ELSEVIER

Journal of Molecular Structure: THEOCHEM 718 (2005) 31–47

THEO CHEM

www.elsevier.com/locate/theochem

A theoretical study on hydration of alanine zwitterions

Kritsana Sagarik*, Supaporn Dokmaisrijan

School of Chemistry, Institute of Science, Suranaree University of Technology, 111 University Avenue, Nakhon Ratchasima 30000, Thailand

Received 6 September 2004; accepted 25 October 2004

Abstract

Structures and energetic of hydrogen bond (H-bond) networks of water at the charged functional groups of two forms of alanine zwitterions were studied using various theoretical methods. The present investigation started with a construction of intermolecular potentials between the zwitterions and water followed by Molecular Dynamics (MD) simulations of aqueous solutions. The three-dimensional structures and the average potential energy landscapes of the H-bond networks of water were analyzed and visualized using various probability distribution (PD) maps. Although the conformation with planar skeleton possesses larger overall stabilization by hydration, the conformation with the COO^- plane being 90° with respect to the $NC^{\alpha}C$ backbone seems to be more accessible by water. The MD analyses revealed that, although the shapes of the average potential energy landscapes at the H-bond networks are highly irregular, they can help characterize the dynamic behavior of water molecules especially at the functional groups the solutes.

Keywords: Alanine zwitterion; Hydrogen bond; Aqueous solution; Molecular dynamics simulations

1. Introduction

It is well known from X-ray diffraction experiments that structures of proteins in crystals depend largely on their conformations in aqueous solution from which they are grown [1]. Several X-ray structural analyses also revealed that large solvent regions separate individual protein in crystal [2–4]. Since various experimental evidence has shown that solution properties can be consistently explained by crystal structures [5–8], it has been proposed that protein structures in crystal are essentially the same as in solution [9,10] and, in most cases, general phenomena of hydration as well as local hydration patterns could be discussed more accurately in the context of water distributions rather than individual water molecules [5].

Attempt has been made to investigate hydration structures and energetic of amino acids in aqueous solution using various theoretical and experimental techniques [11–26]. One of the most pioneering computer simulation on amino acids in aqueous solution was put forward by Clementi et al. [16]. In Ref. [16], intermolecular potentials for twenty-one amino acids were derived from SCF-MO calculations and some of them were applied in Monte Carlo (MC) simulations of aqueous solution [27]. Clementi [27] suggested various possibilities to visualize and analyze hydration structures at functional groups of amino acids from SCF-MO calculations and MC simulations. It was also proposed for the first time that the detailed solvation structure is governed by a subtle balance between solute–solvent and solvent–solvent interactions and the hydrogen bond (H-bond) filaments or H-bond networks of water are most likely the key aspect for fast and long distance deprotonation at one site and protonation at another site of biological molecules [27]. These imply the necessity to include explicitly water molecules in theoretical models and partly form the basis for the present investigation.

Glycine (Gly) and alanine (Ala) have been frequently chosen as model molecules in the study of amino acid in aqueous solution [11–22]. It is well known that Gly and Ala exists in zwitterionic forms (Glyz and Alaz) in polar solvents and in the crystalline state [28–30]. At the earliest stage of theoretical studies, ab initio calculations with restricted basis sets were applied on the Glyz–H_2O 1: n complexes [11–13,27], from which structural properties in aqueous solution were anticipated. Based on

* Corresponding author. Tel./fax: +66 44 224 635.
E-mail address: kritsana@ccs.sut.ac.th (K. Sagarik).

0166-1280/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.theochem.2004.10.091

the semiempirical AM1 and PM3 as well as SCF-MO calculations, it was suggested that the microsolvated species of amino acids with fifteen-water supermolecules can yield solvation energies close to the bulk solvation limit [13]. The most probable microstructures of Glyz–H_2O 1: 7 and 1: 15 complexes were found in Ref. [13] to be represented by H-bond network of up to two water molecules linking together the NH_3^+ and COO^- groups. However, numerous theoretical and experimental investigations have shown that the structures of such microsolvated species in the gas phase and in aqueous solution can be completely different [31]. The complicated H-bond networks in aqueous solution were pointed out to be responsible for the discrepancy [31].

Alagona et al. [17] conducted a systematic analysis of water structures in the vicinities of the functional groups of Glyz in aqueous solution ($[Glyz]_{aq}$). They suggested four types of water molecules in $[Glyz]_{aq}$ namely, those tightly bound to the oxygen atoms of the carboxyl group, those tightly bound to the ammonium group, those hydrophobically localized at the methylene group and those in the bulk. It was also concluded in Ref. [17] that the strong intramolecular H-bond between one of the oxygen atom in the COO^- group and the internal hydrogen atom in the NH_3^+ group prevents the formation of an intermolecular H-bond with water for both atoms.

Limited theoretical and experimental information is available for Ala compared to Gly. A detail analysis of the conformations of Glyz and Alaz in aqueous solution was reported by Kikuchi et al. [19–21], based on the results of ab initio MO calculations with a continuum model and MC simulations. The authors concluded that the nearly planar skeleton is to be the most stable structure in aqueous solution and the stabilization by the aqueous solvent is larger in the conformation with the COO^- plane being 90° with respect to the $NC^{\alpha}C$ backbone plane. The latter conformation is regarded as Alaz-R in the present study. In contrast to the conclusion made in Ref. [17], Kikuchi et al. pointed out the possibility for water molecule to H-bond at the internal hydrogen atom of the NH_3^+ group. Additionally, they suggested that the hydrations at the COO^- and NH_3^+ groups are quite independent for Glyz.

The applicability of ab initio calculations with continuum models, such as the self-consistent reaction field (SCRF) method, has been frequently mentioned [32]. This is due mainly to the fact that ab initio calculations with continuum models neglect specific short-range solute–solvent interactions as well as temperature effects. These make them applicable only for the system, in which solvents act only as perturbation on the gas-phase property of the system [32]. The effects of the presence of water molecules on the structure of Alaz were studied using ab initio calculations at B3LYP/6-31G* level of theory, compared to those within Onsager continuum model [15]. It was reported that the conformations of Alaz are strongly influenced by water molecules, mainly through the electrostatic, polarization and H-bond interactions. It was also shown that, in order to hydrate both NH_3^+ and COO^- groups, at least four water molecules have to be included in the model calculations [15]. The three-dimensional structures of the H-bond networks of water in the vicinities of both NH_3^+ and COO^- groups of Alaz could not be presented due to restricted number of water molecules considered in the theoretical investigation [15].

In the present work, structures and energetic of the H-bond networks in aqueous solution of Alaz ($[Alaz]_{aq}$) and Alaz-R ($[Alaz\text{-}R]_{aq}$) were studied using various computational methods. The two forms of alanine zwitterion, Alaz and Alaz-R, were considered to investigate the effects of conformation change on the H-bond networks at the charged functional groups. The investigation started with construction of intermolecular potentials to describe the interaction between Alaz and water, as well as Alaz-R and water, using the Test-particle model (T-model). The T-model potentials were tested in the calculations of the equilibrium structures and interaction energies of Alaz–H_2O and Alaz–R–H_2O 1: n complexes, with $n = 1$–2. The computed T-model potentials were then applied in Molecular Dynamics (MD) simulations of $[Alaz]_{aq}$ and $[Alaz\text{-}R]_{aq}$. In order to obtain information on the three-dimensional structures of the H-bond networks of water in the vicinities of the NH_3^+ and COO^- groups, the MD results were analyzed and visualized based on the oxygen (PDO) and hydrogen probability distribution (PDH) maps [27]. The interaction energy distributions in $[Alaz]_{aq}$ and $[Alaz\text{–}R]_{aq}$ were computed and displayed using the average solute–solvent (AWPD) and average solvent–solvent interaction energy probability distribution (WWPD) maps [27]. In order to provide insight into the stability and hydration dynamics of water molecules in the H-bond networks, the so-called total–average interaction energy probability distribution maps (AW-WWPD) were computed from the AWPD and WWPD maps. The results were discussed in comparison with available theoretical and experimental data of the same as well as similar systems.

2. Methods

In our previous work [33–39], intermolecular potentials to describe interactions between functional groups of biologically active molecules were constructed using the T-model. Based on the T-model potentials, various types of intermolecular interactions were investigated successfully, ranging from H-bonds between small molecules [33–35] to π–π interactions in benzene [36], benzoic acid [37] and phenol [38]. It has been shown that the T-model potentials are suitable for the investigations of structures and energetics of both aqueous and nonaqueous solutions [36,37], by means of statistical mechanical simulations. In the following subsections, some important aspects of the T-model and MD simulations will be summarized.

2.1. T-model potentials

As mentioned earlier that conformations of Alaz are strongly influenced by the presence of water molecules [15] and the results of geometry optimizations could be totally different after the inclusion of explicit water molecules in the calculation [15]. The Alaz-R structure shown in Fig. 1 was found in Ref. [15] to represent the lowest minimum energy geometry when four water molecules were included in ab initio geometry optimizations at the B3LYP/6-31G* level of theory. However, the Alaz structure became more stable when ab initio geometry optimizations in combination with the Onsager continuum model were conducted on Alaz with four water molecules [15]. The structure of Alaz in Fig. 1 is in accordance with the experimental findings that, in the crystalline state and aqueous solution, the most probable structures of Glyz and Alaz consist of bifurcated-intramolecular H-bonds between the N–H groups and the oxygen atom [30,40,41]. We, therefore, adopted both Alaz and Alaz-R in the investigations of the effects of conformation change on the structures and energetic of the H-bond networks of water in aqueous solutions.

In the present investigation, the T-model was applied in the calculations of intermolecular potentials between Alaz and water, as well as between Alaz-R and water. Within the framework of the T-model, the interaction energy ($\Delta E_{\mathrm{T-model}}$) between molecules A and B is written as a sum of the first-order interaction energy ($\Delta E^{\mathrm{I}}_{\mathrm{SCF}}$) and a higher-order energy term (ΔE^{r}).

$$\Delta E_{\mathrm{T-model}} = \Delta E^{\mathrm{I}}_{\mathrm{SCF}} + \Delta E^{r} \tag{1}$$

$\Delta E^{\mathrm{I}}_{\mathrm{SCF}}$ accounts for the exchange repulsion and electrostatic energies. It is computed from ab initio SCF calculations [42] and takes the following analytical form:

$$\Delta E^{\mathrm{I}}_{\mathrm{SCF}} = \sum_{i\in A}\sum_{j\in B}\left[\exp\left[\frac{-R_{ij}+\sigma_i+\sigma_j}{\rho_i+\rho_j}\right] + \frac{q_i q_j}{R_{ij}}\right]. \tag{2}$$

i and j in Eq. (2) label the sites of molecules A and B. σ_i, ρ_i and q_i are the site parameters. R_{ij} is the site–site distance. The exponential term in Eq. (2) represents the size and shape of the interacting molecules A and B. The point charges q_i and q_j are computed from the requirement that a point-charge model reproduces the electrostatic potentials of molecules of interest. In our previous studies [35–37], we showed that the CHelpG charges [43] are also applicable. In the present study, q_i and q_j for Alaz were, therefore, determined by a fit of the electrostatic potentials at points selected according to the CHelpG scheme which has been embedded in the GAUSSIAN 98 package [44]. The electrostatic potentials employed in the fit were derived from the density matrices computed from ab initio calculations at the MP2/6-311++G(2d,2p) level of theory. About nine thousand electrostatic energies were used in the fit of the atomic charges. The dipole moment of Alaz computed from the CHelpG charges is 11.50 D, whereas those obtained from

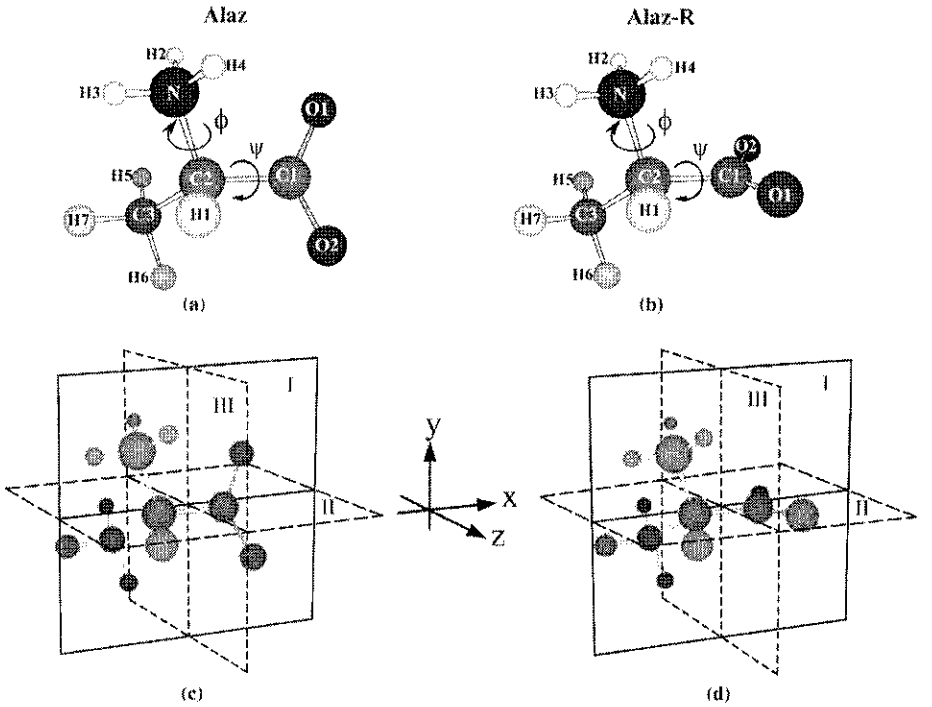

Fig. 1. Two conformations of alanine zwitterion and the reference planes used in MD analyses. (a) Alaz geometry with the nearly planar skeleton. (b) Alaz-R geometry with the COO^- plane being 90° with respect to the NC$^{\alpha}$C backbone plane. (c) Reference planes for $[\mathrm{Alaz}]_{\mathrm{aq}}$. (d) Reference planes for $[\mathrm{Alaz\text{-}R}]_{\mathrm{aq}}$.

ab initio calculations and experiments in aqueous solutions are in the range of 10.8 and 15.7 D [45,46].

The higher-order energy contribution, ΔE^r in Eq. (1), represents the dispersion and polarization contributions of the T-model potential. ΔE^r could be determined from both theoretical and experimental data. Our previous experience has shown that a calibration of the incomplete potential to the properties related to intermolecular interaction energies, such as the second virial coefficients (B(T)), dimerization energies or potential energy of liquid etc. is the most appropriate choice. ΔE^r takes the following form:

$$\Delta E^r = -\sum_{i\in A}\sum_{j\in B} C_{ij}^6 F_{ij}(R_{ij}) R_{ij}^{-6} \tag{3}$$

where

$$F_{ij}(R_{ij}) = \exp\left[-(1.28\, R_{ij}^0/R_{ij} - 1)^2\right], \quad R_{ij} < 1.28\, R_{ij}^0$$
$$= 1, \quad \text{elsewhere} \tag{4}$$

and

$$C_{ij}^6 = C_6 \frac{3}{2} \frac{\alpha_i \alpha_j}{(\alpha_i/N_i)^{1/2} + (\alpha_j/N_j)^{1/2}} \tag{5}$$

R_{ij}^0 in Eq. (4) is the sum of the van der Waals radii of the interacting atoms. Eq. (5) is the Slater-Kirkwood relation. α_i and N_i in Eq. (5) denote the atomic polarizability and the number of valence electrons of the corresponding atom, respectively. $F_{ij}(R_{ij})$ in Eq. (4) is a damping function, introduced to correct the behavior of R_{ij}^{-6} at short R_{ij} distance. Only C_6 in Eq. (5) is unknown.

In our experience, the values of C_6 in Eq. (5) are close to 1 in many cases. The variation of C_6 within the range of 0.8 and 1.5 seems not lead to significant change in the Potential Energy Surface (PES). For most of the microsolvated systems considered, the values of C_6 were determined to be 1.43 [39]. We, therefore, adopted the same value of C_6 for the Alaz–H_2O complexes. Our experience has also demonstrated that the repulsion parameters and C_6 are not very sensitive to a slight conformational change, compared to the point charges. Therefore, only the point charges of Alaz-R were recomputed using ab initio calculations at MP2/6-311++G(2d,2p) level of theory. It turned out that the point charges of Alaz-R from the CHelpG scheme were not substantially different from those of Alaz, with the dipole moment of 11.52 D. Thus, in order to keep our T-model potential simple for further applications, we adopted the same point charges for both Alaz and Alaz-R. The T-model parameters for Alaz, Alaz-R and water can be sent on request.

2.2. Equilibrium structures of Alaz–H_2O and Alaz–R–H_2O 1: n complexes

The T-model potential constructed in the previous section was applied in the calculations of the equilibrium structures and interaction energies of the Alaz–H_2O and Alaz–R–H_2O 1: n complexes, with n = 1–2. Rigid Alaz was placed at the origin of the Cartesian coordinate system and the coordinates of water were randomly generated in the vicinities of Alaz. Based on the T-model potential, the equilibrium structures of the Alaz–H_2O complexes were searched using a minimization technique [47]. A hundred starting configurations were generated for each intermolecular geometry optimization. Only some lowest-lying minimum energy geometries were discussed in details. The same procedure was applied to determine the equilibrium structure of the Alaz–R–H_2O complexes.

2.3. Molecular dynamics simulations

Theoretical methods applied in the study of solvent effects on the static and dynamic properties of amino acids fall into two categories, depending on the treatment of the solvent molecules [48–50]. Microscopic methods treat solvent molecules and their interactions with the solute explicitly, whereas macroscopic methods consider solvent as a continuous medium characterized by a dielectric constant [49,50]. Both approaches seem to have advantages and disadvantages [51]. In aqueous solution, for example, the former can yield a deep insight into microscopic solvation structures, such as the three-dimensional structures of the H-bond networks of water in the vicinities of solute molecules. The latter on the other hand yields the solvation free energy, which is estimated in relation to the dielectric constants. Since the three-dimensional structures of the H-bond networks of water in the first hydration sphere of the solutes were one of our prime interests, the former approach was adopted in the present work.

Based on the T-model potentials, NVE-MD simulations were performed on $[Alaz]_{aq}$ and $[Alaz\text{-}R]_{aq}$ at 298 K. In MD simulations, a rigid solute and 300 rigid water molecules were put in a cubic box subject to periodic boundary conditions. The center of mass of solute was placed at the center of the simulation box. In order to simplify the analysis of the hydration structures, the C1–C2–N backbone of Alaz was assumed to coincide with the XY plane of the box, with Z=0.20 Å. The densities of $[Alaz]_{aq}$ and $[Alaz\text{-}R]_{aq}$ were maintained at 1.0 g cm^{-3}. The cut-off radius was half of the box length. The long-range Coulomb interaction was taken into account by means of the Ewald summations. The time step used in solving the equations of motions was 0.5 fs. In MD simulations, 100,000 timesteps were devoted to the equilibration and additional 200,000 timesteps to property calculations. The latter corresponds to the simulation time of 100 ps.

General energetic results were computed from MD simulations namely, the average potential energy of aqueous solution ($\langle E_{aq}^{pot}\rangle$) and the average solute–solvent interaction energies ($\langle E_{aq}^{solu-solv}\rangle$), as well as the average solvent–solvent interaction energies ($\langle E_{aq}^{solv-solv}\rangle$). These energy values were the results of the average over the timesteps and

the number of water molecules. The structures of water molecules in the H-bond networks were initially characterized based on the average H-bond distances ($\langle R_{A-H\cdots B}\rangle$) and angles ($\langle\theta_{A-H\cdots B}\rangle$). $\langle\theta_{A-H\cdots B}\rangle$ represent the angle between the A–H bond and the line connecting atoms A and B. Since NMR experiments [52,53] suggested that hydration water can mediate protein–DNA recognition through specific H-bond formations which depend on the hydration dynamics of water molecules at particular hydration sites, it is interesting to estimate the duration of the H-bonding between individual water molecules and the NH_3^+ and COO^- groups from MD simulations. Due to the fact that the H-bond formations and disruptions take place quite often and very rapidly at the first hydration shell of proteins [54], the residence times derived from MD simulations could vary in a wide range. Since NMR experiments can effectively detect the long-lived hydration water, it is reasonable to compare the longest H-bond lifetimes ($\tau_{A-H\cdots B,\,max}$) obtained from MD simulations with the NMR average residence times [55]. $\tau_{A-H\cdots B,\,max}$ were approximated from the percentage of simulation steps, during which a specific pair of H-bond donor and acceptor were coming close enough to continuously engage in H-bonds. H-bond donor and acceptor were considered to engage in H-bond formation when the donor–acceptor distance was shorter than 4 Å [56].

The hydration structures and the H-bond networks of water around Alaz and Alaz-R were further analyzed in details using the atom–atom pair correlation functions ($g(R)$) and the average running coordination number ($n(R)$), as well as the PDO and PDH maps [27]. The PDO and PDH maps show the average three-dimensional structures of the H-bond networks at the functional groups of Alaz and Alaz-R. In the present work, three sets of the PDO and PDH maps were constructed, using the predefined reference planes I, II and III in Fig. 1. In order to view the overall picture of hydration structures of Alaz and Alaz-R, the C1–C2–N backbone was chosen to form reference plane I. Since the second set of the PDO and PDH maps was aiming at the hydration structures at the NH_3^+ group, the XZ plane with $Y=0.0$ Å was defined as reference plane II. The YZ plane with $X=0.0$ Å was chosen as reference plane III to view of the hydration structures at the COO^- group.

In the calculations of the PDO and PDH maps, the volumes above and below the reference planes were divided into layers with the thickness of 1.0 Å. In each layer, the PDO and PDH maps were computed at the 61×61 grid intersections, by following the trajectories of the oxygen and hydrogen atoms of water in the course of MD simulations. The PDO and PDH maps were represented by contour lines constructed using the SURFER program [57]. For simplicity, the maximum and minimum of the contour lines, as well as the contour interval, were the same for all the PDO and PDH maps.

In order to obtain insight into the interaction energy distributions in aqueous solutions, a similar approach was adopted in the analysis of the average solute–solvent and average solvent–solvent interaction energies. The AWPD and WWPD maps [27] for $[Alaz]_{aq}$ and $[Alaz\text{-}R]_{aq}$ were constructed with respect to reference planes I, II and III. The AWPD maps account for the average interaction energy between water molecule at the grid intersection and the solute molecule, whereas the WWPD maps reveal the average interaction energy between water molecule at the grid intersection and all other water molecules in aqueous solution. Only negative interaction energies were employed in the calculations of the AWPD and WWPD maps. The contour intervals were the same for all the AWPD and WWPD maps. In order to obtain information on the hydration dynamics of water molecules and to view the average potential energy landscapes at the H-bond networks, the AW-WWPD maps were computed by combination of the AWPD and WWPD maps. Since in general the rate of water exchange and the mobility of water molecules depends on the transition energy barriers, the shapes of the average potential energy landscapes at the H-bond networks were analyzed in details using cross section plots. Various cross section plots were generated by taking vertical slices along predefined profile lines through the surfaces of the AW-WWPD, AWPD and WWPD maps. In the present work, the cross sections derived from the longitudinal profile lines on the AW-WWPD maps could be associated with the transition energy barriers to water exchange within the H-bond network ($\langle\Delta E_{aq}^{L}\rangle$). Whereas those computed from the transverse profile lines are attributed to the transition energy barriers to water exchange between the H-bond network and the outside ($\langle\Delta E_{aq}^{T}\rangle$). It should be noted that, when a particular water molecule leaves a hydration site, its place will be occupied nearly simultaneously by another water molecule. And since the rate of water exchange depends on the transition energy barriers, which is inversely proportional to $\tau_{A-H\cdots B,\,max}$, the hydration dynamics of water molecules at the hydration sites are discussed based on $\tau_{A-H\cdots B,\,max}$, $\langle\Delta E_{aq}^{L}\rangle$ and $\langle\Delta E_{aq}^{T}\rangle$.

3. Results and discussion

3.1. The Alaz–H_2O and Alaz–R–H_2O 1: n complexes

The absolute and some lowest-lying minimum energy geometries of the Alaz–H_2O and Alaz–R–H_2O 1: n complexes, with $n=1$–2 are illustrated in Fig. 2, together with $\Delta E_{T-model}$.

The absolute minimum energy geometry of the Alaz–H_2O 1: 1 complex from the T-model potential is structure **a** in Fig. 2. Structure **a** consists of a cyclic H-bond in which water acts simultaneously as proton acceptor and donor towards the NH_3^+ and COO^- groups of Alaz, respectively. $\Delta E_{T-model}$ of structure **a** is -82.65 kJ mol^{-1}, with the N–H4$\cdots$Ow and Ow–Hw$\cdots$O1 H-bond distances of 2.65 and 2.71 Å, respectively. The cyclic H-bond in structure **b** is similar to

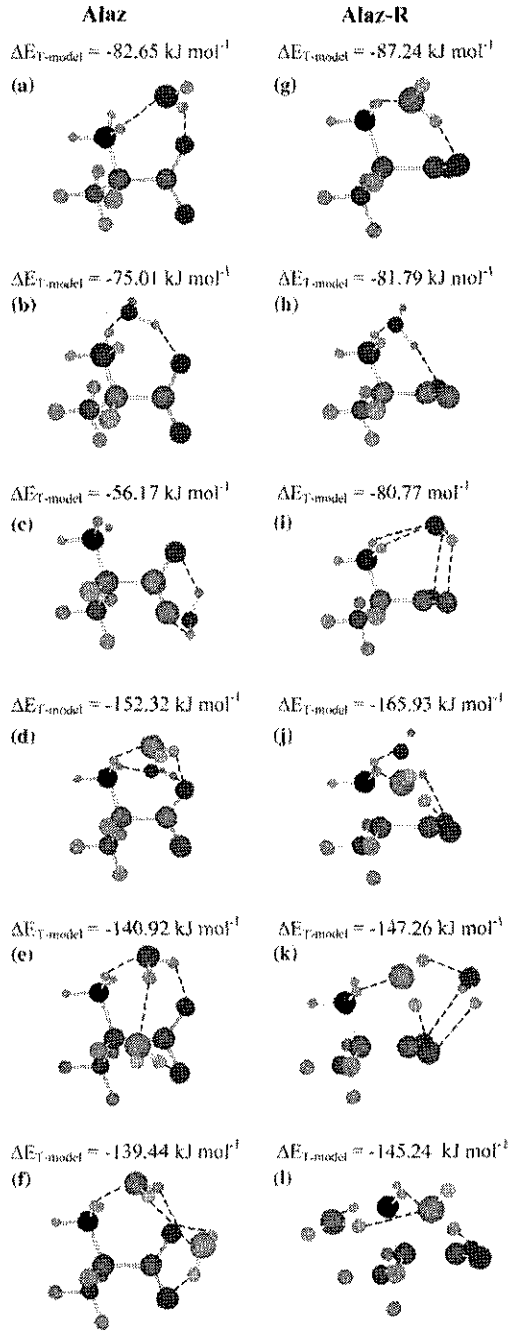

Fig. 2. The absolute and some local minimum energy geometries of Alaz-H_2O and Alaz–R–H_2O 1: *n* complexes computed from the T-model potentials. (a)-(c) Alaz-H_2O 1: 1 complexes. (d)–(f) Alaz–H_2O 1: 2 complexes. (g)–(i) Alaz-R–H_2O 1: 1 complexes. (j)–(l) Alaz–R–H_2O 1: 2 complexes.

structure **a**, with water molecule H-bonding at the H2 atom of Alaz. $\Delta E_{T-model}$ of structure **b** is 7.64 kJ mol^{-1} higher than structure **a**. This could be attributed to weak repulsion between the CH_3 group of Alaz and water in structure **b**. The N–H2···Ow and Ow–Hw···O1 H-bond distances are 2.81 and 2.69 Å, respectively. Structures **c** is considerably less stable than structures **a** and **b**. Water molecule in structures **c** acts only as proton donors and H-bonds simultaneously at O1 and O2 of Alaz, with $\Delta E_{T-model}$ of -56.17 kJ mol^{-1}.

For the Alaz–H_2O 1: 2 complex, both water molecules prefer to bridge between the NH_3^+ and COO^- groups of Alaz, structure **d** in Fig. 2. Structure **d** possesses $\Delta E_{T-model}$ of -152.32 kJ mol^{-1}. Structures **e** and **f** seem to have comparable stability. For structure **e** and **f**, H-bonds between water molecules extend to O_2, with the interaction energies of -140.92 and -139.44 kJ mol^{-1}, respectively.

The situations in the Alaz–R–H_2O complexes are similar to those in the Alaz–H_2O complexes. In general, the interaction energies of the Alaz–R–H_2O complexes are lower than the Alaz–H_2O complexes. The T-model potentials predicted the structure in which water molecule H-bonds simultaneously at H4 and O1, structure **g** in Fig. 2, to be the absolute minimum energy geometry of the Alaz–R–H_2O 1: 1 complex. In this case, the interaction energy amounts to -87.24 kJ mol^{-1}, with the Ow-Hw···O1 and N–H4···Ow distances of 2.72 and 2.74 Å, respectively. The H-bond distances are slightly longer than those in structure **a**. The interaction energies of structures **h** and **i** are comparable and slightly higher than structure **g**. The H-bond features in structure **h** are similar to structure **b**, in which water molecule bridges between H2 and O1. For structure **i**, water molecule acts simultaneously as proton acceptor and donor forming four H-bonds with Alaz-R, two N–H···Ow H-bonds and two Ow–Hw···O H-bonds.

The absolute minimum energy geometry of the Alaz-R-H_2O 1: 2 complex is structure **j** in Fig. 2. Structure **j** is about 14 kJ mol^{-1} more stable than structure **d**. It is characterized by four H-bonds between water molecules bridging between the NH_3^+ and COO^- groups of Alaz-R, similar to that in the Alaz–H_2O 1: 2 complex. The interaction energy of structure **j** amounts to -165.93 kJ mol^{-1}. Structures **k** and **l** possess comparable interaction energies, with rather complicate H-bond network. $\Delta E_{T-model}$ for structures **k** and **l** are -147.26 and -145.24 kJ mol^{-1}, respectively.

In order to check the T-model results on the Alaz–H_2O and Alaz–R–H_2O 1: *n* complexes, the lowest-lying energy geometries in Fig. 2 were reoptimized using MP2 calculations, with quite large basis sets. The MP2 results confirmed that they are all reasonable and the T-model potentials could be further applied in MD simulations with confidence. The MP2 results could be sent on request.

3.2. MD simulations on $[Alaz]_{aq}$ and $[Alaz\text{-}R]_{aq}$

The MD simulation parameters employed in the present work are given in Table 1, together with $\langle E_{aq}^{pot}\rangle$, $\langle E_{aq}^{solu-solv}\rangle$ and $\langle E_{aq}^{solv-solv}\rangle$. Selected PDO maps for Alaz and Alaz-R, computed with respect to reference planes I, II and III, are displayed in Figs. 3 and 4, respectively. The corresponding AWPD and AW-WWPD maps are also included in Figs. 3 and 4. The PDH and WWPD maps are not presented

Table 1
MD simulation parameters and the results on $[Alaz]_{aq}$ and $[Alaz\text{-}R]_{aq}$

	L (Å)	$\langle E_{aq}^{pot}\rangle$	$\langle E_{aq}^{solu-solv}\rangle$	$\langle E_{aq}^{solv-solv}\rangle$
$[Alaz]_{aq}$	20.8888	-31.72 ± 0.22	-1.7326	-29.0697
$[Alaz\text{-}R]_{aq}$	20.8888	-31.58 ± 0.22	-1.7633	-28.9103

The number of water molecules in all MD simulations is three hundreds. Energies are in kJ mol^{-1}. L, simulation box length; $\langle E_{aq}^{pot}\rangle$, average potential energy of aqueous solution; $\langle E_{aq}^{solu-solv}\rangle$, average solute–solvent interaction energy; $\langle E_{aq}^{solv-solv}\rangle$, average solvent–solvent interaction energy.

to limit the number of figures. Some high-density contour areas on the PDO, AWPD and AW-WWPD maps in Figs. 3 and 4 are labeled with letters. The values of the highest probabilities at the labeled areas on the PDO maps, as well as the corresponding lowest average interaction energies on the AWPD, WWPD and AW-WWPD maps, are summarized in Table 2. They are denoted by $\langle P^{PDO}\rangle_{max}$, $\langle \Delta E_{aq}^{AWPD}\rangle_{min}$, $\langle \Delta E_{aq}^{WWPD}\rangle_{min}$ and $\langle \Delta E_{aq}^{AW-WWPD}\rangle_{min}$, respectively. The cross section plots for selected H-bond networks at the NH_3^+ and COO^- groups are displayed in Figs. 5 and 6, respectively, with the lowest energy minima set to 0 kJ mol^{-1} to compare $\langle \Delta E_{aq}^{L}\rangle$ and $\langle \Delta E_{aq}^{T}\rangle$. Since the structures of $g(R)$ for $[Alaz]_{aq}$ and $[Alaz\text{-}R]_{aq}$ are not substantially different, only some characteristic peak positions directly related to the H-bonds between water and the NH_3^+, CH_3, CH and COO^- groups are given in Table 3, together with $n(R)$. Table 4 lists $\langle R_{A-H\cdots B}\rangle$, $\langle \theta_{A-H\cdots B}\rangle$ and $\tau_{A-H\ldots B,\,max}$.

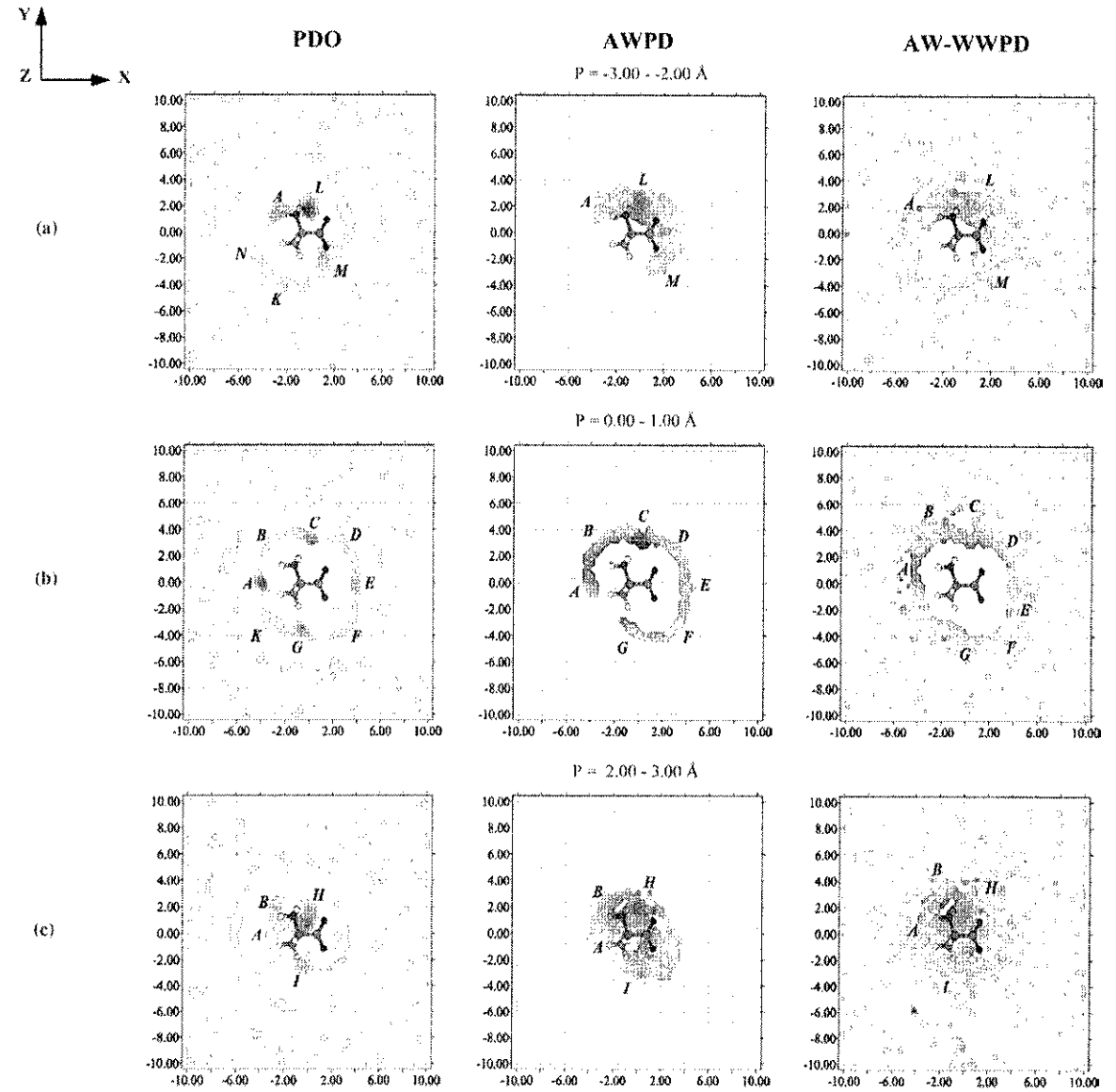

Fig. 3. Selected PDO, AWPD and AW-WWPD maps for $[Alaz]_{aq}$ obtained from MD simulations. X-, Y- and Z-axis are in Å. (a)–(c) Results with respect to reference plane I. (d)–(e) Results with respect to reference plane II. (f)–(g) Results with respect to reference plane III.

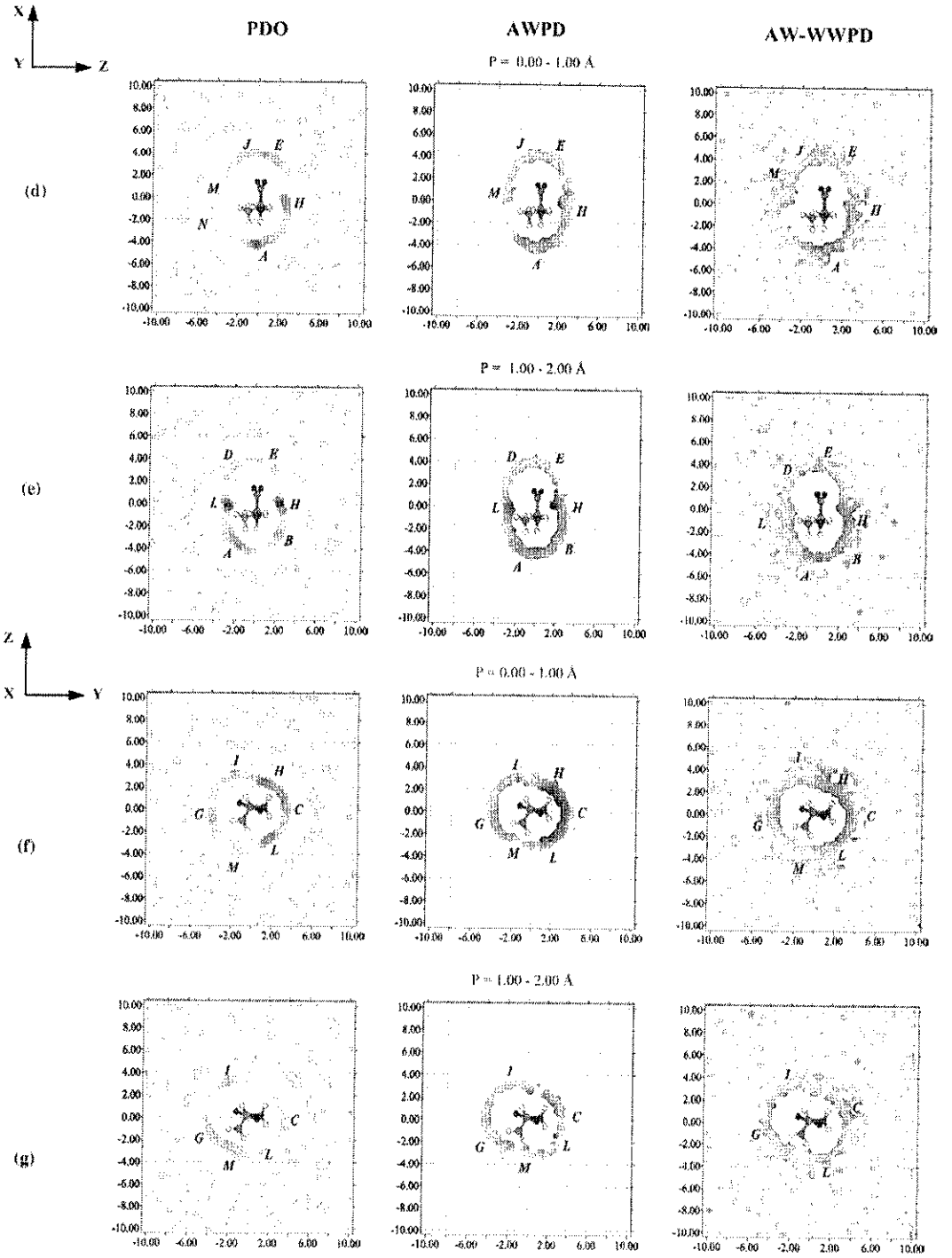

Fig. 3 (*continued*)

3.2.1. Hydration structures and H-bond networks in aqueous solutions

3.2.1.1. [Alaz]$_{aq}$. For [Alaz]$_{aq}$, the preferential hydration sites are labeled with **A**–**N** on the PDO maps in Fig. 3. At least nine well defined hydration sites are observed on the PDO maps of Alaz, five at the NH_3^+ group and four at the COO^- group. The hydration sites labeled with **A**, **B**, **C**, **H** and **L** involve the NH_3^+ group, whereas those with **D**, **E**, **F** and **J** are at the COO^- group. The PDO maps reveal that water molecules form more well defined H-bond networks at the NH_3^+ group, compared to the COO^- group.

According to $\langle P^{PDO}\rangle_{max}$ in Table 2, the order of the preferential hydration at the NH_3^+ group is written as:

H ≥ L > A > C > B.

Combination of Fig. 3a–e shows the three-dimensional structures of the H-bond networks at the NH_3^+ group in details. The hydration sites labeled with **L**, **H** and **A** are located near H2, H4 and H3, respectively. Water molecules at **A** seem to bridge the NH_3^+ and CH_3 groups, whereas the ones at **L** and **H** link between the NH_3^+ and COO^- groups. These H-bonding features are similar to those observed in the Alaz-H_2O 1: 1 and 1: 2 complexes, Fig. 2a, b and d, respectively. Fig. 3b and f suggests that water molecules at **C** are located between and slightly above **L** and **H**. They also reveal the possibility for water molecules at **C** to bridge the NH_3^+ and COO^- groups. This rules out the possibility to form intramolecular H-bond between the NH_3^+ and COO^- groups. The result is in accordance with the ^{17}O–NMR relaxation study on $[Gly]_{aq}$ and MC results in Refs. [58] and [20], respectively. Additional information on the hydration structures at the NH_3^+ group can be inferred from Fig. 3a and c–e. The H-bond networks at **A** and **H** are observed to link together in Fig. 3c and d. Whereas the H-bond networks bridging between **A** and **L** are recognized in Fig. 3a and e.

Water molecules at **D**, **E**, **F** and **J** constitute the three-dimensional structures of the H-bond networks at

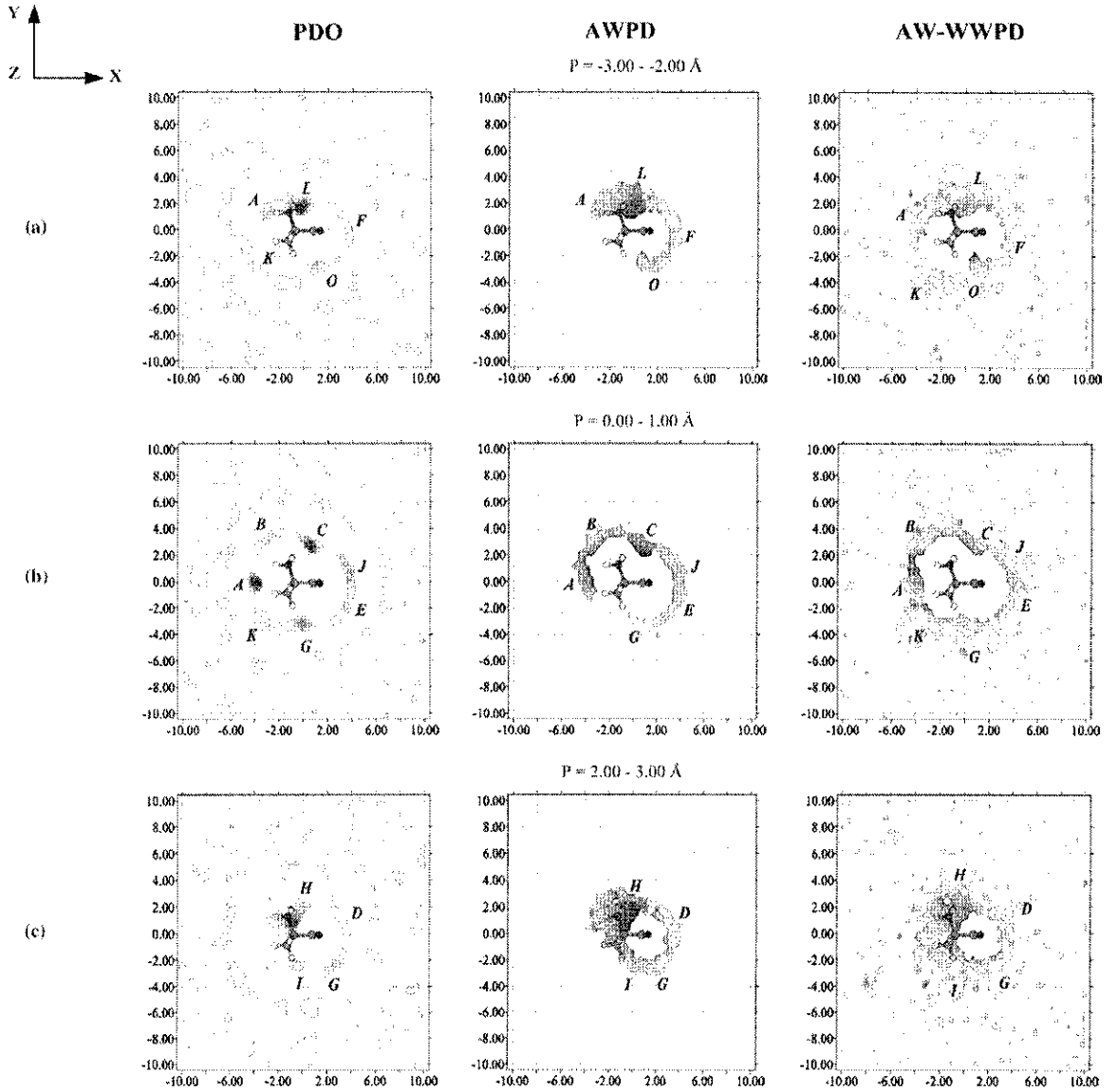

Fig. 4. Selected PDO, AWPD and AW-WWPD maps for $[Alaz\text{-}R]_{aq}$ obtained from MD simulations. *X*-, *Y*- and *Z*-axis are in Å. (a)–(c) Results with respect to reference plane I. (d)–(e) Results with respect to reference plane II. (f)–(g) Results with respect to reference plane III.

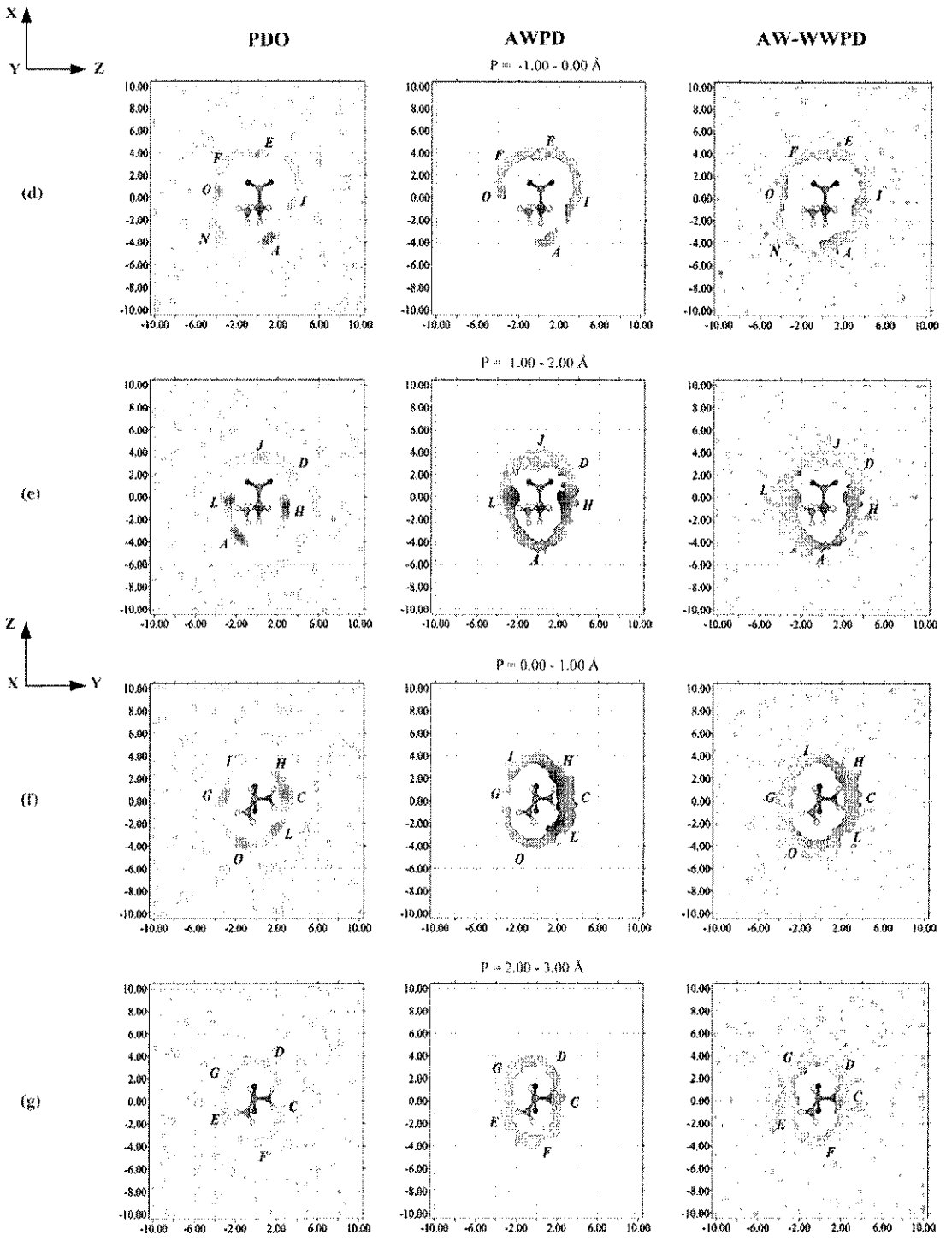

Fig. 4 (*continued*)

the COO^- groups. $\langle P^{PDO}\rangle_{max}$ in Table 2 reveal the highest probability at **J**, followed by those at **E**, **D** and **F**, respectively. From Table 2, the order of the preferential hydration at the COO^- group is written as:

$$\mathbf{J} \geq \mathbf{E} > \mathbf{D} \approx \mathbf{F}.$$

Fig. 3b suggests weak H-bond networks spanning from **D** to **E** to **F** in the vicinity of the COO^- groups. Whereas Fig. 3d suggests the possibility for a weak H-bond network linking between **J** and **E**. Weak H-bond networks are also recognized at the CH group. They span from **G** to **K**, as seen in Fig. 3b.

Table 2
The highest probabilities ($\langle P^{PDO}\rangle_{max}$) at the labeled areas on the PDO maps and the corresponding lowest average interaction energies ($\langle \Delta E^X_{aq}\rangle_{min}$) on the X maps in Figs. 3 and 4

	$\langle P^{PDO}\rangle_{max}$	$\langle \Delta E^{AWPD}_{aq}\rangle_{min}$	$\langle \Delta E^{WWPD}_{aq}\rangle_{min}$	$\langle \Delta E^{AW-WPD}_{aq}\rangle_{min}$
[Alaz]$_{aq}$				
NH_3^+				
A	0.066	−47.55	−77.58	−91.44
B	0.049	−47.27	−74.55	−90.25
C	0.051	−71.06	−74.77	−95.64
H	0.086	−67.27	−74.11	−87.74
L	0.084	−62.06	−79.92	−87.07
$\langle \Delta E^X_{aq}\rangle_{min,\,av}$		−59.04	−76.18	−90.43
COO^-				
D	0.022	−37.60	−70.63	−88.86
E	0.034	−48.78	−84.31	−87.99
F	0.022	−36.92	−69.48	−86.06
J	0.036	−41.53	−75.30	−82.34
$\langle \Delta E^X_{aq}\rangle_{min,\,av}$		−41.21	−74.93	−86.31
[Alaz-R]$_{aq}$				
NH_3^+				
A	0.094	−52.64	−86.92	−96.07
B	0.041	−51.38	−73.74	−84.05
C	0.074	−70.84	−74.93	−87.15
H	0.097	−70.63	−76.01	−91.87
L	0.105	−68.32	−67.50	−86.40
$\langle \Delta E^X_{aq}\rangle_{min,\,av}$		−62.76	−75.82	−89.11
COO^-				
D	0.024	−52.26	−65.55	−81.47
E	0.027	−36.94	−73.49	−84.01
F	0.022	−36.37	−68.65	−85.54
J	0.032	−37.58	−73.72	−82.11
O	0.042	−47.37	−63.31	−89.17
$\langle \Delta E^X_{aq}\rangle_{min,\,av}$		−42.10	−68.94	−84.46

X is AWPD, WWPD or AW-WWPD. Energies are in kJ mol^{-1}. $\langle \Delta E^X_{aq}\rangle_{min,\,av} = \sum_{i=1}^{N} \frac{\langle E^X_{aq}\rangle_i}{N}$: N, number of the labeled area.

Further information on the hydration structures of Alaz can be obtained from $g(R)$ and $n(R)$ in Table 3. For $[Alaz]_{aq}$, the main peak of $g(R_{N-Ow})$ is located at $R=2.76$ Å, with $n(R_{N-Ow})$ of about two (2.17). $n(R_{N-Ow})$ in this case represents the average number of water molecules in close contact with the NH_3^+ group. The integration of $g(R_{N-Ow})$ to the first minimum at $R=3.61$ Å yielded more than six (6.34) water molecules in the first hydration shell of the NH_3^+ group. These could be attributed to the water molecules at **L**, **H**, **A**, **B** and **C** on the PDO maps. The structures of the main peaks of $g(R_{O1-Ow})$ and $g(R_{O2-Ow})$ are quite similar. The main peaks of both $g(R_{O1-Ow})$ and $g(R_{O2-Ow})$ are located at $R=2.81$ Å, with $n(R_{O1-Ow})$ and $n(R_{O2-Ow})$ of 1.49 and 1.38, respectively. Therefore, on average, more than one water molecule is in close contact with O1 and O2, and the degree of hydration at O1 is roughly 8% higher than at O2. The latter could result from the formation of well defined H-bond networks linking between O1 and the NH_3^+ group. Combination of $g(R_{O1-Ow})$ and $g(R_{O2-Ow})$ yielded $g(R_{O-Ow})$, with the main peak position at 2.81 Å and $n(R_{O-Ow})$ of 1.44. Since the C–H···O H–bond is weak in general, the hydration structures at the CH_3 and CH groups were not easy to elucidate. However, $n(R_{C2-Ow})$ and $n(R_{C3-Ow})$ indicate more than three (3.60) water molecules at the CH group and about three (2.95) water molecules at the CH_3 group.

3.2.1.2. [Alaz-R]$_{aq}$. The hydration structures in $[Alaz\text{-}R]_{aq}$ are slightly different from $[Alaz]_{aq}$. Since the COO^- plane in Alaz-R is 90° with respect to the $NC^{\alpha}C$ backbone plane, the NH_3^+ and COO^- groups are more distant compared to Alaz. The O1···H4 and O2···H2 distances in Alaz-R are almost identical (about 2.98 Å). Thus, the steric effects at O1 due to H2 and H4 in Alaz are reduced in Alaz-R. The MD results show that the rotation of the COO^- plane allows the charged functional groups to expose more with water, leading to an increase in the structure and the degree of hydration, especially at the NH_3^+ group. The following discussion supports this scenario.

An increase in the degree of hydration at the NH_3^+ group is recognized from the PDO maps in Fig. 4a, b and e. Table 2 confirms that the probability distributions at **L**, **H**, **A** and **C** are higher than $[Alaz]_{aq}$. Compare to $[Alaz]_{aq}$, $\langle P^{PDO}\rangle_{max}$ at **C** and **A** are increased by about 45 and 42%, respectively. At the NH_3^+ group, only $\langle P^{PDO}\rangle_{max}$ at **B** is decreased, about 16%. Thus, the order of the preferential hydration at the NH_3^+ group in $[Alaz\text{-}R]_{aq}$ is written as:

$$\mathbf{L} > \mathbf{H} \geq \mathbf{A} > \mathbf{C} > \mathbf{B}.$$

The degrees of hydration at the COO^- group are both increased and decreased due to the conformation change. $\langle P^{PDO}\rangle_{max}$ in Table 2 reveal that the degrees of hydration at **E** and **J** are decreased by about 21 and 11%, respectively, whereas at **D** increased by about 9%. The rotation of the $NC^{\alpha}C$ backbone plane creates a new H-bond network at **O**. Table 2 suggests that **O** possesses the highest $\langle P^{PDO}\rangle_{max}$ among the H-bond networks at the COO^- group. Water molecules at **O** link between the O2 atom and the CH_3 group of Alaz-R. Comparison of Fig. 4a and b suggests that the H-bond network at **O** is similar to **G**, which links between the O2 atom and the CH group. According to $\langle P^{PDO}\rangle_{max}$, the order of the preferential hydration at the COO^- group in $[Alaz\text{-}R]_{aq}$ is written as:

$$\mathbf{O} > \mathbf{J} > \mathbf{E} \geq \mathbf{D} \geq \mathbf{F}.$$

The information in Table 3 also confirms the increase in the degrees of hydration at the NH_3^+ and COO^- groups. For $[Alaz\text{-}R]_{aq}$, the position of the main peak of $g(R_{N-Ow})$ is at $R=2.76$ Å, with $n(R_{N-Ow})$ of 2.32. The value is about 7% higher than $[Alaz]_{aq}$. The number of water molecules in close contact with the COO^- group seems to experience larger effect upon the rotation of the $NC^{\alpha}C$ backbone plane. $n(R_{O-Ow})$ at the first maximum of $g(R_{O-Ow})$ amounts to 1.83, increased by about 27%. Therefore, on average, two water molecules are in close contact with the COO^- group in $[Alaz\text{-}R]_{aq}$.

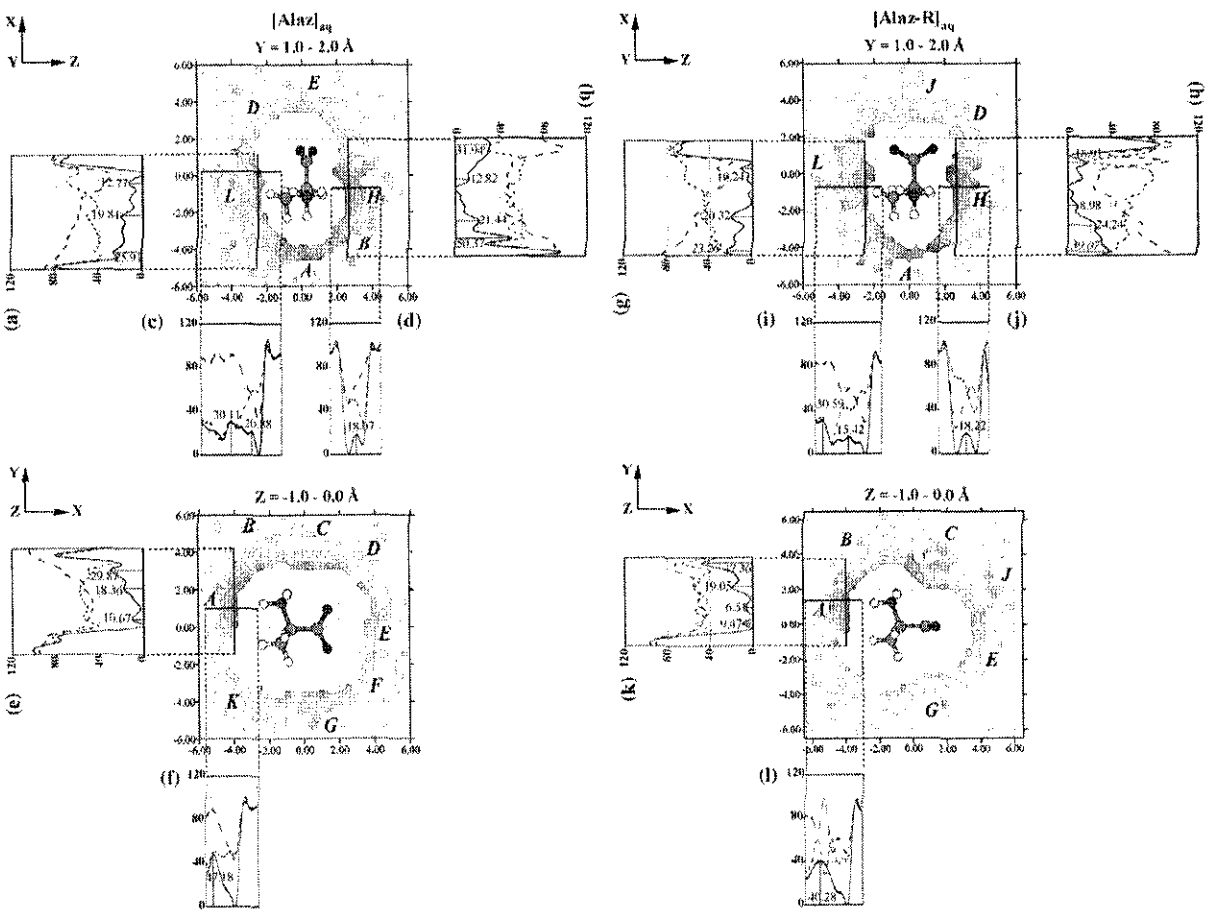

Fig. 5. Cross section plots for the H-bond networks at the NH_3^+ group of $[Alaz]_{aq}$ and $[Alaz\text{-}R]_{aq}$. *X*-, *Y*- and Z-axis are in Å. The lowest energy minima of the cross sections were set to 0 kJ mol^{-1} for comparison. ——— for the AW-WWPD maps, – – – for the AWPD maps, - - - - for the WWPD maps.

3.2.2. Average potential energy landscapes at the H-bond networks

General trends of the interaction energies of polar solute in aqueous solutions were observed in $[Alaz]_{aq}$ and $[Alaz\text{-}R]_{aq}$. The preferential hydrations were resulted more or less from the combined effects of solute–solvent and solvent–solvent interactions. The trend of $\langle E_{aq}^{pot} \rangle$ in Table 1 is in accordance with Ref. [21], in which the overall stabilization by hydration was reported to be slightly larger for $[Alaz]_{aq}$ than $[Alaz\text{-}R]_{aq}$. $\langle E_{aq}^{solu-solv} \rangle$ in Table 1 suggested that the functional groups of Alaz-R is more accessible by water compared to Alaz.

It should be noted that the orders of the preferential hydration derived from the PDO maps and the stability orders from the AW-WWPD maps are not necessarily the same. This is due to the fact that the PDO maps represent the local probability distributions, which could be compared with the electron density maps obtained from the X-ray diffraction experiment. Whereas the minima on the AW-WWPD maps are associated with the interaction energy states, which might be occupied or unoccupied at a given MD timestep. Moreover, the occupancies of these interaction energy states depend primarily on the hydration dynamics of water molecules and the transition energy barriers connecting these states. Therefore, it is inappropriate to directly correlate these structural and energetic properties.

Similar argument was made on the hydration sites inferred from X-ray crystallography and NMR experiment. X-ray crystallography measures the extent to which a given hydration site is occupied by water molecules. It cannot distinguish a long-lived water molecule occupying a site from that involving rapid exchange. Whereas NMR experiment monitors a particular water molecule, which resides sufficiently long at a hydration site before being replaced by another water molecule [4]. NMR experiment is, therefore, more appropriate in the investigation of hydration dynamics of water molecules, such as the rate of water exchange or the residence times of water at specific functional groups of proteins.

Additional remarks should be made on the energy values in Table 2 and the cross sections in Figs. 5 and 6.

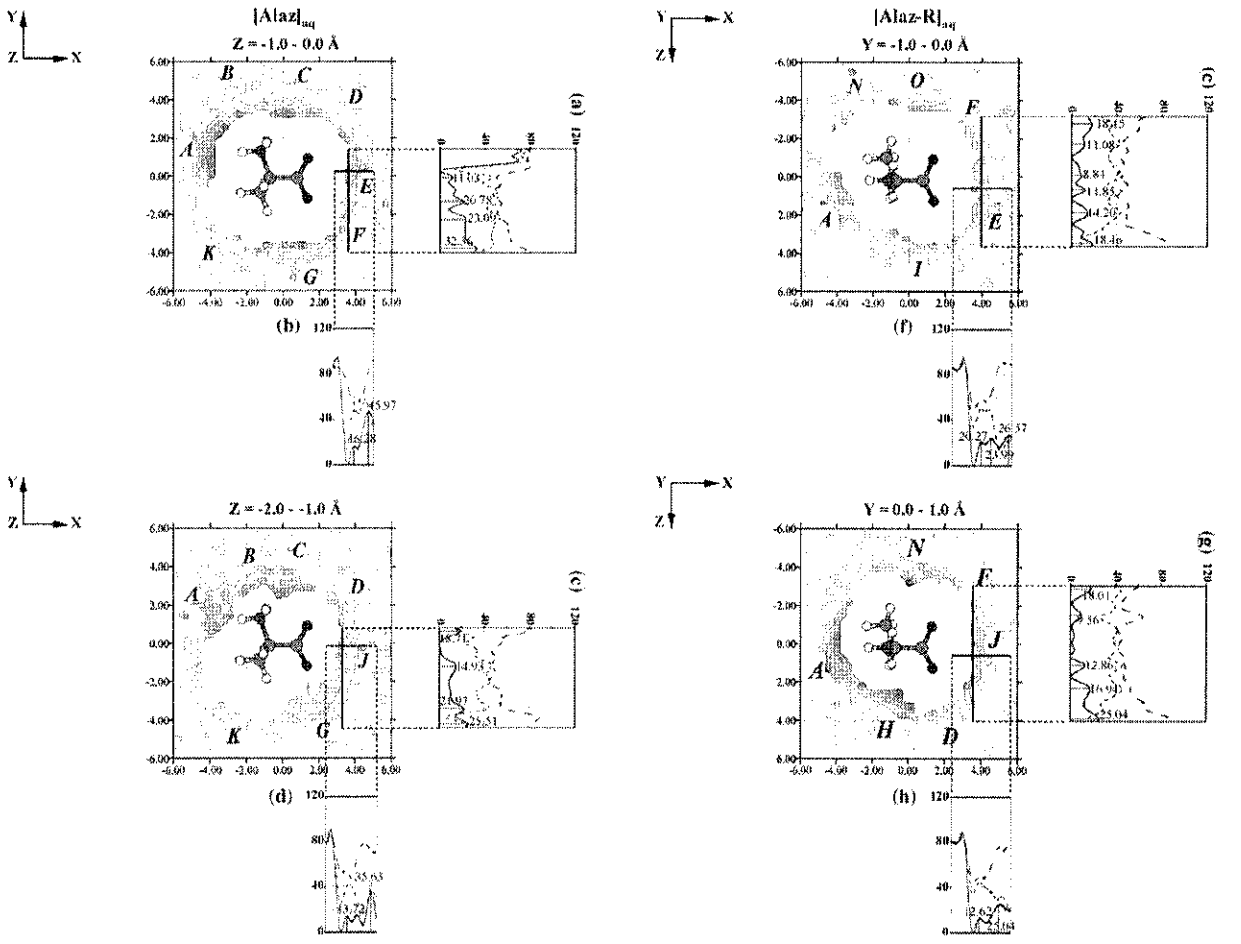

Fig. 6. Cross section plots for the H-bond networks at the COO^- group of $[Alaz]_{aq}$ and $[Alaz\text{-}R]_{aq}$. *X*-, *Y*- and *Z*-axis are in Å. The lowest energy minima of the cross sections were set to 0 kJ mol^{-1} for comparison. ——— for the AW-WWPD maps, – – – for the AWPD maps, - - - - for the WWPD maps.

Since the solute–solvent interactions are quite strong, especially at the NH_3^+ group, the minima on the AWPD and AW-WWPD maps are seen nearly at the same positions. In contrast, the minima on the WWPD maps are located at the boundary or outside the first hydration shells. These were recognized from the cross section plots derived from the transverse profile lines in Fig. 5. The observations support the statement in Ref. [27] that, in the first hydration shell, the stabilization by the solute–solvent interactions is accompanied by the destabilization of the solvent–solvent interactions and vice versa. In addition, it was also noticeable from the same cross section plots that the shapes of the average potential landscapes, especially in the first hydration shell of the NH_3^+ group, are determined by the solute–solvent interactions. Whereas Fig. 6 suggested that the shapes of the average potential landscapes at the COO^- group are influenced by the solvent–solvent interactions. The following discussion will focus on the energetic of particular H-bond networks mentioned in the previous subsections.

Table 3
Characteristic peak positions (R) related to H-bonds between water and the NH_3^+, CH_3, CH and COO^- groups, together with the corresponding running coordination number, $n(R)$, obtained from MD simulations

	R_{max}	$n(R_{max})$	R_{min}	$n(R_{min})$
[Alaz]aq				
$g(R_{N-Ow})$	2.76	2.17	3.61	6.34
$g(R_{O1-Ow})$	2.81	1.49	3.81	6.82
$g(R_{O2-Ow})$	2.81	1.38	3.96	7.25
$g(R_{O-Ow})$	2.81	1.44	3.96	7.35
$g(R_{C2-Ow})$	3.46	3.60	4.30	10.68
$g(R_{C3-Ow})$	3.21	2.95	4.01	8.54
[Alaz-R]aq				
$g(R_{N-Ow})$	2.76	2.32	3.66	6.43
$g(R_{O1-Ow})$	2.81	1.46	4.06	8.09
$g(R_{O2-Ow})$	2.81	1.39	3.91	6.80
$g(R_{O-Ow})$	2.86	1.83	3.91	7.11
$g(R_{C2-Ow})$	3.41	3.52	3.81	6.48
$g(R_{C3-Ow})$	3.21	2.92	4.10	9.15

Distances are in Å. R_{max}, position of the maximum of the main peak. R_{min}, position of the minimum of the main peak.

3.2.2.1. [Alaz]$_{aq}$. Table 2 shows that, at the NH_3^+ group, water molecules at **C** possess the strongest solute–solvent interaction, followed by **H**, **L**, **A** and **B**, respectively. The solvent–solvent interactions are, however, strongest at **L**, followed by **A**, **C**, **B** and **H** respectively. Based on $\langle \Delta E_{aq}^{AW-WWPD} \rangle_{min}$, the interaction energies at **A** and **B**, as well as **H** and **L**, are comparable and the stability order for the hydration at the NH_3^+ group is written as:

$$\mathbf{C > A \geq B > H \geq L}$$

At the NH_3^+ group, $\langle \Delta_{aq}^{AWPD} \rangle_{min,\,av}$, $\langle \Delta E_{aq}^{WWPD} \rangle_{min,\,av}$ and $\langle \Delta E_{aq}^{AW-WWPD} \rangle_{min,\,av}$ are -59.04, -76.18 and -90.43 kJ mol^{-1}, respectively.

Although the probability of finding water molecules at **H** is only slightly higher than **L** according to the PDO maps, $\tau_{A-H\cdots B,\,max}$ in Table 4 reveal that a particular water molecule stays at **H** much longer than **L**. This indicates that water exchange takes place more often at **L** compared to **H**. $\tau_{N-H2\cdots Ow,\,max}$ and $\tau_{N-H4\cdots Ow,\,max}$ are approximately 3 and 21 ps, respectively. The latter is comparable with the average residence time for charged atoms reported in Ref. [55], approximately 19 ps. The shapes of the average potential energy landscapes in Fig. 5 could help provide insight into the discrepancy between the H-bond lifetimes at **L** and **H**.

Table 4
Selected average H-bond distances ($\langle R_{A-H\cdots B} \rangle$) and angles ($\langle \theta_{A-H\cdots B} \rangle$), as well as the longest H-bond lifetimes ($\tau_{A-H\cdots B,\,max}$) derived from MD simulations

			SD
[Alaz]$_{aq}$			
NH_3^+			
N–H2···Ow	$\langle R_{N-H2\cdots Ow} \rangle$	2.94	0.20
	$\langle \theta_{N-H2\cdots Ow} \rangle$	15.23	5.20
	$\tau_{N-H2\cdots Ow,\,max}$	2.70	–
N–H3···Ow	$\langle R_{N-H3\cdots Ow} \rangle$	2.88	0.23
	$\langle \theta_{N-H3\cdots Ow} \rangle$	29.79	12.27
	$\tau_{N-H3\cdots Ow,\,max}$	19.44	–
N–H4···Ow	$\langle R_{N-H4\cdots Ow} \rangle$	2.90	0.23
	$\langle \theta_{N-H4\cdots Ow} \rangle$	27.24	10.76
	$\tau_{N-H4\cdots Ow,\,max}$	21.34	–
COO^-			
Ow–Hw···O1	$\langle R_{Ow-Hw\cdots O1} \rangle$	2.95	0.26
	$\langle \theta_{Ow-Hw\cdots O1} \rangle$	29.31	13.52
	$\tau_{Ow-Hw\cdots O1,\,max}$	5.08	–
Ow–Hw···O2	$\langle R_{Ow-Hw\cdots O2} \rangle$	3.12	0.29
	$\langle \theta_{Ow-Hw\cdots O2} \rangle$	31.20	13.91
	$\tau_{Ow-Hw\cdots O2,\,max}$	4.90	–
[Alaz-R]$_{aq}$			
NH_3^+			
N-H2···Ow	$\langle R_{N-H2\cdots Ow} \rangle$	2.86	0.21
	$\langle \theta_{N-H2\cdots Ow} \rangle$	29.11	15.52
	$\tau_{N-H2\cdots Ow,\,max}$	3.95	–
N–H3···Ow	$\langle R_{N-H3\cdots Ow} \rangle$	2.85	0.25
	$\langle \theta_{N-H3\cdots Ow} \rangle$	43.60	8.65
	$\tau_{N-H3\cdots Ow,\,max}$	17.93	–
N–H4···Ow	$\langle R_{N-H4\cdots Ow} \rangle$	2.87	0.28
	$\langle \theta_{N-H4\cdots Ow} \rangle$	34.74	12.21
	$\tau_{N-H4\cdots Ow,\,max}$	13.87	–
COO^-			
Ow–Hw···O1	$\langle R_{Ow-Hw\cdots O1} \rangle$	2.96	0.26
	$\langle \theta_{Ow-Hw\cdots O1} \rangle$	31.65	13.54
	$\tau_{Ow-Hw\cdots O1,\,max}$	4.77	
Ow–Hw···O2	$\langle R_{Ow-Hw\cdots O2} \rangle$	3.11	0.32
	$\langle \theta_{Ow-Hw\cdots O2} \rangle$	31.59	13.94
	$\tau_{Ow-Hw\cdots O2,\,max}$	2.98	–

SD, Standard Deviation. A–H···B, H-bond donor–acceptor pair between molecules A and B. Distances, angles and times are in Å, degree and ps, respectively.

Investigation of Fig. 5 and Table 4 suggested that the mobility of water molecules at specific hydration sites depend on the structures of the energy valleys of the H-bond networks. Fig. 5a–d clearly show that the shapes of the average potential energy landscapes at **L** and **H** are different in details. Within the range from $X=0.0$ to -4.0 Å, for example, the transition energy barriers ($\langle \Delta E_{aq}^{L} \rangle$) at **L** vary approximately from 0 to 13 to 20 to 26 kJ mol^{-1}, whereas at **H** approximately from 13 to 0 to 50 kJ mol^{-1}, respectively. Fig. 5c and d also reveal that $\langle \Delta E_{aq}^{T} \rangle$ at the boundary of **L** is only about 30 kJ mol^{-1}, whereas at **H** amounts to about 100 kJ mol^{-1}. Since the transition energy barriers within and at the boundary of the H-bond network at **L** are lower, the mobility of water molecules at **L** is expected to be higher than at **H**. This should allow water exchanges with the bulk to take place easier and faster at **L** compared to **H**. The discussions on the transition energy barriers explain why the H-bond lifetime at **L** is considerably shorter than **H**.

Fig. 5b and d also suggested that the motion of water molecules at **H** is rather restricted, within a narrow energy valley of about 3 Å widths and about 4 Å from the Alaz molecular plane. The latter supports the rough estimation in Ref. [27] that the amino acid–water interaction drops sharply within 5 Å, and the thickness of the perturbed water layer is about one water molecule. Similar explanation could be made on the average potential energy landscape at **A**. The shapes of the average potential energy landscape at **A** are between **L** and **H**. The cross section derived from the longitudinal profile line in Fig. 5e is similar to Fig. 5b, whereas that obtained from the transverse profile line in Fig. 5f is similar to Fig. 5c. This suggested that specific water molecules at **A** are more localized compared to **L**, but less localized compared to **H**. $\tau_{N-H3\cdots Ow,\,max}$ in Table 4 confirms this suggestion.

Due to weak H-bond interaction with water, the situation at the COO^- group is rather complicated. In Table 2, $\langle \Delta E_{aq}^{AWPD} \rangle_{min,\,av}$ for the interaction between the COO^- group and water is -41.21 kJ mol^{-1}, about 18 kJ mol^{-1} higher than the NH_3^+ group. Whereas $\langle \Delta E_{aq}^{WWPD} \rangle_{min,\,av}$ is not substantially different from the NH_3^+ group. According to $\langle \Delta E_{aq}^{AW-WWPD} \rangle_{min}$, the stability order for the hydration at the COO^- group is written as:

$$\mathbf{D > E > F > J}.$$

Some difficulties were encountered in the analysis of the average potential energy landscapes at the COO^- group. Fig. 6 shows energy valleys, which are not very well defined

on the AW-WWPD maps. Due to the fact that the O1 and O2 atoms being adjacent makes it difficult to specify the boundaries of the hydration shells. Therefore, attempt will not be made to directly correlate $\tau_{A-H\cdots B,\,max}$ with the size and shape of the cross sections at the COO^- group. It appears in general in Fig. 6a–d that the energy valleys at the COO^- group are shallower compared to the NH_3^+ group. Hence, the shapes of the cross sections at the COO^- group allow water molecules to move in a wider range, especially within the areas between the O1 and O2 atoms. These could help promote the water exchanges at the COO^- group, both within and between the H-bond networks, as well as between the H-bond networks and the bulk. For examples, the transition energy barrier to the water exchange between the H-bond networks at **E** and **F** in Fig. 6a is approximately 32 kJ mol^{-1}, and that between **J** and **G** about 26 kJ mol^{-1}. These are in line with the values of $\tau_{Ow-Hw\cdots O1,\,max}$ and $\tau_{Ow-Hw\cdots O2,\,max}$, being only about 5 ps.

3.2.2.2. [Alaz-R]$_{aq}$. In general, the rotation of the COO^- plane led to slightly stronger solute–solvent interactions both at the NH_3^+ and COO^- groups. This is evident in Table 2, in which $\langle \Delta E_{aq}^{AWPD} \rangle_{min,\,av}$ at the NH_3^+ and COO^- groups are decreased by about 4 and 1 kJ mol^{-1}, respectively. $\langle \Delta E_{aq}^{WWPD} \rangle_{min,\,av}$ is nearly unchanged at the NH_3^+ group, whereas at the COO^- group increased by about 6 kJ mol^{-1}.

Water molecules at **A** possess the lowest $\langle \Delta E_{aq}^{AW-WWPD} \rangle_{min}$, followed by **H**, **C**, **L** and **B**, respectively. Based on $\langle \Delta E_{aq}^{AW-WWPD} \rangle_{min}$, the stability order at the NH_3^+ group is written as:

$$\mathbf{A > H > C > L > B}$$

The rotation of the COO^- plane brought about little change at the energy valleys at **L** and **A**. It, however, created visible changes on the shapes of the cross section and $\langle \Delta E_{aq}^{L} \rangle$ at **H**. Within the range from $X=0.0$ to -4.0 Å, the highest $\langle \Delta E_{aq}^{L} \rangle$ at **H** reduces from approximately 50 kJ mol^{-1} in $[Alaz]_{aq}$ to 39 kJ mol^{-1} in $[Alaz\text{-}R]_{aq}$. This increases the mobility of water molecules at **H**, as evident from the reduction of $\tau_{N-H4\cdots Ow,\,max}$, from about 21 ps in $[Alaz]_{aq}$ to 14 ps in [Alaz-R].

Based on $\langle \Delta E_{aq}^{AW-WWPD} \rangle_{min}$, the stability order at the COO^- group is written as:

$$\mathbf{O > F > E \geq J > D}.$$

Fig. 6a–h show that the rotation of the COO^- plane brought about remarkable changes in the cross sections at the COO^- group. The cross sections at **E** and **J** are discussed as examples. The energy barriers to the water exchange between the H-bond networks, discussed in the previous subsections, are considerably reduced in general. In Fig. 6e, $\langle \Delta E_{aq}^{L} \rangle$ at **E** are about 18 kJ mol^{-1} at most. Fig. 6d and h also illustrate that, upon rotation, $\langle \Delta E_{aq}^{T} \rangle$ at **J** is virtually reduced, from about 36 to 25 kJ mol^{-1}. This should increase the rate of water exchange between the H-bond network and the bulk, as well as allow water molecules to move in a quite wider range in the area of the COO^- group, compared to $[Alaz]_{aq}$.

4. Conclusion

Structures and energetic of the H-bond networks of water molecules at the charged functional groups of two forms of alanine zwitterions (Alaz and Alaz-R) were investigated, using intermolecular potentials derived from the T-model and MD simulations. In order to study the effects of conformation change on the structures and energetic of the H-bond networks, the MD results on $[Alaz]_{aq}$ and $[Alaz\text{-}R]_{aq}$ were analyzed extensively. General trends of the interaction energies of polar solute in aqueous solutions were observed from the MD results. The preferential hydrations resulted more or less from combined effects of solute–solvent and solvent–solvent interactions, as well as the hydration dynamics of water molecules in the first hydration shell. The PDO maps clearly illustrated the three-dimensional structures of the H-bond networks of water at both NH_3^+ and COO^- groups. It was recognized that water molecules establish more well defined H-bond networks at the NH_3^+ group, compared to the COO^- group. For both $[Alaz]_{aq}$ and $[Alaz\text{-}R]_{aq}$, the PDO maps confirmed that water molecules form H-bond networks between the NH_3^+ and COO^- groups, which rules out the possibility for the two charged functional groups to form intramolecular H-bond. This is in accordance with the previous ^{17}O-NMR relaxation study and MC simulations on $[Gly]_{aq}$.

For $[Alaz]_{aq}$, at least five H-bond networks were observed at the NH_3^+ group and more than four at the COO^- group. It was recognized that, the orders of the preferential hydration derived from the PDO maps and the stability orders inferred from the AWPD, WWPD and AW-WWPD maps are not the same. This is due to the fact that the minima on the AW-WWPD maps are associated with the interaction energy states, which might be occupied or unoccupied at a given MD timestep. The occupancies of the interaction energy states depend on the hydration dynamics of individual water molecule, as well as the transition energy barriers interconnected these states.

Attempt was made in the present work to correlate the sizes and shapes of the average potential energy landscapes at the H-bond networks with the H-bond lifetimes. To serve this purpose, cross section plots at the H-bond networks were constructed from the AW-WWPD maps. In the present work, the mobility of water molecules and the possibilities for the water exchanges within and between the H-bond networks, as well as between the H-bond networks and the outsides, were discussed based on $\tau_{A-H\cdots B,\,max}$, $\langle \Delta E_{aq}^{L} \rangle$ and $\langle \Delta E_{aq}^{T} \rangle$. The structures of the energy valleys suggested that, at the NH_3^+ group, water exchanges within the H-bond networks seem to take place easier and faster than between

the H-bond networks and the bulk. Whereas at the COO^- group, water molecules could move or exchange very rapidly within a wider range.

The rotation of the COO^- plane 90° with respect to the $NC^{\alpha}C$ backbone seems to create changes more or less in both structures and energetic of the H-bond networks. It was confirmed in the present work that, although the functional groups of Alaz-R are more accessible by water, the overall stabilization by hydration is larger for $[Alaz]_{aq}$ than $[Alaz\text{-}R]_{aq}$. On average, the solute–solvent interactions are stronger and the solvent–solvent interactions are weaker at the H-bond networks of both NH_3^+ and COO^- groups. The rotation of the COO^- plane created an additional well defined H-bond network at the COO^- group, and brings about changes in $\langle \Delta E_{aq}^{L} \rangle$ and $\langle \Delta E_{aq}^{T} \rangle$. They are considerably reduced at the COO^- group, allowing water molecules to move or exchange within a wider range. The present results implied that, complete information on molecular hydration could be obtained only when explicit water molecules, together with their hydration dynamics at the hydration sites, are considered in the model calculations.

Acknowledgments

All calculations were performed at the School of Chemistry, Suranaree University of Technology. The authors would like to acknowledge the financial support from the Thailand Research Fund (TRF) through the Royal Golden Jubilee (RGJ) Ph.D. Program (Grant no. PHD/00155/2541) to Supaporn Dokmaisrijan and Prof. Dr. Kritsana Sagarik. This work was supported in part by the ASEA-UNINET and the University of Innsbruck, Austria.

References

[1] E.N. Baker, in: R.B. Gregory (Ed.), Protein–Solvent Interactions, Marcel Dekker, New York, 1994 Chapter 2.
[2] M. Levitt, B.H. Park, Structure 1 (1993) 223.
[3] N. Thanki, J.M. Thornton, J.M. Goodfellow, J. Mol. Biol. 202 (1988) 637.
[4] M. Billeter, Prog. Nucl. Magn. Reson. Spectrosc. 27 (1995) 635.
[5] V. Makarov, B.M. Pettitt, M. Feig, Acc. Chem. Res. 35 (2002) 376.
[6] B.P. Schoenborn, A. Garcia, R. Knott, Prog. Biophys. Mol. Biol. 64 (1995) 105.
[7] J.-S. Jiang, A.T. Brünger, J. Mol. Biol. 243 (1994) 100.
[8] M. Feig, B.M. Pettitt, Structure 6 (1998) 1351.
[9] J.A. Rupley, in: S.N. Timasheff, G.D. Fasman (Eds.), Structure and Stability of Biological Macromolecules, Marcel Dekker, New York, 1969, pp. 291–352.
[10] B.W. Mathews, in: H. Neurath, R.L. Hill (Eds.), The Proteins vol. III, Academic Press, New York, 1977, pp. 403–590.
[11] J.H. Jensen, M.S. Gordon, J. Am. Chem. Soc. 117 (1995) 8159.
[12] Y. Ding, K. Krogh-Jespersen, J. Comput. Chem. 17 (1996) 338.
[13] H.S. Rzepa, M.Y. Yi, J. Chem. Soc., Perkin Trans. 2 (1991) 531.
[14] S.-W. Park, D.-S. Ahn, S. Lee, Chem. Phys. Lett. 371 (2003) 74.
[15] E. Tajkhorshid, K.J. Jalkanen, S. Suhai, J. Phys. Chem. B 102 (1998) 5899.
[16] E. Clementi, F. Cavallone, R. Scordamaglia, J. Am. Chem. Soc. 99 (1977) 5531.
[17] G. Alagona, C. Ghio, P.A. Kollman, J. Mol. Struct. (Theochem) 166 (1988) 385.
[18] I. Tuñón, E. Silla, C. Millot, M.T.C. Martins-Costa, M.F. Ruiz-López, J. Phys. Chem. A 102 (1998) 8673.
[19] O. Kikuchi, T. Matsuoka, H. Sawahara, O. Takahashi, J. Mol. Struct. (Theochem) 305 (1994) 79.
[20] T. Watanabe, K. Hashimoto, H. Takase, O. Kikuchi, J. Mol. Struct. (Theochem) 397 (1997) 113.
[21] O. Kikuchi, T. Watanabe, Y. Ogawa, H. Takase, O. Takahashi, J. Phys. Org. Chem. 10 (1997) 145.
[22] M. Mezei, P.K. Mehrotra, D.L. Beveridge, J. Biomol. Struct. Dyn. 2 (1984) 1.
[23] S.G. Kalko, E. Guàrdia, J.A. Padró, J. Phys. Chem. B 103 (1999) 3935.
[24] P. Mark, L. Nilsson, J. Phys. Chem. B 105 (2001) 8028.
[25] G. Castronuovo, V. Elia, F. Velleca, J. Chem. Soc., Faraday Trans. 92 (1996) 4215.
[26] M. Suzuki, J. Shigematsu, Y. Fukunishi, T. Kodama, J. Phys. Chem. B 101 (1997) 3839.
[27] E. Clementi, Computational Aspects for Large Chemical Systems, Lecture Notes in Chemistry, vol. 19, Springer, Berlin, 1980.
[28] J.G. Gaffney, R.C. Pierce, L. Friedman, J. Am. Chem. Soc. 99 (1977) 4293.
[29] H. Kimura, K. Nakamura, A. Eguchi, H. Sugisawa, K. Deguchi, K. Ebisawa, E. Suzuki, A. Shoji, J. Mol. Struct. 447 (1998) 247.
[30] M.S. Lehmann, T.F. Koetzle, W.C. Hamilton, J. Am. Chem. Soc. 94 (1972) 2657.
[31] C. Desfrançois, S. Carles, J.P. Schermann, Chem. Rev. 100 (2000) 3943.
[32] D.A. Smith, Modeling the Hydrogen-Bond, American Chemical Society, Washington DC, 1994.
[33] K.P. Sagarik, R. Ahlrichs, J. Chem. Phys. 86 (1987) 5117.
[34] K.P. Sagarik, E. Spohr, Chem. Phys. 199 (1995) 73.
[35] K.P. Sagarik, J. Mol. Struct. (Theochem) 465 (1999) 141.
[36] K.P. Sagarik, S. Chaiwongwattana, P. Sisot, Chem. Phys., 306 (2004) 1.
[37] K.P. Sagarik, B.M. Rode, Chem. Phys. 260 (2000) 159.
[38] K.P. Sagarik, P. Asawakun, Chem. Phys. 219 (1997) 173.
[39] K.P. Sagarik, V. Pongpituk, S. Chiyapongs, P. Sisot, Chem. Phys. 156 (1991) 439.
[40] S. Vishveshwara, J.A. Pople, J. Am. Chem. Soc. 99 (1977) 2422.
[41] H.A. Levy, R.B. Corey, J. Am. Chem. Soc. 63 (1941) 2095.
[42] H.J. Boehm, R. Ahlrichs, J. Chem. Phys. 77 (1982) 2028.
[43] C.M. Breneman, K.B. Wiberg, J. Comput. Chem. 11 (1990) 361.
[44] M.J. Frisch, G.W. Trucks, H.B.Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, V.G. Zakrzewski, J.A. Montgomery, Jr., R.E. Stratmann, J.C. Burant, S. Dapprich, J.M. Millam, A.D. Daniels, K.N. Kudin, M.C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G.A. Petersson, P.Y. Ayala, Q. Cui, K. Morokuma, P. Salvador, J.J. Dannenberg, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J. Cioslowski, J.V. Ortiz, A.G. Baboul, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, J.L. Andres, C. Gonzalez, M. Head-Gordon, E.S. Replogle, and J.A. Pople, GAUSSIAN 98 (Revision A.5), Gaussian, Inc., Pittsburgh PA, 2001.
[45] J. Voogd, J.L. Derissen, F.B. Van Duijneveldt, J. Am. Chem. Soc. 103 (1981) 7701.
[46] R. Destro, P. Roversi, M. Barzaghi, R.E. Marsh, J. Phys. Chem. A 104 (2000) 1047.
[47] H.B. Schlegel, J. Comput. Chem. 3 (1982) 214.
[48] P.A. Kollman, Chem. Rev. 93 (1993) 2395.
[49] M. Orozco, F.J. Luque, Chem. Rev. 100 (2000) 4187.
[50] C.J. Cramer, D.G. Truhlar, Chem. Rev. 99 (1999) 2161.

[51] A.R. Leach, Molecular Modelling, Pearson Education, England, 2001.
[52] P.B. Sigler, in: S.L. McKnight, K.R. Yamamoto (Eds.), The Molecular Mechanism of trp repression, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1992.
[53] Z. Otwinowski, R.W. Schevitz, R.G. Zhang, C.L. Lawson, A. Joachimiak, R.Q. Marmorstein, B.F. Luisi, B.P. Sigler, Nature 335 (1988) 321.
[54] K. Wüthrich, N.M.R. in, structural biology, World Scientific Series in 20th Century Chemistry 5 (1993) 647–655.
[55] R.M. Brunne, E. Liepinsh, G. Otting, K. Wüthrich, W.F. Van Gunsteren, J. Mol. Biol. 231 (1993) 1040.
[56] D.A. Case, D.A. Pearlman, J.W. Caldwell, T.E. Cheatham III, W.S. Ross, C.L. Simmerling, T.A. Darden, K.M. Merz, R.V. Stanton, A.L. Cheng, J.J. Vincent, M. Crowley, V. Tsui, R.J. Radmer, Y. Duan, J. Pitera, I. Massova, G.L. Seibel, U.C. Singh, P.K. Weiner and P.A. Kollman, AMBER 6, University of California, San Francisco, 1999.
[57] SURFER for Windows, V. 6.04, Golden Software, Inc., USA, 1997.
[58] I.P. Gerothanassis, R. Hunston, J. Lauterwein, Helv. Chim. Acta 65 (1982) 1774.