

Glycine Conformational Analysis

Ching-Han Hu, Mingzuo Shen, and Henry F. Schaefer III*

Contribution from the Center for Computational Quantum Chemistry, University of Georgia, Athens, Georgia 30602

Received September 2, 1992

Abstract: Ab initio quantum mechanical methods, including the self-consistent field (SCF), single and double excitation configuration interaction (CISD), the single and double excitation coupled cluster (CCSD), and the single, double, and perturbative triple excitation coupled cluster [CCSD(T)] have been applied to five C_s conformers and four of their C_1 counterparts on the potential energy hypersurface of glycine. A large basis set TZ2P+f designated H(5s2p1d/3s2p1d) and C,N,O(10s6p2d1f/5s3p2d1f) was chosen to evaluate the importance of d functions on hydrogen and f functions on carbon, nitrogen, and oxygen. A very subtle feature of the glycine potential energy hypersurface is the out-of-plane bending of the C_s structures II (to C_1 structure III) and IV (to C_1 structure V). Conformer I is the global minimum at all levels of theory. Two of our results are different from previous ab initio predictions: structure II/III is lower in energy than structure IV/V at our highest level of theory, and should be identified as the second minimum. Secondly, although the C_1 structure V lies slightly below its C_s counterpart IV at the DZP SCF and TZ2P SCF levels of theory, this ordering is reversed with the TZ2P+f SCF and DZP CISD methods.

Introduction

As the smallest amino acid, the glycine molecule polymerizes to form peptides, the backbone of proteins. Furthermore, glycine is of great importance for radio astronomers in the identification of interstellar amino acids, which may provide insights to the formation of large molecules in space.

The crystal structure of glycine was found in a zwitterionic form.¹ In the gas phase, the covalently bound form was found to be more stable. The different conformations of the glycine molecule result from rotating the three internal axes, i.e., the N-C, C-C, and C-O bonds seen in Figure 1. Early microwave spectroscopic investigations by Brown, Godfrey, Storey, and Bassez² indicated the presence of conformation II (See Figure 1) and led to the conclusion that conformation II is the dominant conformation in the gas phase. This was in conflict with Vishveshwara and Pople's ab initio study which led to the conclusion that conformation I is the most stable.³ At almost the same time, Suenram and Lovas⁴ also observed conformation II, but concluded that the exclusive observation of II did not allow any inference regarding its energy relative to I. Sellers and Schäfer⁵ in their 1978 ab initio study suggested that I is indeed the lowest energy conformation of glycine. Two years later the microwave spectrum of I was observed by Suenram et al.^{6,7} It is more difficult to observe structure I because it has a smaller dipole moment, while the microwave transition intensities are proportional to the square of the dipole moment. Recently Iijima, Tanaka, and Onuma reported electron diffraction results for gaseous glycine,⁸ and they concluded I is the most stable conformer of glycine in the gas phase, while the minor component is probably a mixture of II and IV. Schäfer et al.⁹ also reported a thorough theoretical examination of five conformers (I-V in Figure 1) with a series of basis sets, and two C_1 conformers III and V were

found to be more stable than their C_s counterparts II and IV, respectively. Energies obtained by Schäfer et al. using all basis sets indicated that structure V is more stable than III. Schäfer et al. thus suggested that V might exist in the glycine vapor. However, some variation in the order of relative energy of the different conformers was reported.⁹ Thus far, the correlated theoretical methods applied for studying glycine are the MP2/6-31G**/RHF/6-31G* results reported by Gordon et al.¹⁰ and the MP2/6-311G** study of conformers I-III by Schäfer et al.¹¹

In this research we used four basis sets for the computations on glycine with the self-consistent field (SCF) method. We have also applied the single and double excitation configuration interaction (CISD), the single and double excitation coupled cluster (CCSD), and the single, double, and perturbative triple excitations coupled cluster [CCSD(T)] method with one of the basis sets (DZP). The goal of this study is to obtain precise knowledge about the relative stabilities of different conformers (see Figure 1) on the energy surface, and to provide theoretical results such as rotational constants and vibrational frequencies that may be helpful to experimentalists.

Theoretical Approach

Geometries were optimized using the self-consistent-field (SCF),¹² the single and double excitations configuration interaction (CISD),¹³ and the single and double excitations coupled cluster (CCSD)¹⁴ analytic gradient methods. Harmonic vibrational frequencies were obtained via SCF analytic second derivative methods.¹⁵ For several conformers we applied the single, double, and perturbative triple excitations coupled cluster [CCSD(T)]¹⁶ method at the respective CCSD-optimized structures using the DZP basis set.

(1) Almlöf, J.; Kuick, A.; Thomas, J. O. *J. Chem. Phys.* **1973**, *59*, 3901.
 (2) Brown, R. D.; Godfrey, P. D.; Storey, J. W. V.; Bassez, M. P. *J. Chem. Soc., Chem. Commun.* **1978**, 547.
 (3) Vishveshwara, S.; Pople, J. A. *J. Am. Chem. Soc.* **1977**, *99*, 2422.
 (4) Suenram, R. D.; Lovas, F. J. *J. Mol. Spectrosc.* **1978**, *72*, 372.
 (5) Sellers, H. L.; Schäfer, L. *J. Am. Chem. Soc.* **1978**, *100*, 7728.
 (6) Suenram, R. D.; Lovas, F. J. *J. Am. Chem. Soc.* **1980**, *102*, 7180.
 (7) Schäfer, L.; Sellers, H. L.; Lovas, F. J.; Suenram, R. D. *J. Am. Chem. Soc.* **1980**, *102*, 6566.
 (8) Iijima, K.; Tanaka, K.; Onuma, S. *J. Mol. Struct.* **1991**, *246*, 257.
 (9) Ramek, M.; Cheng, V. K. W.; Frey, R. F.; Newton, S. Q.; Schäfer, L. *J. Mol. Struct.* **1991**, *245*, 1.

(10) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1991**, *113*, 3917.
 (11) Frey, R. F.; Coffin, J.; Newton, S. Q.; Ramek, M.; Cheng, V. K. W.; Momany, F. A.; Schäfer, L. *J. Am. Chem. Soc.* **1992**, *114*, 5369.
 (12) Pulay, P. *Modern Theoretical Chemistry*; Schaefer, H. F., Ed.; Plenum: New York, 1977; pp 153-185.
 (13) (a) Brooks, B. R.; Laidig, W. D.; Saxe, P.; Goddard, J. D.; Yamaguchi, Y.; Schaefer, H. F. *J. Chem. Phys.* **1980**, *76*, 4625. (b) Rice, J. E.; Amos, R. D.; Handy, T. C.; Lee, T. J.; Schaefer, H. F. *J. Chem. Phys.* **1986**, *85*, 963.
 (14) Scheiner, A. C.; Scuseria, G. E.; Lee, T. J.; Rice, J. E.; Schaefer, H. F. *J. Chem. Phys.* **1987**, *87*, 5361.
 (15) Saxe, P.; Yamaguchi, Y.; Schaefer, H. F. *J. Chem. Phys.* **1982**, *77*, 5674.
 (16) Scuseria, G. E. *Chem. Phys. Lett.* **1991**, *176*, 27.

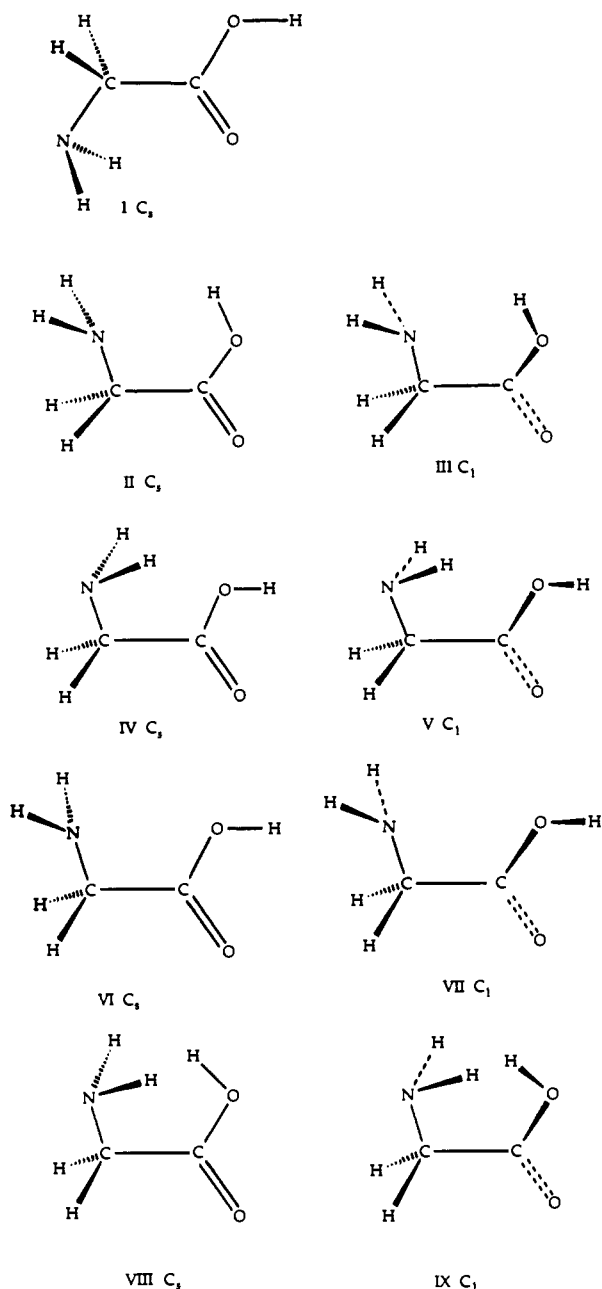


Figure 1. Glycine conformers examined in this study.

Basis sets used in this study are the double-zeta plus polarization (DZP, Huzinaga–Dunning^{17,18}) and triple-zeta plus double polarizations (TZ2P, Huzinaga–Dunning^{17,19}); the most complete basis set employed in this research, obtained by adding a set of higher order polarization functions to the TZ2P basis set, is denoted by TZ2P+f. The DZP basis set is designated as C,N,O(9s5p1d/4s2p1d), H(4s1p/2s1p). Orbital exponents for the polarization functions are $\alpha_d(\text{C}) = 0.75$, $\alpha_d(\text{N}) = 0.8$, $\alpha_d(\text{O}) = 0.85$, $\alpha_p(\text{H}) = 0.75$. The TZ2P basis set is designated as C,N,O(10s6p2d/5s3p2d), H(5s2p/3s2p), with polarization function orbital exponents $\alpha_d(\text{C}) = 1.5$, 0.375 , $\alpha_d(\text{N}) = 1.6$, 0.4 , $\alpha_d(\text{O}) = 1.7$, 0.425 , $\alpha_p(\text{H}) = 1.5$, 0.375 . The TZ2P+f basis set is designated as C,N,O(10s6p2d1f/5s3p2d1f), H(5s2p1d/3s2p1d), and includes a set of *f* functions on C, N, O and a set of *d* functions on H. Orbital exponents for these functions are $\alpha_f(\text{C}) = 0.8$, $\alpha_f(\text{N}) = 1.0$, $\alpha_f(\text{O}) = 1.4$, $\alpha_d(\text{H}) = 1.0$. Pure sets of *d* and *f* functions (i.e., five *d* functions and seven *f* functions) were used in this study.

With the DZP basis set we applied the CISD, CCSD, and CCSD(T) methods, with which five lowest occupied molecular orbitals (C, N, O 1s-like orbitals) were held doubly occupied and the five highest lying

virtual orbitals (antibonding core counterparts) were deleted in all configurations. The DZP CISD wave functions for C_2 and C_1 conformers include 323 008 and 634 501 configurations, respectively.

Computations at the TZ2P SCF and TZ2P+f SCF levels were carried out using the TURBOMOL suite of programs.²⁰ Other computations were carried out using the PSI suite of programs.²¹

Results and Discussion

Relative energies of each conformer with respect to conformer I (see Figure 1) are summarized in Table I. In Figure 1, conformers III, V, VII, and IX are the C_1 counterparts of the C_2 conformers II, IV, VI, and VIII, respectively. Hessian indices (i.e., the number of imaginary harmonic vibrational frequencies) obtained at DZP SCF level are also indicated in Table I. Six conformers are local minima on the DZP SCF energy hypersurface. Geometrical parameters of the eight structures I–VIII are summarized in Tables II–VI. Several bond distances between hydrogen (carboxyl or amino functional groups) and oxygen (carboxyl functional group) or nitrogen (amino functional group) are included in these tables.

Consistent with earlier work, conformer I is found to be the global minimum for glycine at all levels of theory employed in this research. Schäfer et al. have applied the SCF method to I–V using 13 different basis sets.⁹ They found the relative stabilities of conformers II, III and IV, V to vary with several basis sets, most of which occurred when basis sets without polarization functions were used.

Previous ab initio computations^{5,9} showed that conformer V is the second genuine local minimum in the energy hypersurface, where its C_2 counterpart IV is a saddle point; III is the third true minimum, with its C_2 counterpart II a saddle point. However, the results from our study are different from the former computational studies in two ways: first, the C_1 conformer III is more stable than conformer IV at the DZP CCSD and DZP CCSD(T) levels of theory; second, the C_2 conformer IV, rather than its C_1 counterpart V, is the local minimum in their region of configuration space.

In our SCF studies, the relative energies of conformers II and III with respect to conformer I show a slightly decreasing trend when larger basis sets are chosen. At the CISD level of theory, conformers II and III are stabilized to a larger extent than the other conformers (by about 1 kcal/mol). The recent MP2/6-311G** study by Schäfer et al.¹¹ on conformers I, II, and III previously showed that the relative energies of II and III were lowered due to the correlation energy. This effect is more significant with respect to the coupled cluster method. At the DZP CCSD and DZP CCSD(T) levels of theory, conformer III is lower in energy than conformer IV. We have also found that the C_2 conformer IV lies lower in energy than its C_1 counterpart, conformer V.

Conformer III is lower in electronic energy than conformer II at all levels of theory; however, this barrier for the interconversion of the two rotational isomers is extremely small. Experimental results for the moment of inertia linear combination ($I_a + I_b - I_c$), which is more sensitive to the geometry of the molecules than the rotational constants, apparently exclude the possibility of the nonplanar conformation,⁴ thus the potential energy surface is of a double well in this region. Our results are in good agreement with the experimental results⁶ for the relative stability of II with respect to I ($490 \pm 150 \text{ cm}^{-1}$ or $1.4 \pm 0.4 \text{ kcal/mol}$).

Conformer VII should be less abundant than I, III, and IV in the gas phase at low to moderate temperatures. Since the C_1 conformer VII is at least 3 kcal/mol more stable than its C_2 counterpart VI, a nonplanar conformation is expected. Conformers VIII (C_2) and IX (C_1) are the least favorable energetically among

(17) Huzinaga, S. *J. Chem. Phys.* **1965**, *42*, 1293.
 (18) Dunning, T. H. *J. Chem. Phys.* **1970**, *53*, 2823.
 (19) Dunning, T. H. *J. Chem. Phys.* **1980**, *72*, 5639.

(20) (a) Häser, M.; Ahlrichs, R. *J. Comput. Chem.* **1989**, *10*, 104. (b) Ahlrichs, R.; Bar, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165.

(21) PSI 1.1 1990, PSITECH Inc., Watkinsville, GA.

Table I. Theoretical Relative Energies (kcal/mol) of Structures II–IX with Respect to the Global Minimum Glycine Structure I^a

	DZP SCF ^c	TZ2P SCF	TZ2P+f SCF	DZP CISD	DZP CCSD	DZP CCSD(T)	expt ^b
I	0.000 (0i)	0.000	0.000	0.000	0.000	0.000	0.0
II	3.109 (1i)	2.975	2.908	1.901	1.431	1.060	1.4 ± 0.4
III	2.900 (0i)	2.777	2.740	1.769	1.330	1.009	
IV	1.864 (1i)	1.985	1.979	1.699	1.589	1.508	
V	1.862 (0i)	1.973	1.980	1.706			
VI	5.776 (2i)	5.463	5.338	5.804			
VII	2.777 (0i)	2.821	2.764	2.711			
VIII	8.698 (0i)						
IX	8.698 (0i)						

^a The numbers of DZP SCF imaginary vibrational frequencies are included in parentheses. The total energies of conformer I (in au) at all levels of theories are the following: DZP SCF, -282.980 360; TZ2P SCF, -282.950 040; TZ2P+f SCF, -282.959 694; DZP CISD, -283.589 729; DZP CCSD, -283.723 865; DZP CCSD(T), -283.748 007. ^b The experimental energy separation between I and II is 1.4 ± 0.4 kcal/mol according to ref 6. ^c The imaginary vibrational frequencies obtained at the DZP SCF level are the following: II, 101i cm⁻¹; IV, 20i cm⁻¹; VI, 57i cm⁻¹, and 254i cm⁻¹.

Table II. Theoretical Equilibrium Geometrical Parameters for Glycine Conformer I, with Bond Distances in Angstroms and Angles in Degrees

	DZP SCF	TZ2P SCF	TZ2P+f SCF	DZP CISD	DZP CCSD	expt ^a
N–H	1.002	0.998	0.998	1.011	1.021	1.001
N–C	1.440	1.438	1.436	1.445	1.458	1.467
C–H	1.086	1.082	1.083	1.090	1.098	1.081
C–C	1.517	1.514	1.514	1.518	1.525	1.526
C–O	1.331	1.329	1.326	1.344	1.359	1.355
O–H	0.951	0.945	0.945	0.962	0.972	0.966
C=O	1.191	1.181	1.180	1.204	1.216	1.205
H–N–H	106.4	106.5	106.8	105.3	104.7	(110.3)
H–N–C	110.5	110.6	111.2	109.5	108.8	(113.3)
N–C–C	115.3	115.5	115.4	115.2	115.4	112.1
H–C–H	106.1	106.0	106.0	106.2	106.3	(107.0)
H–C–C	107.3	107.3	107.3	107.4	107.5	
C–C–O	111.7	111.6	111.8	111.5	111.5	111.6
C–O–H	108.6	108.7	109.1	107.1	105.9	(112.3)
C–C=O	125.6	125.6	125.4	125.6	125.7	125.1
C=O···H	2.280	2.272	2.277	2.283	2.286	
C=O···H–N	2.794	2.795	2.802	2.771	2.782	

^a Reference 8; the "experimental" geometrical parameters in parentheses were assumed from low-level theoretical predictions (at the 4-21G SCF level, ref 7). See text for discussion.

Table III. Theoretical Geometrical Parameters for Glycine Conformers II and III, with Bond Distances in Angstroms and Angles in Degrees

	DZP SCF		TZ2P SCF		TZ2P+f SCF		DZP CISD		DZP CCSD	
	II	III	II	III	II	III	II	III	II	III
N–H	1.000	1.002	0.995	0.997	0.995	0.997	1.008	1.010	1.018	1.019
N–C	1.458	1.456	1.456	1.454	1.451	1.452	1.463	1.461	1.475	1.472
C–H	1.084	1.085	1.080	1.081	1.081	1.082	1.088	1.089	1.097	1.098
C–C	1.528	1.526	1.525	1.523	1.525	1.523	1.529	1.527	1.538	1.537
C–O	1.320	1.322	1.316	1.319	1.315	1.318	1.332	1.334	1.347	1.350
O–H	0.955	0.954	0.949	0.948	0.949	0.947	0.970	0.969	0.982	0.980
C=O	1.188	1.187	1.178	1.178	1.178	1.178	1.201	1.201	1.214	1.211
H–N–H	108.1	107.8	108.1	107.8	108.5	108.0	107.4	107.2	106.9	107.2
H–N–C	112.3	111.7	112.3	111.8	112.6	112.1	111.6	111.1	111.0	110.7
N–C–C	112.6	112.1	113.0	112.4	113.1	112.7	111.7	111.3	111.2	110.9
H–C–H	106.7	107.2	106.7	107.2	106.6	107.0	106.9	107.3	107.1	107.3
H–C–C	106.5	107.3	106.4	107.3	106.4	107.2	106.8	107.6	107.0	107.7
C–C–O	115.4	115.2	115.5	115.3	115.5	115.3	114.4	114.3	114.1	114.1
C–O–H	108.1	108.5	108.3	108.8	108.5	108.9	105.7	106.2	104.4	104.9
C–C=O	121.6	122.0	121.7	122.1	121.7	122.1	122.1	122.5	122.3	122.7
N–C–C=O	180.0	165.1	180.0	165.4	180.0	165.5	180.0	165.4	180.0	167.3
N–C–C–O	0.0	-16.5	0.0	-16.1	0.0	-16.1	0.0	-16.1	0.0	-14.1
C–C–O–H	0.0	2.9	0.0	3.3	0.0	3.1	0.0	3.3	0.0	3.1
O–H···N	2.024	2.053	2.039	2.063	2.063	2.073	1.940	1.972	1.912	1.944

the glycine structures considered because of the absence of any intramolecular hydrogen bonding. The local minimum has C₂ symmetry, and the geometrical parameters of conformer IX collapse to those of conformer VIII.

The geometry of the global minimum I agrees well with the experimental electron diffraction results of Iijima et al.⁸ (see Table II). Generally the theoretical bond lengths decrease with larger basis sets, and the bond lengths increase when the DZP

CISD and DZP CCSD methods are used. Notice that (see Table II) four of the experimental geometrical parameters (in parentheses) are taken from the 4-21G SCF predictions by Schäfer et al.⁷ Some of these values, namely, the H–N–H, H–N–C, and C–O–H angles, differ from our theoretical predictions by about 5°. As we can see from Table II, these angles decrease with the correlated levels of theory (CISD and CCSD); for this reason the assumed angles are probably too big. Since these angles were

Table IV. Theoretical Equilibrium Geometrical Parameters for Glycine Conformers IV and V, with Bond Distances in Angstroms and Angles in Degrees

	DZP SCF		TZ2P SCF		TZ2P+f SCF		DZP CISD		DZP CCSD
	IV	V	IV	V	IV	V	IV	V	IV
N-H	1.002	1.001	0.997	0.997	0.996	0.997	1.010	1.010	1.019
	1.002	1.002	0.997	0.997	0.996	0.997	1.010	1.010	1.019
N-C	1.442	1.442	1.440	1.440	1.437	1.440	1.447	1.447	1.458
C-H	1.085	1.087	1.081	1.083	1.082	1.083	1.090	1.091	1.098
	1.085	1.084	1.081	1.081	1.082	1.080	1.090	1.088	1.098
C-C	1.521	1.521	1.519	1.519	1.519	1.519	1.521	1.520	1.529
C-O	1.331	1.331	1.327	1.328	1.325	1.327	1.345	1.345	1.361
O-H	0.951	0.951	0.945	0.945	0.945	0.945	0.963	0.963	0.972
C=O	1.191	1.191	1.182	1.181	1.182	1.181	1.204	1.204	1.216
H-N-H	107.2	107.2	107.3	107.3	107.7	107.2	106.1	106.1	105.7
H-N-C	111.3	111.3	111.4	111.4	112.0	111.5	110.3	110.2	109.9
	111.3	111.3	111.4	111.4	112.0	111.5	110.3	110.3	109.9
N-C-C	108.3	118.2	118.6	118.4	118.8	118.5	118.3	118.2	118.4
H-C-H	106.0	106.1	106.0	106.2	105.9	106.2	106.0	106.1	106.2
H-C-C	105.9	105.8	105.9	105.7	105.8	105.6	106.1	105.9	106.2
	105.9	106.2	105.9	106.2	105.8	105.2	106.1	106.3	106.2
C-C-O	114.3	114.1	114.5	114.1	114.5	114.1	113.6	113.4	113.2
C-O-H	108.1	108.1	108.4	108.3	108.4	108.4	106.5	106.6	105.6
C-C=O	123.3	123.5	122.9	123.4	122.9	123.4	123.9	124.0	124.2
N-C-C=O	180.0	169.9	180.0	168.6	180.0	168.5	180.0	168.9	180.0
N-C-C-O	0.0	-12.3	0.0	-12.6	0.0	-12.6	0.0	-12.1	0.0
C-C-O-H	180.0	-178.6	180.0	-178.6	180.0	-178.6	180.0	-178.6	180.0
C=O...H	2.266	2.267	2.257	2.260	2.261	2.261	2.269	2.269	2.278
N-H...O-C	2.705	2.621	2.712	2.630	2.730	2.632	2.667	2.582	2.669
	2.705	2.810	2.712	2.814	2.730	2.817	2.667	2.774	2.669

Table V. Theoretical Equilibrium Geometrical Parameters for Glycine Conformer VI and VII, with Bond Distances in Angstroms and Angles in Degrees

	DZP SCF		TZ2P SCF		TZ2P+f SCF		DZP CISD	
	VI	VII	VI	VII	VI	VII	VI	VII
N-H	0.998	1.002	0.994	0.997	0.994	0.997	1.006	1.010
	0.998	1.001	0.994	0.997	0.994	0.997	1.006	1.009
N-C	1.444	1.451	1.444	1.449	1.442	1.447	1.450	1.456
C-H	1.086	1.093	1.082	1.089	1.083	1.091	1.090	1.097
	1.086	1.082	1.082	1.097	1.083	1.078	1.090	1.086
C-C	1.520	1.511	1.519	1.509	1.518	1.509	1.520	1.510
C-O	1.322	1.332	1.319	1.319	1.317	1.327	1.335	1.345
O-H	0.951	0.951	0.945	0.945	0.945	0.945	0.963	0.962
C=O	1.194	1.190	1.185	1.181	1.184	1.180	1.207	1.203
H-N-H	109.1	108.2	108.7	107.8	108.9	108.2	108.1	107.3
H-N-C	112.7	111.2	112.4	111.1	112.8	111.5	111.7	110.1
	112.7	110.3	112.4	110.1	112.8	110.6	111.7	109.6
N-C-C	115.9	111.7	116.0	112.1	116.1	112.1	115.5	111.3
H-C-H	105.6	107.8	105.8	107.7	105.6	107.6	105.6	107.8
H-C-C	105.0	106.0	105.0	105.9	105.0	105.6	105.0	106.1
	105.0	106.7	105.0	106.7	105.0	106.7	105.0	106.7
C-C-O	115.8	112.6	115.7	112.6	115.6	112.5	115.2	112.0
C-O-H	107.5	108.4	107.7	108.5	108.0	108.8	105.9	106.9
C-C=O	121.6	124.8	121.6	124.7	121.7	124.6	122.0	125.3
N-C-C=O	180.0	139.8	180.0	139.8	180.0	139.8	180.0	139.1
N-C-C-O	0.0	-43.0	0.0	-42.9	0.0	-42.9	0.0	-42.9
C-C-O-H	180.0	-179.0	180.0	-178.9	180.0	-178.9	180.0	-179.0
C=O...H	2.254	2.273	2.249	2.266	2.252	2.273	2.255	2.277

used in obtaining the geometrical parameters in the least-squares analysis for the electron diffraction data, there might exist some uncertainty in the experimental results.

For conformers II (C_2) and III (C_1) (see Table III), the predicted bond distances and bond angles differ slightly among the various theoretical methods. The carboxyl functional group is distorted by about 15° from the C_2 plane of structure II, with the carboxyl functional group approximately planar. Increasing the number of polarization functions does not change the geometrical parameters significantly. An interesting feature in Table III is that, with the correlated levels of theory, the geometrical parameters of the intramolecularly bound five-membered ring (see Figure 1) change more significantly than the other parameters. The N-C, C-C, C-O, and O-H bond distances increase

Table VI. Theoretical Equilibrium Geometrical Parameters for Glycine Conformer VIII, with Bond Distances in Angstroms and Angles in Degrees

	DZP SCF	DZP SCF	
N-H	0.999	H-N-C	115.7
N-C	1.442	N-C-C	119.0
C-H	1.084	H-C-H	106.2
C-C	1.531	H-C-C	105.9
C-O	1.331	C-C-O	117.2
O-H	0.946	C-O-H	110.7
C=O	1.186	C-C=O	121.5
H-N-H	112.3		

by about 0.02 Å in average, the N-C-C and C-C-O bond angles decrease by about 1° , and the C-O-H bond angle decreases by

Table VII. Theoretical Rotational Constants (in MHz) and Theoretical Dipole Moments (in Debye) for Structure I Obtained from Optimized Equilibrium Geometries Compared with the Microwave Experimental Data^a

	DZP SCF	TZ2P SCF	TZ2P+f SCF	DZP CISD	DZP CCSD	expt ^b
A_c	10605 (1.07)	10712 (0.95)	10736 (0.95)	10417 (1.15)	10229 (1.14)	10342 (1.0)
B_c	3917 (0.80)	3927 (0.85)	3930 (0.87)	3904 (0.61)	3841 (0.48)	3876
C_c	2955 (0.00)	2968 (0.00)	2972 (0.00)	2933 (0.00)	2884 (0.00)	2912
dipole moment	1.34	1.27	1.29	1.31	1.24	~1.00

^a The dipole moments along the principal axes are included in parentheses. ^b Reference 6.

Table VIII. Theoretical Rotational Constants (in MHz) and Theoretical Dipole Moments (in Debye) for Structures II and III Obtained from Optimized Equilibrium Geometries, Compared with the Microwave Experimental Data^a

	DZP SCF	TZ2P SCF	TZ2P+f SCF	DZP CISD	DZP CCSD	expt
Conformer II (C_s)						
A_c	10297 (5.94)	10364 (5.72)	10366 (5.52)	10232 (5.82)	10069 (5.69)	10129 (4.5) ^b 10130 (≥ 3) ^c
B_c	4092 (0.71)	4104 (0.58)	4104 (0.63)	4085 (0.93)	4036 (1.06)	4071 (≤ 1) ^b 4071 ^c
C_c	3028 (0.00)	3040 (0.00)	3040 (0.00)	3019 (0.00)	2980 (0.00)	3007 (0.0) ^b 3008 (0.0) ^c
dipole moment	5.98	5.75	5.56	5.89	5.79	4.5–4.6 ^b ≥ 3 ^c
Conformer III (C_1)						
A_c	10196 (5.76)	10271 (5.54)	10275 (5.54)	10126 (5.63)	10015 (5.53)	
B_c	4110 (0.64)	4125 (0.53)	4118 (0.49)	4101 (0.86)	4050 (0.93)	
C_c	3054 (0.05)	3066 (0.03)	3063 (0.04)	3043 (0.01)	3001 (0.07)	
dipole moment	5.79	5.57	5.56	5.70	5.61	

^a The dipole moments along the principal axes are included in parentheses. ^b Reference 2. ^c Reference 4.

Table IX. Theoretical Rotational Constants (in MHz) and Theoretical Dipole Moments (in Debye) for Structures IV and V Obtained from Optimized Equilibrium Geometries^a

	DZP SCF	TZ2P SCF	TZ2P+f SCF	DZP CISD	DZP CCSD
Conformer IV					
A_c	10179 (0.58)	10242 (0.68)	10249 (0.72)	10070 (0.38)	9906 (0.27)
B_c	4038 (2.02)	4054 (1.96)	4052 (1.98)	4019 (1.94)	3960 (1.85)
C_c	2988 (0.00)	3001 (0.00)	3000 (0.00)	2968 (0.00)	2923 (0.00)
dipole moment	2.10	2.08	2.11	1.98	1.87
Conformer V					
A_c	10140 (0.56)	10201 (0.66)	10209 (0.82)	10029 (0.36)	
B_c	4036 (2.01)	4052 (1.95)	4051 (1.77)	4017 (1.93)	
C_c	2995 (0.04)	3009 (0.03)	3009 (0.12)	2975 (0.05)	
dipole moment	2.08	2.06	1.95	1.96	

^a The dipole moments along the principal axes are included in parentheses.

about 4° at the DZP CISD and DZP CCSD levels. This effect is most clearly shown by the last parameter in Table III, where the OH...N bond distance decreases by about 0.11 Å at the DZP CCSD level, and the distance 1.912 Å is rather short for two nonbonded atoms. As we have discussed above, correlated levels of theory stabilize conformers II and III more significantly (by about 1 kcal/mol), and the five-membered ring structures of II and III should be considered as an important intramolecular hydrogen bonding in glycine.

Conformer V (C_1) differs from conformer IV (C_s) by distorting its carboxyl functional group about 12° from the C_s symmetry plane of conformer IV. The energy surface along the heavy atom (N–C–O) framework torsion is very flat, and the relative energies of structures IV and V at all levels of theory are almost identical. Conformer IV is slightly lower in energy at the TZ2P+f SCF and DZP CISD levels of theory. In Table IV we see that at the DZP CISD and DZP CCSD levels of theory the intramolecular hydrogen bonding distances are shorter than those obtained at the SCF level, though the effect is not as significant as for conformers II and III.

The geometrical parameters of conformer VII (C_1) are very different from those of conformer VI (about 40° off the C_s symmetry plane of VI; see Table V), as listed in Table I. The difference between the relative energies of conformers VI and VII is also the largest among the different (C_s , C_1) pairs.

The theoretical equilibrium rotational constants and dipole moments of the different conformers are summarized in Tables

Table X. Theoretical Rotational Constants (in MHz) and Theoretical Dipole Moments (in Debye) for Structures VI and VII Obtained from Optimized Equilibrium Geometries^a

	DZP SCF	TZ2P SCF	TZ2P+f SCF	DZP CISD
Conformer VI (C_s)				
A_c	10276 (2.42)	10343 (2.38)	10355 (2.37)	10153 (2.28)
B_c	3959 (0.03)	3975 (0.03)	3978 (0.06)	3941 (0.17)
C_c	2953 (0.00)	2976 (0.00)	2969 (0.00)	2934 (0.00)
dipole moment	2.42	2.38	2.38	2.28
Conformer VII (C_1)				
A_c	9872 (1.90)	9943 (1.86)	9945 (1.87)	9742 (1.78)
B_c	4093 (1.21)	4105 (1.22)	4115 (1.22)	4082 (1.15)
C_c	3090 (1.66)	3101 (1.61)	3104 (1.57)	3074 (1.68)
dipole moment	2.80	2.74	2.73	2.70

^a The dipole moments along the principal axes are included in parentheses.

VII–XI At the SCF level of theory, the equilibrium rotational constants increase as the number of basis functions increases. At the CISD and CCSD levels of theory, the rotational constants are lower than those obtained at the SCF level. The theoretical results are generally in good agreement with experiment.^{2,4,6}

The high molecular weight and low volatility of glycine has caused difficulties in obtaining and assigning the transitions in the microwave spectrum. Even though conformer I is the global minimum of glycine in the gas phase, it was observed later than

Table XI. Theoretical Rotational Constants (in MHz) and Theoretical Dipole Moments (in Debye) for Structure VIII Obtained from Optimized Equilibrium Geometries^a

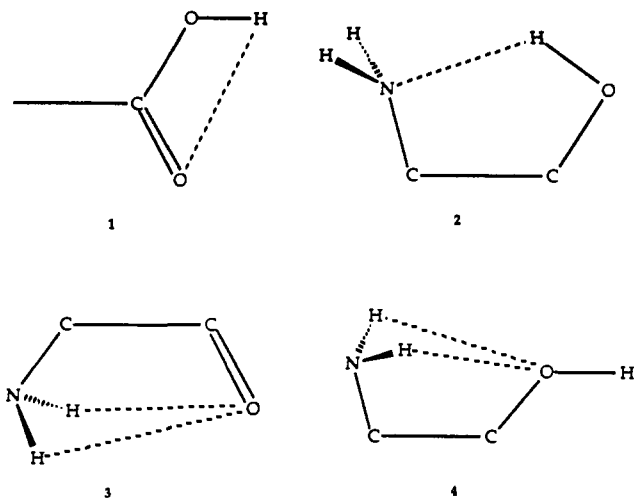
	DZP SCF		DZP SCF
<i>A</i>	10044 (4.30)	<i>C</i>	2969 (0.00)
<i>B</i>	4020 (1.60)	dipole moment	4.59

^a The dipole moments along the principal axes are included in parentheses.

conformer II, because of structure I's much smaller dipole moment.⁶ The dipole moment of conformer IV (~2 D; see Table IX) is somewhat larger than that of conformer I, and the former is probably of comparable abundance to conformer II in the gas phase. Experiments have not been able to distinguish IV from II so far; hopefully spectrometers with greatly improved S/N ratios will advance the opportunities for observation of conformers IV and VII.²²

Harmonic vibrational frequencies and IR intensities for all minima were computed at the DZP SCF level of theory and are listed in Table XII. Theoretical harmonic vibrational frequencies for the other stationary points are not explicitly listed.

The conformation analysis of glycine in the gas phase is of great interest since it generate different conformers by forming the cis carboxylic functional group (1), and glycine displays three different types of intramolecular hydrogen bonding (2–4). The harmonic vibrational frequencies in Table XII enable us to understand the role (importance) of these different types of hydrogen bonding. Here type 1 is a cis carboxyl functional group.



Although we are hesitant to impute hydrogen-bonding properties to the cis carboxyl functional group, our DZP SCF results for formic acid predict that *cis*-formic acid is 5.7 kcal/mol lower in energy than *trans*-formic acid. Type 2 involves hydrogen bonding between the nitrogen lone pair and the carboxylic hydrogen resulting in a five-membered ring. Type 3 and type 4 incorporate bifurcated arrangements between the hydrogen atoms on the amino group and the oxygen atom on the carbonyl and hydroxyl groups, respectively.

The highest five harmonic vibrational frequencies for all conformers in Table XII are pure stretching modes: ω_1 , O–H stretch; ω_2 , asymmetric N–H stretch; ω_3 , symmetric N–H stretch; ω_4 , asymmetric C–H stretch; and ω_5 , symmetric C–H stretch. These pure stretching modes provide useful information about the intramolecular hydrogen bonding in glycine. For comparison we predicted the harmonic vibrational frequencies of both *cis*- and *trans*-formic acid (see Table XIII). In Table XIII, ω_1 is the

pure O–H stretching mode, and ω_2 the pure C–H stretching mode. Notice that the *cis* O–H mode has a lower fundamental frequency than that in *trans*-formic acid. In Table XII we see that the O–H frequencies (conformers I, V, and VII) of the carboxyl group are very close, and are lower than that of conformer IX. This is similar to what happens in the isolated formic acid molecule. The decrease in O–H frequency of conformer III, which is caused by type 2 hydrogen bonding, is much more significant. By comparing the energy difference between the *cis*- and *trans*-formic acid conformations (5.7 kcal/mol), and noting that structure III lies 2.9 kcal/mol above I at the DZP SCF level (structure III sacrifices type 1 and type 3 bonding), type 2 bonding must be an important pattern among the four, at least as important as type 1. Comparing the relative stabilities of structures III (type 2) and VII (type 1) with IX, which could be considered as a nonhydrogen bonded conformation, the stabilizing effects of types 1 and 2 are similar. At the correlated level of theory (DZP CISD and DZP CCSD), type 2 is relatively more effective.

The N–H frequencies ω_2 and ω_3 of conformers I and V are lower than those for conformer VII, and are of similar magnitude. Comparing ω_2 and ω_3 of structures I and V to conformer VII, one can hardly tell which of the bifurcated hydrogen bonding arrangements is more important, even though the harmonic N–H vibrational frequencies of I are slightly lower than those of conformer V. Noticing that the stabilities of these conformers differ by type 3 bonding (conformers I and VII, ~2.7 kcal/mol), and by only type 4 (conformers V and VII, ~1.0 kcal/mol), type 3 and type 4 bonding appear of lesser importance than type 1 and type 2. Type 3 hydrogen bonding might be slightly more important than type 4.

The relative importance of bonding pattern types 1 and 2 over types 3 and 4 suggests that the former two types of bonding should be favored in the conformations of amino acids in the gas phase. In the case of alanine, the conformer involving type 2 bonding was reported to be more stable than that involving type 1 bonding from earlier theoretical studies.^{27,28} For serine, previous theoretical studies^{27,29} of the two conformers displaying types 1 and 2 bonding showed that the relative stability of the two conformers are very close; moreover, these two studies^{27,29} were in contradiction with each other. We would expect that theoretical methods at correlated levels will clarify this problem, and that the conformer of serine with type 2 bonding will be favored by the inclusion of such higher level electronic effects of the type explicitly considered in this research.

The preference of type 3 bonding over type 4 bonding is not decisive, as discussed above. However, we can derive two features from this result: first, the amino functional group is a poor hydrogen donor and a good hydrogen acceptor; second, the carbonyl oxygen is a better hydrogen acceptor than the carboxyl oxygen. These features might affect the structures of peptides in both solid state and solution, where the intramolecular and intermolecular hydrogen bonds are important.³⁰

Finally, the bond distances between hydrogen atoms and their nonbonded electronegative heavy atoms (nitrogen and oxygen) summarized in Tables II–VI facilitate a systematic understanding of the appropriateness of the different theoretical approaches in studying molecular systems containing local structures 1–4, three of which are attributed to hydrogen bonding. The O...H distances of type 1 showed a consistent trend with all levels of theory applied here for conformers I, IV, V, VI, and VII; the bond distances are shorter at the TZ2P SCF level than those obtained at the DZP

(24) Bertie, J. E.; Michaelian, K. H. *J. Chem. Phys.* **1982**, *76*, 886.

(25) Millikan, R. C.; Pitzer, K. S. *J. Chem. Phys.* **1957**, *27*, 1305.

(26) Hisatsune, I. C.; Heicklen, J. *Can. J. Spectrosc.* **1973**, *18*, 135.

(27) Sellers, H. L.; Schäfer, L. *Chem. Phys. Lett.* **1979**, *63*, 609.

(28) Van Alsenoy, C.; Scarsdale, J. N.; Sellers, H. L.; Schäfer, L. *Chem. Phys. Lett.* **1981**, *80*, 124.

(29) Masamura, M. *J. Mol. Struct.* **1987**, *152*, 293.

(30) Di Blasio, B.; Rossi, F.; Benedetti, E.; Pavone, V.; Saviano, M.; Pedone, C.; Zanotti, G.; Tancredi, T. *J. Am. Chem. Soc.* **1992**, *114*, 8277.

(22) Brown, R. D.; Crofts, J. G.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. *J. Mol. Struct.* **1988**, *190*, 185.

(23) Chang, Y.-T.; Yamaguchi, Y.; Miller, W. H.; Schaefer, H. F. *J. Am. Chem. Soc.* **1987**, *109*, 7245.

Table XII. Harmonic Vibrational Frequencies and IR Intensities (in parentheses) for the Five Local Minima of Glycine^a

mode	I	III	V	VII	IX
ω_1	3713 (124.4)	3638 (225.4)	3712 (129.4)	3715 (132.8)	3762 (111.4)
ω_2	3457 (8.3)	3473 (17.6)	3467 (8.3)	3472 (9.1)	3509 (17.1)
ω_3	3385 (4.5)	3394 (9.6)	3390 (3.8)	3391 (3.7)	3407 (7.5)
ω_4	2945 (17.8)	2955 (18.8)	2959 (15.1)	2973 (18.3)	2968 (11.4)
ω_5	2905 (24.6)	2909 (21.6)	2910 (22.1)	2839 (61.7)	2923 (16.6)
ω_6	1818 (440.5)	1842 (461.8)	1814 (435.9)	1824 (462.0)	1842 (408.4)
ω_7	1633 (23.0)	1623 (48.5)	1630 (43.1)	1615 (43.6)	1609 (38.2)
ω_8	1428 (11.0)	1438 (4.6)	1427 (2.1)	1458 (3.2)	1421 (4.8)
ω_9	1409 (46.4)	1395 (402.0)	1363 (87.6)	1421 (82.9)	1346 (286.2)
ω_{10}	1347 (0.0)	1352 (140.5)	1346 (4.6)	1338 (72.2)	1327 (1.0)
ω_{11}	1289 (23.6)	1296 (14.4)	1331 (49.0)	1269 (8.6)	1327 (215.8)
ω_{12}	1171 (286.6)	1213 (2.3)	1197 (211.9)	1220 (125.5)	1167 (6.5)
ω_{13}	1157 (1.9)	1147 (3.6)	1157 (29.7)	1128 (132.6)	1136 (2.2)
ω_{14}	1107 (57.3)	1055 (17.0)	1103 (17.5)	1050 (22.4)	1089 (28.8)
ω_{15}	923 (163.2)	944 (86.5)	900 (87.9)	1010 (11.9)	882 (9.2)
ω_{16}	902 (6.5)	873 (92.2)	891 (99.0)	834 (72.9)	836 (42.4)
ω_{17}	816 (87.5)	823 (23.1)	797 (83.0)	816 (120.1)	646 (131.9)
ω_{18}	626 (13.4)	705 (141.1)	653 (127.5)	695 (93.4)	566 (5.5)
ω_{19}	622 (127.7)	635 (13.7)	587 (74.3)	572 (87.8)	515 (42.9)
ω_{20}	493 (45.0)	548 (16.5)	511 (35.3)	531 (49.6)	507 (109.1)
ω_{21}	454 (32.2)	494 (6.1)	480 (18.2)	526 (6.4)	491 (71.0)
ω_{22}	252 (10.8)	298 (16.8)	262 (50.3)	280 (26.5)	332 (123.2)
ω_{23}	226 (55.1)	246 (17.2)	259 (7.0)	231 (31.5)	276 (19.9)
ω_{24}	72 (5.7)	87 (5.0)	16 (0.1)	68 (2.2)	62 (6.1)

^a Frequencies (scaled by a factor of 0.9) are in cm^{-1} , and IR intensities in km/mol from the DZP SCF method.

Table XIII. Harmonic Vibrational Frequencies and IR Intensities (in Parentheses) for *cis*- and *trans*-Formic Acid^a

mode	<i>cis</i> ^b	<i>trans</i>	expt ^c
A'			
ω_1	3705 (117.0)	3764 (121.8)	3569 ^d
ω_2	2965 (54.9)	2887 (89.4)	2942 ^d
ω_3	1816 (537.5)	1856 (455.2)	1777 ^d
ω_4	1384 (10.5)	1404 (6.0)	1381 ^d
ω_5	1284 (26.2)	1268 (434.3)	1223 ^e
ω_6	1142 (301.8)	1107 (27.5)	1104 ^d
ω_7	622 (61.3)	455 (113.0)	625 ^d
A''			
ω_8	1066 (0.3)	1049 (3.1)	1033 ^{e,f}
ω_9	626 (201.4)	651 (11.5)	642 ^d

^a Frequencies (scaled by a factor of 0.9) in cm^{-1} , IR intensities in km/mol with the DZP SCF method. ω_1 mode is pure O-H stretching; for assignments of the other harmonic frequencies, see ref 23. ^b The DZP SCF total energy for formic acid (*cis*) is $-188.813\,655\text{ au}$; the *trans* form is 5.678 kcal/mol higher in energy. ^c For *cis*-formic acid, note that the theoretical predictions are harmonic frequencies, while the experimental values are the observed (anharmonic) fundamentals. ^d Reference 24. ^e Reference 25. ^f Reference 26.

SCF level of theory; the distances obtained at TZ2P+f SCF level are longer than those obtained with TZ2P SCF but are shorter than those obtained at the DZP SCF level. At the DZP CISD level of theory, the O...H distances are longer than those obtained at the DZP SCF level. The N...H distances in conformers II and III (type 2) lengthen with respect to the increase of basis functions at the SCF method, and are shortened significantly (by 0.11 \AA , compared to the DZP SCF and DZP CCSD N...H distances of conformer II) with correlated levels of theory. Correlated levels of theories (CISD and CCSD) are most important for type 2 hydrogen bond structures. The O...H distances in the type 3 structure (conformer I) and type 4 structure (conformer IV and V) showed similar trends; distances increase with the expansion of the basis set and shorten with correlated levels of theory, though this effect is not as significant as for hydrogen bond type 2. The journey from DZP to TZ2P+f basis sets within the SCF method does not improve the theoretical predictions of structures, since the major effect of including more basis functions is to shorten

the distances between bound atoms in glycine. The use of correlated levels of theory in the interpretation of the properties of glycine or other molecules with intramolecular bonding is thus essential for truly quantitative predictions.

Concluding Remarks

The relative stabilities of nine conformers of the glycine molecule were examined with large basis sets and with correlated levels of quantum mechanical theory. Conformer I is the global minimum at all levels of theory, lying $\sim 1.0\text{ kcal/mol}$ below structure II/III. Conformer II/III was found to lie $\sim 0.5\text{ kcal/mol}$ lower in energy than conformer IV at the DZP CCSD and DZP CCSD(T) levels of theory; thus conformer II/III should be identified as the second lowest lying structure of glycine. Structure IV/V may be nearly as abundant as II/III in the gas phase. The theoretical rotational constants may be helpful to experimentalists in identifying the presently unknown conformer IV/V.

The energy surface near structure II/III is a shallow double well, while the barrier is probably too small for the C_1 conformer (III) to be distinguished experimentally. The C_s conformer IV was found to be a local minimum at the TZ2P+f SCF and DZP CISD levels of theory. Conformer VII lies about 1 kcal/mol above II/III and IV/V; thus it should be less abundant in the gas phase.

We have summarized the theoretical equilibrium rotational constants and dipole moments in Tables VII–XI. The third minimum, C_s conformer IV, hitherto not identified in the laboratory, possesses a dipole moment of about 2 D and may be observable via microwave spectroscopy.

Correlation effects included in the DZP CISD, DZP CCSD, and DZP CCSD(T) methods stabilize structure II/III significantly more than the other structures. The intramolecular hydrogen bonding associated with conformers II and III plays an important role in this regard. It is essential that correlated levels of theory (such as the CCSD(T) method) be used in interpreting the properties of glycine.

Acknowledgment. This research was supported by the U.S. National Science Foundation, Grant CHE-8718469. We thank Professor Richard J. Saykally for many helpful discussions.