J|A|C|S
ARTICLES

Incremental Solvation of Nonionized and Zwitterionic Glycine

Christine M. Aikens[§] and Mark S. Gordon*

Contribution from the Department of Chemistry, Iowa State University, Ames, Iowa 50011

Received May 3, 2006; E-mail: mark@si.fi.ameslab.gov

Abstract: Microsolvation and combined microsolvation−continuum approaches are employed in order to examine the structures and relative energies of nonionized (N) and zwitterionic (Z) glycine clusters. Bridging structures are predicted to be the global minima after 3−5 discrete water molecules are included in the calculations. Calculations incorporating electron correlation stabilize the zwitterionic structures by about 7−9 kcal/mol relative to the N structures regardless of the number of discrete water molecules considered. Continuum calculations stabilize the Z structures relative to N structures; this effect decreases as the number of discrete water molecules is increased. Eight water molecules do not appear to fully solvate glycine.

Introduction

Interactions between water molecules and amino acids and the effects of hydrogen bonding between them are of significant interest, particularly with regard to structures and vibrational spectra. Depending on the pH, amino acids such as glycine may be neutral, anionic, or cationic. Neutral amino acids exist predominantly in their nonionized form (N) in the gas phase, while in solution or in crystals they occur in the zwitterionic form (Z).[1,2] The relative stability of these two forms of glycine is of increasing interest. The intermolecular hydrogen-bonding interaction leads to a considerably larger stability of the Z form compared to the N form; consequently, the experimental free energy and enthalpy for the process Z(aq) → N(aq) are 7.3 and 10.3 kcal/mol, respectively.[3]

Neutral complexes of glycine and water molecules may be isolated in a matrix environment. Related ionic complexes may be segregated using mass spectrometric techniques. Xu et al. suggest that the formation of a zwitterionic species occurs for $[Gly(H_2O)_n]^-$ with $n \geq 5$ based on mass spectrometry and photoelectron spectroscopy data.[4] Diken et al. were later able to produce the bare glycine anion and the smaller hydrates $Gly^-{\cdot}(H_2O)_{1-2}$, and they report photoelectron spectra that are consistent with a nonionized glycine core.[5] A number of experimental studies have examined the formation of the zwitterion of other amino acids. Some of the earliest work in this area was performed by Peteanu and Levy, who investigated a resonant two-photon ionization spectrum of tryptophan.[6] Shortly thereafter, Sulkes et al. used laser-induced fluorescence to examine the excitation spectrum of the tryptophan−single water complex, and determined that they were observing a slightly different complex than the one examined by Peteanu and Levy.[7] More recently, Simons et al. investigated hydrated complexes of tryptophan using a combination of calculation, ion dip infrared spectroscopy, UV hole burning, and resonant two-photon ionization (R2PI).[8−10] Other studies by Williams et al. focused on the examination of hydrated lithium-cationized and sodium-cationized glutamine and valine using blackbody infrared radiative dissociation (BIRD) in addition to computation.[11−13]

Recently, Nonose et al. used electrospray ionization to study the structures and incremental binding energies of glycine and tryptophan.[14] Kresin et al. investigated the hydroxyl loss fractions for glycine and tryptophan molecules picked up by water clusters $(H_2O)_n$ or $(D_2O)_n$ ($n \leq 15$) in a supersonic expansion and analyzed by electron bombardment mass spectrometry.[15] Other studies have focused on the role of hydrogen-bonded bridges in the conformational preferences of biomolecules, such as the work by Zwier et al. who studied the IR and UV spectroscopy of tryptamine and indole compounds.[16−18]

An accurate treatment of solvent effects is increasingly important in quantum chemistry, since most chemical reactions and biological processes take place in solution and many gas-phase processes involve microsolvation. The two main ap-

---

[§] Current address: Department of Chemistry, Northwestern University, Evanston, IL 60208.

(1) Albrecht, G.; Corey, R. B. *J. Am. Chem. Soc.* **1939**, *61*, 1087.
(2) Jonsson, P. G.; Kvick, A. *Acta Crystallogr., Sect. B* **1972**, *28*, 1827.
(3) Wada, G.; Tamura, E.; Okina, M.; Nakamura, M. *Bull. Chem. Soc. Jpn.* **1982**, *55*, 3064.
(4) Xu, S.; Nilles, J. M.; Bowen, K. H., Jr. *J. Chem. Phys.* **2003**, *119*, 10696.
(5) Diken, E. G.; Hammer, N. I.; Johnson, M. A. *J. Chem. Phys.* **2004**, *120*, 9899.
(6) Peteanu, L. A.; Levy, D. A. *J. Phys. Chem.* **1988**, *92*, 6554.
(7) Teh, C. K.; Sipior, J.; Sulkes, M. *J. Phys. Chem.* **1989**, *93*, 5393.
(8) Carcabal, P.; Kroemer, R. T.; Snoek, L. C.; Simons, J. P.; Bakker, J. M.; Compagnon, I.; Meijer, G.; von Helden, G. *Phys. Chem. Chem. Phys.* **2004**, *6*, 4546.
(9) Snoek, L. C.; Kroemer, R. T.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2002**, *4*, 2130.
(10) Robertson, E. G.; Simons, J. P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 1.
(11) Lemoff, A. S.; Bush, M. F.; Wu, C.-C.; Williams, E. R. *J. Am. Chem. Soc.* **2005**, *127*, 10276.
(12) Lemoff, A. S.; Bush, M. F.; Williams, E. R. *J. Am. Chem. Soc.* **2003**, *125*, 13576.
(13) Jockusch, R. A.; Lemoff, A. S.; Williams, E. R. *J. Am. Chem. Soc.* **2001**, *123*, 12255.
(14) Nonose, S.; Iwaoka, S.; Mori, K.; Shibata, Y.; Fuke, K. *Eur. Phys. J. D.* **2005**, *34*, 315.
(15) Moro, R.; Rabinovitch, R.; Kresin, V. V. *J. Chem. Phys.* **2005**, *123*, 074301.
(16) Carney, J. R.; Dian, B. C.; Florio, G. M.; Zwier, T. S. *J. Am. Chem. Soc.* **2001**, *123*, 5596.
(17) Carney, J. R.; Fedorov, A. V.; Cable, J. R.; Zwier, T. S. *J. Am. Chem. Soc.* **2001**, *105*, 3487.
(18) Carney, J. R.; Zwier, T. S. *J. Phys. Chem. A* **2000**, *104*, 8677.

10.1021/ja062842p CCC: $33.50 © 2006 American Chemical Society

proaches to solvation include dielectric continuum methods and discrete microsolvation methods, and each has its own advantages and disadvantages. Most common solvation models are based on a continuum approach, in which the solvent is treated in bulk as a polarizable medium.[19] These models include the self-consistent reaction field (SCRF) or Onsager cavity model,[20,21] the polarizable continuum model (PCM),[22] the COSMO[23] and GCOSMO[24] models, the solvation model 5 (SM5),[25,26] and other variants of these methods. Continuum methods are fast, relatively simple, and designed to satisfactorily reproduce bulk properties of the solvent. However, they can be very sensitive to the choice of the parameters such as the shape of the cavity in which the solute is embedded. Their description of important electronic effects is generally not adequate. In addition, they cannot describe the individual interactions between solute and solvent molecules, which may be particularly important when the solvent and solute can form hydrogen bonds.

In a discrete approach, individual solvent molecules are treated explicitly. This successfully accounts for the intermolecular solute−solvent interactions. However, if ab initio solvent molecules are used, the calculation may be very computationally demanding. A number of quantum mechanical/molecular mechanical (QM/MM) methods have been developed in order to reduce the computational cost associated with discrete solvation approaches.[27] A generalized potential derived from first principles known as the effective fragment potential (EFP) method has been developed by Stevens, Jensen, Gordon, and coworkers.[28,29] The original method (EFP1/HF) was designed specifically for water and is represented by a set of one-electron potentials that is added to the ab initio electronic Hamiltonian.[28,29] The EFP1/HF method contains three energy terms, including Coulombic interactions, induction/polarization interactions, and charge-transfer/exchange−repulsion terms between solvent molecules and between solvent and solute molecules. Charge penetration is included to correct for the pointwise nature of the electrostatic expansion. The first two terms are determined from ab initio calculations on the water monomer, and the third term is fit to a quantum mechanical water dimer potential, so the method can be systematically improved. The EFP1/HF method has been found to closely reproduce Hartree−Fock results for a variety of systems. Other effective fragment potential methods have been developed, including a density functional theory-based model (EFP1/DFT) that includes some short-range correlation effects,[30] and an EFP1/MP2 model that incorporates dispersion effects and a second-order correction to polarization.[31] A general effective fragment potential model (EFP2) has also been developed.[32−35] This method is applicable to any solvent and may be applied across covalent bonds. No fitting is used in this procedure, and the existing model includes terms for exchange repulsion, electrostatic interaction and charge penetration, polarization, dispersion,[36] and charge transfer.[37]

As the number of discrete solvent molecules increases, the number of degrees of freedom in the system also increases. Since numerous minima exist in the configuration space, methods such as molecular dynamics or Monte Carlo simulations must be employed in order to extensively sample this space. Some three-layer approaches have been developed that combine the microsolvation and continuum approaches.[38−44] These methods may reduce the number of explicit solvent molecules required (and hence the computational cost) and may allow for an accurate description of long-range interactions.

As the smallest amino acid, glycine has been the object of many theoretical investigations. The N structure for glycine has three internal rotational degrees of freedom: the rotation of the amino group around the C−N bond, the rotation about the C−C bond, and the rotation of the hydroxyl group around the C−O bond. Initial studies examined various C_s and C_1 stationary points on the conformational potential energy surface of N gas-phase glycine (refs 45, 46, and references therein). Continuum solvent models have been employed in several studies. Bonaccorsi et al. showed that electrostatic solute−solvent interactions stabilize Z glycine with respect to the N form using continuum-only calculations.[47] Tortonda et al. found that the combination of correlation energy and continuum calculations makes the Z form lower in energy than the N form.[48] In addition, they found that solvent effects change the relative stabilities of the gas-phase conformers and that the most stable conformer in solution is well suited for intramolecular proton transfer.[48]

Previous theoretical studies in the gas phase predicted that the zwitterion is not a minimum if basis sets including polarization functions on the hydrogen atoms are used.[49] A complex of Z glycine with one discrete water molecule is a minimum at the RHF level of theory but is not a minimum at the MP2 level of theory.[50,51] Two discrete water molecules

---

(19) Cramer, C. J.; Truhlar, D. G. In *Reviews in Computational Chemistry*; Boyd, D. B., Lipkowitz, K. B., Eds.; VCH: New York, 1995; Vol. 6.
(20) Kirkwood, J. G. *J. Chem. Phys.* **1934**, *2*, 351.
(21) Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486.
(22) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, *55*, 117.
(23) Klamt, A.; Schüürmann, G. *J. Chem. Soc., Perkin Trans.* **1993**, *II*, 799.
(24) Truong, T. N.; Stefanovich, E. V. *Chem. Phys. Lett.* **1995**, *240*, 253.
(25) Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *Chem. Phys. Lett.* **1998**, *288*, 293.
(26) Hawkins, G. D.; Zhu, T.; Li, J.; Chambers, C. C.; Giesen, D. J.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. In *Combined Quantum Mechanical and Molecular Mechanical Methods*, Gao, J.; Thompson, M. A., Eds. American Chemical Society: Washington, D.C., 1998; Vol. 712, p 201.
(27) Gao, J. In *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH: New York, 1996; Vol. 7.
(28) Jensen, J. H.; Day, P. N.; Gordon, M. S.; Basch, H.; Cohen, D.; Garmer, D. R.; Krauss, M.; Stevens, W. J. Effective Fragment Method for Modeling Intermolecular Hydrogen-Bonding Effects on Quantum Mechanical Calculations. In *Modeling the Hydrogen Bond*; Smith, D. A., Ed.; ACS Symposium Series 569; American Chemical Society: Washington D.C., 1994; p 139.
(29) Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D. R.; Basch, H.; Cohen, D. *J. Chem. Phys.* **1968**, *1996*, 105.
(30) Adamovic, I.; Freitag, M. A.; Gordon, M. S. *J. Chem. Phys.* **2003**, *118*, 6725.
(31) Song, J.; Gordon, M. S. Manuscript in preparation.
(32) Jensen, J. H. *J. Chem. Phys.* **1996**, *104*, 7795.
(33) Jensen, J. H.; Gordon, M. S. *Mol. Phys.* **1996**, *89*, 1313.
(34) Jensen, J. H.; Gordon, M. S. *J. Chem. Phys.* **1998**, *108*, 4772.
(35) Gordon, M. S.; Freitag, M. A.; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293.
(36) Adamovic, I.; Gordon, M. S. *Mol. Phys.* **2005**, *103*, 379.
(37) Li, H.; Gordon, M. S.; Jensen, J. H. *J. Chem. Phys.* **2006**, *124*, 214108.
(38) Russell, S. T.; Warshel, A. *J. Mol. Biol.* **1985**, *185*, 389.
(39) Rzepa, H. S.; Yi, M. *J. Chem. Soc., Perkin Trans.* **1991**, *2*, 531.
(40) De Vries, A. H.; Van Duijnen, P. T.; Juffer, A. H.; Rullmann, J. A. C.; Dijkman, J. P.; Merenga, H.; Thole, B. T. *J. Comp. Chem.* **1995**, *16*, 37.
(41) Bandyopadhyay, P.; Gordon, M. S. *J. Chem. Phys.* **2000**, *113*, 1104.
(42) Fernández-Ramos, A.; Smedarchina, Z.; Siebrand, W.; Zgierski, M. Z. *J. Chem. Phys.* **2000**, *113*, 9714.
(43) Bandyopadhyay, P.; Gordon, M. S.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **2002**, *116*, 5023.
(44) Cui, Q. *J. Chem. Phys.* **2002**, *117*, 4720.
(45) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1991**, *113*, 7917.
(46) Hu, C.-H.; Shen, M.; Schaefer, H. F., III. *J. Am. Chem. Soc.* **1993**, *115*, 2923.
(47) Bonaccorsi, R.; Palla, P.; Tomasi, J. *J. Am. Chem. Soc.* **1945**, *1984*, 106.
(48) Tortonda, F. R.; Pascual-Ahuir, J. L.; Silla, E.; Tuñón, I. *Chem. Phys. Lett.* **1996**, *260*, 21.
(49) Ding, Y.; Krogh-Jespersen, K. *Chem. Phys. Lett.* **1992**, *199*, 261.
(50) Jensen, J. H.; Gordon, M. S. *J. Am. Chem. Soc.* **1995**, *117*, 8159.
(51) Ding, Y.; Krogh-Jespersen, K. *J. Comp. Chem.* **1996**, *17*, 338.

are required in order for the zwitterion to be a minimum at the MP2 level of theory.[50] Low-energy structures for the zwitterion with one or two water molecules have also been of interest to other researchers.[52,53] In addition, low-energy structures have been examined for the N form with one or two water molecules.[50,52,54−56]

The lowest-energy Z structure with three waters was first reported by Kassab et al. in 2000 using DFT.[57] This group also reported one N structure relating to proton transfer. Bandyopadhyay et al. examined the energy difference between this structure and the global minima for Z and N species using ab initio RHF and MP2 calculations, EFP1/HF potentials, and PCM.[43] Adamovic et al. extended this work to consider the effects of DFT and EFP1/DFT potentials.[30] Other Z and N trihydrated structures were examined by Ramaekers et al.[56]

A few other studies have explored supermolecular complexes with more than three discrete water molecules, but many of these did not fully sample the configuration space. Ramaekers et al. examined N and Z structures with four and five water molecules.[56] Yamabe et al. examined intermolecular proton-transfer pathways for glycine$(H_2O)_n$ ($n = 1-4$, 11, 17) using density functional theory (DFT) and DFT+Onsager methods.[58] They considered structures based on three conformers (N3, N5, and N8 defined below), which included the lowest-energy structure in solution but did not include the lowest-energy gas-phase structure. Only bridged structures were examined, and no configurational sampling was reported.

Fernández-Ramos et al. examined the proton-transfer reaction for a six-water complex in an Onsager dielectric continuum.[42] They started with several arrangements of the six-water complex and used the lowest-energy configuration they found upon optimization at the HF/6-31G(d) level.

Rzepa and Yi found starting structures for complexes with 7 and 15 waters using Monte Carlo searches with a molecular mechanics force field, and then optimized these structures with semiempirical models.[39] They found that the Z form is more stable than the N form for the glycine$(H_2O)_{15}$ supermolecular complex and for the glycine$(H_2O)_7$ complex in a self-consistent reaction field.

Bandyopadhyay et al. considered the glycine$(H_2O)_8$ complex using three-layer models that combined the effective fragment potential with the Onsager and PCM models.[41,43] In the first study, eight water molecules in the solvation shell were selected from a molecular dynamics calculation, and this structure was optimized.[41] In the second study, MC with local minimization was used in order to find a low-energy configuration.[43] Cui used an approach similar to the former in order to examine the effects of TIP3P waters.[44] Using the three-layer approach, the zwitterion is predicted to be more stable than the nonionized species when a reliable continuum method is used.[41,43,44]

Theoretically predicted spectra may be useful in the analysis of experimentally determined spectra of small hydrogen-bonded complexes such as the glycine−water complex. Theoretical methods have previously been applied to Raman[42] and IR[56] spectra, and anharmonic effects on the spectrum of the glycine-(H_2O) complex have also been considered.[59]

The goal of the present work is to determine the equilibrium structures for small glycine−water clusters and to determine which conformers dominate as the size of the cluster increases. In addition, the present work will determine at what point the zwitterionic form is predicted to be more stable than the nonionized form, and how the inclusion of correlation energy affects this result. The convergence of the (Z − N) energy difference will be considered. The effect of adding a bulk solvent will be discussed, as well as how this is affected by increasing the number of discrete water molecules. The number of waters needed to completely solvate glycine will be considered in an effort to determine the number of waters in the first solvent shell. Previous studies have investigated a first solvation shell consisting of 5,[60] 6,[39,42,61] 8,[41,44] or 11[58] water molecules.

Computational Methods

Molecular structures for water and eight nonionized glycine minima were determined at the restricted Hartree−Fock (RHF) level of theory using the 6-31G(d,p) basis set. Four of the nonionized glycine structures have C_s symmetry, and four structures have C_1 symmetry (see ref 45). MP2/6-31++G(d,p) single-point energies were used to refine the relative energies. For zwitterionic glycine, the polarizable continuum model (PCM) was employed with a geometry optimization at the RHF/6-31G(d,p) level to calculate a minimum energy structure. General effective fragment potentials (EFP2s) that include exchange repulsion, induction, electrostatic interaction, and charge penetration effects were constructed for water, the eight nonionized minima, the zwitterionic minimum, and three other zwitterionic structures constructed by introducing torsion around the C−C bond.

For 1−3 discrete water molecules, a Monte Carlo (MC) simulation with local minimization[62] “basin-hopping” approach was used to find minima on the potential energy surface. Starting structures for MC runs were constructed by adding a water molecule in various positions relative to structures determined from previous runs with one fewer water molecule. For complexes involving one water molecule, 30 steps were allowed in each local optimization. For two and three water molecules, 60 steps were allowed. General effective fragment potentials (EFP2s) were used for glycine and water in order to reduce the computational cost. Structures obtained from the MC runs were used as a starting point for geometry optimizations with glycine treated at the RHF/6-31++G(d,p) level of theory and water molecules treated with the EFP1/HF potential (this will be referred to as an efp−ab optimization). These structures were then fully optimized at the RHF/6-31++G(d,p) level of theory (this will be referred to as an ab−ab optimization). At each step, duplicate minima were removed. As a test of this two-part optimization, full ab initio RHF/6-31++G(d,p) optimizations were performed on selected structures from the MC runs with one discrete water molecule. Single-point calculations incorporating the conductor-like polarizable continuum model (C-PCM)[63] were performed on the efp−ab and ab−ab optimized structures in order to calculate the free energies of solvation for these structures. For one and two water molecules, optimizations incorporating C-PCM were also run. Single-point MP2 calculations were performed on the ab−ab optimized structures. In addition, single-point MP2 energies were

---

[52] Kwon, O. Y.; Kim, S. Y.; No, K. T. *Bull. Korean Chem. Soc.* **1995**, *16*, 410.

[53] Wang, W.; Pu, X.; Zheng, W.; Wong, N.-B.; Tian, A. *J. Mol. Struct. (THEOCHEM)* **2003**, *626*, 127.

[54] Basch, H.; Stevens, W. J. *Chem. Phys. Lett.* **1990**, *169*, 275.

[55] Wang, W.; Zheng, W.; Pu, X.; Wong, N.-B.; Tian, A. *J. Mol. Struct. (THEOCHEM)* **2002**, *618*, 235.

[56] Ramaekers, R.; Pajak, J.; Lambie, B.; Maes, G. *J. Chem. Phys.* **2004**, *120*, 4182.

[57] Kassab, E.; Langlet, J.; Evleth, E.; Akacem, Y. *J. Mol. Struct. (THEOCHEM)* **2000**, *531*, 267.

[58] Yamabe, S.; Ono, N.; Tsuchida, N. *J. Phys. Chem. A* **2003**, *107*, 7915.

[59] Chaban, G. M.; Gerber, R. B. *J. Chem. Phys.* **2001**, *115*, 1340.

[60] Kokpol, U.; Doungdee, P. B.; Hannongbua, S. V.; Rode, B. M. *J. Chem. Soc., Faraday Trans. 2* **1988**, *84*, 1789.

[61] Förner, W.; Otto, P.; Bernhart, J.; Ladik, J. *Theor. Chim. Acta* **1981**, *60*, 269.

[62] Li, Z.; Scheraga, H. A. *Proc. Nat. Acad. Sci. U.S.A.* **1987**, *84*, 6611.

[63] Barone, V.; Cossi, M. *J. Phys. Chem. A* **1995**, *1998*, 102.

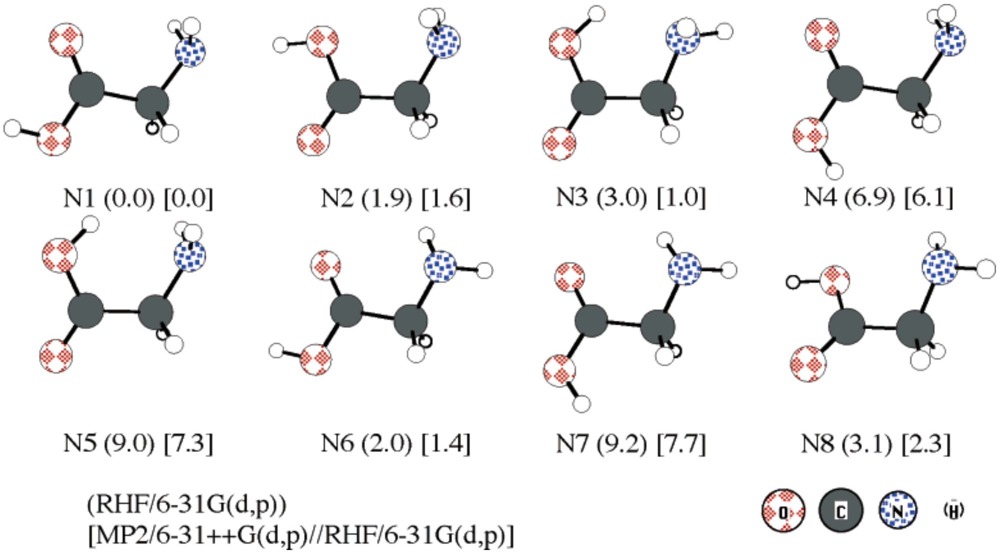


Figure 1. N glycine conformer structures and relative energies (kcal/mol) at the RHF/6-31G(d,p) and MP2/6-31++G(d,p)//RHF/6-31G(d,p) levels of theory.

calculated using C-PCM orbitals and orbital energies to integrate correlation effects with bulk solvent effects.[64−67] MP2 and B3LYP geometry optimizations were performed for selected ab−ab structures. Imaginary frequencies were detected by diagonalizing the energy second-derivative (Hessian) matrix. Relative energies in this work are not corrected for zero-point vibrational energy. Initial calculations indicate that these corrections have negligible effects on the relative energies of nonionized and zwitterionic structures, although they may destabilize the zwitterionic structures by up to 1 kcal/mol with respect to the nonionized structures.

For four discrete water molecules, a simulated annealing Monte Carlo with local minimization approach was used to find minima on the potential energy surface.[68] Ab initio atoms were not allowed to move during the MC steps but were allowed to move during the minimization procedure. Starting structures were generated by adding a water molecule to unique structures with three water molecules. The water molecules were represented by the EFP1/HF potential, and the RHF/6-31++G(d,p) level of theory was used for glycine in order to allow the molecule to sample torsional degrees of freedom. The initial and final temperatures used in the simulation were 10000 K and 200 K, respectively. Fifty temperatures were used in the simulation, and 300 geometries were evaluated at each temperature. Local optimizations were carried out after 150 MC steps, with a maximum of 150 optimization steps. The maximum step size was held constant and the ab initio atoms were not allowed to move during the MC steps. All structures within 3.0 kcal/mol of the global N and Z minima were subjected to a full RHF/6-31++G(d,p) geometry optimization. Single-point MP2 and C-PCM calculations were performed on the unique ab−ab minima.

For five or more discrete waters, a similar simulated annealing MC with local minimization approach was employed. The initial temperature was set at 5000 K. Two hundred temperatures were used in the simulation, and 200 geometries were evaluated at each temperature. Ten translational steps were permitted in each block. Three fragments were allowed to move during each MC step. The box dimensions for periodic boundary conditions were slightly increased to 23 × 19.5 × 20. Local optimizations were performed after 100 MC steps, with a maximum of 200 steps in each. Full RHF/6-31++G(d,p) geometry optimizations were performed on a few lowest-energy N and Z structures. Single-point MP2 and C-PCM calculations were performed on the ab−ab minima.

All calculations were performed with the electronic structure code GAMESS,[69] which is freely available from Iowa State University at www.msg.ameslab.gov. Structures were visualized with MacMolPlt,[70] a graphical interface to GAMESS.

Results and Discussion

Glycine. The eight nonionized glycine conformers determined in the geometry optimizations correspond to the eight minima found in ref 45. These gas-phase structures and their relative energies are shown in Figure 1.

Glycine(H_2O). A single water molecule can interact with each glycine conformer to form a wide variety of structures. It may interact with the carboxylic acid group, the amino group, or the saturated carbon, or it may act as a bridge between the carboxylic acid and amino groups. Sixty-six (25) distinct structures were observed for N (Z) EFP2 glycine interacting with a single EFP2 water molecule. Upon efp−ab optimization, 47 unique N structures survive, ranging from 0.0 to 13.7 kcal/mol relative to the lowest-energy structure, and there are 6 unique Z structures, ranging from 24.2 to 28.2 kcal/mol above the lowest-energy N glycine minimum (see Table 1 and Figures 2 and 3). The reduction in the number of isomers may be due in part to relaxation from the internally frozen EFP2 geometries. After ab−ab optimization, 44 unique N structures and 5 unique Z structures remain. These unique structures are named according to the following convention. The name of each structure has the form *m*N*n*-alpha or *m*Z-alpha, where *m* designates the number of water molecules surrounding the glycine core, N*n* represents the *n*th nonionized glycine conformer, Z indicates the zwitterion, and “alpha” is an alphabetical letter used to label

---

(64) Olivares del Valle, F. J.; Tomasi, J. *Chem. Phys.* **1991**, *150*, 139.
(65) Aguilar, M. A.; Olivares. del Valle, F. J.; Tomasi, J. *Chem. Phys.* **1991**, *150*, 151.
(66) Olivares del Valle, F. J.; Aguilar, M. A.; Tolosa, S. *J. Mol. Struct. (THEOCHEM)* **1993**, *279*, 223.
(67) Olivares del Valle, F. J.; Aguilar, M. A. *J. Mol. Struct. (THEOCHEM)* **1993**, *280*, 25.
(68) Kirkpatrick, S.; Gelatt, C. D.; Vecci, M. P. *Science* **1983**, *220*, 671.
(69) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Gordon, M. S.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., Jr. *J. Comp. Chem.* **1993**, *14*, 1347.
(70) Bode, B. M.; Gordon, M. S. *J. Mol. Graphics Model.* **1998**, *16*, 133.

Table 1. Structures and Relative Energies (in kcal/mol) for Glycine(H_2O)[a] at Various Levels of Theory

structure	refs	EFP1/HF and RHF/6-31++G(d,p)	full RHF/ 6-31++G(d,p)	MP2//RHF/ 6-31++G(d,p)	PCM+ RHF//RHF/ 6-31++G(d,p)	PCM+ MP2//RHF/ 6-31++G(d,p)	description of water interaction	imaginary frequencies
1N1-a	*b**,*c*,*d*,*e*,*f*,*h*	0.0	0.0	0.0	0.0	0.0	COOH	
1N6-a		1.5	1.5	1.2	1.7	1.4	COOH	
1N2-a	*b**,*c*,*e*	1.5	1.5	1.2	1.3	1.1	COOH	
1N6-b		1.7	1.7	1.4	1.7	1.5	COOH	
1N8-a		2.3	2.4	2.0	2.3	2.0	COOH	
1N8-b		2.5	2.6	2.2	2.4	2.1	COOH	
1N1-b	*d*,*e*	2.6	3.3	3.2	−0.2	0.1	NH_2	
1N1-c	*b*,*c*,*d*,*e*,*f*,*h*	3.1	3.1	3.6	1.6	2.1	bridge between C═O and N—H	
1N6-c	*c*	3.8	4.4	3.8	1.3	1.1	bridge between C═O and N	
1N6-d		4.4	4.5	4.6	2.6	2.9	bridge between C═O and N—H	
1N2-b	*e*	4.8	5.4	5.0	1.2	1.4	NH_2	
1N3-b		5.0	5.0	3.9	1.2	0.8	CH_2	6
1N3-a	*e*	5.0	5.0	3.9	1.2	0.8	CH_2	
1N2-c		5.1	5.2	5.0	4.4	4.2	bridge between C═O and N—H	
1N3-c		5.1	5.0	3.8	1.2	0.8	CH_2	
1N3-d	*c*,*e*	5.4	5.8	4.0	2.9	2.0	bridge between O═H and N—H	
1N2-d	*e*	5.5	5.5	6.2	2.4	3.2	CH_2	
1N2-e		5.5	5.6	6.3	2.3	3.2	CH_2	21
1N8-c		5.8	6.4	6.0	2.0	2.2	NH_2	
1N1-d	*d*,*e*,*h*	5.9	6.1	6.4	2.5	3.1	CH_2	
1N8-d		6.0	6.0	6.5	3.5	4.1	CH_2	
1N3-e	*c*,*e*	6.3	6.1	5.5	0.7	1.0	C═O only	
1N3-f	*b*	6.3	6.2	5.5	0.6	1.0	C═O only	64
1N3-g		6.4	6.3	5.9	0.3	0.9	C═O only	67
1N3-h		6.6	6.9	5.8	1.8	1.9	NH_2	
1N2-f	*e*	7.0	7.2	6.8	4.0	4.1	bridge between O—H and N—H	
1N4-a		7.1	7.7	7.3	1.1	1.5	OH	33
1N6-e		7.1	7.4	8.0	3.7	4.8	NH_2	
1N8-e		7.2	7.4	7.0	4.6	4.6	bridge between O—H and N—H	
1N6-f		7.4	7.6	7.9	3.8	4.5	CH_2	
1N8-f		7.9	8.1	8.6	4.5	5.5	NH_2	
1N7-a		8.8	9.4	8.7	3.1	3.3	OH	
1N5-a		9.1	10.0	9.1	3.3	3.4	bridge between O—H and N—H	14
1N4-b		9.2	9.7	8.9	2.8	3.0	NH_2	
1N4-c		9.4	9.3	9.0	4.1	4.5	bridge between C═O and N—H	
1N7-b	*c*,*e*	10.3	7.1	4.7	2.7	1.4	bridge between O—H and N	
1N4-e		11.0	10.9	11.5	2.9	4.3	C═O only	
1N4-d		11.0	10.9	11.4	3.2	4.6	C═O only	
1N7-c		11.1	11.2	10.6	5.4	5.6	bridge between C═O and N—H	
1N7-d		11.7	11.2	10.6	5.5	5.7	bridge between C═O and N—H	
1N7-e		12.8	12.7	13.1	4.9	6.2	C═O only	
1N7-f		12.9	12.7	12.9	5.4	6.5	C═O only	
1N5-b		12.9	13.4	12.6	3.7	4.5	NH_2	
1N7-g		13.7	14.0	13.9	6.6	7.6	NH_2	
1Z-a	*c*,*f*,*g*,*i*,*j*	24.2	23.5	15.3	2.4	−3.6	bridge between C═O and N—H	
1Z-b	*c*	25.9	25.9	16.9	4.9	−2.0	bridge between C═O and N—H	
1Z-d	*c*,*j*	27.2	26.7	17.5	4.8	−2.2	CH_2	6
1Z-c	*c*,*g*	27.2	26.7	17.5	4.9	−2.2	CH_2	
1Z-e	*c*,*f*,*g*,*j*	28.2	28.0	19.4	4.5	−1.7	COO^-	

[a] See **Glycine(H_2O)** section in the text for the definitions of the nomenclature. [b] Reference 54. [c] Reference 50. [d] Reference 55. [e] Reference 56. [f] Reference 52. [g] Reference 53. [h] Reference 59. [i] Reference 58. [j] Reference 51. *Related C_s structure.

the structures. The complexes are labeled alphabetically from the lowest-energy structure at the mixed efp−ab level of theory; thus, structures close to "a" are relatively low in energy.

EFP1/HF waters do a good job of reproducing the RHF relative energy ordering, particularly for the lowest-energy structures. In the six lowest-energy structures (1N1-a, 1N2-a,

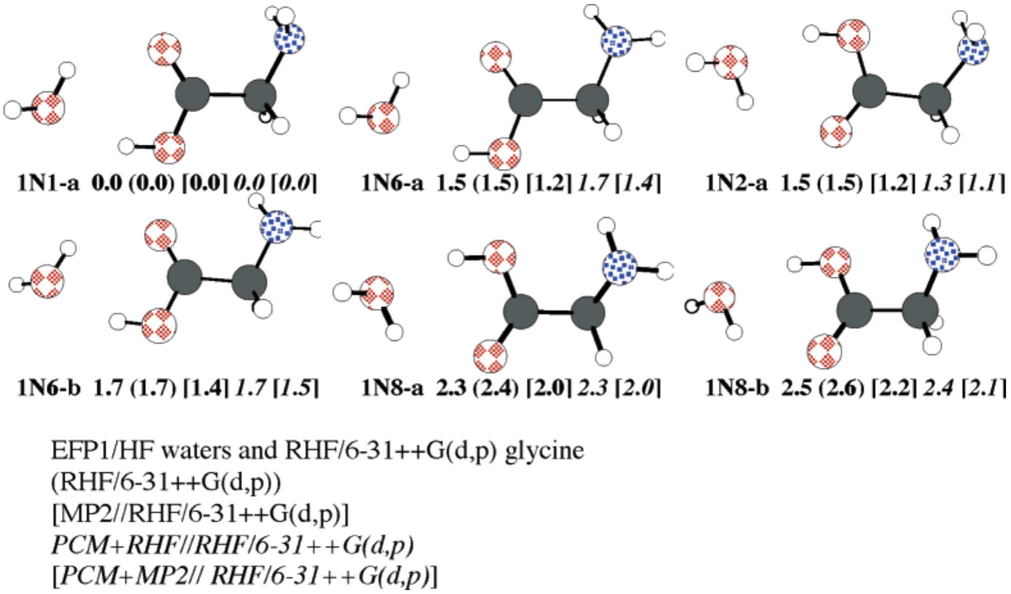


Figure 2. Six lowest-energy conformers and relative energies (kcal/mol) of N glycine(H_2O).

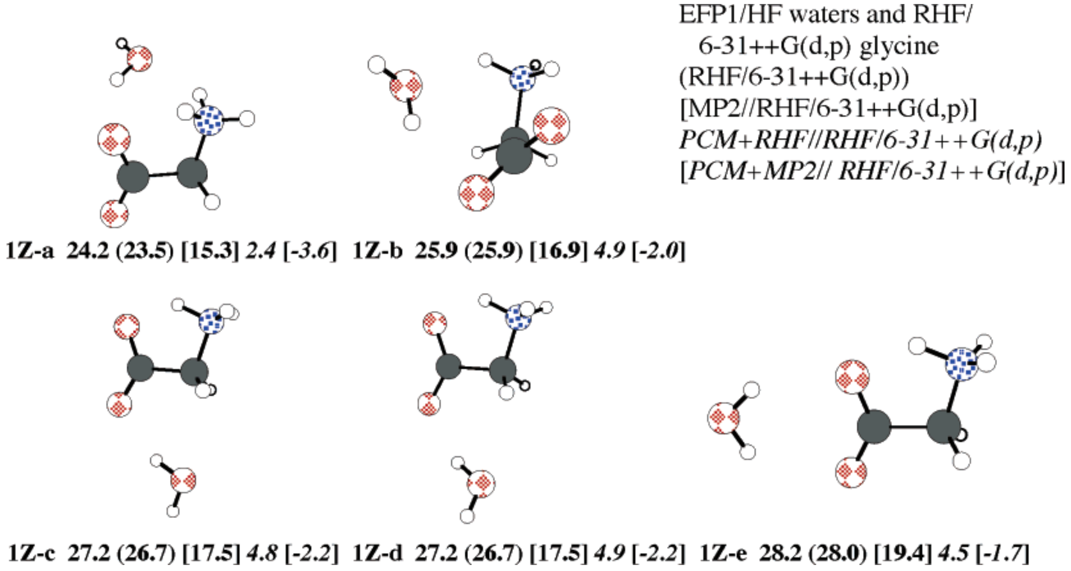


Figure 3. Five lowest-energy conformers and relative energies (kcal/mol) of zwitterionic glycine(H_2O).

1N6-a, 1N6-b, 1N8-a, and 1N8-b) (shown in Figure 2), the water interacts with the carboxylic acid functionality of the glycine molecule. The five zwitterionic structures found in this work (1Z-a, 1Z-b, 1Z-c, 1Z-d, and 1Z-e in Figure 3) are similar to those found in ref 50 (Z1a, Z1b, and i, iv, and iii in Figure 2 of that reference). A frequency analysis shows that the four lowest-energy Z structures are minima on the RHF/6-31++G(d,p) potential energy surface. The fifth structure has an imaginary frequency of 6i cm^{-1}. The two lowest-energy structures have the water molecule bridging the protonated amino group and the carboxylate group. They differ in that the water molecule in 1Z-a bonds to the O that is cis to the N, whereas in 1Z-b, the water is bonded to the O that is trans to the N.

Some of the N structures found here have been previously reported in the literature. Most of these studies were largely concerned with a few conformers. Basch and Stevens examined five structures with C_s symmetry, and as a result a few of their structures are saddle points.[54] Their structures came from four gas-phase conformers. Kwon et al. examined two monohydrated minima that were derived from the most stable gas-phase conformer.[52] Jensen and Gordon found eight N minima from multiple conformers.[50] Wang et al. examined four monohydrated structures based on a single gas-phase conformer.[55] Rameakers et al. determined 16 N structures that were derived from three gas-phase conformers.[56] The present and previous studies agree that cyclic 1N1-a is the most stable structure. However, most 1:1 glycine−water complexes examined in this work have not been reported previously, including four of the five structures that lie within 3.0 kcal/mol of 1N1-a at the RHF/6-31++G(d,p) level of theory. Six of the 47 N structures found in this study have imaginary frequencies ranging from 5−67i cm^{-1}, so they are saddle points at this level of theory. Since the potential energy surface appears to be shallow and other levels of theory may predict slightly different results (as previously reported for gas-phase glycine), these structures are reported here.

Single-point calculations using C-PCM+MP2 with the RHF/6-31++G(d,p) structures simulate the effect of adding a bulk

Table 2. Structures and Relative Energies (in kcal/mol) for Glycine(H_2O)[a] at Various Levels of Theory

structure	refs	EFP1/HF and RHF/6-31++G(d,p)	full RHF/ 6-31++G(d,p)	MP2//RHF/ 6-31++G(d,p)	PCM+ RHF//RHF/ 6-31++G(d,p)	PCM+ MP2//RHF/ 6-31++G(d,p)	description of water interaction
2N1-a	*b*	0.0	0.0	0.0	0.0	0.0	water ring on COOH
2N1-b		0.5	0.4	0.6	0.3	0.6	water ring on COOH
2N2-a	*b*	1.4	1.3	1.1	1.3	1.1	water ring on COOH
2N6-a		1.5	1.5	1.2	1.6	1.3	water ring on COOH
2N6-b		1.6	1.6	1.3	1.7	1.4	water ring on COOH
2N2-b		1.9	1.7	1.7	1.5	1.6	water ring on COOH
2N6-c		1.9	1.8	1.7	1.8	1.8	water ring on COOH
2N8-a		2.3	2.3	2.0	2.4	2.1	water ring on COOH
2N8-b		2.7	2.6	2.5	2.6	2.6	water ring on COOH
2N8-c		3.0	2.4	2.0	2.4	2.1	water ring on COOH
2N1-c	*c*	4.2	4.5	4.8	1.9	2.4	COOH, NH_2
2N1-e	*b*	5.0	4.8	5.6	3.6	4.2	COOH, bridge
2N3-a	*b*	5.1	5.2	3.6	4.5	3.4	two-water bridge between NH and C═O
2N1-f	*c*,*d*	5.1	4.8	5.5	3.6	4.2	COOH, bridge
2N6-e	*b*	5.3	5.6	5.4	3.2	3.1	COOH, NH_2
2N2-d	*b*	6.6	6.6	7.0	5.5	5.7	two water bridge between NH and C═O
2N1-n	*c*	7.1	7.3	8.0	3.5	4.4	bridge, NH_2
2N7-b	*b*	8.8	9.4	7.0	7.4	5.7	bridge, bridge
2Z-a	*b*,*c*	20.2	19.8	10.8	2.4	−4.7	two water bridge between NH and C═O
2Z-b		20.6	20.6	11.3	3.0	−4.4	two water bridge between NH and C═O
2Z-c		20.9	20.6	11.4	3.4	−3.9	two water bridge between NH and C═O
2Z-e		21.3	19.7	11.4	2.8	−3.7	bridge, bridge
2Z-d	*b*	21.3	19.7	11.4	2.7	−3.7	bridge, bridge
2Z-f		23.3	21.9	12.5	5.9	−1.8	bridge, bridge
2Z-g		25.4	24.5	17.6	3.8	−1.2	bridge, NH_2
2Z-h		26.0	24.6	17.6	3.8	−1.2	bridge, NH_2
2Z-i		27.2	27.0	19.2	3.9	0.5	bridge, NH_2
2Z-j		29.3	29.0	20.9	7.3	1.2	water ring on C═O
2Z-k		29.7	28.7	21.2	6.5	0.9	C═O, CH_2

[a] See **Glycine(H_2O)** section in the text for the definitions of the nomenclature. See Table 1S for full glycine(H_2O)$_2$ table. [b] Reference 50. [c] Reference 56. [d] Reference 52.

solvent around the glycine(H_2O) complex. These calculations, which include the effects of electron correlation and the free energy of solvation, predict that the zwitterionic species is 3.6 kcal/mol lower in energy than the nonionized species when one discrete water is considered inside the bulk. The resultant energy lowering of the Z species relative to that of the N species may be partitioned into the effects of electron correlation (MP2) and bulk solvation (PCM). Electron correlation accounts for a relative stabilization of Z structures by about 8−9 kcal/mol, while the inclusion of the bulk effects further stabilizes the Z structures relative to the N structures by about 21−24 kcal/mol.

Glycine(H_2O)$_2$. For glycine with two water molecules, 238 N and 86 Z structures were found in the EFP2 Monte Carlo simulations. After efp−ab optimization, 155 N and 19 Z structures remain. Upon ab−ab optimization, 132 unique N structures and 11 unique Z structures were found. Ten of the N and two of the Z structures were reported in previous studies (see Table 2, 1S, and Figures 4 and 5). As found in earlier studies,[50,56] the water molecules bridge the COO^- and NH_3^+ groups in the lowest-energy zwitterionic structures (Figure 5). In general, a complex with a bridge consisting of two water molecules (such as 2Z-a, 2Z-b, and 2Z-c) is slightly lower in energy than a complex with two single-water bridges (such as 2Z-d and 2Z-e). The (N − Z) energy difference is 20.2 kcal/mol at the efp−ab level of theory, in very good agreement with the RHF energy difference.

Combined C-PCM+MP2 single-point energy calculations predict that the lowest-energy Z species is 4.7 kcal/mol lower in energy than the lowest-energy N species, so the consideration of a second explicit water molecule inside the bulk increases the (Z − N) energy difference by 1.1 kcal/mol. The effects of the C-PCM+MP2 calculations can be analyzed in terms of contributions from electron correlation (7−9 kcal/mol stabilization of Z relative to N) and bulk solvation (16−23 kcal/mol stabilization of Z). The effects of the continuum solvent are smaller when a second discrete water molecule is introduced.

In the lowest-energy N complexes (2N1-a, 2N1-b, 2N2-a, 2N6-a, 2N6-b, 2N2-b, 2N6-c, 2N8-a, 2N8-b, and 2N8-c), the water molecules form a ring with the carboxylic acid end of glycine (Figure 4). This motif was found in an earlier study by Jensen and Gordon.[50] This structure may be visualized with the two oxygen atoms in glycine essentially substituting for two of the oxygen atoms in a cluster of four water molecules. The dihydrated N structure found by Kwon et al. is either 2N1-e or 2N1-f; these structures lie 5.5−5.6 kcal/mol above the global minimum at the MP2// RHF/6-31++G(d,p) level of theory. The lowest-energy structure reported by Rameakers et al. is 2N1-c, which is 4.8 kcal/mol above the global minimum at the MP2//RHF/6-31++G(d,p) level of theory. At the B3LYP/6-31++G(d,p) level of theory, 2N1-c is 5.2 kcal/mol higher in energy than 2N1-a.

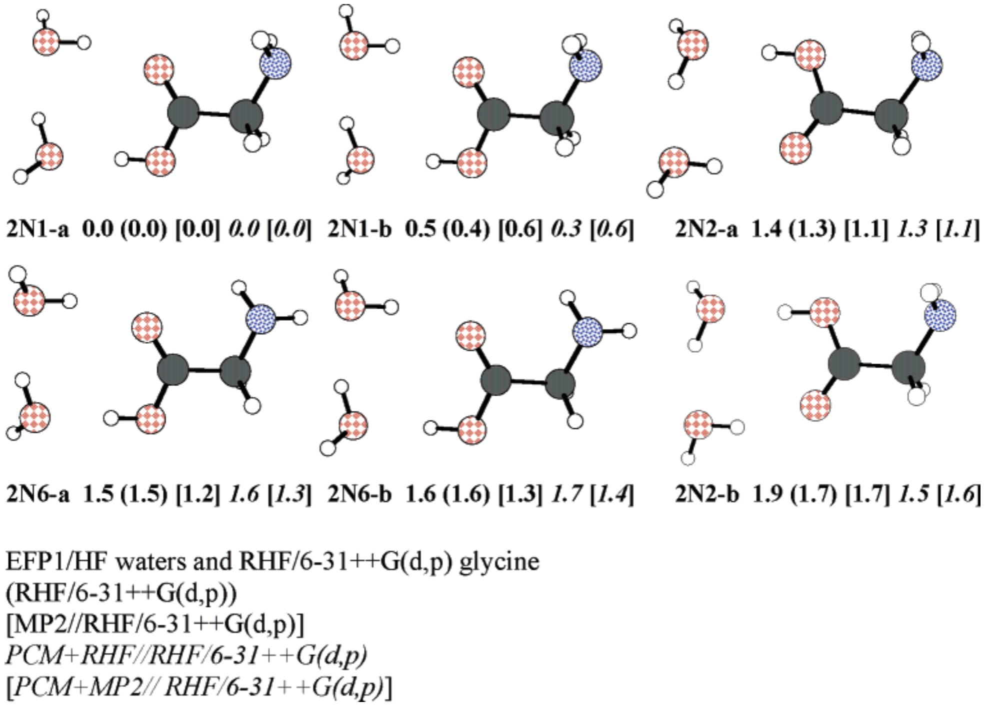


Figure 4. Six lowest-energy conformers and relative energies (kcal/mol) of nonionized glycine$(H_2O)_2$.

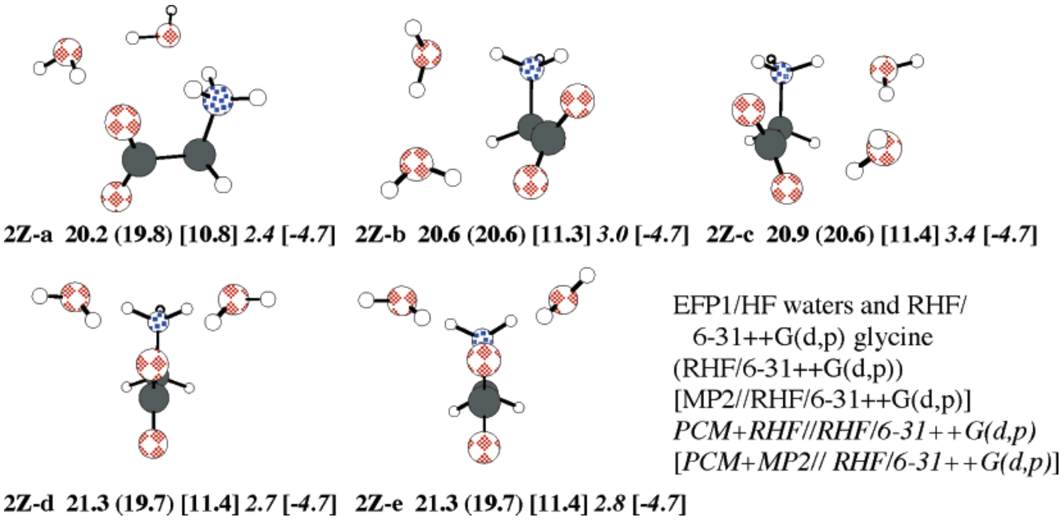


Figure 5. Five lowest-energy conformers and relative energies (kcal/mol) of zwitterionic glycine$(H_2O)_2$.

Glycine$(H_2O)_3$. For the glycine complex with three water molecules, 824 N and 212 Z structures were found in the EFP2 Monte Carlo calculations. After optimization with ab initio glycine and EFP1/HF waters, 450 N and 42 Z complexes were found. Full ab initio optimizations lead to 349 N and 24 Z structures. Twenty-four of the N structures are predicted to lie within 3.0 kcal/mol of the global minimum (Table 3, 2S). The lowest-energy structures fall into two groups. The global minimum structure 1N1-a (see Figure 6) consists of three water molecules interacting with the carboxylic acid end of glycine in a ring similar to that of the water pentamer. This global minimum was previously reported by Bandyopadhyay and Gordon.[43] The other primary motif seen in the low-energy structures is a chain of three waters that bridges the carboxylic acid and amino groups (see 3N6-a and 3N6-c in Figure 6). A similar cyclic bridging structure was reported by Yamabe et al.[58] This type of structure is stabilized by 2.0−3.0 kcal/mol relative to the ring structures when electron correlation is included via MP2 single-point energies. Optimization at the MP2 level of theory further stabilizes the cyclic structures. At the MP2/6-31++G(d,p) level of theory, 3N6-c is predicted to lie 0.8 kcal/mol lower in energy than 3N1-a, so it is predicted to be the global minimum

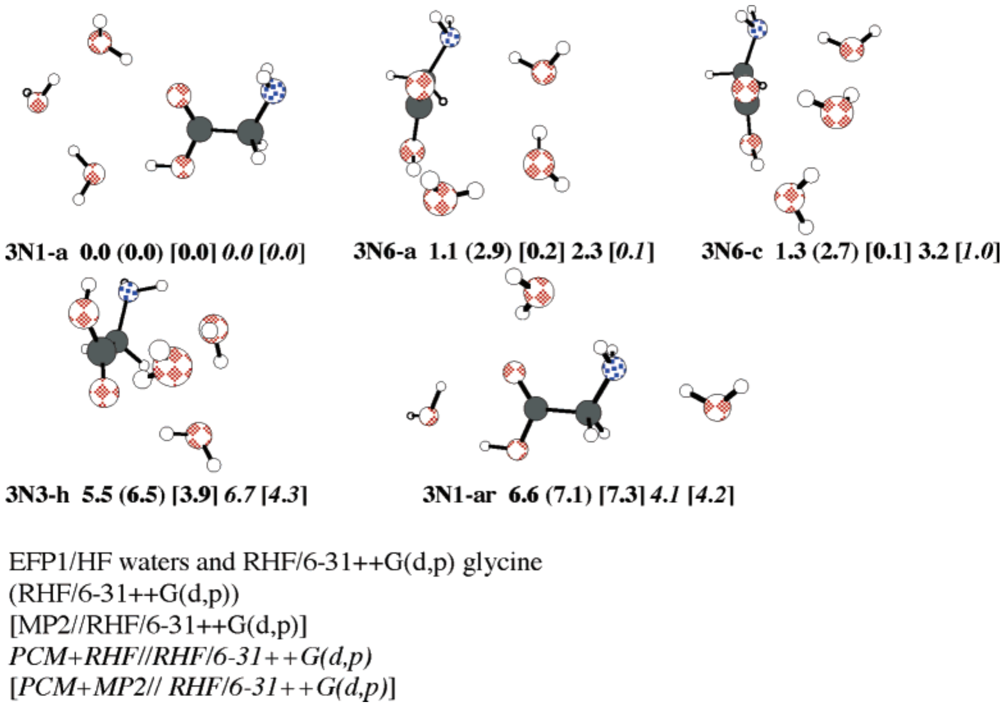
The relative energies for the cyclic and the ring structures with ab initio glycine and EFP1/HF waters are somewhat different than the relative energies from full ab initio calculations. Interestingly, the EFP1/HF water relative energies are closer to the MP2 single-point values. For most other structures, the EFP and RHF relative energies agree to within 1.0 kcal/mol.

Two other previous studies have examined minima for complexes with N glycine and three water molecules. Kassab et al. were interested in the minimum that resulted from intramolecular proton transfer rather than the global minimum.[57]

Table 3. Structures and Relative Energies (in kcal/mol) for Glycine$(H_2O)_3$[a]

structure	refs	EFP1/HF and RHF/6-31++G(d,p)	full RHF/ 6-31++G(d,p)	MP2//RHF/ 6-31++G(d,p)	PCM+RHF//RHF/ 6-31++G(d,p)	PCM+MP2//RHF/ 6-31++G(d,p)
3N1-a	*d*	0.0	0.0	0.0	0.0	0.0
3N1-b		0.4	1.3	1.0	1.9	1.6
3N6-a		1.1	2.9	0.2	2.3	0.1
3N6-b		1.3	2.6	0.1	3.2	1.0
3N8-a		1.3	3.1	0.3	2.0	−0.1
3N6-c		1.3	2.7	0.1	3.2	1.0
3N6-d		1.4	2.9	0.9	3.7	2.0
3N6-e		1.5	3.2	0.5	2.9	0.8
3N2-a		1.5	2.4	1.9	3.0	2.5
3N8-b		1.7	3.5	0.5	3.2	0.7
3N6-f		1.8	2.7	2.1	3.4	2.9
3N1-c		1.9	2.7	2.4	0.5	0.3
3N8-c		2.0	3.7	0.7	3.2	0.8
3N1-d		2.3	3.0	2.3	2.6	1.7
3N6-g		2.4	1.4	1.0	1.5	1.2
3N1-e		2.4	3.1	3.0	0.7	0.8
3N8-d		2.4	2.4	2.0	2.4	2.1
3N1-f		2.4	3.0	3.3	0.9	1.4
3N1-g		2.5	3.2	2.8	2.9	2.6
3N8-e		2.6	3.5	2.8	4.2	3.6
3N2-b	*d*	2.8	1.2	1.0	1.1	1.0
3N1-h		2.8	3.3	3.7	1.0	1.5
3N6-h		2.8	4.3	3.3	4.9	3.9
3N1-i		2.9	3.3	3.6	1.0	1.5
3N2-g	*d*	4.6	5.5	4.7	5.1	4.1
3N3-h	*c,d*	5.5	6.5	3.9	6.7	4.3
3N1-ar	*b*	6.6	7.1	7.3	4.1	4.2
3Z-a	*c,d*	14.7	15.6	5.2	1.9	−6.9
3Z-b		15.7	17.2	6.8	3.4	−5.3
3Z-c		15.8	15.9	6.1	1.7	−6.4
3Z-d		16.3	15.5	6.0	1.2	−6.6
3Z-e		16.3	16.6	7.1	3.7	−4.5
3Z-f		16.4	15.5	6.0	1.5	−6.4
3Z-g		16.5	16.3	6.0	2.4	−6.2

[a] See **Glycine(H_2O)** section in the text for the definitions of the nomenclature. See Table 2S for full glycine$(H_2O)_3$ table. [b] Lowest-energy structure in ref 56. [c] Reference 57. [d] Reference 43.



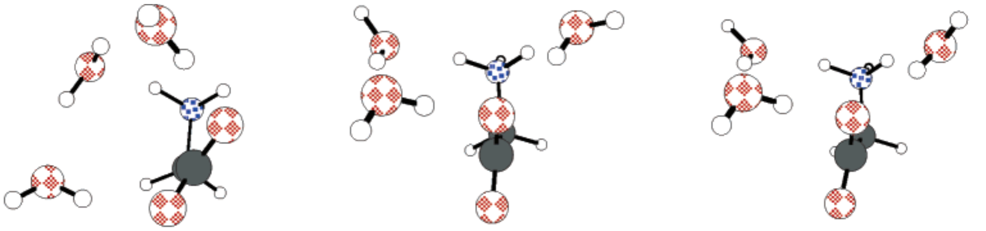
3N1-a 0.0 (0.0) [0.0] *0.0* [*0.0*] **3N6-a 1.1 (2.9) [0.2] 2.3 [*0.1*]** **3N6-c 1.3 (2.7) [0.1] 3.2 [*1.0*]**

3N3-h 5.5 (6.5) [3.9] *6.7* [*4.3*] **3N1-ar 6.6 (7.1) [7.3] *4.1* [*4.2*]**

EFP1/HF waters and RHF/6-31++G(d,p) glycine
(RHF/6-31++G(d,p))
[MP2//RHF/6-31++G(d,p)]
PCM+RHF//RHF/6-31++G(d,p)
[*PCM+MP2// RHF/6-31++G(d,p)*]

Figure 6. Conformers and relative energies (kcal/mol) of N glycine$(H_2O)_3$.

This structure (3N3-h in Figure 6) is 3.9 kcal/mol higher in energy than the global minimum at the MP2//RHF/6-31++G(d,p) level of theory. Rameakers et al. purport to find the global minimum.[56] However, this structure (3N1-ar in Figure 6) is 7.3 kcal/mol higher than the global minimum at the MP2//RHF/6-31++G(d,p) level of theory. After optimizations at the B3LYP/



3Z-a 14.7 (15.6) [5.2] *1.9* **[*-6.9*]** **3Z-d 16.3 (15.5) [6.0]** *1.2* **[*-6.6*]** **3Z-f 16.4 (15.5) [6.0]** *1.5* **[*-6.4*]**

EFP1/HF waters and RHF/6-31++G(d,p) glycine
(RHF/6-31++G(d,p))
[MP2//RHF/6-31++G(d,p)]
PCM+RHF//RHF/6-31++G(d,p)
[*PCM+MP2// RHF/6-31++G(d,p)*]

Figure 7. Conformers and relative energies (kcal/mol) of Z glycine$(H_2O)_3$.

Table 4. Structures and Relative Energies (in kcal/mol) for Glycine$(H_2O)_4$[a]

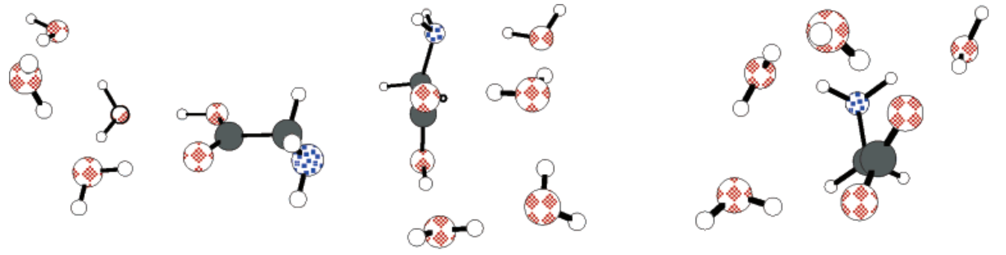
structure	EFP1/HF and RHF/6-31++G(d,p)	full RHF/ 6-31++G(d,p)	MP2//RHF/ 6-31++G(d,p)	PCM+RHF//RHF/ 6-31++G(d,p)	PCM+MP2//RHF/ 6-31++G(d,p)
4N1-a	0.0	0.0	0.0	0.0	0.0
4N1-b	0.3	0.4	0.5	0.0	0.1
4N1-c	0.7	1.0	0.8	0.4	0.3
4N1-d	0.9	0.3	0.4	0.3	0.5
4N1-e	1.2	1.4	1.3	0.5	0.5
4N1-f	1.5	1.0	1.2	0.4	0.7
4N1-g	1.6	1.8	1.8	2.6	2.7
4N1-h	1.7	1.6	1.2	2.1	1.7
4N1-i	1.7	1.4	1.1	0.7	0.5
4N1-j	1.8	2.4	2.3	2.1	1.8
4N1-k	1.8	1.4	2.1	1.4	2.1
4N1-l	1.9	1.4	2.3	1.4	2.3
4N1-m	2.0	1.8	1.6	0.7	0.7
4N6-a	0.3	0.8	−1.0	0.4	−1.0
4Z-a	12.1	11.7	1.6	1.2	−7.7
4Z-b	12.2	13.4	2.7	1.9	−7.4
4Z-c	12.3	13.3	3.0	0.9	−7.8
4Z-d	12.3	13.1	2.6	0.9	−8.0
4Z-e	12.5	13.6	3.2	1.2	−7.5
4Z-f	12.8	14.9	3.2	4.9	−5.3
4Z-g	13.1	13.2	3.9	0.2	−7.6
4Z-h	13.3	13.1	3.1	1.0	−7.4
4Z-i	13.5	12.0	2.9	0.6	−7.1
4Z-j	13.5	14.0	3.6	1.8	−7.0
4Z-k	13.6	13.3	3.3	2.8	−5.8
4Z-l	13.6	14.8	4.8	3.2	−5.3
4Z-m	13.7	14.6	4.8	1.8	−6.2
4Z-n	13.7	12.5	3.4	-0.5	−8.0
4Z-o	13.7	13.0	3.3	1.6	−6.7

[a] See **Glycine(H_2O)** section in the text for the definitions of the nomenclature. See Table 3S for full glycine$(H_2O)_4$ table.

6-31++G(d,p) level of theory, 3N1-ar is 8.5 kcal/mol higher in energy than 3N1-a.

The lowest-energy efp−ab zwitterionic glycine complex with three water molecules (3Z-a) is shown in Figure 7. At the RHF/6-31++G(d,p) level of theory, two structures found in this work (3Z-d and 3Z-f in Figure 7) are predicted to be 0.1 kcal/mol lower in energy than 3Z-a. Neither of these structures has been previously reported. Structure 3Z-a was previously reported in the literature,[43,57] as were three higher-energy structures.[56,58] The energy ordering predicted by the MP2 single-point energies suggests that 3Z-a is the lowest-energy Z minimum for the gas-phase cluster with three water molecules. However, the global minimum for the gas-phase three-water system is predicted to be an N structure. C-PCM+MP2 calculations predict that the lowest-energy Z form is 6.9 kcal/mol lower in energy than the lowest-energy N form, so the inclusion of a third explicit water molecule increases the (Z − N) energy difference by 2.2 kcal/mol. In general, electron correlation effects stabilize the Z structures by 6.7−10.4 kcal/mol, while the continuum solvation effects stabilize the Z structures with respect to the N structures by 12.1−18.7 kcal/mol.

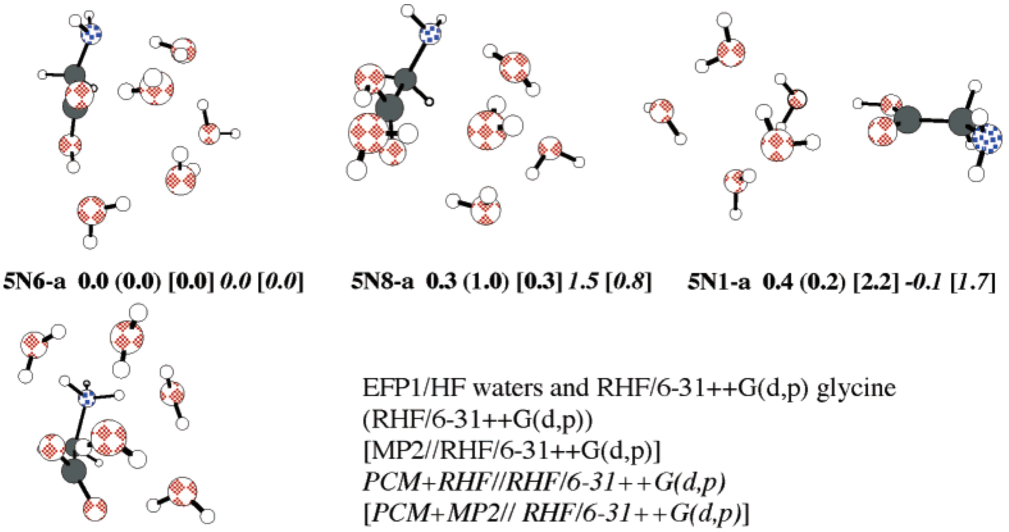
Glycine(H_2O)$_4$. In a glycine$(H_2O)_4$ complex, the number of low-energy local minima increases substantially. This prompted the procedural change described in the Computational Methods section above. At least 26 N complexes of gas-phase glycine conformer 1 with four water molecules are predicted to lie within 3.0 kcal/mol of the global minimum based on the Monte Carlo calculations (Table 4, 3S). The lowest-energy RHF/6-31++G-(d,p) N structure (4N1-a in Figure 8) is analogous to the water hexamer “book” structure. The lowest-energy structure based



4N1-a 0.0 (0.0) [0.0] *0.0* [*0.0*] **4N6-a 0.3 (0.8) [-1.0]** *0.4* [*-1.0*] **4Z-a 12.1 (11.7) [1.6]** *1.2* [*-7.7*]

EFP1/HF waters and RHF/6-31++G(d,p) glycine (kcal/mol)
(RHF/6-31++G(d,p))
[MP2//RHF/6-31++G(d,p)]
PCM+RHF//RHF/6-31++G(d,p)
[*PCM+MP2// RHF/6-31++G(d,p)*]

Figure 8. Conformers and relative energies (kcal/mol) of glycine$(H_2O)_4$.



5N6-a 0.0 (0.0) [0.0] *0.0* [*0.0*] **5N8-a 0.3 (1.0) [0.3]** *1.5* [*0.8*] **5N1-a 0.4 (0.2) [2.2]** *-0.1* [*1.7*]

EFP1/HF waters and RHF/6-31++G(d,p) glycine
(RHF/6-31++G(d,p))
[MP2//RHF/6-31++G(d,p)]
PCM+RHF//RHF/6-31++G(d,p)
[*PCM+MP2// RHF/6-31++G(d,p)*]

5Z-a 8.6 (9.5) [0.4] *0.8* [*-7.3*]

Figure 9. Conformers and relative energies (kcal/mol) of glycine$(H_2O)_5$.

on gas-phase conformer 6 (4N6-a in Figure 8) has a cyclic structure similar to 3N6-a. At the RHF/6-31++G(d,p) level of theory, this structure is 0.8 kcal/mol higher in energy than 4N1-a. However, MP2 single-point energies predict that it is 1.0 kcal/mol more stable. MP2 optimizations on these structures increase this energy difference in favor of the cyclic isomer to 1.7 kcal/mol. Bond lengths for both structures decrease upon MP2 optimization.

The lowest-energy Z complex is shown in Figure 8. C-PCM+MP2 single-point energy calculations predict that 4Z-a is 7.7 kcal/mol lower in energy than 4N1-a and 6.7 kcal/mol lower in energy than 4N6-a. The electron correlation effects stabilize Z structures with four waters by 8.5−11.7 kcal/mol, while bulk solvent effects stabilize the Z structures by 10.0−14.7 kcal/mol. The (Z − N) energy difference for four explicit waters is within 1 kcal/mol of the difference for three explicit waters, which suggests that the C-PCM+MP2 Z-N energy difference may be converging with respect to the number of discrete water molecules considered inside the bulk.

A previous investigation of the energy difference between the N and Z structures predicted the zwitterion to be more stable when four discrete water molecules were considered in the calculations.[56] However, this study did not sample the configuration space, so the "global minima" reported are in fact high-energy structures. At the B3LYP/6-31++G(d,p) level of theory, 4N6-a is predicted to be 16.3 kcal/mol lower in energy than the structure reported in ref 56; 4Z-a is predicted to lie 5.0 kcal/mol above 4N6-a, and the zwitterionic structure reported in ref 56 is predicted to lie 9.6 kcal/mol above 4N6-a. In addition, the structures reported in the Yamabe et al. study on intermolecular proton transfer[58] are somewhat higher in energy than the global minima reported here. The calculations reported in this work predict that the global minimum is an N structure when it is not embedded in bulk water.

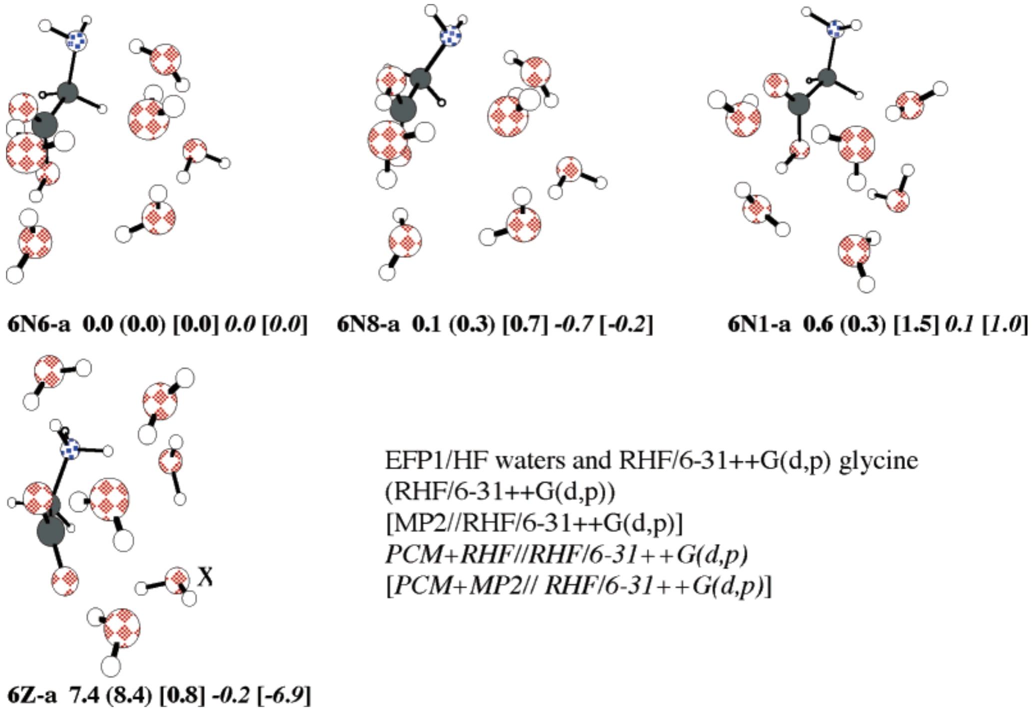
Glycine$(H_2O)_5$. With five water molecules, the number of low-energy local minima escalates further. Only the global minimum and a few selected low-energy minima from the Monte Carlo "basin-hopping" procedure are discussed here.

For the glycine$(H_2O)_5$ complex, the efp−ab method predicts relative energies that are in good agreement with the fully ab initio calculations. The RHF/6-31++G(d,p) global minimum N structure is predicted to be a bridge-like species (see 5N6-a in Figure 9) similar to 3N6-a and 4N6-a, rather than a structure consisting of a water cluster at the carboxylic acid end of glycine

Table 5. Structures and Relative Energies (in kcal/mol) for Glycine$(H_2O)_5$[a]

structure	EFP1/HF and RHF/6-31++G(d,p)	full RHF/ 6-31++G(d,p)	MP2//RHF/ 6-31++G(d,p)	PCM+RHF//RHF/ 6-31++G(d,p)	PCM+MP2//RHF/ 6-31++G(d,p)
5N6-a	0.0	0.0	0.0	0.0	0.0
5N6-b	0.1	0.2	0.3	0.2	0.3
5N6-c	0.3	0.3	0.5	0.1	0.4
5N8-a	0.3	1.0	0.3	1.5	0.8
5N6-d	0.4	0.8	0.8	0.3	0.4
5N1-a	0.4	0.2	2.2	-0.1	1.7
5Z-a	8.6	9.5	0.4	0.8	−7.3
5Z-b	8.7	9.9	1.0	0.6	−7.1

[a] See **Glycine(H_2O)** section in the text for the definitions of the nomenclature.


6N6-a 0.0 (0.0) [0.0] ***0.0 [0.0]*** **6N8-a 0.1 (0.3) [0.7]** ***-0.7 [-0.2]*** **6N1-a 0.6 (0.3) [1.5]** ***0.1 [1.0]***

6Z-a 7.4 (8.4) [0.8] ***-0.2 [-6.9]***

EFP1/HF waters and RHF/6-31++G(d,p) glycine
(RHF/6-31++G(d,p))
[MP2//RHF/6-31++G(d,p)]
PCM+RHF//RHF/6-31++G(d,p)
[PCM+MP2// RHF/6-31++G(d,p)]

Figure 10. Conformers and relative energies (kcal/mol) of glycine$(H_2O)_6$.

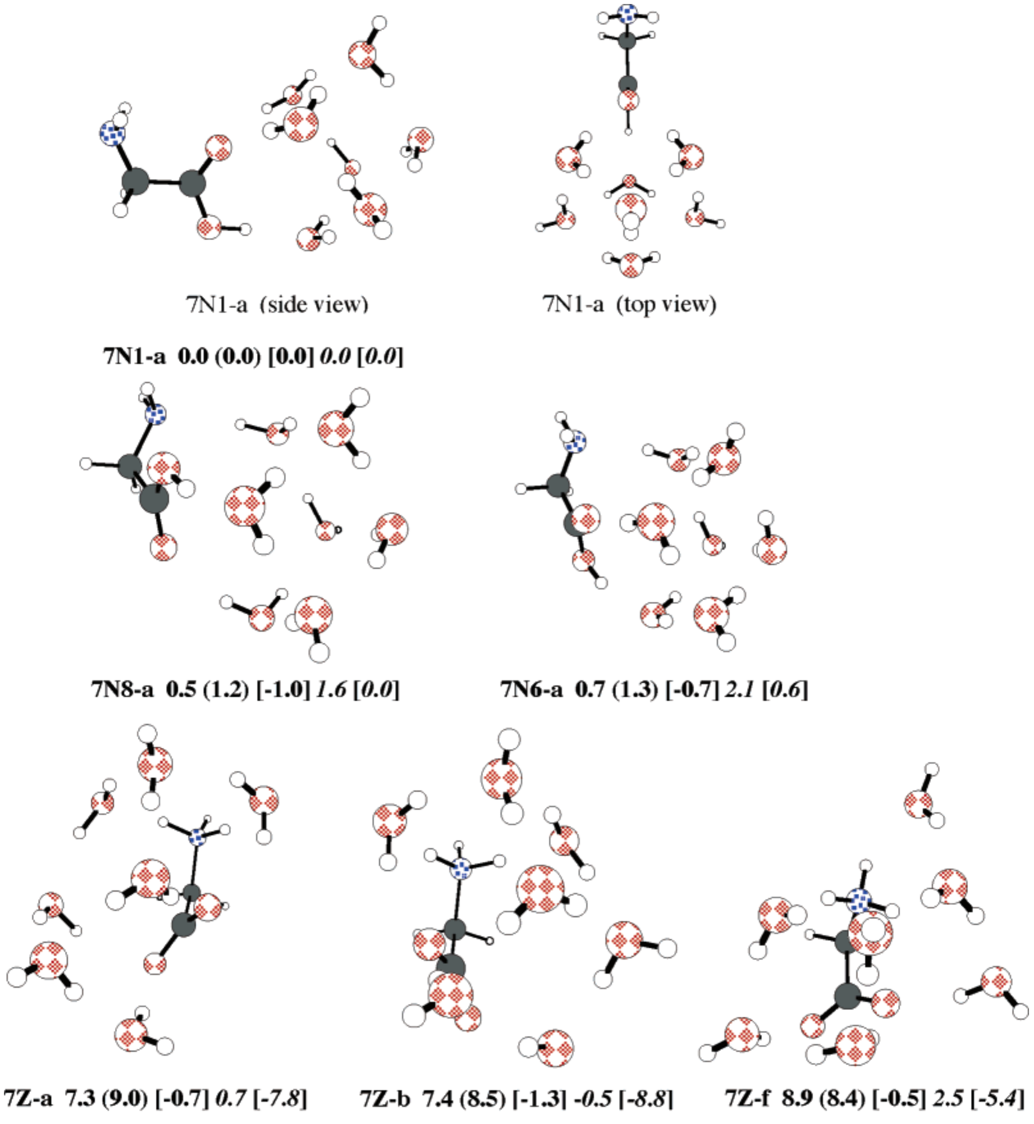
Table 6. Structures and Relative Energies (in kcal/mol) for Glycine$(H_2O)_6$[a]

structure	EFP1/HF and RHF/6-31++G(d,p)	full RHF/ 6-31++G(d,p)	MP2//RHF/ 6-31++G(d,p)	PCM+RHF//RHF/ 6-31++G(d,p)	PCM+MP2//RHF/ 6-31++G(d,p)
6N6-a	0.0	0.0	0.0	0.0	0.0
6N8-a	0.1	0.3	0.7	−0.7	−0.2
6N8-b	0.2	1.5	1.4	0.4	0.3
6N8-c	0.3	1.3	1.3	0.3	0.3
6N6-b	0.6	1.7	1.3	1.1	0.2
6N1-a	0.6	0.3	1.5	0.1	1.0
6N6-c	0.7	0.6	1.2	−0.7	−0.1
6N8-d	0.8	0.7	1.1	−0.5	0.1
6N1-b	1.0	0.7	2.1	0.5	1.5
6N1-c	1.1	0.6	0.7	−0.1	0.1
6N1-d	1.1	0.8	2.3	0.2	1.4
6N1-e	1.2	0.3	3.4	−1.2	1.9
6N8-e	1.4	1.7	1.7	0.7	0.6
6N1-f	1.5	0.5	3.2	−1.0	1.4
6N1-g	1.6	0.8	2.1	0.5	1.5
6Z-a	7.4	8.4	0.8	−0.2	−6.9

[a] See **Glycine(H_2O)** section in the text for the definitions of the nomenclature.

(see 5N1-a in Figure 9). This 5N6-a structure can be considered to be more solvated, as the waters are in closer proximity to the glycine molecule. The nonplanar gas-phase conformer 6 can accept hydrogen bonds, and this appears to encourage bridge-like structures. The lowest-energy Z structure is also a bridged arrangement. C-PCM+MP2 calculations predict that 5Z-a is 7.3 kcal/mol lower in energy than 5N6-a. Incorporation of electron correlation reduces the (Z − N) energy difference to 0.4 kcal/mol, and the addition of continuum solvent further stabilizes the zwitterion by 8.7 kcal/mol (Table 5).

The only previous calculation we are aware of on this system predicted the zwitterion to be more stable by 6.6 kcal/mol,[56] but the minima reported are not the lowest-energy structures. The N structure reported in ref 56 is predicted to lie



7N1-a 0.0 (0.0) [0.0] *0.0* **[*0.0*]**

7N8-a 0.5 (1.2) [-1.0] *1.6* **[*0.0*]** **7N6-a 0.7 (1.3) [-0.7]** *2.1* **[*0.6*]**

7Z-a 7.3 (9.0) [-0.7] *0.7* **[*-7.8*]** **7Z-b 7.4 (8.5) [-1.3]** *-0.5* **[*-8.8*]** **7Z-f 8.9 (8.4) [-0.5]** *2.5* **[*-5.4*]**

EFP1/HF waters and RHF/6-31++G(d,p) glycine (kcal/mol)
(RHF/6-31++G(d,p))
[MP2//RHF/6-31++G(d,p)]
PCM+RHF//RHF/6-31++G(d,p)
[PCM+MP2// RHF/6-31++G(d,p)]

Figure 11. Conformers and relative energies (kcal/mol) of glycine(H_2O)$_7$.

13.2 kcal/mol higher in energy than 5N6-a at the RHF/6-31++G(d,p) level of theory, and 18.0 kcal/mol higher in energy at the B3LYP/6-31++G(d,p) level of theory. The Z structure reported in ref 56 is predicted to lie 9.5 kcal/mol higher in energy than 5N6-a at the B3LYP/6-31++G(d,p) level of theory.

Glycine(H_2O)$_6$. The lowest-energy N complexes with six water molecules found in this study have a cubic structure (Figure 10). The lowest-energy Z structure corresponds to the addition of the water labeled "x" to 5Z-a with subsequent expansion of the structure. At the RHF/6-31++G(d,p) level of theory, the zwitterion is predicted to be 8.4 kcal/mol higher in energy than the global minimum (Table 6). Integration of correlation energy and bulk solvation effects via C-PCM+MP2 single-point energy calculations reverses the energy ordering to 6.9 kcal/mol in favor of the Z structure. This stabilization may be divided into correlation effects of about 7 kcal/mol and bulk solvent effects of about 8 kcal/mol.

Fernández-Ramos et al. examined the proton-transfer energy difference for the glycine(H_2O)$_6$ complex using RHF and MP2 methods with and without an Onsager SCRF, and found that their zwitterionic structure lies lower in energy than their N structure for all levels of theory.[42] However, they only probed a few starting structures and did not report cubic structures. In contrast, the present study predicts that the global minimum for the cluster of glycine with six water molecules will be a nonionized structure in the gas phase.

Glycine(H_2O)$_7$. Low-energy N and Z structures for the glycine(H_2O)$_7$ supermolecular complex are shown in Figure 11. The efp−ab method and the RHF/6-31++G(d,p) level of theory

Table 7. Structures and Relative Energies (in kcal/mol) for Glycine$(H_2O)_7$[a]

structure	EFP1/HF and RHF/6-31++G(d,p)	full RHF/ 6-31++G(d,p)	MP2//RHF/ 6-31++G(d,p)	PCM+RHF//RHF/ 6-31++G(d,p)	PCM+MP2//RHF/ 6-31++G(d,p)
7N1-a	0.0	0.0	0.0	0.0	0.0
7N1-b	0.2	0.1	0.2	0.1	0.2
7N8-a	0.5	1.2	−1.0	1.6	0.0
7N8-b	0.6	1.3	−0.7	1.7	0.3
7N6-a	0.7	1.3	−0.7	2.1	0.6
7N6-b	0.7	1.3	−0.6	2.0	0.4
7N8-c	0.9	1.5	−0.6	1.7	0.3
7Z-a	7.3	9.0	−0.7	0.7	−7.8
7Z-b	7.4	8.5	−1.3	−0.5	−8.8
7Z-c	8.0	9.5	0.1	1.4	−6.7
7Z-d	8.6	9.9	0.9	1.5	−6.4
7Z-e	8.7	9.5	0.3	1.0	−7.0
7Z-f	8.9	8.4	−0.5	2.5	−5.4
7Z-g	8.9	9.0	0.1	0.4	−7.4

[a] See **Glycine(H_2O)** section in the text for the definitions of the nomenclature.

Table 8. Structures and Relative Energies (in kcal/mol) for Glycine$(H_2O)_8$[a]

structure	EFP1/HF and RHF/6-31++G(d,p)	full RHF/ 6-31++G(d,p)	MP2//RHF/ 6-31++G(d,p)	PCM+RHF//RHF/ 6-31++G(d,p)	PCM+MP2//RHF/ 6-31++G(d,p)
8N8-a	0.0	0.0	0.0	0.0	0.0
8N8-b	0.0	0.0	0.0	0.3	0.3
8N8-c	0.1	0.0	0.1	−0.1	−0.1
8N8-d	0.2	−0.3	**−0.1**	6.2	6.3
8N8-e	0.4	0.2	0.0	2.4	2.2
8N6-a	0.8	0.3	0.3	−2.4	−2.3
8N6-b	1.0	0.3	1.5	−2.1	−1.1
8N6-c	1.0	0.3	1.5	−3.0	−1.6
8N1-a	1.3	**−0.7**	2.6	−6.2	−3.3
8N8-g	1.3	0.4	1.6	−4.2	−2.8
8N1-c	1.8	−0.2	3.0	−7.7	−4.8
8N1-d	1.8	−0.2	3.1	**−8.1**	**−5.1**
8Z-a	5.3	5.9	−1.6	−4.3	−11.0
8Z-b	5.8	6.6	−0.5	−4.1	−10.5
8Z-c	6.2	5.9	−1.0	**−4.7**	−10.9
8Z-e	6.6	**5.3**	**−1.9**	−4.6	**−11.1**

[a] See **Glycine(H_2O)** section in the text for the definitions of the nomenclature. The lowest-energy N and Z species at each level of theory are in bold. See Table 4S for full glycine$(H_2O)_8$.

predict the global minimum to be an N prismatic structure based on gas-phase conformer 1 (7N1-a). However, MP2 single-point energy calculations predict 7N8-a and 7N6-a to lie 1.0 and 0.7 kcal/mol lower in energy than 7N1-a, respectively.

With seven discrete water molecules, many structures fall within 1 kcal/mol of the lowest-energy zwitterionic structure; thus, it may be necessary to examine multiple structures in order to interpret experimental results for this system.

Gas-phase MP2 single-point energy calculations stabilize the Z structures by about 9 kcal/mol relative to the lowest-energy N species. These calculations predict that the global minimum structure will be 7Z-b and anticipate that it lies 0.3 kcal/mol lower in energy than 7N8-a. The integration of correlation energy and C-PCM bulk solvation effects further stabilizes the Z structures. At this level of theory, 7Z-b is predicted to lie 8.8 kcal/mol lower in energy than both 7N1-a and 7N8-a. Electron correlation effects again stabilize the Z structures by about 9 kcal/mol relative to the lowest-energy N species. The addition of a bulk solvent stabilizes the zwitterionic structures by about 9 kcal/mol (Table 7).

The zwitterionic structures seen in this work appear to be qualitatively similar to the PM3 and AM1 structures reported previously.[39] The structures show a high degree of hydrogen bonding. In addition, the solvent molecules are asymmetrically clustered around the Z glycine solute, in contrast to the highly symmetrical ordering seen around N glycine. The N−H···O bond is effectively broken, and the CO_2^- group is twisted with respect to the NH_3^+ group. In addition, the semiempirical calculations on the glycine$(H_2O)_7$ system predicted the enthalpy difference between the N and Z structures to be 4.0 kcal/mol; when solvation energies were considered, the combined semiempirical/SCRF approach predicted the zwitterionic structure to be lower in energy than the nonionized structure by 1.4 kcal/mol.[39] This is in good agreement with our combined RHF/C-PCM approach, which predicts 7Z-b to be 0.5 kcal/mol lower in energy than the lowest-energy N structures.

Glycine$(H_2O)_8$. As noted above for the glycine$(H_2O)_7$ complex, the supermolecular N glycine$(H_2O)_8$ complexes are well ordered (Figure 12). The low-energy Z complexes are structured similarly to the N bridge structures and contain rings of four or five water molecules. Upon inclusion of electron correlation effects, the lowest-energy N structure is predicted to be 8N1-d. Five other N minima within 0.4 kcal/mol at the MP2//RHF/6-31++G(d,p) level of theory were found in this study (Table 8).

Correlation stabilizes the Z minima by 7−8 kcal/mol relative to the Hartree−Fock results, so that these minima are predicted to lie lower in energy than the N minima. When both correlation and bulk solvent effects are included, the Z minima are predicted to lie 6 kcal/mol lower in energy than the lowest-energy N structures.

A summary of the (Z − N) energy differences for complexes with 1−8 discrete water molecules is presented in Table 9. The second and third columns in Table 9 summarize the results from

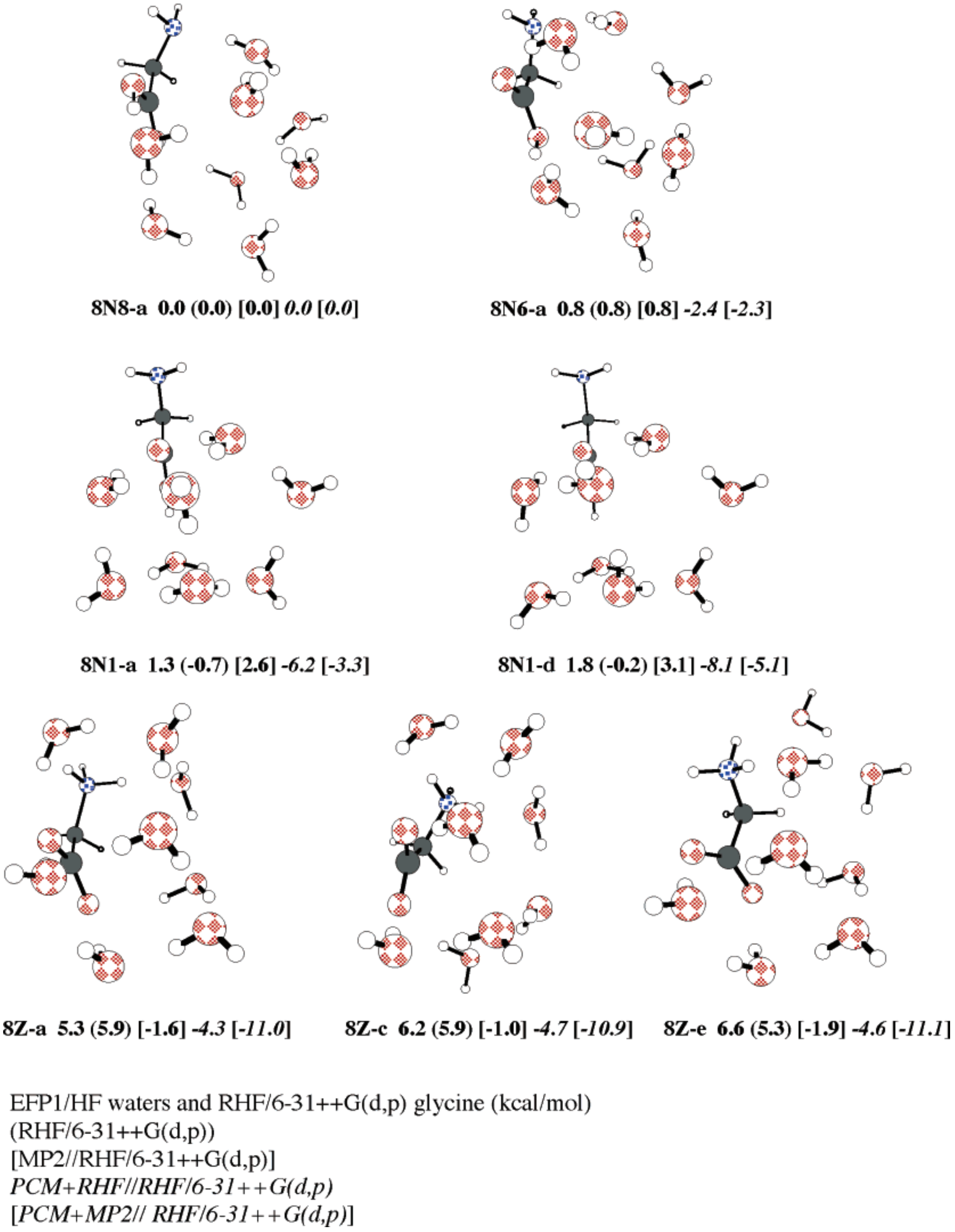


Figure 12. Conformers and relative energies (kcal/mol) of glycine(H_2O)$_8$.

Table 9. Convergence of the Nonionized-Zwitterion Energy Difference (kcal/mol) between the Lowest-Energy N and Z Structures (shown in Parentheses) at Various Levels of Theory[a]

number of waters	EFP1/HF and RHF/6-31++G(d,p)	Full RHF/ 6-31++G(d,p)	MP2//RHF/ 6-31++G(d,p)	PCM+RHF//RHF/ 6-31++G(d,p)	PCM+MP2//RHF/ 6-31++G(d,p)
1	24.2 (1N1-a, 1Z-a)	23.5 (1N1-a, 1Z-a)	15.3 (1N1-a, 1Z-a)	2.6 (1N1-b, 1Z-a)	−3.6 (1N1-a, 1Z-a)
2	20.2 (2N1-a, 2Z-a)	19.7 (2N1-a, 2Z-d)	10.8 (2N1-a, 2Z-a)	2.4 (2N1-a, 2Z-a)	−4.7 (2N1-a, 2Z-a)
3	14.7 (3N1-a, 3Z-a)	15.5 (3N1-a, 3Z-d)	5.2 (3N1-a, 3Z-a)	1.2 (3N1-a, 3Z-d)	−7.0 (3N8-a, 2Z-a)
4	12.1 (4N1-a, 4Z-a)	11.7 (4N1-a, 4Z-a)	2.6 (4N6-a, 4Z-a)	1.2 (4N1-a, 4Z-a)	−7.0 (4N6-a, 4Z-d)
5	8.6 (5N6-a, 5Z-a)	9.5 (5N6-a, 5Z-a)	0.4 (5N6-a, 5Z-a)	0.9 (5N1-a, 5Z-a)	−7.3 (5N6-a, 5Z-a)
6	7.4 (6N6-a, 6Z-a)	8.4 (6N6-a, 6Z-a)	0.8 (6N6-a, 6Z-a)	−0.2 (6N6-a, 6Z-a)	−7.1 (6N8-a, 6Z-a)
7	7.3 (7N1-a, 7Z-a)	9.0 (7N1-a, 7Z-a)	−0.3 (7N8-a, 7Z-b)	−0.5 (7N1-a, 7Z-b)	−8.8 (7N1-a, 7Z-b)
8	5.3 (8N8-a, 8Z-a)	6.0 (8N1-a, 8Z-e)	−1.8 (8N8-d, 8Z-e)	3.4 (8N1-d, 8Z-c)	−6.0 (8N1-d, 8Z-e)

[a] See the respective sections in the text for an analysis of the lowest-energy structures at each level of theory. See the **Glycine(H_2O)** section in the text for the definitions of the nomenclature

the two types of optimizations performed in this study: in the first, glycine was treated at the RHF/6-31++G(d,p) level of theory, and water molecules were treated with the EFP1/HF potential; in the second, both species were treated with RHF/6-31++G(d,p). The last three columns list the lowest-energy structures according to the relevant single-point energy calculations. It should be stressed that global optimizations at these latter three levels of theory will almost certainly yield other low-energy structures.

Conclusions

In general, the EFP1/HF water potentials track RHF structures and energy orderings well. The global minimum structure is the same for both methods for the structures with 1−7 water molecules, and vary by less than 1 kcal/mol for the cluster with 8 water molecules. Since there are a large number of low-energy structures, configurational sampling is important. Multiple gas-phase conformers must be considered. These species are highly fluxional, so multiple structures will likely be important for a physical description of the system. As the number of water molecules in the cluster increases, the identification of one global minimum becomes less significant, and statistical averaging will be essential.

The effect of systematically increasing the level of solvation is observed. The convergence of each method as the number of water molecules increases is summarized in Table 9. As the number of discrete water molecules increases, a different, “more solvated,” structure becomes the global minimum N structure. The internal structure of glycine in this complex is not the lowest-energy gas-phase conformer. This change occurs for three water molecules at the MP2 level of theory and for five water molecules at the RHF level of theory.

The integration of correlation with the continuum solvent effects through C-PCM+MP2 single-point energy calculations suggests that the lowest-energy Z species is 6.7−8.8 kcal/mol lower in energy than the lowest-energy N species in solution, in reasonable agreement with experiment. At least three discrete water molecules are needed in order to obtain converged results within this level of theory. Electron correlation effects typically stabilize the lowest-energy Z structures by about 7−9 kcal/mol with respect to the lowest-energy Z species. Even so, an N isomer is lowest in energy for complexes with up to six water molecules, in the absence of the continuum solvent. The Z structure is the global minimum for a supermolecular complex with seven or eight discrete water molecules at the MP2//RHF level of theory. As the number of discrete water molecules increases, the effect of a continuum solvent on the energy difference between the N and Z forms decreases. However, it still has a significant effect for clusters with eight waters, so these clusters do not completely recover long-range interactions with bulk solvent. Eight water molecules do not appear to fully solvate the glycine molecule.

Acknowledgment. We are grateful to the Air Force Office of Scientific Research for funding. We also thank Dr. Jamie Rintelman and Jonathan Mullin for helpful discussions.

Supporting Information Available: Full tables of relative energies; geometries and absolute energies of minima. This material is available free of charge via the Internet at http://pubs.acs.org.

JA062842P