

# L-Alanine in a Droplet of Water: A Density-Functional Molecular Dynamics Study

Ivan M. Degtyarenko,<sup>\*,†</sup> Karl J. Jalkanen,<sup>‡</sup> Andrey A. Gurtovenko,<sup>§</sup> and Risto M. Nieminen<sup>†</sup>

Laboratory of Physics, Helsinki University of Technology, P.O.B. 1100, Helsinki FIN-02015 HUT, Finland, Nanochemistry Research Institute, Department of Applied Chemistry, Curtin University of Technology, G.P.O. Box U1987, Perth 6845, Western Australia, and Computational Laboratory, Institute of Pharmaceutical Innovation, University of Bradford, Bradford, West Yorkshire BD7 1DP, U.K.

Received: November 20, 2006; In Final Form: February 21, 2007

We report the results of a Born–Oppenheimer molecular dynamics study on an L-alanine amino acid in neutral aqueous solution. The whole system, the L-alanine zwitterion and 50 water molecules, was treated quantum mechanically. We found that the hydrophobic side chain ( $R = \text{CH}_3$ ) defines the trajectory path of the molecule. Initially fully hydrated in an isolated droplet of water, the amino acid moves to the droplet's surface, exposing its hydrophobic methyl group and  $\alpha$ -hydrogen out of the water. The structure of an L-alanine with the methyl group exposed to the water surface was found to be energetically favorable compared to a fully hydrated molecule. The dynamic behavior of the system suggests that the first hydration shell of the amino acid is localized around carboxylate ( $\text{CO}_2^-$ ) and ammonium ( $\text{NH}_3^+$ ) functional groups; it is highly ordered and quite rigid. In contrast, the hydration shell around the side chain is much less structured, suggesting a modest influence of the methyl group on the structure of water. The number of water molecules in the first hydration shell of an alanine molecule is constantly changing; the average number was found to equal 7. The molecular dynamics results show that L-alanine in water does not have a preferred conformation, as all three of the molecule's functional sites (i.e.,  $\text{CH}_3$ ,  $\text{NH}_3^+$ ,  $\text{CO}_2^-$ ) perform rotational movements around the  $\text{C}_\alpha$ -site bond.

## 1. Introduction

The most biologically relevant form of amino acids is the zwitterion, as this is their most abundant form in living organisms: amino acids are synthesized in the zwitterionic form and are essentially always in this form at neutral pH.<sup>1,2</sup> Zwitterionic species of amino acids have both a positively charged ammonium group ( $\text{NH}_3^+$ ) and a negatively charged carboxylate group ( $\text{COO}^-$ ). They are the dominant form in aqueous solution over a wide range of pH.<sup>3</sup> In contrast, in the gas phase, where interactions with environment are not present, amino acids are mostly in their neutral nonionic form.<sup>4–6</sup> The fact that amino acids are not zwitterions in the gas phase, but are in water, implies that interactions with the water environment are a key determinant of the stable zwitterionic structure. Water is essential for all known forms of life, and a complete understanding of hydration effects is one of the major problems in biology. In this work, we focus on one example of the dramatic effect of water on amino acids, namely, the solvation of a single L-alanine (LA) molecule in a droplet of water. Alanine is the smallest chiral  $\alpha$ -amino acid, with a nonreactive methyl group ( $-\text{CH}_3$ ) as the side chain. It is one of the aliphatic amino acids that has the important property of not interacting favorably with water.<sup>7</sup> It is considered to be hydrophobic and is normally placed on the hydrophobicity scale after glycine, leucine, isoleucine, and valine.<sup>8</sup> L-Alanine conformers in aqueous media, isolated, and crystalline are shown in Figure 1.

To study the effect of hydration, we used the density-functional theory Born–Oppenheimer molecular dynamics

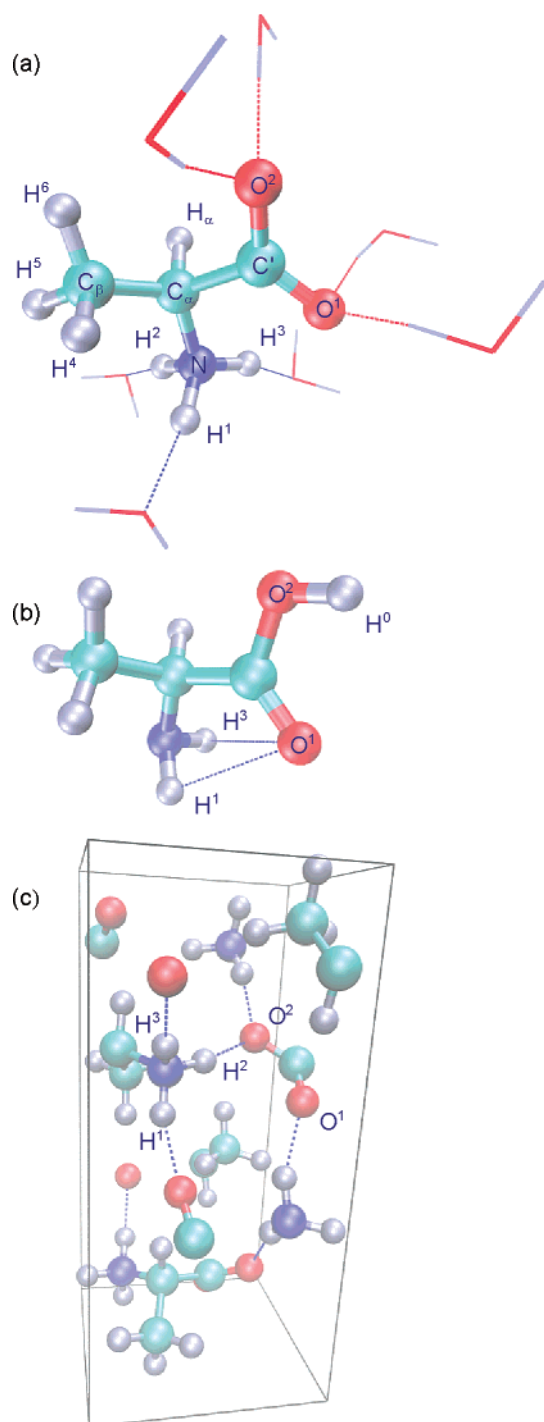
(DFT–BOMD) approach. The whole system, alanine and water molecules, was treated fully quantum mechanically. Within the scope of this work, we determine the structure and properties of the hydration shell around alanine and describe the zwitterion dynamics and its structural stability in a polar water environment. The motivation for this study arose for several reasons. It is now well-established that the effect of solvation has to be included in simulations to achieve physical meaningful results,<sup>9</sup> especially for vibrational spectra.<sup>10–13</sup> Nevertheless, the L-alanine zwitterion in aqueous solution still remains poorly studied. Its structural parameters are known mainly from computational studies.<sup>10,12,14</sup> The widely used experimental zwitterionic structure of alanine, which is normally assumed to also be the structure in aqueous media, is derived from solid-state crystallographic data. However, the natures and origins of the zwitterionic form of alanine in a crystal and in water are different.<sup>15</sup> Infrared, Raman, and neutron scattering spectroscopic techniques have been used to provide information concerning the structure of the hydration shell and interactions between water and amino acid residues in alanine aqueous solutions, but the results are ambiguous. Hetch et al. observed correlations between solute-induced perturbations of the solvent structure and amino acid hydrophobicity.<sup>16</sup> Later, Ide et al. concluded, also on the basis of Raman spectroscopy, that the structure of water in solutions of various amino acids at neutral pH does not depend on the nature of the amino acid side chains.<sup>17</sup> Recently, Kameda et al. obtained data on the number of water molecules inside the hydration shell and nearest-neighbor distances by means of neutron diffraction with the isotopic substitution technique.<sup>18</sup> They determined that the hydration numbers for the ammonium group differ for glycine and alanine, being 3.0 and 2.4, respectively,<sup>18a,19</sup> which means that amino acid hydration structure depends on the type of

\* To whom correspondence should be addressed. E-mail: imd@fyslab.hut.fi.

<sup>†</sup> Helsinki University of Technology.

<sup>‡</sup> Curtin University of Technology.

<sup>§</sup> University of Bradford.



**Figure 1.** Ball-and-stick representations of L-alanine: (a) An L-alanine molecule within its first hydration shell; dotted lines indicate hydrogen bonds. (b) An isolated nonionic form of L-alanine, the most stable gas-phase conformer;<sup>4</sup> the COOH group adopts a cis configuration, and an intramolecular hydrogen bond is formed between the amino group and the carbonyl oxygen. (c) L-Alanine in crystal form, showing one unit cell with the species constrained to periodic boundary conditions, adopted from Lehmann et al.<sup>15b</sup>

molecule. Several other experimental studies on the structural properties of water in aqueous solutions of amino acids exist.<sup>20</sup>

The remaining information related to the structure and dynamic properties of amino acids in aqueous solutions comes predominantly from computer simulations. Structural parameters, vibrational spectra, and stabilities of zwitterions in aqueous media have been exhaustively studied using *ab initio* methods.<sup>10–12,14,21–23</sup> Empirical approaches have been used to study

the dielectric properties of alanine in aqueous solution,<sup>24</sup> the hydration properties of the alanine side-chain analogue methane,<sup>25</sup> and the water distribution at the charged functional groups of the alanine zwitterion.<sup>26</sup> The lack of a first-principles study, which would fully cover the aspects of the dynamics of hydrated alanine amino acids, can be explained by the computational complexity of *ab initio* methods required to simulate systems with large numbers of atoms (Ehrenfest, Car–Parrinello, or first-principles Born–Oppenheimer MD simulations). Recently, an analysis of the conformational dynamics of an alanine dipeptide analogue in the gas phase<sup>27</sup> and a Car–Parrinello molecular dynamics study of the effect of protonation in a hydrated glycine molecule have been reported.<sup>28</sup>

Apart from the computational complexity, explicit incorporation of limited quantities of solvent molecules into simulations entails a principle difficulty: one has to guess the number of water molecules and their positions and orientations around an amino acid.<sup>29</sup> Even for small systems, there must be a very large number of structures having local minima. One way to make an average over these orientations is to perform a full molecular dynamics simulation. Although first-principles molecular dynamics simulations are computationally costly, they have several advantages over empirical potential methods. Concerning amino acids, by treating the system quantum mechanically, we are able to directly observe a proton transfer, which can occur from the ammonium group to the carboxylate group or via water molecule(s).<sup>21a</sup> The second advantage is related to the accuracy of first-principles methods. Instantaneous quantum electronic effects, such as polarization, correlation, and charge distortion, have a significant influence on the dynamic behavior. The widely used empirical force fields are not able to reproduce alanine dipeptide conformational changes<sup>27</sup> and fail to predict the hydration structure and dynamic properties of the COO<sup>−</sup> functional group, which is a crucial part of zwitterion amino acids.<sup>30</sup>

Here, we present Born–Oppenheimer molecular dynamics simulations of an L-alanine zwitterion solvated in water. This study is the first in which the whole system of relatively large size (the L-alanine amino acid and 50 water molecules) has been treated fully quantum mechanically. We performed simulations of the amino acid molecule in an isolated water droplet rather than in the bulk water; this allows the hydrophobic effect to be studied. In this work, the number of solvent molecules was sufficient to accommodate two solvation shells of an L-alanine amino acid molecule. The first hydration shell stabilizes the structure of a zwitterionic amino acid, and the second fully covers the amino acid molecule along with its first hydration shell.

## 2. Computational Methods and Structure

In this study, we considered the L isomer of alanine because this is the only natural form of alanine found in proteins. For the zwitterionic form of alanine, we started with the most energetically favorable conformer, as reported in recent computational studies.<sup>31</sup>

Simulations were performed in several steps: (i) initial system relaxation with a large number of water molecules using classical molecular dynamics (MD) simulations, (ii) structure optimization of an extracted alanine molecule with a limited number of the water molecules nearest to the alanine molecule, and (iii) density-functional theory Born–Oppenheimer molecular dynamics simulations starting from the minimum-energy structure of step ii.

As a first step, we performed classical atomic-scale MD simulations of an L-alanine molecule in a box of water with the

use of an empirical force field to obtain an equilibrated starting structure for ab initio molecular dynamics simulations. Force-field parameters for the L-alanine amino acid were taken from the full-atom Gromacs force field supplied within the GROMACS suite (ffgmx2 set).<sup>32</sup> The molecule of L-alanine was solvated in a box with 506 water molecules. Water was modeled using the simple point charge (SPC) model.<sup>33</sup> The MD simulations were performed for a total time of 100 ps in the  $NpT$  ensemble with temperature being kept constant using the Berendsen thermostat<sup>34</sup> with a coupling constant of 0.1 ps; temperature was set to 300 K. Pressure with a reference value of 1 bar was controlled by the Berendsen barostat with a coupling constant of 1.0 ps.<sup>34</sup>

The Lennard-Jones interactions were cut off at 1 nm. For the long-range electrostatic interactions, we used the particle-mesh Ewald (PME) method.<sup>35</sup> Both a neighbors list and the long-range interactions were updated every five time steps. All bond lengths of the amino acid were constrained using the LINCS algorithm,<sup>36</sup> and the SETTLE scheme<sup>37</sup> was used to constrain the water geometry. The time step used was 1 fs, and the simulations were performed using GROMACS.<sup>32</sup>

The final structure after the 100-ps MD simulation run was then used for preparation of the initial structure of an L-alanine molecule in a droplet of water by removing water molecules that were farther than 6 Å from any atom of the alanine molecule. The radial distribution functions extracted from classical MD simulations showed that the structural influence of the zwitterionic form of alanine extends up to 5–5.5 Å from the molecule. In total, only 50 water molecules were eventually left around the L-alanine amino acid, which was a sufficient amount to form two hydration shells.

The first-principles calculations were carried out using a numerical atomic orbitals density-functional theory (DFT) approach<sup>38</sup> that was recently developed and designed for efficient linear scaling calculations in large systems and implemented in the SIESTA code.<sup>39</sup> We used the generalized gradient approximation (GGA) to DFT and, in particular, the exchange-correlation functional of Perdew, Burke, and Ernzerhof (PBE).<sup>40</sup> Mainly, the choice of this functional was based on its reliability in the description of hydrogen bonds.<sup>41,42</sup> Only the valence electrons were considered in the calculation, with the core being replaced by norm-conserving scalar relativistic pseudopotentials factorized in the Kleinman–Bylander form.<sup>43,44</sup> We used a split-valence double- $\zeta$  basis set including polarization orbitals for all atoms, as obtained with an energy shift of 50 meV.<sup>45</sup> The integrals of the self-consistent terms of the Kohn–Sham Hamiltonian were obtained with the help of a regular real-space grid on which the electron density was projected. The grid spacing is determined by the maximum kinetic energy of the plane waves that can be represented in that grid. In the present work, we used a cutoff of 150 Ry, which yields a spacing between the grid points of around 0.13 Å. We checked that the results were well converged with respect to the real-space grid and the range of the atomic orbitals.

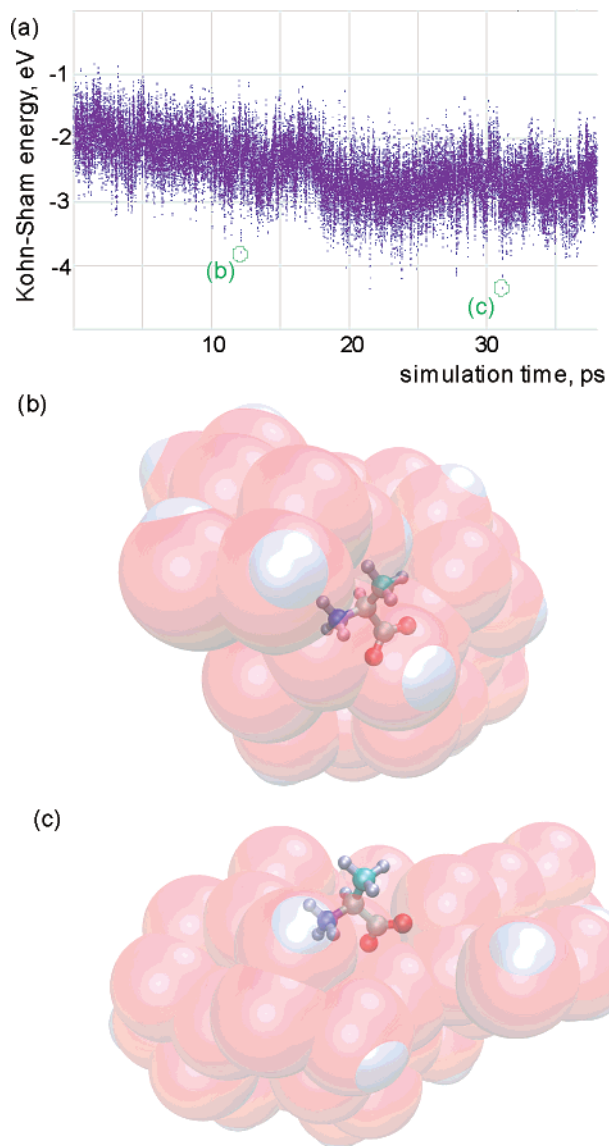
To avoid interactions between a water droplet and its nearest periodic images, we used a large unit cell of size  $19 \times 22 \times 24$  Å (the supercell approach). The density-functional molecular dynamics simulation was extended for 40 ps, using a time step of 1 fs. Only the last 38 ps were then used for analysis. The simulations were performed in the microcanonical ensemble. After initial equilibration at 300 K, the temperature was no longer controlled to avoid nonphysical effects of a thermostat on the dynamics of the simulations. During the production run, the temperature of the system was  $336.8 \pm 18.4$  K. The same

methodology has successfully been used in recent studies of problems of biological relevance, including the modeling of proteins, DNA, and liquid water.<sup>39b,42,46</sup> For visualization and trajectory analysis, we used the Visual Molecular Dynamics package (VMD).<sup>47</sup>

### 3. Results and Discussion

**Zwitterion Trajectory and Dynamics.** The total simulation time of 38 ps can be divided into two periods. The first 14 ps correspond to the situation when the alanine molecule is fully hydrated and stays inside the water droplet. In the second phase (last 22 ps), the alanine molecule is located at the droplet's surface, with its side chain exposed out of the droplet. The time interval of 2 ps in the middle can be considered as a transient state. The methyl group (the side chain of alanine) does not interact favorably with water because its free energy of solvation is positive: The entropy is large and negative because water is more ordered around the methyl group, and the enthalpy is small and also negative because of the formation of additional hydrogen bonds between water molecules surrounding the methyl group. Therefore, the system will try to minimize its free energy by pushing the hydrophobic side chain out of the water phase; this hydrophobic effect gives rise to the net force that defines the molecule's behavior inside the water droplet. The Kohn–Sham energy plot in Figure 2a justifies the existence of the above two parts of the simulation trajectory. The initially fully hydrated amino acid rotates and moves to minimize contacts of its hydrophobic methyl group and  $\alpha$ -hydrogen with water. The energy of the first part steadily decreases as the molecule seeks the most favorable energetic state. The  $\alpha$ -hydrogen is released first from the water droplet, with a corresponding energy drop at 14 ps. Then, shortly after the  $\alpha$ -hydrogen, the methyl group is fully released at 16 ps. The energy peak at 17 ps corresponds to the  $\alpha$ -hydrogen partially dipping back into the water shell, which is energetically costly. Eventually, the molecule adjusts itself in such a way that only the hydrophobic part is mainly out of the water droplet, while the charged carboxylate and ammonium groups remain fully in the water. Two minimum-energy states extracted from the simulation trajectory are shown in Figure 2b,c (their energies are shown by the letters b and c in Figure 2a). The energy of the first minimum (the alanine molecule is fully covered by water) was found to be up to  $\sim 5$  kcal/mol higher than that of the second minimum (the methyl group and  $\alpha$ -hydrogen are exposed at the droplet's surface). Thus, the system state with hydrophobic groups exposed out of the water was found to be the most energetically favorable.

In general, we observed a rapidly fluctuating molecule, constantly changing its conformation. The  $O^1C^1C_\alpha H_\alpha$ ,  $H^3NC_\alpha H_\alpha$ , and  $H^2C_\beta C_\alpha H_\alpha$  dihedral angles, plotted as a function of time in Figure 3, indicate rotational motions of all functional sites, i.e., methyl, ammonium, and carboxylate groups, of L-alanine. Jump-like rotations of  $120^\circ$  were observed at least once for  $NH_3$  and  $CH_3$  sites over 38 ps of simulation time, whereas the  $CO_2$  group was found to rotate mostly gradually, with a  $60^\circ$  jump occurring only once; the gradual rotational motions occurred in both the clockwise or counterclockwise directions. Furthermore, our simulations suggest that, at room temperature, none of the observed conformations of the carboxylate group can be considered as more preferable than the others. However, the ammonium and methyl group prefer to stay in an almost staggered conformation with respect to  $H_\alpha$  and in an unstaggered (eclipsed) conformation with respect to each other.

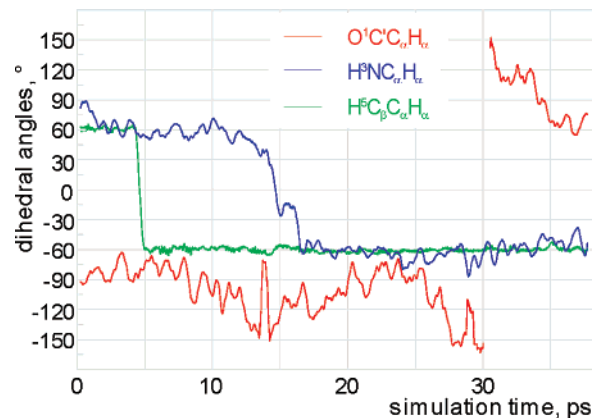


**Figure 2.** (a) Time evolution of the potential energy during the course of the simulation and (b,c) the two lowest-energy structures of L-alanine in a droplet of water: (b) fully hydrated state and (c) state with the methyl group and  $H_{\alpha}$  exposed to the air/water surface. The water shell is drawn as the van der Waals surface, with radii of 1.52 Å for O and 1.2 Å for H.

**First Hydration Shell.** We define the first hydration shell as a set of solvent molecules explicitly interacting with a solvated molecule. For an L-alanine zwitterion, the first hydration shell is defined as a network of hydrogen-bonded water molecules explicitly interacting with carboxylate, ammonium, and methyl sites. Hydration shells around the three groups constitute the first hydration shell of an alanine amino acid molecule.

**Water Structure around Charged Groups: Stable Zwitterion.** An alanine zwitterion in the gas phase is unstable. We were not able to witness a stable zwitterion form in the absence of water molecules: For all starting geometries, the L-alanine amino acid adopted conformations in which the acidic carboxylate group received a proton from the basic ammonium group and the molecule became nonionic ( $\text{NH}_3^+\cdots\text{COO}^- \rightarrow \text{NH}_2\cdots\text{COOH}$ ). Similar results were found in previous ab initio simulations<sup>10,21a</sup> and were also reported for glycine.<sup>48</sup>

In contrast, molecular transformations from the zwitterion to the neutral form were not observed throughout the entire 38 ps

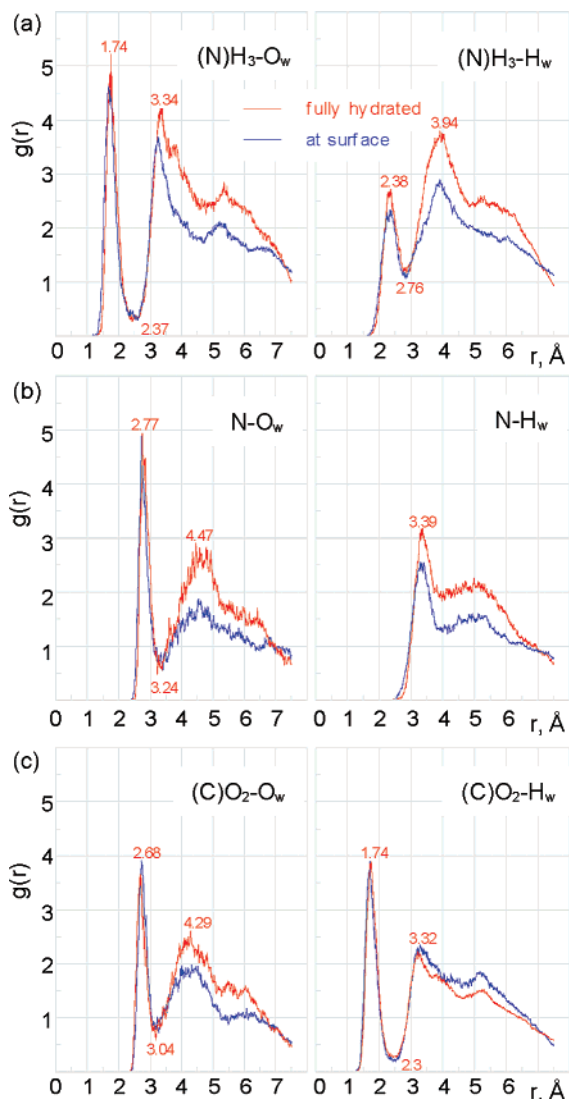


**Figure 3.** Time evolution of the dihedral angles of an alanine zwitterion over the course of the simulation. Shown are the  $\text{O}^1\text{C}_{\alpha}\text{H}_{\alpha}$  (carboxylate group, red),  $\text{H}^2\text{N C}_{\alpha}\text{H}_{\alpha}$  (ammonium group, blue), and  $\text{H}^2\text{C}_{\beta}\text{C}_{\alpha}\text{H}_{\alpha}$  (methyl group, green) dihedral angles. The average was taken every 500 fs.

of the production run of our simulations of an alanine molecule in a water droplet. Essentially, the polar water molecules around the charged groups stabilize the L-alanine zwitterion, preventing a proton of the ammonium group from transferring to the carboxylate group. The fact that an alanine zwitterion is stable in the presence of water implies that interactions with the solvent are the key determinant of its stability.

The minimum number of water molecules required to stabilize a zwitterion of an amino acid has been predicted to range from 2 to 6 for glycine<sup>49</sup> and to be 2 or 4 for alanine.<sup>10,21a</sup> According to a recent neutron diffraction study of aqueous alanine solutions, the total number of water molecules within the first hydration shell of an alanine molecule can be estimated to be roughly 8–9;<sup>18a</sup> most of the water molecules (6–7 of the 8–9) hydrate charged groups. Thus, for full hydration of the carboxylate and ammonium functional groups of an alanine amino acid molecule, one needs more water molecules than is required just to keep the zwitterion stable.

The preferred number of water molecules in the amino acid's hydration shell can be attributed to the propensity of its functional groups to form hydrogen bonds. The radial distribution functions (RDFs) calculated from the first 14 ps of the molecular dynamics trajectory (when the alanine molecule is fully hydrated) for the  $(\text{N})\text{H}_3\text{--O}_w$ ,  $(\text{N})\text{H}_3\text{--H}_w$ ,  $\text{N--O}_w$ ,  $\text{N--H}_w$ ,  $(\text{C})\text{O}_2\text{--O}_w$ , and  $(\text{C})\text{O}_2\text{--H}_w$  pairs are shown in Figure 4. The high first peaks and low first minima of the RDFs for  $(\text{N})\text{H}_3\text{--O}_w/(\text{N})\text{H}_3\text{--H}_w$  and  $(\text{C})\text{O}_2\text{--O}_w/(\text{C})\text{O}_2\text{--H}_w$  (see Figure 4a,c) suggest a highly structured first solvation shell around the charged groups. The water molecules around both charged sites take preferred orientations in which an oxygen atom of each water molecule near the ammonium group faces toward the ammonium hydrogen atom(s) and each hydrogen atom of the water molecules near the carboxylate group faces toward the carboxylate oxygen atom(s). The distances between the  $(\text{N})\text{H}_3$  and  $\text{O}_w$  atoms and between the  $(\text{C})\text{O}_2$  and  $\text{H}_w$  atoms, the values at the peaks of the  $(\text{N})\text{H}_3\text{--O}_w$  and  $(\text{C})\text{O}_2\text{--H}_w$  RDFs, were found to equal 1.74 Å. The results of our simulations are in good agreement with data obtained in several neutron diffraction experiments.<sup>18,50</sup> Corresponding values of the peaks and minima are summarized in Table 1; available experimental values are included in parentheses. A water molecule is defined to be inside the first hydration shell if the distance between its oxygen atom and the ammonium hydrogen or carboxylate oxygen atom falls within the first peaks of the corresponding  $(\text{N})\text{H}_3\text{--O}_w$  or  $(\text{C})\text{O}_2\text{--O}_w$  RDF, respectively. The first minima of these distributions are considered as hydration radii. Integration over



**Figure 4.** Radial distribution functions for distances between the atoms of the ammonium and carboxylate groups and those of the water molecules ( $O_w$ ,  $H_w$ ): (a)  $(N)H_3-O_w$  and  $(N)H_3-H_w$ , (b)  $N-O_w$  and  $N-H_w$ , and (c)  $(C)O_2-O_w$  and  $(C)O_2-H_w$ . Two simulation periods of are shown: fully hydrated L-alanine (red line) and L-alanine with its methyl group exposed at the surface (blue line).

the first peaks, up to hydration radius, yields the hydration numbers, i.e., the numbers of water molecules inside the hydration shells: 2.76 for the  $NH_3^+$  group [integrated over the  $(N)H_3-O_w$  RDF to  $r = 2.37$  Å] and 4.06 for the  $CO_2^-$  group [integrated over the  $(C)O_2-O_w$  RDF to  $r = 3.04$  Å]. The hydration number for the ammonium group is slightly larger than the value reported for alanine (2.4);<sup>18a</sup> on the other hand, this value is smaller than that reported for glycine (3.0).<sup>19</sup>

In addition to the RDFs for a fully hydrated molecule, we calculated RDFs for the second phase of the molecular dynamics simulation (16–38 ps), when the alanine side chain is exposed at the surface of a droplet, depicted with blue lines in Figure 4.

**TABLE 1: Nearest-Neighbor Distances ( $r_{nb}$ ), Donor–Acceptor Distances ( $r_{DA}$ ), Hydration Radii (Integration Limits,  $r_{hdr}$ ) and Hydration Numbers (hdn) of L-Alanine Sites<sup>a</sup>**

	$r_{nb}$ (Å)	$r_{DA}$ (Å)	$r_{hdr}$ (Å)	hdn
$NH_3^+$	$(N)H_3-O_w$ , 1.74	$N-O_w$ , 2.77 (2.88)	$(N)H_3-O_w$ , 2.37	2.76 (2.4)
$CO_2^-$	$(C)O_2-H_w$ , 1.74 (1.89)	$(C)O_2-O_w$ , 2.68 (2.78)	$(C)O_2-O_w$ , 3.04	4.06 (4.0)
$CH_3$	$(C)H_3-O_w$ , 3.18 (2.58)	$C_\beta-O_w$ , 3.82	$C_\beta-O_w$ , 4.60	12.27 (1.86)

<sup>a</sup> Available neutron diffraction measurements data are in parentheses.<sup>18</sup>

the first peaks and minima of the RDFs for the ionic ends with water practically coincide with those of a fully hydrated molecule. This leads us to the conclusion that an alanine molecule, being a zwitterion, always has a highly ordered hydration shell around its charged groups, regardless of whether its side chain is surrounded by solvent molecules. The difference between the two phases becomes visible in the second hydration shell around the charged groups, corresponding to the decrease in the number of surrounding water molecules due to the proximity of the water droplet surface.

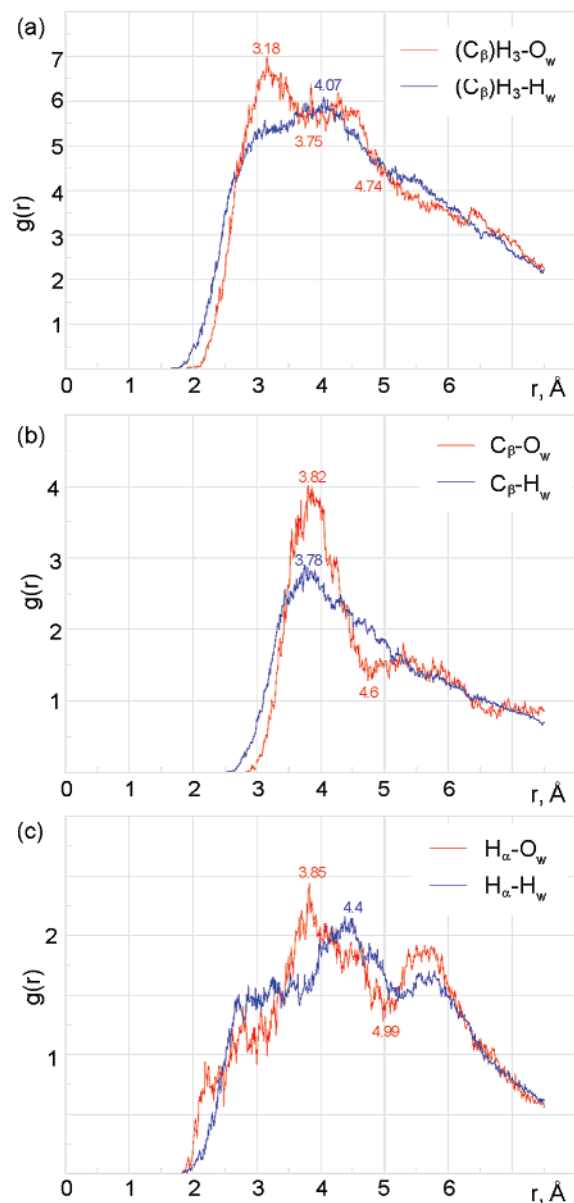
Note that, in the molecular conformation shown in Figure 1a, atoms  $H^3$  and  $O^1$  often have the possibility of sharing the same water molecule. We found that ionic groups were bridged by such “shared” water molecules for about 12% of the full simulation time. Thus, the total average number of water molecules involved in the carboxylate–ammonium hydration shell was found to be 6.71, which is smaller than the simple sum of the corresponding hydration numbers of the two charged groups (6.82) because of the existence of shared water molecules.

**Hydration Structure around the Methyl Group and  $\alpha$ -Hydrogen.** To investigate the hydration structure around the methyl side chain, we used only the first 14 ps of trajectory (i.e., when the molecule is fully in water). The corresponding RDFs for the  $C_\beta-O_w$ ,  $C_\beta-H_w$ ,  $(C)H_3-O_w$ ,  $(C)H_3-H_w$ ,  $H_\alpha-O_w$ , and  $H_\alpha-H_w$  pairs are shown in Figure 5. In contrast to the carboxylate and ammonium groups, the  $(C)H_3-O_w$  and  $(C)H_3-H_w$  RDFs demonstrate less pronounced, wider peaks and degradation of minima, which are an indication of rather disordered shells of water.

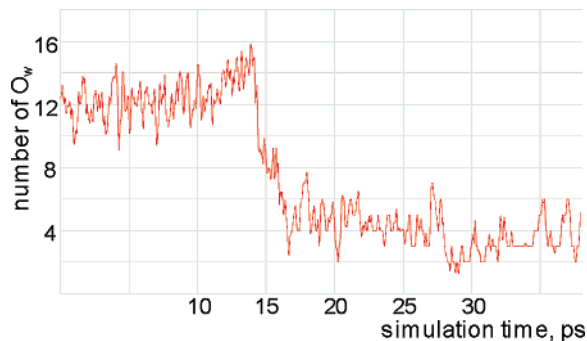
The RDFs for  $H_\alpha$  and the water atoms show practically random hydration shells (Figure 5c). However, in both cases, our results indicate a preferred orientation of the water molecules in which the oxygen atoms face toward the  $\alpha$ -hydrogen and the methyl hydrogen atoms. The number of possible water molecules inside the methyl hydration shell is 12.27 (integration of the  $C_\beta-O_w$  distance distribution to 4.6 Å), which is close to the value calculated for the methyl group of a alanine dipeptide.<sup>51</sup> In experimental studies, the number of water molecules restrained on methyl side chain was found to be about 2 (we note, however, that, in ref 20b, the methyl side chain size was claimed to be too small for the microwave dielectric method).<sup>18a,20b</sup>

Figure 6 presents the evolution of the number of water molecules within a distance of 4.6 from the carbon atom of the methyl group. The distance is considered as a hydration radius. One can see a sharp drop in the number of water molecules at about 14–16 ps, which corresponds to the time at which the methyl group and  $\alpha$ -hydrogen are exposed at the droplet’s surface.

The  $(C)H_3-O_w$  nearest-neighbor distance of 3.18 Å (see Table 1) prevents the existence of hydrogen bonding mentioned by Kameda et al.<sup>18a</sup> A corresponding distance predicted in neutron diffraction experiments<sup>18a</sup> was found to be 2.58(1) Å, i.e., close to the sum of the van der Waals radii of hydrogen and oxygen atoms (2.6 Å); this suggests a weak interaction between the methyl group and the solvent molecules. The

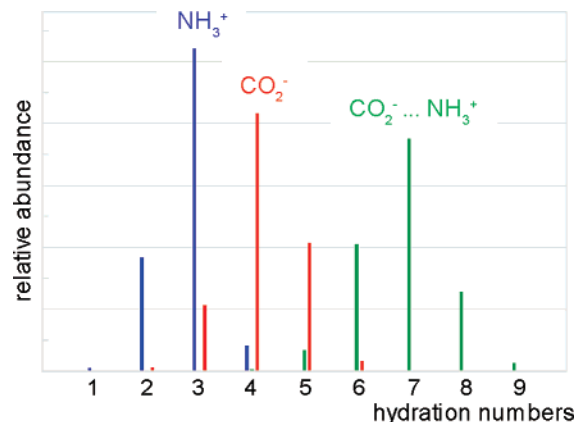


**Figure 5.** Radial distribution functions for the distances between the atoms of the methyl group ( $C_\beta$ , H) and the  $\alpha$ -hydrogen atom ( $H_\alpha$ ) and the atoms of the water molecules ( $O_w$ ,  $H_w$ ); (a)  $(C_\beta)H_3-O_w$  and  $(C_\beta)H_3-H_w$ , (b)  $C_\beta-O_w$  and  $C_\beta-H_w$ , and (c)  $H_\alpha-O_w$  and  $H_\alpha-H_w$ .



**Figure 6.** Number of water molecules around the methyl group of L-alanine over the course of the simulation. The average was taken every 100 fs, and the number of water molecules corresponds to the number of water oxygen atoms within a sphere of radius 4.6 Å around the carbon atom ( $C_\beta$ ) of the methyl group (the radius corresponds to the minimum of the  $C_\beta-O_w$  RDF).

hydrogen-bond properties of the  $C-H\cdots O$  type of interactions were established long ago, although a methyl group is classified

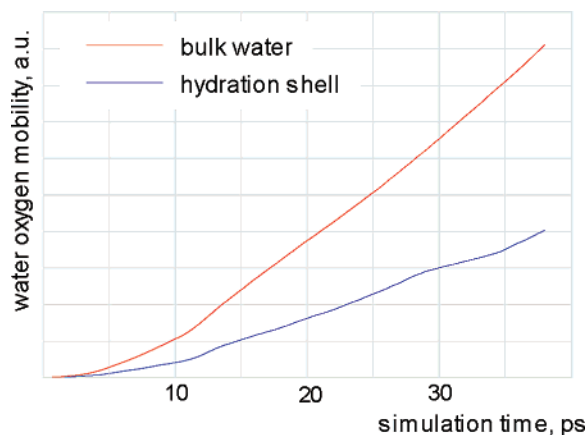


**Figure 7.** Histogram of the hydration numbers of the L-alanine carboxylate (red) and ammonium (blue) functional groups; the histogram for both groups is shown in green.

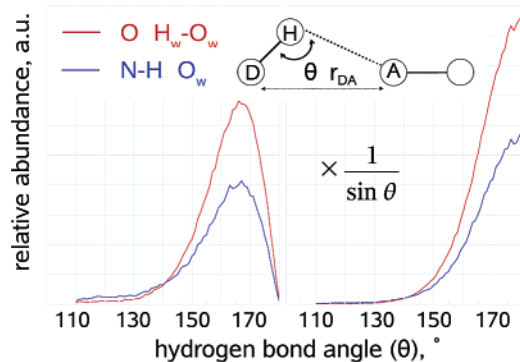
as one of the weakest  $C-H$  donors.<sup>52</sup> Accurate MP2 simulations of the methane $\cdots$ water complex give two minimum-energy structures that can be identified as the  $C(sp^3)-H\cdots O_w$  and  $O_w-H_w\cdots C(sp^3)$  forms with corresponding binding energies of  $-0.49$  and  $-0.78$  kcal/mol, respectively.<sup>53</sup> Having all this in mind, one can expect some weak ordering in the hydration shell around the methyl group within the hydrogen-bonding range (up to  $\sim 3.2$  Å).<sup>52</sup> However, with the exception of a preferred orientation (water oxygen atoms face toward methyl hydrogen atoms), we were not able to observe any highly structured hydration shells around the L-alanine methyl group. Furthermore, the  $(NH_3-O_w)$  and  $(CO_2-O_w)$  RDFs presented in Figure 3 indicate hydration structure perturbations in the second hydration shell. Our results support the Raman spectroscopy study performed by Ide et al. on aqueous solutions of various amino acids (including L-alanine),<sup>17</sup> which suggested that perturbations of the structure of water due to interactions between amino acids and water molecules were localized around the ammonium and carboxylate functional groups and that the structure of water was only slightly affected by the side chains. Also, neutron diffraction measurements on a solution of tertiary butanol in water demonstrated that there was no significant perturbation of the water structure in the first hydration shell of a nonpolar group as compared to that in liquid water.<sup>54</sup>

In general, the results presented for the methyl functional site are less definitive than those for ionic moieties. The computational technique used in this study, particularly the PBE exchange-correlation functional, is definitely adequate for describing mutual interactions of water molecules and interactions with ionic ends. However, it is known to be less accurate for the weak water–methyl interactions. For instance, for *N,N*-dimethylformamide dimers that are bonded only by  $CH\cdots O$  hydrogen bonds, the PBE functional tends to underestimate the association energy in the range of  $\sim 1.5$ – $5$  kcal/mol as compared to the results of MP2 calculations with a large basis set.<sup>41a</sup> Similarly, the energy difference between the fully hydrated state and the state with a methyl group exposed at the surface cannot be defined exactly. Nevertheless, this does not change the essence of dynamics.

**Hydration Shell Properties.** The distribution of the hydration numbers (see Figure 7) shows that the possible numbers of water molecules within the hydration shells of ammonium and carboxylate groups can vary from 1 to 4 and from 2 to 6, respectively. We also observed the ability of a water molecule from the first hydration shell to be trapped in the second hydration shell and then to be replaced by another water molecule. Frimand et al. made similar observation;<sup>12</sup> however,



**Figure 8.** Mean-square displacements of water molecules inside the first hydration shell (blue) and in the bulk (red) as a function of time.



**Figure 9.** Distribution of hydrogen-bond angles ( $\theta$ ) in the hydration shells of the ammonium and carboxylate groups. The plot on the right-hand side is after applying the cone correction [weighting by  $1/(\sin \theta)$ ].

because of the small number of water molecules in the hydration shell, those authors were not able to detect the replacement. During the course of the simulation, every *second* water molecule (26 of the total of 50) entered the first hydration shell at least once. Moreover, all water molecules that were initially in the first hydration shell moved to bulk water.

To characterize the dynamic properties of water molecules in the first hydration shell, we calculated their diffusion coefficients. In Figure 8, we present mean-square displacements calculated over the full trajectory for water molecules within the first hydration shell and in the bulk. The diffusion coefficient of the water molecules in the first hydration shell was found to be about 2.2 times smaller than that of the water molecules in the bulk.

These results strongly suggest that water inside the first hydration shell is no more than half as mobile as water in the bulk. Thus, the hydration shells around the ionic ends of an alanine molecule are relatively rigid.

**Hydrogen-Bonding Interactions.** Hydrogen-bonding interactions between an L-alanine zwitterion molecule and its hydration shell are the main factor that makes the zwitterion stable. We adopted a geometric definition of a hydrogen bond according to which a hydrogen bond between a water molecule and an amino acid site exists if the donor (D)–acceptor (A) distance is no larger than the corresponding hydration radius (the first minimum of the corresponding RDF) and if the angle D–H...A is simultaneously in the range of 120–180° (see Figure 9). The hydration radii for the ammonium and carboxylate groups are reported in Table 1. We intentionally used a wider than generally accepted angle range (150–180°) to take

into account the case when a donor is hydrogen-bonded to two acceptors (a bifurcated bond) or each acceptor is bonded to two donors. The distribution of hydrogen-bond angles shown in Figure 9 demonstrates a number of possibilities. To interpret this distribution properly, one must apply the “cone correction.”<sup>55</sup> The plot on the right-hand side of Figure 9, where the weighting function  $1/(\sin \theta)$  was employed as a correction, shows that linear contacts are preferred, as one must expect from electrostatic interactions.<sup>56</sup>

By employing the above-mentioned hydrogen-bond parameters, we found that a water molecule inside the first hydration shells of the corresponding charged groups is hydrogen-bonded on average for 96.5% and 98.2% of the full simulation time for  $\text{NH}_3^+$  and  $\text{CO}_2^-$ , respectively.

#### 4. Conclusions

In this work, we focused on one of the most intriguing examples of the dramatic effect of water on biological molecules, namely, the behavior of a stable zwitterion of the amino acid L-alanine in aqueous media. We employed ab initio molecular dynamics simulations in which a fully hydrated amino acid with a solvation shell of 50 water molecules was treated quantum mechanically. In contrast to a number of static ab initio calculations of amino acids surrounded by a limited amount of water molecules, the molecular dynamics approach allows us to avoid the problem of “intuitive” positioning of the solvent species. Our simulations showed that, inside the water droplet, the hydrophobic side chain (methyl group) and  $\alpha$ -hydrogen atom of L-alanine define the trajectory path: an initially fully solvated molecule tends to find its way to the surface of the hydration shell. Correspondingly, the structure of an alanine molecule with its methyl group and  $\text{H}_\alpha$  exposed at the droplet surface was found to be more energetically favorable than the fully hydrated state. Our simulations suggest that, at room temperature, none of the observed conformations of the carboxylate group can be considered as more preferable than the others. One of the major results of this work is that the structure of the first hydration shell of an alanine amino acid molecule is defined by the ammonium and carboxylate functional groups. The hydration structure around the methyl group has a rather limited influence on nearby water molecules. The first hydration shell was found to consist of seven water molecules (on average) explicitly bonded to the ammonium and carboxylate functional groups by hydrogen bonds. The water molecules of the first hydration shell were found to be considerably less mobile than those in the bulk because of explicit interactions with the amino acid. These interactions make the first hydration shell rather rigid and well structured. In general, the long molecular dynamics simulations have supported the known fact that interactions with water are the key determinant of the stable zwitterionic structure of amino acids in aqueous solution.

Undoubtedly, computer simulation methods currently have important predictive power and play a key role in understanding biological processes. Although many of these processes involve long time and length scales and therefore can be handled only by simulations at the empirical potential level, first-principles simulations are very much needed in this area. Our work is an example of how linear scaling quantum mechanical techniques can successfully be applied to biologically relevant problems in rather large and complex systems.

**Acknowledgment.** This work was supported by the Academy of Finland (Center of Excellence Grant 2006-2011). Computer resources were provided by the Laboratory of Physics

in Helsinki University of Technology through the M-grid project. K.J.J. thanks the Western Australian government for the Premier Fellows Program for financial support and J. D. Gale for interesting and thoughtful discussions.

## References and Notes

- (1) Voet, D.; Voet, J. G. *Biochemistry*, 3rd ed.; Wiley: New York, 2004.
- (2) Creighton, T. E. *Proteins: Structure and Molecular Properties*, 2nd ed.; W. H. Freeman: New York, 1993.
- (3) The actual pH range is 3.5–8.0. Above 3.5, the  $\alpha$ -carboxylic acid groups of 19  $\alpha$ -amino acids and 1  $\alpha$ -imino acid of proteins are almost entirely in their carboxylate forms, and below 8.0, the  $\alpha$ -amino groups are in their ammonium ion forms. See, for instance, a discussion in ref 1; most other biochemical textbooks contain similar material.
- (4) (a) Blanco, S.; Lessari, A.; Lopez, J. C.; Alonso, J. L. *J. Am. Chem. Soc.* **2004**, *126*, 11675–11683. (b) Császár, A. G. *J. Phys. Chem.* **1996**, *100*, 3541–3551 and references therein.
- (5) The former known exception of a stable arginine zwitterion in the gas phase, predicted by Price et al., has not been confirmed by other studies. See: (a) Price, W. D.; Jockush, R. A. *J. Am. Chem. Soc.* **1997**, *119*, 11988–11989. (b) Ling, S.; Yu, W.; Huang, Z.; Lin, Z.; Haranczyk, M.; Gutowski, M. *J. Phys. Chem. A* **2006**, *110*, 12282–12291.
- (6) Additionally, the zwitterionic forms of amino acids have been shown to be stabilized by the presence of a “solvating” electron. See: (a) Gutowski, M.; Skurski, P.; Simons, J. *J. Am. Chem. Soc.* **2000**, *122*, 10159–10162. (b) Skurski, P.; Rak, J.; Simons, J.; Gutowski, M. *J. Am. Chem. Soc.* **2001**, *123*, 11073–11074.
- (7) The aliphatic residues alanine, valine, leucine, and isoleucine have methylene (CH<sub>2</sub>) and methyl groups on their side chains. They interact more favorably with each other and with other nonpolar atoms than with water, which is one of the main factors in stabilizing the folded conformation of proteins. See, for instance, a discussion in ref 2.
- (8) (a) Radzicka, A.; Wolfenden, R. *Biochemistry* **1988**, *27* (5), 1664–1670. (b) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B. *Biochemistry* **1981**, *20* (4), 849–855.
- (9) Sicinska, D.; Paneth, P.; Truhlar, D. G. *J. Phys. Chem. B* **2002**, *106*, 2708–2713.
- (10) Tajkhorshid, E.; Jalkanen, K. J.; Suhai, S. *J. Phys. Chem. B* **1998**, *102*, 5899–5913.
- (11) Ellzy, M. W.; Jensen, J. O.; Hameka, H. F.; Kay, J. G. *Spectrochim. Acta A* **2003**, *59* (11), 2619–2633.
- (12) Frimand, K.; Bohr, H.; Jalkanen, K. J.; Suhai, S. *Chem. Phys.* **2000**, *255*, 165–194.
- (13) It has been shown that the presence of explicit water molecules is important in stabilizing the dominant species of *N*-acetyl *L*-alanine *N*-methylamide in aqueous solution. The fact that surrounding water molecules dictate the structure of a model dipeptide was later confirmed in nuclear magnetic resonance experiments. See: (a) Han, W.; Jalkanen, K. J.; Elstner, M.; Suhai, S. *J. Phys. Chem. B* **1998**, *102*, 2587–2602. (b) Poon, C.-D.; Samulski, E. T.; Weise, C. F.; Weisshaar, J. C. *J. Am. Chem. Soc.* **2000**, *122*, 5642–5643. (c) Weise, C. F.; Weisshaar, J. C. *J. Phys. Chem. B* **2003**, *107*, 3265–3277.
- (14) Kikuchi, O.; Watanabe, T.; Ogawa, Y.; Takase, H.; Takahashi, O. *J. Phys. Org. Chem.* **1997**, *10*, 145–151.
- (15) (a) Destro, R.; Marsh, R. E.; Bianchi, R. *J. Phys. Chem.* **1988**, *92*, 966–973. (b) Lehmann, M. S.; Koetzle, T. F.; Hamilton, W. C. *J. Am. Chem. Soc.* **1972**, *94*, 2657–2660. (c) Simpson, H. J.; Marsh, R. E. *Acta Crystallogr.* **1966**, *20*, 550–555.
- (16) Hetch, D.; Tadesse, L.; Walters, L. *J. Am. Chem. Soc.* **1993**, *115*, 3336–3337.
- (17) Ide, M.; Maeda, Y.; Kitano, H. *J. Phys. Chem. B* **1997**, *101*, 7022–7026.
- (18) (a) Kameda, Y.; Sugawara, K.; Usuki, T.; Uemura, O. *Bull. Chem. Soc. Jpn.* **2003**, *76*, 935–943. (b) Kameda, Y.; Sasaki, M.; Yaegashi, M.; Tsuji, K.; Oomori, S.; Hino, S.; Usuki, T. *J. Solution Chem.* **2004**, *33* (6–7), 733–745.
- (19) Kameda, Y.; Ebata, H.; Usuki, T.; Uemura, O.; Misawa, M. *Bull. Chem. Soc. Jpn.* **1994**, *67*, 3159–3164.
- (20) (a) Fischer, W. B.; Eysel, H.-H. *J. Mol. Struct. (THEOCHEM)* **1997**, *415*, 249–257. (b) Suzuki, M.; Shigematsu, J.; Fukunishi, Y.; Kodama, T. *J. Phys. Chem. B* **1997**, *101*, 3839–3845.
- (21) (a) Park, S.-W.; Ahn, D.-S.; Lee, S. *Chem. Phys. Lett.* **2003**, *371*, 74–79. (b) Selvarengan, P.; Kolandaivel, P. *J. Mol. Struct. (THEOCHEM)* **2004**, *671*, 77–86.
- (22) Ahn, D.-S.; Park, S.-W.; Jeon, I.-S.; Lee, M.-K.; Kim, N.-H.; Han, Y.-H.; Lee, S. *J. Phys. Chem. B* **2003**, *107* (50), 14109–14118.
- (23) Jalkanen, K. J.; Nieminen, R. M.; Frimand, K.; Bohr, J.; Bohr, H.; Wade, R. C.; Tajkhorshid, E.; Suhai, S. *Chem. Phys.* **2001**, *265*, 125–151.
- (24) Boresch, S.; Willensdorfer, M.; Steinhauser, O. *J. Chem. Phys.* **2004**, *120*, 3333–3347.
- (25) Hess, B.; van der Vegt, N. F. A. *J. Phys. Chem. B* **2006**, *110*, 17616–17626.
- (26) Sagarik, K.; Dokmaisrijan, S. *J. Mol. Struct. (THEOCHEM)* **2005**, *718*, 31–47.
- (27) Wei, D.; Guo, H.; Salahub, D. R. *Phys. Rev. E* **2001**, *64*, 11907–11911.
- (28) Leung, K.; Rempe, S. B. *J. Chem. Phys.* **2005**, *122*, 184506–184518.
- (29) In addition to including solvent molecules explicitly, one can also employ the reaction field continuum model or use both approaches and model the solute molecule in its first hydration shell surrounding this “supermole” with continuum solvent. See: Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999–3093 and references therein.
- (30) Leung, K.; Rempe, S. *J. Am. Chem. Soc.* **2004**, *126*, 344–351.
- (31) As an initial conformation, we used the F' structure from ref 10, which is an optimized structure of an *L*-alanine zwitterion with four water molecules within the Onsager model. See the original publication for references and structure details.
- (32) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306–317.
- (33) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. Interaction models for water in relation to protein hydration. In *Intermolecular Forces*; Pullman, B., Ed.; Reidel: Dordrecht, The Netherlands, 1981.
- (34) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- (35) (a) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089–10092. (b) Essman, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8592.
- (36) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *J. Comput. Chem.* **1997**, *18*, 1463–1472.
- (37) Miyamoto, S.; Kollman, P. A. *J. Comput. Chem.* **1992**, *13*, 952–962.
- (38) (a) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, 864–871. (b) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, 1133–1138.
- (39) (a) Soler, J. M.; Artacho, E.; Gale, J. D.; García, A.; Junquera, J.; Ordejón, P.; Sánchez-Portal, D. *J. Phys.: Condens. Matter* **2002**, *14*, 2745–2779. (b) Ordejón, P. *Phys. Status Solidi B* **2000**, *217*, 335–356. (c) Artacho, E.; Sánchez-Portal, D.; Ordejón, P.; García, A.; Soler, J. M. *Phys. Status Solidi B* **1999**, *215*, 809–817. (d) Sánchez-Portal, D.; Ordejón, P.; Artacho, E.; Soler, J. M. *Int. J. Quantum Chem.* **1997**, *65*, 453–461. (e) Ordejón, P.; Artacho, E.; Soler, J. M. *Phys. Rev. B* **1996**, *53*, 10441–10444.
- (40) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- (41) (a) Ireta, J.; Neugebauer, J.; Scheffler, M. *J. Phys. Chem. A* **2004**, *108*, 5692–5698. (b) Xu, X.; Goddard, W. J. *Phys. Chem. A* **2004**, *108*, 2305–2313.
- (42) Fernandez-Serra, M. V.; Artacho, E. *J. Chem. Phys.* **2004**, *121*, 11136–11144.
- (43) (a) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993–2006. (b) Kleinman, L.; Bylander, D. M. *Phys. Rev. Lett.* **1982**, *48*, 1425–1428.
- (44) The pseudopotentials were generated using the following reference configurations: 2s<sup>2</sup>2p<sup>2</sup> for C, 3s<sup>2</sup>3p<sup>3</sup> for N, 3s<sup>2</sup>3p<sup>4</sup> for O, and 1s<sup>2</sup> for H. The cutoff radii were 1.25 for both the s and p components in C, 1.24 for the s and p components in N, 1.14 for the s and p components in O, and 1.25 for the s component in H.
- (45) Junquera, J.; Paz, O.; Sánchez-Portal, D.; Artacho, E. *Phys. Rev. B* **2001**, *64*, 235111–235120.
- (46) (a) Sánchez-Portal, D.; Ordejón, P.; Canadell, E. *Struct. Bonding* **2004**, *113*, 103 and references therein. (b) Degtyarenko, I.; Nieminen, R. M.; Rovira, C. *Biophys. J.* **2006**, *91* (6), 2024–2034.
- (47) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
- (48) Ding, Y.; Krogh-Jespersen, K. *Chem. Phys. Lett.* **1992**, *199* (3,4), 261–266.
- (49) Xu, S.; Nilles, M.; Bowen, K. H. *J. Chem. Phys.* **2003**, *119* (20), 10696–10701 and references therein.
- (50) Naganuma, H.; Kameda, Y.; Usuki, T.; Uemura, O. *J. Phys. Soc. Jpn., Suppl. A* **2001**, *70*, 356–358.
- (51) Kalko, S. G.; Guardia, E.; Padro, J. A. *J. Phys. Chem. B* **1999**, *103* (19), 3935–3941.
- (52) Steiner, T. *Chem. Commun.* **1997**, *8*, 727–734.
- (53) Rovira, C.; Novoa, J. J.; Whangbo, M.-H.; Williams, J. M. *Chem. Phys.* **1995**, *200*, 319–335.
- (54) Bowron, D. T.; Soper, A. K.; Finney, J. L. *J. Chem. Phys.* **2001**, *114* (14), 6203–6219.
- (55) Kroon, J.; Kanters, J. A. *Nature* **1974**, *248*, 667–669.
- (56) The linear bonds are structurally significant because of the dipole–monopole and dipole–dipole contributions to the electrostatic energy, which has a maximum at  $\theta = 180^\circ$  and vanishes at  $\theta = 90^\circ$ .