Tetrahedron Vol. 44, No 3, pp. 999 to 1006, 1988
Printed in Great Britain.

0040-4020/88 $3.00+.00
Pergamon Journals Ltd

AB INITIO CALCULATIONS FOR N-ACETYLALANYLGLYCINE AMIDE

Anne-Marie Sapse
Department of Chemistry, City University of New York,
New York, New York 10019 U.S.A.

Scott B. Daniels and Bruce W. Erickson
Department of Chemistry, University of North Carolina,
Chapel Hill, North Carolina 27599-3290 U.S.A.

(*Received in USA* 19 *August* 1987)

ABSTRACT Ab initio calculations using the STO-3G basis set were used to obtain the optimal geometry and total energy for several conformations of N-acetyl-L-alanylglycine amide. The most stable structure 1 has two C_7 rings, each of which includes a 1:3 hydrogen bond. Structure 2 having a Type-I beta turn with a 1:4 hydrogen bond is only 0.9 kcal/mol above 1. Structure 3 having a Type-II beta turn with a 1:4 hydrogen bond is 1.8 kcal/mol above 1. In agreement with experiments and molecular mechanics calculations, all three of these structures are expected to co-exist as a statistical ensemble. Structure 4 having one C_7 ring with a 1:3 hydrogen bond and structure 5′ in an extended conformation with no hydrogen bond are both 3.2 kcal/mol above 1 and thus contribute less to the ensemble. Partially optimized structures having a Type-I′, Type-II′ or Type-III beta turn are more than 5 kcal/mol above 1 and probably do not contribute to the ensemble of interconverting structures for Ac-Ala-Gly-NH_2.

Structures of small peptides have been explored experimentally by X-ray crystallography and IR, NMR and CD spectroscopy and theoretically by molecular mechanics, PCILO and ab initio calculations. Zimmerman and Scheraga[1,2] have used an empirical conformational energy program for peptides (ECEPP) to define the low-energy structures for several N-acetyl dipeptide methylamides. The ECEPP method keeps bond lengths and angles fixed and optimizes geometry only by variation of the dihedral angles. In contrast, the ab initio method used in this paper can optimize not only dihedral angles but also bond lengths and bond angles in order to reach true minima on the energy hypersurface. This method permits study of the interplay between short-range (intraresidue) forces and medium or long-range (interresidue) forces, including small changes in angles or bond lengths that can produce significant energy differences. Since the computational effort required for ab initio calculations is much larger than for the corresponding molecular mechanics calculations, only one peptide is studied at a time.

We report here results of ab initio calculations for N-acetyl-L-alanylglycine amide (Ac-Ala-Gly-NH_2) as part of a project to redesign betabellin,[3] a nonbiological protein that contains six beta turns. In our recent ab initio study[4] of N-acetyproline amide (Ac-Pro-NH_2), the minimum-energy structure has a C_7 ring containing the 1:3 hydrogen bond characteristic of a gamma turn. Specifically, the 1:3 hydrogen bond joins the carbonyl oxygen of one residue (Ac, truncated residue 1) and the amino hydrogen of the second following residue (NH_2, truncated residue 3). Since Ac-Pro-NH_2 contains only two peptide bonds, a 1:4 hydrogen bond is not possible. But Ac-Ala-Gly-NH_2 contains three peptide bonds, so it can form either a 1:4 hydrogen bond or one or two 1:3 hydrogen bonds. As noted by Venkatachalam,[5] it is not possible to form simultaneously both types of hydrogen bonds since they require different values for the backbone dihedral angles Φ ($CNC^{\alpha}C$) and ψ ($NC^{\alpha}CN$).

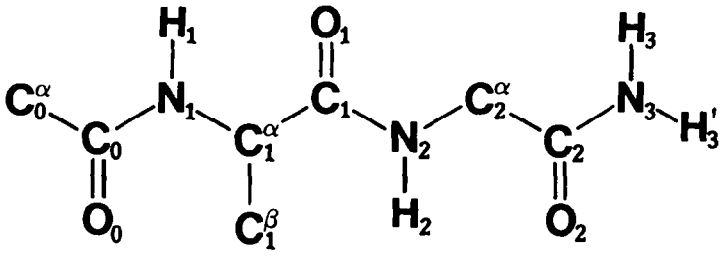

Figure 1. Atomic nomenclature for N-acetylalanylglycine amide.

METHODS

Hartree-Fock calculations were performed on N-acetylalanylglycine amide using the Gaussian-80 computer program.[6] The Roothaan equations were solved to obtain the total energy of the molecule as a function of the geometric parameters. For each atom, one Slater orbital of each kind was used, expanding it in a series of three Gaussians (STO-3G basis set[7] with standard coefficients). The optimization procedure was the Berny gradient method.[8] For each optimized geometry, the matrix of second derivatives of energy with the geometric parameters was examined. All eigenvalues were found to be positive, which showed the geometry to represent a true minimum on the energy hypersurface.

Initial geometries were chosen as follows. The bond lengths and angles were given standard values, while the dihedral angles were chosen to describe the Type I, I', II, II' and III beta turns of Venkatachalam,[5] the double C_7 ring, the single C_7 ring and the fully extended structure. A few other conformations were examined, such as a non-classic C_7 structure (featuring the standard Φ_1 and ψ_1 but different Φ_2 and ψ_2) and a few randomly chosen twisted structures lacking hydrogen bonds. Since the optimization procedure involves a search on the energy hypersurface by varying the geometric parameters in small increments, each conformation includes many others that differ slightly from the original.

The energy of each conformation was minimized within approximately 0.002 and 0.0001 Hartree by varying simultaneously all the geometric parameters of the molecule subject to the following constraints. (1) All NH bonds were kept equal except those involved in hydrogen bonds. (2) For the two methyl groups the six CH bond lengths were kept equal and the six HCC angles were kept equal. Rotation of the methyl groups was allowed to occur independently. The CH bond lengths of the Gly methylene group were kept equal and were allowed to relax independently of the CH bond length of the methyl groups. (3) Except for those involved in hydrogen bonding, all CO bond lengths were kept equal and all OCC angles were kept equal. These constraints are reasonable because the angles and the bond lengths of the two methyl groups should not differ substantially. The same is true for NH and CO bonds not involved in hydrogen bonding. These parameters do differ between conformations. (4) In structure 5, the backbone dihedral angles (Φ, ψ, ω) were kept at 180°. This constraint was used to measure the energy of a fully extended conformation, although structure 5 is not at an energy minimum. When just the dihedral angles of 5 were allowed to relax, structure 5' was obtained having slightly different values of these angles from those of structure 5.

After the energy of a structure was obtained within approximately 0.001 Hartree (0.6 kcal/mol), the lowest conformation was further optimized to within 0.000,01 Hartree (0.006 kcal/mol). Further optimization of just the dihedral angles was performed to another order of magnitude (0.000,6 kcal/mol).

Table I. Optimized Geometric Parameters of Five Structures for N-Acetylalanylglycine Amide

Parameter	1	2	3	4	5'
			Bond lengths, Angstrom		
$C_1^{\alpha}N_1$	1.476	1.472	1.468	1.478	1.462
$C_1^{\alpha}C_1$	1.571	1.564	1.568	1.572	1.561
N_1C_0	1.409	1.427	1.408	1.408	1.409
$C_0^{\alpha}C_0$	1.545	1.542	1.543	1.545	1.547
N_1H_1	1.021	1.020	1.018	1.021	1.015
$C_1^{\alpha}C_1^{\beta}$	1.556	1.548	1.544	1.555	1.555
C_1O_1	1.228	1.220	1.220	1.220	1.221
C_2O_2	1.220	1.220	1.220	1.220	1.221
C_1N_2	1.393	1.406	1.410	1.404	1.402
$N_2C_2^{\alpha}$	1.468	1.458	1.458	1.456	1.457
$C_2^{\alpha}C_2$	1.564	1.568	1.566	1.568	1.554
C_2N_3	1.426	1.426	1.420	1.430	1.401
N_3H_3	1.029	1.026	1.024	1.021	1.015
N_2H_2	1.029	1.020	1.018	1.025	1.017
C_0O_0	1.228	1.222	1.224	1.228	1.222
O_0H_2	1.65	---	---	1.69	---
O_0H_3	---	1.95	1.93	---	---
O_1H_3	1.84	---	---	---	---
HC_1^{α}	1.091	1.096	1.095	1.091	1.096
HC_2^{α}	1.092	1.092	1.092	1.092	1.095
			Bond angles, degree		
$C_1C_1^{\alpha}N_1$	113.4	111.5	109.6	113.5	105.8
$O_1C_1C_1^{\alpha}$	122.5	122.4	122.6	122.3	122.4
$O_2C_2C_2^{\alpha}$	122.5	122.4	122.5	122.3	122.4
$N_2C_1C_1^{\alpha}$	113.5	114.3	114.5	114.0	115.1
$C_2^{\alpha}N_2C_1$	122.9	122.2	122.6	121.0	122.3
$C_2C_2^{\alpha}N_2$	113.5	114.9	114.1	115.2	107.5
$C_1^{\beta}C_1^{\alpha}N_1$	111.4	109.0	109.4	111.1	111.7
$N_3C_2C_2^{\alpha}$	113.2	115.0	114.7	115.3	114.2
$C_0N_1C_1^{\alpha}$	126.9	121.3	123.1	127.2	124.3
$O_0C_0N_1$	122.9	121.9	122.2	123.3	122.7
$C_0^{\alpha}C_0O_0$	122.5	123.9	124.3	122.5	123.7
$H_2N_2C_1$	117.7	118.4	119.0	118.0	122.2
$H_3N_3C_2$	111.1	114.2	115.5	113.0	121.7
$HC_1^{\alpha}N_1$	107.3	111.6	109.5	107.3	111.0
$HC_2^{\alpha}N_2$	109.2	109.3	109.3	109.5	111.5

Table I. (continued)

Parameter	1	2	3	4	5′
			Dihedral angles,[a] degree		
$C_0{}^{\alpha}C_0N_1C_1{}^{\alpha}(\omega_0)$	159.4	-145.0	-179.1	162.1	177.8
$C_0N_1C_1{}^{\alpha}C_1(\Phi_1)$	74.7	-63.7	-65.6	72.8	-168.2
$N_1C_1{}^{\alpha}C_1N_2(\psi_1)$	-47.9	-35.6	100.9	-50.4	172.3
$C_1{}^{\alpha}C_1N_2C_2{}^{\alpha}(\omega_1)$	-182.8	-177.0	-176.5	-191.8	181.4
$C_1N_2C_2{}^{\alpha}C_2(\Phi_2)$	69.2	-97.7	115.6	97.7	182.2
$N_2C_2{}^{\alpha}C_2N_3(\psi_2)$	-78.7	-13.4	-28.4	-11.8	179.0
$C_2{}^{\alpha}C_2N_3H_3{}^{1}(\omega_2)$	161.8	163.4	165.0	155.8	
$C_1{}^{\beta}C_2{}^{\alpha}C_1O_1$	-107.7	-94.7	47.4	-107.6	
$HC_1{}^{\beta}C_1{}^{\alpha}C_1$	85.3	100.0	104.9	85.4	81.9
$HC_0{}^{\alpha}C_0O_0$	58.5	60.0	56.8	59.2	57.9

[a] For structure 5 the first seven dihedral angles listed were kept at 180°. Structure 5′ was derived from 5 by minimizing only the dihedral angles. Structures 5 and 5′ share the listed bond lengths and bond angles.

RESULTS

Eleven conformational structures of Ac-Ala-Gly-NH_2 were examined. Five were found to be relatively stable, being within 5 kcal/mol of the most stable structure. The latter are numbered 1 to 5 in order of increasing energy. Figure 1 shows the atomic nomenclature for Ac-Ala-Gly-NH_2. Table I shows the optimized geometric parameters of structures 1-5. The nomenclature used for the dihedral angles and the sign convention are described by Richardson.[9] Table II shows the total energy and relative energy for each structure. Table III shows some net atomic charges for atoms involved in hydrogen bonding.

Figure 2 shows the molecular geometry of optimized structures 1-5′. Structure 1 has two C_7 rings with one 1:3 hydrogen bond (1.65 A) between O_0 and H_2 and another (1.84 A) between O_1 and H_3. Structure 2 has a Type-I beta turn with a 1:4 hydrogen bond (1.95 A) between O_0 and H_3. Structure 3 has a Type-II beta turn with a 1:4 hydrogen bond (1.93 A) between O_0 and H_3. Structure 4 has a single C_7 ring with a 1:3 hydrogen bond (1.69 A) between O_0 and H_2. Structure 5 was constrained in the fully extended conformation with no hydrogen bonds. The less extended structure 5′ was generated from 5 by minimizing the energy as the dihedral angles were varied.

The other six structures examined were even higher in energy. They have energies relative to structure 1 of 5.5 kcal/mol for a nonclassical C_7 structure and for the structure with a Type-I′ beta turn, 7 kcal/mol for the structure with a Type-II′ beta turn, about 7.5 kcal/mol for a structure without hydrogen bonds, and greater than 10 kcal/mol for the structure with a Type-III beta turn. These numbers are rough approximations because energy optimization was only carried out to within about 0.6 to 0.06 kcal/mol of the minimum. Since Zimmerman and Scheraga[1] showed that structures higher in energy than the minimum by more than 3 kcal/mol contribute little to the partition function, further minimization of these six higher-energy structures was not done.

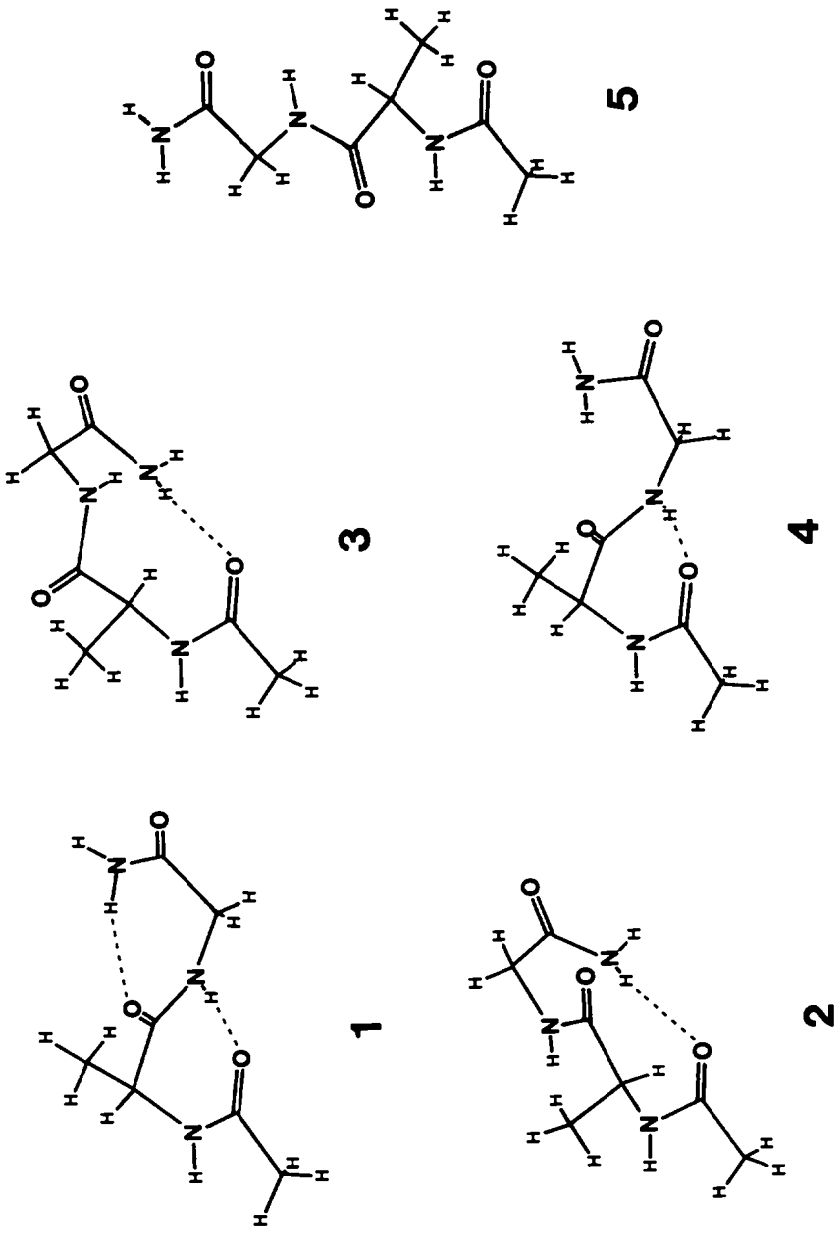

Figure 2. Molecular geometry of five structures for N-acetylalanylglycine amide.

Table II. Total and Relative Energies of Five Structures for N-Acetylalanylglycine Amide

Structure	Total energy (Hartree)	Relative energy (kcal/mol)
1	-652.1335	0.0
2	-652.1320	0.9
3	-652.1306	1.8
4	-652.1283	3.2
5	-652.1275	3.8
5′	-652.1283	3.2

DISCUSSION

Geometry. The bond lengths and angles listed in Table I are similar for the five structures of Ac-Ala-Gly-NH_2. Exceptions include the longer bond lengths of the C=O and N-H bonds adjacent to the hydrogen bonds. Typical values for bond lengths without versus with hydrogen bonding are 1.220 A versus 1.228 A for C=O and 1.020 A versus 1.029 A for N-H. This lengthening is more pronounced for bonds involved in 1:3 hydrogen bonds than those involved in 1:4 hydrogen bonds. In addition, the 1:3 hydrogen bonds between O_0 and H_2 (1.65, 1.69 A) are shorter than that between O_1 and H_3 (1.84 A), which are shorter than the 1:4 hydrogen bonds between O_0 and H_3 (1.93, 1.95 A). Careful optimization of geometry is necessary because small variations in the bond lengths and angles can produce significant energy changes relative to the energy differences between different conformations of the same peptide. For example, when the three C=O bond lengths of structure 1 were set at 1.220 A, which is typical bond length of C_2=O_2 when not hydrogen bonded, the total energy increased about 0.15 kcal/mol.

Energy. Structure 1 of Ac-Ala-Gly-NH_2 with two C_7 rings has the lowest calculated energy (ΔE) in the absence of solvent. This result agrees with the result of Zimmerman and Scheraga[2] for Ac-Ala-Gly-$NHCH_3$ using the ECEPP method. Studies[10,11] of Ac-Ala-NH_2 and our study[3] of Ac-Pro-NH_2 have also established the stability of the C_7 ring containing a 1:3 hydrogen bond.

Structure 2 with Φ and ψ angles characteristic of the Type-I beta turn and a 1:4 hydrogen bond between O_0 and H_3 is higher than 1 by only 0.9 kcal/mol. The energy difference between this Type-I structure 2 and Type-II beta-turn structure 3 is 0.9 kcal/mol. Structures 4 and 5 are higher in energy than structure 1 by more than 3 kcal/mol. The fully extended structure 5 is not at an energy minimum. When the dihedral angles of 5 were allowed to relax, it adopted structure 5′ that is less extended and 0.6 kcal/mol more stable than 5. Thus 5′ is as stable as 4, at 3.2 kcal/mol above 1. Since structures more than 3 kcal/mol above the most stable structure contribute little to the partition function,[1] structures 1-3 are expected to predominate in the ensemble of conformations adopted by Ac-Ala-Gly-NH_2 in the gas phase or a nonpolar liquid phase.

Table III. Some Net Atomic Charges (eu) for Five Structures of N-Acetylalanylglycine Amide

Atom	1	2	3	4	5
O_0	-0.290	-0.275	-0.300	-0.296	-0.289
H_2	0.241	0.202	0.193	0.232	0.213
H_3	0.214	0.207	0.209	0.181	0.198

Librational Entropy. Zimmerman and Scheraga[2] showed that librational entropy tends to stabilize nonturn structures and destabilize turn structures. For several conformations of Ac-Ala-Gly-$NHCH_3$ they found that the entropic $T\Delta S$ term contributed as much as 2 kcal/mol to the Gibbs free energy ($\Delta G = \Delta E - T\Delta S$). The larger $T\Delta S$ values were calculated for non-rigid conformations. The librational entropy of structure 4 may be higher than that of 1-3 because its C-terminal amide bond ($CONH_2$) is not involved in a hydrogen bond. Thus, structures 4 and 5 may co-exist with structures 1-3 in the statistical ensemble of structures adopted by Ac-Ala-Gly-NH_2.

Zimmerman and Scheraga[2] have calculated that a Type-II beta-turn structure for Ac-Ala-Gly-$NHCH_3$ has the lowest free energy. The Φ and ψ angles for Ala in this structure are close to the standard Type-II values but those for Gly are closer to a C_7 ring. These authors described another structure with a Type-II beta turn and a definite 1:4 hydrogen bond to have slightly higher free energy. This latter structure has Φ and ψ angles closer to those of structure 3, but those of 3 are closer to the standard Type-II values. Structure 2 has Φ and ψ angles quite close to the standard values for a Type-I beta turn. Zimmerman and Scheraga[2] reported a Type-I beta turn structure to be less stable than the Type-II structures, but its Φ and ψ angles for Gly are quite difference from the standard values because Gly adopts a C_7 ring, which reduces the librational entropy. Structure 2 should have a larger entropy and higher stability than the Type-I beta turn structure of Zimmerman and Scheraga.[2]

Solvent Effect. The equilibrium ratio of Type-I structure 2 and the Type-II structure 3 should be near one because the small calculated difference in energy (0.9 kcal/mol) may be compensated by solvation entropy effects. NMR studies[12] of cyclo-(Ala-Gly-Aca), where Aca is ϵ-aminocaproic acid, showed that the ratio of Type-I to Type-II beta-turn structures depends on the solvent. In water, for instance, Type I comprises about 35% of the total, while in dimethylsulfoxide it is less. Infrared and Raman studies[13] suggest that the conformationally flexible cyclo-(Ala-Gly-Aca) exists mainly as a mixture of hydrogen-bonded structures with small contributions from structures lacking hydrogen bonds.

Net Atomic Charges. The net atomic charges calculated for O_0, H_2 and H_3 (Table III) show that oxygen O_0 acquires a larger negative charge in the Type-II structure 3 than in the more stable Type-I structure 2. Hydrogen H_2 is more positive than H_3 when neither is hydrogen bonded, and the positive charge of H_2 increases more than that of H_3 by hydrogen bonding.

Hydrogen Bond Strength. To estimate the strength of the hydrogen bond of Type-I structure 2, an ab initio calculation was performed starting with the optimized geometry of structure 2 and setting H_3 at 180° to O_0 instead of in the $C_2N_3O_0$ plane. The energy of the resulting structure 2′ was higher than that of structure 2 by 6 kcal/mol. This value is a reasonable estimation of the strength of the hydrogen bond for the following reasons. The strength of the hydrogen bond is underestimated because hydrogen-bond energy is mostly electrostatic and STO-3G calculations underestimate charge separation. When the H_3 hydrogen is displaced from its position in structure 2, besides breaking the hydrogen bond, other local changes in this geometry of 2′ can raise the energy, which overestimates the strength of the hydrogen bond and compensates for the first effect.

Conclusion. Ac-Ala-Gly-NH_2 is a flexible molecule that can readily adopt hydrogen-bonded conformations having a Type-I or Type-II beta turn or one or two C_7 rings. Future work will explore two aspects of these results. First, the solvent effect on the stability of these structures will be investigated using a modified Born equation. Second, librational entropy will be estimated for these structures through ab initio calculation of the second-derivative matrix. Inclusion of the entropy terms will give access to the free energy of these structures, which is the proper measure of their relative stability.

ACKNOWLEDGEMENTS

This work was supported by the City University of New York through funds for computer time and a PSC grant to A.M.S., and by research contracts to B.W.E. from the Office of Naval Research and the Army Research Office. We thank Mrs. Jo Ann McPherson for excellent editorial assistance.

REFERENCES

1. Zimmerman, S.S. and Scheraga, H.A., Biopolymers 16, 811-843 (1977).
2. Zimmerman, S.S. and Scheraga, H.A., Biopolymers 17, 1849-1869 (1978).
3. Erickson, B.W., Daniels, S.B., Reddy, P.A., Unson, C.G., Richardson, J.S., and Richardson, D.C., in "Computer Graphics and Molecular Modeling," Zoller M. and Fletterick, R., Ed; Cold Spring Harbor Laboratory; Cold Spring Harbor, NY, pp 53-57 (1986).
4. Sapse, A.M., Mallah-Levy, L., Daniels, S.B., and Erickson, B.W., J. Am. Chem. Soc. 109, 3526-3529 (1987).
5. Venkatachalam, C.M., Biopolymers 6, 1425-1436 (1968).
6. Binkley, J.S., Whiteside, R.A., Krishnan, R., Seegar, R., DeFrees, D.J., Schlegel, H.B., Topiol, S., Kahn, L.R., and Pople, J.A., Gaussian 80, Carnegie-Mellon University, Pittsburg, PA (1981).
7. Hehre, W.J., Stewart, R.F., and Pople, J.A., J. Chem. Phys. 51, 2657-2664 (1969).
8. Schlegel, H.B., J. Comput. Chem. 3, 214 (1982).
9. Richardson, J.S., Advan. Prot. Chem. 34, 167-339 (1981).
10. Schafer, L., Klimkowski, V.J., Momany, F.A., and Chuman, H., Biopolymers 23, 2335-2347 (1984).
11. Beveridge, D.L., Ravishanker, G., Mezei, M., and Gedulin, B. in "Biomolecular Stereodynamics 3," Sarma, R.H. and Sarma, M.H., Ed., Adenine Press, Guilderland, N.Y. (1986).
12. Deslauriers, R., Evans, D.J., Leach, S.J., Meinwald, Y.C., Minasian, E., Nemethy, G., Rae, I.D., Scheraga, H.A., Somorjai, R.L., Stimson, E.R., Van Nispen, J.W., and Woody, R.W., Macromolecules 14, 985-996 (1981).
13. Maxfield, F.R., Bandekar, J., Krimm, S., Evans, D.J., Leach, S.J., Nemethy, G., and Scheraga, H.A., Macromolecules 14, 997-1003 (1981).