Calculation of the Conformation of the Pentapeptide cyclo-(Glycylglycylglycylprolylprolyl). I. A Complete Energy Map[1,2]

Nobuhiro Gō[3] and Harold A. Scheraga[4]

Department of Chemistry, Cornell University, Ithaca, New York 14850.
Received December 4, 1969

ABSTRACT: The complete energy surface of cyclo-(glycylglycylglycylprolylprolyl) is computed. With rigid geometry (fixed bond lengths and bond angles, and planar amide groups), the condition of exact ring closure makes two dihedral angles (taken as the angles ψ between Pro-Pro and Pro-Gly) independent; the remaining six dihedral angles are computed as a function of the two independent ones. For given values of the two independent angles, there are from zero to eight sets of the other six angles which are compatible with the ring-closure condition. Therefore, the complete conformational energy surface may be represented by several sheets of two-dimensional contour diagrams. Thirteen local minima are found and located, in that region of the energy surface which is within 100 kcal of the global minimum. The relative order of the energies of the two lowest minima changes when a reasonable change is made in the energy parameters. The conformational energies of the "dipeptides" Pro-Pro and Pro-Gly are computed in order to show that the puckering of the pyrrolidine ring at the C^γ atom is important in determining the conformation of the cyclic pentapeptide. The computation of the statistical weights of the various local minima, and the problem of predicting the most stable conformation, will be treated in the next paper of this series.

I. Introduction

The most stable conformation of a molecule is the one which corresponds to that minimum of the conformational energy surface which has the largest statistical weight. The statistical weight of an arbitrary conformation *with small conformational fluctuations* is given by[5]

$$Z = \left(\frac{1}{\beta\hbar}\right)^m \left[\frac{1}{\det \mathbf{GF}}\right]^{1/2}_{Q=Q_1} \exp[-\beta F_1(Q_1)] \qquad (1)$$

as a good approximation, where m is the number of variable dihedral angles, $\beta = 1/kT$, $F_1(Q_1)$ is the value of the conformational energy[6-8] at a minimum point $Q = Q_1$, the elements of the matrix **G** are the coefficients of the kinetic energy of the molecule in the canonical expression for the Hamiltonian, and the elements of the matrix **F** are the second derivatives of the conformational energy function $F_1(Q)$ at the minimum point. The factor $[1/\det \mathbf{GF}]^{1/2}_{Q=Q_1}$ of eq 1 is the contribution to the statistical weight from the librational degrees of freedom in the molecule. The product $(\det \mathbf{F})^{-1/2} \exp[-\beta F_1(Q_1)]$ was calculated by Gibson and Scheraga[9] for 21 local minima of deca-L-alanine (obtained by an energy minimization procedure[10]), and by Gō, *et al.*,[11] for the right- and left-handed α-helical conformations of polyglycine and poly-L-alanine; the factor $(\det \mathbf{G})^{-1/2}$ was omitted from these computations. From these calculations,[9,11] it was found that the factor $\exp[-\beta F_1(Q_1)]$ was a major portion of Z, even though the factor $(\det \mathbf{F})^{-1/2}$ can play an important role in determining the relative stabilities of two different conformations whose energies $F_1(Q_1)$ differ by an amount of the order of 1 kcal/residue. The factor $(\det \mathbf{G})^{-1/2}$ was calculated for helical conformations of a polyethylene-type molecule,[5] and it was found that the contribution of this factor to Z was of about the same order as that of the factor $(\det \mathbf{F})^{-1/2}$. Therefore, the first step in the determination of the most stable conformation of a molecule is the location of all minima (within about 1 kcal/residue of the global minimum) on the conformational energy surface of the molecule. Then the contributions from the librational degrees of freedom to the statistical weight for each of these minima must be calculated. In the present paper, we are concerned with the first step; the second step will be considered in paper II of this series. Various energy minimization procedures have been used to locate minima on the energy surface,[10,12-15] and it has been found that there are many local minima on the energy surface of a complex molecule. The problem of locating the global minimum (and those within 1 kcal/residue of this minimum) is a nontrivial one, and is receiving much attention.[10,15]

(1) This work was supported by research grants from the National Science Foundation (GB-7160 and GB-7571X), from the National Institute of General Medical Sciences of the National Institutes of Health, U. S. Public Health Service (GM-14312), from the Eli Lilly, Hofmann-La Roche, and Smith Kline and French Grants committees, and from Walter and George Todd.

(2) Presented, in part, before the Division of Biological Chemistry at the 158th Meeting of the American Chemical Society, Sept 1969, New York, N. Y.

(3) On leave of absence from the Department of Physics, Faculty of Science, University of Tokyo, Tokyo, Japan, 1967–1970.

(4) To whom requests for reprints should be addressed.

(5) N. Gō and H. A. Scheraga, *J. Chem. Phys.*, **51**, 4751 (1969).

(6) When the effects of solvent are included in $F_1(Q)$, it is a free energy rather than an energy.[5,7,8]

(7) N. Gō, M. Gō, and H. A. Scheraga, *Proc. Nat. Acad. Sci., U. S.*, **59**, 1030 (1968).

(8) K. D. Gibson and H. A. Scheraga, *Physiol. Chem. Phys.*, **1**, 109 (1969).

(9) K. D. Gibson and H. A. Scheraga, *Proc. Nat. Acad. Sci., U. S.*, **63**, 242 (1969).

(10) K. D. Gibson and H. A. Scheraga, *Proc. Nat. Acad. Sci., U. S.*, **63**, 9 (1969).

(11) M. Gō, N. Gō, and H. A. Scheraga, *J. Chem. Phys.*, in press.

(12) K. D. Gibson and H. A. Scheraga, *Proc. Nat. Acad. Sci., U. S.*, **58**, 420 (1967).

(13) K. D. Gibson and H. A. Scheraga, *ibid.*, **58**, 1317 (1967).

(14) R. A. Scott, G. Vanderkooi, R. W. Tuttle, P. M. Shames, and H. A. Scheraga, *ibid.*, **58**, 2204 (1967).

(15) G. M. Crippen and H. A. Scheraga, *ibid.*, **64**, 42 (1969).

On the other hand, for a molecule like the cyclic pentapeptide

Gly-Gly-Gly-Pro-Pro

it is possible to calculate the complete energy surface, as asserted in the accompanying paper,[16] even though this molecule is fairly complex. The purpose of this paper is to present the procedures and results of this calculation. All local minima within 100 kcal of the global minimum were located on the complete conformational energy surface. This complete energy surface will serve as a useful model for future tests of procedures for passing from one minimum to another (lower) one. In the next paper[17] of this series, the factor $(\det \mathbf{GF})^{-1/2}$ for each local minimum reported here will be computed; thus, the contribution of the entropy of librational motion (together with the influence of water in solution, or of intermolecular interactions in crystals) in determining the most stable conformation of this cyclic pentapeptide will be considered.

II. Calculational Procedures

A. Geometry of Structural Units. Both the glycyl and prolyl residues are assumed to have fixed bond lengths and bond angles, and *trans*-peptide units. Glycyl residues are assigned standard Pauling–Corey geometry,[18] whereas the geometry of prolyl residues is taken from the crystallographic analysis of L-leucyl-L-prolylglycine;[19] these values of the bond lengths and bond angles in the backbone of the prolyl residue differ slightly from those of Pauling–Corey geometry. In using the raw data of Leung and Marsh,[19] we note that the C^γ atom has two possible positions lying on either side out of the plane of the pyrrolidine ring (defined by the four atoms C^β, C^α, N, and C^δ). We designate the conformations corresponding to these two positions of the C^γ atoms as Pro(I) and Pro(II), respectively, depending on whether or not the C^γ atom is on the same side of the plane of the pyrrolidine ring as the C′ atom of the same residue. The fact that the C^γ atom is out of the plane of the pyrrolidine ring turns out to be of essential importance for conformations of the cyclic pentapeptide under consideration here in which the atoms are very crowded (see below). Since puckering was observed only at the C^γ atom in X-ray structures,[19–21] no puckering was allowed at other positions of the pyrrolidine ring in the present calculations. While there is no *a priori* basis for excluding *cis*-peptide units from consideration, we have nevertheless assumed that all peptide units are in the *trans* conformation; in this respect, the present calculations may be incomplete. In the geometry of Leung and Marsh,[19] the dihedral angle ϕ is approximately 122°.

B. Generation of Conformations of Cyclic Structures.

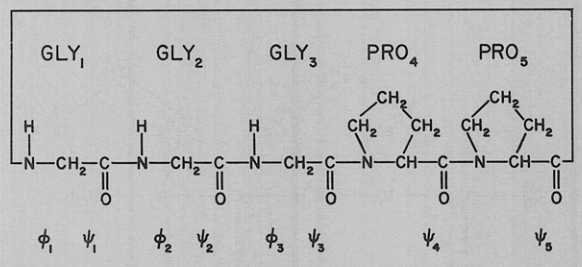

Figure 1. Definition of dihedral angles for rotation in the pentapeptide cyclo-(glycylglycylglycylprolylprolyl).

The five residues and eight dihedral angles for rotation about the backbone of the cyclic pentapeptide

Gly-Gly-Gly-Pro-Pro

are illustrated in Figure 1. Two of these eight dihedral angles, chosen as ψ_4 and ψ_5, are independent because of the condition of ring closure;[16] the other six are determined for a given set of ψ_4 and ψ_5 by solving a set of algebraic equations.[16] The number of solutions for the six dependent variables for a given set of values of ψ_4 and ψ_5 ranges from zero to eight, and is shown in Figure 2. It can be seen that ψ_4 and ψ_5 are restricted to the ranges $43° < \psi_4 < 137°$ and $36° < \psi_5 < 259°$, respectively.

C. Conformational Energy Functions and Parameters. The total conformational energy consists of four different kinds of contributions, *viz.*, torsional energies, nonbonded interactions, electrostatic interactions, and hydrogen bonding. The functions and parameters for the torsional energies and the nonbonded interactions are the same as those used by Ooi, *et al.*[22] The energy functions and parameters proposed by Poland and Scheraga[23] are used for electrostatic interactions and hydrogen bonding. In the treatment of Poland and Scheraga,[23] the parameters for the nonbonded interaction between pairs of atoms participating in the formation of a hydrogen bond are modified so as to account for the energy of the hydrogen bond. Therefore, in the breakdown of the total conformational energy into constituent terms in Table II, the energy of hydrogen bond formation is included inseparably partially in the nonbonded energy and partially in the electrostatic energy.

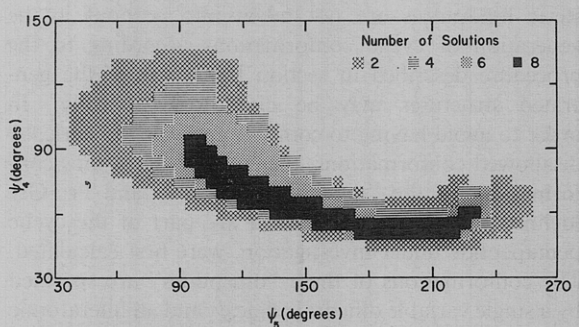

Figure 2. Number of solutions of a set of algebraic equations for ring closure.

(16) N. Gō and H. A. Scheraga, *Macromolecules*, **3**, 178 (1970).
(17) N. Gō, P. N. Lewis, and H. A. Scheraga, *ibid.*, to be submitted.
(18) See H. A. Scheraga, *Advan. Phys. Org. Chem.*, **6**, 103 (1968), for a summary of geometrical data.
(19) Y. C. Leung and R. E. Marsh, *Acta Cryst.*, **11**, 17 (1958).
(20) A. McL. Mathieson and H. K. Welsh, *ibid.*, **5**, 599 (1952).
(21) A. Rich and F. H. C. Crick, *J. Mol. Biol.*, **3**, 483 (1961).
(22) T. Ooi, R. A. Scott, G. Vanderkooi, and H. A. Scheraga, *J. Chem. Phys.*, **46**, 4410 (1967).
(23) D. Poland and H. A. Scheraga, *Biochemistry*, **6**, 3791 (1967).

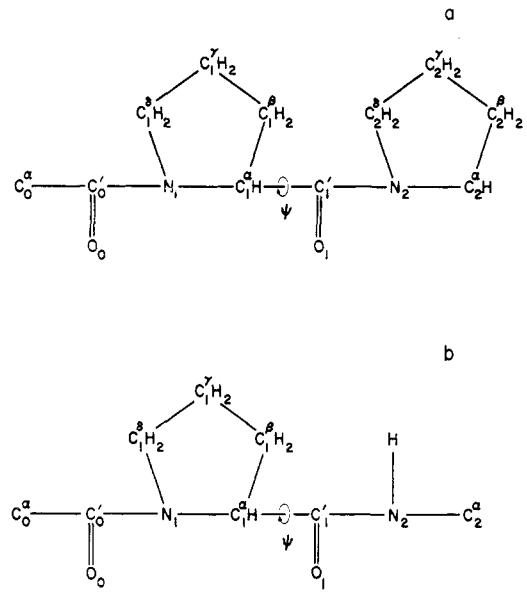

Figure 3. Definition of "dipeptides" (a) Pro-Pro and (b) Pro-Gly. The two hydrogen atoms bonded, for example, to the C_1^{β} atom are designated in the text and in Figure 4 as $H_1^{\beta 1}$ and $H_1^{\beta 2}$, respectively, depending on whether or not the hydrogen atom is at the opposite side of the pyrrolidine ring from the one at which the C′ atom of the same residue (in this case, C_1') is located.

Ooi, *et al.*,[22] used two different sets of values for the parameters of the repulsive potential of the nonbonded interaction, each of which corresponds to a hydrogen atom radius R_H of 1.20 and 1.275 Å, respectively. In the present paper, the set corresponding to R_H = 1.20 Å is used for most of the calculations. However, the other set is also used in some places in order to obtain information about how dependent the calculated results are on a particular choice of energy parameters. Among many other possible changes which could have been made in the parameters, the change of R_H is of particular interest, because (a) in this molecule where atoms are very crowded the steric effect is one of the most important factors in determining the conformation, and (b) being located in the outer portions of the peptide chain, the hydrogen atoms are the most susceptible to steric effects.

D. Calculation of Energies for Dipeptides. Since steric hindrances are not taken into account in the generation of cyclic conformations according to the procedure described in section IIB, some of the generated structures may be disallowed sterically. In order to avoid having to compute energies for sterically disallowed conformations, the energies of various conformations of the "dipeptides" Pro-Pro and Pro-Gly (defined in Figure 3a,b), which are part of the cyclic pentapeptide under investigation, were first calculated. The conformations of these "dipeptides" are specified by a single variable dihedral angle ψ, and all interatomic distances therein are either constant or a function of ψ only. If a particular conformation of these dipeptides is sterically disallowed, then any conformation of the cyclic pentapeptide having this dipeptide conformation is also sterically disallowed, and its energy need not be calculated. The conformational energies of Pro-Pro and Pro-Gly, for R_H = 1.20 Å, are given as functions of

TABLE I
CONTACT DISTANCE BETWEEN PAIRS OF ATOMS IN ÅNGSTRÖM UNITS

	H	C	N	O
H	1.86	2.11	2.11	2.10
C		2.77	2.70	2.55
N			2.50	2.50
O				2.48

ψ in Figures 4 and 5, respectively. Since each prolyl residue can assume the Pro(I) or Pro(II) geometry, the "dipeptides" Pro-Pro and Pro-Gly can assume four and two, respectively, different combinations of side-chain geometry. In Figures 4 and 5, we have also indicated the pairs of atoms and the range of ψ for each such pair for which the interatomic distance is smaller than the selected contact distance given in Table I. The values of Table I were selected on the basis of the following consideration. In the steric diagrams of Ramachandran, *et al.*,[24] and Leach, *et al.*,[25] the two regions in which the β structure and the right-handed α-helical structure, respectively, occur are separated by an intervening disallowed region. However, from X-ray studies of globular proteins, it appears that a fairly large number of residues lie in this disallowed region.[26] Therefore, the values of Table I, which are smaller than those used by Ramachandran, *et al.*,[24] and by Leach, *et al.*,[25] were selected so that the bridging region would become sterically allowed.[27,28]

Consider first the conformations of the "dipeptide" Pro-Pro, whose conformational energies and interatomic distances are given in Figure 4. Variation in the position of the C^{γ} atom does not alter the allowed region in the range between ψ = 280 and 360° significantly, but does influence the allowed region in the range between ψ ~95° and ~175° very strongly. Energy minima in the latter range are assigned the serial numbers I to VI in Figure 4; all minima in this range are higher than those in the range between ψ = 280 and 360°. Therefore, in general, conformations corresponding to minima I to VI would be less stable than those in the range between ψ = 280 and 360°. However, as shown in section IIB, the dihedral angle ψ_4 between prolyl residues in the cyclic pentapeptide under consideration (Figure 1) can assume values only in the range between 43 and 137° (independent of the energy of the corresponding conformation) because of the geometrical condition of ring clos-

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

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

(26) G. N. Ramachandran and V. Sasisekharan, *Advan. Protein Chem.*, **23**, 283 (1968).

(27) When hard-sphere potentials are employed, the bridging region is disallowed for the contact distances used by Ramachandran, *et al.*,[24] and Leach, *et al.*[25] This bridging region can be made allowed either by admitting shorter contact distances (retaining the hard-sphere potential) or by using a Lennard-Jones nonbonded potential (with some softness in the repulsive part) or by introducing flexibility of bond lengths and bond angles.[28] The small contact distances used here should be regarded as effective ones to account for the softness of the repulsive potential and/or the flexibility of bond lengths and bond angles. These contact distances are used only as criteria for *large* steric overlaps; their use does *not* imply that the energy parameters discussed in section IIC have been altered.

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

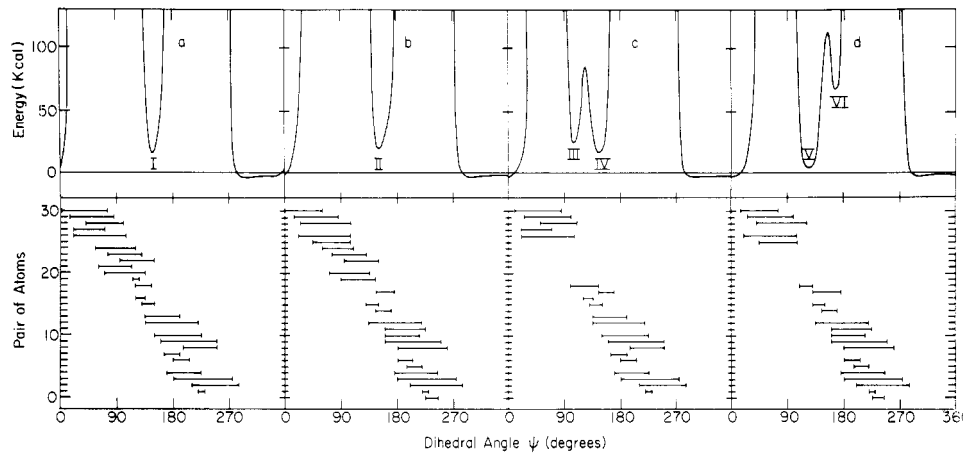

Figure 4. Conformational energy and steric hindrances in "dipeptide" (a) Pro(I)-Pro(I), (b) Pro(I)-Pro(II), (c) Pro(II)-Pro(I), and (d) Pro(II)-Pro(II). Serial numbers I–VI are assigned to each minimum in the range of $90° < \psi < 180°$. Serial numbers assigned to pairs of atoms refer to the following: 0. O_0–C_2^{γ}; 1. O_0–N_2; 2. O_0–$H_2^{\delta 1}$; 3. O_0–C_2^{δ}; 4. O_0–$H_2^{\delta 2}$; 5. $C^{0\alpha}$–$H_2^{\delta 1}$; 6. C_0^{α}–C_2^{δ}; 7. C_0^{α}–$H_2^{\delta 2}$; 8. C_0'–$H_2^{\delta 1}$; 9. C_0'–C_1^{δ}; 10. C_0'–$H_2^{\delta 2}$; 11. N_1–$H_2^{\delta 1}$; 12. N_1–C_2^{δ}; 13. N_1–$H_2^{\delta 2}$; 14. C_1^{δ}–$H_2^{\delta 1}$; 15. C_1^{δ}–C_2^{δ}; 16. C_1^{δ}–$H_2^{\delta 2}$; 17. $H_1^{\delta 2}$–$H_2^{\delta 1}$; 18. $H_1^{\delta 2}$– $H_2^{\delta 2}$; 19. C_1^{γ}–$H_2^{\delta 1}$; 20. C_1^{γ}–C_2^{δ}; 21. C_1^{γ}–$H_2^{\delta 2}$; 22. $H_1^{\gamma 2}$–$H_2^{\delta 1}$; 23. $H_1^{\gamma 2}$–C_2^{δ}; 24. $H_1^{\gamma 2}$–$H_2^{\delta 2}$; 25. C_1^{β}–$H_2^{\delta 1}$; 26. C_1^{β}–C_2^{δ}; 27. C_1^{β}–$H_2^{\delta 2}$; 28. $H_1^{\beta 2}$–$H_2^{\delta 1}$; 29. $H_1^{\beta 2}$–C_2^{δ}; 30. $H_1^{\beta 2}$–$H_2^{\delta 2}$. A pair of atoms is sterically hindered in the range indicated.

ure. Hence, the question arises as to whether the slight steric hindrances in conformations I to VI (see below and Figure 4) are weak enough to enable this cyclic pentapeptide to be synthesized. Since the energy of the conformation at, say, minimum V is higher than the one in the range of ψ between 280 and 360° by about 8 kcal/mole, we shall assume that the cyclic pentapeptide *can* be synthesized, especially since this energy difference would be expected to be smaller if flexibility of bond lengths and bond angles were considered. Hence, we focus our attention on the dipeptide in the range of ψ between 43 and 137°, where the conformational energy depends markedly on the position of the C^{γ} atom because of the steric hindrances indicated in the lower portion of Figure 4.

The energy barrier between minima III and IV in Pro(II)-Pro(I) arises mainly from steric overlaps between $H_1^{\delta 2}$ and $H_2^{\delta 2}$ (pair 18) and C_1^{δ} and $H_2^{\delta 2}$ (pair 16), as indicated in Figure 4c. The energy barrier between minima V and VI in Pro(II)-Pro(II) arises mainly from steric overlaps between $H_1^{\delta 2}$ and $H_2^{\delta 1}$ (pair 17) and C_1^{δ} and $H_2^{\delta 1}$ (pair 14), as indicated in Figure 4d. In conformations corresponding to minima IV and VI, the $(CH_2)_2^{\delta}$ group lies in a shallow hole formed by the $(CH_2)_1^{\delta}$ group and the group of atoms N_1, C_0', O_0 and C_0^{α} (see Figures 3a, 4c, 4d). In conformations corresponding to minima III and V, the $(CH_2)_2^{\delta}$ group lies in a shallow hole formed by the $(CH_2)_1^{\beta}$ and $(CH_2)_1^{\delta}$ groups (see Figures 3a, 4c, 4d). The pyrrolidine ring *must* be puckered at the C^{γ} atom in order for this latter shallow hole to exist. In Pro(II), the C_1^{γ} atom is on the opposite side of the plane of the pyrrolidine ring from the $(CH_2)_2^{\delta}$ group; the latter is accommodated in the shallow hole formed by removal of the C_1^{γ} atom from the plane of the pyrrolidine ring by puckering. However, in Pro(I), the C_1^{γ} atom is on the same side of the plane of the pyrrolidine ring as the $(CH_2)_2^{\delta}$ group; the latter then encounters the C_1^{γ} atom which is effectively a bump on the same side of the pyrrolidine ring (and, in fact, prevents the occurrence of minima in Figures 4a and 4b corresponding to minima III and V of Figures 4c and 4d). We thus see that the puckering of the pyrrolidine ring plays a very important role in influencing the conformational energy of the "dipeptide" Pro-Pro in the range of ψ between 95 and 175°. This accounts for the fact that Schimmel and Flory[29] (who assumed that the pyrrolidine ring is planar) concluded that this range of ψ is disallowed, while Leach, *et al.*[25] (who assumed that the pyrrolidine ring is puckered), found that the range of ψ between 130 and 150° is allowed.

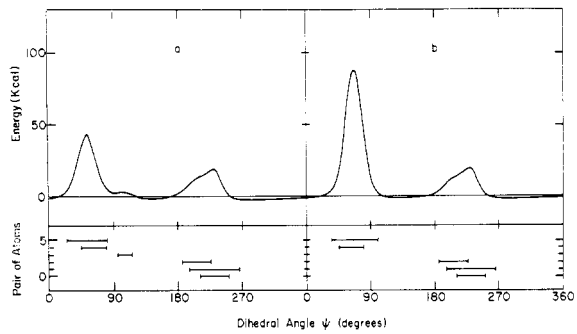

Figure 5. Conformational energy and steric hindrances in "dipeptide" (a) Pro(I)-Gly, and (b) Pro(II)-Gly. Serial numbers assigned to pairs of atoms refer to the following: 0, O_0–N_2; 1, O_0–H_2^N; 2. C_0'–H_2^N; 3, $H_1^{\gamma 2}$–H_2^N; 4. C_1^{β}–N_2; 5. $H_1^{\beta 2}$–H_2^N.

The calculations of the conformational energies of the "dipeptide" Pro-Pro were repeated using energy parameters corresponding to R_H = 1.275 Å, without observing any appreciable changes in the positions of the energy minima and maxima from those of Figure 4; also, the energy values in the range of ψ between 280 and 360° did not change appreciably. The values at minima I to VI increased by amounts which are roughly proportional to the energy value of each minimum (measured from the value of the minimum in the range of ψ between 280 and 360°), the proportionality constant being $^2/_3$

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

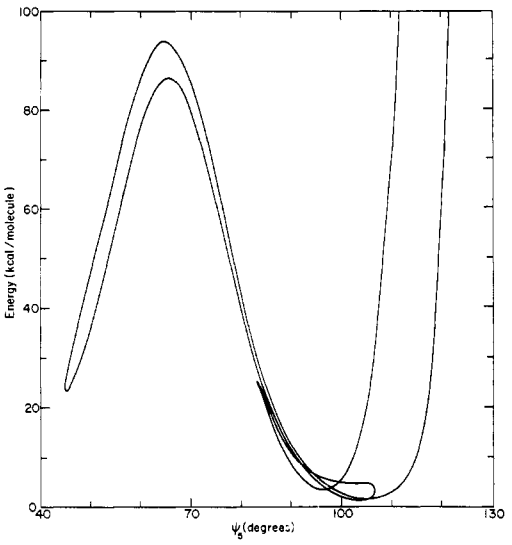

Figure 6. Conformational energy of the pentapeptide cyclo(Gly-Gly-Gly-Pro-Pro) with Pro_4(II) and Pro_5(II) conformations, for ψ_4 fixed at 120° and ψ_5 varied. The multiplicity of conformational energies corresponds to the multiplicity of possible cyclic conformations of the pentapeptide for the given values of ψ_4 and ψ_5.

to $^1/_4$. For example, minimum V was found to be 13 kcal/mole higher than the minimum in the range 280° < ψ < 360°, *i.e.*, there was an increase of 5 kcal when R_H was changed from 1.20 to 1.275 Å. We assume that this difference is still small enough to assure that the cyclic pentapeptide can be synthesized. In brief, the positions and relative order of the energy minima in the "dipeptide" are found to be stable for the change in the value of R_H examined here.

In section III, where we will discuss the conformational energy of the cyclic pentapeptide, we will not consider sterically hindered conformations with energies larger than 100 kcal/mole of the pentapeptide. Since we can certainly expect the energy of the pentapeptide to be greater than 100 kcal if the "dipeptide" Pro-Pro portion has an energy greater than 150 kcal, we will vary ψ_4 only in the range where the energy of the "dipeptide" Pro-Pro is less than 150 kcal. Combining this energy restriction on ψ_4 with the ring-closure restriction, we may omit from consideration side-chain conformations Pro_4(I)-Pro_5(I) and Pro_4(I)-Pro_5(II) (Figures 4a and 4b, respectively), and consider Pro_4(II)-Pro_5(I) and Pro_4(II)-Pro_5(II) only in the range of 94° ≤ ψ_4 ≤137° and 104° ≤ ψ_4 ≤ 137°, respectively.

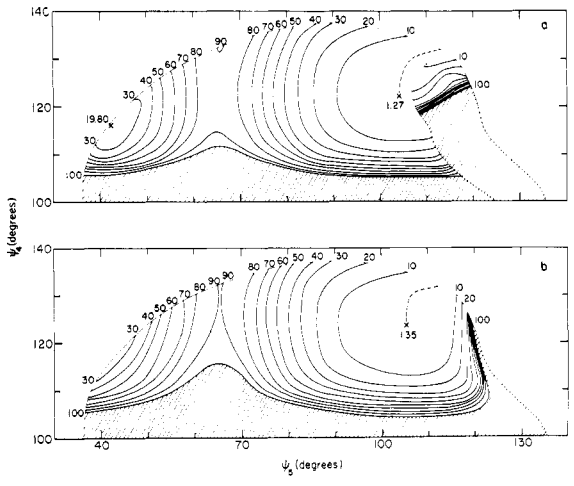

Figure 7. Two of the contour diagrams showing the conformational energy of the pentapeptide cyclo(Gly-Gly-Gly-Pro-Pro), with Pro_4(II) and Pro_5(II) conformations, as a function of the independent dihedral angles ψ_4 and ψ_5. Energies are shown in kilocalories per molecule. The three energy minima in these two diagrams are marked by an ×, and their energy values are given. A dotted line shows the easiest path to go from the minimum with the least energy (1.27 kcal) to the one with the next least energy (1.35 kcal) in the entire energy surface. The shaded region has energies higher than 100 kcal. See text for further description.

We turn next to a discussion of the conformations of the "dipeptide" Pro-Gly given in Figure 5. The conformational energy depends on the geometry of the pyrrolidine ring only in the range of ψ between 20 and ~120°, where the $(NH)_2$ group overlaps the $(CH_2)_1^\beta$ group [in both Pro(I)-Gly and Pro(II)-Gly] and the $H_1^{\gamma 2}$ atom [in Pro(I)-Gly only]. Since the energy never exceeds 150 kcal, we will consider any value of ψ_5 in the pentapeptide which satisfies the ring-closure condition.

III. Results and Discussion

Cyclic conformations were generated for values of the independent variables ψ_4 and ψ_5 indicated at the end of the previous section, and the corresponding conformational energies were calculated. It is a bit difficult to display the results because of the multiplicity of solutions for any given set of values of the independent variables. Figure 6 shows the conformational energy [with Pro_4(II) and Pro_5(II) conformations] as a function of ψ_5, with ψ_4 kept fixed at 120°. In the range ψ_5 ≅ 46–83° and ψ_5 ≅ 107–122°, two cyclic conformations are possible for the given values of ψ_5 (and ψ_4 = 120°); hence, the energy is double-valued in these ranges. In the range ψ_5 ≅ 83–107°, four cyclic conformations are possible for the given values of ψ_5 (and ψ_4 = 120°), and there are therefore four corresponding energy values. Since the curve in Figure 6 is actually a single continuous line, it is possible to change conformations (whose energies are given by the curve) continuously from one to another by changing ψ_5, with ψ_4 fixed at 120°, without breaking, stretching, or bending chemical bonds; of course, as ψ_5 changes, there are accompanying changes in ϕ_1, ψ_1, ϕ_2, ψ_2, ϕ_2, and ψ_3.

Because of the multiplicity of conformations (and corresponding conformational energies) for given values of the independent variables, we need more than one sheet of paper to represent the complete energy surface by a two-dimensional (corresponding to the two independent variables ψ_4 and ψ_5) contour diagram. Two of such two-dimensional contour diagrams are shown in Figure 7. The energies along the line at ψ_4 = 120° in Figures 7a and 7b correspond to *part* of the curve of Figure 6. Figure 7a corresponds to two branches of the curve in Figure 6, one of which is the lower one in the range of ψ_5 ≅ 46–83° and continues up to ψ_5 ≅ 107° where it meets another branch and terminates (*i.e.*, turns toward lower values of ψ_5), and the other is the upper branch in the range of ψ_5 ≅ 107–122°. The termination of the former branch at ψ_5 ~ 107° is ex-

TABLE II
CHARACTERIZATION OF THIRTEEN LOCAL ENERGY MINIMA FOR R_H = 1.20 Å

	A	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*A′*	*B′*	*C′*	*D′*	*E′*
Pro_4[a]	II	II	II	II	II	II	II	II	II	II	II	II	II
Pro_5	I	I	I	I	I	I	I	I	II	II	II	II	II
ϕ_1[b]	147.15	72.22	99.72	73.53	103.03	43.81	52.17	73.38	57.60	135.30	64.67	47.69	85.81
ψ_1	124.34	148.23	76.31	298.24	46.27	311.16	115.65	109.60	305.06	126.28	105.13	127.69	63.13
ϕ_2	96.01	74.73	100.49	267.25	124.01	251.17	133.68	112.84	263.71	97.66	99.26	114.37	127.05
ψ_2	63.08	238.11	249.53	121.13	156.03	185.67	148.88	132.17	141.28	74.02	222.05	160.77	150.92
ϕ_3	129.42	290.71	325.61	82.59	81.56	9.74	9.73	75.15	58.39	122.04	339.06	13.25	78.93
ψ_3	81.81	88.07	60.11	62.86	55.02	59.55	102.38	54.40	60.98	75.31	56.17	81.32	46.04
ψ_4	106.32	106.30	106.52	106.22	106.46	105.58	106.42	136.92	109.02	116.22	124.88	122.50	123.50
ψ_5	37.30	86.50	90.86	93.70	110.74	110.82	111.40	91.26	101.72	42.16	98.66	104.06	105.36
CO_1[c]	↓	↓	↓	↑	↓	↑	↓	↓	↑	↓	↓	←	↓
CO_2	↓	↑	↑	↓	↙	←	←	←	↓	↓	↑	←	←
E_{tot}[d]	42.52	20.41	18.93	19.08	23.40	21.90	20.82	27.40	33.40	19.80	2.09	1.27	1.35
E_{tor}	−0.39	0.85	−0.23	0.29	−0.38	−0.42	0.13	0.29	0.48	−0.61	0.18	−0.11	−0.06
E_{nb}	30.63	12.85	12.20	12.72	13.41	16.75	12.03	17.41	27.26	8.42	−4.60	−7.07	−8.53
E_{es}	12.28	6.71	6.95	6.08	10.37	5.57	8.67	9.70	5.66	11.99	6.51	8.45	9.94

[a] Geometry of pyrrolidine ring in residues Pro_4 and Pro_5, respectively. [b] Dihedral angles are given in degrees. [c] Rough description of orientation of CO bond in residues Gly_1 and Gly_2 relative to those in residue Gly_3, Pro_4, and Pro_5 which are roughly parallel to each other: ↓ parallel, ↙ almost parallel but with O atom being slightly away from the ring, ←perpendicular with O atom away from the ring, ↑ antiparallel. [d] Energies are given in kilocalories/molecule. E_{tot}, total energy; E_{tor}, torsional energy; E_{nb}, nonbonded interaction; E_{es}, electrostatic interaction.

pressed by the cut (which passes through the point $\psi_5 \cong 107°$ for $\psi_4 = 120°$) in Figure 7a. The cut terminates at $\psi_4 \cong 128°$; this means that the two branches mentioned above become a single continuous branch for $\psi_4 \gtrsim 128°$. Figure 7b corresponds to the branch of the curve in Figure 6 which is the upper one in the range of $\psi_5 \cong 46$–$83°$ and continues through the range of $\psi_5 \cong 83$–$107°$ to become the lower one in the range of $\psi_5 \cong 107$–$122°$. At the upper boundaries in Figures 7a and 7b, the two conformations whose energies are expressed by the two contour diagrams, respectively, become identical with each other; no cyclic conformations are possible outside of this boundary. Three energy minima (indicated by ×'s) are found to exist in that part of the energy surface shown in Figure 7. A complete energy map, consisting of 14 such sheets of contour diagrams, was computed for the region where the energy does not exceed 100 kcal. Thirteen minima were found and located in this region of the energy surface, eight for cyclic pentapeptides with Pro_4(II) and Pro_5(I) conformations and five for those with Pro_4(II) and Pro_5(II) conformations. The characteristic parameters for these thirteen minimum-energy conformations are given in Table II.

In conformations A–G, ψ_4 has the value corresponding to energy minimum III in Figure 4, whereas in conformations H and A′–E′, ψ_4 corresponds to energy minima IV and V, respectively, in Figure 4. In all conformations except A and B′, ψ_5 has values corresponding to the right tail of the peak of the energy curve in the range of $\psi \cong 30$–$90°$ (in Figures 5a and 5b), whereas, in conformations A and B′, ψ_5 has values corresponding to the left tail of this peak. In all conformations other than A and B′, $\phi_4 = \phi_5 \cong 122°$ (the value given by Leung and Marsh[19]), and ψ_4 and ψ_5 have values corresponding to the region of the right-handed α-helix. As a result, the CO bonds in residues Gly_3, Pro_4, and Pro_5 are roughly parallel to each other as they are in the α-helix. However, in conformations A and B′, the oxygen atom of the CO bond in residue Pro_5 points somewhat toward the "inside" of the ring structure, while the CO bonds in residues Pro_4 and Pro_5 are parallel to each other. The conformations of the other parts of the cyclic molecule, *i.e.*, of Gly_1 and Gly_2, can be described roughly by the directions of their CO bonds (given in Table II). The obvious relation between the relative alignment of the amide dipoles and the electrostatic energy is indicated in Table II, *viz.*, in conformations A, E and B′, in which five dipoles are nearly parallel and side by side, the electrostatic energy is high, whereas, in conformations B, C, D, A′, C′, in which one of the dipoles assumes an antiparallel position, the electrostatic energy is relatively low, etc. Even though the contribution from the electrostatic interactions to the relative stabilities of the conformations corresponding to the local minima is not negligible, the nonbonded interactions play the most dominant role in this molecule. Thus, the three conformations of lowest energy, C′, D′, and E′, have the lowest nonbonded energy (least steric overlap). Hydrogen bonds are formed only in conformations C and C′, between O_1 and H_3^N in both conformations.

In order to examine the effects of changing R_H, the calculations for the three conformations of lowest energy, C′, D′ and E′, were repeated with R_H = 1.275 Å, and the results are shown in Table III. Since the dihedral angles did not change much, for this change in R_H, the rough description of the conformations in terms of the directions of CO_1 and CO_2, and E_{tor} and E_{es} are not changed much. However, large changes are observed for E_{nb}, and hence E_{tot}. In particular, it should be noted that the relative order of the energies of these three conformations changes for this change in R_H, *viz.*, for R_H = 1.20 Å, conformation *D′* is 0.08 kcal lower in energy than conformation *E′*, whereas, for R_H = 1.275 Å, conformation *E′* is 1.07 kcal lower in energy than conformation *D′*. The implications of this result for the problem of the prediction of the most stable conformation will be discussed in paper II, in which the contributions from the librational degrees of

TABLE III
CHARACTERIZATION OF THREE LOWEST ENERGY MINIMUM CONFORMATIONS FOR R_H = 1.275 Å

	C′	D′	E′
Pro_4	II	II	II
Pro_5	II	II	II
φ_1	66.27	47.10	84.37
ψ_1	104.09	127.42	63.37
φ_2	100.07	113.63	127.43
ψ_2	221.33	162.11	151.55
φ_3	340.17	13.15	77.75
ψ_3	55.83	79.77	45.89
ψ_4	124.96	123.36	123.88
ψ_5	99.28	104.30	105.90
CO_1	↓	←	↓
CO_2	↑	←	←
E_{tot}	11.42	10.36	9.29
E_{tor}	0.16	−0.12	−0.21
E_{nb}	4.76	2.07	−0.57
E_{es}	6.51	8.41	9.88

freedom to the statistical weights will also be discussed.

The complete energy surface could be calculated for the pentapeptide considered here because it is cyclic and it contains only eight variable dihedral angles in the backbone. For more complex molecules, it is impossible to calculate the complete energy surface, and other methods than those used here must be employed to find local minima within a reasonable energy range of the global minimum. The complete energy surface obtained in the present calculation can now be used to test the efficiency of such methods.[29a]

(29a) NOTE ADDED IN PROOF. A reviewer has raised the point that the introduction of the possibility for flexibility of bond angles and bond lengths would greatly expand the domain of the ψ_1, ψ_5 space within which ring closure could be effected. This is correct, even though we have no quantitative information as to how the domain would be expanded. As was discussed elsewhere,[3] the conformation of a flexible molecule can be treated by a two-step procedure, the first step being the minimization of the conformational energy of the rigid molecule having fixed standard bond lengths and bond angles, with the second one being the energy minimization in a larger space corresponding to the flexibility of bond lengths and bond angles *starting* from the conformation obtained by the first step. In this paper, we are concerned with the first step.

Conformational Studies on Polypeptides. Circular Dichroism Properties of Random Copolymers of Lysine and Phenylalanine in Aqueous Solutions at Various pH Values

E. Peggion, A. S. Verdini, A. Cosani, and E. Scoffone

Institute of Organic Chemistry, University of Padova, 35100 Padova, Italy.
Received November 17, 1969

ABSTRACT: Circular dichroism (CD) measurements have been performed on random copolymers of lysine and phenylalanine in 0.1 *M* KCl at various pH values. On increasing the pH, copolymers containing 3:1 and 1:1 molar ratios of lysine to phenylalanine undergo a conformational transition from a random coil to an ordered structure. The shape of the CD patterns, which are affected by the presence of side-chain aromatic chromophores, suggest that the ordered forms do correspond to β structures. In the limits of the examined copolymer compositions, the higher is the phenylalanine content of the copolymers, the higher is the stability of the β form in alkaline aqueous solutions.

It is well known that the overall stability of any conformation of natural and synthetic polypeptides is due to contributions from different forces like inter- or intramolecular hydrogen bonds, hydrophobic bonds, van der Waals interactions, dipole–dipole interactions, and so on.[1]

The relative importance of such forces are strongly dependent on the nature of the solvent. Thus, in aqueous solutions, hydrophobic bonds among nonpolar side chains are mostly responsible for the tendency of peptides and proteins to assume definite conformations.[2]

In organic solvents like 1,2-dichloroethane, chloroform, etc., hydrogen bonds are suggested to play an important role in stabilizing ordered conformations. Strong interacting organic solvents like dichloroacetic or trifluoroacetic acids compete successfully with amide groups of peptides in the formation of hydrogen bonds, and then in such media occurrence of ordered structures is generally prevented.

In a previous investigation[3] we studied the conformational properties of random copolymers of N^{ϵ}-carbobenzoxy-L-lysine (Z-Lys) and L-phenylalanine (Phe) in tetrahydrofuran as the solvent; it was found that increasing proportions of aromatic residues apparently do not perturb the α-helical conformation of poly-N^{ϵ}-carbobenzoxy-L-lysine (PCBL). However the presence of aromatic residues in the copolymer weakened the conformational stability of the α-helical form, owing to steric interference between side chains and peptide

(1) For a comprehensive review on this topic, see "Structure and Stability of Biological Macromolecules," G. N. Timasheff and G. D. Fasman, Ed., Marcel Dekker, Inc., New York, N. Y., 1969.

(2) H. A. Scheraga in "The Proteins," Vol. 1, 2nd ed, H. Neurath, Ed., Academic Press, New York, N. Y., 1963.

(3) E. Peggion, A. S. Verdini, A. Cosani, and E. Scoffone, *Macromolecules*, 2, 170 (1969).