Imidazole in Aqueous Solution: Hydrogen Bond Interactions and Structural Reorganization with Concentration

Marco Pagliai,[*,†] Giada Funghi,[†] Dario Vassetti,[†,‡] Piero Procacci,[*,†] Riccardo Chelli,[†] and Gianni Cardini[†]

†*Dipartimento di Chimica "Ugo Schiff", Universit degli Studi di Firenze, via della Lastruccia 3, I-50019 Sesto Fiorentino, Italy*

‡*Current address: Chemie ParisTech, PSL Research University, 11 Rue Pierre et Marie Curie, F-75005 Paris, France*

E-mail: marco.pagliai@unifi.it; piero.procacci@unifi.it

Abstract

The structural and dynamic properties of imidazole in aqueous solution have been studied by means of classical and *ab initio* molecular dynamics simulations. We developed a new force field for the imidazole molecule with improved modelling of the electrostatic interactions, specifically tailored to address the well known drawbacks of existing force fields based on the atomic fractional charges approach. To this end, we reparametrized the charge distribution on the heterocyclic ring, introducing an extra site accounting for the lone pair on the deprotonated nitrogen. The accuracy of the model in describing the hydrogen bond pattern in aqueous solvent has been confirmed by comparing the classical results on imidazole-water interactions to accurate Car-Parrinello molecular dynamics simulations. The proposed classical model reproduces satisfactorily the experimental water/octanol partition coefficient of imidazole, as well as the structure of the imidazole molecular crystal. The force field has been finally applied to simulate aqueous solutions at various imidazole concentrations to obtain information on both imidazole-water and imidazole-imidazole interactions, providing a description of the different molecular arrangements in solution.

Introduction

Imidazole is an aromatic heterocycle with a 5-membered ring characterized by two non equivalent nitrogen atoms, as shown in Figure 1. The imidazole ring occurs, as molecule or building block, in systems with important chemical, biological, pharmacological and technological applications.[1–4] In water, at physiological pH ($\sim$ 7.4), imidazole is present both as neutral and protonated species and interacts with the solvent forming hydrogen bonds.[1]
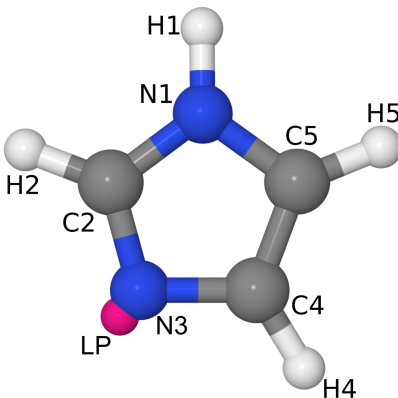
Figure 1: Molecular structure of imidazole. Atom colors: nitrogen, blue; carbon, gray; hydrogen, white. The violet sphere represents the electron pair (i.e. the Pipek-Mezey centroid[5]).

The electronic structure of neutral and cationic imidazole in aqueous solution has been recently investigated using soft X-ray core level spectroscopies[6] in combination with *ab initio* calculations on the isolated molecule. Microscopic details of the structural properties of hydrated imidazole in aqueous solution have been recently revealed by way of neutron scattering experiments[7], interpreted with the aid of *ab initio* and classical molecular dynamics (MD) simulations. CHARMM[8] and OPLS[9] force fields have been used in classical calculations, while Born-Oppenheimer molecular dynamics simulations with the Quickstep module of the CP2K package[10] have been adopted to obtain the *ab initio* data. It has been observed that the OPLS model, contrarily to the CHARMM force field, significantly overestimates the imidazole-imidazole association, as revealed by the comparison of computed and experimental differential second order scattering function $\Delta\Delta S(Q)$. On the other hand, at variance with the *ab initio* data, the CHARMM force field predicts a poorly oriented hydrogen bond between the nitrogen hydrogen bond acceptor (N3) on the ring and the hydrogen bond donor water molecule, underestimating the bias towards the expected orientation along the CNC bisector due to the sp^2 hybridisation of N3 (see Figure 3 in Ref. 7).

In previous studies[11,12] the hydration of neutral and cationic imidazole was investigated by photoelectron spectroscopy. In Ref.,[11] in particular, the comparison between high level quantum chemical calculations and experimental vertical ionization potential (VIP) was

3

used as probe for identifying the most likely structures of neutral imidazole$(H_2O)_n$ clusters in solution, showing the importance of the hydrogen bonds lying on the ring plane.

"Molecule speciation" during imidazole crystallisation from supersaturated aqueous solution was studied recently using X-ray Raman scattering (XRS).[13] Near-edge X-ray absorption fine structure (NEXAFS) has been also employed recently[14] as a probe for imidazole-imidazole self-association in a series of samples where the heterocyclic compound concentration was varied from ∼0.5 M to ∼8 M. Negligible changes are observed in the C and N K-edge XRS spectra during cooling, indicating that the average solvation structure around imidazole molecules does not change significantly while traversing the thermodynamically metastable supersaturated liquid phase until the onset of crystallisation. These results are consistent with those obtained by NEXAFS experiments[14] where it has been observed that the NEXAFS signal does not appreciably change with concentration, while the spectra of species in solution substantially differ from those of imidazole monomers in gas phase. This suggests that self-association occurs not only for the system at the lowest studied concentration (∼0.5 M), but also for more diluted solutions.[14]

The self-association of imidazole in aqueous solutions in the wide concentration range employed in the same range of the NEXAFS experiments was previously studied with classical molecular dynamics simulations,[15] by comparing the results obtained with a point charge model (based on the Amber/TIP3P force field[16,17]) to those obtained using atomic high-rank electrostatic multipole moments[18] (EMM) fitted on high quality *ab initio* electron densities of the isolated imidazole and water[19] molecules. The study showed that a more realistic representation of imidazole electron density via EMM leads to appreciable differences in the local structure of the solutions, as showed by the 3D spatial distribution function for imidizole-imidazole self association and imidazole-water interactions.

In the present study, given the deficiencies of the current AMBER, CHARMM and OPLS force field implementations evidenced in the previously referenced studies,[7,14,15] we have developed a new force field for the imidazole molecule based on the fixed charges paradigm

augmented by the introduction of an extra site (EP)[20] accounting for the lone pair (LP) on the deprotonated nitrogen atom on the ring. To this end, we followed a computational protocol to model the electrostatic potential of the imidazole ring similar to that adopted with success to simulate pyridine[21] and pyrrole.[22]

In order to assess the structural properties of the hydrogen bond pattern in solvated imidazole, the imidazole-water Radial Distribution Function (RDF) from MD simulations, were compared to *ab initio* molecular dynamics simulations using the Car-Parrinello method (CPMD).[23–25] The presented *ab initio* results on solvated imidazole are not only important as a valuable reference for assessing the reliability of fixed charges force fields, but they are interesting *per se*, since CPMD is known to be particularly effective for the description of hydrogen bonded systems, realistically encompassing charge transfer and polarisation effects.[23,26–30]

Our force field for imidazole was devised initially by assigning the atomic types according to the recently released GAFF2[31] protocol using the web interface PrimaDORAC.[32] Such GAFF2 potential, called the *standard* model, was further improved by carefully re-parametrizing the electrostatic part of the potential using accurate QM calculations based on the C-PCM method[33,34] and introducing a negatively charged EP along the bisector of the CNC angle, accounting for the lone pair on the N3 nitrogen atom. The latter augmented fixed charge approach (called *improved* model) reproduces with accuracy the dipole and quadrupole moment and, as shown by the comparison of classical and CPMD imidazole-water RDF's, is apparently able to significantly alleviate the problems observed in the *standard* model as well as in the OPLS and CHARMM models tested in the study of Dubou-Dijon *et al.*,[7] providing a more realistic description the hydrogen bond interactions involving the nitrogen atom N3 of heterocyclic molecule.

The *improved* model has been further validated performing molecular dynamics simulations of imidazole in aqueous solutions with the concentration ranging from 0.5 M to 8.22 M, analogously to previous theoretical[15] and experimental[14] studies, allowing to charac-

terise the structural arrangements of imidazole molecules in solution. Our results regarding the structural arrangement of interacting imidazole molecules are essentially in accord with Liem *et al.*[15] conclusion and with the NEXAFS experiments,[14] evidencing in general a weak dependency of imidazole association on concentration and a surprising small incidence of stacked dimers at all concentrations.

As a further check of the reliability of the *improved model*, we have computed the octanol-water partition coefficient of imidazole. In this regard, we stress that solvation free energies differences of biologically relevant molecules in aqueous and hydrophobic phases are essential for predicting the distribution of drugs within the body.[35,36] Binding free energies themselves can be viewed[35,36] as the difference between the solvation free energy of the ligand in bulk solvent and in the composite environment of the protein binding pocket.

Finally, it has been verified that the *improved* model allows for a correct description of the intermolecular interactions in the crystal, yielding cell parameters matching satisfactorily the experimental counterpart.

The paper is organised as follows. Computational details for CPMD, force field parametrization and classical MD simulations are given in Section 2. In Section 3, we first discuss the structural and dynamic properties on solvated imidazole obtained with the CPMD approach. Next, we compare the CPMD imidazole-water RDF's to those computed via MD, using both the *standard* and *improved* model. We then present the results obtained for the LogP coefficient of imidazole via fast switching alchemical calculations[37] using the *improved* model. Finally we discuss imidazole-imidazole association in aqueous solution as predicted by the *improved* model in a series of simulations with increasing imidazole concentration. In Section 4, some conclusive remarks are presented.

Computational Details

Car-Parrinello Molecular Dynamics Simulations

Car-Parrinello molecular dynamics simulations[23–25] have been carried out on a fully deuterated system made up by 1 imidazole and 64 water molecules with the CPMD code.[38] The potential has been described employing the BLYP[39,40] exchange and correlation functional in the DFT calculations.

The effects of the van der Waals potential term on the structural properties of imidazole in water has been verified performing *ab initio* molecular dynamics simulations with and without the corrections proposed by Grimme.[41] Although Grimme and coworkers have proposed improved version for the van der Waals corrections,[42,43] it has been shown[44] that the method employed in the present work is effective in the description of the structural properties of liquid water. In fact, comparing the results of *ab initio* molecular dynamics simulations obtained by Dubou-Dijon *et al.* [[7]], which employed the BLYP exchange and correlation functional in conjunction with both the D2 or D3(BJ) Grimme correction scheme,[42,43] it is evident that the hydrogen bond interactions between imidazole and solvent are almost insensitive to the different parameters of dispersion term.

In all the simulations, the plane wave expansion has been truncated at 85 Ry[30] modelling the core region through Martins-Troullier norm-conserving pseudopotentials[45] within the Kleinman-Bylander decomposition,[46] while the electron fictitious mass has been set to 700 au. The systems have been thermalized at 300 K, by a velocity rescaling algorithm, integrating the equations of motion with a time step of 5 au (~ 1.2 fs) in the NVE ensemble in a periodic cubic box with 12.6632 Å of side. NVE ensemble has been adopted to avoid a series of problems related to the dimension of the sample. The production runs have been carried out for 30 ps.

Force Field parametrization

The initial parameters for imidazole molecules have been obtained by using the web interface PrimaDORAC.[32] PrimaDORAC allows the assignment of atomic charges at AM1/BCC level[47] through the MOPAC7 program[48] and chemical topology and bond parameters following the convention adopted by the Generalized Amber Force Field[16] (GAFF2). The atomic charges of the force field, called *standard*, have been collected in Table 1.

In studying other heterocyclic molecules, such as pyridine[21,49] and pyrrole,[22] it has been observed that a correct description of intermolecular interactions in pure liquid and in aqueous solution requires a particular attention in modelling the electrostatic part of the potential. In fact, it has been observed that the implicit inclusion of polarisation effects due to the environment, performing DFT calculations at B3LYP/6-31+G(d) level of theory and employing the C-PCM method,[33,34] allows to obtain more effective atomic charges for molecular dynamics simulations. A further improvement is achieved adjusting the atomic charges through a fitting procedure using as reference the molecular dipole and quadrupole moments.[22] Moreover, as in pyridine,[21,49] the explicit description of the sp^2 lone pair of the nitrogen atom (through a centroid of a localised molecular orbital computed with the Pipek-Mezey procedure[5]) allows to correctly describe the directional character of the hydrogen bond interactions, which are of fundamental importance for imidazole. The atomic charges of this force field, called *improved* and obtained adopting the above discussed schema (using Mulliken charges only as initial guess in the fitting procedure) are reported in Table 1 (the topological and parameter files are provided as Supporting Information). The DFT calculations have been carried out with the Gaussian 09 suite of programs.[50]

Table 1: Atomic charges of the *standard* and *improved* imidazole models. The numbering of hydrogen atoms follows that of the atoms on the molecular ring, as shown in Figure 1 . LP represents the charge on the electron pair (i. e. the Pipek-Mezey centroid[5]).

label	*standard*	*improved*
C5	0.1185	0.0074
H5	0.0406	0.0727
C4	-0.2710	-0.0970
H4	0.1758	0.1585
N1	-0.3175	-0.4460
H1	0.2997	0.3820
C2	0.2041	0.2330
H2	0.0589	0.0465
N3	-0.3090	0.0000
LP	n/a	-0.3571

Classical molecular dynamics setup

The samples adopted to validate the force field for imidazole-water interactions are made up by 1 imidazole and 338 water molecules and have been simulated employing the *standard* and *improved* force field for imidazole. Water has been described with the TIP3P model.[17] The systems have been preliminary thermalized for 300 ps, while the time length of the production runs are 3 ns. The samples with increasing imidazole concentration have been initially thermalized for 500 ps followed by a production run of 5 ns. The composition of the systems for the different concentration values is reported in Table 2.

All the simulations have been performed in the NPT ensemble. The external pressure was set to 1 atm using a Parrinello-Rahman Lagrangian[51] with isotropic stress tensor while temperature was held constant at 300 K using three Nosé Hoover-thermostats coupled to the translational degrees of freedom of the systems and to the rotational/internal motions of the solute and of the solvent. The equations of motion were integrated using a multiple time-step r-RESPA scheme[52] with a potential subdivision specifically tuned for biomolecular systems in the NPT ensemble.[53] The long range cut-off for Lennard-Jones interactions was set to 13 Å in all cases. Long range electrostatic were treated using the Smooth Particle

Mesh Ewald method,[54] with an α parameter of 0.38 Å^{-1}, a grid spacing in the direct lattice of about 1 Å and a fourth order B-spline interpolation for the gridded charge array.

Table 2: Number of imidazole and water molecules,volumes and average density of the samples for each simulation at the respective concentration value.

Conc. (M)	# imidazole	# water	volume (nm^3)	Density (g/cm^3)
0.49	16	1724	54.168	0.9855 ± 0.0065
0.95	32	1726	55.967	0.9872 ± 0.0064
1.79	64	1728	59.504	0.9903 ± 0.0064
2.17	80	1730	61.309	0.9917 ± 0.0061
2.52	96	1732	63.111	0.9930 ± 0.0061
2.88	112	1724	64.595	0.9944 ± 0.0062
3.20	128	1730	66.480	0.9962 ± 0.0059
3.50	144	1733	68.335	0.9969 ± 0.0059
3.80	160	1730	69.971	0.9982 ± 0.0059
4.09	176	1724	71.505	0.9995 ± 0.0058
4.35	192	1728	73.351	1.0007 ± 0.0059
4.61	208	1724	74.969	1.0016 ± 0.0057
4.84	224	1730	76.883	1.0025 ± 0.0057
5.08	240	1727	78.484	1.0040 ± 0.0056
5.29	256	1728	80.287	1.0043 ± 0.0056
5.72	256	1528	74.169	1.0065 ± 0.0057
6.24	256	1328	68.100	1.0083 ± 0.0061
6.85	256	1128	62.003	1.0110 ± 0.0065
7.60	256	928	55.907	1.0142 ± 0.0069
8.53	256	728	48.797	1.0185 ± 0.0072

All simulations have been carried out with the ORAC program.[51,53,55]

Results and Discussion

Car-Parrinello Molecular Dynamics Simulations

The structural and dynamic properties of imidazole in water have been characterized through the analysis of the hydrogen bond interactions. We first compared the RDF's involved in hydrogen bond as acceptor or donor with solvent, with and without van der Waals corrections. As shown in Figure 2, only slight differences can be observed by comparing these functions, although the simulations with van der Waals corrections provide a stronger imidazole-solvent

interaction. These results faithfully reproduce those obtained by *ab initio* molecular dynamics simulations[7] employing the BLYP exchange and correlation functional and both the D2 and D3(BJ) Grimme correction scheme.[42,43] Some differences can be observed in the RDF's related to distances between the centre of mass of imidazole and hydrogen atoms of solvent. Shorter distances are obtained with van der Waals corrections and this reflects the propensity of imidazole to form improper hydrogen bond with water molecules.
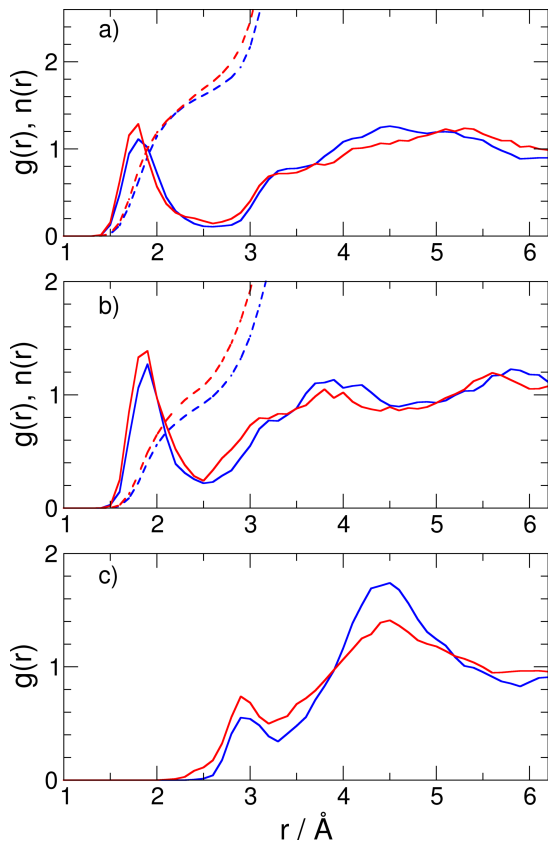
Figure 2: Pair radial distribution function, $g(r)$, and coordination number, $n(r)$. Labeling with CM the centre of mass of imidazole and with HW and OW the hydrogen and oxygen atom of water, respectively, the reported interactions refer to: a) N3···HW; b) H1···OW; c) CM···HW. The radial distribution functions have been reported with solid lines, whereas the coordination numbers with dashed lines. The results with and without van der Waals corrections are reported in red and blue, respectively.

The insensitivity of structural properties to the inclusion of the van der Waals corrections is also observed for the hydrogen bond dynamics, which has been characterized employing the continuous function proposed by Pagliai *et al.*:[29,56]

$$F_j^{HB} = A_j(r(t)) \cdot B_j(\theta(t)) \tag{1}$$

with $A_j(r(t))$ and $B_j(\theta(t))$ given by:

$$\begin{cases} A_j(r(t)) = e^{-(r_e-r_j(t))^2/(2\sigma_r^2)} & \text{if } (r_e - r_j(t)) < 0 \\ A_j(r(t)) = 1 & \text{if } (r_e - r_j(t)) \geq 0 \\ \\ B_j(\theta(t)) = e^{-(\theta_e-\theta_j(t))^2/(2\sigma_\theta^2)} & \text{if } (\theta_e - \theta_j(t)) < 0 \\ B_j(\theta(t)) = 1 & \text{if } (\theta_e - \theta_j(t)) \geq 0 \end{cases}$$

where the F_j^{HB} function represents a time dependent measure of the strength of the hydrogen bond interaction between water molecule j and imidazole, assuming values in the range between 1 and 0, depending of the deviation of $r_j(t)$ and $\theta_j(t)$ from the reference values r_e and θ_e. In fact, the values of the parameters r_e, θ_e, σ_r and σ_θ are obtained from the histograms of the studied hydrogen bond distance, $h(r)$, and angle, $h(\theta)$. r_e and θ_e represent the position of the first peak in $h(r)$ and $h(\theta)$, respectively, while σ_r and σ_θ are the half widths at half maximum in $h(r)$ and $h(\theta)$, respectively. $r_j(t)$ and $\theta_j(t)$ are the instantaneous distance and angle involved in the interaction between solvent and imidazole. This function provides a geometrical description of the hydrogen bond interactions analogous to standard geometry criteria, but with the advantage to correctly describe fast oscillations of interacting molecule in and out the region adopted to state the presence of the hydrogen bond interaction. No arbitrary time intervals are needed to follow the hydrogen bond evolution, allowing to determine the hydrogen bond life time in good agreement with experiments, as reported by Candelaresi et al.[28] for methyl acetate in water.

To follow the water molecule j involved in the interaction during the simulation, the hydrogen bond function for a value of F_j^{HB} greater than 10^{-4} is reported in Figure 3. Although imidazole forms at least one hydrogen bond with solvent both as acceptor and donor, the j-th

water molecule involved in the interaction is not always the same but a series of exchanges occurs during the simulation runs. Similar dynamics have been observed in studying the interactions between amino acids, in particular histidine, and water.[57]
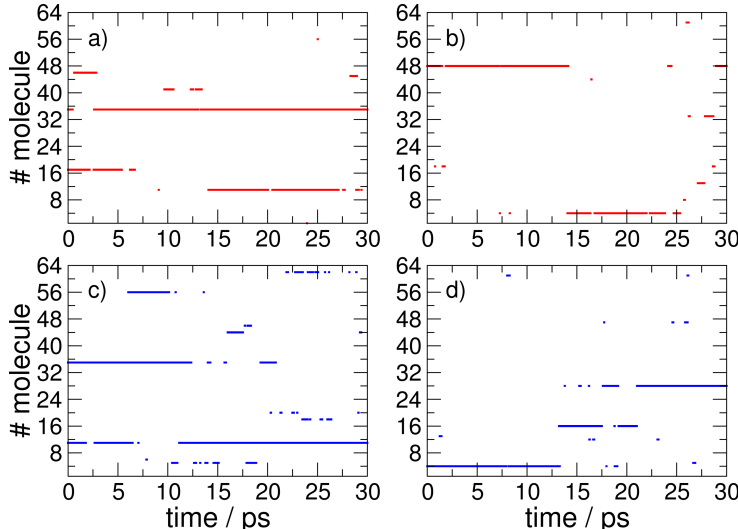

Figure 3: $j-th$ molecule involved in the hydrogen bond interaction with imidazole during the CPMD simulations. In a) and b) are reported the CPMD results for the sample without van der Waals corrections, while these are taken into account in c) and d). In a) and c) imidazole molecule acts as acceptor, while in b) and d) as donor.

Comparison of Imidazole Force Fields for Classical Molecular Dynamics Simulations

It has been shown in the interpretation of neutron scattering experiments[7] that *ab initio* molecular dynamics simulations of imidazole in water provide an accurate description of the structural properties and of the hydrogen bond interactions. Dubou-Dijon *et al.*[7] have employed the OPLS[9] and the CgenFF[8] classical force field to simulate imidazole in aqueous solution, obtaining for CHARMM-based CgenFF "excellent agreement with experimental neutron scattering data". However, it has also been reported by the same authors that there are problems in the description of the hydrogen bond interaction involving the nitrogen atom N3 of heterocycle with both the OPLS and CHARMM force fields. More in detail, quoting Dijon *et al.*[7] "the picture of the H-bonded structure for imidazole [emerging from classical

13

simulations] is quite different from what one would expect based on the N3 sp^2 hybridisation, with the nitrogen lone pair receiving a single H-bond in the imidazole plane and on the CNC bisector line. Description of this kind of properties for aromatic compounds is a well-known weak point of classical force field since no potential term imposes a sp^2 hybridisation at the N3 nitrogen." This structural deficiency of the force fields with atomic charges is reflected (for both OPLS and CHARMM) in a higher coordination number of N3 with respect to that obtained from AIMD simulations performed with the BLYP exchange and correlation functional in conjunction with both the D2 and D3(BJ) Grimme correction scheme.[42,43]
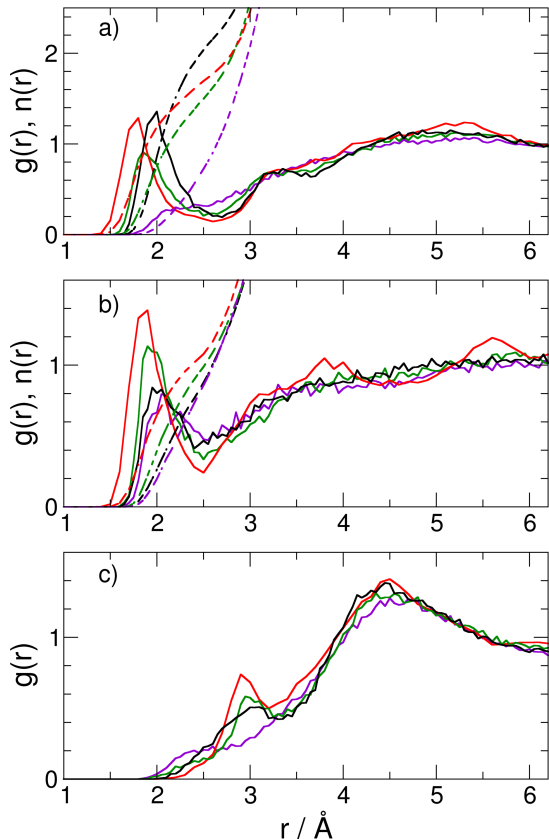

Figure 4: Pair radial distribution function, $g(r)$. Labelling with CM the center of mass of imidazole and with HW and OW the hydrogen and oxygen atom of water, respectively, the reported interactions refer to: a) N3···HW; b) H1···OW; c) CM···HW. In red, violet and green are reported the results obtained with CPMD (with Grimme van der Waals correction) and classical molecular dynamics simulation employing the *standard* and *improved* model, respectively. In black are reported the results obtained with CHARMM[8] force field. The coordination number has been reported with dashed lines.

Figure 4 shows the comparison of the RDF's related to the acceptor and donor hydrogen bond interactions and those between the centre of mass of imidazole with the solvent. Although the two developed force fields are either characterized by fixed atomic charges, the *improved* model takes into account implicitly the average polarisation effects due to the solvent and reproduces the directional character of the hydrogen bond interaction through the virtual site, providing a definitely more accurate description of the structural properties of the N3 hydration with respect to the *standard* model. This result can be appreciated by the comparison of CPMD and classical (*improved* model) spatial distribution functions (SDF) for imidazole-water interactions reported in Figure 5.
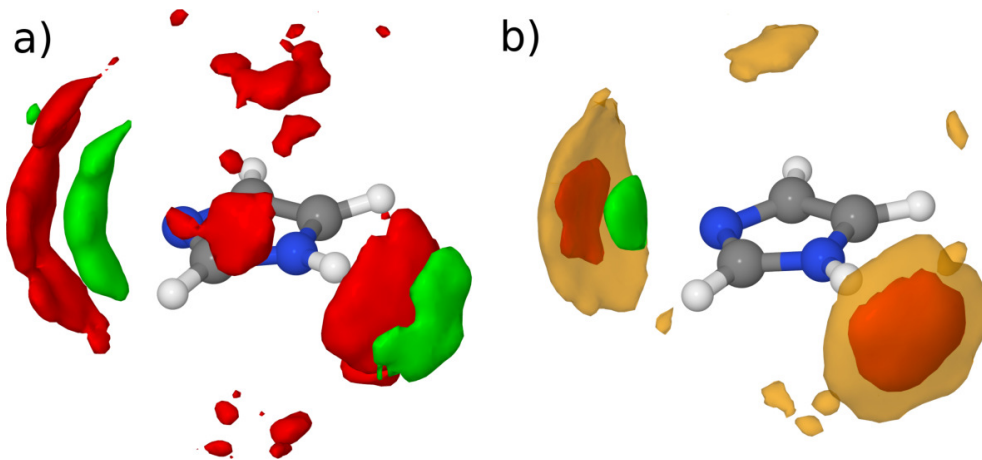

Figure 5: Spatial distribution function for the imidazole-water interactions. Isosurfaces of water oxygen (red, isovalue 5) and hydrogen (green, isovalue 2.5) atoms close to imidazole molecule. a) and b) refers to SDF for CPMD and classical MD (with *improved* model) simulations, respectively. In orange translucent are depicted the surfaces of water oxygen with isovalue 2.5. The SDF surfaces have been obtained analysing MD trajectories with the Travis program.[58]

Furthermore, the *improved* model yields a first shell coordination number of N3 (see Figure 4) that is much closer to the CPMD counterpart with respect to the *standard* model.

These are important results because they show that a slight modification of the simple atomic charges approach in classical force fields of heteroaromatic compounds, consisting in the introduction of EP accounting for the LP and computed according the Pipek-Mezey procedure[5]), overcomes the well-known deficiencies in describing solute hydration that are

shared by the most popular force fields. Most importantly, the *improved* model fully preserves the computational efficiency of the simple atomic charge approach.

To further validate the new developed force field, the water/octanol partition coefficient for imidazole, LogP, has been computed using both the improved model and the CHARMM force field.[8] LogP measures the partition of a molecule between a hydrophobic (e.g. lipid bilayer) and a cytosolic environment and as such is a key quantity in the assessment of drug-likeness of drug candidates.[59] The LogP for the two model is computed by evaluating independently the solvation free energy in both water and 1-octanol in standard conditions. The 1-octanol potential model is based on the atomic types and charges defined in the improved *slipids* force field,[60] while the TIP3P model[17] has been adopted for water.

Solvation free energies have been computed with ORAC[51,53,55] program using the fast switching annihilation method.[37,61–63] Sampling of the initial state in 1-octanol has been performed using a Hamiltonian Replica Exchange technique[64] by scaling the torsional potential according to the scheme proposed by Cardelli *et al.*[65]

Figure 6 shows the work distributions obtained in the fast non-equilibrium annihilation work of the imidazole molecule in water (upper panel) and in 1-octanol (lower panel) for the *improved* model and CHARMM force field. Normality of the distribution is assessed using the Anderson Darling (AD) test.[66,67] For normal distributions, the solvation free energies are computed from the mean and variance of the work distributions using the Crooks theorem-based[68] unbiased estimate[61,62] $\Delta G = \langle W \rangle - \frac{1}{2}\beta\sigma^2$, where $\beta = 1/k_B T$ while σ is the variance. In case of AD failure, the estimate of the solvation free energy was done by exploiting again the Crooks theorem for Gaussian mixtures as described by Procacci *et al.*[69] with the parameters of the N_g components mixture being determined via the expectation-maximisation algorithm.[70,71] Based on the solvation energies reported in Figure 6, the LogP value obtained using the *improved* model and the CHARMM force field is 0.30±0.19 and -0.65±0.29, respectively. The experimental value,[72] LogP=-0.08, indicate that imidazole is essentially equally miscible in water and 1-octanol. The CHARMM force field significantly overesti-

mates the solvation free energy of imidazole in water, in line with its enhanced hydrogen bond interaction (as shown in Figure 4). The improved model is closer to the experimental value, although somewhat overestimating imidazole miscibility in 1-octanol.

These results neatly show the importance of the polarisation effects, which are taken into account in CPMD simulations and have been partially introduced also in the *improved* force field. Further improvement could be achieved by using polarizable force fields, which could be interesting also to obtain information related to vibrational spectroscopy.[73,74]
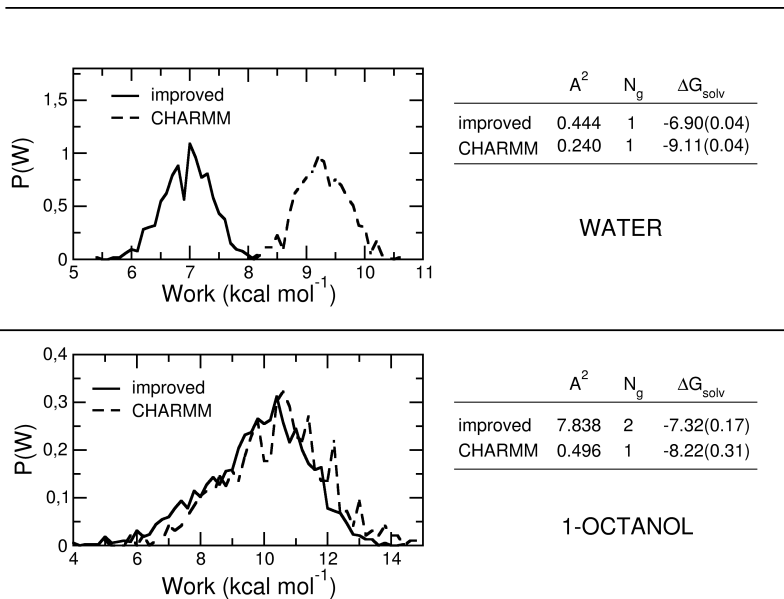

Figure 6: Work distribution in 1-octanol and in water using the *improved* model and the CHARMM[8] force field for imidazole.

Classical Molecular Dynamics Simulations: Imidazole-Water and Imidazole-Imidazole Interactions

The *improved* force field has been applied also to the study of a series of aqueous samples with different imidazole concentrations, as summarised in Table 2. These simulations allow to validate the effectiveness of the *improved* force field in reproducing both the imidazole-water and imidazole-imidazole interactions. The density values have been obtained performing molecular dynamics simulations in the NPT ensemble. As shown in Figure 7 (empty symbols),

the density variation with concentration has a trend similar to that obtained by molecular dynamics simulations with the *improved* force field, exhibiting an almost perfect linear dependency and consistently lying below the experimental data (data taken from Table 3 of Ref. 15) due to the well-known underestimation of the liquid density of the TIP3P model.[75]
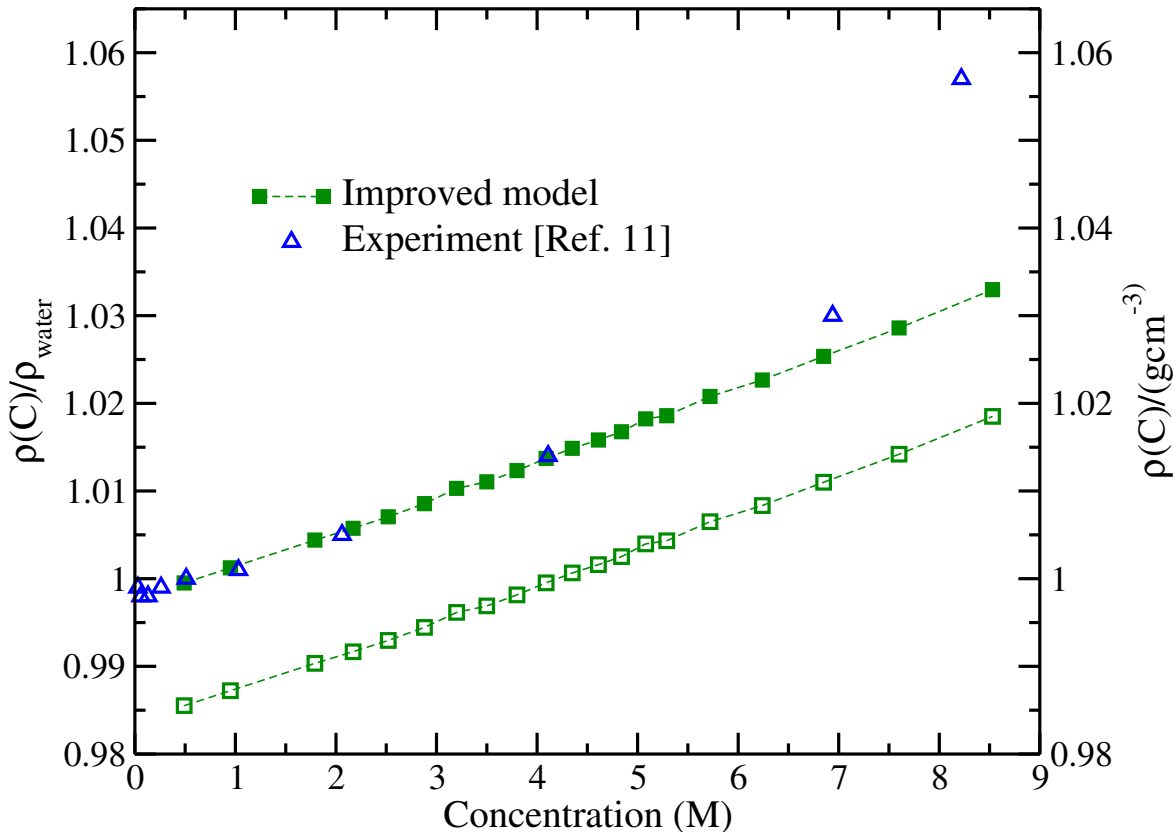

Figure 7: Density of aqueous imidazole solution as a function of molar concentration. Left and right y-scales are for filled and empty symbols, respectively.

If we divide the *improved* model densities and the by the corresponding values of the density of pure water, i.e. $\rho_{TIP3P} = 0.986$ (Figure 7, filled symbols), the ratio ρ/ρ_w from simulation is found to reproduce with remarkable accuracy the experimental density up to concentration of 4M (see Figure 7). At higher concentrations, the experimental ratio apparently ceases to behave linearly with the density ratio at C = 8 M noticeably diverging from the corresponding simulated value. The observed discrepancy at high imidazole concentration could be either due to force field deficiencies or to inaccuracies in the density

measurement. In this regard we point out that in Ref. 15, it was explicitly stated that "the experimental measurements of the densities have not been carried out that accurately, such as with a modern densitometer."

Pure imidazole ($C \simeq 16$ M)[15] is a molecular crystal in standard conditions. The imidazole crystal structure has been simulated with both the *improved* model and CHARMM force field, observing a similar agreement between experimental and computed lattice parameters,[76] as reported in Table 3, with exception of β angle, which shows a stronger deviation in the case of CHARMM force field. The imidazole-imidazole interactions through hydrogen bond are an important feature in the stabilisation of the crystal.[77] An attempt to simulate the crystal has been carried out also with the *standard* GAFF2 model, but the crystal turned out to be not stable during the simulation, very likely due to the incorrect orientational distribution of the N3-H1 hydrogen bonds.

Table 3: **Parameters for the simulation box, which is made up by** $5 \times 4 \times 2$ **conventional cells.**[76]

Model	a (Å)	b (Å)	c(Å)	α (°)	β (°)	γ (°)
exp.[76]	38.66	32.75	29.34	90.00	117.26	90.00
improved	40.90	30.41	29.96	90.00	117.00	90.00
CHARMM	39.91	30.05	29.41	89.99	112.68	90.01

As already done in Ref. 15 for the AMBER and EMM models we computed the relevant RDF's to characterise both the imidazole-water and imidazole-imidazole interactions. The analysis has been carried out on the trajectories of all the simulated samples reported in Table 2, but only the results for selected concentration values and for selected imidazole-imidazole and imidazole-water contacts are shown in Figure 8. In fact, imidazole-imidazole (Figure 8 a,b) and imidazole-water (Figure 8 c,d) RDF's do not show substantial variations increasing imidazole concentration, in agreement with NEXAFS findings[14] and with the MD simulation of Liem *et al.*[15]
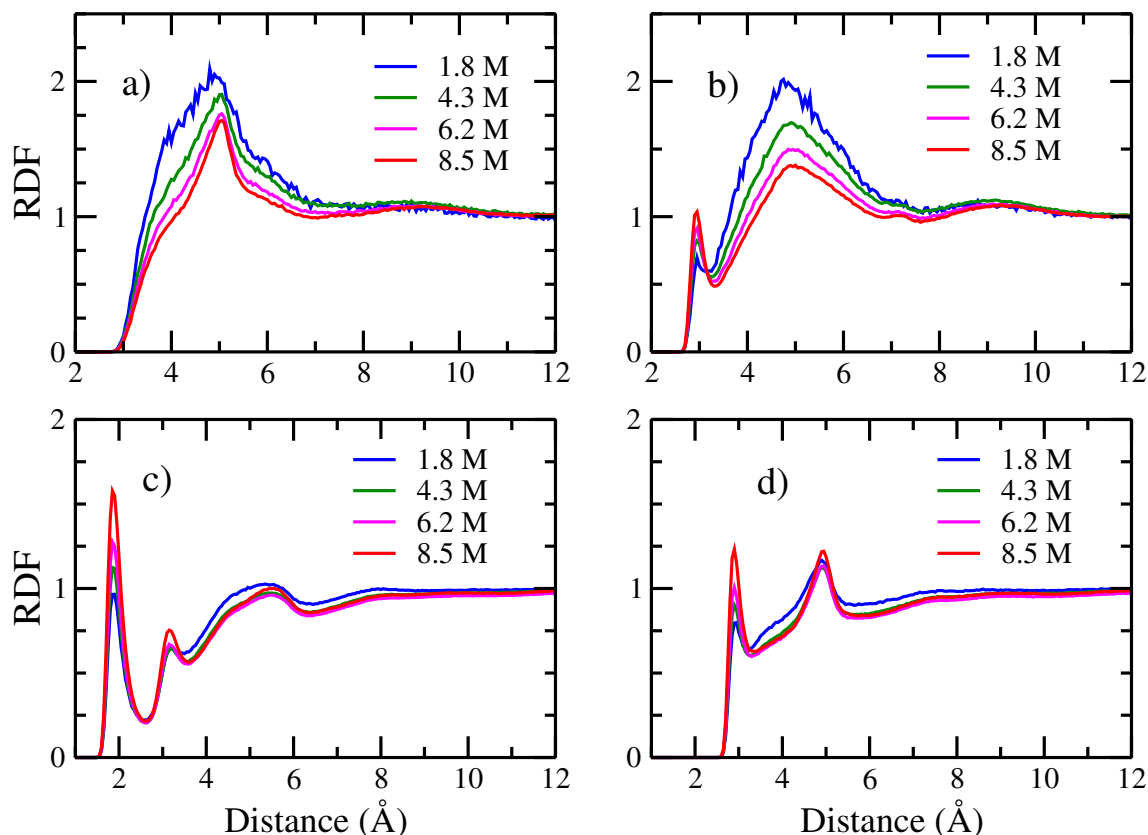
Figure 8: Radial distribution functions of imidazole aqueous solutions with increasing imidazole concentration. a) N1···N1; b) N1···N3; c) N3···HW; d) N1···OW

It should first be remarked that the N3···HW and N1···OW RDF's, related to the imidazole-water interactions, are similar to those obtained in the sample with only one imidazole molecule, as it can be appreciated comparing Figure 4 and Figure 8, confirming, as suggested in Ref. 14, that imidazole-imidazole association occurs already in dilute solutions.

In regard of the imidazole-imidazole interactions, the *improved* model N3···N1 and N3···N1 RDF's exhibit a trend that is consistent to those reported in Figures 3a, 3b of Ref. 15, referred to the EMM model. In particular, it indeed of note that, while the position of the major peaks is independent of concentration, their height steadily *decreases* in going from lower to higher imidazole concentration, with a trend that is similar for both N1···N1 and N3···N1 RDF's. The opposite behaviour is observed for the imidazole-water RDF's (Figure 8 c and Figure 8 d), where again the position of peaks is concentration invariant whereas their height *increases* with concentration. As suggested in Ref. 15, this observation

is due to the water-imidazole and imidazole-imidazole competition which, quite counter-intuitively, tends to favour imidazole-water hydrogen bond association with increasing ratio $N_{\text{imidazole}}/N_{\text{water}}$. It should be also noted that the imidazole-imidazole RDF's exhibit, for all reported concentrations, a weak but clearly discernible secondary peak at a distance of $\simeq 9$Å. Unfortunately, this feature went undetected in Ref. 15, where all the reported RDF's were truncated at $r = 8$Å, probably due to the small size of the MD sample used with the computationally expensive EMM approach. However, the signature of long range imidazole-imidazole association in concentrated imidazole samples observed with the improved model is consistent with the higher tendency of the EMM model to yield extended structured chains in pure liquid imidazole, a feature that the *standard AMBER approach is not able to reproduce.*[15] We conclude this section by reporting an analysis, obtained from the simulations using the *improved* model, of the incidence of stacked (S), T-shaped (T) and H-bonded (H) dimer structures in all examined imidazole solutions (see Table 2). The criteria for discerning S, T or H dimeric structures are based on the angle between the normal to the ring planes, on the distance between the centres of mass, and on the N3-H1 intermolecular distance, with tolerance ranges for distances and angles similar[78] to those given in Table 4 of Ref. 15. As done by Liem *et al.*,[15] the incidence of the dimeric structures is reported in Figure 9 in terms of normalised occurrence. The latter is obtained by dividing the total number of contacts matching the given criteria found in a sample of N_{conf} MD snapshots by the product of $N_{\text{imida}} \times N_{\text{conf}}$. The results of this analysis are shown in Figure 9.
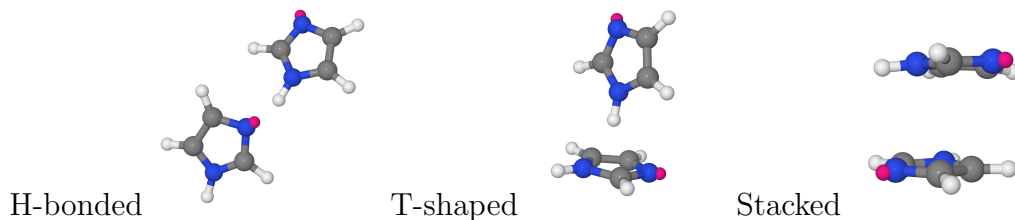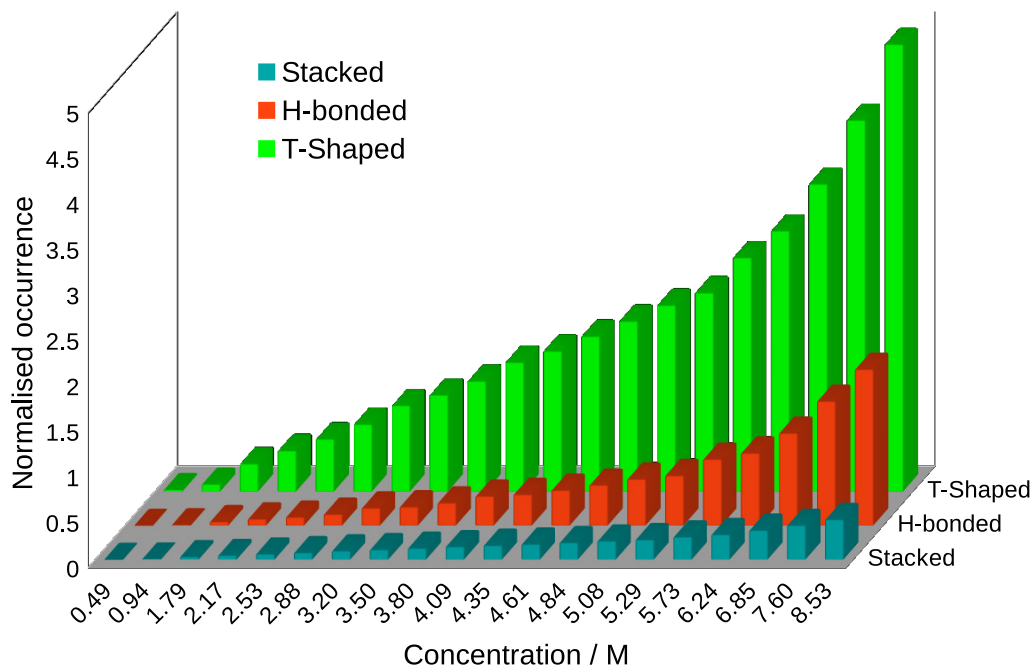
Figure 9: Normalised occurrence[15,78] of hydrogen bond, T-shaped and stacked configurations as a function of concentration. A graphical representation of the three different orientations is also reported. Atom colours: nitrogen, blue; carbon, gray; hydrogen, white. The small violet sphere represents the electron pair (i.e. the Pipek-Mezey centroid[5]).

Briefly, the analysis confirms the findings of Ref. 15 where S structures were found far less abundant with respect to hydrogen bonded structures, with the ratio N_H/N_S steadily increasing with increasing concentration. The most abundant dimeric structures, however, were of the T type, increasing super-linearly with concentration. In Ref. 15 the incidence of T-shaped structures, which in aromatic-aromatic interactions are stabilised mostly by electrostatic interactions,[79] were not investigated.

Conclusions

The hydrogen bond structure and dynamics of imidazole in water has been studied by *ab initio* molecular dynamics simulations with the Car-Parrinello method. It has been observed that van der Waals corrections to the BLYP exchange and correlation functional only slightly affect the hydrogen bond interactions. The CPMD simulations provide results similar to that reported by Dubou-Dijon *et al.*[7] employing the BLYP exchange and correlation functional and the D2 and D3(BJ) methods[42,43] to describe the dispersion part of the potential. The CPMD simulations have been adopted as reference to validate the developed force field for classical molecular dynamics simulations. The introduction in the *improved* model of a negatively charged Extra Point along the bisector of the CNC angle, accounting for the lone pair on the N3 nitrogen atom, allows to describe the imidazole water hydrogen bond interaction in solution in good agreement with the CPMD counterpart. Remarkably the *improved* model faithfully reproduces the spatial distribution function for the N3-OW pair interaction observed in the *ab initio* simulations, exhibiting the expected preferential orientation of the N3-OW hydrogen bond along the bisector of the CNC angle due to N3 sp_2 hybridisation.[7,13] To our knowledge, no other available classical potential model for imidazole, based on the fixed atomic charges paradigm, can produce a correct spatial orientation of the N3-OW directional hydrogen bond in the plane of the imidazole ring. As shown in Ref. 7 for the OPLS and CHARMM potential model, in Ref. 15 for the GAFF/AMBER potential model, and in the present study for the GAFF2/AMBER force field, all these standard models significantly overestimate the probability of the perpendicular H-π bond in water-imidazole interactions. The simple EP approach used in the improved model, when tested in a series of solutions of imidazole in water with increasing concentration up to 8.5 M, appears to be in good agreement with available experimental results and with existing data from previous simulation using the accurate EMM approach[15,18] for electrostatic interactions, confirming that the intermolecular interactions are accompanied by a variety of structural arrangements, the most important of which is of T-shaped type, characteristic of the quadrupole-quadrupole

interaction. The protocol for the parametrization of the GAFF2-based *improved* model for imidazole could be straightforwardly applied for the parametrization of other popular force field such as the OPLS or the CgenFF/CHARMM force fields and can be extended to other hetero-aromatic compounds with lone pairs. We finally mention that the satisfactory behaviour of our force field for imidazole-water and imidazole-imidazole association have been obtained with minimal differences with respect to a simple atom-based charge model and hence with a very limited cost, yielding computational efficiency that are comparable to those of the popular atomic charges models.

https://www.overleaf.com/project/5c612bf784fe0970e345ee10

Acknowledgement

Supporting Information Available

Topological and parameter les for improved model of imidazole (in ORAC[55] txt format). Compressed archive containing the atomic coordinates of aqueous imidazole solutions at various concentrations.

References

(1) Gilchrist, T. L. *Heterocyclic Chemistry*; Longman Scientific & Technical: Essex (UK), 1985.

(2) Eduok, U.; Faye, O.; Szpunar, J. Corrosion Inhibition of X70 Sheets by a Film-Forming Imidazole Derivative at Acidic pH. *RSC Adv.* **2016**, *6*, 108777–108790.

(3) Camacho-Mendoza, R. L.; Gutiérrez-Moreno, E.; Guzmán-Percástegui, E.; Aquino-Torres, E.; Cruz-Borbolla, J.; Rodríguez-Ávila, J. A.; Alvarado-Rodríguez, J. G.; Olvera-Neria, O.; Thangarasu, P.; Medina-Franco, J. L. Density Functional Theory and Electrochemical Studies: Structure-Efficiency Relationship on Corrosion Inhibition. *J. Chem. Inf. Model.* **2015**, *55*, 2391–2402.

(4) Narasimhan, B.; Sharma, D.; Kumar, P. Biological Importance of Imidazole Nucleus in the New Millennium. *Med. Chem. Res.* **2011**, *20*, 1119–1140.

(5) Pipek, J.; Mezey, P. G. A Fast Intrinsic Localization Procedure Applicable for *Ab Initio* and Semiempirical Linear Combination of Atomic Orbital Wave Functions. *J. Chem. Phys.* **1989**, *90*, 4916–4926.

(6) Meyer, F.; Blum, M.; Benkert, A.; Hauschild, D.; Jeyachandran, Y. L.; Wilks, R. G.; Yang, W.; Br, M.; Reinert, F.; Heske, C. et al. Site-specific electronic structure of imidazole and imidazolium in aqueous solutions. *Phys. Chem. Chem. Phys.* **2018**, *20*, 8302–8310.

(7) Dubou-Dijon, E.; Mason, P. E.; Fischer, H. E.; Jungwirth, P. Changes in the Hydration Structure of Imidazole upon Protonation: Neutron Scattering and Molecular Simulations. *J. Chem. Phys.* **2017**, *146*, 185102.

(8) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvenche, O.; Lopes, P.; Vorobyov, I. et al. CHARMM General Force Field: A Force Field for Drug-Like Molecules Compatible with the CHARMM All-Atom Additive Biological Force Fields. *J. Comput. Chem.* **2010**, *31*, 671–690.

(9) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.

(10) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. CP2K: Atomistic Simulations of Condensed Matter Systems. *WIREs Comput. Mol. Sci.* **2014**, *4*, 15–25.

(11) Jagoda-Cwiklik, B.; Slavek, P.; Cwiklik, L.; Nolting, D.; Winter, B.; Jungwirth, P. Ionization of Imidazole in the Gas Phase, Microhydrated Environments, and in Aqueous Solution. *The Journal of Physical Chemistry A* **2008**, *112*, 3499–3505.

(12) Nolting, D.; Ottosson, N.; Faubel, M.; Hertel, I. V.; Winter, B. Pseudoequivalent Nitrogen Atoms in Aqueous Imidazole Distinguished by Chemical Shifts in Photoelectron Spectroscopy. *J. Am. Chem. Soc.* **2008**, *130*, 8150–8151.

(13) Al-Madhagi, L. H.; Chang, S.-Y.; Balasubramanian, M.; Kroner, A. B.; Shotton, E. J.; Willneff, E. A.; Mishra, B.; Schroeder, S. L. M. X-ray Raman Scattering: S New *in Situ* Probe of Molecular Structure During Nucleation and Crystallization from Liquid Solutions. *CrystEngComm* **2018**, *20*, 6871–6884.

(14) Thomason, M. J.; Seabourne, C. R.; Sattelle, B. M.; Hembury, G. A.; Stevens, J. S.; Scott, A. J.; Aziz, E. F.; Schroeder, S. L. M. Self-Association of Organic Solutes in Solution: a NEXAFS Study of Aqueous Imidazole. *Faraday Discuss.* **2015**, *179*, 269–289.

(15) Liem, S. Y.; Shaik, M. S.; Popelier, P. L. A. Aqueous Imidazole Solutions: A Structural Perspective from Simulations with High-Rank Electrostatic Multipole Moments. *J. Phys. Chem. B* **2011**, *115*, 11389–11398.

(16) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General AMBER Force Field. *J. Comput. Chem.* **2004**, *25*, 1157–1174.

(17) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.

(18) Cardamone, S.; Hughes, T. J.; Popelier, P. L. A. Multipolar Electrostatics. *Phys. Chem. Chem. Phys.* **2014**, *16*, 10367–10387.

(19) Liem, S. Y.; Popelier, P. L. A.; Leslie, M. Simulation of Liquid Water Using a High-Rank Quantum Topological Electrostatic Potential. *Int. J. Quantum Chem.* **2004**, *99*, 685–694.

(20) Kolář, M. H.; Hobza, P. Computer Modeling of Halogen Bonds and Other σ-Hole Interactions. *Chem. Rev.* **2016**, *116*, 5155–5187, PMID: 26840433.

(21) Macchiagodena, M.; Mancini, G.; Pagliai, M.; Frate, G. D.; Barone, V. Fine-Tuning of Atomic Point Charges: Classical Simulations of Pyridine in Different Environments. *Chem. Phys. Lett.* **2017**, *677*, 120–126.

(22) Macchiagodena, M.; Mancini, G.; Pagliai, M.; Cardini, G.; Barone, V. New Atomistic Model of Pyrrole with Improved Liquid State Properties and Structure. *Int. J. Quantum Chem.* **2018**, *118*, e25554.

(23) Marx, D.; Hutter, J. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods*; Cambridge University Press: Cambridge (UK), 2009.

(24) Car, R.; Parrinello, M. Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.

(25) Parrinello, M. From Silicon to RNA: the Coming of Age of *Ab Initio* Molecular Dynamics. *Solid State Commun.* **1997**, *102*, 107–120.

(26) Laasonen, K.; Sprik, M.; Parrinello, M.; Car, R. "Ab Initio" Liquid Water. *J. Chem. Phys.* **1993**, *99*, 9080–9089.

(27) Muniz-Miranda, M.; Pagliai, M.; Muniz-Miranda, F.; Schettino, V. Raman and Computational Study of Solvation and Chemisorption of Thiazole in Silver Hydrosol. *Chem. Commun.* **2011**, *47*, 3138–3140.

(28) Candelaresi, M.; Pagliai, M.; Lima, M.; Righini, R. Chemical Equilibrium Probed by Two-Dimensional IR Spectroscopy: Hydrogen Bond Dynamics of Methyl Acetate in Water. *J. Phys. Chem. A* **2009**, *113*, 12783–12790.

(29) Pagliai, M.; Cardini, G.; Righini, R.; Schettino, V. Hydrogen Bond Dynamics in Liquid Methanol. *J. Chem. Phys.* **2003**, *119*, 6655–6662.

(30) Kuo, I.-F. W.; Mundy, C. J.; McGrath, M. J.; Siepmann, J. I.; VandeVondele, J.; Sprik, M.; Hutter, J.; Chen, B.; Klein, M. L.; Mohamed, F. et al. Liquid Water from First Principles: Investigation of Different Sampling Approaches. *J. Phys. Chem. B* **2004**, *108*, 12990–12998.

(31) GAFF and GAFF2 are public domain force fields and are part of the AmberTools16 distribution, available for download at http://amber.org internet address (accessed March 2017). According to the AMBER development team, the improved version of GAFF, GAFF2, is an ongoing poject aimed at "reproducing both the high quality interaction energies and key liquid properties such as density, heat of vaporization and hydration free energy". GAFF2 is expected "to be an even more successful general purpose force field and that GAFF2-based scoring functions will significantly improve the successful rate of virtual screenings.".

(32) Procacci, P. PrimaDORAC: A Free Web Interface for the Assignment of Partial Charges, Chemical Topology, and Bonded Parameters in Organic or Drug Molecules. *J. Chem. Inf. Model.* **2017**, *57*, 1240–1245.

(33) Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102*, 1995–2001.

(34) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, Structures, and Electronic

Properties of Molecules in Solution with the C-PCM Solvation Model. *J. Comput. Chem.* **2003**, *24*, 669–681.

(35) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. The Statistical-Thermodynamic Basis for Computation of Binding Affinities: A Critical Review. *Biophys. J.* **1997**, *72*, 1047–1069.

(36) Procacci, P.; Chelli, R. Statistical Mechanics of Ligand-Receptor Noncovalent Association, Revisited: Binding Site and Standard State Volumes in Modern Alchemical Theories. *J. Chem. Theory Comput.* **2017**, *13*, 1924–1933.

(37) Procacci, P.; Cardelli, C. Fast Switching Alchemical Transformations in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2014**, *10*, 2813–2823.

(38) CPMD, Copyright IBM Corp 1990-2008, Copyright MPI für Festkörperforschung Stuttgart 1997-2001. http://www.cpmd.org, version 3.15.1.

(39) Becke, A. D. Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098–3100.

(40) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785–789.

(41) Grimme, S. Accurate Description of van der Waals Complexes by Density Functional Theory Including Empirical Corrections. *J. Comput. Chem.* **2004**, *25*, 1463–1473.

(42) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

(43) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate *Ab Initio* Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(44) Schmidt, J.; VandeVondele, J.; Kuo, I.-F. W.; Sebastiani, D.; Siepmann, J. I.; Hutter, J.; C. J, M. Isobaric-Isothermal Molecular Dynamics Simulations Utilizing Density Functional Theory: An Assessment of the Structure and Density of Water at Near-Ambient Conditions. *J. Phys. Chem. B* **2009**, *113*, 11959–11964.

(45) N. Troullier and J. L. Martins, Efficient Pseudopotential for Plane-Wave Calculations. *Phys. Rev. B* **1991**, *43*, 1993–2006.

(46) L. Kleinman and D. M. Bylander, Efficacious Form for Model Pseudopotentials. *Phys. Rev. Lett.* **1982**, *48*, 1425–1428.

(47) Araz, J.; B., J. D.; I., B. C. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comput. Chem.* **2002**, *23*, 1623–1641.

(48) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. Development and Use of Quantum Mechanical Molecular Models. 76. AM1: a New General Purpose Quantum Mechanical Molecular Model. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.

(49) Pagliai, M.; Mancini, G.; Carnimeo, I.; De Mitri, N.; Barone, V. Electronic Absorption Spectra of Pyridine and Nicotine in Aqueous Solution with a Combined Molecular Dynamics and Polarizable QM/MM Approach. *J. Comput. Chem.* **2017**, *38*, 319–335.

(50) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A. et al. *Gaussian 09, Revision C.01*; Gaussian, Inc., Wallingford CT (USA), 2010.

(51) Marsili, S.; Signorini, G. F.; Chelli, R.; Marchi, M.; Procacci, P. ORAC: A Molecular Dynamics Simulation Program to Explore Free Energy Surfaces in Biomolecular Systems at the Atomistic Level. *J. Comput. Chem.* **2010**, *31*, 1106–1116.

(52) Tuckerman, M.; Berne, B. J. Reversible Multiple Time Scale Molecular Dynamics. *J. Chem. Phys.* **1992**, *97*, 1990–2001.

(53) Procacci, P.; Paci, E.; Darden, T.; Marchi, M. ORAC: A Molecular Dynamics Program to Simulate Complex Molecular Systems with Realistic Electrostatic Interactions. *J. Comput. Chem.* **1997**, *18*, 1848–1862.

(54) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577–8593.

(55) Procacci, P. Hybrid MPI/OpenMP Implementation of the ORAC Molecular Dynamics Program for Generalized Ensemble and Fast Switching Alchemical Simulations. *J. Chem. Inf. Model.* **2016**, *56*, 1117–1121.

(56) Pagliai, M.; Muniz-Miranda, F.; Cardini, G.; Righini, R.; Schettino, V. Hydrogen Bond Dynamics of Methyl Acetate in Methanol. *J. Phys. Chem. Lett.* **2010**, *1*, 2951–2955.

(57) Sterpone, F.; Stirnemann, G.; Hynes, J. T.; Laage, D. Water Hydrogen-Bond Dynamics Around Amino Acids: The Key Role of Hydrophilic Hydrogen-Bond Acceptor Groups. *J. Phys. Chem. B* **2010**, *114*, 2083–2089.

(58) Brehm, M.; Kirchner, B. TRAVIS - A free Analyzer and Visualizer for Monte Carlo and Molecular Dynamics Trajectories. *J. Chem. Inf. Model.* **2011**, *51*, 2007–2023.

(59) Wishart, D.; Feunang, Y.; Guo, A.; Lo, E.; Marcu, A.; Grant, J.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z. et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. *Nucl. Acids Res.* **2018**, *4*, D1074–D1082.

(60) Jämbeck, J. P. M.; Lyubartsev, A. P. An Extension and Further Validation of an All-Atomistic Force Field for Biological Membranes. *J. Chem. Theory Comput.* **2012**, *8*, 2938–2948.

(61) Sandberg, R. B.; Banchelli, M.; Guardiani, C.; Menichetti, S.; Caminati, G.; Procacci, P. Efficient Nonequilibrium Method for Binding Free Energy Calculations in Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2015**, *11*, 423–435.

(62) Procacci, P. I. Dissociation Free Energies of Drug-Receptor Systems via Non-Equilibrium Alchemical Simulations: a Theoretical Framework. *Phys. Chem. Chem. Phys.* **2016**, *18*, 14991–15004.

(63) Nerattini, F.; Chelli, R.; Procacci, P. II. Dissociation Free Energies in Drug-Receptor Systems via Nonequilibrium Alchemical Simulations: Application to the FK506-Related Immunophilin Ligands. *Phys. Chem. Chem. Phys.* **2016**, *18*, 15005–15018.

(64) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141–151.

(65) Cardelli, C.; Barducci, A.; Procacci, P. Lipid Tempering Simulation of Model Biological Membranes on Parallel Platforms. *Biochim. Biophys. Acta, Biomembr.* **2018**, *1860*, 1480 – 1488.

(66) Anderson, T. W.; Darling, D. A. A Test of Goodness of Fit. *J. Am. Stat. Ass.* **1954**, *49*, 765–769.

(67) Razali, N. M.; Wah, Y. B. Power Comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling Tests. *J. Stat. Model. Anal.* **2011**, *2*, 21–33.

(68) Crooks, G. E. Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems. *J. Stat. Phys.* **1998**, *90*, 1481–1487.

(69) Procacci, P.; Guarrasi, M.; Guarnieri, G. SAMPL6 Host-Guest Blind Predictions Using a Non Equilibrium Alchemical Approach. *J. Comput. Aided Mol. Des.* **2018**, *32*, 965–982.

(70) Dempster, A.; Laird, N.; Rubin, D. Maximum Likelihood from Incomplete Data via the EM Algorithm. *J. Royal Stat. Soc. B* **1977**, *39*, 1–38.

(71) Gupta, M. R.; Chen, Y. Theory and Use of the EM Algorithm. *Found. Trends Signal Process.* **2011**, *4*, 223–296.

(72) DrugBank, Record Name: Imidazole, Description: The DrugBank database is a unique bioinformatics and cheminformatics resource that combines detailed drug (i.e. chemical, pharmacological and pharmaceutical) data with comprehensive drug target (i.e. sequence, structure, and pathway) information. http://www.drugbank.ca/drugs/DB03366, access date: 13/02/2019.

(73) Chelli, R.; Ciabatti, S.; Cardini, G.; Righini, R.; Procacci, P. Calculation of Optical Spectra in Liquid Methanol Using Molecular Dynamics and the Chemical Potential Equalization Method. *J. Chem. Phys.* **1999**, *111*, 4218–4229.

(74) Chelli, R.; Pagliai, M.; Procacci, P.; Cardini, G.; Schettino, V. Polarization Response of Water and Methanol Investigated by a Polarizable Force Field and Density Functional Theory Calculations: Implications for Charge Transfer. *J. Chem. Phys.* **2005**, *122*, 074504.

(75) Izadi, S.; Onufriev, A. V. Accuracy limit of rigid 3-point water models. *J. Chem. Phys.* **2016**, *145*, 074501.

(76) Martinez-Carrera, S. The Crystal Structure of Imidazole at -150°C. *Acta Cryst.* **1966**, *20*, 783–789.

(77) Sharma, B. B.; Verma, A. K.; Thomas, S.; Murli, C.; Sharma, S. M. Hydrogen Bonds and Ionic Forms versus Polymerization of Imidazole at High Pressures. *J. Phys. Chem. B* **2015**, *119*, 372–378.

(78) H-bonded (H) criterium: $r(\text{N3} \cdots \text{H1}) \leq 2.2$ Å; Stacked (S) criterium: $r_{\text{COM}\cdots\text{COM}} \leq 4.0$, $|\theta| \leq 25$ deg. T-shaped (T) criterium: $r_{\text{COM}\cdots\text{COM}} \leq 5.6$ Å, $|\theta| - 90 \leq 15$ deg..

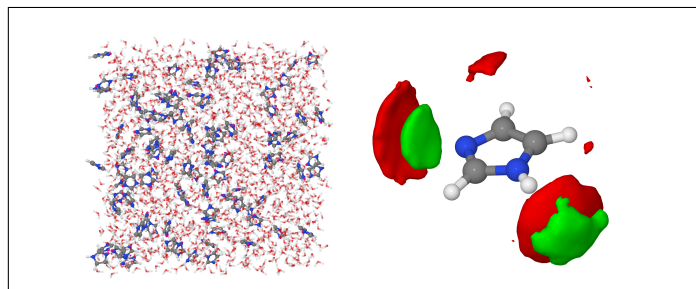(79) Procacci, P. Thermodynamics of Stacking in Proteins. *Annu. Rep. Progr. Chem. Sect. C* **2011**, *107*, 242–262.

Graphical TOC Entry

Table of Contents Graphic