	-	-
	0	190	5798	192	-10.9	212	7.9	-	-
	20	190	5870	192	-13.0	212	8.2	-	-
	40	190	5843	190	-15.0	214	11.6	-	-
	60	190	5849	190	-13.8	210	11.6	-	-
A7	-40	190	4856	186	-2.6	208	12.4	-	-
	-20(-)	190	4967	186	-3.7	208	7.3	-	-
	-20(+)	190	4941	188	-5.0	214	13.2	-	-
	0	190	4986	188	-6.1	208	4.1	226	-0.4
	20	190	5023	188	-8.2	208	4.2	-	-
	40	190	5037	188	-9.2	210	7.1	-	-
	60	190	5053	186	-8.0	206	7.1	-	-

[a] χ^2 are in degrees. λ, λ_1, λ_2, and λ_3 are in nm. ε and $\Delta\varepsilon$ are in Lmol^{-1} cm^{-1}. The bandwidths for the calculations are as follows: A1 through A18, 6000 cm^{-1}; $C_3$1, 4000 cm^{-1}; $C_3$14 and $C_3$15, 6000 cm^{-1}.

Table 21. Continued

Form	χ^2	λ_{max}	ε_{max}	λ_1	$\Delta\varepsilon_1$	λ_2	$\Delta\varepsilon_2$	λ_3	$\Delta\varepsilon_3$
A8	-50	188	4312	190	-3.0	206	3.4	250	7.2
	-40	190	4490	190	-2.6	208	10.8	-	-
	-20(-)	190	4606	190	-3.9	210	6.2	-	-
	-20(+)	190	4587	188	-3.8	210	9.4	-	-
	0	190	4502	190	-6.3	210	7.1	-	-
	20(-)	190	4612	190	-7.4	212	3.7	-	-
	20(+)	190	4693	190	-6.8	210	3.4	232	-0.8
	40	190	4633	190	-8.0	210	6.2	-	-
	60	190	4578	188	-7.3	206	6.6	-	-
A9	-50	198	5469	184	-0.4	198	3.0	250	10.2
	-40	202	5800	-	-	202	7.8	216	-1.2
	-20(-)	204	5913	186	-0.3	202	6.8	216	-6.2
	-20(+)	204	6044	-	-	202	5.8	216	-3.6
	0	204	5869	188	-1.5	202	4.7	216	-7.5
	20	204	5901	188	-1.6	202	3.0	214	-6.0
	40	200	6040	188	-0.6	200	2.1	214	-3.3
	60	198	5931	186	-0.3	198	1.1	210	-1.0
A10	-60	198	5360	190	-3.5	204	3.8	-	-
	-40	200	5326	192	-4.0	204	3.0	-	-
	-20(-)	204	5236	192	-4.2	206	2.5	222	-0.3
	-20(+)	204	5341	192	-4.6	208	4.0	-	-
	0	204	5237	192	-4.2	204	2.4	220	-0.8
	20	204	5311	192	-4.1	206	3.3	222	-0.3
	40	202	5446	192	-4.6	206	4.2	-	-
	60	198	5528	190	-5.2	204	4.0	-	-
A14	-60	192	5010	186	-3.3	206	12.0	-	-
	-40	192	5053	188	-3.4	210	15.1	-	-
	-20	194	5037	188	-3.7	212	17.4	-	-
	0	194	4956	188	-4.1	214	18.6	-	-
	20	194	5025	188	-5.3	212	16.8	-	-
	40	194	5258	188	-6.8	210	11.8	-	-
	60	194	5570	188	-6.9	206	8.9	-	-
A15	-40	204	3920	186	-3.9	206	16.4	-	-
	-20	206	4205	186	-3.0	208	15.4	-	-
	0	208	4437	188	-2.9	210	15.6	-	-
	20	206	4526	188	-4.2	208	14.6	-	-
	40	204	4490	188	-6.0	208	10.9	-	-
	60	194	4479	184	-4.9	198	7.2	-	-

[b]This was the last experimental point.

Table 21. Continued

Form	χ^2	λ_{max}	ε_{max}	λ_1	$\Delta\varepsilon_1$	λ_2	$\Delta\varepsilon_2$	λ_3	$\Delta\varepsilon_3$
A18	-41,-44,-43	192	5068	188	13.0	208	-8.8	-	-
Experiment in H_2O	NA	NA	NA	190*	6.1*	213	-6.1	-	-
$C_3$14	42	190	6293	190	-13.8	212	11.4	-	-
$C_3$15	-42	190	5978	190	-3.9	212	8.9	-	-
Experiment in dioxane	NA	NA	NA	NA	NA	212*	0.3*	230	-5.4

Figure 20. π-π^* absorption spectra for uncomplexed forms of cyclo(Pro-Gly)$_3$ treated as cyclo(Gly)$_6$

Units for the scales are as follows: nm for λ and 10^3 Lmol^{-1}cm^{-1} for ε. Each spectrum is labeled by the backbone code. The scale listed for A1 is the same for all other forms. Bandwidths are 6000 cm^{-1} for A1 through A17 and 4000 cm^{-1} for $C_3$1.

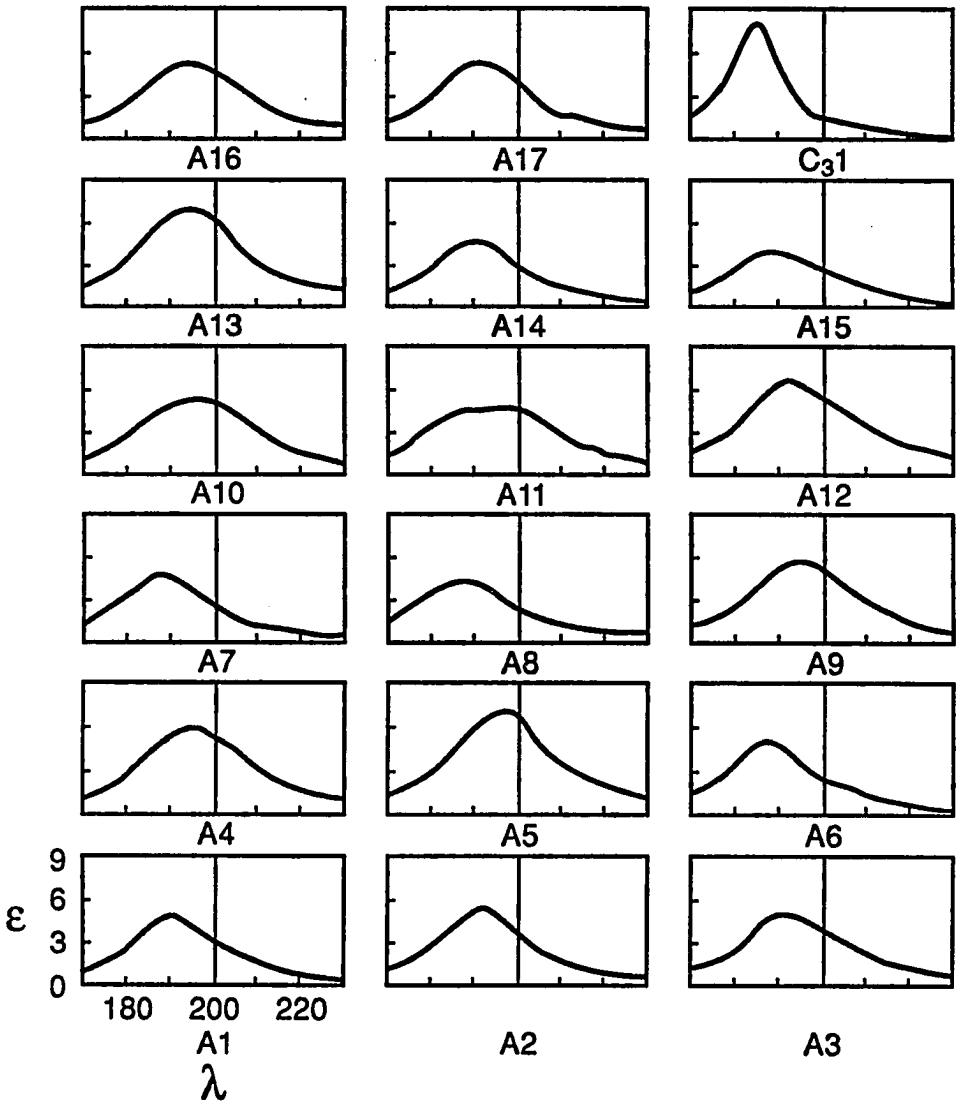
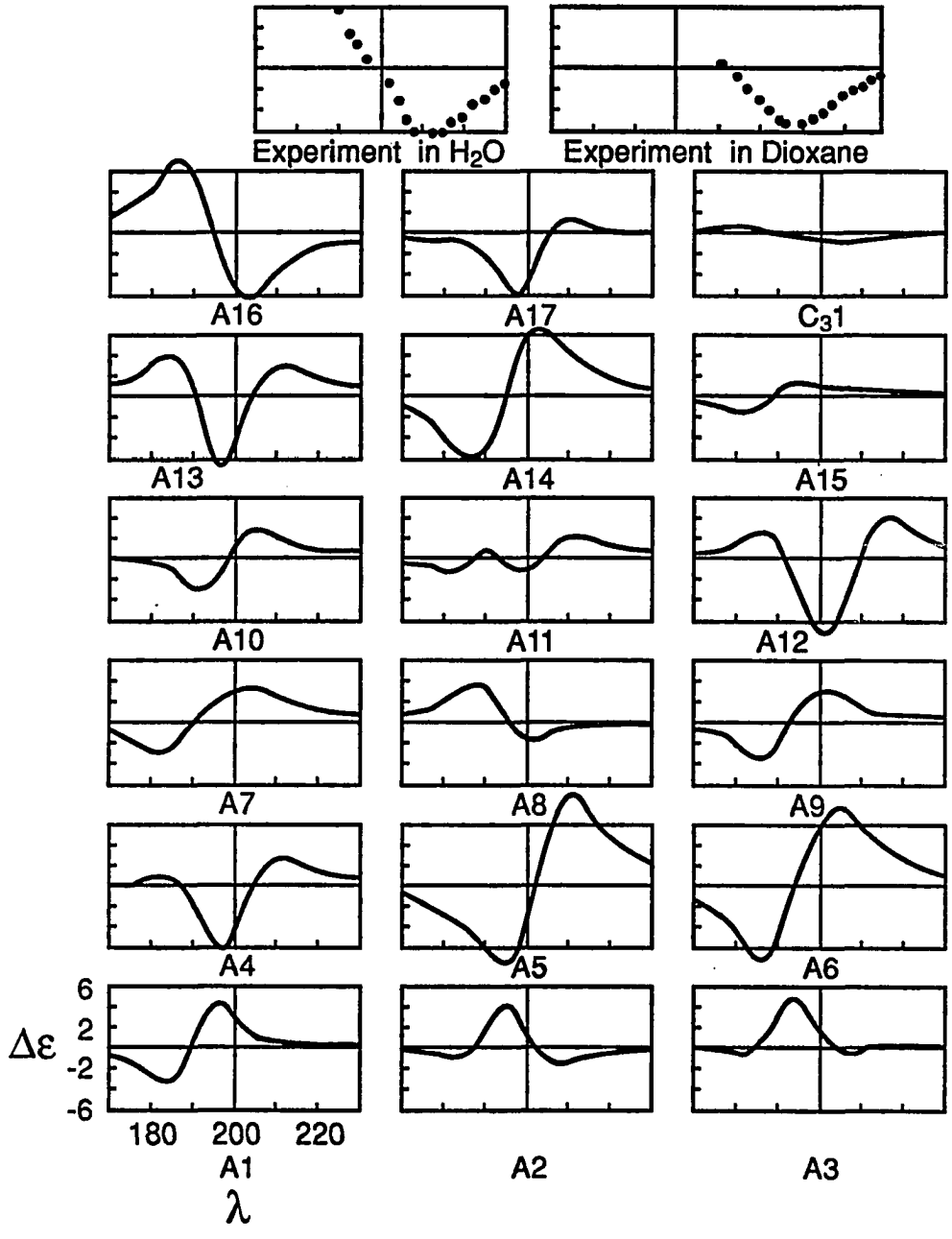

A16
A17
C31
A13
A14
A15
A10
A11
A12
A7
A8
A9
A4
A5
A6
ε
9
6
3
0
180
200
220
A1
λ
A2
A3

Figure 21. $\pi-\pi^*$ CD spectra for uncomplexed forms of cyclo(Pro-Gly)$_3$ treated as cyclo(Gly)$_6$

Units for the scales are as follows: nm for λ and $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$. Each spectrum is labeled by the backbone code. The scale listed for A1 is the same for all other forms. Bandwidths are 6000 cm^{-1} for A1 through A17 and 4000 cm^{-1} for $C_3$1.


Experiment in H_2O
Experiment in Dioxane
A16
A17
C_31
A13
A14
A15
A10
A11
A12
A7
A8
A9
A4
A5
A6
A1
A2
A3
Δε
6
2
-2
-6
180
200
220
λ

structure. The experimental CD spectra were obtained from reference 33. No experimental absorption spectra were available (193). The other Figures, 22-25, and Table 21, include the proline side chain. These figures and table reiterate the sensitivity of the dipole interaction model to changes in the backbone and side chain conformations. Some structures give a reasonably good overall fit to the π-π^* spectra for cyclo(Pro-Gly)$_3$.

D. Discussion

1) Uncomplexed cyclo(Pro-Gly)$_3$ in nonpolar solvents

Predicted π-π^* absorption and CD spectra for the uncomplexed form of cyclo(Pro-Gly)$_3$ in nonpolar solvents can be seen in Figures 20, 21, and 22. For example, $C_3$1 has both predicted absorption and CD which are sensitive to the presence of proline ring structure. There is always a strong absorption band present near 185 to 188 nm irrespective of the presence or conformation of the proline ring; this band is slightly more intense and red when the proline ring is present. The most interesting feature of the absorption spectra is the weak band appearing around 210 nm only when the proline ring is present. This weak absorption band may be responsible for the CD band that appears around 220 nm when proline is included in the calculation. The CD spectra for the cyclo(gly)$_6$ approximation is weak (around $\pm$ 1 L mol^{-1} cm^{-1}). Once the proline ring is included, the CD intensifies considerably to fall within the region of experiment (Figure 22). The greatest difference between experimental and calculated CD spectra is the location of the negative band. The experimental CD band shows a negative maximum at 230 nm; whereas, the predicted CD band falls

Figure 22. π-π^* absorption and CD spectra for cyclo(Pro-Gly)$_3$ C$_3$1, the uncomplexed form in nonpolar solvents

Units for the scales are as follows: nm for λ and $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled according to χ^2 value. All spectra are to the scale listed for the lower left pair. The bandwidth used for all calculated spectra was 4000 cm^{-1}.

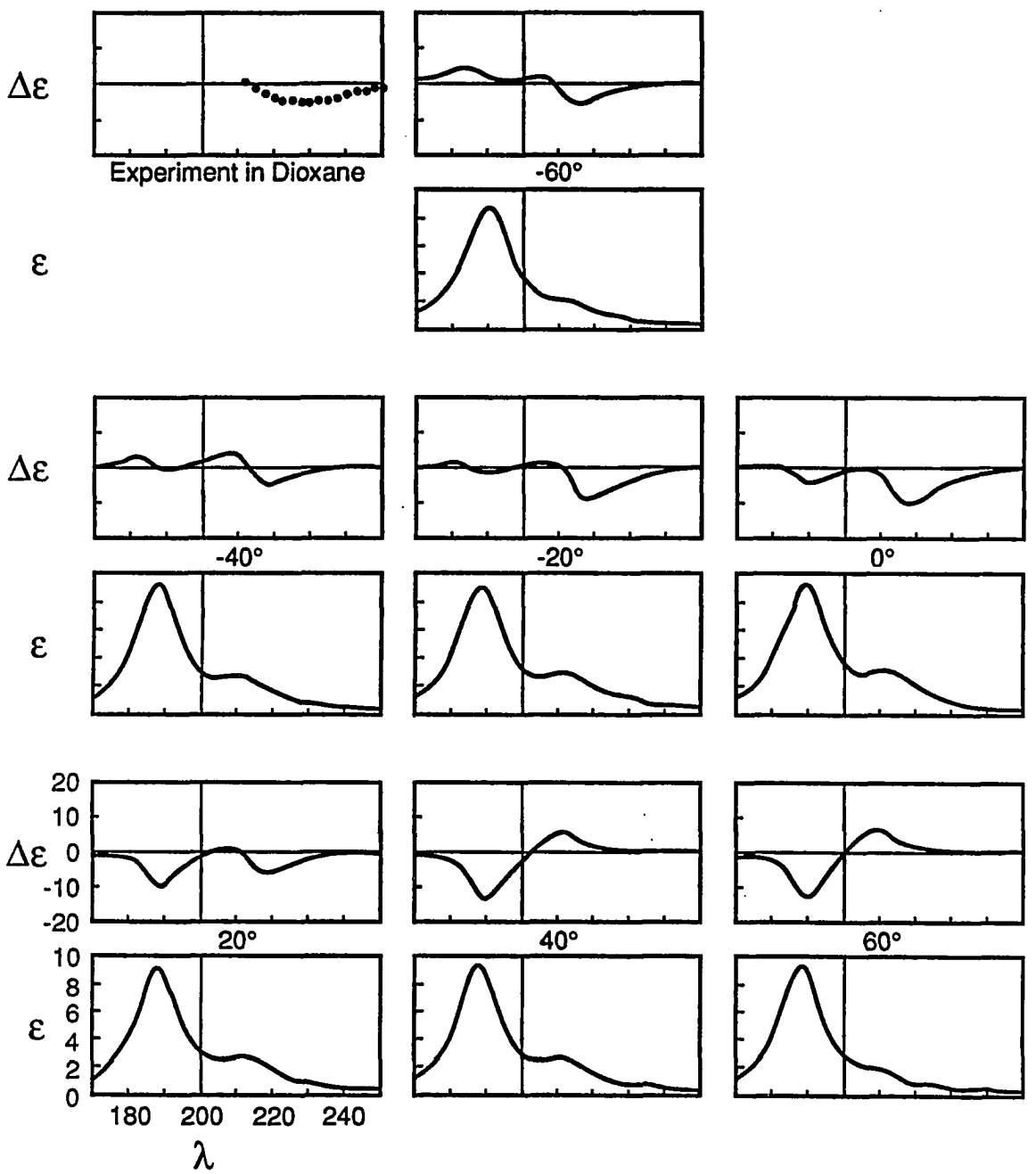

Δε
ε
Experiment in Dioxane
-60°
-40°
-20°
0°
20°
40°
60°
20
10
0
-10
-20
10
8
6
4
2
0
180
200
220
240
λ

Figure 23. π-π^* absorption and CD spectra for cyclo(Pro-Gly)$_3$ uncomplexed in polar solvents backbone A2

Units for the scales are as follows: nm for λ and $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled according to χ^2 value. All spectra are to the scale listed for the lower left pair. The bandwidth is 6000 cm^{-1}.

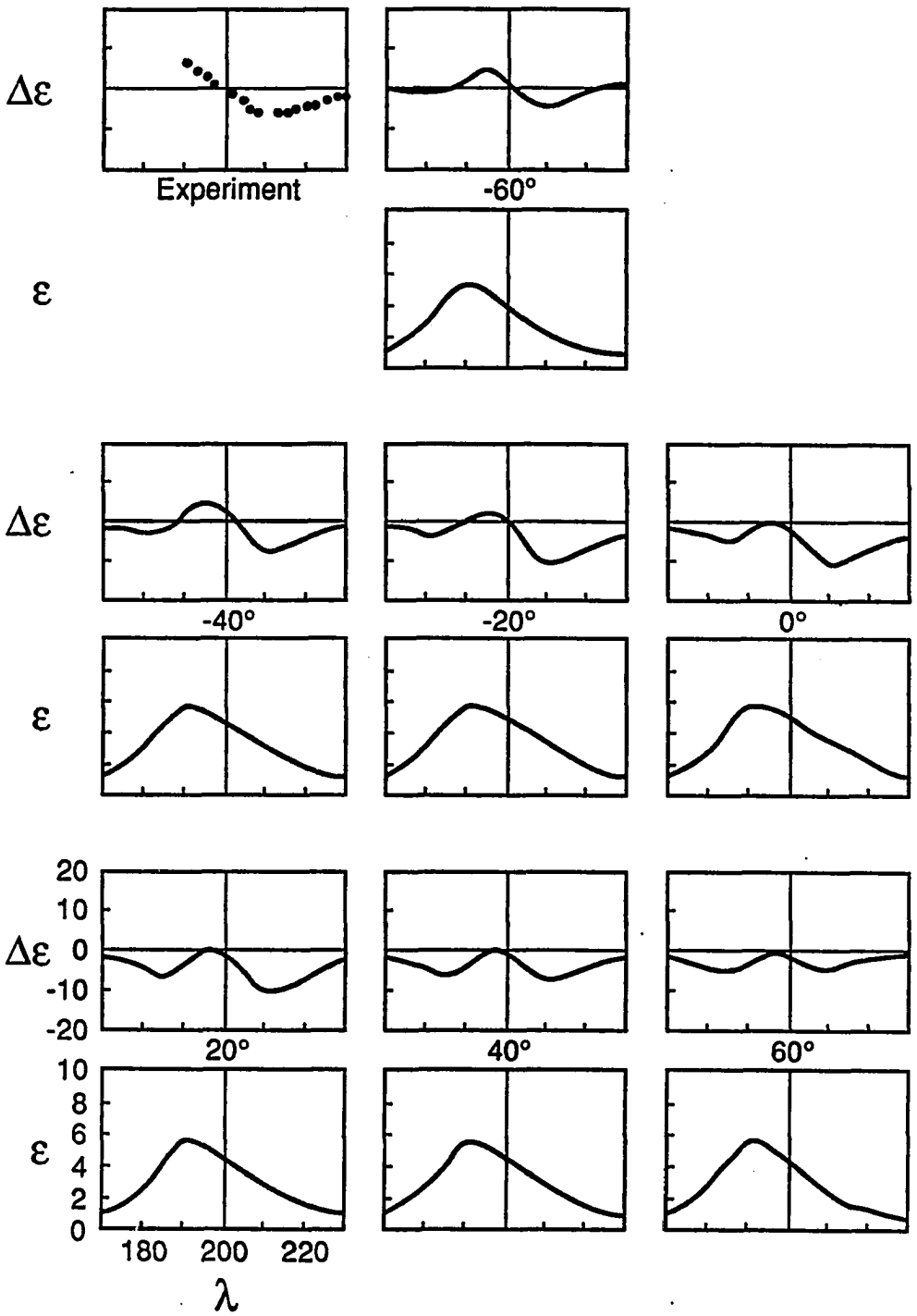

Δε
Experiment
-60°
ε
Δε
-40°
-20°
0°
ε
20
10
Δε 0
-10
-20
20°
40°
60°
10
8
6
ε 4
2
0
180
200
220
λ

Figure 24. $\pi-\pi^*$ absorption and CD spectra for cyclo(Pro-Gly)$_3$ uncomplexed in polar solvents backbone A16

Units for the scales are as follows: nm for λ and $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled according to χ^2 value. The bifurcation at -20° are designated (+) and (-) for solutions $\chi_0^1 = +\chi^2$ and $\chi_0^1 = -\chi^2$ respectively. All spectra are to the scale listed for the lower left pair. The bandwidth is 6000 cm^{-1}.

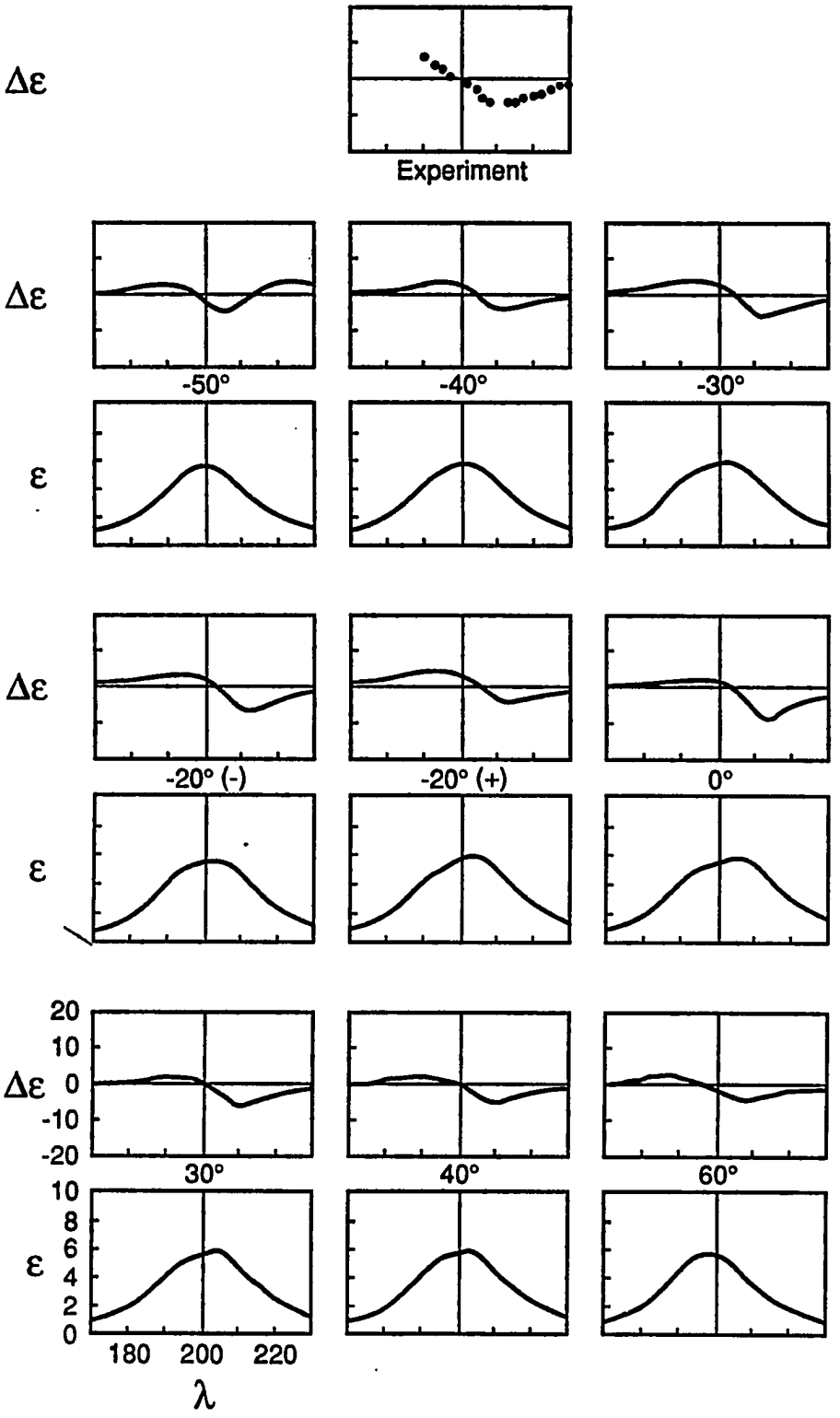
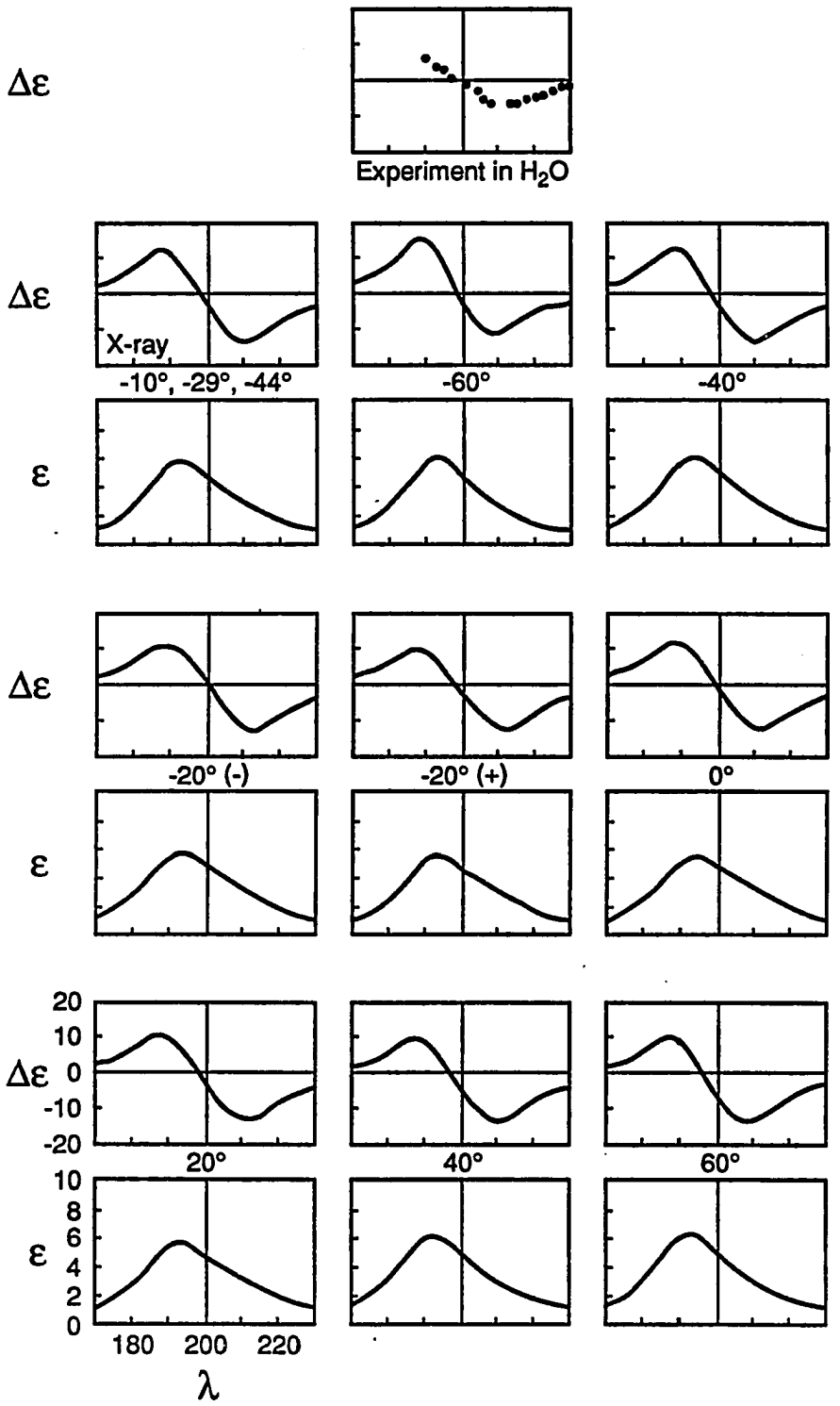

Δε
Experiment
Δε
-50°
-40°
-30°
ε
Δε
-20° (-)
-20° (+)
0°
ε
20
10
Δε 0
-10
-20
30°
40°
60°
10
8
ε 6
4
2
0
180
200
220
λ

Figure 25. π-π^* absorption and CD spectra of cyclo(Pro-Gly)$_3$ uncomplexed in polar solvents backbone A17

Units for the scales are as follows: nm for λ and $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled according to χ^2 value. The bifurcation at -20° is designated (+) and (-) for solutions $\chi_0^1 = +\chi^2$ and $\chi_0^1 = -\chi^2$ respectively. The X-ray values of χ^2 are -10, -29, and -44. All spectra are to the scale listed for the lower left pair. The bandwidth is 6000 cm^{-1}.


Δε
Experiment in H2O
Δε
X-ray
-10°, -29°, -44°
-60°
-40°
ε
Δε
-20° (-)
-20° (+)
0°
ε
20
10
Δε 0
-10
-20
20°
40°
60°
10
8
6
ε 4
2
0
180
200
220
λ

around 220 nm ($-40° \leq \chi^2 \leq 40°$). The discrepancy could be due to the exclusion of the n-π* transition from the present calculation. What is surprising is that the π-π* transition may be considered to contribute in part to the experimental band at 220 nm.

The CD spectra calculated for cyclo(Pro-Gly)$_3$ in nonpolar solvents ($C_3$1) is somewhat dependent on the proline ring conformation. The intensity of the CD band near 220 nm is affected by the proline ring conformation, but more dramatic effects are seen at the bluer CD bands. When $\chi^2 \geq 0°$, the predicted CD band near 190 nm is negative (Figure 22). When $\chi^2 < 0°$, however, the band is bluer (near 182 nm) and positive. It is unfortunate that CD measurements could not be recorded bluer than 212 nm because of solvent interference (33). Without information about this region, few conclusions can be drawn about the proline ring puckering except that the ring is probably puckered in the regions of $\chi^2 = +20°$ to +40° and -20° to -40°.

Hori et al. (36) used the $C_3$1 backbone to produce two energy minimized cyclo(Pro-Gly)$_3$ molecules that were very similar to each other, $C_3$14 and $C_3$15 (Table 19). $C_3$14 and $C_3$15 showed one significant difference between the two: the ring puckered in opposite directions ($\chi^2 = 42°$ and -42° respectively). The predicted CD for these two structures showed little in common with experiment because they did not have a band in the 220 nm region as did $C_3$1 (Table 21). They did produce a strong positive band at 212 nm which corresponded to those of $C_3$1 for $\chi^2 = \pm 40°$ except that those for $C_3$1 were weaker. These differences in predicted CD may be due to the

changes in ϕ, ψ, and ω from the $C_3$1 structure that were at most 20° (Table 19).

2) Uncomplexed cyclo(Pro-Gly)$_3$ in polar solvents

The proposed backbones for cyclo(Pro-Gly)$_3$ uncomplexed in polar solvents are not as clearly defined as in nonpolar solvents. As a result, π-π^* absorption and CD spectra were predicted for all 18 structures whether they were proposed by CD theory, energy minimizations, or X-ray crystallography. The calculated π-π^* absorption and CD spectra for the uncomplexed forms of cyclo(Pro-Gly)$_3$ treated as cyclo(Gly)$_6$ structures can be seen in Figures 20 and 21, respectively. The predicted absorption spectra are not very dependent on the backbone conformations; they show a broad band around 190 to 195 nm that varies in intensity from 4500 L mol^{-1} cm^{-1} to 6900 L mol^{-1} cm^{-1} (Figure 20). The predicted CD spectra, however, are very backbone dependent (Figure 21).

The experimental CD spectrum for cyclo(Pro-Gly)$_3$ uncomplexed in H_2O can be seen in Figure 21 (33). It shows a strong negative band around 210 nm. Dipole interaction predictions for some of the proposed backbone structures (cyclo(Gly)$_6$) show a strong positive band in this region; these include A4, A5, A11, A12, and A13. Therefore, these backbones are probably not representative of the structures in solution. This is not surprising because all are higher energy structures (energies in kcal/mole: A4, 11.2; A5, 35.6; A11, 6.9; A12, 2.9; A13, 10.5 (38)). Because the predicted CD did not resemble experiment and the potential energies were high, it was not considered necessary to pursue these structures further by addition of BORC proline rings.

There were other structures that had a strong positive CD band between 200 and 210 nm. These include A6, A7, A8, A10, and A14 (Figure 21). Because this band was less than 210 nm and because some of these structuers were minimum energy structures (energy in kcal/mole: A6, -6.5; A7, -3.8; A8, 1.0; A10, 1.7; A14, -5.7 (38)), they warranted further study by addition of the proline ring. In all of these cases, however, the proline ring did not perturb the predicted CD spectra enough to produce a negative CD band near 210 nm (Table 21). Generally, the positive band intensified and slightly red shifted. Occasionally, an extremely weak negative band appeared at or past 220 nm which was much too red and weak to resemble experiment. These results indicate that the A6, A7, A8, A10, and A14 backbones are not representative of the uncomplexed form of cyclo(Pro-Gly)$_3$ in H_2O irrespective of the potential energies. This leaves A1, A2, A3, A9, A15, A16, and the X-ray structure, A17, for further investigation.

The CD spectrum predicted for the backbone of A1 shows a positive band at 196 nm (which is near the location to a similar experimental band), but it shows no negative band at 210 nm as experiment (Figure 21). Since A1 has the lowest potential energy of any of the proposed backbones (-7.1 kcal/mole) (38), addition of the proline ring should be of interest. When the proline ring is added a negative band does appear in the region of 214 to 216 nm (Table 21). Moreover, intensities of this band favor χ^2 values around $\pm$ 40°. Furthermore, the preceding band is positive like experiment. The predicted CD for A1 are slightly red of experiment. This indicates that the experimental spectrum cannot be solely due to the A1 structure, but it may contribute as part of a mixture of structures.

The A2 backbone shows a similar positive band to that of the A1 backbone (Figure 21), but A2 shows a weak negative band at 206 nm. When the proline side chain is added, this negative band intensifies and red shifts to resemble the negative band of experiment nicely (Figure 23). The positive band, however, is side chain dependent and only appears when $\chi^2 <$ 0°. Thus, A2 may be considered a partial representative backbone for the uncomplexed form of cyclo(Pro-Gly)$_3$ in H_2O when χ^2 is negative.

The A3 backbone also shows the positive band at 194 nm resembling that for A1 and A2. It also indicates a hint of a negative band at 208 nm (Figure 21). When the proline ring is added, the negative band intensifies, but not as much as it did for A1 (Table 21). The resulting CD spectra are red of experiment (216-220 nm) although the weak maxima at $\chi^2 = \pm$ 40° do resemble the experiment in that region. Thus, just as A1 was a possible contributor to the conformations of cyclo(Pro-Gly)$_3$ in H_2O, so A3 may be considered a possibility. It may not contribute as much as A1 because of its higher energy (0.2 kcal/mole) (38) and the red shift predicted in the CD spectrum.

The A9 backbone shows a weak negative CD band around 200 nm (Figure 21). When the proline side chain is added, this band becomes positive and a negative band appears from 210 to 216 nm depending on χ^2 (Table 21). The intensity of the negative CD band approaches experiment best at $\chi^2 = \pm$ 20°. The positive band near 200 nm, however, is much sharper than experiment. Therefore, the CD of the uncomplexed form of cyclo(Pro-Gly)$_3$ cannot be attributed to A9 alone. Just as A1, A2, and A3 could be considered partial

contributors, so could A9 be considered a possible contributing conformation.

The A15 backbone which was also fairly low in potential energy (-4.6 kcal/mole (38)) produces a theoretical CD curve with a positive band around 194 nm. This band is much broader and weaker than those predicted for A1, A2, and A3. When the proline ring is added to A15, this CD band red shifted to around 208 nm, remained positive, and greatly intensified (Table 21). Since this spectrum is basically of opposite sign to experiment, A15 is not representative of cyclo(Pro-Gly)$_3$ uncomplexed in polar solvents.

The A16 backbone produced a strong CD spectrum that showed a negative band around 204 nm (Figure 21). Addition to the proline ring reduces the intensity of the predicted CD spectrum and red shifts it so it better resembles experiment (Figure 24). Those predicted CD best resembling experiment were for χ^2 = +30° to +40° and -30° to -40°. The positive band near 190 nm is weaker than experiment, but this is not a serious discrepancy considering experiment cuts off at 190 nm. Thus, the A16 backbone is very representative of cyclo(Pro-Gly)$_3$ uncomplexed in nonpolar solvents.

The X-ray structure backbone, A17, does not appear as promising initially because a weak positive band appears at 210 nm (Figure 21). When the proline ring is added, however, this band disappears and a broad, strong, negative band appears near 210 nm (Figure 25). Although the overall predicted CD spectra are stronger than experiment, the sign, location, and shapes of the bands greatly resemble experiment. Changes in χ^2 have little effect on the overall appearance of the CD spectra. Thus,

the A17 backbone can also be considered very representative of cyclo(Pro-Gly)$_3$ uncomplexed in solution.

The energy minimized structure of Hori et al. (36) coded A18 was also used to calculate a CD spectrum. The A18 spectrum resembled experiment by producing a negative band at 208 nm and a positive one at 188 nm (Table 21). All three proline rings were intensely puckered near $\chi^2 = -40°$. This puckering supported some of the earlier puckering predictions for A1, A3, and A16 of $\chi^2 = \pm 40°$ and the negative puckering predictions for A2.

E. Conclusions

The principal finding of this study is that approximately correct π-π^* CD spectra of the uncomplexed forms of cyclo(Pro-Gly)$_3$ are predicted by the dipole interaction model when the proline side chain is included on certain predetermined backbone structures. The experimental CD spectra fall into two classes, one for nonpolar solvents and one for polar solvents. Those backbones in the nonpolar solvents are C_3 symmetric with all trans peptide bonds, and the proline ring may be intensely puckered (χ^2 around $\pm$ 40°).

The CD of uncomplexed cyclo(Pro-Gly)$_3$ in polar solvents may be due to a mixture of structures. Minor contributors include A1, A2, A3, and A9. Major contributors are backbones like A16, A17, and A18. All of these structures support intense proline ring puckering (A1, $\chi^2 = \pm 40°$; A2, $\chi^2 < 0°$; A3, $\chi^2 = \pm 40°$; A9, $\chi^2 = \pm 20°$; A16, $\chi^2 = +30°$ to $+40°$ and $-30°$ to $-40°$; A17, $\chi^2 = -10°, -20°, -44°$ (37); and A18, $\chi^2 = -41°, -44°, -43°$ (36)). All of the major contributors have one interesting structural feature; they all include one cis peptide bond for a glycine residue. Two of the minor contributors, A3 and A9, also include one cis peptide bond for

a glycine residue. Madison et al. (33) also concluded that the CD spectrum was due to asymmetric structures with one cis peptide bond. Moreover, Raman spectroscopic studies indicated the presence of some cis peptide linkages (194). The other minor contributors are all trans conformations. Thus, the predominating structures of cyclo(Pro-Gly)$_3$ in polar solvents have one cis peptide bond, and proline ring puckering is intense with a slight favoring toward $\chi^2 < 0°$.

VII. CYCLO(GLY-PRO-GLY)$_2$

A. Introduction

Many naturally occurring cyclic peptides and proteins have regions where the polypeptide chain reverses direction (39). These regions include various types of turns: β-turns which involve hydrogen bonding occurring between the carbonyl of the first residue of the turn, i, and the NH of the residue located three residues away, i+3, and γ-turns which involve hydrogen bonding occurring between the carbonyl residue i and the NH of the residue i+2 (27). The synthetic peptide cyclo(Gly-Pro-Gly)$_2$ is an excellent model for both β- and γ-turns (1, 27, 39, 41, 42). Although cyclo(Gly-Pro-Gly)$_2$ is of predominant interest as a model for β- and γ-turns, it also binds cations much the way cyclo(Pro-Gly)$_3$ does (1, 195); cyclo(Gly-Pro-Gly)$_2$ has been shown to bind Na^+, Mg^{2+}, Zn^{2+}, and Li^+ (195).

Many previous studies have been done on cyclo(Gly-Pro-Gly)$_2$. These include X-ray crystal structure determination (142), energy minimizations (1, 40), NMR studies (1, 39-41), and CD studies which include experimental (1, 39, 195, 196) and theoretical treatments (1, 39). Moreover, complex formation with metal ions has also been studied (195).

The crystal structure of cyclo(Gly-Pro-Gly)$_2$ is an asymmetric structure with two β-turns; one turn is type I and the other is type II (42). Turn types are defined by Venkatachalam (197). Many of the proposed solution structures also contain β-turns, but solution structures are always C_2 symmetric (1, 39-41). Structures proposed by NMR show cyclo(Gly-Pro-Gly)$_2$ to be C_2 symmetric with two type II β-turns (40, 41). Energy minimizations have produced C_2 conformations that have two type II β-turns, two type II'

β-turns, two type I β-turns, or have the carbonyls located to favor cation binding (1). CD studies which included theoretical predictions also favored C_2 symmetric structures with type II β-turns, but these studies further noted that there was a slight contribution due to γ-turns (39).

Since the dipole interaction model has predicted reasonable CD for helices, cyclic peptides, and β-structures (2, 3, 6-11, Chapters IV, V, and VI), and since it includes the proline side chain, application of this model to the structures of cyclo(Gly-Pro-Gly)$_2$ may provide further insight to the proposed solution backbones and proline ring puckering for this molecule.

B. Structure Generation

Cyclo(Gly-Pro-Gly)$_2$ forms were generated using the Ramachandran-Sasisekharan technique (43) described in the methods chapter. The parameters of the cyclo(Gly-Pro-Gly)$_2$ backbones are derived from the published minimum energy structures (1), the X-ray structure (42), or from the published NMR structures (39-41). Backbone rings were closed as described in the methods chapter. They needed to be closed in order to keep the end of the chain in the proper geometry with as few modifications to the literature φ and ψ as possible. The literature parameters needed to be revised to remove the differences between the Ramachandran-Sasisekharan parameters and those in the literature. Often in the literature, no $\angle NC^{\alpha}C'$ angle was provided, and/or the other bond lengths or angles differed from Ramachandran-Sasisekharan values. Moreover, in some cases closure was needed because the rounded data in the literature were not good enough to give accurate closure by the stepwise method of chain generation.

Furthermore, each residue used in the current study is assumed to have the same bond angles and bond lengths regardless of whether the residue is proline or glycine. Crystal structures often distinguish the two and even have variations among residues of the same kind. After backbone ring closure, BORC was applied to the proline rings for the resulting values of ϕ and τ_1 to produce a wide range of proline rings for optical studies.

π-π^* absorption and CD spectra were calculated for both the C_2 symmetric structures and the asymmetric crystal structure. Half-peak bandwidths were 6000 cm^{-1}. The C_2 symmetric backbone forms were coded C_2 for C_2 symmetry; whereas, the asymmetric form was coded A; all were and arbitrarily numbered. $C_2$1 is the structure derived from NMR conformation of Gierasch et al. (39). $C_2$2 is the solution conformation proposed by Schwyzer et al. from NMR (40). $C_2$3 is the NMR structure of Pease et al. (41). $C_2$4 through $C_2$11 are the energy minimized structures of Deber et al. (1). A1 is the asymmetric crystal structure (42).

C. Results

The backbone parameters of the closed rings used for the various forms of cyclo(Gly-Pro-Gly)$_2$ are listed in Table 22. Backbone ring closure produced changes in the ϕ and ψ angles that were at most 13° (although most changes were less than 6°) with the following exceptions. $C_2$2 had a 39° change in ϕ_p, but this NMR structure had ϕ and ψ listed as approximate values (40). $C_2$3 had a 20° change in ϕ_p, but again this NMR paper listed ϕ and ψ as approximate (41). $C_2$5 had an 18° change in ϕ_p with a 24° change in ψ_g^3 and $C_2$7 had a 40° change in ϕ_p with a 34° change in ψ_g^3. These were

Table 22. Parameters for backbones of cyclo(Gly-Pro-Gly)$_2$[a]

Code	Ref[b]	ϕ_g^1	ϕ_p^2	ϕ_g^3	ϕ_g^4	ϕ_p^5	ϕ_g^6	ψ_g^1	ψ_p^2	ψ_g^3
$C_2$1		168.9	-73.1	148.8	168.9	-73.1	148.8	178.6	88.9	-1.3
	39	170	-70	150	170	-70	150	180	90	0
$C_2$2		-168.1	-30.8	107.0	-168.1	-30.8	107.0	137.3	126.3	-29.6
	40	-165	-70	100	-165	-70	100	150	120	0
$C_2$3		-178.2	-80.2	90.0	-178.2	-80.2	90.0	-178.2	131.2	7.3
	41	180	-60	90	180	-60	90	180	120	0
$C_2$4		158.0	-63.6	65.0	158.0	-63.6	65.0	-178.9	100.0	74.0
	1	160	-68	65	160	-68	65	180	100	74
$C_2$5		78.6	-85.5	85.4	78.6	-85.5	85.4	-159.1	80.8	177.5
	1	80	-68	80	80	-68	80	-160	80	154
$C_2$6		-80.1	-68.2	132.3	-80.1	-68.2	132.3	-179.0	99.9	-80.4
	1	-80	-68	133	-80	-68	133	180	100	-80
$C_2$7		-160.0	-108.4	80.1	-160.0	-108.4	80.1	-81.7	144.4	-39.1
	1	-160	-68	85	-160	-68	85	-80	140	-73
$C_2$8		-60.3	-67.9	100.9	-60.3	-67.9	100.9	-80.0	100.1	-166.7
	1	-60	-68	101	-60	-68	101	-80	100	-166
$C_2$9		120.0	-67.2	-56.0	120.0	-67.2	-56.0	-80.9	169.8	96.7
	1	120	-68	-56	120	-68	-56	-80	170	97

[a]All angles are in degrees. Subscripts p and g stand for proline and glycine, respectively. Superscripts refer to the residue number.

[b]Reference backbone parameters are from the literature; those immediately above them are optimized backbone parameters.

[c]$\angle NC^{\alpha}C'$ is the same for all six residues.

[d]TPE is total potential energy calculated for the backbones by the given reference; they are in kcal/mole.

[e]NA means the parameter was not listed in the reference.

ψ^4_g	ψ^5_p	ψ^6_g	ω^1_g	ω^2_p	ω^3_g	ω^4_g	ω^5_p	ω^6_g	$\langle NC^{\alpha}C'^{c}$	TPE[d]
178.6	88.9	-1.3	180.0	180.0	180.0	180.0	180.0	180.0	109.8	
180	90	0	180	180	180	180	180	180	NA[e]	NA
137.3	126.3	-29.6	180.0	180.0	180.0	180.0	180.0	180.0	110.6	
150	120	0	180	180	180	180	180	180	NA	NA
-178.2	131.2	7.3	180.0	180.0	180.0	180.0	180.0	180.0	109.5	
180	120	0	180	180	180	180	180	180	NA	NA
-178.9	100.0	74.0	180.0	180.0	180.0	180.0	180.0	180.0	109.5	
180	100	74	180	180	180	180	180	180	NA	-6.3
-159.1	80.8	177.5	180.0	180.0	180.0	180.0	180.0	180.0	102.6	
-160	80	154	180	180	180	180	180	180	NA	-6.1
-179.0	99.9	-80.4	180.0	180.0	180.0	180.0	180.0	180.0	109.5	
180	100	-80	180	180	180	180	180	180	NA	-6.6
-81.7	144.4	-39.1	180.0	180.0	180.0	180.0	180.0	180.0	117.9	
-80	140	-73	180	180	180	180	180	180	NA	-5.1
-80.0	100.1	-166.7	180.0	180.0	180.0	180.0	180.0	180.0	109.7	
-80	100	-166	180	180	180	180	180	180	NA	-4.9
-80.9	169.9	96.7	180.0	180.0	180.0	180.0	180.0	180.0	110.6	
-80	170	97	180	180	180	180	180	180	NA	-6.8

Table 22. Continued

Code	Ref	ϕ_g^1	ϕ_p^2	ϕ_g^3	ϕ_g^4	ϕ_p^5	ϕ_g^6	ψ_g^1	ψ_p^2	ψ_g^3
$C_2$10		60.0	-68.4	-70.4	60.0	-68.4	-70.4	-110.0	-50.4	76.8
	1	60	-68	-70	60	-68	-70	-110	-50	77
$C_2$11		149.9	-69.0	-100.1	149.9	-69.0	-100.1	-170.3	-60.1	76.3
	1	150	-68	-100	150	-68	-100	-170	-60	76
Al		-142.0	-72.0	-111.7	-152.3	-48.6	81.0	-173.0	-75.3	-10.2
	42	-142	-66	-115	-150	-53	83	-173	-36	-7

ψ^4_g	ψ^5_p	ψ^6_g	ω^1_g	ω^2_p	ω^3_g	ω^4_g	ω^5_p	ω^6_g	$\langle NC^{\alpha}C'$	TPE
-110.0	-50.4	76.8	180.0	180.0	180.0	180.0	180.0	180.0	109.9	
-110	-50	77	180	180	180	180	180	180	NA	-4.2
-170.3	-60.1	76.3	180.0	180.0	180.0	180.0	180.0	180.0	109.7	
-170	-60	76	180	180	180	180	180	180	NA	-3.7
179.2	117.5	2.3	-177.0	-174.0	-177.0	180.0	179.0	-175.0	112.9	
178	126	-3	-177	-174	-177	180	179	-175	112.6	NA

both energy minimized structures where the proline ϕ was always set to -68° (1). Ring closure allowed the proline ϕ to move so that the change in ψ_g^3 was most likely a compensation for this change. The only other substantial change was in the crystal structure A1; ψ_p^2 was changed by 39°. This change would in part be due to the ring closure problem mentioned earlier. Another contributing factor may have been that the crystal structure could not be determined by direct methods, but was derived from the crystal structure of cyclo(Gly-L-Pro-D-Ala)$_2$. Table 23 contains the proline ring parameters determined by BORC. The C_2 symmetric backbones all had the same proline ring parameters used for both rings in a given backbone because they all had the same ϕ,χ^2 pair. The asymmetric form A1 had different ϕ values, but the same χ^2 values (except when using the X-ray crystal χ^2 values) so that both proline rings could be slightly different.

π-π^* spectra calculated for the various forms of cyclo(Gly-Pro-Gly)$_2$ are given in Figures 26-28 and Table 24. Figures 26 and 27 show the absorption and CD spectra respectively predicted without the proline side chain, i.e., the backbone parameters were used to produce a cyclo(Gly)$_6$ structure. The experimental CD spectra were obtained from Reference 196. No experimental absorption spectra were available (193). The calculations in Figure 28 and Table 24 include the proline side chain. These figures and table reiterate the sensitivity of the dipole interaction model to changes in the backbone and side chain conformations. Some structures give a reasonably good overall fit to the π-π^* spectra for cyclo(Gly-Pro-Gly)$_2$.

Table 23. Proline ring parameters for cyclo(Gly-Pro-Gly)$_2$[a]

Code	ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_5^R	τ_6^R	NC^δ
$C_2$1	-73.1	-46.3	60.0	-49.1	21.6	99.7	97.7	96.0	98.7	110.2	122.3	1.51
Both		-28.2	40.0	-36.3	19.7	103.8	103.3	101.8	102.8	111.7	123.3	1.48
Prolines		-2.3	20.0	-30.0	30.5	104.6	106.3	104.7	103.0	111.6	125.0	1.46
		17.1	0.0	-17.3	29.9	103.9	106.4	106.3	104.4	110.3	123.5	1.46
		28.3	-20.0	4.0	14.8	103.5	105.3	106.0	105.6	111.3	123.2	1.47
		33.8	-40.0	30.5	-9.7	102.9	102.8	102.6	103.6	112.1	124.6	1.48
		48.6	-60.0	46.4	-17.0	98.8	97.4	96.7	98.9	110.9	124.9	1.51
$C_2$2	-30.8	-48.9	60.0	-46.1	16.5	98.8	97.3	96.7	100.0	110.7	124.6	1.51
Both		-36.9	40.0	-27.6	4.5	102.5	102.4	102.8	104.0	111.7	124.1	1.48
Prolines		-29.9	20.0	-2.1	-18.1	102.4	105.2	106.1	105.0	111.6	125.4	1.46
		-17.0	0.0	17.3	-30.1	103.2	106.6	106.3	103.8	111.1	125.8	1.45
		-1.4	-20.0	30.9	-32.0	104.6	106.3	104.6	102.8	111.2	124.6	1.46
		21.2	-40.0	43.6	-32.1	104.5	103.3	100.7	101.2	110.1	121.9	1.47
		39.2	-60.0	56.8	-34.2	101.8	97.8	94.4	96.7	108.2	118.3	1.51
$C_2$3	-80.2	-44.5	60.0	-51.0	24.9	100.2	98.0	95.5	98.3	110.0	121.0	1.51
Both		-26.2	40.0	-38.4	23.3	104.0	103.4	101.5	102.4	111.4	122.9	1.47
Prolines		-2.6	20.0	-29.7	30.0	104.6	106.4	104.8	103.0	111.8	125.2	1.46
		17.9	0.0	-18.1	31.4	103.4	106.4	106.1	103.9	110.4	124.8	1.46
		29.8	-20.0	2.3	17.7	102.8	105.1	106.0	105.3	111.1	124.2	1.47
		34.2	-40.0	30.1	-9.0	102.8	102.8	102.7	103.6	112.2	124.8	1.48
		48.7	-60.0	46.1	-16.6	98.6	97.4	96.8	98.8	111.1	125.6	1.51

[a]All angles are in degrees. The NC^δ bond lengths are in Å.

Table 23. Continued

Code	ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_5^R	τ_6^R	NC^δ
$C_2$4	-63.6	-48.1	60.0	-47.0	18.1	99.1	97.4	96.5	98.9	110.6	124.0	1.51
Both		-31.6	40.0	-32.8	13.7	103.4	103.0	102.3	103.4	112.0	123.9	1.48
Prolines		-3.4	20.0	-29.0	28.5	105.1	106.3	104.8	103.6	111.4	123.7	1.46
		-12.8	0.0	12.9	-22.3	105.0	107.0	106.3	105.6	111.2	121.4	1.46
		19.1	-10.0(-)	-3.0	15.9	105.0	106.4	106.6	106.1	111.4	122.0	1.47
		-5.6	-10.0(+)	21.9	-26.9	105.1	106.9	105.6	104.7	111.1	122.0	1.46
		31.7	-40.0	32.8	-13.6	103.4	103.0	102.4	103.4	112.0	124.0	1.48
		47.5	-60.0	47.8	-19.2	99.4	97.4	96.4	98.8	110.4	123.5	1.51
$C_2$5	-85.5	-43.6	60.0	-52.0	26.6	100.4	98.0	95.3	98.1	109.7	120.5	1.51
Both		-24.9	40.0	-39.7	25.6	104.1	103.4	101.3	102.1	111.2	122.8	1.47
Prolines		-2.6	20.0	-29.7	29.9	104.6	106.4	104.8	103.1	111.8	125.2	1.46
		18.0	0.0	-18.2	31.6	103.2	106.4	106.4	106.1	110.6	125.2	1.45
		30.4	-20.0	1.7	18.7	102.5	105.1	106.0	105.1	111.2	124.6	1.47
		34.4	-40.0	29.9	-8.6	102.7	102.8	102.7	103.6	112.2	124.8	1.48
		48.5	-60.0	46.2	-16.8	98.6	97.4	96.8	98.7	111.2	125.8	1.51
$C_2$6	-68.2	-47.4	60.0	-47.9	19.5	99.4	97.5	96.3	98.8	110.4	123.2	1.51
$C_2$8	-67.9	-30.0	40.0	-34.5	16.5	103.6	103.2	102.1	103.1	111.9	123.6	1.48
$C_2$10	-68.4	-2.6	20.0	-29.8	30.0	104.8	106.3	104.7	103.2	111.4	124.4	1.46
		15.7	0.0	-15.8	27.3	104.5	106.5	106.2	104.9	110.5	122.6	1.46
		25.8	-20.0(-)	6.6	10.3	104.2	105.5	106.0	105.9	111.7	122.8	1.47
		6.6	-20.0(+)	25.8	-22.9	105.4	106.4	105.0	104.5	112.0	122.9	1.46
		32.9	-40.0	31.5	-11.4	103.1	102.9	102.5	103.5	112.0	124.3	1.48
		48.1	-60.0	47.0	-17.9	99.1	97.4	96.6	98.9	110.6	124.2	1.51
$C_2$7	-108.4	-34.2	60.0	-60.7	42.7	100.6	99.4	92.7	94.6	108.1	115.1	1.51
		-15.2	40.0	-49.4	42.5	103.7	103.6	99.3	98.9	108.7	119.8	1.47
		2.7	20.0	-35.2	39.3	104.0	106.0	104.8	101.7	109.4	123.0	1.45

Table 23. Continued

Code	ϕ	χ^1	χ^2	χ^3	χ^4	τ_1^R	τ_2^R	τ_3^R	τ_4^R	τ_5^R	τ_6^R	NC[δ]
		18.8	0.0	-19.0	32.9	103.1	106.3	106.0	103.6	110.1	124.6	1.45
		31.8	-20.0	0.2	21.3	102.1	104.9	106.0	104.9	110.7	124.5	1.46
		41.6	-40.0	22.9	3.6	101.3	101.8	103.0	104.3	110.8	123.1	1.48
		51.2	-60.0	44.2	-12.9	98.5	96.9	96.8	99.7	110.2	122.9	1.51
$C_2$9	-67.2	-47.5	60.0	-47.8	19.3	99.4	97.5	96.3	98.9	110.4	123.3	1.51
		-30.2	40.0	-34.3	16.2	103.6	103.1	102.1	103.2	111.9	123.7	1.48
		-2.6	20.0	-29.8	30.0	104.8	106.3	104.7	103.2	111.4	124.3	1.46
		15.5	0.0	-15.7	27.0	104.6	106.5	106.2	104.9	110.5	122.5	1.46
		6.3	-20.0(+)	26.1	-23.5	105.4	106.4	105.0	104.4	111.9	122.9	1.46
		48.1	-60.0	47.1	-18.1	99.2	97.4	96.5	98.9	110.6	124.1	1.51
Al	-72.0	-46.3	60.0	-49.1	21.6	99.7	97.7	96.0	98.7	110.2	122.2	1.51
Pro(2)		-28.1	40.0[b]	-36.5	20.0	103.8	103.3	101.8	102.8	111.7	123.3	1.48
		-26.4	38.4[b]	-35.5	20.1	104.0	103.6	102.2	103.0	111.8	123.4	1.47
		-2.3	20.0	-30.0	30.4	104.6	106.3	104.7	103.0	111.6	125.0	1.46
		17.2	0.0	-17.4	30.0	104.9	106.4	106.1	104.3	110.3	123.6	1.46
		28.4	-20.0	3.9	15.0	103.5	105.2	106.0	105.6	111.3	123.3	1.47
		33.9	-40.0	30.5	-9.6	102.9	102.8	102.6	103.6	112.1	124.6	1.48
		41.2	-50.0	38.8	-13.7	101.2	100.4	100.0	101.6	111.5	124.9	1.49
Al	-48.6	-48.5	60.0[b]	-46.2	16.8	98.6	97.4	96.8	98.7	111.2	125.7	1.51
Pro(5)		-35.8	42.4[b]	-32.4	10.4	102.4	102.3	102.1	103.1	112.1	124.9	1.48
		-34.2	40.0	-30.2	9.1	102.8	102.8	102.7	103.4	112.2	124.8	1.48
		-29.8	20.0	-2.4	-17.5	102.9	105.2	106.0	105.3	111.1	123.6	1.47
		-17.8	0.0	18.0	-31.2	103.5	106.4	106.0	104.0	110.3	124.0	1.46
		2.4	-20.0	29.9	-30.4	104.6	106.3	104.7	103.0	111.7	125.1	1.46
		27.3	-40.0	37.3	-21.4	104.1	103.2	101.8	102.7	111.5	123.3	1.48
		36.4	-50.0	44.2	-22.6	102.7	100.6	99.0	101.0	110.6	123.4	1.49

Figure 26. π-π^* absorption spectra for cyclo(Gly-Pro-Gly)$_2$ treated as cyclo(Gly)$_6$

Units for the scales are as follows: nm for λ and 10^3 Lmol^{-1}cm^{-1} for ε. Each spectrum is labeled by the backbone code. The scale listed for $C_2$10 is the same for all other forms. Bandwidths are 6000 cm^{-1} for all cases.

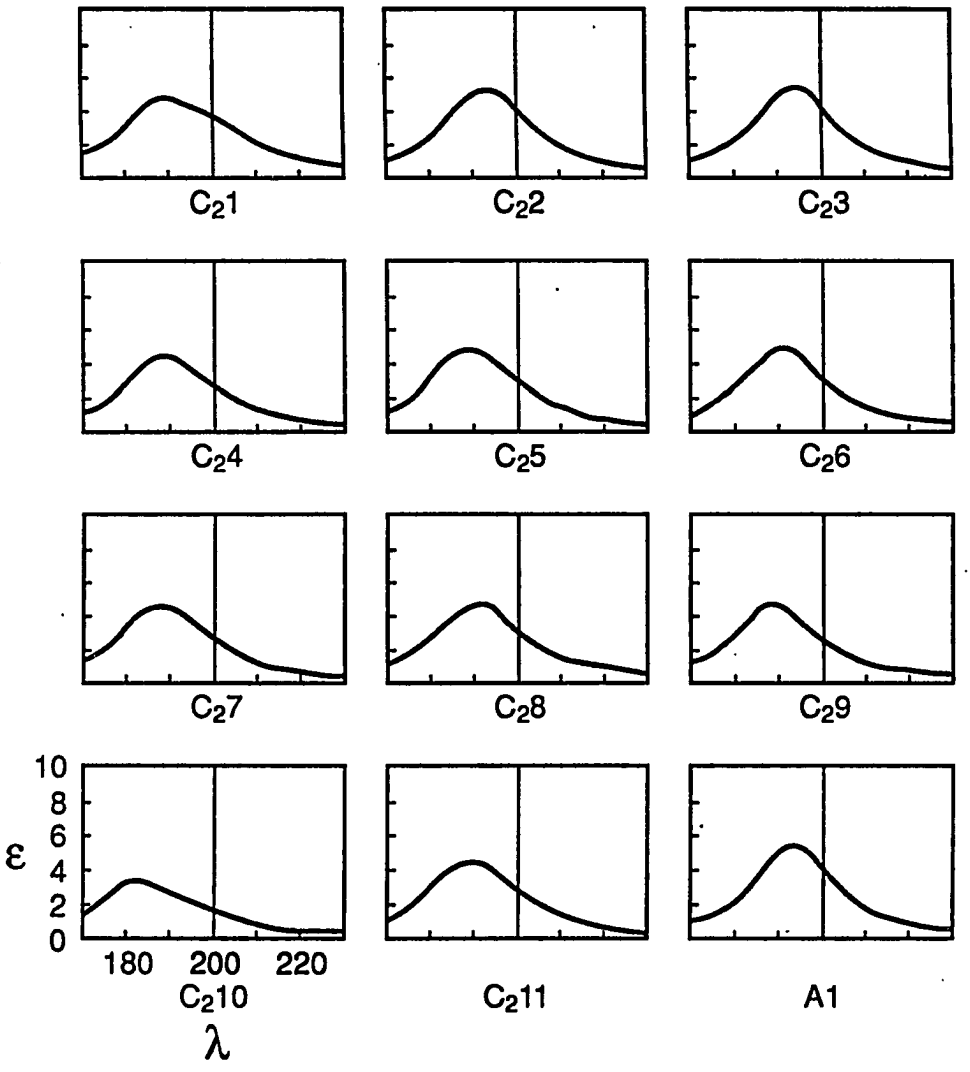

C21
C22
C23
C24
C25
C26
C27
C28
C29
10
8
6
4
2
0
ε
180
200
220
C210
λ
C211
A1

Figure 27. $\pi-\pi^*$ CD spectra for cyclo(Gly-Pro-Gly)$_2$ treated as cyclo(Gly)$_6$

Units for the scales are as follows: nm for λ and $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$. Each spectrum is labeled by the backbone code. The scale listed for $C_2 10$ is the same for all other forms. Bandwidths are 6000 cm^{-1} for all cases.

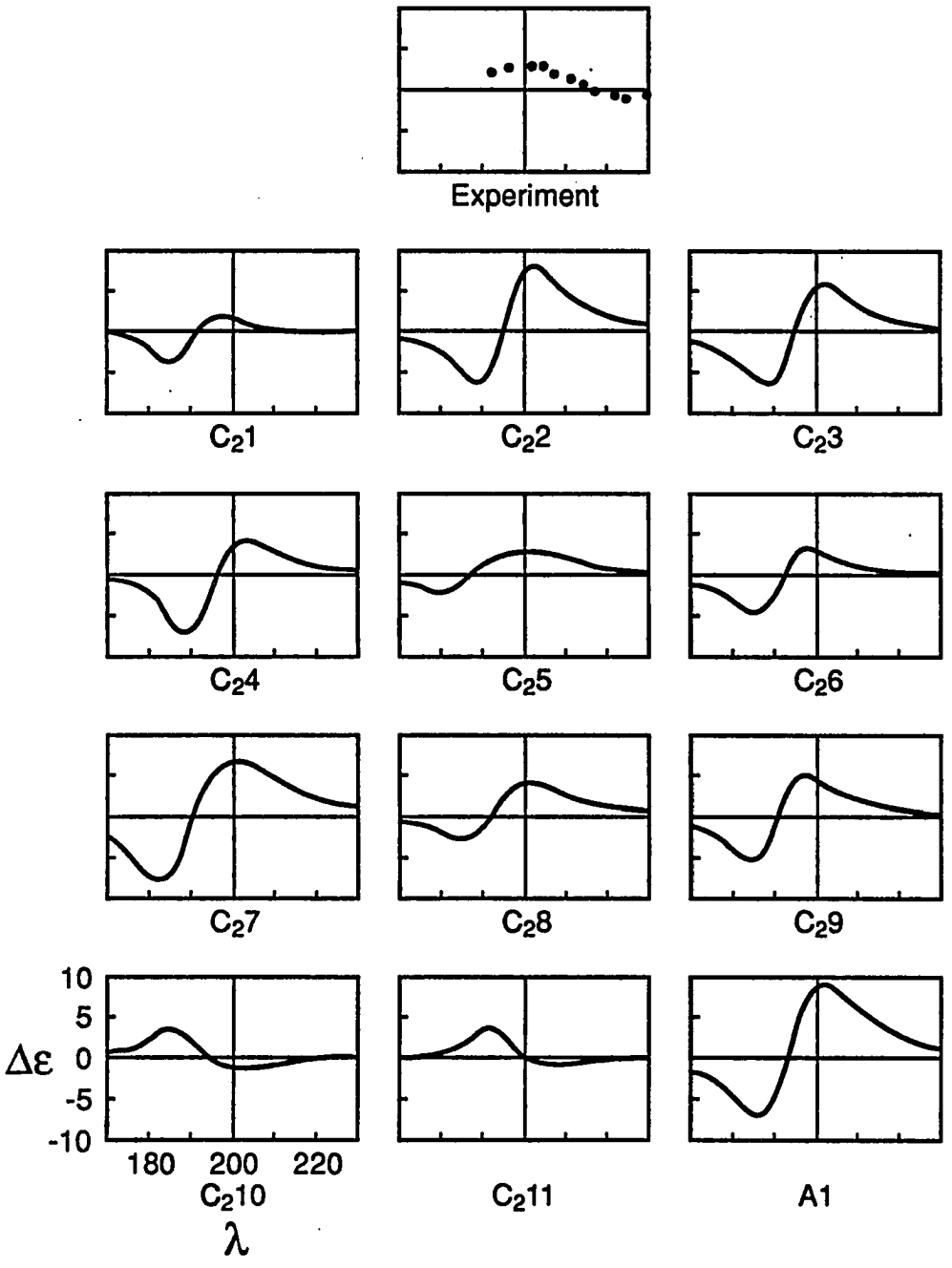

Experiment
C21
C22
C23
C24
C25
C26
C27
C28
C29
Δε
10
5
0
-5
-10
180
200
220
C210
λ
C211
A1

Figure 28. π-π^* absorption and CD spectra for cyclo(Gly-Pro-Gly)$_2$ backbone $C_2 5$

Units for the scales are as follows: nm for λ, $Lmol^{-1}cm^{-1}$ for $\Delta\varepsilon$, and 10^3 $Lmol^{-1}cm^{-1}$ for ε. Each ε, $\Delta\varepsilon$ pair is labeled according to χ^2 values. All spectra are to the scale listed for the lower left pair. The bandwidth is 6000 cm^{-1}.

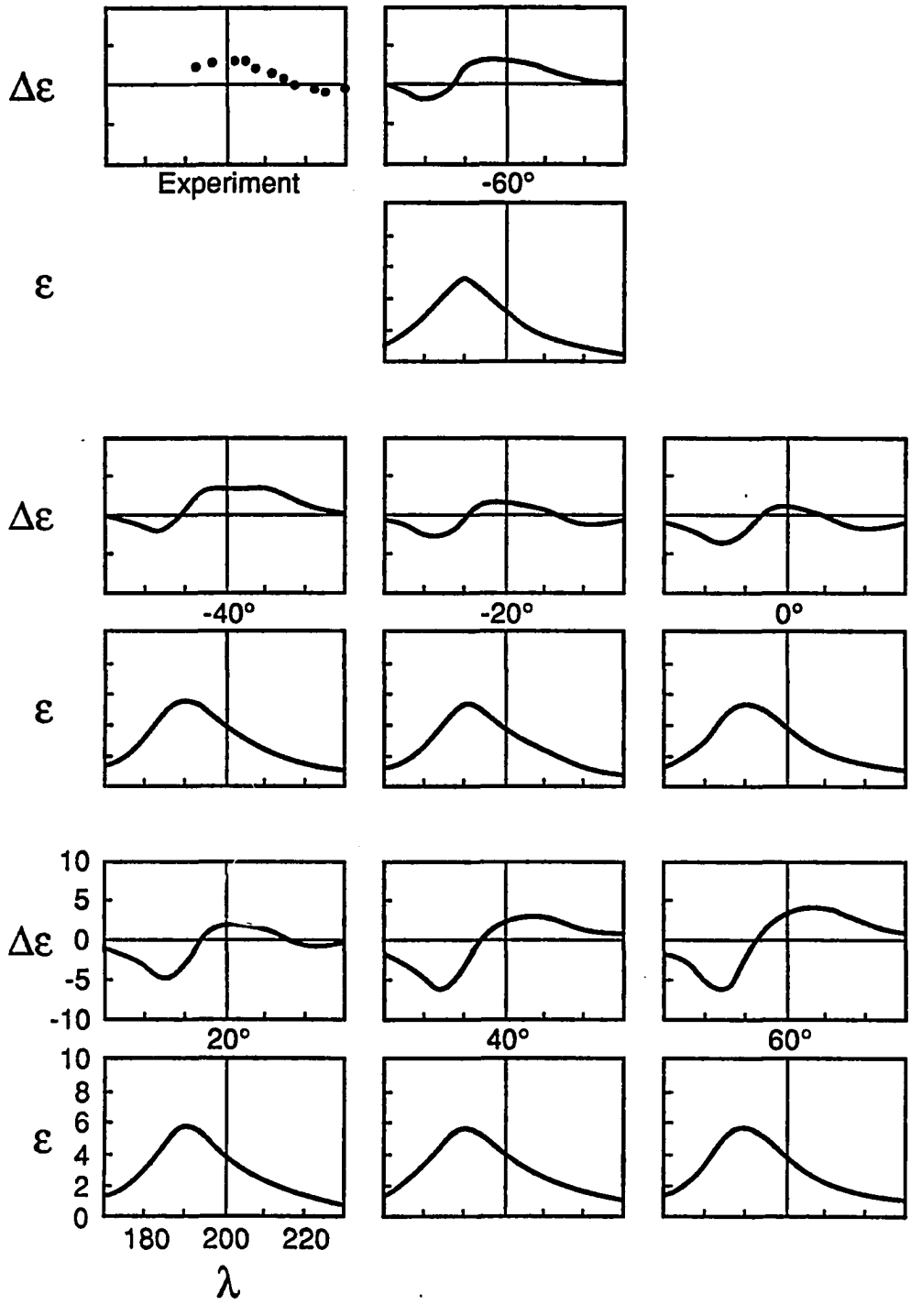

Δε
ε
Experiment
-60°
-40°
-20°
0°
20°
40°
60°
10
5
0
-5
-10
10
8
6
4
2
0
180
200
220
λ

Table 24. π-π* predicted maxima and minima for cyclo(Gly-Pro-Gly)$_2$[a]

Form	χ^2	λ_{max}	ε_{max}	λ_1	$\Delta\varepsilon_1$	λ_2	$\Delta\varepsilon_2$	λ_3	$\Delta\varepsilon_3$
$C_2$1	60.0	190	5331	-	-	188	-7.3	208	3.6
	40.0	190	5318	-	-	188	-8.0	212	3.7
	20.0	190	5277	-	-	190	-6.8	214	2.1
	0.0	190	5254	176	0.3	190	-5.6	214	2.1
	-20.0	190	5264	176	0.8	190	-4.4	212	3.9
	-40.0	190	5312	178	1.3	188	-3.3	210	5.5
	-60.0	188	5225	178	1.6	188	-2.4	204	3.4
$C_2$2	60.0	196	5938	-	-	188	-5.5	204	6.8
	40.0	198	5487	-	-	188	-4.3	202	8.3
	20.0	196	5232	-	-	188	-2.9	204	11.4
	0.0	196	5197	154	0.2	188	-2.4	202	12.8
	-20.0	196	5836	156	0.2	188	-2.6	202	12.2
	-40.0	196	5576	152	0.2	188	-3.0	204	13.5
	-60.0	196	5930	152	0.2	188	-3.0	204	12.3
$C_2$3	60.0	196	6137	-	-	190	-10.0	206	4.3
	40.0	196	5768	-	-	192	-10.5	208	3.3
	20.0	196	5736	-	-	192	-10.9	210	4.0
	0.0	196	5708	-	-	192	-10.7	212	5.0
	-20.0	196	5738	-	-	192	-9.9	210	7.6
	-40.0	196	6002	-	-	190	-8.6	208	11.0
	-60.0	194	2227	-	-	190	-7.1	204	11.4
$C_2$4	60.0	190	5044	-	-	190	-12.0	206	8.2
	40.0	192	5005	-	-	192	-13.2	210	9.4
	20.0	190	4999	-	-	192	-12.5	212	7.7
	0.0	192	4916	-	-	192	-11.8	214	15.3
	-10.0(-)	190	4975	-	-	192	-10.8	212	9.4
	-10.0(+)	192	4936	-	-	192	-11.2	212	15.4
	-40.0	190	5004	-	-	192	-8.5	210	11.5
	-60.0	190	4878	-	-	190	-7.5	206	9.5
$C_2$6	60.0	194	5576	188	-6.8	200	4.0	-	-
	40.0	194	5537	188	-7.1	200	3.7	-	-
	20.0	194	5432	188	-5.5	200	4.1	216	-0.9

[a] χ^2 are in degrees. λ, λ_1, λ_2, and λ_3 are in nm. ε and $\Delta\varepsilon$ are in L mol^{-1} cm^{-1}. The bandwidth is 6000 cm^{-1}.

Table 24. Continued

Form	χ^2	λ_{max}	ε_{max}	λ_1	$\Delta\varepsilon_1$	λ_2	$\Delta\varepsilon_2$	λ_3	$\Delta\varepsilon_3$
$C_2$6 Cont.	0.0	194	5384	188	-4.3	200	5.2	216	-1.1
	-20.0(-)	194	5417	186	-3.5	200	7.0	-	-
	-20.0(+)	194	5471	188	-4.6	202	7.6	-	-
	-40.0	194	5493	186	-2.7	200	8.6	-	-
	-60.0	192	5453	184	-1.5	198	8.7	-	-
$C_2$7	60.0	186	4212	184	-10.0	204	3.7	-	-
	40.0	186	4186	184	-10.6	204	1.8	218	-1.2
	20.0	186	4306	184	-11.6	204	1.5	216	-1.4
	0.0	186	4287	184	-11.7	204	1.1	216	-2.8
	-20.0	188	5173	184	-13.2	200	4.5	-	-
	-40.0	186	4458	184	-10.6	204	6.6	220	-0.7
	-60.0	188	4684	184	-8.9	200	9.5	-	-
$C_2$8	60.0	192	5196	188	-5.3	204	6.7	-	-
	40.0	192	5275	188	-5.4	208	6.5	-	-
	20.0	192	5294	188	-3.3	210	4.0	-	-
	0.0	192	5262	188	-2.07	210	3.5	-	-
	-20.0(-)	192	5252	186	-1.6	210	5.7	-	-
	-20.0(+)	192	5284	188	-4.0	210	10.6	-	-
	-40.0	192	5220	186	-2.0	206	8.1	-	-
	-60.0	192	4924	186	-1.8	202	7.9	-	-
$C_2$9	60.0	190	4404	186	-9.8	198	2.8	212	-.9
	40.0	192	4218	186	-9.7	198	2.8	212	-2.5
	20.0	192	4319	186	-9.8	200	1.9	212	-3.3
	0.0	192	4432	188	-9.1	200	2.7	212	-2.6
	-20.0(+)	192	4206	186	-9.2	200	6.9	-	-
	-60.0	192	5006	186	-14.0	206	7.5	-	-
$C_2$10	60.0	186	2905	-	-	186	4.5	204	-3.2
	40.0	184	2795	174	-.2	188	2.7	204	-3.1
	20.0	186	2766	178	-2.1	190	0.5	202	-2.9
	0.0	186	2850	180	-2.2	190	0.3	202	-2.4
	-20.0(-)	186	2891	176	-.5	188	2.7	200, 214	-1.0, 1.3
	-20.0(+)	184	2757	188	3.7	200	-.5	212	2.0
	-40.0	186	2927	188	6.2	-	-	-	-
	-60.0	192	3069	192	9.7	208	-1.1	250[b]	0.6[b]

[b]Last available point.

Table 24. Continued

Form	χ^2	λ_{max}	ε_{max}	λ_1	$\Delta\varepsilon_1$	λ_2	$\Delta\varepsilon_2$	λ_3	$\Delta\varepsilon_3$
A1	60.0	196	5623	-	-	188	-9.2	204	10.9
	40.0	198	5505	-	-	188	-9.5	206	10.7
	38.4[c], 42.4[c]	198	5498	-	-	188	-9.5	206	10.5
	20.0	198	5356	-	-	188	-8.8	206	10.7
	0.0	198	5515	-	-	188	-7.9	206	11.0
	-20.0	198	5375	-	-	188	-6.5	206	12.5
	-40.0	196	5426	-	-	188	-6.3	206	16.1
	-50.0	196	5270	-	-	188	-6.8	206	12.4
Experiment	NA	NA	NA	192[b]	2.0[b]	202	3.3	225	-0.8

[c]X-ray value of χ^2.

D. Discussion

Predicted π-π^* absorption spectra for the backbones of cyclo(Gly-Pro-Gly)$_2$ can be seen in Figure 26. All the predicted spectra are broad and show minor variations. Most have a peak near 190 nm that varies in intensity from 4500 cm^{-1} to 5700 cm^{-1}. The only exception is $C_2$10 which has a weaker spectrum (around 3500 cm^{-1}) peaking around 184 nm.

The predicted π-π^* CD spectra, however, show more variations than the absorption spectra (Figure 27). The experimental CD spectrum is very broad with a positive peak around 200 nm and a weak negative peak just above 220 nm. Several of the backbone spectra show a positive peak in the region of 200 nm. These include $C_2$1, $C_2$2, $C_2$3, $C_2$4, $C_2$5, $C_2$6, $C_2$7, $C_2$8, $C_2$9, and A1. None of these showed a negative band near 220 nm. Two backbones showed negative peaks in the 200 nm region, $C_2$10 and $C_2$11, which may indicate that these two structures are not representative of cyclo(Gly-Pro-Gly)$_2$ in methanol.

When the side chain was added the predicted absorption intensified and the CD red shifted (Table 24). The CD predicted for $C_2$1 was of similar intensity and sign to experiment for the positive band, but the positive band occurs red of experiment (near 210 nm). For $\chi^2 \leq 0°$ another positive band appears at 176 nm, but there is no experiment in this region for comparison. No small negative band appears anywhere near 220 nm. Thus, $C_2$1 cannot account for the CD spectrum alone. It could contribute as part of a mixture.

$C_2$2 produces CD predictions that are more intense than experiment (Table 24), but the overall shape of the spectrum is similar to that of

$C_2$1. $C_2$2 CD predictions are bluer than those of $C_2$1 (around 204 nm for the intense positive band). The bluer spectrum, although more intense, better resembles experiment in location. Again no negative band appears around 220 nm like experiment. Thus, $C_2$2 may also contribute as part of the mixture.

The CD predictions for $C_2$3 had the same sign and shape as the two preceding forms (Table 24). The intensities and location of the positive band were between the two previous forms. The 190 nm negative band is more intense. This band is too blue of experimental data for comparison. Thus, $C_2$3 may also contribute to the CD spectrum of cyclo(Gly-Pro-Gly)$_2$ as part of a mixture of structures.

The CD predictions for $C_2$4 (Table 24) are very intense compared to the other predictions; however, the red shift is not any greater than for $C_2$1. $C_2$4 also shows a strong negative band at 192 nm similar to that of $C_2$3. There still is no weak negative CD band near 220 nm. Thus, $C_2$4 may be considered part of the mixture just as $C_2$1, $C_2$2, and $C_2$3.

The CD predictions for $C_2$5 are the best of this series (Figure 28). Not only is the positive band at 200 nm broad and weak like experiment, but also a weak negative band appears at 220 nm when $-20° \leq \chi^2 \leq 20°$. Thus, $C_2$5 is a likely representative structure of cyclo(Gly-Pro-Gly)$_2$ in methanol.

The CD predictions for $C_2$6 are similar to those for $C_2$5 except that they are more intense (Table 24). Moreover, the weak negative band appears slightly bluer, 216 nm, and for $0° \leq \chi^2 \leq 20°$. The strong negative band

around 188 nm is also bluer for $C_2$6. Thus, $C_2$6 may also be a major contributor to the conformation mixture.

$C_2$7 had CD predictions similar to $C_2$5 (Table 24). The greatest difference between the two is out of the region of experiment; the negative band at 184 nm is much more intense for $C_2$7 than for $C_2$5. $C_2$7 also produces a weak negative band on or past 216 nm, but it has a wide range of χ^2 values ($40° \geq \chi^2 \geq 0°, -40°$). Thus, just as $C_2$5 was probably a major contributor to the conformation mixture, so may $C_2$7 be a major contributor.

$C_2$8 produces intense CD predictions that resemble those for $C_2$1, $C_2$2, $C_2$3, and $C_2$4 (Table 24). The predictions are generally more intense and red than experiment. No small negative band appears past 220 nm. Thus, $C_2$8 may only be a minor contributing conformation for cyclo(Gly-Pro-Gly)$_2$ in solution.

$C_2$9 produces a weak positive band like experiment near 200 nm (Table 24). It also has a weak negative band around 212 nm; this is blue of experiment. This band only occurs for positive values of χ^2. Therefore, $C_2$9 is a contributing structure to the mixture that probably contributes more than $C_2$1 through $C_2$4, but less than $C_2$5 through $C_2$7.

$C_2$10 and $C_2$11 were the highest energy structures produced by energy minimizations (-4.2 and -3.7 kcal/mole, respectively (1)). They both produced backbone CD spectra that were opposite in sign to experiment (Figure 27). When the proline side chain was added to $C_2$10, the sign of the predicted CD did not change (Table 24). The resulting CD predictions were merely more intense when proline was included. Since $C_2$11 was higher in energy than $C_2$10 and since its backbone spectrum resembled that for

$C_2$10, it was not deemed necessary to pursue $C_2$11. Thus, the two structures $C_2$10 and $C_2$11 probably do not contribute to the possible conformations in solution.

The X-ray structure, A1, produced CD that resembled the predictions for $C_2$2 except that they were more intense (Table 24). The predictions were considerably more intense than experiment. Moreover, no small negative band was seen near 220 nm. Thus, A1 may only be a minor contributor to the solution pot.

E. Conclusions

The principal finding of this study is that approximately correct π-π^* CD spectra for cyclo(Gly-Pro-Gly)$_2$ are predicted by the dipole interaction model when the proline side chain is included on certain predetermined backbone structures. The solution is probably a mixture of structures which include minor contributions from $C_2$1, $C_2$2, $C_2$3, $C_2$4, $C_2$8, and A1 with major contributions from $C_2$5, $C_2$6, $C_2$7, and $C_2$9. Many of these structures contain β-turns. $C_2$1, $C_2$2, $C_2$3, and $C_2$4 all contain type II β-turns (1, 39-41). A1 contains one type II and one type I β-turn (42). $C_2$1 also has a minor contribution for a γ-turn (39). Thus, type II β-turns, type I β-turns, and γ-turns contribute to the solution mixture. Type II' β-turns were eliminated with the $C_2$10 conformation; this agrees with the NMR results (39-41). The experimental and successful by predicted CD spectra resemble those spectra for other molecules with β-turns (6, 39). Moreover, earlier CD predictions by the dipole interaction model also suggested that a mixture of structures was present in solutions for other β-turn models (Ac-L-Ala-Gly-NHMe and cyclo(L-Ala-Gly-Aca)) (6).

Although NMR results concluded that cyclo(Gly-Pro-Gly)$_2$ was C_2 symmetric (and thus the energy minimizations were done for the C_2 forms), asymmetric structures might also occur. The solution mixture contains a variety of structures, and the crystal structure was asymmetric with two different β-turns in it (42). Thus, a search of asymmetric structures containing type II β-turns, type I β-turns or γ-turns in various combinations might be useful.

The conformation that produced CD most like experiment was $C_2$5. This structure has its carbonyls arranged for favorable binding of cations (1). This may be the reason why the CD spectra of various cation complexes of cyclo(Gly-Pro-Gly)$_2$ only show minor variations from the uncomplexed CD spectrum (complexes produce CD that are 2 to 6 nm blue shifted to the uncomplexed form (195)).

Some information about the proline ring puckering was gained by this study. The major contributors $C_2$5, $C_2$6, $C_2$7, and $C_2$9 generally preferred $\chi^2 \geq 0°$. The X-ray structure also had χ^2 values positive around 40° (42). Thus, for cyclo(Gly-Pro-Gly)$_2$ in methanol the proline ring favors $\chi^2 \geq 0°$.

VIII. CONCLUSIONS

A. Overall Quality of the Dipole Interaction Model

There was much previous evidence that the dipole interaction model could predict the main features of π-π^* absorption and circular dichroism spectra for polypeptides (2, 3, 6, 8, 10-12). Results for the current calculations (Chapters IV-VII) also included π-π^* absorption and CD spectra that resembled experiment for poly(L-Proline) I and II, cyclo(Pro-Gly)$_3$ complexed and uncomplexed, and cyclo(Gly-Pro-Gly)$_2$. Thus, spectral properties in the π-π^* region (180-220 nm) can be understood in terms of the dipole interaction model.

B. Sensitivity of the Dipole Interaction Model to Structure

The dipole interaction predictions are extremely sensitive to the structures of polypeptides. This sensitivity was useful for determining not only the backbone structures of cyclo(Pro-Gly)$_3$ and cyclo(Gly-Pro-Gly)$_2$, but occasionally the proline ring puckering. The comparison between theoretical and experimental CD spectra suggest that the cation complexes of cyclo(Pro-Gly)$_3$ are C_3 symmetric and have all trans ($\omega = 180°$) peptide bonds. The uncomplexed forms of cyclo(Pro-Gly)$_3$ are only C_3 symmetric in nonpolar solvents. In polar solvents, there may be a mixture of structures whose major contributors have one cis peptide bond for a glycine residue. Cyclo(Gly-Pro-Gly)$_2$ may also be a mixture of structures several of which contain type II β-turns, γ-turns, and type I β-turns. All structures have all trans peptide bonds, but they may not necessarily be symmetric. Some

cyclo(Gly-Pro-Gly)$_2$ backbones have their carbonyls arranged for favorable binding of cations.

C. Proline Ring Puckering

The proline ring puckering implied using the dipole interaction model supports the structures of proline seen in the literature. The bimodal distribution of χ^2 torsion angles in proline is reflected in those backbones which have rapidly interconverting proline rings like poly(L-proline) I and II and the uncomplexed forms of cyclo(Pro-Gly)$_3$ in nonpolar solvents. Predicted CD spectra for these three molecules are consistent with experiment when the proline ring is puckered $\chi^2 = \pm 40°$. Moreover, the bimodal distribution is also reflected in backbones that prefer one kind of puckering over another because there are many backbone in the literature that pucker preferentially $\chi^2 > 0°$, and there are many other backbones that pucker $\chi^2 < 0°$. The cation complex backbones of cyclo(Pro-Gly)$_3$ with carbonyls pointing to opposite sides of the backbone ring prefer $\chi^2 < 0°$; those cation complex backbones with the carbonyls pointing only to one side of the backbone ring prefer $\chi^2 > 0°$. The uncomplexed forms of cyclo(Pro-Gly)$_3$ in polar solvents favor $\chi^2 < 0°$; whereas, cyclo(Gly-Pro-Gly)$_2$ backbones favor $\chi^2 \geq 0°$. Thus, for most molecules in this study (and perhaps for most in general) proline tends to be intensely puckered, but the direction of puckering is backbone dependent.

The bimodal distribution in proline ring puckering is strongly evident in the crystal literature (Figure 4). The modes for χ^2 (-36°, 39°) agree with the intense puckering implied by the dipole interaction predictions

mentioned above. Moreover, NMR results on poly(L-proline) II also indicated rapid interconversion of proline ring puckering around $\chi^2 = \pm 35°$ (185) which resembles the $\pm 40°$ mentioned above.

IX. BIBLIOGRAPHY

1. Deber, C.; Madison, V.; Blout, E. Acc. Chem. Res. 1976, 9, 106-113.

2. Applequist, J. Biopolymers 1981, 20, 2311-2322.

3. Sathyanarayana, B.; Applequist, J. Int. J. Peptide Protein Res. 1985, 26, 518-527.

4. Applequist, J.; Sundberg, K.; Olson, M.; Weiss, L. J. Chem. Phys. 1979, 70, 1240-1246.

5. Applequist, J.; Sundberg, K.; Olson, M.; Weiss, L. J. Chem. Phys. 1971, 71, 2330.

6. Sathyanarayana, B.; Applequist, J. Int. J. Peptide Protein Res. 1986, 27, 86-94.

7. Applequist, J. Biopolymers 1981, 20, 387-397.

8. Applequist, J. J. Chem. Phys. 1979, 71, 4332-4338.

9. Applequist, J. J. Chem. Phys. 1980, 73, 3521.

10. Applequist, J. Biopolymers 1982, 21, 779-795.

11. Caldwell, J.; Applequist, J. Biopolymers 1984, 23, 1891-1904.

12. Applequist, J. Biopolymers 1982, 21, 703-704.

13. Anteunis, M.; Sleeckx, J. In Molecular Structure and Energetics, Vol. 4, Leibiman, J.; Greenberg, A., Eds.; VCH Publishers: New York, New York, 1987; Chapter 7.

14. Sarkar, S.; Young, P.; Torchia, D. J. Am. Chem. Soc. 1986, 108, 6459-6464.

15. Kessler, H. Agnew. Chem. Int. Ed. Engl. 1982, 21, 512-523.

16. de Leeuw, F.; Altona, C.; Kessler, H.; Bermel, W.; Friedrich, A.; Krack, G.; Hull, W. J. Am. Chem. Soc. 1983, 105, 2237-2246.

17. Sankaram, M.; Easwaran, K. Biopolymers 1982, 21, 1557-1567.

18. Orchinnikov, Yu.; Ivanov, V. Tetrahedron 1975, 31, 2177-2209.

19. Tamaki, M.; Okitsu, T.; Araki, M.; Sakamoto, H.; Takimoto, M.; Muramatsu, I. Bull Chem. Soc. Jpn. 1985, 58, 531-535.

20. Ando, S.; Nishikawa, H.; Takiguchi, H.; Izumiya, N. Bull. Chem. Soc. Jpn. 1986, 59, 1201-1206.

21. Rich, D.; Bhatnagar, P. J. Am. Chem. Soc. 1978, 100, 2212-2218.

22. Swepston, P.; Cordes, A.; Kuyper, L.; Meyer, W. Acta Crystallogr. 1981, B37, 1139-1141.

23. Snyder, J. J. Am. Chem. Soc. 1984, 106, 2393-2400.

24. Yoshioka, H.; Nakatsu, K.; Sato, M.; Tatsuno, T. Chem. Lett. 1973, 1319-1322.

25. Norrestam, R.; Stensland, B.; Branden, C. J. Mol. Biol. 1975, 99, 501-506.

26. Jalal, M.; Mocharla, R.; Barnes, C.; Hossain, M.; Powell, D.; Benson, B.; vander Helm, D. In Peptides Structure and Function Proceedings of the 8th American Peptide Symposium; Hruby, V.; Rich, D., Eds.; Pierce Chemical Company: Rockford, Illinois, 1983; 503-506.

27. Blout, E. Biopolymers 1981, 20, 1901-1912.

28. Ravi, A.; Balaram, P. Biochim. Biophys. Acta 1983, 745, 301-309.

29. Freidinger, R.; Perlow, D.; Randall, W.; Saperstein, R.; Arison, B.; Veber, D. Int. J. Peptide Protein Res. 1984, 23, 142-150.

30. Go, N.; Scheraga, H. Macromolecules 1973, 6, 525-535.

31. Go, N.; Scheraga, H. Macromolecules 1978, 11, 552-559.

32. Deber, C.; Blout, E. Israel J. Chem. 1974, 12, 15-29.

33. Madison, V.; Atreyi, M.; Deber, C.; Blout, E. J. Am. Chem. Soc. 1974, 96, 6725-6734.

34. Kartha, G.; Varughese, K.; Aimoto, S. Proc. Natl. Acad. Sci. USA 1982, 79, 4519-4522.

35. Varughese, K.; Aimoto, S.; Kartha, G. Int. J. Peptide Protein Res. 1986, 27, 118-122.

36. Hori, K.; Kushich, J.; Weinstein, H. Biopolymers 1988, 27, 1865-1886.

37. Kartha, G.; Aimoto, S.; Varughese, K. Int. J. Peptide Protein Res. 1986, 27, 112-117.

38. Madison, V. Biopolymers 1973, 12, 1837-1852.

39. Gierasch, L.; Deber, C.; Madison, V.; Niu, C.; Blout, E. Biochemistry 1981, 20, 4730-4738.

40. Schwyzer, R.; Grathwohl, Ch.; Meraldi, J.; Tun-kyi, A.; Vogel, R.; Wuthrich, K. Helvetica Chimica Acta 1972, 55(7), 253-254, pp 2545-2549.

41. Pease, L.; Deber, C.; Blout, E. J. Am. Chem. Soc. 1973, 95, 258-260.

42. Kostansek, E.; Thiessen, W.; Schomburg, D.; Lipscomb, W. J. Am. Chem. Soc. 1979, 101(19), 5811-5815.

43. Ramachandran, G.; Sasisekharan, V. Advances in Protein Chem. 1968, 23, 283-438.

44. Applequist, J., Biochemistry and Biophysics, Iowa State University, Ames, IA.

45. Go, N.; Scheraga, H. Macromolecules 1970, 3, 178-187.

46. Pauling, L. In The Nature of the Chemical Bond, 3rd ed.; Cornell University Press: Ithaca, New York, 1960; p. 226.

47. Thomasson, K.; Applequist, J. Biopolymers, accepted for publication.

48. Millane, R.; Nzewi, E. J. Appl. Crystallogr. 1989, 22, 138-143.

49. Altona, C.; Geise, H.; Romers, C. Tetrahedron 1968, 24, 13-32.

50. Applequist, J.; Carl, J.; Fung, K. J. Am. Chem. Soc. 1972, 94, 2952-2960.

51. International Mathematical and Statistical Libraries, Inc. IMSL Library Reference Manual, 7th Ed.; Author: Houston, Texas, 1979.

52. Applequist, J. J. Chem. Phys. 1979, 71, 4324-4331.

53. DeTar, D.; Luthra, N. J. Am. Chem. Soc. 1977, 99, 1232-1244.

54. Ashida, T.; Tanaka, I.; Shimonishi, Y.; Kakudo, M. Acta Crystallogr. 1977, B33, 3054-3059.

55. Aubry, A.; Protas, J.; Boussard, G.; Marraud, M. Acta Crystallogr. 1977, B33, 2399-2406.

56. Aubry, A.; Protas, J.; Boussard, G.; Marraud, M. Acta Crystallogr. 1979, B35, 694-699.

57. Aubry, A.; Protas, J.; Boussard, G.; Marraud, M. Acta Crystallogr. 1980, B36, 321-326.

58. Aubry, A.; Protas, J.; Boussard, G.; Marraud, M. Acta Crystallogr. 1980, B36, 2822-2824.

59. Aubry, A.; Protas, J.; Boussard, G.; Marraud, M. Acta Crystallogr. 1980, B36, 2825-2827.

60. Aubry, A.; Vitoux, B.; Boussard, G.; Marraud, M. Int. J. Peptide Protein Res. 1981, 18, 195-202.

61. Aubry, A.; Vitoux, B.; Marraud, M. Cryst. Struct. Commun. 1982, 11, 135-140.

62. Ayato, H.; Tanaka, I.; Ashida, T. J. Am. Chem. Soc. 1981, 103, 5902-5905.

63. Ayato, H.; Tanaka, I.; Ashida, T. J. Am. Chem. Soc. 1981, 103, 6869-6873.

64. Bando, S.; Tanaka, N.; Ashida, T.; Kakudo, M. Acta Crystallogr. 1978, B34, 3447-3449.

65. Bavoso, A.; Benedetti, E.; DiBlasio, B.; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. Macromolecules 1982, 15, 54-59.

66. Benedetti, E.; Bavoso, A.; DiBlasio, B.; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. Int. J. Peptide Protein Res. 1982, 20, 312-319.

67. Benedetti, E.; Ciajolo, A.; DiBlasio, B.; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. Macromolecules 1979, 12, 438-445.

68. Benedetti, E.; Ciajolo, A.; DiBlasio, B.; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. Int. J. Peptide Protein Res. 1979, 14, 130-142.

69. Benedetti, E.; Ciajolo, M.; Maisto, A. Acta Crystallogr. 1974, B30, 1783-1788.

70. Ashida, T.; Kakudo, M. Bull. Chem. Soc. Jpn. 1974, 47, 1129-1133.

71. Benedetti, E.; DiBlasio, B.; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. Macromolecules 1980, 13, 1454-1462.

72. Benedetti, E.; Pavone, V.; Toniolo, C.; Bonora, G.; Palumbo, M. Macromolecules 1977, 10, 1350-1356.

73. Blessing, R.; Smith, G. Acta Crystallogr. 1982, B38, 1203-1207.

74. Brahmachari, S.; Bhat, T.; Sudhakar, V.; Vijayan, M.; Rapaka, R.; Bhatnajar, R.; Ananthanarayanan, V. J. Am. Chem. Soc. 1981, 103, 1703-1708.

75. Cameron, T.; Hanson, A.; Taylor, A. Cryst. Struct. Commun. 1982, 11, 321-330.

76. Druck, U.; Littke, W.; Main, P. Acta Crystallogr. 1979, B35, 253-255.

77. Fridrichsons, J.; Mathieson, A. Acta Crystallogr. 1962, 15, 569-577.

78. Fujimoto, Y.; Irreverre, F.; Karle, J.; Karle, I.; Witkop, B. J. Am. Chem. Soc. 1971, 93, 3471-3477.

79. Fujinaga, M.; James, M. Acta Crystallogr. 1980, B36, 3196-3199.

80. Itoh, H.; Yamane, T.; Ashida, T. Acta Crystallogr. 1978, B34, 2640-2643.

81. Kamwaya, M.; Oster, O.; Bradaczek, H. Acta Crystallogr. 1981, B37, 364-367.

82. Kamwaya, M.; Oster, O.; Bradaczek, H. Acta Crystallogr. 1981, B37, 1564-1568.

83. Kamwaya, M.; Oster, O.; Bradaczek, H. Acta Crystallogr. 1982, B38, 172-176.

84. Kartha, G.; Ashida, T.; Kakudo, M. Acta Crystallogr. 1974, B30, 1861-1866.

85. Kojima, T.; Kido, T.; Itoh, H.; Yamane, T.; Ashida, T. Acta Cryst. 1980, B36, 326-331.

86. Kojima, T.; Tanaka, I.; Ashida, T. Acta Crystallogr. 1982, B38, 221-225.

87. Kojima, T.; Yamane, T.; Ashida, T. Acta Crystallogr. 1978, B34, 2896-2898.

88. Lecomte, C.; Aubry, A.; Protas, J.; Boussard, G.; Marraud, M. Acta Crystallogr. 1974, B30, 1992-1996.

89. Lecomte, C.; Aubry, A.; Protas, J.; Boussard, G.; Marraud, M. Acta Crystallogr. 1974, B30, 2343-2348.

90. Leung, Y.; Marsh, R. Acta Crystallogr. 1958, 11, 17-31.

91. Marsh, R. Acta Crystallogr. 1980, B36, 1265-1267.

92. Marsh, R.; Murthy, M.; Venkatesan, K. J. Am. Chem. Soc. 1977, 99, 1251-1256.

93. Cotrait, M.; Geoffre, S.; Hospital, M.; Precigoux, G. Acta Crystallogr. 1979, B35, 114-118.

94. Donohue, J.; Trueblood, K. Acta Crystallogr. 1952, 5, 419-431.

95. Mathieson, A.; Welsh, H. Acta Crystallogr. 1952, 5, 599-604.

96. Matsuzaki, T. Acta Crystallogr. 1974, B30, 1029-1036.

97. Matsuzaki, T.; Iitaka, Y. Acta Crystallogr. 1971, B27, 507-516.

98. Mitsui, Y.; Tsuboi, M.; Iitaka, Y. Acta Crystallogr. 1969, B25, 2182-2192.

99. Nair, C.; Nagaraj, R.; Ramapasad, S.; Balaram, P.; Vijayan, M. Acta Crystallogr. 1981, B37, 597-601.

100. Nair, C.; Vijayan, M. J. C. S. Perkin II 1980, 1800-1804.

101. Precioux, G.; Geoffre, S.; Hospital, M.; Leroy, F. Acta Crystallogr. 1982, B38, 2172-2176.

102. Prasad, B.; Balaram, H.; Balaram, P. Biopolymers 1982, 21, 1261-1273.

103. Prasad, V.; Shamala, N.; Nagaraj, R.; Chandrasekaran, R.; Balaram, P. Biopolymers 1979, 18, 1635-1646.

104. Reck, G.; Barth, A. Cryst. Struct. Commun. 1981, 10, 1001-1005.

105. Rudko, A.; Low, B. Acta Crystallogr. 1975, B31, 713-725.

106. Karle, I. Acta Crystallogr. 1970, B26, 765-770.

107. Kayushina, R.; Vainshtein, B. Soviet Physics - Crystallography 1966, 10, 698-706.

108. Koetzle, T.; Lehmann, M.; Hamelton, W. Acta Crystallogr. 1973, B29, 231-236.

109. Sabesan, M.; Venkatesan, K. Acta Crystallogr. 1971, B27, 1879-1883.

110. Sasisekharan, V. Acta Crystallogr. 1959, 12, 897-903.

111. Shamala, N.; Venkatesan, K. Cryst. Struct. Commun. 1973, 8, 5-8.

112. Smith, G.; Pletnev, V.; Duax, W.; Balasubramanian, N.; Bosshard, H.; Czerwinski, E.; Kendrick, N.; Mathews, F.; Marshall, G. J. Am. Chem. Soc. 1981, 103, 1493-1501.

113. Sugino, H.; Tanaka, I.; Ashida, T. Bull. Chem. Soc. Jpn. 1978, 51, 2855-2861.

114. Tanaka, I.; Ashida, T. Acta Crystallogr. 1980, B36, 2164-2167.

115. Tanaka, I.; Ashida, T.; Shimonishi, Y.; Kakudo, M. Acta Crystallogr. 1979, B35, 110-114.

116. Tanaka, I.; Kozima, T.; Ashida, T.; Tanaka, N.; Kakudo, M. Acta Crystallogr. 1977, B33, 116-119.

117. Ueki, T.; Ashida, T.; Kakudo, M.; Sasada, Y.; Katsube, Y. Acta Crystallogr. 1969, B25, 1840-1849.

118. Ueki, T.; Bando, T.; Ashida, T.; Kakudo, M. Acta Crystallogr. 1971, B27, 2219-2231.

119. Venkatachalapathi, Y.; Nair, C.; Vijayan, M.; Balaram, P. Biopolymers 1981, 20, 1123-1136.

120. Vitoux, B.; Aubry, A.; Cung, M.; Bossard, G.; Marraud, M. Int. J. Peptide Protein Res. 1981, 17, 469-479.

121. Yagi, Y.; Tanaka, I.; Yamane, T.; Ashida, T. J. Am. Chem. Soc. 1983, 105, 1242-1246.

122. Yamada, Y.; Tanaka, I.; Ashida, T. Acta Crystallogr. 1980, B36, 331-335.

123. Yamanda, Y.; Tanaka, I.; Ashida, T. Bull. Chem. Soc. Jpn. 1981, 54, 69-72.

124. Yamane, T.; Ashida, T.; Shimonishi, K.; Kakudo, M.; Sasada, Y. Acta Crystallogr. 1976, B32, 2071-2076.

125. Bendetti, E.; Goodman, M.; Marsh, R.; Rapoport, H.; Musich, J. Cryst. Struct. Commun. 1975, 4, 641-645.

126. Bhandary, K.; Kartha, G. In Peptides Structure and Function. Proceedings of the Ninth American Peptide Symposium, Deber, C.; Hruby, V.; Kopple, K., Eds.; Pierce Chemical Company: Rockford, Illinois; 1985, pp 163-166.

127. Brown, J.; Teller, R. J. Am. Chem. Soc. 1976, 98, 7565-7569.

128. Brown, J.; Yang, C. J. Am. Chem. Soc. 1979, 101, 446-449.

129. Chiu, Y.; Brown, L.; Lipscomb, W. J. Am. Chem. Soc. 1977, 99, 4799-4803.

130. Cook, W.; Einspahr, H.; Trapane, T.; Urry, D.; Bugg, C. J. Am. Chem. Soc. 1980, 102, 5502-5505.

131. Cook, W.; Trapane, T.; Prasad, K. Int. J. Peptide Protein Res. 1985, 25, 481-486.

132. Czugler, M.; Sasvari, K.; Hollosi, M. J. Am. Chem. Soc. 1982, 104, 4465-4469.

133. Flippen, J.; Karle, I. Biopolymers 1976, 15, 1081-1092.

134. Gierasch, L.; Karle, I.; Rockwell, A.; Yenal; K. J. Am. Chem. Soc. 1985, 107, 3321-3327.

135. Karle, I. Biochemistry 1974, 13, 2155-2162.

136. Karle, I. Int. J. Peptide Protein Res. 1984, 23, 32-39.

137. Karle, I. Int. J. Peptide Protein Res. 1986, 28, 420-427.

138. Karle, I. J. Am. Chem. Soc. 1972, 94, 81-84.

139. Karle, I. J. Am. Chem. Soc. 1978, 100, 1286-1289.

140. Karle, I. J. Am. Chem. Soc. 1979, 101 181-184.

141. Karle, I. In Perspectives in Peptide Chemistry, Eberle, Alex N.; Gerger, Rolf; Nieland, Theodor Eds.; Karger, Basel: New York, 1981; pp 261-271.

142. Karle, I.; Karle, J. Proc. Natl. Acad. Sci. USA 1981, 78, 681-685.

143. Karle, I. J. Am. Chem. Soc. 1974, 96, 4000-4006.

144. Karle, I.; Wieland, T.; Schermer, D.; Ottenheym, J. Proc. Natl. Acad. Sci. USA 1979, 76, 1532-1536.

145. Kartha, G.; Ambady, G. Acta Crystallogr. 1975, B31, 2035-2039.

146. Kartha, G.; Ambady, G.; Shankar, P. Nature 1974, 247, 204-205.

147. Kartha, G.; Bhandary, K.; Kopple, K.; Go, A.; Zhu, P. J. Am. Chem. Soc. 1984, 106, 3844-3850.

148. Kostansek, E.; Lipscomb, W.; Thiessen, W. J. Am. Chem. Soc. 1979, 101, 834-837.

149. Mauger, A.; Stuart, O.; Highet, R.; Silverton, J. J. Am. Chem. Soc. 1982, 104, 174-180.

150. Nakashima, T.; Yamane, T.; Tanaka, I.; Ashida, T. In Peptide Chemistry 1982: Proceedings of the 20th Sympos. on Peptide Chemistry, Sakakibara, S. Ed.; Protein Research Foundation: Osaka, Japan, 1983; pp 311-316.

151. Ravi, A.; Prasad, B.; Balaram, P. J. Am. Chem. Soc. 1983, 105, 106-109.

152. Kostansek, E.; Lipscomb, W.; Yocum, R.; Thiessen, W. Biochemistry 1978, 17, 3790-3795.

153. Jain, S.; Sobell, H. J. Mol. Biol. 1972, 68, 1-20.

154. Ueno, K.; Shimizu, T. Biopolymers 1983, 22, 633-641.

155. Von Dreele, R. Acta Crystallogr. 1975, B31, 966-970.

156. Aleksandrov, G.; Struchkov, Yu.; Kurganov, A. Zhurnal Strukturnoi Khimii 1973, 14, 492-501.

157. Aubry, A.; Marraud, M.; Protas, J.; Wyart, J. C. R. Acad. Sci. Paris Seire C. 1975, 280, 509-512.

158. Freeman, H.; Marzilli, L.; Maxwell, I. Inorg. Chemistry 1970, 9, 2408-2415.

159. Galitskii, N.; Deigin, V.; Saenger, W.; Pletnev, V. Bioorganicheskaya Khimiya 1977, 3, 1445-1454.

160. Sasisekharan, V. Acta Crystallogr. 1959, 12, 903-909.

161. Fraser, R.; MacRae, T.; Suzuki, E. J. Mol. Biol. 1979, 129, 463-481.

162. Traub, W.; Shmueli, U. Nature 1963, 198, 1165-1166.

163. Druyan, M.; Coulter, C.; Walter, R.; Kartha, G.; Ambady, G. J. Am. Chem. Soc. 1976, 98, 5496-5502.

164. Karle, I. J. Am. Chem. Soc. 1977, 99, 5152-5157.

165. Karle, I.; Duesler, E. Proc. Natl. Acad. Sci. USA 1977, 74, 2602-2606.

166. Benedetti, E.; Palumbo, M.; Bonora, G.; Toniolo, C. Macromolecules 1976, 9, 417-420.

167. Benedetti, E.; Christensen, A.; Gilon, C.; Fuller, W.; Goodman, M. Biopolymers 1976, 15, 2523-2534.

168. Dolber, M. Ionophores and Their Structures, John Wiley & Sons, New York, 1981, p. 360.

169. Bats, J.; Buess, H. J. Am. Chem. Soc. 1980, 102, 2065-2070.

170. Tanaka, S.; Scheraga, H. Macromolecules 1974, 7, 698-705.

171. Karplus, S.; Lifson, S. Biopolymers 1971, 10, 1973-1982.

172. Lifson, S.; Warshell, A. J. Chem. Phys. 1968, 49, 5116-5129.

173. Weiner, S.; Kollman, P.; Case, D.; Singh, U.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. J. Am. Chem. Soc. 1984, 106, 765-784.

174. Scott, R.; Scheraga, H. J. Phys. Chem. 1966, 45, 2091-2101.

175. Momany, F.; McGuire, R.; Burgess, A.; Scheraga, H. J. Phys. Chem. 1975, 79, 2361-2381.

176. Engler, E.; Andose, J.; Schleyer, P. J. Am. Chem. Soc. 1973, 95, 8005-8025.

177. Nagler, A.; Huler, E.; Lifson, S. J. Am. Chem. Soc. 1974, 96, 5319-5327.

178. Warshell, A.; Levitt, M.; Lifson, S. J. Mol. Spectrosc. 1970, 33, 84-99.

179. Gelin, B.; Karplus, M. Biochemistry 1979, 18, 1256-1268.

180. Bixon, M.; Lifson, S. Tetrahedron 1967, 23, 769-784.

181. Ramachandran, G.; Lakshminarayanan, A.; Balasubramanian, R; Tegoni, G. Biochimica et Biophysica Acta 1970, 221, 165-181.

182. Liquori, A. Q. Rev. Biophys. 2 1969, I, 65-92.

183. Jenness, D.; Sprecher, C.; Johnson, Jr., W. Biopolymers 1976, 15, 513-521.

184. Bovey, F.; Hood, F. Biopolymers 1967, 5, 325-326.

185. Torchia, D. Macromolecules 1971, 4, 440-442.

186. Rosenheck, K.; Miller, H.; Zakaria, A. Biopolymers 1969, 7, 614-618.

187. Mandel, R.; Holzwarth, G. Biopolymers 1973, 12, 655-674.

188. Timasheff, S.; Susi, H.; Townend, R.; Stevens, L.; Gorbunoff, M.; Kumosinski, T. In Conformations of Biopolymers, Ramachandran, G., Ed.; Academic Press: London, 1967; pp 173-196.

189. Moffitt, W. J. Chem. Phys. 1956, 25, 467-478.

190. Sussman, F.; Weinstein, H. Proc. Natl. Acad. Sci. USA 1989, 86, 7880-7884.

191. Deber, C.; Torchia, D.; Wong, S.; Blout, E. Proc. Nat. Acad. Sci. USA 1972, 69, 1825-1829.

192. Bartman, B.; Deber, C.; Blout, E. J. Am. Chem. Soc. 1977, 99, 1028-1033.

193. Madison, V., Hoffman-La Roche Inc., Nutley, New Jersey, personal communication.

194. Asher, I.; Phillies, G.; Geller, R.; Stanley, H. Biochemistry 1980, 19, 1805-1813.

195. Kodaka, M.; Fukaya, T. Polymer Journal 1986, 18, 775-778.

196. Deber, C.; Scatturin, A.; Vaidya, V.; Blout, E. In Peptides: Chemistry and Biochemistry. Proceedings of the First American Peptide Symposium. Weinstein, B., Ed., Marcel Dekker, Inc.: New York, 1970; pp 163-172.

197. Venkatachalam, C. Biopolymers 1968, 6, 1425-1436.

198. Grossman, S. Elementary Linear Algebra. Wadworth Publishing Company: Belmont, California, 1980; p. 116.

199. Burington, R. Handbook of Mathematical Tables and Formulas, 5th Ed., McGraw-Hill Book Company: St. Louis, 1973; Chapter 3.

X. ACKNOWLEDGMENTS

I would like to thank Dr. Jon Applequist for his assitance and guidance throughout the progress of this work. This investigation was supported by a research grant from the National Institute of General Medical Sciences (GM 13684) and a grant-in-aid from Sigma Xi.

XI. APPENDIX A. DERIVATION OF FORMULAE FOR STRUCTURE GENERATION OF PROLINE AND GLYCINE

A. Location of C^{β} or H_1^{α}

A schematic of the geometry about C^{α} is shown in Figure 29. When N, C^{α}, and C' are located in the same plane with C^{α} at the origin, they can be given by the following vectors (144):

$$\underset{\sim}{C}^{\alpha} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \underset{\sim}{C}' = \begin{pmatrix} b_1 \\ 0 \\ 0 \end{pmatrix} \quad \underset{\sim}{N} = \begin{pmatrix} b_2 \cos \tau_1 \\ -b_2 \sin \tau_1 \\ 0 \end{pmatrix} \tag{1}$$

$\underset{\sim}{C}^{\beta}$ or $\underset{\sim}{H}^{\alpha}$ is located tetrahedrally by first determining the x projection of $C^{\alpha}R$ bond (144).

$$R_x = b_1^R \cos \tau_2 \tag{2}$$

For y, recall the definition of the angle between two vectors (198).

$$\cos \tau_1^R = (\vec{C^{\alpha}N} \cdot \vec{C^{\alpha}R})/(|C^{\alpha}N||C^{\alpha}R|)$$

where $|C^{\alpha}N| = b_2$, and $|C^{\alpha}R| = b_1^R$. Using the appropriate bond lengths and vector representations of $C^{\alpha}N$ and $C^{\alpha}R$

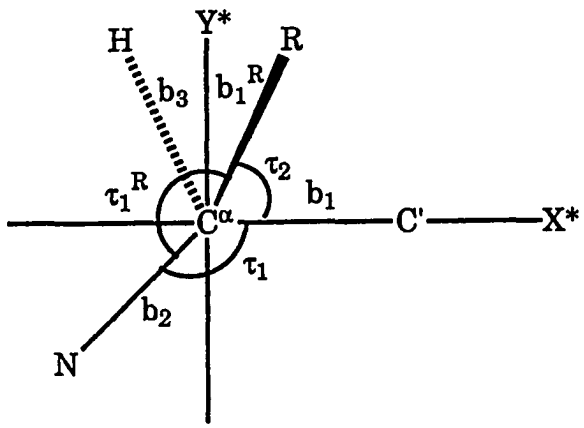
$$\cos \tau_1^R = \frac{N_x R_x + N_y R_y}{b_2 b_1^R} \tag{4}$$

Solving for R_y and using Equation 1

$$R_y = \frac{b_1^R (\cos \tau_1^R - \cos \tau_1 \cos \tau_2)}{-\sin \tau_1} \tag{5}$$

R_z can be found from the direction cosines (199).

$$R_z = (b_1^{R^2} - R_x^2 - R_y^2)^{1/2} \tag{6}$$



$R = C^\beta$ for proline and H for glycine
$b_1 = 1.53$ Å
$b_2 = 1.47$ Å
$b_3 = 1.0$ Å
$b_1^R = 1.54$ Å for proline and 1.0 Å for glycine

*The z axis is orthogonal to the plane of the paper

Figure 29. Geometry about C^α

$\underline{H}^{\alpha}$ is found by reflecting $\underline{R}$ about the $NC^{\alpha}C'$ plane and multiplying the ratio of the bond lengths (44).

$$H_x^{\alpha} = b_3 Rx / b_1^R$$

$$H_y^{\alpha} = b_3 Ry / b_1^R$$

$$H_z^{\alpha} = b_3 R_z / b_1^R$$

B. Location of H on C^{β}, C^{γ}, or C^{δ}

Hydrogens on C^{β}, C^{γ}, and C^{δ} are located by arranging a local twofold symmetry of the bonds attached to the carbon, giving approximately tetrahedral local symmetry (see Figure 30). First, ζ must be found in terms of τ_2^R and θ (44). From the pythagorean theorem

$$b_3^2 = b_p^2 + b_3^2 \sin^2 \zeta \sin^2 \theta$$

and solving for b_p

$$b_p = (b^2_3 - b^2_3 \sin^2 \zeta \sin^2 \theta)^{1/2} \tag{8}$$

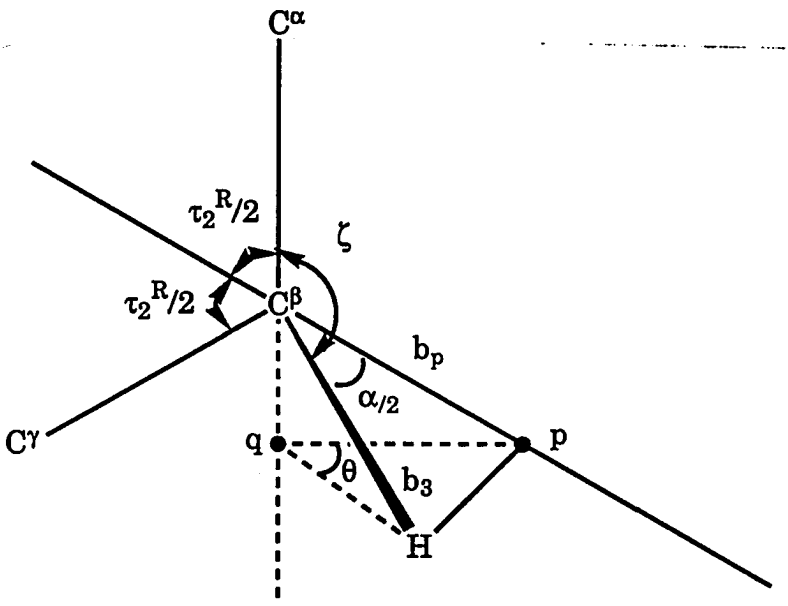
Via trigonometry

$$|pq| = b_p \sin\left(\frac{\tau_2^R}{2}\right)$$

$$= b_3 \sin \zeta \cos \theta$$

Again solving for b_p

$$b_p = \frac{b_3 \sin \zeta \cos \theta}{\sin (\tau_2^R/2)} \tag{9}$$



$\tau_2^R = <C^{\alpha}C^{\beta}C^{\gamma}$

$\zeta = < C^{\alpha}C^{\beta}H$

$\alpha/2 = <HC^{\beta}H/2 = 109.5°/2$

$b_p = |C^{\beta}p|$

$b_3 = |C^{\beta}H|$

Figure 30. Location of H on a proline side chain carbon

Using Equations 8 and 9 and solving for sinζ

$$\sin \zeta = \sin (\tau_2^R /2)/(\cos^2 \theta + \sin^2 \theta \sin^2 (\tau_2^R/2))^{1/2} \quad (10)$$

Next, θ must be found in terms of τ_2^R and α (44). b_p in terms of α is

$$b_p = b_3 \cos (\tfrac{\alpha}{2}) \quad (11)$$

Equations 9 and 11 give

$$\cos (\tfrac{\alpha}{2}) = \sin \zeta \cos \theta / \sin (\tau_2^R/2) \quad (12)$$

Using Equation 12, substituting for sinζ, and solving for cosθ gives

$$\cos \theta = \cos (\tfrac{\alpha}{2}) \sin (\tau_2^R/2)/(1 - \cos^2 (\tfrac{\alpha}{2}) \cos^2 (\tau_2^R/2))^{1/2} \quad (13)$$

Thus, $\underset{\sim}{H}_1^\beta$ and $\underset{\sim}{H}_2^\beta$ can be given by the Ramachandran-Sasisekharan method as

$$\underset{\sim}{H}_1^\beta = \underset{\sim}{C}^\beta + [M^{\chi^1+\pi-\theta}][M^{\pi-\zeta}]b_3\underset{\sim}{u}$$

$$\underset{\sim}{H}_2^\beta = \underset{\sim}{C}^\beta + [M^{\chi^1-\pi+\theta}][M^{\pi-\zeta}]b_3\underset{\sim}{u}$$

where $\underset{\sim}{u}$ is the unit vector between C^α and C^β (44).

XII. APPENDIX B. THE BORC PROGRAM

A. Listing in FORTRAN for the NAS 9160 Computer

```
1.      //          JOB
2.      //S1  EXEC  FORTVCLG
3.      //FORT.SYSIN DD *
4.      C   THIS PROGRAM VARIES THE BOND ANGLES IN THE PROLINE
5.      C RING IN ORDER TO CLOSE THAT RING.  CHI2 IS FIXED.
5.1     C ENERGY OF THE RING IS CALCULATED USE METHOD OF KARPLUS AND LIFSON.
6.      C
6.1     C  COPYRIGHT 1989 KATHRYN A. THOMASSON AND JON B. APPLEQUIST.
6.2     C
7.      C   CHECK LINES 21,25,25.1,28/39,44/49,137/142,344/347,1122 FOR EACH RUN.
8.      C  REMEMBER TO SCRATCH THE FILE WITH THE SAME NAME BEFORE RUNNING.
9.      C
10.     C   DECLARATIONS
11.           IMPLICIT REAL*8(A-H,O-Z)
12.           EXTERNAL FUNC
13.           DIMENSION X(13),F(7),PARM(4),WORK(162),XJAC(7,13)
14.           DIMENSION XJTJ(91),OUTPUT(20,20),A(91),STDEV(10),SIGMA(91)
15.           COMMON/SSQ/C(3,25),ITYPS(25),NUN,CHI(3),PHI,NTYP(20)
16.           COMMON/OUT/CNC(3),SUMOUT,TAUR1,TAUR2,TAUR3,TAUR4,BC
16.5          COMMON/TOR/CHI1,CHI2,CHI3,CHI4
17.     C   FOR DOCUMENTATION ON ZXSSQ TYPE:  USE $SYS1.IMSL.HELP#ZXSSQ
18.     C
19.           TAUR4 = 0.D0
20.           CHI3 = 0.D0
21.           PHI = 102.079D0
22.           L = 1
23.     C   WRITE TITLE
24.           WRITE(2,201) PHI
25.     201   FORMAT('PROLINE RING PHI =',1X,F5.1,1X,'CHI1i = -CHI2 FIXED
25.1        ,  TAU(1) = 110.02 POLY-L-PRO II TAU(6)=121.9')
26.     C
27.     C   INFO FOR ZXSSQ
28.           M=7
29.           N=4
30.           NSIG=4
31.           EPS = 0.000000D0
32.           DELTA = 0.000000D0
33.           MAXFN = 500
34.           IOPT = 1
35.           IXJAC = M
36.           DO 200 K=-60,60,10
37.              CHI(1) = FLOAT(K)
38.              CHI(2) = -FLOAT(K)
39.              CHI(3) = FLOAT(K)
42.           WRITE(2,203) CHI(2)
43.     203   FORMAT('THE RING AT CHI(2) = ',F10.6)
44.     C   INITIAL GUESSES OF THE BOND ANGLES IN THE PROLINE RING:
45.     C     X(1)=TAUR(1), X(2)=TAUR(2), X(3)=TAUR(3), X(4)=CHI(1)
46.           X(1) = 103.0000D0
47.           X(2) = 104.0000D0
48.           X(3) = 104.0000D0
49.           X(4) = CHI(1)
53.             CALL ZXSSQ(FUNC,M,N,NSIG,EPS,DELTA,MAXFN,IOPT,PARM,X,SSQ,F,
54.        &                XJAC,IXJAC,XJTJ,WORK,INFER,IER)
54.05          CALL COOUT(C,ITYPS,NUN)
54.1           WRITE(2,212) 'RESIDUALS FROM 1 TO 7'
54.2    212    FORMAT(' ',A)
55.1              WRITE(2,205) (F(I),I=1,M)
55.2    205       FORMAT(' ',7F15.7)
55.4           WRITE(2,212) 'ANGLES ABOUT NITROGEN: CNC(A),CNC(D),C(D)NC(A)'
```

```
55.5                WRITE(2,206) (CNC(I),I=1,3)
55.6    206         FORMAT(' ',3F15.7)
55.8              N2 = N*(N+1)/2
55.9              WRITE(2,207) (WORK(I),I=1,5), INFER,IER
56.     207       FORMAT('WORK='/5(' ',5X,G15.6/),'INFER=',I1/'IER= ',I3)
56.1    C  FIND THE DEVIATIONS IN THE PARAMETERS.
56.2    C  BEGIN BY DEFINING A AS MATRIX XJTJ WHICH WE WILL INVERT.
56.3          DO 100 I=1,N2
56.4             A(I) = XJTJ(I)
56.5    100   CONTINUE
57.1    C  SET UP A FOR INVERSION.
57.2          CALL DPPFA(A,N,INFO)
57.5          IF(INFO.NE.0)THEN
57.6             WRITE(2,218) INFO
57.7    218      FORMAT('LEADING SUBMATRIX OF A OF ORDER',I4,2X,'IS NOT
57.8           ,            NOT POSITIVE DEFINITE')
57.9             GO TO 104
58.           ENDIF
58.1    C  NOW INVERT A.
58.2          CALL DPPDI(A,N,DET,1)
58.5    C  CALCULATE THE VARIANCE MATRIX SIGMA.
58.6          DO 102 I=1,N2
58.7             SIGMA(I) = SSQ*A(I)/(M-N)
58.8    102   CONTINUE
58.9          WRITE(2,250) 'VARIANCE MATRIX'
58.95   250   FORMAT(' ',A20)
59.           CALL TRIOUT(SIGMA,N)
59.1    C  CALCULATE THE DEVIATION FROM THE VARIANCES ON THE DIAGONAL AND
59.2    C  WRITE OUT THE ANSWERS.
59.3          STDEV(1) = DSQRT(SIGMA(1))
59.4          STDEV(2) = DSQRT(SIGMA(3))
59.5          STDEV(3) = DSQRT(SIGMA(6))
59.6          STDEV(4) = DSQRT(SIGMA(10))
59.7          WRITE(2,219) 'TAUR1 = ',TAUR1,STDEV(1)
59.8          WRITE(2,219) 'TAUR2 = ',TAUR2,STDEV(2)
59.9          WRITE(2,219) 'TAUR3 = ',TAUR3,STDEV(3)
60.           WRITE(2,219) 'CHI1 = ',X(4),STDEV(4)
60.1    219   FORMAT(' ',A10,F10.6,1X,'+-',1X,F10.6)
60.2          WRITE(2,202) SSQ
60.3    202   FORMAT('SSQ = ',G20.6)
65.     C   CALCULATE THE ENERGY AND PRINT OUT COORDINATES.
66.           CALL ENERGY(UBND,UAN,UEL,UNB,UTOR,UTOT)
67.           OUTPUT(L,1) = X(4)
68.           OUTPUT(L,2) = CHI(2)
69.           OUTPUT(L,3) = CHI3
70.           OUTPUT(L,4) =  CHI4
71.           OUTPUT(L,5) = TAUR1
72.           OUTPUT(L,6) = TAUR2
73.           OUTPUT(L,7) = TAUR3
74.           OUTPUT(L,8) = TAUR4
75.           OUTPUT(L,9) = SUMOUT
76.           OUTPUT(L,10) = BC
77.           OUTPUT(L,11) = UBND
78.           OUTPUT(L,12) = UAN
79.           OUTPUT(L,13) = UEL
80.           OUTPUT(L,14) = UNB
81.           OUTPUT(L,15) = UTOR
81.5          OUTPUT(L,16) = UTOT
82.           L = L+1
83.     200   CONTINUE
```

```
C  WRITE SCRIPT COMMANDS FOR DOUBLE SPACING TO BE USED WHEN THE DATA
C  IS PRINTED FROM THE FILE CREATED BY THIS PROGRAM.
      WRITE(2,222)
222   FORMAT('.PA;.LS 1')
      WRITE(2,201) PHI
      WRITE(2,210)
210   FORMAT(' ',1X,'CHI1',5X,'CHI2',4X,'CHI3',4X,'CHI4',3X,'TAUR1',
     , 3X,'TAUR2',2X,'TAUR3',3X,'TAUR4',4X,'SUMN',6X,'BDL',5X,'UBND',
     , 4X,'UAN',5X,'UEL',5X,'UNB',7X,'UTOR',5X,'UTOT')
      DO 90 I=1,L-1
         WRITE(2,211) (OUTPUT(I,J),J=1,16)
211      FORMAT(' ',9F8.3,F6.3,1X,F7.3,1X,F7.3,1X,F10.3,1X,F7.3,1X,
     ,   F9.3,F10.3)
90    CONTINUE
C  RETURN SCRIPT COMMAND TO SINGLE SPACING FOR WHEN THE PRODUCED DATA
C  FILE IS CONNECTED TO OTHER FILES FOR PRINTING.
      WRITE(2,217)
217   FORMAT('.PA;.LS 0')
104   STOP
      END
C
C
      SUBROUTINE FUNC(X,M,N,F)
C
C   THIS SUBROUTINE IS TO BE USED WITH ZXSSQ FOR DETERMINING RING
C CLOSURE FOR A PROLINE RING.
C
C   DECLARATIONS
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION X(N), F(M), COSA(2), COSB(2), COSG(2), A1(3),
     &     A2(3), A3(3), A4(3),BOND(5),TAU(6)
      COMMON/SSQ/C(3,25),ITYPS(25),NUN,CHI(3),PHI,NTYP(20)
      COMMON/OUT/CNC(3),SUMOUT,TAUR1,TAUR2,TAUR3,TAUR4,BC
      COMMON/TOR/CHI1,CHI2,CHI3,CHI4
      CALL PROCRD(X,TAUR4,CHI3,N)
C  COSA ARE THE DIRECTION COSINES IN  THE X DIRECTION.
C  COSB  "   "     "         "     "   "  Y    "     .
C  COSG  "   "     "         "     "   "  Z    "     .
C  BC IS THE CALCULATED LAST BOND LENGTH, N1-C1(DELTA).
C  TAUR4 IS THE CALCULATED LAST BOND ANGLE, N1-C1(DELTA)-C1(GAMMA).
C  CHI3 IS THE CALCULATED TORSIONAL ANGLE: C(BETA)-C(GAMMA)-C(DELTA)-N.
C  CHI4 IS THE CALCULATED TORSIONAL ANGLE: C(GAMMA)-C(DELTA)-N-C(ALPHA).
C  * BL IS THE ACTUAL N-C(DELTA) BOND DISTANCE IN ANGSTROMS.
C  * T4  "  "    "     "    "   ANGLE IN DEGREES.
C * C3  "  "    "     "   TORSIONAL ANGLE C(BETA)-C(GAMMA)-C(DELTA)-N.
C  * C4  "  "    "     "       "       "   C(GAMMA)-C(DELTA)-N-C(ALPHA.
C  CNC ARE THE BOND ANGLES ABOUT N.  CNC(1)=C'-N-C(ALPHA) ANGLE
C                                    CNC(2)=C'-N-C(DELTA) ANGLE
C                                    CNC(3)=C(DELTA)-N-C(ALPHA) ANGLE
C  B3 IS THE CALCULATED N-C' DISTANCE.
C  B5 IS THE CALCULATED N-C(ALPHA) DISTANCE.
C  B1 IS THE CALCULATED C(GAMMA)-C(DELTA) BOND DISTANCE.
C  * MEANS THE VALUE IS INPUT.
C  ** THE COMPUTER WORKS IN RADIANS SO THE CALCULATED ANSWERS ARE
C     CONVERTED TO DEGREES SO THEY CAN BE COMPARED TO THE KNOWN ONES.
C  * WT IS THE WEIGHT FOR THE RESIDUALS FOR THE BOND ANGLES.
      DATA BL/1.470D0/
      T1 = 103.D0
      T2 = 104.D0
      T3 = 104.D0
```

```
141.            T4 = 104.DO
142.            WT = 1.DO
143.      C
144.            PI = DARCOS(-1.DO)
145.            B1 = DSQRT((C(1,6) - C(1,7))**2 + (C(2,6) - C(2,7))**2
146.           &                 + (C(3,6) - C(3,7))**2)
147.            B2 = DSQRT((C(1,4)-C(1,1))**2 + (C(2,4)-C(2,1))**2
148.           ,                  + (C(3,4)-C(3,1))**2)
149.            B3 = DSQRT((C(1,3)-C(1,14))**2 + (C(2,3)-C(2,14))**2
150.           &                  + (C(3,3)-C(3,14))**2)
151.            B4 = DSQRT((C(1,6)-C(1,4))**2 + (C(2,6)-C(2,4))**2
152.           ,                  + (C(3,6)-C(3,4))**2)
153.            B5 = DSQRT((C(1,3)-C(1,1))**2 + (C(2,3)-C(2,1))**2
154.           &                  + (C(3,3)-C(3,1))**2)
155.      C
156.      C     CALCULATION OF THE N1-C1(DELTA) BOND LENGTH.
157.            BC = DSQRT((C(1,3) - C(1,7))**2 + (C(2,3) - C(2,7))**2 +
158.           &           (C(3,3) - C(3,7))**2)
159.            F(1) = (BC - BL)/BL
160.      C  CALCULATION OF THE <N-C(ALPHA)-C(BETA) = TAUR(1).
161.            COSA(1) = (C(1,3)-C(1,1))/B5
162.            COSA(2) = (C(1,4)-C(1,1))/B2
163.            COSB(1) = (C(2,3)-C(2,1))/B5
164.            COSB(2) = (C(2,4)-C(2,1))/B2
165.            COSG(1) = (C(3,3)-C(3,1))/B5
166.            COSG(2) = (C(3,4)-C(3,1))/B2
167.            TAUR1 = DARCOS(COSA(1)*COSA(2) + COSB(1)*COSB(2)
168.           .                  + COSG(1)*COSG(2))
169.            TAUR1 = TAUR1*180.DO/PI
170.            F(2) = (TAUR1 - T1)/(T1*WT)
171.      C  CALCULATION OF THE <C(ALPHA)-C(BETA)-C(GAMMA) = TAUR(2).
172.            COSA(1) = (C(1,1)-C(1,4))/B2
173.            COSA(2) = (C(1,6)-C(1,4))/B4
174.            COSB(1) = (C(2,1)-C(2,4))/B2
175.            COSB(2) = (C(2,6)-C(2,4))/B4
176.            COSG(1) = (C(3,1)-C(3,4))/B2
177.            COSG(2) = (C(3,6)-C(3,4))/B4
178.            TAUR2 = DARCOS(COSA(1)*COSA(2) + COSB(1)*COSB(2)
179.           ,                  + COSG(1)*COSG(2))
180.            TAUR2 = TAUR2*180.DO/PI
181.            F(3) = (TAUR2 - T2)/(T2*WT)
182.      C  CALCULATION OF THE <C(BETA)-C(GAMMA)-C(DELTA) = TAUR(3).
183.            COSA(1) = (C(1,4)-C(1,6))/B4
184.            COSA(2) = (C(1,7)-C(1,6))/B1
185.            COSB(1) = (C(2,4)-C(2,6))/B4
186.            COSB(2) = (C(2,7)-C(2,6))/B1
187.            COSG(1) = (C(3,4)-C(3,6))/B4
188.            COSG(2) = (C(3,7)-C(3,6))/B1
189.            TAUR3 = DARCOS(COSA(1)*COSA(2) + COSB(1)*COSB(2)
190.           ,                  + COSG(1)*COSG(2))
191.            TAUR3 = TAUR3*180.DO/PI
192.            F(4) = (TAUR3 - T3)/(T3*WT)
193.      C
194.      C     CALCULATION OF THE N61-C1(DELTA)-C1(GAMMA) = TAUR(4).
195.            COSA(1) = (C(1,3) - C(1,7))/BC
196.            COSA(2) = (C(1,6) - C(1,7))/B1
197.            COSB(1) = (C(2,3) - C(2,7))/BC
198.            COSB(2) = (C(2,6) - C(2,7))/B1
199.            COSG(1) = (C(3,3) - C(3,7))/BC
200.            COSG(2) = (C(3,6) - C(3,7))/B1
```

```
201.            TAUR4 = DARCOS(COSA(1)*COSA(2) + COSB(1)*COSB(2) +
202.           &        COSG(1)*COSG(2))
203.            TAUR4 = TAUR4*180.DO/PI
204.            F(5) = (TAUR4 - T4)/(T4*WT)
205.      C     CALCULATION OF THE TORSIONAL ANGLE CHI(4);  C(ALPHA)-N-C(DELTA)-C(GAMMA).
206.            DO 30 I=1,3
207.               A1(I) = C(I,1)
208.               A2(I) = C(I,3)
209.               A3(I) = C(I,7)
210.               A4(I) = C(I,6)
211.      30    CONTINUE
212.            CALL TORANG(A1,A2,A3,A4,CHI4)
213.      C CALCULATE THE C'-N'C(ALPHA) ANGLE.
214.            COSA(1) = (C(1,14)-C(1,3))/B3
215.            COSA(2) = (C(1,1)-C(1,3))/B5
216.            COSB(1) = (C(2,14)-C(2,3))/B3
217.            COSB(2) = (C(2,1)-C(2,3))/B5
218.            COSG(1) = (C(3,14)-C(3,3))/B3
219.            COSG(2) = (C(3,1)-C(3,3))/B5
220.            CNC(1) = DARCOS(COSA(1)*COSA(2) + COSB(1)*COSB(2)
221.           &                + COSG(1)*COSG(2))
222.            CNC(1) = CNC(1)*180.DO/PI
223.      C CALCULATE THE C'-N-C(DELTA) ANGLE.
224.            COSA(1) = (C(1,14)-C(1,3))/B3
225.            COSA(2) = (C(1,7)-C(1,3))/BC
226.            COSB(1) = (C(2,14)-C(2,3))/B3
227.            COSB(2) = (C(2,7)-C(2,3))/BC
228.            COSG(1) = (C(3,14)-C(3,3))/B3
229.            COSG(2) = (C(3,7)-C(3,3))/BC
230.            CNC(2) = DARCOS(COSA(1)*COSA(2) + COSB(1)*COSB(2)
231.           &                + COSG(1)*COSG(2))
232.            CNC(2) = CNC(2)*180.DO/PI
233.      C CALCULATE THE C(DELTA)-N-C(ALPHA) ANGLE.
234.            COSA(1) = (C(1,7)-C(1,3))/BC
235.            COSA(2) = (C(1,1)-C(1,3))/B5
236.            COSB(1) = (C(2,7)-C(2,3))/BC
237.            COSB(2) = (C(2,1)-C(2,3))/B5
238.            COSG(1) = (C(3,7)-C(3,3))/BC
239.            COSG(2) = (C(3,1)-C(3,3))/B5
240.            CNC(3) = DARCOS(COSA(1)*COSA(2) + COSB(1)*COSB(2)
241.           &                + COSG(1)*COSG(2))
242.            CNC(3) = CNC(3)*180.DO/PI
243.            F(6) = (CNC(2) - 125.DO)/(125.DO*WT)
244.            F(7) = (CNC(3) - 112.DO)/(112.DO*WT)
245.      C FIND THE SUMS (THE THREE ANGLES AROUND SHOULD ADD TO 360).
246.            SUMOUT = CNC(1) + CNC(2) + CNC(3)
247.      C  CALCULATE THE TORSIONAL ANGLE CHI(3): C(BETA)-C(GAMMA)-C(DELTA)-N
248.            DO 40 I=1,3
249.               A1(I) = C(I,4)
250.               A2(I) = C(I,6)
251.               A3(I) = C(I,7)
252.               A4(I) = C(I,3)
253.      40    CONTINUE
254.            CALL TORANG(A1,A2,A3,A4,CHI3)
255.      C  CALCULATE THE TORSIONAL ANGLE CHI(1): N-C(ALPHA)-C(BETA)-C(GAMMA).
256.            DO 60 I=1,3
257.               A1(I) = C(I,3)
258.               A2(I) = C(I,1)
259.               A3(I) = C(I,4)
260.               A4(I) = C(I,6)
```

```
261.      60     CONTINUE
262.             CALL TORANG(A1,A2,A3,A4,CHI1)
263.      C  CALCULATE THE TORSIONAL ANGLE CHI(2):C(ALPHA)-C(BETA)-C(GAMMA)-C(DELTA).
264.             DO 70 I=1,3
265.                A1(I) = C(I,1)
266.                A2(I) = C(I,4)
267.                A3(I) = C(I,6)
268.                A4(I) = C(I,7)
269.      70     CONTINUE
270.             CALL TORANG(A1,A2,A3,A4,CHI2)
271.      C  CALCULATE THE TORSIONAL ANGLE PHI: C'14-N3-C(ALPHA)-C'2
272.             DO 80 I=1,3
273.                A1(I) = C(I,14)
274.                A2(I) = C(I,3)
275.                A3(I) = C(I,1)
276.                A4(I) = C(I,2)
277.      80     CONTINUE
278.             CALL TORANG(A1,A2,A3,A4,PHI1)
297.             RETURN
298.             END
299.      C
300.      C
301.      C
302.      C
303.             SUBROUTINE PROCRD(A,TAUR4,CHI3,NN)
304.      C   THIS PROGRAM CALCULATES COORDINATES FOR A PROLINE RESIDUE PLUS
305.      C  AN EXTRA C' SO THAT THE ENERGY CAN BE CALCULATED FOR VARIOUS
306.      C  PHI AND CHI(1) VALUES.
307.      C
308.      C   DECLARATIONS
309.                IMPLICIT REAL*8 (A-H,O-Z)
310.                DIMENSION BOND(5),TAU(6),XS(3,25),Z(3),
311.               , BONDR(3),TAUR(4),A(NN),O(3)
312.                COMMON/SSQ/C(3,25),ITYPS(25),NUN,CHI(3),PHI,NTYP(20)
313.      C   *BOND IS THE ARRAY OF BOND LENGTHS b1, b2, b3, b4, b5.
314.      C   *TAU IS THE ARRAY OF  BOND ANGLES t1, t2, t3, t4, t5, t6.
315.      C   *PHI IS THE BACKBONE DIHEDRAL ANGLE.
316.      C   *ITYPS IS THE ARRAY OF TYPE NUMBERS FOR THE SIDE CHAIN UNITS.
317.      C   X ARE THE COORDINATES OF ALL UNITS IN MOLECULAR SYSTEM A.
318.      C   ITYP(1):  THE ARRAY OF TYPE NUMBERS FOR THE ENTIRE MOLECULE.
319.      C                1:  NC'O
320.      C                2:  C(ALIPHATIC)
321.      C                3:  H(ALIPHATIC)
322.      C                4:  H(AMIDE)
323.      C                11: C'
324.      C                12: O
325.      C                13: N
326.      C   * BONDR, TAUR, AND CHI ARE INPUTS FOR XSPRO2.
327.      C   BONDR   RING BOND LENGTHS(A)
328.      C            1:  CA-CB  2:  CB-CG  3:  CG-CD
329.      C   CH      C-H BOND LENGTH(A)
330.      C   CHI     RING TORSIONAL ANGLES (DEGREES) (1969 IUPAC CONVENTION)
331.      C            1:  (N,CA,CB,CG)  2:  (CA,CB,CG,CD)  3:  (CB,CG,CD,N)
332.      C   TAUR    RING BOND ANGLES (DEG)
333.      C            1:  (N,CA,CB)  2:  (CA,CB,CG)  3:  (CB,CG,CD)  4:  (CG,CD,N)
334.      C
335.      C   PROGRAMED 4/20/88 BY K. THOMASSON .
336.      C   * MEANS THAT THESE VARIABLES ARE INPUT.
337.      C
338.      C  FUNCTIONS
```

```
339.            COTOR(X,Y)=X*DSIN(Y)/DSQRT(1.DO-X**2*DCOS(Y)**2)
340.            SIBANG(X,Y)=DSIN(X)/DSQRT(Y**2+(1.DO-Y**2)*DSIN(X)**2)
341.      C
342.      C   DATA STATEMENTS
343.      C
344.            DATA BOND/1.53DO,1.240DO,1.320DO,1.0DO,1.480DO/,NATB/5/
345.            DATA TAU/110.0200DO,112.4DO,121.0DO,114.0DO,123.0DO,121.9DO/
346.            DATA BONDR/1.540DO,1.540DO,1.540DO/
347.            DATA TAUR/103.00DO,104.00DO,104.00DO,104.00DO/
348.            DATA O/O.DO,O.DO,O.DO/
349.            NUN = 14
350.      C   INITIALIZE VALUES FOR TAUR(1), TAUR(2), TAUR(3), TAUR(4),AND CHI(2).
351.            TAUR(1) = A(1)
352.            TAUR(2) = A(2)
353.            TAUR(3) = A(3)
354.            CHI(1) = A(4)
355.            IF(TAUR4.NE.O.DO)THEN
356.               TAUR(4) = TAUR4
357.            ENDIF
358.            IF(CHI3.NE.O.DO)THEN
359.               CHI(3) = CHI3
360.            ENDIF
361.      C
362.      C  LOCATE C ALPHA AT THE ORIGIN
363.            DO 10 I=1,3
364.               XS(I,1) = O(I)
365.      10    CONTINUE
366.            ITYPS(1) = 2
367.            NTYP(1) = 5
368.      C
369.      C  LOCATE C'1.
370.            XS(1,2) = BOND(1)
371.            XS(2,2) = O.DO
372.            XS(3,2) = O.DO
373.            ITYPS(2) = 11
374.            NTYP(2) = 7
375.      C
376.      C   MORE DATA
377.            CH = 1.095DO
378.            PI=DARCOS(-1.DO)
379.            STAU=DARCOS(-1.DO/3.DO)/2.DO
380.            CSTAU=DCOS(STAU)
381.      C
382.      C        LOCATE N.
383.      C
384.           TA=TAU(1)*PI/180.DO
385.           Z(1)=DCOS(TA)*BOND(5)
386.           Z(2)=-DSIN(TA)*BOND(5)
387.           Z(3)=O.DO
388.            XS(1,3) = Z(1)
389.            XS(2,3) = Z(2)
390.            XS(3,3) = Z(3)
391.            ITYPS(3) = 13
392.            NTYP(3) = 2
393.      C
394.      C        LOCATE C(BETA).
395.      C
396.          XS(1,4)=BONDR(1)*DCOS(TAU(2)*PI/180.DO)
397.          XS(2,4)=(BOND(5)*BONDR(1)*DCOS(TAUR(1)*PI/180.DO)-XS(1,4)
398.         & *Z(1))/Z(2)
```

```
      XS(3,4)=DSQRT(BONDR(1)**2-XS(1,4)**2-XS(2,4)**2)
       ITYPS(4) = 2
       NTYP(4) = 12
C
C        LOCATE H(ALPHA).
C
      XS(1,5)=XS(1,4)*CH/BONDR(1)
      XS(2,5)=XS(2,4)*CH/BONDR(1)
      XS(3,5)=-XS(3,4)*CH/BONDR(1)
      ITYPS(5) = 3
      NTYP(5) = 11
C
C        LOCATE C(GAMMA) AND C(DELTA).
C
      CALL CHAIN2(TAUR(2),CHI(1),BONDR(2),Z,0,XS(1,4),XS(1,6))
      CALL CHAIN2(TAUR(3),CHI(2),BONDR(3),0,XS(1,4),XS(1,6),
     & XS(1,7))
      ITYPS(6) = 2
      NTYP(6) = 14
      ITYPS(7) = 2
      NTYP(7) = 16
C
C        LOCATE H(BETA)'S.
C
      TA=TAUR(2)*PI/360.DO
      COT=COTOR(CSTAU,TA)
      TORS=180.DO-DARCOS(COT)*180.DO/PI
      BANG=180.DO-DARSIN(SIBANG(TA,COT))*180.DO/PI
      CALL CHAIN2(BANG,CHI(1)+TORS,CH,Z,0,XS(1,4),XS(1,8))
      CALL CHAIN2(BANG,CHI(1)-TORS,CH,Z,0,XS(1,4),XS(1,9))
      ITYPS(8) = 3
      NTYP(8) = 13
      ITYPS(9) = 3
      NTYP(9) = 13
C
C        LOCATE H(GAMMA)'S.
C
      TA=TAUR(3)*PI/360.DO
      COT=COTOR(CSTAU,TA)
      TORS=180.DO-DARCOS(COT)*180.DO/PI
      BANG=180.DO-DARSIN(SIBANG(TA,COT))*180.DO/PI
      CALL CHAIN2(BANG,CHI(2)+TORS,CH,0,XS(1,4),XS(1,6),
     & XS(1,10))
      CALL CHAIN2(BANG,CHI(2)-TORS,CH,0,XS(1,4),XS(1,6),
     & XS(1,11))
      ITYPS(10) = 3
      NTYP(10) = 15
      ITYPS(11) = 3
      NTYP(11) = 15
C
C        LOCATE H(DELTA)'S.
C
      TA=TAUR(4)*PI/360.DO
      COT=COTOR(CSTAU,TA)
      TORS=180.DO-DARCOS(COT)*180.DO/PI
      BANG=180.DO-DARSIN(SIBANG(TA,COT))*180.DO/PI
      CALL CHAIN2(BANG,CHI(3)+TORS,CH,XS(1,4),XS(1,6),XS(1,7),
     , XS(1,12))
      CALL CHAIN2(BANG,CHI(3)-TORS,CH,XS(1,4),XS(1,6),XS(1,7),
     , XS(1,13))
```

```
459.            ITYPS(12) = 3
460.            NTYP(12) = 17
461.            ITYPS(13) = 3
462.            NTYP(13) = 17
463.      C
464.      C  LOCATE C'2.
465.      C
466.             CALL CHAIN2(TAU(6),PHI+180.DO,BOND(3),XS(1,2),0,XS(1,3),XS(1,14))
467.             ITYPS(14) = 11
468.            NTYP(14) = 7
469.             N= 14
470.            DO 30 I=1,NUN
471.               DO 40 J=1,3
472.                  C(J,I) = XS(J,I)
473.      40        CONTINUE
474.      30     CONTINUE
481.            RETURN
482.            END
483.      C
484.      C
485.      C
486.               SUBROUTINE ENERGY(UBND,UAN,UEL,UNB,UTOR,UTOT)
487.      C    THIS PROGRAM CALCULATES THE POTENTIAL ENERGY OF A PARTICULAR
488.      C    PROLINE FRAGMENT BASED ON ITS CARTESIAN COORDINATES.
489.      C
490.      C
491.      C    DECLARATIONS
492.               IMPLICIT REAL*8 (A-H,O-Z)
493.               DIMENSION R(20,20),Q(20),ALPHA(20),NANECT(20,20),
494.              , EN(20),RG(20),RGKL(20,20),EPSI(20,20),NBNECT(20,20),A(20),Y(20),
495.              , Z(20),AKL(20,20),CKL(20,20),NVRTEX(20,20),THETAK(20,20),
496.              , THETAO(20,20),UVECT(3),VVECT(3),NDNECT(20,20),RK(20,20),RO(20,20)
497.               COMMON/SSQ/C(3,25),ITYPS(25),NUN,CHI(3),PHI,NTYP(20)
497.5             COMMON/TOR/CHI1,CHI2,CHI3,CHI4
498.      C
499.      C    X(3,NATMS) ARE THE ATOMIC COORDINATES.
500.      C    A(NATMS) ARE THE X VALUES.
501.      C    Y(NATMS) ARE THE Y VALUES.
502.      C    Z(NATMS) ARE THE Z VALUES.
503.      C    NATMS IS THE TOTAL NUMBER OF ATOMS IN THE MOLECULE.
504.      C    R(NATMS,NATMS) ARE THE INTERATOMIC DISTANCES.
505.      C    NTYP ARE THE PARTIAL CHARGE TYPES FOR EACH ATOM
506.      C         1:  N(GLY)
507.      C         2:  N(PRO)
508.      C         3:  H(AMIDE)
509.      C         4:  C(ALPHA-GLY)
510.      C         5:  C(ALPHA-PRO)
511.      C         6:  C'(GLY)
512.      C         7:  C'(PRO)
513.      C         8:  O(GLY)
514.      C         9:  O(PRO)
515.      C         10:  H(ALPHA-GLY)
516.      C         11:  H(ALPHA-PRO)
517.      C         12:  C(BETA-PRO)
518.      C         13:  H(BETA-PRO)
519.      C         14:  C(GAMMA-PRO)
520.      C         15:  H(GAMMA-PRO)
521.      C         16:  C(DELTA-PRO)
522.      C         17:  H(DELTA-PRO)
523.      C    Q ARE THE ACTUAL PARTIAL CHARGES IN ECU.
```

```
524.      C   AKL ARE THE REPULSIVE COEFFICIENTS IN KCAL A**12/MOLE.
525.      C   CKL ARE THE ATTRACTIVE COEFFICIENTS IN KCAL A**6/MOLE.
526.      C
527.      C
528.      C  CALCULATE THE CARTESIAN COORDINATES
530.            NATMS=NUN
531.      C
532.      C   DETERMIME INTERATOMIC DISTANCES
533.      C
534.      C
535.      C   SEPARATE THE X, Y, AND Z VALUES FROM X(3,NATMS).
536.      C
537.            DO 100 I=1,NATMS
538.                  A(I) = C(1,I)
539.                  Y(I) = C(2,I)
540.                  Z(I) = C(3,I)
541.      100   CONTINUE
542.      C
543.      C   CALCULATE THE INTERATOMIC DISTANCES IN ANGSTROMS.
544.      C
545.      C       JJ = 1
546.            DO 10 I=1,NATMS
547.               DO 11 J=1,NATMS
548.                     R(I,J) = DSQRT((A(I)-A(J))**2 + (Y(I)-Y(J))**2 +
549.           $         (Z(I)-Z(J))**2)
550.      11       CONTINUE
551.      C            JJ = JJ + 1
552.      10    CONTINUE
553.      C   READ IN THE INTERACTION MATRICES TO DETERMINE WHICH COEFFICIENTS TO
554.      C   CALCULATE. NBNECT IS USED FOR UNB AND UEL, NANECT AND NVRTEX ARE
555.      C   ARE USED FOR UAN, NDNECT IS USED FOR UBND.
556.      C   SINCE THE MATRICES ARE SYMMETRIC, THEY WILL BE INITIALIZED TO 0 AND
557.      C  THEN THE BOTTOM TRIANGLE PUT IN. THAT WILL BE USED TO FILL IN
558.      C  THE TOP TRIANGLE.
559.      C  THE MATRIX OF THE VERTICES FOR THE 1-3 INTERACTIONS, NVRTEX, WILL
560.      C  ALSO BE INITIALIZED TO 0.
561.            DO 301 I=1,NATMS
562.               DO 302 J=1,NATMS
563.                  NANECT(I,J) = 0
564.                  NBNECT(I,J) = 0
565.                  NVRTEX(I,J) = 0
565.5                 NDNECT(I,J) = 0
566.      302      CONTINUE
567.      301   CONTINUE
568.      C
569.      C  INPUT THE NONZERO VALUES OF NBNECT.
570.      C  0 = 1-4 OR GREATER INTERACTIONS
571.      C  1 = 1 TO 1 OR 1 TO 2 INTERACTIONS WHICH ARE NOT INCLUDED IN THE
572.      C  ENERGY CALCULATION.
573.      C  2 = TO 3 ANGLE BENDING INTERACTIONS NOT INCLUDED IN UNB.
574.      C   ROW 1
575.            NBNECT(1,1) = 1
576.      C
577.      C   ROW 2
578.            NBNECT(2,1) = 1
579.            NBNECT(2,2) = 1
580.      C
581.      C   ROW 3
582.            NBNECT(3,1) = 1
583.            NBNECT(3,2) = 2
```

```
584.                NBNECT(3,3) = 1
585.          C
586.          C     ROW 4
587.                NBNECT(4,1) = 1
588.                NBNECT(4,2) = 2
589.                NBNECT(4,3) = 2
590.                NBNECT(4,4) = 1
591.          C
592.          C     ROW 5
593.                NBNECT(5,1) = 1
594.                NBNECT(5,2) = 2
595.                NBNECT(5,3) = 2
596.                NBNECT(5,4) = 2
597.                NBNECT(5,5) = 1
598.          C
599.          C     ROW 6
600.                NBNECT(6,4) = 1
601.                NBNECT(6,6) = 1
602.          C
603.          C     ROW 7
604.                NBNECT(7,3) = 1
605.                NBNECT(7,6) = 1
606.                NBNECT(7,7) = 1
607.          C
608.          C     ROW 8
609.                NBNECT(8,1) = 2
610.                NBNECT(8,4) = 1
611.                NBNECT(8,6) = 2
612.                NBNECT(8,8) = 1
613.          C
614.          C     ROW 9
615.                NBNECT(9,1) = 2
616.                NBNECT(9,4) = 1
617.                NBNECT(9,6) = 2
618.                NBNECT(9,8) = 2
619.                NBNECT(9,9) = 1
620.          C
621.          C     ROW 10
622.                NBNECT(10,4) = 2
623.                NBNECT(10,6) = 1
624.                NBNECT(10,7) = 2
625.                NBNECT(10,10) = 1
626.          C
627.          C     ROW 11
628.                NBNECT(11,4) = 2
629.                NBNECT(11,6) = 1
630.                NBNECT(11,7) = 2
631.                NBNECT(11,10) = 2
632.                NBNECT(11,11) = 1
633.          C
634.          C     ROW 12
635.                NBNECT(12,3) = 2
636.                NBNECT(12,6) = 2
637.                NBNECT(12,7) = 1
638.                NBNECT(12,12) = 1
639.          C
640.          C     ROW 13
641.                NBNECT(13,3) = 2
642.                NBNECT(13,6) = 2
643.                NBNECT(13,7) = 1
```

```
644.            NBNECT(13,12) = 2
645.            NBNECT(13,13) = 1
646.      C
647.      C   ROW 14
648.            NBNECT(14,1) = 2
649.            NBNECT(14,3) = 1
650.            NBNECT(14,7) = 2
651.            NBNECT(14,14) = 1
652.      C
653.      C  NOW INPUT THE INTERACTION MATRIX FOR THE ANGLE BENDING POTENTIAL.
654.      C  2 = <H-C(ALIPHATIC)-H INTERACTION.
655.      C  3 = <H-C(ALIPHATIC)-C(ALIPHATIC)
656.      C  4 = <C(ALIPHATIC)-N-H
657.      C  5 = <C(ALIPHATIC)-N-C'OR C ALIPHATIC.
658.      C  6 = <C'-N-H
659.      C  7 = <C(ALIPHATIC)-C(ALIPHATIC)-N
660.      C  8 = <N-C(ALIPHATIC)-H
661.      C  9 = <N-C'-O
662.      C  10 = <C(ALIPHATIC)-C'-O
663.      C  11 = <C(ALIPHATIC)-C'-N
664.      C  12 = <C'-C(ALIPHATIC)-H
665.      C  13 = <C-C-C
666.      C  14 = <C'-C(ALIPHATIC)-N
667.      C   ROW 1
668.            NANECT(1,1) = 1
669.      C
670.      C   ROW 2
671.            NANECT(2,1) = 1
672.            NANECT(2,2) = 1
673.      C
674.      C   ROW 3
675.            NANECT(3,1) = 1
677.            NANECT(3,3) = 1
678.      C
679.      C   ROW 4
680.            NANECT(4,1) = 1
681.            NANECT(4,2) = 13
682.            NANECT(4,3) = 7
683.            NANECT(4,4) = 1
684.      C
685.      C   ROW 5
686.            NANECT(5,1) = 1
688.            NANECT(5,3) = 8
689.            NANECT(5,4) = 3
690.            NANECT(5,5) = 1
691.      C
692.      C   ROW 6
693.            NANECT(6,1) = 13
694.            NANECT(6,3) = 7
695.            NANECT(6,4) = 1
696.            NANECT(6,6) = 1
697.      C
698.      C   ROW 7
700.            NANECT(7,3) = 1
701.            NANECT(7,4) = 13
702.            NANECT(7,6) = 1
703.            NANECT(7,7) = 1
704.      C
705.      C   ROW 8
706.            NANECT(8,1) = 3
```

```
707.              NANECT(8,4) = 1
708.              NANECT(8,6) = 3
709.              NANECT(8,8) = 1
710.        C
711.        C   ROW 9
712.              NANECT(9,1) = 3
713.              NANECT(9,4) = 1
714.              NANECT(9,6) = 3
716.              NANECT(9,9) = 1
717.        C
718.        C   ROW 10
719.              NANECT(10,4) = 3
720.              NANECT(10,6) = 1
721.              NANECT(10,7) = 3
722.              NANECT(10,10) = 1
723.        C
724.        C   ROW 11
725.              NANECT(11,4) = 3
726.              NANECT(11,6) = 1
727.              NANECT(11,7) = 3
729.              NANECT(11,11) = 1
730.        C
731.        C   ROW 12
732.              NANECT(12,3) = 8
733.              NANECT(12,6) = 3
734.              NANECT(12,7) = 1
735.              NANECT(12,12) = 1
736.        C
737.        C   ROW 13
738.              NANECT(13,3) = 8
739.              NANECT(13,6) = 3
740.              NANECT(13,7) = 1
742.              NANECT(13,13) = 1
743.        C
744.        C   ROW 14
746.              NANECT(14,3) = 1
748.              NANECT(14,14) = 1
749.        C
750.        C  FILL IN BOTTOM HALF OF NVRTEX.
751.        C
752.        C  ROW 3
753.              NVRTEX(3,2) = 1
754.        C
755.        C  ROW 4
756.              NVRTEX(4,2) = 1
757.              NVRTEX(4,3) = 1
758.        C
759.        C  ROW 5
760.              NVRTEX(5,2) = 1
761.              NVRTEX(5,3) = 1
762.              NVRTEX(5,4) = 1
763.        C
764.        C  ROW 6
765.              NVRTEX(6,1) = 4
766.              NVRTEX(6,3) = 7
767.        C
768.        C  ROW 7
769.              NVRTEX(7,1) = 3
770.              NVRTEX(7,4) = 6
771.        C
```

```
772.      C  ROW 8
773.             NVRTEX(8,1) = 4
774.             NVRTEX(8,6) = 4
775.      C
776.      C  ROW 9
777.             NVRTEX(9,1) = 4
778.             NVRTEX(9,6) = 4
779.             NVRTEX(9,8) = 4
780.      C
781.      C  ROW 10
782.             NVRTEX(10,4) = 6
783.             NVRTEX(10,7) = 6
784.      C
785.      C  ROW 11
786.             NVRTEX(11,4) = 6
787.             NVRTEX(11,7) = 6
788.             NVRTEX(11,10) = 6
789.      C
790.      C  ROE 12
791.             NVRTEX(12,3) = 7
792.             NVRTEX(12,6) = 7
793.      C
794.      C  ROW 13
795.             NVRTEX(13,3) = 7
796.             NVRTEX(13,6) = 7
797.             NVRTEX(13,12) = 7
798.      C
799.      C  ROW 14
800.             NVRTEX(14,1) = 3
801.             NVRTEX(14,7) = 3
801.01    C
801.02    C  FILL IN THE BOTTOM HALF OF NDNECT.
801.03    C
801.04           NDNECT(2,1) = 1
801.05           NDNECT(3,1) = 1
801.06           NDNECT(4,1) = 1
801.07           NDNECT(5,1) = 1
801.08           NDNECT(6,4) = 1
801.09           NDNECT(7,3) = 1
801.1            NDNECT(7,6) = 1
801.11           NDNECT(8,4) = 1
801.12           NDNECT(9,4) = 1
801.13           NDNECT(10,6) = 1
801.14           NDNECT(11,6) = 1
801.15           NDNECT(12,7) = 1
801.16           NDNECT(13,7) = 1
801.17           NDNECT(14,3) = 1
802.      C   NOW  FILL IN THE TOP HALF OF THE MATRIX.
803.             JJ = 1
804.             DO 326 I=1,NATMS
805.                DO 327 J=JJ,NATMS
806.                   NANECT(I,J) = NANECT(J,I)
807.                   NBNECT(I,J) = NBNECT(J,I)
808.                   NVRTEX(I,J) = NVRTEX(J,I)
808.5                  NDNECT(I,J) = NDNECT(J,I)
809.      327       CONTINUE
810.                JJ = JJ + 1
811.      326    CONTINUE
812.             JJ = 1
817.             DO 401 I=1,NATMS
```

```
818.                DO 402 J=1,NATMS
819.                   IF(NBNECT(I,J).NE.NBNECT(J,I))THEN
820.                     WRITE(2,212) I,J,J,I
821.     212             FORMAT('NBNECT(',I2,I2,') NE NBNECT(',I2,I2,')')
822.                   ENDIF
823.                  IF(NVRTEX(I,J).NE.NVRTEX(J,I))THEN
824.                    WRITE(2,213) I,J,J,I
825.     213            FORMAT('NVRTEX(',I2,I2,') NE NVRTEX(',I2,I2,')')
826.                  ENDIF
827.     402       CONTINUE
828.     401    CONTINUE
828.01   C
828.02   C  CALCULATE THE BONDING POTENTIAL.
828.03   C
828.04         UBND = 0.D0
828.05   C
828.06   C  ASSIGN THE BONDING COEFFICIENTS RK AND RO.
828.07   C
828.08         DO 130 I=1,NATMS
828.09            DO 140 J=1,NATMS
828.1                IF(NDNECT(I,J).EQ.1.AND.((ITYPS(I).EQ.13.AND.ITYPS(J).EQ.2)
828.11        ,          .OR.(ITYPS(I).EQ.2.AND.ITYPS(J).EQ.13)))THEN
828.12                     RK(I,J) = 810.D0
828.13                     RO(I,J) = 1.458D0
828.14                ELSEIF(NDNECT(I,J).EQ.1.AND.ITYPS(I).EQ.2.AND.
828.15        ,          ITYPS(J).EQ.2)THEN
828.16                     RK(I,J) = 224.D0
828.17                     RO(I,J) = 1.457D0
828.18                 ELSEIF(NDNECT(I,J).EQ.1.AND.((ITYPS(I).EQ.13.AND.ITYPS(J)
828.19        ,           .EQ.11).OR.(ITYPS(I).EQ.11.AND.ITYPS(J).EQ.13)))THEN
828.2                      RK(I,J) = 806.D0
828.21                     RO(I,J) = 1.279D0
828.22                 ELSEIF(NDNECT(I,J).EQ.1.AND.((ITYPS(I).EQ.2.AND.ITYPS(J).
828.23        ,           EQ.11).OR.(ITYPS(I).EQ.11.AND.ITYPS(J).EQ.2)))THEN
828.24                     RK(I,J) = 374.D0
828.25                     RO(I,J) = 1.47D0
828.26                 ELSEIF(NDNECT(I,J).EQ.1.AND.((ITYPS(I).EQ.2.AND.ITYPS(J).
828.27        ,           EQ.3).OR.(ITYPS(I).EQ.3.AND.ITYPS(J).EQ.2)))THEN
828.28                     RK(I,J) = 574.D0
828.29                     RO(I,J) = 1.1D0
828.3                  ELSE
828.31                     RK(I,J) = 0.D0
828.32                     RO(I,J) = 0.D0
828.33                 ENDIF
828.34   140          CONTINUE
828.35   130    CONTINUE
828.36   C
828.37   C  NOW CALCULATE THE BONDING POTENTIAL.
828.38   C
828.39         DO 150 I=1,NATMS
828.4             DO 160 J=1,NATMS
828.41               IF(NDNECT(I,J).EQ.0)THEN
828.42                  UBND = UBND
828.43               ELSE
828.44                  UBND = UBND + RK(I,J)*(R(I,J)-RO(I,J))**2
828.45               ENDIF
828.46   160       CONTINUE
828.47   150    CONTINUE
828.5          UBND = (UBND/2.D0) - 7.548
829.     C  CALCULATE THE ELECTROSTATIC POTENTIAL ENERGY.
```

```
C
      UEL = 0.DO
C  ASSIGN PARTIAL CHARGES:
      DO 20 I=1,NATMS
         IF(NTYP(I).EQ.2)THEN
            Q(I) = -.305DO
         ELSEIF(NTYP(I).EQ.5)THEN
            Q(I) = 0.144DO
         ELSEIF(NTYP(I).EQ.7)THEN
            Q(I) = 0.449DO
         ELSEIF(NTYP(I).EQ.11)THEN
            Q(I) = 0.0DO
         ELSEIF(NTYP(I).EQ.12)THEN
            Q(I) = 0.000DO
         ELSEIF(NTYP(I).EQ.13)THEN
            Q(I) = 0.000DO
         ELSEIF(NTYP(I).EQ.14)THEN
            Q(I) = 0.00DO
         ELSEIF(NTYP(I).EQ.15)THEN
            Q(I) = 0.000DO
         ELSEIF(NTYP(I).EQ.16)THEN
            Q(I) = 0.0DO
         ELSEIF(NTYP(I).EQ.17)THEN
            Q(I) = 0.00DO
         ENDIF
20    CONTINUE
C  CALCULATION OF THE ELECTROSTATIC ENERGY IN KCAL/MOLE:
       DO 35 I=1,NATMS
          DO 40 J=1,NATMS
             IF(NBNECT(I,J).EQ.1.OR.NBNECT(I,J).EQ.2)THEN
                UEL = UEL
             ELSEIF(NBNECT(I,J).EQ.0)THEN
                UEL = UEL + 332.DO*Q(I)*Q(J)/(1.DO*R(I,J))
             ELSEIF(NBNECT(I,J).GT.2)THEN
                WRITE(2,214)
             ENDIF
40        CONTINUE
35     CONTINUE
       UEL = UEL/2.DO
C  CALCULATE THE NONBONDED ENERGY.
      UNB = 0.DO
C  ASSIGN ATTRACTIVE (CKL) AND REPULSIVE (AKL) COEFFICIENTS.
      DO 110 I=1,NATMS
         DO 120 J=1,NATMS
            IF(NBNECT(I,J).EQ.1.OR.NBNECT(I,J).EQ.2)THEN
               AKL(I,J) = 0.DO
               CKL(I,J) = 0.DO
            ELSEIF(NBNECT(I,J).EQ.0)THEN
             IF(ITYPS(I).EQ.3.AND.ITYPS(J).EQ.3)THEN
                AKL(I,J) = 1841.7DO
                CKL(I,J) = 5.7506DO
             ELSEIF((ITYPS(I).EQ.3.AND.ITYPS(J).EQ.2).OR.
     ,             (ITYPS(I).EQ.3.AND.ITYPS(J).EQ.11).OR.
     ,             (ITYPS(I).EQ.2.AND.ITYPS(J).EQ.3).OR.
     ,             (ITYPS(I).EQ.2.AND.ITYPS(J).EQ.4).OR.
     ,             (ITYPS(I).EQ.11.AND.ITYPS(J).EQ.3).OR.
     ,             (ITYPS(I).EQ.11.AND.ITYPS(J).EQ.4).OR.
     ,             (ITYPS(I).EQ.4.AND.ITYPS(J).EQ.2).OR.
     ,             (ITYPS(I).EQ.4.AND.ITYPS(J).EQ.11))THEN
               AKL(I,J) = 2.98D4
```

```
                    CKL(I,J) = 49.00D0
                   ELSEIF((ITYPS(I).EQ.3.AND.ITYPS(J).EQ.13).OR.
     ,                   (ITYPS(I).EQ.4.AND.ITYPS(J).EQ.13).OR.
     ,                   (ITYPS(I).EQ.13.AND.ITYPS(J).EQ.3).OR.
     ,                   (ITYPS(I).EQ.13.AND.ITYPS(J).EQ.4))THEN
                     AKL(I,J) = 3.98D4
                     CKL(I,J) = 65.30D0
                   ELSEIF(ITYPS(I).EQ.2.AND.ITYPS(J).EQ.2)THEN
                     AKL(I,J) = 4.26D5
                     CKL(I,J) = 391.8D0
                   ELSEIF((ITYPS(I).EQ.2.AND.ITYPS(J).EQ.11).OR.
     ,                   (ITYPS(I).EQ.11.AND.ITYPS(J).EQ.2))THEN
                     AKL(I,J) = 4.26D5
                     CKL(I,J) = 391.8D0
                   ELSEIF((ITYPS(I).EQ.2.AND.ITYPS(J).EQ.13).OR.
     ,                   (ITYPS(I).EQ.13.AND.ITYPS(J).EQ.2))THEN
                     AKL(I,J) = 5.69D5
                     CKL(I,J) = 522.4D0
                   ELSEIF(ITYPS(I).EQ.11.AND.ITYPS(J).EQ.11)THEN
                     AKL(I,J) = 4.26D5
                     CKL(I,J) = 391.8D0
                   ENDIF
214                 FORMAT('YOU HAVE GOOFED AGAIN')
                  ENDIF
120            CONTINUE
110         CONTINUE
C  CALCULATION OF THE NONBONDED ENERGY IN KCAL/MOLE.
            DO 70 I=1,NATMS
               DO 80 J=1,NATMS
                  IF(I.EQ.J)THEN
                     UNB = UNB
                  ELSE
                     UNB = UNB + ((AKL(I,J)/R(I,J)**12) -
     ,                           (CKL(I,J)/R(I,J)**6))
                  ENDIF
80             CONTINUE
70          CONTINUE
            UNB = UNB/2.D0
C
C   CALCULATE THE ANGLE BENDING POTENTIAL
C
C  FIRST ASSIGN VALUES OF THE SPRING CONSTANT, THETAK IN KCAL/MOL/DEG
C  AND MINIMUM ANGLE, THETAO IN DEGREES.
            DO 304 I=1,NATMS
               DO 303 J=1,NATMS
                  IF(NANECT(I,J).EQ.2)THEN
                     THETAK(I,J) = 38.2D0
                     THETAO(I,J) = 109.47D0
                  ELSEIF(NANECT(I,J).EQ.3)THEN
                     THETAK(I,J) = 27.D0
                     THETAO(I,J) = 109.47D0
                  ELSEIF(NANECT(I,J).EQ.4)THEN
                     THETAK(I,J) = 31.4D0
                     THETAO(I,J) = 120.D0
                  ELSEIF(NANECT(I,J).EQ.5)THEN
                     THETAK(I,J) = 54.5D0
                     THETAO(I,J) = 120.D0
                  ELSEIF(NANECT(I,J).EQ.6)THEN
                     THETAK(I,J) = 26.7D0
                     THETAO(I,J) = 120.D0
```

```
1002.                  ELSEIF(NANECT(I,J).EQ.7)THEN
1003.                     THETAK(I,J) = 21.DO
1004.                     THETAO(I,J) = 109.47DO
1005.                  ELSEIF(NANECT(I,J).EQ.8)THEN
1006.                     THETAK(I,J) = 30.1DO
1007.                     THETAO(I,J) = 109.47DO
1008.                  ELSEIF(NANECT(I,J).EQ.9)THEN
1009.                     THETAK(I,J) = 48.5DO.
1010.                     THETAO(I,J) = 120.DO
1011.                  ELSEIF(NANECT(I,J).EQ.10)THEN
1012.                     THETAK(I,J) = 40.9DO
1013.                     THETAO(I,J) = 120.DO
1014.                  ELSEIF(NANECT(I,J).EQ.11)THEN
1015.                     THETAK(I,J) = 33.1DO
1016.                     THETAO(I,J) = 120.DO
1017.                  ELSEIF(NANECT(I,J).EQ.12)THEN
1018.                     THETAK(I,J) = 28.7DO
1019.                     THETAO(I,J) = 109.47DO
1020.                  ELSEIF(NANECT(I,J).EQ.13)THEN
1021.                     THETAK(I,J) = 21.6DO
1022.                     THETAO(I,J) = 109.47DO
1023.                  ELSEIF(NANECT(I,J).EQ.14)THEN
1024.                     THETAK(I,J) = 21.DO
1025.                     THETAO(I,J) = 109.47DO
1026.                  ELSE
1027.                     THETAK(I,J) = 0.DO
1028.                     THETAO(I,J) = 0.DO
1029.                  ENDIF
1030.      303       CONTINUE
1031.      304     CONTINUE
1032.      C  NOW CALCULATE THE ANGLE BENDING ENERGY.
1033.            UAN = 0.DO
1034.            DO 306 I=1,NATMS
1035.               DO 307 J=1,NATMS
1036.                  IF(NANECT(I,J).GE.2.AND.NANECT(I,J).LE.14)THEN
1037.                     UVECT(1) = C(1,I) - C(1,NVRTEX(I,J))
1038.                     UVECT(2) = C(2,I) - C(2,NVRTEX(I,J))
1039.                     UVECT(3) = C(3,I) - C(3,NVRTEX(I,J))
1040.                     UDIST = R(I,NVRTEX(I,J))
1041.                     VVECT(1) = C(1,J) - C(1,NVRTEX(I,J))
1042.                     VVECT(2) = C(2,J) - C(2,NVRTEX(I,J))
1043.                     VVECT(3) = C(3,J) - C(3,NVRTEX(I,J))
1044.                     VDIST = R(J,NVRTEX(I,J))
1047.                     COSTH = (UVECT(1)*VVECT(1) + UVECT(2)*VVECT(2)
1048.                      + UVECT(3)*VVECT(3))/(UDIST*VDIST)
1049.                     THETA = DARCOS(COSTH)
1050.                     PI=DARCOS(-1.DO)
1051.                     THETA = THETA*(180.DO/PI)
1052.                     UAN = UAN + 0.5DO*THETAK(I,J)*(THETA - THETAO(I,J))**2
1055.                  ELSE
1056.                     UAN = UAN
1057.                  ENDIF
1058.      307       CONTINUE
1059.      306    CONTINUE
1060.            UAN = UAN*3.04617D-4
1060.05    C  CALCULATION OF THE TORSIONAL POTENTIAL, UTOR.
1060.1           UTOR = 0.DO
1060.15    C  CONVERT EACH  CHI FROM DEGREES TO  RADIANS.
1060.2            CHI1R = CHI1*PI/180.DO
1060.25           CHI2R = CHI2*PI/180.DO
```

```
1060.3            CHI3R = CHI3*PI/180.DO
1060.31           CHI4R = CHI4*PI/180.DO
1060.32           PHIR = PHI*PI/180.DO
1060.35     C  NOW CALCULATE THE INDIVIDUAL TORSIONAL ENERGIES FOR X1, X2, X3.
1060.4            UX1 = 2.836DO*DCOS(3.DO*CHI1R)
1060.45           UX2 = 2.836DO*DCOS(3.DO*CHI2R)
1060.5            UX3 = 2.836DO*DCOS(3.DO*CHI3R)
1060.51           UX4 = -1.5DO*DCOS(3.DO*CHI4R)
1060.52           UX5 = -1.5*DCOS(3.DO*PHIR)
1060.55     C  NOW CALCULATE THE TORSIONAL POTENTIAL.
1060.6            UTOR = UX1 + UX2 + UX3 + UX4 + UX5
1061.       C   CALCULATE THE TOTAL POTENTIAL ENERGY.
1079.             UTOT = UBND + UAN + UEL + UNB + UTOR
1097.             RETURN
1098.             END
1099.       C
1100.       C
1101.             SUBROUTINE COOUT(X,ITYP,NUNITS)
1102.             REAL*8 X(3,NUNITS)
1103.             DIMENSION ITYP(NUNITS)
1104.             WRITE(2,215)
1105.       215   FORMAT('COORDINATES INPUT'//' ',T4,'UNIT',T9,'TYPE',
1106.          $          T22,'X',T37,'Y',T53,'Z'/)
1107.             WRITE(2,216) (I,ITYP(I),(X(M,I),M=1,3),I=1,NUNITS)
1108.       216   FORMAT (' ',2I5,3F15.6)
1109.             RETURN
1110.             END
1110.005    C
1110.006    C
1110.01           SUBROUTINE TRIOUT(A,N)
1110.02           IMPLICIT REAL*8(A-H,O-Z)
1110.03           DIMENSION A(1)
1110.04     C  THIS ROUTINE OUTPUTS THE TRIANGULAR MATRIX A.
1110.05     C
1110.06           DO 105 J1=1,N,9
1110.07              J2=MINO(J1+8,N)
1110.08              DO 106 I=J1,N
1110.09                 JT=MINO(J2,I)
1110.1      106        WRITE(2,202) I,(A((I*I-I)/2+J),J=J1,JT)
1110.11               WRITE(2,203) (J,J=J1,JT)
1110.12     105       WRITE(2,203)
1110.13           RETURN
1110.14     202   FORMAT('  ',I3,9F13.6)
1110.15     203   FORMAT(' ',9(8X,I5))
1110.16           END
1111.       //LKED.SYSLIB  DD
1112.       //    DD
1113.       //    DD
1114.       //    DD
1115.       //    DD
1116.       //    DD
1118.       //    DD  DSN=S1$APP.LOAD,DISP=SHR
1119.       //    DD  DSN=S1$APP.IMSL9,DISP=SHR
1120.       //    DD  DSN=SYSU.LINPACK.SUBLIB,DISP=SHR
1121.       //GO.SYSIN  DD  *
1122.       //GO.FT02F001 DD DSN=S1$KAT.T3,DISP=(NEW,CATLG),
1123.       // SPACE=(TRK,(1,1),RLSE),DCB=(RECFM=FB,LRECL=150,BLKSIZE=6150),
1124.       // UNIT=DISK
1125.       //
```