Determination of Intermolecular Potentials from Crystal Data. II. Crystal Packing with Applications to Poly(amino acids)

ARTICLE *in* MACROMOLECULES · JANUARY 1971

Impact Factor: 5.8 · DOI: 10.1021/ma60019a025

9 AUTHORS, INCLUDING:

Frank A Momany
United States Department of Agriculture
122 PUBLICATIONS **7,000** CITATIONS

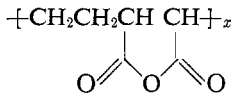
SEE PROFILE

Gordon Crippen
University of Michigan
150 PUBLICATIONS **5,810** CITATIONS

SEE PROFILE

most of the carbons in both the concentrated and dilute gels are contributing to the observed nmr signals.

Figure 2a shows the noise-decoupled, natural abundance ^{13}C nmr spectrum of the carbonyl region of a solution of 20% (by weight) ethylene–maleic anhydride (A) in dimethylformamide. Only a single, narrow line



A

is observed corresponding to the single type of carbonyl carbon in the completely alternating copolymer. If this material is cross-linked through the ring by a difunctional monomer, it is no longer soluble in dimethylformamide (in any concentration) but rather forms a swollen gel or slurry which remains visibly inhomogeneous despite heating to elevated temperatures. Nevertheless, a narrow carbonyl line is still observed from the gel, containing 20% (by weight) polymer and 80% dimethylformamide (Figure 2b). The level of signal intensity obtained from the gel is comparable to that obtained from the true solution. (The new lines at low field are probably the result of rearrangements during the heat treatment, but may contain information about the cross-linking sites.)

Apparently the segmental chain motion in these gels is sufficient to eliminate much of the dipolar interaction responsible for the broadening of the nmr spectra of polymers, at least to the extent that high-resolution ^{13}C nmr spectra can be obtained. This is true even for polymers which are not rubbery and are, in fact, characterized by a modest degree of crystallinity. Proton nmr spectra of the gels provided no usable, high-resolution signal.

Determination of Intermolecular Potentials from Crystal Data. II. Crystal Packing with Applications to Poly(amino acids)[1]

R. F. McGuire,[2a] G. Vanderkooi,[2b] F. A. Momany,[2c] R. T. Ingwall, G. M. Crippen,[2d] N. Lotan,[2e] R. W. Tuttle, K. L. Kashuba, and H. A. Scheraga*[2f]

Department of Chemistry, Cornell University, Ithaca, New York 14850.
Received July 27, 1970

ABSTRACT: A technique used to study the crystal packing of both small molecules and biopolymers is presented, and the development of an internally consistent set of intermolecular potentials is described. Using this set of potentials for all intermolecular interactions together with a previously reported set of intramolecular potentials, the conformation and crystal packing energies of the β structure of poly-L-alanine and the ω-helix of poly(β-benzyl-L-aspartate) were calculated. The internal torsional energies, as well as both the intramolecular and intermolecular nonbonded, hydrogen-bonded, and electrostatic energy contributions, were included in these computations. Energy minimization was carried out with respect to both the intramolecular variables (*i.e.*, the backbone and side-chain dihedral angles) and also variations in the intermolecular orientations and crystal packing of the homopolymers. In general, the conformation and crystal packing of both polymers, predicted from the computed energies, were in agreement with the reported crystallographic results. For poly-L-alanine, the observed packing, as calculated here, is caused primarily by intermolecular interactions, *viz.*, the intermolecular hydrogen bonds and the nonbonded side-chain interactions, with very little influence from the intramolecular interactions. Not only were the conformation and the lattice constants in good agreement with results from crystal studies, but, further, from a comparison of the calculated energies for a number of "similar" crystal structures, a statistical model as proposed from the X-ray diffraction studies was verified. For the packing of the ω-helix of poly(β-benzyl-L-aspartate), the major contribution to the total interaction energy arises from the intramolecular interactions, although the observed conformation and packing results from contributions from both the intermolecular and intramolecular interactions.

It is necessary to have accurate potential functions for calculating conformational energies of macromolecules.[3,4] For this purpose, much effort has been devoted in recent years[5–9] to calculating the forces

(1) This work was supported by research grants from the National Science Foundation (GB-7571X and GB-17388), 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, Hoffmann-LaRoche, and Smith Kline and French Grants Committees, and from Walter and George Todd.

(2) (a) NIH Postdoctoral Trainee, 1968–1969; Postdoctoral Fellow of the National Institute of General Medical Sciences, National Institutes of Health, 1969–1971; (b) Postdoctoral Fellow of the National Institute of General Medical Sciences, National Institutes of Health, 1965–1968; (c) Special Fellow of the National Institute of General Medical Sciences, National Institutes of Health, 1968–1969; (d) Predoctoral Fellow of the National Institute of General Medical Sciences, National Institutes of Health, 1967–1971; (e) on leave from the Weizmann Institute of Science, 1966–1969; (f) to whom requests for reprints should be addressed.

(3) H. A. Scheraga, *Advan. Phys. Org. Chem.*, **6**, 103 (1968).

(4) H. A. Scheraga, Nobel Symposium 11 on "Symmetry and Function of Biological Systems at the Macromolecular Level," Stockholm, Aug 26–29, 1968, A. Engstrom and B. Strandberg, Ed., Almqvist and Wiksell, Stockholm, 1969, p 43.

(5) (a) A. I. Kitaigorodskii, *Acta Crystallogr.*, **18**, 585 (1965); *J. Chim. Phys. Physicochim. Biol.*, **63**, 9 (1966); (b) A. I. Kitaigorodskii and K. V. Mirskaya, *Sov. Phys.-Crystallogr.*, **6**, 408 (1962); **9**, 137 (1964).

(6) (a) D. E. Williams, *Science*, **147**, 605 (1965); (b) D. E. Williams, *J. Chem. Phys.*, **45**, 3770 (1966); **47**, 4680 (1967); (c) D. E. Williams, *Acta Crystallogr., Sect. A*, **25**, 464 (1969).

(7) D. P. Craig, R. Mason, P. Pauling, and D. P. Santry, *Proc. Roy. Soc., Ser. A*, **286**, 98 (1965).

(8) F. A. Momany, G. Vanderkooi, and H. A. Scheraga *Proc. Nat. Acad. Sci. U. S.*, **61**, 429 (1968), paper I of this series,

(9) A. Warshel and S Lifson, *J. Chem. Phys.*, **53**, 582 (1970.

involved in determining the conformation and packing of small molecules in crystals. Thus far, most attention has been devoted to the packing of nonpolar molecules,[5-9] using simple empirical equations to represent central forces, which are considered to be most important in determining conformation and packing arrangements. However, for use in conformational studies of polypeptides, other interactions (including electrostatic forces and other forces of a quantum mechanical nature such as hydrogen bonding[10,11]) must be included.

In this paper, we present a technique to study the crystal packing of small (nonpolar and polar) molecules as well as the intramolecular and intermolecular conformation and packing of biopolymers. This procedure was used earlier[4] with a preliminary set of intermolecular potential functions, to compute the lattice constants of crystals of small molecules. This same procedure is applied here together with these intermolecular potential functions and the intramolecular potential functions of Scott and Scheraga,[12] Ooi, *et al.*,[13] and Yan, *et al.*,[14] to compute the conformation and packing of the β structure of poly-L-alanine and the ω-helix of poly(β-benzyl-L-aspartate); the results are compared with X-ray diffraction data on fibers of these poly(amino acids).

Theory

A. General. In the first paper of this series,[8] we showed how to obtain the parameters of empirical potential functions for pairwise intermolecular interactions by application of the equations of static equilibrium (together with heat of sublimation data) to the single-crystal coordinates and cell parameters found from high-resolution X-ray diffraction data. The potentials used were of the Lennard-Jones 6–12 form, and the parameters were obtained for the nonpolar molecule benzene. The crystal structures at three different temperatures were studied, and a preliminary analysis of thermal vibrational effects was given. Further work has expanded our knowledge of the symmetry conditions and packing criteria of atom-centered spherically symmetric potentials, and this information will be reported elsewhere.[15]

A different procedure was used in another paper[4] to obtain a preliminary set of intermolecular potential functions. The parameters of individual pairwise potentials for each atom type found in hydrocarbon crystals such as pentane, hexane, etc., were varied until a least-squares best fit of all computed independent crystal lattice constants and sublimation energies to the corresponding experimental values was obtained. This initial set of parameters was then expanded to include other types of interactions (*viz.*, electrostatic and hydrogen bonding) by using crystal data from heteroatomic molecules, gradually increasing the data set, until an internally self-consistent set of intermolecular parameters of interest for the packing of polypeptides was obtained. The complete parameter set was tested by determining the lattice constants of 21 different molecules of known crystal structure, at each structure of minimized energy, and comparing the computed lattice constants with those obtained experimentally.[4] The standard deviations reported[4] for the 21 molecules were within the uncertainty associated with thermal motion (*i.e.*, differences in vibrational and librational terms for the different molecules).

The procedure for treating packed molecules in a crystal allows for the generation of the coordinates of symmetry-related molecules from the coordinates of a given reference molecule by the symmetry operators corresponding to the space group of the crystal. In turn, the conformation and position of the reference molecule may be changed by internal rotations about its bonds and by external rigid-body motions. Thus, the energy may be computed for an assembly of molecules in a crystal in any given conformation and relative positions. The equations for the symmetry operators, transformation of coordinate systems, and translation and rotation in a unit cell are all similar to those used by Williams[6] (see especially ref 6c and also ref 8), and will not be given here.

Our procedure differs from that of Williams[6] in that the functional form of the potentials is different. Further, each molecular crystal was examined individually, and, most importantly, *polar* molecules as well as nonpolar ones were considered; in contrast, Williams[6] and also Kitaigorodskii[5] have considered only nonpolar molecules in any detail.

A more comprehensive treatment, including thermal effects on the intermolecular potentials, will be presented in a later paper,[15] together with a complete listing of the refined parameters of the potential functions.

The same procedure, applied earlier[4] to the packing of small molecules, is used here for calculations on the crystal structures of two poly(amino acids). Specifically, we select a starting conformation of an isolated poly(amino acid) (obtained either by energy minimization or, in some cases, by selection of a proposed conformation), generate the assembly of chains in the crystal, and then minimize the energy of the crystal with respect to various parameters of the system (*e.g.*, such as the lattice constants).

B. Geometry. The geometry used for the small molecules[4] is that given in the structural literature for each molecule except for the hydrogen atoms which, for internal consistency, were positioned to best represent the refined hydrogen positions as determined from the literature on electron and neutron diffraction. For the polymers considered here, the geometry was the same as that described earlier,[12,13] except that the length of the C–H bond in aliphatic side chains was taken as 1.07 Å; the value of 1.00 Å for the C^{α}-H bond length was retained. The coordinates of a given (reference) molecule derived from the above-mentioned geometry may be used with the symmetry operators to generate a crystal structure, as indicated in section A.

(10) (a) J. F. Yan, F. A. Momany, R. Hoffmann, and H. A. Scheraga, *J. Phys. Chem.*, **74**, 420 (1970); (b) F. A. Momany, R. F. McGuire, J. F. Yan, and H. A. Scheraga, *ibid.*, **74**, 2424 (1970).

(11) R. G. C. Arridge and C. G. Cannon, *Proc. Roy. Soc., Ser. A*, **278**, 91 (1964).

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

(13) T. Ooi, R. A. Scott, G. Vanderkooi, and H. A. Scheraga, *ibid.*, **46**, 4410 (1967).

(14) J. F. Yan, G. Vanderkooi, and H. A. Scheraga, *ibid.*, **49**, 2713 (1968).

(15) F. A. Momany and H. A. Scheraga, in preparation.

C. Energy Functions. The functional forms of the various potentials are similar to those used by Scott and Scheraga,[12] Ooi, *et al.*,[13] and Yan, *et al.*,[14] The nonbonded terms were taken to be of the form

$$U(r) = A/r^{12} - C/r^{6} \quad (1)$$

The parameters A and C of ref 14 were used for intramolecular interactions, while those used to compute the crystal structures of small molecules[4] were taken for intermolecular interactions. Torsional potentials were identical with those used previously;[12-14] in addition, the potential barrier for rotation about the peptide bond, *i.e.*, U_ω in the equation

$$U(\omega) = (U_\omega/2)(1 - \cos 2\omega) \quad (2)$$

was taken as $U_\omega = 20.0$ kcal/mol.

The monopole approximation was used for the calculation of both inter- and intramolecular electrostatic energies. Partial charges were assigned to the atoms of polar groups so as to reproduce correctly the magnitude and orientation of the group dipole moment. A refined method of assigning charges has recently been presented;[10] however, we have retained the charges of Ooi, *et al.*,[13] for consistency with previous results on homopolymers. The Coulomb potential used for inter- and intramolecular calculations was

$$(U_{ij})_{el} = 332.0 q_i q_j / D r_{ij} \quad (3)$$

where q_i and q_j are the partial electronic charges on the ith and jth atoms, D is the dielectric constant (set equal to 1.0 in all calculations presented here), and r_{ij} is the interatomic distance; $(U_{ij})_{el}$ is given in kilocalories per mole when r_{ij} is in ångström units.

The hydrogen bond potential[4] was taken as a sum of Coulombic and nonbonded energy terms;[16] the nonbonded parameters for the atoms C=O···H–N (*i.e.*, those involved in the hydrogen bond) were previously[4] assigned empirically to yield experimentally reasonable hydrogen bond lengths and energies. In this work, the parameters were further refined to reproduce lattice constants whose values were solely dependent on intermolecular hydrogen bonding. When used with the previously described intramolecular nonbonded, torsional, and electrostatic potentials, these refined values were tested by correctly predicting the known screw sense of the infinitely long, regular α-helical forms of isolated chains of poly-L-alanine, poly(β-methyl-L-aspartate), poly(β-benzyl-L-aspartate), and poly(γ-methyl-L-glutamate). No explicit directional properties were included in the hydrogen bond function,[16] in contrast to the equations used by Ooi, *et al.*,[13] and Yan, *et al.*,[14] where directional terms were included. This exclusion of a directional dependence in the hydrogen bond function conforms to recent results from quantum mechanical calculations on model peptide systems.[10b]

D. Energy Calculations. The total energy of interaction in the crystals of small molecules was computed, for any given rigid conformation, as the sum of all intramolecular torsional, nonbonded, hydrogen bonding, and electrostatic contributions, as well as of the intermolecular nonbonded, hydrogen bonding, and electrostatic contributions. These energy terms were summed over all pairwise atomic intramolecular interactions involving atoms separated by three or more bonds, over all pairwise atomic intermolecular interactions among the molecules of an asymmetric unit, and from the molecules of an asymmetric unit to the surrounding molecules in the crystal lattice. Special care was taken in order to maintain a balance of the interactions between asymmetric units independent of the number of cells generated, by proper treatment of surface and edge effects. For crystals with water of hydration, the water molecules were treated as independent rigid bodies rather than as a continuum or high-dielectric medium.

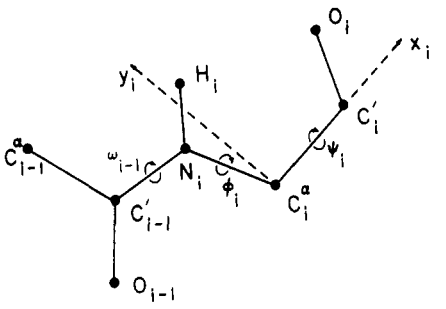


Figure 1. Coordinate system for atoms of ith peptide unit. See text for description. The notation for the backbone atoms and bond rotations is that of Edsall, *et al.*[17]

For the case of regular homopoly(amino acid) helices, the intramolecular energy was calculated as the sum of interactions between the atoms of one peptide unit plus the interactions from one peptide unit to the other peptide units in one direction along the chain; in this way, the result approaches the energy per residue for an infinite helix. Solvent was not included in the helix calculations.

E. Transformation of Coordinates. In our earlier work on poly(amino acids),[12-14] the backbone (ϕ and ψ) and side-chain (χ) dihedral angles were varied, but the dihedral angle ω was held fixed at 0°, corresponding to the planar trans conformation of the amide group.[17] The restriction that ω be held fixed has now been removed, and, for this reason, we will present in some detail the complete coordinate transformation equations necessary for the internal rotational problems in poly(amino acids).

The coordinates of the atoms of peptide unit i were expressed in a local coordinate system defined, as in Ooi, *et al.*,[13] as follows (see Figure 1): the origin is taken to be the C_i^α atom, with the x axis pointing toward the C_i' atom along the C_i^α–C_i' bond; the y axis is perpendicular to the C_i^α–C_i' bond and in the plane formed by the C_i^α, C_i', and O_i atoms, with the O_i atom having positive x and y coordinates; the z axis is located so as to form a right-handed orthogonal coordinate system.

The coordinates $\mathbf{r}^{(i)}$ of any atom in the coordinate system of the ith unit were expressed as $\mathbf{r}^{(i-1)}$ in the

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

(17) J. T. Edsall, P. J. Flory, J. C. Kendrew, A. M. Liquori, G. Nemethy, G. N. Ramachandran, and H. A. Scheraga, *Biopolymers*, **4**, 121 (1966); *J. Biol. Chem.*, **241**, 1004 (1966); *J. Mol. Biol.*, **15**, 399 (1966). The conventions of Edsall, *et al.*, are used in the present paper.

$(i-1)$th coordinate system by the following equation

$$\mathbf{r}^{(i-1)} = \mathbf{T}_{i\to i-1}\mathbf{r}^{(i)} + \mathbf{p}_i^{(i-1)} \quad (4)$$

where $\mathbf{p}_i^{(i-1)}$ is a vector, in the $(i-1)$th system, connecting the C_{i-1}^{α} and C_i^{α} atoms. In order to allow for rotation about the peptide bond, *i.e.*, for changes in the dihedral angle ω, the matrix product $\mathbf{T}_{i\to i-1}$ given by Ooi, *et al.*,[13] was modified to

$$\mathbf{T}_{i\to i-1} = \mathbf{T}_{\alpha}\mathbf{T}_{\omega}\mathbf{T}_{\gamma}\mathbf{T}_{\phi}\mathbf{T}_{\beta}\mathbf{T}_{\psi} \quad (5)$$

The matrices $\mathbf{T}_{\psi}$, $\mathbf{T}_{\beta}$, and $\mathbf{T}_{\phi}$ are identical with those used previously,[13] but the remaining matrices (including the subscripts) replace the $\mathbf{T}_{\alpha}$ matrix of Ooi, *et al.*[13] After the transformation $\mathbf{T}_{\phi}\mathbf{T}_{\beta}\mathbf{T}_{\psi}$, the resulting coordinate system has its origin at the C_i^{α} atom, with the x axis pointing along the extension of the N_i–C_i^{α} bond in the $N_i\to C_i^{\alpha}$ direction, and the xy plane containing the C_{i-1}', N_i, and C_i^{α} atoms; the positive y direction is such that the C_{i-1}' atom has a negative x and a positive y coordinate system to make it parallel to the local one in the $(i-1)$th system (whose origin is at the C_{i-1}^{α} atom). The following transformations are then carried out.

(a) Rotate about the z axis by an angle γ until the new y axis is parallel to the direction of the C_{i-1}'–N_i bond, where the angle γ is equal to $\{\tau[C_{i-1}'N_iC_i^{\alpha}] - 90°\}$; this involves the matrix $\mathbf{T}_{\gamma}$.

(b) Rotate about the new y axis by an angle ω until the xy plane is in the plane defined by the C_{i-1}^{α}, C_{i-1}', and N_i atoms; this involves the matrix $\mathbf{T}_{\omega}$.

(c) Rotate about the new z axis by an angle α $\{=\tau[C^{\alpha}C'N] - 90°\}$ until the x axis points along the C_{i-1}^{α}–C_{i-1}' bond; this involves the matrix $\mathbf{T}_{\alpha}$. The matrices $\mathbf{T}_{\gamma}$, $\mathbf{T}_{\omega}$, and $\mathbf{T}_{\alpha}$ are given by

$$\mathbf{T}_{\gamma} = \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6)$$

$$\mathbf{T}_{\omega} = \begin{bmatrix} \cos\omega & 0 & -\sin\omega \\ 0 & 1 & 0 \\ \sin\omega & 0 & \cos\omega \end{bmatrix} \quad (7)$$

$$\mathbf{T}_{\alpha} = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (8)$$

Because of the variation in the value of ω, $\mathbf{p}_i^{(i-1)}$ is no longer a constant; therefore, its dependence on ω is calculated in the following way. The position of the C_i^{α} atom is expressed in the intermediate coordinate system that is obtained after rotation of the ith local coordinate system by the angles ψ, β, ϕ, and γ (involving the matrix product $\mathbf{T}_{\gamma}\mathbf{T}_{\phi}\mathbf{T}_{\beta}\mathbf{T}_{\psi}$) and translation of the origin to the N_i atom. The coordinates of the C_i^{α} atom written in this intermediate coordinate system can then be expressed in the $(i-1)$th local coordinate system by rotating the coordinate system by the angles ω and α (involving the matrix product $\mathbf{T}_{\alpha}\mathbf{T}_{\omega}$) and translating the origin from the N_i to the C_{i-1}^{α} atom. If $\mathbf{p}_{N_i}^{(i-1)}$ represents the coordinates of the N_i atom in the $(i-1)$th system, and $\mathbf{s}_{C_i^{\alpha}}^{\text{int}}$ the coordinates of the C_i^{α} atom in the above intermediate system, then $\mathbf{p}_i^{(i-1)}$ (for use in eq 4) may be written as

$$\mathbf{p}_i^{(i-1)} \equiv \mathbf{p}_{C_i^{\alpha}}^{(i-1)} = \mathbf{T}_{\alpha}\mathbf{T}_{\omega}\mathbf{s}_{C_i^{\alpha}}^{\text{int}} + \mathbf{p}_{N_i}^{(i-1)} \quad (9)$$

The coordinates of the H_i^N atom in the $(i-1)$th system also depend on ω, and were computed in the same way as were the coordinates of the C_i^{α} atom.

The coordinates of the side-chain atoms were first obtained, using the procedure of Ooi, *et al.*,[13] in a coordinate system having its origin at the C_i^{α} atom and its x axis along the C_i^{α}–C_i^{β} bond, with the C_i^{β} atom having a positive x coordinate. The xy plane of this coordinate system contains the N_i, C_i^{α}, and C_i^{β} atoms, with the positive y direction chosen so that the N_i atom has a positive y coordinate; this is the $j = 0$ side-chain system of Ooi, *et al.*[13] The following operations were then applied to this system to make it coincident with the local coordinate system whose origin is at the C_{i-1}^{α} atom.

(a) Rotation around the z axis (of the $j = 0$ coordinate system) by an angle of $180° - \beta_0$, where β_0 is the $\tau[N_iC_i^{\alpha}C_i^{\beta}]$ bond angle, until the x axis points along the N_i–C_i^{α} bond; this involves the matrix $\mathbf{T}_{\beta_0}$ of Ooi, *et al.*[13]

(b) Rotation about the new x axis by the angle $\chi_0 = \phi + 60°$ (for L residues) or $\chi_0 = \phi - 60°$ (for D residues) until the xy plane is in the plane of the C_{i-1}', N_i, and C_i^{α} atoms; this involves the matrix[18] $\mathbf{T}_{\chi_0}$ which is given in eq 11. This system is now identical with the intermediate coordinate system obtained by applying the transformation $\mathbf{T}_{\phi}\mathbf{T}_{\beta}\mathbf{T}_{\psi}$ to the ith coordinate system with its origin at the C_i^{α} atom.

(c) Rotation by $\mathbf{T}_{\alpha}\mathbf{T}_{\omega}\mathbf{T}_{\gamma}$ to make this intermediate system parallel to the local system with its origin at the C_{i-1}^{α} atom.

(d) Translation of the origin from the C_i^{α} to the C_{i-1}^{α} atom, using the vector $\mathbf{p}_i^{(i-1)}$ given by eq 9.

Therefore, if $\mathbf{r}_{sc}^{\,i}$ represents the coordinates of the side-chain atoms of residue i in the $j = 0$ side-chain system of Ooi, *et al.*,[13] the following equations will generate the coordinates of these atoms in the local coordinate system with its origin at the C_{i-1}^{α} atom

$$\mathbf{r}_{sc}^{(i-1)} = \mathbf{T}_{\alpha}\mathbf{T}_{\omega}\mathbf{T}_{\gamma}\mathbf{T}_{\chi_0}\mathbf{T}_{\beta_0}\mathbf{r}_{sc}^{\,i} + \mathbf{p}_i^{(i-1)} \quad (10)$$

where

$$T_{\chi_0} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi \pm 60°) & -\sin(\phi \pm 60°) \\ 0 & \sin(\phi \pm 60°) & \cos(\phi \pm 60°) \end{bmatrix} \quad (11)$$

F. Generation of Helical Structures. In treating helical structures, the method of Sugeta and Miyazawa[19] was used to transform the Cartesian coordinates of the atoms from a system in which the origin is at a C^{α} atom to one in which the origin is on the helix axis, with one of the coordinate axes coincident with the helix axis, and another of the axes pointing toward a C^{α} atom, *viz.*, the (ξ, η, ζ) system of Sugeta and Miyazawa (see Figure 2). This method was also used to calculate the values of t and h, t being the angle of rotation around the helix axis from the origin of one residue to the origin

(18) For $j \geq 1$, eq 11 of Ooi, *et al.*,[13] was used for $\mathbf{T}_{\chi_j}$. However, eq 11 of the present paper should replace the $\mathbf{T}_{\chi_0}$ matrix of Ooi, *et al.*;[13] while these authors used the correct matrix in their calculations, their reported matrix had a typographical error in the published paper for the case of $j = 0$.

(19) H. Sugeta and T. Miyazawa, *Biopolymers*, **5**, 673 (1967).

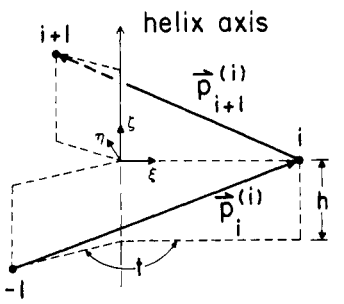


Figure 2. A three-dimensional representation of the vectors connecting three C^α atoms (shown as the points $i - 1$, i, and $i + 1$) in the ith coordinate system. Also shown are the helix coordinates ξ, η, and ζ, as well as the helix parameters h and t.

of the succeeding residue, and h being the corresponding translation along the helix axis. The Sugeta–Miyazawa equations are applicable to polymer chains with any number of dihedral angles per monomer unit.

In order to make the Sugeta–Miyazawa equations relating ϕ, ψ, and ω to h and t compatible with our equations for transformation of coordinates, several modifications were necessary. The first modification is to rewrite the unit vectors of Sugeta and Miyazawa in a slightly different form since our transformation matrix $\mathbf{T}_{i\to i-1}$ is the inverse of their matrix A. To include this modification in our system of nomenclature, we define (in the ith coordinate system) the vector connecting the origins of the coordinate systems based on two successive α-carbon atoms, C_i^α and C_{i+1}^α, as $\mathbf{p}_{i+1}^{(i)}$ (the same as B of Sugeta and Miyazawa) (see Figure 2); the vector $\mathbf{p}_i^{(i)}$ points from the C_{i-1}^α to the C_i^α atom. The unit vectors $\mathbf{e}_\xi$ and $\mathbf{e}_\eta$ are perpendicular to each other and to the helix axis with $\mathbf{e}_\xi$ pointing toward the C_i^α atom; $\mathbf{e}_\zeta$ is perpendicular to both $\mathbf{e}_\xi$ and $\mathbf{e}_\eta$ to form a right-handed coordinate system, and is directed along the helix axis. From Figure 2, these vectors are written as

$$\mathbf{e}_\xi = \frac{\mathbf{p}_i^{(i)} - \mathbf{p}_{i+1}^{(i)}}{|\mathbf{p}_i^{(i)} - \mathbf{p}_{i+1}^{(i)}|} \tag{12}$$

$$\mathbf{e}_\zeta = \frac{(\mathbf{p}_i^{(i)} - \mathbf{p}_{i+1}^{(i)}) \times (\mathbf{p}_{i+1}^{(i)} - \mathbf{p}_{i+2}^{(i)})}{|\mathbf{p}_i^{(i)} - \mathbf{p}_{i+1}^{(i)}|^2 \sin t} \tag{13}$$

$$\mathbf{e}_\eta = \mathbf{e}_\zeta \times \mathbf{e}_\xi \tag{14}$$

and

$$\mathbf{p}_i^{(i)} - \mathbf{p}_{i+1}^{(i)} = (\mathbf{T} - \mathbf{E})\mathbf{p}_{i+1}^{(i)} \tag{15}$$

$$\mathbf{p}_{i+1}^{(i)} - \mathbf{p}_{i+2}^{(i)} = (\mathbf{E} - \mathbf{T}^+)\mathbf{p}_{i+1}^{(i)} \tag{16}$$

where $\mathbf{T}^+$ is the transpose of the $\mathbf{T}_{i\to i-1}$ matrix of eq 5 and $\mathbf{E}$ is the 3×3 identity matrix.

The second modification of the Sugeta–Miyazawa equations results from the indeterminate nature of eq 13 for the cases $t = 0°$ or $180°$. A discontinuity would arise in studies of β conformations (for which $t = 180°$) if eq 13 were used. Instead, as can be seen from Figures 2 and 3, as t goes to 0 or 180°, eq 13 becomes

$$\mathbf{e}_\zeta = \frac{\mathbf{p}_i^{(i)} + \mathbf{p}_{i+1}^{(i)}}{|\mathbf{p}_i^{(i)} + \mathbf{p}_{i+1}^{(i)}|} \tag{17}$$

and

$$\mathbf{p}_i^{(i)} + \mathbf{p}_{i+1}^{(i)} = (\mathbf{T} + \mathbf{E})\mathbf{p}_{i+1}^{(i)} \tag{18}$$

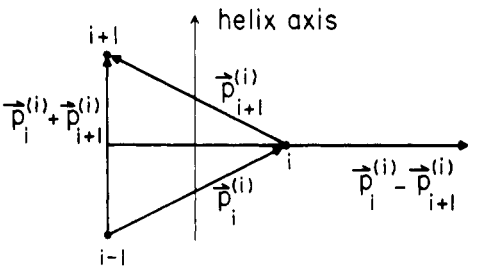


Figure 3. Coordinate system of Figure 2, for the special case in which $t = 0°$ or $180°$ (*i.e.*, when all three C^α atoms lie in the same plane with the helix axis).

In practice, one cannot simply substitute eq 17 for eq 13 at $t = 0$ and 180° without creating a discontinuity in the values of ϕ and ψ. To retain a near-continuity in ϕ and ψ, we have chosen to use eq 17 in a small region about the discontinuous points, the size of this region being such that the solutions of eq 13 approximate the solutions of eq 17 at the boundary (within 0.1° in ϕ and ψ).

The unit vectors $\mathbf{e}_\xi$, $\mathbf{e}_\eta$, and $\mathbf{e}_\zeta$ were then used in the manner described by Sugeta and Miyazawa to form the matrix (eq 23 of ref 19) for rotation of the coordinate system whose origin is at the C_i^α atom so that it is parallel to the one whose origin is on the helix axis; eq 16 and 24 of ref 19 were used to obtain the vector to translate from one of these coordinate systems to the other.

The complete poly(amino acid) helix, for either β or non-β structures, is generated in the (ξ, η, ζ) coordinate system by successive rotations of the first (reference) residue about the helix axis by the calculated value of t and translations along this axis by the calculated value of h (eq 25 of ref 19).

The above operations can be carried out in two ways. In one way, the values of ϕ, ψ, and ω are specified, and a helix is generated in which each residue has this specified set of diheral angles. In the other way, h, t, and ω are specified, and the values of ϕ and ψ are computed for the given helix. The latter option is used when it is desired to generate helical structures consistent with known values of h and t (from X-ray diffraction studies of poly(amino acid) fibers) and assumed values of ω. In this option, since ϕ and ψ are degenerate as determined from h and t (for the assumed value of ω), the two simultaneous nonlinear equations must be solved by an iterative procedure from an initial estimate of ϕ and ψ. It was found that, if the initial estimate is close enough to the final answer, *i.e.*, within 20 or 30°, the iterative procedure finds the desired solution and ignores the second one.

G. Energy Minimization. The energy minimization procedure employed for both intermolecular and intramolecular interactions is the same as the one described by Scott, *et al.*[20] That is, the complete list of variables is divided into subsets, and for each subset the energy is minimized. A system of line searches and quadratic interpolations is used to find the approximate energy minimum. In general, symmetry elements are used

(20) R. A. Scott, G. Vanderkooi, R. W. Tuttle, P. M. Shames, and H. A. Scheraga, *Proc. Nat. Acad. Sci. U. S.*, **58**, 2204 (1967).

such that not all energy terms need be calculated during the course of energy minimization; rather, only those energy terms are summed which are dependent upon the variables included in the subset to be minimized.

The crystal energy may be minimized with respect to any or all of several geometrical parameters. These include internal rotations about bonds, rigid-body motion of the molecules in the asymmetric unit relative to each other and with respect to the fixed lattice coordinate system (maintaining the constraints of the crystal symmetry), and variation of the six lattice constants a, b, c, α, β, and γ.

For regular poly(amino acid) helices, the energy may be minimized with respect to the dihedral angles for internal rotations about the three backbone bonds, ϕ(N–C$^\alpha$), ψ(C$^\alpha$–C′), and ω(N–C′); alternatively, it may be minimized with respect to only one of these three angles, but with the constraint that h and t must remain constant, which thereby determines the values for the other two backbone dihedral angles.

Minimization may also be carried out on the side-chain dihedral angles, χ_i, in any order or any one at a time.

Packed Structures of Poly(amino acid) Helices

β-Poly-L-alanine. Although the most stable conformation of the isolated poly-L-alanine molecule is the α-helix,[21] X-ray diffraction studies[22,23] indicate that the molecule exists in the antiparallel β conformation in fibers stretched in steam.

In the antiparallel β structure, there are two identical helical polymers with their twofold screw axes (t = 180°) lying parallel to each other but pointed in opposite directions. Many such pairs of molecules form an antiparallel pleated sheet, and the crystal consists of a definite arrangement of such sheets.[24]

In order to position the polymer chains in the unit cell so that the crystal lattice of an antiparallel β structure is formed, a Cartesian coordinate system is affixed to the reference chain as follows: for each backbone conformation (*i.e.*, values of ϕ, ψ, and ω) the positions of three successive C$^\alpha$ atoms may be located and the helix axis determined using eq 12–14 and 17. The original coordinates are assigned so that the ith C$^\alpha$ atom lies on the positive x axis which is perpendicular to the helix axis and the z axis is along the helix axis (taken positive from the N to C terminus), with the y axis chosen to complete a right-handed orthogonal coordinate system. The remainder of the atoms in this chain are then determined using eq 22 and 25 of ref 19. The orientation of the reference chain within the unit cell is then determined by aligning the z axis of the molecule to coincide with the c axis of the crystal, with the origin of the coordinate system of the reference molecule coincident with that of the unit cell and the ith C$^\alpha$ atom lying on the a axis. This molecule may then be translated in the positive c direction by Δz and its x axis rotated about the c axis by an amount $z_{rot.}$, the positive direction of rotation being the right-handed one when looking from the N to the C terminus. The original position of the reference chain in the unit cell corresponds to the values Δz = 0.0 Å and $z_{rot.}$ = 0°. The antiparallel chain is then constructed from this reference chain by symmetry operators, *i.e.*, translation in the a direction by $a/2$, translation in the c direction by $+c$, and inversion of the helix axis (z direction) so that the chain is pointing in the $-c$ direction. Figure 4 shows the relative alignment of a short section of this pleated sheet in the a–c and a–b projections with Δz = 0.0 Å and $z_{rot.}$ = −106.0°. As shown, the helix axis of molecule I points in the $+c$ direction and that of molecule II in the $-c$ direction. In Figure 4A, the direction of translation along the helix axis, Δz, is shown for both chains I and II, with Δz being

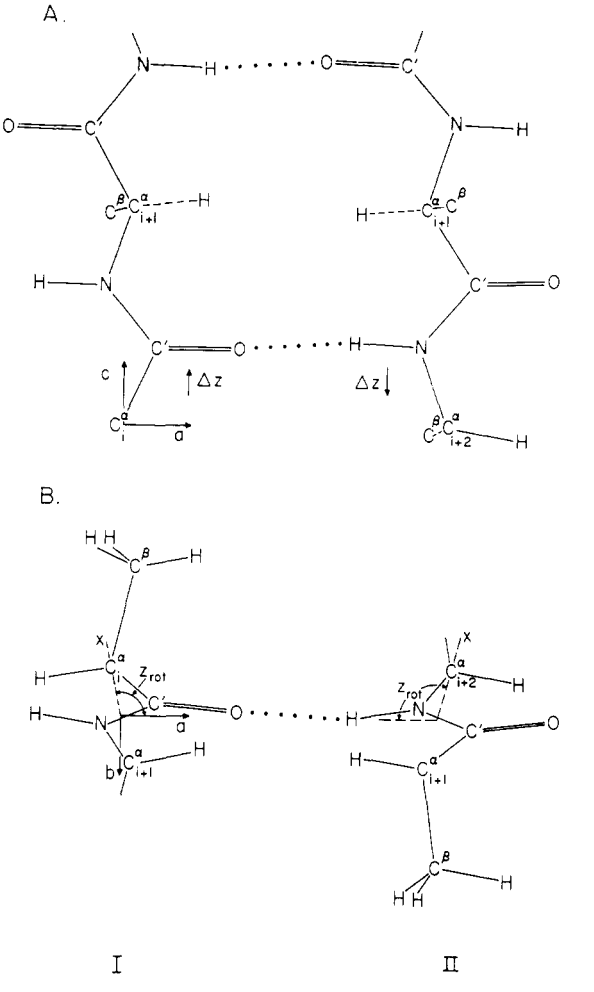


Figure 4. Two poly L-alanine chains in the β conformation of the antiparallel pleated sheet. The axes of molecules I and II lie parallel to the c axis but point in opposite directions. In this drawing, $z_{rot.}$ = −106.0° and Δz = 0.0 Å: A, the a–c projection, showing the intermolecular hydrogen bonds and the translational freedom of each chain (Δz), defined as the linear displacement of each chain along its helix axis relative to a common origin; B, the a–b projection, showing the side-chain orientations of the hydrogen-bonded chains ($z_{rot.}$ is a right-hand rotation about the helix axis when looking along the chain from the N to the C terminus. In Figure 4B, the chain is viewed from the C to the N terminus in I, and in the opposite direction in II).

(21) A. R. Downie, A. Elliott, W. E. Hanby, and B. R. Malcolm, *Proc. Roy. Soc., Ser. A*, **242**, 325 (1957).
(22) L. Brown and I. F. Trotter, *Trans. Faraday Soc.*, **52**, 537 (1956).
(23) S. Arnott, S. D. Dover, and A. Elliott, *J. Mol. Biol.* **30**, 201 (1967).
(24) L. Pauling and R. B. Corey, *Proc. Roy. Soc., Ser. B*, **141**, 21 (1953).

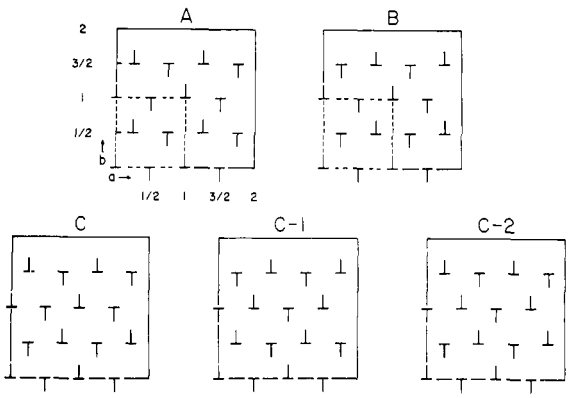


Figure 5. Schematic projection showing the a and b lattice constants for several different unit cells of β-poly-L-alanine packed in antiparallel pleated sheets. A is the expanded model of Tussah silk fibroin, the dotted lines representing the original crystal lattice of Marsh, *et al.*[25] B is the alternate to structure A proposed by Arnott, *et al.*[23] C is the combination model of Arnott, *et al.*,[23] incorporating both structures A and B. C-1 and C-2 are two possible statistical equivalents to structure C, obtained by translating different rows alone the lattice by $a/2$. As shown, the two β-poly-L-alanine chains in the antiparallel pleated sheet are parallel to the c axes, one with its helix axis pointing out of the paper (⊥), and the other with its helix axis into the paper (⊤).

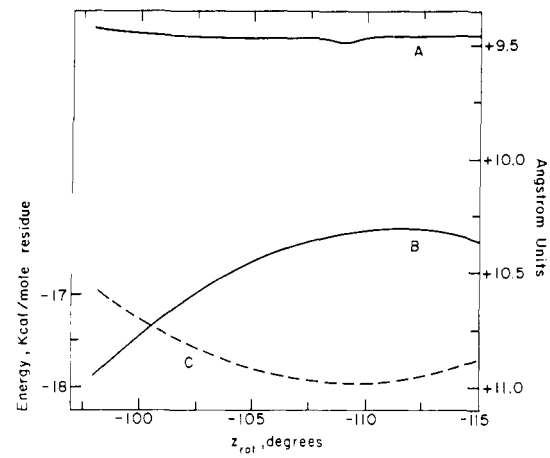


Figure 6. Variation of unit cell dimensions a (curve A) and b (curve B) to minimize the total energy (curve C) of structure C of Figure 5, as a function of the rotation of the poly(amino acid) chains, z_{rot}. The third lattice constant c, and also the internal dihedral angles ϕ, ψ, ω, and χ, were held fixed, and Δz was kept at 0.0 Å.

positive in the direction of the helix axis. In Figure 4B, the rotational displacement, $z_{rot.}$, is shown. All other molecules in the unit cell were generated in a similar fashion. In order to retain the initial symmetry of the unit cell, the same values of Δz and $z_{rot.}$ apply to all chains in the structure.

From the work of Brown and Trotter[22] and Arnott, *et al.*,[23] the unit cell was found to be orthorhombic ($\alpha = \beta = \gamma = 90°$) with unit cell dimensions a = 9.47 Å, b = 10.54 Å, and c = 6.89 Å. Although it is customary to define the unit cell in terms of a, b, and c, the combination of two types of intersheet packing proposed by Arnott, *et al.*,[23] (see below) requires that the a and b dimensions be doubled.

The different lattice arrangements of β-poly-L-alanine considered in this work are shown in Figure 5. The symbols ⊥ and ⊤ represent antiparallel chains with opposite alignment of the side-chain methyl groups (see Figure 4B), the horizontal bar designates the alignment of the C=O and NH groups, and the vertical tail simulates the orientation of the side-chain methyl group for this same residue in the helix. Structure A is the packing model of Marsh, *et al.*,[25] assumed for Tussah silk. Arnott, *et al.*,[23] suggested that structure B is quite similar to structure A, differing only by a translation of rows 2 and 4 by $+a/2$. However, since neither A nor B could account for the observed X-ray reflections, Arnott, *et al.*, proposed a combination of these two, as shown in structure C; but, since even structure C did not fit the X-ray data, Arnott, *et al.*, assumed a statistical random displacement by $\pm a/2$ in the direction of the intermolecular hydrogen bond for their best model. Since the computational procedure used here cannot treat a fluctuating system, we shall consider the static structures A and B and also the static structures C, C-1, and C-2 which (together) may be considered as being equivalent to the statistical model of Arnott, *et al.*,[23] (*i.e.*, if C, C-1, and C-2 turn out to have comparable energies, then the actual system can be considered to be a thermal fluctuation among these three). Structure C-1 is obtained from C by translating rows 2 and 4 by $a/2$, and structure C-2 by translating only the third row of structure C by $a/2$.

In calculating the most stable conformation and crystal packing arrangement of β-poly-L-alanine, we will examine all of the structures of Figure 5 to see which one optimizes the packing, *i.e.*, minimizes the intramolecular and intermolecular interactions. The reference chain, used to generate all molecules in the unit cell, was assumed to be regular, *i.e.*, all dihedral angles[17] were assumed to be the same in every residue, and the initial chain conformation was determined by minimizing the intramolecular energy of a single isolated regular homopolymer, including the constraint that the values of h and t remain fixed at the experimentally observed X-ray values (3.445 Å and 180.0°, respectively). This restriction confines the conformations to a narrow path through the β region of a ϕ–ψ map.[26] Beginning with χ = 60.0° and ω = 0.0° (with the initial values of ϕ and ψ determined by the given values of h, t, and ω), and minimizing the energy with respect to χ (keeping ϕ, ψ, and ω fixed), the resulting dihedral angles, used as the starting conformation of the reference molecule, are given in footnote a of Table I. This reference chain, in the computed initial β conformation, was placed at the origin of the unit cell in all of the structures of Figure 5. The various structures of Figure 5 were then generated from the reference chain, as described previously. Of course, only a finite number of unit cells constitute a crystal in the computations. Each unit cell consists of eight

(25) R. Marsh, R. B. Corey, and L. Pauling, *Acta Crystallogr.*, **8**, 710 (1955).

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

TABLE I

CONFORMATIONS AND ENERGIES OF β-POLY-L-ALANINE IN THE CRYSTAL LATTICES OF MINIMUM ENERGY[a]

Crystal structure	a, Å	b, Å	c, Å	$z_{rot.}$,[b,d] deg	E_{INTER},[e] kcal/mol res	E_{TOTAL}, kcal/mol res
Exptl[c]	9.47	10.54	6.89			
A	9.50[b]	10.66[b]	6.89	−106.0	−10.53	−15.88
B	9.47[b]	10.27[b]	6.89	−109.0	−11.11	−16.46
C	9.47[b]	10.33[b]	6.89	−110.0	−12.60	−17.95
C-1	9.50[b]	10.87[b]	6.89	−105.0	−11.85	−17.20
C-2	9.47[b]	10.33[b]	6.89	−110.0	−12.57	−17.92

[a] The values $\phi = 40.3°$, $\psi = 315.0°$, and $\chi = 55.0°$, for the isolated molecule, were determined from the experimental values of $h = 3.445$ Å and $t = 180.0°$ and the *assumption* that $\omega = 0.0°$. These values of ϕ, ψ, ω, and χ were kept constant in all calculations reported in this table. The intramolecular energy, E_{INTRA}, was −5.35 kcal/mol residue. [b] The parameters which were allowed to vary during this energy minimization were a, b, and $z_{rot.}$, with Δz maintained fixed at 0.0 Å and c fixed at 6.89 Å. [c] These were the initial values for the energy minimization. [d] See text for definition of $z_{rot.}$. [e] The crystal packing energy arising from all intermolecular interactions.

pairs of antiparallel chains of two residues each, for a total of 32 residues per unit cell. The packed structures of the crystal were then generated by translating the unit cells by multiples of $2a$, $2b$, and c in the appropriate directions. Most calculations were carried out with a $3 \times 3 \times 3$ array (*i.e.*, 27) of unit cells; the extension to a $5 \times 5 \times 5$ array was found not to affect the results significantly. In order to maintain the proper symmetry balance of interactions, it was necessary to include some interactions beyond the $3 \times 3 \times 3$ array in computing the nonbonded and electrostatic interactions over all pairs of atoms in the array.

The energy of the $3 \times 3 \times 3$ crystalline array was minimized, in turn, with respect to selected sets of variables, *e.g.*, internal variables (ϕ, ψ, ω, χ), lattice constants (a, b, c), etc., in order to determine which of the structures of Figure 5 is the most stable (and most compatible with the X-ray data), and also in order to assess the relative importance of intra- and intermolecular interactions in determining the conformations of the packed chains.

In the initial calculations, the conformations of the chains were held fixed at the values obtained by minimizing the intramolecular energy of the isolated molecule (see footnote *a* of Table I), and the energy was minimized with respect to the unit cell dimensions a and b for several values of $z_{rot.}$ (with Δz fixed at 0.0 Å and the lattice constant c maintained constant). Figure 6 shows the region of $z_{rot.}$ in which the total energy for the packing of structure C becomes a minimum. For each value of $z_{rot.}$, the corresponding values of a (curve A) and b (curve B) resulting from the minimization of the energy (curve C) for that configuration are also given. As the energy decreases to its minimum value, the value of a remains constant and the value of b decreases considerably, but becomes fairly constant in the region of the energy minimum. Since, as shown in Figure 4B for this backbone conformation, the atoms involved in the intermolecular hydrogen bonds between chains (*i.e.*, the C=O of molecule I and the H–N of molecule II) become aligned along the a axis in the region of $z_{rot.}$ between −100 and −115°, the contribution of the hydrogen bond to the total energy is relatively constant over this range of $z_{rot.}$. This implies that other nonbonded interactions (*i.e.*, methyl–methyl and methyl–backbone) are primarily responsible for the relative orientations of the chains within the region of $z_{rot.}$ where the magnitude of the hydrogen bond energy is dominating the total energy. Thus, although the intermolecular hydrogen bonds are very important for the formation of the antiparallel pleated sheet structure, the arrangement of these sheets in the unit cell is very sensitive to the other nonbonded intermolecular interactions (as is shown by the variation in the b parameter as $z_{rot.}$ varies).

The results for structures A, B, C, C-1, and C-2 are similar, and are shown in Table I. Since the intramolecular energy, E_{INTRA}, of the regular homopolymer was kept constant in this initial calculation, the only variation in the total energy, E_{TOTAL}, arises from E_{INTER}, the intermolecular crystal packing energy. From Table I, it can be seen that structures C, C-1, and C-2 have a more favorable packing energy than structures A and B; whereas structures C and C-2 differ by only 0.03 kcal/mol residue in energy, structure C-1 is only slightly less favorable (by 0.75 kcal/mol residue). Another important observation is the constant value of the a dimension for all five structures and its excellent agreement with the experimentally observed value.

In a second set of calculations, a and b were held fixed at their experimental values, χ was held at 58.5° and ω at 0°,[27] and t maintained at 180° (for the chain to repeat every two residues in the c direction). For each value of c (which, together with the fixed values of t and ω, define h, ϕ, and ψ), the energy was minimized with respect to $z_{rot.}$ and Δz. The results for structure C are shown in Figure 7. It can be seen that, even though ϕ and ψ vary with c, E_{INTRA} is essentially independent of c; however, the intermolecular energy, E_{INTER}, decreased slightly for c values somewhat shorter than the experimental value. Since it was also observed that $z_{rot.}$ and Δz did not change much in the minimization procedure, the energy is dominated by the shortening of the c axis. The correspondence of our calculated

(27) These values of χ and ω were obtained by minimizing the total packing energy, starting with the conformation given in footnote *a* of Table I, and holding the parameters a, b, c, h, and t fixed at their experimental values. $z_{rot.}$ and Δz were fixed at the values corresponding to the lowest energy conformations as given in Table I. The independent variables in this minimization were χ and ω, and the dependent variables (to satisfy the geometry of h, t, and ω) were ϕ and ψ.

TABLE II
CONFORMATIONS AND ENERGIES OF β-POLY-L-ALANINE IN THE OBSERVED CRYSTAL LATTICE[a]

Crystal structure	ϕ,[c] deg	ψ,[c] deg	ω,[b] deg	χ,[b] deg	$z_{rot.}$,[b,d] deg	Δz,[b,e] Å	E_{INTRA},[f] kcal/mol res	E_{INTER},[g] kcal/mol res	E_{TOTAL}, kcal/mol res
A	38.8	314.9	2.9	58.2	−106.0	0.17	−5.41	−11.43	−16.84
B	42.7	317.5	−0.7	58.5	−106.1	0.16	−5.43	−11.85	−17.28
C	40.3	315.0	0.0	58.2	−106.2	0.16	−5.41	−13.41	−18.82
C-1	41.4	316.1	2.2	54.8	−105.0	0.17	−5.24	−12.87	−18.11
C-2	40.3	315.0	0.0	58.3	−106.1	0.16	−5.40	−13.39	−18.79

[a] The lattice constants were held fixed at a = 9.47 Å, b = 10.54 Å, and c = 6.89 Å, and the constraints h = 3.445 Å and t = 180.0° were maintained. [b] The parameters which were allowed to vary during the energy minimization. [c] The backbone dihedral angles were determined from the minimized energy value of ω and the constraint that h = 3.445 Å and t = 180.0°. [d] See text for definition of $z_{rot.}$. [e] See text for definition of Δz. [f] The intramolecular energy for each homopolymer. [g] The crystal packing energy arising from all intermolecular interactions.

parameters to experimental values is thus quite good and certainly within the uncertainty associated with the experimental procedures (*e.g.*, the effect of moisture content on the experimental values, etc.)

The third and final set of calculations considers the complete lattice (*a*, *b*, and *c*) as fixed by the experimental data, and finds the minimum energy conformation by varying the intramolecular parameters ω (and thus ϕ and ψ) and χ as well as the two intermolecular parameters $z_{rot.}$ and Δz. Because of the constraint that h and t remain fixed, ϕ and ψ assume different values for each value of ω, according to the Sugeta–Mizayawa equations. The five structures of minimum energy are given in Table II. Again, we predict that the most energetically favorable crystal packing occurs in structures C and C-2, with C-1 being less favorable by only ~0.7 kcal/mol residue. Structures C and C-2 have planar amide groups (ω = 0.0°), whereas structure C-1 exhibits a slight departure from planarity (~2°) without much increase in energy. Table II indicates that E_{INTER} rather than E_{INTRA} determines the order for the most favorable packing of β-poly-L-alanine.

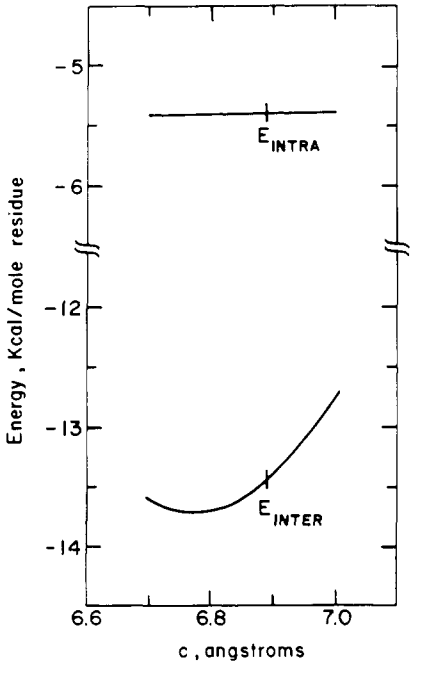


Figure 7. Values of E_{INTRA} and E_{INTER} found when E_{TOTAL} is minimized as a function of the parameters $z_{rot.}$ and Δz, at each value of the lattice constant c for structure C of β-poly-L-alanine. The values of a, b, χ, ω, and t were held fixed. The vertical line on each curve indicates the experimental value of c.

Examination of the most preferable values for $z_{rot.}$ and Δz leads to similar conclusions concerning the conformation of the C=O and H–N groups involved in the intermolecular hydrogen bonds. As indicated in Figure 4B, $z_{rot.} \simeq -106.0°$ corresponds to an almost linear (C=O···H–N) hydrogen bond. In fact, the translation by Δz of ~0.15–0.20 Å (as shown in Table II) makes these groups very linear. Since the observed variation in b and the constancy of a in this region of $z_{rot.}$, as discussed earlier, suggest that the nonbonded rather than the hydrogen bond intermolecular interactions are more important in determining the crystal structure, and further since our hydrogen bond function has no directional character, we can conclude that these same nonbonded intermolecular interactions determine this linear hydrogen bond arrangement.

In conclusion, the energy calculations on the crystal structure of β-poly-L-alanine give excellent agreement with the statistical model proposed by Arnott, *et al.*[23] Energetically, the combination of the two simpler crystal structures A and B, to form structure C, is much more favorable. Furthermore, structures C and C-2 have the same energy, and only a slight preference is shown for these structures over structure C-1. Therefore, these calculations suggest that the statistical model of Arnott, *et al.*,[23] is the most favorable of those considered, and accounts for the observed crystal X-ray reflections.

Poly(β-benzyl-L-aspartate). It has been shown that, in solution, poly(β-benzyl-L-aspartate) (PBLA) is a left-handed α-helix[28–30] but, in stretched fibers and films, it exists as a left-handed ω-helix,[31,32] packed in a tetragonal unit cell of dimensions a = b = 13.85 Å

(28) R. H. Karlson, K. S. Norland, G. D. Fasman, and E. R. Blout, *J. Amer. Chem. Soc.*, **82**, 2268 (1960).
(29) E. M. Bradbury, A. R. Downie, A. Elliott, and W. E. Hanby, *Proc. Roy. Soc., Ser. A*, **259**, 110 (1960).
(30) G. D. Fasman in "Poly-α-amino Acids," G. D. Fasman, Ed., Marcel Dekker, New York, N. Y., 1967, p 551.
(31) E. M. Bradbury, L. Brown, A. R. Downie, A. Elliott, W. E. Hanby, and T. R. R. McDonald, *Nature (London)*, **183**, 1736 (1959).
(32) E. M. Bradbury, L. Brown, A. R. Downie, A. Elliott, R. D. B. Fraser, and W. E. Hanby, *J. Mol. Biol.*, **5**, 230 (1962).

TABLE III
INITIAL AND MINIMUM-ENERGY CONFORMATIONS OF ISOLATED RIGHT- AND LEFT-HANDED PBLA HELICES

	ϕ, deg	ψ, deg	ω,[a] deg	χ_1, deg	χ_2, deg	χ_3, deg	χ_4, deg	χ_5, deg	$-E_{INTRA}$, kcal/mol res E_R	E_H	Total energy, kcal/mol res
α-Helix											
Right-handed											
Initial	130.1	124.1	0.0	305.7	306.3	161.6	228.8	159.6			
Minimized	132.2	123.2	0.0	306.7	317.5	158.8	237.5	157.3	−1.29	−29.19	−30.48
Left-handed											
Initial	228.6	238.1	0.0	174.0	20.9	203.4	146.3	24.7			
Minimized	230.3	236.4	0.0	175.3	19.8	201.6	145.7	26.7	−2.35	−28.57	−30.92
ω-Helix											
Left-handed											
Initial	244.1	234.1	−5.0	264.1	256.4	123.3	264.6	178.3			
Minimized	243.6	242.9	−5.0	300.1	228.3	141.7	260.5	215.9	−2.23	−23.37	−25.55

[a] ω was held fixed in the computations.

and c = 5.30 Å. The ω-helix differs from the α-helix in that it has 4.0 instead of ~3.6 residues per turn. We investigate here the relative roles of the crystal packing energy compared to the intramolecular energy associated with helix stability in order to account for the observed conformation in the fiber.

The energies of the regular homopolymer for the isolated α- and ω-helices of PBLA were initially computed to determine their lowest energy conformation. Both the initial and minimum-energy conformations and the minimized energies for the right- and left-handed α-helices and the left-handed ω-helix[33] are given in Table III. The initial conformations chosen for the right- and left-handed α-helices were those of Yan, *et al.*,[14] with ω = 0.0°. The initial backbone conformation for the left-handed ω-helix was obtained by computing ϕ and ψ from the Sugeta–Miyazawa equations, using the parameters found by Bradbury, *et al.*,[32] *viz.*, h = 1.325 Å, t = −90.0°, and ω = −5.0°;[34] the initial side-chain conformation for this helix was that proposed by Bradbury, *et al.*,[32] based on X-ray diffraction and infrared spectral data for the *packed* structure. For comparison of the different isolated helices, the values of ω were not varied in the computations. During the minimization there was no restriction imposed on the values of t or h; however, the minimum-energy values of the number of residues per turn ($360/t$) were within 0.1 of the usually quoted values of 3.67 and 4.0 for the α- and ω-helix, respectively. As shown in Table III, our minimum-energy conformations for the α-helices differed slightly from the starting conformations because a different hydrogen bond potential function was used in the present work, as discussed earlier. The total energy, being only an intramolecular one, E_{INTRA}, for an isolated helix, is divided into its intra-residue (E_R) and inter-residue or helix (E_H) components. In agreement with Ooi, *et al.*,[13] and Yan, *et al.*,[14] the primary difference in stability between the right- and left-handed α-helices of PBLA is attributable to E_R (which includes the amide–ester dipole–dipole interaction), even though the other interactions (nonbonded, electrostatic, and hydrogen bond) differ by varying amounts, so that the lowest energy left-handed α-helix results from all of these differences. It can be seen that, for the isolated polymer, the left-handed α-helix is the most stable of the three shown in Table III.

It would have been desirable to demonstrate here that, whereas the left-handed α-helix is the most stable form of the isolated polymer, intermolecular packing energies make the left-handed ω-helix the most stable one in the crystal (or fiber). However, while calculations of packed ω-helices can be carried out, because we know the symmetry of the unit cell from the experimental work of Bradbury, *et al.*,[31,32] such calculations cannot be carried out for the α-helix, since we do not know in advance what the symmetry of its packed structure would be, and we cannot, at the present stage of development of the computational procedure, predict the symmetry. We will, therefore, confine our calculations here to obtaining the energetically most favorable conformation of the packed left-handed ω-helix, and showing how the various components of the total energy contribute to the stability of this conformation. From the point of view of the applicability of the technique developed here to the packing of macromolecules in general, in order to compute conformations of biopolymers in the crystal, it will be necessary to know at least the symmetry and number of molecules per unit cell, which is the minimal information necessary from X-ray diffraction studies.

The calculations for packed PBLA ω-helices were carried out for a tetragonal lattice, generated by placing the reference helix parallel to the c axis, at the origin of the unit cell, and then constructing other unit cells by translating this reference helix by multiples of 13.85 Å in the directions of the a and b axes. In this model, there is only one chain per unit cell, and we will omit consideration of any statistical model by assuming that all chains run in the same direction.

Two initial conformations (structures A and B of Table IV) were selected; structure A is that proposed by Bradbury, *et al.*,[32] and structure B has the side-chain conformation of Yan, *et al.*,[14] for the α-helix and a

(33) These calculations consume a large amount of computer time in order to explore all possible initial combinations of side-chain dihedral angles. Since such calculations had already been carried out by Yan, *et al.*,[14] for both the left- and right-handed α-helices, both helix senses were considered here for the α-helix. However, it was not felt necessary to invest the computer time to explore the right-handed ω-helix, since an initial (left-handed) conformation was available from the results of Bradbury, *et al.*[32]

(34) In the nomenclature of Edsall, *et al.*,[17] the ω values given by Bradbury, *et al.*,[32] become negative.

TABLE IV
INITIAL AND MINIMUM-ENERGY CONFORMATIONS OF PACKED LEFT-HANDED ω-HELICES OF PBLA

Structure	h, Å	t, deg	ϕ,[a] deg	ψ,[a] deg	ω, deg	χ_1, deg	χ_2, deg	χ_3, deg	χ_4, deg	χ_5, deg
A										
Initial[b]	1.325	−90.0	244.1	234.1	−5.0	264.1	256.4	123.3	264.6	178.3
Minimized	1.325	−90.0	230.7	246.1	−3.0	300.7	223.9	178.0	272.2	201.5
B										
Initial	1.325	−90.0	217.5	258.7	0.0	174.0[c]	20.9[c]	203.4[c]	146.3[c]	24.7[c]
Minimized	1.325	−90.0	230.7	246.1	−3.0	190.3	17.3	195.7	141.4	24.7
C[d]										
Initial	1.600	−98.6	228.6	238.1	0.0	174.0	20.9	203.4	146.3	24.7
Minimized	1.610	−98.4	230.3	236.4	0.0	175.3	19.8	201.6	145.7	26.7

[a] Obtained from Sugeta–Miyazawa equations for the given value of ω and the experimental values of t and h. [b] Conformation proposed by Bradbury, *et al.*[32] (same as that shown in Table III). [c] Side-chain conformation of Yan, *et al.*[14] (same as that shown in Table III). [d] *Isolated* left-handed α-helix of Yan, *et al.*[14] (same as that shown in Table III).

TABLE V
COMPARISON OF ENERGIES OF THE MINIMUM-ENERGY CONFORMATIONS OF PACKED PBLA HELICES

Structure[a]	ϕ, deg	ψ, deg	ω, deg	$z_{rot.}$,[b] deg	$-E_{INTRA}$, kcal/mol res— E_R	E_H	E_{INTER}, kcal/mol res	E_{TOTAL}, kcal/mol res
A	230.7	246.1	−3.0	5.4	−5.43	−20.89	−3.68	−30.00
B	230.7	246.1	−3.0	−41.4	−2.19	−22.45	0.33	−24.32

[a] Minimum-energy structures of Table IV. The crystal lattices were fixed at their experimental values, *i.e.*, $a = b = 13.85$ Å, $c = 5.30$ Å. [b] See text for definitions of $z_{rot.}$ and Δz, the latter being held fixed at 0.0 Å since there is only one molecule per unit cell.

backbone conformation for an ω-helix, determined from the Sugeta–Miyazawa equations for $\omega = 0.0°$ and the experimental values of t and h.

The energies of the packed structures were minimized with respect to the dihedral angles of the backbone and side chain, and also with respect to the relative orientation about the c axis, $z_{rot.}$, for fixed a, b, and c. The translational displacement along the c axis, Δz, was held constant at 0.0 Å, since, in a crystal with one molecule per unit cell, the relative positions of equivalent atoms in corresponding molecules are unaffected by a variation in Δz. Also, to retain the symmetry of the ω-helix (*i.e.*, four residues per turn), the value of t was held fixed at −90.0°. Hence, as in the case of poly-L-alanine, the values of ϕ and ψ of PBLA were calculated from the independent variables ω, h, and t, using the Sugeta–Miyazawa equations.

In order to obtain an improved set of side-chain dihedral angles, the energies of structures A and B were first minimized with respect to the five χ_i's and $z_{rot.}$ for fixed a, b, c, ω, h, and t. Using the resulting side-chain conformations, the energies were then minimized with respect to the five χ_i's and $z_{rot.}$, for various values of ω, for fixed a, b, c, h, and t. The results for structures A and B are shown in Figures 8 and 9, respectively, as the total energy, E_{TOTAL}, the intermolecular crystal packing energy, E_{INTER}, and E_R and E_H, for various values of ω. It can be seen that, in both structures, E_{TOTAL} is a minimum for $\omega \sim -3°$, *i.e.*, for a nonplanar peptide bond, with E_H being the dominant term favoring this conformation. This value of ω is slightly less than the −5° suggested by Bradbury, *et al.*,[32] and, hence, results in an intramolecular hydrogen bond (*i.e.*, an N···O distance) of ~2.75 Å, which is slightly shorter than the 2.90 Å given by Bradbury, *et al.*[32] The structures corresponding to the minimum in E_{TOTAL} are given in Table IV.

From the energies of the minimum-energy structures given in Table V it can be seen that structure A is 5.7 kcal/mol residue more stable than structure B. From Figures 8 and 9, it appears that, while E_H is approximately ten times greater than E_{INTER} or E_R and forces the minimum in E_{TOTAL} to occur at $\omega \sim -3°$, E_H is similar for structures A and B (being ~1.6 kcal/mol residue lower for structure B at $\omega = -3°$) and is therefore *not* the factor making structure A more stable than structure B. Instead, the stability of structure A at $\omega = -3°$ arises from ~4.0 kcal/mol residue difference in E_{INTER} which favors A over B; *i.e.*, it is the crystal packing energy which is most important in making A the most stable structure.

An *a–b* projection of one complete turn of one chain of structure A of minimum energy is shown in Figure 10. The *a–c* and *a–b* projections of the same structure A of minimum energy, showing the relative side-chain orientations, are given in Figure 11.

It should be noted that the isolated ω-helix is less favorable compared to the isolated α-helix by ~5 kcal/mol residue (see Table III). However, the ω-helix (structure A) acquires additional stabilization from intermolecular interactions in the crystal; in fact, the energy of an ω-helix in a crystal (Table V) becomes comparable to that of an isolated α-helix (Table III). Because we do not know the symmetry of a packed structure for the α-helix, we cannot compare the stabilities of the crystalline forms of the α- and ω-helices,[35] nor can we attribute any significance

(35) Calculations with α-helices packed in the tetragonal cell of the ω-helix (with one molecule per unit cell) indicate that the unit cell would have to expand to $a = b = 15.0$ Å and $c = 28.9$ Å to give a packed structure of α-helices (*i.e.*, 18 residues per repeat along the c axis) with a minimized energy comparable to that of a packed structure found for the ω-helices. This calculation does not provide any information about the actual crystalline form which packed α-helices might take, other than the tetragonal one, but it is at least in agreement with the conclusions of Bradbury, *et al.*,[32] that PBLA is not α-helical under the conditions of their experiments.

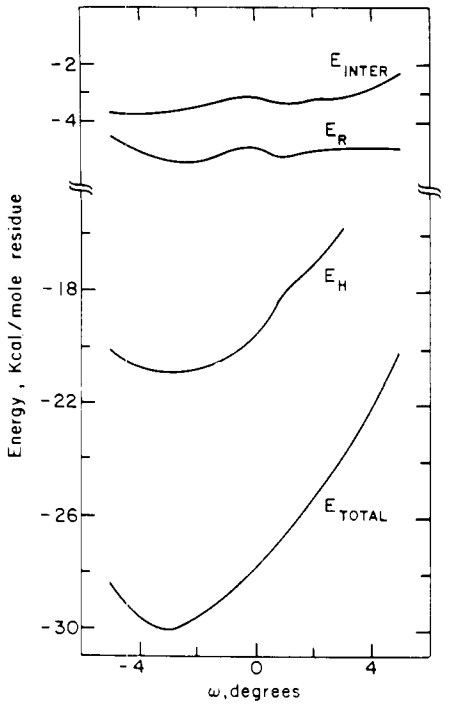


Figure 8. The total minimized energy (E_{TOTAL}) and its components E_{INTER}, E_R, and E_H for structure A of PBLA at selected values of ω. At each value of ω, the total energy was minimized with respect to the five side-chain dihedral angles and the packing parameter z_{rot}. The crystal parameters Δz, a, b, c, h, and t were held fixed.

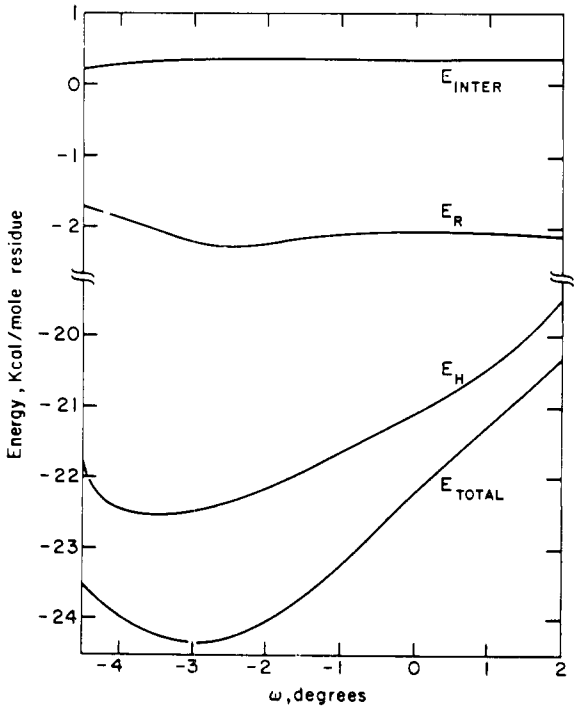


Figure 9. The total minimized energy (E_{TOTAL}) and its components E_{INTER}, E_R, and E_H for structure B of PBLA at selected values of ω. At each value of ω, the total energy was minimized with respect to the five side-chain dihedral angles and the packing parameter z_{rot}. The crystal parameters Δz, a, b, c, h, and t were held fixed.

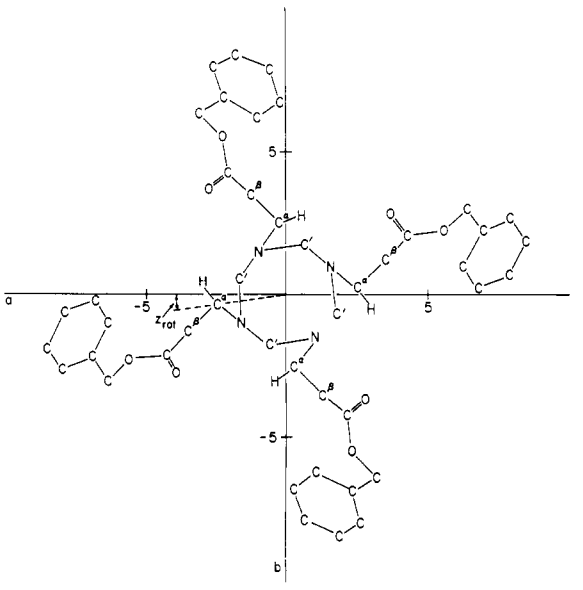


Figure 10. An a–b projection of one complete turn in the minimum-energy conformation of structure A for the left-handed ω-helix (four residues) of PBLA. The helix axis is perpendicular to the plane of the paper. (The scale from 0 to 5 is in ångströms). z_{rot} is a right-handed rotation about the helix axis when looking from the N to the C terminus, and is 5.4° in this figure.

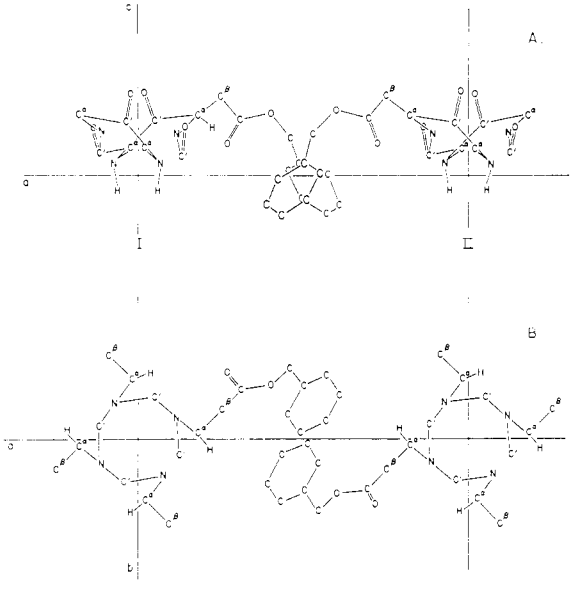


Figure 11. a–c (A) and a–b (B) projections of two packed ω-helices (I and II) in their minimum-energy conformation for PBLA (structure A of Table V), showing the arrangement of the side chains and the overlapping packing. For clarity, only the backbone atoms of one complete helix turn, plus one side chain in each chain, are shown for each ω-helix.

to the difference in energy between the packed ω-helix and the isolated α-helix. However, the difference in energy (~5.7 kcal/mol residue) between structures A and B is sufficiently large to enable us to select the former as the more stable of the two, despite the small uncertainties in the intermolecular potential functions.[4]

Finally, the energies of packed structures A and B were minimized with respect to the five χ_i's and z_{rot}.

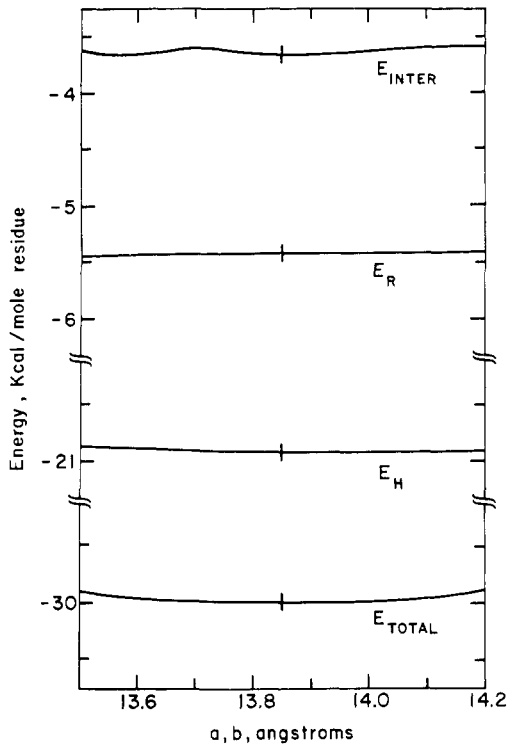


Figure 12. The total minimized energy (E_{TOTAL}) and its components E_{INTER}, E_R, and E_H for structure A of PBLA at selected values of a and b. For each value of a and b, the total energy was minimized with respect to the five side-chain dihedral angles and the packing parameter z_{rot}. The parameters Δz, ω, c, h, and t were held fixed. The vertical line on each curve indicates the experimental value of a and b.

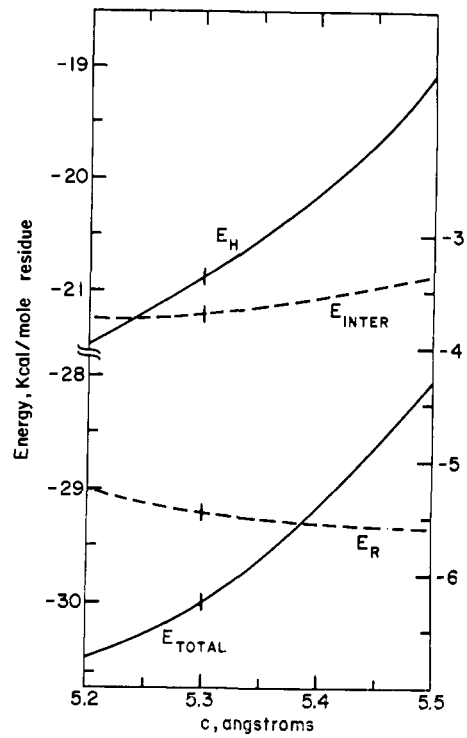


Figure 13. The total minimized energy (E_{TOTAL}) and its components E_{INTER}, E_R, and E_H for structure A of PBLA at selected values of c. For each value of c, the total energy was minimized with respect to the five side-chain dihedral angles and the packing parameter z_{rot}. The parameters Δz, a, b, ω, and t were held fixed. The right-hand ordinate pertains to the dashed curves, and the vertical line on each curve indicates the experimental value of c.

at various values of a and b for fixed c, ω, h, and t (starting with the minimum-energy structures of Table IV). The results for structure A are shown in Figure 12. Over this range of a and b there is very little variation in the various components of the energy, with E_{TOTAL}, showing a slight minimum at the observed values of these lattice constants. These results, and also the variation (not shown here) of the five χ_i's as a and b change, suggest that structure A is determined by intermolecular interactions which avoid close contacts between neighbors and, as shown in Figure 11, which maintain good nonbonded interaction (*i.e.*, stacking) between the phenyl groups. Furthermore, this insensitivity of E_{TOTAL} to a and b, because of the absence of close atom–atom contacts, should allow considerable rotational freedom of the side chains; this might be a possible cause of the more diffuse side-chain reflections, instead of the statistical "up and down" model proposed by Bradbury, *et al.*[32]

The energy does vary considerably with c (for structure A), as shown in Figure 13, but no minimum is observed over this range of c. These calculations were carried out for various values of c (with h restricted to $c/4$ to ensure that there are 4.0 residues per turn), starting with the minimum-energy structure of Table V, and minimizing the energy with respect to the five χ_i's and z_{rot} for fixed Δz, a, b, t, and ω ($-3°$). Whereas E_R becomes less negative as c is decreased, both E_H and E_{INTER} become more negative. At c = 5.2 Å, where E_{INTER} appears to have reached a minimum, E_H is still becoming more negative, tending to distort the ω-helix into some other conformation. We believe that these discrepancies in the c parameter arise from influences of solvent and fiber preparation which were not taken into account in these calculations. However, both the calculated internal homopolymer conformation and also the calculated crystal packing, with the constraint that c remain fixed at the observed value, agree very well with the model proposed by Bradbury, *et al.*,[32] to account for the observed X-ray reflections.

Conclusions

The technique used here to study the crystal packing of two poly(amino acids) is applicable not only to small molecules[4] and homopolymers, but also to heteropolymers and proteins. Thus, it should be applicable to crystals of, say, gramicidin-S, whose conformation as an isolated molecule is being computed. Extension of the calculations to crystals of this decapeptide should enable us to assess the influence of intermolecular packing energies on the conformation. Also, in general, such calculations can then be tested directly by comparison with structural information from X-ray crystallography.

Finally, by applying this technique (together with thermodynamic and spectroscopic data) to small molecules,[4,15] the potential functions are being refined to make them more accurate for conformational energy calculations of macromolecules.