The 1 : 1 Glycine Zwitterion–Water Complex: An *Ab Initio* Electronic Structure Study

YANBO DING and KARSTEN KROGH-JESPERSEN*

Department of Chemistry, Rutgers, The State University of New Jersey, New Brunswick, New Jersey, 08903

ABSTRACT

More than a dozen stationary points on the potential energy surface for the 1 : 1 glycine zwitterion–water complex have been investigated at Hartree-Fock or MP2 levels of theory with basis sets ranging from split valence (4-31G) to split valence plus polarization and diffuse function (6-31 + + G**) quality. Only one true minimum (GLYZWM, C_1 symmetry) could be located on the potential energy surface. GLYZWM features a bridged water molecule acting as both a hydrogen bond acceptor and donor with the NH_3^+ and CO_2^- units of the glycine zwitterion. The total hydrogen bond energy in GLYZWM is computed as 16 kcal/mol (MP2/6-31 + + G**//6-31 + + G**, including corrections for basis set superpositions errors). The computed vibrational frequencies and normal mode forms of the GLYZWM complex resemble in many cases experimental assignments made for the glycine zwitterion in bulk water on the basis of Raman spectroscopy. © 1996 by John Wiley & Sons, Inc.

Introduction

Glycine (NH_2CH_2COOH, GLY) is the smallest amino acid, and its interactions with surrounding media (in particular, water) are of considerable biological interest. In the gas phase, only nonionized glycine conformers have been detected[1]; however, glycine exists singularly as a zwitterion ($NH_3^+CH_2COO^-$, GLYZW) in polar so-lution or in the crystalline phase.[2] Several research groups have applied high-level *ab initio* molecular orbital techniques to elucidate the properties of nonionized GLY, and five local minima have been identified on the potential energy surface.[3] The computational[3] and experimental[2] studies agree that GLY(I) represents the global minimum in the gas phase. Recently, we concluded from a detailed *ab initio* electronic structure investigation that there is no minimum on the gas phase potential energy surface corresponding to a glycine zwitterion.[4] With basis sets which contain polarization functions on hydrogen atoms (e.g., 6-31G**), the

*Author to whom all correspondence should be addressed.

intramolecularly hydrogen-bonded zwitterion species GLYZW(cis) collapses to glycine conformer GLY(II) in calculations at both independent particle and electron correlated levels; the GLY(II) local minimum lies 1.0–1.5 kcal/mol higher in energy than GLY(I).[2,3] The GLYZW(trans) rotamer may be viewed as a transition state for hydrogen exchange between the carboxylic acid and the amino groups, located approximately 20 kcal/mol higher in energy than GLY(II).[4] Similar conclusions regarding the glycine zwitterion potential energy surface were reach simultaneously and independently by Rauk et al.[5]

GLY(I)

GLY(II)

GLYZW(cis)

GLYZW(trans)
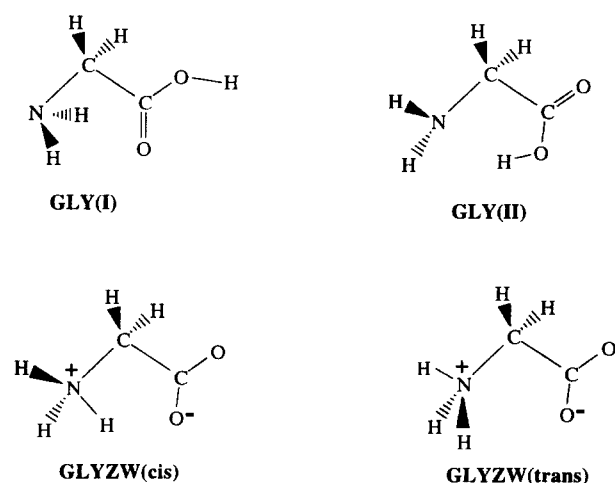
In this article we examine the potential energy surface for the 1:1 complex between the glycine zwitterion and water (GLYZW:H_2O). Complexation of the strongly dipolar zwitterion with a single hydrogen-bonding water molecule provides information about specific solute–solvent interactions and serves as an initial indicative step of the full solvation process. We are not aware of any quantitative experimental information regarding the properties of a monosolvated glycine zwitterion complex; computational results on the interaction between the glycine zwitterion and water molecules have been reported, however.[6–12] Several studies have employed classical molecular mechanics or molecular dynamics simulations.[6,7] A few studies have used self-consistent reaction field methods[8] or a supermolecule approach within the framework provided by semiempirical or ab initio electronic structure theory.[9–12] The highest-level ab initio study to date on the GLYZW:H_2O complex has been carried out by Langlet et al.[10] These authors imposed molecular symmetry constraints which, as we will show in this article, lead to the location of stationary points which are saddle points, rather than minima, on the potential energy

surface. A small number of structures representing 1:1 nonionized glycine–water complexes have recently been investigated by Basch and Stevens using ab initio electronic structure techniques.[13]

We have investigated a large number of possible conformers for the GLYZW:H_2O complex using an ab initio supermolecule approach. Only one true minimum could be located on the potential energy surface, a low-symmetry species not identified in previous studies. The computed vibrational frequences and normal mode forms of the minimum energy complex resemble in many cases experimental assignments made for the glycine zwitterion in bulk water on the basis of Raman spectroscopy.[14]

Computational Details

Ab initio calculations were carried out with the Gaussian 92[15] series of electronic structure programs on Hewlett-Packard 9000/735 and Convex C220 computers. A large number of stationary points were located at the Hartree-Fock (HF) level and characterized by harmonic normal mode analysis as minima, transition states, or higher-order saddle points. The geometries of a few selected conformations were optimized with electron correlation incorporated through Møller-Plesset perturbation theory at the level of second order (MP2). Due to extreme shallowness of the potential energy surface with respect to several rocking and twisting degrees of freedom, some geometry optimizations required the use of an analytically computed second derivative Hessian matrix at every optimization cycle in order to reach convergence. Optimization in Cartesian rather than internal coordinates proved beneficial in a number of cases as well. The calculations made use of standard split valence (4-31G), split valence plus polarization (6-31G*), and split valence plus polarization and diffuse function (6-31 + G*, 6-31 + + G**) basis sets.[16a] In a few instances, a hybrid basis set which included only a set of p-type polarization functions on a specific hydrogen atom (α_p = 1.10) was used in conjunction with the 4-31G basis set, 4-31G(H*). A special basis set (Basis A) was assembled for use in geometry optimizations at the MP2 level. Basis A was designed to provide a flexible description of the charged groups on the zwitterion and their interaction with the solvating water. Basis A consisted of the 6-31G** basis set for the NH_3^+ unit, 6-31 + G for the CO_2^- unit, 6-31G for

the CH_2 group, and 6-31 + G** for the H_2O molecule. GLYZW-H_2O hydrogen bond interaction energies were evaluated and corrected for basis set superposition errors (BSSE).[16b] Finally, a few calculations employed the self-consistent reaction field (SCRF) model incorporated into the Gaussian 92 program.

To assist in the vibrational assignments, a program was developed locally to compute the normal mode potential energy distributions in terms of force constants corresponding to local internal coordinates. The input consisted of structural information and Cartesian force constants from the Gaussian 92 program.
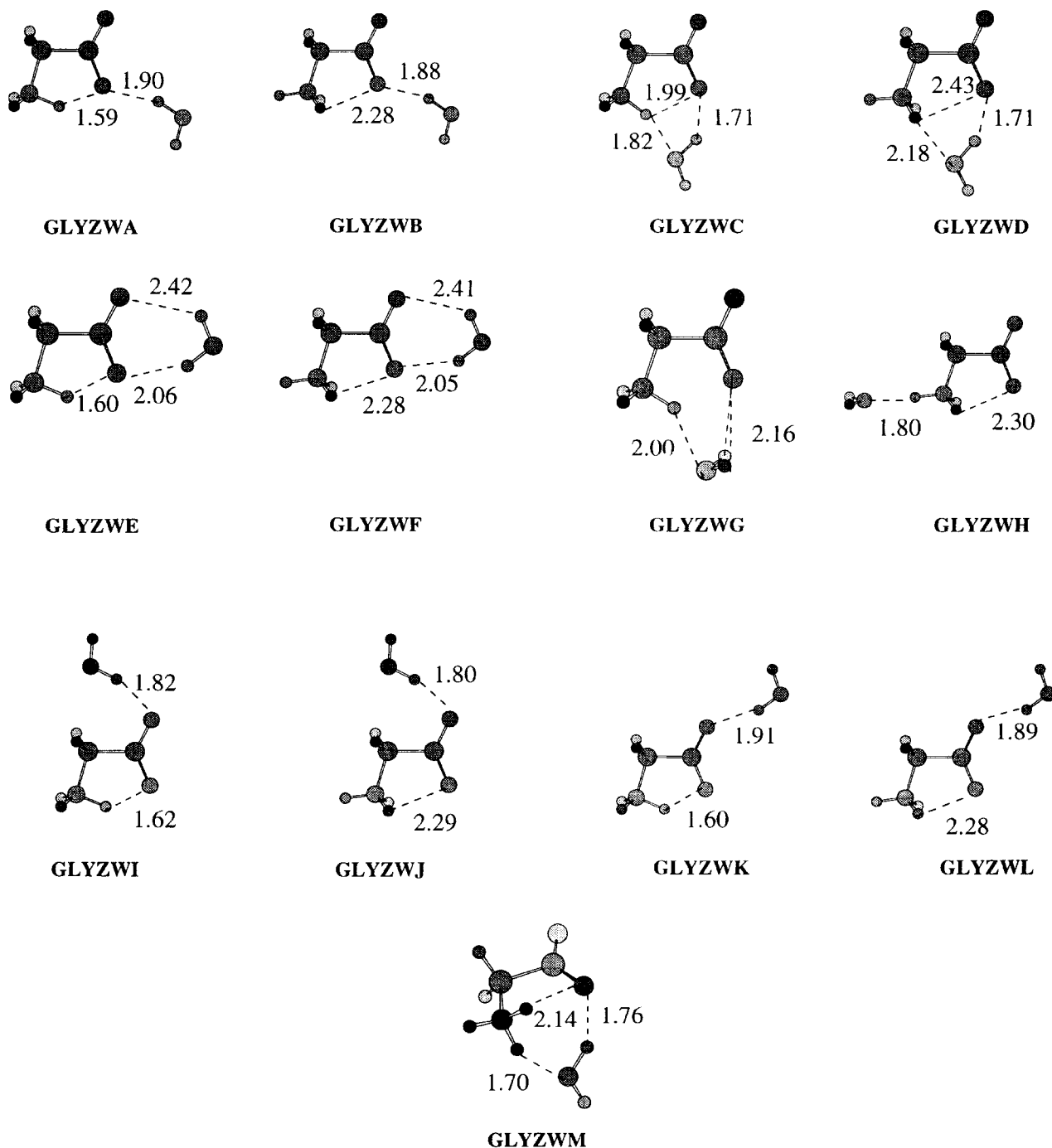

FIGURE 1. Schematic structures of the 1 : 1 glycine zwitterion –water complexes considered in this article.

Results and Discussion

GEOMETRIES AND RELATIVE STABILITIES OF 1 : 1 GLYZW:H₂O COMPLEXES

A number of possible structures for the $1:1$ GLYZW: H_2O complex (GLYZWA-GLYZWM) are shown schematically in Figure 1. Total and relative energies for stationary points representing the GLYZWA-GLYZWM structures located at the HF/4-31G level are given in Table I. Point group symmetries (C_s) were imposed in the optimization of all conformers except GLYZWM (C_1), and the magnitudes of any imaginary frequencies obtained from the normal mode analyses are also shown in Table I. The conformers examined feature several potential types of hydrogen bonding between the glycine zwitterion and the water molecule: a single hydrogen bond with the CO_2-group as the acceptor and the H_2O as the donor (GLYZWA, GLYZWB, GLYZWI, GLYZWJ, GLYZWK, GLYZWL); a single hydrogen bond with the NH_3^+ group as the donor and H_2O as the acceptor (GLYZWH); two nonsymmetrical hydrogen bonds with the CO_2^- group as the acceptor and H_2O acting as a double proton donor (GLYZWE, GLYZWF) and multiple hydrogen bonds involving both charged glycine terminal groups and the water molecule acting as both a proton donor and acceptor (GLYZWC, GLYZWD, GLYZWG, GLYZWM). Several conformers form rotameric pairs in which the individual conformers differ primarily in the relative orientation of the NH_3^+ group (GLYZWA, GLYZWB; GLYZWC, GLYZWD; GLYZWE, GLYZWF; GLYZWI, GLYZWJ; GLYZWK, GLYZWL) and hence permit linear, in-plane (GLYZWA, GLYZWC, GLYZWE, GLYZWI, GLYZWK), or bifurcated (GLYZWB, GLYZWD, GLYZWF, GLYZWJ, GLYZWL) intramolecular NH_3^+—CO_2^- hydrogen bonding.

Only two local minima (GLYZWE and GLYZWM) could be located at the HF/4-31G level; all other structures considered are either transition states (GLYZWA, GLYZWF, GLYZWG, GLYZWI, GLYZWJ, and GLYZWK) or second-order saddle points (GLYZWB, GLYZWC, GLYZWD, GLYZWH, GLYZWL) on the GLYZW:H_2O potential energy surface. The two minima are almost 9 kcal/mol apart in energy, and the GLYZWM conformer is computed as the global minimum. A pair of rotameric structures (GLYZWC, GLYZWD) is located some 5 kcal/mol above GLYZWM as second-order saddle points. Two transition state structures are computed to be approximately 7 kcal/mol (GLYZWI) or 9 kcal/mol (GLYZWG) above GLYZWM, whereas the remaining structures shown in Figure 1 are more than 10 kcal/mol higher in energy than GLYZWM. A few common characteristics may be drawn from the data listed in Table I. Thus, as expected, two-site binding, which engages both water protons in interactions with the NH_3^+/COO^- units, is far preferable to single-site binding, and single-site binding structures (GLYZWA, GLYZWB, GLYZWH, GLYZWI, GLYZWJ, GLYZWK, and GLYZWL) are all computed well above the global minimum in energy. With the exception of the (GLYZWC, GLYZWD)

TABLE I.

Total energies (au), relative energies (kcal / mol), and imaginary frequencies (cm⁻¹) of the 1 : 1 glycine zwitterion – water complexes GLYZWA-GLYZWM (HF / 4-31G // 4-31G).

Conformer	E	E_{rel}	Imag. Freqs.
GLYZWA	−358.29278	11.5	120i
GLYZWB	−358.28664	15.3	233i, 126i
GLYZWC	−358.30357	4.7	147i, 45i
GLYZWD	−358.30401	4.4	148i, 41i
GLYZWE	−358.29701	8.8	None
GLYZWF	−358.29101	12.6	231i
GLYZWG	−358.29724	8.7	292i
GLYZWH	−358.29350	11.0	136i, 48i
GLYZWI	−358.30070	6.5	11i
GLYZWJ	−358.29479	10.2	230i
GLYZWK	−358.28401	10.7	10i
GLYZWL	−358.28785	14.6	235i, 15i
GLYZWM	−358.31107	0.0	None

pair of rotamers, the NH_3^+—CO_2^- intramolecular hydrogen bond interactions preferentially stabilize the in-plane hydrogen bonded conformers by 3.8 ± 0.1 kcal/mol (GLYZWA vs. GLYZWB, etc.), a computed difference which is remarkably constant through the sets of rotamers. Nuclear magnetic resonance (NMR) experiments show nearly free rotation of the glycine zwitterion NH_3^+ unit in water (i.e., a barrier of only 1–2 kcal/mol), indicating that in bulk water, additional water molecules are solvating the hydrogens of this structural unit.[17] Imaginary frequency values near $230i$ cm^{-1} are associated with internal rotation of the NH_3^+ group around the C—N bond, while most frequencies near $130i$ cm^{-1} correspond to rotation of the OH bond around the OH(water) $\cdots$ O(glycine) axis. The structures identified at the HF/4-31G level as less than 10 kcal/mol in energy above the global minimum (GLYZWM)—that is, the second minimum (GLYZWE) and the low-energy transition states (GLYZWI, GLYZWG) or second-order saddle points (GLYZWC, GLYZWD)— were selected for additional calculations.

Relative energy and vibrational frequency results obtained on the GLYZWC, GLYZWD, GLYZWE, and GLYZWM conformers at the HF level with larger basis sets are presented in Table II. The GLYZWE species features two intermolecular hydrogen bonds between the H_2O molecule and the CO_2^- unit as well as intramolecular hydrogen bonding between NH_3^+ and one of the oxygens of the CO_2^- group. Although normal mode analysis establishes GLYZWE as a local minimum at the HF/4-31G level, we did notice that GLYZWE contains the intramolecular hydrogen bond type also displayed by the zwitterion structure, GLYZW(cis).[4] As we have shown, in the gas phase the bridging hydrogen atom in GLYZW(cis) will transfer back to the CO_2^- unit to form GLY(II) without an activation energy barrier when a p-type polarization function is included in the basis set

covering the migrating hydrogen atom (e.g., 6-31G* → 6-31G**). Thus, it would perhaps be surprising if GLYZWE survived as a minimum at higher calculational levels.

GLYZWE   **TS**
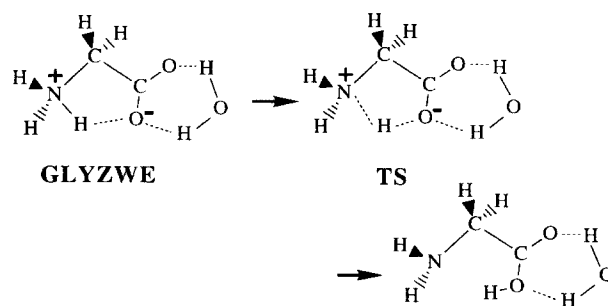
We have located the transition state for the NH_3^+—CO_2^- single hydrogen transfer reaction at the HF level using the 4-31G and 4-31G(H*) basis sets and imposed C_s symmetry constraints. The magnitude of the single imaginary frequency in the TS structure is $953i$ cm^{-1} at the HF/4-31G level [$584i$ cm^{-1}, HF/4-31G(H*)] and the corresponding normal mode clearly promotes in-plane hydrogen transfer from NH_3^+ to CO_2^-. However, the computed activation energy barriers are only 0.44 kcal/mol and 0.09 kcal/mol at the HF/4-31G and HF/4-31G(H*) levels, respectively. Single-point calculations at the correlated MP2 level using the optimized HF geometries of GLYZWE and TS (MP2/4-31G//4-31G) predict that TS, the transition state structure, is 4.3 kcal/mol lower in energy than the GLYZWE reactant; the difference is 0.5 kcal/mol with the 4-31G(H*) basis set. These energy separations increase by an additional 2.1 kcal/mol [1.5 kcal/mol for 4-31G(H*)] when vibrational zero-point energy corrections are made, reaching a total of 6.4 kcal/mol (4-31G) and 2.0 kcal/mol [4-31G(H*)] in favor of TS. A series of one-dimensional potential energy surface scans along the N—H internal coordinate with all remaining parameters optimized at the HF or MP2 levels [using either the 4-31G or 4-31G(H*) basis

TABLE II.

Relative energies (kcal / mol) and imaginary frequencies (cm^{-1}, in parentheses) of the GLYZWM, GLYZWE, GLYZWC, and GLYZWD conformers.

	HF / 6-31G*	HF / 6-31 + G*	HF / 6-31 ++ G**
GLYZWM	0.00	0.00	0.00
GLYZWE	5.54	5.32	4.52
GLYZWC	4.53 ($194i, 70i, 32i$)	2.97 ($119i, 24i$)	2.81 ($98i, 21i$)
GLYZWD	3.35 ($89i, 18i$)	2.04 ($69i$)	2.17 ($68i$)

sets] show (Fig. 2) that structure TS exists at the HF level, but also that the total energy decreases steadily as a function of the N—H distance at the MP2 correlated level. Furthermore, geometry optimizations of GLYZWE at the HF/6-31G*, HF/6-31 + G*, and HF/6-31 + + G** levels lead to a local minimum in all three cases, but the barrier for hydrogen atom transfer is only on the order of 0.1–0.5 kcal/mol. The zwitterionic structure collapses without a barrier to a neutral glycine : water complex whenever electron correlation is included in the calculations. All these results are similar to the results we obtained previously regarding the stability of GLYZW(cis),[4] and we therefore conclude that the GLYZWE conformer does in fact not represent a stable complex.

At the HF/4-31G level, the dual hydrogen-bonded GLYZWC and GLYZWD conformers are both second-order saddle points approximately 4.5 kcal/mol higher in energy than the GLYZWM conformer (Table I). The GLYZWC structure was calculated by Langlet et al. (their CZ1 structure) as the most stable GLYZW : H_2O conformer at the HF/6-31G* level.[10] We have repeated their calculation, and vibrational analysis unequivocally shows that the purported GLYZWC minimum has three imaginary frequencies (Table II). In addition, the rotamer GLYZWD is nearly 1 kcal/mol lower in energy than GLYZWC (HF/6-31G*), but
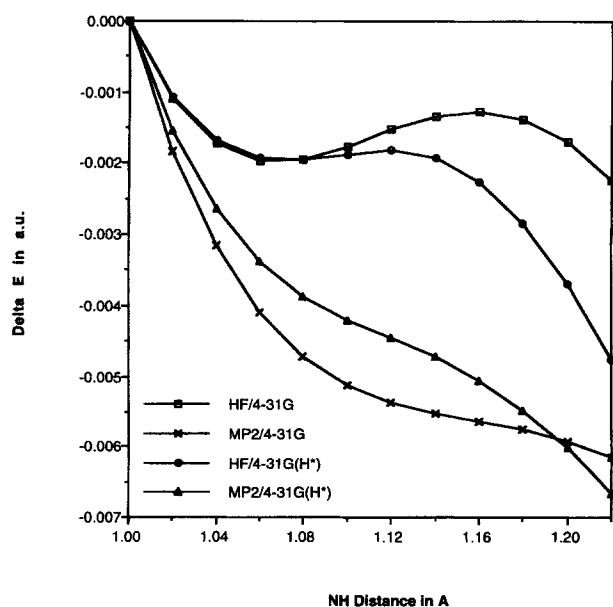

FIGURE 2. Energy of the GLYZWE conformer as a function of the internal N — H coordinate. At each N — H distance, all other geometrical variables were optimized subject to an overall C_s molecular symmetry constraint.

GLYZWD still possesses two imaginary frequencies. The closeness in energy of these two rotameric structures results, judging from the distances between the relevant interacting moieties (Fig. 1), from the combination of a relatively weak in-plane NH_3^+—CO_2^- hydrogen bonding interaction in GLYZWC and a surprisingly strong bifurcated hydrogen bond developing in GLYZWD between the NH_3^+ unit and the out-of-plane $O(2p)$ orbital on the water molecule. The two structures remain saddle points at the HF level with even larger basis sets (6-31 + G* and 6-31 + + G**), although the number of imaginary frequencies decreases to two for GLYZWC and one for GLYZWD. One of the normal modes with an imaginary frequency in the GLYZWC structure describes rotation of the NH_3^+ group, and the frequency of this mode becomes positive in GLYZWD. Upon symmetry relaxation ($C_s \rightarrow C_1$) and reoptimization, both GLYZWC and GLYZWD collapse to the GLYZWM conformer without energy barriers. Thus, the GLYZWD structure may be regarded as the planar transition state connecting GLYZWM with its mirror image.

GLYZWI (alias CZ2 in ref. 10) lies 6.5 kcal/mol above GLYZWM at the HF/4-31G level and has just one imaginary frequency of a very small magnitude ($\sim 10i$ cm^{-1}); consequently, a true minimum could be expected to be located in its vicinity. In fact, symmetry relaxation from C_s to C_1 and reoptimization at the HF/4-31G level lead to a shallow minimum only 0.01 kcal/mol lower in energy. GLYZWI features the intramolecular N—H ⋯ O hydrogen bond type, which collapses in the GLYZWE conformer [and GLYZW(cis) itself] when higher-level calculations are carried out (polarization functions on hydrogen or inclusion of electron correlation). Indeed, at the HF/6-31G** level there is no energy barrier opposing the collapse of GLYZWI to a nonionized glycine–water complex either, and we predict on the basis of the calculations outlined in this article and in our previous work[4] that all other structures (such as GLYZWK) showing intramolecular NH_3^+—CO_2^- hydrogen bonding will also collapse to nonionized glycine–water complexes at higher levels of theory. Finally, the imaginary mode in the GLYZWG structure involves rotation of the bifurcating water hydrogen atoms away from perpendicularity with respect to the glycine plane, and a small distortion along this mode followed by geometry reoptimization leads directly to GLYZWM.

Consequently, we believe that GLYZWM represents the only true minimum on the potential

energy surface for the $1:1$ glycine zwitterion–water complex. The structure of GLYZWM has been optimized at the MP2 level using both the 4-31G basis set and Basis A; essential structural parameters are given in Table III, while two views of the structure (including the atom numbering scheme) are presented in Figure 3. The solvating water is at all levels of calculation centered between the CO_2^- and NH_3^+ groups [O_3—O_{11} = 2.77 Å; N_5—O_{11} = 2.74 Å; $N_5H_{10}\cdots O_{11}$ = 2.06 Å, $O_3\cdots H_{13}(O_{11})$ = 1.93 Å; MP2/Basis A], engaged in bridged hydrogen bonding both as a proton donor to CO_2^- and as a proton acceptor from

NH_3^+. The GLYZWM structure is also stabilized by a seemingly strong interaction between a second hydrogen atom of the NH_3^+ group and an oxygen atom of the CO_2^- group. The presence of the solvating water molecule induces slight torsion to the glycine zwitterion backbone. The optimized GLYZWM structure shows a torsion angle ($\angle N_5C_1C_2O_3$) of approximately 7° at the HF/6-31G*, HF/6-31 + G*, and HF/6-31 + + G** levels. Optimization at the MP2 level leads to a smaller backbone torsion angle of 4° (4-31G and Basis A). The backbone torsion angle is nearly 20° in the crystal structure of zwitterionic glycine, and it can

TABLE III.

Optimized parameters of GLYZWM (distances in Å, angles in degrees).

|	HF / 6-31 ++ G**	HF / 4-31G	MP2 / 4-31G	MP2 / Basis A
Bond				
C1 — C2	1.558	1.551	1.580	1.577
C2 — O3	1.246	1.271	1.320	1.326
C2 — O4	1.208	1.224	1.260	1.267
C1 — N5	1.499	1.509	1.527	1.512
C1 — H7	1.079	1.077	1.091	1.095
C1 — H8	1.081	1.078	1.093	1.094
N5 — H6	1.004	1.002	1.020	1.015
N5 — H9	1.016	1.011	1.054	1.070
N5 — H10	1.014	1.031	1.045	1.024
O3 — O11	2.782	2.621	2.717	2.767
O11 — H12	0.943	0.947	0.971	0.963
O11 — H13	0.957	0.975	0.996	0.980
Bond Angle				
C1 — C2 — O3	113.0	113.3	112.8	112.1
C1 — C2 — O4	114.7	115.3	116.4	117.3
C2 — C1 — N5	107.7	108.0	105.7	104.9
C2 — C1 — H7	111.6	110.6	111.0	110.8
C2 — C1 — H8	110.2	110.3	110.7	110.9
C1 — N5 — H6	114.4	113.8	115.1	115.2
C1 — N5 — H9	105.1	106.4	100.8	99.1
C1 — N5 — H10	110.7	109.3	109.1	112.2
C2 — O3 — O11	104.5	105.9	97.9	101.4
O3 — O11 — H12	131.7	137.2	133.0	132.1
H12 — O11 — H13	108.7	115.1	112.7	107.3
Dihedral Angle				
O3 — C1 — C2 — O4	180.3	180.0	180.1	− 179.9
O3 — C2 — C1 — N5	− 7.3	− 0.5	− 3.8	− 3.7
C2 — C1 — N5 — H6	157.7	164.3	149.5	137.0
N5 — C2 — C1 — H7	− 119.9	− 119.1	− 118.9	− 118.0
N5 — C2 — C1 — H8	118.0	119.2	118.2	119.3
H6 — C1 — N5 — H9	− 123.1	− 123.6	− 124.0	− 122.1
H6 — C1 — N5 — H10	125.4	125.5	128.8	128.7
O4 — C2 — O3 — O11	240.0	232.6	237.4	239.1
C2 — O3 — O11 — H12	96.8	110.8	95.6	104.3
O3 — H12 — O11 — H13	− 11.4	− 8.0	− 11.0	− 7.8

The atom numbering scheme is shown in Figure 3.

thus be expected that some backbone deformation occurs in solution.

The structures of five 1:1 nonionized glycine–water complexes have been optimized by Basch and Stevens at the HF level using effective core potentials, split valence plus polarization function basis sets, and planar (C_s) symmetry constraints.[13] To compare the relative energies of nonionized and zwitterionic glycine–water complexes, we have optimized a single nonplanar nonionized glycine–water complex (GLY:H_2O) in which the water molecule is bridging between the NH_2 and COOH groups, representing the immediate product of sequential proton transfers from the N_5 to the O_3 terminus of GLYZWM through the intervention of the water molecule. The major structural feature of interest in this complex is that the backbone ($\angle N_5C_1C_2O_3(H)$) is twisted by nearly 60° so that a hydrogen atom on C is approximately coplanar with the COOH unit. This GLY:H_2O conformer is a minimum[†] and 6.4 kcal/mol lower in energy than GLYZWM (HF/6-31 + + G**), demonstrating that the monohydrated glycine zwitterion is thermodynamically unstable relative to its nonionized form plus a water molecule. The total hydrogen bond energy for the zwitterion based directly on the supermolecule approach is 15.9 kcal/mol (HF/6-31 + + G**) and 18.3 kcal/mol (MP2/6-31 + + G**//6-31 + + G**);

[†]Our GLY:H_2O conformer is 7.0 kcal/mol higher in energy than the lowest-energy structure located in ref. 13 (I-A). We have investigated the I-A conformer in some detail and have found that it is a minimum at the HF/4-31G level, but I-A has one imaginary frequency of magnitude $35i$ cm^{-1} at the HF/6-31 + G* level. The imaginary frequency is larger, $135i$ cm^{-1} at the HF/6-31 + + G** level. Reoptimization of I-A in C_1 symmetry leads to a nonplanar minimum about 0.1 kcal/mol lower in energy (HF/6-31 + + G**).

after BSSE corrections have been included, the computed interaction energies decrease to 15.0 kcal/mol and 15.9 kcal/mol, respectively. The analogously computed total hydrogen bond energy for our nonplanar GLY:H_2O structure is 9.0 (8.0) kcal/mol at the HF level and 12.5 (9.9) kcal/mol at the MP2 level (BSSE corrected values in parentheses), so the hydrogen bond interaction energies are approximately 6–7 kcal/mol larger in the monosolvated zwitterion than in the nonionized species.
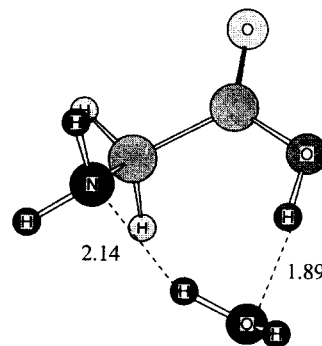
Continuum models represent efficient approaches to the inclusion of solvation effects which are different in nature from the discrete supermolecule model applied above. Bonaccorsi et al. carried out *ab initio* calculations on zwitterionic glycine including a point–dipole self-consistent reaction field (SCRF) model.[8a, 1a] They found that geometries with a fully planar backbone were preferred and that GLYZW(trans) was more stable than GLYZW(cis) by 0.7 kcal/mol at the HF/4-31G level. Calculations using the semiempirical MNDO
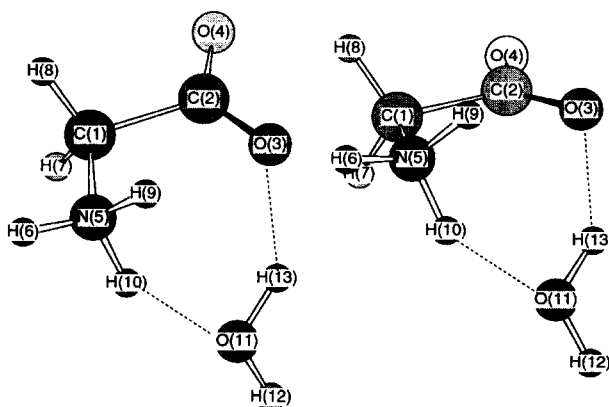

FIGURE 3. Two views of the GLYZWM conformer with complete atom numbering scheme.
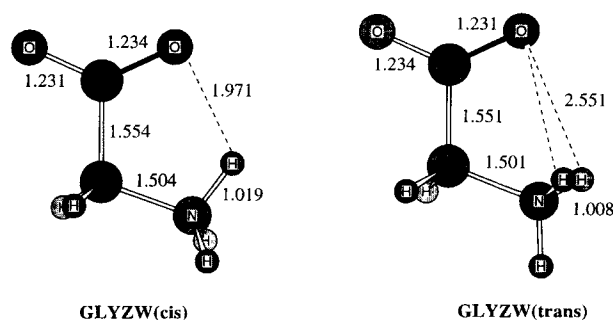
FIGURE 4. Optimized HF / 6-31 + + G** geometries of GLYZW(cis) and GLYZW(trans) from the SCRF method.

Hamiltonian and a generalized Born equation to compute the solvation energy also suggested that planar structures are most stable, although the cis form appeared to be preferred.[8b] To satisfy our curiosity regarding the performance of simple solvation models on the present system, we have applied the SCRF approach[18b] implemented in the Gaussian 92 series of programs to GLYZW (cis and trans) at the HF/6-31 + + G** level. In this method based on the Onsager reaction field model, the solute is placed in a spherical cavity and the surrounding solvent is treated as a uniform dielectric medium. An estimated cavity radius of 3.5 Å (obtained using the "volume" keyword in Gaussian 92[15]) and a dielectric constant appropriate to bulk water (80) were used. The solute–solvent interaction depends on the (square of) the solute dipole moment, and specific solute–solvent interactions are completely ignored. The present calculations, with the model solvent included, find GLYZW(cis) as a minimum of C_s symmetry with a very low-lying normal mode for backbone torsion (~ 35 cm^{-1}), in contrast to the results obtained in the gas phase, where no minimum corresponding to GLYZW(cis) exists at the HF/6-31 + + G** level. GLYZW(cis) in aqueous solvent is calculated to be 0.6 kcal/mol more stable than GLYZW(trans), but GLYZW(trans) is only a transition state for internal NH$_3{}^+$ rotation. Thus, even this crude solvation model appears capable of capturing some of the stabilizing effects exerted by a highly polar solvent on the zwitterionic form of glycine.

VIBRATIONAL ANALYSIS OF THE MINIMUM ENERGY COMPLEX (GLYZWM)

It is not possible to carry out a meaningful harmonic vibrational analysis on isolated GLYZW since this species does not represent a minimum. Alper et al.[7] have presented a detailed analysis of

the vibrational spectrum of the glycine zwitterion in aqueous solution on the basis of molecular dynamics simulations in which water molecules (~ 250) surrounded the glycine zwitterion. They determined the effective forces and force constants in solution, adopted the average glycine zwitterion geometry from the dynamics calculations (similar to the crystal structure geometry) as the equilibrium geometry, and assigned the observed vibrational spectra using 13 scale factors. Our computed minimum energy 1:1 complex, GLYZWM, incorporates some of the effects of water solvent and, importantly, since this conformer represents a minimum on the potential energy surface, it is in fact possible to perform a normal mode analysis directly based on a quadratic expansion. In this section we assign the normal modes for GLYZWM and make comparisons to the experimental aqueous solution spectra for the glycine zwitterion.

The vibrational frequencies computed at the HN/6-31 + + G** level for GLYZWM are given in Table IV along with a description of the normal modes in terms of their potential energy distributions. The choices of the $3N - 6$ ($N =$ number of atoms) nonredundant local internal coordinates are based on the recommendations made by Pulay et al.,[19] except that we define an equivalent bending rather than scissoring mode [e.g., the CH$_2$ bending coordinate is defined as

$$4\Delta\Theta(H_7C_1H_8) - \Delta\Theta(H_7C_1C_2) - \Delta\Theta(H_8C_1C_2)$$
$$- \Delta\Theta(H_7C_1N_5) - \Delta\Theta(H_8C_1N_5)].$$

Theoretical frequencies derived from *ab initio* calculations at the HF level are typically 10% larger than experimentally measured frequencies due to the neglect of anharmonicity and electron correlation. Indeed, a comparison (not shown) of the HF/4-31G and MP2/4-31G frequencies in GLYZWM finds the latter to be on the average 7% smaller in magnitude. The theoretical frequencies listed in Table IV have therefore been scaled down by 10% from their actual values computed at the HF/6-31 + + G** level. Several experimental vibrational spectroscopic studies have been performed on the glycine zwitterion and its isotopomers in water solution.[14,20] The most recent report by Furic et al.[14] presented the vibrational Raman spectroscopy data on NH$_3{}^+$CH$_2$CO$_2{}^-$ in H$_2$O reproduced in Table IV.

All the fundamental stretching modes involving a hydrogen atom fall into the spectral region above 2900 cm^{-1}. These frequencies are well separated

into noninteracting groups corresponding to internal O—H, N—H, and C—H displacements, with the water and methylene group modes at the highest and lowest energies, respectively. The only observed glycine zwitterion bands in this region correspond to the methylene modes. The relative Raman activity of the calculated CH_2 symmetric stretch at 2938 cm^{-1} (exp.: 2972 cm^{-1}) is twice as strong as that of the antisymmetric stretch at 3002

cm^{-1} (exp.: 3016 cm^{-1}), in correct order according to experimental observations, although the measured intensity ration appears closer to four.[20b] The internal water modes are computed at 3808 cm^{-1} (antisymmetric stretch) and 3578 cm^{-1} (symmetric stretch) in the complex, a decrease of 35–50 cm^{-1} from the values computed in a free H_2O molecule (3842 and 3723 cm^{-1}, respectively; HF/6-31 + + G** frequencies reduced by 10%).

TABLE IV.
Normal mode frequencies and assignments for GLYZWM.

No.	Description According to the Calculations[a]	Calc.[b]	Exp.[c]	Raman[d]	IR[e]
1	HOH as(100)	3808		613	74
2	HOH ss(100)	3578		639	179
3	NH3 as(90)	3406		571	48
4	NH3 as(104)	3263		257	101
5	NH3 ss(88), NH3 as(17)	3212		610	70
6	CH2 as(97)	3002	3016 vw	436	3
7	CH2 ss(97)	2938	2972 s	887	5
8	COO as(86)	1744	1620 m, br	41	370
9	NH3 ad(84)	1643		46	43
10	NH3 ad(58)	1578		50	13
11	H2O b(89)	1572		29	89
12	NH3 sd(70), NH3 as(16)	1462	1514 vw	29	133
13	NH2 b(83)	1441	1444 m	46	8
14	COO ss(46), CH2 wag(21)	1347	1412 s	48	122
15	CH2 wag(50), COO ss(31)	1300	1330 s	61	39
16	CH2 tw(53), NH3 r(23)	1274	1320 w, sh	32	10
17	NH3 r(55)	1061	1121 w, br	19	16
18	NH3 r(29), CH2 tw(27), CH2 r(15)	1051		17	8
19	CN s(67), CC s(15)	1006	1031 m	43	2
20	CH2 r(35), NH3 r(32), COO wag(20)	892		6	16
21	OCO b(35), CC s(28)	834	897 s	87	61
22	glyzw : water int(34), OCO b(24)	652	671 w	11	44
23	glyzw : water int(56), OCO b(16), CC s(16)	629		22	53
24	COO wag(57), CH2 r(31)	564	585 w	12	12
25	CCN b(20), CC s(22), OCO r(19)	478	507 m	33	22
26	glyzw : water int(77), OCO r(17)	443		3	62
27	CCN b(51), OCO r(46)	327		2	16
28	glyzw : water int(51), NH3 tors(45)	221		4	50
29	glyzw : water int(55), COO tors(54)	190		1	27
30	glyzw : water int(100)	167		10	29
31	glyzw : water int(47), NH3 tors(37)	136		5	27
32	COO tors(52), glyzw : water int(46)	102		4	7
33	glyzw : water int(100)	53		14	1

[a] Normal mode potential energy distribution; contributions of less than 15% have not been included. The nomenclature used for the internal coordinates is s = stretch, as = antisymmetric stretch, ss = symmetric stretch, ad = antisymmetric deformation, sd = symmetric deformation, tors = torsion, b = bend, tw = twist, wag = wag, r = rock, glyzw : water int = internal coordinates describing GLYZW: H_2O interactions.

[b] Computed values (cm^{-1}) at the HF/6-31 + + G** level diminished by 10% (see text).

[c] Experimental Raman frequencies (cm^{-1}) from ref. 14. The nomenclature used for the intensities is w = weak, vs = very strong, m = medium, vw = very weak, s = strong, sh = shoulder, br = broad.

[d] Computed relative Raman intensities. The intensity of the weakest mode (number 29) has arbitrarily been set to unity.

[e] Computed relative IR intensities. The intensity of the weakest mode (number 33) has arbitrarily been set to unity.

Two intense bands observed at 3415 and 3235 cm^{-1} in the Raman spectra of the zwitterion in water are associated with bulk water.[14]

The 1000–2000 cm^{-1} spectral region contains 12 fundamentals in the GLYZWM complex. These modes correspond mostly to stretch coordinates involving nonhydrogen atoms only or a variety of angle bending motions. A total of eight frequencies have been assigned experimentally to the glycine zwitterion in the 1000–1650 cm^{-1} region, and most of these are readily recognized in the computed spectrum of GLYZWM. For example, the weak and broad band observed at 1121 cm^{-1} must correspond to the NH_3^+ rocking modes computed at 1061 and 1051 cm^{-1}. Likewise, an absorption feature observed at 1330 and 1320 cm^{-1} as a strong line with a weak shoulder matches well with the computed CH_2 wag and twist fundamentals at 1300 and 1274 cm^{-1} of relative Raman intensities 2:1. In some cases, computed values for the complex and experimental assignments from solution data form a nearly perfect match. Thus, the computed mode at a frequency of 1006 cm^{-1} is a mixture of CN and CC stretches, in complete agreement with the experimental assignment at a fundamental frequency of 1031 cm^{-1}. An almost pure CH_2 bending mode is computed at 1441 cm^{-1} and observed at 1444 cm^{-1}. However, on the basis of our computed frequencies and mode forms, we question the assignment made in ref. 14 of the observed band at 1620 cm^{-1} as solely arising from the CO_2^- antisymmetric stretch. Considering the frequency, intensity, and pronounced broadness of this band, we find it more likely that several absorption lines are contributors, including the NH_3^+ antisymmetric deformations, the H_2O bending mode, and, perhaps, the CO_2^- antisymmetric stretch. The bending mode of the water molecule is present at a frequency near 1630 cm^{-1} in H_2O solutions,[20b] and an NH_3^+ degenerate deformation mode has been assigned near 1600 cm^{-1}.[20c]

Four fundamental frequencies have been observed in the 1000–500 cm^{-1} region. The most intense feature is at 897 cm^{-1} and is from our calculation assigned to a mixed OCO bending and CC stretching mode. Two weak bands observed at 671 and 585 cm^{-1} have been experimentally assigned to CO_2^- bending and wagging modes, respectively. The lower-energy mode has an analogue computed in GLYZWM at 564 cm^{-1} (primarily CO_2^- wag), but there is no immediate analogue to the higher-frequency mode. A glycine

zwitterion–water interaction coordinate mixes with the CO_2^- rocking coordinate, and two normal modes of frequencies 652 and 629 cm^{-1} result. The lowest observed fundamental frequency for glycine in water (507 cm^{-1}) has a computed match in GLYZWM at 478 cm^{-1}, a mixed mode with CCN bending, CC stretch, and OCO rocking character. According to our calculations, the primary spectroscopic effects of glycine zwitterion–water interactions should be found in the fundamental modes below 500 cm^{-1}, but they are, unfortunately, all computed to possess low Raman scattering intensities.

Conclusions

We have located the minimum structure for the 1:1 complex formed between the glycine zwitterion and water. In this conformer, GLYZWM, the backbone of the glycine unit is slightly twisted and the water molecule acts as both a hydrogen bond donor and acceptor, positioned to prevent the transfer of a hydrogen atom between the zwitterion-forming NH_3^+ and CO_2^- units. The total hydrogen bond energy appears to be substantial ($\sim$ 15–16 kcal/mol) but is not sufficient to render the 1:1 glycine zwitterion–water complex lower in energy than a 1:1 nonionized glycine–water complex. Additional solvent molecules are necessary to stabilize the zwitterionic species and the oxygens of the COO^- unit, and the hydrogens attached to nitrogen appear to be the most obvious binding sites for the next few waters. The complete vibrational analysis of GLYZWM provides some analogies with experimental Raman spectroscopic data for the glycine zwitterion in bulk water and may prove beneficial in verifying experimental detection of the GLYZWM species alone or as the principal component of a larger GLYZW: nH_2O cluster.[11,12] The simple neutral and zwitterionic glycine–water complexes provide information on pairwise solute–solvent interactions, which could be useful in calibrating potential energy functions employed in future molecular mechanics/dynamics simulations on the properties of the glycine zwitterion in water.

Acknowledgments

References

1. (a) R. D. Suenram and F. Lovas, *J. Am. Chem. Soc.*, **102**, 7180 (1980); (b) K. Iijima, K. Tanaka, and S. Onuma, *J. Mol. Struct.*, **246**, 257 (1991).

2. (a) G. Albrecht and R. B. Corey, *J. Am. Chem. Soc.*, **61**, 1087 (1939); (b) R. E. Marsh, *Acta Crystallogr.*, **11**, 654 (1958); (c) P.-G. Jonsson and A. Kvick, *Acta Crystallogr. Sect. B*, **28**, 1827 (1972); (d) W. Huckel, *Theoretical Principles of Organic Chemistry*, Vol. II, Elsevier, New York, 1958, p. 150.

3. (a) J. H. Jensen and M. S. Gordon, *J. Am. Chem. Soc.*, **113**, 7917 (1991); (b) D. Yu, A. Rauk, and D. A. Armstrong, *J. Am. Chem. Soc.*, **117**, 1789 (1995); (c) M. Ramek, V. K. W. Cheng, R. F. Frey, S. Q. Newton, and L. Schaefer, *J. Mol. Struct. (Theochem)*, **235**, 1 (1991); (d) R. F. Frey, J. Coffin, S. Q. Newton, M. Ramek, V. K. W. Cheng, F. A. Momany, and L. Schaefer, *J. Am. Chem. Soc.*, **114**, 9369 (1992); (e) A. G. Csaszar, *J. Am. Chem. Soc.*, **114**, 9568 (1992); (f) C. Hu, M. Shen, and H. F. Schaefer III, *J. Am. Chem. Soc.*, **115**, 2923 (1993).

4. Y. Ding and K. Krogh-Jespersen, *Chem. Phys. Lett.*, **119**, 261 (1992).

5. D. Yu, D. A. Armstrong, and A. Rauk, *Can. J. Chem.*, **70**, 1792 (1992).

6. (a) J. A. C. Rullmann and P. Th. van Duijnen, *Mol. Phys.*, **63**, 451 (1988); (b) E. Clementi, F. Cavallone, and R. Scordamaglia, *J. Am. Chem. Soc.*, **99**, 5531 (1977); (c) L. Carozzo, G. Corogiu, C. Petrongolo, and E. Clementi, *J. Chem. Phys.*, **68**, 787 (1978); (d) S. Ramano and E. Clementi, *Intern. J. Quantum Chem.*, **14**, 839 (1978).

7. (a) J. S. Alper, H. Dothe, and D. F. Coker, *Chem. Phys.*, **153**, 51 (1991); (b) J. S. Alper, H. Dothe, and M. A. Lowe, *Chem. Phys.*, **161**, 199 (1992).

8. (a) R. Bonaccorsi, P. Palla, and J. Tomasi, *J. Am. Chem. Soc.*, **106**, 1945 (1984); (b) O. Kikuchi, T. Natsui, and T. Kozaki, *J. Mol. Struct. (Theochem)*, **207**, 103 (1990).

9. S. U. Kokpol, P. B. Doungdee, S. V. Hannongbua, B. M. Rode, and J. Limtrakul, *J. Chem. Soc., Faraday Trans.*, **84**, 1789 (1988).

10. J. Langlet, J. Caillet, E. Evleth, and E. Kassab, In *Studies in Physical and Theoretical Chemistry*, Vol. 71, J.-L. Rivail, Ed., Elsevier Amsterdam, 1990, p. 345.

11. W. Forner, P. Otto, J. Bernhardt, and J. Ladik, *Theoret. Chim. Acta (Berl.)*, **60**, 269 (1981).

12. H. S. Rzepa and M. Yi, *J. Chem. Soc., Perkin Trans.*, 531 (1991).

13. H. Basch and W. J. Stevens, *Chem. Phys. Lett.*, **169**, 275 (1990).

14. K. Furic, V. Mohacek, M. Bonifacic, and I. Stefanic, *J. Mol. Struct. (Theochem)*, **267**, 39 (1992).

15. M. J. Frisch, G. W. Trucks, M. Head-Gordon, P. M. W. Gill, M. W. Wong, J. B. Foresman, B. G. Johnson, H. B. Schlegel, M. A. Robb, E. S. Replogle, R. Gomperts, J. L. Andres, K. Raghavachari, J. S. Binkley, C. Gonzalez, R. L. Martin, D. J. Fox, D. J. Defrees, J. Baker, J. J. P. Stewart, and J. A. Pople, Gaussian 92, Revision C.4, Gaussian Inc. Pittsburgh, PA, 1992.

16. (a) W. J. Hehre, L. Radom, J. A. Pople, and P. v. R. Schleyer, *Ab Initio Molecular Orbital Theory*, Wiley, New York, 1986; (b) S. F. Boys and F. Bernardi, *Mol. Phys.*, **19**, 553 (1970).

17. C. L. Perrin and B. H. Yang, In *Water–Biomolecule Interactions*, M. U. Palma, M. B. Palma-Vitorelli, and F. Parak, Eds., Italian Physical Society Conference Proceedings **43**, 622 (1993). We thank a referee for pointing out this reference.

18. (a) L. Onsager, *J. Am. Chem. Soc.*, **58**, 1486 (1936); (b) M. W. Wong, M. J. Frisch, and K. B. Wiberg, *J. Am. Chem. Soc.*, **113**, 4776 (1991).

19. P. Pulay, G. Fogarasi, F. Pang, and J. E. Boggs, *J. Am. Chem. Soc.*, **101**, 2550 (1979).

20. (a) M. Takeda, R. E. S. Iavazzo, D. Garfinkel, I. H. Scheinberg, and J. T. Edsall, *J. Am. Chem. Soc.*, **80**, 3813 (1958); (b) S. A. S. Ghazanfar, D. V. Mayers, and J. T. Edsall, *J. Am. Chem. Soc.*, **86**, 3439 (1964); (c) M. Tsuboi, T. Onishi, I. Nakagawa, T. Shimanonchi, and S. Mizushima, *Spectrochim Acta*, **12**, 253 (1958).