(9) (a) J. A. Ross, R. P. Seiders, and D. M. Lemal, *J. Am. Chem. Soc.*, **98**, 4325 (1976); (b) C. H. Bushweller, J. A. Ross, and D. M. Lemal, *ibid.*, **99**, 629 (1977).
(10) H. Kwart and T. J. George, *J. Am. Chem. Soc.*, **99**, 5214 (1977).
(11) E. L. Allred, B. R. Beck, and N. A. Mumford, *J. Am. Chem. Soc.*, **99**, 2694 (1977), and references cited therein.
(12) K. B. Lipkowitz and R. Larter, *Tetrahedron Lett.*, 33 (1978).
(13) (a) Carbon monoxide bond length: 1.128 Å from G. Herzberg, "Spectra of Diatomic Molecules", Van Nostrand, Princeton, N.J., 1966. (b) Hydrogen cyanide bond lengths: C–H 1.068 Å, C–N 1.156 Å from G. Herzberg, "Electronic Spectra and Electronic Structure of Polyatomic Molecules", Van Nostrand, Princeton, N.J., 1945. (c) D_{6h} benzene bond lengths: C–C 1.390 Å, C–H 1.100 Å, R. T. Morrison and R. N. Boyd, "Organic Chemistry", 3rd ed., Allyn and Bacon, Boston, Mass., 1973.
(14) P. A. Dobosh, QCPE 223, 1968.
(15) It would be to our advantage to use a more sophisticated basis set and include CI but the cost of such a treatment, particularly for the larger systems, is unduly prohibitive. A more serious concern, though, is that in this paper when the wave functions for monomer A and B strongly interact, we treat the complex AB as a single molecule. Here, the basis set for the complex is the union of the basis set of the two monomers. If the monomer basis sets are small or incomplete a basis set extension phenomenon should be of concern. That is, if small basis sets are used, as the two monomers are brought together the orbitals on B improve the monomer properties of A and the orbitals on A improve the monomer properties of B. Since the change in any property P as a function of intermolecular separation is the calculated property of the dimer AB minus the sum of the calculated properties of the two monomers, $\Delta P = P_{AB} - (P_A + P_B)$, a poor representation of the interaction results. Elimination of basis set extension effects is being investigated by using ghost orbitals. N. S. Ostlund and D. L. Merrifield, *Chem. Phys. Lett.*, **39**, 612 (1976), and references cited therein.
(16) D. Goutier and R. Macaulay, QCPE 241, 1974.
(17) J. S. McKennis, L. Brener, J. S. Ward, and R. Pettit, *J. Am. Chem. Soc.*, **93**, 4957 (1971).
(18) (a) For leading references see A. R. Gregory and M. N. Paddon-Row, *J. Am. Chem. Soc.*, **98**, 7521 (1976). Of interest, though, is that a CNDO calculation of two eclipsed benzenes shows a repulsive interaction (as anticipated) when they are brought together. (b) D. B. Chestnut and P. E. S. Wormer, *Theor. Chim. Acta*, **20**, 250 (1971). (c) It was suggested that the perturbative procedure, PCILO, may be better adapted to the study of molecular complexes than CNDO. See, for example, R. Lochman and H. J. Hofmann, *Int. J. Quantum Chem.*, **11**, 427 (1977); R. Arnaud, D. Faramond-Baud, and M. Gelus, *Theor. Chim. Acta*, **31**, 335 (1973). (d) We are now trying the MNDO method (M. J. S. Dewar and W. Thiel, *J. Am. Chem. Soc.*, **99**, 4399 (1977)) in hopes that we can accurately reproduce the ab initio results with an approximate molecular orbital technique.
(19) G. Maier, U. Schäfer, W. Sauer, R. Matusch, and J. Oth, *Tetrahedron Lett.*, In press. We thank Professor Maier for a preprint of this work.
(20) The work cited in ref 19 indicates a carbonyl IR shift to lower frequency upon complexation with tri-*tert*-butylcyclobutadiene. This corresponds to a 0.22 kcal mol^{-1} change in bond energy between free and complexed carbon monoxide. We calculate (Figure 5) an attractive interaction between carbon monoxide and cyclobutadiene of 0.30 kcal mol^{-1} for geometry III.
(21) Since this work was completed, a communication has appeared confirming the destabilizing effect of C_4H_4 with carbon dioxide as we had originally reported in ref 12. Also reported is an energy minimum for cyclobutadiene with carbon monoxide. These results were obtained using MNDO type calculations and suggest that MNDO mimics ab initio results quite well. A. Schweig and W. Thiel, *Tetrahedron Lett.*, 1841 (1978).

Quantum Mechanical Partitioning of the Energies of Polypeptides. Conformational Study of Polyglycine

Steve Scheiner*[1] and C. W. Kern

Contribution from the Department of Chemistry, The Ohio State University, Columbus, Ohio 43210. *Received May 2, 1978*

Abstract: The total energy of a polypeptide is partitioned into the interaction energies between pairs of residues. The interactions between adjacent residues are evaluated by quantum mechanical methods. An empirical hydrogen bond potential based on the quantum mechanical interaction between amide units is formulated and used to calculate the interaction energy between nonadjacent residues. This treatment is used to study regular helical conformations of an isolated single strand of polyglycine. The α helix is found most stable while the 3_{10} and 2_7 helices are also low-energy structures. The calculated results are compared to those obtained via classical partitioning procedures as well as experiment.

I. Introduction

Reliable means of evaluating energies of proteins are needed for predictive and interpretative studies of three-dimensional structure. Although quantum mechanical procedures have yielded results in excellent agreement with experiment when applied to many chemical systems, including oligopeptides,[2-4] these methods are subject to severe size limitations. An ab initio calculation of even the smallest naturally occurring protein is well beyond the range of modern computers and programming techniques. Pentamers are the largest polypeptide units that have been studied so far by quantum mechanical techniques.[3,4]

More approximate approaches[5-20] have therefore been devised to obtain the total energy of polypeptides. Most of these use simple expressions for the potential energy in which the parameters are evaluated empirically for each term that is thought to be physically significant (e.g., electrostatic terms, bond stretching terms). These partitioning methods can be applied to very large systems, since the energy of any conformation is computed in a fraction of the time required for an analogous quantum mechanical calculation. However, such partitioning is rather arbitrary, as are the mathematical forms chosen for the potential functions and the values of the empirical parameters.

We report here an alternate method of energy partitioning in which all contributions are evaluated quantum mechanically.[21] This treatment also employs empirical parameters, but only in identifying those high-energy conformations in which certain residues are in close proximity.

II. Methodology

Previous quantum mechanical studies of polypeptides indicate that the total energy of certain conformations can be expressed approximately as the sum of pairwise interactions between residues. Using ab initio and approximate ab initio methods, respectively, Shipman and Christoffersen[3] and Kleier and Lipscomb[4] examined oligomers of glycine ranging in size up to the pentamer level. Only regular helices were considered in which the pair of dihedral angles ϕ and Ψ about the C^α atom were identical for each residue. In such structures, the relative orientation of any two peptide units is dependent only on the values chosen for the dihedral angles ϕ and Ψ, and on the number of additional units that separate the units of interest along the polypeptide chain. The interaction energy between each pair of peptide units was assumed to be dependent only on their relative orientation. Numbering peptide units consecutively along the chain, the interaction energy, ΔE_m, of two peptide units, i and $(i + m + 1)$, in a regular helical n-peptide

0002-7863/78/1500-7539$01.00/0

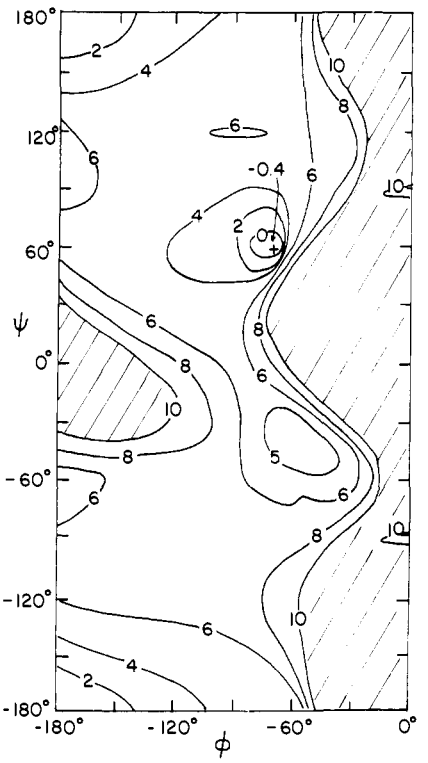


Figure 1. PRDDO conformational energy map of *N*-formylglycylamide. The geometry was chosen using the standard values of Scheraga.[19] Peptide groups were held in planar trans configurations. Energies (kcal mol^{-1}) are relative to the fully extended conformation $(\Phi, \psi) = (-180°, -180°)$. Shaded areas represent regions of energy greater than 10 kcal mol^{-1}. Energies were computed at 30° intervals except in the vicinities of the absolute minimum (−70°, 60°) and the α helix (−48°, −57°), where a finer grid was used. A center of symmetry is present at (0°, 0°). The angles ϕ and Ψ are defined according to ref 31.

was thus assumed to be independent of the position of the pair along the chain, i, and of the total chain length, n.

Since the total number of pairs of units separated by m other units in an n-peptide is equal to $n - m - 1$, the total energy of such a polymer, E_n, can be expressed as[4]

$$E_n = \sum_{m=0}^{n-2} [(n - m - 1)\Delta E_m] \qquad (1)$$

For each of several conformations, as defined by the angles ϕ and Ψ, the interaction energy between adjacent peptide units, ΔE_0, was evaluated as the difference in energy between that conformation and the fully extended (FE) reference configuration as calculated quantum mechanically at the dimer level. After similar calculations were performed on the trimers, ΔE_1 was obtained by substituting the previously determined value of ΔE_0 into eq 1. This procedure was continued until values were obtained for all ΔE_m up to $m = 3$ for several different conformations. For both parallel and antiparallel pleated sheet structures, ΔE_m $(m > 0)$ was found to be of much smaller magnitude than ΔE_0. This result implies that the extended nature of these strands precludes any strong interactions between nonadjacent units. The results for the α and 3_{10} helices were found to be consistent with the hydrogen-bonding patterns predicted for these structures. A large negative energy was obtained for ΔE_2 for the α helix reflecting the hydrogen bond between units i and $i + 3$. The proposed bond between units i and $(i + 2)$ in the 3_{10} helix was manifested in the value found for ΔE_1 for that structure. Other values of ΔE_m $(m \neq 0)$ were calculated to be of considerably smaller magnitude.

In these studies[3,4] little interaction was found between those pairs of residues that are neither adjacent nor in positions suitable for hydrogen bonding. This suggests that the total energy might be approximated as the sum of two contributions. The first is the "adjacent energy", which is the sum of interactions between all pairs of adjacent peptide units ($\Sigma\Delta E_0$). This term alone is sufficient to approximate the energies of the extended pleated sheet structures. For more compact conformations such as the α and 3_{10} helices, the "total nonadjacent" energy must be added to the adjacent energy. This contribution is dominated by the interactions between those pairs of peptide units suitably positioned for hydrogen bonding, e.g., pairs (1,4) and (2,5) in the α helical pentamer. (Non-hydrogen-bonding type interactions may be important for certain nonhelical conformations.)

The conformation of an n-peptide may be described partially by the $(n - 1)$ pairs of dihedral angles ϕ and Ψ. Each pair of angles, ϕ_i and Ψ_i, determines the relative orientation of adjacent peptide units i and $(i + 1)$. In order to evaluate the adjacent energy of a polypeptide, ϵ_{adj}, the interaction energy of a pair of adjacent peptide units must therefore be known as a function of ϕ and Ψ. The necessary conformational energy maps are obtained in our method via quantum mechanical calculations performed on the dipeptide. Such maps of the glycyl residue have been obtained previously using various ab initio and semiempirical techniques.[22-29] We have chosen to apply our procedure to the map in Figure 1 of *N*-formylglycylamide[29] as calculated by the partial retention of diatomic differential overlap (PRDDO)[30] method.

In order to evaluate the total hydrogen bond energy, the interaction between peptide units must be known as a function of their relative orientation. It is important that such a function reflect properly both the radial and angular dependence of the N-H···O=C interpeptide hydrogen bond. An ab initio study[32] has found most stable an alignment of the NH group approximately along a lone pair of the carbonyl oxygen. This conclusion is consistent with crystal structures of various amides and peptides.[11b,33,34] In addition, a linear hydrogen bond, i.e., $\angle$NHO $\sim$ 180°, has been found most stable by quantum mechanical techniques.[35,36]

Previous classical energy partitioning methods have made use of various hydrogen bond potential functions. DeSantis et al.[5] used a potential originally proposed by Stockmayer[37] which consisted of the sum of a 6-12 radial potential and a term containing angular dependence. The latter term was formulated as the electrostatic interaction of two dipoles centered on the O and H atoms involved in the H bond. The minimum in this potential occurs when the C=O and N-H bonds are collinear, thus neglecting the directionality of the lone pairs of the oxygen. Poland and Scheraga[6] modified the electrostatic term eliminating the explicit angular dependence which is absent also in the modifiedMorse potential of Popov et al.[7] Scott and Scheraga[8] used a form of the Schroeder-Lippincott potential[38] as modified by Moulton and Kromhaut.[39] The minimum in their potential, however, occurs for $\angle$COH = 90° whereby the NH group is 30° from one lone pair and 150° from the other. This function, furthermore, makes no provision for the nonlinearity of the H bond. The potentials proposed by Gibson and Scheraga,[10] by Ramachandran et al.,[11] and by Singh and Ferro[40] consider this nonlinearity but not the angular dependence about the carbonyl oxygen. The potential of Brant[12] is a function of both $\angle$NHO and $\angle$HOC but contains cylindrical symmetry about the CO and NH bonds.

Potentials that contain no explicit angular dependence have reappeared more recently[33] and are currently used in the classical energy partitioning scheme of Scheraga et al.[14] The electrostatic and nonbonded energy contributions to the total energy are instead relied on to simulate the angular dependence of hydrogen-bonding interactions and are claimed[33] to reproduce accurately the results of quantum mechanical calculations. To test this, we have examined the N-H···O=C hy-

drogen bond of the linear formamide dimer. Idealized geometries were used for the formamides and the planes of the two molecules were chosen to be mutually perpendicular. A linear hydrogen bond ($\angle OHN = 180°$) was assumed and the angle $\theta = \angle COH$ varied. The electrostatic and nonbonded energies were computed by the classical partitioning method of ref 33. The nonbonded energy parameters were taken from Scott and Scheraga.[8] Two different sets of partial atomic point charges were used to evaluate the electrostatic energy. These charges were taken from CNDO/2[33] and PRDDO calculations on the isolated molecule. Energies were calculated,both by the PRDDO and classical partitioning methods, at three values of the hydrogen bond length, r_{OH} = 1.4, 1.7, and 2.0 Å. According to PRDDO, the preferred conformation for all three distances is the one in which the NH group is collinear with a lone pair of the carbonyl oxygen ($\theta = 120°$), in agreement with theoretical and experimental predictions.[11b,32-34] By contrast, the minimum in the classical partitioned energy was found to be a completely linear arrangement ($\theta = 180°$) for all three values of r_{OH}. (An "apparent dielectric constant" of 2 is assumed by McGuire et al.[33] Changing this constant to unity had no effect on this result.) The minimum in the classical electrostatic contribution to the total energy was also found at $\theta = 180°$ whereas the minimum in the quantum mechanical electrostatic energy[32] was found at $\theta \sim 135°$. The empirical formulation of Scheraga et al. therefore fails to reproduce the angular features of the quantum mechanical interaction for the formamide dimer.

In this study, the interaction energy between nonadjacent peptide units was evaluated using an empirical function which was fit to the quantum mechanical interaction energy between two amide units. Because the number of possible geometries is very large, we kept the number of atoms in the amide-amide system to a minimum by using formamide as a model. It was expected that the smaller formamide system will yield the essential features of the interpeptide interaction energy surface. Each formamide molecule was taken to be fully planar with idealized bond angles and bond lengths. [All bond angles were set equal to 120°, r(CO) = 1.24 Å, r(CH) = 1.10 Å, r(CN) = 1.32 Å, r(NH) = 1.00 Å.] The formamides were positioned such that the CO group of one molecule (f_1) could interact with an NH group of the second (f_2) (see Figure 2).

We take the carbonyl oxygen atom of f_1 as the origin of a polar coordinate system. The polar axis is taken to be perpendicular to the plane of f_1 and the C=O bond axis is the x axis. The position of the hydrogen bonding proton of f_2 may be described by the spherical polar coordinates (r, θ_1, ϕ_1) as shown in Figure 2a. When the proton of f_2 lies along the direction of either lone pair of an sp^2-hybridized carbonyl oxygen of f_1, the angles θ_1 and ϕ_1 take the values of ±60 and 90°, respectively. Similarly, we define a second coordinate system with the hydrogen bonding proton of f_2 as its origin (Figure 2b). The polar axis is perpendicular to the plane of f_2 and the NH axis is taken as the x axis. The position of the carbonyl oxygen of f_1 is then specified by the coordinates (r, θ_2, ϕ_2). A linear hydrogen bond would thus be denoted by (θ_2, ϕ_2) = (0, 90°). Note that r represents the internuclear distance between the oxygen of f_1 and the hydrogen bonding proton of f_2 in both coordinate systems. A sixth geometrical parameter is needed to fix completely the relative position and orientation of the two amide groups. In the case of a fully linear C=O···H-N arrangement ($\theta_1 = \theta_2 = 0°$, $\phi_1 = \phi_2 = 90°$), this parameter specifies the dihedral angle made by the two molecular planes about the hydrogen bond axis. The energy was found to be rather insensitive to these rotations in agreement with an ab initio study.[41] We therefore do not include this parameter in our empirical hydrogen bond function.

The PRDDO method was used to compute the interaction energy surface. The dominant effects for many configurations

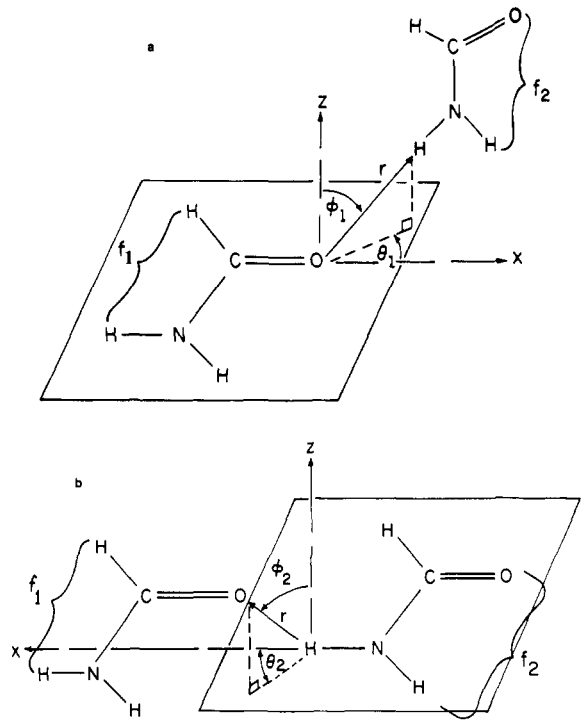


Figure 2. Spherical polar coordinate systems used to define the geometry of the hydrogen bond between two amide units.

were found to be steric repulsions between the two molecules. Since we are interested primarily in the attractive part of the total hydrogen bond interaction, it was convenient to subtract this repulsion energy, ϵ_{rep}, from the quantum mechanical interaction to obtain an "adjusted hydrogen bond energy" ϵ_{HB} to which the empirical function was fitted. (This repulsion energy must, of course, be added back to obtain the full interaction in the polypeptide.) The repulsion energy was identified with the sum of all positive pairwise energies between atoms on different molecules (exclusive of the H···O pair involved in the hydrogen bond) and calculated via Lennard-Jones potentials.[8] The two pairs of atoms of the formamide dimer that consistently yielded the greatest steric repulsions (as indicated by Mulliken overlap populations) were the CH and ON pairs of the C=O···H-N hydrogen bond. These overlaps were found to be in poor correlation with the repulsion energies calculated via the parameters of Scott and Scheraga.[8] In addition, the locations of the minima of the adjusted hydrogen bond energies fluctuated erratically as r was varied. Much better correlation was obtained between overlap populations and repulsion energies when the values of r_{min} in Table IV of ref 8 were altered to 2.50 and 3.15 Å for the CH and NO pairs, respectively. With these changes, the minima in the angular part of the adjusted hydrogen bond energy were independent of r for 1.4 Å $< r <$ 2.4 Å. The adjusted values of the repulsive parameters were respectively 1.56×10^4 and 1.78×10^5 for the CH and NO pairs while the attractive parameters were unchanged.

The functional form

$$\epsilon_{HB} = R(r)\,\Theta_1(r, \theta_1, \phi_1)\,\Phi_1(r, \phi_1)\,\Theta_2(r, \theta_2)\,\Phi_2(r, \phi_2) \quad (2)$$

was used to fit the calculated adjusted hydrogen bond energies ϵ_{HB}. Here,

$$R(r) = \frac{A_l}{r^{m_l}} - \frac{B_l}{r^{n_l}} \quad (l = 1, 2, 3) \quad (3)$$

was constructed by dividing r into three regions such that m_l and n_l are constant integers throughout a given range, l. The parameters A_l and B_l were chosen not only to reproduce the

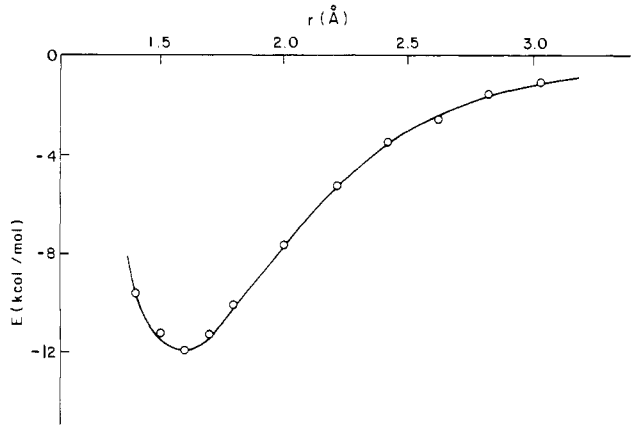


Figure 3. Comparison of the radial dependence of quantum mechanical results (circles) and the empirical function (solid curve) for the coplanar formamide dimer. $\phi_1 = \phi_2 = 90°$, $\theta_2 = 0°$, $\theta_1 = 60°$.

data but also to ensure continuity of both $R(r)$ and dR/dr at the boundaries between the three regions. The values of each of these constants are given in Table I.

To obtain the angular part of eq 2, we used the fact that minima occur in the adjusted hydrogen bond energy at $\phi_1 = \pi/2$, $\phi_2 = \pi/2$, and $\theta_2 = 0$ regardless of the values of r and θ_1. The functions Φ_1, Φ_2, and Θ_2 were therefore taken as

$$\Gamma_k(r, \gamma_k) = 1 - \frac{a_k}{\alpha_k(r)} \gamma_k{}^{b_k} \quad (k = 1, 2, 3) \tag{4}$$

where γ_k refers to the angular parameters defined in Table II. (Here, angles are in units of degrees and distances in Å.) Note that $\Phi_1(r, \phi_1)$ and $\Phi_2(r, \phi_2)$ are symmetric about $\phi_1 = 90°$ and $\phi_2 = 90°$, respectively; that is, these functions take identical values for $\phi_m = 90° \pm \delta$. This symmetry is a consequence of the planar conformation of the amide groups.

Both lone pairs on the carbonyl oxygen are available for hydrogen bonding. The function $\Theta_1(r, \theta_1, \phi_1)$ was thus found to have the characteristic shape of a double-well potential. The minima, $\theta_1{}^0$, occur at positive and negative values of θ_1. We express Θ_1 conveniently in the form

$$\Theta_1(r, \theta_1, \phi_1) = \begin{cases} 1 - \alpha_4(r, \phi_1) \sin^2\left[\dfrac{90}{\theta_1{}^0}(\theta_1{}^0 - \theta_1)\right] & \theta_1 < \theta_1{}^0 \\ 1 - \alpha_5(r, \phi_1)(\theta_1 - \theta_1{}^0)^{5/2} & \theta_1 \geq \theta_1{}^0 \end{cases} \tag{5}$$

where

$$\theta_1{}^0 = \begin{cases} 60[1 - \sin^4(1.2\,\bar{\phi}_1)] & \bar{\phi}_1 \leq 75 \\ 0 & \bar{\phi}_1 > 75 \end{cases} \tag{6}$$

is a function of $\bar{\phi}_1 = |\phi_1 - 90|$.

The functions α_4 and α_5 are expressed by

$$\alpha_4(r, \phi_1) = \begin{cases} 0.15[60 - \bar{\phi}_1] \exp(-2.3r) & \bar{\phi}_1 \leq 60° \\ 0 & \bar{\phi}_1 > 60° \end{cases} \tag{7}$$

$$\alpha_5(r, \phi_1) = 3.9 \times 10^{-5}\left[1 + \left(\frac{\bar{\phi}_1}{40}\right)^{3.4}\right]^{-1} [r - 1.1]^{-1}$$

Equations 5-7 are applicable for positive values of θ_1. The expressions for negative values of θ_1 are similar but slightly more complicated, as follows:

$$\Theta_1(r, \theta_1, \phi_1) = \begin{cases} T_0 - T_1 \sin^2\left[\dfrac{90}{\theta_1{}^0}(\theta_1{}^0 - \theta_1)\right] & \theta_1 > \theta_1{}^0 \\ T_0 - \alpha_6(\phi_1)(\theta_1 - \theta_1{}^0)^2 & \theta_1 \leq \theta_1{}^0 \end{cases} \tag{8}$$

Table I. Parameters Used to Evaluate $R(r)$

l	range, Å	m_l	n_l	A_l, kcal mol^{-1} Å^{m_l}	B_l, kcal mol^{-1} Å^{n_l}
1	$1.4 < r \leq 1.6$	5	3	187.2	121.9
2	$1.6 < r \leq 2.0$	13	3	1608	63.37
3	$2.0 < r$	13	5	17 818	316.8

Table II. Parameters Used in Equation 4

k	γ_k	$a_k \times 10^5$	$\alpha_k(r)$	b_k
1	$\lvert\phi_1 - 90\rvert$	7.0	$1 + [(r - 1.4)/0.6]^{0.85}$	2.2
2	$\lvert\phi_2 - 90\rvert$	16.7	$r - 1.1$	2.0
3	θ_2 ($\theta_2 \geq 0$)	3560	$\exp(2.3r)$	1.8
	$-\theta_2$ ($\theta_2 < 0$)	254		2.5

where

$$\alpha_6(\phi_1) = 5 \times 10^{-4}\left[1 + \left(\frac{\bar{\phi}_1}{35}\right)^{3.2}\right]^{-1} \tag{9}$$

$$T_0 = \begin{cases} 0.92 + 1.33 \times 10^{-3}\bar{\phi}_1 & \bar{\phi}_1 \leq 60° \\ 1 & \bar{\phi}_1 > 60° \end{cases} \tag{10}$$

$$T_1 = \begin{cases} -0.08 + 1.33 \times 10^{-3}\bar{\phi}_1 \\ \quad + 0.15[60 - \bar{\phi}_1] \exp(-2.3r) & \bar{\phi}_1 < 60° \\ 0 & \bar{\phi}_1 \geq 60° \end{cases} \tag{11}$$

$$\theta_1{}^0 = \begin{cases} -45 & \bar{\phi}_1 < 30 \\ \frac{3}{2}\bar{\phi}_1 - 90 & 30 \leq \bar{\phi}_1 \leq 60 \\ 0 & \bar{\phi}_1 > 60 \end{cases} \tag{12}$$

If the value of any angular functions in eq 2 becomes negative, the hydrogen bond energy for that interaction is set equal to 0. Equations 2-12 are applicable only for $r \geq 1.4$ Å. In all cases where $r < 1.4$ Å, ϵ_{HB} was also set equal to 0. The resulting discontinuity in the potential had no effect in every conformation considered below because large nuclear repulsions are found to dominate all terms in which there was present a value of $r < 1.4$ Å.

The correlation between the quantum mechanical interaction energies (circles) and the values obtained via the empirical function (solid curve) is exhibited in Figures 3-5. The radial function is shown in Figure 3 to be in excellent agreement with the calculated adjusted hydrogen bond energies. The relative orientation of the two amides corresponds to a linear hydrogen bond positioned along a lone pair of the carbonyl oxygen. The absolute minimum in the empirical potential ϵ_{HB} occurs at $r = 1.6$ Å, for which $\epsilon_{HB} = -11.9$ kcal mol^{-1}. The adjusted interaction energies are shown as a function of θ_1 for three values of r in Figure 4a. For each value of r, two minima are found at about -45 and 60°, reflecting the hydrogen bonding to each of the two lone pairs of the carbonyl oxygen. The dependence of the positions of the minima on the value of ϕ_1 may be seen in Figure 4b. A decrease of ϕ_1 below 90° corresponds to a rotation of formamide f_2 up out of the amide plane of f_1 (Figure 2). This rotation removes the NH group of f_2 from the plane in which lie the two lone pairs of the carbonyl oxygen. A decrease in ϕ_1 consequently results in a shift in the positions of the minima toward smaller values of θ_1. The agreement between quantum mechanical and empirical energies is quite good for $\phi_1 \geq 45°$ with an exception occurring for large negative values of θ_1. The empirical potential predicts accurately the shapes of the curves as well as the positions of the minima for all values of ϕ_1. The correlation is shown as a function of ϕ_1 in Figure 5a. The minima are found to occur at $\phi_1 = 90°$, i.e., when the two molecules are coplanar. Several of these

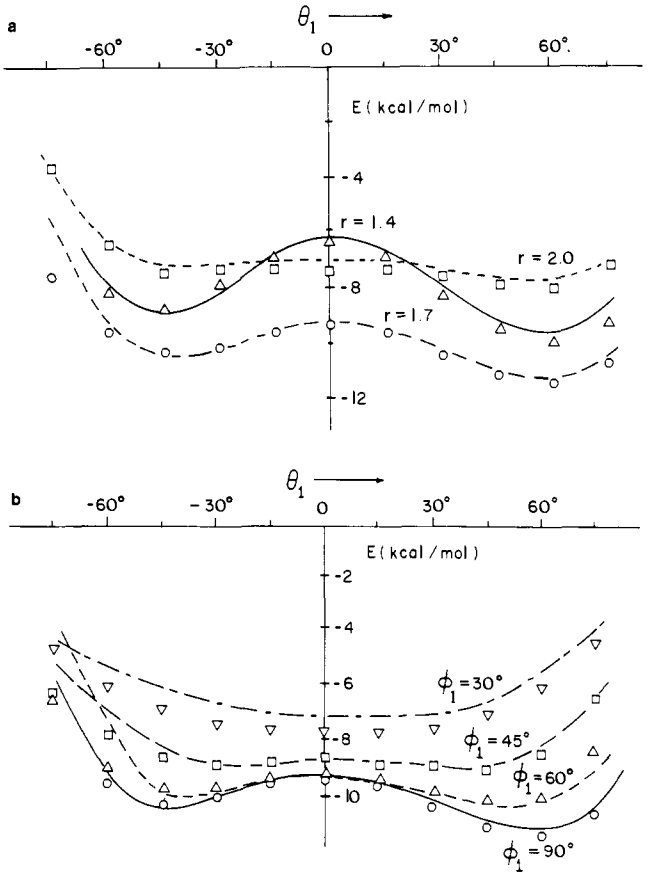


Figure 4. Interaction energies as a function of θ_1. Curves represent empirical values and points (circles, squares, etc.) indicate quantum mechanical results. (a) $\phi_1 = \phi_2 = 90°$, $\theta_2 = 0°$. (b) $r = 1.7$ Å, $\phi_2 = 90°$, $\theta_2 = 0°$.

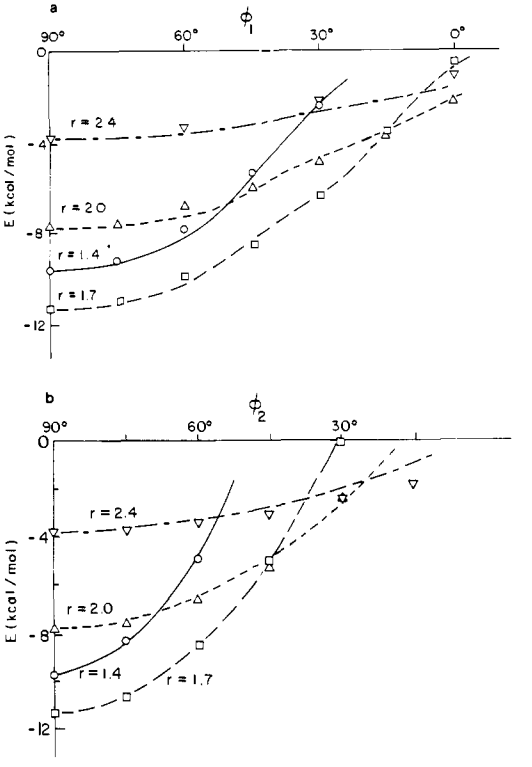


Figure 5. Quantum mechanical and empirical interaction energies as a function of (a) ϕ_1. $\theta_1 = 60°$, $\theta_2 = 0°$, $\phi_2 = 90°$. (b) ϕ_2. $\phi_1 = 90°$, $\theta_1 = 60°$, $\theta_2 = 0°$.

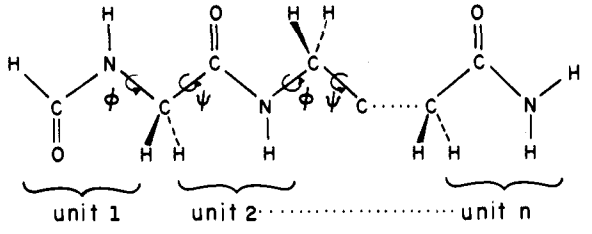


Figure 6. n-Mer of glycine. Geometries were obtained as described in the caption to Figure 1.

curves, particularly those for small values of r, are shaped rather irregularly since the ϕ_1 dependence is distributed between Θ_1 and Φ_1 (see eq 5-12). The empirical potential nonetheless reproduces the quantum mechanical results rather well. The interaction energy as a function of ϕ_2 is shown in Figure 5b. The empirical potential, which is a quadratic function of ϕ_2, once again comes quite close to the quantum mechanical one. The Θ_2 function is not shown but these results are also in good agreement. The minimum in the Θ_2 potential occurs at $\theta_2 = 0°$ regardless of the values of the other parameters, and is asymmetric about this point owing to the absence of a plane of symmetry perpendicular to the plane of the formamide molecule.

Although most of these curves were obtained by maintaining several geometrical parameters at their optimum values and varying one other, the results are essentially the same when the parameters are varied simultaneously. The empirical potential was found to agree with calculated quantum mechanical interaction energies within 1 kcal mol^{-1} in the great majority of the configurations examined. Considerably greater accuracy was found in those conformations likely to be of importance in a polypeptide structure.

The energy, E, of a given conformation of polyglycine is therefore computed by our procedure using the equation

$$E = E_{adj} + E_{HB} + E_{rep}$$

$$= \sum_i \sum_{j=i+1} \epsilon_{adj}^{ij} + \sum_i \sum_{j>i+1} \epsilon_{HB}^{ij} + \sum_i \sum_{j>i+1} \epsilon_{rep}^{ij} \quad (13)$$

The values of ϕ and Ψ for each pair of adjacent residues are used to obtain a value of ϵ_{adj} from the conformational map of Figure 1. The adjacent energy E_{adj} is equal to the sum of ϵ_{adj} for all such adjacent pairs. The relative orientation of each pair of nonadjacent units is then calculated in terms of the geometrical parameters defined in Figure 2. The values of these parameters are substituted into eq 2 to obtain the adjusted hydrogen bond energy, ϵ_{HB}, of that pair. The total hydrogen bond energy, E_{HB}, is calculated as the sum of ϵ_{HB} over all nonadjacent pairs. Finally, the total repulsion energy, E_{rep}, is computed as the sum of all positive interaction energies between pairs of atoms on nonadjacent units using the Lennard-Jones potentials of Scott and Scheraga[8] modified as described above. (The interactions between the carbonyl O of one peptide unit and the amide H of the other are excluded. The repulsion energy between adjacent units is included implicitly in the value obtained for E_{adj}.)

III. Results

Our procedure was tested first on the small n-mers ($3 \leq n \leq 6$) of glycine shown in Figure 6. The total energies of α, 3_{10}, π, and 2_7 helices as well as parallel (P) and antiparallel (AP) chain pleated sheet structures were calculated relative to the fully extended conformation for each value of n. These energies were then compared to those obtained by quantum mechanical (PRDDO) calculations on the same molecules. The maximum difference and the average deviation between the partitioned and PRDDO energies were found to be 1.0 and 0.3 kcal mol^{-1}

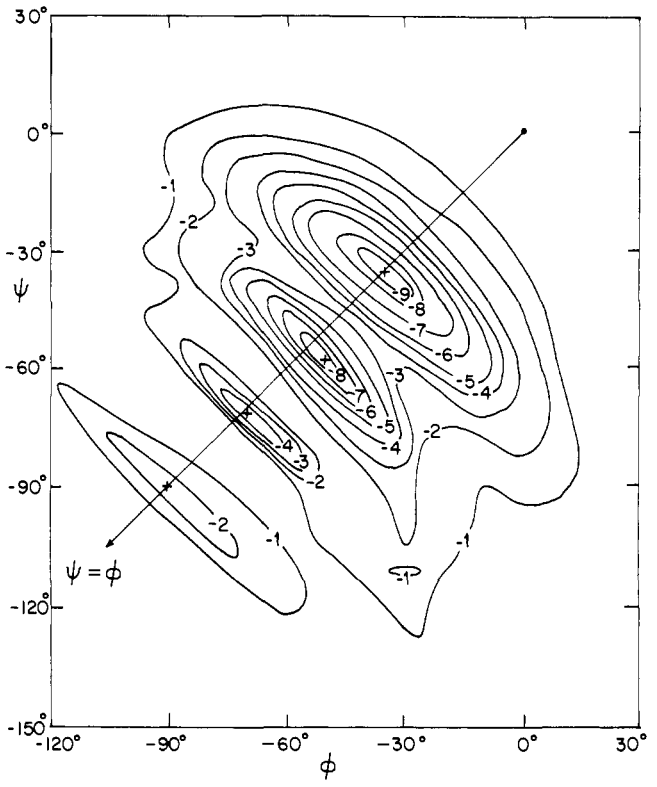


Figure 7. Conformational map of the total hydrogen bond energy, E_{HB} (kcal mol^{-1} residue^{-1}), for the 20-mer of glycine. $|E_{HB}| < 1$ for regions not included in this map, e.g., $\phi < -120°$. E_{HB} was computed at 10° increments of ϕ and Ψ except near minima (+ sign) where a finer grid was used.

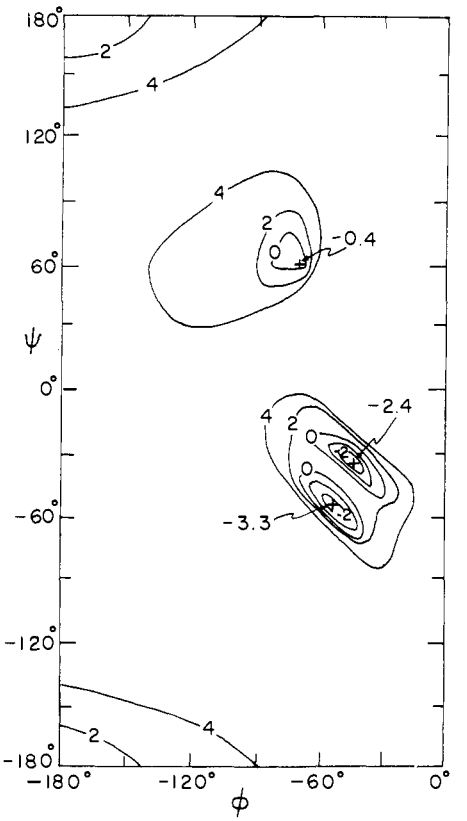


Figure 8. Total energy $E = E_{adj} + E_{HB} + E_{rep}$ of 20-mer of glycine. All energies are in units of kcal mol^{-1} residue^{-1} and relative to the fully extended structure (−180°, −180°). Contours greater than 4 are not shown.

residue^{-1}, respectively. Typical results are given in Table III for the hexamer.

The method was applied next to polymers of glycine containing 20 peptide units, representative of the longer chains found in globular proteins. Energies for regular helical conformations of the 20-mer were determined as a function of the dihedral angles ϕ and Ψ (Figure 6). A conformational energy map of the total hydrogen bond energy, E_{HB}, is presented in Figure 7. (Since a center of symmetry must occur at $(\phi, \Psi) = (0°, 0°)$ in conformational maps of polyglycine, only half-contours are shown.) Significant attractive interaction between nonadjacent residues is found when $-120° < (\phi, \Psi) < 0°$ (and the symmetric region). The majority of proposed helices have been predicted to occur in these regions.[42]

Let us take (0°, 0°) as a starting point and proceed along the path $\Psi = \phi$ toward the arrowhead. As we proceed, the helix is seen to reorganize itself as different peptide units come into proximity. The number of repeating units per turn of helix, m, takes the value 2 at (0°, 0°) and increases along the path. At (−35°, −35°) the deepest minimum is encountered in which $m \sim 3$ and there exists a strong hydrogen bond between peptide units separated by one other unit (the 1-3 bond of the 3_{10} helix). The next minimum occurs at (−50°, −60°) where $m \sim 3.6$ and the 1-4 bond characteristic of the α helix is present. The hydrogen bond energy of the latter conformation is calculated to be slightly smaller than that of the first minimum. Proceeding further along the path brings us to a region in which the helices are of very low pitch. Consequently, no minimum is found with a stabilizing 1-5 bond because the appropriate residues are in extremely close proximity when they are suitably positioned for hydrogen bonding. The third minimum at (−70°, −70°) is a result, instead, of interactions between residues on alternate turns of the helix (1-10 and 1-9 interactions). (The presence of the intervening turn of the helix would, of course, prevent any such hydrogen bond formation, which is eliminated in our algorithm when nuclear repulsions are added.) As we continue along the path the pitch of the helix begins to increase again as we switch from right- to left-handed helices. The fourth and final minimum is encountered at (−90°, −90°) in which $m \sim -5$ and corresponds to a weak 1-6 (and weaker 1-5) hydrogen bond. The depths of the four minima are seen to become progressively smaller as we proceed along the path from beginning to end.

Table III. Energies of Hexamers Relative to the Fully Extended Structure

configuration	(ϕ, Ψ)	E, kcal mol^{-1} residue^{-1} partitioned	E, kcal mol^{-1} residue^{-1} PRDDO
α	(−48°, −57°)[a]	−0.7	0.0
3_{10}	(−49°, −16°)[a]	−0.6	−0.4
π	(−45°, −70°)[a]	8.0	8.0
2_7	(−75°, 70°)[a]	−0.1	−0.3
P[c]	(−119°, 113°)[b]	5.1	5.2
AP[d]	(−142°, 145°)[b]	2.9	3.0

[a] Reference 42. [b] Reference 3. [c] Parallel chain pleated sheet. [d] Antiparallel chain pleated sheet.

The total energy E of the 20-mer of polyglycine, shown in Figure 8, is obtained by adding the "adjacent energy" to $E_{HB} + E_{rep}$. Three minima are obtained, the deepest occurring at (−48°, −57°), which is the configuration expected for an α helix.[42] Lying close to this minimum but separated from it by an energy ridge is a second minimum at (−45°, −35°), which is close to the position expected for a 3_{10} helix.[42] A third and rather shallow minimum of 0.4 kcal mol^{-1} residue^{-1} is found in the region of the 2_7 helix or ribbon structure at (−70°, +60°). This minimum is due solely to interactions between adjacent peptide units as can be seen from the fact that it is

virtually unchanged in the dipeptide conformational map (Figure 1).

IV. Comparison with Other Theoretical Results and Experiment

Several schemes have been formulated that partition the total energy of a polypeptide into physically significant sources.[5-20] To analyze the results obtained with our procedure we use the method of Scheraga et al.[14] in which contributions are identified with electrostatic, nonbonded, hydrogen bond, and torsional energy terms. A partial charge is assigned to each atom and the sum of all pairwise electrostatic interactions is obtained via Coulomb's law including an "effective dielectric constant" (assigned the value of 2 by Scheraga et al.). The nonbonded energy is similarly computed as the sum of pairwise interactions between atoms. A Lennard-Jones 6-12 potential is used with a different set of parameters for each pair of atoms. The hydrogen bond energy is calculated for each pair of atoms capable of forming hydrogen bonds, i.e., amide H and carbonyl O in polyglycine. Scheraga's "general hydrogen bond" potential is a 10-12 type function of the internuclear distance. All torsional energy terms for rotations about the ϕ and Ψ dihedral angles are assumed equal to zero.[14] The conformational energy map shown in Figure 9 was obtained when this treatment was applied to the 20-mer of polyglycine. The deepest minimum, 1.5 kcal mol^{-1} residue^{-1} more stable than the FE conformation, occurs in the α helical region at (−52°, −52°). Secondary minima are found at (−95°, −85°) and (−85°, +80°).

Crystalline polyglycine is known to exist in two different forms.[43] Polyglycine I contains a good deal of β sheet structure while three strands are wound around each other to form a triple helix in polyglycine II. The system that we have examined in this study, however, is an isolated single strand of polyglycine. There is no possibility in this structure of interchain hydrogen bonding, which is a necessary ingredient of both the β sheet and the triple helix. It is therefore not surprising that neither of the configurations of polyglycine I nor II is found at an energy minimum in our calculations.

Structures such as the α helix, on the other hand, are stabilized by hydrogen bonding between residues on the same chain. Both our procedure and that of Scheraga et al.[14] predict the α helix to be the most stable conformation for a single strand of polyglycine. The replacement of the hydrogen of glycine with the side chains of many of the amino acid residues would not be expected to drastically alter this result. For example, Scott and Scheraga[8] have found the α helical structure to be most stable for poly-L-alanine as well. The α helix has been observed in the X-ray crystallographic structures of numerous globular proteins and is by far the most commonly observed helical formation. The minimum energy conformation predicted by our method coincides with the classical α helical structure,[42] while that found by Scheraga's procedure lies several degrees away and is somewhat less stable than is ours. The dominant feature of our α helical structure is a nearly linear 1-4 hydrogen bond of length 1.84 Å. The geometry of the H bond in Scheraga's α helix is similar with a length of 1.90 Å.

Another type of helix with intrachain hydrogen bonding is the 3_{10} helix.[42] The 1-3 bond found in our 3_{10} helix is of length 1.66 Å and is strained somewhat by a value of $|\phi_1 - \pi/2| = 43°$. Short segments of the 3_{10} helix have been observed on occasion in globular proteins,[11b,44] although its occurrence is much less frequent than that of the α helix. This fact is reflected by the greater stability of the α helix (0.9 kcal mol^{-1} residue^{-1}) in our calculations. The 3_{10} helix is found by the procedure of Scheraga et al.[14] to lie in a region of very high energy. The 3_{10} structure (− 49°, −26°) is calculated to be 12.5 kcal mol^{-1} residue^{-1} higher in energy than the α helix

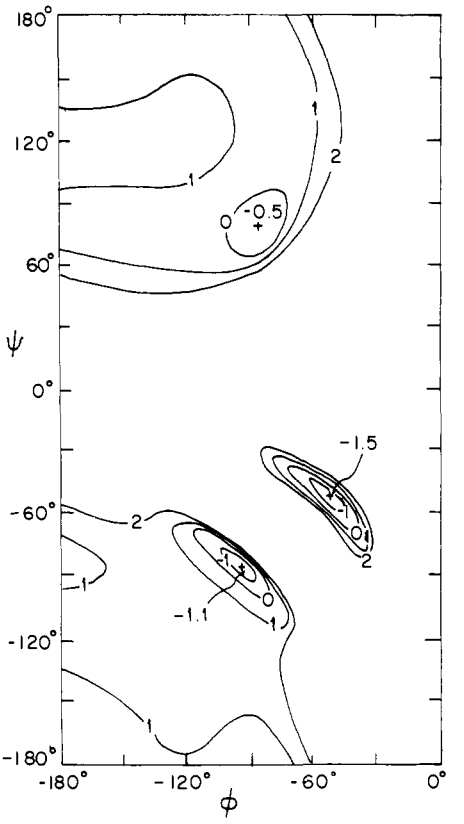


Figure 9. Conformational map of 20-mer of polyglycine obtained using the procedure of Scheraga et al. Energies (kcal mol^{-1} residue^{-1}) are shown relative to the FE structure. Contours of energy greater than 2 are not shown.

(−48°, −57°). An analysis reveals that the major portion of this energy difference (90%) stems from the calculated interactions between adjacent peptide units. By contrast, quantum mechanical calculations[3,4,23,25,29] show a much smaller difference in energy between the two conformations. At the dipeptide level, the 3_{10} conformation is found to be higher in energy than the α structure by only 0.8-1.8 kcal/mol (depending on the standard geometry used for the glycine residue) by the PRDDO method[4,29] while the former structure is computed to be *lower* in energy by 1.9 kcal/mol by the ab initio molecular fragment approach.[3] Similar results are obtained at the pentamer level, where Scheraga finds the 3_{10} structure less stable than the α conformation by 47 kcal/mol. The molecular fragment and PRDDO techniques predict the 3_{10} conformation to be *more* stable by 14.2 and 1.2-5.8 kcal/mol, respectively. We see that Scheraga's classical partitioning method yields a significantly different stability for the 3_{10} helix than either quantum mechanical results or our partitioning procedure.

The second most stable minimum found by Scheraga's method occurs instead at (−95°, −85°) and is calculated to be 0.4 kcal mol^{-1} residue^{-1} higher in energy than the absolute minimum. However, very few residues within globular proteins have been found in this region of configuration space nor have there been found any helices with such parameters.[28] Furthermore, the shortest hydrogen bond found in this structure (a 1-6 interaction) is 2.67 Å and is therefore of very limited strength. The presence of this minimum thus appears to be an artifact of the classical partitioning procedure.

The third minimum found by both methods is in the region of the 2_7 ribbon structure. Although observed for dipeptides, this configuration is uncommon in proteins.[28] The minimum found by our method at (−70°, +60°) is 2.9 kcal mol^{-1} residue^{-1} less stable than our α helical global minimum. This

Table IV. Partial Charges

atom	CNDO/2	STO-3G[a]	PRDDO[b]
C	0.450	0.317	0.220
O	−0.384	−0.298	−0.240
H	0.176	0.226	0.250
N	−0.344	−0.374	−0.310
C^{α}	−0.008	−0.018	−0.190
H^{α}	0.055	0.079	0.140

[a] From Mulliken population analysis of central portion of *N*-formylglycylamide in fully extended conformation. [b] Calculated charge from hexaglycine averaged over several conformations.

minimum is found by Scheraga's method at (−85°, +80°) and is only 1.0 kcal mol^{-1} $residue^{-1}$ higher in energy than the α helix. The two procedures find nearly identical differences in energy between the 2_7 and FE structures.

Analysis of the individual components of the classical energy partitioning scheme reveals that nonbonded interactions are principally responsible for the appearance of the three minima. Both the shape and position of the α helical minimum are determined mostly by minimizing internuclear repulsions. A contour map of repulsion energy[45] only shows this minimum virtually unchanged from that of Figure 9 except that it is 1.9 kcal mol^{-1} $residue^{-1}$ higher than the FE structure. No other minima are present. Addition of the dispersion energy[45] lowers the nonbonded energy of the α helix relative to that of the FE structure. A second minimum at (−95°, −85°) is present in the conformational map of nonbonded energy as the absolute minimum, which is 1.5 kcal mol^{-1} $residue^{-1}$ more stable than the α helix. The 2_7 minimum is present in this map also. Contributions from the electrostatic and hydrogen bond energies stabilize preferentially the α helix to yield the map of total energy shown in Figure 9.

Since the classical partitioning method relies to a great extent on empirical parameters of a rather arbitrary nature, we have investigated the sensitivity of the results to different assignments of these parameters. The partial atomic charges used in computing the electrostatic energy are assigned by Scheraga et al. via CNDO/2 calculations. As alternatives, we have reassigned these charges using both the STO-3G[46] and PRDDO techniques (Table IV). The maps obtained using all choices of charges were quite similar to Figure 9 yielding approximately the same three minima. The relative order of stability of the minima was identical for all three charge assignments and the positions and shapes of the minima were similar to those in Figure 9. As the charges were varied from CNDO/2 to STO-3G to PRDDO, the magnitudes of the energies of the three minima relative to the FE structure all decreased uniformly. The energies calculated for the PRDDO charges were ~65% those for CNDO/2 charges.

Scott and Scheraga[8] have found approximately the same three minima as those described above in a study of decaglycine using a different form of the hydrogen bond function, idealized partial atomic charges, and nonzero torsional energy terms. The classical partitioning schemes thus lead to the following general results. The absolute minimum is in the α helical region. A second minimum which is nearly equal in energy to the α helix occurs at about (−100°, −80°). This structure has not been observed experimentally and is presumably an artifact. The 2_7 ribbon structure, which is quite uncommon in proteins, is calculated to lie within 0.7–1.0 kcal mol^{-1} $residue^{-1}$ of the α helix. By contrast, our method predicts two principal minima, the α and 3_{10} helices, both of which have been observed. The ribbon structure is calculated to be considerably less stable than either principal minimum.

V. Discussion

The method outlined in this paper is based on a partitioning of the total energy of a polypeptide into contributions from interactions between pairs of residues. Unlike classical partitioning procedures, little attempt is made here to subdivide further these total interaction energies into their various physical components. The required partitioning of the interaction energy between nonadjacent residues into the two components of attractive hydrogen bonding and repulsions is an arbitrary one dependent on the particular choice of nonbonded potentials. The potentials used do appear to be consistent with quantum mechanical results and allow the attractive hydrogen bonding energy to be cast in a convenient form. Regardless of the choice of nonbonded potentials, the true quantum mechanical interaction energy is obtained when the two components are added together as in the calculation of the total energy of the polypeptide. The quantum mechanical interaction energy between amide units contains electrostatic contributions as well as the stabilization energy which arises from the direct interaction of the amide proton with a lone pair orbital of the carbonyl oxygen. The "hydrogen bond energy", E_{HB}, as formulated by our procedure, thus includes implicitly both of these contributions (as well as parts of those nonbonded components which are inadequately described by the Lennard-Jones potentials used). Similarly, torsional angle twisting terms as well as electrostatic, nonbonded, and other classical contributions are included in the "adjacent energy" term E_{adj}.

Dispersion energy, on the other hand, is not accounted for by self-consistent field calculations unless one goes beyond the Hartree-Fock level to include electron correlation. The nonbonded potentials used by most classical partitioning treatments do, on the other hand, include dispersion terms. Since this is very difficult to do in any ab initio treatment, we have investigated the effects of dispersion forces on our results by resorting to the procedure of Scheraga et al.[14,45] We find that the map represented in Figure 8 would be essentially unaffected. The stabilities of all three minima would be slightly increased as would the differences in energy between them. No additional minima would result. As described above, the minimum observed at about (−95°, −85°) via Scheraga's procedure is due principally to dispersion terms. These terms, when introduced into our formalism, are not of large enough magnitude to produce such a minimum.

Our considerations so far have been concerned only with enthalpy and have neglected the important role which entropy may play in determining preferred conformations. Whereas our results indicate that the α helical configuration is favored by considerations of enthalpy, polyglycine has not in fact been observed in such a structure in either the solid state or solution.[47] In addition, the presence of a glycyl residue in an amino acid sequence has been found to decrease the probability that the sequence will adopt an α helical structure.[48] The helix breaking nature of the glycyl residue may be accounted for by entropy considerations as follows.[49] When the polypeptide backbone takes on a helical structure, all residues, including glycine, are restricted to a narrow range of conformations. Randomization of the backbone configuration allows all residues a wider range of conformations and increases the entropy of the system. The glycyl residue, with its small side chain, enjoys a particularly wide range of allowed conformations and hence suffers the largest loss in entropy when the α helix is formed. On this basis, our calculated low enthalpy of the polyglycine α helix is consistent with the fact that no such structure has been observed experimentally.

The calculations reported in this paper were performed under several restrictions. First, all structures considered were limited to regular helices. A relaxation of this requirement would allow all angles ϕ_i and Ψ_i to be independent variables and result in a large number of degrees of freedom. (There are 38 such angles in our 20-mer.) We cannot exclude the possi-

bility that such an extension to all of configurational space might lead to additional minima. For example, a sharp turn near the middle of the chain, allowing hydrogen bonding between the two sections, could lead to an antiparallel pleated sheet arrangement. The possibility of β sheet structure might also be realized by the introduction into our calculations of additional polypeptide chains in which interchain hydrogen bonding could compete effectively with intrachain bonding of the α helix. In addition, if the polypeptide chain occurs in an aqueous environment, hydrogen bonding between the waters and the peptide might also compete with peptide–peptide binding since quantum mechanical studies find the two types of bonds to be of comparable strength.[50-52] Partially extended conformations which allow a considerable degree of solvation would consequently be stabilized relative to the more compact helical structures.

Although we have applied the PRDDO method here, *our energy partitioning procedure is formulated such that any quantum mechanical method may be substituted in its place.* Incorporation of an alternative method requires both the calculation of a conformational map analogous to Figure 1 and the formulation of a hydrogen bond potential similar to eq 2. As an example of how different procedures can affect the results, we consider the 2_7 minimum of Figure 8. Since this structure is stabilized, as we have seen, by the interactions between adjacent residues, we concern ourselves with the corresponding C_7 structure of the dipeptide. Infrared and NMR studies have found evidence that both the C_7 and fully extended (or C_5) configurations of dipeptides are present in dilute solutions of nonpolar solvents but various quantum mechanical studies yield conflicting results on the relative stabilities of the two structures.[29] While the PCILO method is in agreement with PRDDO, the ab initio STO-3G basis set finds the C_5 structure more stable.[26] Thus if the dihedral energy were computed at the STO-3G level, the 2_7 ribbon structure of Figure 8 would be found less stable than the FE conformation, suggesting that the former structure is less likely to be observed.

In order to see how the full conformational space of Figure 8 may be altered when another quantum mechanical method is employed, one must consider the different amide–amide hydrogen bond energies. The PRDDO procedure appears to slightly overestimate this interaction energy while underestimating the corresponding bond length in comparison to ab initio procedures and experiment.[32,52,53] (Further improvements in the calculations to include a more flexible basis set and correlation effects have been found to produce no significant increase in the accuracy of the calculated interaction energies.[53]) The calculated stabilities of the two principal minima of Figure 8, the α and 3_{10} helices, are highly dependent on hydrogen bonding interactions. On this basis, one might expect the use of a higher quality quantum mechanical procedure to result in a relative destabilization of these two structures. However, the change in adjacent energy must also be taken into account and it is difficult to predict the effects of this change. For example, the ab initio STO-3G and molecular fragment[3] methods yield higher adjacent energies of both the α and 3_{10} helices, relative to the FE structure, than does PRDDO. By contrast, the PCILO procedure, which employs limited configuration interaction, yields opposite results. The energy obtained by our procedure is an approximation to the total energy which would be obtained by a full SCF calculation using a minimal Slater basis set. Just as these same methods are frequently at variance for small molecules, one must expect similar differences to occur for very large systems.

The procedure described in this paper is strictly applicable only to polyglycine. An extension of the method to include amino acids other than glycyl residues would necessitate a conformational map for each additional residue. Such computations are within the range of present quantum mechanical procedures, and have, in fact, been performed previously for most residues using PCILO.[28] Nonglycyl residues such as serine and lysine lead to hydrogen bonds of a type different from the peptide–peptide bonds treated here. Each type of hydrogen bond may be dealt with in the same fashion as the amide–amide bond treated in this paper. The formulation of empirical hydrogen bond functions via quantum mechanical calculations on suitable model systems is feasible at the present time. With these extensions, our formulation is quite general and may be used to study large, complex proteins containing several hundred residues. It would be very worthwhile to examine, for example, the folding of a protein in vitro or the factors responsible for the packing observed in crystals.

Acknowledgment. We gratefully acknowledge a grant of computer time from the Instruction and Research Computer Center of The Ohio State University.

References and Notes

(1) Weizmann Postdoctoral Fellow. Address correspondence to author at the Department of Chemistry, Southern Illinois University, Carbondale, Ill. 62901.
(2) A. Pullman and B. Pullman, *Q. Rev. Biophys.*, 7, 505 (1975); L. Radom and J. A. Pople, *MTP Int. Rev. Sci.*, **1**, 71 (1972); M. J. S. Dewar, *Science*, **187**, 1037 (1975).
(3) L. L. Shipman and R. E. Christoffersen, *J. Am. Chem. Soc.*, **95**, 4733 (1973).
(4) D. A. Kleier and W. N. Lipscomb, *Int. J. Quantum Chem., Quantum Biol. Symp.*, **No. 4**, 73 (1977).
(5) P. DeSantis, E. Giglio, A. M. Liquori, and A. Ripamonti, *Nature (London)*, **206**, 456 (1965).
(6) D. Poland and H. A. Scheraga, *Biochemistry*, **6**, 3791 (1967).
(7) E. M. Popov, G. M. Lipkind, S. F. Arkhipova, and V. G. Dashevskii, *Mol. Biol. (Moscow)*, **2**, 498 (1968); E. M. Popov, V. G. Dashevskii, G. M. Lipkind, and S. F. Arkhipova, *ibid.*, **2**, 491 (1968).
(8) R. A. Scott and H. A. Scheraga, *J. Chem. Phys.*, **45**, 2091 (1966).
(9) T. Ooi, R. A. Scott, G. Vanderkooi, and H. A. Scheraga, *J. Chem. Phys.*, **46**, 4410 (1967).
(10) K. D. Gibson and H. A. Scheraga, *Proc. Natl. Acad. Sci. U.S.A.*, **58**, 420 (1967).
(11) (a) G. N. Ramachandran in "Conformation of Biological Molecules and Polymers", E. D. Bergmann and B. Pullman, Ed., Israel Academy of Sciences and Humanities, Jerusalem, Israel, 1973, p 1; (b) G. N. Ramachandran, R. Chandrasekharan, and R. Chidambaram, *Proc. Indian Acad. Sci., Sect. A*, **74**, 270, 284 (1971); (c) G. N. Ramachandran, A. V. Lakshminarayanan, R. Balasubramanian, and G. Tegoni, *Biochem. Biophys. Acta*, **221**, 165 (1970); R. Balasubramanian, R. Chidambaram, and G. N. Ramachandran, *ibid.*, **221**, 196 (1970).
(12) D. A. Brant, *Macromolecules*, **1**, 291 (1968).
(13) D. A. Brant, W. G. Miller, and P. J. Flory, *J. Mol. Biol.*, **23**, 47 (1967).
(14) F. A. Momany, R. F. McGuire, A. W. Burgess, and H. A. Scheraga, *J. Phys. Chem.*, **79**, 2361 (1975).
(15) M. Levitt, *J. Mol. Biol.*, **82**, 393 (1974); **104**, 59 (1976).
(16) C. M. Venkatachalam and S. Krimm in "Conformation of Biological Molecules and Polymers", E. D. Bergmann and B. Pullman, Ed., Israel Academy of Sciences and Humanities, Jerusalem, Israel, 1973, pp 141–154.
(17) P. K. Ponnuswamy and V. Sasisekharan, *Biopolymers*, **10**, 565 (1971).
(18) J. L. DeCoen, *J. Mol. Biol.*, **49**, 405 (1970).
(19) V. Renugopalakrishnan, S. Nir, and R. Rein in "Environmental Effects on Molecular Structure and Properties", B. Pullman, Ed., Israel Academy of Sciences and Humanities, Jerusalem, Israel, 1975, pp 109–133.
(20) H. A. Scheraga, *Adv. Phys. Org. Chem.*, **6**, 103 (1968), and references cited therein.
(21) A preliminary account has appeared: S. Scheiner and C. W. Kern, *Proc. Natl. Acad. Sci. U.S.A.*, **75**, 2071 (1978).
(22) R. Hoffmann and A. Imamura, *Biopolymers*, 7, 207 (1969).
(23) G. Govil, *J. Chem. Soc. A*, 2464 (1970).
(24) F. A. Momany, R. F. McGuire, J. F. Yan, and H. A. Scheraga, *J. Phys. Chem.*, **75**, 2286 (1971).
(25) B. Maigret, B. Pullman, and M. Dreyfus, *J. Theor. Biol.*, **26**, 321 (1970).
(26) A. Pullman and H. Berthod, *C. R. Hebd. Seances Acad. Sci., Ser. D*, **277**, 2077 (1973).
(27) L. L. Shipman and R. E. Christoffersen, *J. Am. Chem. Soc.*, **95**, 1408 (1973).
(28) B. Pullman and A. Pullman, *Adv. Protein Chem.*, **28**, 347 (1974), and references cited therein.
(29) S. Scheiner and C. W. Kern, *Biochim. Biophys. Acta*, submitted.
(30) T. A. Halgren and W. N. Lipscomb, *J. Chem. Phys.*, **58**, 1569 (1973); T. A. Halgren, D. A. Kleier, J. H. Hall, Jr., L. D. Brown, and W. N. Lipscomb, *J. Am. Chem. Soc.*, **100**, 6595 (1978).
(31) IUPAC–IUB Commission on Biochemical Nomenclature, *Biochemistry*, **9**, 3471 (1970).
(32) M. Dreyfus and A. Pullman, *Theor. Chim. Acta*, **19**, 20 (1970).
(33) R. F. McGuire, F. A. Momany, and H. A. Scheraga, *J. Phys. Chem.*, **76**, 375 (1972).

(34) G. C. Pimentel and A. L. McClellan in "The Hydrogen Bond", W. H. Freeman, San Francisco, Calif., 1960, Chapter 10.
(35) F. A. Momany, R. F. McGuire, J. F. Yan, and H. A. Scheraga, *J. Phys. Chem.*, **74**, 2424 (1970).
(36) P. Kollman and L. C. Allen, *Chem. Rev.*, **72**, 283 (1972).
(37) W. H. Stockmayer, *J. Chem. Phys.*, **9**, 398 (1941).
(38) R. Schroeder and E. R. Lippincott, *J. Phys. Chem.*, **61**, 921 (1957); E. R. Lippincott and R. Schroeder, *J. Chem. Phys.*, **23**, 1099 (1955).
(39) W. G. Moulton and R. A. Kromhaut, *J. Chem. Phys.*, **25**, 34 (1956).
(40) R. D. Singh and D. R. Ferro, *J. Phys. Chem.*, **78**, 970 (1974).
(41) H. Berthod and A. Pullman, *Chem. Phys. Lett.*, **14**, 217 (1972).
(42) R. E. Dickerson and I. Geis in "The Structure and Action of Proteins", W. A. Benjamin, Menlo Park, Calif., 1969, pp 24–43.
(43) A. Elliot in "Poly-α-amino Acids", G. D. Fasman, Ed., Marcel Dekker, New York, N.Y., 1967, pp 48–57.
(44) C. H. Carlisle, R. A. Palmer, S. K. Muzumdar, B. A. Gorinsky, and D. G. R. Yeates, *J. Mol. Biol.*, **85**, 1 (1974).
(45) The repulsion energy is defined here as the sum of all interatomic nonbonded interactions calculated to be greater than zero. Similarly, the dispersion energy is the sum of all negative interactions.
(46) R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.*, **54**, 724 (1971); W. J. Hehre, R. Ditchfield, M. D. Newton, and J. A. Pople, GAUSSIAN 70, Program No. 236, Quantum Chemistry Program Exchange, Indiana University, 1974.
(47) C. M. Venkatachalam and G. N. Ramachandran, *Annu. Rev. Biochem.*, **38**, 45 (1969).
(48) R. D. B. Fraser, B. S. Harrap, T. P. MacRae, F. H. C. Stewart, and E. Suzuki, *Biopolymers*, **5**, 251 (1967).
(49) G. Némethy, S. J. Leach, and H. A. Scheraga, *J. Phys. Chem.*, **70**, 998 (1966).
(50) S. Scheiner and C. W. Kern, *J. Am. Chem. Soc.*, **99**, 7042 (1977).
(51) J. F. Hinton and R. D. Harpool, *J. Am. Chem. Soc.*, **99**, 349 (1977).
(52) A. Johansson, P. Kollman, S. Rothenberg, and J. McKelvey, *J. Am. Chem. Soc.*, **96**, 3794 (1974).
(53) J. D. Dill, L. C. Allen, W. C. Topp, and J. A. Pople, *J. Am. Chem. Soc.*, **97**, 7220 (1975).

On the Synchronous 1,4-Addition of Methylene to *cis*-Butadiene

W. W. Schoeller* and E. Yurtsever

Contribution from the Fakultät für Chemie der Universität Bielefeld, Postfach 8640, 48 Bielefeld, West Germany. *Received May 10, 1978*

Abstract: The concerted 1,4-addition of methylene to *cis*-butadiene has been investigated by the semiempirical MINDO/3 method. The orbital symmetry allowed reaction is impeded by repulsion between the orbital σ (of the methylene) and the subfrontier MO π_1 (of the diene). An energy barrier of 28 kcal/mol is obtained. Trajectories requiring less energy (by about 5 kcal/mol) can be traced from a two-dimensional plot of the computed energy for the concerted reaction path.

Although the addition of singlet carbenes to olefins has been under active investigation for about 20 years,[1] very few 1,4-additions have been reported.[2-13] Most of them turned out to be two-step processes. Either the singlet carbene underwent relaxation to the triplet state, or a cyclopropane adduct was initially formed, and subsequently underwent a vinylcyclopropane rearrangement. Only in one case, the homo 1,4-addition of difluorocarbenes to norbornadiene, has a one-step process been established.[14]

On the basis of an orbital correlation diagram the reaction is symmetry allowed (Figure 1). Transfer of electron density can occur from (a) the HOMO π_2 to the empty p orbital of the methylene (type I interaction) and (b) the σ orbital (of the methylene) into the LUMO π_3* (type II interaction). In this respect the carbene can act as an *electrophilic* ($\pi_2 \rightarrow$ p) *and nucleophilic* ($\sigma \rightarrow \pi_3$*) species toward the diene.[15]

Results and Discussion

In order to deepen the understanding of the mechanism of the concerted 1,4-addition we have performed a theoretical study on this reaction employing the semiempirical MINDO/3 method.[18] All calculated geometries were optimized for a single Slater determinant wave function with the gradient procedure.[19]

As a model reaction the approach of methylene in its energy lowest σ^2 state[16a,20] to *cis*-butadiene was investigated.

The selection of the reaction coordinate is shown in Figure 2. For the computation of the pathway directing the 1,4 adduct, C_s symmetry had to be imposed ($\beta = 90°$). All other parameters were optimized.

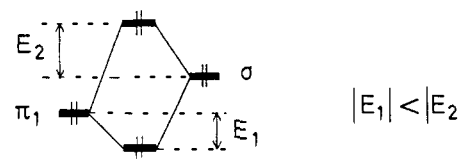
The calculated energy path as a function of the reaction coordinates R is plotted in Figure 3. A sizable energy barrier of 28 kcal/mol is predicted for the reaction path. This is in contrast to the findings on the 1,2-addition of methylene to ethylene, where no activation energy is required for the process.[16]

In view of the fact that the 1,4-addition of singlet methylene to *cis*-butadiene has been classified as a concerted process (Figure 1)[21] the magnitude of the energy barrier seems to be unexpectedly high[22] (compared with the corresponding symmetry-forbidden 1,2-addition).

A first analysis which helps to explain this anomaly is provided by an inspection of the energy hypersurface. In Figure 4 the reaction path obtained from the complete energy optimization is summarized in a series of snapshots, the methylene approaching the butadiene unit.

With decreasing values of R the methylene tends to avoid the σ approach ($R < 2.7$ Å). When R is further reduced a sudden change in the geometry of the butadiene unit takes place. The methylene groups at C_1 and C_4 in butadiene start to rotate (disrotatory). At this point of the reaction path the overlap of the π MOs of the diene with the orbitals of the methylene is maximized. Hence the concerted 1,4-addition occurs in two crucial different stages.

The effect which counterparts the σ approach with maximum overlap involves *repulsion* between the electrons in the σ orbital of the methylene and the subfrontier[23] π_1 MO of the butadiene (Figure 1).



Since these orbitals possess like symmetry they will interact. The σ orbital will be raised in energy more as the π_1 MO is

0002-7863/78/1500-7548$01.00/0