Energetic Approach to the Packing of α-Helices. 2. General Treatment of Nonequivalent and Nonregular Helices

Kuo-Chen Chou,[1] George Némethy, and Harold A. Scheraga*

Contribution from the Baker Laboratory of Chemistry, Cornell University, Ithaca, New York 14853. *Received September 16, 1983*

Abstract: Conformational energy calculations have been carried out in order to determine energetically favorable ways of packing two α-helices. A generalized mathematical formulation of the selection of helical coordinate systems and of coordinate transformations used in packing calculations has been developed. It is suitable for the description of any interacting helical assembly in proteins and any other macromolecules. The packing was investigated for two CH_3CO-(L-Ala)$_{10}$-$NHCH_3$ α-helices, as well as for a pair consisting of one such α-helix and a CH_3CO-(L-Leu)$_{10}$-$NHCH_3$ α-helix. The helices were allowed to be nonequivalent (in terms of sequence and dihedral angles) and nonregular (i.e., dihedral angles of residues along each helix could be different). Comparison of these two pairs of helices indicates the effect of introducing a bulky side chain. Furthermore, Leu was chosen because it occurs frequently in α-helices of globular proteins. The most favorable packing arrangements for the poly(L-Leu)/poly(L-Ala) helix pair have nearly antiparallel orientation of the helix axes, with orientational torsion angles of −154 and 144°, respectively, between the helix axes. Nearly antiparallel arrangements, with closely similar values of the orientational torsional angles, are favored for the poly(L-Ala)/poly(L-Ala) pair as well. In these two packings, the substitution of Leu for Ala does not significantly affect the interhelix energies and the geometrical arrangement of residues in contact. In other low-energy packings, however, with nearly perpendicular orientation of the helix axes or with a torsion angle near −35°, differences occur in the geometry of packing and in the relative energies. The results indicate that details of the manner of packing depend sensitively on side-chain interactions in packings with intermediate energies. All of these computed packing arrangements also occur frequently in globular proteins; thus the packing preferences can be accounted for in terms of noncovalent interhelix interactions.

Introduction

Interactions between regularly folded segments of the polypeptide chain are among the important features of the conformation of globular proteins[2,3] at the level of medium- and long-range interactions. The packing of α-helices is one of the main examples of such interactions. Recently, we have developed a computational method to calculate and optimize the interaction energy between two α-helices.[4] The analysis of the packing of two *equivalent* poly(L-Ala) α-helices indicated that there is a limited number of low-energy packing arrangements, with well-defined orientational torsion angles (to be defined below) between the helix axes.[4] The results are consistent with earlier studies of the geometry of helix–helix packing[5-11] and of electrostatic interactions between helices.[12-14] The computations of the potential energy,[4] however, permit a distinction between more favored and less favored packing arrangements and therefore give information that could not be obtained from the geometrical models. The most stable computed orientations of two poly(L-Ala) α-helices[4] correspond to frequently observed packing arrangements in globular proteins.[2,8,14-18] Thus, the computations provide an energetic explanation for the observed preferences.

The mathematical formulation developed earlier[4] was limited to an idealized model, viz., to a pair of identical α-helices that were regular, i.e., with the backbone dihedral angles constrained to be the same for each residue along the α-helix. In order to deal with actual α-helices occurring in proteins, it is necessary to generalize the methodology to the treatment of two or more nonequivalent and nonregular helices, i.e., ones for which the amino acid sequences can be different and in which all dihedral angles along each polypeptide chain can vary independently. This generalization required several nontrivial modifications of the mathematical formulation. The reference frame used to correlate the relative positions and orientations of helices and the method used to generate coordinates of the various helices had to be sufficiently general to deal with nonequivalent helices, while a simple formulation, based on symmetry operations, was sufficient[4] for equivalent regular α-helices. The general formulation developed here also permits the analysis of other interacting structural elements in proteins, such as those between α-helices and β-sheets, as well as within and between β-sheets.[19] Furthermore, this formulation can be used to analyze geometrical and energetic relationships in *nonregular* helical structures of fibrous proteins,[3] as well as other biological and synthetic macromolecules.

In the present work, the effect of side-chain interactions on the interaction between α-helices is studied by analyzing the packing of poly(L-Leu) and poly(L-Ala) α-helices, and comparing it with the packing of two poly(L-Ala) α-helices computed earlier.[4] Leucine was chosen for two reasons. It occurs frequently in α-helices in globular proteins,[20-22] and it has a bulky side chain,

(1) Visiting Professor from Shanghai Institute of Biochemistry, Chinese Academy of Sciences.
(2) Richardson, J. S. *Adv. Protein Chem.* **1981**, *34*, 167-339.
(3) Scheraga, H. A.; Chou, K. C.; Némethy, G. In "Conformation in Biology"; Srinivasan, R. N., Sarma, R. H., Eds.; Adenine Press: Guilderland, N.Y., **1982**; pp 1-10.
(4) Chou, K. C.; Némethy, G.; Scheraga, H. A. *J. Phys. Chem.* **1983**, *87*, 2869-2881.
(5) Crick, F. H. C. *Acta Crystallogr.* **1953**, *6*, 689-697.
(6) Ptitsyn, O. B.; Rashin, A. A. *Dokl. Biochem. (Engl. Transl.)* **1973**, *213*, 473-475.
(7) Chothia, C.; Levitt, M.; Richardson, D. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 4130-4134.
(8) Chothia, C.; Levitt, M.; Richardson, D. *J. Mol. Biol.* **1981**, *145*, 215-250.
(9) Richmond, T. J.; Richards, F. M. *J. Mol. Biol.* **1978**, *119*, 537-555.
(10) Efimov, A. V. *Dokl. Biochem. (Engl. Transl.)* **1977**, *235*, 699-702.
(11) Efimov, A. V. *J. Mol. Biol.* **1979**, *134*, 23-40.
(12) Wada, A. *Adv. Biophys.* **1976**, *9*, 1-63.
(13) Hol, W. G. J.; Halie, L. M.; Sander, C. *Nature (London)* **1981**, *294*, 532-536.
(14) Sheridan, R. P.; Levy, R. M.; Salemme, F. R. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 4545-4549.
(15) Hendrickson, W. A.; Klippenstein, G. L.; Ward, K. B. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 2160-2164.
(16) Clegg, G. A.; Stansfield, R. F. D.; Bourne, P. E.; Harrison, P. M. *Nature (London)* **1980**, *288*, 298-300.
(17) Argos, P.; Rossmann, M. G.; Johnson, J. *Biochem. Biophys. Res. Commun.* **1977**, *75*, 83-86.
(18) Weber, P. C.; Salemme, F. R. *Nature (London)* **1980**, *287*, 82-84.
(19) Chou, K. C.; Némethy, G.; Scheraga, H. A., to be submitted for publication.

0002-7863/84/1506-3161$01.50/0

so that its use indicates the effect of increased size of side chains on the packing of α-helices.

Mathematical Formulation

Determination of the Helix Axis. As we have shown earlier,[4] it is convenient to use a coordinate system (f, g, h) for which the h axis coincides with the axis of one of the helices. With this choice, it is easy to describe the relative position and orientation of two or more helices.

The method developed earlier (Appendix A of ref 4; based on ref 23 and 24) defines the helix axis in terms of given atomic coordinates for regular helices only. The method is not directly applicable to nonregular helices. In the latter, the nonregularity of structure can be a consequence of a nonregular amino acid sequence (with consequent variation of residue geometry along the chain). In the analysis of α-helices that occur in experimentally observed protein structures, further nonregularity can arise from experimental uncertainty of atomic coordinates. For nonregular helices, the helix axis is defined here (Figure 1) as a least-squares line. It is computed from the coordinates of *all* C^α atoms of the helix in such a manner that the sum of the squares of the distances of all C^α atoms from this line, $\mathcal{S}$, is a minimum.

Let the coordinates of the ith C^α atom of an α-helix (where $i = 1, 2, ..., n_r$, and n_r is the number of residues) in any given (x, y, z) coordinate system be (x_i, y_i, z_i). Then the equation of the least-squares axis can be written as

$$(x - x^*)1/l = (y - y^*)/m = (z - z^*)/n \qquad (1)$$

where

$$x^* = \frac{1}{n_r}\sum_{i=1}^{n_r} x_i, \qquad y^* = \frac{1}{n_r}\sum_{i=1}^{n_r} y_i, \qquad z^* = \frac{1}{n_r}\sum_{i=1}^{n_r} z_i \qquad (2)$$

and l, m, and n are the direction cosines, obtained as the solution of the equations

$$\begin{aligned} (\delta_{xx} + \lambda)l + \delta_{xy}m + \delta_{xz}n &= 0 \\ \delta_{xy}l + (\delta_{yy} + \lambda)m + \delta_{yz}n &= 0 \\ \delta_{xz}l + \delta_{yz}m + (\delta_{zz} + \lambda)n = 0 \qquad l^2 + m^2 + n^2 &= 0 \end{aligned} \qquad (3)$$

where

$$\delta_{xx} = \sum_{i=1}^{n_r}(x_i - x^*)^2, \qquad \delta_{yy} = \sum_{i=1}^{n_r}(y_i - y^*)^2,$$

$$\delta_{zz} = \sum_{i=1}^{n_r}(z_i - z^*)^2$$

$$\delta_{xy} = \sum_{i=1}^{n_r}(x_i - x^*)(y_i - y^*), \qquad \delta_{xz} = \sum_{i=1}^{n_r}(x_i - x^*)(z_i - z^*) \qquad (4)$$

$$\delta_{yz} = \sum_{i=1}^{n_r}(y_i - y^*)(z_i - z^*)$$

and λ is determined by solving the equation:

$$\det(\Delta_{xyz} + \lambda \mathbf{I}) = 0 \qquad (5)$$

In eq 5,

$$\Delta_{xyz} = \begin{bmatrix} \delta_{xx} & \delta_{xy} & \delta_{xz} \\ \delta_{xy} & \delta_{yy} & \delta_{yz} \\ \delta_{xz} & \delta_{yz} & \delta_{zz} \end{bmatrix} \qquad (6)$$

and $\mathbf{I}$ is the 3×3 unit matrix. Since Δ_{xyz} is a matrix with real and symmetric elements, eq 5 has three real roots λ. Substitution of these in eq 3 gives three sets of values of (l, m, n), respectively.

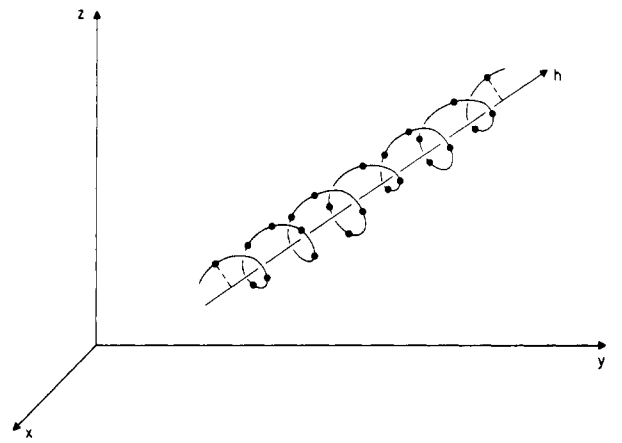

Figure 1. The helix axis h, determined as the line from which the sum of the squares of the distances of all C^α atoms (schematically represented as heavy dots on a helical line) is a minimum. The dashed lines indicate the distances of C^α atoms from the helix axis.

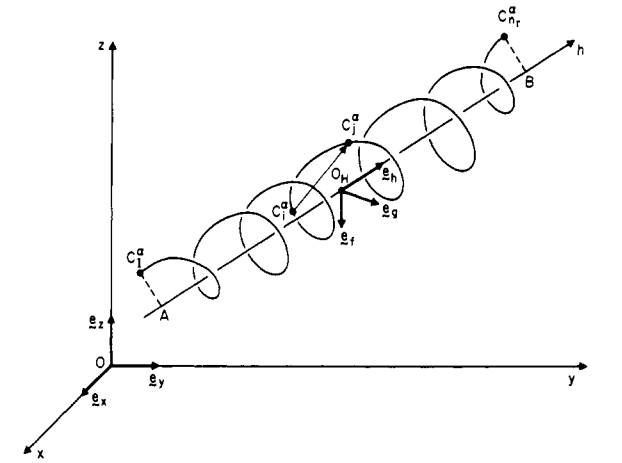

Figure 2. Location of an α-helix in a general (x, y, z) coordinate system, and the definition of a helical (f, g, h) coordinate system, as described in the text. Axis h coincides with the helix axis. Axis f (whose direction is defined by the unit vector $\mathbf{e}_f$) is perpendicular to the h axis and to the $\overrightarrow{C_i^\alpha C_j^\alpha}$ vector connecting the C^α atoms of residues i and j. Axis g (whose direction is defined by the unit vector $\mathbf{e}_g$) is perpendicular to both the f and h axes. Points A and B are the projections of the C_α atoms of the first and last (n_r) residues of the α-helix on the h axis. The distance AB is defined as the length of the α-helix. The origin of the (f, g, h) coordinate system, denoted O_H, is the midpoint of the AB line segment.

That set which results in the smallest value of $\mathcal{S}$ upon substitution in eq 7 should then be used in eq 1 to describe the least-squares axis.

$$\mathcal{S} = \sum_{i=1}^{n_r}\{(x_i - x^*)^2 + (y_i - y^*)^2 + (z_i - z^*)^2 - [l(x_i - x^*) + m(y_i - y^*) + n(z_i - z^*)]^2\} = (1 - l^2)\delta_{xx} + (1 - m^2)\delta_{yy} + (1 - n^2)\delta_{zz} - 2[lm\delta_{xy} + ln\delta_{xz} + mn\delta_{yz}] \qquad (7)$$

Helical Coordinate System (f, g, h). The coordinate system is based on the definition of the helix axis as the h axis, as described above. The origin of the (f, g, h) coordinate system is chosen as the midpoint of the helix. It is defined as the point O_H on the axis that is in the middle between points A and B, the projections of the first and last C^α atom of the helix on the h axis (Figure 2).[25] In the (x, y, z) coordinate system, the coordinates of the point O_H can be written as

(20) Chou, P. Y.; Fasman, G. D. *Biochemistry* **1974**, *13*, 211–222.
(21) Tanaka, S.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 142–159.
(22) Maxfield, F. R.; Scheraga, H. A. *Biochemistry* **1976**, *15*, 5138–5153.
(23) Sugeta, H.; Miyazawa, T. *Biopolymers* **1967**, *5*, 673–679.
(24) McGuire, R. F.; Vanderkooi, G.; Momany, F. A.; Ingwall, R. T.; Crippen, G. M.; Lotan, N.; Tuttle, R. W.; Kashuba, K. L.; Scheraga, H. A. *Macromolecules* **1971**, *4*, 112–124.

(25) It should be noted that the location of the origin O_H and the definition of the f axis adopted here differ from that used in the first part of this series[4] in which an arbitrary reference atom (chosen as the peptide nitrogen of the fourth residue) was used to define the origin and the f axis. The present method is more general because it is adaptable to other structures as well, such as the β-sheet, and because it is less sensitive to irregularities of the helix and to changes of dihedral angles of the helix in the course of energy minimization.

$$x_H^0 = x^* + lq \qquad y_H^0 = y^* + mq \qquad z_H^0 = z^* + nq \tag{8}$$

where

$$q = [(x_1 + x_{n_r})/2 - x^*]l + [(y_1 + y_{n_r})/2 - y^*]m + [(z_1 + z_{n_r})/2 - z^*]n \tag{9}$$

The direction of the h axis in a general (x, y, z) coordinate system is defined by the unit vector $\mathbf{e}_h$, given in the equation:

$$\mathbf{e}_h = l\mathbf{e}_x + m\mathbf{e}_y + n\mathbf{e}_z \tag{10}$$

where $\mathbf{e}_x$, $\mathbf{e}_y$, and $\mathbf{e}_z$ are unit vectors pointing along the coordinate axes x, y, and z, respectively. The direction of the f axis of the helical coordinate system[25] is defined by the unit vector $\mathbf{e}_f$, given by:

$$\mathbf{e}_f = (\mathbf{e}_h \times \overrightarrow{C_i^\alpha C_j^\alpha})/|\mathbf{e}_h \times \overrightarrow{C_i^\alpha C_j^\alpha}| \tag{11}$$

where

$$\overrightarrow{C_i^\alpha C_j^\alpha} = (x_j - x_i)\mathbf{e}_x + (y_j - y_i)\mathbf{e}_y + (z_j - z_i)\mathbf{e}_z \tag{12}$$

is a vector pointing from the ith to the jth C^α atom. In principle, any pair of residues along the helix can be chosen as i and j. It is preferable, however, to choose both of them near the middle of the α-helix, as shown in Figure 2. If they are chosen near the two ends of the helix, there is the possibility that a slight distortion of the helix (arising from variation of the backbone dihedral angles ϕ, ψ, ω) would cause an abrupt change of the direction of the coordinate axes. This should be avoided. In the work reported here, $i = 6$ and $j = 7$. Finally, the direction of the g axis is defined by the unit vector $\mathbf{e}_g$, given by

$$\mathbf{e}_g = \mathbf{e}_h \times \mathbf{e}_f \tag{13}$$

so as to form a right-handed coordinate system.

The coordinate transformation from the (x, y, z) system to the (f, g, h) system is given by

$$\begin{bmatrix} f \\ g \\ h \end{bmatrix} = \mathcal{L}_H \begin{bmatrix} x - x_H^0 \\ y - y_H^0 \\ z - z_H^0 \end{bmatrix} \tag{14}$$

where

$$\mathcal{L}_H = \begin{bmatrix} (\mathbf{e}_f)_x & (\mathbf{e}_f)_y & (\mathbf{e}_f)_z \\ (\mathbf{e}_g)_x & (\mathbf{e}_g)_y & (\mathbf{e}_g)_z \\ (\mathbf{e}_h)_x & (\mathbf{e}_h)_y & (\mathbf{e}_h)_z \end{bmatrix} \tag{15}$$

with $(\mathbf{e}_f)_x$ denoting the projection of the $\mathbf{e}_f$ unit vector on the x axis, etc. The inverse transformation, from the (f, g, h) system to the (x, y, z) system, is given by

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathcal{L}_H^{-1} \begin{bmatrix} f \\ g \\ h \end{bmatrix} + \begin{bmatrix} x_H^0 \\ y_H^0 \\ z_H^0 \end{bmatrix} \tag{16}$$

where $\mathcal{L}_H^{-1}$ is the inverse matrix of $\mathcal{L}_H$.

Throughout this paper, the subscript H denotes transformations referring to the coordinate system based on a helix axis, as described above. This notation is introduced here in anticipation of its extension to transformations referring to β-sheet structures.[19]

Location of the Second Helix in a General Position. As described in ref 4, the coordinates of all atoms of a given helix in a *general position* can be generated, starting from the coordinates of the helix placed in the *reference position*, i.e., along the h axis of the helical (f, g, h) coordinate system (Figure 3), by a rotational operation followed by a translational operation, given by

$$\mathbf{r}_2 = \mathbf{T}_H + \mathbf{R}_H\mathbf{r}_2' \tag{17}$$

In eq 17, the vector $\mathbf{r}_2'$ denotes the coordinates of any atom of the helix before rotation and translation, i.e., in the reference position as described in the next paragraph (helix H_2' of Figure 3), $\mathbf{r}_2$ denotes the coordinates of the same atom in the general position of the helix (helix H_2 of Figure 3) in the (f, g, h) coordinate system, and $\mathbf{R}_H$ is the Euler rotational operator[4,26,27] given in eq 18, where α, β, γ are the Euler angles. The translational

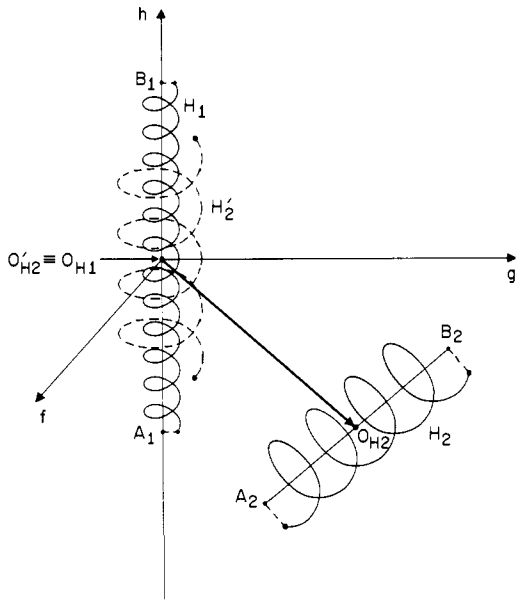

Figure 3. Determination of the position and orientation of a second helix in the (f, g, h) coordinate system. This coordinate system is defined in terms of helix H_1, as shown in Figure 2. The second helix is in the reference position, denoted by the dotted line H_2', when $\alpha = \beta = \gamma = t_f = t_g = t_h = 0$. In this case, the midpoints O_{H1} and O'_{H2}, and the axes of helix H_1 and H_2' coincide. When the second helix is in a general position, denoted by H_2, the translational operation $\mathbf{T}_H$ is given by the vector $\overrightarrow{O_{H1}O_{H2}}$ connecting the midpoints of the two helices H_1 and H_2, and the rotational operation $\mathbf{R}_H$ is defined in terms of the Euler angles in the (f, g, h) coordinate system (see Figure 2 of ref 4). The lengths of helices 1 and 2 are given by the distances A_1B_1 and A_2B_2, respectively (see Figure 2).

$$\mathbf{R}_H = \begin{bmatrix} \cos\alpha\cos\gamma - \sin\alpha\cos\beta\sin\gamma & -\cos\alpha\sin\gamma - \sin\alpha\cos\beta\cos\gamma & \sin\alpha\sin\beta \\ \sin\alpha\cos\gamma + \cos\alpha\cos\beta\sin\gamma & -\sin\alpha\sin\gamma + \cos\alpha\cos\beta\cos\gamma & -\cos\alpha\sin\beta \\ \sin\beta\sin\gamma & \sin\beta\cos\gamma & \cos\beta \end{bmatrix} \tag{18}$$

operator $\mathbf{T}_H$ is given by the vector $\overrightarrow{O'_{H1}O_{H2}}$ connecting the midpoint O'_{H2} of the helix H_2' in the reference position with the midpoint of helix H_2 in the general position (Figure 3). By definition, O'_{H2} coincides with the origin O_H of the (f, g, h) coordinate system (given by the point O_{H1}) and therefore $\mathbf{T}_H$ is given by eq 19, where t_f, t_g, and t_h are the components, along the

$$\mathbf{T}_H = \overrightarrow{O_{H1}O_{H2}} = \begin{bmatrix} t_f \\ t_g \\ t_h \end{bmatrix} \tag{19}$$

three-coordinate axes, of the translation operator $\mathbf{T}_H$ (cf. Figure 2 of ref 4).[28]

When helices H_1 and H_2 are identical, as in the derivation given in our first paper,[4] helix $\mathbf{H}_2'$ in the reference position is identical with helix H_1; i.e., atomic coordinates of H_2 in the general position can be generated from the atomic coordinates of helix H_1, by applying eq 17. When the two helices are not identical in structure, this procedure is not feasible. Instead, it is first necessary to generate the second helix in the reference position H_2', by placing

(26) Chou, K. C.; Pottle, M. S.; Némethy, G.; Ueda, Y.; Scheraga, H. A. *J. Mol. Biol.* **1982**, *162*, 89–112.

(27) Chou, K. C.; Scheraga, H. A. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 7047–7051.

(28) The numerical values of the components of the rotation and translation operators, i.e., the elements $\mathbf{R}_H$ and $\mathbf{T}_H$, depend on the choice of the coordinate system. Therefore, for a given relative orientation of two helices,[25] the values of α, β, γ, t_f, t_g, and t_h obtained by the transformations described in the present paper differ numerically from the corresponding parameters based on the definitions used in ref 4.

this helix into the coordinate system (f, g, h) (defined in terms of helix 1) in such a way that the least-squares axis h_2 of H_2' coincides with the h axis, and axes f_2 and g_2, defined for this helix in the manner described in the preceding section, coincide with the corresponding axes of the (f, g, h) system. Once this has been done, it is possible to generate helix H_2 in the desired position and orientation (defined by α, β, γ, t_f, t_g, t_h) by applying eq 17 to the coordinates of H_2'.

Parameters Defining the Relative Position and Orientation of Two Helices. In paper 1,[4] we have shown that the *relative* position and orientation of the *axes* of two helices can be specified (except for the displacement of each helix along its own axis) by using only two parameters, viz., the distance D of closest approach between the two helix axes[8] and an angle describing the relative orientation of the two helix axes. The latter was expressed in two alternate forms, either as Ω_o, the angle of orientation between the axes of the two helices, or as Ω_p, the angle between the projections of the axes of the two helices onto the "contact plane"[8] between them (defined in ref 4). We give here the definitions of these parameters, with the necessary modifications from the earlier definitions,[4] in order to render them applicable to both identical and nonidentical helices. D, Ω_o, and Ω_p do not depend on the choice of coordinate system used to define the helices.

(i) **Orientation Angle Ω_o.** This parameter is defined, as before,[4] as the angle between the two helix axes (see Figure 2b of ref 4).

$$\begin{aligned}\Omega_o &= \beta \text{ for } p > 0\\ &= -\beta \text{ for } p < 0\end{aligned} \tag{20}$$

where

$$\begin{aligned}p &= (t_f \cos\alpha + t_g \sin\alpha)/(t_f^2 + t_g^2)^{1/2}\\ &= (t_f \cos\alpha + t_g \sin\alpha)/t_r\end{aligned} \tag{21}$$

is the projection factor, required to specify the sign of Ω_0 (as described in Appendix B of ref 4), and $t_r = (t_f^2 + t_g^2)^{1/2}$.

(ii) **Projected Torsion Angle Ω_p.** This quantity can be used[4] as an alternative to Ω_o for comparisons with earlier work.[8] It has been defined and described in physical terms earlier.[4] Its definition is the same for identical and nonidentical helices,[29] and is given by eq 22. The numerical values of Ω_p and Ω_o differ from each other, except when the projection factor $p = 1$.

$$\begin{aligned}\Omega_p &= 0 \quad \text{for } p = 0\\ &= \tan^{-1}(p\tan\beta) \quad \text{for } -90° \le \beta \le 90°\\ &= \tan^{-1}(p\tan\beta) + \frac{p}{|p|}180° \quad \text{for } 90° < \beta \le 180°\\ &= \tan^{-1}(p\tan\beta) - \frac{p}{|p|}180° \quad \text{for } -180° \le \beta < 90°\end{aligned} \tag{22}$$

(iii) **The Distance, D, of Closest Approach between the Helix Axes.** The length L of an α-helix is given by the distance between the projections of the first and last C^α atoms on the helix axis, i.e., the distance AB (Figure 2). If the lengths of helices H_1 and H_2 are L_1 and L_2, respectively, then it is possible to define[4] a function $f(S_1,S_2)$ that can be used to calculate D, as shown by

$$\begin{aligned}f(S_1,S_2) = \{&[t_f + (S_2L_2 - L_1/2)\sin\alpha\sin\beta]^2 +\\ &[t_g - (S_2L_2 - L_1/2)\cos\alpha\sin\beta]^2 + [S_1L_1 - L_1/2 - t_h -\\ &(S_2L_2 - L_1/2)\cos\beta]^2\}^{1/2}\end{aligned} \tag{23}$$

where S_1 and S_2 are variable parameters,[4] describing the positions of points along the helix axes. $S_1 = 0$ corresponds to point A_1, $S_1 = 1$ corresponds to point B_1, $0 \le S_1 \le 1$ describes points lying on the line segment A_1B_1, i.e., points defined to fall within the helix, and $S_1 < 0$ or $S_1 > 1$ describe points located on the axis outside the helix, i.e., on extensions of the line segment A_1B_1 (Figure 3). Analogous relations hold for S_2 with respect to the axis of helix 2. Equation 23 corresponds to eq 22 of ref 4, and it has been modified only to take into account that L_1 and L_2 generally differ from each other.

When the axes of H_1 and H_2 are parallel or antiparallel, the distance of closest approach is shown in eq 24. When the axes

$$\begin{aligned}D = (t_f^2 + t_g^2)^{1/2} = t_r \quad &\text{for } -(L_1 + L_2)/2 \le t_h \le (L_1 + L_2)/2, \text{ and } \beta = 0°, \text{ or } \beta = \pm 180°\\ = f(1,0) \quad &\text{for } t_h > (L_1 + L_2)/2 \text{ and } \beta = 0°\\ = f(0,1) \quad &\text{for } t_h < -(L_1 + L_2)/2, \text{ and } \beta = 0°\\ = f(1,1) \quad &\text{for } t_h > (L_1 + L_2)/2, \text{ and } \beta = \pm 180°\\ = f(0,0) \quad &\text{for } t_h < -(L_1 + L_2)/2, \text{ and } \beta = \pm 180°\end{aligned} \tag{24}$$

of H_1 and H_2 are not parallel or antiparallel, D is given by eq 25.

$$\begin{aligned}D = f(S_1^0,S_2^0) \quad &\text{for } (0 \le S_1^0 \le 1, 0 \le S_2^0 \le 1)\\ = f(0,0) \quad &\text{for } (\delta_1 < 0, \delta_2 < 0)\\ = f(0,\delta_2) \quad &\text{for } (S_1^0 < 0, 0 \le \delta_2 \le 1)\\ = f(0,1) \quad &\text{for } (\Delta_1 < 0, \delta_2 > 1)\\ = f(\delta_1,0) \quad &\text{for } (S_2^0 < 0, 0 \le \delta_1 \le 1)\\ = f(\Delta_1,1) \quad &\text{for } (S_2^0 > 1, 0 \le \Delta_1 \le 1)\\ = f(1,0) \quad &\text{for } (\delta_1 > 1, \Delta_2 < 0)\\ = f(1,\Delta_2) \quad &\text{for } (S_1^0 > 1, 0 \le \Delta_2 \le 1)\\ = f(1,1) \quad &\text{for } (\Delta_1 > 1, \Delta_2 > 1)\end{aligned} \tag{25}$$

Terms are defined in eq 26. Equations 24 to 26 are analogous

$$\begin{aligned}S_1^0 &= \left[(t_g\cos\alpha - t_f\sin\alpha)\cos\beta + \left(t_h + \frac{L_1}{2}\right)\sin\beta\right] \Big/ (L_1\sin\beta)\\ S_2^0 &= \left(t_g\cos\alpha - t_f\sin\alpha + \frac{L_2}{2}\sin\beta\right) \Big/ (L_2\sin\beta)\\ \delta_1 &= \left(\frac{L_1}{2} - \frac{L_2}{2}\cos\beta + t_h\right) \Big/ L_1\\ \Delta_1 &= \left(\frac{L_1}{2} + \frac{L_2}{2}\cos\beta + t_h\right) \Big/ L_1\\ \delta_2 &= \left[\frac{L_2}{2} - \frac{L_1}{2}\cos\beta + (t_g\cos\alpha - t_f\sin\alpha)\sin\beta - t_h\cos\beta\right] \Big/ L_2\\ \Delta_2 &= \left[\frac{L_2}{2} + \frac{L_1}{2}\cos\beta + (t_g\cos\alpha - t_f\sin\alpha)\sin\beta - t_h\cos\beta\right] \Big/ L_2\end{aligned} \tag{26}$$

to eq 26 to 32 of ref 4. Certain terms in the equations differ from those given earilier,[4] because the α-helices may have different lengths and because of the difference in the choice of origin[25] between the earlier[4] and present work. The present definition of the origin along the helix axis results in simpler and more generally useful expressions for the distance D.

Definition of Contacts. The presence of contacts between atoms or residues, respectively, is defined on the basis of sums of atomic van der Waals radii, as described in ref 4.

Computational Methods

The atomic coordinates of each α-helix were generated and its intrachain energy was computed by means of the ECEPP (Empirical Conformational Energy Program for Peptides) algorithm.[30]

(29) In ref 4, eq 12, defining Ω_p, contained some typographical errors and did not include the definition for $p = 0$. The corrected form, shown here as eq 22, applies to both equivalent and nonequivalent helices.

(30) Momany, F. A.; McGuire, R. F.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1975**, *79*, 2361–2381.

Some of the parameters used in this algorithm have been updated recently.[31] The revised version, ECEPP/2, has been used in the computations reported here.

The intrachain energy of each polypeptide chain is computed in both ECEPP and ECEPP/2 as the sum of nonbonded, electrostatic, hydrogen bond, and torsional energies.[30-32] The interchain energy between two α-helices is computed as the sum of pairwise nonbonded and electrostatic energies between all atoms of the two helices. The parameters for these pairwise energies are the same as those used in ECEPP/2. The total energy E_{tot} is the sum of the intrachain energies of each of the helices and of the interchain energy. The computer program used here is based on the ECEPP/2 algorithm,[30,31] augumented by an algorithm to generate two helices in given relative positions and orientations (see preceding section and refs. 4 and 26). The program computes and minimizes E_{tot} with respect to all backbone and side-chain dihedral angles of both helices as well as the six external parameters α, β, γ, t_f, t_g, t_h.[33] Optionally, it is possible to keep any set of the dihedral angles constant, or to restrain corresponding dihedral angles in each residue of a given helix to vary synchronously with each other (in computations on regular helices). Energy minimization was carried out using the nonlinear least-squares algorithm of Dennis et al.[34] The computations were carried out on a Prime 350 minicomputer with an attached Floating Point Systems array processor.[35]

The standard conventions for the nomenclature of polypeptide conformations are followed.[36] For brevity, an α-helix is symbolized here by $\alpha\{X-(R)_{n_r}-Y\}$, where X and Y denote end groups, R represents any amino acid residue, and n_r is the number of residues in the helix.

Selection of Starting Points for Energy Minimization. In the work reported here, the energy was minimized (a) for a pair of $\alpha\{CH_3CO-(L\text{-}Ala)_{10}-NHCH_3\}$ molecules and (b) for an $\alpha\{CH_3CO-(L\text{-}Ala)_{10}-NHCH_3\}$ molecule interacting with an $\alpha\{CH_3CO-(L\text{-}Leu)_{10}-NHCH_3\}$ molecule.

In case a, only one starting orientation was used here, viz., the lowest energy packing arrangement of two identical poly(L-Ala) α-helices, as computed in ref 4. The energy was minimized with respect to various sets of the dihedral angles and external parameters used as variables, as described in the Results section.

In case b, energy minimization was carried out in three sucessive steps.

(i) The intrachain energy of each isolated α-helix was minimized with respect to the dihedral angles ϕ, ψ, and χ^1 for both helices, as well as χ^2 for Leu, while ω was fixed at 180°, and the side-chain dihedral angles defining the rotation of CH_3 groups (viz. χ^1 in Ala, $\chi^{3,1}$ and $\chi^{3,2}$ in Leu and those in the end groups) were fixed at 60°. The chain was constrained to be regular; i.e., the variable dihedral angles ϕ, ψ, and, in the case of Leu, χ^1 and χ^2, were the same in all residues throughout the minimization.

(ii) The interchain energy of the two α-helices was minimized with respect to the six external variables, keeping the dihedral angles fixed at the values obtained in step (i) above. $\alpha\{CH_3CO-(L\text{-}Leu)_{10}-NHCH_3\}$ was considered as helix H_1; i.e., it was used to define the (f, g, h) coordinate system (Figure 3). $\alpha\{CH_3CO-(L\text{-}Ala)_{10}-NHCH_3\}$ was considered as helix H_2. It was placed in 48 different starting orientations, chosen as follows, and the energy was minimized from each of these orientations. In all starting orientations, $t_g = t_h = 0.0$ Å, $t_f = 14.28$ Å which is the sum of the maximal peripheral radii of the two α-helices,[37] $\alpha = 0°$, while γ was varied in 30° steps from 0 to 330°, and β was chosen to be 0, 180, 90, and −90°, corresponding to the parallel, the antiparallel, and the two perpendicular orientations of the helix axes, respectively.

(iii) Starting from the 10 lowest energy packing arrangements found in step ii, the energy was further minimized, by considering all backbone and side-chain dihedral angles including all ω's (except those in the end groups) and the external parameters as independent variables. Thus, the energy was a function of 116 variables: 70 dihedral angles in $\alpha\{CH_3CO-(L\text{-}Leu)_{10}-NHCH_3\}$, 40 dihedral angles in $\alpha\{CH_3CO-(L\text{-}Ala)_{10}-NHCH_3\}$, as well as the 6 external variables.

Results and Discussion

Packing of Poly(L-Ala) α-Helices. In the first part of this study,[4] the packing of two equivalent α-helices with fixed backbone dihedral angles had been investigated. Before carrying out the studies reported below, it was necessary to assess whether the restrictions, viz., equivalence and fixed backbone conformation, affect the packing significantly. Furthermore, the effect of the recent changes in the ECEPP/2 parameters[31] had to be assessed. Computational tests were carried out on the computed lowest energy packing state of two equivalent $\alpha\{CH_3CO-(L\text{-}Ala)_{10}-NHCH_3\}$ helices reported earlier.[4] This packing state had been characterized[4] by $\Omega_o = 153.5°$, an intrachain energy of 15.30 kcal/mol per chain,[32] and an interchain energy of −17.23 kcal/mol, with the backbone dihedral angles fixed at the values (ϕ, ψ, ω) = (−57, −47, 180°) for all residues in both helices. This is a helical conformation often quoted as a reference state for poly(L-Ala) α-helices,[36] but it is not an energy-minimized single-chain structure.

The recalculation of the energy for the same dihedral angles, using the revised (ECEPP/2) parameters, changes the computed numerical value of the intrachain energy to −11.54 kcal/mol per chain (a change that has no physical significance).[32,38] The interchain energy and the distance of closest approach are not altered at all. Thus, none of the conclusions reported earlier[4] are modified by the updating of the ECEPP parameters to ECEPP/2.

After energy minimization on an isolated regular poly(L-Ala) α-helix, its dihedral angles become (ϕ, ψ, ω) = (−68.1, −38.3, 180°) and the intrachain energy is lowered to −21.47 kcal/mol. These dihedral angles represent the energetically most favorable α-helix with the ECEPP/2 geometry and energy parameters. This α-helix is used henceforth as the reference helix in these studies. The lowest minimum-energy packing of two such α-helices occurs with an orientational torsion angle $\Omega_o = -149.5°$ and an interchain energy of −17.50 kcal/mol; i.e., these parameters differ only slightly from those reported earlier[4] and summarized in the first paragraph of this section. If, in this structure, the backbone dihedral angles (ϕ, ψ, ω) are allowed to vary but the condition of regularity is maintained (i.e., all residues in a given chain have the same dihedral angles), intra- and interchain energies change only by 0.2 kcal/mol and the dihedral angles by 0.1°; i.e., the packing is virtually the same as before.

Finally, the condition of regularity was relaxed; i.e., all dihedral angles of both α-helices as well as the external variables were allowed to vary freely during energy minimization. This condition

(31) Némethy, G.; Pottle, M. S.; Scheraga, H. A. *J. Phys. Chem.* **1983**, *87*, 1883-1887.

(32) Because of the emission of certain energy terms that are independent of the conformation in ECEPP and ECEPP/2, only the *relative* values of the intrachain energy for a given polypeptide have physical significance.[30] Intrachain energies for polypeptides with different amino acid sequences cannot be compared with each other. For the same reason, the conformationally independent change of the zero level of the intrachain energy between ECEPP and ECEPP/2 is of no consequence.[31]

(33) The actual computer operations during energy minimization are carried out in a fixed (x, y, z) coordinate system, as defined in ECEPP,[30] for computational ease, because the helical coordinate system, described in the section "Mathematical Formulation" depends on the (variable) dihedral angles of helix 1, and hence its use would entail unnecessary added computational steps. After minimization, the results are expressed in terms of the helical coordinate system (f, g, h) described here.

(34) Dennis, J. E.; Gay, D. H.; Welsch, R. E. *ACM Trans. Math. Software* **1981**, *7*, 369-383.

(35) Pottle, C.; Pottle, M. S.; Tuttle, R. W.; Kinch, R. J.; Scheraga, H. A. *J. Comput. Chem.* **1980**, *1*, 46-58.

(36) IUPAC-IUB Commission on Biochemical Nomenclature: *Biochemistry* **1970**, *9*, 3471-3479.

(37) The peripheral radius r_p of any atom j of an α-helix is defined as $r_p = R_j + r_j$, where R_j is the distance of atom j from the helix axis and r_j is its van der Waals radius (given in Table II of ref 4). The maximal peripheral radius of an α-helix is defined as the highest value taken by any r_p.

(38) The zero of the intrachain energy scale is different in the earlier[4] and present studies, because of the change in the ECEPP/2 parameters.[30,31] Therefore, values of E_{intra} given here cannot be compared with those of ref 4.

resulted in essentially the same packing, with only small changes in the parameters characterizing the structure. The dihedral angles are slightly different for various residues along the chains, but they deviate by less than ±2° from $(\phi, \psi, \omega) = (-68, -38, 180°)$, i.e., the values cited above for the energy-minimized isolated regular α-helix. the orientational and projected torsional angles are $\Omega_o = -151.3°$ and $\Omega_p = -151.5°$, respectively (as compared with $\Omega_o = -153.5°$ and $\Omega_p = -155.0°$ for the constrained equivalent structures reported earlier[4]). The intrachain energies are −21.69 and −21.67 kcal/mol for the two helices; i.e., they are only 0.2 kcal/mol lower than for regular α-helices, and the interchain energy is −18.10 kcal/mol, i.e., 0.9 kcal/mol lower than in the structure reported earlier.[4] The distance of closest approach between the helix axes is $D = 7.7$ Å, nearly the same as the value of 7.6 Å in the regular structure reported earlier.[4] The effect of relaxation of the condition of regularity was tested also on several of the other packing arrangements reported earlier.[4] In all of them, the relative energies ΔE, the torsional angles, and the distances between the helical axes (not shown here) were closely similar to those computed earlier (Table III of ref 4) for the regular α-helices.

In summary, the relaxation of the conditions of equivalence and regularity of two poly(L-Ala) α-helices does not alter the geometry and energetics of their packing arrangements significantly. In particular, the most favorable packing arrangement computed earlier[4] is unchanged. The conclusions reached in our previous study[4] can therefore be generalized to the packing of nonequivalent poly(L-Ala) α-helices.

The Packing of a Poly(L-Ala) and a Poly(L-Leu) α-Helix. The effect of substitution of a large side chain on an α-helix was investigated by computing the stable packing arrangements of $\alpha\{CH_3CO-(L\text{-}Ala)_{10}-NHCH_3\}$ with $\alpha\{CH_3CO-(L\text{-}Leu)_{10}-NHCH_3\}$. Energy minimization was carried out in three steps, as described in the section on Computational Methods.

In the first step, the minimization of the energy of an isolated *regular* poly(L-Leu) α-helix results in the lowest energy conformation $(\phi, \psi, \omega, \chi^1, \chi^2, \chi^{3,1}, \chi^{3,2}) = (-66.5, -39.6, 180, 176.7, 54.5, 60, 60°)$, with an intrachain energy of −41.66 kcal/mol. This conformation was used in the subsequent steps, in order to obtain various packing arrangements. If all Leu side chains are placed simultaneously into any other rotameric state with respect to χ^1 and χ^2, and the energy is minimized, then the energy of the resulting minimized conformations of the isolated α-helix is at least 20 kcal/mol of helix above the one reported here (i.e., at least 2 kcal/mol per residue). Since the most favorable total interchain energy, gained from packing, is of the order of −17 kcal/mol and its variation for different packings is less than 5 kcal/mol (Table I), it is unlikely that other Leu side-chain conformations would result in packings that have lower energies than those reported in Table I, and hence they can be disregarded here. This conclusion was verified by means of computational tests, as described below.

After the application of steps ii and iii, 10 low-energy packing arrangements were obtained within an energy interval of 5 kcal/mol. The parameters characterizing these packing states are summarized in Table I. The backbone and side-chain dihedral angles vary slightly along each helix and in the various packing states, but they remain within a few degrees of those found for the isolated α-helices, reported above. The ranges of the dihedral angles are as follows. For $\alpha\{CH_3CO-(L\text{-}Leu)_{10}-NHCH_3\}$, $\phi = -66 \pm 4°$, $\psi = -39 \pm 4°$, $\omega = 180 \pm 3°$, $\chi^1 = 176 \pm 3°$, $\chi^2 = 65 \pm 4°$, $\chi^{3,1} = 60 \pm 6°$, $\chi^{3,2} = 60 \pm 5°$; for $\alpha\{CH_3CO-(L\text{-}Ala)_{10}-NHCH_3\}$, $\phi = -68 \pm 3°$, $\psi = -38 \pm 4°$, $\omega = 180 \pm 3°$. As an illustration of the extent of variation occurring within one packed structure, all dihedral angles of the lowest energy packing state (line 1 of Table I) are listed in Table II. The lengths of the two helices also change slightly in the various packing states, corresponding to the small variation of the dihedral angles. The lengths are $L_1 = 15.9 \pm 0.2$ Å and $L_2 = 16.0 \pm 0.1$ Å for the poly(L-Leu) and poly(L-Ala) α-helices, respectively.

As shown in Table I, most of the energy differences between the various packing states arise from differences in the interchain

Table I. Parameters Characterizing the Computed Low-Energy Packing States[a] after Minimization of the Total Energy of an $\alpha\{CH_3CO-(L\text{-}Leu)_{1,0}-NHCH_3\}$ and an $\alpha\{CH_3CO-(L\text{-}Ala)_{10}-NHCH_3\}$ α-Helix Pair[b]

Euler angles (deg)			Translational displacements (Å)			Torsion angles (deg)		Distance of closest approach between the helix axes (Å)	Energy (kcal/mol)					No. of atom pairs in contact[g]	No. of residue pairs in contact[g]
									Total		Intrahelix[e]		Interhelix		
α	β	γ	t_f	t_g	t_h	Ω_o	Ω_p	D	E_{tot}^c	ΔE_{tot}^d	$E_{intra,1}$	$E_{intra,2}$	E_{inter}[f]	Electrostatic E_{ES}	Nonbonded E_{NB}		
-117.0	154.4	-21.1	9.0	-0.9	-2.7	-154.4	-170.1	6.2	-83.08	0.00	-43.47	-21.87	-17.74	-1.25	-16.49	126	9
-76.1	144.2	-81.2	10.2	-0.2	-1.0	144.2	169.5	5.8	-81.73	1.35	-43.04	-21.59	-17.11	-1.14	-15.97	116	9
8.3	80.9	156.1	7.0	0.5	-6.7	80.9	80.8	7.0	-81.63	1.45	-43.24	-21.58	-16.81	0.02	-16.83	128	10
-171.6	111.8	168.9	8.4	-1.4	-4.4	-111.8	-112.8	8.1	-81.36	1.72	-43.61	-21.47	-16.28	-0.50	-15.78	118	10
-75.1	32.5	-173.2	8.8	3.2	1.7	-32.5	-3.1	5.2	-81.14	1.94	-43.01	-21.61	-16.51	1.43	-17.94	141	11
23.1	27.4	102.0	8.3	-2.9	-3.4	27.4	20.9	7.2	-79.82	3.26	-43.64	-22.04	-14.13	0.75	-14.88	113	9
-7.3	127.9	109.4	8.9	0.8	-1.8	127.9	128.6	8.7	-79.71	3.37	-43.25	-21.93	-14.53	-0.82	-13.71	105	9
179.1	65.5	-34.5	8.6	1.7	-1.2	-65.0	-65.0	8.6	-79.43	3.65	-43.14	-21.87	-14.42	0.42	-14.84	104	9
173.1	145.4	-61.5	8.0	3.5	3.3	-145.4	-149.3	7.6	-79.39	3.69	-43.50	-22.00	-13.89	-0.84	-13.05	88	9
-150.8	129.6	-42.0	9.3	-2.7	-3.2	-129.6	-139.9	6.8	-78.66	4.42	-43.70	-21.90	-13.07	-0.60	-12.47	76	6

[a] Arranged in order of increasing energy. [b] The ranges of dihedral angles in each of the helices are given in the text. [c] Total energy of the system consisting of the two helices: $E_{tot} = E_{intra,1} + E_{intra,2} + E_{inter}$. [d] $\Delta E_{tot} = E_{tot} - E_{tot}^o$, where $E_{tot}^o = -83.08$ kcal/mol. [e] Total intramolecular energy of each helix. $\alpha\{CH_3CO-(L\text{-}Leu)_{10}-NHCH_3\}$ is helix 1, $\alpha\{CH_3CO-(L\text{-}Ala)_{10}-NHCH_3\}$ is helix 2. For each helix, only the relative values of this energy have a physical significance (see footnote 32). [f] Total interaction energy between the two α-helices. It is zero at infinite separation of the two α-helices. [g] These quantities were defined in ref. 4. See the discussion of their physical implication in ref. 4.

Table II. Dihedral Angles[a] of the Lowest Energy Packing State[b] of an α {CH_3CO-(L-Leu)$_{10}$-$NHCH_3$} and an α {CH_3CO-(L-Ala)$_{10}$-$NHCH_3$} α-Helix Pair

	α {CH_3CO-(L-Leu)$_{10}$-$NHCH_3$}							α {CH_3CO-(L-Ala)$_{10}$-$NHCH_3$}			
	dihedral angles (deg)							dihedral angles (deg)			
residue	ϕ	ψ	ω	χ^1	χ^2	$\chi^{3,1}$	$\chi^{3,2}$	ϕ	ψ	ω	χ^1
1	−64.1	−38.9	−179.1	177.2	63.4	54.9	59.8	−66.8	−36.3	179.1	60.7
2	−65.7	−39.9	−178.5	177.2	63.6	55.0	59.8	−65.1	−41.3	−177.2	60.4
3	−68.7	−38.4	178.6	177.9	62.3	54.3	58.8	−69.1	−39.1	−179.7	61.2
4	−64.6	−42.6	−178.9	178.8	66.5	55.7	60.5	−67.2	−39.1	−179.5	60.5
5	−65.8	−40.3	−178.8	176.9	63.9	55.1	59.8	−66.9	−40.9	−179.0	60.4
6	−66.0	−40.8	−179.3	177.2	64.0	55.1	59.8	−65.9	−41.5	−177.6	60.6
7	−66.3	−40.1	−179.7	178.3	67.4	55.8	60.0	−67.6	−39.4	−177.8	60.6
8	−66.1	−40.9	−178.8	177.4	64.9	55.3	59.9	−70.1	−39.4	−178.5	62.0
9	−66.5	−39.6	179.6	177.5	65.0	55.4	59.9	−67.3	−39.1	178.5	60.3
10	−66.0	−41.0	178.9	177.6	65.1	55.3	59.9	−68.2	−39.0	178.4	60.5
average	−66.0	−40.3	−179.6	177.6	64.6	55.2	59.8	−67.4	−39.5	−179.3	60.7

[a] The standard definitions for dihedral angles in peptides are used.[36] [b] Line 1 of Table I.

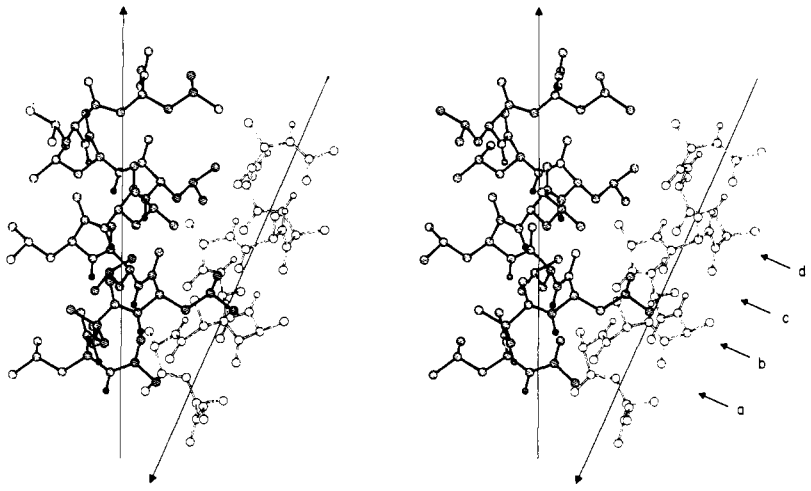

Figure 4. Stereoscopic pictures of an α{CH_3CO-(L-Leu)$_{10}$-$NHCH_3$} (shaded atoms and bonds) and an α{CH_3CO-(L-Ala)$_{10}$-$NHCH_3$} (open atoms and bonds) α-helix in the lowest energy packing state (line 1 of Table I), with $\Omega_0 = -154.4°$. The two helices are nearly antiparallel. The helix axes are indicated by arrows, with the head of the arrow pointing in the direction of the C terminus of each helix. Hydrogen atoms are omitted, except for the amide hydrogens. The arrows marked a, b, c, and d indicate regions of intercalation of the two α-helices, as described in the text.

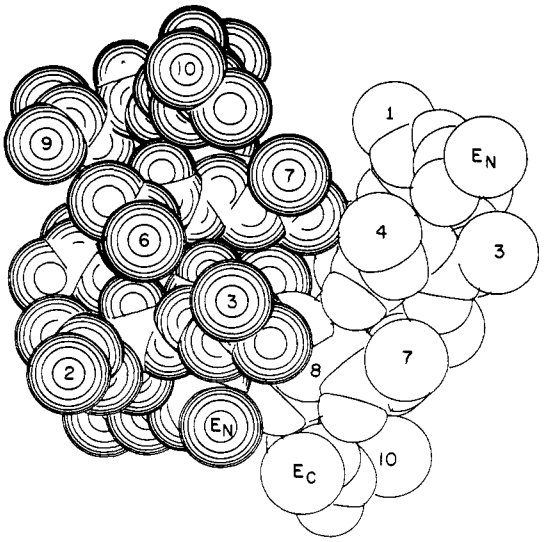

Figure 5. Space-filling representation[39] of the structure shown in Figure 4, indicating the contacts between the surfaces of the two α-helices. The shading corresponds to that of Figure 4. The direction of viewing is nearly the same in Figures 4 and 5. Numbers denote the residues along the α-helix which are visible in this view. The N- and C-terminal groups (where they are visible) are denoted E_N and E_C, respectively. Only heavy atoms are shown, with approximate van der Waals radii.

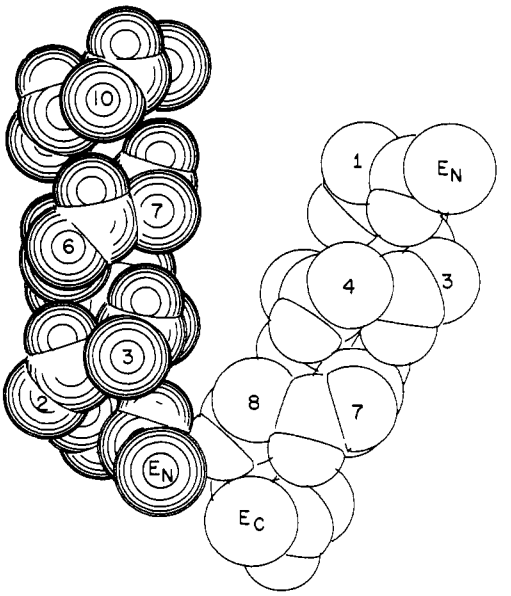

Figure 6. Space-filling representation[39] of the backbone heavy atoms of the structure shown in Figures 4 and 5. The direction of viewing is the same as in Figure 5. All side-chain atoms have been omitted in order to indicate the distance and relative orientation of the backbones of the two helices.

energy E_{inter}. The intrachain energy (E_{intra}) for each of the two helices varies much less, staying within a range of only 0.7 kcal/mol.

The structures of the lowest energy packing state, with a nearly antiparallel orientation of the helix axes (line 1 of Table I), and one low-energy packing with nearly perpendicular orientation of

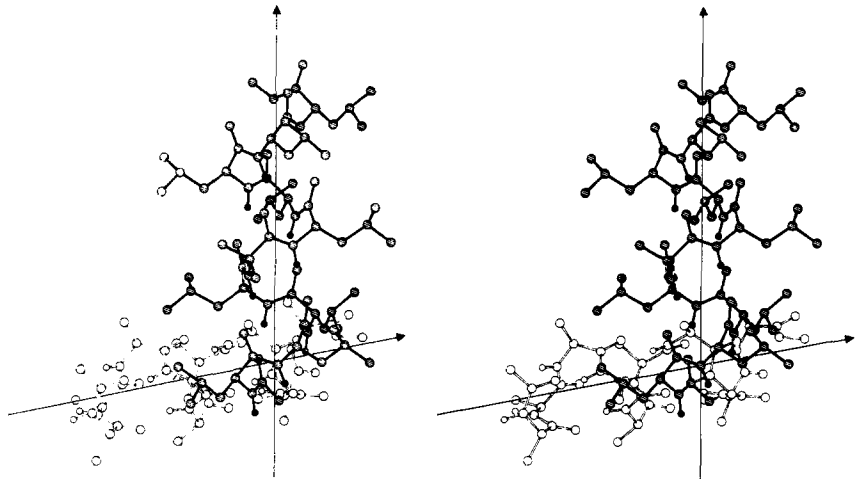

Figure 7. Stereoscopic picture of an α-{CH_3CO-(L-Leu)$_{10}$-$NHCH_3$} (shaded atoms and bonds) and an α-{CH_3CO-(L-Ala)$_{10}$-$NHCH_3$} (open atoms and bonds) α-helix in a low-energy near-perpendicular packing state (line 3 of Table I), with $\Omega_0 = 80.9°$. See Figure 4 for further explanation.

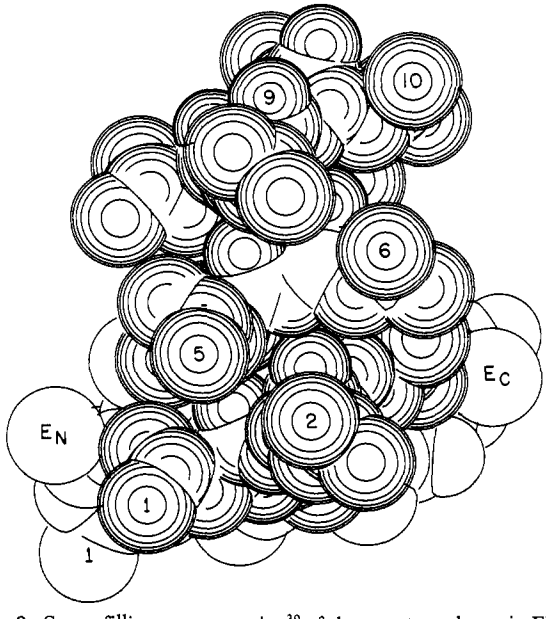

Figure 8. Space-filling representation[39] of the structure shown in Figure 7. See Figure 5 for further explanation.

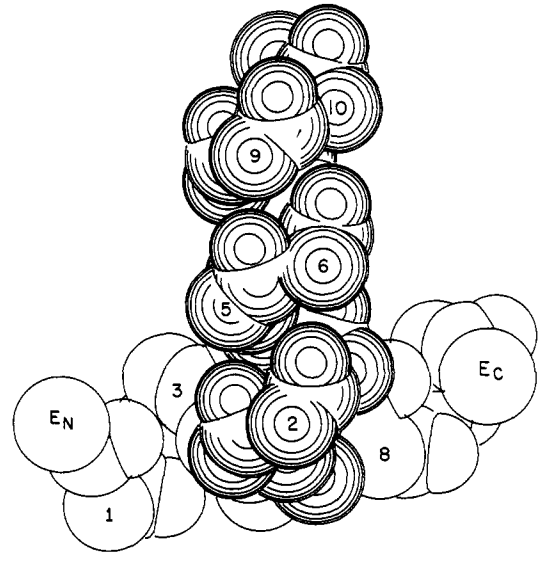

Figure 9. Space-filling representation of the backbone heavy atoms and C^β atoms of the structure shown in Figures 7 and 8. See Figure 6 for further explanation.

the helix axes (line 3 or Table I), are shown in Figures 4–9. Figures 4 and 7 are stereoscopic ORTEP representations of the molecular structures. For the near-antiparallel arrangement, it is discernible at the bottom of Figure 4 how the two helices pack approximately in the general manner of "ridges into grooves", as described by geometrical models of packing;[5,7] the last two turns of the poly(Ala) α-helix (marked a and c in the figure) are located near the "grooves" between successive turns of the poly(Leu) α-helix. Conversely, successive turns of the poly(Leu) α-helix, with the side chains extending to the right, are located near the "grooves" between successive turns of the poly(Ala) α-helix (b and d in the figure). It should be noted, however, that the "ridges into grooves" concept indicates only the approximate orientation of the axes, and the actual contact between the helices is governed by the balance of the noncovalent interaction energies between the two α-helices, as discussed earlier.[4] In Figure 5, the same packing is shown in an approximate space-filling representation,[39] in order to indicate the actual surface of the two helices and the close packing between their atoms. The packing is determined by favorable nonbonded interactions, occurring to a considerable extent between side-chain atoms in or near contact, as seen in a comparison of Figures 5 and 6. Figure 6 represents the same structure as Figure 5, but the atoms of the side chains have been omitted. Figure 6 shows that the backbone atoms of the two helices are not in contact with each other, so that most contacts involve side-chain atoms. The relative positioning of the two backbones are seen more clearly from Figure 6 than from Figure 5.

Figures 7, 8, and 9 are similar ORTEP and space-filling representations, respectively, of the near-perpendicular packing. It is seen from both Figures 4 and 7 that many of the contacts between the two helices involve the side-chain atoms of the Leu residues with the atoms of the poly(Ala) α-helix.

The tight packing of the two α-helices is indicated by the low values of the distance, D, between the axes, ranging from 5.2 to 8.7 Å. This is comparable to the range of 6.7 to 7.6 Å computed for two poly(L-Ala) α-helices,[4] and is much less than the distance when the side chains are in peripheral contact only, as discussed earlier.[4]

The conformations of the Leu side chain in all packing states listed in Table I is the same as in the most favorable side-chain conformation in an isolated poly(L-Leu) α-helix, with (χ^1, χ^2) near (180, 60°). All other side-chain conformations in the isolated α-helix have much higher energy, as described above. It might have been possible, however, that one of these other side-chain conformations would result in better packing of the two α-helices, with a favorable total energy, in spite of the higher intra-chain energy. This possibility was checked by repeating the energy

(39) Zientara, G. P.; Nagy, J. A. *Comput. Chem.* **1983**, *7*, 67–74.

Table III. Effect of Rotating Leu Side Chains on the Packing Energy of an α{CH_3CO-(L-Leu)$_{10}$-$NHCH_3$} and an α{CH_3CO-(L-Ala)$_{10}$-$NHCH_3$} α-Helix Pair

	altered side-chain conformation after minimization					
		dihedral angles[b] (deg)		energy (kcal/mol)		
Ω_0 (deg)	in residues[a]	χ^1	χ^2	E_{tot}	$E_{intra,i}$[c]	E_{inter}
−154.4[d]	3,4,7	178	65	−83.08	−43.47	−17.74
−155.6		175	140	−75.98	−37.31	−16.88
−159.0		−66	165	−77.77	−41.29	−14.68
80.9[e]	1,3,4	178	65	−81.63	−43.24	−16.81
80.8		−179	140	−73.26	−37.03	−14.80
122.7[f]		−60	171	−72.25	−40.82	−9.49

[a] Residues selected as described in the text. These residues are in contact with the poly(L-Ala) α-helix. For these residues, three different starting conformations were selected, as described in the text. After minimization, they gave the three conformations listed in the table. The side chains of all other residues were kept in the lowest energy rotameric state, i.e., were minimized from the starting conformation with $(\chi^1, \chi^2) = (-176, 65°)$. [b] Average value for the three side chains after minimization. [c] Intrachain energy of the α{CH_3CO-(L-Leu)$_{10}$-$NHCH_3$} α-helix after minimization. [d] Same as line 1 of Table I. [e] Same as line 3 of Table I. [f] Minimization resulted in a different packing arrangement in this case.

minimization with altered initial side-chain conformations, for two of the packing arrangements (viz., those of lines 1 and 3 in Table I). The backbone dihedral angles and external variables used as starting points for the energy minimization were those given in Table I, i.e., the minima obtained earlier. The initial side-chain conformations were altered for the three Leu side chains that are in contact with the poly(L-Ala) α-helix (see Figures 4 and 7), viz., residues 3, 4, and 7 for the nearly antiparallel packing (line 1 of Table I and Figure 4) and residues 1, 3, and 4 for the nearly perpendicular packing (line 3 of Table I and Figure 7). In both packings, two new conformations were used for these Leu residues, with starting side-chain dihedral angles of either $(\chi^1, \chi^2) = (176, 140°)$ or $(-54, 173°)$ obtained by energy minimization of an isolated α-helix (step i of the Methods). Minimization was then carried out with respect to the dihedral angles and external variables, according to step iii of the Methods. The results are summarized in Table III. The total energies of the resulting four new packing arrangements were all higher than those with the original choice of side-chain conformations (lines 1 and 3 of Table I). The energy differences between the new arrangements and the original ones range from 5.3 to 9.4 kcal/mol (Table III). Not only are the intrahelix energies higher (by 2.2 to 6.2 kcal/mol) in all of the new packing arrangements (as expected for energetically less favorable side-chain orientations) but the interhelix energies are also higher (by 0.9 to 7.3 kcal/mol); i.e., both intra- and interhelix energies favor the packings with side-chain conformations having (χ^1, χ^2) near (177, 65°). This is the side-chain conformation that was favored in the isolated poly(L-Leu) α-helix as well. Thus, for the pair of α-helices studies here, only one side-chain conformation results in favorable packing arrangements. It is possible, however, for the packing of α-helices containing other amino acids, that the favorable side-chain conformations might differ from those in isolated α-helices.

Comparison of Poly(L-Ala)/Poly(L-Ala) and Poly(L-Leu)/Poly(L-Ala) α-Helix Packing Arrangements. The effect of introducing a bulky alkyl side chain on an α-helix can be seen by comparing the results reported here for the poly(L-Leu)/poly(L-Ala) pair of α-helices with our earlier analysis[4] of the packing between two poly(L-Ala) α-helices.[38] The lowest energy packing arrangement in both cases is a nearly antiparallel one, with an orientational torsion angle $\Omega_0 = -154.4$ and −153.5°, respectively, i.e., with a nearly identical relative orientation of the helix axes. The projected torsion angle Ω_p differs somewhat in the two cases, changing from −155.0° for two equivalent poly(L-Ala) helices[4] to −170.1° for the packing of a poly(L-Leu)/poly(L-Ala) pair of helices. Both values are near the averaged observed angle[14] of −162° between helix axes in the bundle of four interconnected nearly antiparallel α-helices, an arrangement frequently observed in proteins.[2,8,14-18] As discussed earlier,[4] this arrangement is favored by both electrostatic (dipole–dipole) and nonbonded interactions over other packings. The introduction of the large Leu side chains does not alter this preference because the Leu side chains can be accommodated near the poly(L-Ala) α-helix without unfavorable interactions, as seen from Figure 4. The values of Ω_0 for *some* of the other packing arrangements (reported in Table I) are similar to the corresponding packing arrangements of two equivalent poly(L-Ala) helices (Table III of ref 4). For these arrangements, the relative energies, ΔE, also are similar. For example, the second-lowest packing reported here (line 2 of Table I) also is near antiparallel, with $\Omega_0 = 144.2°$, and is similar to a low-energy packing of two equivalent poly(L-Ala) α-helices. A third nearly antiparallel low-energy packing of the latter is not of low energy for the poly(L-Leu)/poly(L-Ala) pair. Apparently, it does not lead to good packing.

On the other hand, near-perpendicular packing arrangements, found occassionally in proteins,[8] are of relatively low energy (ΔE = 1.45 and 1.72 kcal/mol) for the packing involving Leu and Ala residues (lines 3 and 4 of Table I), but they are of high energy ($\Delta E > 4.5$ kcal/mol) for the packing of two poly(L-Ala) helices.[4] It appears, therefore, that these packing arrangements are more sensitive to the amino acid sequence and hence to details of the interhelix contacts than the nearly antiparallel arrangements described above.

As in the case of a pair of poly(L-Ala) α-helices,[4] the packing of the α-helices considered here is influenced by *both* nonbonded and electrostatic interactions (Table I). The magnitude of the interchain electrostatic energy E_{ES} and its variation with changing Ω_0 are similar to those seen earlier for two poly(L-Ala) α-helices (cf. the corresponding columns of Table I here and Table III in ref 4). The relative contributions of the electrostatic energy and of the nonbonded energy E_{NB} to the *energy differences* between various packing arrangements computed here are also similar in magnitude to those reported earlier.[4] Therefore, the discussion of the role of the electrostatic and nonbonded energies in the packing of poly(L-Ala) α-helices, presented earlier,[4] generally holds for the α-helix pair of this paper as well. The relative contributions of these two energy terms generally depend on the amino acid sequence.[40]

The interchain nonbonded energy for the poly(L-Leu)/poly(L-Ala) packings generally is about 1 to 4 kcal/mol lower than that of comparable packings[4] of two poly(L-Ala) α-helices. This is due to the increased extent of interactions of the Leu side chains with the poly(L-Ala) α-helix, as compared with Ala side chains. The improvement of interactions between the helices is also reflected qualitatively by the higher number of pairs of atoms in contact in the present case than for two poly(L-Ala) α-helices (cf. the last two columns of Table I here and Table III in ref 4). It must be noted, however, that the number of atoms in contact is not a quantitative measure of the strength of nonbonded interactions. The latter, E_{NB}, generally is not proportional to the number of contacts. This finding reinforces the conclusion, described earlier,[4] that an assessment of the relative stability of various polypeptide conformations cannot be based on comparisons of simple geometrical criteria (such as numbers of atoms in contact), but must use comparisons of the interaction energies.

Conclusions

The method developed in this paper can be used to describe any interacting helical assembly in proteins and other macromolecules. Thus, it applies not only to the packing of nonequivalent α-helices, but also to the packing of polypeptide strands in a β-sheet.[19]

The packing energy and geometry for two nonequivalent poly(L-Ala) α-helices in which all dihedral angles can adjust freely

(40) Gerritsen, M.; Chou, K.-C.; Némethy, G.; Scheraga, H. A., to be submitted for publication.

are closely similar to the results for two poly(L-Ala) α-helices that are identical with each other and constrained to be regular in terms of backbone dihedral angles (studied earlier[4]). This indicates that the constraints of equivalence and regularity are not severe restrictions for poly(L-Ala) model α-helices, and the conclusions obtained with the use of these restrictions can be extended to the more realistic case of poly(L-Ala) α-helices not contrained to be regular.

The introduction of bulky Leu side chains on one of the two α-helices results in slight adjustments of the relative orientation of the two helices in the low-energy packing states. The orientational torsion angles Ω_0 remain nearly the same as for poly(L-Ala). The packing is very tight, as indicated by the low values of the distance between the helix axes. In the energetically most favorable packing arrangements, the helix axes are nearly antiparallel for both the poly(Ala)/poly(Ala) and the poly(Leu)/poly(Ala) α-helix pairs. The geometrical parameters and the interhelix energies are nearly the same. Thus, the change from Ala to Leu appears to have little effect on *these* packing arrangements.

Upon comparing the results for the two kinds of pairs of α-helices, other packing arrangements can be found which are similar in terms of Ω_0. In some of these arrangements the relative energies, the values of Ω_0, or those of D are different, however. This indicates that the details of the nonantiparallel arrangements depend more sensitively on the nature of residues in contact than do those of the nearly antiparallel poly(L-Leu)/poly(L-Ala) α-helix pair. Other side chains may influence the manner of packing of α-helices differently. The packing of α-helices with actual amino acid sequences that occur in globular proteins is being investigated in more detail.[40]

The same side-chain conformation of the Leu residue is favored in an individual α-helix and in the poly(L-Leu)/poly(L-Ala) helix pair. This side-chain conformation of the bulky branched Leu side chain is sufficiently favored by intrahelix interactions (including both steric hindrance and attractive nonbonded interactions), so that added interactions with the second α-helix do not change the preferred side-chain conformation. The orientation of the side chains in this conformation does not hinder the close approach of the second α-helix, as indicated in Figures 4 and 7. The preferred side-chain conformation may change in the case of polar residues which interact with functional groups on the other α-helix.[40] Nevertheless, many basic features of the interaction and packing of α-helices in proteins can be derived and explained in terms of the interactions between pairs of poly(amino acid) α-helices. The results reported here provide further support to the principle proposed earlier,[3] viz., that it is possible to account for the main features of frequently occurring packing arrangements of regular polypeptide structures in terms of local interaction energies, without requiring the inclusion of long-range interactions.

Note Added in Proof. In a recent survey of four highly refined protein crystal structures, Blundell et al.[41] have pointed out that the mean values of the backbone dihedral angles in α-helices are close to $(\phi, \psi) = (-63°, -42°)$, with a variance of 6°, instead of the usually cited[36] reference state obtained from crystalline α-helical poly(L-Ala), viz., (-48°, -57°). The observed mean values are very close to those that were obtained in this paper by energy minimization on isolated α-helices and a pair of interacting α-helices (-66 to -68°, -38 to -40°).

Acknowledgments. We thank M. Gerritsen for helpful discussions, and M. S. Pottle, S. Rumsey, and R. W. Tuttle for assistance with computer programming and for preparation of the diagrams. This work was supported by grants from the National Institute of General Medical Sciences (GM-14312) and the National Institute on Aging (Ag-00322) of the National Institutes of Health, U.S. Public Health Service, from the National Science Foundation (PCM79-20279), from the Mobile Foundation, and from the National Foundation for Cancer Research.

Registry No. CH_3CO-(L-Ala)^{10}-$NHCH_3$, 85251-49-6; CH_3CO-(L-Ala)$_{10}$-$NHCH_3$·CH_3CO-(L-Ala)^{10}-$NHCH_3$, 89399-16-6.

(41) Blundell, T.; Barlow, D.; Borkakoti, N.; Thornton, J. *Nature (London)* **1983**, *306*, 281-283.

Bond-Deficient Molecules

Nicolaos D. Epiotis

Contribution from the Department of Chemistry, University of Washington, Seattle, Washington 98195. *Received May 20, 1983*

Abstract: Bond-diagrammatic molecular orbital–valence bond theory suggests that molecules comprised of an axial C_2 fragment and a set of equatorial ligands are bond deficient; i.e., they have one bond less than expected by drawing a Lewis formula satisfying the octet rule. Molecules of this type can be isolable when no significantly better geometric alternative exists.

Valence bond (VB) theory[1] provides a very simple picture of the electronic structure of H_2O: Water is bent because, in this geometry, *ground-state* triplet O ($s^2p^2p^1p^1$) can be coupled to two ground-state doublet H's into an overall singlet.[2] This is no longer possible in the linear geometry. In this case, an *excited* triplet O ($s^1p^2p^2p^1$) combines with two ground-state doublet H's to produce an overall singlet species. In recent times, chemists became fascinated with molecular orbital (MO) theory and tried to reinterpret phenomena for which a VB explanation existed, as well as other trends discovered by recent experimentation, only to find themselves enmeshed in a web of confusion from which only few perceptive theoreticians could escape. Thus, for example, controversies arose as to what are "lone pairs" and "bond pairs" in H_2O, why H_2O is bent, why H_2S is more bent than H_2O, etc. Thes ambiguities have now been largely resolved by Hall[3] and Shustorovich,[4] who provided a largely correct interpretation of

(1) (a) Heitler, W.; London, F. *Z. Phys.* **1927**, *44*, 455. (b) Slater, J. C. "Quantum Theory of Molecules and Solids"; McGraw-Hill: New York, 1963; Vol. 1.

(2) Murrell J. A.; Kettle, S. F. A.; Tedder, J. M. "Valence Theory"; 2nd ed.; Wiley: New York, 1970.

(3) (a) Hall, M. B. *J. Am. Chem. Soc.* **1978**, *100*, 633. (b) Hall, M. B. *Inorg. Chem.* **1978**, *17*, 2261.

(4) Shustorovich, E.; Dobosh, P. A. *J. Am. Chem. Soc.* **1979**, *101*, 4090.

0002-7863/84/1506-3170$01.50/0