Molecular Dynamics Simulation Studies of the Interactions between Ionic Liquids and Amino Acids in Aqueous Solution

Luciana I. N Tomé[‡], *Miguel Jorge*[ϒ], *José R. B. Gomes*[‡] *and João A. P. Coutinho*[‡]*

[‡]CICECO, Departamento de Química, Universidade de Aveiro, 3810-193 Aveiro, Portugal

[ϒ] LSRE – Laboratory of Separation and Reaction Engineering – Associate Laboratory

LSRE/LCM, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, 4200-465 Porto, Portugal

*Corresponding author

Tel: +351-234-370200; Fax: +351-234-370084; E-mail address: jcoutinho@ua.pt

Abstract:

Although the understanding of the influence of ionic liquids (ILs) on the solubility behavior of biomolecules in aqueous solutions is relevant for the design and optimization of novel biotechnological processes, the underlying molecular-level mechanisms are not yet consensual or clearly elucidated. In order to contribute to the understanding of the molecular interactions established between amino acids and ILs in aqueous media, classical molecular dynamics (MD) simulations were performed for aqueous solutions of five amino acids with different structural characteristics (glycine, alanine, valine, isoleucine and glutamic acid) in the presence of 1-butyl-3-methylimidazolium bis(trifluoromethyl)sulfonyl imide. The results from MD simulations enable to relate the properties of the amino acids, namely their hydrophobicity, to the type and strength of their interactions with ILs in aqueous solutions, and provide an explanation for the direction and magnitude of the solubility phenomena observed in [IL+amino acid+water] systems by a mechanism governed by a balance between competitive interactions of the IL cation, IL anion and water with the amino acids.

Keywords: amino acids, ionic liquids, molecular interactions, molecular dynamics

Introduction

During the past few years, ionic liquids (ILs) have emerged as an exceptionally interesting alternative to common solvents in a wide range of processes with industrial and biotechnological relevance, and research on this field has grown to be of recognized importance and applicability. In this context, the replacement of organic solvents by ILs in the separation and purification of biomolecules such as amino acids, proteins, carbohydrates, lactic acid, antibiotics and alkaloids [1-8], their use as reaction media in biocatalysis [9,10], biosynthesis [11] and in the kinetic resolution of racemates [12-14], as well as their role in the stabilization and activity of enzymes [15-19] have become crucial and urgent subjects in chemical and biochemical research.

The implementation of IL-based media in biotechnology has enabled to overcome environmental, operational and efficiency problems associated with the conventional application of liquid-liquid extraction techniques [20], and has shown great potential in other related central issues in the domains of biocatalysis [9-14,16-18]. Actually, due to their remarkable and advantageous properties, in conjunction with the possibility of adjusting their parameters through the adequate manipulation of their constituting ions [21,22], ILs provide unique environments for (bio)chemical processes to take place. For this reason, there are currently strong demands for exploring the potentialities of ILs in order to design task-specific solvents with chemical and physical versatility to substantially improve the success of bioseparations and enhance the stability and activity of enzymes and the yields of biocatalyzed reactions in environmentally-friendly media.

For the optimization and control of biotechnological processes through the design of adequate ILs, a detailed knowledge of the factors that influence the solvation of biocompounds in ILs and aqueous phases, as well as of their underlying molecular level mechanisms, is essential. In this respect, solubility data and other results for aqueous solutions of ILs and salts, sugars, fermentation metabolites, amino acids, enzymes and proteins have been precious sources of information and have enabled the establishment of correlations between the structure of the solvent and biomolecules, the behavior of these systems and the success of the processes [16-18,20,23-26]. Recently, additional information from molecular simulation has become widespread and proved to be extremely useful [27-32]. Despite all the efforts, however, the basic knowledge indispensible to achieve a comprehensive picture of the molecular level phenomena occurring in aqueous solutions of ILs and biomolecules is still lacking, and thus further investigation on this subject is required.

In previous studies, thermodynamic and spectroscopic data were used by us to develop a well-supported mechanism to interpret the solubility behavior of (IL+salts+water) [24,25] and (IL+amino acids+water) mixtures [26]. In the present work, molecular dynamics (MD) simulation methods are used to study the interactions between ILs and amino acids in aqueous media and to evaluate their dependence on the structural characteristics and properties of the species involved. With that aim, MD simulations were performed for aqueous solutions of an imidazolium-based IL (1-butyl-3-methylimidazolium bis(trifluoromethyl)sulfonyl imide, [C4mim][NTF_2], represented in Figure 1) with five different amino acids - glycine (Gly), alanine (Ala), valine (Val), isoleucine (ILe) and glutamic acid (Glu), all depicted in Figure 2. The imidazolium ion has been the most appealing and most frequently considered cation in the IL field and is known for its non-denaturing nature, low toxicity and reduced water pollution risk, when comprising a short alkyl chain [33,34], and [NTF_2] has been selected by IUPAC as a benchmark anion. [C4mim][NTF_2] is commonly used as a model ionic liquid while amino acids can be regarded as model biomolecules. With the choice of these specific amino acids we intend to cover a wide range of different hydrophobicities and polarities, and span the solubility behaviors observed experimentally – from salting-in to salting-out effects on the IL aqueous solubility [26]. Moreover, because the available experimental data [26] for all the amino acids except Glu refers to pH=7, only the zwitterionic forms were selected for the simulations. Unlike the other biomolecules under study, glutamic acid possesses an acidic character [23,35], and therefore the simulations of Glu were performed considering its neutral zwitterionic form, in order to reproduce the conditions (pH = 3) at which the reference experimental results [26] were obtained.

MD simulation methods have proved to be a valuable tool for the study of the molecular interactions in biochemical systems, including aqueous solutions and aqueous saline solutions of amino acids, peptides, proteins, lipid bilayers and hydrophobic solutes [28-32,36-43], and have lately been playing an important role in the understanding of the fundamental chemistry of ILs and of their mixtures with water, enzymes and hydrocarbons [27,44,45]. We have previously applied with success MD calculations to characterize the interactions established in aqueous saline solutions of amino acids [28] and in aqueous solutions of imidazolium-based ILs in the presence of salts [46], but we are not aware of any theoretical investigation on the interactions between ILs and amino acids in aqueous solution. In this work, the analysis of the radial distribution functions (RDFs) of the various groups and moieties, estimated by MD, will provide an explanation for the solubility

behavior experimentally observed for aqueous solutions of imidazolium-based IL in the presence of amino acids [26].

It is worth to notice that the choice of the force field to be employed in MD simulations is a crucial aspect since it often has repercussions on the accuracy of the results obtained [47-52]. For ILs in particular, the formulation of force fields capable of representing their energetics and structure has been a non-trivial and challenging subject mainly due, on the one hand, to the fact that these solvents are neither simple molecular fluids nor common (high-molten) salts but exhibit complex interactions between their constituting ions, and, on the other hand, to the paucity of experimental data published for the almost infinite number of possible ILs [50-52]. Despite such problems, significant progress has been made in the application of MD methods to IL systems, and force field parameters have already been developed and validated for a large number of cations and anions [50-52], yielding reliable results. The force fields selected in this work for the cation [53] and anion [54] have been successfully parameterized and tested, providing accurate descriptions of ILs and their mixtures [46,55,56].

Computational methods

MD calculations were performed for aqueous solutions of [C4mim][NTf_2] at a concentration of approximately 0.25 mol dm^{-3} in the presence of five amino acids in their zwitterionic forms (pH=3 for Glu and pH=7 for the rest of the amino acids considered). The simulations were carried out using the isothermal-isobaric *NpT* (T = 298.15 K and p = 1 bar) ensemble and the GROMACS 4.07 molecular dynamics package [57]. The equations of motion were integrated with the Verlet-Leapfrog algorithm [58] and a time step of 2 fs. The Nosé-Hoover thermostat [59,60] was used to fix the temperature while the Parrinello-Rahman barostat [61] was employed to fix the pressure. Starting configurations were generated in cubic boxes with lateral dimensions of 45 Å, and periodic boundary conditions were applied in three dimensions. The systems were prepared by randomly placing all species in the simulation box. Four [C4mim][NTf_2] ion pairs and 900 water molecules were incorporated in each box. Nine amino-acid molecules were also included to obtain a concentration of about 0.55 M. A 10 000 step energy minimization was performed followed by two simulations, the first one with 50 000 steps for equilibration and the final one with 5 000 000 steps (10 ns) for production. After

equilibration, the values of the box volume ranged between 29.3 and 30.2 nm^3, depending on the amino acid considered. Equilibration was checked by ensuring that all observables (including the RDFs) fluctuated around their equilibrium values during the production state.

The intermolecular interaction energy between pairs of neighboring atoms was calculated using the Lennard-Jones potential to describe dispersion/repulsion forces and the point-charge Coulomb potential was used for electrostatic interactions. Long-range electrostatic interactions were accounted for using the particle-mesh Ewald method [62], with a cutoff of 1.0 nm for the real-space part of the interactions. A cutoff radius of 1.2 nm was used for the Lennard-Jones potential, and long-range dispersion corrections were added to both energy and pressure. All bond lengths were held rigid using the LINCS constraint algorithm [63], while angle bending was modeled by a harmonic potential and dihedral torsion was described (where appropriate) by a Ryckaert-Bellemans function. Potentials available in the literature were taken for all the species considered in the simulations. Water was described by the rigid SPC/E model [64], while the OPLS all-atom potential was used for the amino acids [65]. All-atom force field parameters for [C4mim][NTf_2] were obtained from the works of Cadena and Maginn [53] for the cation and of Canongia Lopes and Pádua [54] for the anion. The charges for the cation and the anion were recalculated in this work with the CHelpG scheme using an optimized geometry for the C4mim-NTf_2 dimer in the gas-phase. The calculations considered the B3LYP/6-311+G(d) approach as included in the Gaussian 03 code [66], i.e., using the same computational strategy employed by Morrow and Maginn for the [C4mim][PF_6] ionic liquid [67]. The total charges on the cation and anion are +0.797 a.u. and -0.797 a.u., respectively. The estimation of partial charges for an IL from calculations of an ion pair in vacuum can be a problematic treatment and this issue has been addressed and discussed in other works [55,68]. Nevertheless, it has been demonstrated that models with total charges on each ion in the range ±0.7 to 0.8 yield a better description of both structural and (most noticeably) dynamic properties of ionic liquids [68-70]. Furthermore, we have also performed calculations considering a pair of ions surrounded by other IL ions and no significant differences were found in the results obtained. The full set of atomic charges is supplied as Supporting Information (Table S1).

Coordination numbers (CN) were calculated for the interactions between selected atoms. For that purpose, the function $N(r)$ was obtained by integrating the corresponding RDFs ($g(r)$):

$$N(r) = 4\pi\rho_B \int_0^r \left(r^2 g(r)\right) dr \quad (1)$$

where ρ_B is the number density of each atom in the bulk.

Results and Discussion

In recent investigations carried out by us [26], it was observed that the interactions between amino acids and an imidazolium-based IL in aqueous solution are essentially dependent on the nature and properties of the amino acid side chain, namely its size, polarity and charge, and can be understood in terms of a delicate balance between direct or indirect (water-mediated) interactions established between the amino acids and the ionic liquid. In this work, MD calculations are used to study the molecular interactions that occur in IL+amino acids+water ternary systems and their dependence on the hydrophobicity of the biomolecules, in order to obtain further insight into the molecular mechanism controlling the solubility behavior of these mixtures.

The experimental data available [26] for this sort of systems were obtained for 1-butyl-3-methylimidazolium tricyanomethane, [C4mim][$C(CN)_3$], at T = 298.15 K, pH = 3 for Glu and pH = 7 for the other amino acids. While polar amino acids, such as Gly and Ala, induce salting-out effects, hydrophobic molecules, such as Val and Ile, increase slightly the IL/water mutual solubilities. The observed trend is Gly > Ala > Glu > Val > Ile, from salting-out to salting-in, with Glu having a negligible influence on the solubility. These data will be used as a reference in the discussion below.

To infer about the molecular origin of the interactions between the biomolecules and the ILs in aqueous solution, the RDFs calculated from the MD simulations of [C4mim][NTf_2] in aqueous solutions of a series of amino acids with different polarities were considered. These RDFs provide a quantitative description of enhancement (values larger than 1) or depletion (values smaller than 1) of densities of species around a selected moiety, and are presented in Figures 3 to 13. The positions and intensities of the RDF peak maxima are given in tables S2 to S7 (*cf.* Supp. Inf.). Table 1 displays the literature values for the solubility [35] of the amino acids studied and molar Gibbs energy of hydration [71] of the amino acids' side chains, which will be used to rank their hydrophobic/hydrophilic character.

Solute-Water Interactions. We first analyze the RDFs of water around the amino acid molecules and the IL ions. The RDFs displayed in Figure 3 show that Gly, Ala, Val and Ile have very similar interactions with the water molecules. Due to the absence of an alkyl side chain, Gly

has slightly more pronounced peaks for the first and second solvation shells and, consequently, is the most hydrated of the amino acids studied. This observation is confirmed by the analysis of the coordination numbers (CN) of the water hydrogen and oxygen atoms around the charged groups of the amino acids, presented in Table 2, which are larger for Gly than for the other aminoacids. As shown by the solubility values summarized in Table 1, the water affinity of the amino acids decreases in the order Gly > Ala > Val > Ile, reflecting an increasing hydrophobicity as their side chains become longer.

As shown in Figure 4, no significant differences are observed as well in the water distribution around the terminal carbon atoms of the cation and anion of the IL due to the presence of the amino acids. Only a small, yet noticeable, decrease in the intensity of the RDF peak on going from Gly and Ala to Val, Ile and Glu is observed around the Ct_b and C atoms of the IL cation and anion. The same trend is observed for the water distribution around other moieties of the IL (Figure S1, *cf* Supp. Inf.). This molecular information is quantitatively corroborated by the values of the CN calculated for the interactions of the terminal carbon atoms of the IL cation and the C atom of the anion with water (Table 2) which are higher for Gly and Ala than for Val and Ile. The evidence obtained for the interaction patterns of the IL ions and of the amino acids with water does not actually show significant differences in the hydration of the species among the different systems studied. However, in the presence of hydrophobic biomolecules, the IL is less available to interact with water molecules because it is being co-solvated by the amino acid. In other words, because hydrophobic amino acids are able to bind to the IL cation and anion (since that association is more favorable than that with water molecules), some water molecules are excluded from the vicinity of the IL, resulting in a weaker hydration. On the contrary, hydrophilic biomolecules, which are not capable to interact with the IL, will leave the latest more available to be solvated by water and will have themselves a stronger hydration.

Further evidence can be obtained from the analysis of the RDFs represented in Figure 5, which show that the strength of the interactions between the terminal carbon atoms of the IL cations increase with the size of the non polar chains of the amino acids, indicating also a less effective hydration of the cation (and thus a more favorable $Ct_b \cdots Ct_b$ association) in systems containing more hydrophobic amino acids. Notice that the position of the first peak (approximately 0.4 nm) reflects a direct interaction between the cation chains, without other

mediating species. On the other hand, the intense peak observed in the second hydration shell in the presence of Ile, the most hydrophobic biomolecule studied, suggests that in this particular case the interaction between the terminal carbon atoms of the cation can be mediated by the amino acid, probably due to its binding to the IL cation alkyl chain. This is not observed for the other amino acids since Ile is the only amino acid whose alkyl chain is long enough to allow for such a pronounced interaction with the hydrophobic moieties of the IL cation. These conclusions are supported by the CN (Table 2) determined for the Ct_b(cation)···Ct_b(cation) interaction which are slightly larger for Val than for Gly and Ala and present the smallest value in the case of Ile. These results picture a system where there is no withdrawal of the water molecules by the most hydrophilic compound but instead the water solvation of the hydrophobic moieties is only affected by the presence of the more hydrophobic compounds that will act as co-solvents, as observed before [24-26,28]. This is in stark contradiction with the conventional models [72,73] which would explain the observed amino acid effects as resulting from a modification of the water structure around the solute, but is consistent with the most recent theories which underline the central role of ionic polarizabilities and of ion size in the interpretation of Hofmeister effects [74].

An additional information concerns the solvation state of the IL ions. According to experimental data for the equilibrium concentration of [C4mim][NTf2] in water provided in a recent study [75], a complete dissociation of the IL only occurs for very dilute solutions and above 0.1 M (which includes the ~0.25 M concentration considered in this work) it will exist majorly as ion pairs. This information is further supported by the results obtained in this work for the RDFs corresponding to the interaction between the IL cation and IL anion (Figure S2, *cf.* Supp. Inf.) which show a strong first peak and a less intense second peak for the contact pair C_3(cation)···O(anion), definitely excluding the possibility of the existence of a complete dissociation of the IL.

Ionic Liquid-Amino Acid Interactions. If the differences in IL solubility are not arising mainly from differences in the interactions between the solutes and water, the interactions between the amino acids and the IL ions are likely to play an important role. In order to identify the most important interactions between those species, we have analyzed a large number of RDFs for the IL cation and IL anion atoms around Val. In Figures 6, 7 and 9, we present the most relevant RDFs, which reveal the existence of interactions of the amino acid with both ions of the ionic

liquid. As shown in Figure 6 the IL anion exhibits an appreciable affinity for the non polar groups of the amino acid, as suggested by the intense peaks of the RDFs of the F atoms of the anion around the CG and CB moieties of Val, but is weakly associated to its charged parts. The exception is the particularly strong interaction established between the amino group of the biomolecule and the negatively charged oxygen of the anion, as indicated by the peaks corresponding to N(Val)···O(anion) and H(Val)···O(anion), which are quite intense and occur at short distances (the short distance of the latter peak identifies a direct amine-anion interaction, without intervening water molecules). The interactions of the IL anion with the other selected moieties of Val, are much weaker and occur at larger distances.

As far as the IL cation is concerned, the RDFs displayed in Figure 7 show some binding to the amino acid, but this only occurs between the Ct_b (and, to a smaller extent, the Ct_m) atoms of the former and the apolar groups of the latter (Figures 7 (a), 7 (b) and S3, *cf* Supp. Inf.). Indeed, the intensity of the peak corresponding to Ct_b···CG contact pairs is quite significant. The Ct_b···CB peak, located well beyond the Ct_b···CG peak, is merely a reflection of the latter interaction, which is indicative of the preference of the cation for the least polar moieties of the amino acid. Similar but less intense peaks are observed in the RDFs of Ct_m, revealing the occurrence of minor, yet still favorable, interactions. On the other hand, the binding of Val to the charged moieties of the cation is not favorable, as shown by the RDFs calculated for the interactions of C_3 around the different groups of Val presented in Figure 7 (c). A tridimensional picture of the interactions between the IL anion (O and F atoms) and the IL cation (Ct_b atoms) with the valine amino acid can be obtained from the spatial distribution functions (SDF) calculated for these atoms around valine which are depicted in Figure 8. The SDF for the O atoms of the IL anion are between the region comprehended by the amino group and the apolar groups of valine. The SDFs for the F atoms of the IL anion and for the Ct_b atoms of the IL cation are located near to the apolar regions of this amino acid (CG_x and CB atoms). Water molecules solvate the polar regions around the carboxylate group or between the latter and the amino group. As expected, it is clearly seen that the H atoms of water are closer to the COO^- group than the O atoms of water while opposite configuration is found for the regions around NH_3^+. The information retrieved from the SDFs is consistent with the results obtained from the RDFs discussed above.

It seems therefore that the most important interactions taking place in the systems containing the more hydrophobic species are those established between the non polar groups of the amino acids and the alkylic moities of the IL cation. This observation is further supported by the RDFs depicted in Figure 9, which reveal an enhanced binding of CG(Val) to the CB_x atoms that are further away from the cation charged moieties, also observed in the case of Ile (Figure S4 *cf* Supp. Inf.). In order to better visualize the molecular picture described above, snapshots from a simulation of Val and of Ile mixtures showing the relative positions and the distances between the amino acids' side chain and the carbon atoms of the cation's alkyl chain are displayed in Figure 10 and Figure S5 (*cf* Supp. Inf.), respectively.

Effect of Amino Acid Hydrophobicity. Having identified the most important interactions in the system and the RDFs that best characterize them, we proceed to compare their relative intensity for different amino acids (Figure 11). The results obtained for the more hydrophobic amino acid, Ile, are similar to those described for Val, but in the former case somewhat more significant interactions with the IL cation are observed. In fact, as can be seen from the RDFs displayed in Figure 11 (a), the peak referring to Ct_b···CD(Ile) contacts is more intense than that of Ct_b···CG(Val), while the interactions Ct_b···CG_2(Ile) and Ct_b···CG(Val) are similar. Accordingly, the values of the CN are ~ 0.10 for the Ct_b···CD(Ile) association, while for Ct_b···CG_2(Ile) and for Ct_b···CG(Val) they are ~ 0.13 (Table 2). Comparing to Val, however, Ile has a less remarkable association to both non polar and polar regions of the IL anion, as shown by the RDFs represented in Figure 11(c) and 10(d), respectively, and confirmed by the CN presented in Table 2 for the interactions of the amino acids with the F and O atoms of the anion, which are smaller for Ile than for Val. The cation also does not seem to establish any interactions with the charged or polar parts of the amino acids, as shown in Figure 11(b). These results are consistent with the SDFs calculated for Ile that are shown in Figure S6 (*cf*. Supp. Inf.).

When analyzing the more hydrophilic amino acids Gly and Ala, the peaks corresponding to the interactions between the terminal carbon atoms of the IL cation and the biomolecules become much reduced. Actually, the RDFs of Gly and Ala depicted in Figure 11(a) show that there is little interaction between the IL cation and the amino acids, although the depletion in Ct_b density around Ala's CB atom is less pronounced. Furthermore, the association of the amino acids with the charged groups of the cation does not occur either, as suggested by the data presented in Figure 11(b). Concerning the interactions with the IL anion, the RDFs displayed in

Figures 11(c) and 11(d) indicate that the binding of the amino group of the amino acids to the anion is very significant in the case of Val, but less important as far as Gly and Ile are concerned. On the other hand, the strength of the interactions between the terminal carbon atoms (Figure 2) of the biomolecules and the IL anion decreases in the order Val > Ile > Ala > Gly (Figure 11(c)), almost following the decrease of their hydrophobic character (Table 1). It is also worth noticing that while the interaction of Ile with the IL anion is mainly established at the level of the apolar moieties of both species, the interaction of Gly with the IL anion has, comparatively, a more electrostatic character, occurring between their charged groups (Figures 11(c), 11(d) and S7, *cf* Supp. Inf.). The concentration profiles of the IL cation and anion around the amino acids are consistent with the molecular interpretation derived from the structural data. While the CN calculated for the Ct(amino acid) $\cdots Ct_b$(cation) association is ~ 0.14 for Val and Ile, the values obtained for Gly and Ala were ~ 0.01 and 0.05, respectively. As far as the N(amino acid)$\cdots$O(anion) interaction is concerned, the largest CN is found in the case of Val and the number of F atoms of the IL anion surrounding each terminal carbon atom of the amino acids decreases in the order Val>Ala≈Ile>Gly (Table 2). The distribution of the different species around Gly and Ala, shown in Figure S6 (*cf*. Supp. Inf.), provide further support for these results. It seems therefore that, contrary to hydrophobic species, amino acids with a more polar character do not interact with the IL cation and are less effectively bound to the hydrophobic moieties of the IL anion, showing thus a clear preference for polar environments. This behavior is analogous to that observed in ternary salt+IL+water systems, where the more hydrophilic salts were strongly hydrated and interacted weakly with the IL while the more hydrophobic salts interacted strongly with the apolar moieties of the IL [24,25,46].

Other Structural Effects. To further evaluate the influence of the structure of the amino acids on their interaction with ILs in aqueous solution, another structural aspect, the presence of a polar terminal group on the amino acid side chain, was considered. The characteristic RDFs calculated from the MD simulations of Glu that are also displayed in Figure 11 show the existence of a strong association of both the IL cation and anion with the amino acid's hydrophobic regions (Figures 11(a) and 11(c), respectively), a weak binding of the IL cation to the charged COO^- group (Figure 11(b)), and a relatively strong N(Glu)$\cdots$O(anion) interaction (Figure 11(d)). On the other hand, Glu's charged groups are strongly hydrated, as indicated by the intense peaks of the RDFs shown in Figure 3 and by the CN calculated for the interactions

with water molecules (Table 2). Perhaps more importantly, the presence of a carboxyl group at the end of the alkyl chain of Glu introduces new interactions relative to the other, simpler, amino acids. The RDFs corresponding to these new interactions are depicted in Figures 12 and 13. The RDFs involving the F and O atoms of the anion (Figure 12(a)) indicate a large presence of the IL anion around the COOH group of Glu which, together with the strong interactions with CG (Figure 11(c)), imply very strong interactions between the IL anion and the uncharged chain of the amino acid. Similarly, the clear and intense peaks of the RDFs depicted in Figure 12(b) suggest a strong association of the COOH group of Glu with both polar and non polar moieties of the IL cation. Quantitative evidence for this qualitative insight provided by the RDFs can be obtained from the analysis of the CN calculated for the interactions of the atoms of the carboxyl group of Glu with the IL cation and anion presented in Table 3, and further support is obtained from the results depicted in Figure S6 (*cf.*, Supp. Inf.) with the SDFs calculated for various groups around Glu.

In face of the RDFs of Figures 11 and 12, and of the arguments presented above to explain the solubility effects of the other amino acids, Glu would be expected to behave as a salting-in inducing ion, because of its clear binding to the IL. That is not the case, however. In fact, despite possessing two methylene groups in its side chain, Glu also comprises a polar functional group (COOH), which confers polar character and hydrophilicity to the amino acid side chain, allowing it to be highly solvated by water molecules. Actually, the RDFs represented in Figure 13 show that the ordering of the water oxygen atom around the HE_2 atom of Glu and of the hydrogen atom of water around the O atoms of Glu are strong and occur for very close distances, which is further confirmed by the values of the CN determined for the water molecules around the atoms of the COOH group of Glu (Table 3). Moreover, as observed in Figure 3, strong interactions are observed between water and the charged NH_3^+ and COO^- moieties of Glu. This evidence suggests that the interactions between Glu's side chain and the IL cation and anion, and those between the amino acid and water are balanced, resulting in the negligible effect upon the IL solubility observed experimentally [26]. The existence of a balance between competitive interactions is in fact an important aspect to take into account when dealing with the molecular interpretation of the behavior of these systems, and has been suggested before [26]. The solubility data reported previously [26] and the MD results discussed above suggest the existence of a relation between the properties of the amino acids, namely the size and polarity of their side

chain, and the predominance of either (amino acid-IL) interactions (hydrophobic amino acids) or (amino acid-water) interactions (hydrophilic amino acids) or even a balance between multiple viable associations, as supported by the MD results obtained in this work for Glu.

Effect of the IL anion. An open question, unanswered in previous works addressing the molecular mechanism behind the solubility effect of ions and amino acids in aqueous IL solutions, is the role of the IL anion and the interactions that take place therewith [25,26]. As the results obtained in the current work show, there are actually interactions not only with the IL cation but also with the anion. The anion, unlike the cation, establishes interactions with all the amino acids studied here. In general, the main characteristics of the IL anion-amino acid interactions, derived from the analysis of the corresponding RDFs, are: (i) a preferential binding of the IL anion to the non polar moieties of the hydrophobic amino acids (e.g. Val and Ile) and to the charged groups of the most hydrophilic biomolecules (e.g. Gly and Ala); (ii) a significant association of the anion to both the charged and non polar parts of amino acids with an intermediate hydrophobic character (e.g. Glu). Our results suggest that, along with the cation, the anion of the IL also plays an important role in the mechanism by which the amino acids influence the IL solubility, and, together, they will determine the direction and magnitude of the observed solubility phenomena. These arguments are consistent with the trend observed for amino acid effects on aqueous solutions of imidazolium-based ILs - Ile>Val>Ala>Gly, from salting-in to salting-out [26]. In fact, while Gly, the simplest amino acid, with no side chain, a highly hydrophilic character and no affinity to the apolar parts of the IL cation and anion, induces the most significant salting-out influence as a result of a preferential hydration, the polar but less hydrophilic Ala is, to some extent, able to establish interactions with the apolar groups of the IL anion and produces a less pronounced decrease of the IL solubility. On the other hand, Ile, with a long apolar chain, has a strong association with the hydrophobic moieties of the IL cation and anion, and therefore promotes a pronounced salting-in effect as a consequence of those direct IL-amino acid side chain interactions. Val, in turn, exhibits also a significant binding to both the anion and cation of the IL, and the balance between these interactions is responsible for a less remarkable salting-in influence. It seems therefore that it is the competition and compensation effect between the strength of the interactions of the amino acids with the IL cation and anion, and with water, that will determine the direction and magnitude of the solubility phenomena observed experimentally [26].

To conclude, the MD results reported in this work provide a view on the relation between the structural characteristics and properties of the amino acids and the nature of their interactions with ILs in aqueous solution, showing how differences in amino acid hydrophobicity (derived either from different sizes of the apolar side chain or the presence of a polar functional group) affect the type and strength of the interactions established with the ILs in aqueous mixtures. As shown by the MD data reported here, while amino acids with longer alkyl chains strongly bind to the hydrophobic moieties of the cation and anion of the IL, hydrophilic biomolecules have an unfavorable binding to those moieties, being preferentially hydrated. The strength of the interactions between the amino acids and the IL cation increases with the increase of the amino acid side chain, i.e., with the increase of the hydrophobic character, and the association to the apolar moieties of the IL anion is less significant for hydrophilic amino acids. The introduction of a terminal (polar) functional group on the amino acid alkyl chain enhances the number and intensity of the interactions established with the cation, the anion and water. These conclusions are further strengthened considering the following quantitative evidence obtained from the analysis of the interaction energies calculated from the simulations and compiled in Table 4: (i) the Lennard-Jones terms for the interaction of Gly and Ala with the IL cation and anion are smaller (less negative, implying less favorable interactions) than those of Val and Ile, following the trend Ile>Val>Ala>Gly in the case of the IL cation; (ii) the Coulomb term is ~ 0 for the interaction of Val and Ile with the IL cation and much smaller for the interactions with the cation than with the anion; (iii) globally, the interaction energies obtained for Glu are the most negative (most favorable) of all the amino acids; (iv) Gly and Glu present the most negative (amino acid-water) Coulomb terms.

The evidences gathered allow for an interpretation of amino acid effects on aqueous IL solubility according to which salting-in phenomena are the consequence of direct interactions between the alkyl side chains of hydrophobic amino acids and the apolar moieties of the IL cation and anion, and salting-out inducing behavior results from a preferential hydration and an unfavorable binding of hydrophilic biomolecules to the IL. The balance between those competitive interactions is ultimately the factor responsible for the direction and magnitude of the solubility effect.

These conclusions support and provide a refined version of the empirical model proposed earlier to explain the influence of amino acids on the liquid-liquid equilibria of amino

acid+IL+water ternary systems [26]. Differences in amino acid hydrophobicities were also shown to be responsible for their ability to induce salting-out of imidazolium-based ILs from aqueous solutions, leading to the formation of aqueous biphasic systems (ABS). The ability of the amino acids to form ABS follows the order of their hydrophobicity [2]. Analogous MD studies concerning amino acids+salts+water solutions [28] have provided evidence for a mechanism based on the presence/absence of interactions between the ions and the non-polar moieties of the amino acids, instead of an indirect effect mediated by the water structure. Other experimental and theoretical works on aqueous saline solutions of amino acids, peptides, proteins and lipidic surfaces have provided similar evidence on interactions of the ions with the hydrophobic fragments of the biomolecules [30,76-78]. The results reported here for ionic liquids are in good agreement with these results and interpretations. Further consistent conclusions have been provided recently by MD studies of protein-ILs systems [32], according to which the strength of the interactions between the protein and the IL is primarily determined by Coulomb interactions with the IL anions and dispersion interactions with the IL cations, and is dependent on the ion size and magnitude of the ion surface charge – non polar regions of the protein surface were shown to interact with non polar domains of the IL, while charged regions were shown to attract the anions.

Conclusions

MD simulations were performed in order to investigate the interactions between ILs and amino acids in aqueous solutions and the dependence of their type and strength on the hydrophobicity of the biomolecules. The RDFs show clear evidence for the existence of important interactions of the non polar moieties of hydrophobic amino acids (Val and Ile) with the less polar groups of the IL cation, but no signs of interaction between the hydrophilic species (Gly and Ala) and this ion. Although all the amino acids studied exhibit in some way affinity to the IL anion, the most important IL anion-amino acid interactions occur mainly at the level of the apolar moieties of the anion and the side chains of the hydrophobic biomolecules. Strong interactions with both ions of the IL and with water are observed for Glu, as a consequence of the presence of a polar functional group on the alkyl chain of the amino acid (decrease of the hydrophobic character).

The data gathered allows for a molecular interpretation of the experimentally observed behavior of aqueous solutions of ILs with amino acids and provides evidence against classical interpretations which would relate the solubility effects to changes in water structure. Instead, a refined version of a mechanism earlier proposed is suggested. Our results support a mechanism of salting-in based on the direct binding of the non-polar moieties of hydrophobic amino acids to the hydrophobic groups of the IL cation and anion, a salting-out influence governed by a preferential hydration of hydrophilic amino acids, and the existence of a competition between multiple viable interactions (strictly dependent on the structure of the amino acid), whose balance dictates the direction and magnitude of the solubility phenomena.

Since amino acids can be taken as model systems, the molecular mechanism reported here can be helpful to understand the solubility and stability behavior of other more complex biomolecules, in particular drugs and proteins, in aqueous solutions with ILs and thus be relevant to develop and improve IL-based biotechnological processes.

Acknowledgment: The authors thank financial support from Fundação para a Ciência e a Tecnologia for Programa Ciência 2007 and the post-doctoral grant SFRH/BPD/44926/2008 awarded to Luciana I. N. Tomé. This work is partially supported by projects PEst-C/EQB/LA0020/2011 and PEst-C/CTM/LA0011/2011, financed by FEDER through COMPETE - Programa Operacional Factores de Competitividade and by FCT - Fundação para a Ciência e a Tecnologia.

References:

(1) Zhao, H.; Xia, S.; Ma, P., *J. Chem. Technol. Biotechnol.* **2005**, *80*, 1089.

(2) Domínguez-Pérez, M.; Tomé, L. I. N.; Freire, M. G.; Marrucho, I. M.; Cabeza, O.; Coutinho, J. A. P., *Sep. Purif. Technol.* **2010**, *72*, 85.

(3) Tomé, L. I. N.; Catambas, V. R.; Teles, A. R. R.; Freire, M. G.; Marrucho, I. M.; Coutinho, J. A. P., *Sep. Purif. Technol.* **2010**, *72*, 167.

(4) Freire, M. G.; Louros, C. L. S.; Rebelo, L. P. N.; Coutinho, J. A. P., *Green Chem.* **2011**, *13*, 1536.

(5) Neves, C. M. S. S.; Granjo, J. F. O.; Freire, M. G.; Robertson, A.; Oliveira, N. M. C.; Coutinho, J. A. P., *Green Chem.* **2011**, *13*, 1517.

(6) Cláudio, A. F. M.; Freire, M. G.; Freire, C. S. R.; Silvestre, A. J. D.; Coutinho, J. A. P., *Sep. Purif. Technol.* **2010**, *75*, 39.

(7) Freire, M. G.; Neves, C. M. S. S.; Marrucho, I. M.; Canongia Lopes, J. N.; Rebelo, L. P. N.; Coutinho, J. A. P., *Green Chem.* **2010**, *12*, 1715.
(8) Ventura, S. P. M.; Sousa, S. G.; Freire, M. G.; Serafim, L. S.; Lima, A. S.; Coutinho, J. A. P., *J. Chromatogr. B* **2011**, *879*, 2679.
(9) Kim, K.-W.; Song, B.; Choi, M.-Y.; Kim, M., *Org. Lett.* **2001**, *3*, 1507.
(10) Erbeldinger, M.; Mesiano, A. J.; Russel, A. J., *Biotechnol. Prog.* **2000**, *16*, 1129.
(11) Vallette, H.; Ferron, L.; Coquerel, G.; Guillen, F.; Plaquevent, J.-C., *ARKIVOC* **2006**, *iv*, 200.
(12) Malhotra, S. V.; Zhao, H., *Chirality* **2005**, *17*, S240.
(13) Zhao, H.; Luo, R. G.; Malhotra, S. V., *Biotechnol. Prog.* **2003**, *19*, 1016.
(14) Zhao, H.; Jackson, L.; Song, Z.; Olubajo, O., *Tetrahedron (Asymmetry)* **2006**, *17*, 2491.
(15) Zhao, H., *J. Mol. Catalysis B: Enzymatic* **2005**, *37*, 16.
(16) Zhao, H.; Olubajo, O.; Song, Z.; Sims, A. L.; Person, T. E.; Lawal, R. A.; Holley, L. A., *Bioorg. Chem.* **2006**, *34*, 15.
(17) Shan, H.; Li, Z.; Li, M.; Ren, G.; Fang, Y., *J. Chem. Technol. Biotechnol.* **2008**, *83*, 886.
(18) Féher, E.; Major, B.; Bélafi-Bakó, K.; Gubicza, L., *Biochem. Soc. Trans.* **2007**, *35*, 1624.
(19) Rantwijk, F.; Secundo, F.; Sheldon, R. A., *Green Chem.* **2006**, *8*, 282.
(20) Smirnova, V. S.; Torocheshnikova, I. I.; Formanovsky, A. A.; Pletnev, I. V., *Anal. Bioanal. Chem* **2004**, *378*, 1369.
(21) Wasserscheid, P.; Keim, W., *Angew. Chem. Int. Ed.* **2000**, *39*, 3722.
(22) Earle, M. J.; Seddon, K. R., *Pure Appl. Chem.* **2000**, *72*, 1391.
(23) Wang, J.; Pei, Y.; Zhao, Y.; Zhiguo, H., *Green Chem.* **2005**, *7*, 196.
(24) Freire, M. G.; Carvalho, P. J.; Silva, A. M. S.; Santos, L. M. N. B. F.; Rebelo, L. P. N.; Marrucho, I. M.; Coutinho, J. A. P., *J. Phys. Chem. B* **2009**, *113*, 202.
(25) Tomé, L. I. N.; Varanda, F. R.; Freire, M. G.; Marrucho, I. M.; Coutinho, J. A. P., *J. Phys. Chem. B* **2009**, *113*, 2815.
(26) Tomé, L. I. N.; Domínguez-Pérez, M.; Cláudio, A. F. M.; Freire, M. G.; Marrucho, I. M.; Cabeza, O.; Coutinho, J. A. P., *J. Phys. Chem. B* **2009**, *113*, 13971.
(27) Micaêlo, N. M.; Soares, C. M., *J. Phys. Chem. B* **2008**, *112*, 2566.
(28) Tomé, L. I. N.; Jorge, M.; Gomes, J. R. B.; Coutinho, J. A. P., *J. Phys. Chem. B* **2010**, *114*, 16450.
(29) Heyda, J.; Vincent, J. C.; Tobias, D. J.; Dzubiella, J.; Jungwirth, P., *J. Phys. Chem. B* **2010**, *114*, 1213.
(30) Lund, M.; Vrbka, L.; Jungwirth, P., *J. Am. Chem. Soc.* **2008**, *130*, 11582.
(31) Vrbka, L.; Jungwirth, P.; Bauduin, P.; Touraud, D.; Kunz, W., *J. Phys. Chem. B* **2006**, *110*, 7036.
(32) Klahn, M.; Lim, G. S.; Seduraman, A.; Wu, P., *Phys. Chem. Chem. Phys.* **2011**, *13*, 1649.
(33) Romero, A.; Santos, A.; Tojo, J.; Rodríguez, A. J., *Haz. Materials* **2008**, *151*, 268.
(34) Couling, D. J.; Bernot, R. J.; Docherty, K. M.; Dixon, J. K.; Maginn, E. J., *Green Chem.* **2006**, *8*, 82.
(35) *CRC Handbook of Chemistry and Physics*; CRC Press: Boca Raton, FL, 1982.
(36) Heyda, J.; Hrobárik, T.; Jungwirth, P., *J. Phys. Chem. A* **2009**, *113*, 1969.
(37) Sagarik, K.; Dokmaisrijan, S., *J. Mol. Struct. (THEOCHEM)* **2005**, *718*, 31.
(38) Fujita, T.; Watanabe, H.; Tanaka, S., *Chem. Phys. Lett.* **2007**, *434*, 42.
(39) Hess, B.; van der Vegt, N. F. A., *PNAS* **2009**, *106*, 13296.
(40) Vrbka, L.; Vondrásek, J.; Jagoda-Cwiklik, B.; Vácha, R.; Jungwirth, P., *PNAS* **2006**, *103*, 15440.
(41) Fedorov, M. V.; Goodman, J. M.; Schumm, S., *J. Am. Chem. Soc.* **2009**, *131*, 10854.
(42) Cordomi, A.; Edholm, O.; Perez, J. J., *J. Phys. Chem. B* **2008**, *112*, 1397.

(43) Heyda, J.; Lund, M.; Oncák, M.; Slavícek, P.; Jungwirth, P., *J. Phys. Chem. B* **2010**, *114*, 10843.
(44) Moreno, M.; Castiglione, F.; Mele, A.; Pasqui, C.; Raos, G., *J. Phys. Chem. B* **2008**, *112*, 7826.
(45) Harper, J. B.; Lynden-Bell, K. M., *Molec. Phys.* **2004**, *102*, 85.
(46) Freire, M. G.; Neves, C. M. S. S.; Silva, A. M. S.; Santos, L. M. N. B. F.; Marrucho, I. M.; Rebelo, L. P. N.; Shah, J. K.; Maginn, E. J.; Coutinho, J. A. P., *J. Phys. Chem. B* **2010**, *114*, 2004.
(47) Weerasinghe, S.; Smith, P. E., *J. Chem. Phys.* **2003**, *119*, 11342.
(48) Hess, B.; Holm, C.; van der Vegt, N. F. A., *J. Chem. Phys.* **2006**, *124*, 164509.
(49) Hess, B.; van der Vegt, N. F. A., *Proc. Natl. Acad. Sci. USA* **2010**, *106*, 13296.
(50) Hunt, P. A., *Mol. Simul.* **2006**, *32*, 1.
(51) Bhargava, B. L.; Balasubramanian, S.; Klein, M. L., *Chem. Commun.* **2008**, 3339.
(52) Maginn, E. J., *J. Phys.: Condens. Matter* **2009**, *21*, 373101.
(53) Cadena, C.; Maginn, E. J., *J. Phys. Chem. B* **2006**, *110*, 18026.
(54) Canongia Lopes, J. N.; Pádua, A. A. H., *J. Phys. Chem. B* **2004**, *108*, 16893.
(55) Logotheti, G.-E.; Ramos, J.; Economou, I. G., *J. Phys. Chem. B* **2009**, *113*, 7211.
(56) Lynden-Bell, K. M.; Del Pópolo, M. G.; Youngs, T. G. A.; Kohanoff, J.; Hanke, C. G.; Harper, J. B.; Pinilla, C. C., *Acc. Chem. Res.* **2007**, *40*, 1138.
(57) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E., *J. Chem. Theory Comput.* **2008**, *4*, 435.
(58) Hockney, R. W.; Goel, S. P. J., *J. Comput. Phys.* **1974**, *14*, 148.
(59) Nosé, S., *Mol. Phys.* **1984**, *52*, 255.
(60) Hoover, W. G., *Phys. Rev. A* **1985**, *31*, 1695.
(61) Parrinello, M.; Rahman, A., *J. Appl. Phys.* **1981**, *52*, 7182.
(62) Essman, U.; Perela, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pederson, L. G., *J. Chem. Phys.* **1995**, *103*, 8577.
(63) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M., *J. Comput. Chem.* **1997**, *18*, 1463.
(64) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P., *J. Phys. Chem.* **1997**, *91*, 6269.
(65) Aqvist, J., *J. Phys. Chem.* **1990**, *94*, 8021.
(66) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian 03, Revision C.01; Gaussian, Inc.: Wallingford CT, 2004.
(67) Morrow, T. I.; Maginn, E. J., *J. Phys. Chem. B* **2002**, *106*, 12807.
(68) Zhong, X.; Liu, Z.; Cao, D., *J. Phys. Chem. B* **2011**, *115*, 10027.
(69) Youngs, T. G. A.; Hardacre, C., **2008**, *9*, 1548.
(70) Wendler, K.; Dommert, F.; Zhao, Y. Y.; Berger, R.; Holm, C.; Delle Site, L., *Faraday Discuss.* **2011**, *154*, 1.

(71) Wolfenden, R.; Andersson, L.; Cullis, P. M.; Southgate, C. C. B., *Biochemistry* **1981**, *20*, 849.
(72) Washabaugh, M. W.; Collins, K. D., *J. Biol. Chem.* **1986**, *261*, 2477.
(73) Batchelor, J. D.; Olteanu, A.; Tripathy, A.; Pielak, G. J., *J. Am. Chem. Soc.* **2004**, *126*, 1958.
(74) Parsons, D. F.; Bostrom, M.; Lo Nostro, P.; Ninham, B. W., *Phys. Chem. Chem. Phys.* **2011**, *13*, 12352.
(75) Lee, S. H.; Lee, S. B., *J. Chem. Technol. Biotechnol.* **2009**, *84*, 202.
(76) Zangi, R.; Hagen, M.; Berne, B. J., *J. Am. Chem. Soc.* **2007**, *129*, 4678.
(77) Khoskbarchi, M. K.; Vera, J. H., *Ind. Eng. Chem. Res.* **1997**, *36*, 2445.
(78) Mason, P. E.; Neilson, G. W.; Dempsey, C. E.; Barnes, A. C.; Cruickshank, J. M., *PNAS* **2003**, *100*, 4557.

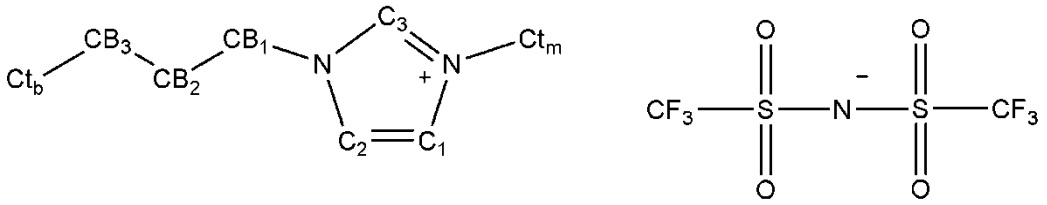


Figure 1. Structure of the ions constituting the IL studied in this work and corresponding atom labeling. Ct_b and Ct_m stand, respectively, for the terminal carbon atoms of the butyl and methyl side chains of the cation, while CB_x (x=1,2,3) is used to denote the other carbon atoms of the butyl chain. Hydrogen atoms of the cation are omitted for clarity.

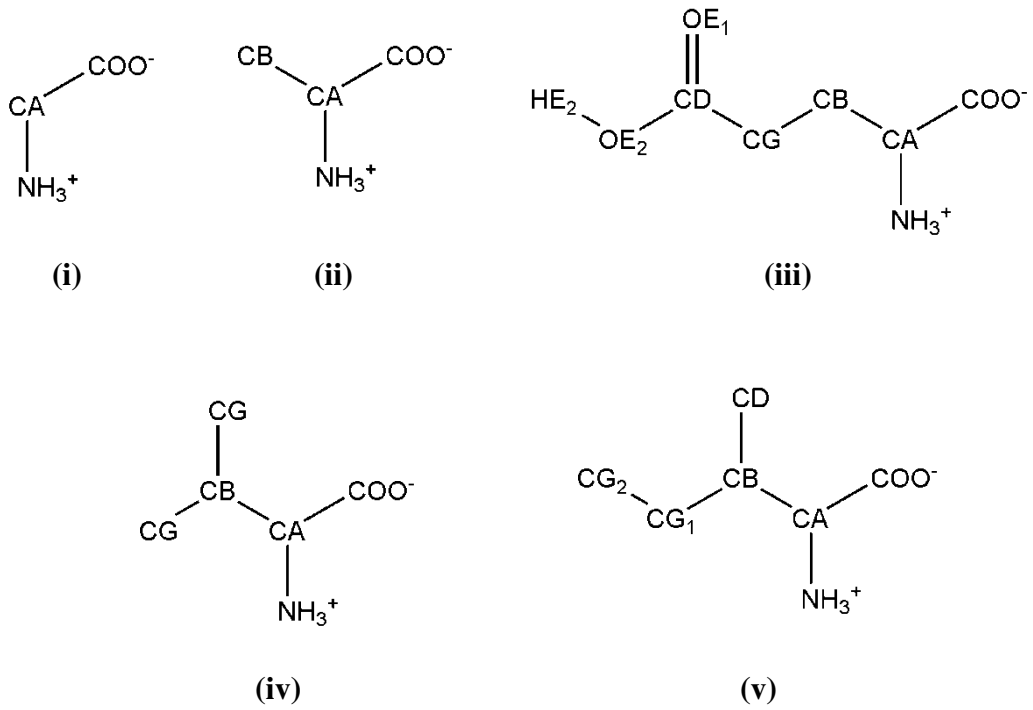
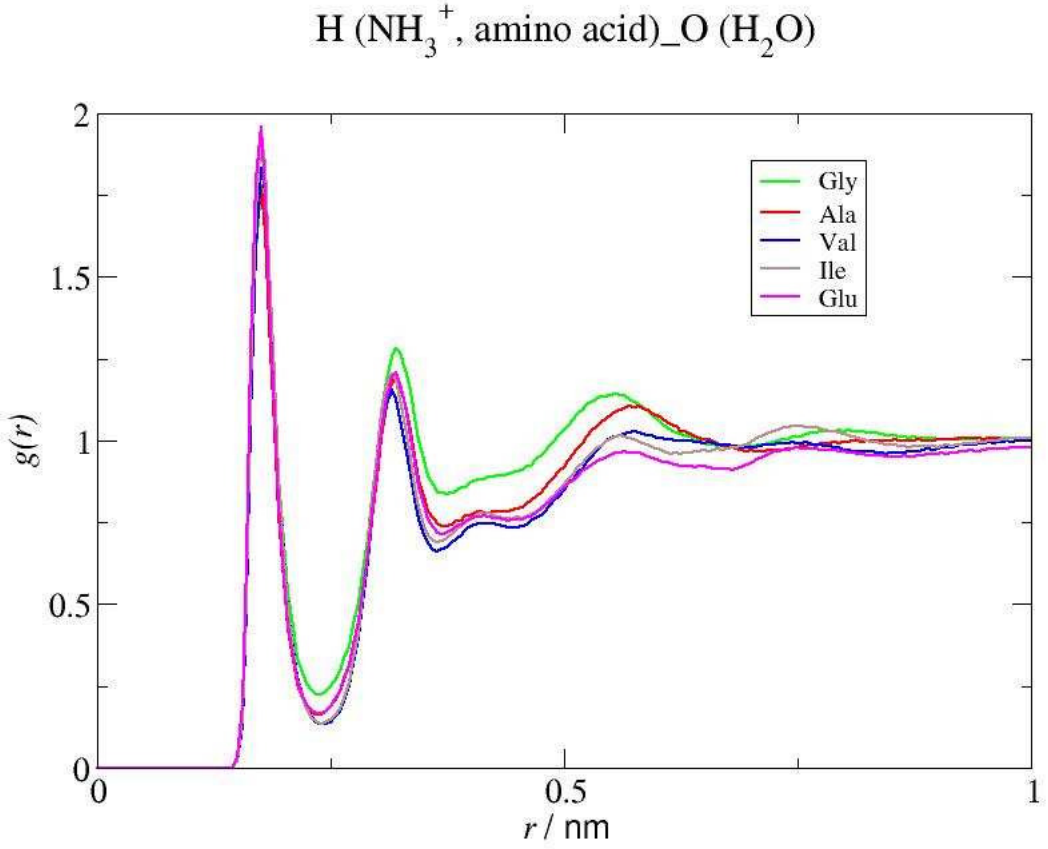


Figure 2. Structure and atom labeling of the amino acids studied in this work: (i) glycine (Gly); (ii) alanine (Ala); (iii) glutamic acid (Glu); (iv) valine (Val) and (v) isoleucine (Ile). $C_{terminal}$ in these amino acids correspond to CA (Gly), CB (Ala), CG (Val) and CG_2 (Ile). Hydrogen atoms of the alkyl groups are omitted for clarity.



(a)

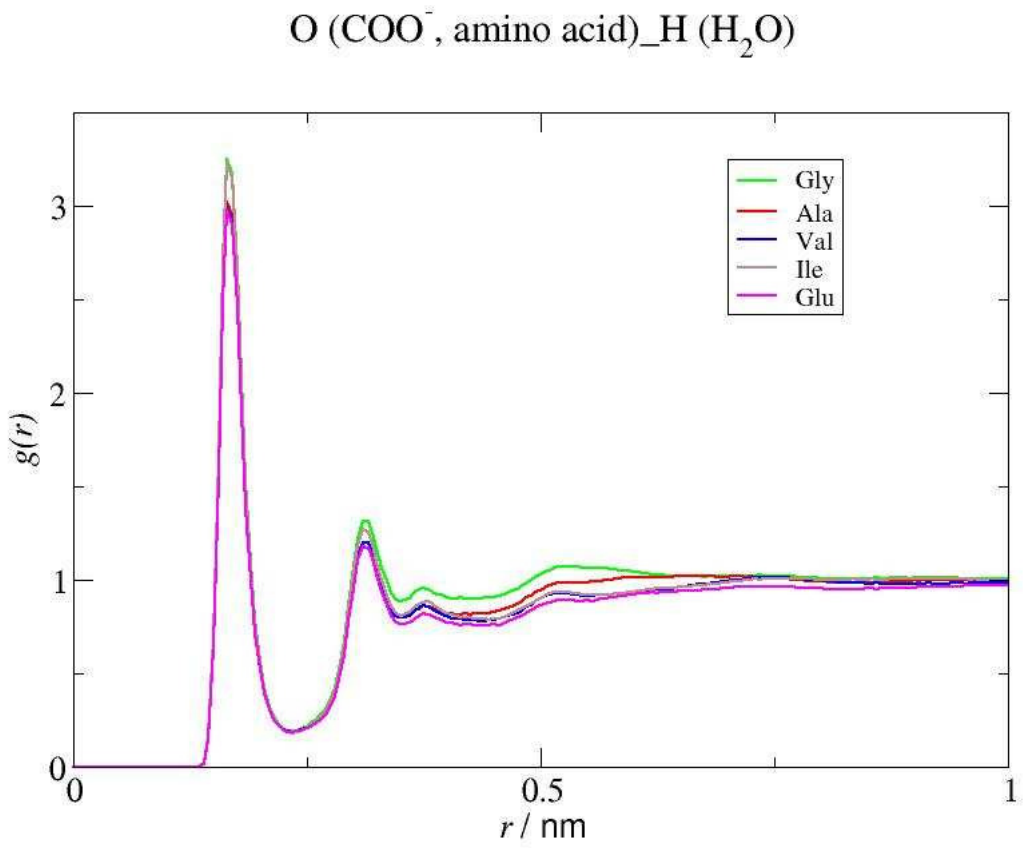
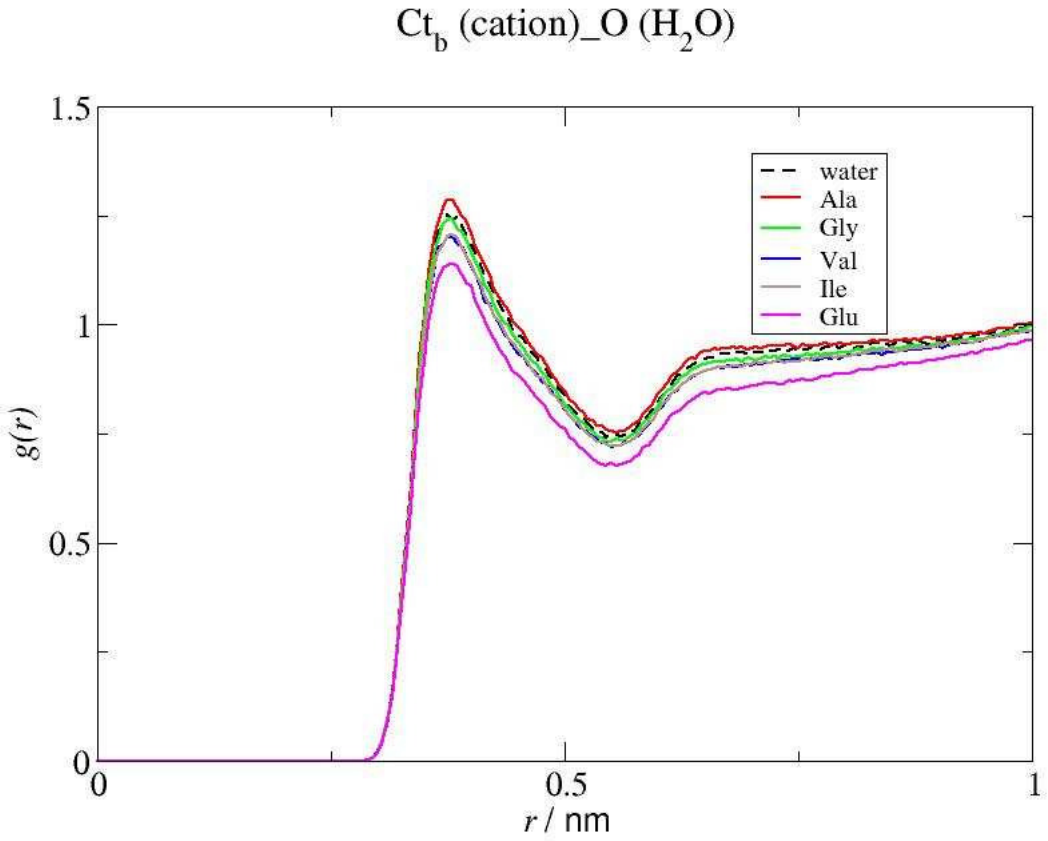


Figure 3. Radial distribution functions of (a) the O atom of water and the amino hydrogens of the amino acids, and (b) the H atom of water and the carboxylic oxygens of the amino acids.



(a)

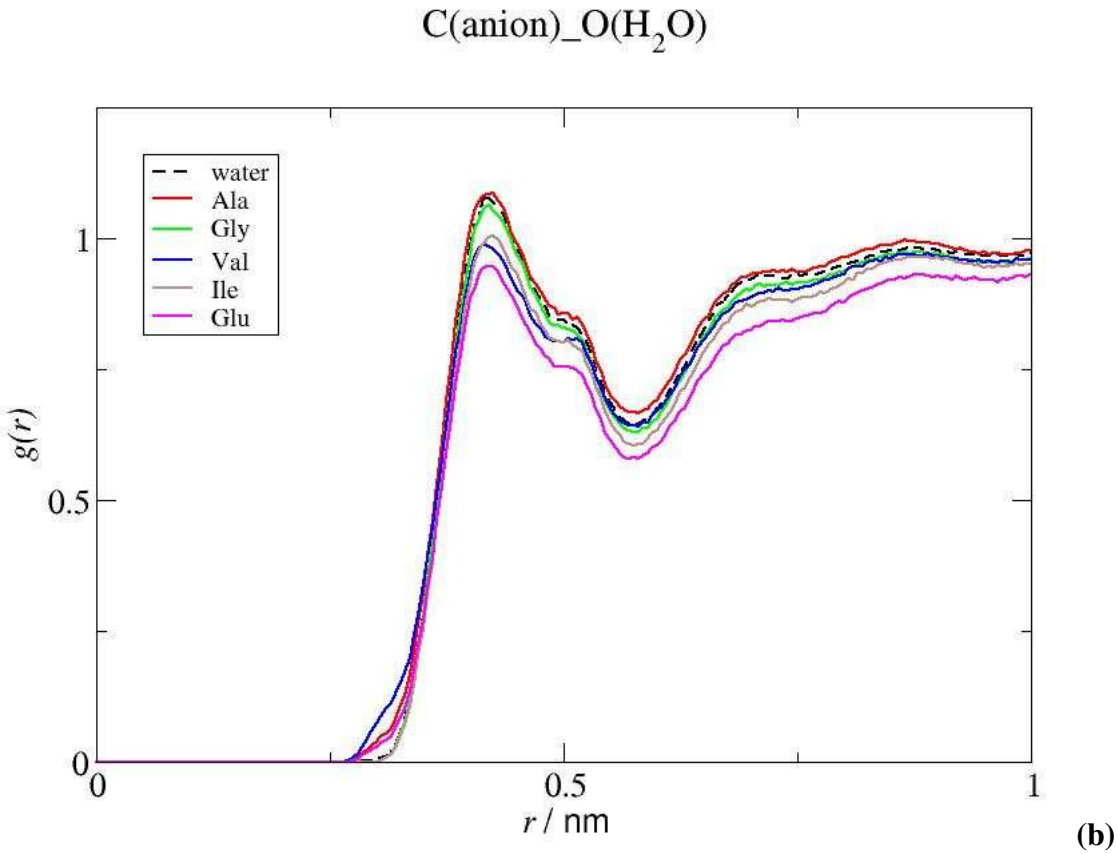


Figure 4. Radial distribution functions of the water oxygen atoms around (a) the terminal carbon atom (Ct_b) of the butyl chain of the cation and (b) the carbon atom (C) of the anion, in the IL+water binary system or in the IL+water+amino acid ternary systems.

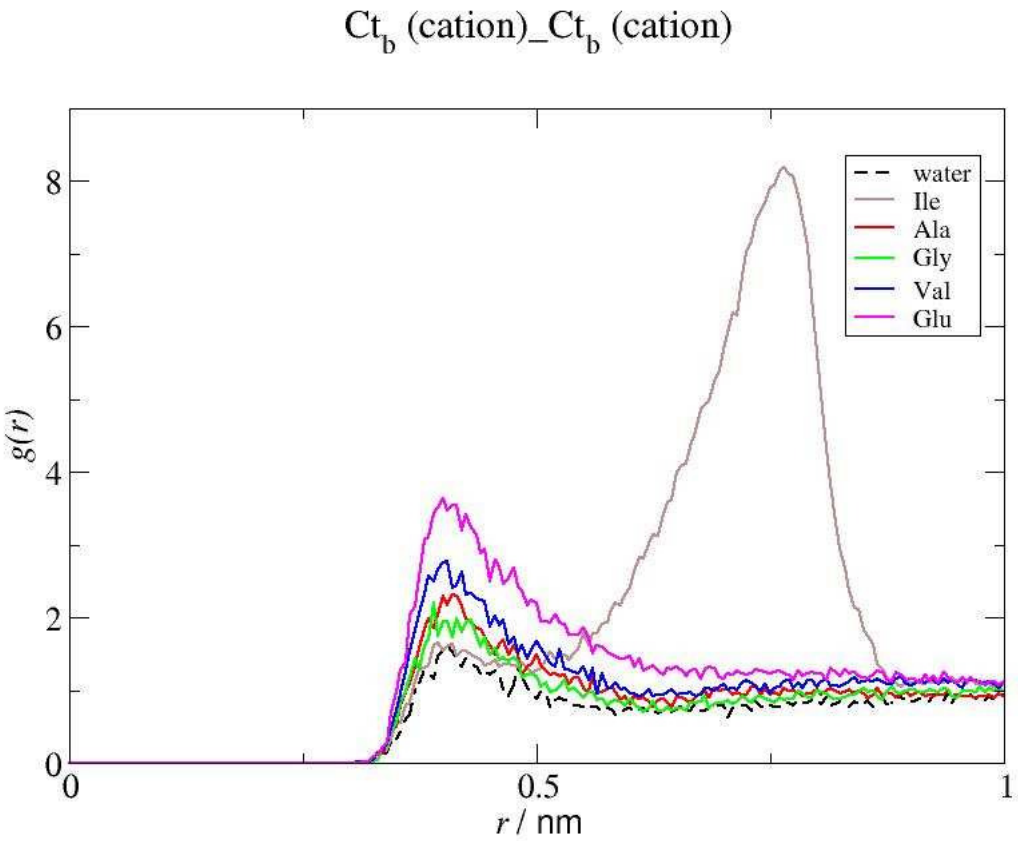
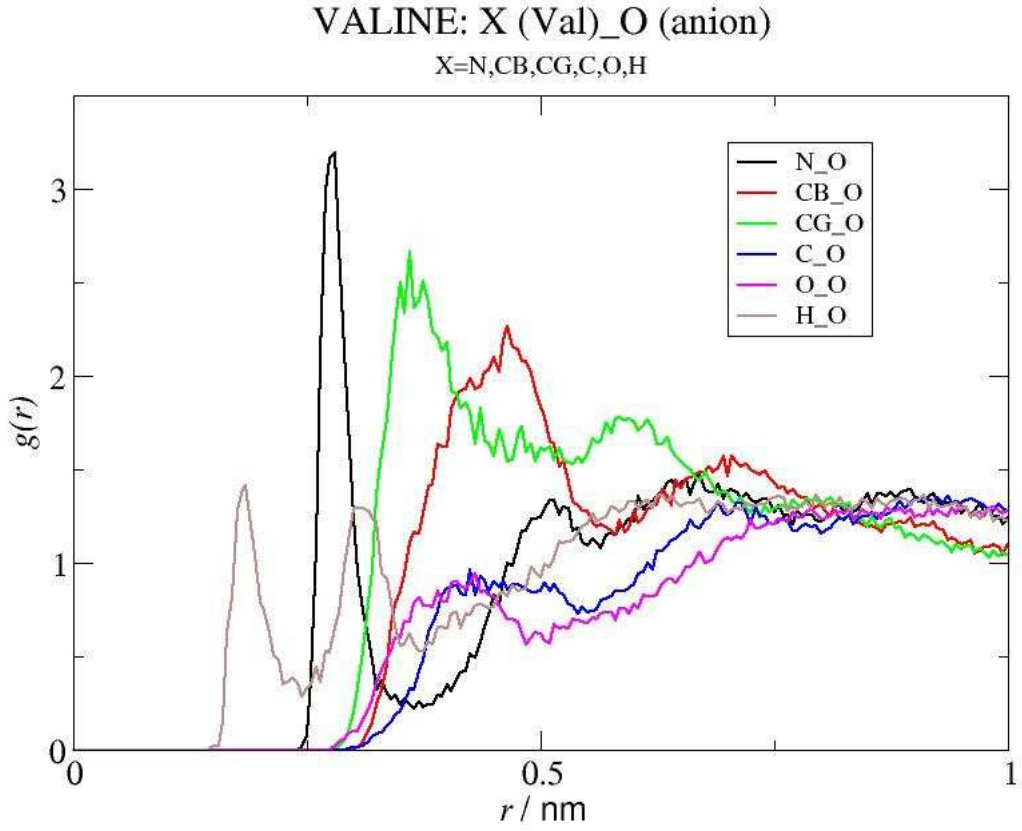
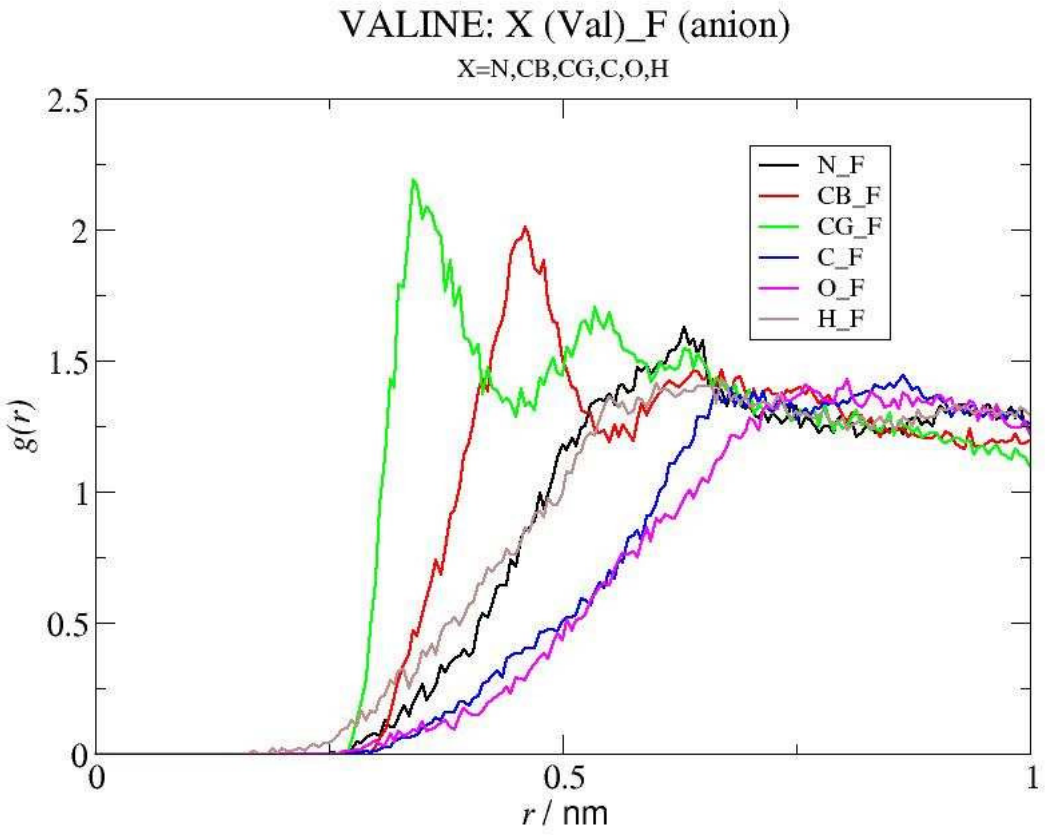


Figure 5. Radial distribution functions of the interactions between the terminal carbon atoms (Ct_b) of the butyl chain of the IL cations.

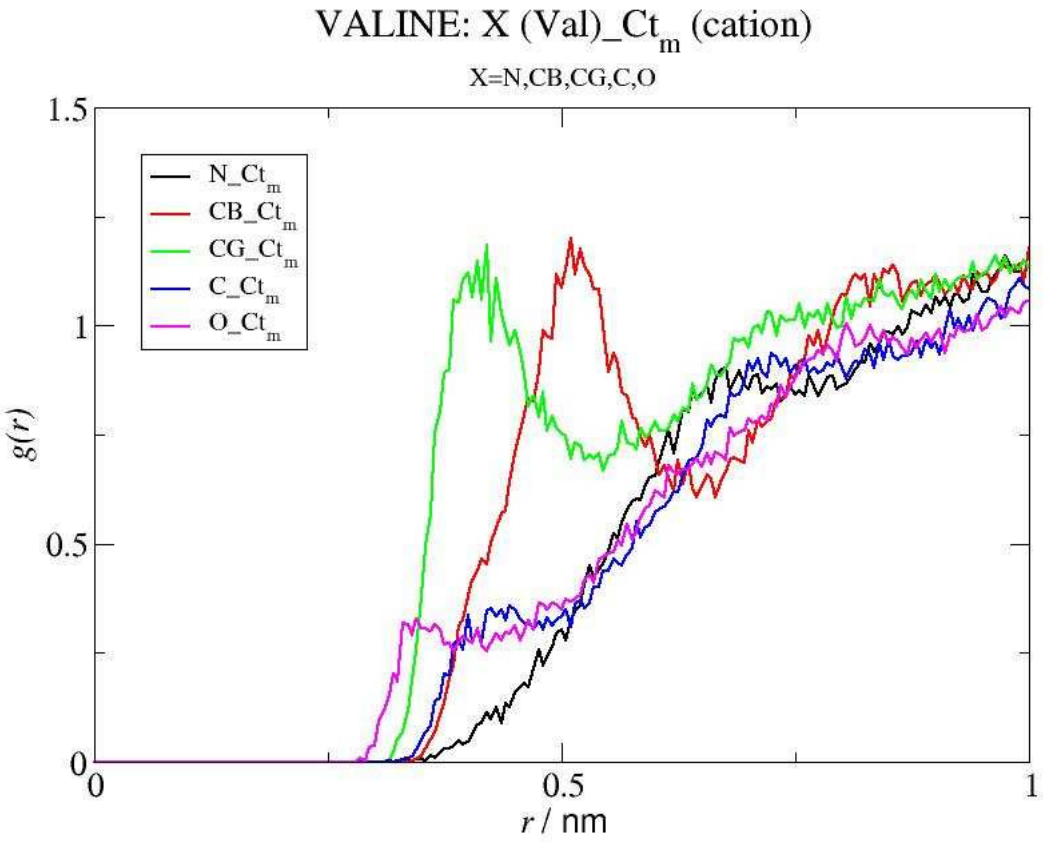

VALINE: X (Val)_O (anion)
X=N,CB,CG,C,O,H
N_O
CB_O
CG_O
C_O
O_O
H_O
g(r)
r / nm
0
0.5
1
2
3

(a)

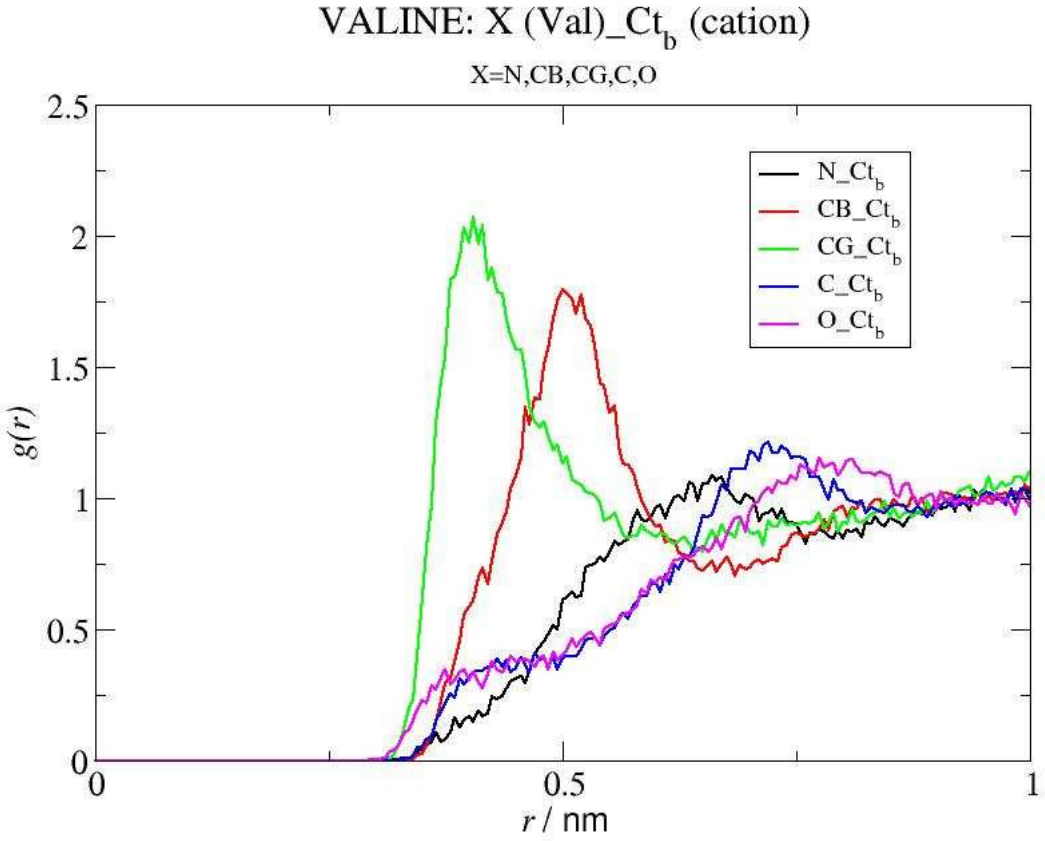


(b)

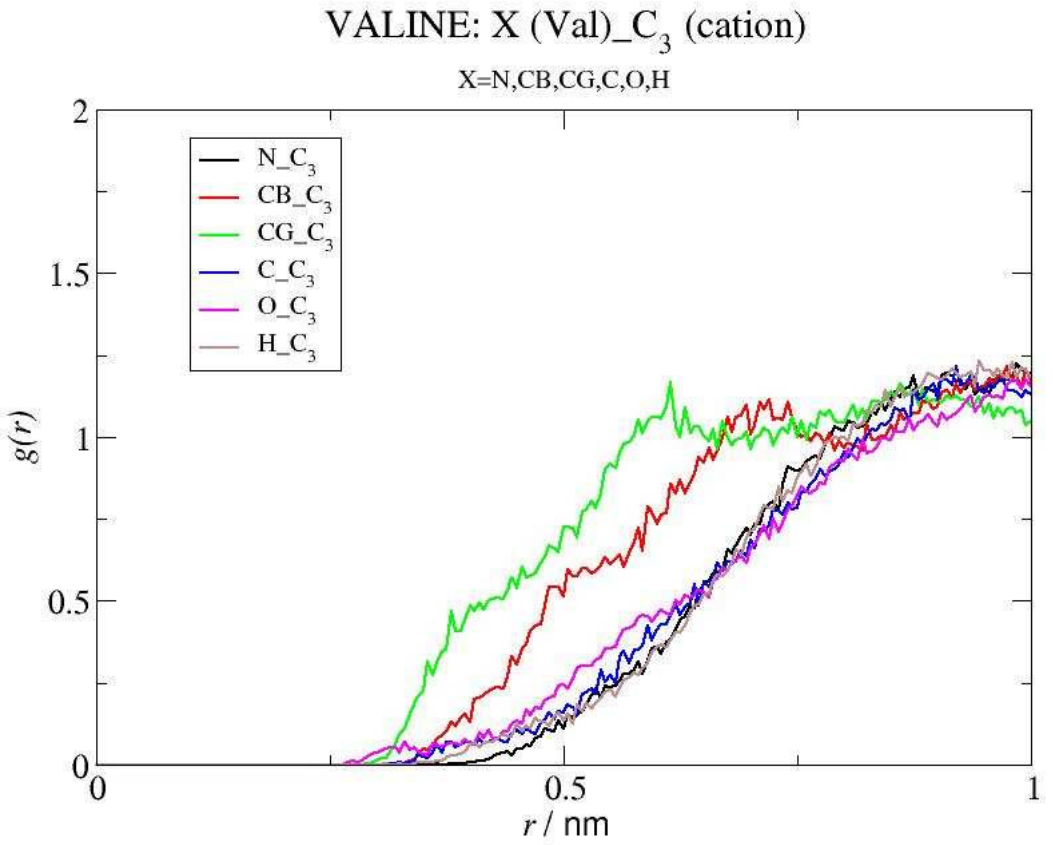
Figure 6. Radial distribution functions between different molecular regions of Val and selected atoms of the IL anion.


VALINE: X (Val)_Ct$_m$ (cation)
X=N,CB,CG,C,O
N_Ct$_m$
CB_Ct$_m$
CG_Ct$_m$
C_Ct$_m$
O_Ct$_m$
g(r)
r / nm
0
0.5
1
1.5

(a)


VALINE: X (Val)_Ct$_b$ (cation)
X=N,CB,CG,C,O
N_Ct$_b$
CB_Ct$_b$
CG_Ct$_b$
C_Ct$_b$
O_Ct$_b$
g(r)
r / nm
0
0.5
1
1.5
2
2.5
0
0.5
1

(b)



(c)

Figure 7. Radial distribution functions between different molecular regions of Val and selected atoms of the IL cation.

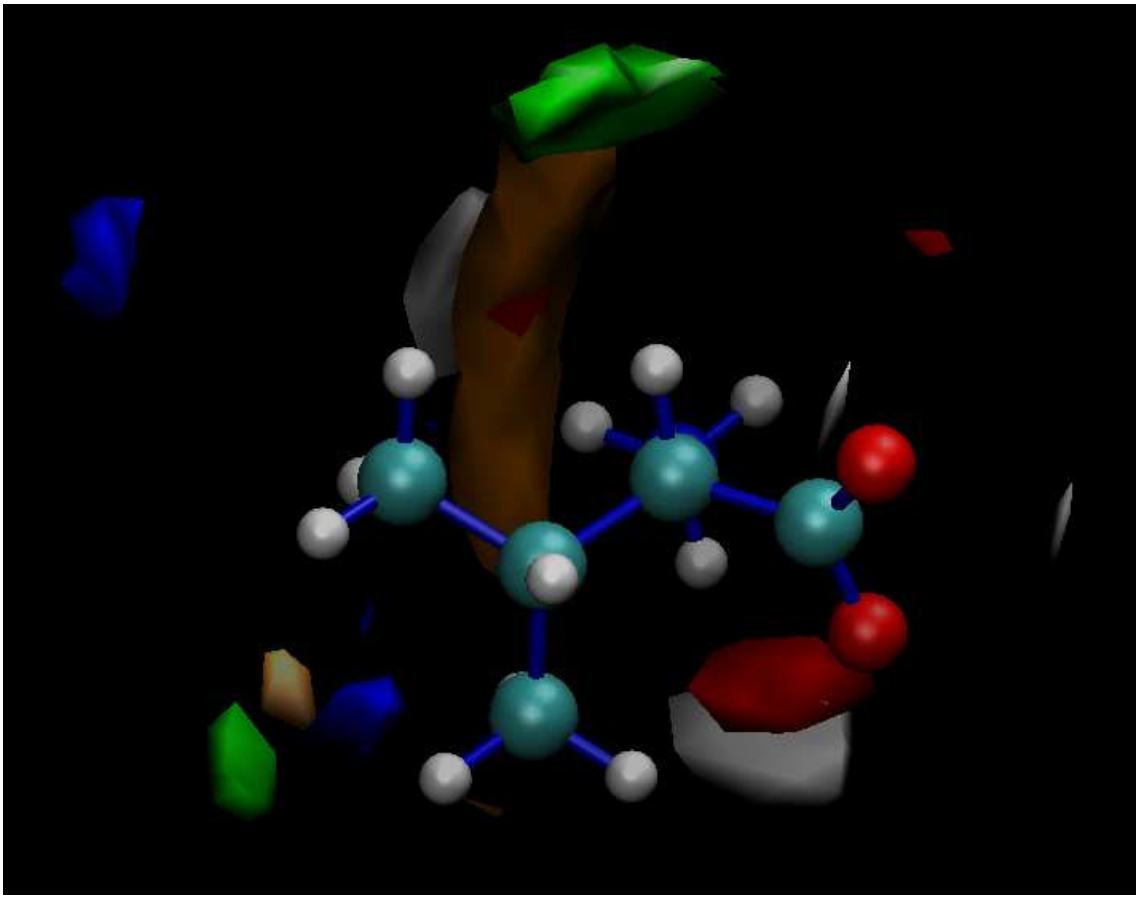


Figure 8. Spatial distribution functions (SDF) for different groups around valine. Orange: O atoms of the IL anion; blue: Ct_b atom of the IL cation; green: F atoms of the IL anion; red: O atoms of water; white: H atoms of water.

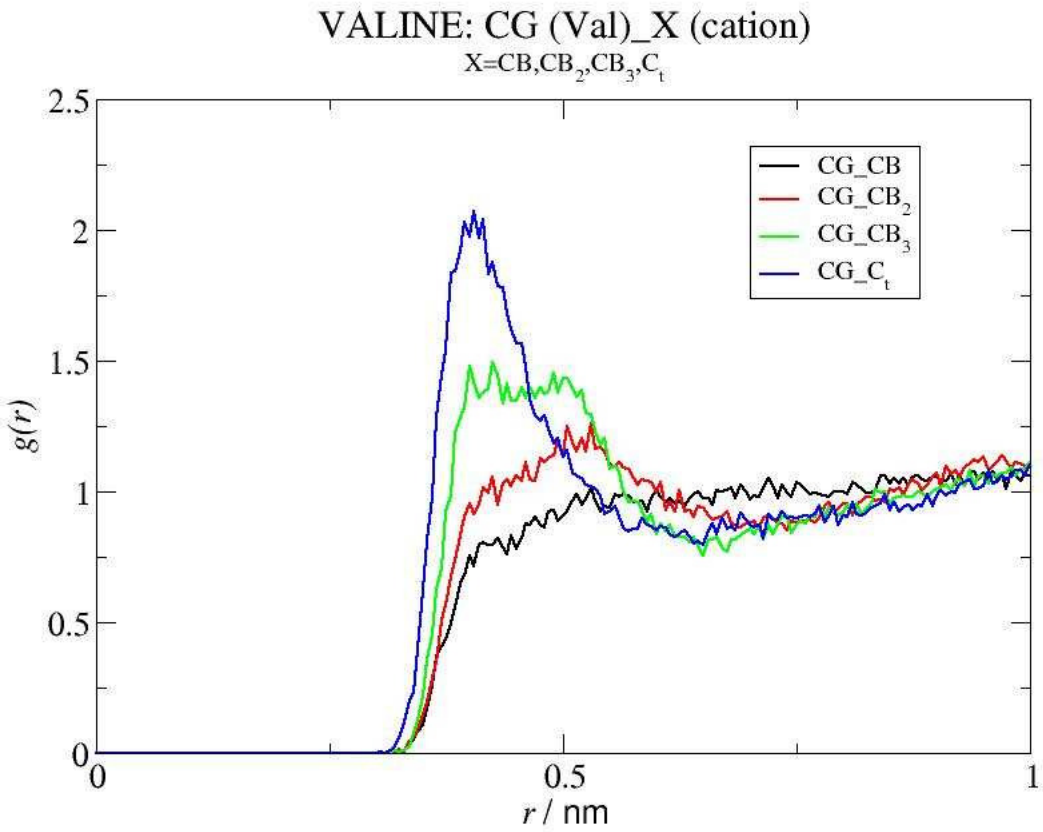


Figure 9. Radial distribution functions of the carbon atoms of the alkyl chain of the IL cation (Ct_b and CB_x) around the CG atom of Val.

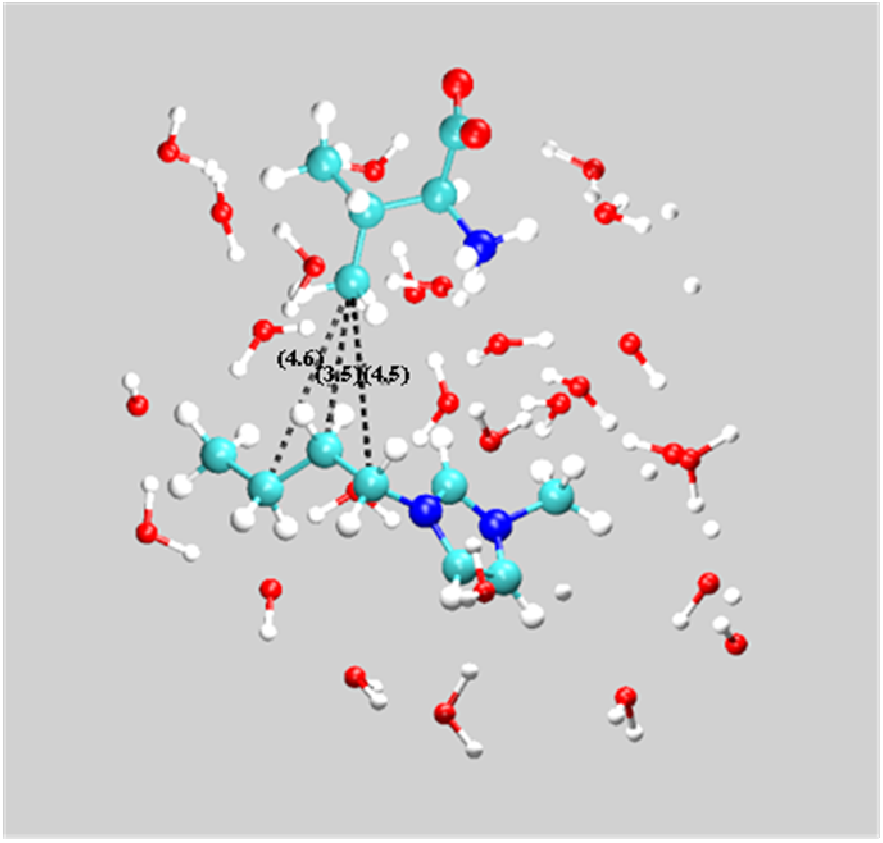
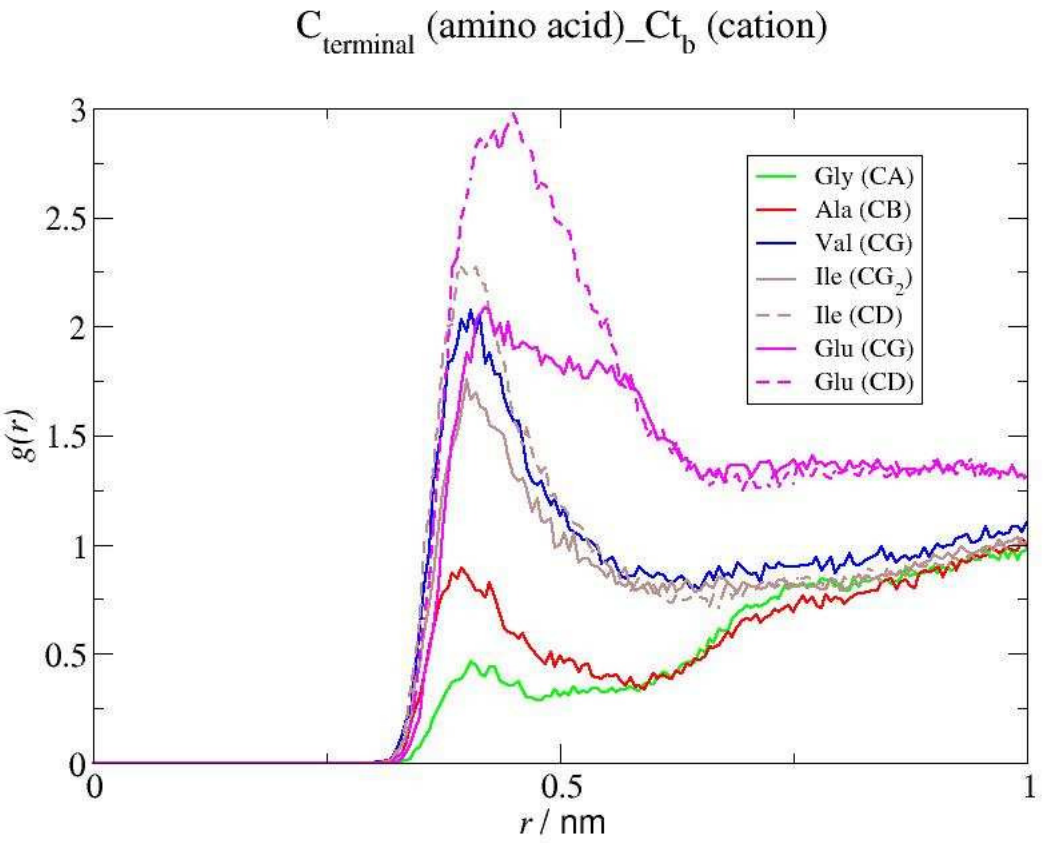
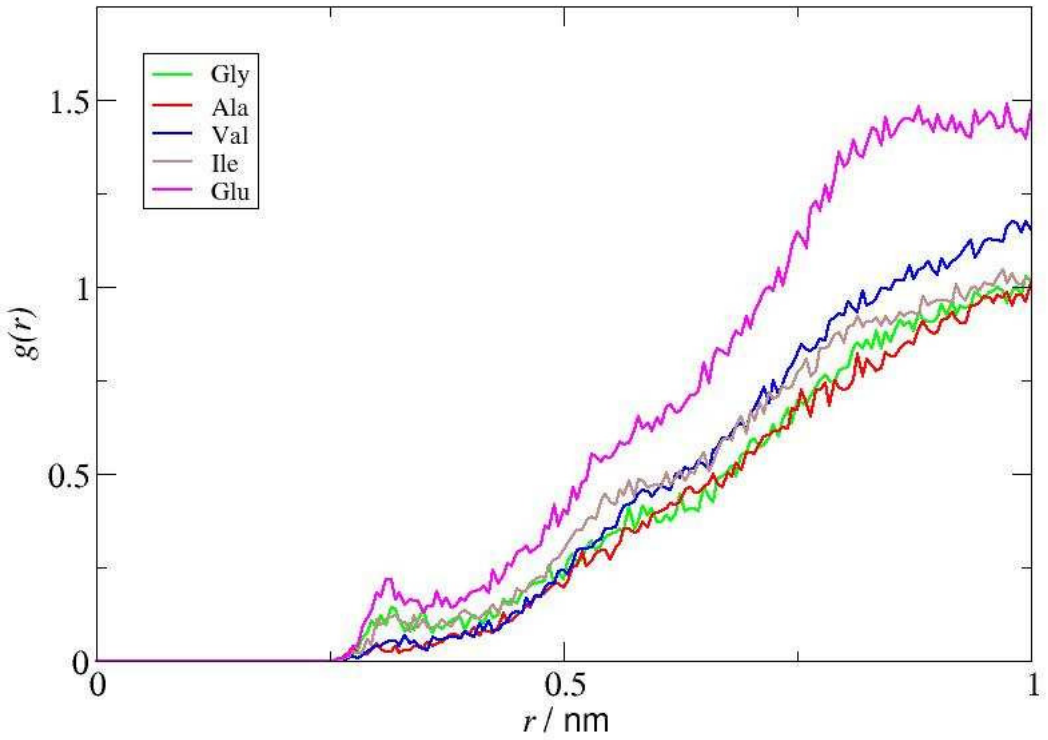


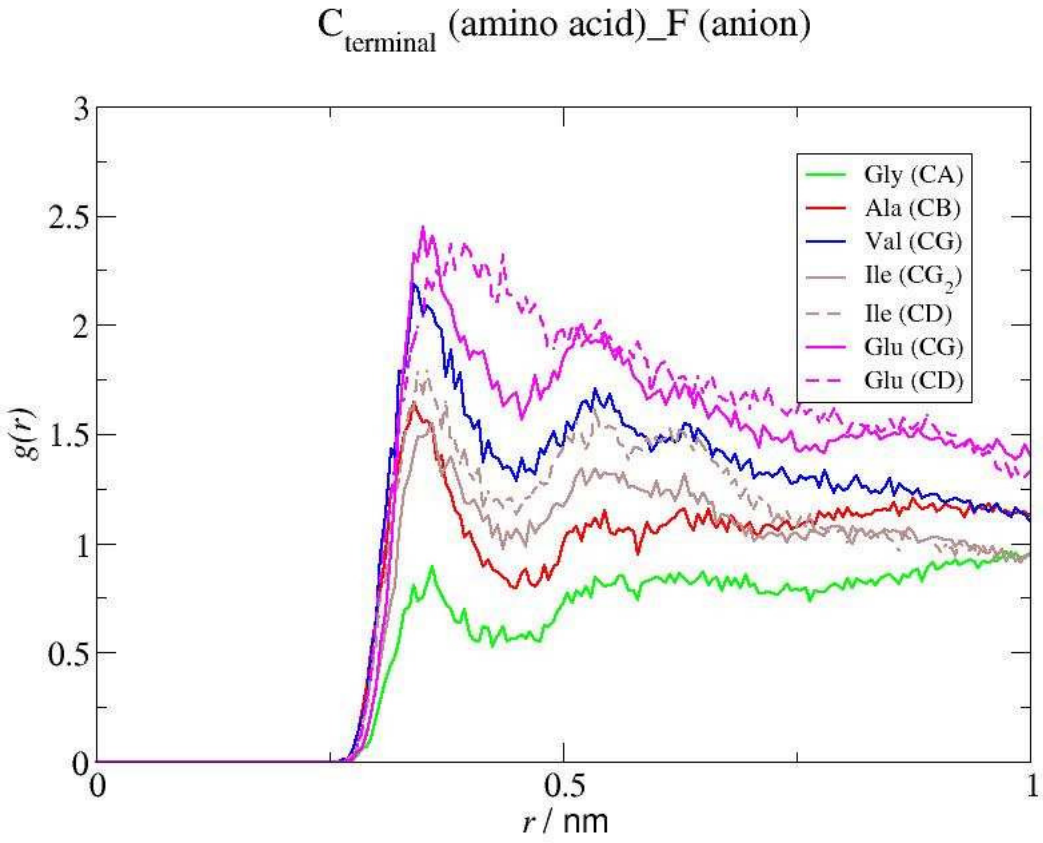
Figure 10. Snapshot from a simulation of Val mixtures, showing the distances (Å) between the CG atom of Val and the carbon atoms of the cation's alkyl chain. Only water molecules within a distance of 4Å from either [C4mim] or Val are displayed.


C terminal (amino acid)_Ct b (cation)
g(r)
r / nm
0
0.5
1
1.5
2
2.5
3
Gly (CA)
Ala (CB)
Val (CG)
Ile (CG2)
Ile (CD)
Glu (CG)
Glu (CD)

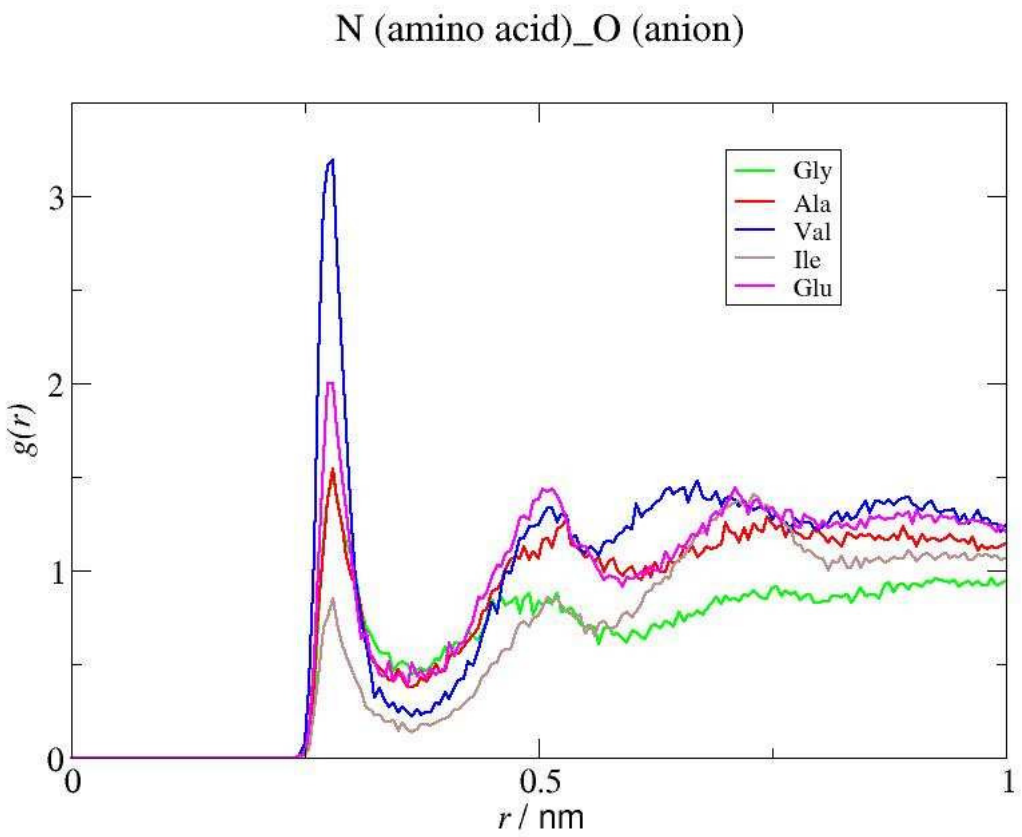
(a)



(b)

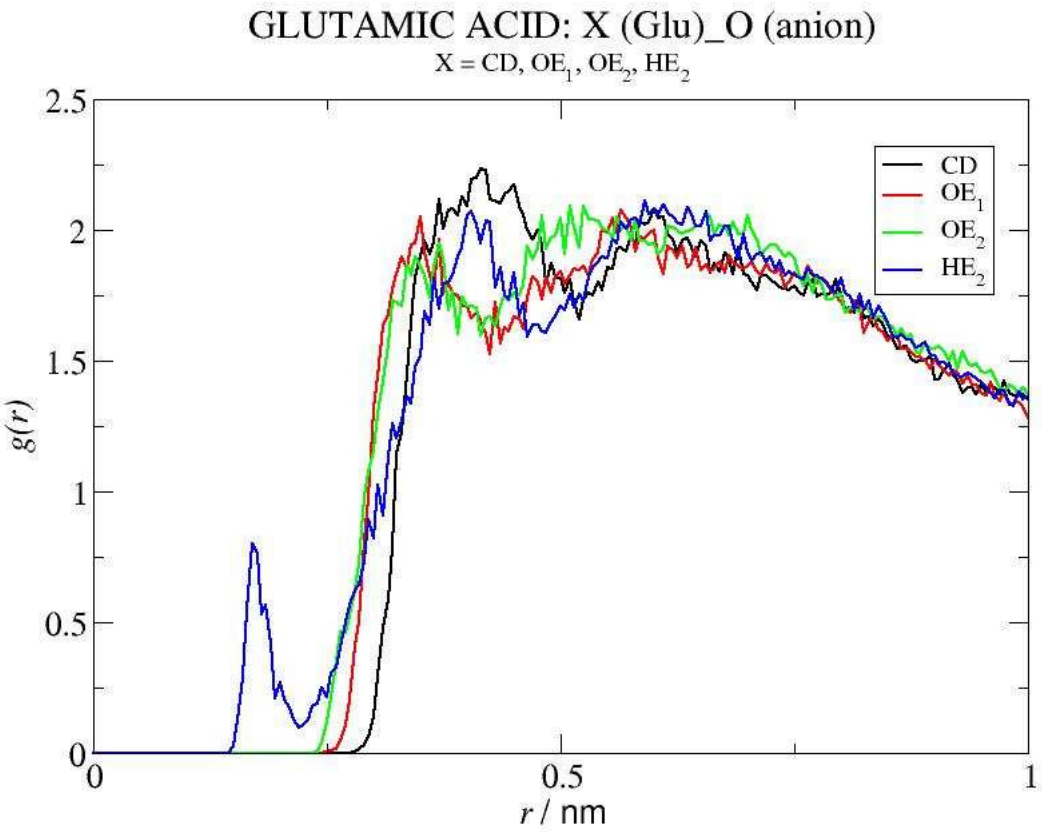

C terminal (amino acid)_F (anion)
g(r)
r / nm
0
0.5
1
1.5
2
2.5
3
0
0.5
1
Gly (CA)
Ala (CB)
Val (CG)
Ile (CG2)
Ile (CD)
Glu (CG)
Glu (CD)

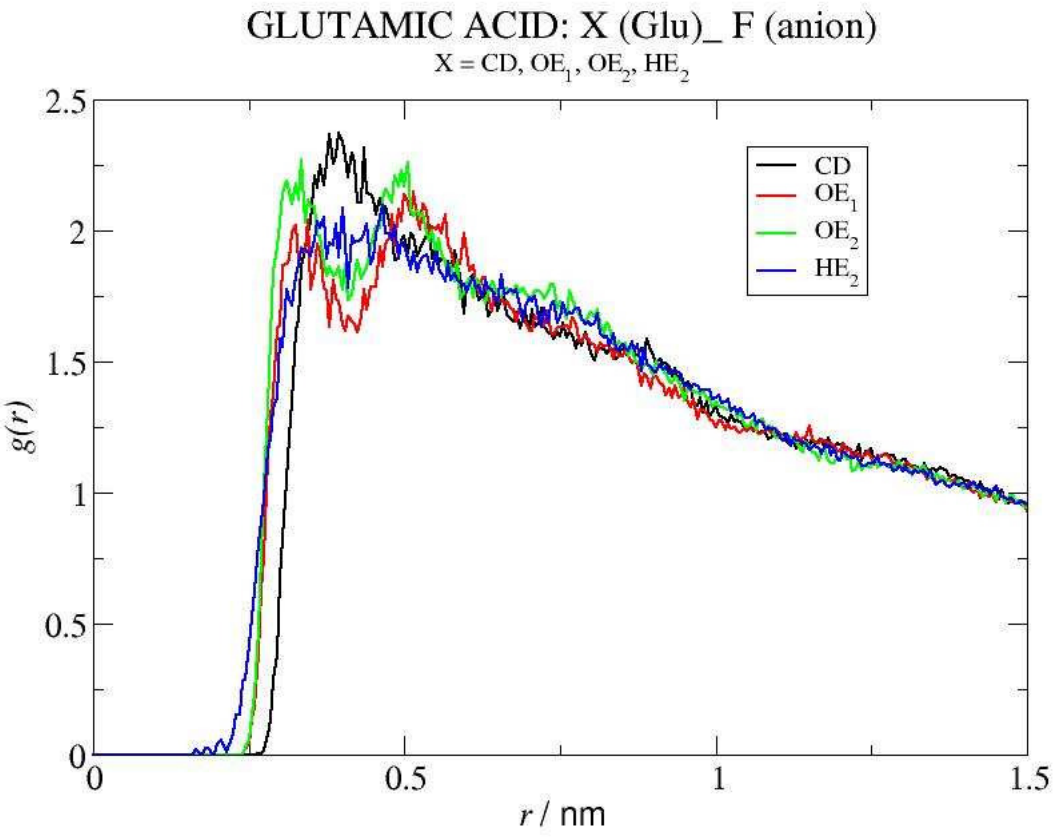
(c)



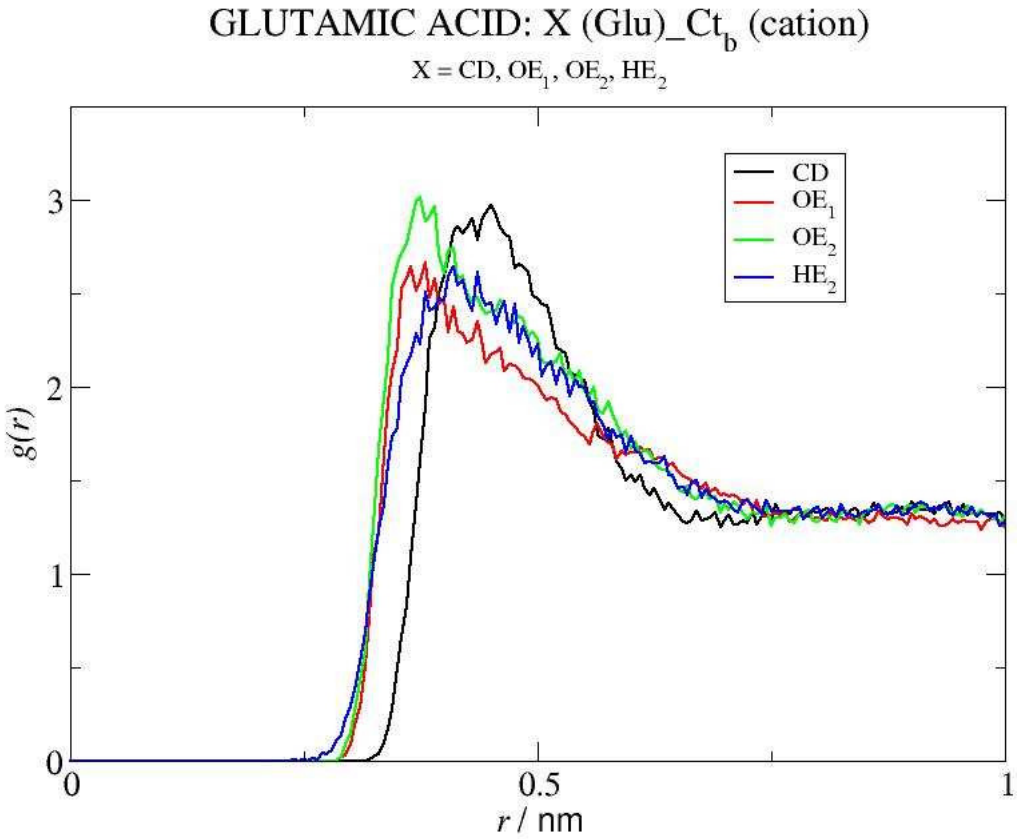
(d)

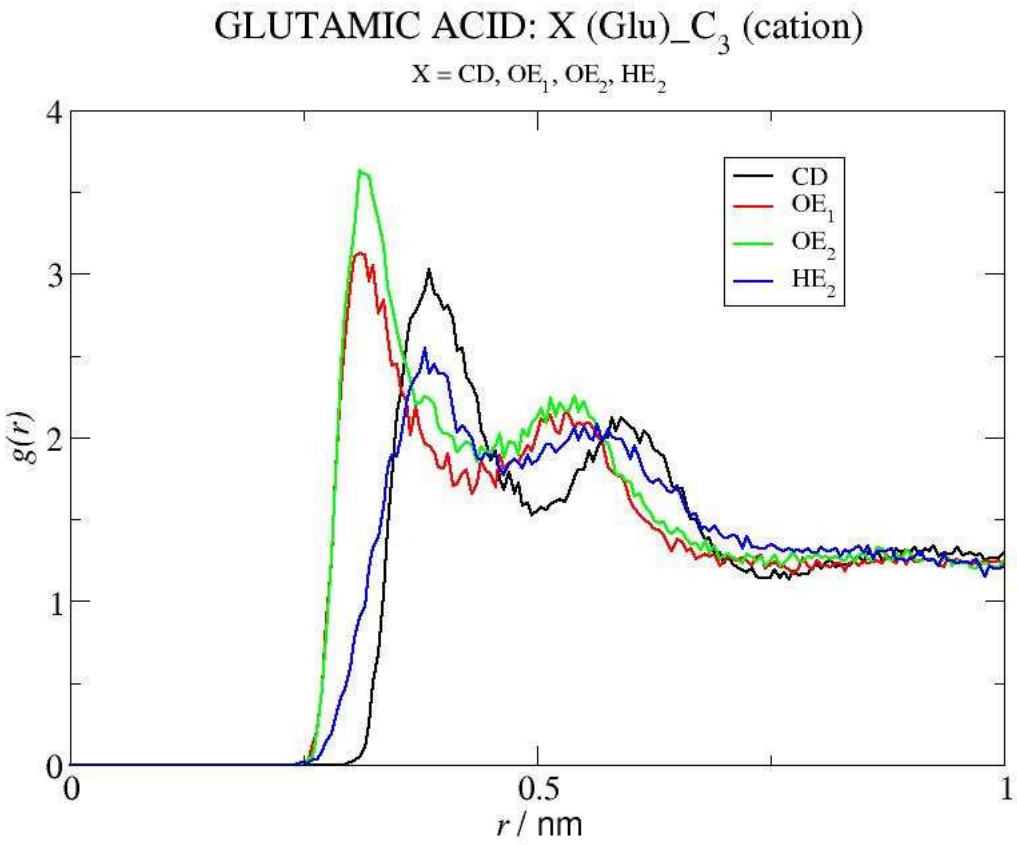
Figure 11. Radial distribution functions for the interactions between different molecular regions of the amino acids and selected atoms of the IL: (a) $C_{terminal}$ (amino acid)···Ct_b(cation); (b) O(amino acid)···C_3(cation); (c) $C_{terminal}$ (amino acid)···F(anion) and (d) N(amino acid)···O(anion).


GLUTAMIC ACID: X (Glu)_O (anion)
X = CD, OE1, OE2, HE2
g(r)
r / nm
CD
OE1
OE2
HE2
0
0.5
1
1.5
2
2.5
0
0.5
1


GLUTAMIC ACID: X (Glu)_ F (anion)
X = CD, OE1, OE2, HE2
CD
OE1
OE2
HE2
g(r)
r / nm
0
0.5
1
1.5
2
2.5

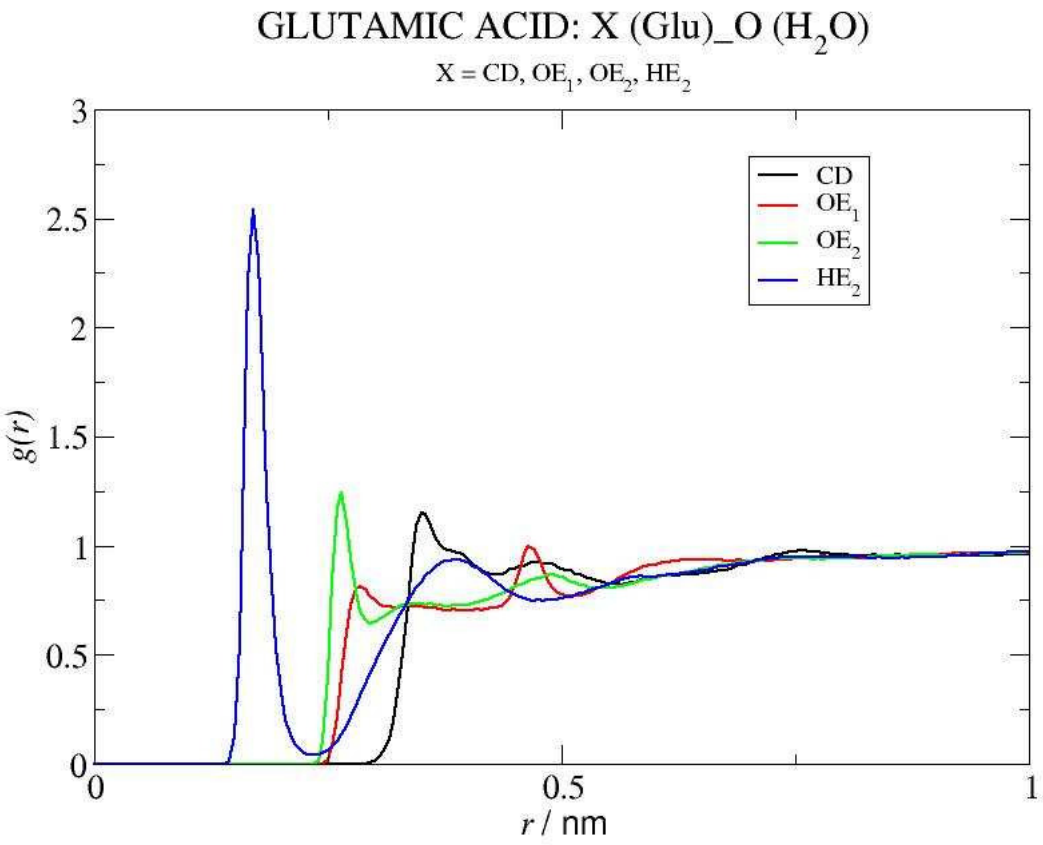
(a)


GLUTAMIC ACID: X (Glu)_Ct$_b$ (cation)
X = CD, OE$_1$, OE$_2$, HE$_2$
CD
OE$_1$
OE$_2$
HE$_2$
g(r)
r / nm
0
0.5
1
2
3

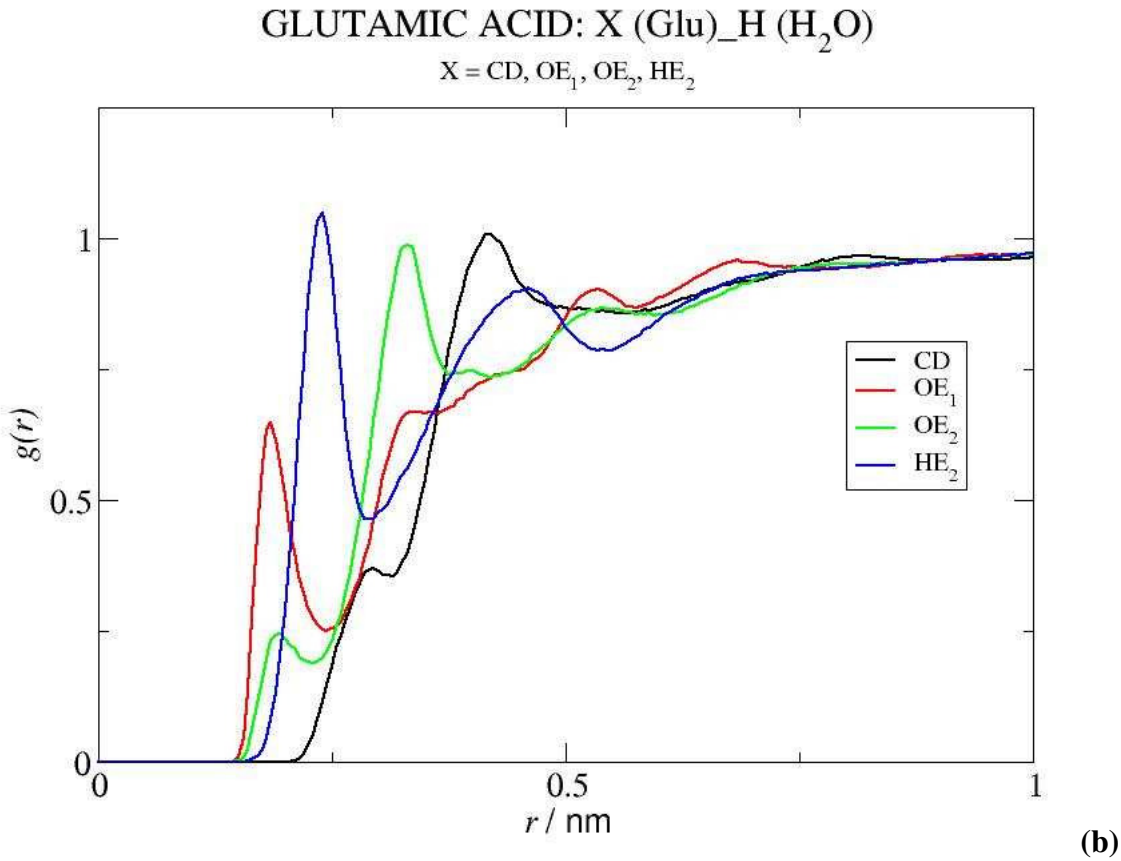


(b)

Figure 12. Radial distribution functions between the carboxyl group of Glu and selected atoms of the (a) IL anion and (b) IL cation.


GLUTAMIC ACID: X (Glu)_O (H2O)
X = CD, OE1, OE2, HE2
CD
OE1
OE2
HE2
g(r)
r / nm
0
0.5
1
1.5
2
2.5
3
0
0.5
1

(a)



(b)

Figure 13. Radial distribution functions between the water O and H atoms and the carboxyl group of Glu.

Table 1. Literature Values for Amino Acid Solubilities in Water (*s*) [35] and for Molar Gibbs Energy of Hydration ($\Delta_{hyd}G_m$) [71] of the Amino Acids' side chains, at pH=7 and 298.15 K.

amino acid	*s* / g (100 g^{-1})	$\Delta_{hyd}G_m$ / kJ.mol^{-1}
Gly	24.99	10.0
Ala	16.65	8.1
Glu	0.864	-42.7
Val	8.850	8.3
Ile	4.117	9.0

Table 2. Calculated Coordination Numbers (CN) for the Interactions between Selected Atoms in IL/water/amino acid Ternary Systems. [a]

Interaction	Gly	Ala	Glu	Val	Ile	Water
$H(NH_3^+)$-$O(H_2O)$	2.74	2.44	2.55	2.50	2.59	-
$O(COO^-)$-$H(H_2O)$	5.31	4.92	4.75	4.94	5.03	-
Ct_b(cation)-$O(H_2O)$	15.33	15.98	13.78	15.06	14.97	15.54
C(anion)-$O(H_2O)$	31.81	32.63	26.82	30.89	28.92	32.09
Ct_b(cation)-Ct_b(cation)	0.13	0.15	0.28	0.16	0.11	0.09
$C_{terminal}$(aa)-Ct_b(cation) [b]	0.01	0.05	0.27	0.14	0.13/0.11	-
$O(COO^-)$-C_3(cation)	NP[c]	NP[c]	NP[c]	NP[c]	NP[c]	-
$C_{terminal}$(aa)-F(anion) [b]	0.12	0.25	NP[c]	0.35	0.22	-
N(aa)-O(anion)	0.04	0.04	0.05	0.06	0.04	-

[a] The values of *r* at which the RDFs used for the calculation of the coordination numbers were truncated are presented in Table S8 of the Supporting Information. [b] $C_{terminal}$=CA for Gly; =CB for Ala; =CG for Val; =CG2 and =CD for Ile; =CG and =CD for Glu. [c] Not possible to identify the first peak.

Table 3. Calculated Coordination Numbers (CN) for the Interactions between the CD, OE_1, OE_2 and HE_2 atoms of the carboxyl group of Glu and Selected Atoms in IL/water/amino acid Ternary Systems. [a]

	O(anion)	F(anion)	Ct_b(cation)	C_3(cation)	O(H_2O)	H(H_2O)
CD	0.45	NP	0.27	0.10	5.41	1.39/17.80
OE_1	0.21	0.34	0.37	0.07	1.37	1.07
OE_2	0.15	0.34	0.39	0.08	1.14	0.39/7.45
HE_2	0.01	NP	0.34	0.08	0.84	3.14

[a] The values of *r* at which the RDFs used for the calculation of the coordination numbers were truncated are presented in Table S9 of the Supporting Information.

Table 4. Values (kJ mol^{-1}) of the Lennard-Jones (LJ) and Coulomb (Coul) terms of the energies calculated for the interactions (amino acid-IL cation), (amino acid- IL anion) and (amino acid-water).

	amino acid-IL cation		**amino acid-IL anion**		**amino acid-water**	
	LJ	Coul	LJ	Coul	LJ	Coul
Gly	-4.6	-1.3	-9.3	-18.9	306.7	-3841.8
Ala	-6.1	-0.9	-16.6	-19.1	213.8	-3490.9
Glu	-29.9	-13.1	-41.5	-26.2	182.3	-4047.1
Val	-14.1	-0.5	-30.0	-29.1	113.7	-3517.3
Ile	-15.0	-0.5	-28.9	-21.8	68.0	-3624.0

TOC