Low-Energy Structures of Two Dipeptides and Their Relationship to Bend Conformations[1]

K. Nishikawa,[2a] F. A. Momany,[2b] and H. A. Scheraga*

Department of Chemistry, Cornell University, Ithaca, New York 14850. *Received May 21, 1974*

ABSTRACT: The low-energy conformations of two dipeptides are examined using "empirical" energy calculations and the virtual-bond method. This method enables the complete conformational space, within which bend structures occur, to be represented in tabular form. The complete conformational spaces of the dipeptides, *N*-acetyl-*N'*-methylglycyl-glycineamide and *N*-acetyl-*N'*-methyl-L-alanyl-L-alanineamide, were searched systematically, and all conformations of minimum energy were found. The interaction energy between the first and second amino acid units is generally small compared to the total energy of the dipeptide at a local minimum-energy conformation. Thus, the dipeptide energy can be represented, to a first approximation, as a sum of the energies of its constituent amino acid units. Most combinations of single-residue energy minima correspond to local minima on the energy surface of the dipeptide, and the *global* minimum of both Gly-Gly and L-Ala-L-Ala is simply the combination of the global minima for each single residue. However, for minimum-energy bend conformations, there is a significant departure from additivity of single-residue energies. The low-energy conformation of the alanine dipeptide, which closely approximates a type II bend (standard values of $\phi_i = -60°$; $\psi_i = 120°$; $\phi_{i+1} = 80°$; $\psi_{i+1} = 0°$), does not correspond to a combination of single-residue minima. Conformations similar to a type I bend (standard values $\phi_i = -60°$; $\psi_i = -30°$; $\phi_{i+1} = -90°$; $\psi_{i+1} = 0°$) are not found to be of minimum energy for either the glycine or alanine dipeptides.

The question of how a polypeptide chain folds in order to become the observed native conformation of a protein can be answered, to some degree, by the assumption that short-range interactions are dominant, *i.e.*, that each amino acid behaves, to some extent, independently of its surrounding residues and is not strongly affected by longer range interactions.[3] To test this approximation, it is necessary to investigate the influence of next nearest neighbor interactions on the conformations of small segments of polypeptides. Since the conformational energy surfaces of the *N*-acetyl-*N'*-methylamides of the 20 naturally occurring amino acids ("single residues"[4]) have been reported previously,[5] we can use these results and extend them to "dipeptides,"[4] looking in particular for how the minimum energies of a dipeptide deviate from the sums of the energies of the two single-residue units from which it is composed and observing whether deviations in conformation result from these deviations in energy.

If the single-residue units at a low-energy minimum of the dipeptide exist in conformations not observed for the single residues, the stability of this dipeptide conformation must arise from the interaction between the single-residue units.[6,7] Bend or "chain reversal" conformations[8-13] constitute examples where stabilization of the dipeptide conformation may arise from the interaction between the single-residue units. It should be possible to decide whether bend conformations in proteins arise from short-range interactions, *i.e.*, within the single-residue units, or require longer range interactions (between the two single-residue units, or possibly beyond) for stability. The possible contribution that solvent may make toward stabilizing bend structures will not be included here. However, we recognize that solvent interactions may play a significant role in conformational preference; to minimize this effect we have selected dipeptides with nonpolar side chains for this study.

In this context, conformational energy calculations on *N*-acetyl-*N'*-methylglycyl-glycineamide and *N*-acetyl-*N'*-methyl-L-alanyl-L-alanineamide have been carried out to identify the interaction energy necessary to form bends as dipeptides, compared to the behavior of glycine and alanine as two isolated single-residue structures. The results of these calculations are reported here.

Since the conformational space of a dipeptide is at least four dimensional, existing methods for describing its energy surface are not adequate. A convenient description of the conformational space of a dipeptde, using the virtual-bond method, has been formulated and is presented herein.

I. Representation of Dipeptide Conformations

Ignoring the rotation of the end methyl groups, only two variables, the dihedral angles[14] ϕ and ψ, are required to specify the conformations of the backbone of a single residue with planar trans peptide groups, and four (*viz.*, ϕ_1, ψ_1, ϕ_2, and ψ_2) are required for a dipeptide. In order to deduce a basis for a two-dimensional representation of the conformational energy surface of a dipeptide, we will introduce a set of new variables, which depend on ϕ_1, ψ_1, ϕ_2, and ψ_2.

It should be noted that no allowance is made for variation of ω in the virtual bond approach developed here. For the small dipeptides studied here, there are no forces which could make the nonplanar amide the preferred conformation, as shown in previous papers.[5,13] Of course, this condition might not hold for larger polypeptides for which long-range forces could cause some strain on the backbone, which could result in nonplanar peptide bonds.

A. The Virtual-Bond Method (General). If the peptide group is planar, the $C^{\alpha}_i \cdots C^{\alpha}_{i+1}$ distance is independent[15] of ϕ and ψ; this invariance of the $C^{\alpha}_i \cdots C^{\alpha}_{i+1}$ distance allows the backbone conformation of a dipeptide to be specified by three variables, *viz.*, the angles between the virtual bonds $C^{\alpha}_0 \cdots C^{\alpha}_1$ and $C^{\alpha}_1 \cdots C^{\alpha}_2$, and $C^{\alpha}_1 \cdots C^{\alpha}_2$ and $C^{\alpha}_2 \cdots C^{\alpha}_3$, designated as θ_1 and θ_2, respectively, and the dihedral angle around the $C^{\alpha}_1 \cdots C^{\alpha}_2$ bond, designated as γ (see Figure 1). The angle γ may take on any value from -180 to $+180°$ (with $\gamma = 0°$ being chosen when C^{α}_0 is cis to C^{α}_3, and the clockwise rotation of $C^{\alpha}_2 \cdots C^{\alpha}_3$ being positive when looking from C^{α}_1 to C^{α}_2). The angle θ, on the other hand, is restricted to some range of values

$$\theta_{\min} \leq \theta \leq \theta_{\max} \tag{1}$$

because of the geometry of the peptide chain. The values of $\theta_{\min}$ and $\theta_{\max}$, which depend on the geometry of the peptide chain, are discussed below (see eq 16 and 17). The reason why four variables (ϕ_1, ψ_1, ϕ_2, ψ_2) may be reduced to three (γ, θ_1, θ_2) is that the description of the conformation in terms of γ, θ_1, and θ_2 concerns only successive C^{α} atoms and contains no information concerning the other atoms of the chain (*e.g.*, the rotational positions of the planes of the peptide groups or the orientation of the side chains with respect to the backbone). Thus, a point on the γ, θ_1, and θ_2

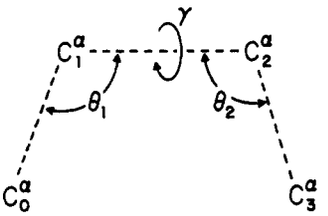

Figure 1. Schematic representation of a dipeptide, showing the definitions of the angles θ_1 and θ_2 between the virtual bonds and the dihedral angle γ.

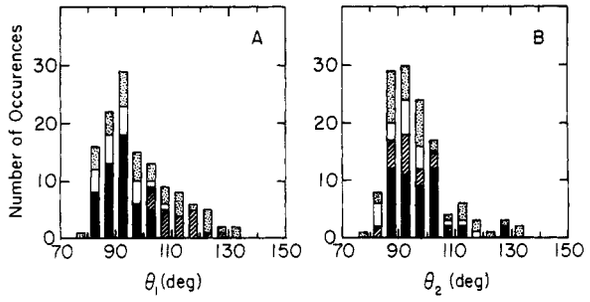

Figure 2. Distribution of values of θ_1 and θ_2 in 129 bends in the eight proteins considered by Lewis, *et al.*[13] The symbols (▬, ▨, ▭, ⁙) refer to bends of type[13] I, II, III, and other types, respectively.

conformational surface may be degenerate with respect to ϕ_1, ψ_1, ϕ_2, and ψ_2. On the other hand, it is easier to visualize the backbone conformation in terms of γ, θ_1, and θ_2 rather than in terms of ϕ_1, ψ_1, ϕ_2, and ψ_2. This is especially true for bend conformations.

B. Virtual-Bond Description of Bends. The virtual-bond method provides a convenient basis for considering bend conformations, for which the values of θ_1 and θ_2 can be expected to lie in a somewhat smaller range than that indicated by eq 1, *viz.*

$$\theta_{\min} \leq \theta_i \leq \theta_{\text{bend}} \tag{2}$$

with $i = 1, 2$ and $\theta_{\text{bend}} < \theta_{\max}$; *i.e.*, the condition $\theta_i \leq \theta_{\text{bend}}$ brings C^α_0 closer to C^α_3 than for values of θ_i between θ_{bend} and $\theta_{\max}$, as would be expected for bend conformations. Similarly, γ would be expected to lie in a restricted range (within γ_{bend}) around $\gamma = 0°$, *viz.*

$$|\gamma| < \gamma_{\text{bend}} \tag{3}$$

Using the experimental data on bends observed in proteins to set the limits, θ_{bend} and γ_{bend}, we can then regard eq 2 and 3 as general conditions for bend formation.

The observed values of θ_1, θ_2, and γ for the 129 observed bends discussed by Lewis, et al.,[13] were calculated from the X-ray coordinates of eight proteins and are shown in Figures 2 and 3 for the various types of bends considered.[13] It appears that each type of bend may be characterized by a particular range of values of θ_1, θ_2, and γ, *viz.*

$$\text{type I} \quad \sim 80° < \theta_1 < \sim 105°; \sim 85° < \theta_2 < \sim 105°; \sim 30° < \gamma < \sim 70° \tag{4}$$

$$\text{type II} \quad \sim 100° < \theta_1 < \sim 125°; \sim 80° < \theta_2 < \sim 105°; \sim -20° < \gamma < \sim 40° \tag{5}$$

$$\text{type III} \quad \sim 80° < \theta_1 < \sim 105°; \sim 80° < \theta_2 < \sim 105°; \sim 40° < \gamma < \sim 90° \tag{6}$$

For type I′, II′, and III′ bends (even though only a few examples of these appear in the 129 bends considered[13]), the ranges of θ and γ are the same as those of type I, II, and III, respectively, except that the sign of γ is reversed.

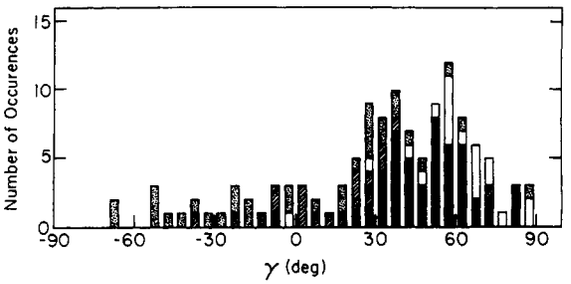

Figure 3. Same as Figure 2, but for γ.

Type I and III bends are observed to have similar values of θ_1 and θ_2, generally <105°. However, type II bends are observed to have differing values of θ_1 and θ_2, with θ_1 tending to be larger, but <125°. The values of γ for type II bends distribute closely around $\gamma = 0°$, but the ranges of γ observed for type I and III bends are positive and large, which means that the peptide chain twists in the right-handed direction (in the sense that the four successive C^α atoms constitute a right-handed screw). The observed values agree with the standard values of θ_1, θ_2, and γ given in Table I. It should be pointed out that the definition of bend types I and II by Venkatachalam[8] includes the criterion that four successive C^α atoms are coplanar (*i.e.*, $\gamma \sim 0°$), with the difference being that the middle peptide group is flipped over. However, in Table I we have used the "standard" values of the four dihedral angles as defined by Venkatachalam,[8] and we find that γ for a type I bend deviates considerably (*i.e.*, $\gamma = 48°$) from zero. The value of θ_1 for a type II bend is also large compared to the other θ values. The reason for this lies in the choice of dihedral angles which meet the condition that an N—H···O=C (4 → 1) hydrogen bond be formed.[8]

In general, for all of the bends in Figures 2 and 3, the values of θ are related to the value of γ. When γ is close to zero, θ becomes large, as illustrated for a type II bend (in eq 5), and *vice versa*. In addition, when θ_1 or θ_2 is greater than 105°, the other is always smaller than 105° (only one exception to this observation was noted in all of the bends examined).

From the observed ranges of Figures 2 and 3, and from the discussion above, the limiting expressions of eq 2 and 3 can now be rewritten as follows for all types of bends.

$$\theta_1 \text{ and } \theta_2 < 105° \tag{7}$$

$$\text{or} \quad \theta_i < 105° \text{ and } \theta_j < 125° \; (i, j, = 1 \text{ or } 2) \tag{8}$$

$$|\gamma| < 90° \tag{9}$$

Equations 7–9 will be used in section ID to formulate the conditions for occurrence of a bend.

C. General Relation between (γ,θ) and (ϕ,ψ). For planar trans peptide groups, an angle θ_i depends only on ϕ_i and ψ_i of the same residue, but the angle γ depends on all four dihedral angles, ϕ_1, ψ_1, ϕ_2, and ψ_2, of the dipeptide. In order to obtain the dependence of γ on these four dihedral angles, we define the auxiliary variable λ as that dihedral angle which defines the orientation of the planar peptide group (see Figure 4). λ_1 and λ_2 are the angles for rotation of the peptide planes about the virtual bonds $C^\alpha_0 \cdots C^\alpha_1$ and $C^\alpha_1 \cdots C^\alpha_2$, respectively, with respect to the plane containing the three C^α atoms of the single-residue unit (C^α_0, C^α_1, and C^α_2). The positive rotation is the clockwise one for either λ_1 or λ_2 (with $\lambda_1 = \lambda_2 = 0°$ when $\phi = \psi = 0°$) when looking from the central C^α atom (*i.e.*, C^α_1) to C^α_0 (for λ_1)

Table I
Characteristics of Various Standard Types of Bends

Type of bend[a,b]	Dihedral angles, deg						
	ϕ_1	ψ_1	ϕ_2	ψ_2	θ_1[c]	θ_2[c]	γ[c]
I	−60	−30	−90	0	87	89	48
II	−60	120	80	0	107	87	1
III	−60	−30	−60	−30	87	87	68

[a] For Type I′, II′, and III′ bends the values of ϕ_1, ψ_1, ϕ_2, ψ_2, and γ are the inverse of, and those of θ_1 and θ_2 are the same as, those for types I, II, and III, respectively. [b] The values of ϕ_1, ψ_1, ϕ_2, and ψ_2 are taken from ref 13 (and/or ref 8). [c] The values of θ_1, θ_2, and γ depend not only on ϕ_1, ψ_1, ϕ_2, and ψ_2 but also on the bond lengths and bond angles (the bond angle, $\tau(NC^{\alpha}C')$, differs[16] for glycine and alanine). Thus, the values listed in the table represent averages for the computed θ's and γ for glycine and alanine, using the dihedral angles listed.

and to $C^{\alpha}{}_2$ (for λ_2); therefore, λ_1 and λ_2 will move in opposite rotational directions when moving along the backbone chain from the ith to the $(i + 1)$th residues. Thus at $\phi = \psi = 0°$, the two peptide planes are in the same plane as defined by the $C^{\alpha}{}_0$, $C^{\alpha}{}_1$, and $C^{\alpha}{}_2$ atoms; *i.e.*, all C^{α}, C', and N atoms are in the same plane as shown in Figure 4. However, when ϕ and/or ψ (and hence λ_1 and λ_2) depart from 0°, the N and C' atoms of both peptide groups depart from the $C^{\alpha}{}_0$–$C^{\alpha}{}_1$–$C^{\alpha}{}_2$ plane, and the angle θ changes; the values of λ_1 and λ_2 are then given by the angles between the respective peptide planes and the new $C^{\alpha}{}_0$–$C^{\alpha}{}_1$–$C^{\alpha}{}_2$ plane.

A backbone conformation of any single-residue unit is uniquely determined in terms of θ, λ_1, and λ_2. However, it should be noted that these are not independent variables. Rather, when θ and λ_1 are held constant at certain values, an arbitrary change in λ_2 will give rise to a change in the angle α (see Figure 4); since α is fixed (for a system with rigid geometry), an arbitrary variation of λ_2 (with θ and λ_1 held fixed) is therefore impossible. This result follows naturally from the fact that only two of the variables θ, λ_1, and λ_2 can be independent, since only two variables (such as ϕ and ψ) are sufficient to define the backbone conformation of a single residue with fixed geometry. In this study, we will use λ_1 and λ_2 as auxiliary parameters to combine the expressions for θ and γ with the expressions for ϕ and ψ. The value of each angle (θ, λ_1, and λ_2) can be expressed as a function of ϕ and ψ, as will be shown below.

The angle γ (Figure 1), λ_2 of the first single-residue unit $[(\lambda_2)_{1st}]$, and λ_1 of the second single-residue unit $[(\lambda_1)_{2nd}]$ are all defined about the same virtual bond, $C^{\alpha}{}_1 \cdots C^{\alpha}{}_2$. Hence, they are related as follows

$$\gamma = (\lambda_2)_{1st} + (\lambda_1)_{2nd} + 180° \qquad (10)$$

where the constant 180° arises from the definition of the zero positions of $(\lambda_2)_{1st}$ and $(\lambda_1)_{2nd}$. Since θ, λ_1, and λ_2 are all defined within one single-residue unit, as shown in Figure 4, then changes of conformation may be carried out with transformation matrices expressed in terms of (θ, λ_1, λ_2) or, equivalently, (ϕ, ψ), *i.e.*

$$\mathbf{T}_{\lambda_1}{}^{x}\mathbf{T}_{(\pi-\theta)}{}^{z}\mathbf{T}_{\lambda_2}{}^{x} = \mathbf{T}_{\xi}{}^{z}\mathbf{T}_{\phi}{}^{x}\mathbf{T}_{(\pi-\alpha)}{}^{z}\mathbf{T}_{\psi}{}^{x}\mathbf{T}_{\eta}{}^{z} \qquad (11)$$

where ξ, η, and α are the constant angles shown in Figure 4, and the matrices $\mathbf{T}_{\beta}{}^{x}$ and $\mathbf{T}_{\beta}{}^{z}$ (β being an arbitrary angle) are general matrices for rotation about the x and z axis, respectively. The coordinate system used here is the same as Figure 3 of ref 15, so that the right-hand side of eq 11 is the same as the right-hand side of eq 2 of ref 15, except for a change of definition of ϕ and ψ to the new IUPAC-IUB convention.[14] Multiplying both sides of eq 11 by the unit row vector, $\mathbf{v} = (1,0,0)$, and the unit column vector, $\mathbf{u}$, whose transpose is identical with $\mathbf{v}$, we obtain

$$\mathbf{v}\mathbf{T}_{\lambda_1}{}^{x}\mathbf{T}_{(\pi-\theta)}{}^{z}\mathbf{T}_{\lambda_2}{}^{x}\mathbf{u} = \mathbf{v}\mathbf{T}_{\xi}{}^{z}\mathbf{T}_{\phi}{}^{x}\mathbf{T}_{(\pi-\alpha)}{}^{z}\mathbf{T}_{\psi}{}^{x}\mathbf{T}_{\eta}{}^{z}\mathbf{u} \qquad (12)$$

Here $\mathbf{u}$ and $\mathbf{v}$ are invariant with respect to a rotation about the x axis, *i.e.*, $\mathbf{v}\mathbf{T}_{\lambda_1}{}^{x} = \mathbf{v}$ and $\mathbf{T}_{\lambda_2}{}^{x}\mathbf{u} = \mathbf{u}$. Thus, eq 12 may be simplified to give θ as a function of ϕ and ψ, *viz.*

$$\cos\theta = a\cos\xi + b\sin\xi\cos\phi - \sin\xi\sin\eta\sin\phi\sin\psi \qquad (13)$$

where

$$a = \cos\eta\cos\alpha + \sin\eta\sin\alpha\cos\psi \qquad (14)$$

and

$$b = \cos\eta\sin\alpha - \sin\eta\cos\alpha\cos\psi \qquad (15)$$

In eq 13, only positive values of θ are considered. For any value of ϕ and ψ, θ may be obtained uniquely from eq 13, within the range $0° < \theta < 180°$. The maximum and minimum values of θ are obtained by substituting $(\phi,\psi) = (180°,180°)$ and $(0°,0°)$, respectively, in eq 13, *i.e.*

$$\theta_{max} = \alpha + (\xi + \eta) \qquad (16)$$

$$\theta_{min} = \alpha - (\xi + \eta) \qquad (17)$$

Equations 16 and 17 also follow obviously from Figure 4.

In order to obtain the relation between λ_1 and ϕ and ψ, eq 11 is rewritten as follows

$$\mathbf{T}_{(\pi-\theta)}{}^{z}\mathbf{T}_{\lambda_2}{}^{x} = \mathbf{T}_{-\lambda_1}{}^{x}\mathbf{T}_{\xi}{}^{z}\mathbf{T}_{\phi}{}^{x}\mathbf{T}_{(\pi-\alpha)}{}^{z}\mathbf{T}_{\psi}{}^{x}\mathbf{T}_{\eta}{}^{z} \qquad (18)$$

where $\mathbf{T}_{-\lambda_1}{}^{x}$, a matrix for rotation about the x axis by an angle of $-\lambda_1$, is the same as $(\mathbf{T}_{\lambda_1}{}^{x})^{-1}$, the inverse of $\mathbf{T}_{\lambda_1}{}^{x}$. Then both sides of eq 18 are multiplied by the row vector $\mathbf{v}' = (0,0,1)$ and the column vector $\mathbf{u}$ defined above to make the left-hand side of eq 18 vanish, because $\mathbf{T}_{\lambda_2}{}^{x}\mathbf{u} = \mathbf{u}$, $\mathbf{v}'\mathbf{T}_{(\pi-\theta)}{}^{z} = \mathbf{v}'$ (*i.e.*, $\mathbf{v}'$ is invariant with respect to a rotation about the z axis), and $\mathbf{v}'\mathbf{u} = 0$. The result is

$$\tan\lambda_1 = (-b\sin\phi - \sin\eta\cos\phi\sin\psi)/(a\sin\xi - b\cos\xi\cos\phi + \cos\xi\sin\eta\sin\phi\sin\psi) \qquad (19)$$

where a and b are defined in eq 14 and 15, respectively. Equation 19 has two solutions for λ_1, but one of them can be eliminated because it gives a negative value of θ when applied to the equation relating ϕ, λ_1, and θ (which is not shown here but which is easily obtained by eliminating λ_2 and ψ from eq 11 in a similar way to that shown above).

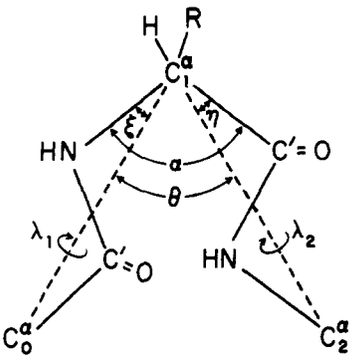

Figure 4. Schematic representation of a single residue (with $\phi = \psi \simeq 0°$), showing the definitions of the constant bond angles α, ξ, and η, the virtual bond angle θ, and the auxiliary variables λ_1 and λ_2.

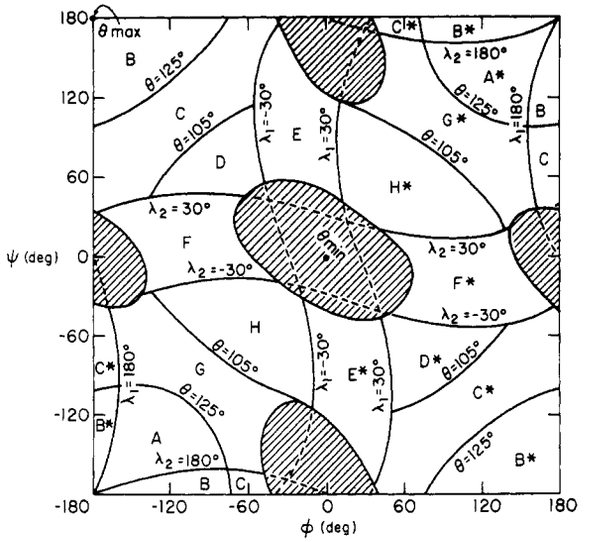

Figure 5. The (ϕ,ψ) space of a single residue, divided into 16 areas A to H and A* to H*. The boundaries between the various areas are lines of constant θ, λ_1, or λ_2. The shaded regions are those of very high energy.[5] The λ_1 scale almost coincides with the ϕ scale, and the λ_2 scale almost coincides with the ψ scale; $\lambda_1 = \phi$ and $\lambda_2 = \psi$ at the nine points ($\phi = 0°, \pm180°$, and $\psi = 0°, \pm180°$).

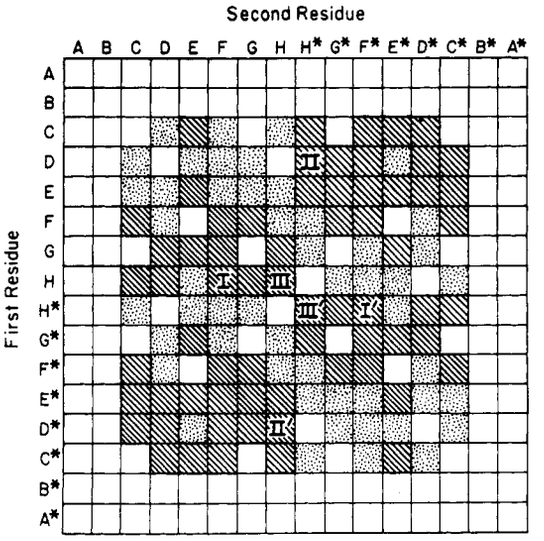

Figure 6. Area table, showing the various regions in which bend conformations can occur, based on the condition that θ_1 and θ_2 satisfy eq 7 and 8. The Roman numerals describe particular types[8,13] of bends. ▨ is an area in which there is at least one backbone conformation with $\gamma = 0°$; ⁘ is an area in which there is at least one backbone conformation with $|\gamma| < 90°$.

In the same manner, the following relation between λ_2 and ϕ and ψ is also obtained

$$\tan \lambda_2 = (-e \sin \psi - \sin \xi \cos \psi \sin \phi)/(d \sin \eta - e \cos \eta \cos \psi + \cos \eta \sin \xi \sin \phi \sin \psi) \quad (20)$$

where

$$d = \cos \xi \cos \alpha + \sin \xi \sin \alpha \cos \phi \quad (21)$$

and

$$e = \cos \xi \sin \alpha - \sin \xi \cos \alpha \cos \phi \quad (22)$$

As a result, θ, λ_1, and λ_2 can be computed for a given value of (ϕ,ψ), and then γ is determined by applying eq 10 to successive single-residue units.

D. Representation of the General Conformation of a Dipeptide by an Area Table. A sequence of two single-residue units makes a dipeptide whose conformation may be expressed by the virtual-bond method, as discussed in section IA. Therefore, for the study of dipeptides, it is convenient to divide the (ϕ,ψ) space of a single residue into different areas based on the values of θ, λ_1, and λ_2, *i.e.*, by drawing lines of constant values of θ, λ_1, and λ_2 using eq 13, 19, and 20, respectively. The choice of values of θ, λ_1, and λ_2, for this purpose, is arbitrary, but the choice shown in Figure 5 (*viz.*, θ = 105 and 125°, and λ_1 and $\lambda_2 = \pm30$ and 180°) divides the single-residue (ϕ,ψ) space into 16 areas (of roughly equal size) A to H and A* to H* (the latter being the inverse of the former about the point $\phi = \psi = 0°$). The shaded regions in Figure 5, which correspond to single-residue conformations of very high energy for glycine,[5] were omitted from consideration; this includes the area for θ_{min} (which is ~73°, based solely on geometrical considerations according to eq 17, and corresponds to $\phi = \psi = 0°$). Thus, the actual minimum value of θ (*viz.*, $\theta \sim 80°$) would arise at the margin of this high-energy region; this value of ~80° is indicated clearly as the lower limit of the distributions of θ_1 and θ_2 in Figure 2.

Most regions of the area map of Figure 5 correspond to known conformations of a single residue; *e.g.*, areas B and C correspond to the extended or β structures, and areas D, F, H, and H* correspond to the C_7^{eq}, bridge region, and α_R and α_L conformations, respectively. The areas may be placed in a linear sequence in the order A to H and H* to A* by moving continuously from one area to the next on the ϕ,ψ map. However, there are some exceptions to this contiguity criterion, *viz.*, E to F, H to H*, and F* to E* (see Figure 5). With this sequence, the whole conformational space of a dipeptide may then be represented in an "area table" as a combination of the conformations of the first and second single-residue units, in which the areas (A to A*) of the first single-residue unit are placed along the columns and the areas of the second single-residue unit are placed along the rows of the table (see Figure 6).

The cross-hatched and dotted areas in this tabulation are those satisfying the general bend condition of section IB. First of all, the θ values of areas A, B, B*, and A* are always greater than 125° (Figure 5), so that the outermost two rows and columns of the table are excluded from being bends. Some of the remaining areas are also excluded because they do not satisfy the inequalities of eq 7 and 8. However, it is more complicated to apply the condition of eq 9 for γ. If we were to compute γ for each dipeptide area of Figure 6 by means of eq 10 [with the values of λ_1 and λ_2 of the constituent single-residue areas (interpolated between the limits given in Figure 5)], in most cases we would obtain a very wide range of values of γ (both $|\gamma| < 90°$ and >90°); in other words, *each* dipeptide area of Figure 6 pertains to various kinds of conformations with different values of γ. We may express the condition on γ (for bend formation) in two ways and, together with eq 7 and 8, identify the dipeptide areas in which bend conformations arise, *viz.*, (1) those areas which involve *at least* one conformation with $\gamma = 0°$ (cross-hatched areas in Figure 6), and (2) those areas which involve at least one conformation with a value of γ within the range $|\gamma| < 90°$ (dotted areas in Figure 6). The areas for bends of types I, II, and III are HF, DH*, and HH, respectively; the symmetrically arranged areas for bends of type I′, II′, and III′ are also shown. The cross-hatched plus dotted areas do not correspond exactly to the conditions for general bend formation of eq 7–9 (in particular, the dotted areas pertain to many nonbend conformations, as described above). It should be noted that, since no interatomic interactions are included in this geometrical representation, even though some single-residue energy restrictions are built in (see Figure 5), all possible dipeptide bends appear on the area map. This gives the im-

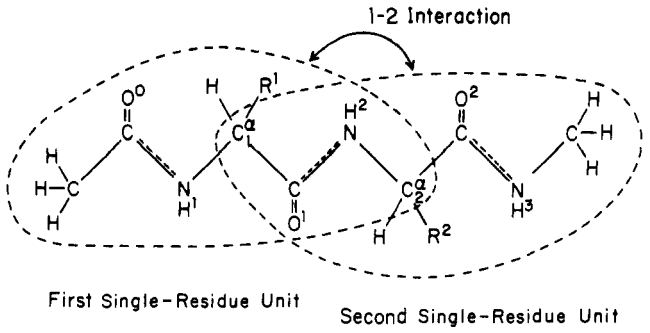

Figure 7. Illustration of partitioning of the energy of a dipeptide between the energy of the two separate single-residue units, E_{res}, and the inter-residue energy, E_{12}; the central peptide group contributes the constant energy, E_{pep}, for each dipeptide.

pression that a very large volume of conformational space corresponds to some type of bend structure. Although we cannot determine the precise proportion of bend conformations from Figure 6 without including energetic considerations, which would depend upon the amino acid sequence of the first and second residues, the result shown in Figure 6 may correspond to the observations that the proportion of residues in bends (counting four residues in each bend) is 31% of the total number of residues in eight proteins[13] and 33% in seven proteins,[11] when the bend is defined quite generally (*i.e.*, ignoring the 4 → 1 hydrogen bond criterion[8]) in terms of the $C^{\alpha}_{i} \cdots C^{\alpha}_{i+3}$ distance[13] or similar criteria.[10,11] The blank areas of Figure 6 cannot pertain to general bend conformations (in the sense of eq 7-9); *e.g.*, areas DD, DH, HH*, and H*H, etc., which are surrounded by cross-hatched or dotted areas, can never pertain to bends since their values of γ are always in the range $|\gamma| > 90°$.

In section II, the procedure for calculating the conformational energies of dipeptides will be presented. The area table of Figure 6 will be useful in summarizing the results of these energy calculations.

II. Conformational Energies of the Dipeptides Gly-Gly and L-Ala-L-Ala

A. Components of the Dipeptide Energy. The total energy of a dipeptide was calculated as the sum of contributions of the nonbonded, electrostatic, and hydrogen-bonding energies.[16] One particular difference from the earlier calculations[16] is that a hydrogen-bonding function[17] is used in place of a nonbonded interaction between the H atom of the backbone peptide group and the N atom of another backbone peptide group; this interaction provides some additional stability in the region near $\psi \sim 0°$ of the single-residue map. The bond lengths and bond angles of Gly and L-Ala were those of Momany, *et al.*[16]

As discussed in section I, a dipeptide is considered to be made up of two single residue units. Hence, it is convenient to partition the total energy of a dipeptide (E_{dipep}) between the energy of two separate single-residue units (E_{res}) and the interaction energy between the first and second single-residue units (E_{12}), as shown in Figure 7.

$$E_{dipep} = E_{res} + E_{12} \quad (23)$$

where

$$E_{res} = E_1 + E_2 - E_{pep} \quad (24)$$

and E_1 and E_2 are the total energies of the first and second single-residue units, respectively, which differ because of the different end groups on each single residue (*i.e.*, the left-hand C^{α} carries three hydrogens in the first single residue, whereas the right-hand C^{α} carries the three hydrogens

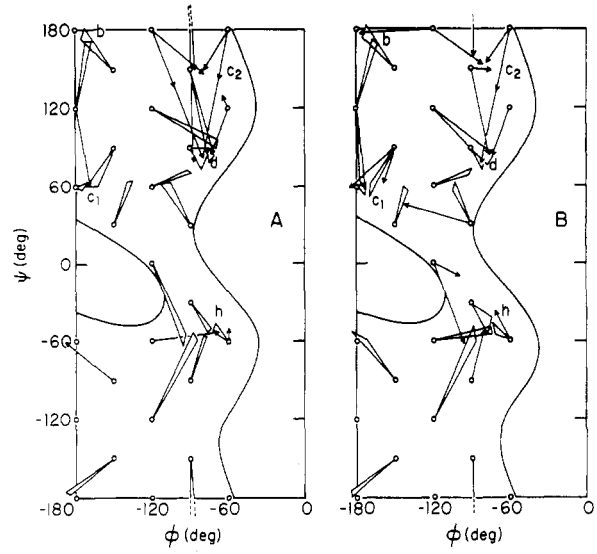

Figure 8. The (ϕ,ψ) space of the first (A) and second (B) single-residue units of the dipeptide Gly-Gly. The energy contour shown is 5 kcal/mol above the global minimum (at point d). The circles represent starting conformations for energy minimization. The lower-case letters indicate single-residue minima corresponding to the capital letters designating the area in Figure 5. The arrows and wedges indicate the direction of movement after one cycle of energy minimization. Where subscripts 1 and 2 appear, 1 refers to a minimum, but 2 refers to a low-energy region that is not a minimum. In the case of glycine, the other half of each diagram, A and B, is obtained by a 180° rotation of each diagram around the point $\phi = \psi = 0°$.

in the second single residue); E_{pep}, the energy of the central peptide group (Figure 7), is subtracted because it is calculated twice in E_1 and E_2. Since the peptide group is fixed in the planar trans conformation ($\omega = 180°$) in this study, E_{pep} is constant (for a given dipeptide), so that E_{res} of eq 24 is essentially the sum of the single-residue energies (aside from this constant value), and includes the minor differences arising from the different end groups as noted above. As a result, E_{12} of eq 23 is a crucial term for expressing the dependence of E_{dipep} on conformation. In other words, if any local minima in the energy space of a dipeptide differ from combinations of the local minima of its constituent single residues, these additional minima can arise only from the E_{12} term. The E_{12} term also involves the so-called 4 → 1 hydrogen bond[8] (between the NH^3 of residue 3 and the CO^0 of the zeroth residue; see Figure 7) that is characteristic of some bend conformations.

B. Location of Low-Energy Minima in the Dipeptide Space. Forty-one starting single-residue conformations were chosen for minimizing the energy of the Gly-Gly dipeptide. These covered the space of the Gly single-residue map within 5 kcal of the global minimum uniformly (see Figure 8). Similarly, 34 starting conformations were chosen for the L-Ala single residue (see Figure 9). Thus, 943 (*i.e.*, 41 × 23) starting conformations (*i.e.*, ϕ_1, ψ_1, ϕ_2, ψ_2) were taken for Gly-Gly and 1156 (*i.e.*, 34 × 34) for L-Ala-L-Ala (the former being less than 41 × 41 because the symmetrically equivalent conformations were eliminated). The side-chain dihedral angles of L-Ala-L-Ala were also treated as variables and were set initially at 60°. However, not all of these starting conformations were subjected to energy minimization (although, the initial energy of each one was computed); instead, this number was reduced by using the following criteria to decide which starting conformations of Gly-Gly (and, in parentheses, of L-Ala-L-Ala) should be subjected to energy minimization.

(i) The value of E_{res} for each starting conformation was compared with all other values of E_{res}. If $E_{res} \leq 2$ kcal (3

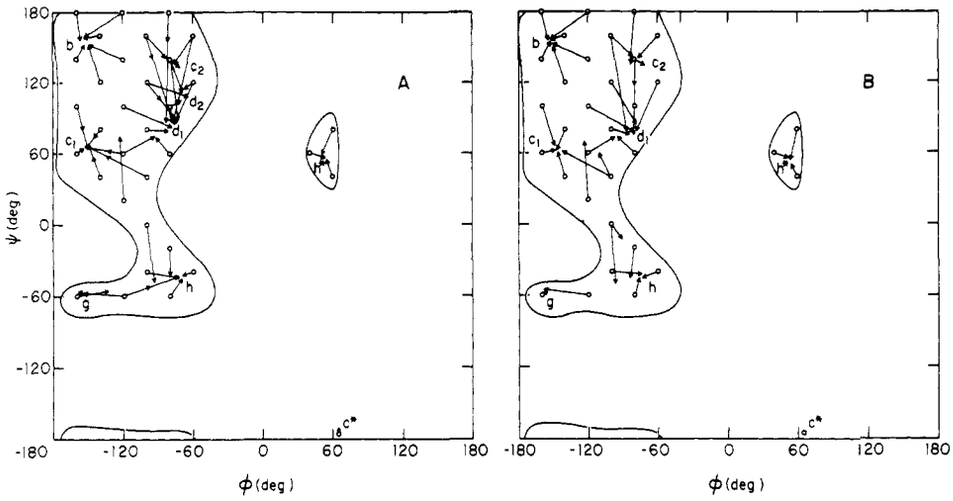

Figure 9. Same as Figure 8 but for L-Ala-L-Ala.

kcal) above the lowest value found, the conformation was minimized.

(ii) If (i) was not satisfied, but $E_{12} \leq -1.5$ kcal (−2 kcal) and in some cases $E_{12} > 0$, the conformation was minimized.

(iii) If (i) and (ii) were not satisfied, and all combinations of one particular initial point on the single-residue map had been eliminated, the combinations giving the lowest values of E_{res} and E_{12}, respectively, were selected for minimization. In this way each initial point on the single-residue map was examined by energy minimization.

By using these criteria, we see that (ii) will select conformations for minimization which have strong attractive E_{12} interactions and some repulsive ($E_{12} > 0$) interactions. The conformations with repulsive interactions are included since it might be that only very slight changes in conformation would reduce E_{12} to a negative stabilization energy. Using the above criteria, the original number of starting conformations was reduced to about one-third, and these were put through one cycle of energy minimization (using the minimization procedure proposed by Powell[18] and modified by Zangwill[19]). The purpose of this was to determine the initial directions of movement that each point followed on the energy surface of the single residue.

At this point, many different conformations appeared to be converging to a single local minimum; thus, in order to reduce the computational time, only the conformation of lowest energy (in the neighborhood of each local minimum) was selected for subsequent cycles of minimization. The final minimization was carried out in each case until the total energy decreased by less than 0.002 kcal/mol for each cycle of minimization.

III. Results

The results of the energy minimizations are presented below. The alanine dipeptide will be discussed first, since it is simpler than Gly-Gly, all conformations of which have symmetry-related ones.

A. L-Ala-L-Ala. Using the three rules described in section IIB, the energies of 420 starting conformations of L-Ala-L-Ala were minimized with respect to the six variables ϕ_1, ψ_1, χ_1, ϕ_2, ψ_2, and χ_2. The movements of the backbone dihedral angles (ϕ_i, ψ_i) of the first and second residues after one cycle of minimization are shown in Figures 9A and 9B, respectively. While a single arrow in Figure 9A represents the movement of ϕ_1,ψ_1, it may represent a family of conformations, each of which has different values of ϕ_2,ψ_2; similarly, a single arrow in Figure 9B represents the movement of ϕ_2,ψ_2 and may represent a family of conformations each of which has different values of ϕ_1,ψ_1. Although some cases arise in which two or more arrows emanate in different directions from a single starting point (depending on the different conformations of the other residue), the general feature of these figures is the appearance of a simple tendency for each starting point to move directly to the vicinity of one of the single-residue minima (even after only one cycle of minimization); each single-residue minimum is indicated by a lower case letter which corresponds to an area (with capital letters) of Figure 5. The subscripts on c and d distinguish between actual minima (*viz.*, c_1 and d_1) and local low-energy regions that are not minima (*viz.*, c_2 and d_2). These differences will be described further in this section.

As mentioned in section IIA, E_{12} is the only term which distinguishes a dipeptide from two single residues. If E_{12} is close to zero, each single residue unit of the dipeptide behaves not only independently of the other single-residue unit but also exactly the same as would be expected for the corresponding single residue. A good example of this is shown around the minimum b in Figures 9A and 9B. A given arrow in this region (in Figure 9A) indicates that ϕ_1,ψ_1 (as well as several different values of ϕ_2,ψ_2 in this region of Figure 9B) all move in the same direction. In other words, the movements of the points of one single-residue unit are almost independent of the movements of the points of the other single-residue unit because E_{12} is small in that region of conformational space where the chain is almost fully extended. On the other hand, several arrows diverge from a given point around, say, minima d_1 and h because of a large negative E_{12} term; *i.e.*, in these regions the energy deviates from the sum of the single-residue energies.

Further cycles of minimization were then carried out in the region of each local minimum, as described in section IIB, starting from the lowest energy conformations obtained after the first cycle of minimization; in addition, all the other intermediate conformations far from any of the single-residue minima were included at this stage of the minimization. The energies of the resulting minima of the alanine dipeptide are given in Table II as combinations of conformations of the first and second single-residue units. The minimum-energy conformation of each single residue is listed in Table III. The exact position of each minimum in the dipeptide deviated by only a few degrees in ϕ and ψ from the corresponding minimum of the single residue (those with deviations of more than 10° are indicated in

Table II
Energy Minima[a,b] of L-Ala-L-Ala

First residue	Second residue: b	c_1	c_2	d_1	g	h	h*	c*
b	0.83	1.24	0.97	0.48	1.76	1.45	2.48	*
	(−0.60)	(−0.82)	(−0.79)	(−0.90)	(−0.93)	(−1.11)	(−1.16)	
c_1	1.16	1.44	1.48	0.86	2.04	1.80	2.78[c]	*
	(−0.62)	(−0.97)	(−0.62)	(−0.86)	(−1.00)	(−1.11)	(−1.21)	
d_1	0.59	0.57[d]	0.58	0.00[e]	0.49[c]	0.71		*
	(−0.75)	(−1.52)	(−1.13)	(−1.33)	(−2.21)	(−1.80)		
d_2							1.27[c,d]	
							(−3.57)	
g	1.49	2.03	1.69	1.24[c]	2.67	2.41[c]	*	*
	(−0.90)	(−0.97)	(−1.03)	(−1.10)	(−0.98)	(−1.10)		
h	1.10	1.11[c,d]		0.42[c]	2.56[c]	1.21[c]	*	*
	(−1.38)	(−2.31)		(−2.14)	(−1.17)	(−2.53)		
h*	2.12	2.73		2.10	*	*	*	*
	(−1.56)	(−1.56)		(−1.52)				
c*	*	*		*	*	*	*	*

[a] The total energy (relative to zero for d_1d_1) in kcal/mol is given for each minimum of energy lower than 3 kcal/mol; higher energy minima are indicated by asterisks. The calculated values of E_{12} (unscaled) are given in parentheses. [b] A blank space means that the conformation is not a minimum-energy one. [c] These are general bend conformations, satisfying eq 7-9. [d] One of the values of ϕ_1, ψ_1, ϕ_2, and ψ_2 deviates more than 10° from the location listed in Table III, *i.e.*, (−84°, 77°, −153°, 62°) for d_1c_1; the ϕ_i and ψ_i values of d_2h^* and hc_1 are given in Table IV. [e] The total energy of the minimum, d_1d_1, was found to be 0.44 kcal/mol.

Table II). The energy values in Table II are on a scale in which the energy is taken as zero for the global minimum of the dipeptide, which is simply the combination of the two global minima d_1d_1 (corresponding to two C_7^{eq} conformations) of the single residue. Almost all of the local minima of the dipeptide are combinations of single-residue minima, and, conversely, most combinations of single-residue minima turn out to be minima in the dipeptide space, regardless of the value of the energy, an exception being the d_1h^* combination, which shifts to a new minimum d_2h^*. Dipeptide minima with energies greater than 3 kcal/mol above the global minimum are indicated by asterisks in Table II. The only dipeptide minima which correspond to nonminimum conformations of the single residue are c_2 ($\phi = -75°$, $\psi = 140°$) and d_2 ($\phi = -66°$, $\psi = 110°$). While the c_2 conformation of a single residue is not at a minimum, it is in a broad low-energy region, and (in combination with another residue) can lead to minima in the dipeptide space, *e.g.*, bc_2, c_1c_2, d_1c_2, and gc_2. On the other hand, the point d_2 is a special case, the ϕ,ψ value of which differs from that at minimum d_1 and lies in between those at d_1 and c_2 (see Table III). Thus it is reasonable that the point d_2 be regarded as a new conformation specific to the dipeptide and occurs only with h* in the d_2h^* conformation. This new minimum arises because it has the largest negative value of E_{12} of all the minima shown in Table II.

The minima which satisfy the condition for general bend formation (eq 7-9) are indicated by superscript *c* in Table II. Those with fairly large negative values of E_{12} (<−2 kcal/mol) belong to this set of bends, *viz.*, d_1g, d_2h^*, hc_1, hd_1, and hh. This is reasonable, since we would expect a bend conformation to be the only one in which two successive residues are close enough to interact strongly with each other. This is in contrast to nonbend conformations for which the energy of the dipeptide is almost the simple sum of the individual single-residue energies (*e.g.*, bc_1, bd_1, bg, etc., have almost the same total energies as their inverse combinations, c_1b, d_1b, gb, etc., respectively). On the other hand, there are significant differences in energy between the bend conformations described above and their inverse (nonbend) combinations, *e.g.*, hc_1 and c_1h, hd_1 and d_1h, and also d_2h^* and h^*d_1.

Table III
Location and Energy[a] of Minima for L-Ala Single Residue[5]

Minimum	ϕ, deg	ψ, deg	χ, deg	E, kcal/mol
b	−154	153	58	0.39
c_1	−151	74	56	0.72
c_2	−75	140	65	0.72[b]
d_1	−84	79	61	0.00[c]
d_2	−66	110	61	0.79[b]
g	−160	−60	42	1.37
h	−74	−45	62	1.14
h*	56	55	73	2.22
c*	65	−177	90	4.00

[a] These minima are for *N*-acetyl-*N'*-methyl-L-alanineamide. However, the values of ϕ, ψ, and χ, and E differ slightly from those of ref 5 because of the inclusion of the NH···N hydrogen bond described in section IIA. [b] Not a minimum. [c] $E_{d1} = -3.28$ kcal/mol.

In addition, the minima corresponding to bend conformations are included in the cross-hatched area of Figure 6 (except for d_1g which is in a dotted area). Although the correspondence is not exact, one may regard the cross-hatched areas of Figure 6 not only as ones of general bend conformation but also ones in which strong E_{12} interactions may arise.

Of these general bends with large negative values of E_{12}, d_2h^* (which is the only one specific to the dipeptide) has values of ϕ_1, ψ_1, ϕ_2, and ψ_2 equal to (−66°, 110°, 58°, and 44°). Although ϕ_2 and ψ_2 deviate from the standard value of a type II bend (80°,0°), the values of ϕ_1 and ψ_1 are close enough to the standard values (−60°,120° of Table I) for this conformation to be considered a type II bend.[13] The dipeptide is stabilized by the largest value of E_{12}, arising from the formation of a 4 → 1 type hydrogen bond. Conformation hd_1 may be regarded as a distorted type I bend since only a single dihedral angle, ψ_2, differs by more than 50° from the standard dihedral angles.[13] However, the value of E_{12} is much less negative than that of d_2h^*. No true minima of Table II correspond to type I bends, apparently because the values of E_{12} are not sufficiently nega-

Table IV
Dihedral Angles[a] and Energies[b] at Representative Minimum-Energy Conformations of L-Ala-L-Ala and Gly-Gly

		Φ_1	ψ_1	χ_1	Φ_2	ψ_2	χ_2	θ_1	θ_2	γ	E_{total}	E_{12}	E_{HB}[c]
L-Ala-L-Ala Conformation													
Global minimum	d_1d_1	−83	79	61	−84	79	61	97	97	179	0.00	−1.33	−0.21
Extended	bb	−154	154	58	−154	154	59	139	139	−178	0.83	−0.60	−0.29
Helical (or type III bend)	hh	−69	−42	61	−72	−42	62	92	92	40	1.21	−2.53	−0.27
Type I bend (distorted)	hd_1	−69	−46	61	−86	75	62	93	96	59	0.42	−2.14	−0.37
Type II bend	d_2h^*	−66	110	61	58	44	78	104	89	−17	1.27	−3.57	−1.06
Other bends with low	hc_1	−69	−41	62	−144	60	60	91	106	4	1.11	−2.31	−0.81
E_{12}(<−2 kcal/mol)	d_1g	−81	79	61	−161	−58	42	96	112	61	0.49	−2.21	−0.17
Gly-Gly[d] Conformation													
Global minimum	dd*	−82	76		83	−75		97	97	−64	0.00	−1.00	−0.51
Extended	bb	−180	180		−180	180		147	147	180	1.75	−0.15	−0.29
Helical (or type III bend)	hh	−66	−46		−70	−44		94	94	39	1.63	−1.91	−0.27
Type I bend (distorted)	hd	−67	−55		−85	73		96	97	53	0.58	−1.61	−0.40
Type II bend	dh*	−78	82		77	46		98	96	−27	0.24	−2.19	−0.79
Other bends with low	hc_1	−68	−43		−151	59		93	109	−5	1.86	−1.59	−0.71
E_{12} (<−1.5 kcal/mol)	dc_1*	−80	78		171	−53		97	111	31	0.36	−1.79	−0.15

[a] All angles are in degrees; θ_1, θ_2, and γ are defined in Figure 1. [b] E_{total} is relative to the global minimum; E_{12} and E_{HB} are the absolute values; all are in kcal/mol. [c] E_{HB} pertains to the 4 → 1 hydrogen bond energy for $CO^{\delta}\cdots H^{\delta}N$ (Figure 7): the sum of the general hydrogen bond term[17] between O and H, the nonbonded energies for the other three interactions, and the electrostatic energies of all the four atom pairs. [d] The inverse conformation, with the same energy, is possible for each of the Gly-Gly conformations (with the signs of ϕ_1, ψ_1, ϕ_2, ψ_2, and γ reversed).

Table V
Energy Minima[a,b] of Gly-Gly

First residue	Second residue				
	b (b*)	c_1	c_2	d	h
b	1.75	2.04		1.05	2.11
(b*)	(−0.15)	(−0.22)		(−0.23)	(−0.34)
c_1	1.84	2.11		1.24	2.31
	(−0.27)	(−0.36)		(−0.24)	(−0.34)
d	1.22	0.76[c,d]	1.35	0.21	0.96
	(−0.38)	(−1.30)	(−0.82)	(−0.78)	(−1.20)
h	1.81	1.86[c,d]		0.58[c]	1.63[c]
	(−0.87)	(−1.59)		(−1.61)	(−1.91)
h*		2.42		1.23	2.65
		(−0.60)		(−0.82)	(−0.56)
d*		0.36[c]	1.16[c]	0.00[c,e]	0.24[c]
		(−1.79)	(−1.02)	(−1.01)	(−2.19)
c_1*		2.20	2.04	1.07[c]	2.19[c]
		(−0.27)	(−0.61)	(−0.40)	(−0.47)

[a] The total energy (relative to zero for d*d) in kcal/mol is given for each energy minimum. The absolute values of E_{12} are given in parentheses. [b] A blank space means that the conformation is not a minimum-energy one or is excluded by symmetry considerations. [c] These are general bend conformations, satisfying eq 7-9. [d] One of the values of ϕ_1, ψ_1, ϕ_2, and ψ_2 deviates more than 10° from the location listed in Table VI, *i.e.* (−83°, 73°, 176°, 60°) for dc_1; see Table IV for the ϕ_i and ψ_i values of hc_1. [e] The global minimum energy (at d*d) was −2.61 kcal/mol.

Table VI
Location and Energy[a] of Minima for Gly Single Residue[5]

Minimum	ϕ, deg	ψ, deg	E, kcal/mol
b (b*)	−180	180	0.82
c_1	−173	62	1.03
c_1*	173	−62	1.03
c_2	−74	149	1.37[b]
c_2*	74	−149	1.37[b]
d	−83	76	0.00[c]
d*	83	−76	0.00[c]
h	−72	−53	1.19
h*	72	53	1.19

[a] These minima are for *N*-acetyl-*N'*-methylglycineamide. However, the values of ϕ, ψ, and E differ slightly from those of ref 5 because of the inclusion of the NH···N hydrogen bond described in section IIA. [b] Not a minimum. [c] $E_d = E_{d^*} = -4.78$ kcal/mol.

tive to stabilize these conformations. Conformation hh is a type III bend, or part of an α helix. Conformations d_1g and hc_1 do not fall into any particular bend classification.

The dihedral angles and related data for the minimum-energy conformations described here are summarized in Table IV, together with those of the glycine dipeptide.

B. Gly-Gly. The movements of the dihedral angles in one cycle of minimization on Gly-Gly are shown in Figures 8A and 8B for the first and second residue, respectively. Since the movements are more spread out than in the case of L-Ala-L-Ala, wedges are sometimes used instead of arrows in Figure 8. Because of the plane of symmetry at the C^{α} atom in the glycine molecule, only half of the ϕ,ψ space is required to show the results of the energy minimization for each residue. The center of symmetry on the ϕ,ψ map is not only at (0°,0°) and (±180°,±180°) but also at (±180°,0°) and (0°,±180°). For example, the conformation (−180°,−60°) is the same as both (180°,60°) and (−180°,60°). Therefore, movement from the starting points (−180°,−60°) and (−180°,−120°) can be obtained by rotating the figure 180° around the position (−180°,0°). Thus the movements at (−180°,60°) and (−180°,120°) now represent those at (−180°,−60°) and (−180°,−120°), respectively. Although fewer starting points were taken for Gly-Gly than for L-Ala-L-Ala, the movements appear to be so regular that the number of points taken should suffice to locate all of the minima for Gly-Gly.

All of the low-energy minima of Gly-Gly, obtained by further cycles of minimization, are summarized in Table V, where they are expressed in terms of the glycine single-residue minima of Table VI. Since minimum b (−180°,180°) is located exactly at one of the symmetric centers mentioned above, point b is equivalent to its inverse b*; hence the bottom three entries of the column b(b*) become unnecessary; *e.g.*, h*b = hb* = hb. There are 28 different minima, with the global minimum at d*d (or dd*) instead of at dd; the

former (called a type V bend in ref 13) has a more negative value of E_{12} than the latter. Although conformation c_2 (c_2*) is not a minimum on the single-residue map, it is in a broad low-energy region and leads to dipeptide minima dc_2 (d*c_2*), d*c_2 (dc_2*), and c_1*c_2 (c_1c_2*). One particular case, which differs from the results on the alanine dipeptide, is the occurrence of a minimum at dh* (d*h). This minimum is a simple combination of single-residue minima in the glycine dipeptide but does not appear at all in the alanine dipeptide data of Table II. Those minima which satisfy the conditions for general bend formation (eq 7–9) are indicated by a superscript *c* in Table V. Those with large negative values of E_{12} (<−1.5 kcal/mol) are hc_1 (h*c_1*), hd (h*d*), hh (h*h*), d*c_1 (dc_1*), and d*h (dh*). All of these correspond to general bends and are included in the cross-hatched areas of Figure 6 (see Table IV for their dihedral angles). dh* is a type II bend with very low E_{12} energy, and d*h is of type II′. Similar to results for the alanine dipeptide, hd (h*d*) may be regarded as distorted type I (type I′) bends, and hh (h*h*) is a type III (type III′) bend or part of an α helix. Other general bend conformations with moderately large (negative) values of E_{12} are dc_1* (d*c_1) and hc_1 (h*c_1*). They do not fall into the usual bend categories (see Table IV), but hc_1 (h*c_1*) is rather close to a type I (type I′) bend. In the type II bend, the low-energy region of the second glycine residue extends from the left-handed α helix continuously toward the c_1* minimum, making this region special to the glycine dipeptide, and indicating that there would be a larger region for formation of a type II bend available to L-Ala-Gly (*i.e.*, to a dipeptide in which the second residue is Gly) than for L-Ala-L-Ala, as was reported by Chandrasekaran, *et al.*[12]

IV. Discussion

In the particular dipeptides studied here, the values of E_{12} fall in a small range. For example, the values of E_{12} for the glycine dipeptide (including all 943 starting points) fall in a range $-3.1 \leq E_{12} < 0$ kcal/mol in absolute energy and, for all starting points for the alanine dipeptide, fall in the range $-3.7 \leq E_{12} \leq +3.8$ kcal/mol (with only three out of 1156 conformations having positive values). The magnitudes of these terms may be compared to the total single-residue energy which was occasionally found to be hundreds of kilocalories per mole or greater above the global minimum (which is −2.6 kcal/mol for the glycine dipeptide and 0.44 kcal/mol for the alanine dipeptide). Thus, when all of conformational space is considered, the single-residue energy is the one that is primarily responsible for determining the accessible regions of low energy. This result supports the assumption made heretofore[3] that the influence of nearest neighboring residues on the conformation of a particular residue is, in general, relatively weak. However, at or near the minima, the E_{12} term may determine which total energy minimum is lowest. It remains to be seen whether these conclusions will hold for other dipeptides.

Those structures, for which the energy deviates from the sum of the single-residue energies (because of a significantly large and negative value of E_{12}), generally are bends, as shown in Tables II and V. However, in order to obtain a new minimum-energy conformation, which is specific to the dipeptide, a low value of E_{res} is also necessary. It seems that these conditions are difficult to fulfill since the only new minimum found in this study is the type II bend conformation, d_2h* for the alanine dipeptide. This conformation seems to be shifted from a position made up of single residue minima (*i.e.*, d_1h*) by the large negative value of E_{12}. The type II bend of the glycine dipeptide, dh*, is a simple combination of single-residue minima, which implies that, in this case, the E_{12} energy is not sufficient to force the conformation away from the single-residue minima.

The other typical bend conformation, type I, which would be an HF conformation (Figure 6), did not appear as a minimum energy one for either dipeptide and may require side-chain interactions of a type not found in alanine or glycine (or possibly longer range interactions or specific solvent effects). It was found that energy minimization starting from the HF region always resulted in an hd_1 conformation, which is a distorted form of type I bend, or to the hh conformation, a type III bend, or sometimes to the hc_1 conformation. We subsequently allowed the central dihedral angle ω to vary during the minimization, and the results remained the same. Thus, at this time we cannot speculate further as to this result, even though various amino acid sequences in proteins[20,21] (but not the Gly-Gly or L-Ala-L-Ala sequences) have been reported to have conformations in the type I bend regions. This problem is presently under investigation in our laboratory. Although the type I bend is obtained only in the distorted form, hd_1, Table IV shows that the relative stabilities of types I and II bends agree with the results of other authors;[12,13] *viz.*, comparing the total energies, type I is more stable for L-Ala-L-Ala, but type II is preferred for Gly-Gly (or, in general, when the second residue is Gly).

The method developed here can be applied to the search for the energy minima of a longer peptide chain such as Gly-Gly-Gly.[22] The new interaction energy term for such a tripeptide would be E_{13} between the first and third residues. If E_{13} were always zero, all of the energy minima of Gly-Gly-Gly would correspond to simple combinations of the dipeptide minima found in this study. Therefore, conformations that were specific to a tripeptide would have to have strong E_{13} interactions; moreover, such conformations can be expected to be structures in which the first and third residues come close to each other, as in the bend conformations of dipeptides. Thus, we could choose in advance the possible conformations with strong E_{13} interactions simply by geometric considerations, *e.g.*, by choosing conformations with small end-to-end distances, without consuming much time in minimizing the energy. Crippen and Scheraga[22] have already tried to locate the global minimum of Gly-Gly-Gly with a different method, *viz.*, one that excludes the higher energy regions of the multidimensional space.

The virtual-bond method described in section IA may be suitable for treating the skeletal shape of a peptide backbone. Since it is concerned only with the C^α atoms of the peptide chain, it is a simpler procedure than that in which ϕ's and ψ's are computed. However, it cannot treat the fine structure, *i.e.*, the orientation of the peptide plane, the direction of the side chain with respect to the backbone, etc. On the contrary, the virtual-bond method is independent of all parameters except the $C^{\alpha}_{i} \cdots C^{\alpha}_{i+1}$ distance, while the (ϕ,ψ) expression depends on many factors, such as bond angles, as seen in eq 11. This is important if X-ray crystallographic data, which are reported in Cartesian coordinates, are to be expressed in terms of other variables such as dihedral angles. Much of the error in the values of ϕ,ψ from experimental X-ray data can be considered to arise from ambiguities in the choice of bond angles, in addition to the experimental errors in the Cartesian coordinates. Another factor, which also introduces ambiguity into the (ϕ,ψ) values, is the degeneracy of the (ϕ,ψ) expression; *i.e.*, there is always one more variable in the (ϕ,ψ) expression than in the (θ,γ) expression as described in section IA. Therefore, there may exist a continuous set of values of (ϕ,ψ) for the same skeletal structure of a protein,[23] while

only a single set of (θ,γ) should satisfy the structure. For these reasons, we can regard the virtual bond method, as applied in the analysis of bends (Figures 2 and 3), as more reliably reflecting the original experimental data than the (ϕ,ψ) expression as far as only the skeletal backbone is concerned. The conversion from the virtual bond to the (ϕ,ψ) expression is possible by use of eq 10 and 11, with the aid of the standard geometry of the peptide unit.

Acknowledgment. We are indebted to Drs. A. W. Burgess and S. S. Zimmerman for helpful comments on this paper.

References and Notes

(1) This work was supported by research grants from the National Science Foundation (GB-28469X3), from the National Institute of General Medical Sciences of the National Institutes of Health, U.S. Public Health Service (GM-14312), and from Walter and George Todd.
(2) (a) Financial support from a Japanese Naito Research grant; (b) NIH Special Fellow, 1968–1969.
(3) (a) H. A. Scheraga, *Pure Appl. Chem.*, **36**, 1 (1973); (b) H. A. Scheraga, "Currents Topics in Biochemistry, 1973" C. B. Anfinsen and A. N. Schechter, Ed., Academic Press, New York, N. Y., 1974, pp 1–42.
(4) As used here, the terms "single residue" and "dipeptide" refer to the structures CH_3CO–X–$NHCH_3$ and CH_3CO–X–Y–$NHCH_3$, respectively, where X and Y are amino acid residues.
(5) P. N. Lewis, F. A. Momany, and H. A. Scheraga, *Isr. J. Chem.*, **11**, 121 (1973).
(6) The importance of longer (medium) range interactions in oligopeptides of various lengths has been discussed elsewhere.[7]
(7) P. K. Ponnuswamy, P. K. Warme, and H. A. Scheraga, *Proc. Nat. Acad. Sci. U. S.*, **70**, 830 (1973).
(8) C. M. Venkatachalam, *Biopolymers*, **6**, 1425 (1968).
(9) P. N. Lewis, F. A. Momany, and H. A. Scheraga, *Proc. Nat. Acad. Sci. U. S.*, **68**, 2293 (1971).
(10) I. D. Kuntz, *J. Amer. Chem. Soc.*, **94**, 4009 (1972).
(11) J. L. Crawford, W. N. Lipscomb, and C. G. Schellman, *Proc. Nat. Acad. Sci. U. S.*, **70**, 538 (1973).
(12) R. Chandrasekaran, A. V. Lakshminarayanan, U. V. Pandya, and G. N. Ramachandran, *Biochim. Biophys. Acta*, **303**, 14 (1973).
(13) P. N. Lewis, F. A. Momany, and H. A. Scheraga, *Biochim. Biophys. Acta*, **303**, 211 (1973).
(14) IUPAC-IUB Commission on Biochemical Nomenclature, *Biochemistry*, **9**, 3471 (1970).
(15) G. Nemethy and H. A. Scheraga, *Biopolymers*, **3**, 155 (1965).
(16) F. A. Momany, R. F. McGuire, and H. A. Scheraga, *J. Phys. Chem.*, to be submitted for publication. A *preliminary* listing of these geometric and energy parameters was given in ref 5.
(17) R. F. McGuire, F. A. Momany, and H. A. Scheraga, *J. Phys. Chem.*, **76**, 375 (1972).
(18) M. J. D. Powell, *Comput. J.*, **7**, 155 (1964).
(19) W. I. Zangwill, *Comput. J.*, **10**, 293 (1967).
(20) A. W. Burgess, P. K. Ponnuswamy, and H. A. Scheraga, *Isr. J. Chem.*, **12**, 239 (1974).
(21) F. M. Pohl, *Nature (London), New Biol.*, **234**, 277 (1971).
(22) G. M. Crippen and H. A. Scheraga, *Arch. Biochem. Biophys.*, **144**, 453 (1971).
(23) K. Nishikawa and T. Ooi, *J. Phys. Soc. Jap.*, **32**, 1338 (1972).

Dielectric Behavior of DL Mixtures of Poly(γ-benzyl glutamates)

Tohru Takahashi, Akihiro Tsutsumi, Kunio Hikichi,* and Motozo Kaneko

Department of Polymer Science, Faculty of Science, Hokkaido University, Sapporo 060, Japan.
Received April 12, 1974

ABSTRACT: Dielectric measurements were carried out for mixtures of poly(γ-benzyl L-glutamate (PBLG) and poly(γ-benzyl D-glutamate) (PBDG) over a frequency range from 110 Hz to 1 MHz and a temperature range from 20 to 130°. Mixtures of L and D molecules were found to show a dielectric dispersion (β dispersion) due to side-chain motion above room temperature as well as PBLG. It was found that the β dispersion for mixtures is reduced to a great extent and shifted to higher temperatures compared to PBLG. This fact was interpreted in terms of regular stacks of benzyl groups occurring in mixtures, which presumably restrict side-chain motion. A transition was observed for the mixtures in the vicinity of 90° in dielectric properties. The results obtained in the present work are consistent with a previous proposal that the transition is caused by the breakdown of regular stacks of benzyl groups.

Molecular motions of poly(γ-benzyl L-glutamate) (PBLG) have been investigated by means of various methods. Kail, *et al.*,[1] and one of the present authors (K. H.)[2] observed changes in the second moment of nuclear magnetic resonance spectra for PBLG with temperature, which could be accounted for wholly in terms of side-chain motion. The mechanical and dielectric relaxations due to the side-chain motion have also been studied.[3-10]

Recently, attention has been directed to the study of the racemic form of poly(γ-benzyl glutamate) (PBG), which is a mixture of equal parts by weight of poly(γ-benzyl L-glutamate) (PBLG) and poly(γ-benzyl D-glutamate) (PBDG).

X-Ray diffraction patterns of racemic PBG show the appearance of "extra" layer lines which cannot be interpreted in terms of the α helix.[11-15] Squire and Elliott[15] suggested that these layer lines are caused by regular stacks of benzyl groups from about five side chains which are formed by interaction between adjacent molecules of opposite screw sense.

On the basis of X-ray diffraction, differential thermal analysis, infrared absorption, and dilatometric measurements on the racemic form of PBG, Yoshikawa, *et al.*,[16] reported that the racemic form of PBG exhibits a reversible first-order solid–solid phase transition at a temperature of about 100°. They suggested that the phase transition is caused by the breakdown of regular stacks of benzyl groups.

Our earlier dielectric studies of PBLG showed a dispersion due to side-chain motion above room temperature.[4] Tsuchiya, *et al.*,[17] have carried out dielectric measurements for the racemic form of PBG. They reported that the dispersion due to side-chain motion is much broader and is shifted to higher temperatures compared to PBLG.

However, no detailed dielectric investigation has been made in view of the phase transition. It is of interest to examine the relationship of the dielectric properties of the racemic form of PBG and the phase transition. In this work dielectric investigations were carried out for mixtures of PBLG and PBDG. Results obtained were interpreted in terms of stacking of benzyl groups.

Experimental Section

PBLG and PBDG used in this work were prepared by polymerization of the respective *N*-carboxyanhydrides (NCA's) in dioxane. A trace of triethylamine was used as an initiator. Polymers were precipitated from the polymerization solution with ethanol and