

# Simulation of liquid imidazole using a high-rank quantum topological electrostatic potential†

Majeed S. Shaik,<sup>ab</sup> Steven Y. Liem,<sup>ab</sup> Yongna Yuan<sup>ab</sup> and Paul L. A. Popelier<sup>\*ab</sup>

Received 1st May 2010, Accepted 16th September 2010

DOI: 10.1039/c0cp00417k

Rigid body molecular dynamics simulations were carried out on pure liquid imidazole at four different temperatures and at 1 atm. Imidazole, which is important both in life science and materials science, is one of the simplest molecules to possess both a lone pair and a  $\pi$  system. These two features are known to benefit from multipolar electrostatics. Here the electrostatic interaction is governed by atomic multipole moments obtained from topologically partitioned *ab initio* electron densities. The non-electrostatic terms are modeled with Lennard-Jones parameters adjusted to fit the experimental liquid density. All  $\sigma$  values are incrementally increased by one single scaling factor. We report on how the presence of multipolar electrostatics influences the local structure, dynamics and thermodynamics of the liquid compared to electrostatics by atomic point charges. The point charge force field exaggerates the number of  $\pi$ -stacked dimers in the liquid, and underestimates the number of hydrogen-bonded dimers. The effect of the temperature on the local structure of liquid imidazole was analysed using radial and spatial distribution functions.

## 1. Introduction

Knowledge of the intermolecular interactions, structure and behaviour of organic systems is important to understand the microstructure and properties of materials. Various computational studies<sup>1–7</sup> have probed intermolecular interactions and the structures they generate, for both liquids and crystals. The last decade or so<sup>8–20</sup> has seen a growth in efforts to develop methods with an increased accuracy and completeness in describing intermolecular interactions. For example, proton transport has been simulated in liquid using a multistate empirical valence bond approach allowing proton transfer to simultaneously occur on both donor and acceptor reaction sites of imidazole.<sup>21</sup> Molecular simulation that uses electrostatic potentials incorporating multipole moments is one route to examine molecular systems with increased accuracy, compared to the traditional point charge models. Here we contrast the liquid structure of imidazole obtained from a Molecular Dynamics (MD) simulation using point charges with one using high rank multipole moments.

Imidazole ( $C_3N_2H_4$ ) is a five-membered heterocyclic aromatic compound of relevance to both biochemistry and materials sciences. Azoles are an interesting class of compounds with two different structural characteristics. They possess a  $\pi$ -system and at least one lone pair as well. Both features are known<sup>22</sup> to benefit from multipole moments in order to be accurately described at short range. The  $\pi$ -system and the hydrogen-bond donor and acceptor groups are the basis for

the biological behaviour of nucleotides and enzymes.<sup>23</sup> Much attention, experimentally as well as theoretically, has been directed toward the structure and vibrational spectra of imidazole, due to its small size and rigidity.<sup>24</sup> Imidazole is found in many important biological systems. Its most prominent appearance is undoubtedly in the amino acid histidine (and the related hormone histamine). Histidine is present in a large number of proteins and enzymes, and plays a fundamental role in the binding behaviour of haemoglobin,<sup>25</sup> as well as in the structure of purine nucleobases<sup>23</sup> and several other naturally occurring compounds. Imidazole can be found in drugs, such as anti-fungal drugs,<sup>26</sup> and features in many industrial and technological processes. For example, imidazole featured for a long time as a corrosion inhibitor on metals like copper.<sup>27</sup> It has applications as a fire retardant and appears in many compounds that are used for photography<sup>28</sup> and electronics. Finally, imidazole appears in ionized form in ionic liquids, for example, in the cation 1-ethyl-3-methylimidazolium ([EMIM]).<sup>29</sup> Ionic liquids are novel solvents composed of organic cations and inorganic anions.<sup>30</sup> Their negligible vapour pressure makes them promising replacements for volatile organic solvents in industrial processes, which is why they are regarded as environmentally friendly. Simulations on ionic liquids are dominated by the use of point charges.<sup>31</sup>

There are a number of ways to represent electrostatic interactions by means of multipole moments<sup>32–36</sup> (for more details see ref. 37). One possible avenue is Quantum Chemical Topology (QCT),<sup>38–40</sup> which offers an unambiguous way of partitioning an electron density into chemically meaningful<sup>41,42</sup> atomic fragments. Note that QCT is more general than the Quantum Theory of Atoms in Molecules as explained in ref. 43 and 44. The central idea of QCT is to use the gradient vector field of a quantum mechanical function to obtain local (*i.e.* point) and global (*i.e.* space partitioned) properties. Unambiguous atomic multipole moments can be calculated

<sup>a</sup> Manchester Interdisciplinary Biocentre (MIB), 131 Princess Street, Univ. of Manchester, Manchester M1 7DN, UK.

E-mail: pla@manchester.ac.uk

<sup>b</sup> School of Chemistry, Univ. of Manchester, Oxford Road, Manchester M13 9PL, UK

† Electronic supplementary information (ESI) available: Fig. S1, Tables S1–S3. See DOI: 10.1039/c0cp00417k

by means of 3D integration over the volume of the QCT atoms.<sup>45</sup> In view of their compactness, we employ spherical tensor harmonics<sup>46,47</sup> to express the moments, instead of Cartesian tensors. The first step towards a QCT force field<sup>43</sup> is to represent the electrostatic part of the potential realistically at short range. Multipole moments are good at achieving this, provided that the multipole expansion converges.<sup>48,49</sup> QCT multipole moments have been used to explore potential energy surfaces<sup>50,51</sup> with analytical first and second derivatives.<sup>52</sup> Distributed polarisabilities of topological atoms can also be obtained as demonstrated for the water dimer.<sup>53</sup> They have also been used in molecular simulations.<sup>4,54,55</sup> Extensive comparisons were made<sup>56</sup> of local energy minima, generated at *ab initio* level, by multipole moments and point charges for pure water clusters, and hydrated serine and tyrosine clusters. High rank multipole models (including quadrupole–quadrupole interactions) performed consistently better than point charge models in predicting the geometries. In the current study, MD simulations were carried out on pure liquid imidazole, at four different temperatures. The focus of this study is to reveal the difference in structural information as predicted from QCT multipole moments as opposed to point charges. This work serves as a precursor for the study of the self-association of imidazole in aqueous solution, which will be reported on elsewhere.

## 2. Background and method

### 2.1 The QCT potential

A single imidazole molecule was geometry optimised at nine different levels of theory, exhaustively combining three correlation methods (B3LYP, MP2, CCSD) and three basis sets (6-311+G(2d,p), aug-cc-pVDZ and aug-cc-pVTZ) using GAUSSIAN03.<sup>57</sup> Table 1 shows the molecular dipole moment at each level of theory, and compares it to experiment. An experimental value of  $3.8 \pm 0.4$  D, determined by Stark measurements, was reported<sup>58</sup> in 1967. A more recent experimental value of  $3.66 \pm 0.06$  D was published<sup>59</sup> in 1982 and mentioned in a review<sup>60</sup> on molecular electric moments. Averaging the discrepancy with experiment over the three basis sets reveals that MP2 performs worse than both B3LYP and CCSD. B3LYP performs remarkably well, with an average deviation of experiment of 2.6%. We decided to

represent the electrostatic interaction with a monomeric wave function obtained at CCSD/aug-cc-pVTZ//CCSD/aug-cc-pVDZ level.

Using the default parameters, the program MORPHY98<sup>61,62</sup> computed the multipole moments (up to hexadecapole) centered on the nucleus of each atom. This was achieved by volume integration over each atomic basin  $\Omega$ . An atomic basin is a portion of 3D space that is crossed by gradient vector trajectories terminating at the atom's nucleus. As shown in eqn (1), multipole moments are then obtained by integrating the appropriate spherical tensor  $C_{\ell m}(\theta, \varphi)$ , weighted by the total charge density  $\rho_{\text{tot}}(\mathbf{r})$ ,

$$Q_{\ell 0} = \int_{\Omega} d\mathbf{r} \rho_{\text{tot}}(\mathbf{r}) r^{\ell} C_{\ell 0}(\theta, \varphi)$$

$$Q_{\ell m c} = \int_{\Omega} d\mathbf{r} \rho_{\text{tot}}(\mathbf{r}) r^{\ell} C_{\ell m c}(\theta, \varphi)$$

$$Q_{\ell m s} = \int_{\Omega} d\mathbf{r} \rho_{\text{tot}}(\mathbf{r}) r^{\ell} C_{\ell m s}(\theta, \varphi) \quad (1)$$

The total charge density is given by  $\rho_{\text{tot}}(\mathbf{r}) = \sum_{\text{A}} Z_{\text{A}} \delta(\mathbf{r} - \mathbf{R}_{\text{A}}) - \rho(\mathbf{r})$ , where  $Z_{\text{A}}$  is the charge of nucleus A located at position  $\mathbf{R}_{\text{A}}$ . The functions  $C_{\ell 0}(\theta, \varphi)$ ,  $C_{\ell m c}(\theta, \varphi)$ , and  $C_{\ell m s}(\theta, \varphi)$  are real normalized spherical harmonics of rank  $\ell$ , and depend on the usual angular coordinates  $\theta$  and  $\varphi$ . At rank  $\ell = 0$  one obtains the atomic charge  $Q_{00}$ . There are three dipole moments ( $\ell = 1$ ), denoted by  $Q_{10}$ ,  $Q_{11c}$  and  $Q_{11s}$ , which correspond to the Cartesian components  $\mu_z$ ,  $\mu_x$  and  $\mu_y$ , respectively. Instead of the six familiar Cartesian quadrupole moments, there are only five quadrupole moments ( $\ell = 2$ ) in the spherical-tensor formalism we use here, and they are denoted by  $Q_{20}$ ,  $Q_{21c}$ ,  $Q_{21s}$ ,  $Q_{22c}$  and  $Q_{22s}$ . Similarly, there are seven octopole moments ( $\ell = 3$ ) and nine hexadecapole moments ( $\ell = 4$ ).

The total potential energy ( $E_{\text{p}}$ ) between any two atoms A and B has two contributions,

$$E_{\text{p}} = E_{\text{elec}} + 4\epsilon \left[ \left( \frac{\sigma}{r_{\text{AB}}} \right)^{12} - \left( \frac{\sigma}{r_{\text{AB}}} \right)^6 \right] \quad (2)$$

**Table 1** A comparison with experiment of the gas phase dipole moment of imidazole with that calculated at all levels of theory

Method/Basis set <sup>a</sup>	Dipole moment (D)	% Difference with experiment
B3LYP/6-311+G(2d,p)	3.77	3.0
B3LYP/aug-cc-pVDZ	3.73	1.6
B3LYP/aug-cc-pVTZ	3.78	3.3
MP2/6-311+G(2d,p)	3.27	-11.0
MP2/aug-cc-pVDZ	3.38	-7.7
MP2/aug-cc-pVTZ	3.82	4.4
CCSD/6-311+G(2d,p)	3.81	4.1
CCSD/aug-cc-pVDZ	3.76	2.7
CCSD/aug-cc-pVTZ <sup>b</sup>	3.76	2.7
Experiment <sup>59</sup>	$3.66 \pm 0.06$	

<sup>a</sup> This is an abbreviation for the full notation, typically written as “Method A/Basis set A//Method B/Basis set B” in which the same level of theory generates both the electron density and the optimized geometry. <sup>b</sup> Although the electron density was obtained at this level, the geometry was optimized at a slightly lower level, to save computer time, which is CCSD/aug-cc-pVDZ.

where  $E_{\text{elec}}$  corresponds to the electrostatic interaction energy between two atoms, and  $r_{\text{AB}}$  is the distance between their respective nuclei. The other contribution is the Lennard-Jones(LJ) term, which introduces the parameters  $\epsilon_{\text{AB}}$  and  $\sigma_{\text{AB}}$ , corresponding to the well depth and collision diameter, respectively. These are the only two parameters in  $E_{\text{pot}}$ , and they are fitted to the density of liquid imidazole at 1 atm and 383 K (112 °C), which is 20 K above its melting point of approximately 363 K (90 °C). Note that imidazole's boiling point of 536 K (263 °C) is anomalously high. The  $E_{\text{elec}}$  term can be evaluated from eqn (3),

$$E_{\text{elec}} = \frac{1}{2} \sum_A \sum_{B \neq A} E_{\text{elec}}(\text{A}, \text{B}) \quad (3)$$

where A and B are any two atoms, each belonging to different imidazole molecules. Details on the precise way in which  $E_{\text{elec}}(\text{A}, \text{B})$  is calculated have been given before, for example, in the Appendix of ref. 50. Here we just reiterate that the underpinning multipole expansion is

$$E_{\text{elec}}(\text{A}, \text{B}) = \sum_{\ell_A \ell_B m_A m_B} Q_{\ell_A m_A}(\Omega_A) T_{\ell_A m_A \ell_B m_B}(\mathbf{R}_{\text{AB}}) Q_{\ell_B m_B}(\Omega_B) \quad (4)$$

where  $T_{\ell_A m_A \ell_B m_B}$  is a purely geometric interaction tensor, which is dependent on the distance between the two atoms and the relative orientation of their local frames. In eqn (4),  $\mathbf{R}_{\text{AB}}$  is a vector connecting the two nuclear positions and  $Q_{\ell m}$  are the  $2\ell + 1$  multipole moments of rank  $\ell$  with respect to the local frames of each atom.

The electrostatic interaction is usually associated by an overall rank  $L$ , which is defined as:

$$L = \ell_A + \ell_B + 1 \quad (5)$$

where  $\ell_A$  and  $\ell_B$  represent the rank of multipole moments on atoms A and B, respectively. The interaction between two atomic quadrupole moments, for example, is associated with an overall rank of  $L = 2 + 2 + 1 = 5$ . The convergence of the electrostatic interaction energy between two atoms is typically monitored by  $L$ . In the current study,  $L$  is set to 5, because of observations made on liquid water simulations,<sup>4</sup> carried out in our lab some years ago. We showed, using the oxygen–oxygen radial distribution function (RDF) at 1 atm and 300 K, that  $L = 5$  is essential to recover the typical features of a liquid-like structure, such as the mandatory appearance of a second peak, right of the first and major peak. In spite of this observation it is important that the full convergence behaviour is properly understood. This is why we followed the convergence behaviour, up to very high rank  $L$ , of selected atom–atom interactions in the imidazole dimer. We also made comparisons with exact *non-multipolar* electrostatic energies. The exact non-multipolar interaction energy can be obtained *via* a six-dimensional (6D) integration over the two participating atoms,

$$E^{\text{AB}} = \int_{\Omega_A} d\mathbf{r}_A \int_{\Omega_B} d\mathbf{r}_B \frac{\rho_{\text{tot}}(\mathbf{r}_A) \rho_{\text{tot}}(\mathbf{r}_B)}{r_{\text{AB}}} \quad (6)$$

where  $r_{\text{AB}}$  is the distance between two infinitesimally small charge elements and  $\rho_{\text{tot}}$  is the total charge density. All atoms of both the imidazole monomer and dimer (supermolecule) were integrated with MORPHY01, which was also used to

perform the 6D integrations. The quadrature inside the  $\beta$ -sphere was set to  $(n_r, n_\theta, n_\phi) = (90, 30, 50)$ , which corresponds to the number of radial and angular quadrature points, respectively. For the rest of the atom, we used  $(n_r, n_\theta, n_\phi) = (70, 53, 87)$ . The atomic volumes were capped at the  $\rho = 10^{-6}$  a.u. isodensity envelope. The 6D integrations were carried out between a pair of atoms in a supermolecule (*i.e.* a single wave function), and not between atoms in different monomers (*i.e.* each monomer being described by its own wave function). The quadrature grid of the 6D integration (over a pair of atoms) is the same as that of the 3D integration over single atoms.

The multipolar electrostatic energy was computed up to rank  $L = 20$ . We define a multipole expansion to be convergent at rank  $L$  if the difference between the expansion at rank  $L$  and at rank  $L - 1$ , as well as the difference between rank  $L$  and rank  $L - 2$  are both less than  $0.1 \text{ kJ mol}^{-1}$ . This value embodies a rather severe convergence criterion. This value also represents the accuracy of the converged multipole expansion. If the criterion is satisfied then rank  $L$  energy can be regarded as the “exact” multipolar reference energy. The exact energy obtained by 6D integration is computationally very expensive and can therefore be obtained only sparingly. The issue of convergence is investigated in Section 3.3, which contains the results.

## 2.2 Simulation details

We carried out MD simulations on pure liquid imidazole using multipole moments of topological atoms (the “QCT” model) on one hand, and using traditional point charges (“AMBER”) on the other hand. We discuss the details of each type in turn.

The multipolar simulations were performed by the program DLMULTI<sup>63</sup> version 3.0, which is a heavily modified version of the DLPOLY code.<sup>64,65</sup> A major feature of DLMULTI is the implementation of an Ewald summation generalized for high-rank multipole moments, so that the long ranged contribution of the electrostatic interactions is accounted for. All simulations were carried out using an isothermal–isobaric ensemble ( $NpT$ ) using the Berendsen thermostat,<sup>66</sup> and barostat,<sup>67</sup> always at a pressure of 1 atm. A first batch of simulations (at 383 K only, 20 K above melting point) focuses on the influence of the level of theory on the electrostatic potential. A second batch fixes the level of theory but optimizes the LJ parameters, against the experimental density of the liquid. The third and final batch is performed at four different temperatures, with a fixed potential, both in terms of electrostatic and LJ parameters. In all simulations, a cubic primary cell with full periodic conditions containing 125 ( $= 5^3$ ) imidazole molecules was used. Preliminary simulations tackled the influence of the system size, investigating a larger system with  $6^3 = 216$  molecules. Already between a 125- and 216-molecule system, there is only a  $0.1 \text{ kJ mol}^{-1}$  difference in the potential energy and the RDFs (or  $g(r)$  profiles) are visually identical. In all simulations, the integration time-step used was set to 0.5 fs. For the first and second batch, the system was simulated with canonical ensemble (NVT) for 150 ps and then equilibrated with  $NpT$  ensemble for 100 ps and, once equilibrated, the simulations with  $NpT$  ensemble

were continued for 1 ns. In preparation of the third batch, the system was simulated by an NVT ensemble for 150 ps and then equilibrated for 250 ps as an  $NpT$  ensemble. Finally,  $NpT$  simulations were carried out for at least 2 ns to collect data for analysis. For all QCT based simulations, the relaxation time,  $\tau$ , for the thermostat and barostat is set to be 0.1 ps and 2.0 ps, respectively.

All simulations in the first batch used LJ parameters from the AMBER ff94<sup>68</sup> parameter set. The LJ parameters for each atom type in imidazole are given in Table 2. The van der Waals parameters for interactions between different atom types were obtained from the Lorentz–Berthelot mixing rule.<sup>69</sup>

For the point charge based simulations, the all-atom AMBER ff99SB force field<sup>70</sup> was used. MD simulations are carried out using periodic boundary conditions and a cubic simulation cell. Electrostatic interactions are calculated with the Ewald particle-mesh method.<sup>71</sup> We used a cut-off of 8 Å for the real-space direct sum part of the Ewald summation and for the van der Waals interactions. The Berendsen coupling algorithm<sup>66</sup> was used to control the temperature and pressure of the simulation system. The relaxation time for the thermostat and barostat is chosen to be 0.1 and 1.0 ps, respectively.

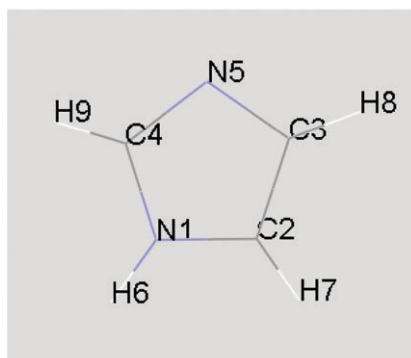
In the first simulation stage, the system was gas-phase optimized, using steepest descent, with all atoms restrained except hydrogen atoms. In the second stage the system was relaxed as an NVT ensemble for 100 ps with all the atoms except hydrogen atoms restrained. The time integration step is set to 2 fs for simulations at 298 K. In the third and final stage, the system was simulated as an  $NpT$  ensemble at 298 K and 1 atm for 2 ns, now without any restraints. The integration time-step was set to 0.5 fs for this final stage.

### 2.3 Optimisation of Lennard-Jones parameters

Once the best level of theory is decided, the LJ parameters are optimised for that level in order to reproduce liquid imidazole's experimental density<sup>72</sup> of 1.0257 g cm<sup>-3</sup> at 383 K (110 °C) and 1 atm. The LJ parameter values are optimised by carrying out simulations, in which all  $\sigma$  values are incrementally increased by one single scaling factor. More details are given in the Results section. The density of liquid imidazole obtained from

**Table 2** Lennard-Jones parameters  $\epsilon$  and  $\sigma$  for all atom types in imidazole as taken from the AMBER ff94 parameter set<sup>a</sup>

Atom	$\epsilon/\text{kJ mol}^{-1}$	$\sigma/\text{Å}$
N1	0.7113	3.25
C2	0.360	3.40
C3	0.360	3.40
C4	0.360	3.40
N5	0.7113	3.25
H6	0.066	1.07
H7	0.063	2.51
H8	0.063	2.51
H9	0.063	2.42



<sup>a</sup> Ref. 68.

these simulations is then compared to the experimental density. The LJ parameters that agree best with experiment are adopted to carry out simulations at different temperatures to look at changes in local structure and bulk property values.

### 2.4 Simulations at different temperatures

The optimised LJ parameters are used to carry out MD simulations at 1 atm and four different temperatures: 298 K, 368 K, 378 K and 383 K. The latter three temperatures are above the melting point of 363 K (90 °C), while the first temperature corresponds to the supercooled liquid at 298 K (25 °C). We include this temperature in order to compare with the RDF calculated by McDonald and Jorgensen,<sup>73</sup> who used their OPLS-AA force field. All the simulations were carried out with multipole moments calculated at the mixed level of CCSD/aug-cc-pVTZ//CCSD/aug-cc-pVDZ.

### 2.5 Bulk properties

We calculated three bulk properties: the isobaric heat capacity ( $C_p$ ), the coefficient of thermal expansion ( $\alpha$ ) and the self-diffusion coefficient ( $D$ ). The isobaric heat capacity is calculated from eqn (7),

$$C_p = \left( \frac{\partial H}{\partial T} \right)_p \quad (7)$$

where  $H$ ,  $p$ , and  $T$  are the enthalpy, pressure and temperature of the system, respectively. We note that the internal energy, which contributes to the enthalpy, consists of both potential and kinetic energy. The coefficient of thermal expansion  $\alpha$  is evaluated from eqn (8),

$$\alpha = \frac{1}{V} \left( \frac{\partial V}{\partial T} \right)_p \quad (8)$$

$D$  is obtained from  $NpT$  simulations and evaluated from the Einstein relationship,<sup>74</sup>

$$D = \frac{1}{2t} \lim_{t \rightarrow \infty} \frac{1}{3} \langle |r(t) - r(0)|^2 \rangle \quad (9)$$

where  $t$  is the simulation time. The position vectors  $r(t)$  and  $r(0)$  represent the coordinates of the atoms at times 0 and  $t$ , respectively. The current four different temperatures suffice to obtain the required derivatives (eqn (7) and (8)) by curve fitting and subsequent analytical differentiation with respect to  $T$ . The set of  $NpT$  simulations carried out at four different temperatures yielded  $C_p$  and  $\alpha$ .

### 2.6 Analysis of local structure

The short-range structure in liquid imidazole has been examined by one-dimensional RDFs, and 3D spatial distribution functions (SDFs). A SDF<sup>75</sup> measures the distribution of the neighbouring molecules (or a specific atom type) around a central molecule (or an atom). For an  $A \cdots B$  RDF or  $A \cdots B$  SDF, the first atom ( $A$ ) in the  $A \cdots B$  notation is always an atom in the central imidazole molecule while the second atom ( $B$ ) belongs to a neighbouring imidazole. Here we describe the procedure to evaluate the SDF of nitrogen–nitrogen separation, where one nitrogen has a hydrogen attached to it [N(H) or N1 in Table 2] and the other nitrogen has a lone pair

(N or N5 in Table 2). The SDF is evaluated for a system of 125 imidazoles in the simulation cell. Each imidazole molecule in this simulation cell is designated as the “central molecule”, one at a time. It is necessary to translate and rotate the whole system such that the central molecule of choice always has the same absolute position with respect to the global frame. This guarantees that the neighbours around the central molecule are always perceived in the same perspective. The examination of the immediate environment of the central molecule requires a distance based criterion to identify the neighbours around the central molecule. A cut-off distance of 5.5 Å for the N...N(H) separation is taken in the construction of the current SDF. A “snapshot” of the simulation cell is sampled every 1000 time steps (0.5 ps), which delivers approximately 2 ns/0.5 ps = 4000 snapshots. Each such snapshot contributes 125 configurations of neighbouring molecules around the central molecule. Hence, there are about  $4000 \times 125 = 600\,000$  configurations.

To analyse the distribution of neighbouring N atoms, a spherical volume with a diameter of 11.0 Å is mapped onto a cubic 3D grid of the same dimension (*i.e.* 11 Å) and with a resolution of 0.1 Å. Hence, there are  $110 \times 110 \times 110 = 1\,331\,000$  grid elements. For each configuration, the corresponding location of the neighbouring N atoms in the cubic grid can be identified based on its position relative to the central molecule's N(H) atom. Consequently, when all 600 000 configurations have been examined, each grid element has accumulated the frequency of occurrence of neighbouring N atoms,  $F_N$ , in that particular location. These raw data are then normalised with the number of configurations used (*i.e.*  $N_{\text{snapshot}} \times N_{\text{N(H)}} = 600\,000$ ). The latter number is then further normalised with the averaged number of N atoms in each grid element,  $\rho_N$ , which is simply given by:

$$\rho_N = N_N \frac{V_{\text{grid}}}{V_{\text{sys}}} \quad (10)$$

where  $N_N$  is the number of N atoms in the simulation cell while  $V_{\text{grid}}$  and  $V_{\text{sys}}$  are the volumes of a grid element and MD cell, respectively. Hence, numerical values for the SDF are obtained from the following equation:

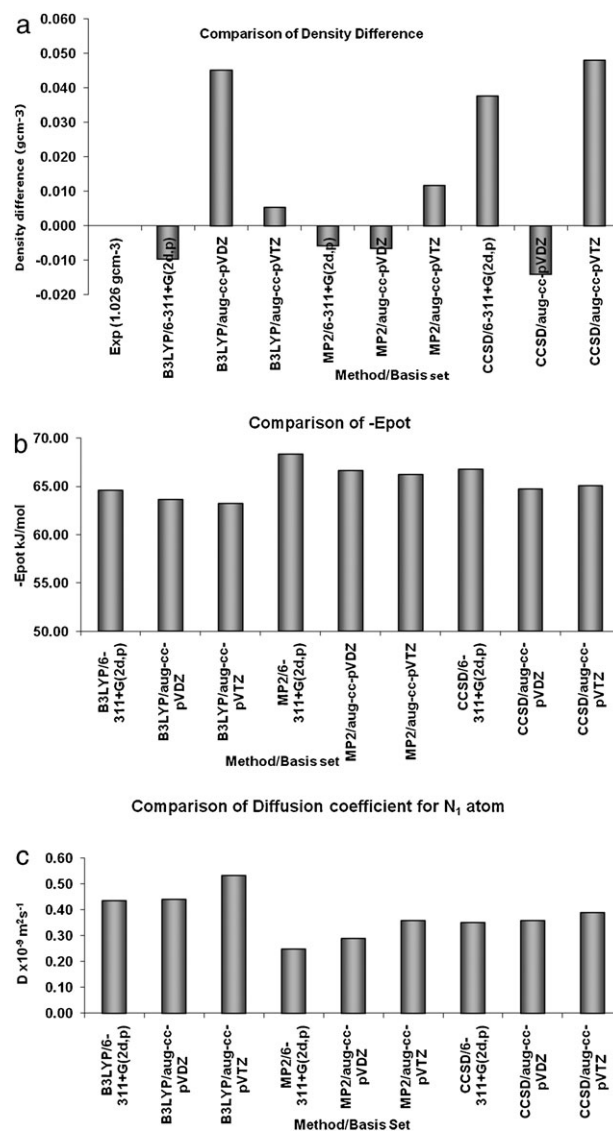
$$\text{SDF}_{\text{N...N(H)}} = \frac{F_N}{N_{\text{snapshot}} N_{\text{N(H)}}} \frac{V_{\text{sys}}}{N_N V_{\text{grid}}} \quad (11)$$

The same analysis can be carried out for other pairs of atom types and the corresponding SDF equation derived.

### 3. Results and discussion

#### 3.1 Comparison between levels of theory

**3.1.1 Properties of liquid imidazole.** The liquid density, potential energy and diffusion coefficients of pure liquid imidazole were evaluated at 1 atm and 383 K. Fig. 1a compares the experimental liquid density with the values obtained from QCT simulations using nine different levels of theory. Tables S1 and S2 (ESI†) display the same information. Fig. 1a highlights MP2/6-311+G(2d,p) and B3LYP/aug-cc-pVTZ as the two best levels. From Table S1 (ESI†) it is clear that the liquid density obtained using multipole moments calculated at



**Fig. 1** A comparison between experiment and QCT simulations (at 1 atm and 383 K) using multipole moments computed at all nine levels of theory of (a) the density of liquid imidazole, (b) the potential energy and (c) the diffusion coefficient of the hydrogen bearing nitrogen N(H).

B3LYP/aug-cc-pVTZ level (and the AMBER ff94 parameter set of LJ parameters) agrees slightly better with experiment than MP2/6-311+G(2d,p). Hence we proceed with B3LYP/aug-cc-pVTZ. For the same LJ parameter set, the simulations carried out at the most CPU time consuming level of theory (*i.e.* CCSD/aug-cc-pVTZ//CCSD/aug-cc-pVDZ), the density agreed poorly with experiment compared to the other levels of theory.

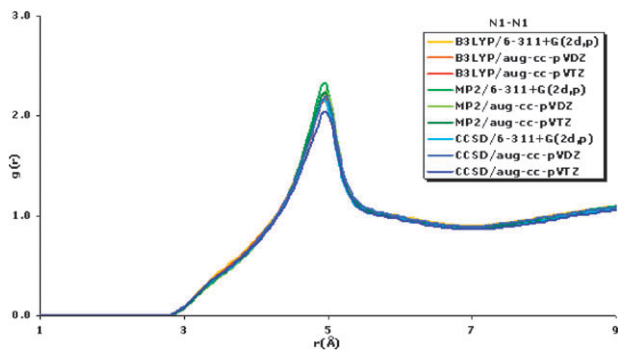
The potential energies obtained for the liquid imidazole simulations using multipole moments from different levels of theory are shown in Table S1 of the ESI† and illustrated in Fig. 1b. This figure shows that the absolute value of the potential energy obtained for liquid imidazole decreases for a given method from 6-311+G(2d,p) to aug-cc-pVTZ basis set. A similar trend was observed with QCT simulations on liquid water carried out in our lab. In those simulations we also

observed that the 6-311+G(2d,p) basis set causes interaction energies to be slightly stronger (*i.e.*  $E_{\text{pot}}$  more negative) than the interaction energies observed using aug-cc-pVDZ and aug-cc-pVTZ.

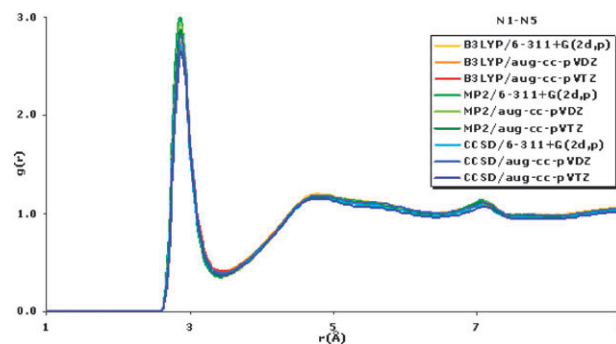
A similar analysis is carried out for the diffusion coefficients. A comparison of diffusion coefficients obtained for each atom type of imidazole is shown in Table S2 of the ESI.† Fig. 1c illustrates the data for the hydrogen-bearing nitrogen only (*i.e.* N(H)). For a given method (B3LYP, MP2, CCSD) the diffusion coefficient increases from 6-311+G(2d,p) to aug-cc-pVTZ. It turns out that every atom type shows this tendency, which was observed in QCT simulations on liquid water as well. From the above observations one deduces that the simulated diffusion coefficients are more dependent on the basis set than on the method.

**3.1.2 Local structure.** The local structure of liquid imidazole was examined by inspecting both RDFs and SDFs, calculated at 383 K and 1 atm. The computed N(H)··N(H), N(H)··N, and N··H(N) RDFs are shown in Fig. 2–4, respectively, for all nine levels of theory. Note that H(N) is the hydrogen atom bonded to N (see numbering scheme of Table 2). An A··B RDF describes the likelihood of finding a neighbouring atom B at a distance  $r$  from a central atom A in the liquid. Fig. 2 compares intermolecular N(H)··N(H) RDFs from QCT simulations at all levels of theory. The first conclusion from Fig. 2 is that the level of theory barely influences the shape of the RDFs. The peak is most pronounced for MP2 and least for B3LYP. Secondly, Fig. 2 shows a large peak in the N(H)··N(H) RDF, corresponding to an internuclear distance ( $r$ ) of just under 5 Å. This value can be mapped onto the N(H)··N(H) distance of 4.90 Å found in a hydrogen-bonded energy minimum of the gas phase imidazole dimer calculated at BHandH/6-311+G(2d,p), which is shown in Fig. 5a. Fig. 2 also shows a weak shoulder at distances of *ca.* 3–4 Å in the RDFs. A pronounced shoulder was observed in the work<sup>73</sup> carried out by McDonald and Jorgensen on super-cooled imidazole at 25 °C (or 298 K). They suggested that this shoulder was due to a significant population of both T-shaped and  $\pi$ -stacked dimers. The authors proposed this interpretation based on a visual inspection of the simulated system.

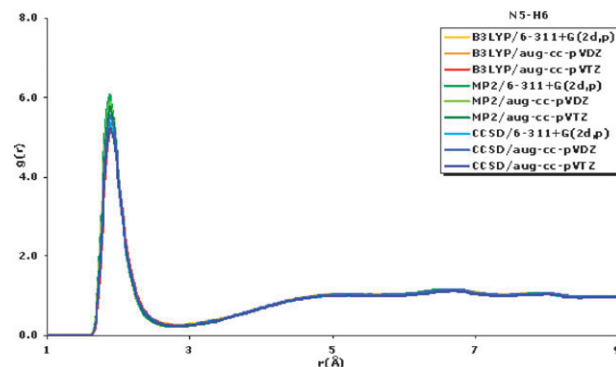
We aimed at retrieving guidance from *ab initio* calculations as to the origin of the N(H)··N(H) distances of 3 to 4 Å. This



**Fig. 2** A comparison of computed intermolecular N(H)··N(H) radial distribution functions for liquid imidazole from QCT simulations (at 1 atm and 383 K) at nine different levels of theory.



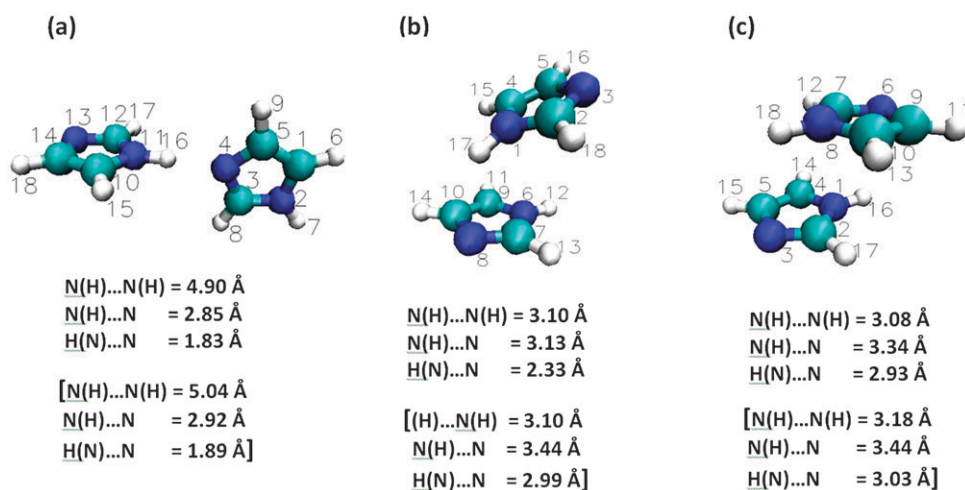
**Fig. 3** A comparison of computed intermolecular N(H)··N radial distribution functions for liquid imidazole from QCT simulations (at 1 atm and 383 K) at nine different levels of theory.



**Fig. 4** A comparison of computed intermolecular N··H(N) radial distribution functions for liquid imidazole from QCT simulations (at 1 atm and 383 K) at nine different levels of theory.

is why we searched for both T-shaped and  $\pi$ -stacked dimers at BHandH/6-311+G(2d,p) level, with a further check at MP2/6-311+G(2d,p) level. The hydrogen-bonded (or chain) configuration was readily found at both levels of theory, and confirmed as a minimum by a frequency calculation, using Gaussian 03<sup>57</sup> without any geometry constraints. From Fig. 5a one learns that the N(H)··N(H) distance differs by only 3% between MP2 and BHandH. The  $\pi$ -stacked dimer configuration proved harder to locate, and its existence may appear surprising in the light of the strong propensity of the nitrogen lone pair to form a hydrogen bond. We managed to find two (frequency-confirmed)  $\pi$ -stacked minima at BHandH/6-311+G(2d,p) level, again without any geometry constraints. One minimum consists of a nearly parallel arrangement (Fig. 5c) while the other minimum shows a tilt (Fig. 5b) between the two imidazole planes, which is less stable by 1.3 kJ mol<sup>-1</sup>. In spite of several attempts we could not recover the tilted structure at MP2 level, which always reverted to the parallel structure. The T-shaped configuration proved elusive with both BHandH and MP2. However, a T-shaped second order transition state could be located with BHandH. To the best of our knowledge no T-shaped configuration of the imidazole dimer has ever been reported.

The BHandH functional improves upon the ability of B3LYP to calculate  $\pi$ -stacked geometries.<sup>76</sup> However, this functional contains a fortuitous cancellation<sup>77,78</sup> in the

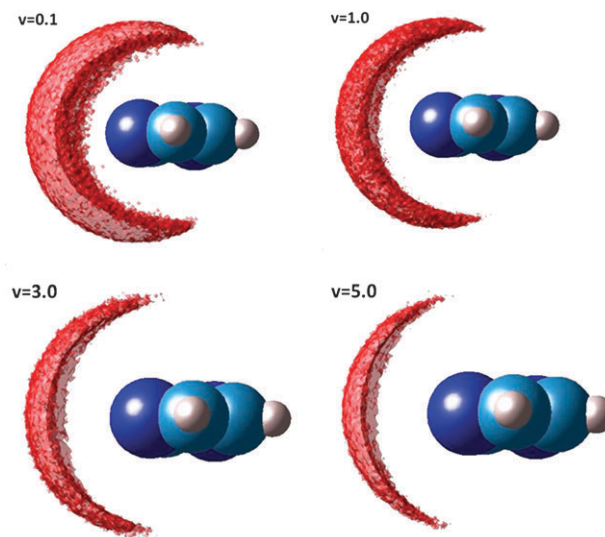


**Fig. 5** Optimised geometries of two configurations of the imidazole dimer in the gas phase, obtained at BHandH/6-311+G(2d,p) level. (a) Hydrogen-bonded (or “chain”) configuration, (b) tilted “ $\pi$  stacked” and (c) parallel “ $\pi$  stacked”. Selected distances are shown under each configuration with the corresponding values obtained at MP2/6-311+G(2d,p) level in square brackets.

exchange–correlation functional that mimics the dispersive part of methods such as MP2. As seen from Fig. 5b and c, the  $\text{N(H)}\cdots\text{N(H)}$  distance for a  $\pi$ -stacked dimer energy minimum is  $\sim 3.10 \text{ \AA}$ . In Fig. 2, these distances fall in the region of the very weak shoulder left of the peak and we also observe  $\pi$ -stacked configurations from visual examination of the trajectories of the full simulation.

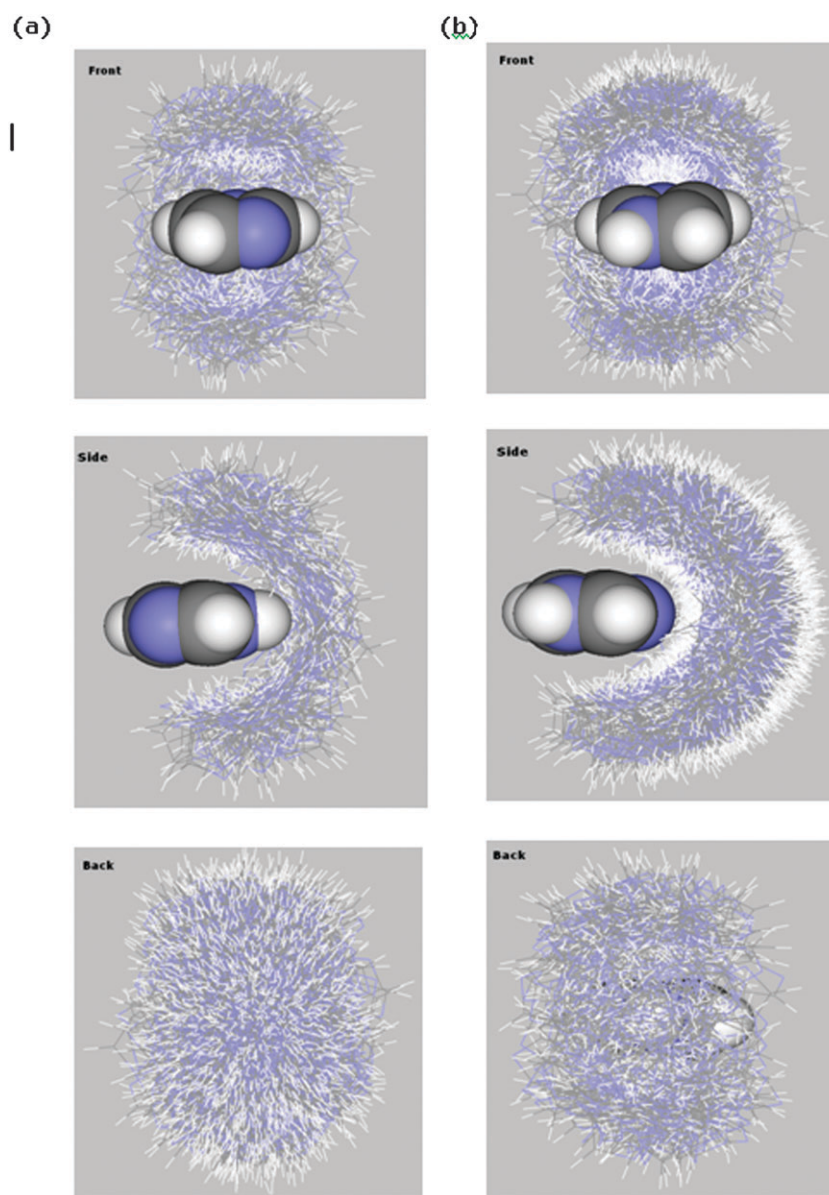
Fig. 3 shows  $\text{N(H)}\cdots\text{N}$  RDFs in which a peak appears at about  $2.9 \text{ \AA}$ . This value can be mapped onto the  $\text{N(H)}\cdots\text{N}$  distance of  $2.85 \text{ \AA}$  (and  $2.92 \text{ \AA}$  at MP2) in the hydrogen-bonded configuration of the gas phase dimer, shown in Fig. 5a. From the  $\text{N}\cdots\text{H(N)}$  RDFs in Fig. 4 one observes that the distance of the peak also corresponds closely to the  $\text{N}\cdots\text{H(N)}$  distance in the optimised imidazole dimer in Fig. 5a. The length of the hydrogen bond from the simulations is  $\sim 1.88 \text{ \AA}$ , which is closely mapped onto the optimised  $\text{N}\cdots\text{H(N)}$  distance in the gas phase dimer at  $1.83 \text{ \AA}$  (and  $1.89 \text{ \AA}$  at MP2).

An alternative way of examining the local structure is to use the SDFs where the distribution and number of adjacent molecules are represented by isosurfaces of specific values. A high isovalue can be interpreted as a high likelihood of finding a particular atom (or molecule) at a specific location relative to the central molecule. Fig. 6 shows the SDF for  $\text{N}\cdots\text{N(H)}$  at four different isovalues, *i.e.*, 0.1, 1.0, 3.0, and 5.0. The SDFs show that the neighbouring atoms ( $\text{N(H)}$ ) are most likely to be found in a region of space that resembles a 3D crescent (in red) near the N atom from the central molecule. The height, the depth and the arc length of the crescent decrease as the isovalue increases. At the lowest isovalue, the tips of the crescent extend to cover the top and bottom side of the central molecule. This confirms the existence of  $\pi$ -stacked dimers in our simulations, albeit with a low frequency of occurrence. At the highest isovalue, most of the neighbouring  $\text{N(H)}$  atoms can be regarded as hydrogen bonded to the N atom. Now the two “tips” of the crescent only reach to the region immediately above and below the N atom.



**Fig. 6** Spatial distribution functions for  $\text{N}\cdots\text{N(H)}$  interactions in liquid imidazole at 383 K and 1 atm, at four different isovalues ( $v = 0.1, 1.0, 3.0$  and  $5.0$ ) illustrating the variation in the immediate environment of the central imidazole molecule (plotting scales adjusted visually and a cut-off of  $7 \text{ \AA}$  is applied).

Fig. 7 shows plots of explicit cluster configurations, from which the SDFs are calculated. In order to focus on a particular atom of interest it is best to look at the immediate molecular surroundings of this atom, cutting out molecules further away than  $5.5 \text{ \AA}$ . In this way we obtain Fig. 7a, where  $\text{N(H)}$  is the atom of interest, and Fig. 7b where it is N. The existence of  $\pi$ -stacked dimers can be readily inferred from Fig. 7. The side views of Fig. 7a and b clearly show the presence of neighbouring imidazole molecules above and below the ring, with their hydrogen atoms pointing toward the central ring. The “front” and “side” views of Fig. 7b also demonstrate the strong presence of hydrogen-bonded hydrogen atoms from neighbouring molecules close to the N atom. However, this feature is not observed in the



**Fig. 7** Location of neighbouring molecules adjacent to a central imidazole molecule in the simulation cell for (a)  $\text{N(H)}\cdots\text{N(H)}$  and (b)  $\text{N}\cdots\text{N(H)}$ , where the first atom is the central atom. Three different views are presented: front, side and back view.

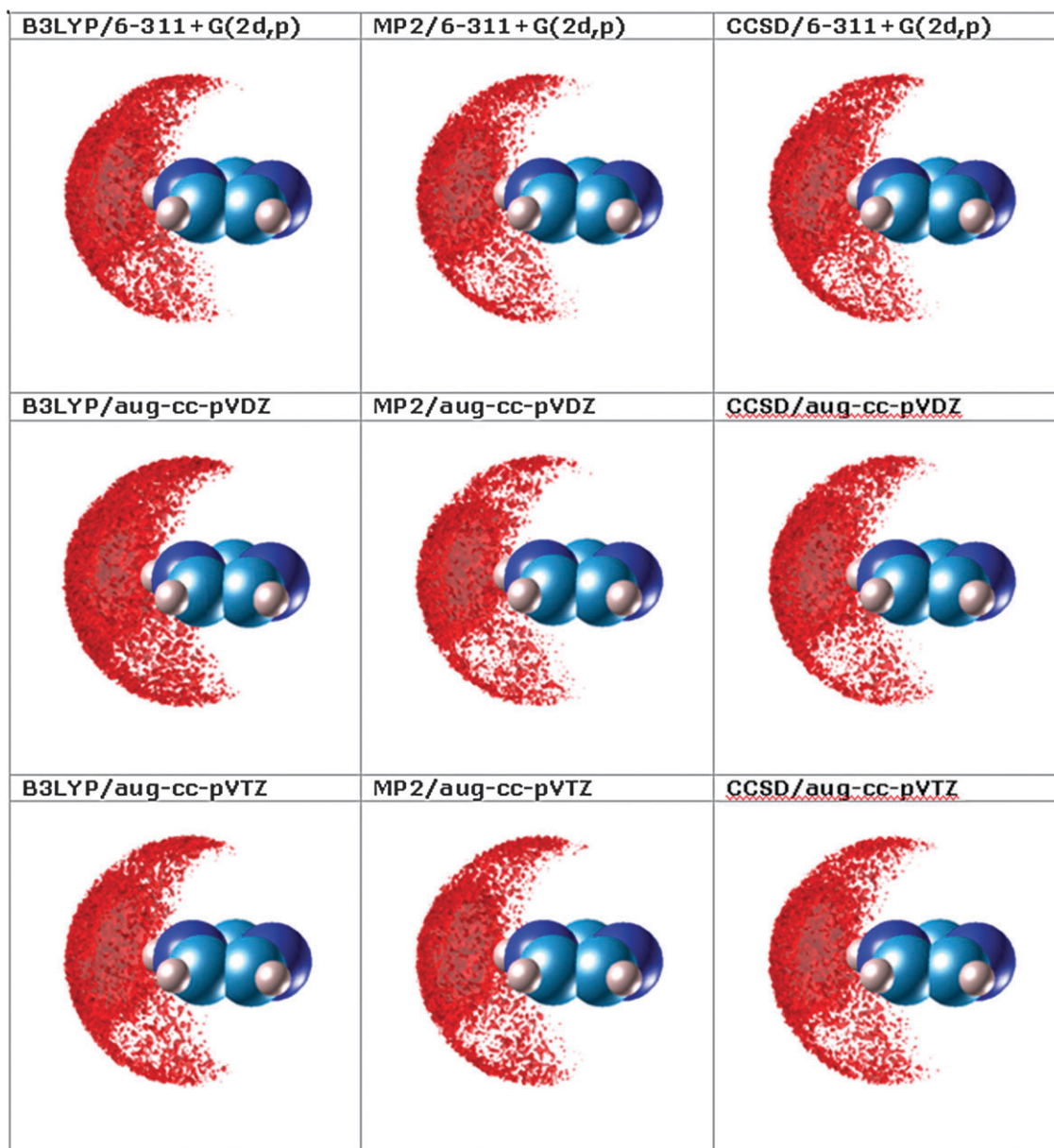
$\text{N(H)}\cdots\text{N(H)}$  clusters (Fig. 7a) where the orientation of neighbouring molecules seems to be more random. This observation is consistent with these molecules not necessarily being hydrogen-bonded to the central molecule (through the  $\text{N(H)}$  atom).

The SDFs obtained from the simulations with multipole moments from all levels of theory are shown in Fig. 8. The SDFs obtained from simulations using B3LYP, MP2 and CCSD methods with the 6-311+G(2d,p) basis set gave similar profiles. Overall, the effects of change in level of theory are hard to discern visually. In contrast, in a previous systematic study on water<sup>79</sup> a change in basis set caused a larger change in the SDFs compared to a change in method. However, differences caused by change in level of theory have been identified in other characteristics such as thermodynamic and dynamic properties of the system (see Tables S1 and S2, ESI<sup>†</sup>).

### 3.2 Optimisation of Lennard-Jones parameters

Irrespective of the multipole moment used, simulations using the standard AMBER ff94 LJ parameters cause the liquid's density to deviate somewhat from the experimental value. As a result, the LJ parameters were optimised in order to carry out simulations at densities that are consistent with experiment. For this purpose we simply scaled all  $\sigma$  parameters used in the simulation by the same scaling factor. We systematically increased the  $\sigma$  value by up to 5% in steps of 0.5%. MD simulations were carried out for liquid imidazole at 1 atm and 383 K. Table S3 of the ESI<sup>†</sup> shows the variation in liquid density of the simulated systems using different  $\sigma$  values. As expected, the density decreases as the  $\sigma$  value increases. At 2.5% the density of the simulated system is closest to the experimental density (see Fig. 9). We used these values (see Table 3)





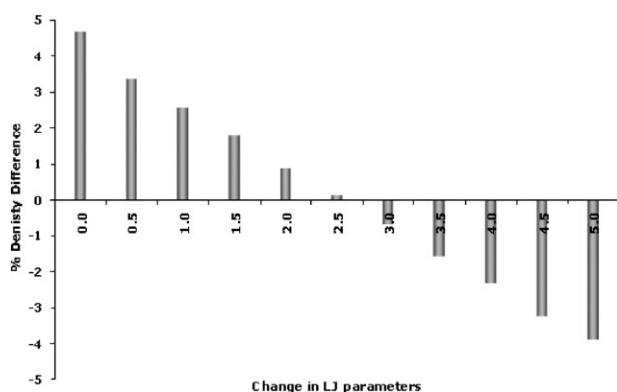
**Fig. 8** Spatial distribution functions for N(H)···N interactions in liquid imidazole at 383 K and 1 atm at isovalue  $v = 1.0$ , illustrating the variation in the immediate environment of the central imidazole molecule with respect to different levels of theory (cut-off of 5.5 Å applied).

to carry out further simulations to examine the behaviour of the system at various temperatures.

### 3.3 Convergence of the multipolar expansion

The issue of convergence cannot be ignored any longer and is addressed here, as announced in Section 2.1. We focus first on the dimer's chain configuration (Fig. 5a) calculated at BHandH/6-311+G(2d,p) level. The atoms most at risk of participating in a divergent electrostatic interaction are “frontier atoms”. Candidates for these “frontier atoms” would be C<sub>3</sub>, N<sub>4</sub> and C<sub>5</sub> in the right monomer, and N<sub>11</sub> and H<sub>16</sub> in the left monomer. In spite of the short range nature of the 6 (= 3×2) possible interactions some interactions converge. It is instructive to monitor the behaviour of the interaction energy of four atom–atom interactions as a function of the

interaction rank  $L$ . Fig. 10 shows the profiles of N<sub>4</sub>H<sub>16</sub>, N<sub>4</sub>N<sub>11</sub>, C<sub>3</sub>H<sub>16</sub> and N<sub>1</sub>H<sub>16</sub>. A glance at Fig. 10 reveals that the N<sub>4</sub>H<sub>16</sub> and N<sub>4</sub>N<sub>11</sub> interactions diverge, as deduced from the strong oscillations at high  $L$  values. The wildest oscillations occur for N<sub>4</sub>H<sub>16</sub> in Fig. 10a, which is expected since this pair has the shortest internuclear distance (1.8 Å). Energies evaluated at higher  $L$  values do not necessarily guarantee a smaller deviation from the exact energy, as is clear from the increasingly dramatic oscillations upon increasing  $L$ . Of great interest is the energy at  $L = 5$  because this is the rank at which the simulation was carried out. Surprisingly and fortuitously, the discrepancy between the exact energy and the energy calculated at  $L = 5$  is only 0.6 kJ mol<sup>-1</sup>. The energy profile of N<sub>4</sub>N<sub>11</sub> shows the beginning of dramatic oscillations, which now occur “later”, *i.e.* at higher rank  $L$ . Among the four



**Fig. 9** A comparison of % density difference of liquid imidazole computed models with increase in Lennard-Jones parameter values by 0.5% increase in  $\sigma$  values in the AMBER ff94 parameter set with experimental density at 383 K and 1 atm.

**Table 3** Lennard-Jones parameters for different atom types of imidazole obtained from optimisation of the AMBER ff94<sup>68</sup> parameter set with a 2.5% increase in  $\sigma$  values (QCT simulations carried out at 383 K and 1 atm)

Atom	$\epsilon/\text{kJ mol}^{-1}$	$\sigma/\text{\AA}$
N1 (or N(H))	0.7113	3.331
C2	0.360	3.485
C3	0.360	3.485
C4	0.360	3.485
N5 (or N)	0.7113	3.331
H6 (or H(N))	0.066	1.096
H7	0.063	2.573
H8	0.063	2.573
H9	0.063	2.482

short-range interactions shown in Fig. 10,  $\text{N}_4\text{N}_{11}$  shows the largest energy discrepancy between the energy at  $L = 5$  and the exact energy, namely  $\Delta E = -2.9 \text{ kJ mol}^{-1}$ . Although the discrepancy between these two N atoms is quite considerable, it is still reasonable in spite of the short distance of 2.9 Å between  $\text{N}_4$  and  $\text{N}_{11}$ . The energy profile of  $\text{N}_4\text{N}_{11}$  is shown in Fig. 10b. Fig. 10c shows an example of so-called pseudo-convergence. Here the  $\text{C}_3\text{H}_{16}$  interaction displays a plateau-like energy profile (between  $L = 6$  and  $L = 18$ ) that oscillates around the exact (6D) energy. At  $L = 5$ , the energy is lower than the exact 6D energy by only  $0.2 \text{ kJ mol}^{-1}$ , which is most encouraging. Fig. 10d shows the  $\text{C}_1\text{H}_{16}$  energy profile, which is an example of a perfectly converged interaction. Overall, Fig. 10 demonstrates that an energy obtained at  $L = 5$  is the best choice for pseudo-convergent and divergent interactions. Fig. 10 covers the behaviour of many possible interactions between all atoms. We note that the  $L = 5$  energies are always below the exact value.

In addition, we also calculated atom–atom interactions between two imidazole molecules starting from their monomeric wave functions rather than the supermolecular dimer wave function used above. Note that the energies have been obtained from a multipole expansion and that no 6D integration was carried out. Qualitatively, we observed the same convergence/divergence patterns as before: the  $\text{N}_4\text{N}_{11}$  and  $\text{N}_4\text{H}_{16}$  interactions are divergent, the  $\text{C}_3\text{H}_{16}$  interaction is pseudo-convergent

while the  $\text{C}_1\text{H}_{16}$  interaction is convergent. The energies of many other interactions (except  $\text{N}_4\text{N}_{11}$ ,  $\text{N}_4\text{H}_{16}$  and  $\text{C}_3\text{H}_{16}$ ) computed from the monomers are numerically similar to the energies computed from the dimer. For example, the monomer and dimer energy profiles of the  $\text{H}_6\text{C}_{14}$  interaction are largely parallel and differ by only  $0.6 \text{ kJ mol}^{-1}$  in their high rank plateaus. Overall, the relative deviation of the energy difference  $100|E_{\text{monomer}} - E_{\text{dimer}}|/E_{\text{dimer}}$  is less than 10%. This deviation was also confirmed for the stacked configuration of the imidazole dimer. We also note that the number of divergent atom–atom interactions in the stacked configuration is larger than in the chain configuration. In the stacked configuration there are eight divergent atom–atom interactions, when computed by both the monomer and dimer route. This is not surprising because the stacked configuration has the highest number of short atom–atom distances.

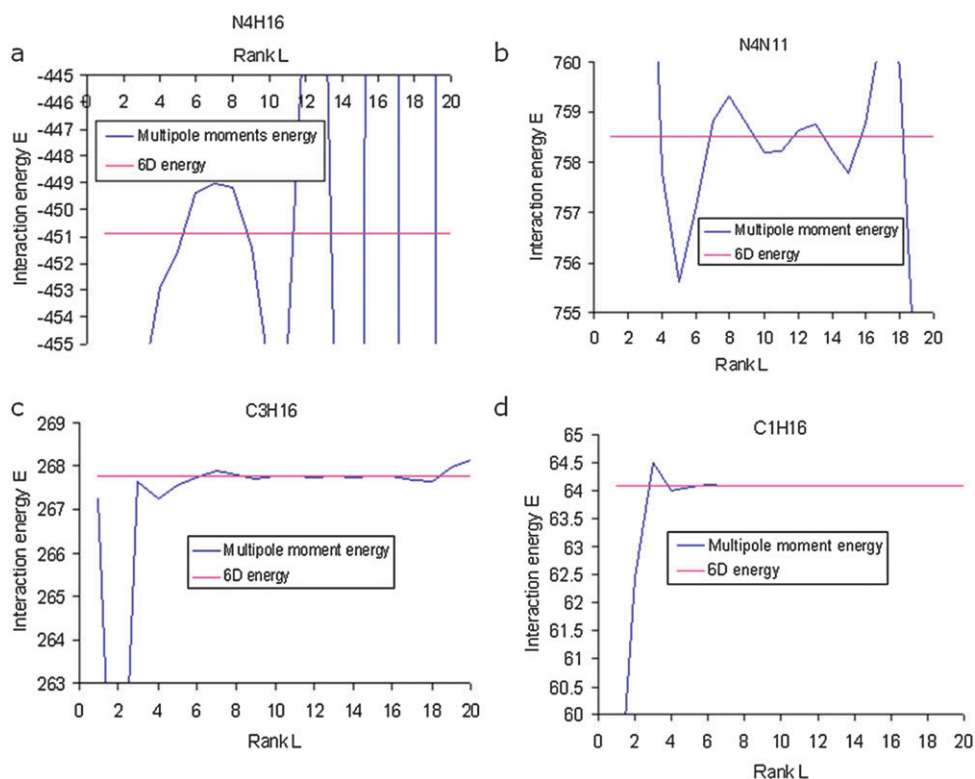
Finally, we recalculated monomeric and dimeric wave functions at MP2/6-311+G(2d,p) level. For the chain configuration, the number of divergent atom–atom interactions at the MP2/6-311+G(2d,p) level is the same than at the BHandH/6-311+G(2d,p) level. A comparison for the stacked configuration is hampered by the fact that the two levels do not yield sufficiently similar geometries. Secondly, the interaction energies of the dimer structures are lower, by 20%, as measured by  $100|E_{\text{BHandH}} - E_{\text{MP2}}|/E_{\text{MP2}}$ .

### 3.4 Simulations at different temperatures

**3.4.1 Bulk properties.** Liquid imidazole simulations were carried out at four different temperatures ( $T = 298, 368, 378$  and 383 K). The various properties calculated from the QCT-simulated imidazole at ambient pressure and different temperatures are summarized in Table 4. The variation of density of the system as a function of temperature is shown in Fig. 11. Clearly, the density of the system decreases as the temperature of the system increases. Experimental data are limited in the literature but our results seem to be in reasonable agreement with the available data. The simulated densities are close to the experimental values at 383 K<sup>72</sup> ( $1.026 \text{ g cm}^{-3}$ ) and 374 K<sup>80</sup> ( $1.030 \text{ g cm}^{-3}$ ). In addition, the density of the system at 298 K differs by only  $\sim 2\%$  from the density ( $1.120 \text{ g cm}^{-3}$ ) obtained using OPLS-AA method.<sup>73</sup>

Table 4 shows that the calculated potential energy of the system at different temperatures increases with the system's temperature. Due to a lack of experimental data no comparison with experiment could be made. The values for  $C_p$ ,  $\alpha$  and the diffusion coefficients also increase with temperature. There are experimental data available for  $C_p$  but the values were obtained at a pressure of 1 kPa,<sup>81</sup> again preventing a direct comparison. However, those experimental values also show an increase of  $C_p$  with temperature. The calculated  $C_p$  value deviates by 34% from the value given by OPLS-AA.

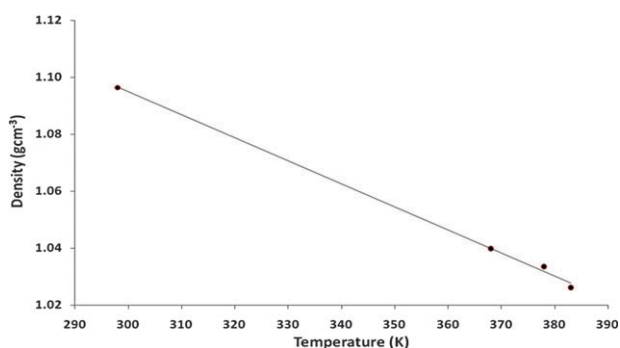
**3.4.2 Local structure.** The local structure of the liquid imidazole was examined at different temperatures by monitoring the RDFs and SDFs. The computed RDFs for  $\text{N(H)} \cdots \text{N(H)}$  and  $\text{N(H)} \cdots \text{N}$  are shown in Fig. 12, and the RDFs of  $\text{N} \cdots \text{H(N)}$  in Fig. 13, all at different temperatures and at 1 atm. The  $\text{N(H)} \cdots \text{N(H)}$  RDFs in Fig. 12a reveal that the position of the peak ( $\sim 5.03 \text{ \AA}$ ) corresponds closely to the



**Fig. 10** Atom-atom interaction energies *versus* interaction rank  $L$  in the chain structure of the imidazole dimer. Three kinds of interaction profiles occur: (a)  $N_4H_{16}$  and (b)  $N_4N_{11}$  are divergent, (c)  $C_3H_{16}$  is pseudo-convergent and (d)  $C_1H_{16}$  interaction is convergent.

**Table 4** Calculated properties of imidazole at different temperatures and 1 atm

Temp./K	Density/g cm <sup>-3</sup>	Potential energy/kJ mol <sup>-1</sup>	$C_p$ /J K <sup>-1</sup> mol <sup>-1</sup>	$\alpha \times 10^5$ K <sup>-1</sup>	$D_{N(H)} \times 10^{-9}$ m <sup>2</sup> s <sup>-1</sup>	$D_N \times 10^{-9}$ m <sup>2</sup> s <sup>-1</sup>
298	1.096	-68.6	109.0	76.0	0.02	0.02
368	1.040	-62.6	112.2	77.2	0.40	0.40
378	1.034	-62.0	112.7	77.4	0.46	0.46
383	1.026	-61.3	112.9	77.2	0.63	0.62

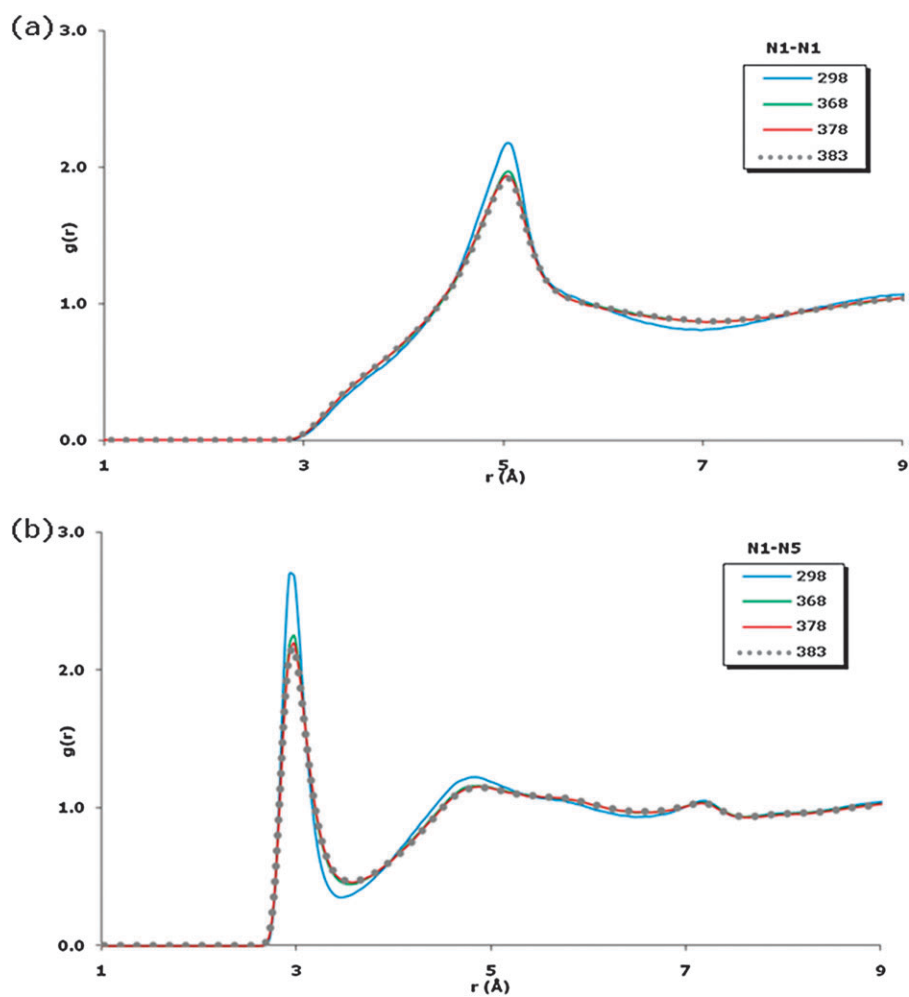


**Fig. 11** Calculated density of liquid imidazole at different temperatures and 1 atm.

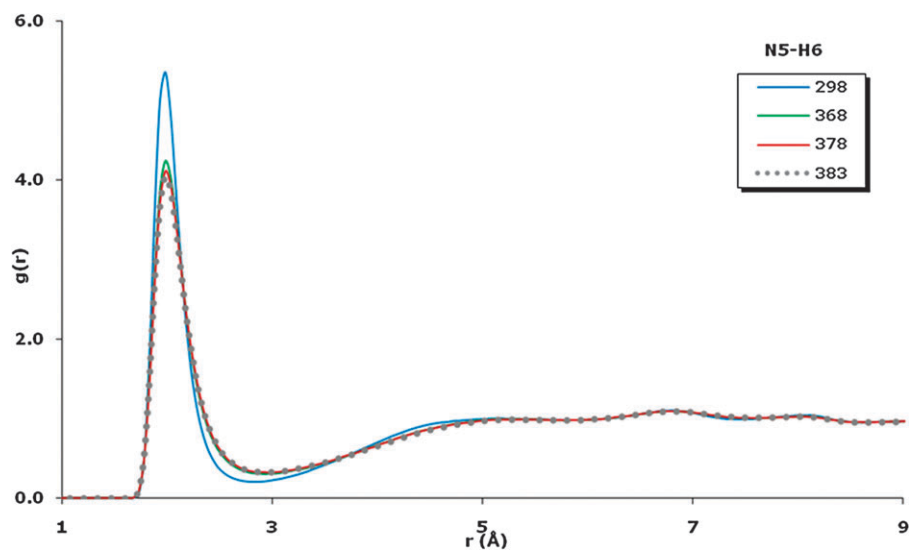
$N(H) \cdots N(H)$  distances (4.90 Å with BHandH and 5.04 Å with MP2) in the hydrogen-bonded dimer shown in Fig. 5a. A small shoulder, seen in earlier calculations, can also be observed in the RDFs (at all temperatures) at distances of  $\sim 3$ –4 Å. In contrast, a more pronounced shoulder between 3.0 and 4.5 Å was observed in the work of McDonald and Jorgensen<sup>82</sup> using the OPLS-AA force field on supercooled imidazole (at 298 K).

As explained in the previous section, this shoulder is due to the presence of  $\pi$ -stacked dimers. The  $N(H) \cdots N(H)$  distance in the  $\pi$ -stacked dimer optimised at BHandH/6-311+G(2d,p) level is  $\sim 3.10$  Å (Fig. 5b), which falls in the region of the shoulder in the RDFs of Fig. 12a. The snapshots of Fig. 7 corroborate this finding because a direct analysis reveals a certain amount of  $\pi$ -stacked dimers characterised by  $N(H) \cdots N(H)$  distances in a range of 3.0 to 4.0 Å.

Fig. 12a also shows that the height of the peak decreases with increasing temperature. The well near 7.0 Å becomes less pronounced when the temperature is increased. In Fig. 12b one observes three peaks, at 2.98 Å, 4.83 Å and 7.13 Å. The maximum peak occurs at 2.98 Å at all the temperatures except 298 K, where it occurs at 2.93 Å. This is in agreement with the  $N(H) \cdots N$  distance observed in the hydrogen-bonded dimer shown in Fig. 5a, where the distance is 2.85–2.92 Å. At higher temperatures, the first (and major) peak broadens while its height decreases. The second and third peaks are weak and defied attempts to be assigned to interatomic distances occurring in known gas phase *ab initio* dimer configurations. The well between the first and second peak also becomes less pronounced when the temperature is increased. Fig. 13 demonstrates that



**Fig. 12** (a) A comparison of computed intermolecular N(H)··N(H) radial distribution functions for liquid imidazole from QCT simulations at different temperatures and 1 atm. (b) A comparison of computed intermolecular N(H)··N radial distribution functions for liquid imidazole from QCT-simulations at different temperatures and at 1 atm.



**Fig. 13** A comparison of computed intermolecular N··H(N) radial distribution functions for liquid imidazole from QCT simulations at different temperatures and at 1 atm.

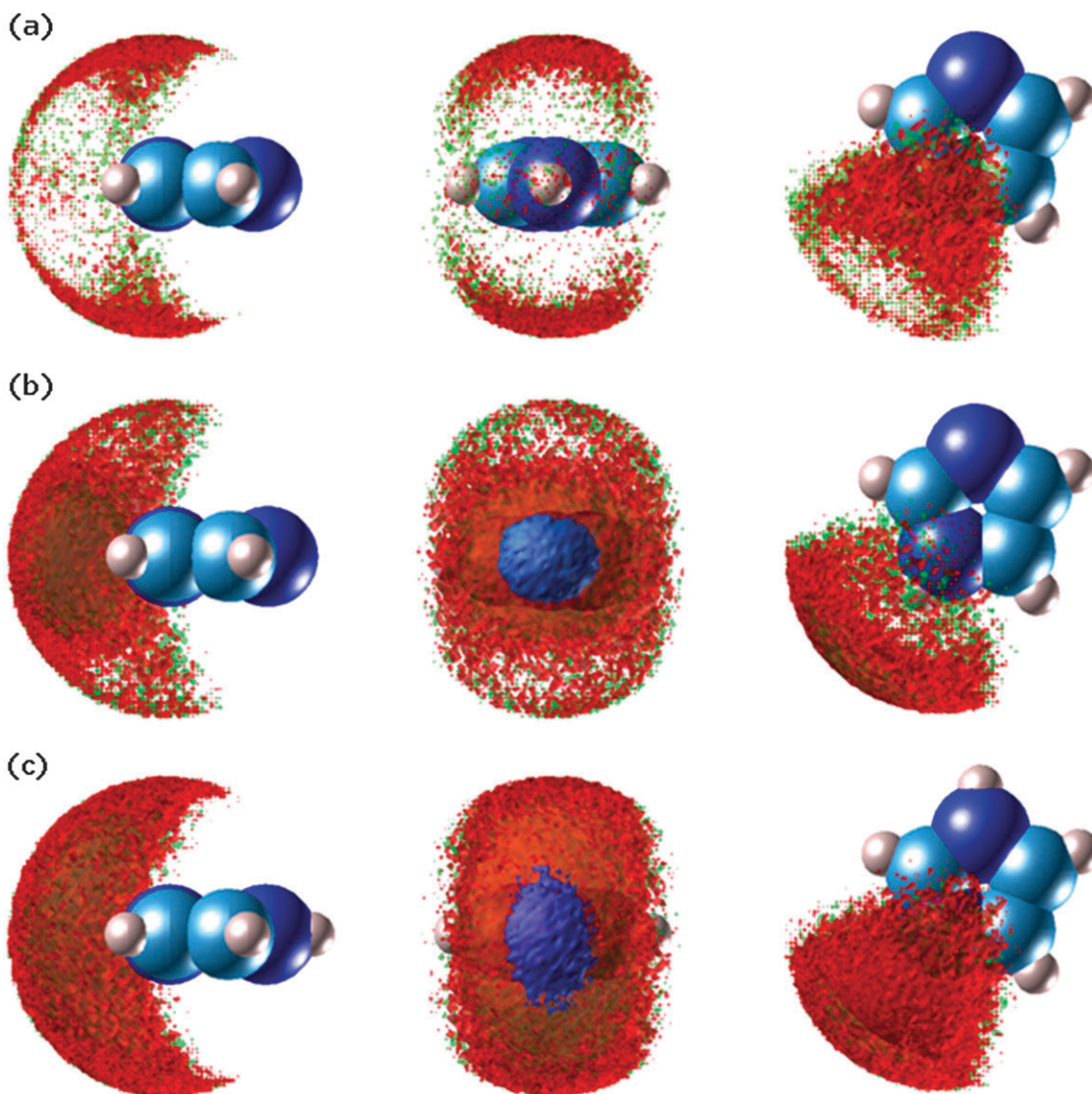
in the  $N \cdots H(N)$  RDFs the location of the first peak at 1.98 Å corresponds closely to the  $N \cdots H(N)$  distance of 1.83 Å, in the optimised imidazole dimer shown in Fig. 5a.

The SDFs acquired at different temperatures are shown in Fig. 14. These SDFs were obtained for a 5.5 Å cut-off distance between the central imidazole and its neighbours and the value for the isosurface in the SDFs is 1.0. Overall the SDFs reveal a picture that is consistent with the findings from the RDFs. As in the RDFs, the effect of temperature is limited to a small expansion in the SDF profile from 298 K (green) to 383 K (red). However, the shape of the isosurfaces does not change significantly with temperature. One unexpected observation in the SDF of  $N(H) \cdots N$  (Fig. 14b) and  $N \cdots N(H)$  (Fig. 14c) is that the width of the crescent is slightly larger for  $N(H) \cdots N$ . In the  $N(H) \cdots N$  SDF, a secondary feature appears in the center of the crescent, which is circular in shape and located opposite to the hydrogen atom, *i.e.*  $H(N)$ . This feature coincides with the region of high hydrogen density in Fig. 7b, and

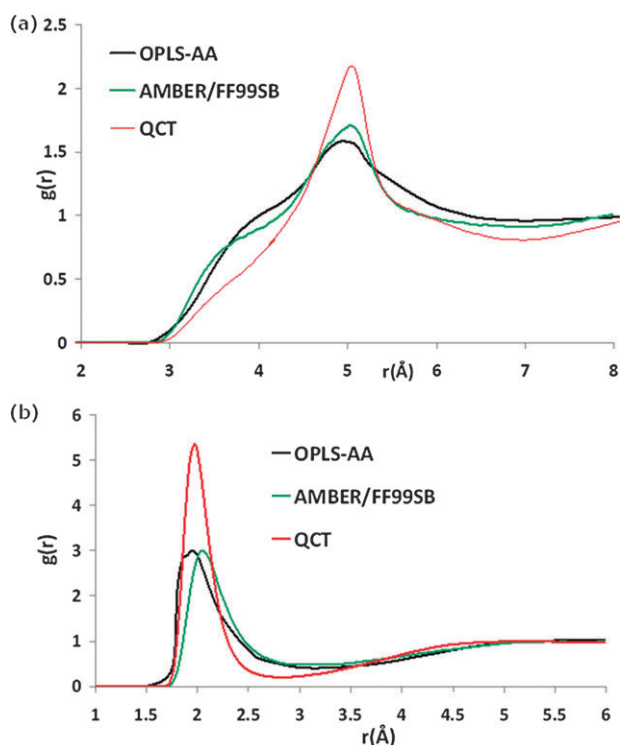
expresses the directional character of the hydrogen bond. In contrast, there is no such secondary feature in the  $N \cdots N(H)$  SDF (Fig. 14c). Here the lone pair of the nitrogen (N) in the central imidazole does not direct and therefore diffuses the hydrogen-bonded hydrogen atoms of the neighbouring imidazoles. In addition, we have included the SDF isosurface with an isovalue of 20 in Fig. 14b and c. The choice of this very high contour value helps in revealing the location of the maxima in the SDFs.

### 3.4.3 Comparison with OPLS-AA and AMBER/ff99SB.

An important aim of this study is comparing the QCT results with those produced using a traditional point-charge force field. To this end, we compare with the study of McDonald and Jorgensen,<sup>73</sup> who used the OPLS-AA force field. We also performed a new simulation using the AMBER/ff99SB force field. Fig. S1 of the ESI† shows the deformation of imidazole due to intramolecular flexibility inherent in the AMBER



**Fig. 14** Spatial distribution functions for different interactions in liquid imidazole at 298 K (green) and 383 K (red) at isovalue  $\nu = 1.0$  and 3.5 Å cut-off for (a)  $N(H) \cdots N(H)$ , (b)  $N(H) \cdots N$  and (c)  $N \cdots N(H)$ . From left to right, the views are: side, front and top. The blue region in the middle figure of (b) and (c) corresponds to the isovalue  $\nu = 20$ .



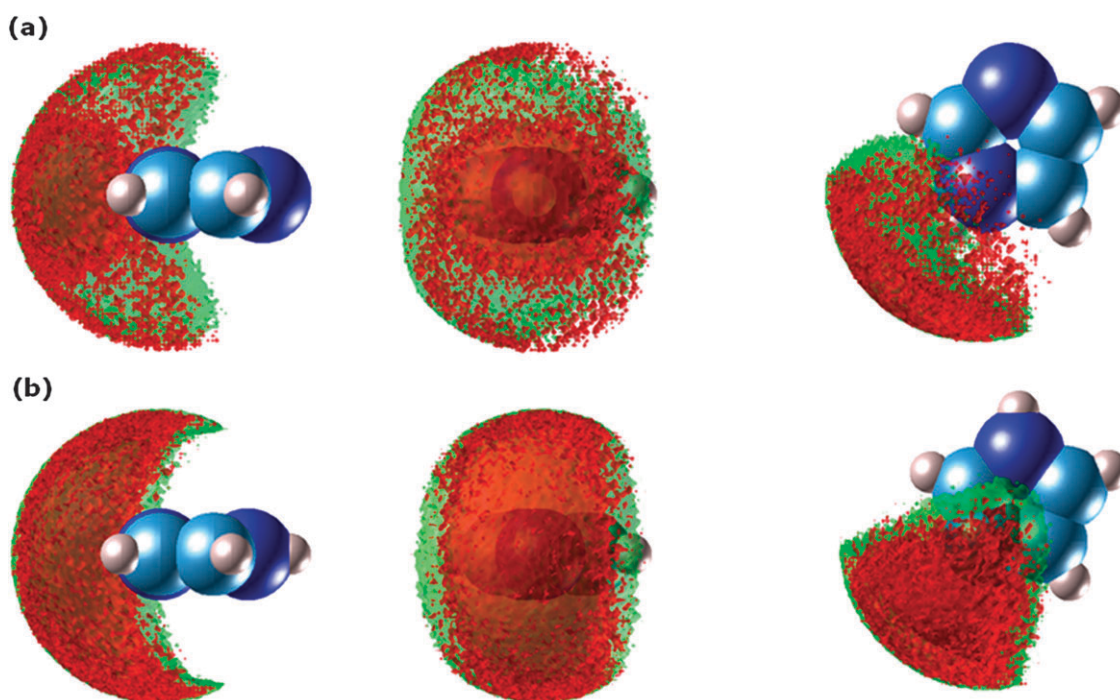
**Fig. 15** A comparison of computed intermolecular radial distribution functions for liquid imidazole from QCT, AMBER/FF99SB and OPLS-AA simulations at 298 K and 1 atm, (a)  $\text{N(H)}\cdots\text{N(H)}$  and (b)  $\text{N}\cdots\text{H(N)}$ .

potential. The ring deformations are fairly small, as well as the out-of-plane bending motions of the C and H atoms. Hence we believe that it is safe enough to compare the results from

the (flexible) point charge model with those of the (rigid body) multipolar model.

The results for the  $\text{N(H)}\cdots\text{N(H)}$  and  $\text{N}\cdots\text{H(N)}$  RDFs are presented in Fig. 15. For  $\text{N(H)}\cdots\text{N(H)}$ , the overall appearance of the RDF from OPLS and AMBER is broadly similar, with small differences in the height of the curve at several locations. One common feature between the two RDFs is the existence of a pronounced shoulder between 3 and 4.5 Å, which is barely identifiable in the QCT simulation. Nevertheless, the three force fields seem to be in agreement on the location of the first peak. However, the peak height of the QCT simulation is significantly higher than that of the point charge simulations. From our previous analysis, this indicates that there are more  $\pi$ -stacked dimers in the OPLS and AMBER simulations. On the other hand, the hydrogen-bonded dimers are more frequently in the QCT simulation. This is corroborated by the  $\text{N}\cdots\text{H(N)}$  RDF (Fig. 15b). Here, the OPLS and AMBER results are in agreement with each other apart from the location of the first peak.

Contrasting the SDFs generated from QCT and AMBER simulations allows a more detailed comparison. The  $\text{N(H)}\cdots\text{N}$  SDF comparison (Fig. 16a) shows that the distribution is quite different between QCT (red) and AMBER (green). The AMBER distribution is more homogeneous than the QCT distribution. Secondly, the QCT distribution resembles a “crescent” with a non-even spread of neighbouring N atoms. The “front” view of Fig. 16a shows that the N atoms are concentrated in the circular region directly opposite to the H atom in  $\text{N(H)}$ . Again, we attribute this distinction to the directionality of the hydrogen bond, a characteristic not reproduced by the AMBER force field. The  $\text{N}\cdots\text{N(H)}$  SDFs (Fig. 16b) show the existence of a crescent-like distribution



**Fig. 16** Spatial distribution functions for different interactions in liquid imidazole at 298 K using AMBER/ff99SB (green) and QCT (red) at isovalue  $\nu = 1.0$  and  $5.5$  Å cut-off for (a)  $\text{N(H)}\cdots\text{N}$ , (b)  $\text{N}\cdots\text{N(H)}$ . From left to right, the views are: side, front and top.

directly opposite to the central N atom for both AMBER and QCT. However, the width of the crescent is wider for AMBER (“front” view of Fig. 16b). We also find that for AMBER, the extremities of the crescent extend to regions above and below the centre of the aromatic ring. In contrast, the extremities of the QCT crescent only reach the N atom (“top” view of Fig. 16b). We believe that this demonstrates the higher likelihood for  $\pi$ -stacked dimers to occur in the AMBER system.

#### 4. Conclusion

For the first time a potential based on high-rank atomic multipole moments computed according to quantum chemical topology (QCT) has been used to study liquid imidazole by molecular dynamics simulations. In this study only LJ parameters were adjusted while multipole moments for the electrostatic interaction were obtained from *ab initio* wave functions of a single gas phase imidazole molecule. We carried out a systematic comparison of the effects of both the method (B3LYP, MP2 and CCSD) and the basis set (6-311+G(2d,p), aug-cc-pVDZ and aug-cc-pVTZ) on the macroscopic behaviour of the simulated system. In total, nine sets of multipole moments were generated at all possible levels of theory. The thermodynamic and dynamic properties studied were affected by these two factors but the local structure (as seen in SDFs) was almost indifferent to the level of theory. The LJ parameter values were optimised by carrying out simulations, in which all  $\sigma$  values were incrementally increased by one single scaling factor. Using optimised LJ parameters, we have carried out MD simulations at four different temperatures. The behaviour of the liquid’s density against temperature was in reasonable agreement with the available experimental data. The analysis of the local environment surrounding an imidazole molecule revealed the existence of hydrogen-bonded dimers and to a much lesser extent the presence of  $\pi$ -stacked dimers. This conclusion was supported by the geometries of such dimers optimised in the gas phase using *ab initio* calculations. Finally, we contrasted the QCT force field with two point-charge force fields (OPLS-AA and AMBER/FF99SB). This comparison reinforced the importance of using multipole moments because the QCT force field successfully reproduced the directional nature of the hydrogen bond. Consequently, *the local environment as predicted by the point-charge force field is qualitatively different from the QCT force field.* The point charge force field also exaggerated the number of  $\pi$ -stacked dimers appearing in the liquid, and it underestimated the number of hydrogen-bonded dimers. This was also reflected in the differences observed in the RDFs of the three systems.

#### Acknowledgements

We thank Prof. I. Alkorta for his kind help in the gas phase optimisations of the dimer.

#### References

- W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
- W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225.
- W. L. Jorgensen and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1988, **110**, 1657.
- S. Y. Liem, P. L. A. Popelier and M. Leslie, *Int. J. Quantum Chem.*, 2004, **99**, 685.
- C. Sagui, L. G. Pedersen and T. A. Darden, *J. Chem. Phys.*, 2004, **120**, 73.
- S. L. Price, S. Hamad, A. Torrisi, P. G. Karamertzanis, M. Leslie and C. R. Catlow, *Mol. Simul.*, 2006, **32**, 985.
- J. Pranata, S. G. Wierschke and W. L. Jorgensen, *J. Am. Chem. Soc.*, 1991, **113**, 2810.
- P. Ren and J. W. Ponder, *J. Phys. Chem. B*, 2003, **107**, 5933.
- P. J. Winn, G. G. Ferenczy and C. A. Reynolds, *J. Phys. Chem. A*, 1997, **101**, 5437.
- S. W. Brodersen, F. J. J. Leusen and G. Engel, *Phys. Chem. Chem. Phys.*, 2003, **5**, 4923.
- G. G. Ferenczy, P. J. Winn and C. A. Reynolds, *J. Phys. Chem. A*, 1997, **101**, 5446.
- U. Koch and E. Egert, *J. Comput. Chem.*, 1995, **16**, 937.
- T. A. Halgren and W. Damm, *Curr. Opin. Struct. Biol.*, 2001, **11**, 236.
- D. E. Williams, *J. Comput. Chem.*, 2001, **22**, 1154.
- S. L. Price, *CrystEngComm*, 2004, **6**, 344.
- V. Marcon and G. Raos, *J. Phys. Chem. B*, 2004, **108**, 18053.
- P. A. Hunt and T. Welton, *Chem.–Eur. J.*, 2006, **12**, 6762.
- P. Karamertzanis, A. V. Kazantsev, G. A. Welch, C. S. Adjiman, C. C. Pantelides and S. L. Price, *J. Chem. Theory Comput.*, 2009, **5**, 1432.
- P. Karamertzanis, P. Raiteri and A. Galindo, *J. Chem. Theory Comput.*, 2010, **6**, 1590.
- P. Karamertzanis, P. Raiteri, M. Parrinello, M. Leslie and S. L. Price, *J. Phys. Chem. B*, 2008, **112**, 4928.
- H. Chen, T. Yan and G. A. Voth, *J. Phys. Chem. A*, 2009, **113**, 4507.
- A. J. Stone and S. L. Price, *J. Phys. Chem.*, 1988, **92**, 3325.
- L. Gontrani, R. Caminiti, L. Bencivenni and C. Sadun, *Chem. Phys. Lett.*, 1999, **301**, 131.
- A. A. El-Azhary, *Spectrochim. Acta, Part A*, 2003, **59**, 2009.
- J. Ghosh, *A Textbook of Pharmaceutical Chemistry*, S. Chand & Co Ltd, New Delhi, India, 2003.
- P. Edwards, *Drug Discovery Today*, 2001, **6**, 1072.
- H. Kim and J. Jang, *Polymer*, 2000, **41**, 6553.
- E. S. Selezneva, Z. P. Belousova, L. A. Gusak, E. A. Zvyagina and P. P. Purygin, *Pharm. Chem. J.*, 1992, **26**, 259.
- A. R. Porter, S. Y. Liem and P. L. A. Popelier, *Phys. Chem. Chem. Phys.*, 2008, **10**, 4240.
- T. Welton, *Chem. Rev.*, 1999, **99**, 2071.
- P. A. Hunt, *Mol. Simul.*, 2006, **32**, 1.
- P. Karamertzanis and C. C. Pantelides, *Mol. Simul.*, 2004, **30**, 413.
- S. L. Price, *Phys. Chem. Chem. Phys.*, 2008, **10**, 1996–2009.
- J.-P. Piquemal, R. Chelli, P. Procacci and N. Gresh, *J. Phys. Chem.*, 2007, **111**, 8170.
- C. Sagui, L. G. Pedersen and T. A. Darden, *J. Chem. Phys.*, 2004, **120**, 73.
- P. Ren and J. W. Ponder, *J. Phys. Chem. B*, 2003, **107**, 5933.
- M. G. Darley, C. M. Handley and P. L. A. Popelier, *J. Chem. Theory Comput.*, 2008, **4**, 1435.
- R. F. W. Bader, *Atoms in Molecules. A Quantum Theory*, Oxford University Press, Oxford, Great Britain, 1990.
- P. L. A. Popelier, *Atoms in Molecules. An Introduction*, Pearson Education, London, Great Britain, 2000.
- C. F. Matta and R. J. Boyd, *The Quantum Theory of Atoms in Molecules. From Solid State to DNA and Drug Design*, Wiley-VCH, Weinheim, Germany, 2007.
- P. L. A. Popelier, F. M. Aicken and S. E. O’Brien, in *Chemical Modelling: Applications and Theory*, ed. A. Hinchliffe, Royal Society of Chemistry, Specialist Periodical Report, 2000, vol. 1, ch.3, pp. 143–198.
- C. F. Matta, *Quantum Biochemistry: Electronic Structure and Biological Activity*, Wiley-VCH, Weinheim, Germany, 2010.
- P. Popelier, M. Rafat, M. Devereux, S. Y. Liem and M. Leslie, *Lect. Ser. Comput. Comput. Sci.*, 2005, **4**, 1251.
- P. L. A. Popelier and E. A. G. Bremond, *Int. J. Quantum Chem.*, 2009, **109**, 2542.
- P. L. A. Popelier, *Mol. Phys.*, 1996, **87**, 1169.

- 46 A. J. Stone, *The Theory of Intermolecular Forces*, Clarendon, Oxford, Great Britain, 1996.
- 47 P. L. A. Popelier, A. J. Stone and D. J. Wales, *Faraday Discuss.*, 1994, **97**, 243.
- 48 D. S. Kosov and P. L. A. Popelier, *J. Chem. Phys.*, 2000, **113**, 3969.
- 49 D. S. Kosov and P. L. A. Popelier, *J. Phys. Chem. A*, 2000, **104**, 7339.
- 50 P. L. A. Popelier, L. Joubert and D. S. Kosov, *J. Phys. Chem. A*, 2001, **105**, 8254.
- 51 L. Joubert and P. L. A. Popelier, *Phys. Chem. Chem. Phys.*, 2002, **4**, 4353.
- 52 P. L. A. Popelier and A. J. Stone, *Mol. Phys.*, 1994, **82**, 411.
- 53 M. in het Panhuis, P. L. A. Popelier, R. W. Munn and J. G. Angyan, *J. Chem. Phys.*, 2001, **114**, 7951.
- 54 S. Y. Liem and P. L. A. Popelier, *J. Chem. Theory Comput.*, 2008, **4**, 353.
- 55 S. Liem and P. L. A. Popelier, *J. Chem. Phys.*, 2003, **119**, 4560.
- 56 M. S. Shaik, M. Devereux and P. L. A. Popelier, *Mol. Phys.*, 2008, **106**, 1495–1510.
- 57 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr, T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.
- 58 J. H. Griffiths, A. Wardley, V. E. Williams, N. L. Owen and J. Sheridan, *Nature*, 1967, **216**, 1301.
- 59 D. Christen, J. H. Griffiths and J. Sheridan, *Z. Naturforsch., A: Phys. Phys. Chem. Kosmophys.*, 1982, **37**, 1378.
- 60 M. A. Spackman, *Chem. Rev.*, 1992, **92**, 1769.
- 61 P. L. A. Popelier, *Comput. Phys. Commun.*, 1996, **93**, 212.
- 62 P. L. A. Popelier, *Comput. Phys. Commun.*, 1998, **108**, 180.
- 63 M. Leslie, *Mol. Phys.*, 2008, **106**, 1567.
- 64 W. Smith and T. R. Forester, *J. Mol. Graphics*, 1996, **14**, 136.
- 65 W. Smith, M. Leslie and T. R. Forester, *DLPOLY*, CCLRC: Daresbury Lab, Daresbury, Warrington WA4 4AD, England, 2003.
- 66 H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684.
- 67 H. J. C. Berendsen and W. F. Van Gunsteren, in *MD simulations: techniques and approaches in molecular liquids, dynamics and interactions*, ed. W. J. Orville-Thomas, A. J. Barnes and J. Yarwood, Reidel, New York, USA, 1984, pp. 475–500.
- 68 W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 5179.
- 69 M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Clarendon, Oxford, Great Britain, 1987.
- 70 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins: Struct., Funct., Bioinf.*, 2006, **65**, 712.
- 71 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089.
- 72 I. INCHEM, Imidazole, <http://www.inchem.org/pages/about.html>, 2008.
- 73 N. A. McDonald and W. L. Jorgensen, *J. Phys. Chem. B*, 1998, **102**, 8049.
- 74 M. P. Allen and D. J. Tildesley, *Computer Simulations of Liquids*, Oxford University Press, Oxford, Great Britain, 1993.
- 75 I. M. Svishchev and P. G. Kusalik, *J. Chem. Phys.*, 1993, **99**, 3049.
- 76 M. P. Waller, A. Robertazzi, J. A. Platts, D. E. Hibbs and P. A. Williams, *J. Comput. Chem.*, 2006, **27**, 491.
- 77 M. O. Sinnokrot and C. D. Sherrill, *J. Phys. Chem. A*, 2006, **110**, 10656.
- 78 T. van Mourik and R. J. Gdanitz, *J. Chem. Phys.*, 2002, **116**, 9620.
- 79 M. S. Shaik, S. Y. Liem and P. L. A. Popelier, *J. Chem. Phys.*, 2010, **132**, 174504, 174501–174513.
- 80 D. R. Lide and H. P. R. Frederikse, *CRC Handbook of Chemistry and Physics*, CRC press, Boca Raton, FL, USA, 78th edn, 1997.
- 81 H. G. M. de Wit, C. G. de Kruif and J. C. van Miltenburg, *J. Chem. Thermodyn.*, 1983, **15**, 891.
- 82 N. A. McDonald and W. L. Jorgensen, *J. Phys. Chem. B*, 1998, **102**, 8049.