



Published in final edited form as:

*J Phys Chem B*. 2011 January 20; 115(2): 249–261. doi:10.1021/jp108642r.

## Development of a ReaxFF Reactive Force Field for Glycine and Application to Solvent Effect and Tautomerization

Obaidur Rahaman<sup>a</sup>, Adri C.T. van Duin<sup>b</sup>, William A. Goddard III<sup>c</sup>, and Douglas J. Doren<sup>a,\*</sup>

<sup>a</sup>Department of Chemistry and Biochemistry, University of Delaware, Newark DE 19716, USA

<sup>b</sup>Department of Mechanical and Nuclear Engineering, The Pennsylvania State University, University Park, PA 16802, USA

<sup>c</sup>Material and Process Simulation Center, California Institute of Technology, Pasadena CA 91125, USA

### Abstract

Tautomerization of amino acids between the neutral form (NF) and the zwitterionic form (ZW) in water has been extensively studied, often using glycine as a model to understand this fundamental process. In spite of many advanced studies, the tautomerization reaction remains poorly understood because of the intrinsic complexities of the system, including multiple accessible reaction pathways, charge transfer, and variations of solvation structure. To establish an accurate model that can be used for molecular dynamics simulations, a ReaxFF reactive force field has been developed for glycine. A training set for the ReaxFF hydrocarbon potential was augmented with several glycine conformers and glycine-water complexes. The force field parameters were optimized to reproduce the quantum mechanically derived energies of the species in the training set. The optimized potential could accurately describe the properties of gas-phase glycine. It was applied to investigate the effect of solvation on the conformational distribution of glycine. Molecular dynamics simulations indicated significant differences in the dominant conformers in the gas phase and in water. This suggests that the tautomerization of glycine occurs through a conformational isomerization followed by the proton transfer event. The direct reaction mechanism of the NF→ZW proton transfer reaction in water, as well as mechanisms mediated by one or two water molecules, were investigated using molecular dynamics simulations. The results suggest that the proton transfer reaction is most likely mediated by a single water molecule. The ReaxFF potential developed in this work provides an accurate description of proton transfer in glycine and thus provides a useful methodology for simulating proton transfer reactions in organic molecules in the aqueous environment.

### Keywords

DFT; quantum mechanics; zwitterion; proton transfer; amino acid

## 1. Introduction

Proton transfer is a fast and common means to transport charges in biological systems. In aqueous environments, the protonation and deprotonation of organic functional groups play pivotal roles in regulating the activities of many enzymes and ion channels.<sup>1-3</sup> A

\*corresponding author; doren@udel.edu.

**Supporting Information Available:** The optimized ReaxFF parameters used in this work (in ffield file format) and the coordinates of the glycine-water and glycine-(water)<sub>2</sub> complexes. This material is available free of charge via the Internet at <http://pubs.acs.org>.

fundamental example is the protonation state of amino acids, which take the neutral form (NF) in the gas phase<sup>4-13</sup> but prefer the zwitterionic form (ZW) in water and crystalline states<sup>14-21</sup>. The transfer of the carboxylic proton of the NF to the amine group, forming the charge-separated ZW in the presence of water, serves to illustrate the mechanism of proton transfer.

Experimental<sup>22, 23</sup> studies of this proton transfer reaction in solution have used the smallest amino acid, glycine, as a model. Titration studies demonstrated that the ZW is more stable than the NF by 7.3 kcal/mol<sup>22</sup>. Chemical relaxation studies estimated the ZW→NF activation energy barrier to be 14.4 kcal/mol<sup>23</sup>. These two observations suggest that the NF→ZW activation energy barrier should be ~7.1 kcal/mol. Many quantum chemical studies have estimated the free energies of the NF-ZW tautomerization reaction.<sup>17,18,21,24,25</sup> Several of these computational studies considered a neutral form of glycine with the amine and the carboxyl group conveniently oriented for a direct proton transfer (NF<sub>cis</sub>). A density functional theory (DFT) study by Tunon et al. challenged this simplistic formulation<sup>24</sup>. They suggested that the proton transfer process was coupled with a conformational equilibrium of the NF. They predicted that the rate limiting step of the tautomerization process was a H-atom reorientation in the carboxyl group and not the proton transfer step. Kassab et al.<sup>17</sup> and Bandyopadhyay et al.<sup>25</sup> applied quantum chemistry to study the tautomerization process using clusters of one glycine with a few water molecules in continuum models of the solvent. These discrete/continuum studies estimated much lower free energy barriers for the NF<sub>cis</sub>→ZW transformation and suggested that NF<sub>cis</sub> might be an intermediate and not the most stable neutral form. Balta et al. used DFT to detect several NF, TS (transition state) and ZW structures of glycine-water complexes containing up to 6 water molecules.<sup>18,21</sup> The results of these quantum chemical studies are not conclusive because of limited configurational sampling, limitations on cluster size and variations in methods and basis sets. The lowest-energy conformers of both NF and ZW need to be identified in order to compare their energies, but this is difficult for well-solvated glycine. For instance, Bachrach detected 35 and 61 different configurations of NF and ZW respectively in glycine-(H<sub>2</sub>O)<sub>7</sub> clusters using DFT<sup>26</sup>. The existence of so many local minima also suggests the importance of statistical averaging of the configurations in order to assess the macromolecular properties reliably.

Molecular dynamics (MD) simulation can be used to derive the thermodynamic properties of a system by adequate statistical sampling of the configuration space. Several MD-based studies have described the proton transfer mechanism in glycine. An Empirical Valence Bond (EVB) model for reactive potentials was developed by Warshel et al.<sup>27-31</sup> and Chang and Miller.<sup>32,33</sup> By applying the EVB formulation in combination with MD simulation and Free Energy Perturbation, Nagaoka et al. estimated the free energy profile of the intramolecular proton transfer of glycine (ZW→NF) in aqueous solution<sup>15</sup>. The free energy change and the activation free energy were predicted to be 8.5 kcal/mol and 16.9 kcal/mol respectively, in good agreement with the experimental values of 7.3 kcal/mol and 14.4 kcal/mol.<sup>22,23</sup> They suggested that the water molecules interact more strongly with the ZW than with TS and NF, especially at the charged amine and carboxyl sites, making the ZW lower in energy.<sup>15</sup> They also suggested that the enthalpy contribution to the free energy barrier was larger than the entropy contribution. This was in contradiction to the experimental observation.<sup>23</sup> The hybrid quantum mechanics/molecular mechanics (QM/MM) method has also been applied to the glycine/water system.<sup>9,34,35</sup> Shoeib et al. used this approach to estimate the ZW-NF free energy difference as 39 kcal/mol,<sup>35</sup> a significant overestimate of the experimental value. Cui used a combination of QM/MM and an implicit model of the solvent, the polarizable continuum model (PCM), to predict that the ZW was more stable than NF by 7.0 kcal/mol, in good agreement with the experimental observation.<sup>34</sup> Leung et al. applied ab initio molecular-dynamics (AIMD) simulation to derive the potential of mean

force of the direct intra-molecular proton transfer event in glycine with 52 H<sub>2</sub>O molecules.<sup>20</sup> They demonstrated that the average hydration number of glycine changes from 5 to 8 when NF transforms into ZW. They obtained a ZW-NF free energy difference of 11.2 kcal/mol as compared to the experimental value of 7.3 kcal/mol. Recently Takayanagi et al. demonstrated that the computationally less expensive semiempirical PM6 method can reproduce the structures and energies of low-energy and proton-transfer-transition-state configurations of glycine-(H<sub>2</sub>O)<sub>n</sub> (n=2-7) clusters that compare well to previous DFT and ab initio studies.<sup>36</sup> They applied MD simulation using PM6 potential surface to the clusters. Due to the high free energy barrier, the NF-ZW transformation was not observed during the simulation, but they observed significant differences in the solvation structures between NF and ZW.

Despite these investigations, the fundamental characteristics of the NF-ZW tautomerization reaction still remain unresolved. Computational studies are very useful in predicting the microstructure of the system and in deriving thermodynamic properties to compare with the experimental observations. However, the NF-ZW isomerization reaction requires effective modeling of the change in bonding and charge separation (which are difficult to model with traditional force fields) and averaging over solvation structure (which is difficult to model with quantum chemistry methods). Furthermore, the lowest energy transition states and the correct reaction pathways are far from being well resolved. Thus, it is essential to develop an accurate but computationally inexpensive method that can provide adequate sampling of the configurations in a large-scale simulation while allowing for bond breaking and formation. Reactive force fields are attractive targets for addressing this problem.

ReaxFF is a bond-order dependent reactive force field that can account for polarization and charge transfer effects in complex systems. This force field is efficient in large-scale simulations of dynamic systems due to its low computational cost but high accuracy. In this formulation, bond energies are calculated from the bond orders, which are continually updated during the MD simulation allowing spontaneous bond formation and bond breaking necessary for a chemical reaction. It is computationally much less expensive than QM or even semi-empirical methods like AM1 or PM6 which are able to describe chemical reactions. This force field has been developed by extensive training against QM-derived data to describe bond formation and dissociation in many hydrocarbon compounds.<sup>37</sup> It can accurately reproduce the geometries and stabilities of several non-conjugated, conjugated and radical-containing compounds. Both the reactive and nonreactive behaviors of these compounds have been successfully characterized by the force field. The development of this empirical potential opened a window toward large scale simulations of various types of chemical reactions in hydrocarbons. For example, ReaxFF was applied to study the cyclization pathways of a diaromatic carotenoid, isorenieratene, via a combination of the Diels–Alder mechanism and intra-molecular hydrogen shifts.<sup>38</sup> The hydrogen shift steps were predicted to have the highest energy barriers, acting as rate-determining steps. Other notable applications of this potential involved gas phase hydrocarbon oxidation at high temperature as a model to understand the mechanisms of hydrocarbon combustion.<sup>39</sup> Several MD simulations were performed with mixtures of O<sub>2</sub> and hydrocarbons, including methane, *o*-xylene, propene and benzene. Valuable insights about the complex reaction mechanisms were obtained from the simulation results. ReaxFF could correctly predict the relative reaction rates, including barrier heights. These studies are the building blocks for developing ReaxFF to simulate complex reaction mechanisms in organic and biological molecules.

Considering its success in describing reactive systems with low computational expense, we selected ReaxFF to simulate the NF→ZW proton transfer in water. The method can model the breaking and formation of covalent bonds, and includes polarization effects needed to

account for the significant differences of charge distribution between NF and polarized ZW (see Computational Methods section for details). The lowest energy transition state of the NF-ZW proton transfer reaction is not known. Most the prior computational studies assume a direct intra-molecular proton transfer mechanism, but mechanisms in which proton transfer is mediated by a water molecule (similar to the Grotthuss mechanism<sup>40,41</sup> in bulk water) have not been extensively explored for this system. In this work, we focus on identifying the correct transition state of the reaction, comparing the energy barriers associated with the direct and water-mediated reactions. This paper describes the development of the ReaxFF potential for glycine and initial applications to the NF-ZW tautomerization reaction in water. This work builds on a recent extension of the ReaxFF description to proton transfer reactions in water. This ReaxFF description for water was trained against binding energies for water clusters, concerted proton transfer reactions in neutral water, water self-ionization reactions and density and cohesive energies for ice crystals water.\* This water description was validated for bulk water by comparison against experimental data on water cohesive energy, diffusion constant and structure. Some properties of this water potential are available in the literature.<sup>70</sup> Using the O/H parameters from this ReaxFF water description, we augmented the training set of the existing potentials that describe hydrocarbons and their oxidation<sup>39</sup> with several glycine conformers and glycine-water complexes. The force field parameters were optimized against quantum mechanically derived geometries and energies of the species in the training set. The potential was examined for its ability to describe the conformational energies of gas-phase glycine and then applied to investigate the NF-ZW tautomerization reaction in water including both direct and water-mediated mechanisms.

## 2. Computational Methods

### 2.1 ReaxFF: A Reactive Force Field

ReaxFF is a reactive force field that can describe bond breaking and bond formation and is useful for large-scale dynamic simulations of reactive systems. Each element is described using a single atom type. Thus, no alterations of atom types are required due to the modification of bonding during a chemical reaction. This also facilitates the transferability of the parameters to a different system. The force field is trained against a QM-derived set of energies for small molecules and clusters. For each element, several parameters to describe valence bond parameters, electronegativity, hardness, and other effects are optimized to reproduce the QM derived energies and charges. ReaxFF can reproduce energies near the accuracy of the training set with much less computational cost.

The reactive site or pathways need not be predefined in a ReaxFF MD simulation and connectivities between the atoms are not predetermined. Instead, bond orders are calculated from the inter-atomic distances  $r_{ij}$ :<sup>37</sup>

$$BO'_{ij} = \exp \left[ p_{bo,1} \cdot \left( \frac{r_{ij}}{r_o} \right)^{p_{bo,2}} \right] + \exp \left[ p_{bo,3} \cdot \left( \frac{r_{ij}^\pi}{r_o} \right)^{p_{bo,4}} \right] + \exp \left[ p_{bo,5} \cdot \left( \frac{r_{ij}^{\pi\pi}}{r_o} \right)^{p_{bo,6}} \right] \quad (1)$$

The parameters ( $p_{bo,1}$  and  $p_{bo,2}$ ), ( $p_{bo,3}$  and  $p_{bo,4}$ ) and ( $p_{bo,5}$  and  $p_{bo,6}$ ) correspond to the sigma bond, first pi bond and second pi bond respectively. The connectivities are updated with the inter-atomic distances at each MD step, allowing for bond breaking and formation. Valence bond contributions, including bond energy, valence angle energy, and torsion angle

\*ACTvD, VSB, S and WAG, in preparation

energy, are calculated and added to determine the total bonding- related energy of the system. For example, the bond energy is calculated from the bond order  $BO'_{ij}$  as <sup>37</sup>

$$E_{\text{bond}} = -D_e \cdot BO_{ij} \cdot \exp[p_{be,1}(1 - BO_{ij}^{p_{be,1}})] \quad (2)$$

where  $D_e$  and  $p_{be,1}$  are bond parameters. According to Equation (2), the bond energy term  $E_{\text{bond}}$  gradually disappears as the bond order  $BO'_{ij}$  approaches zero during the dissociation of a bond. The valence angle and torsion angle energy terms are also functions of bond orders and go to zero upon bond dissociation. This ensures continuity of the energy and forces during bond dissociation/formation.

In order to correctly describe a system with changing connectivity, the nonbonded van der Waals and Coulomb energy are calculated between *every* atom pair. The van der Waals energy is calculated using a distance-corrected Morse potential. A shielded interaction term is required to avoid the excessive high repulsion between two closely interacting bonded atoms. Similarly, a shielded coulomb potential was implemented to calculate the electrostatic energy between all atom pairs. The Electron Equilibration Method (EEM)<sup>42</sup> is implemented to derive the changing atomic charges. This formulation enables polarization and charge transfer effects. More details about the ReaxFF method can be found elsewhere.<sup>37,43</sup>

## 2.2 Quantum Mechanical Calculations

Quantum mechanical calculations were performed using the B3LYP<sup>44,45</sup> hybrid DFT functional with 6-311++G(2df,2p) basis set. The Gaussian 03<sup>46</sup> program package was used for all geometry, energy and frequency calculations. The geometries of glycine, glycine-water and glycine-(water)<sub>2</sub> complexes were fully minimized. Vibrational frequency analysis was performed for each minimum; all of the calculated frequencies were real for the completely minimized geometries (without restraints), confirming the presence of true minima. Angular distortion energies and rotational energy barriers were calculated by fixing the angle of interest at a particular value and optimizing all other molecular coordinates.

## 2.3 Molecular Dynamics Simulations

**NF in gas and water**—We performed NVT-MD simulations with NF in the gas phase as well as one NF solvated by 213 water molecules in a cubic simulation box of length 18.62 Å with periodic boundary conditions. Each simulation was performed for 1 ns with the temperature maintained at 300 K with a Berendsen thermostat and a temperature damping constant of 0.1 ps. The time step was 0.25 fs. The density of the system in water was maintained at 1.01 kg/dm<sup>3</sup>. To allow for equilibration, configurations generated in the first 100 ps were discarded. The dihedral angle distributions were constructed using configurations generated in the next 900 ps of simulation.

**Proton transfer in water**—Free energy perturbation<sup>15,16</sup> and umbrella sampling<sup>20</sup> methods have been previously applied to predict the classical reaction energy barrier for the direct intramolecular proton transfer. The ReaxFF force field allows the comparison of the energy barriers of the direct as well as water-mediated reaction mechanisms. However, methods that require ensemble averaging are computationally complex and expensive for water-mediated reaction mechanisms due to the presence of multiple reaction coordinates. Thus, in this initial study, we have applied a relatively simple and computationally inexpensive method of sliding restraints (as described below) to compare the potential energy barriers of the direct and water mediated reactions instead of calculating the free energies using the above mentioned methods. We note that a rigorous treatment of proton

transfer in this system would also have to account for quantum effects.<sup>69</sup> This is beyond the scope of the present work, though we make some additional comments in the Conclusions.

Bond restraining forces have been previously applied in ReaxFF to derive the lowest energy pathways for chemical reactions like isorenieratene cyclization<sup>38</sup> or oxidative dehydrogenation on vanadium oxide catalysts<sup>47</sup>. To derive the glycine-zwitterion transformation energy in water we augmented the ReaxFF potential with an additional restraint to the covalent O–H bond or the N---H hydrogen bond in order to force the proton transfer reaction,

$$E_{\text{restraint}} = k_1 \{1 - \exp(-k_2(r_{ij} - R_{\text{restraint}})^2)\} \quad (3)$$

The values of the bond restraint force constants  $k_1$  and  $k_2$  were 250 kcal/mol and  $1.0 \text{ \AA}^{-2}$  respectively. In the sliding restraint formulation, the restraint distance  $R_{\text{restraint}}$  was gradually shifted along the reaction coordinate so that the interatomic O–H or N---H bond distance,  $r_{ij}$ , changed from the value corresponding to the reactant to that of the product. The restraints were shifted during a NVT-MD simulation in which sufficient time was allocated to allow the system to equilibrate as the restraints were shifted.

Proton transfer reaction rates are often significant near room temperature. However, thermal fluctuations mask the relatively small reaction energy barriers in simulations at this temperature. On the other hand, simulations at a lower temperature reduce the thermal energy fluctuations, but the rearrangement of the surrounding water molecules is slow at very low temperature. Therefore, an optimum temperature must be determined in order to estimate the energy barriers with reasonable accuracy but low computational expense. We examined the energy profiles of the proton transfer reaction at several temperatures and selected  $T = 100 \text{ K}$  as the optimum temperature based on small thermal fluctuations and reasonable equilibration times.

Three glycine-(water)<sub>2</sub> complexes from the training set were selected as the starting points of direct, one-water mediated and two-water mediated proton transfer reactions. Each conformation was solvated with 98 additional water molecules randomly placed around them, creating a droplet of water. No boundary potentials or periodic boundary conditions were implemented. Then a MD simulation was performed using bond restraints. For the direct reaction, one bond restraint between the nitrogen and the carboxylic proton was required to keep the carboxylic proton oriented toward the amine group. For the one-water mediated reaction, two bond restraints, COOH...OH<sub>2</sub> and HOH...NH<sub>2</sub>, were required to keep the bridging water molecule properly aligned. For the two-water mediated reaction, an additional third bond restraint between the two water molecules, H<sub>2</sub>O...HOH, was required to keep the two bridging water molecules properly aligned. The systems were equilibrated in about 150 ps as indicated by stable energy profiles. Then the bond restraints were slowly and continuously shifted to derive the product from the reactant in conjunction with a MD simulation for 250 ps. The sliding of the restraint gradually shifted the proton to the acceptor (amine) group, forming a bond, while the bond between the donor (carboxyl) group and the proton was cleaved as a consequence. Finally the product was equilibrated for another 625 ps without any bond restraints. A simulation of at least 1 ns was generally required for the system to reach the transition state and gradually transform into a stable product as indicated by a steady energy profile. In a similar fashion, a reverse reaction was also simulated starting from the product and moving toward the reactant. We monitored the potential energy of the system as it evolved in the restrained MD simulation. The additional energy due to the bond restraint was subtracted from the potential energy at each step to finally construct the energy profiles of the proton transfer reactions.

## 3 Results and Discussions

### 3.1 ReaxFF Force Field Parameterization

Training a force field to describe a system in the gas phase as well as in solution is challenging, especially when the system has qualitatively different structures in the two environments. Glycine prefers to stay as NF in the gas phase but as ZW in water. In order to derive the ReaxFF parameters for glycine, we augmented the training set to include valence and dihedral angle distortions of glycine, its local minima in the gas phase and several glycine-water complexes. These structures are described in more detail below. The geometries and energies of the species and complexes were calculated using DFT, and the ReaxFF force field parameters were optimized to minimize the differences between the DFT and ReaxFF energies. These optimized parameters are available in a ffield file (see Supporting Information Available section) that can be directly used in the freely available “Large-scale Atomic/Molecular Massively Parallel Simulator” (LAMMPS) molecular dynamics program from Sandia National Laboratories<sup>48</sup>.

**3.1.1 Glycine Valence and Dihedral Angle Energies**—Quantum mechanical calculations were performed to estimate angular distortion energies of the glycine molecule and this information was used to determine the ReaxFF valence angle parameters. All of the valence angles in glycine, namely H–N–C, H–N–H, C–C–N, C–C=O, C–C–O(H), C–O–H, O–C=O and H–C–H, were considered. Good agreement between the QM and ReaxFF energies at different valence angles was obtained (Figure 1).

ReaxFF dihedral angle parameters were optimized by comparing the QM and ReaxFF energies of the N–C–C–O(H), C–C–O–H and C–C–N–H torsions. ReaxFF is able to closely reproduce the distortion energies of the N–C–C–O(H) and C–C–O–H dihedrals (Figures 2a and 2b). The energy barrier for the C–C–O–H dihedral angle is significantly higher than that of the other two. This is consistent with a previous DFT study on conformational stability of glycine.<sup>19</sup> There is a slight discrepancy between the QM and ReaxFF energies for the C–C–N–H dihedral angle torsion (Figure 2c). While the absolute energies are close (ReaxFF can reproduce the QM distortion energies with an RMSD of 0.79 kcal/mol), QM prefers a symmetric structure containing two intra-molecular N–H...O hydrogen bonds (both involving the same oxygen). Other QM studies also predicted this symmetric structure.<sup>5,6</sup> QM prefers this structure by ~1.5 kcal/mol over an asymmetric structure with one dominant N–H...O hydrogen bond, whereas ReaxFF has a slight preference (~0.5 kcal/mol) for the asymmetric structure.

**3.1.2 Glycine Gas-phase Local Minima**—Glycine is predicted to exist in multiple neutral conformations in the gas phase<sup>4-8,12,19</sup>. We optimized the five lowest-energy glycine conformations that were first predicted by Csaszar<sup>5</sup> and reproduced by Miller and Clary<sup>12</sup>. The conformers nA, nB, nC, nD and nE (the letter ‘n’ is used to indicate the neutral form of glycine) in Figure 3 correspond to the neutral Conf I, II, III, IV, and V in Figure 1 of reference<sup>12</sup> or Gly(1), Gly(2), Gly(3), Gly(4) and Gly(5) in Figure 4 of reference<sup>7</sup>. Experimental studies using matrix-IR spectroscopy,<sup>13,49-51</sup> microwave spectroscopy<sup>52-54</sup> and electron diffraction<sup>55</sup> also detected the three lowest-energy conformers, nA, nB and nC.

The QM results match well with the other high level *ab initio* calculations reported in the literature (Table 1). ReaxFF can reproduce the relative QM energies of the five conformers, correctly predicting nA as the global minimum. All the glycine conformations in the gas phase are neutral, which is consistent with the literature.<sup>4-8,13,19,49-51</sup> Optimizations from an initial geometry with a zwitterionic form ultimately led to one of the neutral forms.

**3.1.3 Glycine –water Complexes**—Several gas-phase glycine-water complexes have been studied by high level *ab initio* methods<sup>9,10,14,56</sup>. We obtained an extensive set of these glycine-water complexes by DFT energy minimizations. The initial configurations for the geometry optimizations were constructed by replicating the structures available in the literature<sup>9,10,14,56</sup> or by placing a water molecule in the vicinity of a glycine local minimum, making hydrogen bonds between them. Configurations of 12 glycine-water complexes were obtained (Figure 4). The coordinates of the complexes are given in the Supplementary material. We developed a systematic nomenclature for the complexes in order to categorize them in groups of similar conformations and facilitate comparisons with the glycine minima (Figure 3) and glycine-(water)<sub>2</sub> complexes (Figure 5). The letters ‘n’ and ‘z’ are used to represent the neutral and zwitterionic forms of glycine respectively. For example: nA-1w-a and nA-1w-b are two neutral glycine-water complexes with glycine molecules that resemble the species nA in Figure 3. However, the water molecules interact with different parts of the glycine molecule. This distinction is represented by letters ‘a’ and ‘b’. Table 2 compares the QM and ReaxFF energies of the glycine-water complexes. The average unsigned error of 2.82 kcal/mol indicates a reasonably good match between them. ReaxFF also satisfactorily reproduces the QM optimized geometries of the complexes (average RMSD = 0.15 Å).

The asymmetrical conformer nA-1w-a is the QM global minimum. In this conformer, the water molecule strongly interacts with the COOH group through two hydrogen bonds. Other computational studies<sup>9,10,56</sup> and a recent microwave spectroscopy study<sup>57</sup> identified this conformer as the global minimum. In the conformer, the water molecule has a specific position with respect to the glycine molecule. The oxygen atom lies in the plane of the heavy atoms of the glycine molecule, accepting a hydrogen bond. One hydrogen atom acts as the donor for the hydrogen bond with the C=O group of the glycine molecule. This arrangement forms a planar ring structure that contributes to the stability of the complex. The other hydrogen atom lies outside this plane. The conformer nA-1w-a is not recognized by ReaxFF as the global minimum but it is only 0.1 kcal/mol higher in energy than the ReaxFF global minimum, nC-1w-b, which is structurally very similar to nA-1w-a.

All complexes in Figure 4 were fully minimized without any restraints except the last two zwitterionic complexes, z-1w-a and z-1w-b. After several trials, it was not possible to get an unrestrained zwitterion-water complex with the selected method and basis set. A prior HF and MP2 study has reported a zwitterion-water complex (GLYZWM) as a true minimum<sup>14</sup>. However, most other theoretical calculations have not detected it.<sup>9,10,56</sup> Jensen and Gordon showed that at least two water molecules are necessary to stabilize the zwitterion and produce a potential energy minimum<sup>56</sup>. Thus, to obtain the two zwitterionic complexes, we restrained the three amide protons at a distance of 1.08 Å from the nitrogen atom during the geometry minimizations. The QM energies of z-1w-a and z-1w-b are respectively 19.9 kcal/mol and 22.5 kcal/mol higher than the QM energy of the global minimum, nA-1w-a. The relative instabilities of the zwitterion-water complexes are reproduced by the ReaxFF force field. The ReaxFF energies of z-1w-a and z-1w-b are respectively 14.6 kcal/mol and 15.7 kcal/mol higher than the energy of the ReaxFF global minimum nC-1w-b (see Table 2).

**3.1.4 Glycine-(water)<sub>2</sub> Complexes**—Many computational studies have investigated the gas-phase complexes of glycine with two or more water molecules in order to understand the neutral to zwitterion proton transfer mechanism.<sup>11,18,56,58-60</sup> We searched for the local minima of glycine-(water)<sub>2</sub> complexes. DFT energy optimizations were initiated with two water molecules placed at favorable locations around a neutral or zwitterionic glycine local minimum. We obtained 24 conformations, including zwitterionic complexes that were fully minimized without any restraints (Figure 5). The coordinates of the complexes are given in the Supplementary Material. We applied a nomenclature similar to that in the previous



section to describe the complexes. Table 3 compares the QM and ReaxFF energies of the glycine-(water)<sub>2</sub> complexes. The average unsigned error of 3.5 kcal/mol suggests a reasonable match between them. The ReaxFF/QM agreement can be improved by giving these glycine-(water)<sub>2</sub> complexes more weight in the force field parameterization, but doing so decreased performance of the force field for the rotational barriers. ReaxFF satisfactorily reproduces the QM optimized geometries of the complexes (average RMSD = 0.16 Å).

The QM global minimum is nA-2w-a, with the two water molecules interacting with each other and with the carboxyl group through hydrogen bonds. The formation of a planar ring structure provides special stability to the complex. The second lowest-energy conformer is nC-2w-c, which is 1.1 kcal/mol higher energy than nA-2w-a. Its geometry is similar with the exception that the glycine molecule acquires a conformation similar to nC instead of nA (Figure 3). Our observations match well with those of Jensen and Gordon<sup>56</sup>. They identified nA-2w-a (their N2d) as the global minimum and nC-2w-c (their N2e) as the second minimum having 1.0 kcal/mol higher energy.<sup>56</sup> ReaxFF successfully identifies nA-2w-a as the lowest-energy conformer and nC-2w-c as the second lowest-energy conformer with 0.4 kcal/mol higher energy (see Table 3).

The 24 complexes in Figure 5 are all distinct. However, conformer nB-2w-c and nB-2w-d are essentially the same, except for slightly different orientations of the water molecules, and differ by only 0.2 kcal/mol in energy. The glycine molecules in the conformer nF-2w-a and nF-2w-b are not analogous to any of the glycine local minima. In these two complexes, the plane of the carboxyl group is nearly perpendicular to the plane formed by the nitrogen and the two carbon atoms. Thus we used the letter “F” to distinguish them from the other conformers. Conformers z-2w-a to z-2w-d are all fully minimized stable complexes of the zwitterionic form. They have higher energy than the neutral conformers (see Table 3). Conformers z-2w-a, z-2w-b and z-2w-d were also found by Jensen and Gordon as Z2b, Z2a and Z2c respectively.<sup>56</sup> We found that ReaxFF adequately reproduced the QM-energy differences between the neutral and the zwitterionic glycine-(water)<sub>2</sub> clusters (Table 3).

**3.1.5 Gas Phase ZW → NF Transformation**—The ZW is unstable in the gas phase and it transforms into NF without any energy barrier<sup>56</sup>. In order to obtain a ReaxFF force field that can correctly describe the gas-phase ZW→NF transformation, we trained the force field parameters against the QM energies of several glycine conformers along the intrinsic reaction coordinate (IRC). The conformer nB was selected for the proximity of the carboxylic proton to the nitrogen atom. The O–H bond was restrained at different values during the geometry optimizations. Figure 6 shows the ReaxFF and QM energies of the ZW→NF proton transfer along the IRC. Both QM and ReaxFF predicted that NF is lower in energy than ZW in the gas phase. QM predicted a barrier-less ZW→NF transformation, which is consistent with the experimental studies.<sup>13,49-51</sup> Although this trend is successfully captured by the ReaxFF force field, it underestimates the energy difference between NF and ZW.

### 3.2 Solvation Effect on the Conformational Distribution of NF

Several experimental<sup>49-51,54,61-65</sup> and theoretical<sup>5,7,12,19,66,67</sup> studies have investigated the conformational behavior of glycine in the gas phase. The three lowest energy conformers: nA, nB and nC are conclusively observed using matrix-IR spectroscopy.<sup>13,49-51</sup> However, the effect of solvation on the distribution of the NF conformations is not well known. While it is evident that these NF conformations are unstable in water and rapidly transform into ZW, an analysis of the solvation of NF can be useful in understanding the NF→ZW transformation mechanism. Tortonda et al.<sup>68</sup> and Ke et al.<sup>67</sup> used quantum chemical methods in conjunction with a continuum model of the solvent to investigate the solvation

effect on neutral glycine. In the present work, we have applied ReaxFF to study this effect using larger sampling space and with a molecular model of the solvent that allows explicit consideration of the specific solute-solvent interactions.

Two NVT-MD simulations, one in the gas phase and one in water, were started from the gas phase global minimum, the nA conformation. After 1 ns of simulation, many inter-conversions between nA, nC, nD and nE could be observed in both simulations. However, transformation to nB (or the zwitterion) was not observed in either simulation. This is consistent with the fact that the energy barrier of the C–C–O–H dihedral angle is much higher than that of the N–C–C–O(H) or C–C–N–H dihedral angles (Figure 2). Experimental studies using matrix-IR spectroscopy also suggested that the intramolecular H bond, N---H–O–C, in nB is much stronger than the NH<sub>2</sub>---O H bond in nA and nC.<sup>50,51</sup> Consequently, only the nC → nA interconversion was observed.<sup>50</sup> Thus, in order to obtain samples of nB conformations, two other simulations, one in the gas phase and one in water, were started from the nB conformation. In both simulations, the glycine molecule remained in the nB conformation throughout the simulation time of 1 ns.

NF acquires its different conformations by changing the dihedral angles, N–C–C–O(H), C–C–O–H and C–C–N–H. Thus, we analyzed the effect of solvation on these three dihedral angle distributions (Figure 7). The N–C–C–O(H) torsional distributions for the gas phase and aqueous phase nB were comparable to each other. However, there is a clear distinction between the two in the case of nACDE. In the gas phase, the population density near  $\angle 180^\circ$  is significantly higher than near  $\angle 0^\circ$ . This indicates a higher population of “nA and nD” than “nC and nE”. On the other hand, in the aqueous phase, a preference of “nC and nE” over “nA and nD” is evident. This indicates that “nC and nE”, rather than “nA and nD”, are the relevant reference states for the transformation of the neutral form to a zwitterion in water. However, multiple interconversions between these states during our simulations in water indicate that these species are in rapid equilibrium. Bandyopadhyay et al. used a discrete/continuum model of the solvent to investigate the NF↔ZW conversion as well as the interconversions between the NF conformers in water.<sup>25</sup> They suggested that the nA→ZW transformation initiates with a nA→nC transformation characterized by a low energy barrier, and that transformation to the nB conformation precedes the final step, a nB→ZW proton transfer. Their calculations suggested that the transformation to nB was rate limiting among the NF interconversions (as well as the overall nA→ZW transformation). In our ReaxFF simulation, the spontaneous interconversions between the nA, nC, nD, nE and the absence of any interconversions involving nB are consistent with this conclusion.

The substantial localization in the C–C–O–H torsional distributions (Figure 7b) can be directly attributed to the high energy barrier of the corresponding distortion (Figure 2b). The absence of any interconversion involving nB is evident. The peak heights in the aqueous phase are higher than the corresponding peak heights in the gas phase in both the nB and nACDE distributions. This is probably due to the intermolecular hydrogen bonds between the glycine and the water molecules in the aqueous phase that impose additional restrictions on the deviations of the glycine molecule from its symmetrical local energy minima. Tortonda et al. suggested a strengthening of the intramolecular hydrogen bond N---H–O–C in nB (their IIp) upon solvation due to an increase in the polarity. This might be another possible explanation in the case of nB.

The C–C–N–X torsional distributions (Figure 7c) also provide useful information about the solvent effect on the conformational distribution of NF. X is an imaginary atom located exactly between the two amide hydrogen atoms. For the case of nB in the gas phase, most of the conformers resembled the nB structure. It is interesting to note that an asymmetric structure was preferred over a symmetric structure, as suggested by the peak near  $120^\circ$

instead of  $180^\circ$ . The QM optimized gas phase structure of nB was also slightly asymmetric as indicated by a C–C–N–X dihedral angle of  $\sim 165^\circ$ . The corresponding value produced by ReaxFF was  $\sim 122^\circ$ . The reason for the preference of the force field for a more asymmetric structure is not apparent. For the case of nB in the aqueous phase, a peak can be observed near  $0^\circ$  in addition to the peak near  $120^\circ$ . This peak near  $0^\circ$  corresponds to the nB' structure (Figure 8) which has previously been predicted as one of the high energy gas-phase glycine conformers.<sup>4-6,67</sup> A few of these structures are also observed in the gas phase (Figure 7b). For the case of nACDE, a wide range of structures were observed in the gas phase. The peaks between  $0^\circ$  and  $60^\circ$  correspond to the nA and nC conformers and the distribution above  $60^\circ$  corresponds to the nD and nE conformers. In the aqueous phase, the distribution indicates significant population of nD and nE in addition to nA and nC conformers.

Two conclusions about the solvent effect on NF can be drawn from the current study. First, the abundance of the structures nB', nD and nE in the aqueous phase, in addition to the dominant nA, nB and nC structures in the gas phase, suggests that the energy differences between the NF structures decrease due to the solvent. This is in good agreement with the quantum-chemical/continuum studies by Tortonda et al.<sup>68</sup> and Ke et al.<sup>67</sup> Second, the population distribution among the NF structures is altered due to the solvent effect. The population distribution involving the nB (or nB') conformer cannot be deduced from the current simulations due to the lack of interconversions to the other conformers. However, it is possible to determine and compare the population distributions among nA, nC, nD and nE in the gas phase and in water.

For the case of nACDE, the glycine molecule acquires its different conformers by changing the N–C–C–O(H) and C–C–N–X dihedral angles. Therefore, we constructed a torsional distribution as a function of these two dihedral angles in order to further illustrate the population distribution (Figure 9). In the gas phase, glycine was detected predominantly in the nA form, along with a few nC and nD conformers. The C–C–N–X vs N–C–C–O(H) torsional space was divided into four subsections corresponding to nA, nC, nD and nE and the number of occurrences were counted. The gas phase population consisted of 50.5% nA, 19.2% nC, 22.7% nD and 7.6% nE. This distribution is consistent with the relative energies of these conformers in the gas phase (Figure 3). The population distribution was significantly changed in the aqueous phase with nE being the dominant conformer. The population consisted of 9.0% nA, 18.2% nC, 16.9% nD and 55.9% nE. While the determination of the most stable NF conformer in water is not possible due to the lack of spontaneous interconversions with nB, we will apply more sophisticated sampling to explore this issue in our future work. However, the significant change in the conformational distribution of NF due to solvation suggests that a conformational isomerization from nA to nE might be the first step in the NF  $\rightarrow$  ZW tautomerization.

### 3.3 NF $\rightarrow$ ZW conversion in Water

We now turn to the proton transfer event that completes the tautomerization process. Which conformer of the NF is involved as the reactant in this step is not clear. The transition state(s) of the NF $\rightarrow$ ZW proton transfer is also not well identified in the literature. In particular, the water-mediated reactions require a proper alignment of the mediating water molecule(s) and the transferable proton(s) between the amine and the carboxyl groups. Such conformations are rare and were not observed in our time-limited MD simulations. Thus, in order to identify possible transition states, we examined the gas phase glycine-(water)<sub>2</sub> complexes. The nB configuration is clearly the most suitable reactant for the direct reaction due to the proximity of the carboxylic proton to the amine group, so we assumed nB as the reactant of the proton transfer reaction. The QM energy of nB-2w-d was the lowest among the eight nB-2w complexes and the proton is aligned for direct transfer, so this configuration was selected as the initial configuration for a simulation of the direct proton transfer

reaction. For the reaction mediated by one water, any one of nB-2w-f, nB-2w-g or nB-2w-h could be used, as they all have the proper alignment of a water molecule between the carboxyl and the amine group. We selected the lowest energy nB-2w-f conformer for this purpose. There was no neutral glycine-(water)<sub>2</sub> complex with two bridging water molecules between the carboxyl and the amine group, though two zwitterionic complexes, z-2w-b and z-2w-c, contain two bridging water molecules. We selected z-2w-c for the proton transfer simulation because the zwitterionic form of z-2w-b was unstable upon desolvation in water.

To estimate the energy profiles for direct, and water-mediated reactions, the procedure described in the Computational Methods section was followed, starting from the three complexes, nB-2w-d, nB-2w-f and z-2w-c, respectively. Both forward (NF→ZW) and reverse (ZW→NF) reactions were simulated. The product state was equilibrated until a steady energy profile was attained. A total simulation time of 1 ns was necessary for this purpose. While we recognize that the resulting energy profiles are distinct from free energy profiles, and that other reaction paths are possible, we have limited this initial study to these few cases.

The potential energy profiles of the three reaction mechanisms are shown in Figure 10. In each case, the potential energy of ZW was less than that of NF and the reactions went over substantial energy barriers, as expected from experimental observations<sup>22,23</sup>. Representative configurations of the reactant, transition state and product are shown in each figure. To maintain clarity, only the glycine and the mediating water molecules are shown.

The direct forward reaction required 2 ns of simulation to attain a steady product state. The reaction led to an intermediate that eventually transformed into the lower-energy product. For the reaction mediated by two water molecules, we started with a zwitterionic form: z-2w-c. Thus, the reverse reaction was simulated first, followed by the forward reaction. The average values for the activation energy and the energy difference between NF and ZW from the forward and reverse reactions were calculated for each case (Table 4). The reactant-product potential energy difference ( $\Delta E$ ) does not depend on the reaction pathway and the calculated  $\Delta E$ s for all three mechanisms agree within the fluctuations. While all of the activation energy barriers ( $\Delta E^\ddagger$ ) are comparable, those for the water-mediated reactions are clearly lower than that for the direct reaction. The one-water mediated mechanism produced the lowest  $\Delta E^\ddagger$  of  $7.04 \pm 0.35$  kcal/mol. The quoted error was estimated by calculating the energy difference of the forward and backward reactions and dividing by two.

Comparing  $\Delta E^\ddagger$  of the three reaction mechanisms suggests competitive reaction pathways, with a preference for the water-mediated pathways. This is similar to the conclusion reached by Balta and Ayivente<sup>18,21</sup> who used a DFT method to study the direct and water-mediated reaction mechanisms in glycine-water clusters containing up to 6 water molecules, and compared to a continuum solvent model. They concluded that the direct reaction mechanism could compete with the water-mediated reaction mechanisms.

The  $\Delta E^\ddagger$  values calculated in this work do not include the entropy contribution ( $T\Delta S^\ddagger$ ) and are not directly comparable to the experimental activation free energy ( $\Delta G^\ddagger$ ). For the reverse reaction (ZW→NF), Slifkin and Ali suggested a higher entropy contribution than the enthalpy contribution, based on a chemical relaxation study<sup>23</sup>. However, two recent computational studies concluded that the entropy contribution to  $\Delta G^\ddagger$  is minor. Using an Empirical Valence Bond (EVB) model in combination with MD simulation, Nagaoka et al. found that the entropy contribution is lower than the enthalpy contribution<sup>15</sup>. Using DFT with glycine-water complexes containing up to 6 water molecules, Balta and Ayivente concluded that the entropy contribution to  $\Delta G^\ddagger$  is rather small.<sup>21</sup> Considering this evidence,

we suggest that the water-mediated pathway is the preferred pathway based on the comparison of  $\Delta E^\ddagger$ . However, we cannot completely rule out the possibility that the inclusion of the entropy contribution might change this conclusion.

The water-mediated pathways (similar to the Grothuss mechanism<sup>40,41</sup>) of  $\text{NF} \rightarrow \text{ZW}$  conversion in bulk water are accessible to simulation because of the reactive nature of ReaxFF. In this initial study, we have explored this mechanism using a particular pathway selected from numerous possible routes of proton transfer. We note that a more realistic simulation of this mechanism would include a large number of possible states and account for the delocalization of the excess proton, similar to the multistate empirical valence bond (MS-EVB) approach by Voth and coworkers.<sup>69</sup> The ReaxFF force field developed in this work can be applied to investigate such a mechanism in the future.

## Conclusion

We have developed a reactive force field description of glycine that can describe the tautomerization process. We extended the training set of a previously developed hydrocarbon potential with several conformations of glycine. Rigorous conformational searches were performed using DFT to identify 12 glycine-water and 24 glycine-(water)<sub>2</sub> minima that were included in the training set. The force field parameters were optimized to reproduce the DFT-derived energies of all the complexes in the training set. The derived ReaxFF potential was tested for its ability to reproduce the gas-phase characteristics of glycine. The potential accurately reproduces different valence and dihedral angle distortion energies of gas-phase glycine. It satisfactorily estimates the relative energies of different conformations of gas-phase glycine and glycine-water complexes. The potential was examined for its ability to describe the gas-phase  $\text{ZW} \rightarrow \text{NF}$  transformation along the intrinsic reaction coordinate. The potential correctly recognizes NF as the lowest energy conformer in the gas phase with a barrierless  $\text{ZW} \rightarrow \text{NF}$  energy profile.

The potential was applied to investigate the effect of solvation on the conformational equilibrium of NF. Molecular dynamics simulation studies suggested that nB has high barriers to interconversion with other conformations in the gas phase and in water. The simulation starting with nA produced an ensemble of nA, nC, nD and nE with many interconversions among them, both in the gas phase and in water. However, nA was the dominant conformer in the gas phase while nE was the dominant conformer in water.

The potential was then applied to derive the energy profiles of the proton transfer reaction in water. The direct mechanism and paths mediated by one or two water molecules were investigated. The simulations correctly predicted the ZW as the lower-energy species in water. The activation energy barrier for the direct mechanism is a few kcal/mol higher than the water-mediated mechanisms. The one-water mediated mechanism had the lowest activation energy barrier,  $7.04 \pm 0.35$  kcal/mol. This suggests that the  $\text{NF} \rightarrow \text{ZW}$  proton transfer reaction in water is most likely mediated by a single water molecule.

The reactive potential developed in this work provides a computationally inexpensive tool to investigate the glycine tautomerization with adequate sampling of the configurations in a large-scale dynamic simulation. This potential can be applied to characterize many aspects of this biologically important phenomenon. For instance, it can be applied to investigate charge separation, variation of solvation structures, identification of lowest energy transition states and reaction pathways etc. This potential can be extended to describe the tautomerization reactions in other amino acids. The force field developed in this work are a first step in constructing a realistic model to accurately describe complex organic reactions in aqueous environment and to assess the protein protonation states during reactive events.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

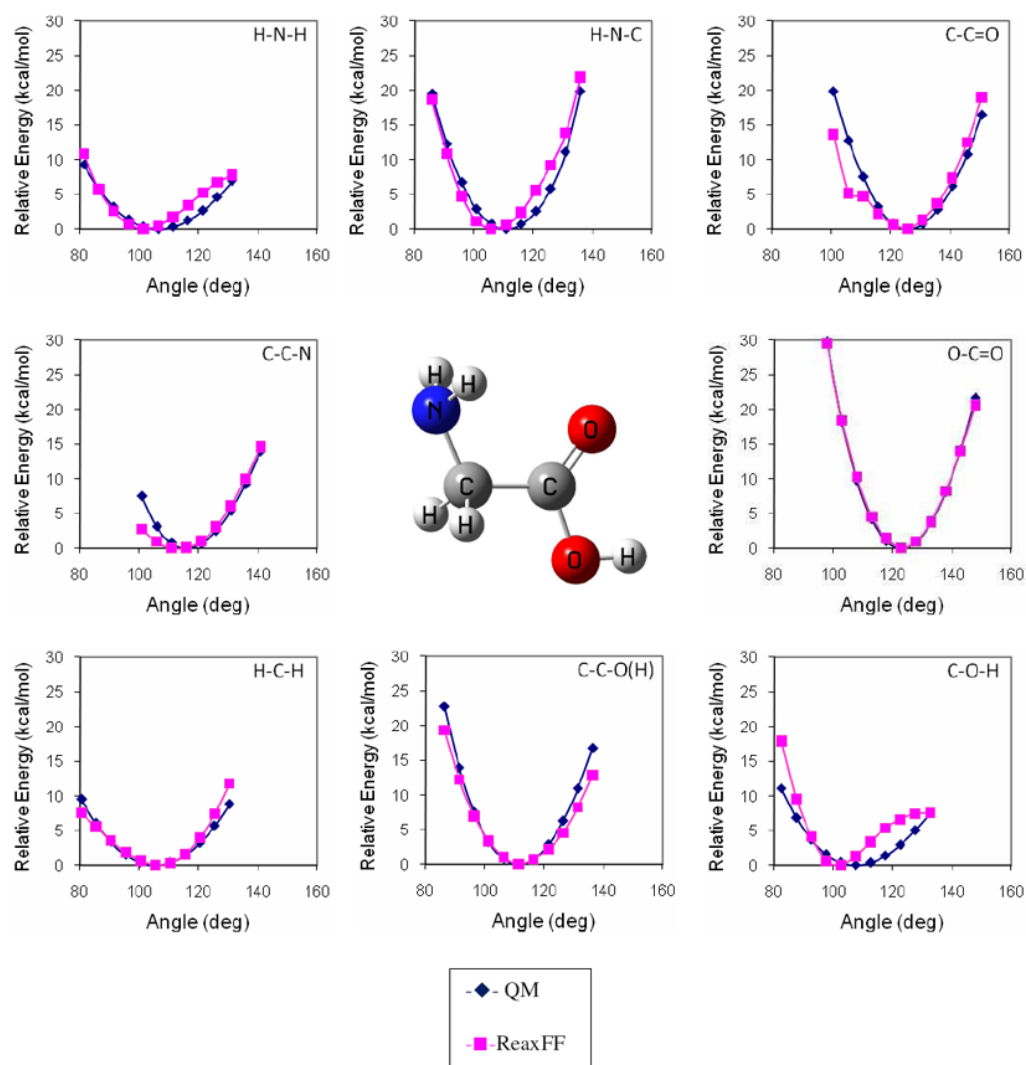
## Acknowledgments

ACTvD and WAG acknowledge NSF-funding DMR 0427177: ITR for the development of the ReaxFF water-potential. ACTvD acknowledges funding from KISK startup grant #C000032472. DJD and OR acknowledge support from NIH under P42 ES010344-06A2.

## References

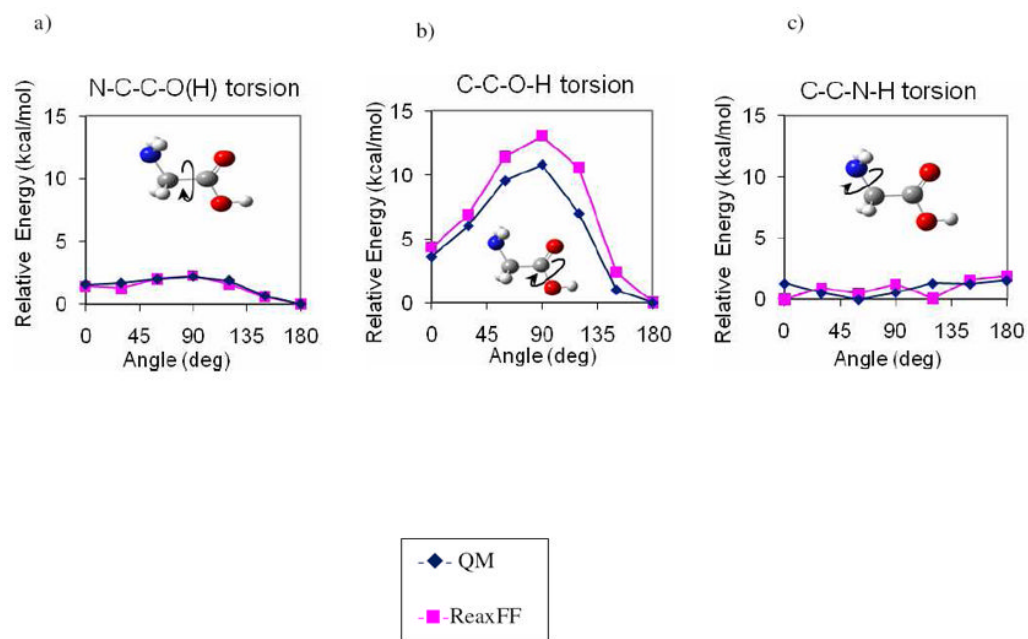
1. Cuello LG, Romero JG, Cortes DM, Perozo E. *Biochemistry-Us*. 1998; 37:3229.
2. DeCoursey TE. *Physiol Rev*. 2003; 83:1067.
3. Nimigean CM, Chappie JS, Miller C. *Biochemistry-Us*. 2003; 42:9263.
4. Jensen JH, Gordon MS. *J Am Chem Soc*. 1991; 113:7917.
5. Csaszar AG. *J Am Chem Soc*. 1992; 114:9568.
6. Hu CH, Shen MZ, Schaefer HF. *J Am Chem Soc*. 1993; 115:2923.
7. Zhang K, Chung-Phillips A. *Journal of Physical Chemistry A*. 1998; 102:3625.
8. McGlone SJ, Elmes PS, Brown RD, Godfrey PD. *J Mol Struct*. 1999; 485:225.
9. Chaban GM, Gerber RB. *Journal of Chemical Physics*. 2001; 115:1340.
10. Wang W, Zheng W, Pu X, Wong N, Tian A. *Journal of Molecular Structure (Theochem)*. 2002; 618:235.
11. Chaudhari A, Sahu PK, Lee SL. *Journal of Chemical Physics*. 2004; 120:170. [PubMed: 15267274]
12. Miller TF, Clary DC. *Phys Chem Chem Phys*. 2004; 6:2563.
13. Reva ID, Plokhotnichenko AM, Stepanian SG, Ivanov AY, Radchenko ED, Sheina GG, Blagoi YP. *Chem Phys Lett*. 1995; 232:141.
14. Ding YB, KroghJespersen K. *J Comput Chem*. 1996; 17:338.
15. Nagaoka M, Okuyama-Yoshida N, Yamabe T. *Journal of Physical Chemistry A*. 1998; 102:8202.
16. Okuyama-Yoshida N, Nagaoka M, Yamabe T. *Journal of Physical Chemistry A*. 1998; 102:285.
17. Kassab E, Langlet J, Evleth E, Akacem Y. *J Mol Struc-Theochem*. 2000; 531:267.
18. Balta B, Aviyente V. *J Comput Chem*. 2003; 24:1789. [PubMed: 12964198]
19. Selvarengan P, Kolandaivel P. *J Mol Struc-Theochem*. 2004; 671:77.
20. Leung K, Rempe SB. *Journal of Chemical Physics*. 2005:122.
21. Balta B, Aviyente V. *J Comput Chem*. 2004; 25:690. [PubMed: 14978712]
22. Wada G, Tamura E, Okina M, Nakamura M. *B Chem Soc Jpn*. 1982; 55:3064.
23. Slifkin MA, Ali SM. *J Mol Liq*. 1984; 28:215.
24. Tunon I, Silla E, Ruiz-Lopez MF. *Chem Phys Lett*. 2000; 321:433.
25. Bandyopadhyay P, Gordon MS, Mennucci B, Tomasi J. *Journal of Chemical Physics*. 2002; 116:5023.
26. Bachrach SM. *Journal of Physical Chemistry A*. 2008; 112:3722.
27. Aqvist J, Warshel A. *Chem Rev*. 1993; 93:2523.
28. Hwang JK, King G, Creighton S, Warshel A. *J Am Chem Soc*. 1988; 110:5297.
29. Warshel, A. *Computer Modeling of Chemical Reactions in Enzymes and Solutions*. Wiley and Sons, Inc.; New York: p. 1991
30. Warshel A, Russell S. *J Am Chem Soc*. 1986; 108:6569.
31. Warshel A, Weiss RM. *J Am Chem Soc*. 1980; 102:6218.
32. Chang YT, Miller WH. *J Phys Chem-Us*. 1990; 94:5884.
33. Miller WH, Chang Y-T, Makri N. *Molecular Structure and Reactivity in Computational Advances in Organic Chemistry*. :1991.

34. Cui Q. *Journal of Chemical Physics*. 2002; 117:4720.
35. Shoeib T, Ruggiero GD, Siu KWM, Hopkinson AC, Williams IH. *Journal of Chemical Physics*. 2002; 117:2762.
36. Takayanagi T, Yoshikawa T, Kakizaki A, Shiga M, Tachikawa M. *J Mol Struct-Theochem*. 2008; 869:29.
37. van Duin ACT, Dasgupta S, Lorant F, Goddard WA. *Journal of Physical Chemistry A*. 2001; 105:9396.
38. van Duin ACT, Damste JSS. *Org Geochem*. 2003; 34:515.
39. Chenoweth K, van Duin ACT, Goddard WA. *Journal of Physical Chemistry A*. 2008; 112:1040.
40. Agmon N. *Chem Phys Lett*. 1995; 244:456.
41. von Grotthuss CJD. *Ann Chim*. 1806:54.
42. Mortier WJ, Ghosh SK, Shankar S. *J Am Chem Soc*. 1986; 108:4315.
43. van Duin ACT, Strachan A, Stewman S, Zhang QS, Xu X, Goddard WA. *Journal of Physical Chemistry A*. 2003; 107:3803.
44. Lee CT, Yang WT, Parr RG. *Phys Rev B*. 1988; 37:785.
45. Becke AD. *Journal of Chemical Physics*. 1993; 98:5648.
46. Frisch, MJ. Gaussian 03, Revision C. 02. Gaussian, Inc.; Wallingford CT: 2004.
47. Chenoweth K, van Duin ACT, Persson P, Cheng MJ, Oxgaard J, Goddard WA. *Journal of Physical Chemistry C*. 2008; 112:14645.
48. <http://lammps.sandia.gov/>
49. Ivanov AY, Plokhhotnichenko AM, Izvekov V, Sheina GG, Blagoi YP. *J Mol Struct*. 1997; 408:459.
50. Stepanian SG, Reva ID, Radchenko ED, Rosado MTS, Duarte MLTS, Fausto R, Adamowicz L. *Journal of Physical Chemistry A*. 1998; 102:1041.
51. Ivanov AY, Sheina G, Blagoi YP. *Spectrochim Acta A*. 1999; 55:219.
52. Sellers HL, Schafer L. *J Am Chem Soc*. 1978; 100:7728.
53. Schafer L, Sellers HL, Lovas FJ, Suenram RD. *J Am Chem Soc*. 1980; 102:6566.
54. Suenram RD, Lovas FJ. *J Am Chem Soc*. 1980; 102:7180.
55. Iijima K, Tanaka K, Onuma S. *J Mol Struct*. 1991; 246:257.
56. Jensen JH, Gordon MS. *J Am Chem Soc*. 1995; 117:8159.
57. Alonso JL, Cocinero EJ, Lesarri A, Sanz ME, Lopez JC. *Angew Chem Int Edit*. 2006; 45:3471.
58. Chaudhari A, Lee SL. *Chem Phys*. 2005; 310:281.
59. Selvarengan P, Kolandaivel P. *Int J Quantum Chem*. 2006; 106:1001.
60. Park SW, Im S, Lee S, Desfrancois C. *Int J Quantum Chem*. 2007; 107:1316.
61. Brown RD, Godfrey PD, Storey JWV, Bassez MP. *J Chem Soc Chem Comm*. 1978:547.
62. Suenram RD, Lovas FJ. *J Mol Spectrosc*. 1978; 72:372.
63. Godfrey PD, Brown RD. *J Am Chem Soc*. 1995; 117:2019.
64. Godfrey PD, Brown RD, Rodgers FM. *J Mol Struct*. 1996; 376:65.
65. Bludsky O, Chocholousova J, Vacek J, Huisken F, Hobza P. *Journal of Chemical Physics*. 2000; 113:4629.
66. Miller TF, Clary DC, Meijer AJHM. *Journal of Chemical Physics*. 2005; 122
67. Ke H-W, Rao L, Xu X, Yan Y-J. *Journal of Theoretical and Computational Chemistry*. 2008; 7:889.
68. Tortonda FR, Pascual-Ahuir JL, Silla E, Tunon I. *J Mol Struct-Theochem*. 2003; 623:203.
69. Voth GA. *Acc Chem Res*. 2006; 39:143. [PubMed: 16489734]
70. Fogarty JC, Aktulga HM, Grama AY, van Duin ACT, Pandit SA. *J Chem Phys*. 2010; 132:174704. [PubMed: 20459180]

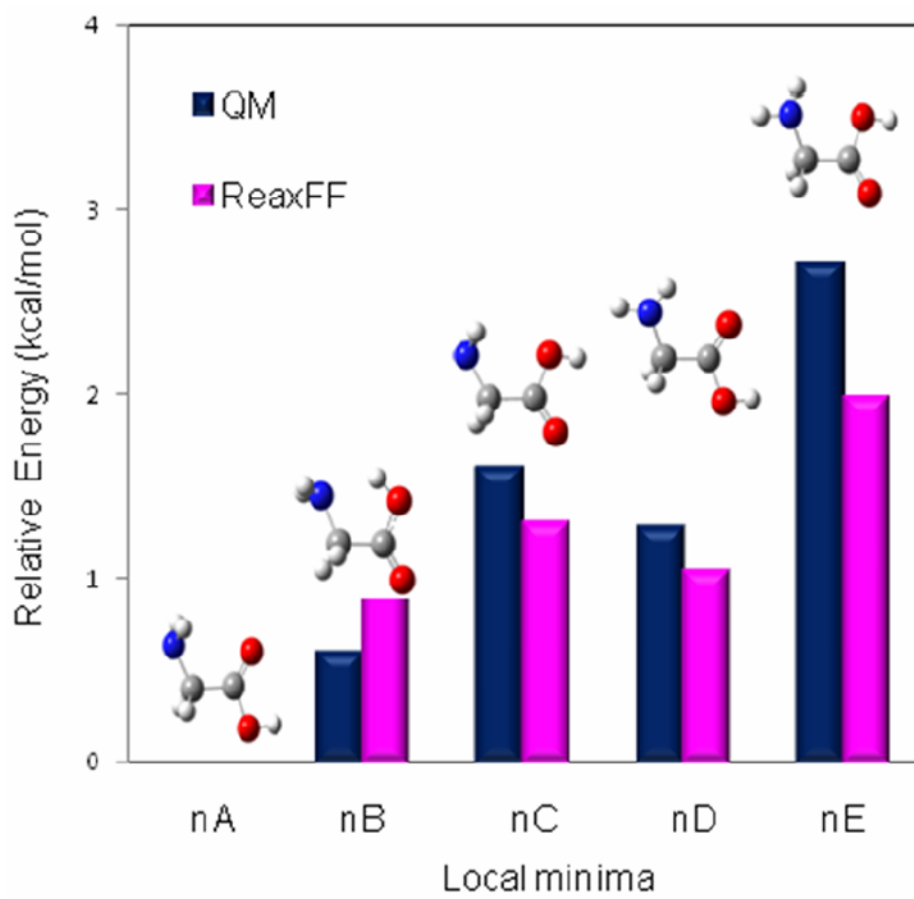


**Figure 1.** QM and ReaxFF energies for the distortion of various angles of glycine in gas phase.

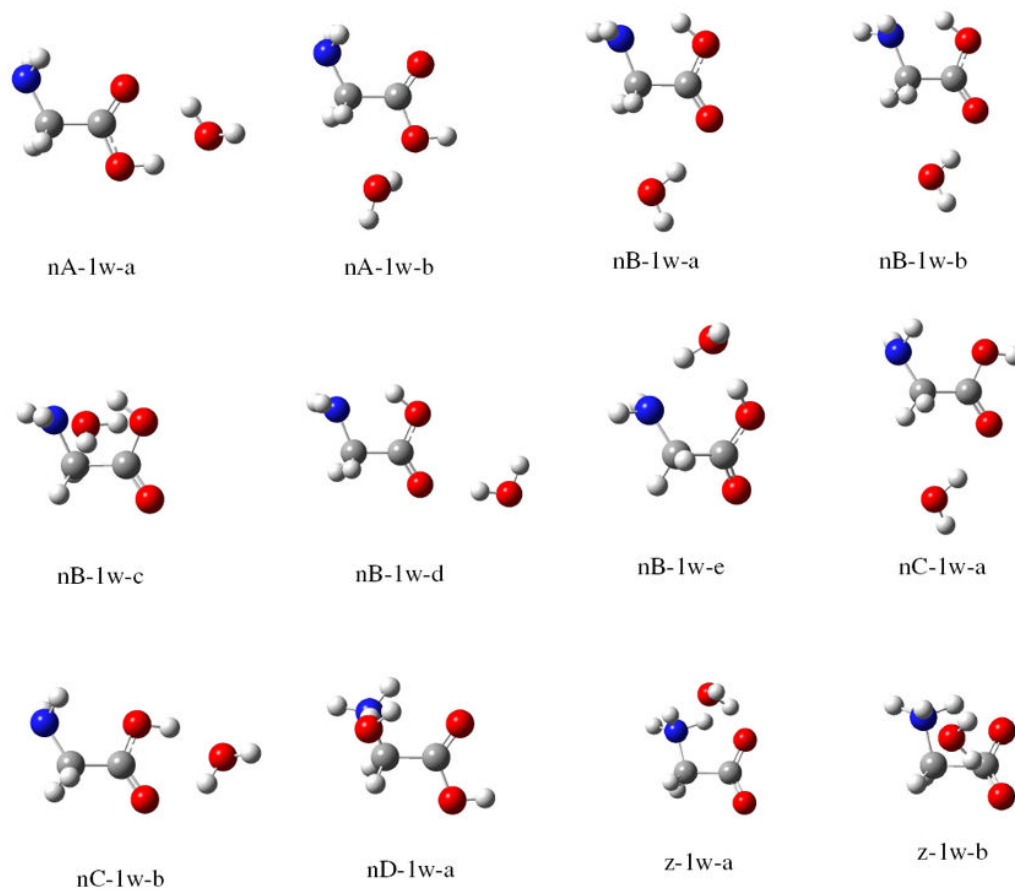




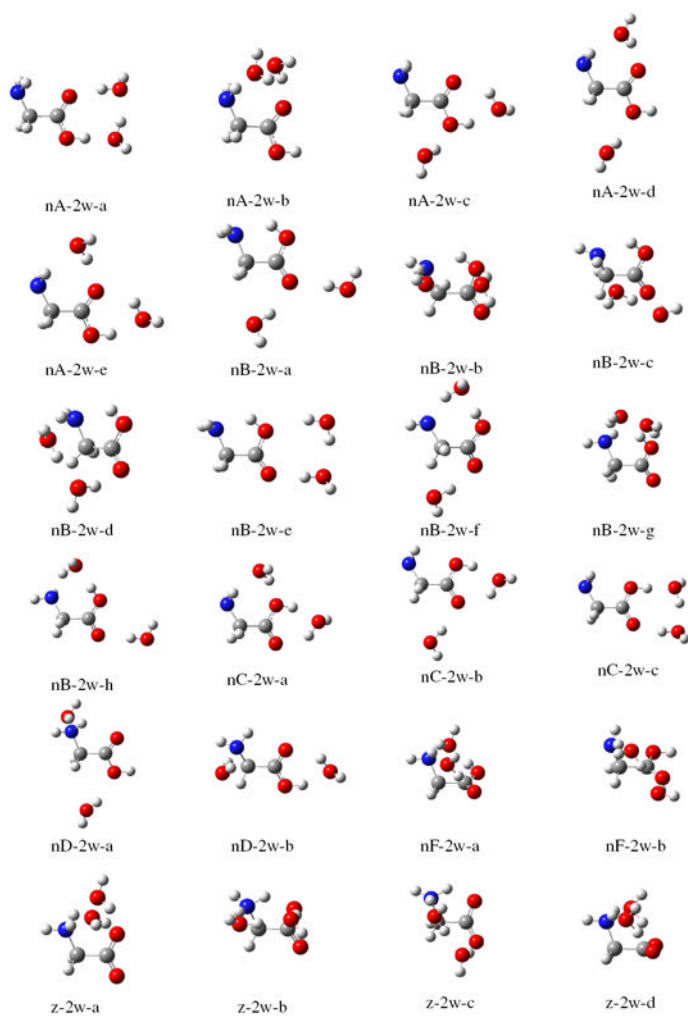
**Figure 2.** QM and ReaxFF energies for different dihedral angles of glycine in gas phase.



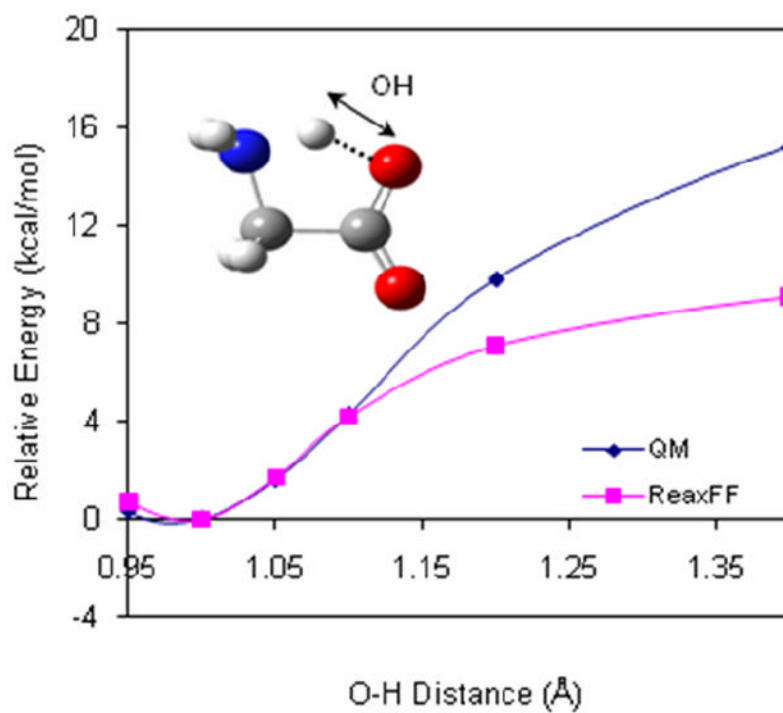
**Figure 3.** QM and ReaxFF energies of the five glycine minima in the gas phase.



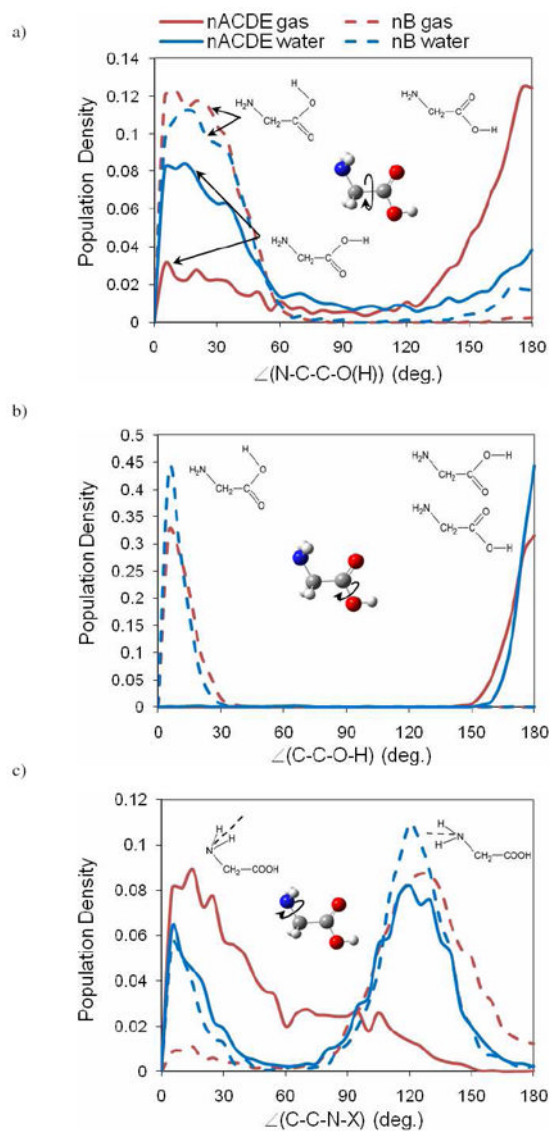
**Figure 4.**  
QM minimized structures of the glycine-water complexes.



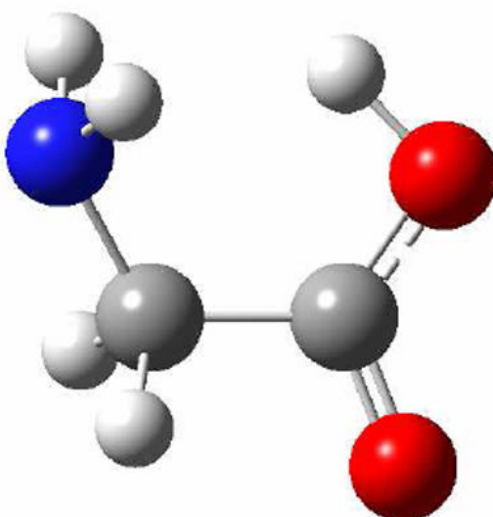
**Figure 5.**  
QM minimized structures of the glycine-(water)<sub>2</sub> complexes.



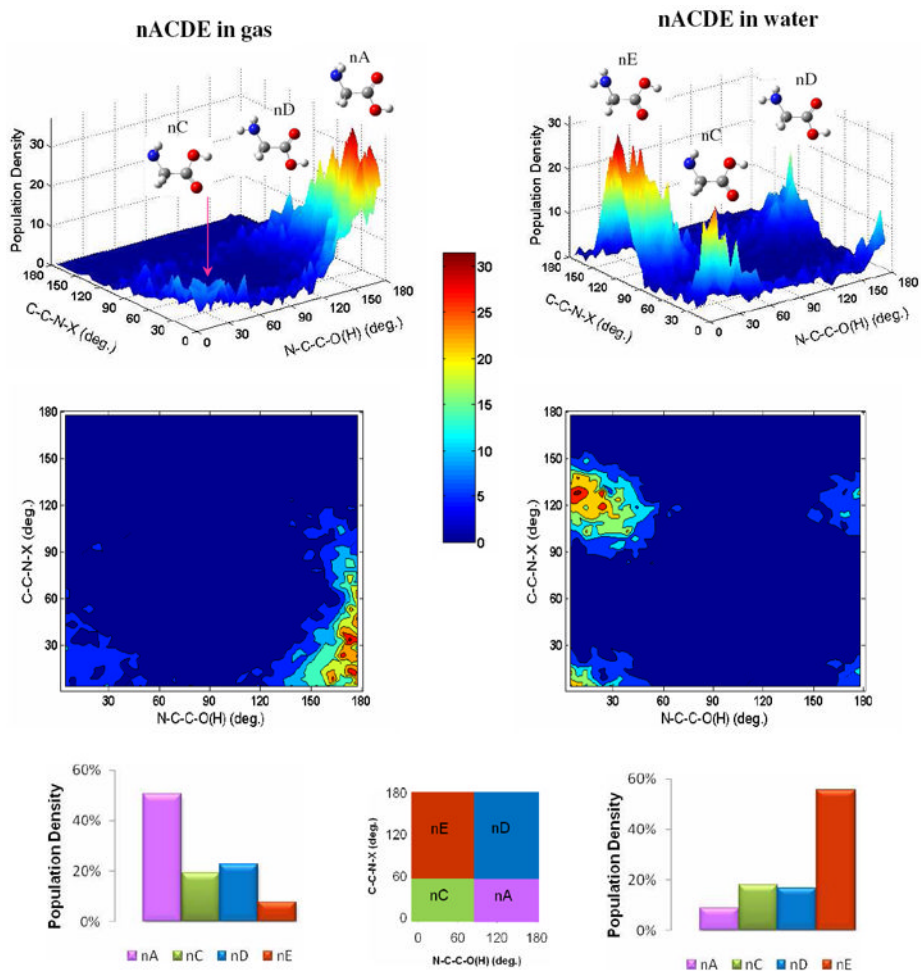
**Figure 6.** QM and ReaxFF relative energies of the gas-phase ZW→NF proton transfer. The glycine conformer changes from NF on the left to ZW on the right.



**Figure 7.** Dihedral angle distributions of NF in the gas phase and in water. nACDE indicates the distribution from a simulation started from nA that shows rapid interconversions among nA, nC, nD and nE. nB indicates the distribution from a simulation started from nB.

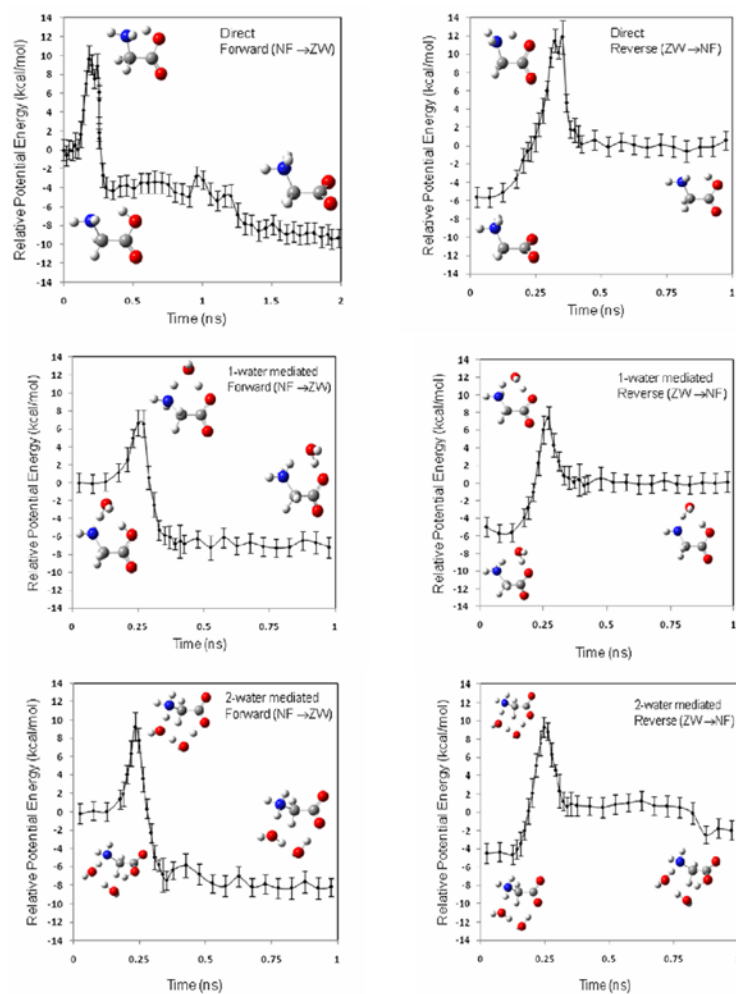


**Figure 8.**  
The structure of  $nB'$ .



**Figure 9.** Effect of solvation on nACDE, the distribution from a simulation started from nA that shows rapid interconversions among nA, nC, nD and nE. nA and nE are the most frequent conformers in the gas phase and aqueous phase, respectively.





**Figure 10.** Energy profiles of the direct, one-water mediated and two-water mediated NF-ZW proton transfer reactions. Both the forward (NF→ZW) and reverse (ZW→NF) reactions were simulated.

**Table 1**

Relative energies of glycine conformers in the gas phase as compared to other studies

Conformer	QM relative energies (kcal/mol)			ReaxFF relative energies (kcal/mol)
	This work ub3lyp/6-311++g(2df,2p)	Ref. <sup>12</sup> MP2/6-311++G**	Ref. <sup>7</sup> MP2/6-31+g(2d,2p)//MP2/6-31+G**	This work
nA	0.00	0.00	0.00	0.00
nB	0.59	0.53	0.39	0.97
nC	1.60	1.60	1.59	1.29
nD	1.28	1.27	1.31	1.03
nE	2.70	2.20	2.49	1.98

**Table 2**

QM and ReaxFF relative energies of glycine-water complexes in the gas phase

Conformer	Relative energies (kcal/mol)		
	QM	ReaxFF	RMSD (Å)
nA-1w-a	0.00	0.13	0.07
nA-1w-b	6.67	2.97	0.06
nB-1w-a	3.41	2.62	0.11
nB-1w-b	3.41	2.29	0.12
nB-1w-c	5.15	1.08	0.27
nB-1w-d	4.91	0.88	0.16
nB-1w-e	5.09	4.21	0.13
nC-1w-a	5.99	1.92	0.16
nC-1w-b	1.21	0.00	0.12
nD-1w-a	4.27	2.55	0.09
z-1w-a	19.86	14.61	0.18
z-1w-b	22.53	15.68	0.37

**Table 3**QM and ReaxFF relative energies of glycine-(water)<sub>2</sub> complexes in the gas phase

Conformer	Relative energies (kcal/mol)		
	QM	ReaxFF	RMSD (Å)
nA-2w-a	0.00	0.00	0.10
nA-2w-b	10.51	3.34	0.22
nA-2w-c	7.93	6.19	0.06
nA-2w-d	12.56	5.97	0.14
nA-2w-e	5.68	0.63	0.17
nB-2w-a	9.72	4.44	0.18
nB-2w-b	6.04	2.26	0.15
nB-2w-c	5.15	3.90	0.24
nB-2w-d	4.99	4.32	0.25
nB-2w-e	9.22	2.52	0.15
nB-2w-f	9.38	6.89	0.11
nB-2w-g	10.02	6.17	0.17
nB-2w-h	10.76	5.14	0.15
nC-2w-a	8.18	2.19	0.18
nC-2w-b	7.40	3.79	0.14
nC-2w-c	1.09	0.39	0.09
nD-2w-a	12.48	7.89	0.10
nD-2w-b	5.70	4.68	0.15
nF-2w-a	7.85	6.38	0.22
nF-2w-b	7.67	2.16	0.16
z-2w-a	13.89	15.96	0.22
z-2w-b	12.92	14.30	0.18
z-2w-c	13.64	15.41	0.19
z-2w-d	13.93	16.82	0.21

**Table 4**

The NF-ZW energy differences ( $\Delta E$ ) and the activation energy barriers ( $\Delta E^\ddagger$ ) for the NF $\rightarrow$ ZW transformation reaction

	$\Delta E$ (kcal/mol)	$\Delta E^\ddagger$ (kcal/mol)
Direct	$7.53 \pm 1.89$	$10.78 \pm 1.11$
One-water mediated	$6.52 \pm 0.79$	$7.04 \pm 0.35$
Two-water mediated ( $\Delta E$ )	$6.85 \pm 1.52$	$8.90 \pm 0.37$