INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106

76-9493

PRICE, Brandon Jay, 1948-
CONFORMATIONAL ENERGIES OF PROLINE AND SUBSTITUTED PROLINE INTERNAL DIPEPTIDES.

The University of Michigan, Ph.D., 1975
Biophysics, general

Xerox University Microfilms, Ann Arbor, Michigan 48106

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.

CONFORMATIONAL ENERGIES OF PROLINE AND SUBSTITUTED PROLINE INTERNAL DIPEPTIDES

by
Brandon Jay Price

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biophysics)
in The University of Michigan
1975

Doctoral Committee:

Assistant Professor C. M. Venkatachalam, Chairman
Professor S. Krimm
Professor J. L. Oncley
Associate Professor R. Zand

Dedicated to my parents

John and Patricia Price

ACKNOWLEDGEMENTS

My sincere thanks go to Dr. C. M. Venkatachalam for his inspiration, guidance and friendship throughout the course of my graduate studies. I am especially grateful for the time and talent which he made so abundantly available.

I am also indebted to Drs. Samuel Krimm and J. L. Oncley for many valuable discussions and helpful suggestions.

Thanks also go to Marion Frost for her kindness and administrative support.

Finally, I wish to acknowledge N. I. H. Training Grant GM-1355 for financial support which greatly facilitated completion of this work.

TABLE OF CONTENTS

		Page
DEDICATION		ii
ACKNOWLEDGEMENTS		iii
LIST OF TABLES		vi
LIST OF ILLUSTRATIONS		x
LIST OF APPENDICES		xiv
Chapter		
I.	INTRODUCTION	1
II.	PUCKERING OF THE PYRROLIDINE RING	10
2.1	Determination of Ring Coordinates	10
2.2	Classification of Ring Geometries	14
2.3	Contributions to the Potential Energy	15
2.4	Details of the Calculation	15
2.5	Energy of the Isolated Ring as a Function of Γ and θ	17
2.6	Analysis of Bond Angles and Torsion Angles	27
III.	RING PUCKERING AND THE PRO-PRO INTERNAL DIPEPTIDE	44
3.1	Structure and Coordinates of the Pro-Pro Dimer	44
3.2	Conformational Energy Maps for *Cis* and *Trans* Dimers	46
3.3	Stable Conformations of the Proline Dimer	55
3.4	The *Cis'* Region in the Proline Internal Dipeptide	55
3.5	Energy Barriers Between Stable Conformations for Ring Geometries with Constant θ	59
3.6	Energy Barriers Between Stable Conformations with no Restrictions on θ	66

Page

3.7 (ω,ψ) Map for the Pro-Pro Dimer with Independently Puckered Rings. 74
3.8 Γ-θ Energy Plots for the Pro-Pro Dimer 83

IV. METHYLATED PROLINE INTERNAL DIPEPTIDES . . 87

4.1 Structure and Coordinates of the 2- and 5-Methylprolines. 87
4.2 Analysis of Torsion Angles and Bond Angles for the 5-Methylprolines . . . 90
4.3 Γ-θ Plots for the 5-Methylprolines. . 92
4.4 (θ,ψ) Conformational Energy Maps for the 5-Methylproline Dimers. 99
4.5 Energy Barriers Between Stable Conformations of the 5-Methylproline Dimers. 105
4.6 (ω,ψ) Energy Maps for the 5-Methylproline Dimers with Independently Puckered Rings. 112
4.7 Γ-θ Energy Plot for 2-Methylproline . 124
4.8 (ω,ψ) Conformational Energy Map for the 2-Methylproline Dimer 129
4.9 Possible Implications of Structural Differences Among the Proline and Methylproline Dimers. 140

V. HYDROGEN BONDING IN THE 4-HYDROXYPROLINE INTERNAL DIPEPTIDE 143

5.1 Structure and Coordinates 143
5.2 Γ-θ Plots for the 4-Hydroxyproline Isolated Rings. 145
5.3 Hydrogen Bonding in the 4-Hydroxyproline Dimer 151
5.4 Distance and Angle Restrictions as Criteria for Hydrogen Bonding 154
5.5 Energy Restrictions as Criteria for Hydrogen Bonding. 167
5.6 Bifurcated Hydrogen Bonding 169
5.7 Summary of Results and Conclusions on Hydrogen Bonding. 172

VI. RESULTS AND DISCUSSION 176

6.1 Summary of Results. 176
6.2 Discussion. 178

APPENDICES 184

BIBLIOGRAPHY 198

LIST OF TABLES

Table		Page
1-1.	Characteristic Ratios for Polyproline Using Square-Well Model Potentials	8
2-1.	Average Values of Proline Residue Bond Lengths Observed in Crystal Structures . . .	16
2-2.	Torsion Angles and Bond Angles (in degrees) of the Isolated Proline Ring Corresponding to Minimum Energy Ring Puckering	29
2-3.	Γ-θ Energies for the Isolated Proline Ring (in kcal/mole)	33
2-4.	Γ-θ Torsion Angles for $\chi^{\alpha\beta}$ (in degrees) . .	38
2-5.	Γ-θ Torsion Angles for $\chi^{\beta\gamma}$ (in degrees) . .	39
2-6.	Γ-θ Torsion Angles for $\chi^{\delta\gamma}$ (in degrees) . .	40
2-7.	Γ-θ Torsion Angles for $\chi^{N\delta}$ (in degrees) . .	41
2-8.	Comparison with Ring Torsions Experimentally Found in Some Crystals	42
3-1.	Location of Minima and Energies at these Minima for the <u>cis</u> and <u>trans</u> Proline Dimers	50
3-2.	Values of (ω,ω') for each (θ,ψ) Corresponding to the Lowest Energy Ring for the <u>trans</u> Pro-Pro Dimer	52
3-3.	Values of (ω,ω') for each (θ,ψ) Corresponding to the Lowest Energy Ring for the <u>cis</u> Pro-Pro Dimer	53
3-4.	Torsion Angles (in degrees) and Minimum Energies (in kcal/mole) for the Pro-Pro Dimer in Regions II and III Including Variation in $\tau^{N'}$, $\tau^{N''}$ and Partial Charges	58

Table		Page
3-5.	The Torsional Angles (in degrees) and Energy Values (in kcal/mole) of Minimum Energy Conformations of the Pro-Pro Internal Dipeptide	64
3-6.	The Torsional Angles (in degrees) and Energy Values (in kcal/mole) of the Conformations at the Barriers Separating the Stable Conformations of the Pro-Pro Internal Dipeptide.	65
3-7.	Value of θ for Each (ω,ψ) Corresponding to the Lowest Energy Ring for the Pro-Pro Dimer with Identical Ring Puckering	70
3-8.	Values of (θ_1,θ_2) for Each (ω,ψ) Corresponding to the Lowest Energy Rings for the Pro-Pro Dimer with Independently Puckered Rings.	80
4-1.	Γ-θ Energies for the cis 5-Methylproline Isolated Ring	96
4-2.	Γ-θ Energies for the trans 5-Methylproline Isolated Ring	97
4-3.	The Torsional Angles (in degrees) and Energy Values (in kcal/mole) of Minimum Energy Conformations of the cis and trans 5-Methylproline Internal Dipeptides.	109
4-4.	The Torsional Angles (in degrees) and Energy Values (in kcal/mole) of the Conformations at the Barriers Separating the Low Energy Conformations of the cis and trans 5-Methylproline Internal Dipeptides.	110
4-5.	Effect of Methyl Group Rotation on the 5-Methylproline Dimer Energies in the cis' Region	113
4-6.	Values of (θ_1,θ_2) for Each (ω,ψ) Corresponding to the Lowest Energy Rings for the cis 5-Methylproline with Independently Puckered Rings	125

Table		Page
4-7.	Values of (θ_1,θ_2) for Each (ω,ψ) Corresponding to the Lowest Energy Rings for the trans 5-Methylproline Dimer with Independently Puckered Rings.	127
4-8.	Γ-θ Energies for the 2-Methylproline Isolated Ring.	130
4-9.	The Torsional Angles (in degrees) and Energy Values (in kcal/mole) of Minimum Energy Conformations of the 2-Methylproline Internal Dipeptide	133
4-10.	The Torsional Angles (in degrees) and Energy Values (in kcal/mole) of Minimum Energy Conformations of the 2-Methylproline Internal Dipeptide Where Backbone Bond Angles are Varied	136
5-1.	Γ-θ Energies and Value of χ for Each Energy in the cis 4-Hydroxyproline Isolated Ring. .	149
5-2.	Γ-θ Energies and Value of χ for Each Energy in the trans 4-Hydroxyproline Isolated Ring	150
5-3.	Type 3 Hydrogen Bonding Conformations in the 4-Hydroxyproline Internal Dipeptide.	157
5-4.	Type 4 Hydrogen Bonding Conformations in the 4-Hydroxyproline Internal Dipeptide for $\psi=-45^{\circ}$	160
5-5.	Type 4 Hydrogen Bonding Conformations in the 4-Hydroxyproline Internal Dipeptide for $\psi=170^{\circ}$	162
5-6.	Type 4 Hydrogen Bonding Conformations in the 4-Hydroxyproline Internal Dipeptide for $\psi=180^{\circ}$	163
5-7.	Type 4 Hydrogen Bonding Conformations in the 4-Hydroxyproline Internal Dipeptide for $\psi=190^{\circ}$	164
5-8.	Bifurcated Hydrogen Bonding Types in the 4-Hydroxyproline Dimer	170
5-9.	Summary of Conclusions for Hydrogen Bonding in the 4-Hydroxyproline Internal Dipeptide .	173

Table		Page
I-1.	Van der Waals Interaction Constants.	187
I-2.	Comparison of Different Sets of Partial Charges (in e.s.u.) for the Proline Internal Unit.	189

LIST OF ILLUSTRATIONS

Figure		Page
1-1.	Cis and Trans forms of the proline residue .	5
2-1.	Proline ring with bond angles and torsion angles marked.	11
2-2.	The geometry of the C^{γ} arcing.	12
2-3.	Γ-θ energy plot for the proline isolated ring with a planar trigonal nitrogen.	18
2-4.	Γ-θ angle plot for τ^{α} in the proline isolated ring with a planar trigonal nitrogen	20
2-5.	Γ-θ angle plot for τ^{N} in the proline isolated ring with a planar trigonal nitrogen	21
2-6.	Γ-θ energy plot for the proline isolated ring with a sp^3 tetrahedral nitrogen.	23
2-7.	Γ-θ angle plot for τ^{α} in the proline isolated ring with a tetrahedral nitrogen	25
2-8.	Γ-θ angle plot for τ^{N} in the proline isolated ring with a tetrahedral nitrogen	26
2-9.	Distribution of bond angles and torsion angles in the set of proline rings with energy less than 1.5 kcal/mole from the global energy minimum	28
2-10.	Γ-θ energy plot for the proline isolated ring	32
2-11.	Self-energy of the variously puckered proline rings as a function of θ	34
2-12.	Puckering in the isolated proline ring . . .	36
3-1.	The Pro-Pro internal dipeptide	45
3-2.	θ-ψ conformational energies for the trans proline dimer.	48
3-3.	θ-ψ conformational energies for the cis proline dimer.	49

Figure		Page
3-4.	Γ-θ energy plot for the proline isolated ring with $\tau^{\alpha}=103^{\circ}$, $\tau^{N}=113^{\circ}$ and $\tau^{N'}=121^{\circ}$. . .	61
3-5.	(ω,ψ) energy contour map for the Pro-Pro internal dipeptide in which θ is held constant at -10°.	62
3-6.	(ω,ψ) energy contour map for the Pro-Pro internal dipeptide in which θ is held constant at $+10^{\circ}$.	67
3-7.	(ω,ψ) energy contour map for the Pro-Pro internal dipeptide in which θ is allowed to vary.	69
3-8.	(ω,ψ) energy contour map for the Pro-Pro internal dipeptide in which θ is allowed to vary and ℓ(C'-N) is a function of ω.	73
3-9.	(ω,ψ) energy contour map for the Pro-Pro internal dipeptide with independent ring puckering.	75
3-10.	$(Pro)_2$ <u>cis</u> (stereoscopic view)	77
3-11.	$(Pro)_2$ <u>trans</u> (stereoscopic view)	78
3-12.	(Pro) <u>cis'</u> (stereoscopic view).	79
3-13.	Γ-θ energy plot for the <u>cis</u> proline dimer. .	84
3-14.	Γ-θ energy plot for the <u>trans</u> proline dimer.	86
4-1.	The 5-methylproline isolated rings	88
4-2.	The 2-methylproline isolated ring.	89
4-3.	Distribution of bond angles and torsion angles in the set of 5-methylproline rings with energy less than 1.5 kcal/mole from the global minimum energy.	91
4-4.	Γ-θ energy plot for the <u>cis</u> 5-methylproline isolated ring.	93
4-5.	Γ-θ energy plot for the <u>trans</u> 5-methylproline isolated ring.	94
4-6.	Self-energy of variously puckered 5-methylproline rings as a function of θ	98

Figure		Page
4-7.	θ-ψ conformational energies for the *cis* 5-methylproline *trans* dimer.	101
4-8.	θ-ψ conformational energies for the *trans* 5-methylproline *trans* dimer.	102
4-9.	θ-ψ conformational energies for the *cis* 5-methylproline *cis* dimer.	103
4-10.	θ-ψ conformational energies for the *trans* 5-methylproline *cis* dimer.	104
4-11.	(ω,ψ) energy contour map for the *cis* 5-methylproline internal dipeptide	107
4-12.	(ω,ψ) energy contour map for the *trans* 5-methylproline internal dipeptide	111
4-13.	(ω,ψ) energy contour map for the *cis* 5-methylproline internal dipeptide with independently puckered rings	115
4-14.	(ω,ψ) energy contour map for the *trans* 5-methylproline internal dipeptide with independently puckered rings	116
4-15.	*Cis* (5-methyl pro)$_2$ *cis* (stereoscopic view)	117
4-16.	*Cis* (5-methyl pro)$_2$ *trans* (stereoscopic view)	118
4-17.	*Cis* (5-methyl pro)$_2$ *cis'* (stereoscopic view)	120
4-18.	*Trans* (5-methyl pro)$_2$ *cis* (stereoscopic view).	121
4-19.	*Trans* (5-methyl pro)$_2$ *trans* (stereoscopic view).	122
4-20.	*Trans* (5-methyl pro)$_2$ *cis'* (stereoscopic view).	123
4-21.	Γ-θ energy plot for the 2-methylproline isolated ring.	128
4-22.	(ω,ψ) energy contour map for the 2-methyl-proline internal dipeptide	131
4-23.	(2-methyl pro)$_2$ *cis* (stereoscopic view). . .	137
4-24.	(2-methyl pro)$_2$ *trans* (stereoscopic view). .	138

Figure		Page
4-25.	(2-methyl pro)$_2$ *cis'* (stereoscopic view) . .	139
5-1.	Two forms of the 4-hydroxyproline isolated ring	144
5-2.	Γ-θ energy plot for the *cis* 4-hydroxyproline isolated ring.	146
5-3.	Γ-θ energy plot for the *trans* 4-hydroxy-proline isolated ring.	147
5-4.	Schematic drawing of the 4-hydroxyproline dimer showing the four possible types of hydrogen bonding	152
5-5.	Two kinds of Type 4 hydrogen bonding	165
6-1.	Γ-θ energy plot for the *cis* proline helix. .	181
6-2.	Γ-θ energy plot for the *trans* proline helix	182

LIST OF APPENDICES

Appendix		Page
I	CALCULATION OF CONFORMATIONAL ENERGIES.	184
II	A FORTRAN PROGRAM FOR Γ-ARCING.	193

CHAPTER I

INTRODUCTION

One of the challenging problems in the study of protein structure is to make an accurate theoretical prediction of secondary or tertiary structure by knowing only the amino acid sequence and molecular environment. To achieve this the conformational energy of the protein as a function of its internal coordinates must be known. The calculation of this energy involves several complex problems. In the first place, accurate energy functions are necessary. Even if these functions are available, one is confronted with an exceedingly large number of conformations which the polypeptide may assume due to rotations about backbone bonds as well as side chain bonds. In the solid state, crystal packing effects are present, while in solution, polymer-solvent interactions may play an important role.

The number of conformations to be studied can be considerably reduced if we restrict ourselves to oligopeptides and small homopolypeptides. Besides the obvious reduction in computing time, these small peptides serve as excellent model systems and the understanding of their conformational stability may be expected to be quite useful in ultimately solving the problem of protein structure. Also, these small

peptides are nowadays readily synthesized and a wealth of experimental data obtained.

In the computation of the conformational energy of the small peptides, two basic problems exist:

1. Only approximate potential functions are available.
2. Polymer-solvent interactions are not well understood, though some attempts have been recently made (1-5) to study the effect of solvation on conformation.

In spite of these difficulties, approximate energy calculations of small peptides in vacuo have been quite useful (6-10). The ultimate objective, however, will be to gradually remove these approximations to the point where the energy of a protein molecule could be properly determined.

The allowed conformations of a molecule are limited, of course, by restrictions on the contact distances between different atoms. In the earliest calculations (11-22), a conformation was deemed acceptable if no single pair of interacting atoms came closer than the sum of their van der Waals radii. This all-or-none criterion was followed by the development of expressions for energy as a function of interatomic distance. Eventually, other energy contributions such as electrostatic energy, bond angle bending energy, bond length stretching energy and intrinsic torsional energy were introduced. These functions and their various forms are discussed in Appendix I.

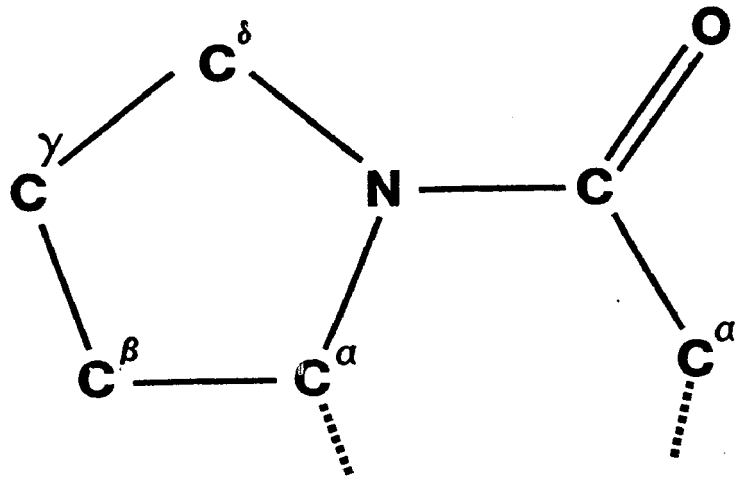
A sensitive test of the validity of a conformational energy map is the successful theoretical prediction of an experimentally measurable quantity. Such a quantity, which depends upon the conformational energies for all portions of the map, is the characteristic ratio (CR) as described by Flory (23). The calculation of CR is dependent upon the conformational energy E of the chain under consideration. If E is separable into a sum of energies, each due to a segment of the polymer chain, both the conformational energy and CR calculations turn out to be relatively simple. The conformational energy map for a dimer yields enough information for calculating the CR for a polymer inasmuch as the nearest neighbor energy contributions can be easily obtained. With the dimer as the segment of interest, this technique has been applied successfully to a host of synthetic homopolymers and has been further adapted to polypeptide chains (7). This theoretical treatment is of interest particularly because it permits an analytical correlation of the CR with the polymer structure. The CR of a linear polymer can be experimentally determined in dilute solution. Excellent agreement between theory and experiment has been found for several synthetic polymers (see ref. 23, chaps. 5-7) as well as for poly(L-lactic acid) (23) and poly(L-alanine) analogs (24).

Making use of the above concepts, the objective of this thesis is to understand the conformational nature of the pyrrolidine ring and its subsequent effects upon the

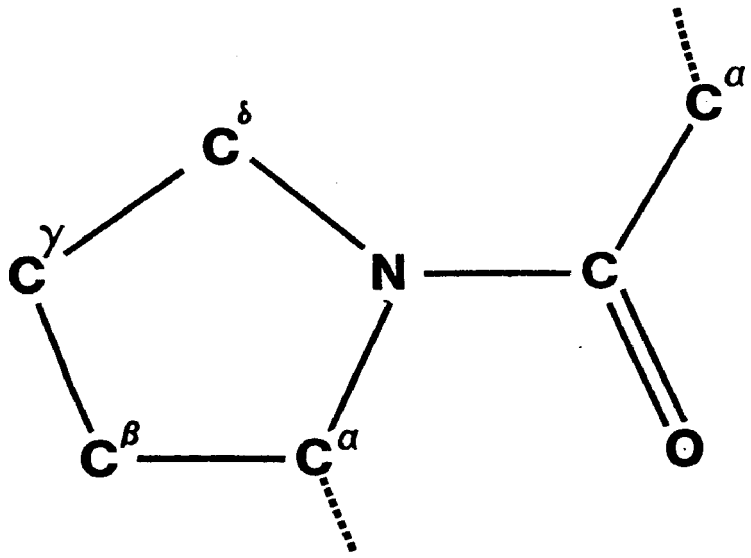
proline and substituted proline dimers. The molecules which have been studied are the internal dipeptides of proline, 2-methylproline, cis and trans 5-methylproline and cis and trans 4-hydroxyproline. (The terms cis and trans refer to the configuration of the methyl or hydroxyl group*.) This thesis follows the conventions of the 1970 IUPAC-IUB Commission on Biochemical Nomenclature (25).

Both proline and hydroxyproline are naturally occurring imino acids. Together they constitute almost one-fourth the residues in collagen, the major constituent of cartilage and other connective tissue. As shown in Figure 1-1, the side chain is attached to the nitrogen of the same residue, supplanting the hydrogen found in all other naturally occurring amino acids. This confers upon proline the unique properties of interrupting an α-helix in a protein and/or forcing a bend in the main chain. Myoglobin is a good example of the effects of proline upon structure. While every bend in that molecule does not have a proline, every proline forces a bend (26). In addition, proline can exist with the peptide bond in either the cis or trans conformation. In poly(L-proline) two forms are known to exist. Form I corresponds to all peptide bonds in the planar cis conformation while Form II has all peptide bonds

*According to the American Chemical Society convention (97), the terms erythro and threo should be employed to designate optical isomers of the proline derivatives containing a second asymmetric carbon atom. Nevertheless, in this thesis, cis is used for erythro and trans refers to threo, in keeping with the current usage in the proline literature (63,64). Admittedly such usage can be confusing!



CIS (Form I)



TRANS (Form II)

Figure 1-1. Cis and Trans forms of the proline residue.

in the planar trans conformation.

The structure of the proline residue in the solid state has been determined by X-ray diffraction studies (27-32). It is on the basis of the structure determined for the internal trans residue by Cowan and McGavin (27), Sasisekharan (29) and Leung and Marsh (28) and for poly(L-proline) I by Traub and Shmueli (30) that many energy calculations have been done for helical poly(L-proline) (6,33-36), for the L-proline interior dipeptide unit (36-41) and for the L-prolyl-L-proline dipeptide (34,41). These studies vary considerably in their use of non-bonded potentials, inclusion of electrostatic and intrinsic torsional potentials and dependence upon dihedral angles.

The CR for poly(L-proline) II has been determined experimentally (42) in several different solvents and has been found to differ by about an order of magnitude from calculations (6,33-39,43). This discrepancy suggests that all aspects of the structure of the proline residue have not been fully understood. Most early calculations involving proline have suffered from two approximations in particular. Firstly, the backbone torsion angle ω has been held constant at its planar cis or trans value. It seems quite reasonable to assume that ω could vary somewhat, perhaps as much as $\pm 20^{\circ}$ from its equilibrium value. A few calculations (37,35) which have taken this into account indicate that ω variation could be important in calculating the CR. The second approximation, which involves holding φ constant

because of ring closure, is more serious in its implications on the polymer structure because it involves the bulky pyrrolidine ring. Model building with ball and stick models or CPK space-filling models shows that the pyrrolidine ring is quite flexible, making it possible for the ring torsion angle θ associated with the four ring atoms C^{β}, C^{α}, N and C^{δ} to rotate about 45° on <u>either</u> side of its planar 0° value. This contrasts sharply with early calculations in which the ring has been held rigid in either one or two puckered conformations or even in the planar form! Despite these approximations, the calculations all predict a low conformational energy for ψ in the region $140^{\circ} \leq \psi \leq 160^{\circ}$.

Poly(L-proline) II is a rather stiff polymer compared to polyglycine or poly(L-alanine). Both ring flexibility and ω variation will alter the conformational energy distribution which could be quite important in assessing the statistical properties of this molecule. To illustrate this point further, we have calculated the CR for an infinitely long poly(L-proline) II chain using model square-well potentials, with three wells of varying depths lying between $\psi=110^{\circ}$ and $\psi=190^{\circ}$ (i.e., the region of allowed ψ for the proline dimer). The three wells, A, B and C were centered at $\psi=120^{\circ}$, 150° and 180°. Each was 20° in width and either 2 or 4 kcal/mole deep. The results are tabulated in order of increasing CR in Table 1-1. The experimentally observed CR for poly(L-proline) II is about 20 (42). Values obtained from Table 1-1 range from 29 to 608! It is obvious

Table 1-1. Characteristic Ratios for Polyproline Using Square-Well Model Potentials

Configuration*			CR
A	B	C	
-2	0	-2	29
-4	0	-4	29
-4	-2	-4	29
-2	-2	-2	43
-4	-4	-4	43
0	0	0	46
0	-2	-2	80
-2	-4	-4	80
0	-4	-4	88
-2	-2	0	102
-4	-4	-2	102
0	0	-2	111
-4	-4	0	113
-2	-2	-4	116
-2	0	-4	130
-2	0	0	177
-4	-2	-2	187
-4	0	-2	212
0	-2	-4	246
0	-2	0	261
-2	-4	-2	265
0	-4	-2	354
-2	-4	0	359
-4	-2	0	408
0	-4	0	542
-4	0	0	608

*Well A is centered at $\psi = 120^{\circ}$; well B at $\psi = 150^{\circ}$; and well C at $\psi = 180^{\circ}$. The well depths are in kcal/mole. For values of ψ between 110° and 190° other than in well regions, the energy is 0 kcal/mole.

from these results that the CR for poly(L-proline) is quite sensitively dependent upon the nature of the energy distribution.

In addition to ring flexibility and ω variation, it is probable that other factors such as bond angle bending, partial charge assignment, etc. may play crucial roles in the conformation and structure of the proline dimer and hence, upon statistical properties deduced for the polymer. It is these theoretical aspects to which this thesis addresses itself. In the next chapter, the concepts of pyrrolidine ring puckering are developed. In succeeding chapters, these notions on ring flexibility are incorporated into the proline dimer, the 2- and 5-methylproline dimers and finally, the 4-hydroxyproline dimers.

CHAPTER II

PUCKERING OF THE PYRROLIDINE RING

2.1 Determination of Ring Coordinates

Consider an isolated pyrrolidine ring as shown in Figure 2-1. We wish to study the puckering of this ring to determine its most stable conformations. For this, the energy of each ring conformation must be computed. There are nine degrees of freedom associated with the five ring atoms. Since the five bond lengths may be held constant while computing the ring energy, only four degrees of freedom are left. The bond angles τ^{α} and τ^{N} along with the ring torsion angle θ associated with the atoms C^{β}-C^{α}-N-C^{δ} completely specify the positions of these four atoms. This leaves one degree of freedom to position the C^{γ} atom. The locus of all positions of the C^{γ} atom maintaining constant $\ell(C^{\beta}\text{-}C^{\gamma})$ and $\ell(C^{\delta}\text{-}C^{\gamma})$ is obviously the intersection of two spheres with centers on the C^{β} and C^{δ} atoms and with radii $\ell(C^{\beta}\text{-}C^{\gamma})$ and $\ell(C^{\delta}\text{-}C^{\gamma})$. Thus the C^{γ} atom will lie somewhere on a circle whose center O is on the line $C^{\beta}\ldots C^{\delta}$. Rotation of the line $O\ldots C^{\gamma}$ about the line $C^{\beta}\ldots C^{\delta}$ produces all possible positions of C^{γ}. As shown in Figure 2-2, we refer to the angle of this pseudorotation as Γ. $\Gamma=0^{o}$ is taken to correspond to the position of C^{γ} lying in the plane bisecting

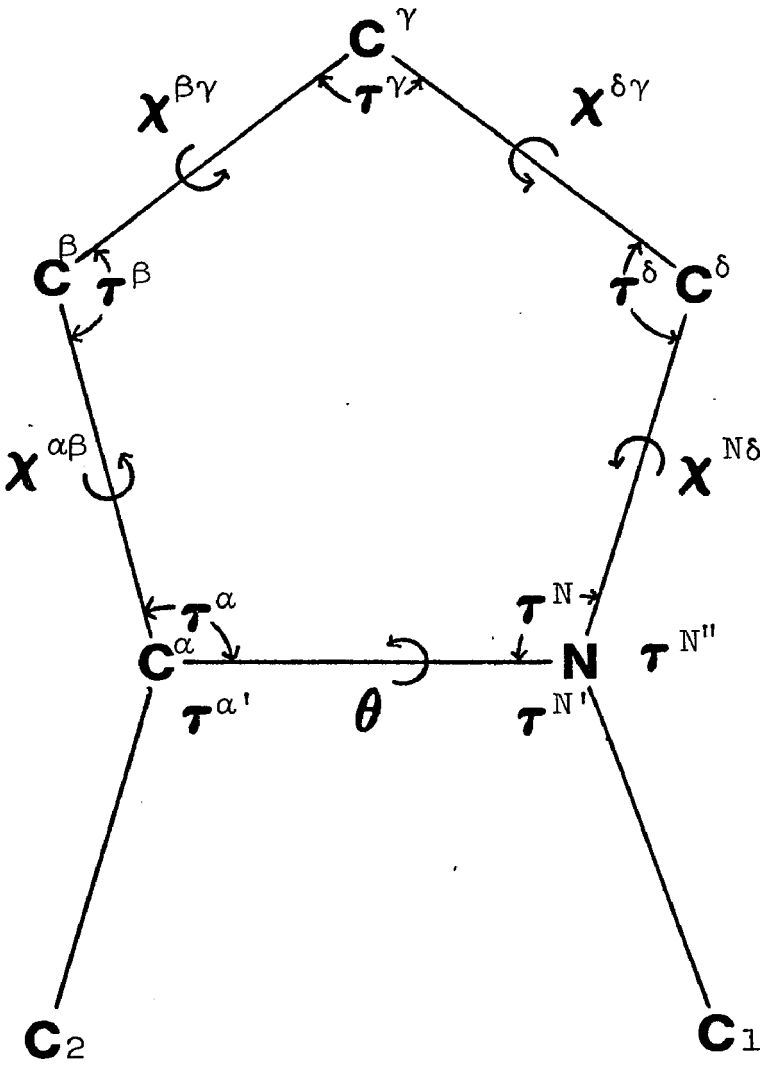


Figure 2-1. Proline ring with bond angles and torsion angles marked. The torsion angle θ refers to the atoms C^{β}-C^{α}-N-C^{δ}. φ refers to C_1-N-C^{α}-C_2 and is linearly related to θ by the relation $\varphi \approx -60^{o}+\theta$

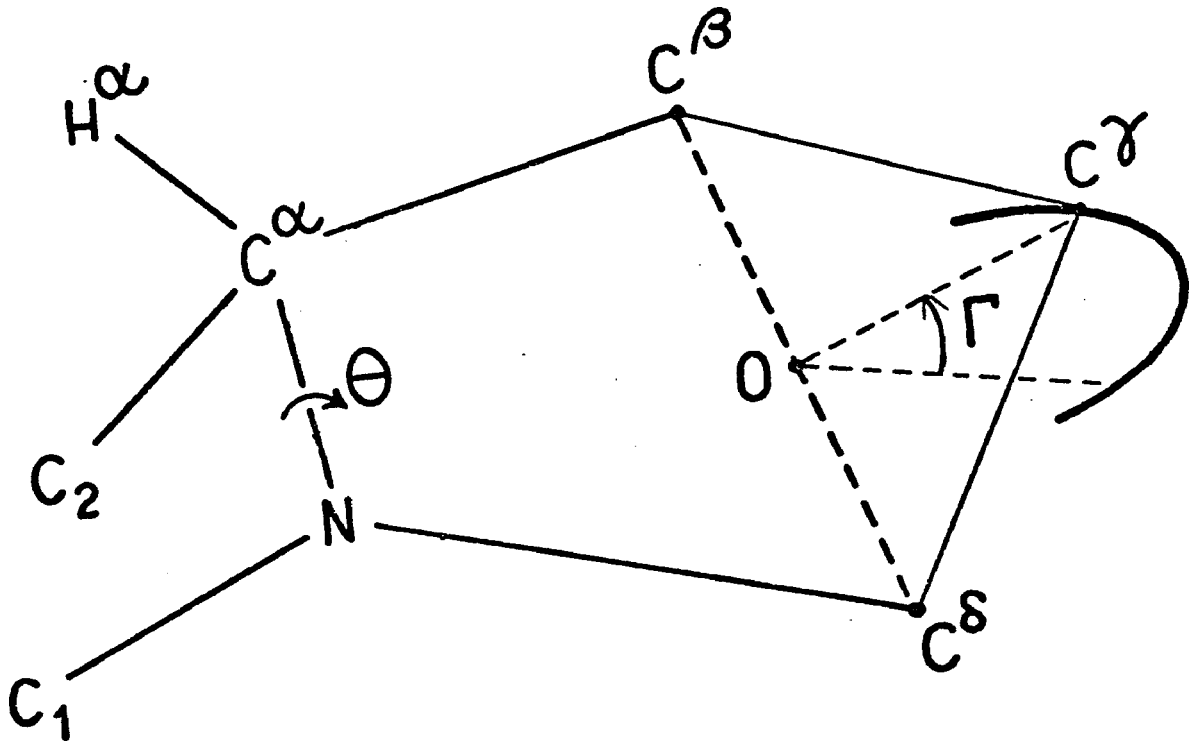


Figure 2-2. The geometry of the C^{γ} arcing. O is the center of the Γ-circle and lies on the line $C^{\delta}...C^{\beta}$. The methylene hydrogens are not shown here, though they have been included in the energy calculations.

the dihedral angle θ. Therefore the conformation of the ring skeleton is completely determined by τ^{α}, τ^{N}, θ and Γ. This 'arcing' of the C^{γ} atom along the prescribed circle was accomplished by a subroutine subprogram in the computer calculations, the mathematics and Fortran coding of which are contained in Appendix II. As one proceeds along the Γ-arcing circle, the bond angles τ^{β} and τ^{δ} continuously vary. This restricts the realizable positions of C^{γ} to only a part of the Γ-circle.

The hydrogens on C^{β}, C^{γ} and C^{δ} were fixed such that the HCH plane bisects the appropriate C-C-C or C-C-N angle. The H^{α} and C_2' were fixed in the same manner. All HCH angles as well as $\tau[H^{\alpha}C^{\alpha}C_2]$ were set at 110^{o}.

The position of the additional atom C_1 was determined by knowing $\tau^{N'}$. Here the question arises as to the orbital configuration around the nitrogen atom. In an interior prolyl residue, the three bonds at the nitrogen may be assumed to be in the same plane, according to the sp^2 hybrid nature of the orbitals. However, if the proline is at the N-terminal, the nitrogen assumes a sp^3 tetrahedral configuration, as in the case of the imino acid proline. The nature of ring puckering is significantly affected by the nitrogen configuration, as discussed in Section 2.5. For all interior dipeptides an sp^2 nitrogen atom will be assumed.

2.2 Classification of Ring Geometries

The C^{γ}-arcing procedure is quite convenient in that ring geometries can be classified in terms of θ rather than any other torsion angle. This is convenient because θ is linearly related to φ, the backbone torsion angle, by the relation $\varphi=-\mu+\theta$, where μ is a positive angle which depends upon $\tau^{N'}$ and $\tau^{\alpha'}$ and is $\sim 60^{o}$. On the other hand, while puckering of the ring will influence the preferred values of θ, non-bonded interactions in the chain will influence the value of φ. For every θ, ring geometries of various kinds of puckering are produced in a specific sequence as the angle Γ is varied. The simplest geometries structurally are those corresponding to puckering at a single atom X. By perfect X puckering in this five-membered ring, we mean that the torsion angle about the bond directly opposite to atom X is zero; i.e., all atoms except X are coplanar. There are two types of puckering, depending upon which side of the plane X is displaced. Using the notation of Torchia (44), if atom X lies on the same side of the ring plane as the carbonyl carbon C_2 (see Fig. 2-2), the ring is X^{-} puckered. Conversely, the ring is X^{+} puckered if X lies on the opposite side. In general, classification of ring geometries that are not single-atom puckered will be more involved. It is found useful to broadly divide this set of geometries into two categories, based on the sign of the ring dihedral angle $\chi^{\delta\gamma}$. (The reason that this particular dihedral angle is convenient is that all low-energy ring geometries are

either largely positive or largely negative in $\chi^{\delta\gamma}$.) Following Ramachandran et al. (10), all rings having negative $\chi^{\delta\gamma}$ are referred to as Type A geometries while those rings with positive values of $\chi^{\delta\gamma}$ are Type B geometries. Of course, this classification breaks down for perfectly α-puckered structures (i.e., when $\chi^{\delta\gamma}=0^{o}$). In low energy ring structures, this problem does not arise for as is shown in Section 2.6, α-puckered structures are highly unlikely. This is especially true for rings which survive in the proline dimer.

2.3 Contributions to the Potential Energy

The total conformational energy or 'self-energy' of the pyrrolidine ring is a function of five parameters:

$$E = E(\tau^{\alpha}, \tau^{N}, \tau^{N'}, \Gamma, \theta) \qquad (1)$$

This energy has been evaluated as a sum of four terms: the van der Waals energy (V_{vw}), electrostatic energy (V_{es}), bond angle deformation energy (V_{τ}) and intrinsic torsional energy (V_{χ}). Each of the terms which contribute to the conformational energy is discussed in Appendix I. For the ring self-energy calculations we have utilized Holzwarth and Chandrasekharan's (33) ring atom charges.

2.4 Details of the Calculation

The self-energy E of the rings was evaluated as a sum of the terms discussed in Appendix I. The bond lengths used are from Ramachandran et al. (10) and are listed in Table 2-1.

Table 2-1. Average Values of Proline Residue Bond Lengths Observed in Crystal Structures

BOND	LENGTH (Å)
$N-C^{\alpha}$	1.47
$C^{\alpha}-C^{\beta}$	1.53
$C^{\beta}-C^{\gamma}$	1.50
$C^{\gamma}-C^{\delta}$	1.52
$C^{\delta}-N$	1.48
N-C	1.32
C-H	1.10

The internal parameters of equation (1) were all varied. At this point, it becomes convenient to describe a notation for parametric variation which will be extensively used throughout the text of this thesis. Each parameter has a certain range over which it is varied and it usually has a series of increments within that range. For example, if the parameter p is varied from value x to value y in increments of length Δ, the notation for this variation will be $\{p: x, y, \Delta\}$. Using this, the ranges of the parameters used in calculating the total pyrrolidine ring energy may be represented as:

$$\{\tau^{\alpha}: 99^{\circ}, 107^{\circ}, 2^{\circ}\}$$
$$\{\tau^{N}: 109^{\circ}, 117^{\circ}, 2^{\circ}\}$$
$$\{\theta: -40^{\circ}, 40^{\circ}, 5^{\circ}\}$$
$$\{\Gamma: 0^{\circ}, 360^{\circ}, 2^{\circ}\}$$

Calculations show that θ could be restricted to between -40° and $+40^{\circ}$, while reasonable values of Γ are between -60° and $+60^{\circ}$.

2.5 Energy of the Isolated Ring as a Function of Γ and θ

Figure 2-3 is the Γ-θ energy plot for the isolated proline ring with a planar trigonal nitrogen. The energy is also a function of the bond angles τ^{α} and τ^{N}. For each (Γ,θ) the values of τ^{α} and τ^{N} giving the lowest ring energy were chosen for the plot. The global energy minimum (E=13.4 kcal/mole) is located at $\Gamma=-32^{\circ}$, $\theta=-10^{\circ}$ and is indicated by the symbol $*$. There is an additional slightly higher

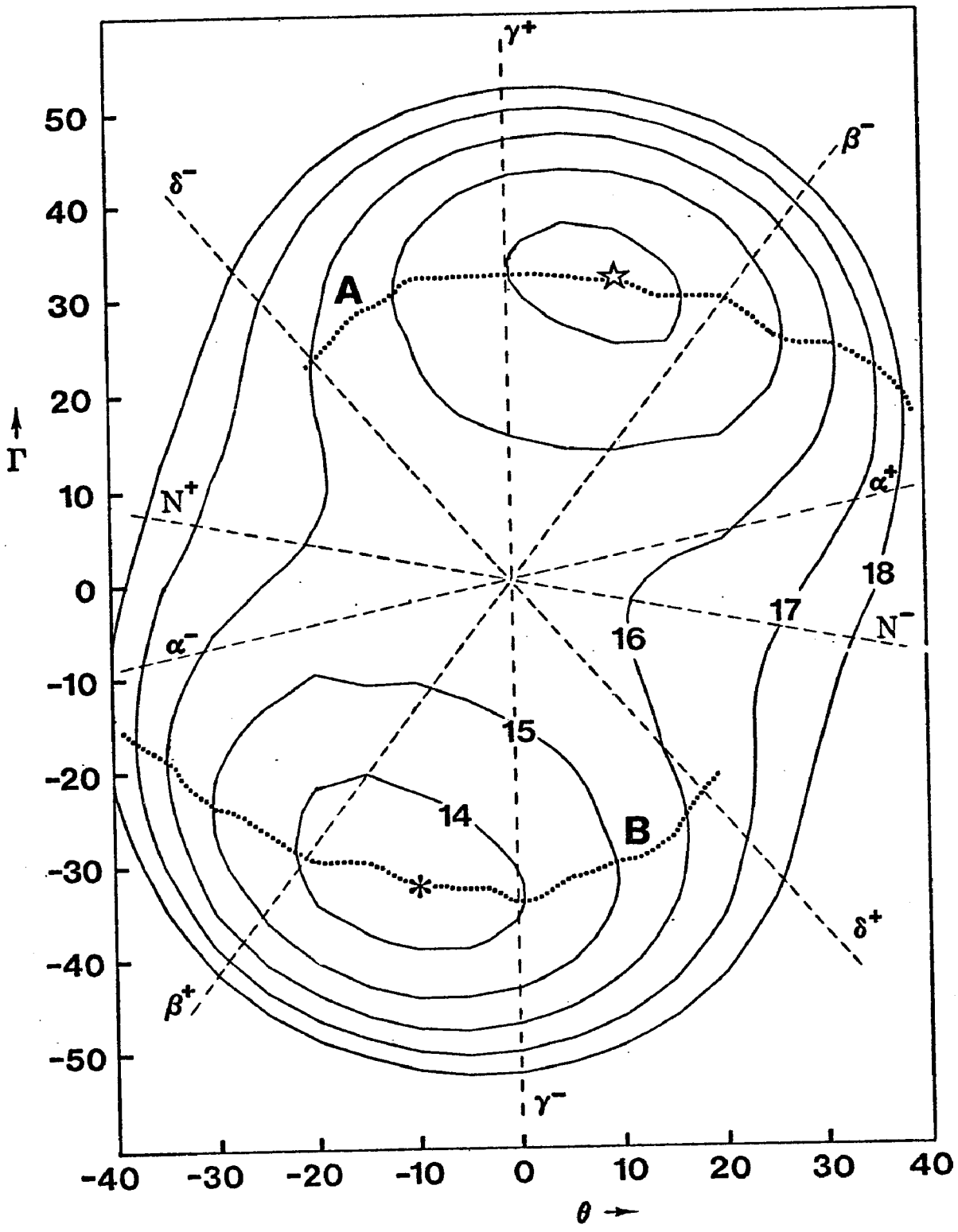


Figure 2-3. Γ-θ energy plot for the proline isolated ring with a planar trigonal nitrogen. Each (Γ,θ) energy has been minimized with respect to τ^{α} and τ^{N}.

local minimum with E=13.7 kcal/mole at $\Gamma=+32^{o}$, $\theta=+10^{o}$, indicated by a ☆ . The contours on the map are spaced at 1 kcal/mole intervals. The dotted lines, marked A and B, designate the (Γ,θ) positions of Type A and Type B minimum energy geometries. Another notation which is useful and will be referred to extensively is the use of 'M1' to designate the lowest energy geometry for each θ. The geometry whose θ corresponds to the secondary energy minimum for each θ is called the 'M2' geometry. Thus, in Figure 2-3, the M1 geometries are B type for $\theta \leq 0^{o}$ and are A type for $\theta > 0^{o}$. The dashed lines in Figure 2-3 indicate positions where the ring is completely puckered at one atom. The lines all converge to a point at $\Gamma=\theta=0^{o}$; that is, where the ring is completely planar. The signs and the type of puckering are marked at the ends of each line. The danger of using a planar ring in energy calculations is particularly evident in Figure 2-3, where the planar ring is 2.4 kcal/mole higher in energy than the global minimum energy ring. Since the division between Type A and Type B geometries is at the C^{α}-puckering line, all (Γ,θ) positions above this line correspond to Type A geometries while all those below the C^{α} line belong to Type B geometries.

In Figure 2-3, the plotted energies are minimized with respect to τ^{α} and τ^{N}. It is useful to see the values of these bond angles corresponding to the minimum energies. Accordingly, Figures 2-4 and 2-5 contain the Γ-θ angle plots for τ^{α} and τ^{N} respectively. The M1 and M2 points are

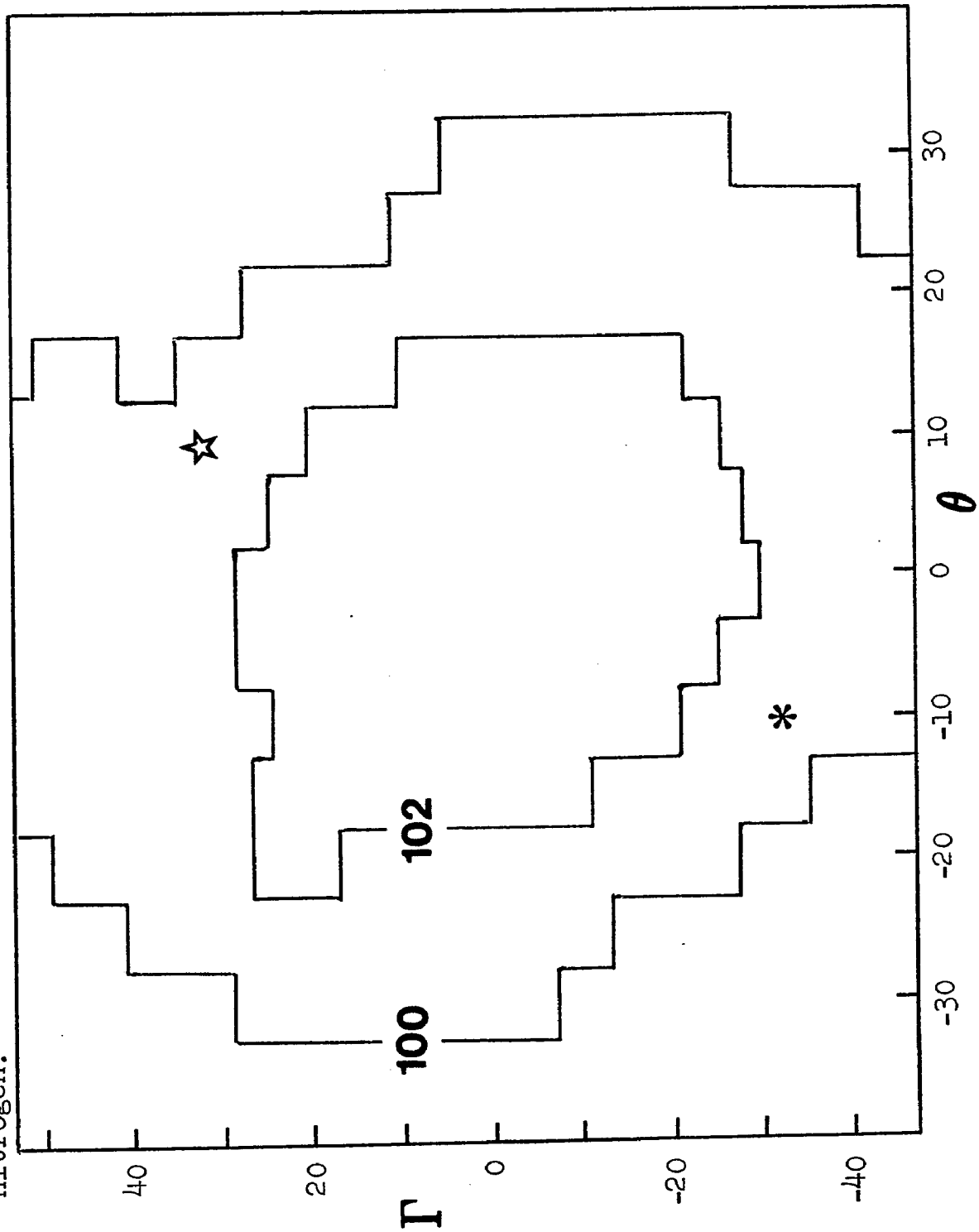


Figure 2-4. Γ-θ Angle Plot for τ^{α} in the proline isolated ring with a planar trigonal nitrogen.

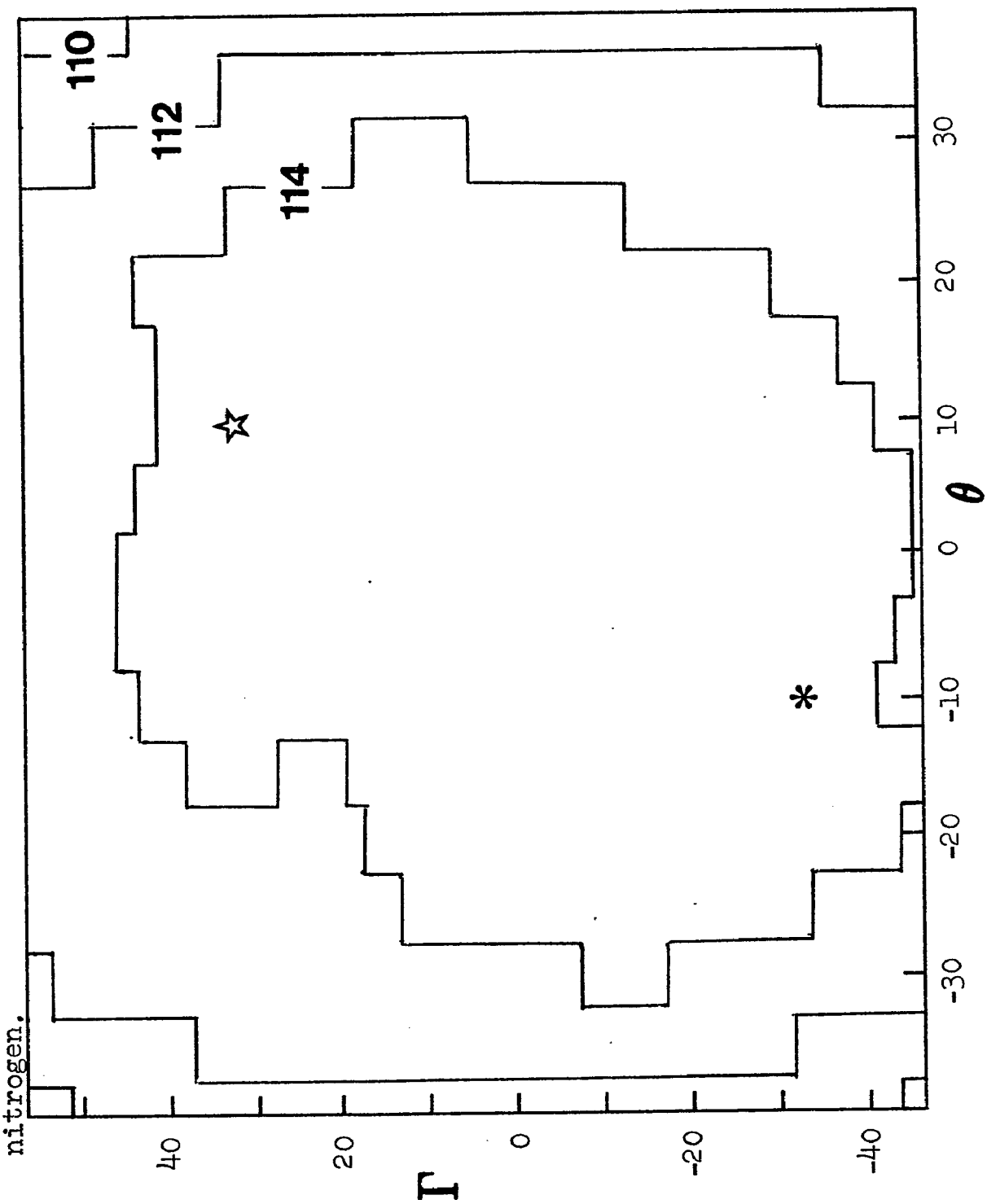


Figure 2-5. Γ-θ Angle Plot for τ^N in the proline isolated ring with a planar trigonal nitrogen.

marked as in Figure 2-3. As seen in Figure 2-4, $\tau^{\alpha}=101^{\circ}$ for both minima. However, as we approach the planar ring structure at $\Gamma=\theta=0^{\circ}$, τ^{α} becomes larger. In Figure 2-5, similar results are obtained. Both M1 and M2 minima prefer $\tau^{N}=115^{\circ}$. τ^{N} varies rapidly with θ only for values of $|\theta|>20^{\circ}$. In the minimum energy region ($-20^{\circ}<\theta<20^{\circ}$), only one value of τ^{N} is found for almost all values of Γ. Thus, it seems that we may, to a good approximation, keep these two bond angles constant in most of the calculations.

A closer inspection of the general shape of the Γ-θ energy plot (Figure 2-3), especially the positions of the puckered geometry lines, seems to suggest an approximate centrosymmetry about the point $\Gamma=\theta=0^{\circ}$; that is, the energies at any points (Γ,θ) and $(-\Gamma,-\theta)$ are very nearly equal. This is a reasonable result to anticipate in view of the fact that trigonality of the nitrogen confers a symmetry about $\theta=0^{\circ}$ and hence, about $\Gamma=0^{\circ}$ also. In fact, if the molecule were completely symmetric, one would expect a completely centrosymmetric set of energy contours. However, the molecule itself is not symmetric. The α-carbon differs from the methylene carbons in that it has two different substituents, namely C_2 and H^{α}, instead of two hydrogens.

If the nitrogen is fixed in a tetrahedral configuration, the global Γ-θ plot of Figure 2-6 is found. The ranges of parameters used in this calculation are:

$$\{\tau^{\alpha}: 103^{\circ}, 109^{\circ}, 2^{\circ}\}$$

$$\{\tau^{N}: 101^{\circ}, 109^{\circ}, 2^{\circ}\}$$

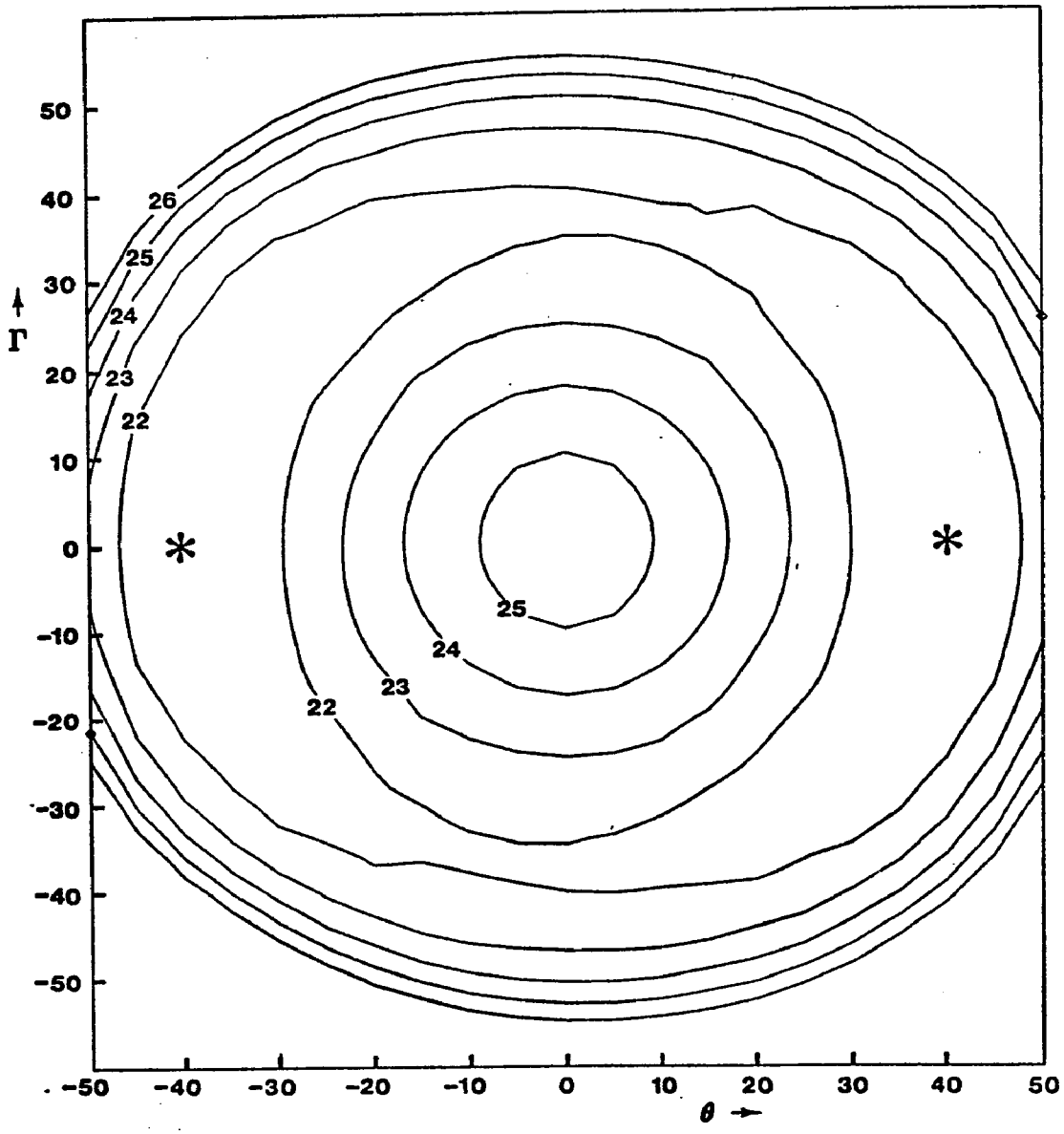


Figure 2-6. Γ-θ energy plot for the proline isolated ring with a sp^3 tetrahedral nitrogen. Each (Γ,θ) energy has been minimized with respect to τ^{α} and τ^{N}. Both minima are denoted by asterisks because they are of equal energy.

$$\{\theta: -50^{\circ}, 50^{\circ}, 5^{\circ}\}$$
$$\{\Gamma: -60^{\circ}, 60^{\circ}, 2^{\circ}\}$$

As before, each (Γ,θ) point is minimized over all τ^{α} and τ^{N} combinations. The value of K_{τ} used for bond angle bending of τ^{N} was 72 kcal/mole. This energy map is considerably different from that of Figure 2-3. The energy contours are essentially a set of concentric circles about the planar ($\Gamma=\theta=0^{\circ}$) structure. Any displacement from the center is a measure of the non-planarity of the ring; therefore, for a given displacement there is an infinite set of (Γ,θ) values of nearly equal energy. Thus, for a given degree of non-planarity, the ring can adopt any type of puckering. The two minima are of equal energy and occur for $\theta=\pm 40^{\circ}$ with $\Gamma=0^{\circ}$. θ can vary from -45° to $+45^{\circ}$ within 1 kcal/mole from the minima. This is considerably more freedom than was found for proline with a trigonal nitrogen.

For the case of the tetrahedral nitrogen, we may also consider how τ^{α} and τ^{N} vary with Γ and θ. Figures 2-7 and 2-8 contain the corresponding Γ-θ angle plots pertaining to Figure 2-6. The results here are very similar to the case of the trigonal nitrogen. τ^{α} prefers 103° while τ^{N} prefers 111° for the minimum energy structures. We also see that as the planar ring structure is approached, both τ^{α} and τ^{N} become larger. This effect is more pronounced here than in the trigonal nitrogen.

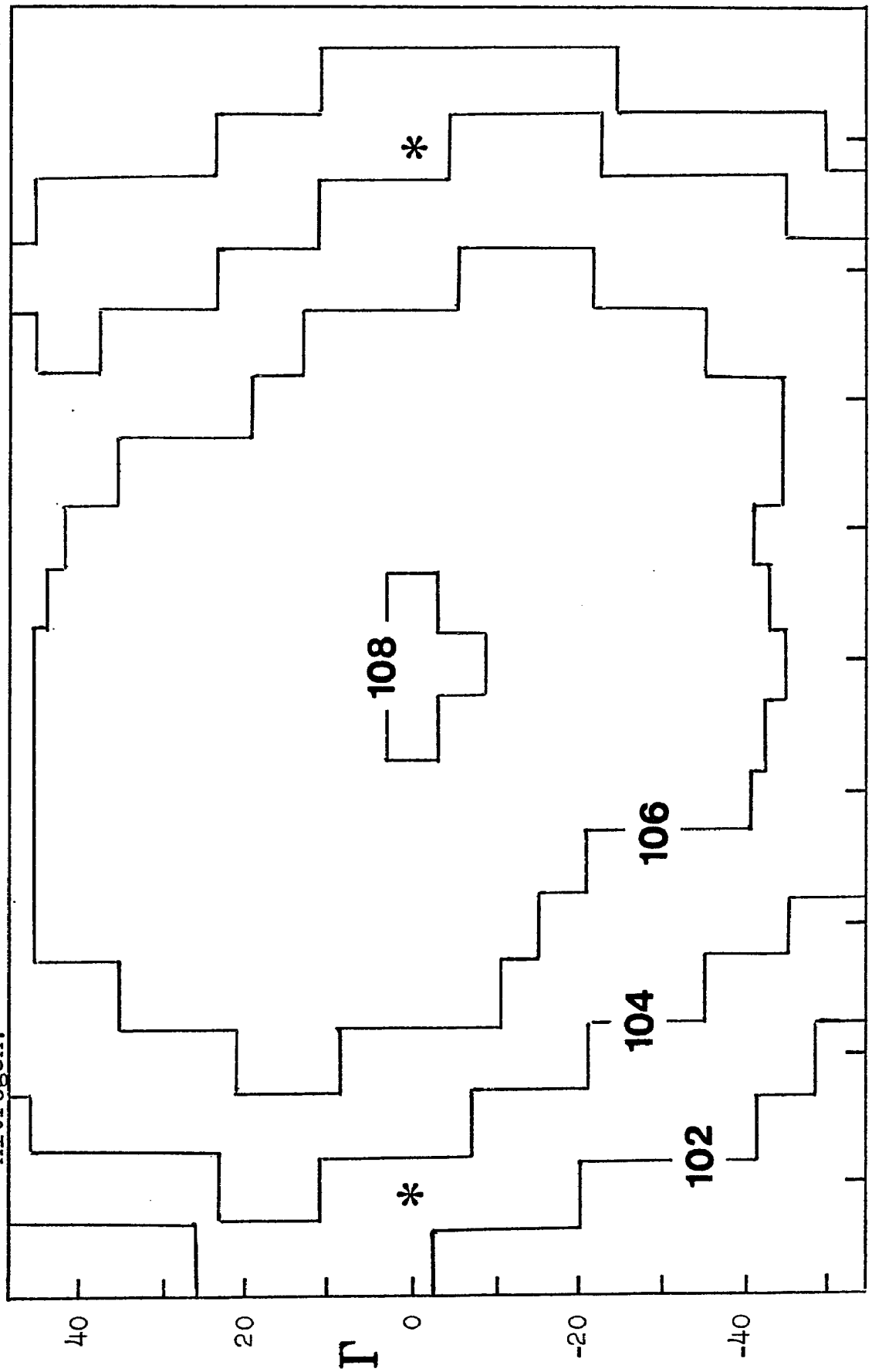


Figure 2-7. Γ-θ Angle Plot for τ^{α} in the proline isolated ring with a tetrahedral nitrogen.

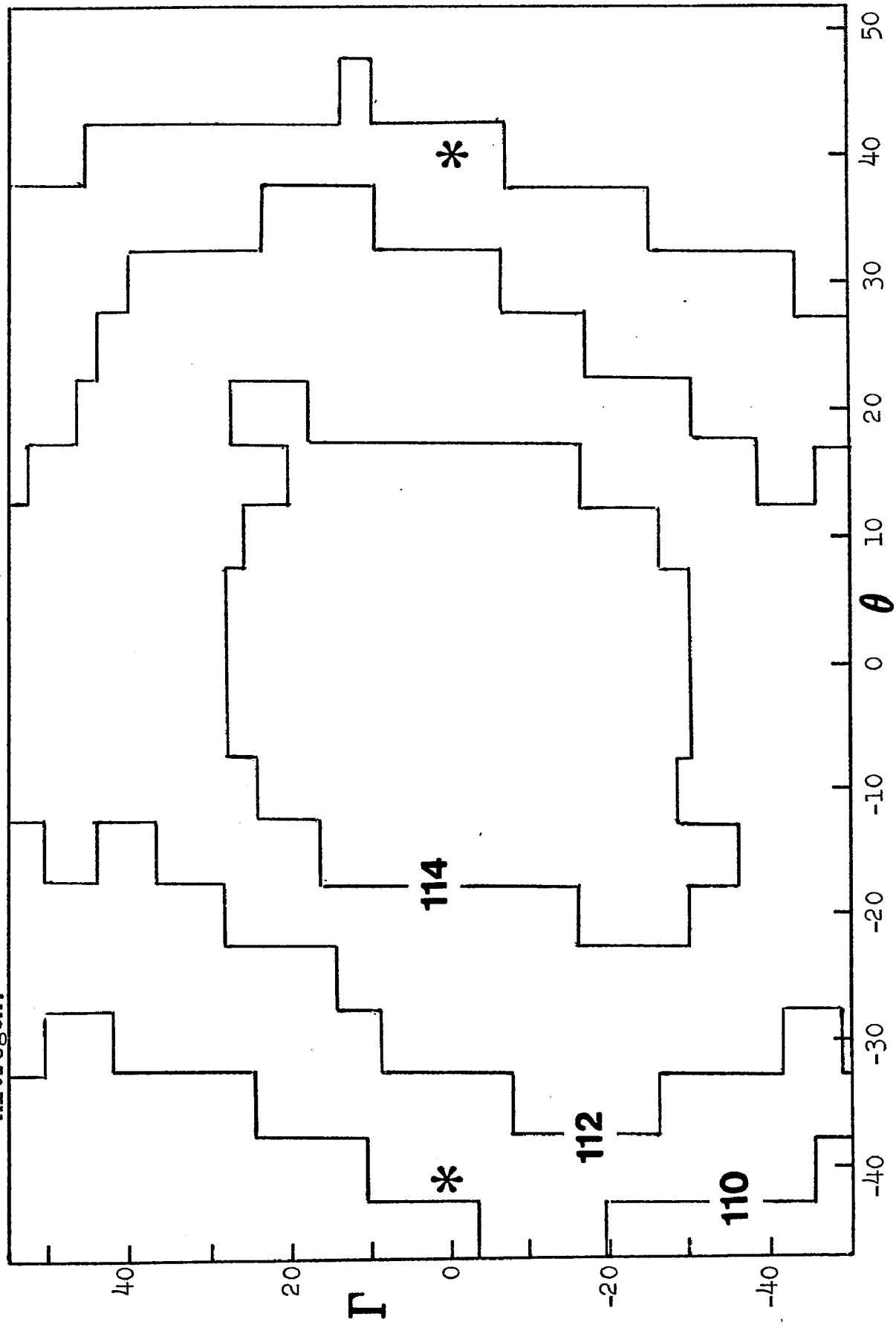


Figure 2-8. Γ-θ Angle Plot for τ^N in the proline isolated ring with a tetrahedral nitrogen.

2.6 Analysis of Bond Angles and Torsion Angles

It is useful to examine the nature of the bond angle and torsion angle distributions for the various single-atom puckered prolyl rings with planar trigonal nitrogen. The distributions for both Type A and Type B geometries having ring energies within 1.5 kcal/mole from the global minimum are shown in detail in Figure 2-9. In these calculations $\tau^{N'}$ was allowed to vary:

$$\{\tau^{N'}: 119^{o}, 127^{o}, 2^{o}\}$$

The broad horizontal lines at the bottom of each distribution show the spread of values of each angle found by X-ray analysis in proline containing peptides (45). In Figure 2-9b, the agreement is good between the experimental values and the peaks of the torsion angle distributions. It is interesting that there is an overlap in the A and B curves only in the case of the θ and $\chi^{N\delta}$ torsions. In other words, γ- and β-puckering are the only single-atom puckerings available to the ring within 1.5 kcal/mole of the global minimum. The calculated distribution peaks for the bond angles also agree well with the experimental spread (Fig. 2-9a). For every bond angle and torsion angle, the number of B type geometries within 1.5 kcal/mole of the minimum is greater than the number of A type geometries.

Table 2-2 contains a list of the torsion angles and bond angles for the lowest Type A and Type B geometries as a function of θ. The most striking feature of this table is that, in spite of an 80^{o} span for θ, the variation in the

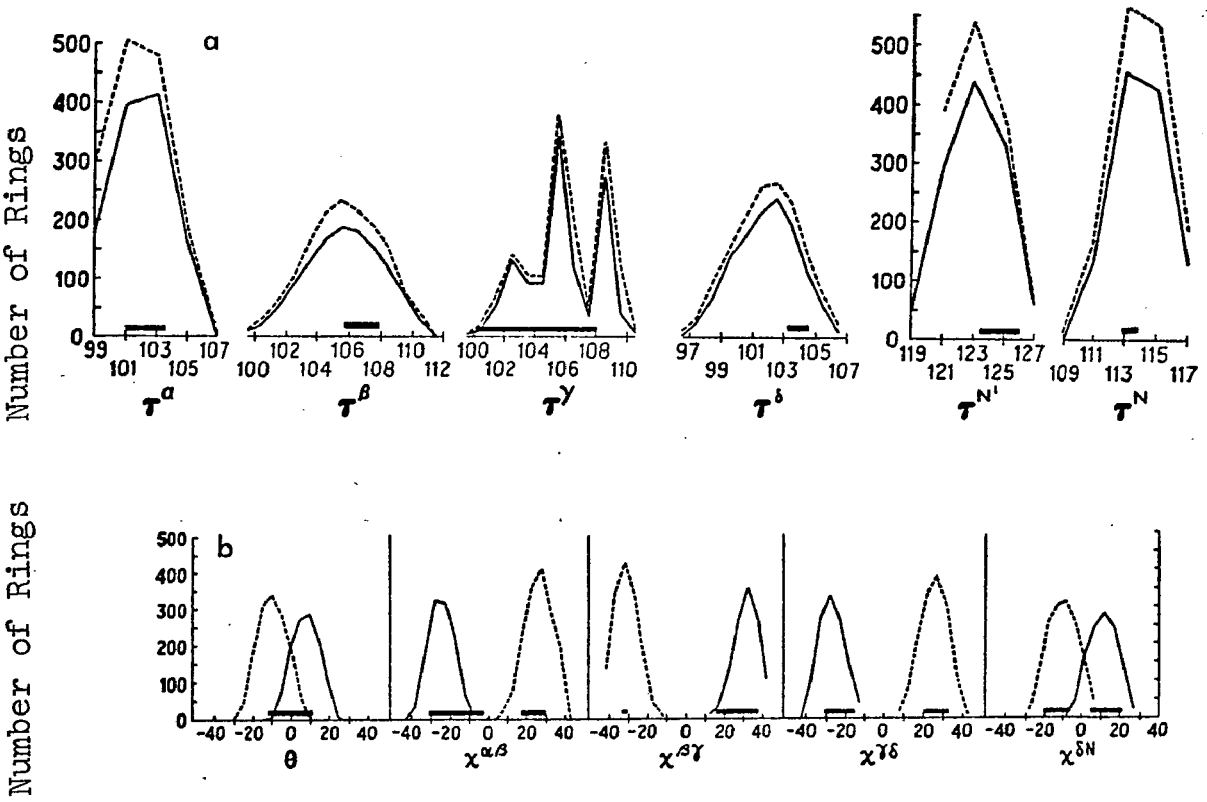


Figure 2-9. Distribution of bond angles and torsion angles in the set of proline rings with energy less than 1.5 kcal/mole from the global minimum energy: (———) Type A geometries; (------) Type B geometries. a) bond angles and b) torsion angles. The horizontal lines above the axis show the spread of values experimentally observed in proline containing crystal structures.

Table 2-2. Torsion Angles and Bond Angles (in degrees) of the Isolated Proline Ring Corresponding to Minimum Energy Ring Puckering

	Type A Geometries									
θ	$\tau^{N'}$	τ^{N}	τ^{α}	τ^{β}	τ^{γ}	τ^{δ}	$\chi^{\alpha\beta}$	$\chi^{\beta\gamma}$	$\chi^{\delta\gamma}$	$\chi^{\delta N}$
-40										
-35					--------*					
-30										
-25										
-20	123	113	101	109	105	100	-2	22	-32	34
-15	123	115	101	109	107	100	-5	23	-29	28
-10	123	115	101	108	106	99	-12	29	-33	27
-5	123	113	103	106	105	101	-17	32	-33	24
0	123	113	103	106	105	102	-21	34	-32	20
5	123	113	103	105	105	102	-24	34	-30	16
10	123	115	101	106	106	101	-27	35	-27	11
15	123	115	101	105	107	101	-30	35	-25	6
20	125	113	101	105	105	103	-35	36	-23	1
25	125	113	101	104	107	102	-38	34	-18	-5
30	125	113	99	104	106	102	-42	35	-16	-10
35	125	113	97	105	105	102	-41	35	-14	-15
40	125	113	97	103	108	100	-41	30	-7	-22

Table 2-2. (continued)

Type B Geometries

θ	$\tau^{N'}$	τ^{N}	τ^{α}	τ^{β}	τ^{γ}	τ^{δ}	$\chi^{\alpha\beta}$	$\chi^{\beta\gamma}$	$\chi^{\delta\gamma}$	$\chi^{\delta N}$
-40	123	113	97	104	108	100	40	-30	5	23
-35	123	113	99	104	108	101	38	-30	9	17
-30	123	113	99	105	106	102	38	-35	16	10
-25	123	113	101	104	107	102	35	-34	18	5
-20	123	113	101	105	105	103	34	-36	23	-1
-15	123	115	101	105	107	101	29	-34	23	-5
-10	123	115	101	106	106	101	27	-35	27	-11
-5	123	113	103	105	105	102	24	-34	30	-16
0	123	113	103	106	105	102	20	-32	31	-19
5	123	113	103	106	105	101	17	-31	33	-24
10	123	115	101	108	106	99	12	-29	33	-27
15	123	115	101	109	107	100	5	-23	30	-28
20	125	113	101	109	105	100	2	-22	32	-33
25										
30					--------*					
35										
40										

*The minimum does not exist.

bond angles is small compared to the torsion angle spread. The effect of $\tau^{N'}$ variation on the puckering characteristics is minor when compared to τ^{α} and τ^{N}. In fact, as τ^{α} and τ^{N} decrease, the torsion angles of the low energy rings increase. Without any loss of generality we can take an average value for the bond angles of the isolated ring from Table 2-2. Specifically, we can set $\tau^{\alpha}=103^{o}$, $\tau^{N}=113^{o}$ and $\tau^{N'}=123^{o}$. If we hold these three bond angles fixed, it is more convenient to study the energy of the ring as a function of just two parameters, Γ and θ. In fact, when energies of the proline dimer are considered in the next chapter, it is advantageous in some cases to continue using just one set of bond angles, especially from the savings in computer time it affords.

Figure 2-10 is the Γ-θ energy contour plot for the isolated proline ring with the bond angles $\tau^{\alpha}=103^{o}$, $\tau^{N}=113^{o}$ and $\tau^{N'}=123^{o}$. It is important to note that this map differs only slightly from the global Γ-θ energy map of Figure 2-3. This justifies our use of constant τ^{α}, τ^{N} and $\tau^{N'}$ in further ring self-energy calculations. The values of the energy for each Γ and θ are contained in Table 2-3.

Another view of the energy distribution of the isolated ring with θ can be seen by plotting total ring energy as a function of θ for the various perfectly puckered geometries as well as Types A and B geometries. This corresponds to taking vertical slices of the global Γ-θ plot (Figure 2-3) along the dashed and dotted lines. Figure 2-11 provides this

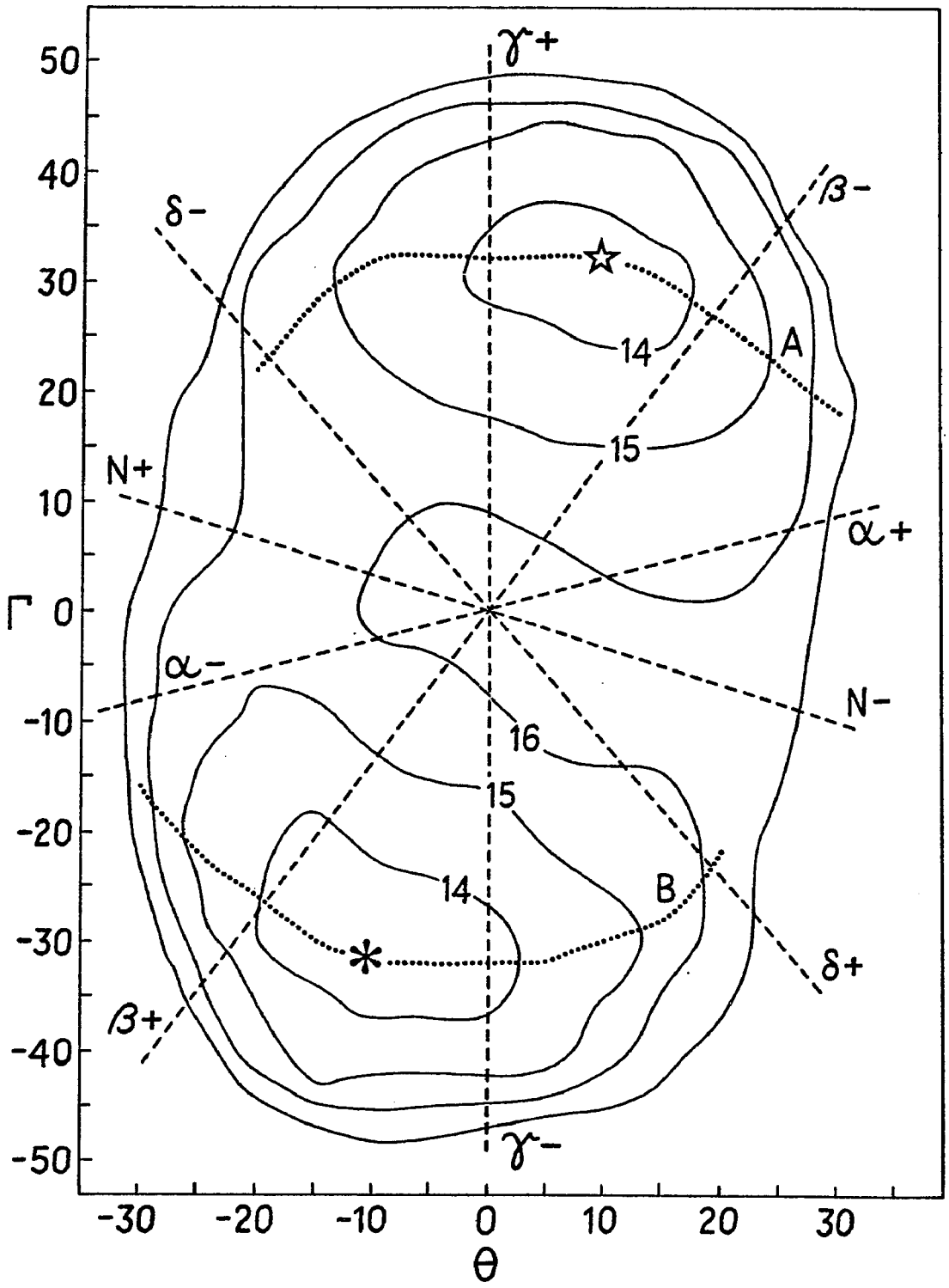


Figure 2-10. Γ-θ energy plot for the proline isolated ring. $\tau^{\alpha}=103^{\circ}$, $\tau^{N}=113^{\circ}$ and $\tau^{N'}=123^{\circ}$.

Table 2-3. Γ-θ Energies for the Isolated Proline Ring (in kcal/mole)

Γ \ θ	-30	-25	-20	-15	-10	-5	0	5	10	15	20	25	30
60	35.8	31.6	28.6	26.4	24.9	24.0	23.7	23.9	24.7	26.2	28.7	32.5	38.2
50	26.9	23.6	21.2	19.5	18.3	17.5	17.2	17.2	17.6	18.5	20.2	22.6	26.3
40	21.9	19.3	17.4	16.1	15.2	14.6	14.2	14.1	14.3	14.8	15.9	17.7	20.5
30	19.4	17.2	16.0	15.1	14.6	14.2	13.9	13.6	13.6	13.8	14.4	15.7	17.8
20	18.1	16.6	15.7	15.3	15.1	15.0	14.8	14.6	14.4	14.3	14.5	15.3	16.9
10	17.5	16.3	15.8	15.8	15.9	16.0	16.0	15.8	15.6	15.3	15.3	15.8	17.1
0	16.9	15.9	15.6	15.7	16.0	16.3	16.5	16.5	16.3	16.1	16.0	16.4	17.6
-10	16.4	15.2	14.9	15.0	15.3	15.7	16.0	16.1	16.1	16.1	16.3	16.8	18.1
-20	16.3	14.8	14.1	13.9	14.1	14.5	14.8	15.1	15.4	15.7	16.2	17.1	18.8
-30	17.3	15.2	14.0	13.5	13.4	13.5	13.8	14.3	14.8	15.5	16.4	17.9	20.0
-40	20.5	17.5	15.7	14.6	14.1	14.0	14.3	14.7	15.5	16.5	17.9	19.9	22.6
-50	28.2	23.5	20.5	18.6	17.6	17.2	17.2	17.7	18.6	19.9	21.7	24.2	27.6
-60	49.8	38.0	31.4	27.6	25.4	24.3	24.0	24.3	25.3	26.8	29.1	32.2	36.5

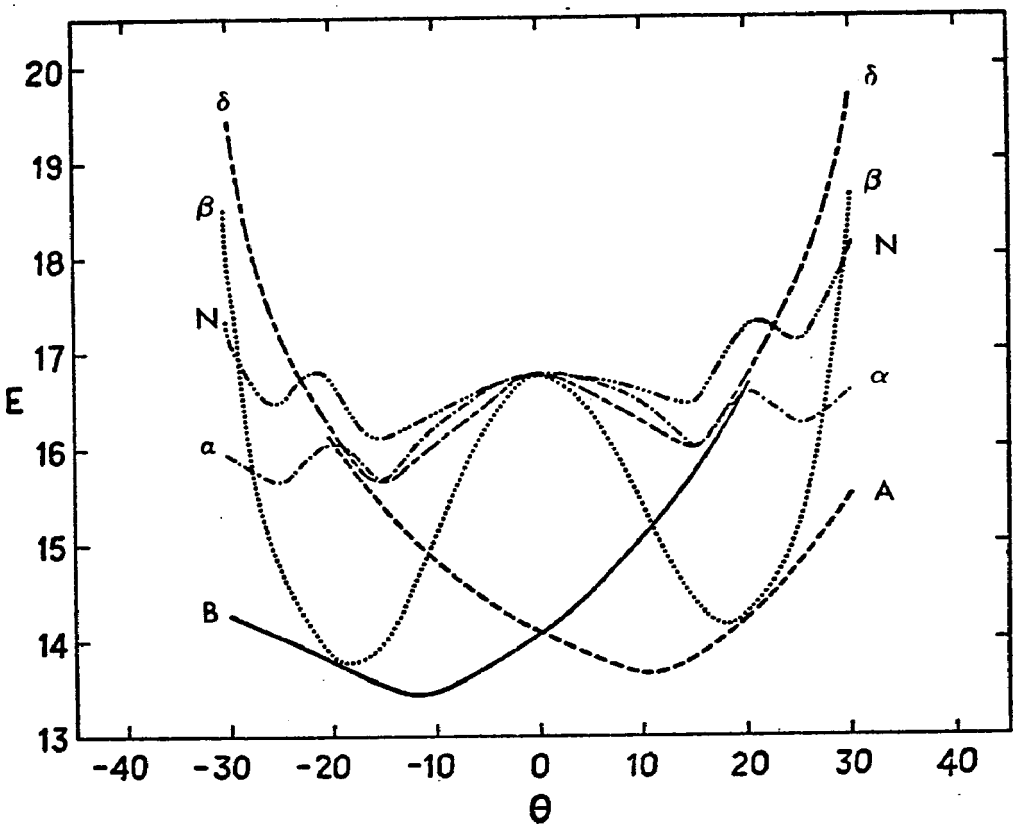


Figure 2-11. Self-energy of the variously puckered proline rings as a function of θ. Puckering types are marked at the end of each line.

information. It is more clearly seen from this figure that there are two well separated Γ values for which the self-energy is a minimum. It is also clear from this figure, as well as Figure 2-3, that the global energy minimum ring structure is not single-atom puckered. This has already been noted by Torchia (44) who has estimated the torsion angles in the prolyl ring from coupling constants obtained from nuclear magnetic resonance spectra of polyproline. Note that α-, δ- and N-puckered structures have ring energies that are at least 2 kcal/mole higher than the minimum energy geometry. It is unlikely, therefore, than an otherwise uncrowded molecule will adopt an α-, δ- or N-puckered ring structure. This is perhaps the most significant aspect of the study of the isolated ring; that is, the ability to predict the type of puckering in an unstrained molecule. As seen from Figure 2-10, for values of θ between -10° and 0°, the unstrained ring will be approximately γ^{-} puckered, while for $\theta \sim -20^{\circ}$, the ring will be close to β^{+} puckered. Following the Type A geometry line for $0^{\circ} < \theta < 10^{\circ}$, the unstrained ring is likely to be γ^{+} puckered and similarly, for $\theta \sim +20^{\circ}$, β^{-} puckering will likely be found. This analysis does not rule out the possibility of α-, δ- or N-puckered rings in an unusually strained system, where the ring might possibly adopt one of these geometries in order to minimize the total conformational energy of the molecule. The types of puckering found in the isolated ring are summarized in Figure 2-12.

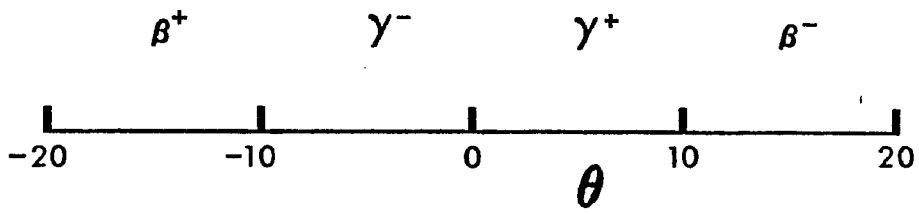


Figure 2-12. Puckering in the isolated proline ring.

The results presented here are in excellent agreement with an earlier study of ring puckering by Ramachandran et al. (10). Another set of ring energy calculations has been reported by Nishikawa and Ooi (46). Their C^{γ} positioning method was noticeably different in that they imposed the condition that either $\chi^{N\delta}=0^{o}$ or $\chi^{\alpha\beta}=0^{o}$. These two C^{γ} positions were termed $\gamma 1$ and $\gamma 2$ respectively. In terms of the present study, this restriction means that only β-puckered and δ-puckered rings were considered. In Figure 2-10, for the regions of low ring energy, i.e., for values of θ around -10^{o} and $+10^{o}$, both β- and δ-puckered structures are at least 1.5 kcal/mole higher in energy than the minimum Types A and B ring geometries. This discrepancy becomes even more pronounced as θ approaches 0^{o}, where Nishikawa and Ooi's restrictions lead to the perfectly planar ring which is more than 3 kcal/mole higher in energy at $\theta=0^{o}$ than our minimum energy structure. Care should be taken in utilizing the Nishikawa-Ooi geometries for energy calculations on polymers containing proline as results could be misleading.

Tables 2-4 through 2-7 contain the values of the torsion angles $\chi^{\alpha\beta}$, $\chi^{\beta\gamma}$, $\chi^{\delta\gamma}$ and $\chi^{N\delta}$ as functions of Γ and θ. Inspection shows the (Γ-θ) line where complete puckering is found, i.e., where each torsion angle is approximately zero, as indicated by dashed lines in Figure 2-10. For $\chi^{\beta\gamma}$ and $\chi^{\delta\gamma}$ (Tables 2-4 and 2-6) all angles above the completely puckered line are negative; all below are positive. The reverse is true for $\chi^{\alpha\beta}$ and $\chi^{N\delta}$ (Tables 2-5 and 2-7). The torsion angles $\chi^{\alpha\beta}$ and $\chi^{N\delta}$ which flank θ, have perfectly puckered lines with large slopes and, at extreme values of Γ and θ on either ends of these lines, the torsion angles are nearly zero. On the other hand, for $\chi^{\beta\gamma}$ and $\chi^{\delta\gamma}$, extreme values of Γ and θ yield correspondingly large values for the torsion angles.

In the context of the above discussion, it is interesting to compare the values of the torsion angles found in X-ray analysis of crystal structures to those calculated for the isolated ring. The planar trigonal nitrogen assumed in these calculations is appropriate for a prolyl ring in a peptide with the partial double bond C-N. The torsion angles from X-ray analysis were taken from the literature of peptides containing proline with a trigonal nitrogen. Table 2-8 provides this comparison for the cases of five structures. In three of them, Leu-Pro-Gly (Type B), cyclo(Pro-Leu) and tosyl Pro-Hyp, the agreement between calculated and observed torsion angles is excellent. In both Gly-Pro-Leu-Gly and Leu-Pro-Gly (Type A), the torsion

Table 2-4. Γ-θ Torsion Angles for $\chi^{\alpha\beta}$ (in degrees)

Γ \ θ	-30	-25	-20	-15	-10	-5	0	5	10	15	20	25	30
60	-8.8	-12.8	-16.9	-20.8	-24.8	-28.5	-32.2	-35.6	-38.8	-41.7	-44.3	-46.7	-48.7
50	-4.7	-8.7	-12.7	-16.8	-20.8	-24.7	-28.4	-31.9	-35.3	-38.3	-41.1	-43.6	-45.8
40	0.1	-3.9	-7.9	-12.0	-16.1	-20.0	-23.9	-27.6	-31.0	-34.3	-37.2	-39.9	-42.2
30	5.3	1.4	-2.5	-6.6	-10.7	-14.7	-18.7	-22.5	-26.1	-29.5	-32.6	-35.5	-38.0
20	10.9	7.2	3.3	-0.7	-4.8	-8.8	-12.9	-16.8	-20.5	-24.1	-27.4	-30.5	-33.2
10	16.7	13.1	9.4	5.6	1.6	-2.5	-6.6	-10.6	-14.5	-18.2	-21.7	-25.0	-28.0
0	22.4	19.2	15.6	11.9	8.1	4.1	0.0	-4.1	-8.1	-11.9	-15.6	-19.2	-22.4
-10	28.0	25.0	21.7	18.2	14.5	10.6	6.6	2.5	-1.6	-5.6	-9.4	-13.1	-16.7
-20	33.2	30.5	27.4	24.1	20.5	16.8	12.9	8.8	4.8	0.7	-3.3	-7.2	-10.9
-30	38.0	35.5	32.6	29.5	26.1	22.5	18.7	14.7	10.7	6.6	2.5	-1.4	-5.3
-40	42.2	39.9	37.2	34.3	31.0	27.6	23.9	20.0	16.1	12.0	7.9	3.9	-0.1
-50	45.8	43.6	41.1	38.3	35.3	31.9	28.4	24.7	20.8	16.8	12.7	8.7	4.7
-60	48.7	46.7	44.3	41.7	38.8	35.6	32.2	28.5	24.8	20.8	16.9	12.8	8.8

Table 2-5. Γ-θ Torsion Angles for $\chi^{\beta\gamma}$ (in degrees)

Γ	θ												
	-30	-25	-20	-15	-10	-5	0	5	10	15	20	25	30
60	46.5	48.3	50.1	51.9	53.6	55.1	56.6	57.9	59.0	60.1	61.0	61.9	62.6
50	38.8	39.7	41.5	43.4	45.1	46.7	48.2	49.6	50.8	51.9	52.9	53.7	54.5
40	28.8	30.7	32.5	34.4	36.2	37.8	39.4	40.9	42.2	43.3	42.6	43.6	44.4
30	19.5	21.3	23.1	25.0	26.8	28.5	30.1	31.7	33.1	34.4	35.5	36.5	37.5
20	9.9	11.6	13.4	15.2	16.9	18.7	20.4	22.0	23.5	24.9	26.2	27.4	28.4
10	0.2	1.8	3.4	5.1	6.8	8.6	10.3	12.0	13.6	15.1	16.6	17.9	19.1
0	-9.5	-8.1	-6.6	-5.1	-3.4	-1.7	0.0	1.7	3.4	5.1	6.6	8.1	9.5
-10	-19.1	-17.9	-16.6	-15.1	-13.6	-12.0	-10.3	-8.6	-6.8	-5.1	-3.4	-1.8	-0.2
-20	-28.4	-27.4	-26.2	-24.9	-23.5	-22.0	-20.4	-18.7	-16.9	-15.2	-13.4	-11.6	-9.9
-30	-37.5	-36.5	-35.5	-34.4	-33.1	-31.7	-30.1	-28.5	-26.8	-25.0	-23.1	-21.3	-19.5
-40	-46.1	-45.3	-44.4	-43.3	-42.2	-40.9	-39.4	-37.8	-36.2	-34.4	-32.5	-30.7	-28.8
-50	-54.5	-53.7	-52.9	-51.9	-50.8	-49.6	-48.2	-46.7	-45.1	-43.4	-41.5	-39.7	-37.8
-60	-62.6	-61.9	-61.0	-60.1	-59.0	-57.9	-56.6	-55.1	-53.6	-51.9	-50.1	-48.3	-46.5

Table 2-6. Γ-θ Torsion Angles for $\chi^{\delta\gamma}$ (in degrees)

Γ \ θ	-30	-25	-20	-15	-10	-5	0	5	10	15	20	25	30
60	-59.1	-58.6	-58.0	-57.3	-56.5	-55.5	-54.5	-53.4	-52.1	-50.7	-49.3	-47.8	-46.3
50	-51.1	-50.6	-49.9	-49.2	-48.4	-47.4	-46.3	-45.1	-43.8	-42.4	-41.0	-39.5	-38.0
40	-42.9	-42.3	-41.6	-40.8	-39.9	-38.9	-37.8	-36.5	-35.2	-33.8	-32.3	-30.8	-29.3
30	-34.4	-33.7	-33.0	-32.1	-31.1	-30.0	-28.8	-27.5	-26.1	-24.7	-23.3	-21.8	-20.4
20	-25.6	-24.8	-24.0	-23.0	-21.9	-20.7	-19.5	-18.1	-16.8	-15.3	-13.9	-12.6	-11.2
10	-16.6	-15.7	-14.7	-13.6	-12.4	-11.1	-9.8	-8.5	-7.1	-5.7	-4.4	-3.2	-1.9
0	-7.4	-6.3	-5.2	-3.9	-2.7	-1.3	0.0	1.3	2.7	3.9	5.2	6.3	7.4
-10	1.9	3.2	4.4	5.7	7.1	8.5	9.8	11.1	12.4	13.6	14.7	15.7	16.6
-20	11.2	12.6	13.9	15.3	16.8	18.1	19.5	20.7	21.9	23.0	24.0	24.8	25.6
-30	20.4	21.8	23.3	24.7	26.1	27.5	28.8	30.0	31.1	32.1	33.0	33.7	34.4
-40	29.3	30.8	32.3	33.8	35.2	36.5	37.8	38.9	39.9	40.8	41.6	42.3	42.9
-50	38.0	39.5	41.0	42.4	43.8	45.1	46.3	47.4	48.4	49.2	49.9	50.6	51.1
-60	46.3	47.8	49.3	50.7	52.1	53.4	54.5	55.5	56.5	57.3	58.0	58.6	59.1

Table 2-7. Γ-θ Torsion Angles for $\chi^{N\delta}$ (in degrees)

Γ	θ: -30	-25	-20	-15	-10	-5	0	5	10	15	20	25	30
60	50.2	47.7	45.0	42.0	38.8	35.3	31.6	27.7	23.7	19.6	15.4	11.2	7.0
50	47.0	44.4	41.6	38.5	35.1	31.5	27.7	23.8	19.7	15.5	11.3	7.1	2.9
40	43.2	40.5	37.5	34.3	30.8	27.1	23.2	19.2	15.0	10.8	6.6	2.4	-1.7
30	38.9	36.0	32.9	29.5	25.9	22.0	18.1	14.0	9.8	5.6	1.4	-2.8	-6.8
20	34.1	31.0	27.7	24.1	20.4	16.4	12.4	8.2	4.0	-0.2	-4.3	-8.3	-12.2
10	28.8	25.6	22.1	18.4	14.5	10.4	6.3	2.1	-2.0	-6.2	-10.2	-14.1	-17.7
0	23.4	19.9	16.2	12.3	8.3	4.2	0.0	-4.2	-8.3	-12.3	-16.2	-19.9	-23.4
-10	17.8	14.1	10.2	6.2	2.0	-2.1	-6.3	-10.4	-14.5	-18.4	-22.1	-25.6	-28.8
-20	12.2	8.3	4.3	0.2	-4.0	-8.2	-12.4	-16.4	-20.4	-24.1	-27.7	-31.0	-34.1
-30	6.8	2.8	-1.4	-5.6	-9.8	-14.0	-18.1	-22.0	-25.9	-29.5	-32.9	-36.0	-38.9
-40	1.7	-2.4	-6.6	-10.8	-15.0	-19.2	-23.2	-27.1	-30.8	-34.3	-37.5	-40.5	-43.2
-50	-2.9	-7.1	-11.3	-15.5	-19.7	-23.8	-27.7	-31.5	-35.1	-38.5	-41.6	-44.4	-47.0
-60	-7.0	-11.2	-15.4	-19.6	-23.7	-27.7	-31.6	-35.3	-38.8	-42.0	-45.0	-47.7	-50.2

Table 2-8. Comparison with Ring Torsions Experimentally Found in Some Crystals

Crystal	θ	Type	Observed				Calculated[a]			
			$\chi^{\alpha\beta}$	$\chi^{\beta\gamma}$	$\chi^{\gamma\delta}$	$\chi^{\delta N}$	$\chi^{\alpha\beta}$	$\chi^{\beta\gamma}$	$\chi^{\gamma\delta}$	$\chi^{\delta N}$
Leu-Pro-Gly[b]	-11.1	A	-3.4	16.1	-21.8	21.0	-10	26	-31	27
		B	26.1	-31.4	23.8	-8.2	27	-34	26	-9
Gly-Pro-Leu-Gly[c]	6.3	A	-17.3	22.1	-17.5	6.2	-24	34	-29	14
cyclo(Pro-Leu)[d]	18.5	A	-31.5	36.0	-25.1	4.5	-31	34	-22	2
Tosyl-Pro-Hyp[e]	0.3	B	20.6	-32.4	31.2	-20.0	21	-34	32	-20

[a]These dihedral angles correspond to the minimum energy structure for the corresponding value of θ.

[b]Reference 28

[c]Reference 31

[d]Reference 32

[e]Unpublished results of M.N. Sabesan and K. Venkatesan reported in reference 10

angles found in the X-ray analysis are considerably different than the calculated angles. For the isolated ring, the minimum energy ring for $\theta=6.3^{\circ}$ has $\Gamma=32^{\circ}$. From the X-ray coordinate data for Gly-Pro-Leu-Gly (31), Γ was calculated to be 18°. The difference in energy between the $\Gamma=18^{\circ}$ and $\Gamma=32^{\circ}$ positions is only about 1 kcal/mole. It was therefore suspected that additional interactions between C^{γ} and other atoms in the crystal could easily force the C^{γ} atom away from its minimum energy position. A calculation was done in which the observed symmetry of the crystal was employed to generate neighboring molecules. The C^{γ} atom was then Γ-arced in the crystal, including all intramolecular and intermolecular interactions. The minimum in total energy was found to be at $\Gamma=18^{\circ}$, in exact agreement with X-ray analysis. This result seems to confirm the validity of the energy calculations as well as to indicate that, for a given θ, crystal interactions can easily perturb the C^{γ} atom from its minimum energy position of Table 2-2. It is apparent therefore, that discrepancies between observed and calculated proline ring structure may be explained by invoking crystal interactions not present in the isolated ring, but it also complicates any comparisons made between calculated and crystal structures.

CHAPTER III

RING PUCKERING AND THE PRO-PRO INTERNAL DIPEPTIDE

3.1 Structure and Coordinates of the Pro-Pro Dimer

The lowest energy prolyl ring geometries of the isolated proline residue may not necessarily be those preferred by the proline internal dipeptide. We have therefore extended the idea of allowing flexibility for the prolyl ring to the dimer. The structure of the Pro-Pro internal dipeptide is shown in Figure 3-1. The conformational energy of the dimer is a function of the parameters ω', φ_1, Γ_1, ψ, ω, φ_2, Γ_2 and the bond angles $\tau[C^{\alpha}C\ N]$, τ^{N}, $\tau^{N'}$ and $\tau^{\alpha'}$. Initially, to simplify the calculation and reduce computing time, we impose the requirement that both rings have the same puckering.* Therefore, both ring structures are specified by one Γ and θ. This means, according to Table 2-2, that $\tau^{\alpha'}$, τ^{N} and $\tau^{N'}$ are also known. It is more convenient to express energies in terms of θ, rather than φ, so that the conformation of the preferred rings found in the proline dimer may be easily compared to isolated ring structures whose energies have already been studied as a function of θ and Γ.

*Requiring identical puckering for both rings is useful for computing the conformational energy of a helix.

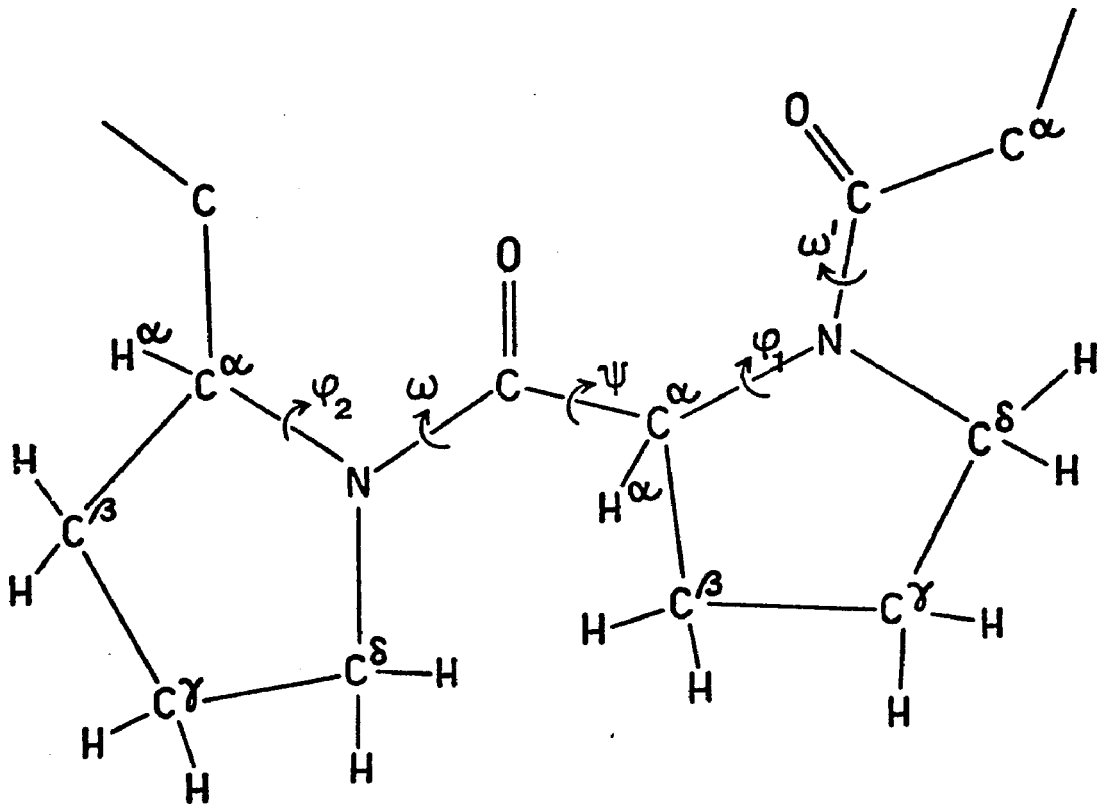


Figure 3-1. The Pro-Pro internal dipeptide.

3.2 Conformational Energy Maps for Cis and Trans Dimers

If the bond angle $\tau[C^{\alpha}C\ N]$ is held constant, the conformational energy of the Pro-Pro dimer is a function of five internal parameters:

$$E_{dimer} = E(\omega', \psi, \omega, \Gamma, \theta) \qquad (1)$$

In the dimer calculations we have allowed these parameters to vary as follows:

$\{\omega,\omega'$: 155°, 205°, $5^{\circ}\}$ for the trans dimer

$\{\omega,\omega'$: -25°, $+25^{\circ}$, $5^{\circ}\}$ for the cis dimer

$\{\theta$: -30°, 30°, $5^{\circ}\}$

$\{\psi$: -180°, $+180^{\circ}$, $10^{\circ}\}$

For each value of θ, six different Γ's were selected, corresponding to α-, β-, δ- and N-puckered ring geometries as well as the M1 and M2 minimum energy geometries. Both $\tau[C^{\alpha}C\ N]$ and $\tau[C^{\alpha}C\ O]$ were set at 119° (28) and $\ell(C'=O) = 1.24$Å (10).

The total conformational energy of the dimer was then calculated for both the cis and trans dimers by summing the non-bonded, electrostatic and torsional potentials along with the self-energy of the two prolyl rings. The non-bonded potential was of the Lennard-Jones type with interaction constants taken from Scott and Scheraga (see Appendix I, Table I-1). The partial charges used in the electrostatic term are from Holzwarth and Chandrasekharan (Appendix I, Table I-2). Included in the torsional potentials is that due to ω-variation (Appendix I, equation 7) for which V_o was taken to be 20 kcal/mole.

Contour plots of the energy distribution in the (θ,ψ) plane for the cis and trans proline dimers appear in Figures 3-2 and 3-3. For each (θ,ψ) the best ring geometry and the values of ω and ω' that give the lowest dimer energy have been plotted. Since τ^{α} and $\tau^{N'}$ differ for the various geometries, each (θ,ψ) contains the best values of τ^{α} and $\tau^{N'}$, as dictated by the prolyl ring, which yields the best energies in the proline dimer. The contours are at 1 kcal/mole intervals. The location of the two local minima of Figure 3-2 and the minimum of Figure 3-3 as well as the energies at these minima are contained in Table 3-1. The ring geometry which gave the lowest energy for each (θ,ψ) point is indicated symbolically. Dots represent the C^{γ} position corresponding to M1 geometries, while M2 geometries are indicated by crosses.* If the ring conformation happens to be single-atom puckered, it is marked so along with the sign of the puckering.

The energy map for the trans dimer, Figure 3-2, shows a broad flexibility, both in θ and ψ. The range of values of ψ available to the dimer is between 100° and 190°. This is in good agreement with results of earlier calculations by Schimmel and Flory (36), Krimm and Venkatachalam (35) and Venkatachalam and Krimm (4). Within 2 kcal/mole of the energy minimum, a range of θ from -25° to $+20^{\circ}$ is possible.

*Referring to Figure 2-5, it is seen that for negative θ, dots represent Type B geometries while crosses mark Type A geometries. The converse is true for positive θ.

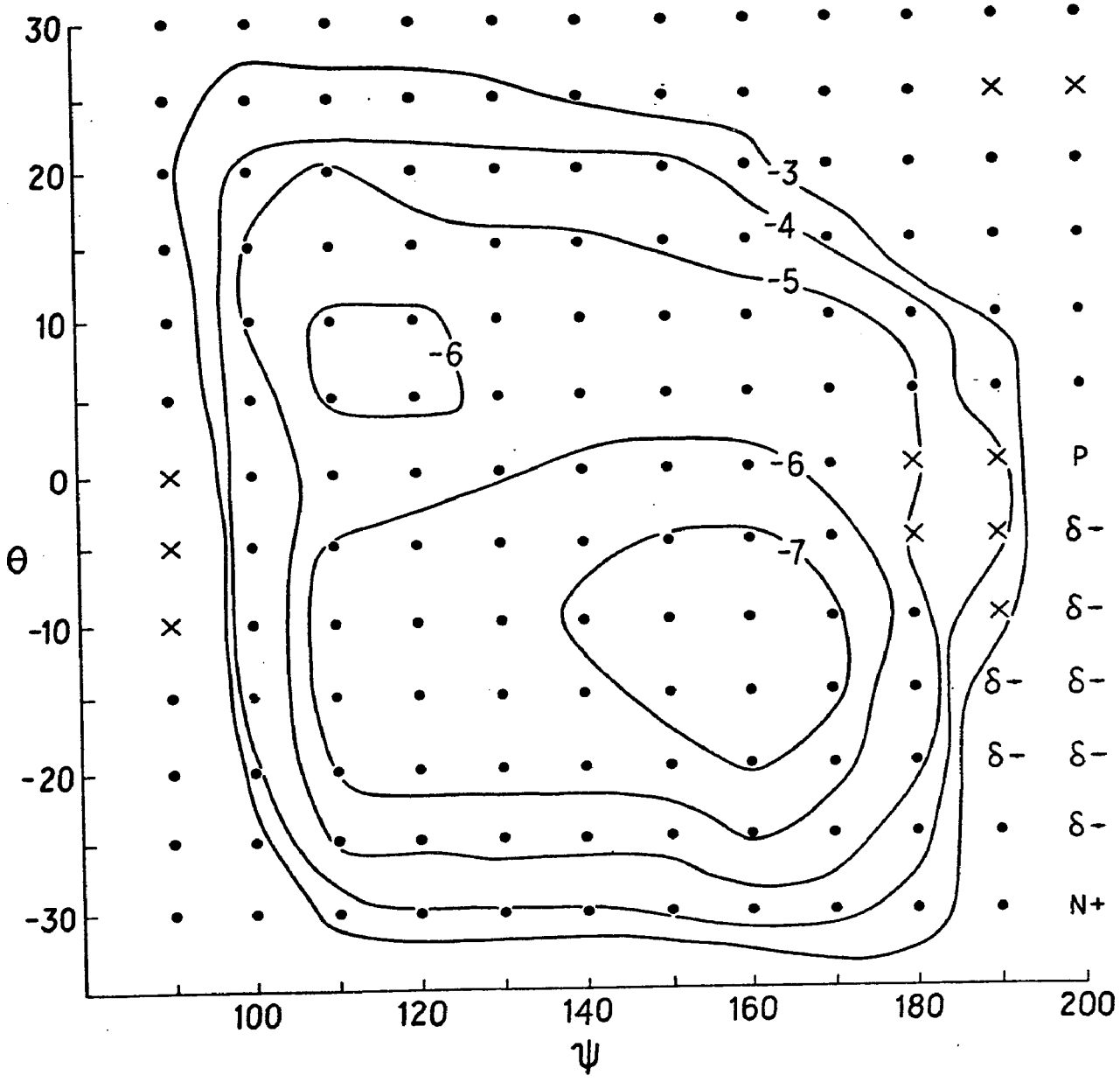


Figure 3-2. θ-ψ Conformational energies for the trans proline dimer. The best ring geometry for each (θ,ψ) is symbolically represented (see text).

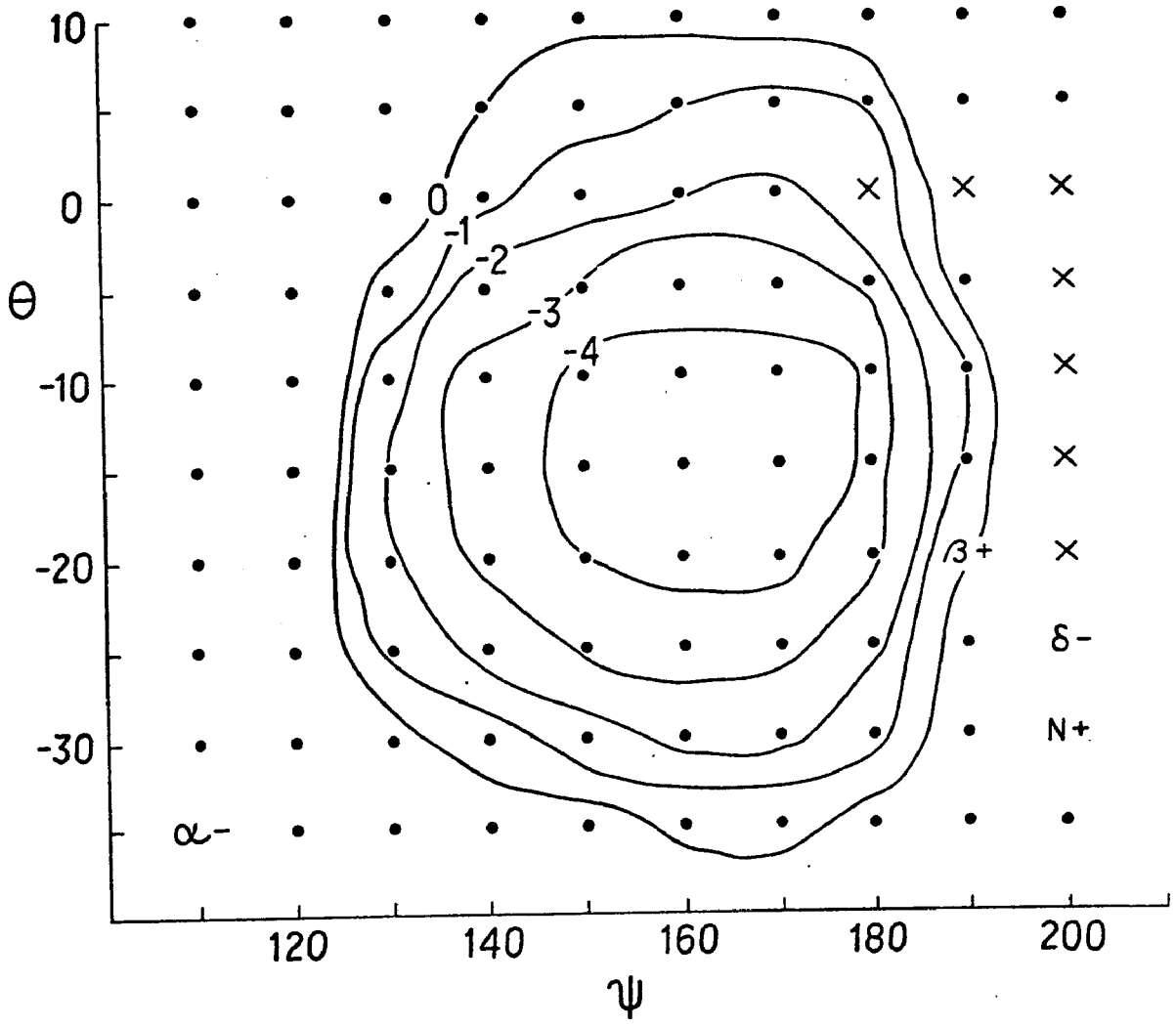


Figure 3-3. θ-ψ Conformational energies for the cis proline dimer. The best ring geometry for each (θ,ψ) is symbolically represented (see text).

Table 3-1. Location of Minima and Energies at these Minima for the cis and trans Proline Dimers

Dimer type	θ	ψ	ω	ω'	Γ	Energy (kcal/mole)
trans[a]	-10°	160°	180°	180°	-32°	-7.5
trans[b]	$+10^{\circ}$	110°	180°	180°	$+32^{\circ}$	-6.1
cis	-15°	170°	-5°	0°	-30°	-4.9

[a] Global minimum

[b] Secondary minimum

In terms of φ, this range is -40° to -85°. This result itself is noteworthy because most earlier proline dimer energy calculations have employed either completely planar rings or have had θ fixed at one or two other non-planar values near $\theta=0^{\circ}$. In the regions of interest the dimer seems to prefer the same Γ value as the isolated ring. In other higher energy regions the ring can become single-atom puckered. Considering the region around $\theta\sim-20^{\circ}$ in the Γ-θ plot of Figure 2-10, the minimum energy Γ-arced geometries (Type B) are closest to β^{+} puckered. For $-15^{\circ}< \theta < +15^{\circ}$, the range in θ where most of the ring geometries lie, the ring will be either γ^{+} or γ^{-} puckered. Therefore, the internal dimers in a trans polyproline are likely to contain primarily γ^{+} and γ^{-} puckered rings and to a smaller extent, β^{+} puckered geometries.

The flexibility of the cis dimer is considerably less, as Figure 3-3 shows. The ranges of θ and ψ have both been shortened. In particular, θ is now restricted to values less than 0°. Again, as in the case of the trans dimer, the rings which give the lowest dimer energy have the same Γ as the isolated ring. Since $\theta<0^{\circ}$, the rings will be predominantly γ^{-} or β^{+} puckered.

Table 3-2 lists the best values of the torsion angles ω and ω' for each (θ,ψ) of the trans Pro-Pro dimer. For almost every (θ,ψ) within the -3 kcal/mole contour, ω' is found to be almost exclusively 180°, even though in the calculation it was allowed a flexibility of $\pm 25^{\circ}$. This is

Table 3-2. Values of (ω,ω') for each (θ,Ψ) Corresponding to the Lowest Energy Ring for the trans Pro-Pro Dimer

θ \ Ψ	90	100	110	120	130	140	150	160	170	180	190	200
+30	(190,185)	(185,180)	(185,175)	(185,175)	(185,175)	(185,175)	(180,170)	(180,165)	(175,160)	(165,160)	(200,160)	(200,160)
+25	(195,180)	(185,180)	(185,175)	(185,175)	(180,175)	(180,175)	(180,170)	(180,170)	(175,165)	(165,160)	(200,165)	(160,160)
+20	(195,185)	(185,180)	(180,180)	(180,180)	(180,175)	(180,175)	(180,175)	(180,170)	(180,170)	(190,165)	(195,160)	(190,160)
+15	(200,185)	(190,180)	(180,180)	(180,180)	(175,180)	(175,175)	(180,175)	(180,175)	(180,170)	(185,165)	(185,160)	(185,160)
+10	(200,185)	(190,185)	(180,180)	(175,180)	(175,180)	(175,180)	(175,175)	(180,175)	(180,175)	(180,170)	(180,170)	(180,165)
+5	(200,180)	(190,180)	(180,180)	(175,180)	(175,180)	(175,180)	(175,180)	(180,180)	(180,175)	(180,175)	(180,170)	(175,170)
0	(200,190)	(195,185)	(190,180)	(190,180)	(185,180)	(185,180)	(185,180)	(180,180)	(180,180)	(180,175)	(180,175)	(160,175)
-5	(200,190)	(200,185)	(190,180)	(185,180)	(185,180)	(185,180)	(180,180)	(180,180)	(175,180)	(180,180)	(180,175)	(160,175)
-10	(200,190)	(200,185)	(190,180)	(185,180)	(185,180)	(185,180)	(180,180)	(180,180)	(175,180)	(165,180)	(175,180)	(160,175)
-15	(200,160)	(200,185)	(190,180)	(185,180)	(180,180)	(180,180)	(180,180)	(180,180)	(175,180)	(165,180)	(175,180)	(160,180)
-20	(200,160)	(200,180)	(185,180)	(180,180)	(180,180)	(180,180)	(180,180)	(180,180)	(175,180)	(165,180)	(175,180)	(165,180)
-25	(200,160)	(200,180)	(185,180)	(180,180)	(180,180)	(180,180)	(180,180)	(180,180)	(180,180)	(170,180)	(160,180)	(170,180)
-30	(200,165)	(195,175)	(185,180)	(180,180)	(175,180)	(175,180)	(180,180)	(180,180)	(180,180)	(170,180)	(170,180)	(160,180)

Table 3-2. Values of (ω,ω') for each (θ,ψ) Corresponding to the Lowest Energy Ring for the trans Pro-Pro Dimer

θ \ Ψ	90	100	110	120	130	140	150	160	170	180	190	200
+30	(190,185)	(185,180)	(185,175)	(185,175)	(185,175)	(185,175)	(180,170)	(180,165)	(175,160)	(165,160)	(200,160)	(200,160)
+25	(195,180)	(185,180)	(185,175)	(185,175)	(180,175)	(180,175)	(180,170)	(180,170)	(175,165)	(165,160)	(200,165)	(160,160)
+20	(195,185)	(185,180)	(180,180)	(180,180)	(180,175)	(180,175)	(180,175)	(180,170)	(180,170)	(190,165)	(195,160)	(190,160)
+15	(200,185)	(190,180)	(180,180)	(180,180)	(175,180)	(175,175)	(180,175)	(180,175)	(180,170)	(185,165)	(185,160)	(185,160)
+10	(200,185)	(190,185)	(180,180)	(175,180)	(175,180)	(175,180)	(175,175)	(180,175)	(180,175)	(180,170)	(180,170)	(180,165)
+5	(200,180)	(190,180)	(180,180)	(175,180)	(175,180)	(175,180)	(175,180)	(180,180)	(180,175)	(180,175)	(180,170)	(175,170)
0	(200,190)	(195,185)	(190,180)	(190,180)	(185,180)	(185,180)	(185,180)	(180,180)	(180,180)	(180,175)	(180,175)	(160,175)
-5	(200,190)	(200,185)	(190,180)	(185,180)	(185,180)	(185,180)	(180,180)	(180,180)	(175,180)	(180,180)	(180,175)	(160,175)
-10	(200,190)	(200,185)	(190,180)	(185,180)	(185,180)	(185,180)	(180,180)	(180,180)	(175,180)	(165,180)	(175,180)	(160,175)
-15	(200,160)	(200,185)	(190,180)	(185,180)	(180,180)	(180,180)	(180,180)	(180,180)	(175,180)	(165,180)	(175,180)	(160,180)
-20	(200,160)	(200,180)	(185,180)	(180,180)	(180,180)	(180,180)	(180,180)	(180,180)	(175,180)	(165,180)	(175,180)	(165,180)
-25	(200,160)	(200,180)	(185,180)	(180,180)	(180,180)	(180,180)	(180,180)	(180,180)	(180,180)	(170,180)	(160,180)	(170,180)
-30	(200,165)	(195,175)	(185,180)	(180,180)	(175,180)	(175,180)	(180,180)	(180,180)	(180,180)	(170,180)	(170,180)	(160,180)

Table 3-3. Values of (ω,ω') for each (θ,ψ) Corresponding to the Lowest Energy Ring for the cis Pro-Pro Dimer

θ \ Ψ	110	120	130	140	150	160	170	180	190	200
+10	(5, 20)	(-5, 20)	(-10, 10)	(-15, 0)	(-20, 0)	(-20, -5)	(-20, -5)	(-15,-10)	(-5,-15)	(-10,-20)
+5	(10, 20)	(0, 20)	(-5, 10)	(-10, 0)	(-15, 0)	(-15, 0)	(-15, -5)	(-10, -5)	(-5,-10)	(-10,-15)
0	(15, 20)	(5, 20)	(0, 10)	(-10, 5)	(-10, 0)	(-10, 0)	(-10, 0)	(-5, -5)	(0, -5)	(-20,-10)
-5	(20, 20)	(10, 20)	(5, 5)	(-5, 0)	(-5, 0)	(-5, 0)	(-5, 0)	(-15, 0)	(-15,-15)	(-20, -5)
-10	(20, 20)	(15, 15)	(5, 5)	(0, 0)	(-5, 0)	(-5, 0)	(-5, 0)	(-15, 0)	(-20, 0)	(-20, 0)
-15	(20, 20)	(20, 5)	(10, 0)	(0, 0)	(0, 0)	(0, 0)	(-5, 0)	(-15, 0)	(-20, 0)	(-20, 0)
-20	(20,-20)	(20, 5)	(10, 0)	(5, 0)	(0, 0)	(0, 0)	(-5, 0)	(-15, 0)	(-20, 0)	(-20, 0)
-25	(20,-20)	(20, 0)	(10, 0)	(5, 0)	(5, 0)	(0, 0)	(-5, 0)	(-15, 0)	(-20, 0)	(-20, 0)
-30	(20,-15)	(20, 0)	(10, 0)	(5, 0)	(5, 0)	(0, 0)	(-5, 0)	(-15, 0)	(-20, 0)	(-20, 0)
-35	(20,-10)	(20, 0)	(15, 0)	(10, 0)	(5, 0)	(5, 0)	(0, 0)	(-10, 0)	(-20, 0)	(-20, 0)

a convenient result as it allows us to neglect ω' rotation in other calculations involving the trans Pro-Pro internal dipeptide. The same conclusion can be made with respect to the cis Pro-Pro dimer also. As Table 3-3 shows, within 4 kcal/mole of the minimum, conformations with $\omega'=0^{o}$ are in the great majority. For both the cis and trans dimers, there is a considerable variation for ω' only in regions where there is already a high dimer energy. This variation is more pronounced in the cis dimer, however, because the molecule itself is relatively constrained.

In comparing the ring puckering of the cis and trans dimers, it is noted that the cis dimer will have more β^{+} puckered and considerably less γ^{+} puckered ring geometries than in the trans dimer. This result is compatable with those of Torchia (44) who has estimated the torsion angles in the prolyl ring from coupling constants obtained from NMR spectra of polyproline. In light of our observations as to the kinds of puckerings associated with the ring in both the cis and trans dimers, it is interesting to note two C^{13} NMR studies (49,50) on molecules containing proline. According to Dorman et. al. (50), the chemical shifts for the C^{α} and C^{δ} atoms in trans polyproline are almost identical to those in cis polyproline. However, the C^{β} and C^{γ} chemical shifts are altered in going from cis to trans polyproline. In NMR, the magnetic environment of each proton is important in determining its chemical shift. The environment is probably affected by ring puckering.

The differences in ring puckering gleaned from Figures 3-2 and 3-3 may provide a clue as to why the chemical shifts of only C^{β} and C^{γ} are altered.

3.3 Stable Conformations of the Proline Dimer

It is useful to consider the possibility of three types of conformations in proline-like molecules. They are the trans'· cis conformation ($\psi \sim 160^{\circ}$, $\omega \sim 0^{\circ}$), the trans'· trans conformation ($\psi \sim 160^{\circ}$, $\omega \sim 180^{\circ}$) and the cis'· trans conformation ($\psi \sim -45^{\circ}$, $\omega \sim 180^{\circ}$). In this notation, trans' and cis' refer to the rotational states about the C^{α}-C' bond while the unprimed terms belong to the C'-N bond rotational states. There is no cis'· cis conformation, so the notations can be shortened from trans'· cis, cis'· trans and trans'· trans to cis, cis' and trans (or trans') respectively. In the next section, the cis' region for the Pro-Pro dimer is discussed.

3.4 The Cis' Region in the Proline Internal Dipeptide

There is a compact conformation, corresponding to $\psi \sim -45^{\circ}$, $\omega \sim 180^{\circ}$ which is possible for the isolated proline residue. As early as 1960, the existence of an occasional cis' conformation in a solution of polyproline was suggested by Steinberg et. al. (51). In subsequent conformational energy calculations (36,52,6) where fixed ring geometries were employed, the cis' structure was completely disallowed due to severe steric interactions between the two rings. Gō and Scheraga (39) using two γ-puckered rings, found a minimum in the cis' region, but the energy at this minimum

was about 8 kcal/mole higher than the global energy at the trans' minimum. They recognized that flexibility of bond lengths and bond angles might reduce the energy difference between these two regions. More recent calculations (40,43) have suggested that this difference is less than 2 kcal/mole.

We have reported (48) that the energy of this compact conformation is remarkably sensitive to the backbone bond angles as well as the ring puckering and partial charges used in calculating the electrostatic energy. We wish to investigate this extreme sensitivity. We have therefore computed the energy differences between the trans' and cis' regions using three sets of bond angles for $\tau^{N'}$ and $\tau^{N''}$ (while holding τ^{N} and τ^{α} constant at 113° and 103° respectively) and using three different sets of partial charges. The three sets of bond angles for $(\tau^{N'},\tau^{N''})$ were $(121^{\circ},126^{\circ})$ as used by Mattice et. al. (43), $(123^{\circ},124^{\circ})$ as used in our previous calculations and $(126^{\circ},121^{\circ})$. The charges were taken from three sources, Holzwarth and Chandrasekharan (33), Mattice et. al. (43) and Poland and Scheraga (47). The ring self-energies were first determined by Γ-arcing in the prescribed manner for each set of bond angles. The torsion angles were varied:

$\{\psi: -70^{\circ}, -30^{\circ}, 5^{\circ}\}$ for the cis' region

$\{\psi: 110^{\circ}, 180^{\circ}, 5^{\circ}\}$ for the trans' region

$\{\theta: -30^{\circ}, 30^{\circ}, 5^{\circ}\}$

$\{\omega: 160^{\circ}, 200^{\circ}, 5^{\circ}\}$

ω' was held constant at 180°. For each ψ, both M1 and M2 minimum energy ring geometries for all θ were employed in the dimer. The minimum energy and the relevant parameters associated with that minimum are shown in Table 3-4. The effect of the bond angles and partial charges used in this study is significant, not so much in the energies of the trans' region, but quite a bit in the cis' region. The energy difference between these regions ranges from 1.7 kcal/mole to 7.2 kcal/mole. It is obvious that, even when the rings in the dimer are allowed freedom to pucker, the energy of the cis' region is quite sensitively dependent upon charge and internal parameters.

As it turns out, the actual energy difference between these two conformations is crucial in calculating the CR of polyproline. While solution measurements of Mattice and Mandelkern (42) show that the CR of poly(L-proline) II is between 14 and 20, calculated values (36,52,4,42,48) have ranged from 45 to 110. A low observed CR suggests that the polymer is in a collapsed state in solution. It also supports the possibility of an occasional cis' or cis conformation in the polymer. A calculation (43) which took into account a 6% population of the cis' conformation successfully predicted both the CR and its temperature dependence. There is nothing to suggest, however, that a small percentage of cis residues in the poly(L-proline) II chain could not give the same results. In fact, recent work by Torchia and Bovey (53) and Wu et. al. (54) has indicated that about 2-3% of

Table 3-4. Torsion Angles (in degrees) and Minimum Energies (in kcal/mole) for the Pro-Pro Dimer in Regions II and III Including Variation in $\tau^{N'}$, $\tau^{N''}$ and Partial Charges

$\tau^{N'}$	$\tau^{N''}$	Charges	θ	ω	ψ	E_1	θ	ω	ψ	E_2	E_2-E_1
121	126	H-C[1]	-10	180	150	-7.5	10	185	-45	-4.7	2.8
123	124	H-C	-10	180	155	-7.5	10	185	-45	-3.4	4.1
126	121	H-C	-10	180	160	-7.0	10	185	-45	+0.2	7.2
121	126	M[2]	-10	180	160	-5.3	10	185	-45	-3.5	1.8
123	124	M	-10	180	160	-5.2	10	185	-45	-2.3	2.9
126	121	M	-10	180	160	-4.9	10	195	-50	+1.5	6.4
121	126	P-S[3]	-10	180	120	-5.7	10	185	-45	-4.0	1.7
123	124	P-S	-10	180	160	-5.6	10	185	-45	-2.8	2.8
126	121	P-S	-10	180	160	-5.2	10	185	-50	+1.0	6.2
				REGION II				REGION III			

[1]Holzwarth-Chandrasekharan (33) [2]Mattice et. al. (43) [3]Poland-Scheraga (47)

the peptide bonds in poly(L-proline) II do exist in the cis conformation. A resolution of this problem may have to await more sensitive experimental evidence.

3.5 Energy Barriers Between Stable Conformations for Ring Geometries with Constant θ

Examination of molecular models with planar imide bonds shows that there are severe steric conflicts in transitions between the trans' and cis' regions by ψ rotation. Conformational energy calculations (39,42,34) and molecular orbital calculations (37) indicate that this barrier is extremely high. Our calculations (48) show that it is at least 100 kcal/mole. However, the cis → trans transition barrier is only about 20 kcal/mole (55). Both cis and cis' residues have similar effects on the conformation of a polypeptide chain; that is, they tend to collapse the molecule. In view of this, it would seem more plausible that the low CR exhibited by poly(L-proline) II in solution might be attributed to a few percent of residues in the cis conformation rather than the cis' conformation.

It seems quite possible that ring puckering could significantly affect the conformational energies in the regions between the cis' and trans' minima. We have therefore computed dimer conformational energies over the entire (ω,ψ) range. Our attempt is to study the influence of prolyl ring flexibility on the transition between the cis' and trans' regions. For these calculations, we chose a suitable set of bond angles that would be conducive to a low energy

cis' region. We have already noted that the dimer energies in the trans' region are considerably less sensitive to small changes in backbone bond angles than energies in the cis' region. τ^{α} was set at 103°, τ^{N}=113° and $\tau^{N'}$=121°. Ring self-energies were once again obtained for this set of bond angles. With these parameters, $\varphi = -61.3^{\circ}+\theta$. The Γ-θ self-energy map so obtained (Figure 3-4) is, of course, little different from Figure 2-10. The partial charges used were from Poland and Scheraga (47). The ranges of parameters were:

$$\{\omega: -180^{\circ}, 180^{\circ}, 20^{\circ}\}$$
$$\{\psi: -180^{\circ}, 180^{\circ}, 20^{\circ}\}$$

For each θ, all four single-atom puckered structures as well as M1 and M2 minimum energy ring structures were considered for the prolyl ring.

We consider first the Pro-Pro internal dipeptide in which both rings have identical puckering with θ fixed at -10°. This is the value of θ for which the ring energy of the isolated ring is a minimum and is like most earlier calculations which held the ring in one rigid conformation. Figure 3-5 is the resulting (ω,ψ) energy map. The contours are at 5 kcal/mole intervals. There are three regions where a minimum is found. These local minima are marked by crosses and are indicated by the numerals I, II and III, corresponding to the cis, trans and cis' conformations respectively. The hatched areas show where the dimer energy is greater

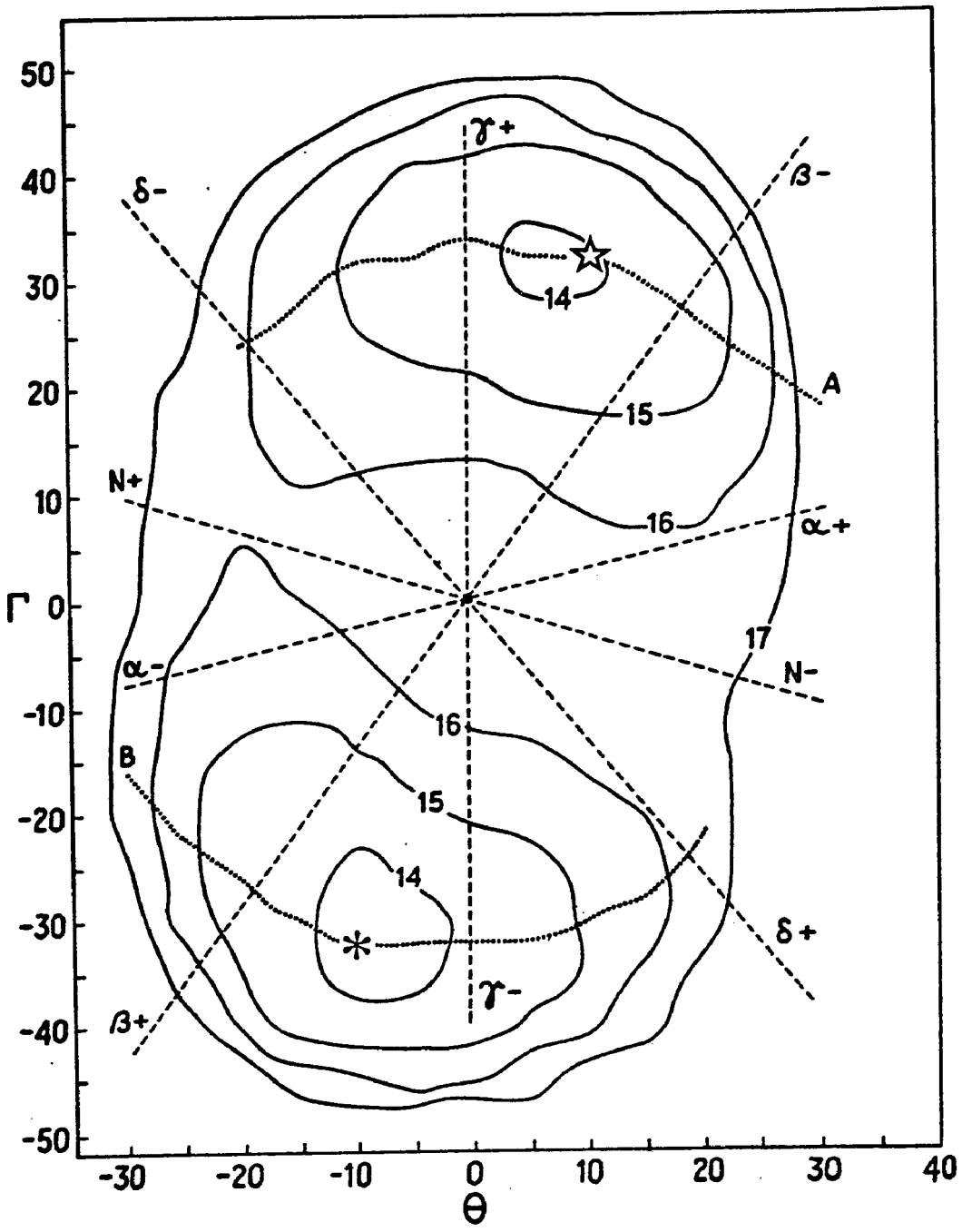


Figure 3-4. Γ-θ energy plot for the proline isolated ring. $\tau^{\alpha}=103^{o}$, $\tau^{N}=113^{o}$ and $\tau^{N'}=121^{o}$.

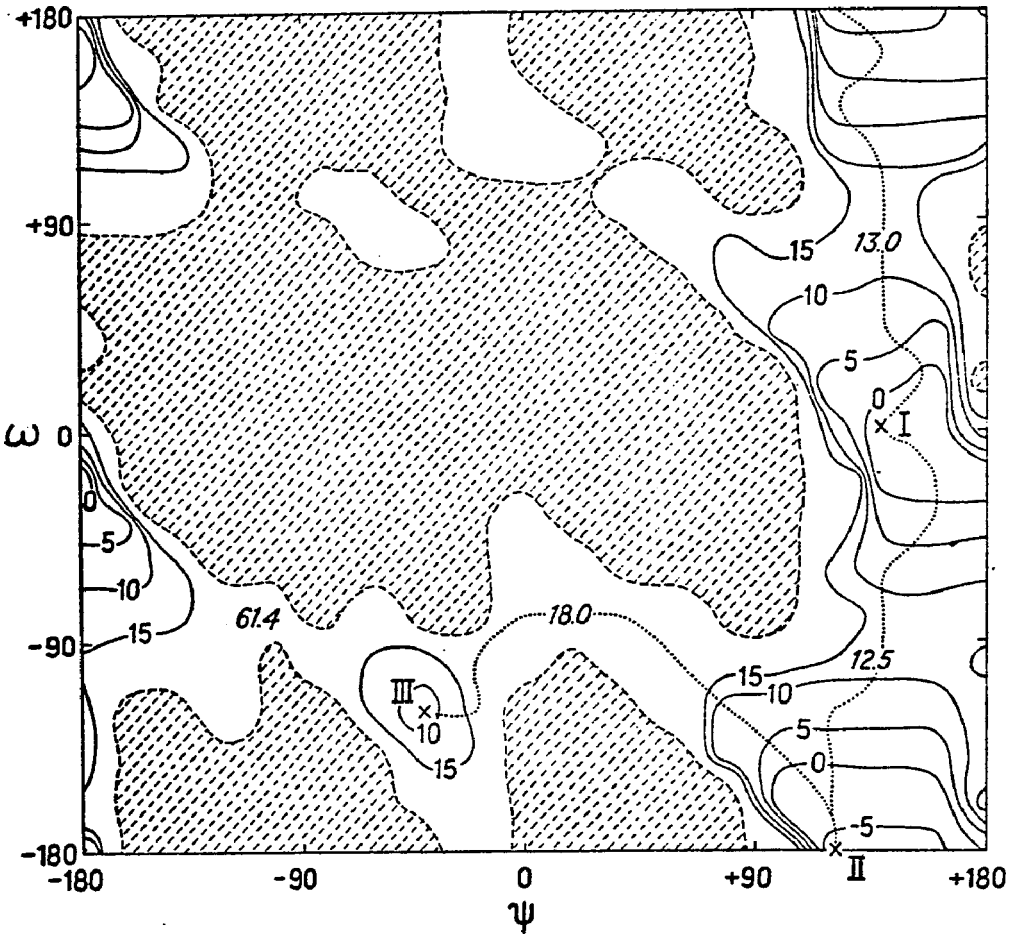


Figure 3-5. (ω,ψ) energy contour map for the Pro-Pro internal dipeptide in which θ is held constant at -10° ($\varphi \approx -71^{\circ}$).

than 75 kcal/mole; i.e., for all purposes impenetrable barriers for the system under consideration. Between the local minima are dotted lines which define the best conformational pathways between the minima. These conformational pathways were determined by taking the line of least resistance between two minima. The line passes through points in the (ω,ψ) map where energies have been computed; that is, no interpolation has been made. Along each of these paths, the energy at the barrier is marked in italics at the (ω,ψ) point where the barrier is found. The values of the torsional angles and the energies at the minima are contained in Table 3-5. The same information for the barrier energies is listed in Table 3-6. By barrier energy, we mean the energy <u>at</u> the barrier rather than the <u>height</u> of the barrier which normally refers to the difference between the energy at a barrier and a minimum. Thus, for example, the <u>trans</u> to <u>cis</u> transition along the pathway $\omega=180^{o} \rightarrow \omega=0^{o}$ involves a barrier height of 18.2 kcal/mole. The barrier height in going the opposite direction along the same pathway is 16.9 kcal/mole. In each case, the barrier <u>energy</u> is only 12.5 kcal/mole. There is a pathway from Region II to Region III with a barrier energy of 18 kcal/mole. While this barrier is certainly not insurmountable, the energy at the <u>cis'</u> minimum is 8.5 kcal/mole, 14.2 kcal/mole above the <u>trans'</u> energy minimum. This structure would not exist in a population of dimers with $\theta=-10^{o}$. In addition, the transition from Region II to Region III involves an ω rotation of 100^{o}.

Table 3-5. The Torsional Angles (in degrees) and Energy Values (in kcal/mole) of Minimum Energy Conformations of the Pro-Pro Internal Dipeptide

Figure	Region I					Region II					Region III				
	ω	ψ	θ_1	θ_2	E	ω	ψ	θ_1	θ_2	E	ω	ψ	θ_1	θ_2	E
3-5	0,	140,	-10,	-10	-4.4	180,	120,	-10,	-10	-5.7	-120,	-40,	-10,	-10	8.5
3-6	0,	140,	10,	10	-0.5	180,	120,	10	10	-4.6	-180,	-40,	10,	10	-3.8
3-7+3-8	0,	160,	-15,	-15	-4.9	180,	120,	-10,	-10	-5.7	-180,	-40,	10,	10	-3.8
3-9	0,	160,	-15,	-15	-4.9	180,	120,	-10,	-15	-5.7	-180,	-40,	10,	5	-4.1

Table 3-6. The Torsional Angles (in degrees) and Energy Values (in kcal/mole) of the Conformations at the Barriers Separating the Stable Conformations of the Pro-Pro Internal Dipeptide

Figure	cis-trans					cis'-trans'					cis-cis'				
	ω	ψ	θ_1	θ_2	E	ω	ψ	θ_1	θ_2	E	ω	ψ	θ_1	θ_2	E
3-5	-100,	140,	-10,	-10	12.5	- 80,	20,	-10,	-10	18.0	-80,	-100,	-10,	-10	61.4
3-6	- 80,	140,	10,	10	13.8	- 60,	-140,	10,	10	30.1			—*		
3-7	-100,	140,	-10,	-10	12.5	-100,	0,	-25,	-25	17.5	140,	0,	-10,	-10	21.4
3-8	-100,	140,	-10,	-10	13.0	- 80,	- 20,	-10,	-10	14.7	100,	40,	-25,	-25	15.7
3-9	-100,	140,	-10,	-10	12.5	-100,	0,	-25,	0	14.5	80,	60,	-15,	10	13.5

*No pathway directly linking cis and cis' is found in this case.

Rotation of ψ along the $\omega=180^{\circ}$ line in either direction is not possible. This explains the difficulty in converting from a trans' to a cis' structure using molecular space filling models with planar imide bonds.

Figure 3-6 shows the same calculation, but the θ restricted to $+10^{\circ}$, the value of θ for which the isolated ring has a secondary energy minimum. The trans-cis barriers are only slightly affected by the change in θ, but the location and characteristics of the cis' region have been drastically altered. Whereas for $\theta=-10^{\circ}$, the imide bond was highly non-planar in the cis' region, the cis' minimum now occurs for a planar imide bond. Most notably, the energy difference between the cis' and trans' conformations is now only 0.8 kcal/mole. Considering that the cis minimum in this case is 4.1 kcal/mole higher in energy than the trans' minimum it would seem that the cis' region would be more favorable than the cis region. However, the pathway found in Figure 3-5 between cis' and trans' is now blocked. An alternate pathway has opened up, but the barrier of over 30 kcal/mole seems prohibitive. Clearly, the flexibility of the proline ring is important in estimating the barrier energies. The value of θ should, therefore, be allowed to vary.

3.6 Energy Barriers Between Stable Conformations with no Restrictions on θ

When θ is allowed to vary over the following range:

$$\{\theta: -25^{\circ}, +25^{\circ}, 5^{\circ}\}$$

and both rings have the same puckering, the energy map in

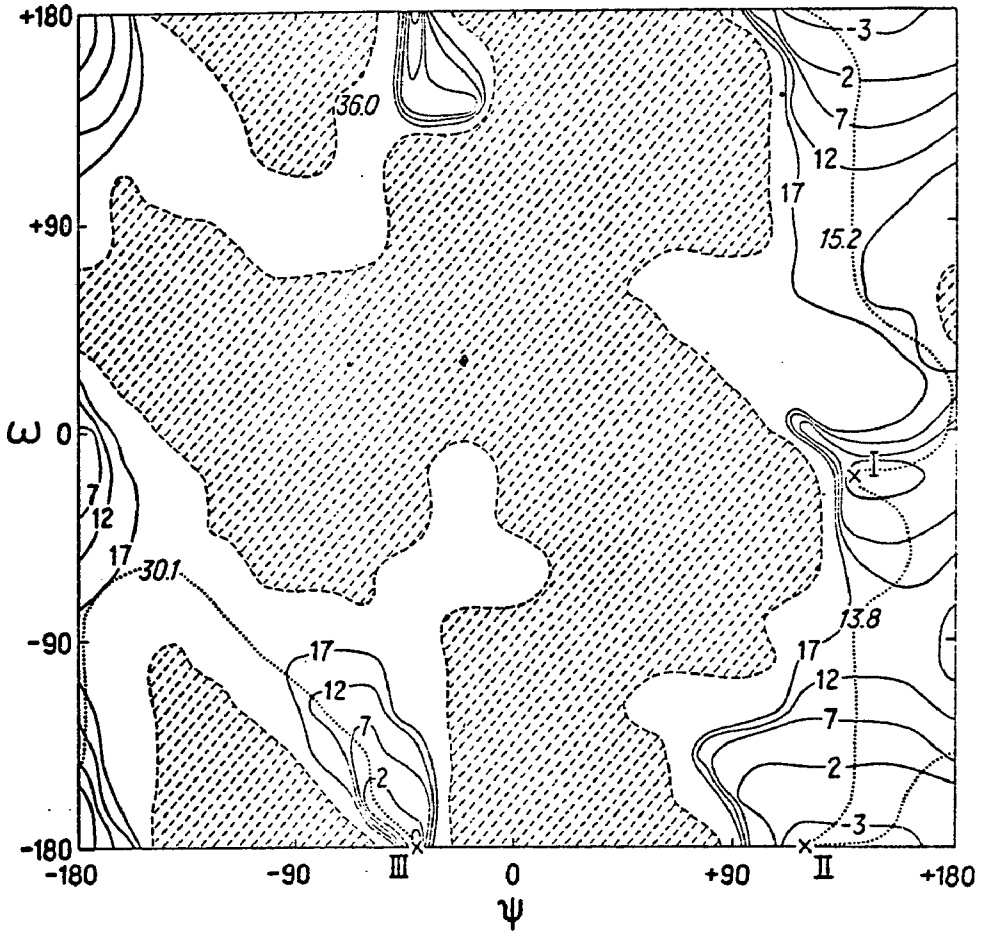


Figure 3-6. (ω,ψ) energy contour map for the Pro-Pro internal dipeptide in which θ is held constant at $+10^{\circ}$ ($\varphi \approx -51^{\circ}$).

Figure 3-7 is obtained. For each (ω,ψ) in this map, the best value of θ giving the lowest dimer energy is plotted. In comparing this map with the previous two, the most obvious feature is the great amount of steric freedom under θ relaxation. This is especially true in the cis' region. Even though the cis and trans regions are opened up somewhat, the barriers between them are affected very little. However, the cis' conformation not only has a low energy (1.9 kcal/mole higher than trans', see Table 3-5), but there are two pathways opened up between these regions (only the best pathway is indicated in the figure) each of which has a barrier energy only about 5 kcal/mole higher than that for the trans → cis conversion. Even though there is a marked increase in steric freedom, the molecule is still not free to rotate about ψ along the ω=180° line from Region II to Region III.

The best values of θ for each (ω,ψ) are contained in Table 3-7. There are θ values only in those regions which are not hatched; that is, whose energies are below 75 kcal/mole. In both Regions I and II, the θ values are almost exclusively negative. In view of Figures 3-2 and 3-3, this is expected. Using Figure 3-2 for comparison in the trans region, the larger number of preferred conformations lie in the region where θ<0° and the global energy minimum is located there. In the case of the cis region, as in Figure 3-3, the argument is even stronger as almost all of the preferred conformations occur for θ<0°. On the other

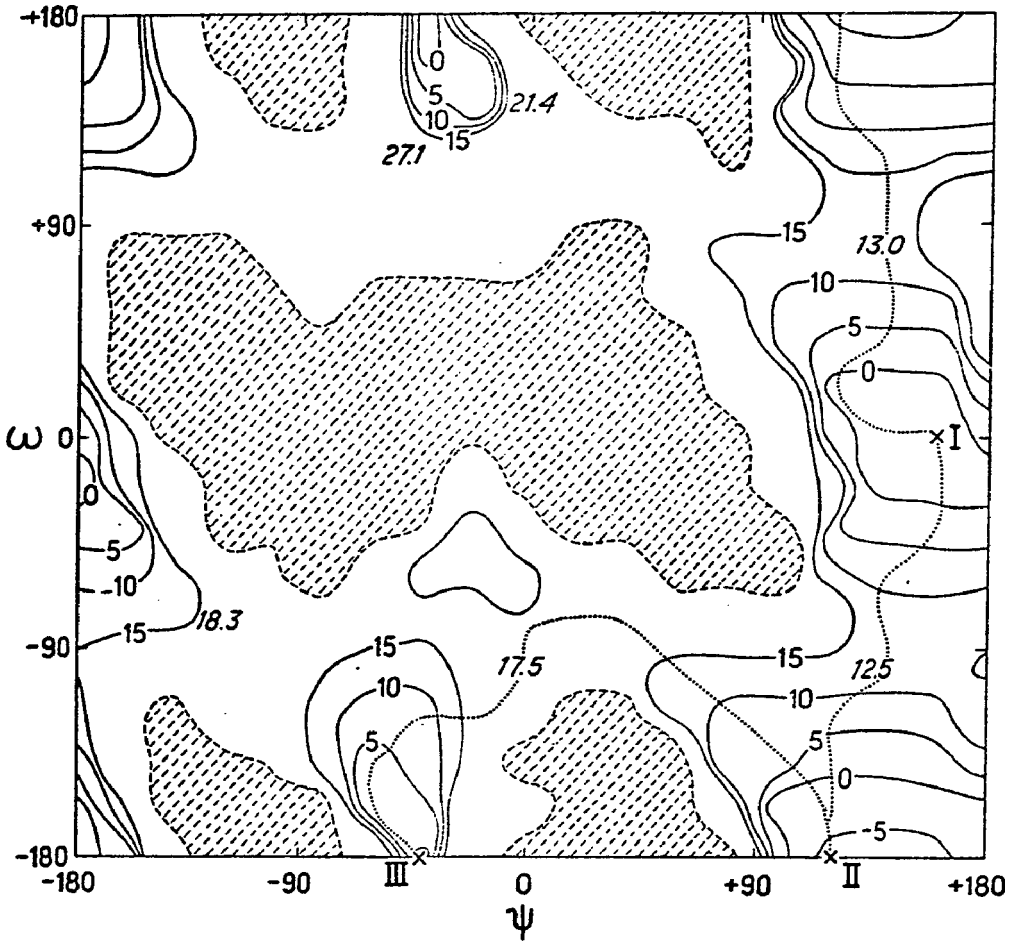


Figure 3-7. (ω,ψ) energy contour map for the Pro-Pro internal dipeptide in which θ is allowed to vary.

Table 3-7. Value of θ for Each (ω,ψ) Corresponding to the Lowest Energy Ring for the Pro-Pro Dimer With Identical Ring Puckering

ω \ ψ	-180	-160	-140	-120	-100	-80	-60	-40	-20	0	20	40	60	80	100	120	140	160	180
180	5	5	5				25	10	-15						20	-10	-10	-10	5
160	-10	5	5				25	15	5	-20					25	-15	-15	-10	-10
140	-10	-10	5	10			15	10	5	-10	-25				-25	-10	-15	-10	-10
120	-15	-15	-10	5	25	-20	5	5	-15	-10	-20	-25	-25		-25	-15	-10	-10	-15
100	5		-25	5	5	20	-5	0	-25	-25	-25	-25	-25	-25	-25	-20	-10	-15	5
80	5			10	15	0	-15	-25					-10	-15	-20	-15	-10	-20	5
60	-25				20								10	-10	-15	-15	-15	-15	-25
40	0													10	-20	-20	-20	-10	0
20	5	0												25	5	-20	-20	-15	5
0	5	5													25	-20	-15	-15	5
-20	-10	5	10					5							-25	-20	-5	-10	-10
-40	-10	-15	5				5	5	-5							-20	-10	-5	-10
-60	-10	-15	-15	5	25		25	5	5	-5	-15	-25				-10	-10	-10	-10
-80	-10	-15	-15	-10	5	25	0	5	0	-15	-15	-20	-25	-25	-25	-25	-10	-10	-10
-100	-5	5	-15	0	5	5	-10	-10	-25	-25	-25	-25	-15	-15	-20	-15	-10	-10	-5
-120	-5	10		-5	5	5	10	-5	-15	-25			-25	-10	-10	-10	-10	-10	-5
-140	5	5			0	5	15	-5	-15					5	-10	-10	-10	-10	5
-160	5	5					10	10	-10	-20				25	-10	-10	-10	-10	5
-180	5	5	5				25	10	-15						20	-10	-10	-10	5

hand, at and near the energy minimum in the cis' region, values of θ greater than 0° are favored. However, as one travels along the best cis' → trans' pathway, the preferred geometries outside of the immediate area of the cis' region become increasingly negative in θ. There are (ω,ψ) points along the pathway and on either side of the pathway where $\theta=-25^{\circ}$. For those particular points, we extended the calculations to $\theta=-30^{\circ}$ and found that the best geometry was indeed the $\theta=-25^{\circ}$ M1 geometry.

As discussed previously, the cis' conformation is an extremely compact structure and the energy is sensitively dependent upon geometrical parameters and partial charges. Because of this, it is informative to study the effect of varying any relevant parameter involved in the problem. In particular, as ω is varied from the planar values of 180° or 0° to the extreme non-planar values of $\pm 90^{\circ}$, it is reasonable to assume that the C-N bond length should be varied from the 1.32Å partial double bond length we have used in the previous calculations, to the single bond value of 1.47Å. It would be expected that this is what might be happening in the NMR studies (56-58) where the energy required to break the double bond has been measured. Therefore, we assume that this energy is reflected in the V_o of 20 kcal/mole that has been used in the ω intrinsic torsion energy. Accordingly, to complete the analogy, we have varied ℓ(C-N) linearly from 1.32Å to 1.47Å as the imide bond is taken from a planar conformation to its extreme

non-planar conformation. The resultant energy map is shown in Figure 3-8. To a certain extent, there is increased flexibility for the dimer in the cis' region. Little change is seen in the contour shapes for both the cis and trans regions. Of course, the energies at the cis and trans minima do not change, because they are located on the $\omega=180^{o}$ or $\omega=0^{o}$ line where ℓ(C-N) is not changed from Figure 3-7. However, the energies at the two barriers between the cis and trans regions are raised about 0.5 kcal/mole and one of their positions is slightly perturbed.

The most important point of this study of ℓ(C-N) bond length variation is that the cis' → trans' barrier energy along the best pathway is brought down to 14.7 kcal/mole; viz., to a value comparable to that for the cis → trans conversion. In fact the best pathway for the cis' → cis transition (not marked) has a low barrier energy of 15.7 kcal/mole. This was not seen in Figure 3-7. It should be emphasized that this map in no way completes the discussion of the effect of geometrical parameter variation on the transitions between stable conformations in the Pro-Pro dimer. Its primary purpose is to illustrate that such variation can improve the conformational pathways leading to the cis' region.

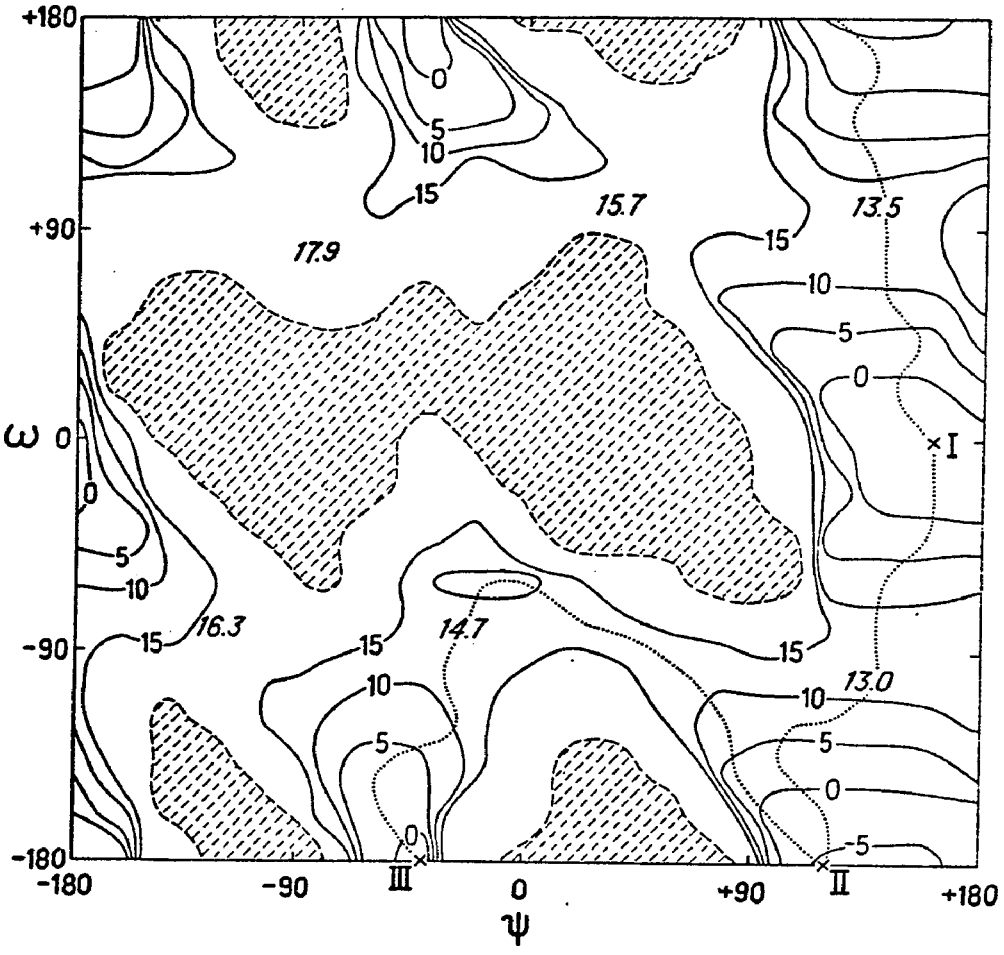


Figure 3-8. (ω,ψ) energy contour map for the Pro-Pro internal dipeptide in which θ is allowed to vary and ℓ(C-N) is a function of ω.

3.7 (ω,ψ) Map for the Pro-Pro Dimer with Independently Puckered Rings

In the previous calculations, we have puckered both the rings in an identical manner. There is nothing to suggest that such is the case in the proline dimer. It is therefore imperative to see the effect of allowing each ring in the dimer to pucker independently. Accordingly, for each (ω,ψ) this was accomplished by varying both θ's independently:

$$\{\theta_1,\theta_2: -30^o, 30^o, 5^o\}$$

Initially, the C'-N bond length was held constant at 1.32Å. In order to keep the computing time within reasonable limits, we considered only the portions of the (ω,ψ) map of Figure 3-8 that are of the most interest. Therefore, we chose Regions I, II and III, along with a corridor of (ω,ψ) points on either side of those which constitute the pathways between the three stable conformations. The energy map for these calculations is shown in Figure 3-9. The energies at the stable conformations are not significantly altered. The increased flexibility is most prominent in the reduction of the barrier energies. It is interesting to note that when the rings are puckered independently, a new pathway opens up between the trans' and cis' regions; namely the pathway which involves a rotation through ω=+80^o in Figure 3-9. We have also increased the C'-N bond length in the areas around the barrier energies but this has no significant effect on these energies.

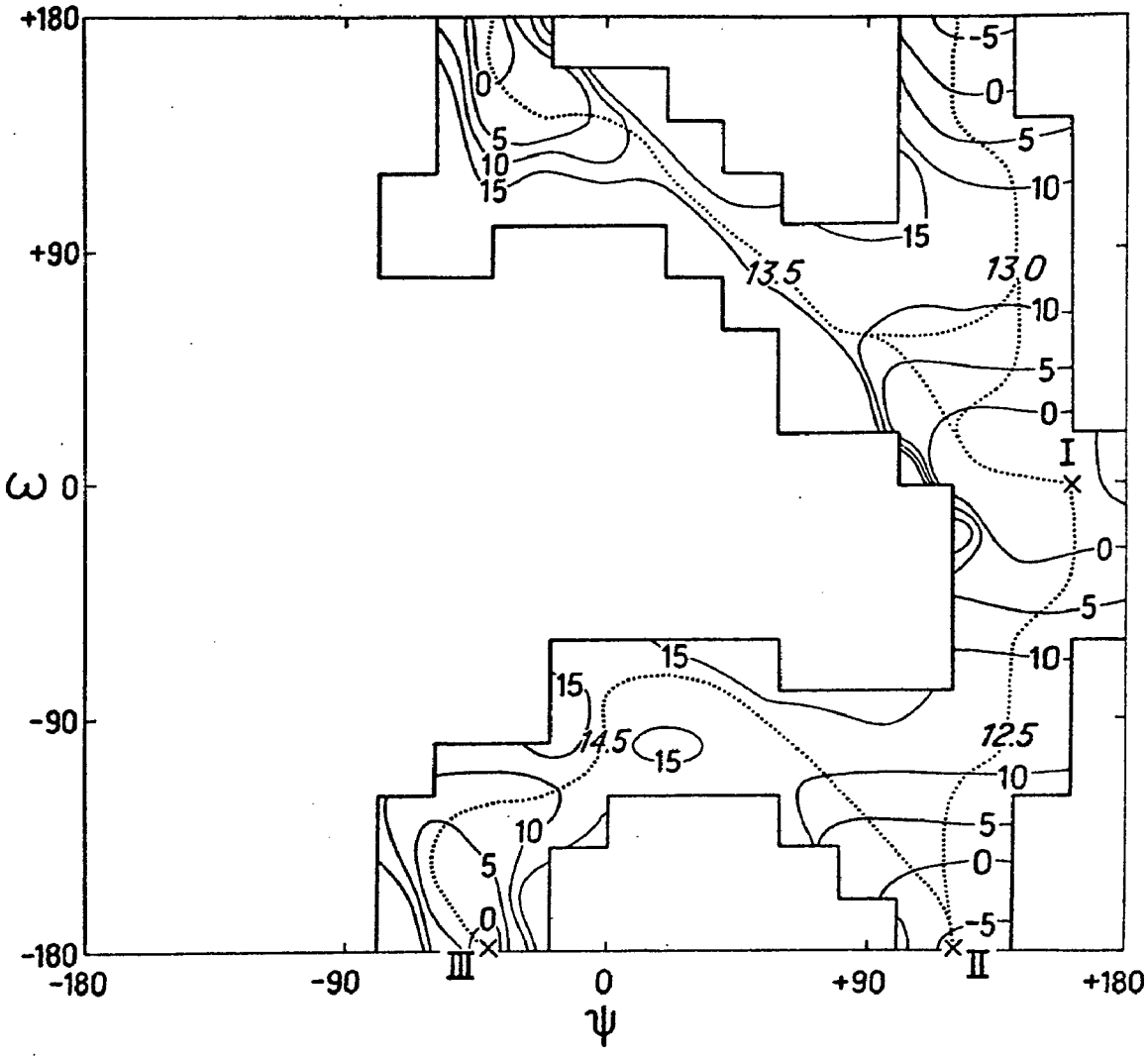


Figure 3-9. (ω,ψ) energy contour map for the Pro-Pro internal dipeptide with independent ring puckering.

Figures 3-10, 3-11 and 3-12 are stereoscopic drawings for the minimum energy cis, trans and cis' Pro-Pro internal dipeptide structures. The coordinates for these molecules were obtained from the program which calculates the dimer conformational energy with independently puckered rings. These coordinates were then utilized by the Oak Ridge Thermal Ellipsoid Plotting Program (59) which produced the stereo diagrams on the University of Michigan CalComp Plotter. The backbone atoms along the chain are lettered: CA=C^{α}, N=nitrogen and CP=C . The chain direction is from top to bottom in each diagram. The ring puckering as well as possible steric conflicts are clearly seen from these figures.

In all three diagrams, the two rings are essentially mutually perpendicular. The interesting point to note, however, is that in both the cis and trans cases, the rings point away from each other, while for the cis' structure, the rings point toward each other. Thus, the steric conflicts associated with the relatively narrow minimum in the cis' region are ring-ring conflicts. In particular, as seen from Figure 3-12, the H^{δ}'s of both rings interact, as well as C of the first ring with the H^{δ}'s of the second ring.

In Table 3-8, the values of θ_1 and θ_2 for each of the independently puckered rings are given for each (ω,ψ) point in the regions for which energy calculations have been done. In the stable regions I and II as well as the pathways connecting them, both rings definitely prefer a negative θ.

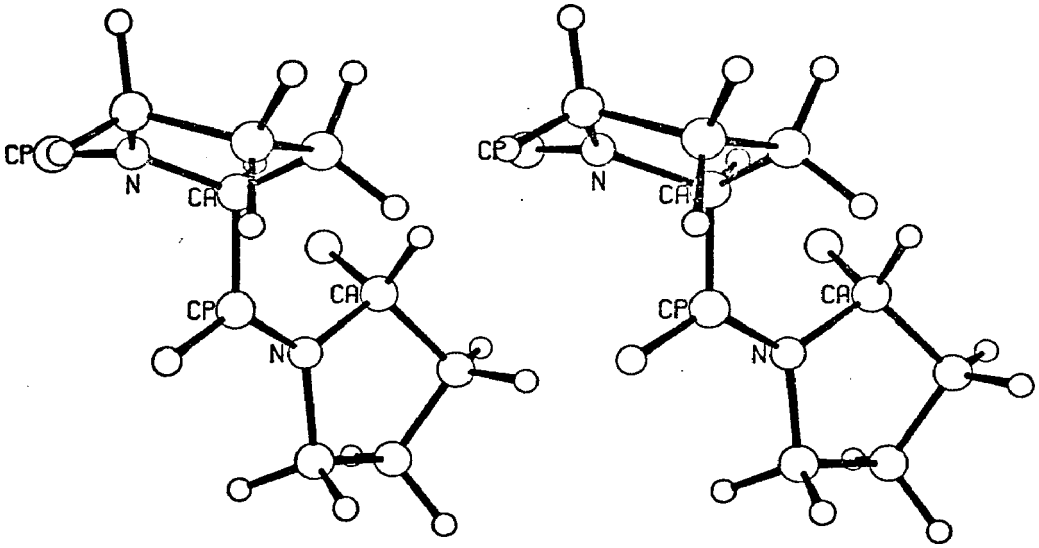


Figure 3-10. $(Pro)_2$ <u>cis</u>

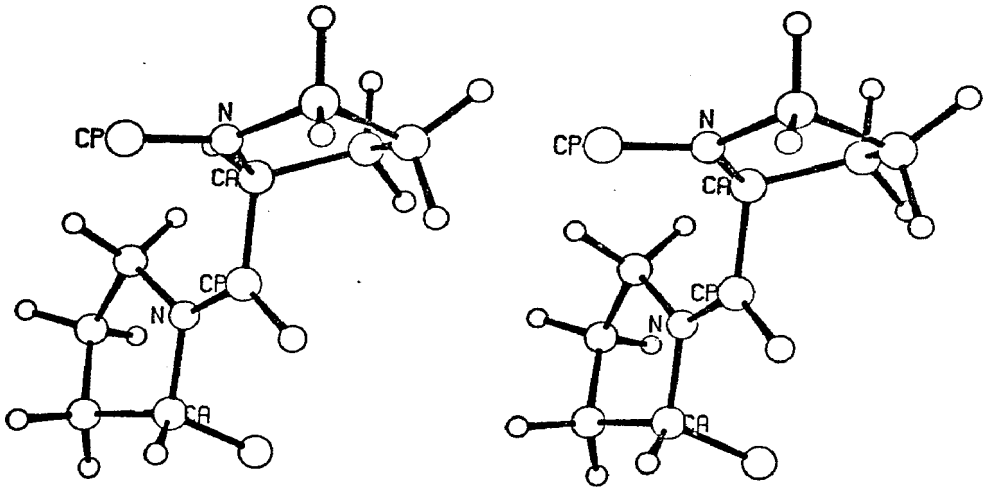


Figure 3-11. $(Pro)_2$ <u>trans</u>

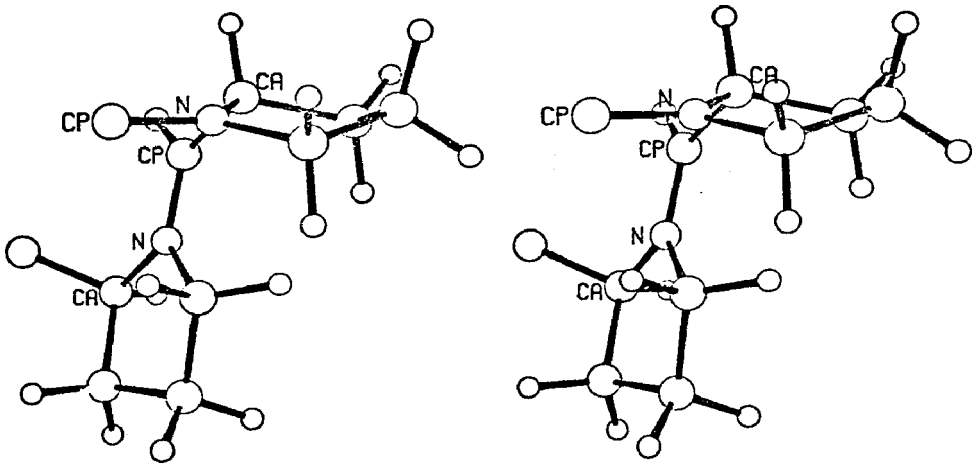


Figure 3-12. $(Pro)_2$ cis'

Table 3-8. Values of (θ_1,θ_2) for Each (ω,ψ) Corresponding to the Lowest Energy Rings for the Pro-Pro Dimer With Independently Puckered Rings

ω \ ψ	-180	-160	-140	-120	-100	-80	-60	-40	-20	0
180							(15, 0)	(10, 5)	(-5, 5)	
160							(5, 0)	(15,-15)	(5, 5)	(-15, 5)
140							(5, 15)	(10,-15)	(5, 5)	(-15, 5)
120						(0, 5)	(5, 5)	(5,-25)	(-10,-25)	(-10, 15)
100						(-10, 5)	(0,-20)	(5,-25)	(-25, 5)	(-25,-25)
80						(15,-15)	(10,-25)	(-5,-25)		
60										
40										
20	(5,-10)									
0	(5,-10)									
-20	(-15,-10)									
-40	(-15,-10)									
-60	(-15, 10)			(5,-20)	(15, 25)				(5,-15)	(0, 5)
-80			(-15,-10)	(-5,-15)	(5,-15)				(-5, 25)	(-15, 5)
-100			(-20, 0)				(-10, 5)	(-10,-15)	(-25, 25)	(-25, 0)
-120						(5,-10)	(10,-10)	(-5, -5)	(-20, 0)	(-25, 0)
-140						(5, 5)	(10,-10)	(0,-15)	(-15,-10)	(-25, 0)
-160							(5, -5)	(10, 10)	(-10,-20)	
-180							(15, 0)	(10, 5)	(-5, 5)	

Table 3-8. (continued)

				(20, 5)	(-10,-15)	(-10,-10)			180
(-25, 5)				(-20, -5)	(-15,-15)	(-15, 10)	(-10,-10)		160
(-25, 5)	(-25, 5)			(-25, 15)	(-10,-10)	(-10, 10)	(-15, 10)		140
(-20, 5)	(-25, 5)	(-25, 5)		(-25,-25)	(-15, 10)	(-10,-10)	(-10,-10)		120
(-20, 25)	(-20, 15)	(-25, -5)	(-25, 0)	(-25,-25)	(-10, 5)	(-10,-10)	(5,-10)		100
	(-25, 25)	(-15, 10)	(-20,-10)	(-20,-10)	(-15,-10)	(-10,-10)	(-10,-20)		80
		(-25, 25)	(-5,-10)	(-10,-20)	(-5,-20)	(-15,-15)	(-15,-15)		60
			(20,-10)	(-5,-20)	(-10,-25)	(-10,-20)	(-10,-10)		40
			(25, 20)	(20,-15)	(-10,-20)	(-10,-20)	(-10,-15)	(5,-10)	20 ω
				(25, 20)	(15,-10)	(-10,-15)	(-15,-15)	(5,-10)	0
					(-20,-20)	(-10, -5)	(-15,-10)	(-15,-10)	-20
					(-25,-15)	(-15; -5)	(-15, 10)	(-15,-10)	-40
(-25, 5)	(-25, 25)	(-25, 25)			(-25, 5)	(-10,-10)	(-15, 10)	(-15, 10)	-60
(-15, 5)	(-20,-15)	(-25,-20)	(-25,-25)	(-25,-25)	(-25, 10)	(-10, 10)	(-10,-10)		-80
(-30, -5)	(-25, 5)	(-15,-15)	(-15,-15)	(-20,-20)	(-15, 15)	(-10,-10)	(-10,-10)		-100
(-25, 0)	(-25, 0)	(-25, 5)	(-10,-10)	(-10,-10)	(-10,-10)	(-10,-10)	(-10,-10)		-120
			(-15, 5)	(-10,-10)	(-10,-10)	(-10,-10)			-140
			(25, 5)	(-10,-10)	(-10,-10)	(-10,-10)			-160
				(20, 5)	(-10,-15)	(-10,-10)			-180
20	40	60	80	100	120	140	160	180	

Ψ

The sign of the two θ angles (θ_1,θ_2) for each (ω,ψ) will be designated as (-,-) in these regions. Using this notation, the θ's at the *cis'* minimum are (+,+) while in the region immediately around the minimum, they are (+,-). However, along the pathways interconnecting the *cis'* region with either the *cis* or *trans* regions, the θ's prefer (-,+). Thus, if one follows along either *cis'* → *trans'* transition pathway, the values of θ change in the sequence (+,+) → (+,-) → (-,+) → (-,-).

The energy map of Figure 3-9 definitely supports the claim of Mattice *et. al.* (43) that the *cis'* conformation is a possible stable structure for the proline dimer. These calculations indicate an energy barrier of 19.2 kcal/mole for *trans'* → *cis'* conversion, in good agreement with the studies of Steinberg *et. al.* (51) who have estimated this barrier to be 20.6 kcal/mole for a dilute aqueous solution of polyproline in 12M LiBr. It should be noted that minimum energies *and* barrier energies estimated in these calculations are applicable only for the dimer depicted in Figure 3-1. Interactions between non-neighboring units in a polymer will be relevant. Just as the introduction of an occasional *cis* residue in the polymer tends to collapse the chain, this is also the case when there is a residue in the *cis'* conformation. Therefore, there will undoubtably be other factors which will influence conformational transitions of polyproline in solution.

There is some significance in relating these calculations to NMR analysis of polypeptides containing proline. The fact that there are three separate stable conformations with barrier heights greater than 15 kcal/mole between them implies that three separate nuclear magnetic resonances should be observed at room temperature. This is a different situation from, say, the case of an isolated proline where the cis' → trans' barrier height has been calculated to be less than 10 kcal/mole by Tonelli (60). Only two separate resonances, however, have been found by Dorman et. al. (50) and by Wu et. al. (54) in their NMR analysis of poly(L-proline) II. Extremely small percentages of non-trans peptide bonds or non-trans' C^{α}-C bonds are very difficult to detect by NMR. It is likely that, should either cis' or cis (or both!) structures exist in a solution of poly(L-proline) II, they will be present only in very minute quantities.

3.8 Γ-θ Energy Plots for the Pro-Pro Dimer

The conformational energy of the proline internal dipeptide may also be represented as a Γ-θ energy plot. Figure 3-13 shows such a calculation for the cis ($\omega=0^{\circ}$) proline dimer. [The backbone bond lengths and bond angles for this calculation were taken from Traub and Shmueli (30)] For each Γ and θ, the best value of ψ giving the minimum dimer energy is employed. This plot is quite different from the one in Figure 2-3 for the isolated proline. Most notably, positive values of θ are energetically unfavorable

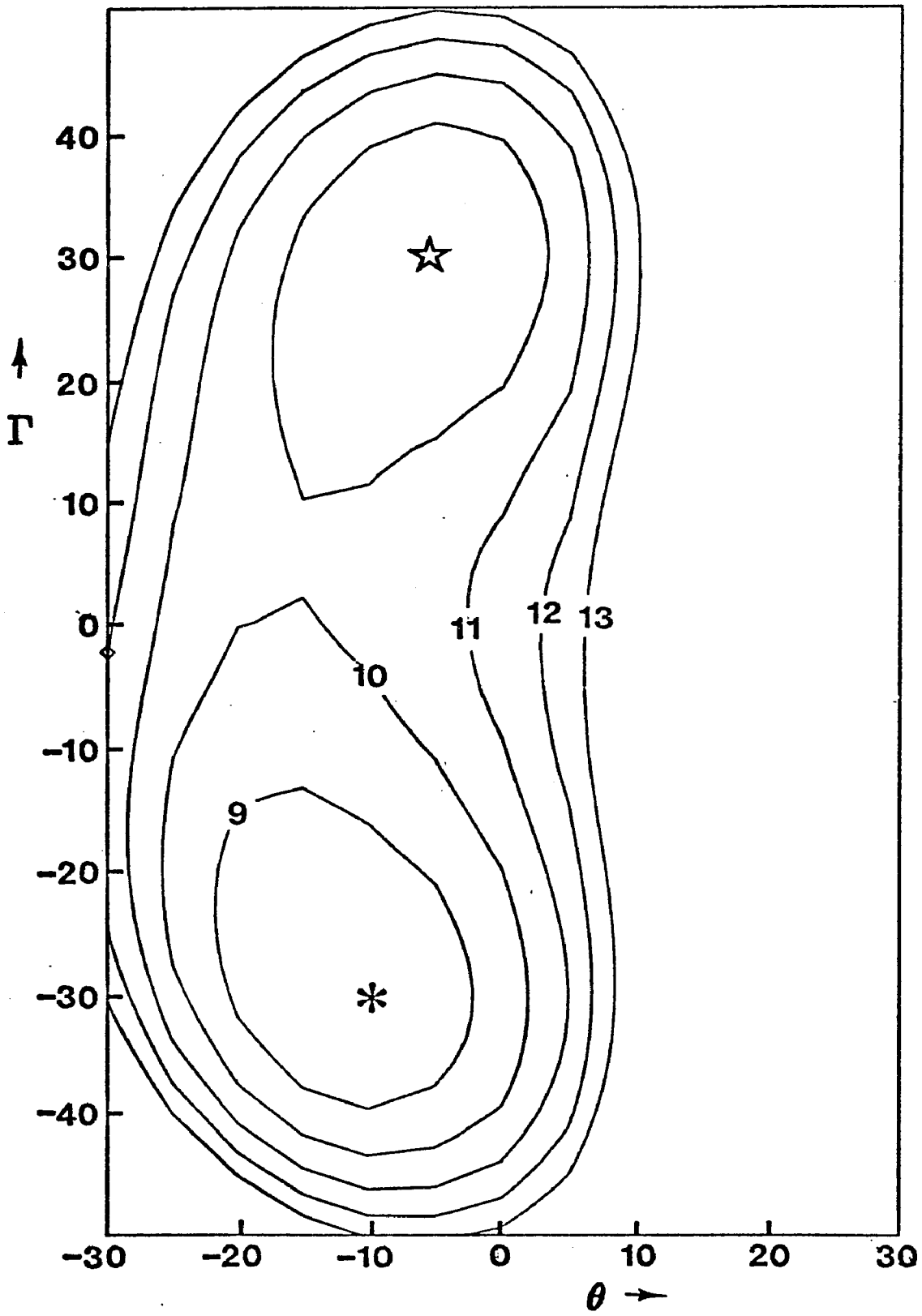


Figure 3-13. Γ-θ energy plot for the cis proline dimer. Each (Γ,θ) energy has been minimized with respect to ψ.

for the cis dimer. Both minimum energy structures occur for $\theta<0^{o}$. The energy difference between these structures is 1.0 kcal/mole, much higher than in the case of the isolated ring. Thus, the flexibility of the ring is markedly altered by the interactions in the dimer.

Figure 3-14 contains the Γ-θ plot for a trans proline dimer. [Here, we employed the backbone bond lengths and angles of Arnott and Dover (82)] The secondary minimum is 1.6 kcal/mole higher in energy than the global energy minimum. The energy contours extend somewhat more into the positive θ region than in the cis dimer; however, it is not likely that the value of θ will exceed 15^{o}.

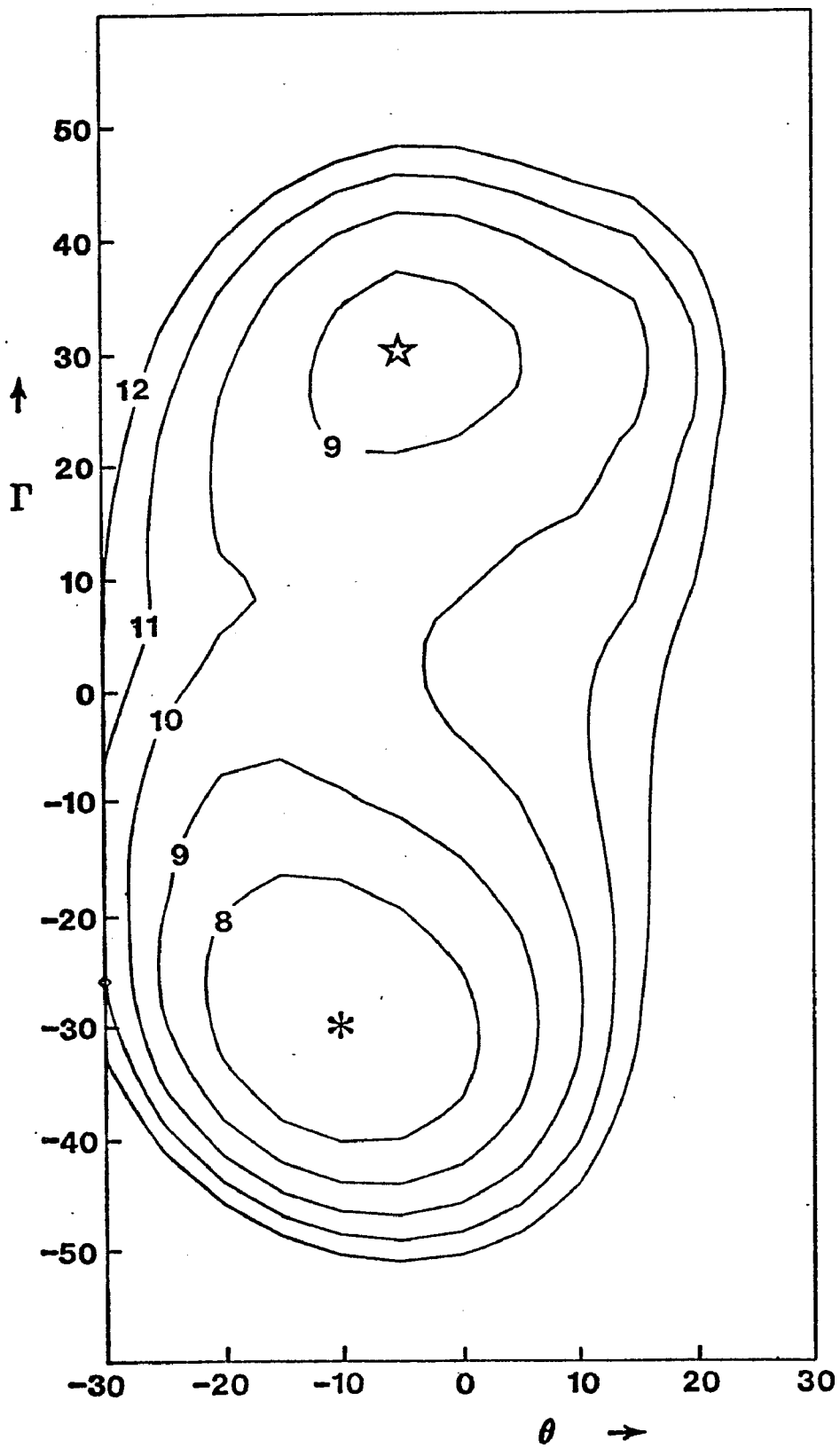


Figure 3-14. Γ-θ energy plot for the trans proline dimer. Each (Γ,θ) energy has been minimized with respect to ψ.

CHAPTER IV

METHYLATED PROLINE INTERNAL DIPEPTIDES

4.1 Structure and Coordinates of the 2- and 5-Methylprolines

Although they do not occur naturally, the 2- and 5-methylprolines are of interest to this study because of the role that the methyl group plays in ring puckering. In 5-methylproline, the methyl group replaces one of the hydrogens attached to the C^{δ} atom in the ring. This substitution may be done in either of two ways, as shown in Figure 4-1. In the cis structure, the methyl group is on the same side of the ring as the carbonyl carbon (C_2) of the same residue, while it is opposite to C_2 in trans 5-methylproline. The methyl group, of course, can rotate. The angle of rotation is denoted by χ. In view of the discussion of Section 3.7 regarding those atoms which influence ring puckering in the dimer, it seems likely that methylating the ring at the 5 position will lead to markedly different conformational features in the dimer.

Figure 4-2 shows the 2-methylproline internal monomer unit. There is one form of this molecule as only the L residue is considered. Model building shows that the 2-methylproline dimer is quite crowded. This is clearly

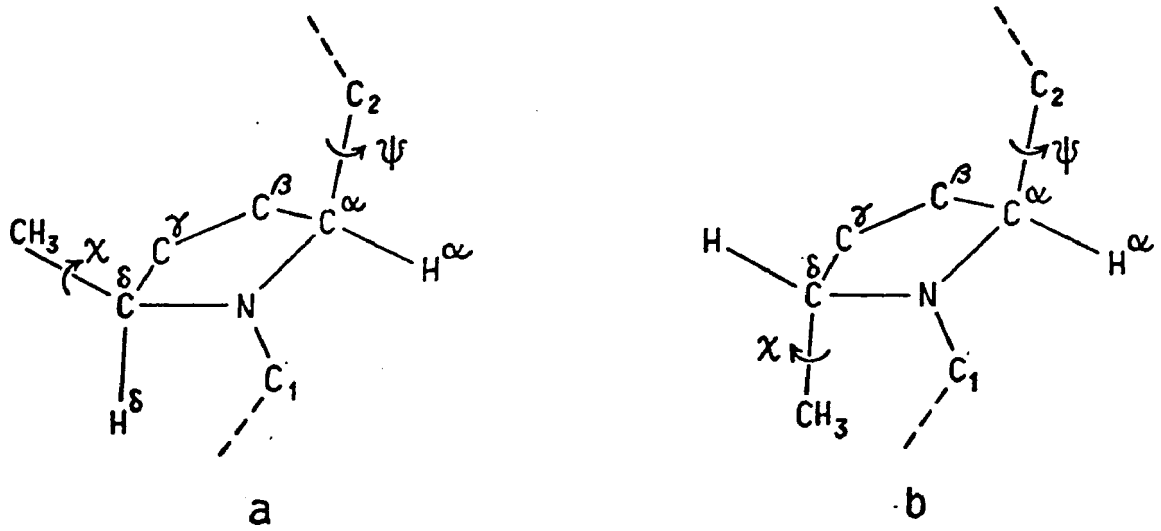


Figure 4-1. The 5-methylproline isolated rings.
(a) cis 5-methylproline
(b) trans 5-methylproline

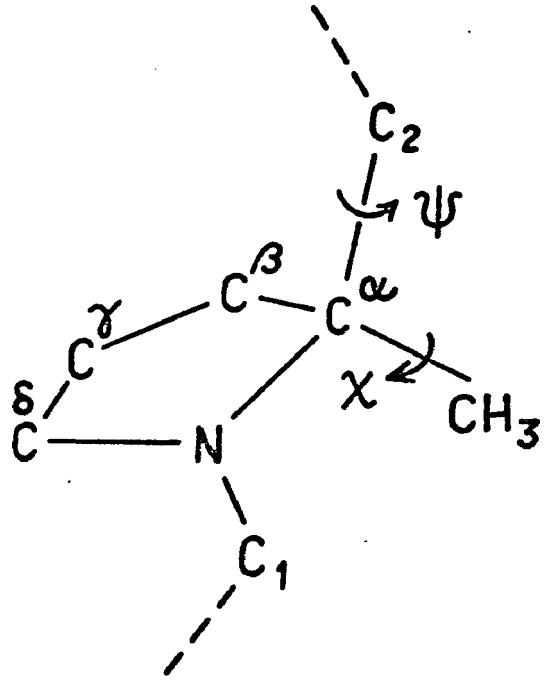


Figure 4-2. The 2-methylproline isolated ring.

the case, as results presented in later sections of this chapter show, but interesting conclusions arise from these results. We first consider the 5-methylproline internal dipeptides.

4.2 Analysis of Torsion Angles and Bond Angles for the 5-Methylprolines

The self-energy of the 5-methylprolines was calculated in an identical manner as for proline (see Section 2.4 for discussion). In addition to the five parameters of equation 1 (Chapter II), the ring energy is a function of χ. The bond length C^{δ}-C^{ϵ} was taken as 1.52Å. Initially, the angle χ (C^{γ}-C^{δ}-C^{ϵ}-H) was set at 60^{o} so as to stagger the methyl group hydrogens with the remaining substituents of C^{δ} (viz., C^{γ}, N and H^{δ}). The nature of the bond angle and torsion angle distribution for both the cis and trans 5-methylprolines is shown in Figure 4-3. There are no experimental data available for molecules containing 5-methylprolines with which to make comparisons. From Figure 4-3, it is seen that the cis ring prefers Type A geometries while Type B geometries are preferred by the trans ring. Note that these preferences are markedly greater than for proline. Also, as in the proline case, the only overlaps occur for the θ and $\chi^{N\delta}$ torsion angles; that is, γ- and β-puckered rings are the only single atom puckered rings available within 1.5 kcal/mole of the minimum. For other torsion angles, the A and B type geometries are well separated. Let us consider the bond

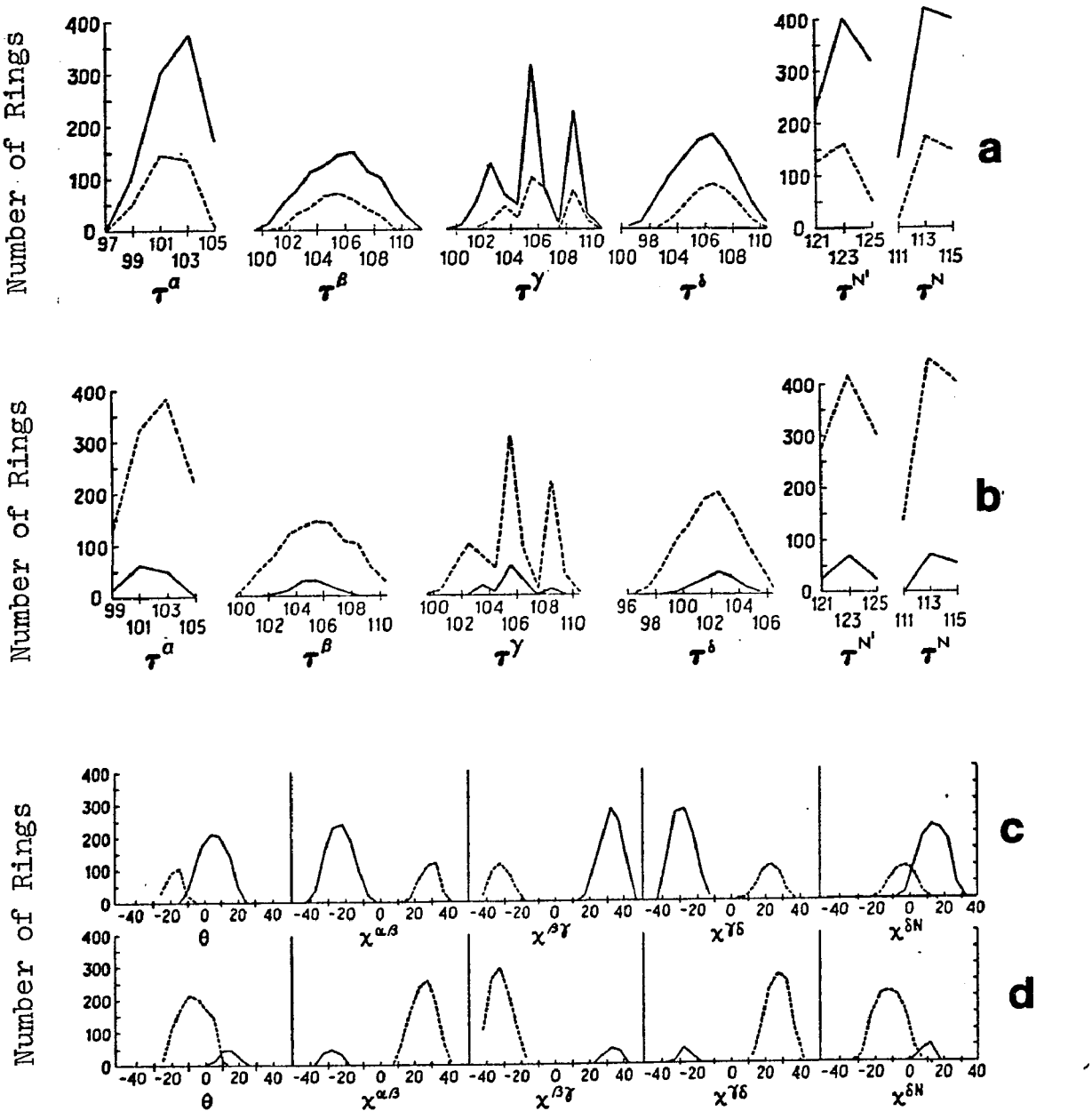


Figure 4-3. Distribution of bond angles and torsion angles in the set of 5-methylproline rings with energy less than 1.5 kcal/mole from the global minimum energy: (———) Type A geometries; (------) Type B geometries. a) bond angles, cis 5-methylproline, b) bond angles, trans 5-methylproline, c) torsion angles, cis 5-methylproline and d) torsion angles, trans 5-methylproline

angle distributions (Figure 4-3a and 4-3b). The preference for the Type A geometry by the cis structure and Type B geometry by the trans ring can be inferred from these distributions. Both cis and trans 5-methylprolines do not seem to show preferences for different bond angles. There are differences in detail, however, in comparing the 5-methylprolines. For example, in cis 5-methylproline, Type B geometry is preferred for $\theta<-5^{\circ}$ and Type A for $\theta>-5^{\circ}$, while for trans 5-methylproline this change from B to A occurs at $\theta=+5^{\circ}$. This influences the puckering near $\theta=0^{\circ}$, where the cis form will prefer γ^{+} puckering and the trans form will prefer γ^{-} puckering.

As mentioned above, the ring energy is also a function of χ. Calculations show that when $\chi=60^{\circ}$, the lowest ring self-energies are found. A departure from the staggered conformation by 30° produces an increase in ring energy of 2.3 kcal/mole. A 60° change in χ increases the energy difference to 4.3 kcal/mole. Therefore, χ was fixed at 60° for the remainder of the 5-methylproline calculations.

4.3 Γ-θ Plots for the 5-Methylprolines

Just as in the case of the proline monomer (discussed in Section 2.5), the cis and trans 5-methylproline Γ-θ plots were produced for the bond angles $\tau^{\alpha}=103^{\circ}$, $\tau^{N}=113^{\circ}$ and $\tau^{N'}=121^{\circ}$. These plots are shown in Figures 4-4 and 4-5. For the cis molecule, the M1, or global, energy minimum (E=12.0 kcal/mole) occurs at $\Gamma=34^{\circ}$, $\theta=5^{\circ}$. The M2 minimum (E=12.2 kcal/mole) is found at $\Gamma=-32^{\circ}$, $\theta=-10^{\circ}$. This is

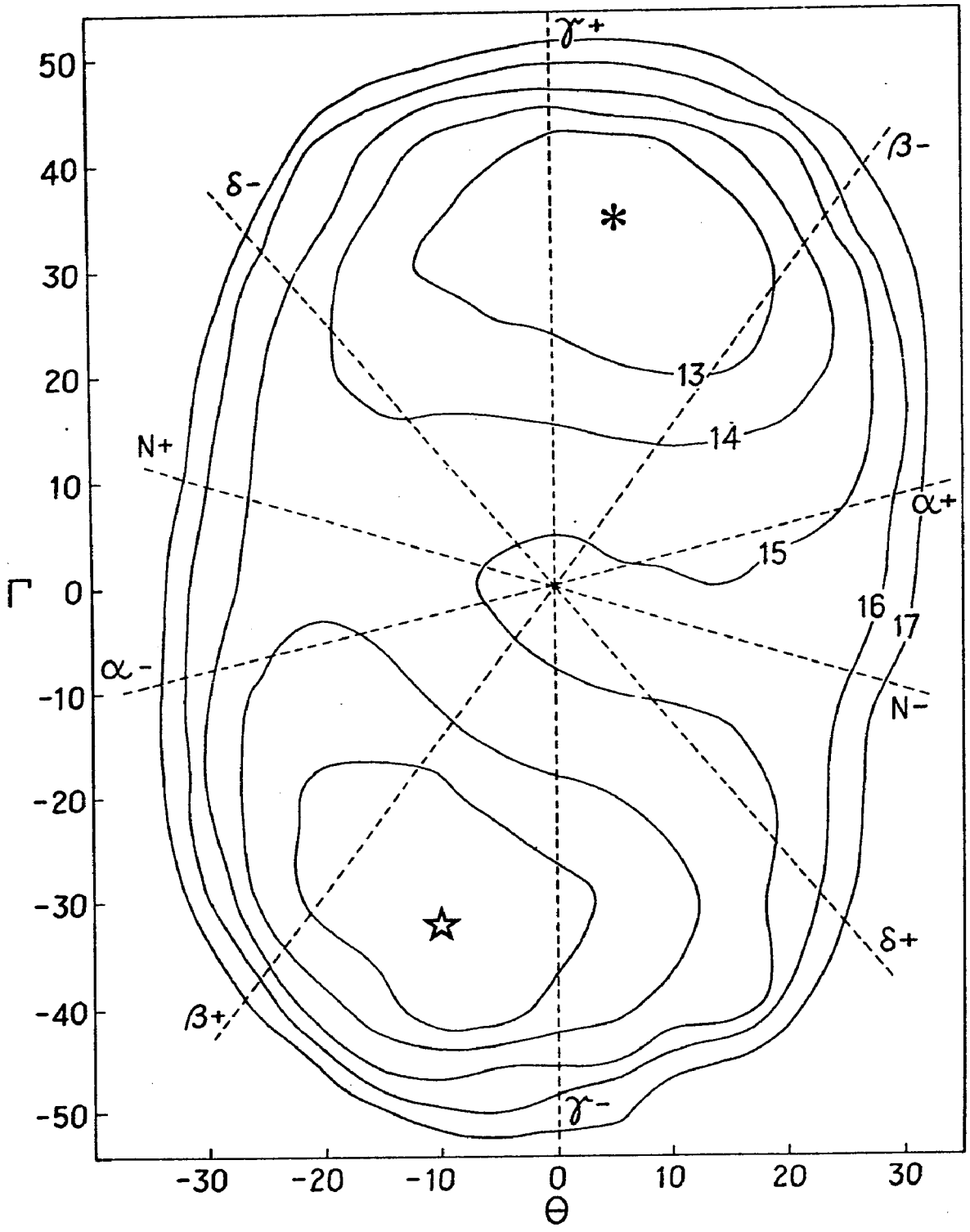


Figure 4-4. Γ-θ energy plot for the cis 5-methylproline isolated ring. $\tau^{\alpha}=103^{\circ}$, $\tau^{N}=113^{\circ}$ and $\tau^{N'}=121^{\circ}$.

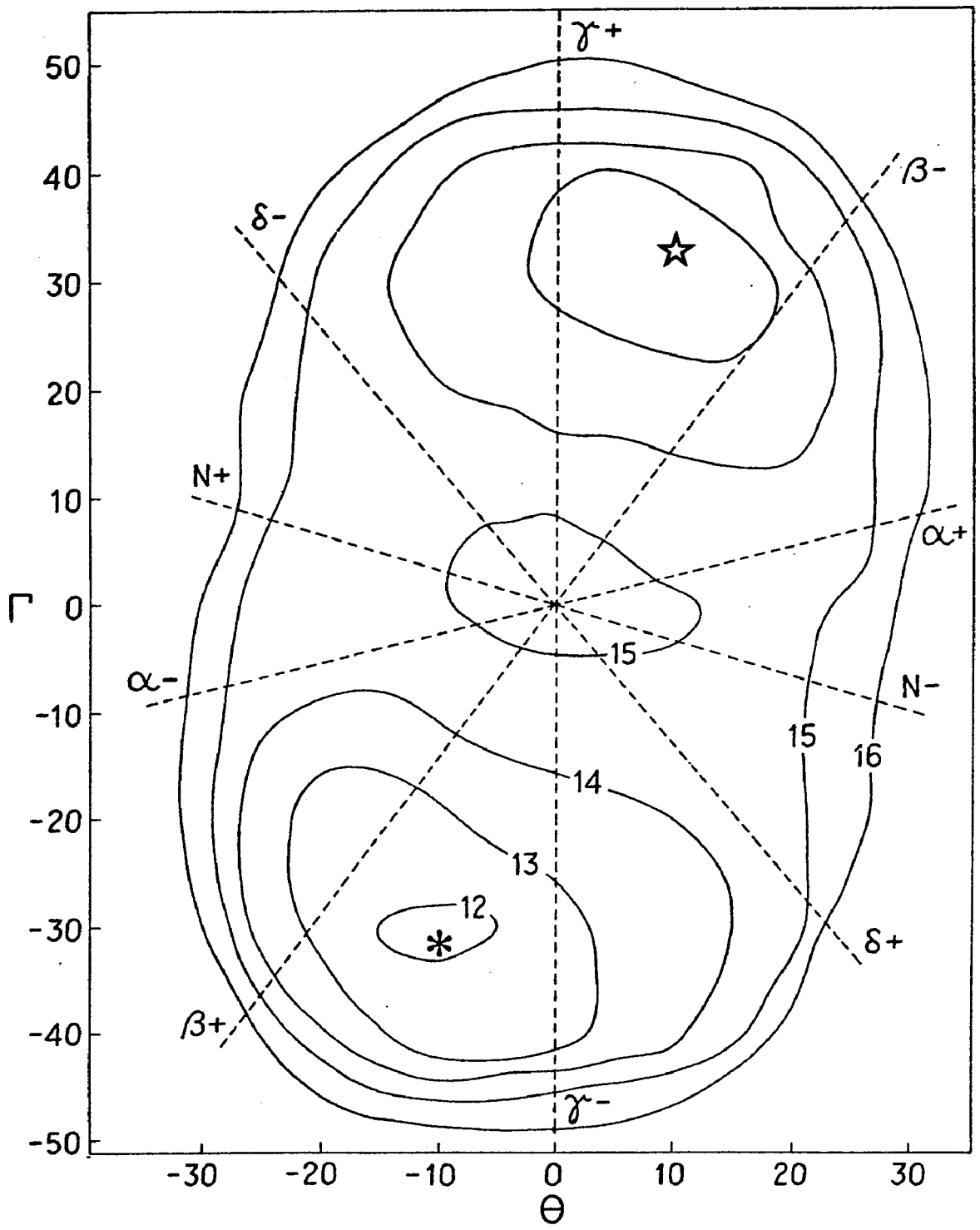


Figure 4-5. Γ-θ energy plot for the trans 5-methyl-proline isolated ring. $\tau^{\alpha}=103^{\circ}$, $\tau^{N}=103^{\circ}$ and $\tau^{N'}=121^{\circ}$.

to be contrasted to the trans molecule with M1 minimum (E=11.9 kcal/mole) located at $\Gamma=-32^{\circ}$, $\theta=-10^{\circ}$, while M2 (E=12.4 kcal/mole) is at $\Gamma=32^{\circ}$, $\theta=10^{\circ}$. The fact that the global minima appear in different portions of the map has already been inferred from the torsion angle analysis in the previous section. The values of the energy for each Γ and θ are contained in Tables 4-1 and 4-2.

The total ring energies as a function of θ for the variously perfectly puckered geometries, including Types A and B geometries are contained in Figure 4-6. As with proline there are two separate Γ values for which the self energy is a minimum. The minimum energy structure is not single atom puckered. It is important to note here that Figure 4-6 is derived from the same data that produced the torsion angle and bond angle distributions of Figure 4-3. Thus, the bond angles τ^{α}, τ^{N} and $\tau^{N'}$ have not been held constant as they were for the Γ-θ plots of Figures 4-4 and 4-5. Consequently, there may be some slight differences in accurate comparison of M1 or M2 energy minima and/or locations between Figures 4-4, 4-5 and Figure 4-6.

Note that, as stated above, the cis 5-methylproline M1 minimum is located at $\theta=+5^{\circ}$ for a Type A geometry while for the trans case (Figure 4-6), M1 occurs at $\theta=-10^{\circ}$ for a Type B geometry. In this respect, the trans 5-methylproline ring is similar to the proline ring. The difference in energy between the global energy minimum structure and the completely planar structure is about 3.3 kcal/mole in

Table 4-1. Γ-θ Energies for the cis 5-Methylproline Isolated Ring

Γ	-30	-25	-20	-15	-10	-5	0	5	10	15	20	25	30
60	34.8	31.0	28.6	27.2	26.6	26.3	26.1	25.8	25.9	26.5	28.2	31.5	37.0
50	25.4	22.2	19.9	18.3	17.3	16.6	16.3	16.2	16.5	17.3	18.8	21.2	24.9
40	20.4	17.7	15.8	14.5	13.6	13.0	12.7	12.6	12.8	13.3	14.5	16.3	19.1
30	17.8	15.6	14.2	13.4	12.9	12.5	12.3	12.1	12.1	12.4	13.1	14.4	16.6
20	16.6	15.0	14.1	13.8	13.6	13.5	13.4	13.2	13.0	13.0	13.3	14.2	15.9
10	16.0	14.8	14.3	14.3	14.5	14.6	14.7	14.6	14.4	14.2	14.3	14.8	16.1
0	15.5	14.5	14.2	14.4	14.7	15.1	15.3	15.3	15.2	15.0	15.1	15.5	16.7
-10	15.1	13.9	13.6	13.7	14.1	14.5	14.9	15.0	15.1	15.1	15.3	16.0	17.3
-20	15.0	13.5	12.8	12.7	12.9	13.3	13.7	14.0	14.3	14.6	15.3	16.2	17.9
-30	16.0	13.9	12.7	12.2	12.2	12.4	12.7	13.2	13.7	14.4	15.5	16.9	19.1
-40	19.2	16.2	14.4	13.4	12.9	12.9	13.1	14.0	14.8	15.9	17.4	19.5	22.4
-50	26.9	22.2	19.2	17.4	16.4	16.0	16.0	16.6	17.5	18.8	20.6	23.2	26.6
-60	48.4	36.7	30.1	26.3	24.1	23.1	22.8	23.2	24.2	25.7	28.0	31.2	35.5

θ

Table 4-2. Γ-θ Energies for the trans 5-Methylproline Isolated Ring

Γ	-30	-25	-20	-15	-10	-5	0	5	10	15	20	25	30
60	34.7	30.5	27.5	25.3	23.8	22.9	22.5	22.7	23.5	25.0	27.5	31.3	37.0
50	25.8	22.5	20.1	18.4	17.2	16.4	16.0	16.0	16.4	17.4	19.0	21.5	25.1
40	20.9	18.2	16.4	15.1	14.1	13.5	13.1	13.0	13.1	13.7	14.8	16.6	19.4
30	18.4	16.3	14.9	14.1	13.5	13.1	12.8	12.5	12.5	12.7	13.3	14.5	16.7
20	17.1	15.6	14.7	14.3	14.1	13.9	13.7	13.5	13.3	13.2	13.4	14.2	15.8
10	16.5	15.3	14.8	14.8	14.9	14.9	14.9	14.7	14.4	14.2	14.2	14.6	15.9
0	15.9	14.9	14.5	14.7	14.9	15.2	15.3	15.3	15.1	14.9	14.8	15.2	16.3
-10	15.4	14.2	13.7	13.8	14.1	14.5	14.7	14.8	14.8	14.8	14.9	15.5	16.8
-20	15.1	13.6	12.8	12.6	12.8	13.1	13.4	13.7	13.9	14.2	14.7	15.7	17.4
-30	16.0	13.8	12.6	12.0	11.9	12.0	12.3	12.7	13.2	13.9	14.9	16.3	18.5
-40	19.0	16.0	14.2	13.1	12.6	12.5	12.8	13.2	13.9	14.9	16.3	18.3	21.1
-50	26.7	22.0	19.1	17.4	16.5	16.3	16.4	16.9	17.6	18.8	20.4	22.8	26.1
-60	48.5	37.0	30.9	27.8	26.6	26.3	26.5	26.7	27.1	27.8	29.2	31.7	35.5

θ

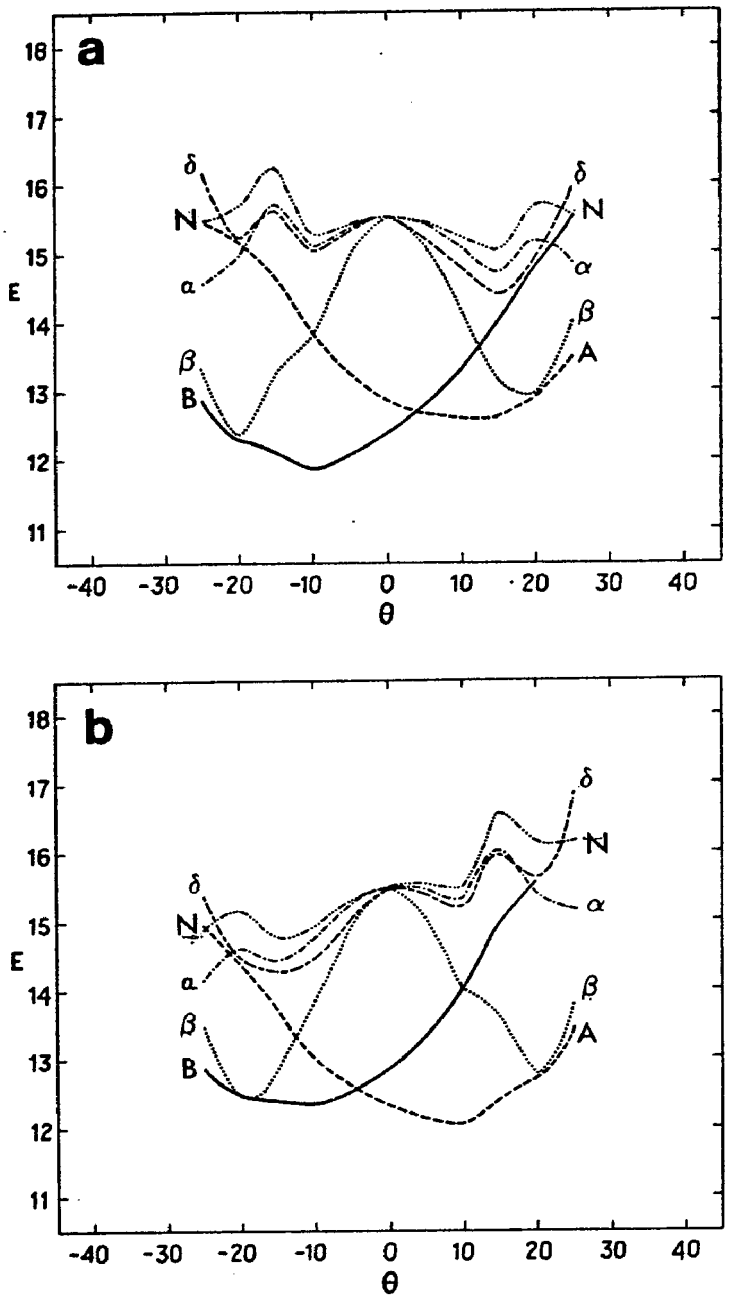


Figure 4-6. Self-energy of variously puckered 5-methylproline rings as a function of θ: a) trans 5-methylproline, b) cis 5-methylproline. Puckering types are marked at the end of each line.

both 5-methylprolines. Figures 4-4 through 4-6 clearly show that the puckering expected in the 5-methylprolines is very similar to that found in proline. At $\theta \sim 0^{\circ}$, the ring is expected to be γ-puckered, while at $|\theta| \sim 20^{\circ}$, it will be β-puckered. Model building indicates that the methyl group should not drastically affect puckering in the isolated ring when in a staggered conformation. The results seem reasonable from that point of view. The dimer, of course, presents a different story.

4.4 (θ,ψ) Conformational Energy Maps for the 5-Methylproline Dimers

Figures 4-7 through 4-10 contain the (θ,ψ) energy contour plots for the two types of 5-methylproline dimers, with the imide bonds in either the cis or trans region. In this section, we refer to a molecule as X 5-methylproline Y where X designates the methyl configuration, either cis or trans and Y designates the ω rotational state, again either cis or trans. Following similar calculations for the proline dimer (Section 3.2), the following parameters were varied:

$\{\omega: 155^{\circ}, 205^{\circ}, 5^{\circ}\}$ for the trans dimer

$\{\omega: -25^{\circ}, +25^{\circ}, 5^{\circ}\}$ for the cis dimer

$\{\theta: -30^{\circ}, 30^{\circ}, 5^{\circ}\}$

$\{\psi: -180^{\circ}, 180^{\circ}, 10^{\circ}\}$

Guided by previous experience, ω' was held constant at 180° or 0° depending upon the dimer under consideration. For

each θ, six Γ's were utilized corresponding to the four single atom puckered geometries as well as A and B type minimum energy ring structures. Initially the χ's for both rings were held constant at 60°. Effect of χ rotation on the dimer energies will be discussed later. Symbols denote the lowest energy ring geometry for each (θ,ψ) point; dots refer to M1 geometries, crosses indicate M2 geometries.

We first consider the 5-methylproline <u>trans</u> dimers. Figure 4-7 contains the <u>cis</u> 5-methylproline <u>trans</u> dimer energy plot. There are two minima; the deeper one is located at $\theta=0^{\circ}$, $\psi=120^{\circ}$ with M1 geometry. From Figure 4-4, this corresponds to γ^{+} puckering. The secondary minimum is at $\theta=-20^{\circ}$, $\psi=130^{\circ}$ with β^{+} puckering. Note that in the high energy region separating these two minima, the dimer prefers the M2 geometry, which for the <u>cis</u> ring is Type A. Thus, according to Figure 4-4, in this region the puckering is closest to γ^{+}. If the <u>trans</u> 5-methylproline <u>trans</u> dimer is considered (Figure 4-8), it is seen that there is a lot less conformational flexibility. There is one minimum at $\theta=-5^{\circ}$, $\psi=150^{\circ}$. It is an M1 minimum and, therefore is γ^{-} puckered. These results indicate that the <u>trans</u> methyl configuration reduces the conformational flexibility in the dimer. In Figure 4-8 the energetically allowed ranges of <u>both</u> θ and ψ are shortened.

For the <u>cis</u> imide bond conformation, the (θ,ψ) maps for both 5-methylprolines (Figures 4-9 and 4-10) are quite

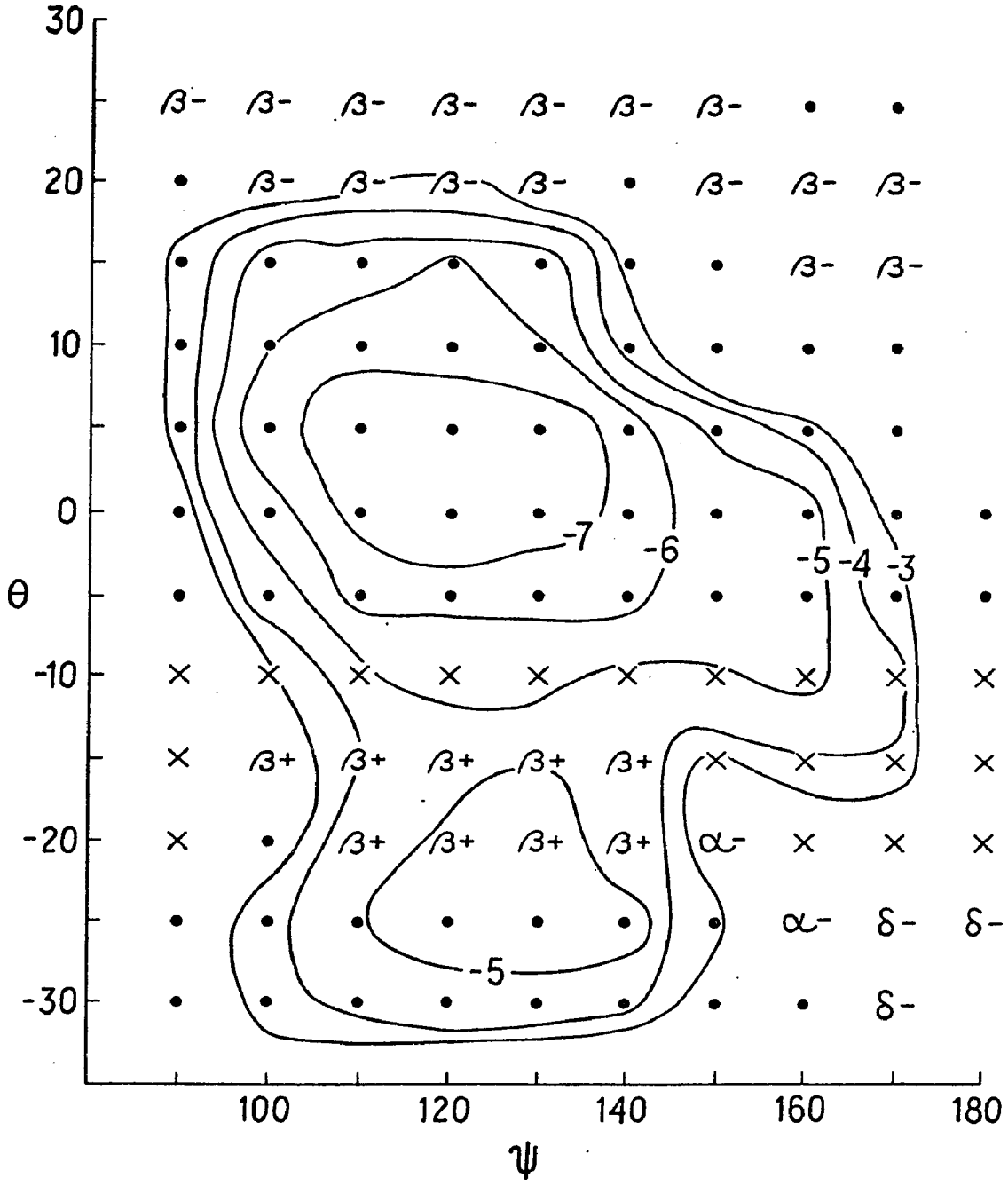


Figure 4-7. θ-ψ conformational energies for the cis 5-methylproline trans dimer. The best ring geometry for each (θ,ψ) is symbolically represented (see text).

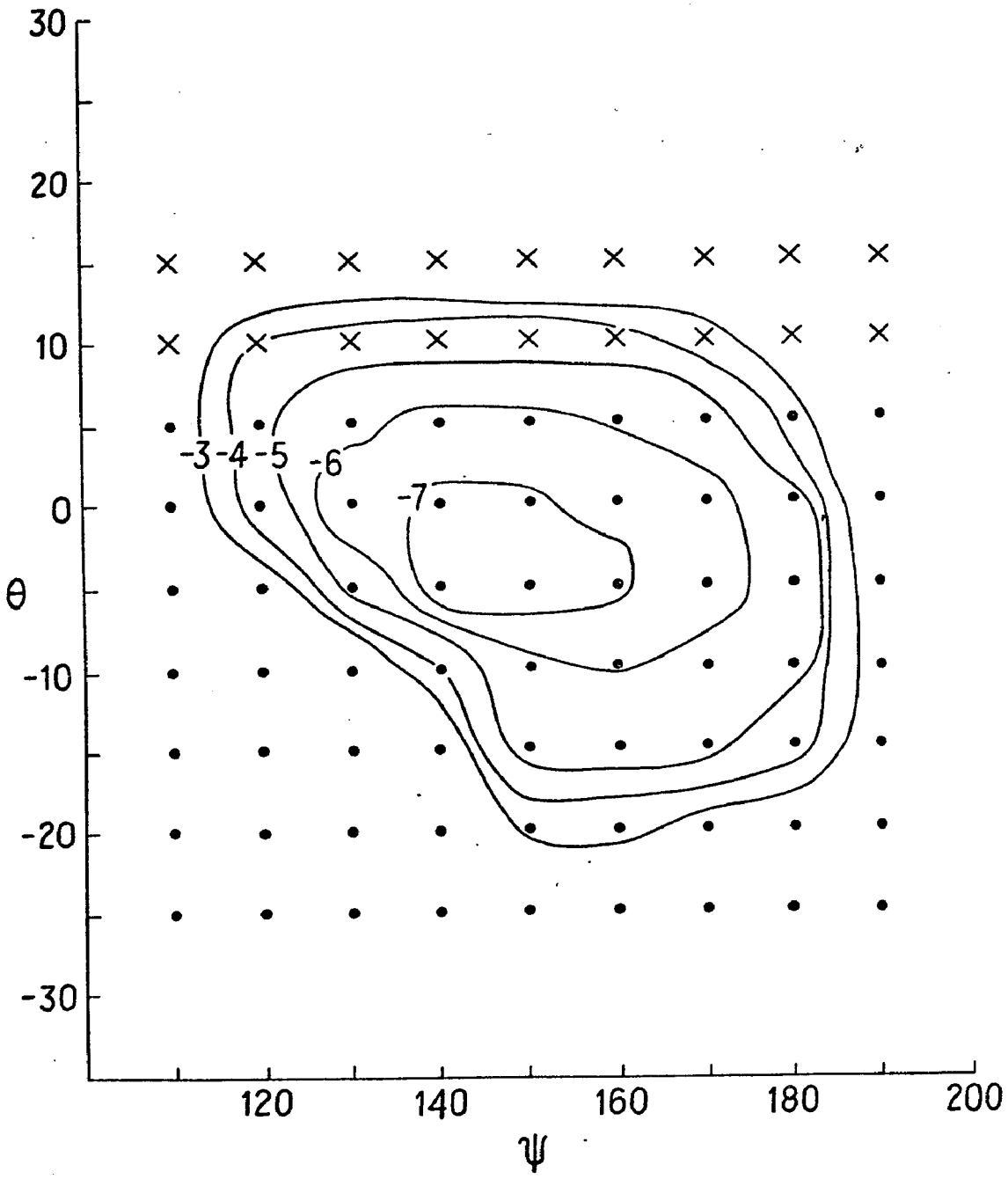


Figure 4-8. θ-ψ conformational energies for the trans 5-methylproline trans dimer. The best ring geometry for each (θ,ψ) is symbolically represented (see text).

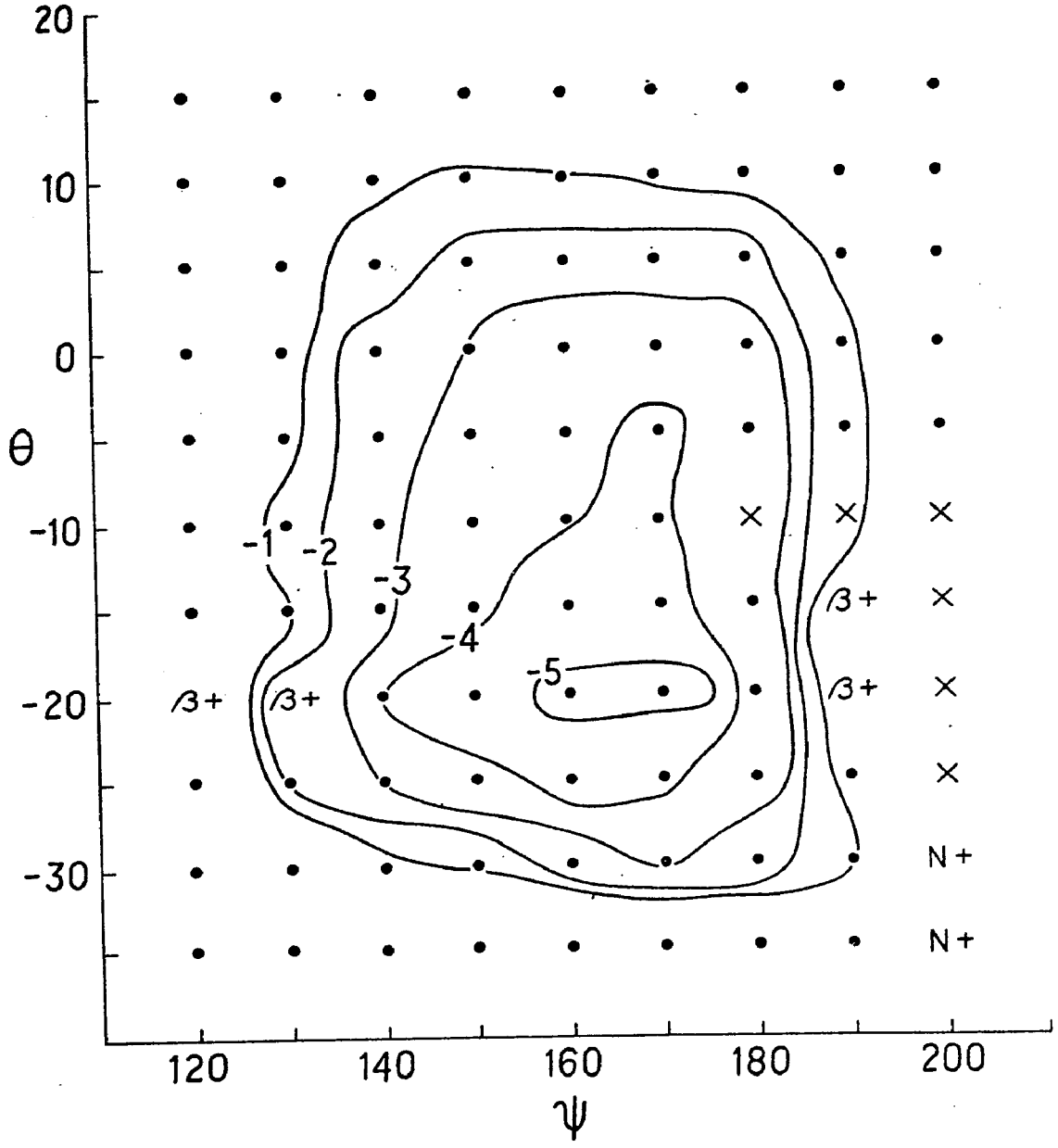


Figure 4-9. θ-ψ conformational energies for the cis 5-methylproline cis dimer. The best ring geometry for each (θ,ψ) is symbolically represented (see text).

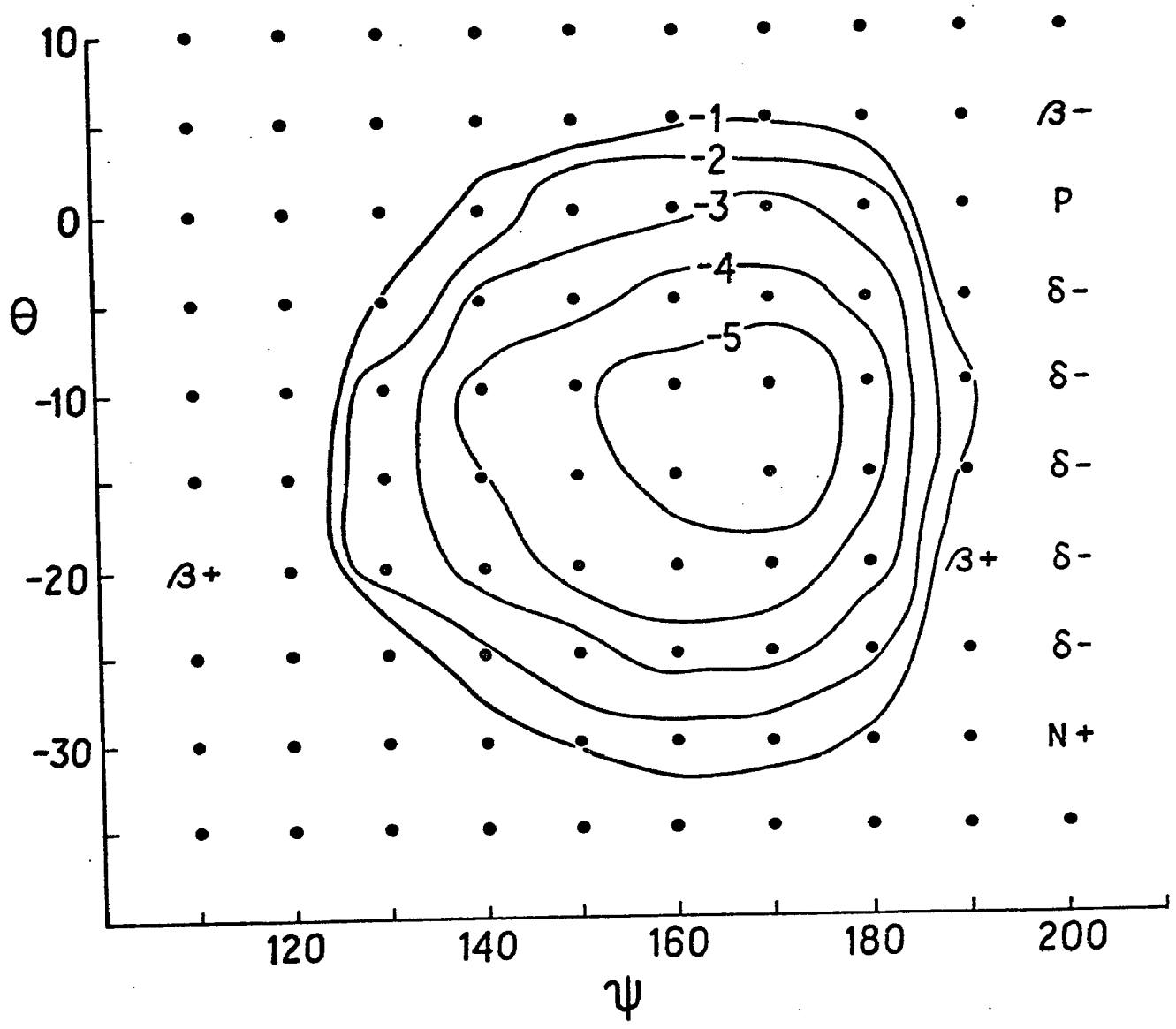


Figure 4-10. θ-ψ conformational energies for the trans 5-methylproline cis dimer. The best ring geometry for each (θ,ψ) is symbolically represented (see text).

similar to the cis proline dimer of Figure 3-3. For the cis 5-methylproline cis dimer, the minimum is M1 type (γ^+ puckered) and is located at $\theta=-20^\circ$, $\psi=160^\circ$. The minimum energy structure for the trans dimer is also M1, but is γ^- puckered. It is located at almost the same spot; $\theta=-15^\circ$, $\psi=160^\circ$ as the cis dimer. It seems, therefore, that the methyl substitution at the C^δ atom has little effect when the imide bond is cis. This is definitely not the case for the trans imide bond.

4.5 Energy Barriers Between Stable Conformations of the 5-Methylproline Dimers

In the calculations of the previous section, no cis' region was found for either of the 5-methylprolines. In view of our earlier discussion of flexibility for the proline dimer in Sections 3.5-3.7, it is of interest to inquire how complete ω rotational freedom would affect the barriers between the allowed conformations in the 5-methylprolines. Accordingly, we have calculated the dimer energies in the (ω,ψ) plane allowing for θ variation, initially requiring both rings to have identical puckering. The energy of the dimer was therefore a function of five parameters:

$$E_{dimer} = E(\Gamma,\theta,\omega,\psi,\chi)$$

ω' was held constant at 180° for all values of ω and ψ. The variation of the other parameters was as follows:

$\{\theta: -25^\circ, 25^\circ, 5^\circ\}$

$\{\omega: -180^\circ, 180^\circ, 20^\circ\}$

$\{\psi: -180^\circ, 180^\circ, 20^\circ\}$

All potentials, charges, etc. are identical to the earlier calculations, except that the methyl group hydrogens were assumed to have zero charge. Initially χ was fixed at 60°. After having obtained the (ω,ψ) maps, χ was varied only in those lowest energy regions of interest. This will be discussed later in this section. For these and subsequent dimer calculations, we have used only the M1 geometry for each θ. This cuts the computing time by a factor of 5/6 and is a reasonable simplification when one considers that the M1 geometries are the exclusive survivors in all low energy regions of the previous dimer energy calculations. Single atom puckered geometries play a role only in high energy regions and in fact, in these regions, there is only a small energy difference between the favored single atom puckered geometries and the corresponding M1 geometries. At each (ω,ψ) point, therefore, the dimer could choose from 11 different geometries corresponding to the M1 geometry for each θ value.

The ring self-energies were taken from Tables 4-1 and 4-2. Therefore, just as in the case of the proline dimer, the bond angles τ^{α}, τ^{N} and $\tau^{N'}$ were held constant at 103°, 113° and 121° respectively. Figure 4-11 contains the result for the <u>cis</u> 5-methylproline dimer. Compared to

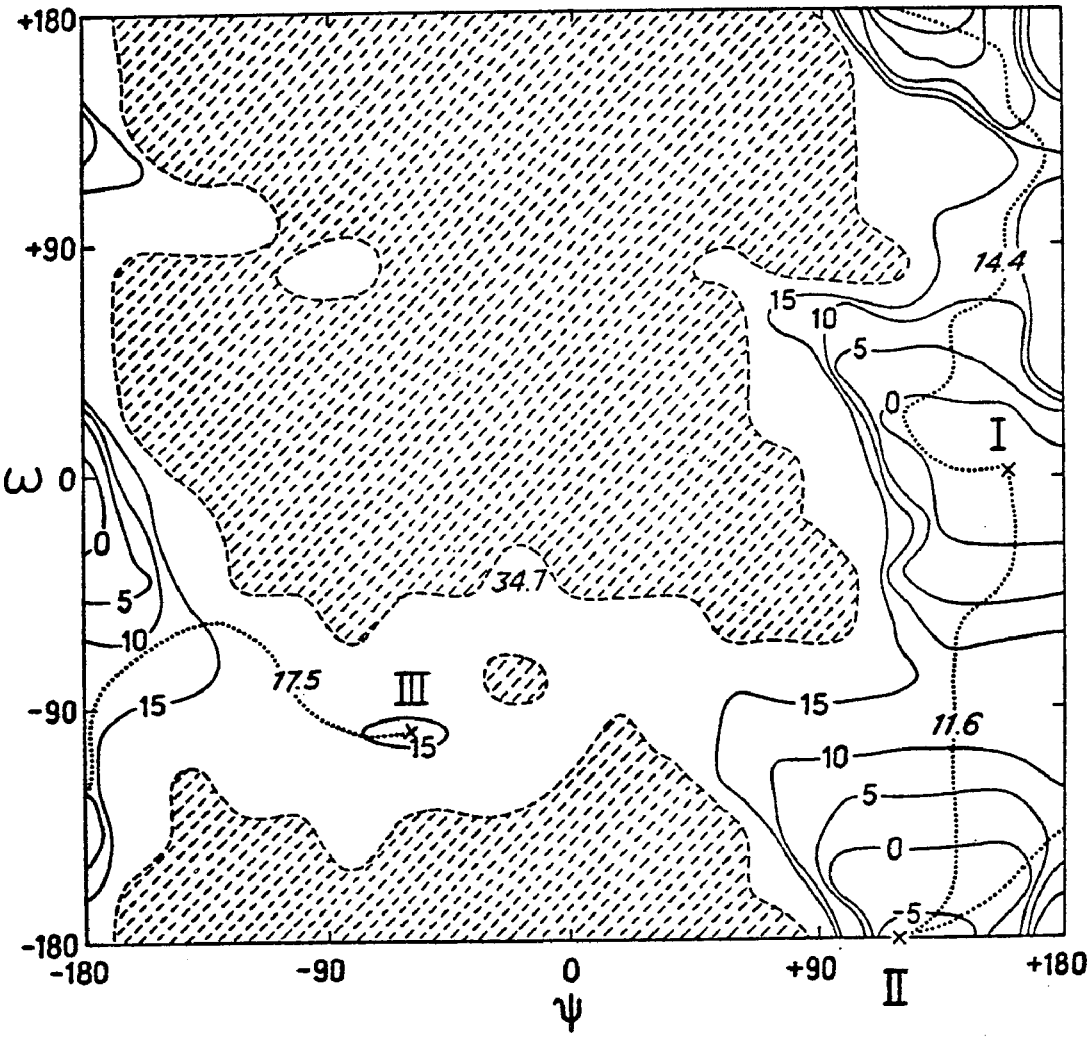


Figure 4-11. (ω,ψ) energy contour map for the cis 5-methylproline internal dipeptide.

Figure 3-7 for the proline dimer, there is much less conformational flexibility, especially in the trans region. Table 4-3 contains the information about the minimum energy points for the three low energy regions. The cis and trans minimum energies are not very different, as in proline. Also, the barrier energies are of the same order of magnitude. These energies and the torsion angles associated with them are shown in Table 4-4. The striking feature of this map is the complete absence of an energetically competitive cis' region. The same is true for the trans 5-methylproline dimer, as Figure 4-12 indicates. Even though the Region III minimum position is shifted by 180° in ω, it is still almost 17 kcal/mole higher in energy than the Region I minimum. It should be noted that the trans 5-methylproline dimer is less sterically hindered in Region II than the cis dimer. In this sense, the trans structure compares more favorably with the proline dimer. An interesting point in this comparison between the 5-methylproline dimers and the proline dimer is that the values of θ for the energy minimum structures of Region I are largely negative in both cases. In the proline dimer, Region II minimum energy θ's remain negative, but for both 5-methylprolines, θ is positive. Insofar as barrier height energies go, the θ values for the trans 5-methylproline conformations at the barrier heights are closer to the proline θ values for both the cis $\rightarrow$ trans and cis' $\rightarrow$ trans' transitions.

Table 4-3. The Torsional Angles (in degrees) and Energy Values (in kcal/mole) of Minimum Energy Conformations of the cis and trans 5-Methylproline Internal Dipeptides

Figure	Region I					Region II					Region III				
	ω	ψ	θ_1	θ_2	E	ω	ψ	θ_1	θ_2	E	ω	ψ	θ_1	θ_2	E
4-11	0,	160,	-20,	-20	-4.4	180,	120,	5,	5	-6.4	-100,	-60,	10,	10	12.4
4-12	0,	160,	-15,	-15	-5.5	180,	140,	0,	0	-6.0	120,	20,	-15,	-15	11.5
4-13	0,	140,	0,	-15	-5.5	180,	120,	0,	5	-6.9	-120,	-60,	15,	0	10.5
4-14	0,	160,	10,	-15	-5.7	180,	140,	-10,	0	-6.3	140,	-40,	10,	0	8.8

Table 4-4. The Torsional Angles (in degrees) and Energy Values (in kcal/mole) of the Conformations at the Barriers Separating the Low Energy Conformations of the cis and trans 5-Methyl-proline Internal Dipeptides

Figure	cis - trans					cis' - trans'					cis - cis'				
	ω	ψ	θ_1	θ_2	E	ω	ψ	θ_1	θ_2	E	ω	ψ	θ_1	θ_2	E
4-11	-100,	140,	5,	5	11.6	-80,	-100,	5,	5	17.5			———	*	
4-12	-100,	140,	-5,	-5	10.8	100,	40,	-20,	-20	16.5			———	*	
4-13	-100,	140,	0,	10	11.2	-80,	-100,	5,	0	16.6			———	*	
4-14	-60,	140,	-5,	-5	10.8	80,	60,	-15,	10	12.9			———	*	

*No pathway directly linking cis with cis' is found.

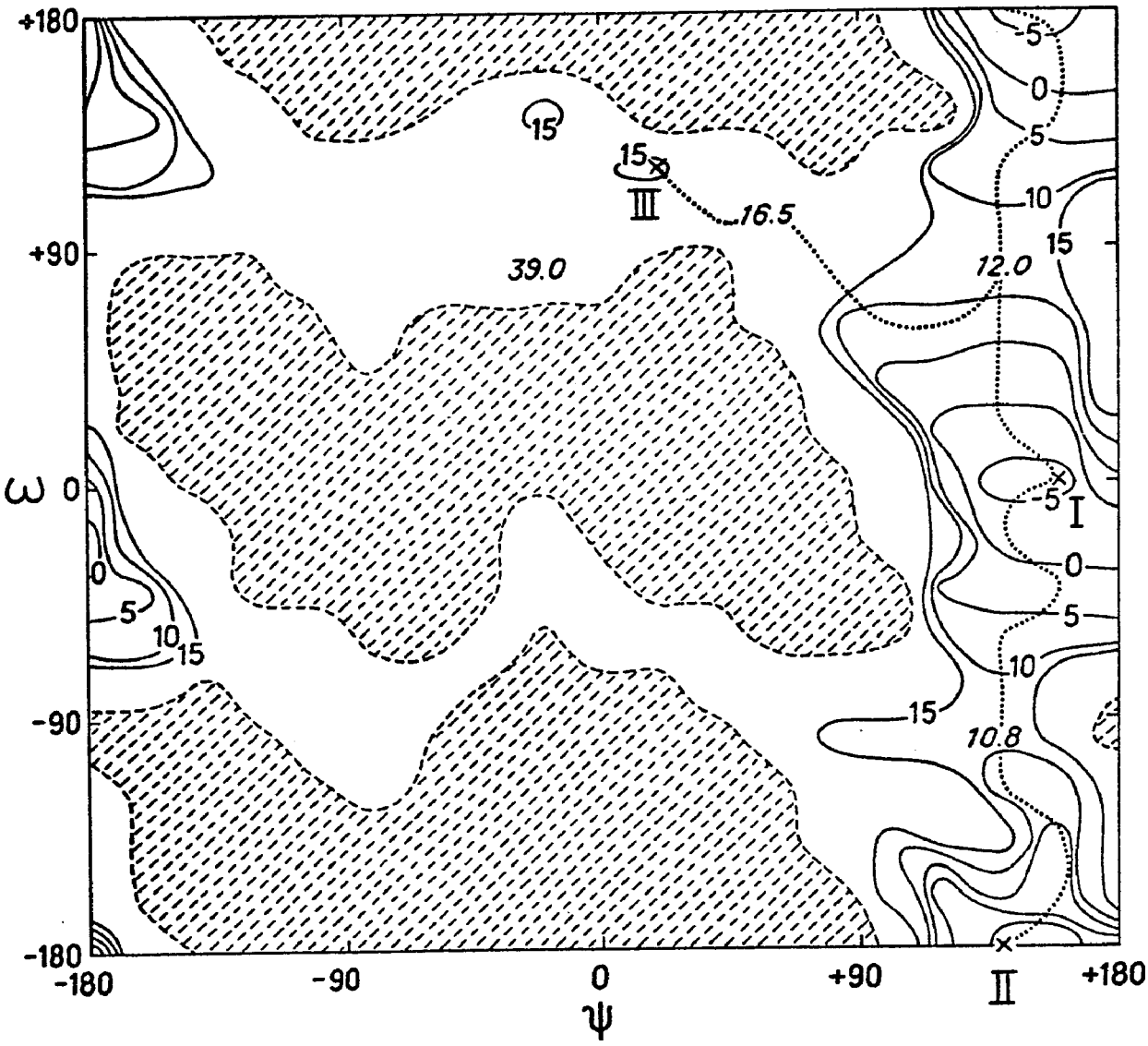


Figure 4-12. (ω,ψ) energy contour map for the trans 5-methylproline internal dipeptide.

As noted above, the energy of the dimer is a function of χ also. The important conclusion to be drawn from Figures 4-11 and 4-12 is that the cis' structure does not exist for either the cis or trans 5-methylproline dimer. It is possible, however, that χ rotation could bring the energy at the cis' minimum down to a level competitive with the trans' minimum. To investigate this, the methyl group hydrogens were rotated in concert:

$$\{\chi_1,\chi_2: \ 60^\circ, \ 160^\circ, \ 20^\circ\}$$

This was done at the (ω,ψ) point where the cis' minimum occurs for both 5-methylproline dimers. In each case, θ was allowed to vary:

$$\{\theta: \ -25^\circ, \ 25^\circ, \ 5^\circ\}$$

All geometries were M1 type. The best energy for each θ and the value of θ for which that energy is found is contained in Table 4-5. Methyl group rotation does nothing to enhance the cis' minimum energy for either of the 5-methylprolines. It does lower the energy by about 1.3 kcal/mole for the trans case, but that is of little significance when compared to the trans' energy minimum.

4.6 (ω,ψ) Energy Maps for the 5-Methylproline Dimers with Independently Puckered Rings

In this section, the effect of independent ring puckering in the 5-methylproline dimers is investigated. For each (ω,ψ), the θ's of both rings were varied

Table 4-5. Effect of Methyl Group Rotation on the 5-Methylproline Dimer Energies in the *cis'* Region

χ	*cis*[a]		*trans*[b]	
	θ	E	θ	E
60	10	12.4	-15	11.5
80	10	12.7	-15	11.4
100	15	13.9	-20	10.4
120	-5	13.4	-20	10.2
140	15	16.0	-20	10.3
160	10	13.1	-20	11.2

[a]For the *cis* 5-methylproline dimer, these calculations were done for $\omega=-100^{\circ}$, $\psi=-60^{\circ}$; that is, at the *cis'* minimum point in Figure 4-11.

[b]For the *trans* 5-methylproline dimer, these calculations were done for $\omega=120^{\circ}$, $\psi=20^{\circ}$; that is, at the *cis'* minimum point in Figure 4-12.

independently:

$$\{\theta_1, \theta_2: -25^\circ, 25^\circ, 5^\circ\}$$

As in the case of the proline dimer, we considered the cis, trans and cis' regions along with the corridors of (ω,ψ) points interconnecting these stable regions. Figures 4-13 and 4-14 show the results for the cis and trans 5-methylprolines respectively. There is almost no change in the energy distribution for Regions I, II and III. As can be seen from Tables 4-3 and 4-4, there is a small drop in the energies at the three minima as well as barrier height energies. Thus it is seen that, except for some fine detail, independent ring puckering does not alter any previous conclusions made about the 5-methylproline dimers. We therefore restate our earlier point; namely, there is no cis' structure available to either the cis or trans 5-methylproline internal dipeptides.

Figures 4-15 through 4-20 are stereoscopic drawings for the minimum energy cis, trans and cis' structures for both the cis and trans 5-methylproline dimers. As for the proline case, the backbone atoms along the chain are indicated by letters. The chain direction is from top to bottom in each diagram. The coordinates are taken from the independent ring puckering program and have been plotted by the University of Michigan CalComp Plotter using the ORTEP2 program (59). It is obvious in Figures 4-15 and 4-16 that the methyl group does not impose any

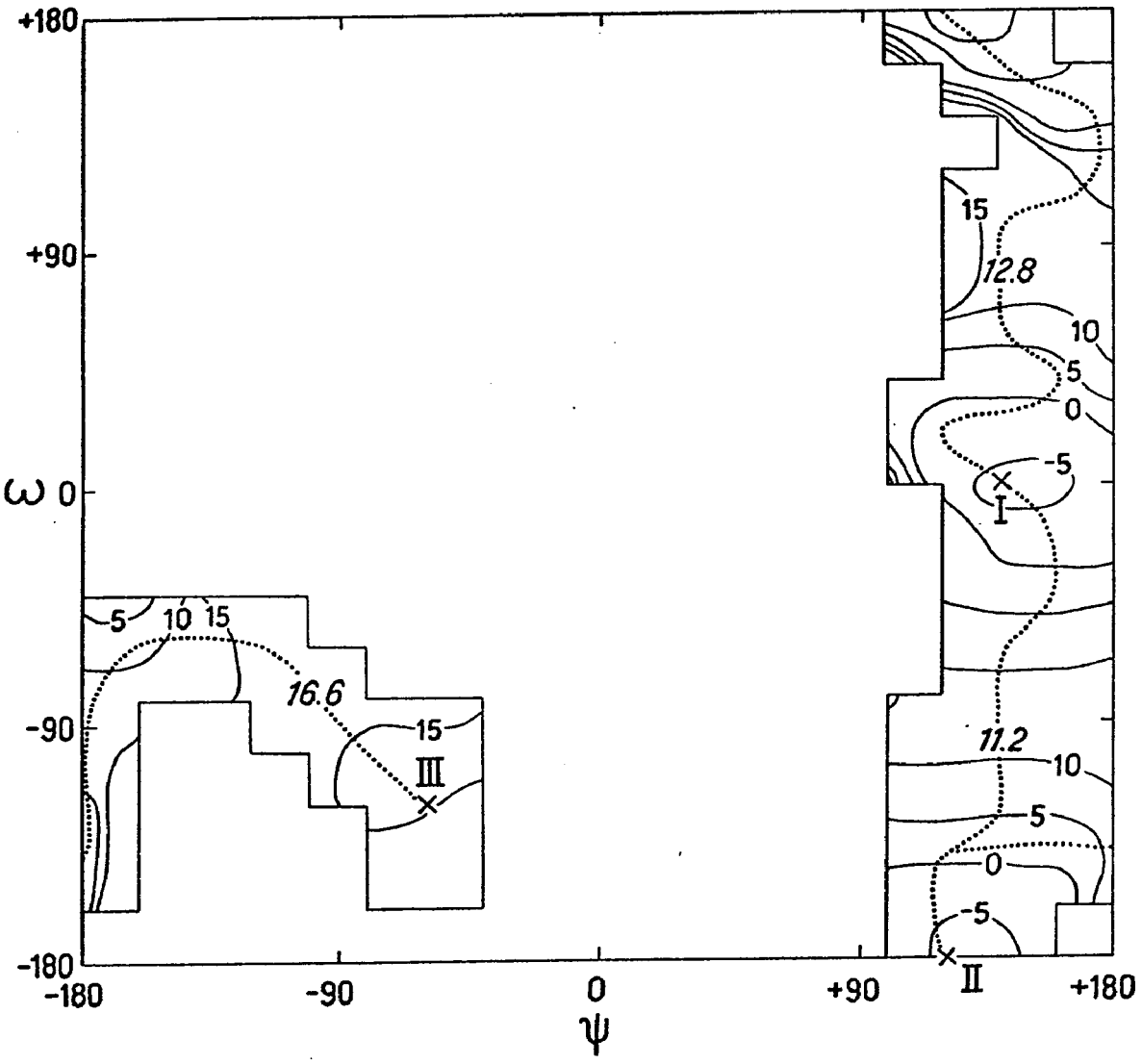


Figure 4-13. (ω,ψ) energy contour map for the cis 5-methylproline internal dipeptide with independently puckered rings.

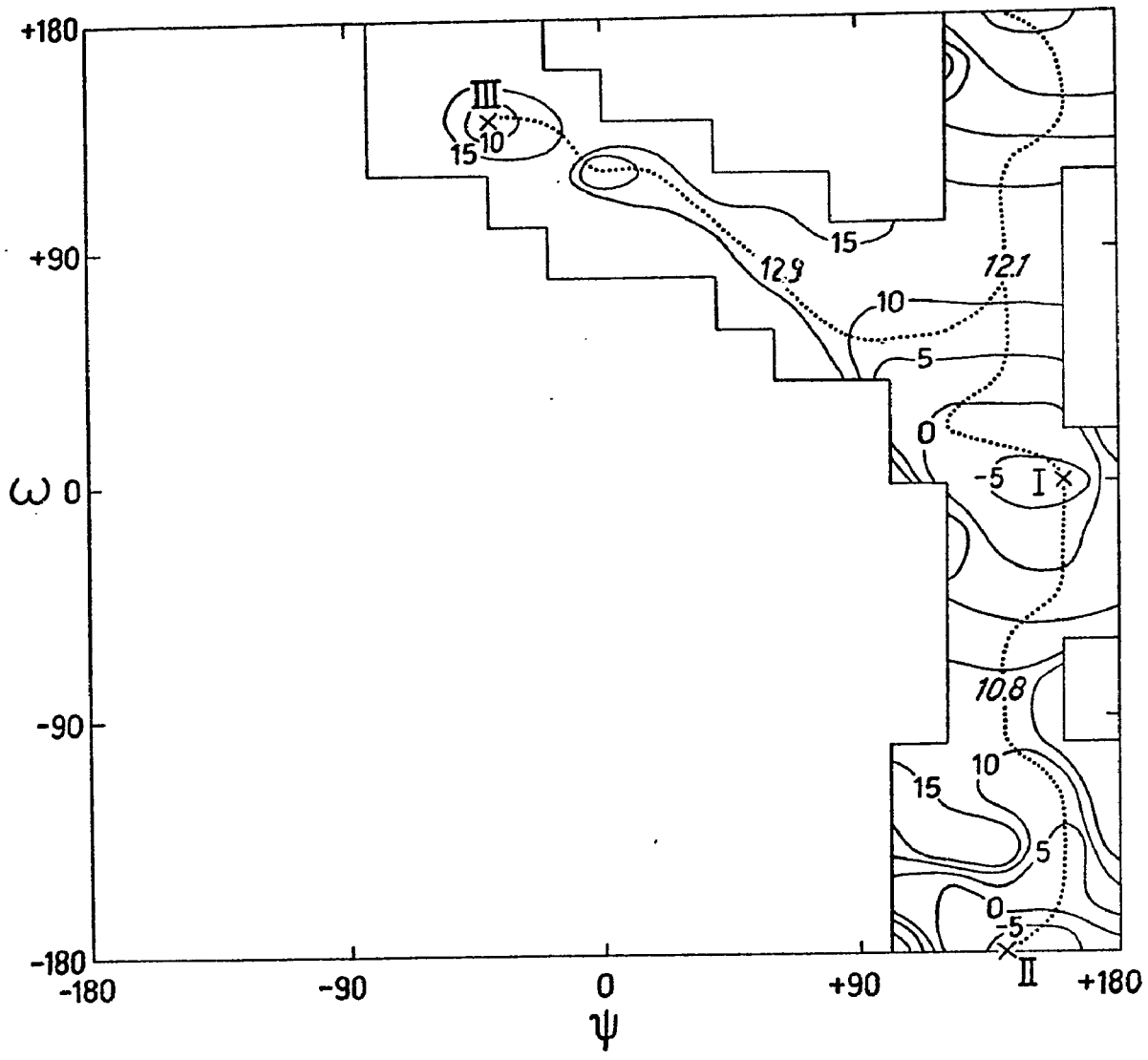


Figure 4-14. (ω,ψ) energy contour map for the trans 5-methylproline internal dipeptide with independently puckered rings.

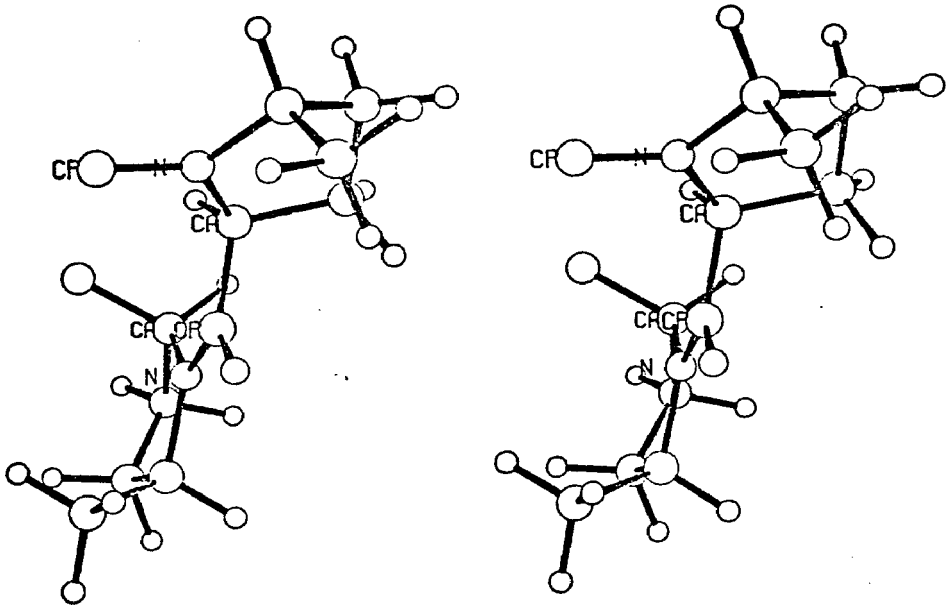


Figure 4-15. Cis (5-methyl pro)$_2$ cis

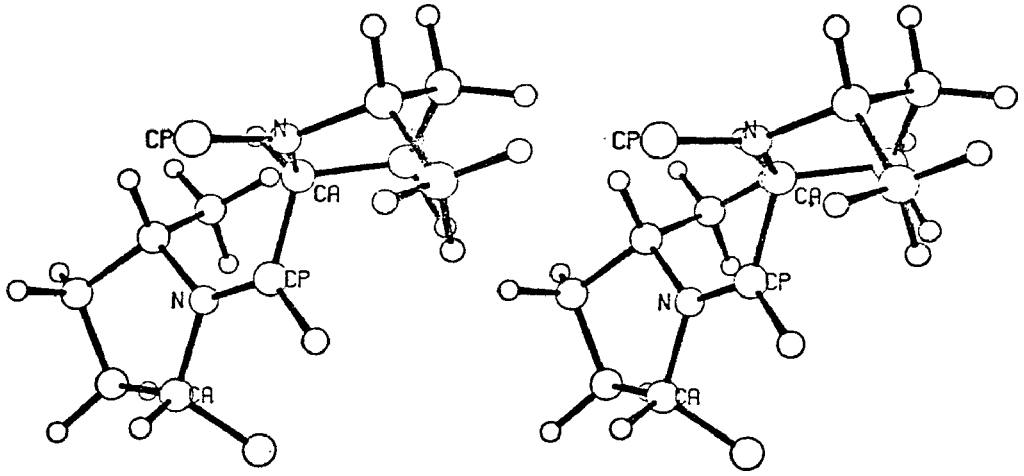


Figure 4-16. Cis (5-methyl pro)$_2$ trans

steric hindrances upon the structure. In each case, the methyl groups point away from each other as well as the pyrrolidine rings. It is equally obvious why the cis 5-methylproline cis' structure (Figure 4-17) does not exist. The methyl groups point toward each other. In particular, the first methyl group is pointing towards the second ring. One of the methyl hydrogens is in close contact with H^{δ} of the second ring while another methyl hydrogen is close to one of the second methyl group hydrogens. In addition, the second methyl group is close to the initial C atom. It can be seen that methyl rotation will not help in alleviating bad steric contacts for this conformation. Only a positive ω or ψ rotation will relieve the congestion. Figures 4-18 and 4-19 depict the minimum energy structures for the trans 5-methylproline cis and trans 5-methylproline trans dimers. Here, as in the cis 5-methylprolines, the two methyl groups point away from each other and away from the rings. In the disallowed conformation of Figure 4-20, the first methyl group points away from the molecule and is not implicated in any steric conflicts. The crucial interactions disallowing this structure involve the second methyl group projecting into the center of the first ring. Rotation of this methyl group would only cause steric conflict between the methyl hydrogens and either H^{δ} or H^{β} of the first ring. It is clear that only an ω or ψ rotation will free the second methyl group from interacting with the first ring.

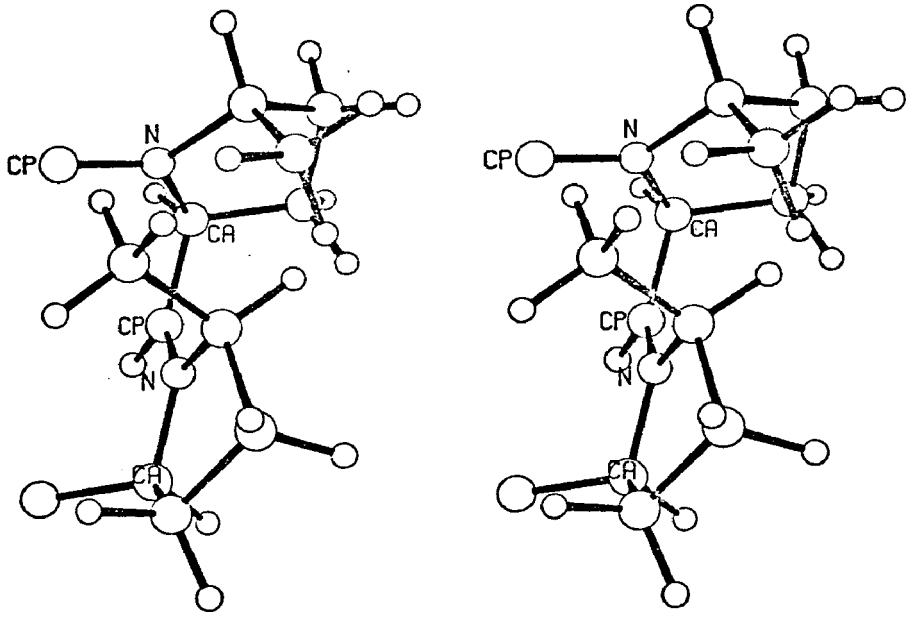


Figure 4-17. Cis (5-methyl pro)$_2$ cis'

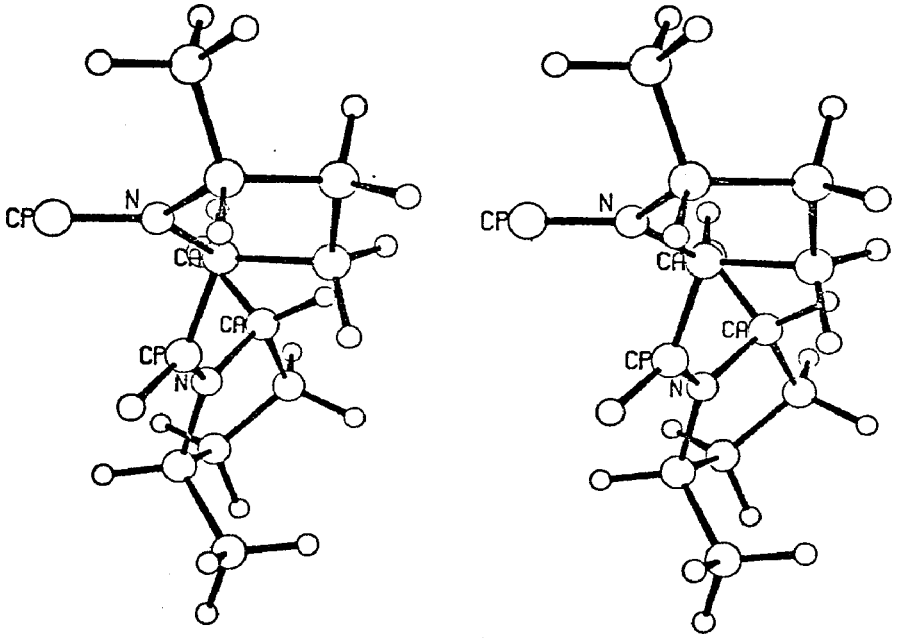


Figure 4-18. Trans (5-methyl pro)$_2$ cis

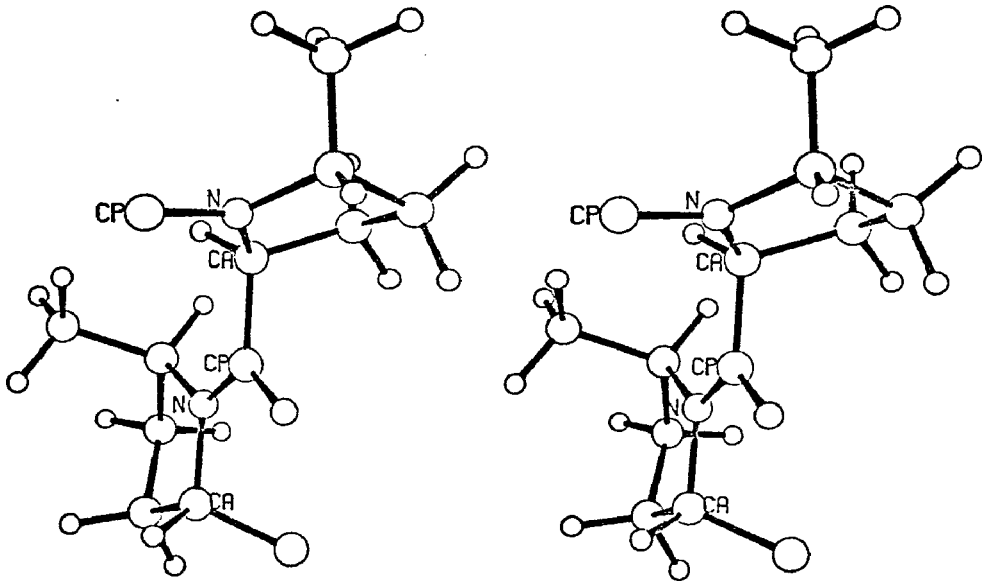


Figure 4-19. Trans (5-methyl pro)$_2$ trans

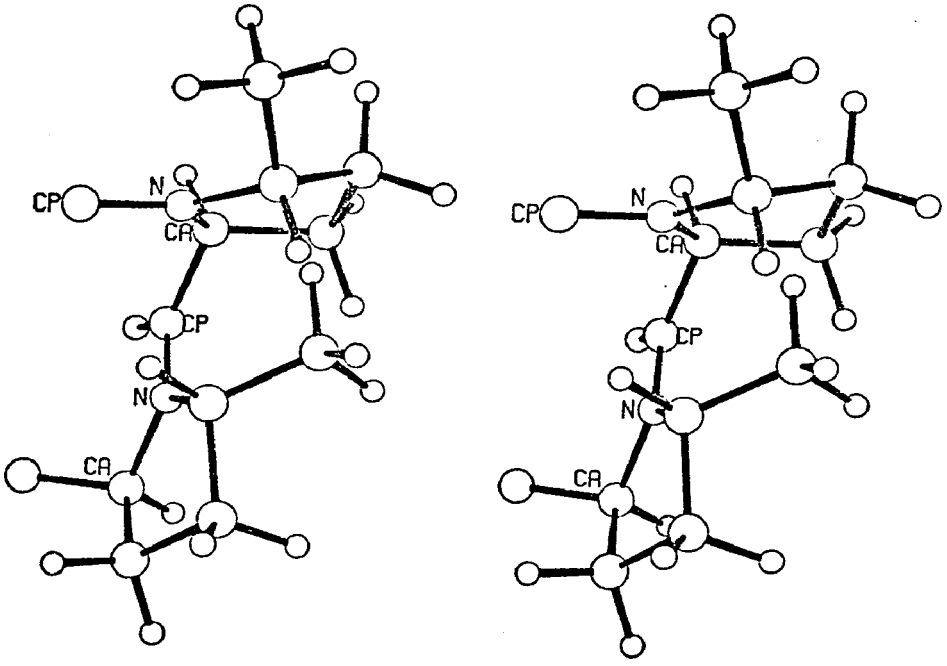


Figure 4-20. Trans (5-methyl pro)$_2$ cis'

Tables 4-6 and 4-7 contain the θ_1 and θ_2 values for each (ω,ψ) point in the regions for which the energy calculations of Figures 4-13 and 4-14 have been done. Using the notation of Section 3.7, Table 4-6 for cis 5-methylproline shows that Region I prefers (+,-) θ's, while (+,+) θ combinations are found in Region II. In the trans 5-methylproline case, there are about equal numbers of (+,-) and (-,-) θ combinations in Region I. In Region II, (-,+) and (+,+) combinations are found. This contrasts with the almost exclusive (-,-) θ's observed in the proline dimer for both Regions I and II.

4.7 Γ-θ Energy Plot for 2-Methylproline

Figure 4-2 shows the 2-methylproline monomer. A Γ-θ energy plot has been produced for the isolated ring. In this calculation, the methyl hydrogens have not been considered. Interactions involving methyl hydrogens are included, however, in the dimer calculations (this procedure facilitates both monomer and dimer calculations!). As in previous ring energy calculations, the bond angles $_{\tau}{}^{\alpha}$, $_{\tau}{}^{N}$ and $_{\tau}{}^{N'}$ were held constant at 103°, 113° and 121° respectively. Figure 4-21 is the Γ-θ plot for the 2-methylproline monomer. the M1 energy minimum (E=13.9 kcal/mole) is located at Γ=-32°, θ=-10°, while M2 (E=14.0 kcal/mole) occurs at Γ=+32°, θ=+10°. The general shape is quite similar to other Γ-θ plots which have been presented. The puckering characteristics of the isolated 2-methylproline

Table 4-6. Values of (θ_1,θ_2) for Each (ω,ψ) Corresponding to the Lowest Energy Rings for the cis 5-Methylproline Dimer with Independently Puckered Rings

ω \ Ψ	-180	-160	-140	-120	-100	-80	-60	-40	-20	0
180										
160	(-10, 25)									
140	(-20, 0)									
120	(-20, 0)									
100	(0,-15)									
80	(0,-25)									
60	(0,-25)									
40	(0,-20)									
20	(0,-10)									
0	(0, 0)									
-20	(-20,-10)									
-40	(-20, 5)	(-20,-15)	(0, 20)	(15, 25)					(20, 5)	
-60	(-20, 10)	(-20, 5)	(-25,-15)	(0,-20)	(10,-15)	(-10, 20)			(15, 25)	(10, 5)
-80	(-20, -5)	(-20, 5)	(-20,-10)	(10,-10)	(5, 0)	(15,-25)	(-5, 5)	(10, 20)	(-5, 25)	(-5,-25)
-100	(0,-10)	(0, -5)		(0, 5)	(10, -5)	(15, 0)	(5,-25)	(0, 15)	(-15, 25)	(-15,-10)
-120	(0,-10)	(0,-15)			(15, 0)	(25, 0)	(15, 0)	(0, 5)	(-20, 20)	(-25,-10)
-140	(0, 0)	(0, 5)				(25, 0)	(20, 0)			
-160	(10, 0)									
-180										

Table 4-6. (continued)

				(20, 0)	(0, 5)	(0, 5)	(0, 0)		180
				(25, 0)	(10, 0)	(0, 0)	(-20, 0)	(-10, 25)	160
					(-25, 0)	(0, 0)	(0, 0)	(-20, 0)	140
					(-20, 0)	(0, 0)	(0, 0)	(-20, 0)	120
					(15, 0)	(0, 5)	(0,-10)	(0,-15)	100
					(-25,-10)	(5, 0)	(0,-20)	(0,-25)	80
					(-25,-20)	(-20,-10)	(0,-20)	(0,-25)	60
				(-20,-25)	(-20,-25)	(0,-20)	(0,-15)	(0,-20)	40
				(20,-15)	(0,-20)	(0,-20)	(0,-20)	(0,-10)	20
				(25, 20)	(15,-10)	(0,-15)	(0,-15)	(0, 0)	0
					(25,-10)	(5, 0)	(0,-10)	(-20,-10)	-20
					(-25,-15)	(-20, 0)	(0, 10)	(-20, 5)	-40
(0,-20)	(-25,-25)				(-25, 5)	(0,-10)	(0, 10)	(-20, 10)	-60
(0, 5)	(-20, 0)	(-25,-20)	(-25,-25)	(-25, 25)	(-20, 15)	(0, 5)	(0, 10)	(-20, -5)	-80
	(-20, 5)	(-20,-15)	(-20,-15)	(-25,-10)	(-20, 10)	(0, 10)	(0,-10)	(0,-10)	-100
(-25, -5)	(-25, 0)	(-25, 0)	(-20,-10)	(-20,-10)	(0,-10)	(0, 10)	(0, 10)	(0, 10)	-120
			(-25, 0)	(0,-10)	(0,-10)	(0, 10)	(0, 5)	(0, 0)	-140
				(5, 10)	(0, 15)	(0, 10)	(0, 10)	(10, 0)	-160
				(20, 0)	(0, 5)	(0, 5)	(0, 0)		-180
20	40	60	80	100	120	140	160	180	

Ψ

ω

Table 4-7. Values of (θ_1,θ_2) for Each (ω,ψ) Corresponding to the Lowest Energy Rings for the trans 5-Methylproline Dimer with Independently Puckered Rings

ω \ Ψ	-60	-40	-20	0	20	40	60	80	100	120	140	160	180
180										(20,-10)	(-10, 0)	(-10, 0)	(0, 0)
160										(25, 10)	(0, 0)	(-10,-10)	(-10,-10)
140	(5,-15)	(10, 0)	(0, 0)	(-5, 0)	(-25, 0)					(-20, 0)	(-15, 0)	(-10,-10)	(-10,-10)
120	(5, 5)	(5,-20)	(-5,-25)	(-5,-25)	(-15, 0)	(-25, 0)	(-25, 5)			(-15,-10)	(-10, -5)	(-10,-15)	(-10, 10)
100		(5,-25)	(-25, 5)	(-25, 25)	(-20, 25)	(-20, 10)	(-20, -5)	(-25, 0)	(-25, 5)	(-15, 10)	(-10,-10)	(-10,-20)	
80		(-5,-25)	(-25,-25)			(-25, 25)	(-15, 10)	(-15,-10)	(-20,-10)	(-10, 5)	(-10,-10)	(-10,-20)	
60							(-25, 25)	(-5,-10)	(10,-20)	(-10,-20)	(-10,-15)	(-10,-15)	
40								(20,-10)	(-5,-20)	(-10,-25)	(-10,-20)	(-10,-10)	
20									(20,-15)	(-10,-20)	(-10,-20)	(5,-15)	(5,-15)
0									(25, 20)	(20,-10)	(-10,-15)	(10,-15)	(0, 0)
-20										(25,-10)	(-10,-15)	(-10,-10)	(-10,-10)
-40										(-10,-10)	(-10, -5)	(-10, 10)	(-10,-10)
-60										(-15, -5)	(-10,-10)	(-10, 10)	(-10,-15)
-80										(-15, 15)	(-5, -5)	(15,-10)	
-100							(-5, 5)	(-20, 5)	(-20,-15)	(-10, 15)	(-10, -5)	(0, 0)	
-120								(0, 0)	(-10, 10)	(-10, 5)	(-10,-10)	(-10, 0)	(0, 0)
-140								(25, 0)	(0, 0)	(-10, 0)	(-10, 0)	(-10, 0)	(-10,-10)
-160									(25, 0)	(0, 0)	(-10, 0)	(-10, 0)	(10, 0)
-180									(25,-15)	(20,-10)	(-10, 0)	(-10, 0)	(0, 0)

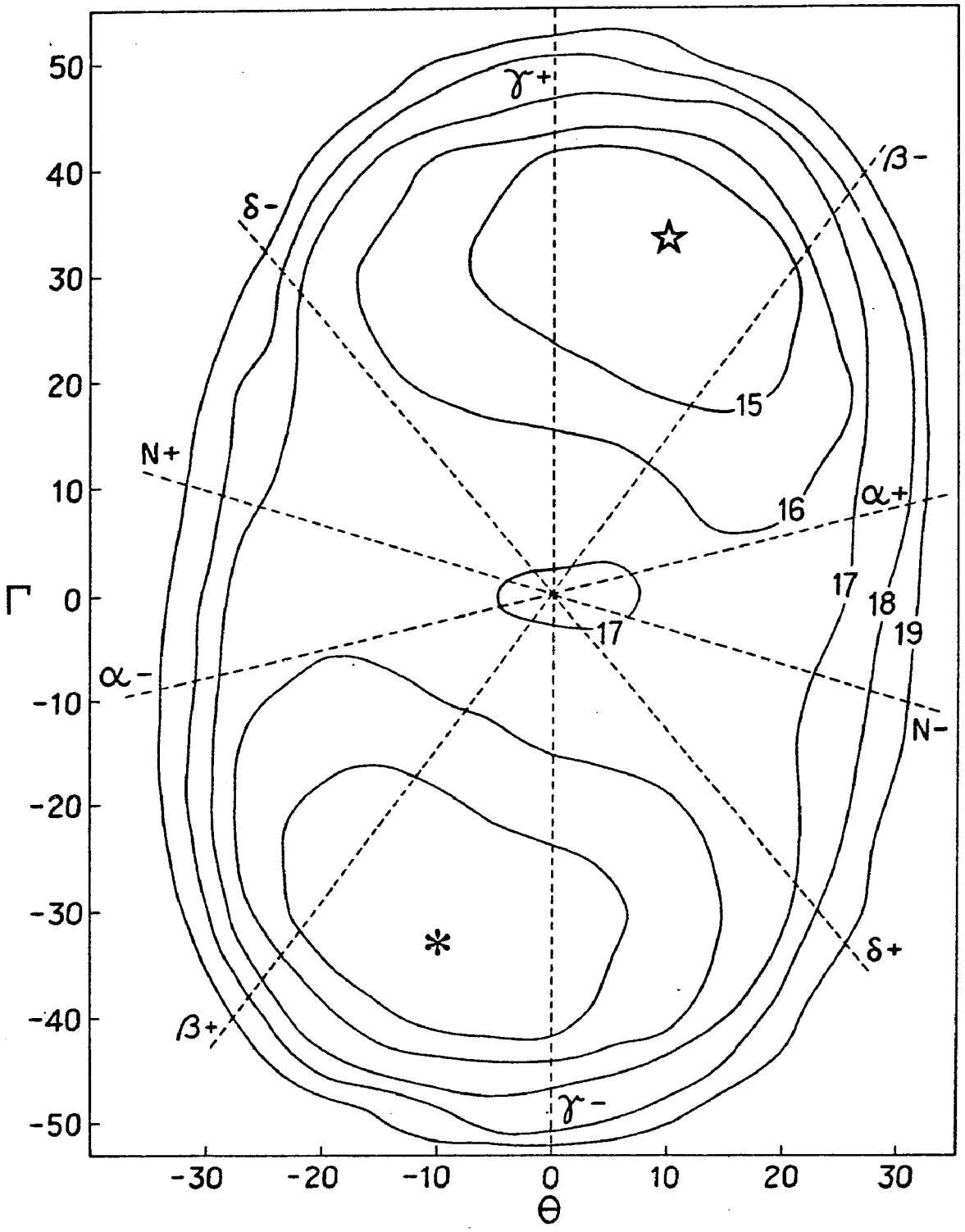


Figure 4-21. Γ-θ energy plot for the 2-methylproline isolated ring. $\tau^{\alpha}=103^{\circ}$, $\tau^{N}=113^{\circ}$ and $\tau^{N'}=121^{\circ}$.

ring are similar to those of the proline ring. Table 4-8 contains the Γ-θ energies from which Figure 4-21 was prepared.

4.8 (ω,ψ) Conformational Energy Map for the 2-Methylproline Dimer

Figure 4-22 shows the (ω,ψ) energy plot for the 2-methylproline dimer. In this calculation, methyl hydrogens are included; they are staggered with respect to the remaining substituents of C^{α} (i.e., $\chi[C^{\beta 1}\text{-}C^{\alpha}\text{-}C^{\beta 2}\text{-}H] = 60^{\circ}$). This energy map is striking in its complete absence of the cis region. In fact, the hatched area extends completely over the normal location of the cis minimum. As yet, this has not been seen in any dipeptide energy calculations. It is also interesting to note the unusual position of the trans' minimum. Though it is located at ω=180°, the ψ position of 200° is somewhat higher than that found in proline or the 5-methylprolines. Furthermore, the global energy minimum occurs in the cis' region, unlike the proline and 5-methylproline dimers, where it occurs in the trans' region. The barrier energy between these two regions is 23.0 kcal/mole, which is not excessive. It would seem, therefore, that by approximately 80° of ω rotation, the trans' $\rightarrow$ cis' transition could occur under suitable conditions.

In view of the fact that Region I does not exist along with the feature that the Region III minimum is lower in

Table 4-8. Γ-θ Energies for the 2-Methylproline Isolated Ring

Γ \ θ	-30	-25	-20	-15	-10	-5	0	5	10	15	20	25	30
60	36.7	32.4	29.3	27.1	25.6	24.7	24.4	24.8	26.0	28.3	32.2	39.0	51.1
50	28.8	24.4	22.0	20.2	18.9	18.1	17.7	17.7	18.2	19.3	21.3	24.4	29.3
40	22.9	20.1	18.2	16.8	15.9	15.2	14.8	14.6	14.7	15.3	16.5	18.5	21.6
30	20.3	18.2	16.8	15.8	15.2	14.7	14.4	14.1	14.0	14.2	14.9	16.2	18.4
20	19.0	17.4	16.5	16.1	15.8	15.6	15.4	15.1	14.9	14.8	15.0	15.8	17.4
10	18.4	17.1	16.6	16.5	16.6	16.7	16.6	16.4	16.1	15.8	15.8	16.3	17.6
0	17.8	16.7	16.4	16.5	16.7	17.0	17.1	17.1	16.8	16.6	16.5	16.9	18.1
-10	17.3	16.0	15.6	15.7	16.0	16.3	16.6	16.7	16.7	16.7	16.8	17.4	18.6
-20	17.2	15.5	14.8	14.6	14.8	15.1	15.4	15.7	15.9	16.2	16.7	17.7	19.3
-30	18.1	15.9	14.7	14.1	13.9	14.1	14.4	14.8	15.3	16.0	16.9	18.4	20.6
-40	21.3	18.2	16.3	15.2	14.7	14.6	14.8	15.2	15.9	17.0	18.4	20.3	23.1
-50	29.0	24.1	21.1	19.2	18.1	17.6	17.7	18.1	19.0	20.3	22.1	24.7	28.1
-60	50.5	38.7	32.0	28.1	25.9	24.7	24.4	24.7	25.7	27.2	29.5	32.7	36.9

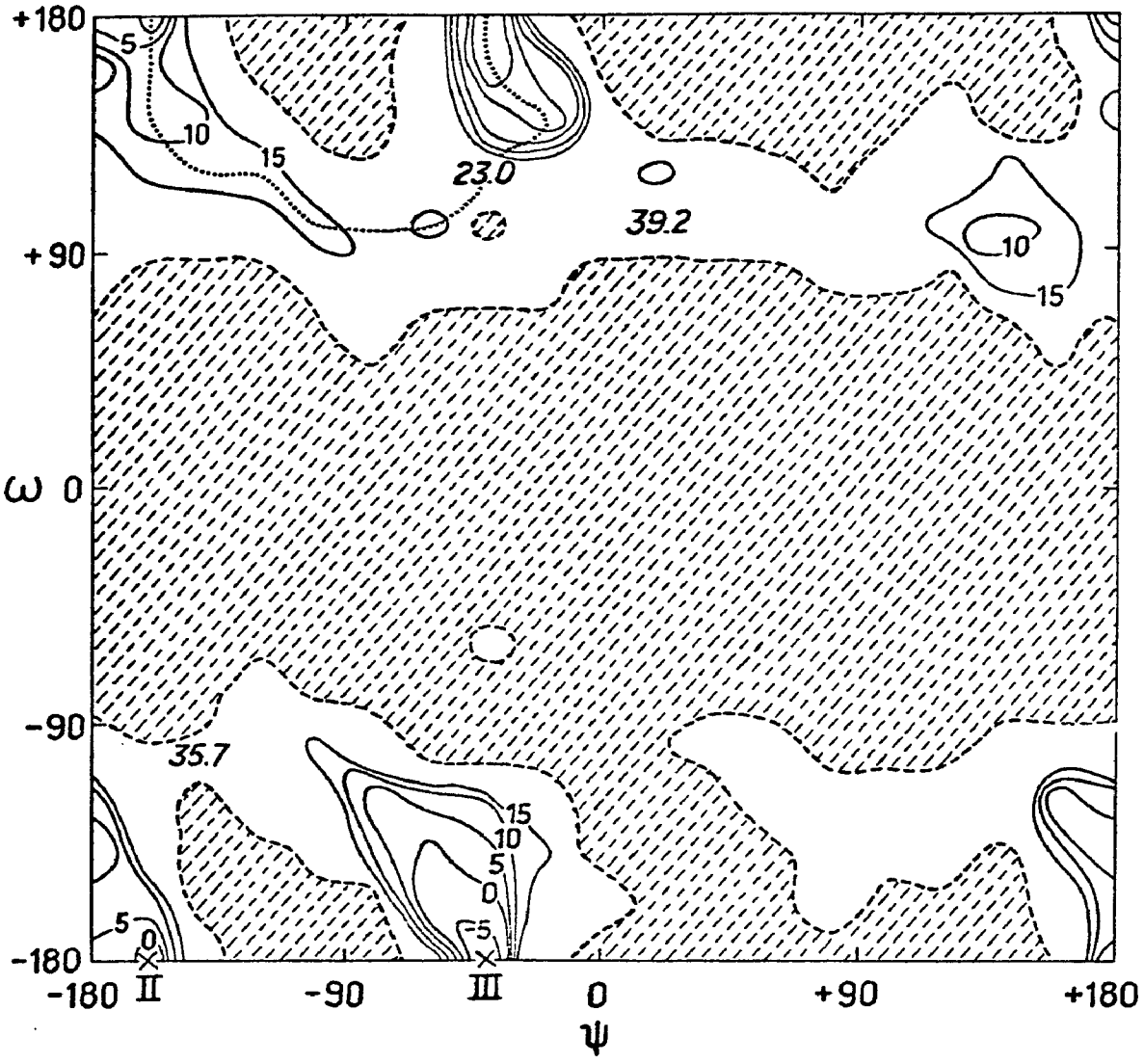


Figure 4-22. (ω,ψ) energy contour map for the 2-methylproline internal dipeptide.

energy than the Region II minimum, it is important to refine these calculations by way of allowing independent ring puckering and including methyl group rotation.

Due to computational expense, all refinement calculations were done only in regions around the trans' and cis' minima. The relevant parameters were varied over the following ranges:

$\{\omega: 175^\circ, 195^\circ, 5^\circ\}$
$\{\psi: -55^\circ, -35^\circ, 5^\circ\}$ for the dimer in the cis' region

$\{\omega: 165^\circ, 185^\circ, 5^\circ\}$
$\{\psi: 165^\circ, 205^\circ, 5^\circ\}$ for the dimer in the trans' region

$\{\chi_1, \chi_2: 60^\circ, 140^\circ, 40^\circ\}$ for the dimers in both regions

Three-fold torsional potentials for χ_1 and χ_2 were included (Equation 6, Appendix II) with V_1=3.4 kcal/mole.

Table 4-9 shows the results of these calculations. Each succeeding row in the table contains a refinement on the calculations from the previous row. Some interesting observations can be made regarding the effects of independent ring puckering and methyl rotation. In Region II, the minimum shifts slightly from that of Figure 4-22 due to the finer increments of ω and ψ. Independent ring puckering apparently has little effect on the energy of Region II. It is obvious that methyl group rotation should be able to relieve some of the steric conflicts which have probably accounted for the unusual location of the minimum. The results for Region II, shown in Row 3 of Table 4-9, bear

Table 4-9. The Torsional Angles (in degrees) and Energy Values (in kcal/mole) of Minimum Energy Conformations of the 2-Methylproline Internal Dipeptide

2-methylproline dimer	Region II							Region III						
	ω	ψ	θ_1	θ_2	χ_1	χ_2	E	ω	ψ	θ_1	θ_2	χ_1	χ_2	E
+ fixed methyl hydrogens	205,	185,	10,	10,	60,	60	-2.2	185,	-45,	10,	10,	60,	60	-9.1
+ independent rings	200,	190,	5,	15,	60,	60	-2.8	185,	-45,	10,	10,	60,	60	-9.1
+ independent methyl rotation	185,	170,	-5,	0,	140,	60	-5.2	180,	-40,	10,	10,	60,	60	-9.1

this out. Both ω and ψ are shifted downward by 15-20°. In addition, θ_1 and θ_2 are more negative, reflecting what was found in proline and trans 5-methylproline.

In Region III, independent ring puckering and methyl group rotation do not play any role in reducing the energy of the cis' minimum structure. In comparing cis' and trans' energies, there is a difference of 6.3 kcal/mole if methyl rotation is not allowed. Independent ring puckering does not affect this result. Under rotation this difference drops to 3.9 kcal/mole, though the cis' minimum still has the lower energy. Thus, it is concluded that methyl group rotation is quite an important factor in evaluating the energy differences between minimum energy structures.

The energy difference between the trans' and cis' structures could be quite significant in assigning a backbone conformational structure to the 2-methylproline polymer. Accordingly, we have made a study in which the three backbone bond angles were independently altered. $\tau[C^{\alpha}C\ N]$, $\tau[C\ NC^{\alpha}]$ and $\tau[NC^{\alpha}C\ ]$ were varied $\pm$ 4° from their equilibrium values of 119°, 121° and 111° respectively. Bond angle bending energies were taken into account according to Equation 8 of Appendix II. The K_{τ}'s used were 72 kcal/mole (78), 72 kcal/mole (78) and 101 kcal/mole (62) for the $\tau[C^{\alpha}C\ N]$, $\tau[C\ NC^{\alpha}]$ and $\tau[NC^{\alpha}C\ ]$ angles respectively. We have investigated the (ω,ψ) region surrounding the minimum energy points of Row 3, Table 4-9. In addition, θ's and χ's were fixed at their minimum energy

values from the same table. The results of the energy minimization analysis involving both the cis' and trans' regions are shown in Table 4-10. In the cis' region, only $\tau[C^{\alpha}C\ N]$ deviates from its equilibrium value, increasing from 119° to 123°. The conformational energy for this structure is lowered only 0.8 kcal/mole. In energy terms, a similar situation is found for the trans' structure where a 1 kcal/mole decrease is found. In this case, only one backbone bond angle changes; $\tau[NC^{\alpha}C\ ]$ contracts from 111° to 107°. There does not appear to be any significant change in the minimum energies for each region. More importantly, the energy difference between Regions II and III remains at nearly 4 kcal/mole. Thus, it is concluded that backbone bond angle variations of the kinds considered here are not especially meaningful insofar as changing relative energies between cis' and trans' minimum structures.

Figures 4-22 through 4-25 contain the stereoscopic views of the minimum energy structures for the cis, trans and cis' regions respectively. The coordinates for these drawings were taken from the structures of Table 4-9 for the trans' and cis' cases. Since, for the non-existant cis structure, all dimers give high energies, the coordinates corresponding to the minimum energy proline structure in that region were chosen. The trans structure (Figure 4-24) shows very little steric interference. The two methyl groups, while on the same side of both rings, point away from the rings as well as away from each other.

Table 4-10. The Torsional Angles (in degrees) and Energy Values (in kcal/mole) of Minimum Energy Conformations of the 2-Methylproline Internal Dipeptide Where Backbone Bond Angles are Varied

Region	ω	ψ	θ_1	θ_2	χ_1	χ_2	$\tau[C^{\alpha}C\ N]$	$\tau[C\ NC^{\alpha}]$	$\tau[NC^{\alpha}C\ ]$	Energy
cis'	180	-40	10	10	60	60	123	121	111	-9.9
trans'	180	175	-5	0	140	60	119	121	107	-6.2

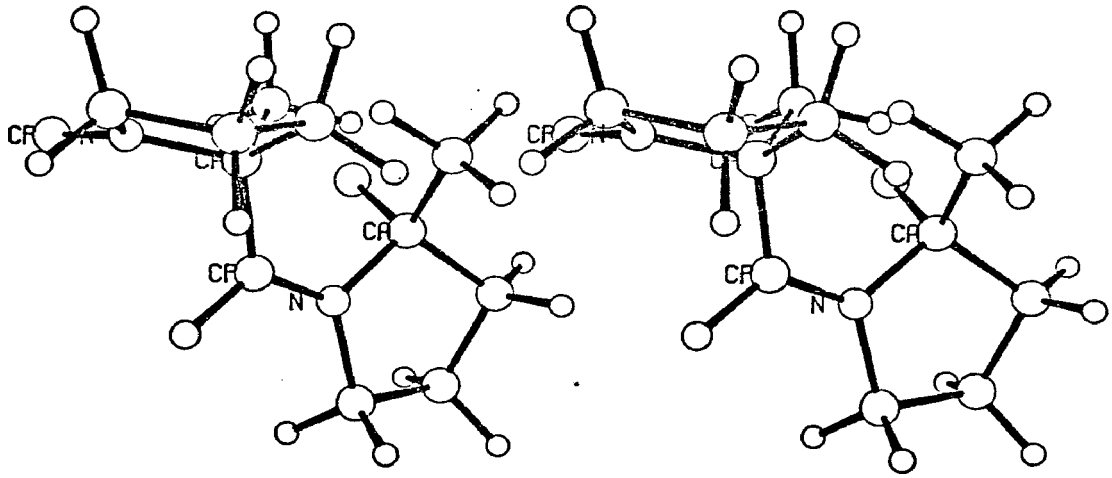


Figure 4-23. $(2\text{-methyl pro})_2$ <u>cis</u>

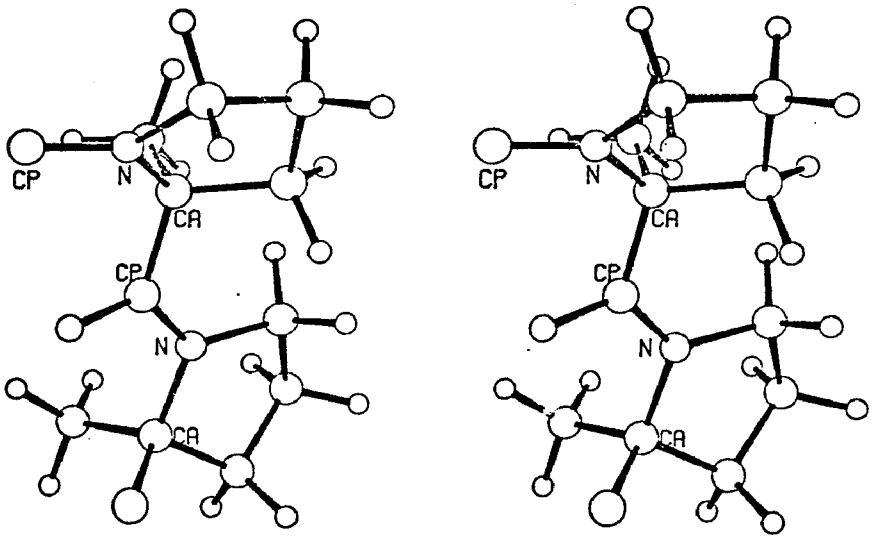


Figure 4-24. (2-methyl pro)$_2$ <u>trans</u>

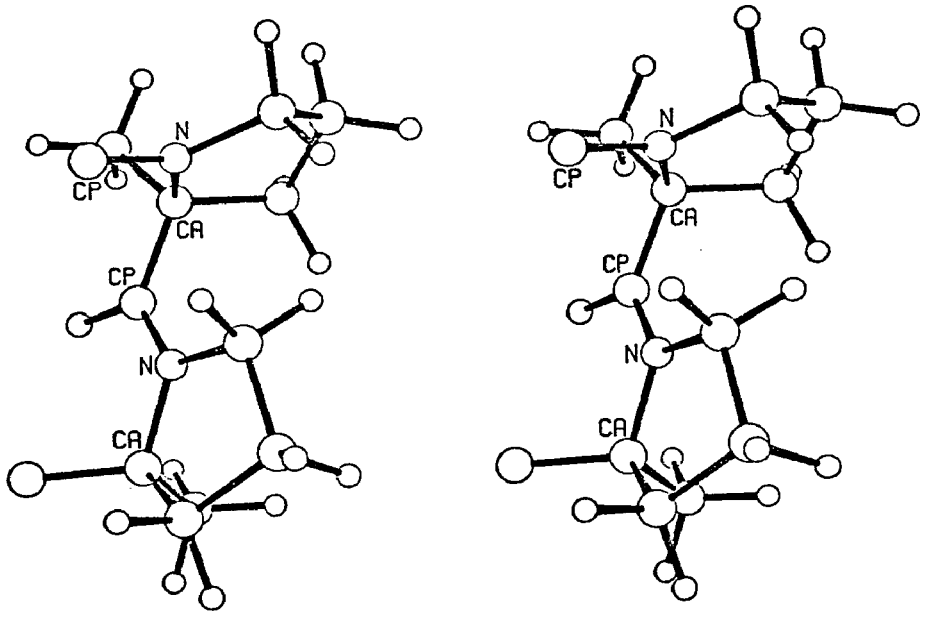


Figure 4-25. (2-methyl pro)$_2$ <u>cis'</u>

The puckering of the first ring is such that the H^{β} atom staggers the H^{δ}'s of ring 2. In the cis' structure (Figure 4-25) the rings are closer together, but not close enough to promote steric conflicts. The methyl groups are, again, well separated. On the other hand, the cis structure of Figure 4-23 quite plainly shows why this particular dimer does not exist. There are no ring-ring interactions, but the two methyl groups are in extremely close contact; so much so that rotation cannot alleviate the bad steric conflicts. In addition, there are possible close contacts between first methyl group hydrogens and the final C' atom. As Figure 4-23 indicates, ψ rotation would not help. More negative values of ψ would increase methyl-methyl interactions. Positive ψ rotation, while relieving these interactions, would bring ring 1 into direct contact with methyl group 2. It is clear that large ω rotations, either negative or positive, would effectively separate all the steric conflicts mentioned above. According to Figure 4-22, such large rotations bring the structure into the trans' region, which is energetically favorable.

4.9 Possible Implications of Structural Differences Among the Proline and Methylproline Dimers

As discussed in Chapter III, the three stable regions for the proline internal dipeptide should lead to three separate NMR resonances at room temperature. In an early study (53), the high resolution proton nuclear resonance

spectrum of poly(L-proline) in salt solutions showed two resonances and was interpreted as a result of a random distribution of a small number of cis imide bonds among the predominantly trans bonds along the chain. In further work on the cis - trans isomerization of poly(L-proline) in D_2O by the same authors, the resonance at 4.3-4.4 ppm was assigned to the cis configuration of the C^{α} protons. This assignment, however, was open to challenge, primarily due to the small size of the cis peak. A more recent highly sensitive Fourier transform proton NMR study by Wu et. al. (54) seems to have confirmed the presence of a small peak at 4.3 ppm and the same conclusion was made with regard to randomly distributed cis peptide bonds. As noted previously in this thesis, even a small percentage of cis residues in a long poly(L-proline) II chain could be quite influential on the chain dimensions. However, the possibility that the observed resonance at 4.3 belongs to the trans C^{α} protons occupying the cis' rotational state has not been completely discounted by either set of authors (50,54). Indeed, a small CR consistent with experimental observations can be accounted for by either hypothesis. It would appear, from our previous discussions of the 5-methylproline and 2-methylproline dimers, that an unambiguous determination of both the cis and cis' C^{α} resonances could be made. According to the calculations presented in this chapter, the 5-methylprolines can exist in either the cis or trans conformation, but not the cis' conformation.

On the other hand, the trans and cis' conformations are available to the 2-methylproline dimer, but not cis.

Thus, NMR analysis of these two methylated molecules should yield proper peak assignments for the cis and cis' C^{α} resonances. In fact, these molecules have been synthesized (63-65) in their polymeric form and their properties are currently being studied (66).

CHAPTER V

HYDROGEN BONDING IN THE 4-HYDROXYPROLINE INTERNAL DIPEPTIDE

5.1 Structure and Coordinates

Our understanding of the puckering of the pyrrolidine ring is quite useful in systems where puckering may influence polymer conformation in ways other than steric interaction. One such system is 4-hydroxyproline, a naturally occurring imino acid which has the capability of hydrogen bonding via the hydroxyl group. Figure 5-1 depicts the chemical structures of the 4-hydroxyproline isolated rings. There are two forms of this molecule, the cis type having the hydroxyl group on the same side of the ring plane as C_2, while in trans 4-hydroxyproline, the hydroxyl group is opposite to C_2. The possibility of hydroxyl group rotation is present. χ denotes this rotation, which is associated with the four atoms C^{δ}-C^{γ}-O-H.

Several conformational studies, both theoretical and experimental, have been reported on molecules involving the 4-hydroxyproline residue (67-73). Of the two theoretical studies (68,73), only Ooi et. al. (73) have considered ring flexibility in their calculations. However, only the $\gamma 1$ and $\gamma 2$ (see discussion in Section 2.6) rings were utilized. This is an artificial restriction on the

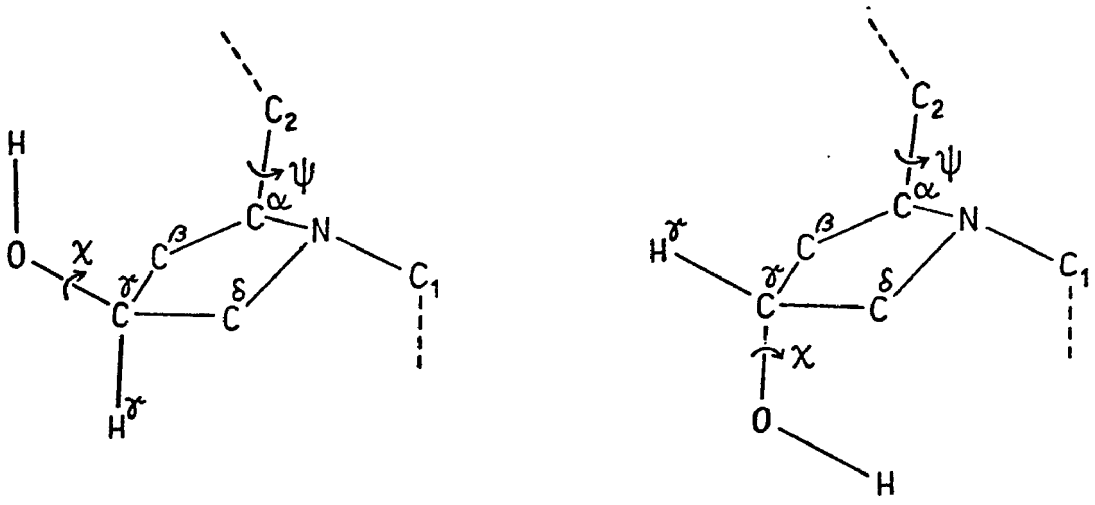


Figure 5-1. Two forms of the 4-hydroxyproline isolated ring; a) cis form, b) trans form.

kinds of puckering available to the ring. These authors also held ω fixed at its planar trans value of 180°. In these aspects, their study was, at best, incomplete.

The objective of this study on 4-hydroxyproline is an evaluation of the influence that various parameters, such as backbone torsional rotation, ring puckering characteristics and hydroxyl group positioning, have on hydrogen bonding within the dimer.

5.2 Γ-θ Plots for the 4-Hydroxyproline Isolated Rings

Figures 5-2 and 5-3 contain, respectively, the Γ-θ energy plots for the cis and trans 4-hydroxyproline isolated rings. As usual, $\tau^{\alpha}=103°$, $\tau^{N}=113°$ and $\tau^{N'}=121°$. The ring puckering parameters and χ were varied:

$$\{\Gamma: -60°, 60°, 10°\}$$
$$\{\theta: -40°, 40°, 10°\}$$
$$\{\chi: 60°, 300°, 120°\}$$

$\ell(C^{\gamma}\text{-}O)$ was taken to be 1.43Å, ℓ(O-H) was 0.96Å and the C^{γ}-O-H angle was fixed at 109°. The same bond lengths as found in proline were used for the ring as well as the backbone. For each (Γ,θ) point in Figures 5-2 and 5-3, the best value of χ giving the lowest ring energy is plotted. The value of χ at that point is indicated symbolically according to the legend in the figure. The global M1 minimum for the cis molecule (E=18.3 kcal/mole) occurs at $\Gamma=30°$, $\theta=10°$. The M2 minimum is at $\Gamma=-30°$, $\theta=-10°$, with an energy of 18.5 kcal/mole. In the trans

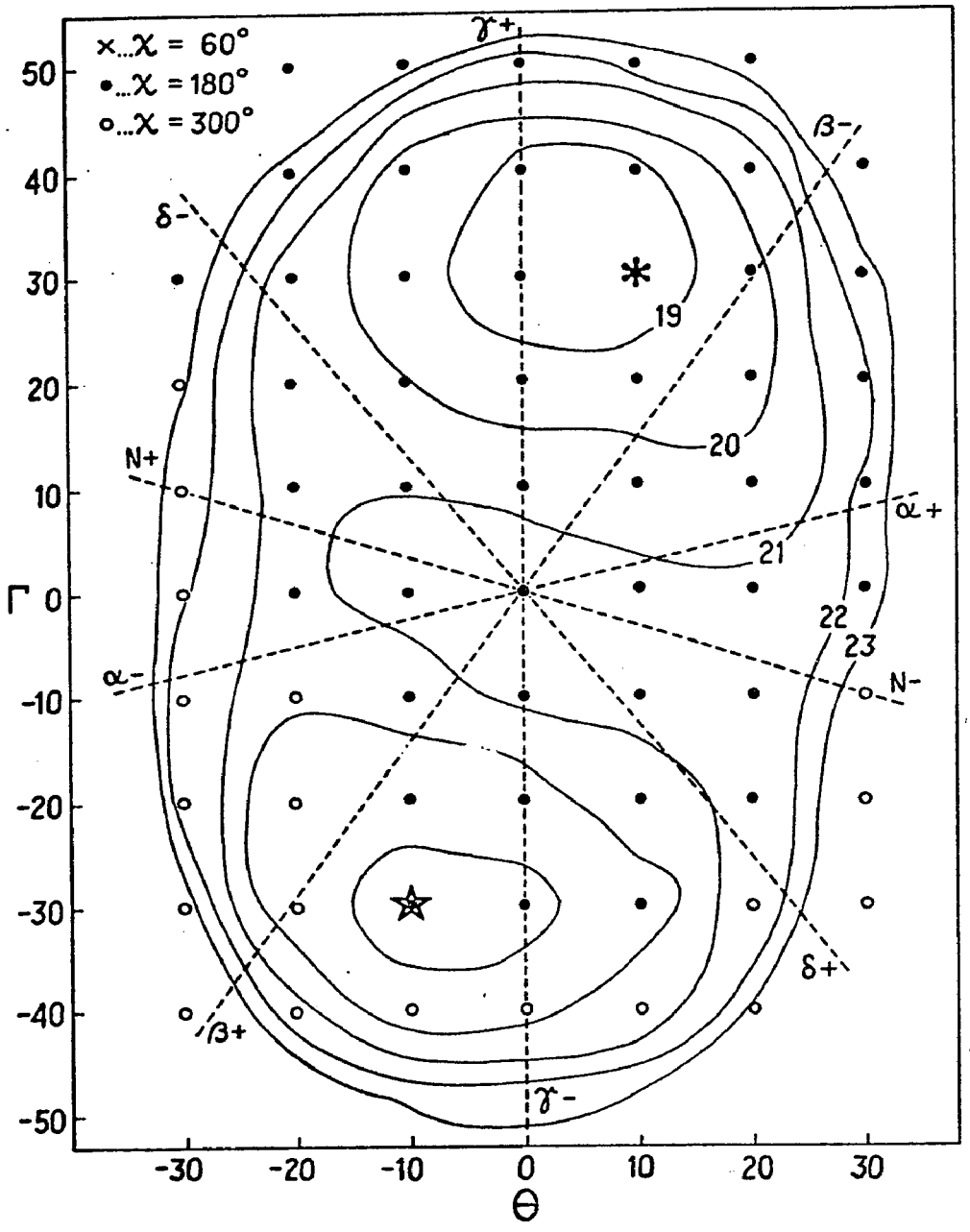


Figure 5-2. Γ-θ energy plot for the cis 4-hydroxyproline isolated ring. The best values of χ for each (Γ,θ) are indicated, according to the legend.

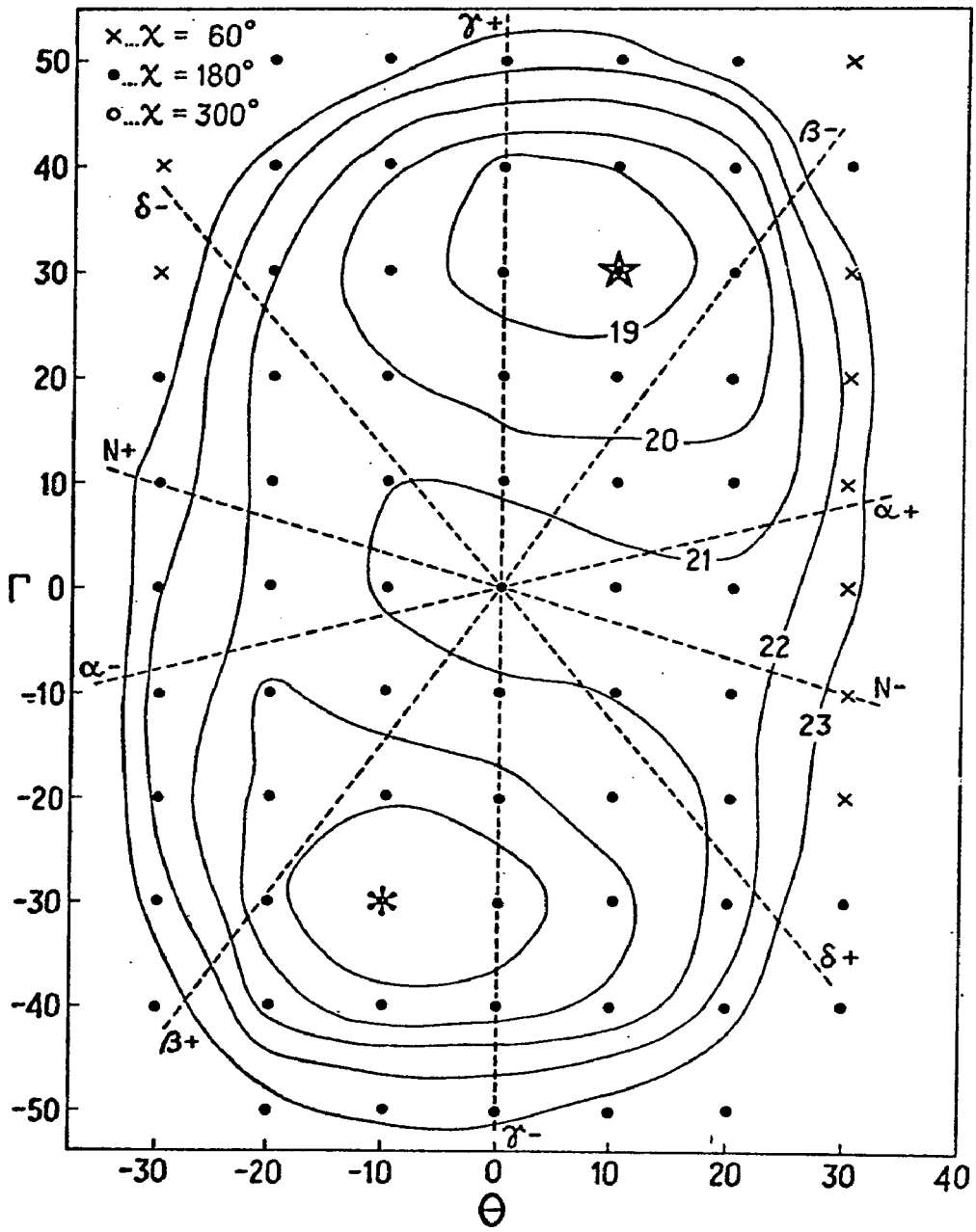


Figure 5-3. Γ-θ energy plot for the trans 4-hydroxyproline isolated ring. The best values of χ for each (Γ,θ) are indicated, according to the legend.

structure, the minima reverse their positions, with M1 (E=18.3 kcal/mole) occurring at $\Gamma=-30^{\circ}$, $\theta=-10^{\circ}$, while M2 (E=18.4 kcal/mole) is found at $\Gamma=30^{\circ}$, $\theta=10^{\circ}$. In both the cis and trans cases, $\chi=180^{\circ}$ is found for the minimum energy structures as well as the regions surrounding these minima. The molecular environment of the hydroxyl group is similar in both molecules; viz., it points away from the interior of the ring, but towards the H^{β} atom on the same side of the ring. For the cis molecule, $\chi=300^{\circ}$ is found for $-30^{\circ} \leqslant \Gamma \leqslant 10^{\circ}$ and extreme θ values. For $\Gamma \leqslant -30^{\circ}$ all θ's have $\chi=300^{\circ}$. In the trans molecule, $\chi=60^{\circ}$ almost exclusively for $\theta=30^{\circ}$. Apparently, for extreme ring puckering, the hydroxyl group prefers to point into the ring, staggering the H^{β} and H^{δ} atoms on its side of the ring, but in the direction of H^{β}.

The puckering characteristics for the isolated 4-hydroxyproline rings are analogous to those of the 5-methylproline rings. In both types of molecules, the M1 positions occupy the same regions of the Γ-θ map, shifting with respect to the cis or trans placement of the methyl (or hydroxyl) group. The difference in energy between the global minimum and completely planar geometries is 3.2 kcal/mole for both ring types. The Γ-θ energies and value of χ corresponding to each energy are found in Tables 5-1 and 5-2 for both ring types.

Table 5-1. Γ-θ Energies and Value of χ for Each Energy in the cis 4-hydroxyproline isolated ring

Γ \ θ	-30	-20	-10	0	10	20	30
60	40.9	33.5	29.7	28.4	29.4	33.5	43.2
	180	180	180	180	180	180	180
50	32.0	26.9	23.0	21.8	22.3	24.9	31.3
	180	180	180	180	180	180	180
40	27.0	22.3	20.0	18.9	19.0	20.7	25.5
	180	180	180	180	180	180	180
30	24.5	20.9	19.4	18.6	18.3	19.2	22.8
	180	180	180	180	180	180	180
20	23.4	20.8	20.0	19.6	19.2	19.4	22.0
	300	180	180	180	180	180	180
10	22.8	21.0	20.9	20.9	20.5	20.3	22.2
	300	180	180	180	180	180	180
0	22.3	20.9	21.2	21.5	21.3	21.1	22.8
	300	180	180	180	180	180	180
-10	21.9	20.2	20.5	21.1	21.3	21.4	23.5
	300	300	180	180	180	180	300
-20	21.8	19.4	19.3	19.9	20.4	21.4	24.1
	300	300	180	180	180	180	300
-30	23.2	19.4	18.5	18.9	19.8	21.6	25.3
	300	300	180	180	180	300	300
-40	28.6	21.6	19.4	19.4	20.5	23.1	28.0
	300	300	300	300	300	300	300
-50	51.1	29.9	23.9	22.8	23.9	27.2	33.7
	300	300	300	300	300	300	300
-60	99.0	63.8	37.8	31.8	32.2	36.4	45.6
	300	300	300	300	300	300	300

Table 5-2. Γ-θ Energies and Value of χ for Each Energy in the trans 4-hydroxyproline isolated ring

Γ	θ						
60	44.7	35.5	31.1	29.7	31.5	38.6	60.1
	180	180	180	180	180	60	60
50	33.1	26.6	23.3	22.1	22.6	25.8	34.0
	180	180	180	180	180	180	60
40	27.4	22.5	20.0	19.0	19.0	20.9	26.0
	60	180	180	180	180	180	180
30	24.8	21.0	19.4	18.6	18.4	19.3	23.0
	60	180	180	180	180	180	60
20	23.6	20.9	20.1	19.7	19.3	19.5	22.1
	180	180	180	180	180	180	60
10	22.9	21.0	20.9	21.0	20.5	20.4	22.3
	180	180	180	180	180	180	60
0	22.3	20.7	21.1	21.5	21.3	21.2	22.8
	180	180	180	180	180	180	60
-10	21.7	20.0	20.3	20.9	21.1	21.3	23.3
	180	180	180	180	180	180	60
-20	21.6	19.1	19.1	19.7	20.3	21.2	24.0
	180	180	180	180	180	180	60
-30	22.6	19.1	18.3	18.7	19.6	21.3	25.1
	180	180	180	180	180	180	180
-40	25.8	20.8	19.1	19.1	20.3	22.8	27.7
	180	180	180	180	180	180	180
-50	33.7	25.7	22.6	22.2	23.5	26.6	32.7
	180	180	180	180	180	180	180
-60	55.3	36.7	30.6	29.1	30.3	34.2	41.7
	300	180	180	180	180	180	180
	-30	-20	-10	0	10	20	30

θ

5.3 Hydrogen Bonding in the 4-Hydroxyproline Dimer

A schematized drawing of the 4-hydroxyproline internal dipeptide is shown in Figure 5-4. The carbonyl oxygens are indicated by the double lines connecting them to the chain backbone. Additionally, the backbone torsion angles and χ's are indicated. There are four types of simple hydrogen bonding (H-bonding) which are possible with this dimer. They are shown as double arrowed lines and are numbered. $O_1 \longleftrightarrow (OH)_1$ is Type 1 hydrogen bonding, $O_2 \longleftrightarrow (OH)_1$ is Type 2 bonding, $O_1 \longleftrightarrow (OH)_2$ is Type 3 bonding and $(OH)_1 \longleftrightarrow (OH)_2$ is Type 4 bonding. Note that $O_2 \longleftrightarrow (OH)_2$ is merely a Type 1 bonding and is not indicated. Since there are two positions the hydroxyl group can assume in each ring, there are four possible dimers that can be formed; cis-cis (CC), cis-trans (CT), trans-cis (TC) and trans-trans (TT). It is the TT dimer that is of primary interest because the trans hydroxyl is what occurs in collagen. The three other dimer types are important since they are illustrative of pyrrolidine ring properties and how they relate to hydrogen bonding conformations within the 4-hydroxyproline dimer.

As seen from Figure 5-4, the existence of different H-bonding types depends upon molecular conformation as expressed by the backbone torsion angles in addition to the Γ and χ values. This functional dependence for the four H-bonding types can be expressed as follows:

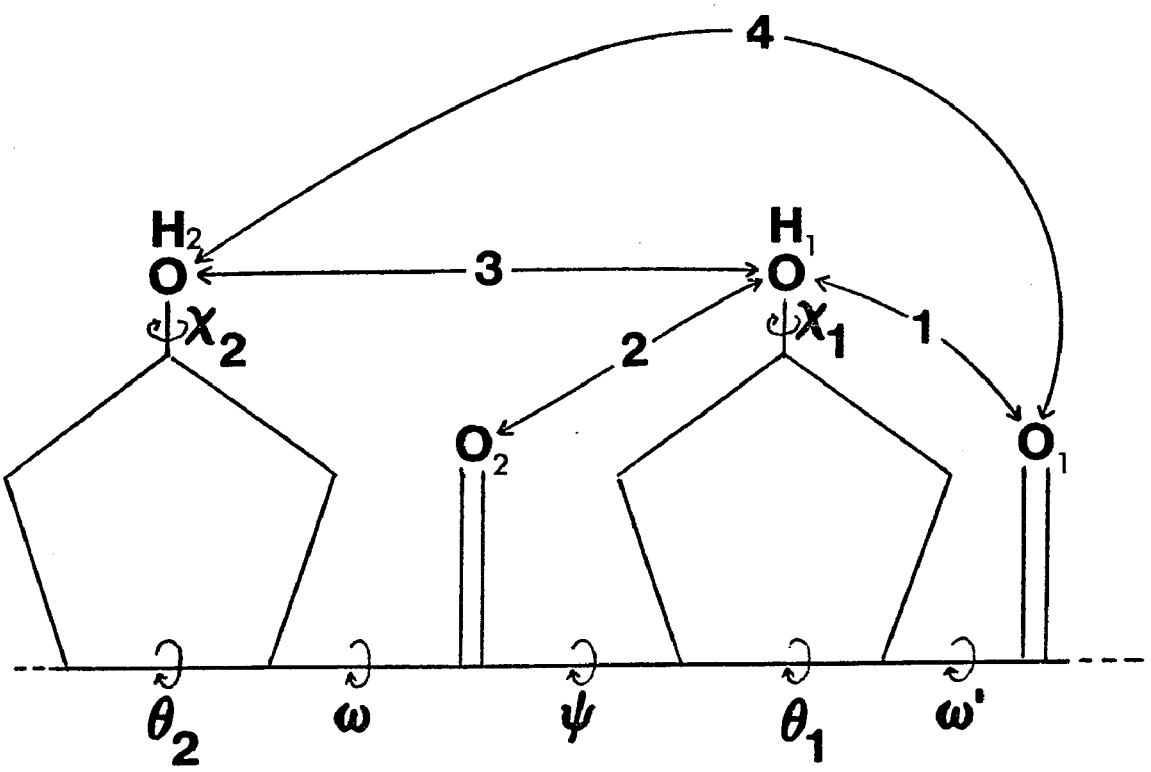


Figure 5-4. Schematic drawing of the 4-hydroxyproline dimer showing the four possible types of hydrogen bonding.

Type 1 is a function of $[\omega', \theta_1, \Gamma_1, \chi_1, p_1]$

Type 2 is a function of $[\psi, \theta_1, \Gamma_1, \chi_1, p_1]$

Type 3 is a function of $[\omega', \theta_1, \psi, \omega, \theta_2, \Gamma_2, \chi_2, p_2]$

Type 4 is a function of $[\theta_1, \Gamma_1, \chi_1, p_1, \psi, \omega, \theta_2, \Gamma_2, \chi_2, p_2]$

where p denotes the hydroxyl configuration, either cis or trans. It is obvious, especially for Types 3 and 4 H-bonding that exceedingly large numbers of conformations would have to be studied in order to completely characterize H-bonding in the dimer. In particular, if parameter ranges used throughout this thesis were applied to Type 4 H-bonding, about 10^{10} conformations would have to be considered. We therefore have performed the analysis of possible H-bonding conformations in a two step procedure:

1. First, we isolated those regions of conformational space which satisfy steric criteria for H-bonding; namely, that the O...O distance be within a certain range conducive to H-bonding and that the H-O...O angle be restricted to within known limits.
2. We further reduced the total number of conformations by additionally requiring that the total self-energy of the ring puckering type(s) needed for a particular H-bond does not exceed the maximum amount of energy gained by hydrogen bond formation.

After having applied these restrictive criteria, the total number of hydrogen bonding conformations will be greatly reduced. We can then more easily understand which geometrical parameters have the most influence on hydrogen bond formation in the 4-hydroxyproline dimer.

5.4 Distance and Angle Restrictions as Criteria for Hydrogen Bonding

As already seen, an excessive number of conformations available to the dimer necessitates a drastic reduction of parameter variation. There are straightforward methods for accomplishing this. It is not fruitful to carry out calculations in those regions of conformational space where the energy of the dimer is known to be large; for example where $-20^{\circ} < \psi < 100^{\circ}$. One can also predict where regions of maximum interest lie on the basis of measurement with accurately scaled molecular models. Using these two methods as a basis, the range chosen for ψ was restricted to one cis' and four trans' values:

$$\psi: -45^{\circ}, 160^{\circ}, 170^{\circ}, 180^{\circ}, 190^{\circ}$$

Both ω and ω' were restricted to $\pm 10^{\circ}$ from their equilibrium cis and trans values:

$$\omega,\omega': -10^{\circ}, 0^{\circ}, 10^{\circ}, 170^{\circ}, 180^{\circ}, 190^{\circ}$$

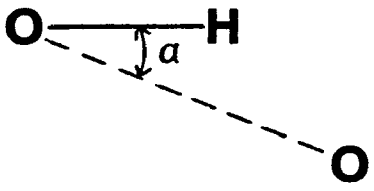
The θ's for both rings were independently varied:

$$\{\theta_1,\theta_2: -30^{\circ}, 30^{\circ}, 10^{\circ}\}$$

The Γ's were also independently varied:

$$\{\Gamma_1,\Gamma_2: -60^\circ, 60^\circ, 10^\circ\}$$

Following Chidambaram *et. al.* (74), only O...O distances between 2.6Å and 3.1Å were considered. In addition, the angle α (H-O...O, see below) was limited to the range 0°-40°.



Employing these limits, the χ angles of the two rings were varied independently:

$$\{\chi_1,\chi_2: 60^\circ, 300^\circ, 120^\circ\}$$

The results for each type of H-bonding are discussed below.

TYPE 1 HYDROGEN BONDING

No Type 1 hydrogen bonding can exist in the 4-hydroxyproline internal dipeptide. The O...O distance criterion alone rules this H-bonding type out.

TYPE 2 HYDROGEN BONDING

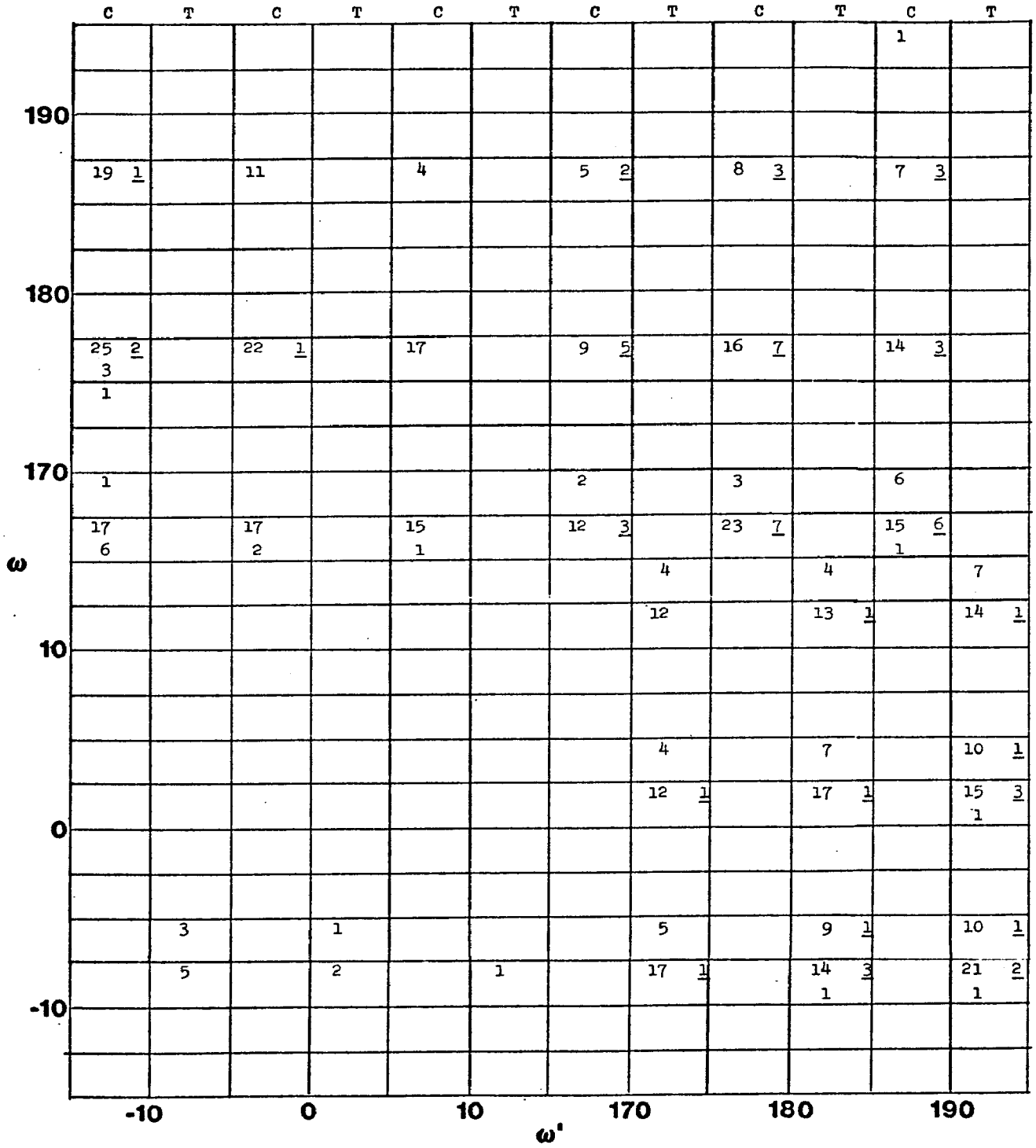
Type 2 hydrogen bonding does exist. In order to satisfy the steric H-bonding criteria, $(OH)_1$ must be in the cis configuration and χ_1 must be 60°. The criteria are satisfied by all values of ψ. However, for $\psi=-45^\circ$, there is only one possible ring puckering conducive to H-bonding and that occurs for $\Gamma_1=-60^\circ$, $\theta_1=-30^\circ$. For this (Γ,θ) the ring is very strained and hence, it is unlikely

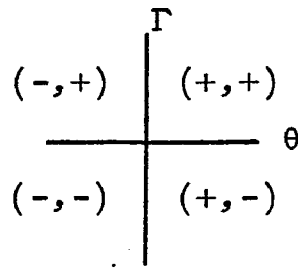
that Type 2 H-bonding will exist for ψ in the cis' rotational state. For ψ in the trans' region, Γ_1 must be less than or equal to -20°, but θ_1 can assume all values. In all cases, α ranges between 4° and 23° for trans' ψ distances. In total, 38 conformations out of a possible 2730 have survived. A priori, this seems a relatively small percentage of survival (1.4%) but it is in fact quite large when compared to Types 3 and 4 H-bonding.

TYPE 3 HYDROGEN BONDING

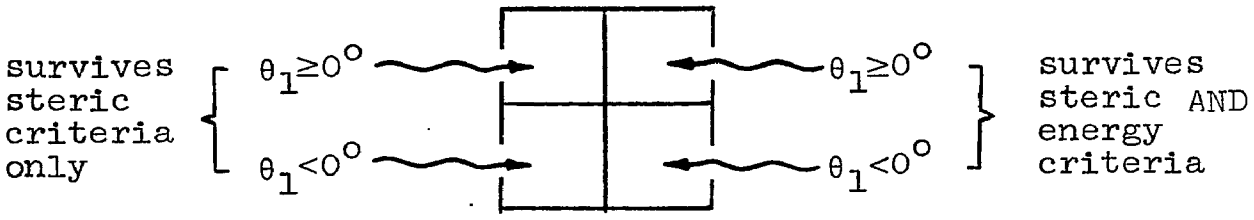
After application of the distance and angle criteria, only 456 conformations are left. Table 5-3 shows the values of the parameters for these surviving conformations. Type 3 hydrogen bonding exists only in the cis' ($\psi \sim -45^\circ$) region. The abscissa and the ordinate of Table 5-3 contain the ω' and ω values respectively. For each (ω',ω) combination, the $(OH)_2$ can be either cis or trans. This is marked as C or T along the top of the table. That leaves three variables, namely θ_1, θ_2 and Γ_2 to be accounted for. For each (ω,ω') and dimer type (C or T), there are four boxes stacked vertically. The position of each box denotes one of the four possible signs of the (θ_2,Γ_2) pair, i.e., the first box (from the top) is (+,+), the second box is (-,+), the third is (-,-) and the fourth is (+,-). This sign designation corresponds to the figure below which shows a θ,Γ coordinate system. The signs for each box are obtained from the signs of each quadrant, reading counterclockwise.

Table 5-3. Type 3 Hydrogen Bonding Conformations in the 4-Hydroxyproline Internal Dipeptide





Each box can contain at most four numbers. The top two numbers denote the number of survivors of a particular H-bonding conformation with $\theta_1 \geq 0^\circ$, the bottom numbers are for $\theta_1 < 0^\circ$. This is illustrated in the diagram of a single box given below:



Let us for the present consider only the left-adjusted numbers in each box. The right-adjusted underlined numbers will be discussed later. Thus, for example, for the cis $(OH)_2$ position with $\omega = 180^\circ$ and $\omega' = -10^\circ$, there are 25 survivors in the $(\theta_2, \Gamma_2) = (+,-)$ box with $\theta_1 \geq 0^\circ$ while only 3 conformations are found with $\theta_1 \leq 0^\circ$. Presentation of data in this manner enables one to quickly see trends in the data even though six independent variables are involved.

From energy considerations in the very similar Pro-Pro internal dipeptide, all data for ω in the cis region can be neglected. It is already known that the cis'·cis ($\psi \sim -45^\circ$, $\omega \sim 0^\circ$) structure cannot exist. Favorable hydrogen bonding conditions are, of course, meaningless

in such high energy regions of the (ω,ψ) map. Concentrating on the trans ω region (upper portion of Table 5-3), the great majority of conformations are found in the (+,-) box, indicating that $\theta_2 \geq 0^o$ and $\Gamma_2 < 0^o$ for Type 3 H-bonding. Also, all conformations are for the hydroxyl group in the cis position. In addition, almost all the surviving conformations prefer $\theta_1 \geq 0^o$. Analysis of the data shows that it is $\theta_1 \geq 20^o$ which is actually preferred.

TYPE 4 HYDROGEN BONDING

The situation for Type 4 H-bonding is somewhat more complicated to assess because various values of ψ must be represented and both (θ_1,Γ_1) and (θ_2,Γ_2) are involved. In this case, for each ψ value, a separate table is needed. Tables 5-4 through 5-7 contain the results for Type 4 hydrogen bonding. The abscissa contains the dimer type information, whether CC, CT, TC or TT. The ordinate contains the ω value. For each (ω, dimer type) position there is a 4 x 4 array of boxes. The horizontal direction, from left to right, contains the (θ_1,Γ_1) information exactly as described for Type 3 H-bonding, while the vertical direction contains the (θ_2,Γ_2) information.

There is an additional consideration which must be taken into account in discussing Type 4 H-bonding. There are two possible methods by which a hydrogen bond may be formed between $(OH)_1$ and $(OH)_2$. These are designated by the direction of proton donation, as shown in Figure 5-5. 'Forward' H-bonding is where H_1 donates a proton to O_2

Table 5-4. Type 4 Hydrogen Bonding Conformations in the 4-Hydroxyproline Internal Dipeptide for $\psi=-45^{\circ}$

ω	CC				CT			
					27 4 22 3			
190	9 1 5 1			41 13 22 5	88 30 53 14	9 2		
	46	93 1			90 5 5	7 1	91	41 1
	6	11 3 7 5			46 10 17 7	58 21 71 21	10 1	7
						5 4		
180		7 5			23 9 10 5	81 27 57 19	7	
	65	105 2		2	89 6 7 1	17 17 6 4	103	63 1
	9	8 1 4 3			34 3 2 1	84 7 48 16	9 9	
170		2 2			6 3	54 13 27 4	2	
	106	115 1	2	2	72 5	27 11 7 3	106	100 2
	20	9		10	27 1	80 4	6	18
10			11 11	21 8				
			15 3 75 14	37 43 9				
			16 3 57 23	45 1 32 14				
0			3 3	4 2				
			12 3 61 12	40 23 9	1	1		
			11 35 12	31 17 8				
-10								
			6 1 48 10	37 22 3	1	1		
			6 15 4	13 2 1				

Table 5-4. (continued)

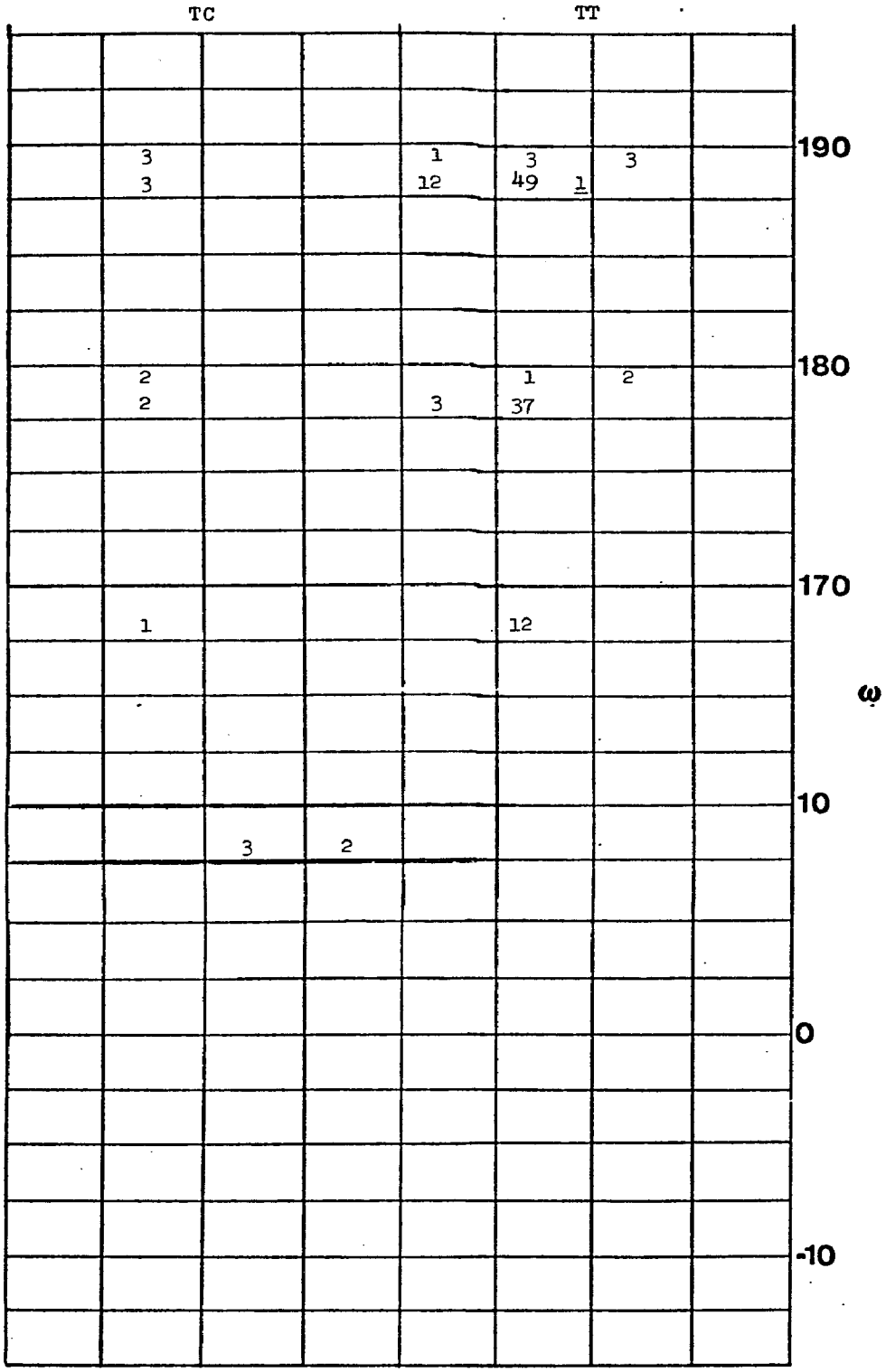
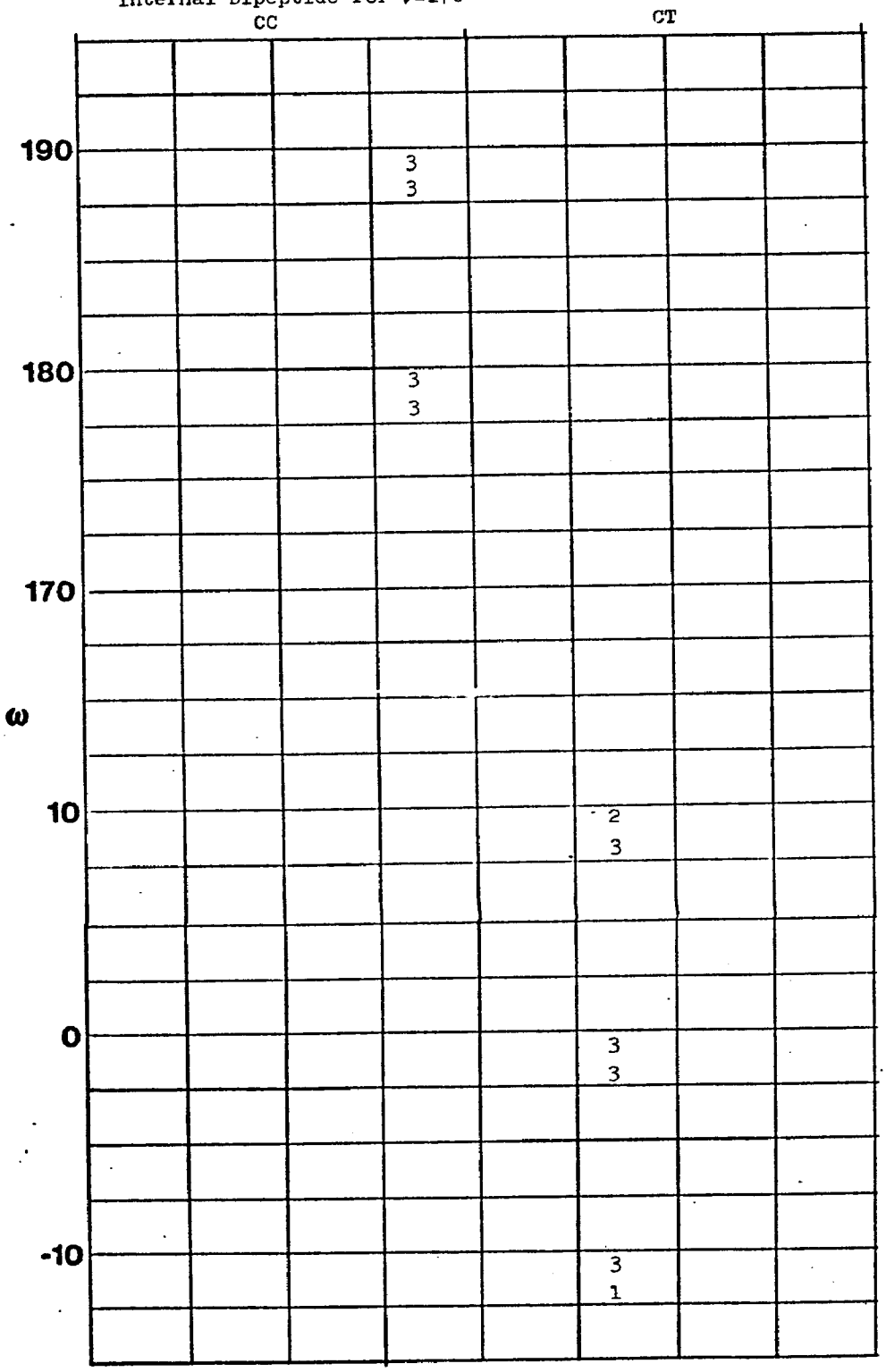


Table 5-5. Type 4 Hydrogen Bonding Conformations in the 4-Hydroxyproline Internal Dipeptide for $\psi=170^{\circ}$ *



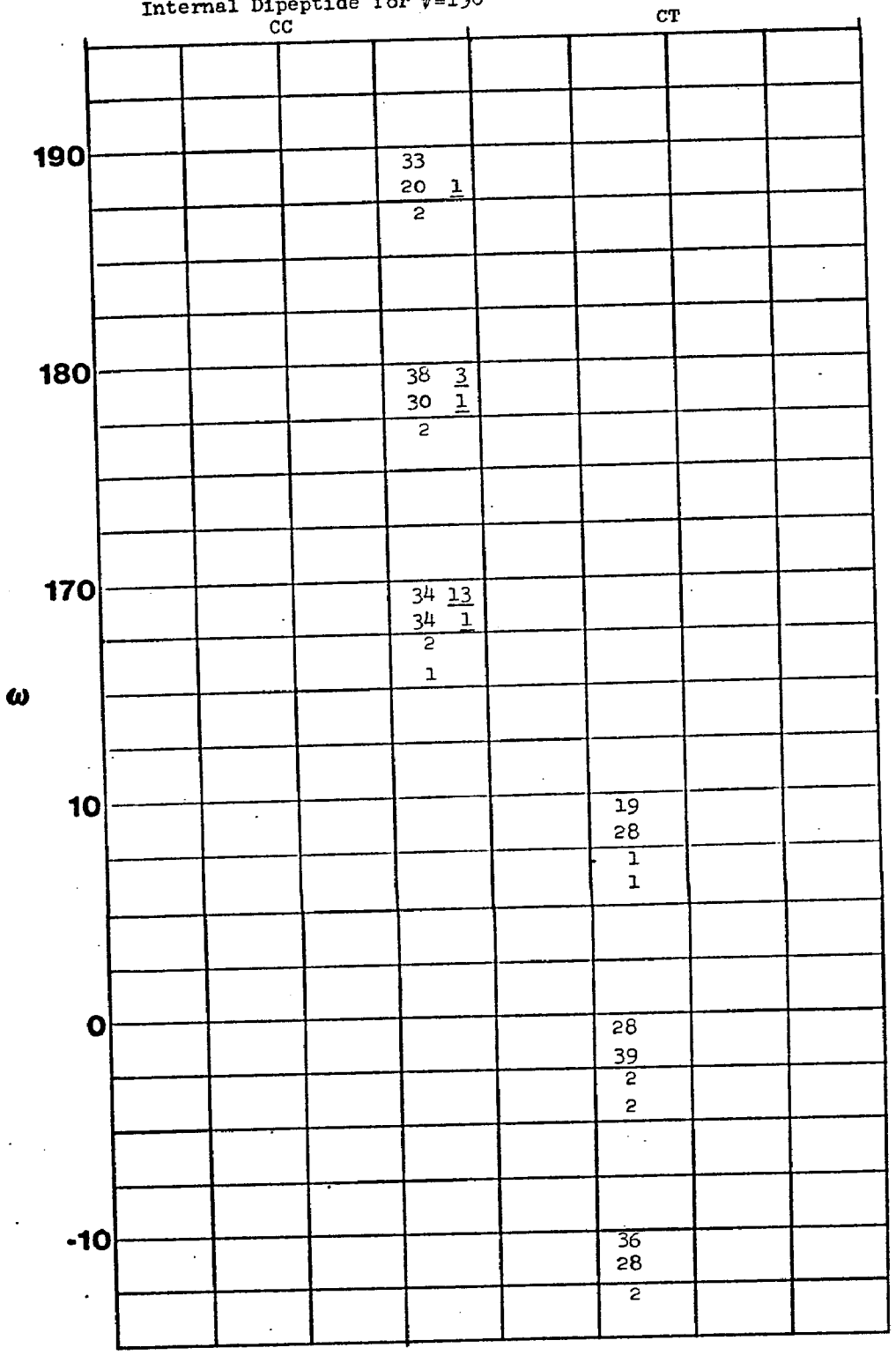
*No survivors are found for the TC or TT dimer.

Table 5-6. Type 4 Hydrogen Bonding Conformations in the 4-Hydroxyproline Internal Dipeptide for $\psi=180^{\circ}$*

ω	CC	CT
190	15 11	
180	14 13	
170	9 10	
10		11 13
0		13 16
-10		14 5

*No survivors are found for the TC or TT dimer.

Table 5-7. Type 4-Hydrogen Bonding Conformations in the 4-Hydroxyproline Internal Dipeptide for $\psi=190^{\circ}$ *



*No survivors are found for the TC or TT dimer.

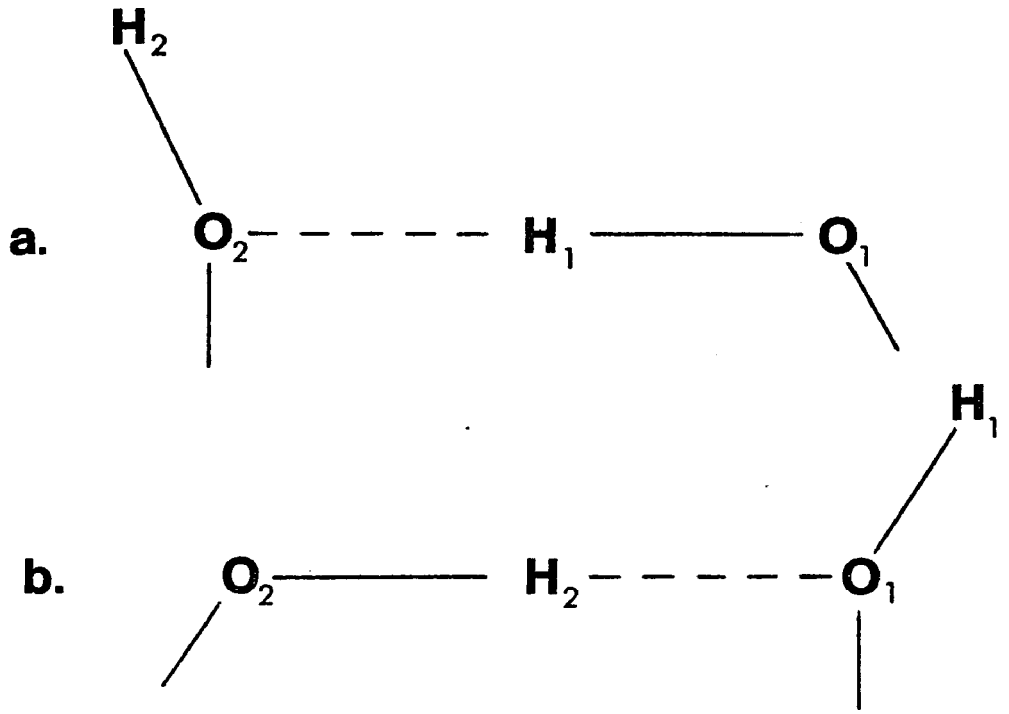
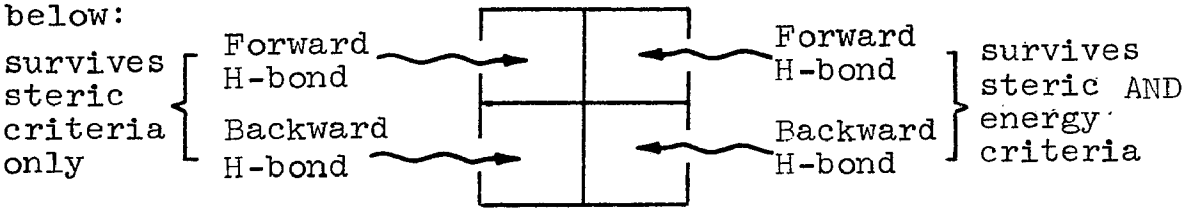


Figure 5-5. Two kinds of Type 4 hydrogen bonding. a) 'Forward' H-bonding; H_1 is the proton donor and b) 'Backward' H-bonding; H_2 is the proton donor.

and 'Backward' H-bonding is where H_2 donates a proton to O_1. The H-bonding direction is indicated in the same manner as the sign of θ_1 was in Type 3 H-bonding; that is, the top numbers in each box of the 4 x 4 array represent Forward proton donation while the bottom numbers denote Backward proton donation, as illustrated in the diagram below:



Matrix notation can be used to denote any element in the 4 x 4 array be referring to the (row,column) for that element. Thus, in Table 5-4, for $\omega=180^{\circ}$, CT dimer type, the (2,1) element contains 23 surviving conformations with Forward bonding, while only 10 Backward H-bonding survivors are found. All the numbers in this box correspond to $(\theta_1,\Gamma_1) = (+,+)$ and $(\theta_2,\Gamma_2) = (-,+)$. As before, we consider only the left-adjusted numbers. We also note that the cis'·cis conformations are not useful to the dimer for hydrogen bonding because of the large dimer energy in that area of conformational space. We therefore neglect these in further analysis. For trans ω, greater than 90% of the surviving conformations are for CC and CT dimers. For the trans'·trans dimer, all surviving conformations require both hydroxyl groups in the cis configuration (CC dimer type), while for the trans'·cis dimer, all conformations require CT dimer types.

5.5 Energy Restrictions as Criteria for Hydrogen Bonding

In the previous energy calculations involving the Pro-Pro internal dipeptide, it was seen that in the minimum energy regions, the M1 geometry was always the preferred geometry for each θ in the dimer. This idea was utilized to simplify the 2- and 5-methylproline calculations. Although dimer energy calculations have not been performed for 4-hydroxyproline, it is expected that similar results would be found in the absence of hydrogen bonding. It is entirely conceivable that the ring could pucker somewhat away from the M1 structure in order to form a hydrogen bond, thereby lowering the total dimer energy. This would occur only if the energy expense incurred in non-M1 puckering did not exceed that gained from hydrogen bond formation. One can as a first approximation, without knowing the energy contribution that dimer interactions might make to the non-M1 geometry dimer, utilize the monomer ring energies as a cutoff point. The energy involved in H-bond formation is about 3-8 kcal/mole (55). Therefore, a cutoff of 5 kcal/mole is employed. In other words, if the total energy of the ring or rings whose hydroxyl group(s) are participating in hydrogen bond formation exceeds 5 kcal/mole above the minimum energy ring(s), that conformation is rejected as a possible H-bonding conformation. Each bonding type is discussed separately.

TYPE 2 HYDROGEN BONDING

Energy restrictions have reduced the number of surviving conformations in half, from 38 to 19. All of these occur for $\Gamma_1 \leq -20^\circ$ and $\theta_1 \leq 0^\circ$. Thus, less than 1% of all possible Type 2 H-bonding conformations have ultimately survived beyond the energy cutoff restriction.

TYPE 3 HYDROGEN BONDING

We consider now the right-adjusted underlined numbers in Table 5-3. They represent the conformations that survive the energy cutoff. There are several conclusions that can be gleaned from these numbers. For Type 3 hydrogen bonding to occur, we must satisfy all the following conditions:

1. ψ must be in the cis' rotational state.
2. $(OH)_2$ must be in the cis configuration.
3. θ_1 must be $\geq 0^\circ$.
4. θ_2 must be $> 0^\circ$.
5. Γ_2 must be $< 0^\circ$.

TYPE 4 HYDROGEN BONDING

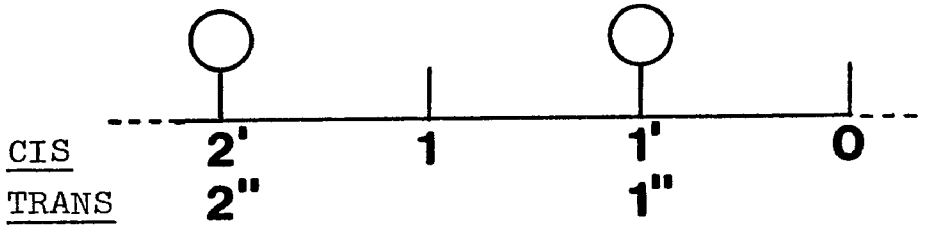
The results for the energy restrictions are shown for the various ψ values as the right-adjusted underlined numbers in Tables 5-4 through 5-7. The H-bonding conformations which survive the energy cutoff are confined almost exclusively to the cis' ψ region (Table 5-4). In fact, there are only 9 conformations outside this region and they occur in the $\psi=190^\circ$, ω=trans region of Table 5-7. At $\psi=190^\circ$, the dimer energy in the Pro-Pro dimer is high

compared to the minimum at $\psi=160^{\circ}$. Therefore, it is doubtful that these conformations would remain viable under a 4-hydroxyproline energy calculation. They are neglected in any further discussion. A summary of all H-bonding results and discussion is contained in Section 5.7.

5.6 Bifurcated Hydrogen Bonding

A bifurcated hydrogen bond can exist when a proper orientation of one hydroxyl and two oxygens exist such that the proton is donated and shared simultaneously by both oxygens. Although little is known regarding relative orientations of the three atoms involved, the results of this study may be used to rule out certain possible types of bifurcated hydrogen bonding. Table 5-8 contains a schematic drawing of the 4-hydroxyproline dimer where the carbonyl oxygens are denoted as vertical straight lines and the hydroxyl groups as vertical lines with open circles attached at the top. The oxygens and hydroxyls are numbered. Hydroxyl groups may be either cis or trans and this is noted by a prime (') or a double prime (") after the appropriate hydroxyl number. There are twenty bifurcated H-bonding types to be considered for the 4-hydroxyproline dimer and these are enumerated in Table 5-8. The notation for bifurcated H-bonds takes the form (X)YZ where X denotes the proton donor (either hydroxyl hydrogen) and YZ are the two proton acceptors (carbonyl

Table 5-8. Bifurcated Hydrogen Bonding Types in the 4-hydroxyproline Dimer. See Section 5.6 for a Description of the Notation.*



1. (1')01	11. (2')01
2. (1")01	12. (2")01
3. (1')02'	13. (2')01'
4. (1")02'	14. (2")01'
5. (1')02"	15. (2')01"
6. (1")02"	16. (2")01"
7. (1')12'	17. (2')11'
8. (1")12'	18. (2")11'
9. (1')12"	19. (2')11"
10. (1")12"	20. (2")11"

*In this diagram, vertical straight lines are carbonyl groups; vertical lines with circles attached at the top are hydroxyl groups.

oxygen and/or hydroxyl oxygen).

Having established the notation and with the results of the previous sections in mind, the possibility of bifurcated H-bonding can be discussed. Since Type 1 hydrogen bonding does not occur, any conformation involving 1' or 1" with O, and 2' or 2" with 1 can be discarded. This immediately eliminates conformation numbers 1-6, 11, 12 and 17-20 from Table 5-8. For Type 2 H-bonding, $(OH)_1$ must be in the cis configuration. This precludes any (1")XY conformations where X or Y equals 1; i.e., conformations 8 and 10 are eliminated. In a similar manner, numbers 14 and 16 are removed from consideration because the cis $(OH)_2$ position is required for Type 3 H-bonding. Conformation 15 can also be dropped because its Type 4 contribution requires a TC dimer type which, according to Table 5-4, does not exist.

Only three bifurcated H-bonding possibilities are left, numbers 7, 9 and 13. To ascertain whether or not any of these are ruled out by previous results, a closer look at the allowed torsion angles must be made. Type 2 H-bonding requires ψ to be in the trans' region, while ψ must be cis' for Type 4 bonding. This effectively eliminates conformations 7 and 9 from contention because they involve simultaneous Type 2 and Type 4 H-bonding. Conformation 13 involves Types 3 and 4 H-bonding. This bifurcated bonding type is not allowed, but the reasoning is more subtle. The Type 3 portion of the bond (c.f.

Table 5-3) requires $\theta_1 \geq 0^\circ$. In fact, it was mentioned that θ_1 must actually be $\geq 20^\circ$. The Type 4 H-bonding portion is a CC dimer type. Consulting Table 5-4 shows that the only survivors occur in the (2,2) and (4,2) boxes for $\omega = 190^\circ$ and only the (4,2) box for $\omega = 170^\circ$. In each of these cases $(\theta_1, \Gamma_1) = (-,+)$; that is, θ_1 must be negative. This contradiction in sign requirement for θ_1 between the Types 3 and 4 H-bonding portions rules out conformation 13 as a possible bifurcated hydrogen bond.

From the foregoing analysis and discussion, it is reasonable to conclude that the possibility of bifurcated hydrogen bonds in the 4-hydroxyproline internal dipeptide is, at best, highly unlikely.

5.7 Summary of Results and Conclusions on Hydrogen Bonding

Table 5-9 is a summary of results from this chapter. These results can be viewed in the light of earlier conclusions drawn from the Pro-Pro dimer energy calculations in Chapter III. For the Pro-Pro dimer where ψ is in the cis' rotational state, θ must be between $+5^\circ$ and $+10^\circ$. Dimer energies in this region increase sharply for other values of θ. Since $\theta_1 \geq 20^\circ$ for Type 3 hydrogen bonding, it is probably unlikely that this bonding type will occur in the 4-hydroxyproline dimer. Since, for Type 4 H-bonding, all θ (particularly $\theta = +10^\circ$) are present, it seems that this is the only type of hydrogen bonding that can

Table 5-9. Summary of Conclusions for Hydrogen Bonding in the 4-Hydroxyproline Internal Dipeptide

H-Bond Type	ω'	θ_1	Γ_1	ψ	ω	θ_2	Γ_2	$(OH)_i$ Position
2		-	-	trans'				cis
3	trans	+($\geq 20^\circ$)		cis'	trans	+	-	cis
4		$\pm$	$\pm$	cis'	trans	$\pm$	+	1st: cis 2nd: cis or trans

occur in the dimer. It should be stressed here that ring puckering of the type discussed in Chapter II plays the prominent role in conferring enough flexibility to the pyrrolidine rings so as to make intrachain hydrogen bonding possible. The fact that planar pyrrolidine rings are not conducive to H-bond formation between adjacent peptide groups has already been noted by Mattice and Mendelkern (71) who have extensively studied the experimental properties of sequential copolypeptides containing proline and 4-hydroxyproline (71,72). These authors have also noted that in the Hyp-Gly-Hyp sequence, a Type 3 hydrogen bond could be formed if ψ associated with the glycine residue were 120°. This is certainly quite possible as the ψ rotational freedom is large in glycine compared to proline. According to Figure 3-9, $\psi=120^{\circ}$ is about 6 kcal/mole higher than the $\psi=160^{\circ}$ minimum point for Region II. Were the glycine residue to be replaced by a proline, the gain in total dimer energy due to H-bond formation for this particular conformation might not be enough to justify formation of an H-bond in the first place.

One interesting aspect in connection with this study is that the cis hydroxyl position is implicated in all the bonding types for the dimer. However, naturally occurring hydroxyproline is found with (OH) in the trans position. Hydroxyproline is not one of the 20 amino (or imino) acids coded for in nucleic acids (26). The hydroxyl group placement is directed by a stereospecific

enzyme. The connection between short-range intramolecular hydrogen bonding and deliberate trans placement of the hydroxyl group in 4-hydroxyproline can be only speculatory at this stage of understanding.

CHAPTER VI

RESULTS AND DISCUSSION

6.1 Summary of Results

This research has demonstrated that puckering of the pyrrolidine ring is extremely important in considering the conformational properties of the proline and proline-like internal dipeptides. The major results may now be summarized:

1. The Γ-θ energy plot for the isolated pyrrolidine ring shows that there are two types of minimum energy structures available for the ring; one of Type A geometry and the other of Type B geometry. These two structures are of approximately the same energy, with Type B being slightly lower than Type A. Strictly speaking, these minimum energy structures are twisted.

2. The planar isolated proline ring is 2.4 kcal/mole higher in energy than the Type B minimum structure and, therefore, is energetically unfavorable.

3. The 2-methylproline, *trans* 5-methylproline and *trans* 4-hydroxyproline isolated rings all have puckering characteristics similar to those of the proline ring. The *cis* forms of 5-methylproline and 4-hydroxyproline are also similar, but have Type A structures slightly lower in energy than Type B structures.

4. There are three stable conformations for the Pro-Pro internal dipeptide; cis, trans and cis'. The flexibility of the pyrrolidine rings within the dimer (see Section 3.8) is different than that of the isolated ring. In particular, a Type B geometry is preferred by the dimer, whereas the isolated ring may have either a Type A or Type B geometry.

5. There are only two stable conformations for both the cis and trans 5-methylproline internal dipeptides; viz., cis and trans. The cis' structure does not occur (See Figures 4-13 and 4-14). A trans → cis transitional pathway exists with barrier height comparable to the same transition in the proline dimer.

6. There are only two stable conformations for the 2-methylproline internal dipeptide; i.e., the trans' and cis' structures. The cis structure is not possible (see Figure 4-22). The barrier height between the two allowed regions is low enough to permit a trans' → cis' transition. In contrast to the other molecules studied, it is the cis' minimum, not the trans' minimum, which has the lowest energy. This result has interesting implications on the NMR assignment of peaks due to the cis' structure.

7. In the 4-hydroxyproline internal dipeptide, there are four kinds of hydrogen bonding which are possible (see Section 5.7). Of these Type 4 H-bonding is most likely to occur. This bonding type is confined to the cis' ψ rotational state. In addition, $(OH)_1$ must be in the cis

configuration. 'Forward' hydrogen bonding is preferable to 'Backward' bonding.

6.2 Discussion

Through calculations presented in this thesis, we have gained a considerable amount of insight into the role pyrrolidine ring puckering plays in the conformation of an interior prolyl residue or dipeptide. There are, however, several aspect of this work that are incomplete. We shall briefly discuss these here.

The validity and usefulness of any energy calculation is, to a large extent, dependent upon the accuracy of the potential functions employed. As was mentioned in Chapter I, more accurate potential functions have been under development recently, particularly by Lewis *et al.* (75), Hagler *et al.* (76) and Hagler and Lifson (77). While use of more refined potential functions would probably not affect any of the general conclusions presented here, they will be important in obtaining accurate conformational energies. In this context, more emphasis could be placed on the normal coordinate analysis of small oligopeptides in an effort to devise a set of reliable self-consistent force fields for these molecules. Accurate values for bond angle bending constants are important in assessing conformational flexibility of the proline ring.

A more subtle aspect in the category of possible improvements to these calculations concerns the nature of

the orbitals about the nitrogen atom. It is entirely possible that the three bonds meeting at the nitrogen may not be coplanar, but instead, somewhat pyramidal. Recent quantum mechanical calculations on N-methyl acetamide, as reported by Ramachandran (79), indicate that a slightly pyramidal nitrogen configuration is preferred by that molecule. Moreover, the further that the torsion angle ω departs from 180^{o}, the greater is the pyramidal nature of the nitrogen atom. It is clear that some account of this behavior should be taken in calculations where ω is allowed to vary. This is particularly true of our barrier calculations (Chapters III and IV).

All calculations in this thesis have been done for molecules in vacuo. In particular, solvent effects have not been considered. Obviously, these must be studied if we wish to make relevant comparisons between theory and experiments involving solutions of these molecules. Two approaches to a theoretical treatment of polymer-solvent interactions have been reported (1-4). One involves explicitly placing the solvent molecule in different orientations around the polymer and evaluating the energy of the system for each solvent orientation. Unfortunately this takes a lot of computing time. The second approach is the hydration shell theory (3,4) which ignores explicit solvent interactions.

In addition to solvation, there are other molecular environments where the ring puckering may be quite

different. For example, we have reported (81) that the flexibility of the proline ring is further modified in the case of helical polyproline. The energy of both cis and trans helical molecules with six units has been computed (see Section 3.8 for discussion on backbone geometries employed). Figure 6-1 shows the Γ-θ energy plot for the cis helix, where the best value of ψ giving the lowest helical energy is employed. In contrast to the isolated ring, only negative values of θ are preferred by the helix. The global minimum structure of Figure 6-1 is Type B geometry and is almost β^+ puckered, while the secondary minimum is close to δ^- puckered. This is different from the isolated ring, where minimum energy structures may involve γ-puckering to a large extent.

In Figure 6-2, the Γ-θ energy plot for the trans helix is shown. It is likely that the puckering in helical trans polyproline will be essentially Type B geometry corresponding to a twisted structure close to γ^- and β^+ puckered structures.

Thus we see that the ring puckering in helical polyproline will be of only one type (namely, Type B) even though the isolated ring may equally well adopt either Types A or B puckering.

We have briefly touched upon some crystal calculations in Chapter III, but nothing has been said about how ring puckering might affect packing of polyproline and the polymethylprolines in the solid state. The crystal

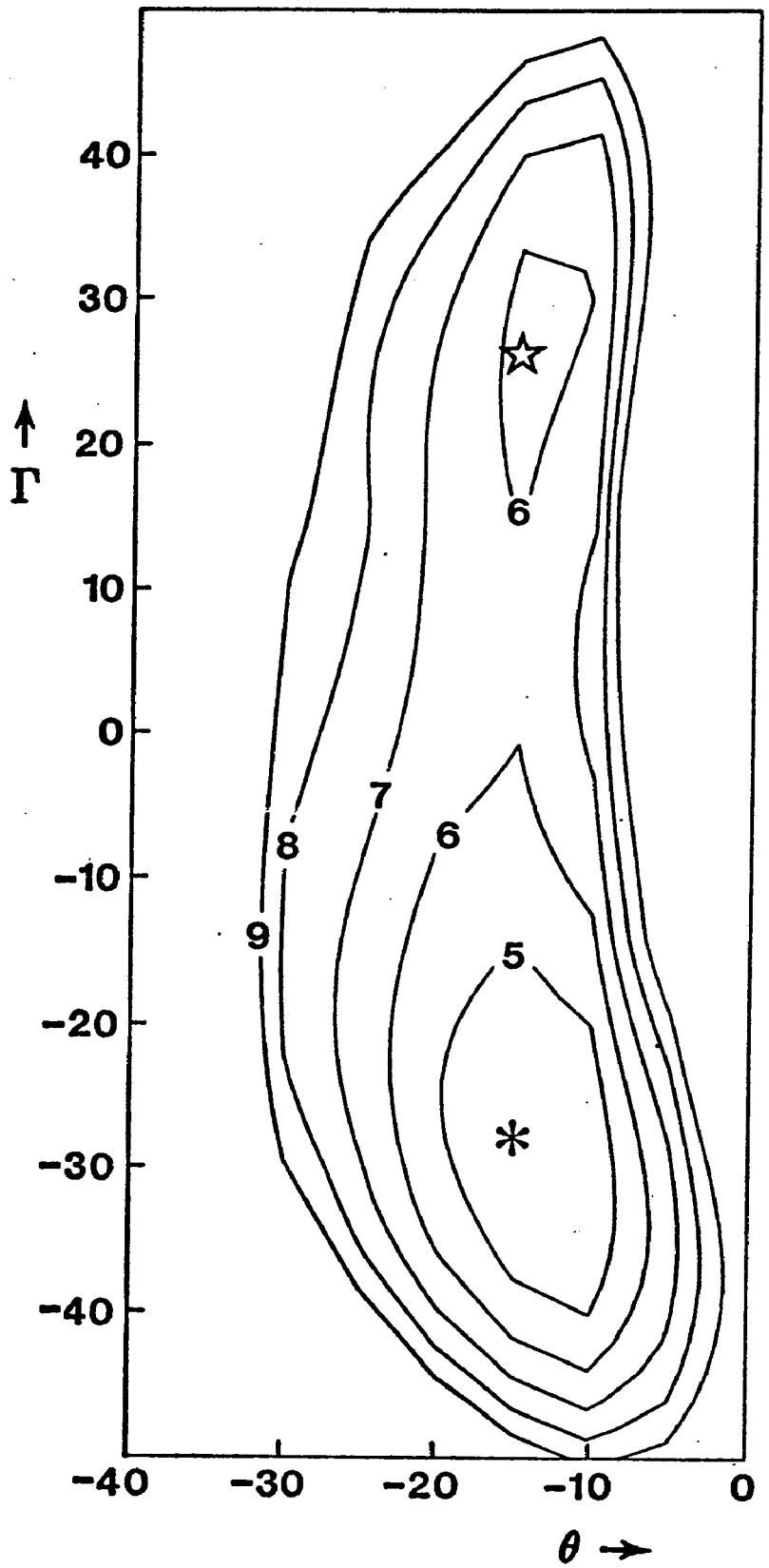


Figure 6-1. Γ-θ energy plot for the cis proline helix. Each (Γ,θ) energy has been minimized with respect to ψ.

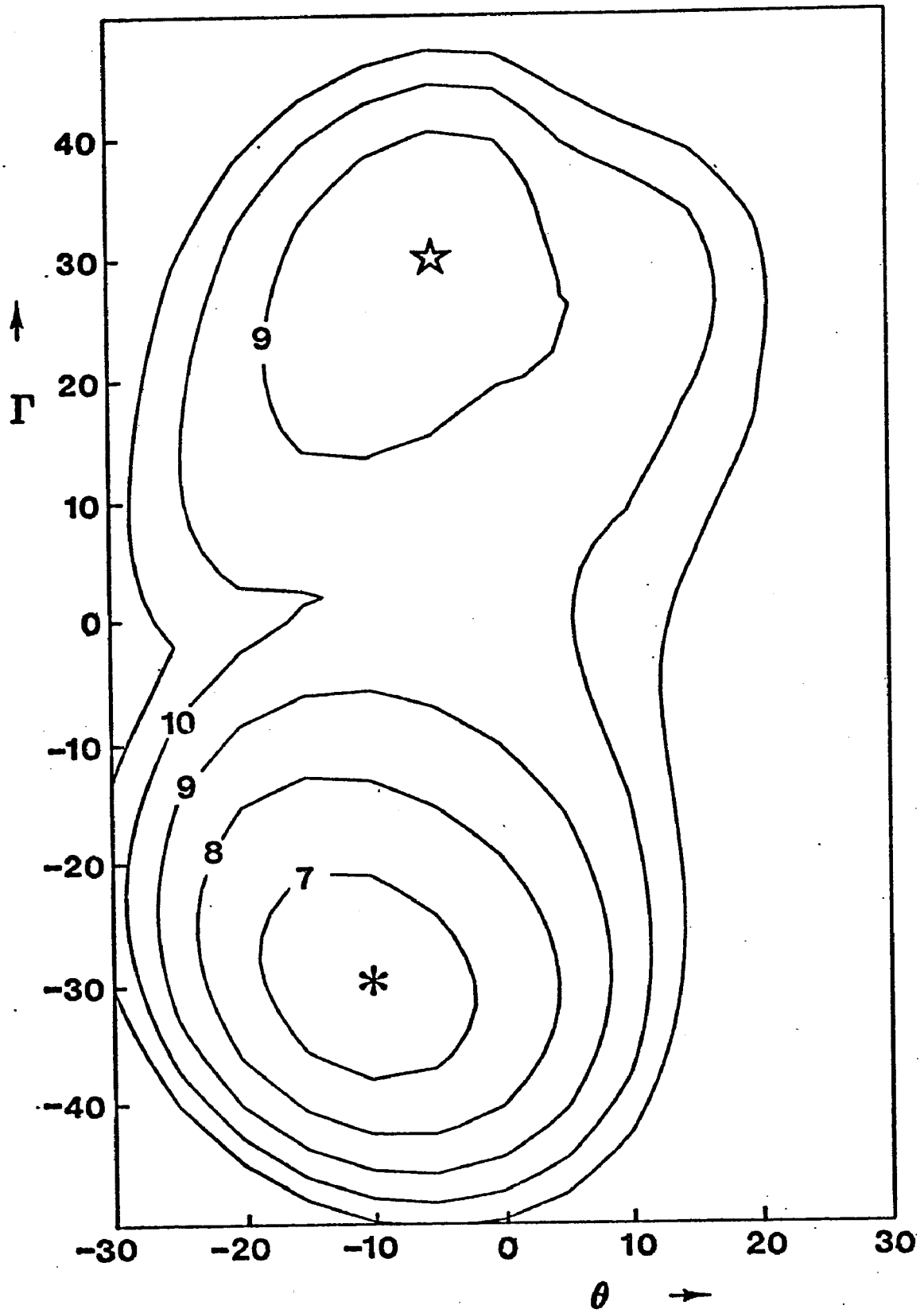


Figure 6-2. Γ-θ energy plot for the trans proline helix. Each (Γ,θ) energy has been minimized with respect to ψ.

structures for the methyl prolines have not been determined. Especially interesting is the effect that a low cis' energy in the 2-methylproline dimer might have on the packing of this molecule in a crystal. Also interesting are copolymers containing both 2- and 5-methylprolines as well as proline. Such sequences are currently under investigation (80) and the results should prove quite interesting.

It is expected that valuable experimental techniques such as NMR and IR should be useful in correlating our theoretical predictions on ring puckering with those found in solution. Recently, the mobilities of the pyrrolidine ring atoms have been inferred from relaxation time measurements (83,84).

Finally, we would like to point out that our procedure for studying the energy of the isolated ring as a function of Γ and θ could be refined by the application of quantum mechanical molecular orbital calculations. These calculations may throw light on the roles that bond angle variation and pyramidal nitrogen configuration play in the flexibility of the isolated ring.

APPENDIX I

CALCULATION OF CONFORMATIONAL ENERGIES

I.1 Terms Contributing to the Total Energy

In the semi-empirical approach to calculating conformational energies, the total energy of a molecule is expressed as a sum of several separable terms. These terms include: (1) van der Waals interactions (V_{vw}); viz., those originating from atomic interactions, (2) electrostatic interactions (V_{es}), (3) intrinsic torsional energies (V_{χ}) due to internal rotation about single or partial double bonds, (4) bond angle deformation energies (V_{τ}), (5) bond length stretching energies (V_{ℓ}), (6) hydrogen bonding (V_{hb}) and (7) solvent-molecule interactions (V_{ms}). Each of the terms in Equation (1) below may be conformation dependent:

$$V_{total} = V_{vw} + V_{es} + V_{\chi} + V_{\ell} + V_{hb} + V_{ms} \qquad (1)$$

Values for each of these terms have been determined from physico-chemical data on low-molecular weight model compounds. These data have been refined by applying the equations to other systems where structural information is available. In this thesis, no calculations involving molecule-solvent interactions are reported. In addition, bond length stretching (except in the case of rotation

about the peptide bond, see Chapter III) has not been considered. Hydrogen bonding energies are not discussed here, although several hydrogen bonding functions have been proposed (85-88). The remaining terms and their functional forms are described below.

I.2 V_{vw} - Van der Waals Interaction Energies

The van der Waals interactions are expressed as a pairwise sum over all interactions between third and higher neighboring pairs of atoms, that is, atoms whose relative positions are determined by at least one torsion angle. There are two factors which contribute to the van der Waals energy. The first, which is attractive, arises from the interaction between transient fluctuating charge distributions. The second is repulsive and involves electronic orbital overlap. There are two functional forms for this repulstion and they are reflected in the two most commonly used equations describing the van der Waals energy:

$$V_{vw} = \sum_{i,j} [a_{ij}/r_{ij}^{12} - b_{ij}/r_{ij}^{6}] \qquad (2)$$

Lennard-Jones "6-12" Potential

$$V_{vw} = \sum_{i,j} [c_{ij}\exp(-Dr_{ij}) - b_{ij}/r_{ij}^{6}] \qquad (3)$$

Buckingham "6-exp" Potential

In each equation, r_{ij} is the distance (in Angstroms) between the interacting pair. The other constants in these equations are independent of r_{ij}. The attractive

term constant, b_{ij}, is calculated for both types of van der Waals functions from the Slater-Kirkwood equation (55,89). Both a_{ij} and c_{ij} are determined by requiring that the minimum value of V_{vw} occurs when the interacting pair of atoms is separated by the sum of their van der Waals radii. D is determined from collision experiments in rare gases (55,89). Several undesirable characteristics preclude the use of the Buckingham potential in this thesis. At extremely short internuclear distances, this function exhibits an unrealistically high maximum, then proceeds to $-\infty$ as the internuclear distance approaches zero. There is also one additional parameter, D, in the Buckingham potential. These problems can be dealt with by careful programming, but with an increase in computing time. Since both functions exhibit nearly the same behavior for moderate r_{ij}, the Lennard-Jones form is preferable and has been used exclusively in this work. There is a range of values reported in the literature for the interaction constants (8, 90,91). For these calculations, Scott and Scheraga's (8) parameters have been used. Table I-1 lists these constants for all interaction types.

I.3 V_{es} - Electrostatic Interaction Energies

Dipole-dipole electrostatic interactions are important in conformational energy calculations. Two interacting dipoles $\underset{\sim}{\mu}_i$ and $\underset{\sim}{\mu}_j$ may be taken into account explicitly as:

Table I-1. Van der Waals Interaction Constants[a]

Interacting pair	r_{min}[b]	d_{ij}[c]	c_{ij}[d]
C···C	3.40	2.86×10^5	370
C···N	3.25	2.16×10^5	366
C···O	3.32	2.05×10^5	367
C···H	2.90	3.80×10^4	128
N···N	3.10	1.61×10^5	363
N···O	3.07	1.53×10^5	365
N···H	2.75	2.70×10^4	125
O···O	3.04	1.45×10^5	367
O···H	2.72	2.51×10^4	124
H···H	2.40	4.46×10^3	46.7

[a]Taken from Scott and Scheraga (8).

[b]Sum of the van der Waals radii in Å.

[c]In units of kcal·$Å^{12}$/mole.

[d]In units of kcal·$Å^6$/mole.

$$V_{es} = \sum_{i,j} [(\underset{\sim}{\mu}_i \cdot \underset{\sim}{\mu}_j)/\epsilon r_{ij}^3 - 3(\underset{\sim}{\mu}_i \cdot \underset{\sim}{r}_{ij})(\underset{\sim}{\mu}_j \cdot \underset{\sim}{r}_{ij})/\epsilon r_{ij}^5] \quad (4)$$

where r_{ij} is the distance between them. This is referred to as the dipole approximation. Alternatively, in the monopole approximation, one can compute a set of residual charges on each atom in the molecule, subject to the condition that dipole moments of bonds and groups of atoms are reproduced. For proline, such a method was used by Holzwarth and Chandrasekharan (33) for the known dipole moment in a number of substituted alkylamides which have in common the group N-(C =O)-C^{α}, analogous to a peptide bond. The resultant charges on these four atoms are contained in the first column of Table I-2.

The most widely used method for calculating the partial charges is attributed to Del Re et al. (93,94). The method is essentially the same technique, for saturated systems, as the Hückel method for π-electron systems. In treating the σ-bonds as localized units rather than delocalized orbitals, this technique is semi-empirical and relies upon a suitable choice of parameters which yields charge distributions corresponding to observed dipole moments. Del Re's method has been applied by Poland and Scheraga (47) as well as Hopfinger and Walton (34) for the internal proline monomer unit. Columns 2 and 3 of Table I-2 contain these two sets of charges. In comparing them, it is interesting to note the similarities

Table I-2. Comparison of Different Sets of Partial Charges (in e.s.u.) for the Proline Internal Unit

Atom	Set A[a]	Set B[b]	Set C[c]
N	+0.220	-0.048	-0.036
C^{α}	+0.110	+0.037	+0.037
C^{β}	0	-0.067	-0.068
C^{γ}	0	-0.066	-0.068
C^{δ}	0	-0.030	-0.021
H^{α}	0	+0.046	+0.045
H^{β}'s	0	+0.038	+0.039
H^{γ}'s	0	+0.038	+0.039
H^{δ}'s	0	+0.045	+0.045
O	-0.480	-0.425	-0.127
C	+0.260	+0.311	+0.270

[a]Holzwarth and Chandrasekharan (33)

[b]Poland and Scheraga (47)

[c]Hopfinger and Walton (34)

except for the C and O atoms. Poland and Scheraga have included π charges from the model compound formamide in their calculations of residual charge for the N, C and O atoms. It is not clear that Hopfinger and Walton have done this. Therefore, the Poland and Scheraga charges have been used in all dimer calculations presented (for ring self-energy calculations, see Chapter II).

Once the net charges are known, the electrostatic energy is calculated from a Coulombic type of potential:

$$V_{es} = 332 q_i q_j / \epsilon r_{ij} \qquad (5)$$

where q_i and q_j are the partial charges (in e.s.u.) of each non-bonded interacting atom pair i and j. r_{ij} is the distance between them and the constant 332 converts the expression to kcal/mole. ϵ is the effective dielectric constant, for which a reliable estimate is difficult to make. Following Brant and Flory (7,95) and Goodman et al. (96), a value of 4.0 is used throughout this thesis.

I.4 V_χ - Intrinsic Torsional Potential Energies

V_χ is the intrinsic torsional energy of rotation about a covalent bond. This energy arises from the exchange interactions between electrons other than in the bond about which rotation is taking place. The general form (55) for a single bond of a tetrahedrally bonded atom may be approximately taken to be:

$$V_\chi = \tfrac{1}{2}V_1(1 + \cos 3\chi) + \tfrac{1}{2}V_2(1 + \cos 6\chi) + \ldots. \qquad (6)$$

V_1 and V_2 are barrier constants for the three-fold and six-fold symmetry barriers and are obtained from microwave spectral data on small molecule analogues. $V_2 \sim 1/100\ V_1$, so the second term has been neglected in these calculations. The minimum potential energy occurs for the staggered relative orientations, i.e., at $\chi=60^\circ$, 180° and 300°. For the C-C bond within the pyrrolidine ring, $V_1=3.4$ kcal/mole (10). For the C-N bonds, a lower value of 1.0 kcal/mole (10) was employed.

For rotation about ℓ(C-N) (ω rotation), a different function form is required, due to the trigonal planar nature of both the C' and N atoms. ℓ(C-N) is a partial double bond, in contrast to a regular double bond like C=C where rotation can be neglected. For possible large variations in ω, such as are considered in this thesis, the torsional potential takes the form (55):

$$V_\omega = \tfrac{1}{2}V_1(1 - \cos 2\omega) \qquad (7)$$

corresponding to the observation that the C'=O prefers to <u>eclipse</u> either $\ell(N\text{-}C^\alpha)$ or $\ell(N\text{-}C^\delta)$. The barrier V_1 as measured from model amide compounds (55) is 20 kcal/mole.

I.5 V_τ - <u>Bond Angle Deformation Energies</u>

The bond angle bending energy takes the form:

$$V_\tau = K_\tau(\tau - \tau_o)^2 \qquad (8)$$

where τ and τ_o are, respectively, the actual and equilibrium bond angles (in radians) for each deformable angle

in the molecule. The force constant K_τ was taken as 72 kcal/mole for angles about the carbon atoms. Following Bixon and Lifson (92), τ_o was set at 112°; i.e., the C-C-C bond angle of n-alkanes. In the case of the three angles around the sp^2 nitrogen atom, τ_o was assigned a value of 120° for each angle, in accordance with a normal coordinate analysis of N,N-dimethylacetamide by Johnston (78). The force constants for the nitrogen angles were calculated from Johnston's results (78). Both exterior angles, $\tau^{N'}$ and $\tau^{N''}$ had K_τ=72 kcal/mole, while for τ^{N}, a value of 107 kcal/mole was calculated and employed. In the case of the sp^3 nitrogen, equilibrium bond angles of 109° around the nitrogen atom were used. A K_τ of 72 kcal/mole was used for all bond angles around the nitrogen.

APPENDIX II

A FORTRAN PROGRAM FOR Γ-ARCING

II.1 Description of FXAT4 Subroutine

This program calculates the position of the C^{γ} atom corresponding to the $\Gamma=0^{o}$ position, consistent with ring closure. It takes as input the Cartesian coordinates of the ring atoms with arbitrary origin as well as the bond lengths between the closure atom (C^{γ}) and its two neighbors. It is described here in terms of the pyrrolidine ring, but FXAT4, as it stands, can be used for any five-membered ring. The program is compact, so no flowchart is included. Instead, the Fortran listing is given along with an explanation of the Γ-arcing mathematics.

II.2 Input and Output Parameters

Input: $\underset{\sim}{XB}$ = (x,y,z) Cartesian coordinates of the C^{β} atom.

$\underset{\sim}{XA}$ = (x,y,z) Cartesian coordinates of the C^{α} atom.

$\underset{\sim}{XN}$ = (x,y,z) Cartesian coordinates of the N atom.

$\underset{\sim}{XD}$ = (x,y,z) Cartesian coordinates of the C^{δ} atom.

ALBG = $\ell(C^{\beta}-C^{\gamma})$

ALDG = $\ell(C^{\delta}-C^{\gamma})$

Output:

$\underset{\sim}{XCNTR}$ = center of circle whose circumference is the locus of all possible C^{γ} positions.

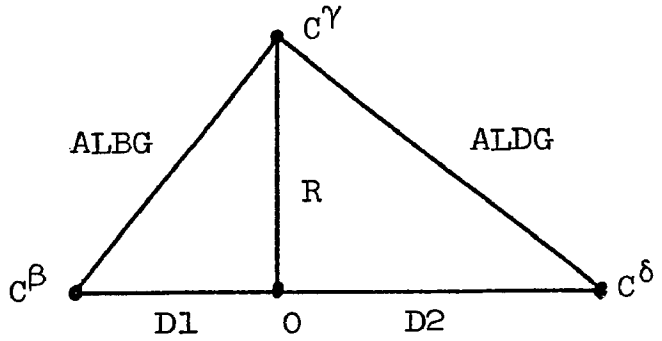
$\underset{\sim}{XG}$ = (x,y,z) Cartesian coordinates of the C^{γ} atom in the $\Gamma=0^{o}$ position.

$\overset{\wedge}{UG}$ = unit vector along the $C^{\beta}...C^{\delta}$ direction, pointing towards C^{δ}.

IG = integer constant which is returned as 0 if ring closure is not possible. When IG=0, FXAT4 returns immediately to the main program.

II.3 Mathematics of Γ-Arcing

Consider the plane formed by the three ring atoms C^{β}, $C^{\delta}$and C^{γ} as shown below. O is the center of the circle which is the intersection of two spheres whose centers are the C^{β} and C^{δ} atoms and whose radii are ALBG and ALDG.



We wish to calculate D1 and R, which are interrelated by the Pythagorean theorem. By eliminating R from the two equations:

$$(D1)^2 + R^2 = (ALBG)^2$$
$$(D2)^2 + R^2 = (ALDG)^2$$

and utilizing the fact that D1+D2 = DBD = $C^{\beta}...C^{\delta}$ distance (calculated by FXAT4 from input data) one can solve for D1:

$$D1 = \frac{(DBD)^2 + (ALBG)^2 - (ALDG)^2}{2(DBD)}$$

and subsequently for R:

$$R = \{(ALBG)^2 - (D1)^2\}^{\frac{1}{2}}$$

What remains is to find $\underset{\sim}{XCNTR}$ and the direction of R corresponding to the $\Gamma=0^o$ position for the C^γ atom. $\underset{\sim}{XCNTR}$ is obtained by adding the two vectors:

$$\underset{\sim}{XCNTR} = \underset{\sim}{XB} + D1*\hat{UG}$$

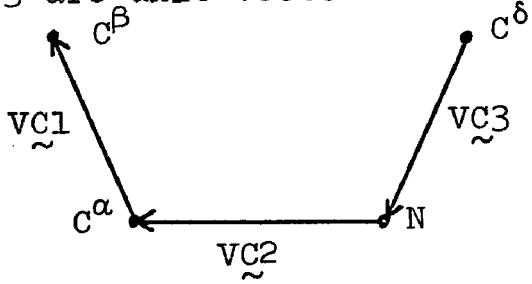
The direction of R is given by the unit vector $\hat{XR}$:

$$\hat{XR} = (\hat{VN} \times \hat{UG})$$

where:

$$\hat{VN} = \{(\hat{VC1} \times \hat{VC2}) + (\hat{VC2} \times \hat{VC3})\}$$

$\hat{VC1}$, $\hat{VC2}$ and $\hat{VC3}$ are unit vectors as defined by the following diagram:



Finally, the position of the C^γ atom corresponding to $\Gamma=0^o$ is given by:

$$\underset{\sim}{XG} = \underset{\sim}{XCNTR} + R*\hat{XR}$$

FXAT4 uses the following standard subroutines:

1. DIST2(A,B): distance between two atoms A and B whose Cartesian coordinates are known.
2. UVEC1(A,B): transforms vector $\underset{\sim}{A}$ into unit vector $\hat{B}$.
3. CRPR1(A,B,C): performs the cross product of two vectors, $\underset{\sim}{A} \times \underset{\sim}{B} = \underset{\sim}{C}$.

To obtain a Γ-arced position for the C^{γ} atom, the line $O...C^{\gamma}$ is rotated about $C^{\beta}...C^{\delta}$. To achieve this, a short subroutine ROTN1 has been written for use with FXAT4. There are three inputs for ROTN1:

R = 3 x 3 rotation matric through an angle Γ about the vector $\hat{UG}$.

XCNTR = output from FXAT4.

$\hat{UG}$ = output from FXAT4.

The output for ROTN1, $\underset{\sim}{XG}$, is the position vector of the C^{γ} atom for the given Γ.

II.4 Fortran Code for FXAT4 and ROTN1

```
      SUBROUTINE FXAT4(XB,XA,XN,XD,ALBG,ALDG,XCNTR,XG,UG,IG)
      DIMENSION XB(3),XD(3),XA(3),XN(3),XCNTR(3),XG(3),UG(3),
     1VC1(3),VC2(3),VC3(3),VN1(3),VN2(3),VN(3),XR(3)
      DBD = DIST2(XB,XD)
      D1 = (DBD*DBD + ALBG*ALBG - ALDG*ALDG)/(2.*DBD)
      RSQ = ALBG*ALBG - D1*D1
      IF(RSQ.GT.0)  GO TO 100
      IG = 0
      RETURN
  100 IG = 1
      R = SQRT(RSQ)
      DO 200 I = 1,3
  200 UG(I) = XD(I) - XB(I)
      CALL UVEC1(UG,UG)
      DO 300 I = 1,3
  300 XCNTR(I) = XB(I)+D1*UG(I)
      DO 400 I = 1,3
      VC1(I) = XB(I) - XA(I)
      VC2(I) = XA(I) - XN(I)
  400 VC3(I) = XN(I) - XD(I)
      CALL CRPR1(VC1,VC2,VN1)
      CALL CRPR1(VC2,VC3,VN2)
      CALL UVEC1(VN1,VN1)
      CALL UVEC1(VN2,VN2)
      DO 500 I = 1,3
  500 VN(I) = VN1(I) + VN2(I)
      CALL UVEC1(VN,VN)
      CALL CRPR1(VN,UG,XR)
      CALL UVEC1(XR,XR)
      DO 600 I = 1,3
  600 XG(I) = XCNTR(I) + R*XR(I)
```

```
      RETURN
      END


      SUBROUTINE ROTN1(R,XCNTR,UG,XG)
      DIMENSION R(3,3),XCNTR(3),UG(3),XG(3)
      DO 10 I = 1,3
      XG(I) = 0.
      DO 20 J = 1,3
   20 XG(I) = XG(I) + R(I,J)*(UG(J)-XCNTR(J))
   10 XG(I) = XG(I) + XCNTR(I)
      RETURN
      END
```

II.5 Example of FXAT4 and ROTN1 Usage

The following sequence of Fortran statements illustrates how these subroutines can be used. All inputs are assumed to be already calculated in the main program.

```
      .
      .
      .
      CALL FXAT4(XB,XA,XN,XD,ALBG,ALDG,XCNTR,XG,UG,IG)
      IF(IG.EQ.0)  GO TO 500
      CALL ROTN1(R,XCNTR,UG,XG)
  500 CONTINUE
      .
      .
      .
```

BIBLIOGRAPHY

1. K.D. Gibson and H.A. Scheraga, Proc. Natl. Acad. Sci. U.S.A. 58, 420 (1967).

2. A.J. Hopfinger, Macromolecules 4, 731 (1971).

3. S. Krimm and C.M. Venkatachalam, Proc. Natl. Acad. Sci. U.S.A. 68, 2468 (1971).

4. C.M. Venkatachalam and S. Krimm, in "Conformation of Biological Molecules and Polymers," The Israel Academy of Sciences and Humanities, p. 141 (1973).

5. G. Alagona, A. Pullman, E. Scrocco and J. Tomasi, Int. J. Peptide Protein Res. 5, 251 (1973).

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

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

8. R.A. Scott and H.A. Scheraga, J. Chem. Phys. 45, 2091 (1966).

9. C.M. Venkatachalam, Biopolymers 6, 1425 (1968).

10. G.N. Ramachandran, A.V. Lakshminarayanan, R. Balasubramanian and G. Tegoni, Biochim. Biophys. Acta 221, 165 (1970).

11. V. Sasisekharan, Proc. Ind. Acad. Sci. 53A, 296 (1961).

12. V. Sasisekharan, in "Collagen" (N. Ramanathan, ed., Interscience Publishers, New York) p. 39 (1962).

13. G.N. Ramachandran, C. Ramakrishnan and V. Sasisekharan, J. Mol. Biol. 7, 95 (1963).

14. G.N. Ramachandran, C. Ramakrishnan and V. Sasisekharan, in "Aspects of Protein Structure" (G.N. Ramachandran, ed., Academic Press, London) p. 121 (1963).

15. J.A. Schellman and C. Schellman, in "The Proteins" (H. Neurath, ed., Academic Press, New York, vol. 2) pp. 14-24 (1964).

16. H.A. Scheraga, in "Molecular Architecture in Cell Physiology" (T. Hayashi and A.G. Szent-Gyorgyi, eds.) p. 39 (1966).

17. G. Nemethy and H.A. Scheraga, Biopolymers 3, 155 (1965).

18. H.A. Scheraga, S.J. Leach, R.A. Scott and G. Nemethy, Discuss. Faraday Soc. 40, 268 (1965).

19. H.A. Scheraga, G. Nemethy, S.J. Leach, R.A. Scott and D. Poland, Fed. Proc. 24, 413 (1965).

20. S.J. Leach, G. Nemethy and H.A. Scheraga, Biopolymers 4, 369 (1966).

21. S.J. Leach, G. Nemethy and H.A. Scheraga, Biopolymers 4, 887 (1966).

22. G. Nemethy, S.J. Leach and H.A. Scheraga, J. Phys. Chem. 70, 998 (1966).

23. P.J. Flory, "Statistical Mechanics of Chain Molecules" (Interscience Publishers, New York) Chapters I, II and III (1969).

24. D.A. Brant, W.G. Miller and P.J. Flory, J. Mol. Biol. 23, 47 (1967).

25. IUPAC-IUB Abbreviations and Symbols for the Description of the Conformation of Polypeptide Chains, J. Mol. Biol. 52, 1 (1970).

26. R.E. Dickerson and I. Geis, "The Structure and Action of Proteins" (Harper and Row, New York, 1969).

27. P.M. Cowan and S. McGavin, Nature 176, 501 (1955).

28. Y.C. Leung and R.E. Marsh, Acta Crystallogr. 11, 17 (1958)

29. V. Sasisekharan, Acta Crystallogr. 12, 897 (1959).

30. W. Traub and U. Shmueli, Nature 198, 4886 (1963).

31. T. Ueki, T. Ashida, M. Kakudo, Y. Sasada and Y. Katsube, Acta Crystallogr. 25B, 1840 (1969).

32. I.L. Karle, J. Am. Chem. Soc. 94, 81 (1972).

33. G. Holzwarth and R. Chandrasekaran, Macromolecules 2, 245 (1969).

34. A.J. Hopfinger and A.G. Walton, J. Macromol. Sci. Phys. B3(1), 171 (1969).

35. S. Krimm and C.M. Venkatachalam, Proc. Natl. Acad. Sci. U.S.A. 68, 2468 (1971).

36. P.R. Schimmel and P.J. Flory, Proc. Natl. Acad. Sci. U.S.A. 58, 52 (1968).

37. B. Maigret, D. Perahia and B. Pullman, J. Theor. Biol. 29, 275 (1970).

38. V. Madison and J. Schellman, Biopolymers 9, 65 (1970).

39. N. Gō and H.A. Scheraga, Macromolecules 3, 188 (1970).

40. K. Nishikawa and T. Ooi, Bull. Inst. Chem. Res., Kyoto Univ. 50, 94 (1972).

41. S. Tanaka and H.A. Scheraga, Macromolecules 7, 698 (1974).

42. W.L. Mattice and L. Mandelkern, J. Am. Chem. Soc. 93, 1769 (1971).

43. W.L. Mattice, K. Nishikawa and T. Ooi, Macromolecules 6, 443 (1973).

44. D.A. Torchia, Macromolecules 4, 440 (1971).

45. R. Balasubramanian, A.V. Lakschminarayanan, M.N. Sabesan, G. Tegoni, K. Venkatesan and G.N. Ramachandran, Int. J. Protein Res. 3, 25 (1971).

46. K. Nishikawa and T. Ooi, Prog. Theoret. Phys. 46, 670 (1971).

47. D. Poland and H.A. Scheraga, Biochemistry 6, 3791 (1967).

48. C.M. Venkatachalam, B.J. Price and S. Krimm, Macromolecules 7, 212 (1974).

49. W.A. Thomas and M.K. Williams, J. Chem. Soc. Chem. Commun. 994 (1972).

50. D.E. Dorman, D.A. Torchia and F.A. Bovey, Macromolecules 6, 80 (1973).

51. I.Z. Steinberg, W.F. Harrington, A. Berger, M. Sela and E. Katchalski, J. Am. Chem. Soc. 82, 5263 (1960).

52. P.R. Schimmel and P.J. Flory, J. Mol. Biol. 34, 105 (1968).

53. D.A. Torchia and F.A. Bovey, Macromolecules 4, 246 (1971).

54. C.C. Wu, R.A. Komoroski and L. Mandelkern, Biophys. Soc. Abstr. 15, 188a (1975).

55. G.N. Ramachandran and V. Sasisekharan, Advan. Prot. Chem. 23, 283 (1968).

56. H.S. Gutowsky and C.H. Holm, J. Chem. Phys. 25, 1228 (1956).

57. M.T. Rogers and J.C. Woodbury, J. Phys. Chem. 66, 540 (1962).

58. R.M. Hammaker and B.A. Gugler, J. Mol. Spectr. 17, 356 (1965).

59. C.K. Johnson, Report ORNL-3794 (modified), Oak Ridge National Laboratory, Oak Ridge, Tennessee (1965).

60. A.E. Tonelli, J. Am. Chem. Soc. 95, 5946 (1973).

61. C.M. Venkatachalam, B.J. Price and S. Krimm, Biopolymers 14, 1121 (1975).

62. Y. Abe and S. Krimm, Biopolymers 11, 1817 (1972).

63. C.G. Overberger, K.-H. David and J.A. Moore, Macromolecules 5, 368 (1972).

64. C.G. Overberger and K.-H. David, Macromolecules, 5 373 (1972).

65. C.G. Overberger, private communication.

66. C.M. Venkatachalam, private communication.

67. V. Sasisekharan, Acta Crystallogr. 12, 903 (1959).

68. A.J. Hopfinger and A.G. Walton, J. Macromol. Sci. Phys. B3(1), 195 (1969).

69. W.L. Mattice and L. Mandelkern, Macromolecules 3, 199 (1970).

70. W.L. Mattice and L. Mandelkern, Biochemistry 9, 1049 (1970).

71. W.L. Mattice and L. Mandelkern, Biochemistry 10, 1926 (1971).

72. W.L. Mattice and L. Mandelkern, Biochemistry 10, 1934 (1971).

73. T. Ooi, D.S. Clark and W.L. Mattice, Macromolecules 7, 337 (1974).

74. R. Chidambaram, R. Balasubramanian and G.N. Ramachandran, Biochim. Biophys. Acta 221, 182 (1970).

75. P.N. Lewis, F.A. Momany and H.A. Scheraga, Israel J. Chem. 11, 121 (1973).

76. A.T. Hagler, E. Huler and S. Lifson, J. Am. Chem. Soc. 96, 5319 (1974).

77. A.T. Hagler and S. Lifson, J. Am. Chem. Soc. 96, 5329 (1974).

78. N. Johnston, Ph.D. Dissertation, University of Michigan (1975).

79. G.N. Ramachandran, in "Proceedings of the Rehovot Symposium on Poly(Amino Acids), Polypeptides, and Proteins and their Biological Implications" pp. 14-34 (1974).

80. C.G. Overberger, private communication to C.M. Venkatachalam.

81. B.J. Price and C.M. Venkatachalam, in "Environmental Effects on Molecular Structure and Properties," the Eighth Annual Jerusalem Symposium (to be published).

82. S. Arnott and S.D. Dover, Acta Crystallogr. B24, 599 (1968).

83. R. Deslauriers, I.C.P. Smith and R. Walter, J. Am. Chem. Soc. 96, 2289 (1974).

84. D.A. Torchia and R. Lyerla, Biopolymers 13, 97 (1974).

85. E.R. Lippincott and R. Schroeder, J. Chem. Phys. 23, 1099 (1955).

86. R.A. Scott and H.A. Scheraga, J. Chem. Phys. 44, 3054 (1966).

87. G.N. Ramachandran, R. Chandrasekharan and R. Chidambaram, Proc. Ind. Acad. Sci. 74, 270 (1971).

88. A.T. Hagler and S. Lifson, in "Proceedings of the Rehovot Symposium on Poly(Amino Acids), Polypeptides, and Proteins and their Biological Implications" pp. 35-48 (1974).

89. H.A. Scheraga, Adv. Phys. Org. Chem. 6, 103 (1968).

90. L.S. Bartell, J. Chem. Phys. 32, 827 (1960).

91. G.N. Ramachandran, C.M. Venkatachalam and S. Krimm, Biophys. J. 6, 849 (1966).

92. M. Bixon and S. Lifson, Tetrahedron 23, 769 (1967).

93. G. Del Re, B. Pullman and T. Yonezawa, Biochem. Biophys. Acta 75, 153 (1963).

94. G. Del Re, J. Chem. Soc. 4031 (1958).

95. D.A. Brant and P.J. Flory, J. Am. Chem. Soc. 87, 2788 (1965).

96. M. Goodman, V. Chen, E. Benedetti, C. Pedone and P. Corradini, Biopolymers 11, 1779 (1972).

97. H.B. Vickery, J. Org. Chem. 28, 291 (1963).