Molecular dynamics simulations of structure and dynamics of organic molecular crystals†

Alexandra Nemkevich,^a Hans-Beat Bürgi,^{ab} Mark A. Spackman^a and Ben Corry*^a

Received 3rd August 2010, Accepted 10th September 2010 DOI: 10.1039/c0cp01409e

A set of model compounds covering a range of polarity and flexibility have been simulated using GAFF, CHARMM22, OPLS and MM3 force fields to examine how well classical molecular dynamics simulations can reproduce structural and dynamic aspects of organic molecular crystals. Molecular structure, crystal structure and thermal motion, including molecular reorientations and internal rotations, found from the simulations have been compared between force fields and with experimental data. The MM3 force field does not perform well in condensed phase simulations, while GAFF, CHARMM and OPLS perform very similarly. Generally molecular and crystal structure are reproduced well, with a few exceptions. The atomic displacement parameters (ADPs) are mostly underestimated in the simulations with a relative error of up to 70%. Examples of molecular reorientation and internal rotation, observed in the simulations, include in-plane reorientations of benzene, methyl rotations in alanine, decane, isopropylcyclohexane, pyramidal inversion of nitrogen in amino group and rotation of the whole group around the C-N bond. Frequencies of such dynamic processes were calculated, as well as thermodynamic properties for reorientations in benzene and alanine. We conclude that MD simulations can be used for qualitative analysis, while quantitative results should be taken with caution. It is important to compare the outcomes from simulations with as many experimental quantities as available before using them to study or quantify crystal properties not available from experiment.

Introduction

Thermal motion, in the form of atomic vibrations and reorientations of part or all of a molecule, affects a broad range of bulk crystal properties, such as specific heat, thermal expansion, conduction of heat, diffusion into and out of porous and void-containing organic and metal-organic materials, flexibility (of polymers), ability to withstand stress, solid-state reactivity etc.

Most of our knowledge of molecular motion in crystals comes from various kinds of diffraction, solid state NMR, IR and Raman experiments. Standard Bragg diffraction experiments give the structure of an average crystal unit cell in the form of mean positions and atomic displacement parameters (ADPs). Spectroscopic experiments provide information on molecular structure and indications of mostly intramolecular deformation and reorientation processes. However, without detailed interpretation of these experimental results neither technique gives direct information on the relative phases of atomic motions.

Molecular dynamics (MD) provides an alternative method of investigating motion in crystals. It does provide direct information on the displacements, types of intramolecular reorientations and frequencies of such processes, since the motion of individual atoms and molecules can be monitored. It has already been applied to a number of tasks in crystallography, such as macromolecular structure refinement,¹⁻³ describing crystal formation^{4–7} and determining the stability of crystal polymorphs.8

There are several reports on using MD trajectories for calculating ADPs and comparing them to experimental values.⁹⁻¹² The systems investigated are as different as deuterated ammonia,¹¹ a crown ether9 and proteins.^{10,12} Computational methods include Car-Parinello MD,¹¹ classical MD using an AMBER force field,⁹ a CHARMM force field¹⁰ or a collection of force fields -AMBER, CHARMM, OPLS and GROMOS.¹²

Classical MD has also been used for simulating reorientations of part of a molecule in crystals. Chatfield and Wong investigated methyl rotation in crystalline alanine¹³ and other amino acids.¹⁴ They found that a modification of a torsion angle parameter in the CHARMM22 force field was needed to reach agreement between MD and NMR correlation times for methyl reorientations. Dittrich and coworkers compared variable-temperature X-ray diffraction data of 2-aminoisobutyrate hydrochloride with MD simulations using NAMD and CHARMM27 force field.¹⁵ They found methyl rotation for one of three methyl groups, the one which is disordered in the crystal. The literature contains many more molecular dynamics studies, often on specific problems and based on force fields tailor-made for the purpose. We are not reviewing this work here.

Given that the number of such studies is comparatively small, it is still difficult to asses the reliability of MD

^a School of Biomedical, Biomolecular and Chemical Sciences, University of Western Australia, Perth, Australia. E-mail ben corry@uwa edu au

^b Laboratory of Crystallography, University of Berne, Berne and Institute of Organic Chemistry, University of Zurich, Zurich, Switzerland

[†] Electronic supplementary information (ESI) available: Further details of the data analysis and data from all simulations; details of sizes of the simulated systems; simulated and experimental cell constants; bond lengths and equivalent ADPs for model compounds; AGA used in the preliminary study. See DOI: 10.1039/c0cp01409e

simulations at reproducing the dynamic aspects of molecules in crystals. Furthermore, the published literature gives no indication of the most appropriate force field to use in various situations nor the best way of doing MD simulations of thermal motion in molecular crystals. Here we aim to begin to address these issues by examining how well classical MD simulations employing four different force fields reproduce structure and dynamics in organic molecular crystals. To the extent that such MD simulations prove to be suitable for simulating organic molecular crystals, they could be used for many purposes. These could include refinement of disordered crystal structures including large biomolecules; obtaining anisotropic ADPs for weakly scattering hydrogens in X-ray structure determination; explaining some interesting physical phenomena, like absorption of gases by crystals having a potential for molecular motion, and aiding the design of materials showing these phenomena; giving insight into processes occurring inside host-guest systems, such as clathrates, and crystals with continuous pores; better understanding of drug-receptor interactions and reactivity in the solid state; helping to design molecular motors and many more.

Simulation details

Simulation setup

The model compounds chosen for this study cover a range of flexibility and polarity (Fig. 1) and feature high quality crystal structure data, taken from the Cambridge Structural Database (CSD) (Table 1).

Unit cells of all model compounds were replicated in three dimensions to make a supercell of approximately cubic shape. All simulations were performed with periodic boundary conditions in the NPT ensemble at atmospheric pressure and the temperature reported in the CIF files of the crystal structures.



Fig. 1 Model compounds used in this study. The molecules range in flexibility and polarity.

The choice of other simulation parameters is described in detail below.

The force fields tested are: GAFF,²⁶ CHARMM22,²⁷ all-atom OPLS²⁸ and MM3.²⁹⁻³¹ GAFF, as seen from the name (General Amber Force Field) was developed to cover a wide range of organic species. CHARMM22 was designed for simulating large biomolecular systems, but has been benchmarked on smaller molecules and dipeptides. The $OPLS^{28}$ (Optimised Potential for Liquid Simulations) force field was parameterised to reproduce various properties of simple organic liquids and is well known for its non-bonded parameters. The MM3 force field was shown to very accurately reproduce structures and energetic properties of small isolated organic molecules, but it has also been used in condensed phase simulations.³² All four force fields are widely used in MD simulations for different purposes and are parameterised for a large variety of organic molecules, making them an ideal starting point for our study. Simulations using GAFF, CHARMM and OPLS were performed using the MD code NAMD, version 2.6,³³ while MM3 force field simulations were performed using the Tinker molecular modelling package, version 4.2.34

All systems were energy-minimised before the simulation. The minimisation part of the trajectory was not used for analysis. In the dynamics part of the trajectory the systems equilibrate (atom positions reach their average *rmsd* per frame) within less than 1 ps, which is so small compared to the full length of the trajectory, that deleting this initial equilibration period was not deemed necessary before analysing the trajectory. Molecular graphics software VMD³⁵ was used to visualise all the trajectories. Any motion of a supercell as a whole was eliminated by aligning all frames of the trajectory to the first one using VMD's RMSD trajectory tool.

Conformational changes and molecular reorientations were studied with the method of umbrella sampling as implemented in Plumed, a plugin for free energy calculations.³⁶ Results were analysed using an implementation of the Weighted Histogram Analysis Method^{37,38} written by Alan Grossfield.³⁹

Data analysis

From the trajectories generated by the MD simulations we obtained average atomic positions, distances between average positions, average bond lengths, average lengths of unit cells and ADPs (given in Tables S1–S11 of ESI) using the methods described below.

Normally average positions of atoms were obtained as the mean of their coordinates throughout the trajectory, however special treatment was required for the trajectories where molecules exhibit reorientations and internal rotations. Each such molecule was followed in each frame of a trajectory to find the angle to which the molecule (benzene), or atoms (methyl hydrogens, amino group hydrogens) rotated, the atoms were renumbered according to this angle and a new trajectory without reorientations was written out. This new trajectory was used to obtain average structures and ADPs.

Bond lengths were obtained in two ways: from average atomic positions and as average distances between instantaneous positions. Bond lengths from average atomic positions

Table 1 Details of the crystal structures used in our simulations, including standard uncertainties (s.u.) in bond lengths and ADPs

Compound	<i>Т</i> , К	Radiation	CSD refcode	R factor	s.u. bonds, Å	s.u. ADPs, Å ²	Reference
Deuterated benzene	123	Neutron	BENZEN07	0.032^{a}	$(1.1-1.3) \times 10^{-3}$	$(3-5) \times 10^{-4}$	16
Benzene	138	Neutron	BENZEN01	0.059	$(7-10) \times 10^{-3}$	<i>b</i>	17
Benzene	218	Neutron	BENZEN	0.078	$(9-17) \times 10^{-3}$	b	17
Isopropylcyclohexane	150	X-ray	PUYWEL	0.0608	3×10^{-3}	$(9-11) \times 10^{-4}$	18
n-decane	150	X-ray	QQQFBG01	0.0541	$(1.5-2.0) \times 10^{-3}$	$(4-5) \times 10^{-4}$	19
Imidazole	103	Neutron	IMAZOL06	0.026	$(1-2) \times 10^{-3}$	$(2-7) \times 10^{-4}$	20 and 21
1,7-diaminoheptane	130	X-ray	ROKZOG01	0.0317	$(0.9-1.1) \times 10^{-3}$	$(2-3) \times 10^{-4}$	22
1,7-diaminoheptane	213	X-ray	ROKZOG	0.0493	4×10^{-3}	$(11-20) \times 10^{-4}$	23
D-alanine	60	Neutron	ALUCAL05	0.0799	$(3-19) \times 10^{-3}$	$(5-19) \times 10^{-4}$	24
D-alanine	295	Neutron	ALUCAL04	0.0844	$(3-12) \times 10^{-3}$	$(3-30) \times 10^{-4}$	24
L-alanyl-glycyl-L-alanine	293	X-ray	WIRNIU	0.0279	$(2-3) \times 10^{-3}$	$(3-6) \times 10^{-4}$	25
^{<i>a</i>} Calculated as $\frac{wR[F^2]}{2}$. ^{<i>b</i>} n.a	. but ADP	s given to 10 $ imes$	10^{-4} .				

are defined as distances between average positions of bonded atoms and what we call average bond lengths are instantaneous bond lengths from each frame, averaged throughout the trajectory. Crystallographically equivalent bond lengths were averaged over all molecules in the system and standard uncertainties were calculated. The standard uncertainties of bond lengths from average positions reflect the spread of values for a particular bond throughout the average simulated structure. Bond lengths from average positions and their standard uncertainties are directly comparable to X-ray and neutron diffraction results. In our simulations the standard uncertainties of bond lengths from average positions are typically around $(4-6) \times 10^{-4}$ Å, which is similar to or smaller than the corresponding standard uncertainties of diffraction data (Table 1). In the case of average bond lengths the standard uncertainty represents the spread of values for a bond length throughout the trajectory, which is due to stretching deformations of the bonds, *i.e.* they represent root mean square amplitudes of stretching vibration. These values can be compared to experimental data from electron diffraction. From the simulations the root mean square amplitudes are \sim (1-3) \times 10⁻² Å. This is 2-3 times smaller than the experimental values. Definitions of average atomic coordinates, bond lengths from average positions and average bond lengths are given in section S.1 of ESI.

Average lengths of unit cell edges were obtained by dividing the average lengths of the supercell by the number of unit cells in the corresponding direction. Standard uncertainties measure the fluctuations of the size of the system during the simulation due to pressure control. In our simulations the standard deviations for unit cell lengths are typically around $(5-15) \times 10^{-3}$ Å. The largest value is found for the largest unit cell dimension of diaminoheptane (26×10^{-3} Å for the 22.271 Å average length of *a* at 130 K as simulated by GAFF, and 34×10^{-3} Å for the 22.234 Å average length at 213 K as simulated by OPLS.)

The atomic displacement parameters (ADPs) for each atom can be calculated from MD trajectories as elements of the covariance matrix:

$$U_{ij} = \langle (X_i - \langle X_i \rangle)(X_j - \langle X_j \rangle) \rangle, \tag{1}$$

where X_i and X_j are the x, y or z coordinates in a Cartesian coordinate system, and angular brackets denote averaging over time, *i.e.* over all frames of the trajectory.

We compared the values of equivalent isotropic ADPs, which in orthogonal coordinate systems are equal to one third of the trace of the atomic displacement matrix U_{ij}^C and represent an average mean square displacement of the atom over all directions.

$$U_{\rm eq} = \frac{1}{3} (U_{11}^C + U_{22}^C + U_{33}^C)$$
(2)

where U_{11}^C , U_{22}^C and U_{33}^C are the mean square displacements in the *x*, *y* and *z* directions, respectively.

Mean square displacements of atoms converge well during our simulations, differences between values for symmetry equivalent atoms in the same molecule do not exceed $3 \times 10^{-4} \text{ Å}^2$ and are typically around $1 \times 10^{-5} \text{ Å}^2$. Standard deviations of such values, averaged over the same type of atoms in different molecules, are typically around $(2-5) \times 10^{-4} \text{ Å}^2$, which is comparable to or smaller than experimental uncertainties from a diffraction experiment.

Average values from simulations determined in our analysis are compared to experimental values in terms of three quantities: $\overline{\Delta}$, σ and δ . The quantity $\overline{\Delta}$ is the mean difference between two sets of N numbers:

$$\Delta_i = a_{i,\text{calc}} - a_{i,\text{exp}} \tag{3}$$

$$\bar{\Delta} = \frac{\sum_{i=1}^{N} \Delta_i}{N} \tag{4}$$

A positive/negative value of $\overline{\Delta}$ indicates that the simulated value of *a* is systematically too large/too small. The quantity σ is the rms deviation of the individual values of Δ_i and serves as a measure of the spread of values of absolute differences:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N} \left(\Delta_i - \bar{\Delta}\right)^2}{N}} \tag{5}$$

The relative percent difference between each pair of numbers from two sets is defined as

$$\delta = 100 \frac{a_{i,\text{calc}} - a_{i,\text{exp}}}{a_{i,\text{exp}}} \tag{6}$$

The percent difference between two sets is then the average of all percent differences between their components. The percent difference preserves the sign of the direction of change (analogous to eqn (4)). We also refer to the quantity *rmsd* as a measure of the difference between two structures:

$$rmsd = \sqrt{\frac{\sum_{i=1}^{N} \| \mathbf{R}_{i}^{1} - \mathbf{R}_{i}^{2} \|^{2}}{N}},$$
 (7)

where N is the total number of atoms in each structure, \mathbf{R}_{i}^{1} and \mathbf{R}_{i}^{2} are atomic positions in two structures being compared.

All trajectories, including those generated by Tinker, were analysed using VMD,³⁵ a set of programs written in Fortran90 and several tel scripts.

Choice of simulation parameters

A preliminary study was performed to establish the simulation parameters that achieve reliable results at a minimal cost of computer power. To this end L-alanyl-glycyl-L-alanine (further referred to as AGA) was simulated using NAMD and CHARMM22 for a range of systems with differing numbers of molecules (*i.e.* system size), different cutoff for non-bonded interactions and different time lengths of simulation. The timestep in all simulations was 1 fs. The results of the simulations were assessed based on the average atomic positions and the average equivalent isotropic ADPs of the atoms (eqn (2)).

To assess the impact of changing the simulation length, a system containing 64 molecules of AGA was simulated for 500 ps, 1 ns, 2 ns and 10 ns. The *rmsd* of atom positions compared to the first time step was found to plateau within 1 ps after minimisation, so the length of the simulation had no effect on either average atomic positions or the equivalent isotropic ADPs of the atoms.

The effect of changing the size of the simulation system was monitored by examining systems containing 12, 64, 96 and 768 molecules, each simulated for 2 ns. In the small 12 molecule system an 8 Å cutoff for non-bonded interactions has been used, while in all other cases it was set to 12 Å, the standard NAMD cutoff for non-bonded interactions. Average atomic positions change upon increasing the system from 12 to 64 molecules (and the cutoff from 8 to 12 Å), but remain the same upon further increase of system size. Similarly, ADPs increase by about 25% on increasing the system from 12 to 64 molecules, but increase only about 10% more on going from 64 to 768 molecules in the system. This small increase is probably due to a better description of phonon modes of the crystal in the larger systems. To test if further increasing the size of the system leads to mean square displacements converging at some point we conducted a simulation of 3192 molecules of AGA for 500 ps: the average values of equivalent

ADPs increase on average by about 0.0005 Å² compared to the 768 molecules system, which is an increase of ~2%. Their standard uncertainties increase 3 to 4 times—from 0.0002 to 0.0008 Å², for example (Fig. S2 of ESI). Thus we consider the ADPs converged and the smallest system size allowing for NAMD's standard cutoff of 12 Å sufficient to assess thermal motion in a system.

Increasing the cutoffs for non-bonded interactions from 12 to 16, 20 and 24 Å in a 768 molecule system had no effect on the average positions and the ADPs.

Unit cell parameters, bond lengths and equivalent ADPs from all simulations in this preliminary study are given in Tables S12–S15 of ESI.

In light of these results, in the remainder of the simulations reported in this study we have used a cutoff of 12 Å, set each system dimension to just over twice the size of the cutoff and used a simulation time of 2 ns unless otherwise specified. The standard uncertainties for unit cell parameters, bond lengths from average positions and ADPs obtained from simulations with these parameters are on the order of magnitude or smaller than the standard uncertainties of experimental values. The comparison between simulated and experimental data is thus limited by the uncertainties in the experimental data.

Results and discussion

Crystal structure

From visual inspection, the packing patterns of the experimental crystal structures were reproduced well in most simulations, e.g. benzene (Fig. 2a and b). In several cases significant differences from the experimental crystal structures were observed. In the case of imidazole simulated by CHARMM, we observed a phase transition. In this new form pairs of planes are shifted relative to each other as shown in Fig. 3. This structure shows the same space group as the known crystal structure ($P2_1/c$), but different cell constants: a = 6.84 Å, b = 5.28 Å, c = 9.62 Å and $\beta = 105.85^{\circ}$. The difference in potential energy between this form and the experimental crystal structure is only $\sim 1 \text{ kJ mol}^{-1}$. We found no evidence of this crystal form of imidazole in the literature, however previous polymorph prediction studies found a large number of structures within 10 kJ mol⁻¹ of the global minimum.⁴⁰ Angles of the simulation supercell were kept fixed in our simulations. In the case of imidazole, simulated by CHARMM, the phase change manifested itself in protrusions into adjoining periodic boxes.



Fig. 2 Comparison of structures of benzene at 218 K. Crystal structure is preserved in (b) and not preserved in (c).



Fig. 3 Comparison of the neutron diffraction crystal structure of imidazole and its average structure from a CHARMM22 simulation. The double layer pattern that is preserved is highlighted with dashed lines.

In the case of AGA the molecules underwent large conformational changes in simulations conducted with GAFF and MM3: in MM3 they curled and in GAFF half of them curled and half stretched (Fig. 4). Some of the average structures were disordered, e.g. AGA, simulated by GAFF (Fig. 4b) and alanine, simulated by MM3 (Fig. 5). In benzene simulated by MM3 the crystal structure was lost completely (Fig. 2c). In many other simulations conducted with MM3 the molecules move considerably, making it very hard or impossible to obtain reasonable average structures. For example in simulations of deuterated benzene, benzene at 138 K and imidazole, the planes of the molecules librate substantially about their centres of mass, the average atomic positions move towards the centre of the molecules and the bond lengths become unreasonably short. We exclude these structures from further analysis.

Simulating molecular packing is primarily a test of the force field's intermolecular interactions. A decreased unit cell size indicates that the intermolecular interactions are too strong, while an enlarged unit cell means the intermolecular interactions are too weak. Table 2 lists mean absolute and relative percent differences between average simulated and experimental unit cell lengths for structures that were reproduced well. The sign indicates the direction of change. In Table 2 no σ s are given as each average involves only 3 unit cell lengths.

The values in Table 2 are mostly negative, except for benzene, alanine simulated by GAFF, AGA simulated by CHARMM and OPLS, decane simulated by CHARMM and all MM3 simulations. Thus the cell size is generally underestimated, implying that on average the distances corresponding to the minima of the intermolecular potentials are too short. Since the absolute differences do not change much with temperature for a given compound, thermal expansion is reproduced well. The expansion or shrinkage of the simulated cell in different directions is quite uniform with only one exception: imidazole simulated by OPLS. In this case the changes in the length of the different cell dimensions even have different signs – the simulated unit cell length is too long along a and too short along b and c. There is no obvious reason for this anisotropy, but it is worth noting that the direction of hydrogen bonding in imidazole is along c.

For a more quantitative comparison between simulated and experimental crystal structures we align each simulated unit cell of the supercell with the experimental unit cell and calculate the *rmsd* of atom positions (Table 3). As the compounds in Table 3 are listed in order of increasing polarity, it can be seen that in the non-polar part of the table, GAFF performs better than the other force fields, but as the polarity of the compounds increases, CHARMM starts performing better. OPLS is somewhere between the two. In the few cases in which MM3 reproduces packing patterns, its performance is similar to that of CHARMM.

Molecular structure in the crystal

To see if the geometry of the molecules is reproduced well, we compared simulated bond lengths from average positions with experimental X-ray and neutron diffraction values (Table 4). Deviations of simulated values from experiment ($\overline{\Delta}$) are often within diffraction experimental error (Table 1).

Average bond lengths are very close to their equilibrium values in the corresponding force field parameter files, which is not surprising. Given the relatively high bond stretching force constants the influence of the intermolecular interactions on bond lengths is expected to be small; we find differences of $\sim (1-7) \times 10^{-3}$ Å (simulated values are larger). The mean amplitudes of vibration of aliphatic C–C and C–H bonds in our simulations depend on temperature, but even at room temperature they are smaller than the values from gas electron diffraction⁴¹ by 0.025 Å and by 0.048 Å respectively and vary insignificantly between bonds involving heavy atoms and hydrogens.

All bond lengths found in the simulations are given in the ESI.

While the comparison of bond lengths gives a good first indication of the accuracy of the force fields, it is also



(a) Crystal structure

(b) Average structure, GAFF

(c) Average structure, MM3

Fig. 4 Comparison of the X-ray crystal structure of AGA and average structures simulated by GAFF and MM3. It is seen that crystal structure is not preserved during these simulations.



Fig. 5 Comparison of crystal structure of alanine at 60 K with average structures at 60 K and at 295 K from simulations using MM3. Translational symmetry is not preserved.

 Table 2
 Absolute and relative differences between average unit cell lengths from MD simulations using GAFF, CHARMM, OPLS and MM3 and average unit cell lengths from diffraction experiments. Only simulations which reproduced packing patterns are included

	$\bar{\Delta}, \text{\AA}$				δ (%)			
Molecule	GAFF	CHARMM	OPLS	MM3	GAFF	CHARMM	OPLS	MM3
d ₆ -benzene 123 K	0.008	0.057	0.016		0.1	0.7	0.2	_
Benzene 138 K	0.019	0.067	0.027	_	0.3	0.9	0.4	_
Benzene 218 K	0.033	0.078	0.037	_	0.4	1.0	0.5	
Isopropylcyclohexane 150 K	-0.048	-0.050	-0.125	0.264	-0.5	-0.5	-1.2	2.5
n-decane 150 K	-0.019	0.002	-0.062	0.214	-0.3	0.0	-0.8	2.9
Imidazole 103 K	-0.040	_	-0.002	_	-0.5	_	-0.03	
Diaminoheptane 130 K	-0.277	-0.149	-0.291	0.200	-2.4	-1.3	-2.5	1.7
Diaminoheptane 213 K	-0.287	-0.148	-0.293	0.222	-2.5	-1.3	-2.5	1.9
Alanine 60 [°] K	0.039	-0.153	-0.163	_	0.5	-1.9	-2.0	
Alanine 295 K	0.067	-0.149	-0.159	_	0.8	-1.9	-2.0	
AGA 293 K	_	0.015	0.042	_	_	0.1	0.4	

Table 3 *rmsd* between atomic positions in the average unit cell from MD simulations using GAFF, CHARMM, OPLS and MM3 and those in experimental structures. For X-ray structures only heavy atoms were taken into account. Values represent averages over all unit cells in a supercell. Standard uncertainties of these averages are on the order of $10^{-4}-10^{-3}$. All values are in Å

	rmsd						
Molecule	GAFF	CHARMM	OPLS	MM3			
d ₆ -benzene 123 K	0.053	0.154	0.119	_			
Benzene 138 K	0.053	0.153	0.116	_			
Benzene 218 K	0.070	0.209	0.155	_			
Isopropylcyclohexane 150 K	0.077	0.093	0.110	0.128			
n-decane 150 K	0.020	0.012	0.013	0.007			
Imidazole 103 K	0.152		0.124	_			
Diaminoheptane 130 K	0.152	0.094	0.162	0.108			
Diaminoheptane 213 K	0.157	0.098	0.164	0.129			
Alanine 60 K	0.290	0.129	0.248	_			
Alanine 295 K	0.312	0.131	0.250	_			
AGA 293 K	_	0.178	0.352	_			

important to consider how the bonds are positioned relative to one another. A simple way to do this is to align each molecule of the average simulated structure to the corresponding molecule from the experimental structure. The *rmsd* of atom positions is a measure of the overall accuracy of the molecular geometry and includes the discrepancies for bonds, angles and dihedrals together (Table 5).

Generally the *rmsds* in Table 5 increase with increase in polarity, and for a given compound there is a slight increase with temperature. GAFF, CHARMM and OPLS mostly perform very similarly, differing in the polar region, where CHARMM

performs better than other force fields. The largest individual discrepancies are discussed below.

The bigger *rmsd* for isopropylcyclohexane in the CHARMM simulation is caused mainly by the sterically crowded ring carbon-isopropyl carbon bond which is calculated to be 1.524 Å rather than 1.557 Å as observed experimentally.

Decane simulated by GAFF has longer C–C bonds, but C–C–C angles smaller by 2° , thus shortening the whole molecule by 0.3 Å. This is not the case with CHARMM, OPLS and MM3. Although the molecular structure is similar with all force fields, differences are apparent in the relative error of the unit cell size (Table 2). This difference must be a result of slight differences in the intermolecular distances created by differing intermolecular interactions.

The main contribution to the molecular *rmsd* for diaminoheptane comes from a change in the N–C1–C2 angle: it is about 5° smaller than the diffraction value for GAFF and OPLS and about 6° smaller for CHARMM, resulting in a shortening of the molecule.

In alanine simulated at 60 K with GAFF, the angle between the carboxyl oxygens changes from the experimental value of 125.77(15)° to 118.68(\pm 1.30)°, the whole group rotates by about 17° around the C–C bond and the angle between the nitrogen, α -C and carboxyl C changes from 109.79(8)° to 119.70(\pm 1.32)° (see Fig. 6a). The differences are smaller for OPLS and even smaller for CHARMM (see Fig. 6) and this is reflected in the relative size of the *rmsd* seen in Table 5. Similar differences are observed for the alanine structure at 295 K.

Table 4Accuracy of reproducing bond lengths from average positions of different molecules by MD simulations using GAFF, CHARMM22,OPLS and MM3. All values are in Å

	$ar{\Delta}$			σ					
Compound	GAFF	CHARMM	OPLS	MM3	GAFF	CHARMM	OPLS	MM3	
d ₆ -benzene 123 K	-0.005	-0.005	-0.002	_	0.002	0.005	0.009	_	
Benzene 138 K	-0.007	-0.008	-0.004	_	0.006	0.009	0.011		
Benzene 218 K	-0.008	-0.010	-0.007	_	0.011	0.008	0.009		
Isopropylcyclohexane 150 K	-0.001	-0.012	-0.008	-0.011	0.002	0.008	0.002	0.002	
n-decane 150 K	0.012	0.001	0.003	0.006	0.003	0.003	0.002	0.003	
Imidazole 103 K	-0.015	-0.019	-0.015	_	0.012	0.012	0.012		
Diaminoheptane 130 K	0.009	0.002	-0.001	0.005	0.006	0.001	-0.001	0.003	
Diaminoheptane 213 K	0.009	0.011	-0.001	0.003	0.012	0.014	0.012	0.011	
Alanine 60 K	-0.001	0.006	-0.006	_	0.026	0.013	0.011	_	
Alanine 295 K	-0.001	-0.003	-0.016	_	0.029	0.014	0.018		
AGA 293 K	—	-0.003	0.003	_	_	0.011	0.012	_	

Table 5 *rmsd* between average molecular structures from MD simulations using GAFF, CHARMM, OPLS and MM3 and experimental structures. Numbers are averages over all molecules in the supercell. Standard deviations are on the order of 10^{-8} – 10^{-6} . For X-ray structures only non-hydrogen atoms were taken into account. All values are in Å

	rmsd					
Molecule	GAFF	CHARMM	OPLS	MM3		
d ₆ -benzene 123 K	0.010	0.010	0.007	_		
Benzene 138 K	0.016	0.016	0.013			
Benzene 218 K	0.028	0.029	0.025	_		
Isopropylcyclohexane 150 K	0.012	0.034	0.022	0.033		
n-decane 150 K ^a	0.020	0.012	0.013	0.007		
Imidazole 103 K	0.020	0.026	0.023			
Diaminoheptane 130 K	0.044	0.052	0.054	0.027		
Diaminoheptane 213 K	0.047	0.055	0.056	0.030		
Alanine 60 K	0.202	0.074	0.144	_		
Alanine 295 K	0.215	0.071	0.147	_		
AGA 293 K		0.146	0.313			

^{*a*} For decane the numbers in Tables 5 and 3 (*rmsd* of the unit cell and the molecular structure) are the same, because the unit cell contains only one molecule.

The ability to reproduce molecular geometry should be a test of intramolecular parameters of the force fields, but since this is a condensed phase simulation, intermolecular interactions affect the structure of a single molecule as well, the softer bond and torsion angles in particular. The deformations are most obvious in the case of the extended floppy molecule AGA. MM3 and GAFF distort the molecule substantially (Fig. 4). The best structure is given by the only biomolecule-specific force field in our work—by CHARMM, with OPLS not much worse. Overall the molecular structures of the compounds tested in the crystalline state are reproduced reasonably well by the force fields tested. The exceptions show, however, that one has to be very careful when using MD for simulating structural aspects of crystals.

Atomic vibrations about mean positions

Simulated ADPs are plotted against experimental equivalent ADPs (U_{eq}) in Fig. 7 and differences between them are given in Table 6.

Examination of Fig. 7 and Table 6 shows that the magnitude of thermal motion calculated with GAFF, CHARMM and OPLS is almost always underestimated, more so for hydrogen than for heavy atoms. We have already seen in the section "Molecular structure in the crystal" that the mean amplitudes of bond stretching are too small compared to experiment. Here we see that the total mean square displacements averaged over all directions are also too small. The underestimation is substantial and can be as high as 70%. The biggest deviations in ADPs from experimental values are found for diaminoheptane (all three force fields) and alanine (CHARMM and OPLS). This provides a surprising, difficult to understand contrast with the smaller difference seen in the simulations for AGA with CHARMM and OPLS and alanine with GAFF. For benzene, diaminoheptane and alanine, examples for which data at several temperatures are available, the differences tend to increase with increasing temperature.

MM3 is the only force field that mostly overestimates thermal motion (Table 6, exception: diaminoheptane). Given its poor performance in reproducing crystal structures and the fact that we obtained ADPs only for 4 systems out of 11, we exclude this force field from further analysis.



Fig. 6 Overlay of a molecule from the crystal structure of D-alanine at 60 K (in colour) and average simulated structure (in orange).



Fig. 7 Simulated U_{eq} for heavy atoms plotted against corresponding experimental U_{eq} . Data points for GAFF are given in red, for CHARMM in green and for OPLS in blue. The points fall mostly below the line U_{eq} (exp) = U_{eq} (sim).

In the section "Choice of simulation parameters" we have already shown that the difference between simulated and experimental ADPs cannot be attributed to the limited size of the system simulated. Part of the discrepancy, especially at low temperatures, may arise from the use of a classical model of atomic motion in the simulations rather than the appropriate quantum expression. The mean square amplitude $\langle u^2 \rangle$ of a mass point with mass μ moving with frequency ω in a harmonic potential is:⁴²

$$\langle u^2 \rangle = \frac{\hbar}{2\mu\omega} \coth \frac{\hbar\omega}{2k_{\rm B}T} \tag{8}$$

At low temperature this expression simplifies to the quantum mechanical, temperature-independent zero-point motion amplitude:

$$\langle u^2 \rangle = \frac{\hbar}{2\mu\omega}.\tag{9}$$

At sufficiently high temperatures it simplifies to the corresponding classical picture:

$$\langle u^2 \rangle = \frac{2k_{\rm B}T}{\mu\omega^2},\tag{10}$$

i.e. at very low temperatures the classical expression underestimates $\langle u^2 \rangle$, whereas at sufficiently high temperatures it is a realistic model of atomic vibrations.

There is also a difference in the behaviour of the mean square displacements of heavy atoms and hydrogens, since for

	C, N, C)				H (D)			
Molecule		GAFF	CHARMM	OPLS	MM3	GAFF	CHARMM	OPLS	MM3
d ₆ -benzene 123 K	$\bar{\Delta}, {\rm \AA}^2$	-0.0011	-0.0003	-0.0009		-0.007	-0.004	-0.005	_
	σ , Å ²	0.0001	0.0002	0.0002		0.0004	0.0007	0.0006	
	δ (%)	-4.7	-1.3	-3.9		-16.5	-9.8	-12.3	
Benzene 138 K	$\bar{\Delta}, {\rm \AA}^2$	0.0012	0.0021	0.0016		-0.007	-0.003	-0.004	
	$\sigma, Å^2$	0.0009	0.0009	0.0009		0.004	0.003	0.003	
	δ (%)	5.0	8.6	6.4		-13.0	-6.5	-8.5	
Benzene 218 K	$\bar{\Delta}, \mathrm{\AA}^2$	0.012	0.014	0.012		0.012	0.021	0.0170	
	$\sigma, Å^2$	0.0004	0.0004	0.0004		0.0018	0.0018	0.0016	
	δ (%)	25.4	30.0	25.4	_	14.5	25.9	20.7	
Isopropylcyclohexane 150 K	$\bar{\Delta}, \mathrm{\AA}^2$	-0.012	-0.006	-0.012	0.014				
	$\sigma, Å^2$	0.001	0.001	0.001	0.004				
	δ (%)	-38.6	-20.7	-39.5	43.9				
n-decane 150 K	$\overline{\Delta}, \overset{\circ}{A}^2$	-0.012	-0.006	-0.011	0.009				
	$\sigma, Å^2$	0.001	0.001	0.001	0.004				
	δ (%)	-41.3	-21.3	-38.5	29.0				
Imidazole 103 K	$\bar{\Delta}, \dot{A}^2$	-0.002	_	-0.003	_	-0.013	_	-0.015	
	$\sigma, Å^2$	0.0009	_	0.0008		0.004		0.003	
	$\delta(\%)$	-11.0	_	-19.3		-31.1		-36.9	
Diaminoheptane 130 K	$\bar{\Delta}, \dot{A}^2$	-0.011	-0.008	-0.010	-0.003				
1	σ . Å ²	0.0008	0.0006	0.0005	0.0018				
	δ (%)	-54.4	-37.4	-47.0	-16.1				
Diaminoheptane 213 K	$\bar{\Delta}$, Å ²	-0.033	-0.026	-0.030	-0.014				
I. I. I. I. I.	σ , Å ²	0.002	0.002	0.002	0.005				
	δ (%)	-65.6	-51.1	-59.7	-28.0				
Alanine 60 K	$\bar{\Delta}$, Å ²	-0.004	-0.006	-0.006		-0.017	-0.019	-0.019	
	σ , Å ²	0.0005	0.0008	0.0007		0.002	0.003	0.003	
	δ (%)	-48.4	-71.6	-69.7	_	-70.8	-80.7	-81.5	
Alanine 295 K	$\bar{\Delta}$, Å ²	-0.006	-0.018	-0.017		-0.015	-0.028	-0.029	
	σ , Å ²	0.0013	0.0024	0.0013	_	0.0015	0.0043	0.0038	
	$\delta(\%)$	-21.2	-57.1	-554		-30.3	-53.8	-56.0	
AGA 293 K	$\bar{\Lambda}$, Å ²		-0.009	-0.010					
	σ Å ²	_	0.006	0.007	_				
	$\delta(\%)$	_	-254	-25.8	_				
	5 (70)		20.1	20.0					

Table 6 Comparison of U_{eq} as simulated by GAFF, CHARMM, OPLS and MM3 with experimental diffraction data. For neutron data heavyatoms and hydrogen (deuterium) are analysed separately. For X-ray structures only differences of ADPs for heavy atoms are shown

a given frequency the mean square displacement of an atom depends on its mass μ (eqn (10)). Thus, one would expect that at low temperatures the difference between classically simulated and experimental ADPs will be larger for hydrogens than for heavy atoms. This is indeed found to be the case as is illustrated in Table 6, which shows that the difference between simulated and observed ADPs for hydrogen atoms is more negative (less positive) than for carbon, nitrogen and oxygen atoms. For C, N and O atoms at temperatures of 200-300 K the quantum effects are expected to be small, so the differences between simulated and experimental ADPs should be regarded as a characteristic of the deficiencies of the force field used in the simulation. Such deficiencies can be expected due to the use of atomic point charges in the Coulomb potential and the approximate form of the 6-12 potential in describing van der Waals interactions.

To further illustrate the difference between quantum and classical behaviour and better understand the temperature dependence of the vibrational motion we ran two additional simulations of benzene at 100 K and 50 K. The experimental temperature dependence and those obtained from simulations using GAFF, CHARMM and OPLS are compared in Fig. 8. At the higher temperatures simulated ADPs are significantly larger than experimental ones, and at lower temperatures, as required by the classical model of motion, they are smaller than experimental by the zero point motion contribution. Apart from these differences simulated and observed ADPs both show positive anharmonicity, *i.e.* the mean square displacements increase faster than linear in temperature (eqn (10)). This is not too surprising as the intermolecular interactions derive from anharmonic electrostatic and van der



Fig. 8 Temperature dependence of mean square displacement for carbon and hydrogen atoms of benzene: experimental and simulated by GAFF, CHARMM and OPLS. Graphs from simulations are extrapolated to zero, as required by the classical model of motion (shown in dashed lines). Experimental data for benzene at temperatures other than 138 and 218 K have been obtained by the procedure outlined in ref. 43.

Waals potentials. At the intersection of the curves for simulated and experimental ADPs quantum and anharmonic effects cancel and the simulated and experimental ADPs are accidentally equal.

In order to simulate vibrational motion realistically, one could potentially scale the simulated ADPs or, preferably, run simulations at an appropriately scaled temperature, $T_{\rm eff}$. Such a temperature could be determined by comparing the temperature dependence of simulated mean square displacement with experimental ADP data measured at the temperature of interest. The difference from the appropriate simulation temperature may be quite large. Assuming for simplicity the classical harmonic model of thermal motion, where the mean square displacement of atoms changes linearly with temperature, a 20% difference in ADPs at a given temperature corresponds to a 20% difference in temperature for a given magnitude of an ADP. For example, ADPs simulated in this work for benzene at 218 K ($U_{eq} = 0.058 \text{ Å}^2$ average over three carbons for GAFF and OPLS, and 0.060 Å² for CHARMM) are close to ADPs observed experimentally at 270 K (0.0704 Å²).⁴⁴ For the case of diaminoheptane it is the other way around - thermal motion simulated for 213 K is very close to experimental values of ADPs for 130 K, a difference of more than 80 degrees.

In spite of the absolute values of the simulated ADPs being so far from experimental data, there is some meaning in their relative values. For most compounds simulated by GAFF, CHARMM and OPLS, the correlation between the magnitude of the mean square displacement and crystallographic atom type follows the experimental trend. Thus, the character of motion is reflected in the simulation. For example, mean square displacements of carbon atoms in decane are largest at the ends of the molecules and decrease slightly towards the centre, a behaviour that is reflected by all force fields. Another example is the inequivalence of the values of the ADPs of the chemically equivalent but crystallographically different atoms in benzene which is reproduced with the same trend in the simulations (Tables S1–S3 of ESI).

ADPs differ in different directions. Here we have considered only their isotropic averages. A more detailed analysis taking into account anisotropies of ADPs has the potential of relating differences of simulated and observed ADPs to the type of interactions dominating in the corresponding directions. Such an analysis might help to improve the parametrization of nonbonded intermolecular interactions.

Molecular reorientation, internal rotation

In the simulations we observed reorientation of benzene molecules around their 6-fold axis with all force fields at all three temperatures; rotation of hydrogens in the CH₃ groups of decane, isopropylcyclohexane, alanine at 295 K and AGA at 293 K (all force fields); rotation of the carboxylate group of AGA as simulated by GAFF and OPLS; inversion and rotation of the NH₂ groups of diaminoheptane around the C–N bond at 213 K (only with CHARMM) and rotation of the NH₃⁺ group of AGA (GAFF and OPLS). Some examples of this motion are pictured in Fig. 9.



Fig. 9 Some examples of molecular and group reorientations observed in the simulations: reorientation of a benzene molecule (a), reorientation of the hydrogen atoms in the methyl group of AGA(b), reorientation of hydrogen atoms in the ammonium group of AGA (c), reorientation of oxygen atoms in the carboxyl group of AGA (d). Shown in the pictures are surfaces of equal probability of finding the hydrogen highlighted in red or one of the oxygens in the carboxyl group.

Some of these types of motion have also been observed experimentally. The reorientation of the benzene molecule has been studied by solid-state NMR and is known to start at about 90 K at a frequency of the order of 10^4 s⁻¹, which increases to the order of 10^{11} s⁻¹ near the melting point at 278.7 K.⁴⁵ Reorientation of hydrogen atoms in methyl groups has been observed in many organic compounds.⁴⁶⁻⁴⁹ Of all methyl-containing compounds investigated in this work we only found experimental information on methyl reorientation in L-alanine.50 In the cases of the relatively loose packing of the hydrocarbon molecules decane and isopropylcyclohexane, methyl reorientation is reasonable to expect, but the rearrangement of the carboxylate and the ammonium groups in AGA is more surprising as they are involved in hydrogen bonding. There is experimental evidence of pyramidal inversion at nitrogen and rotation of amino groups around N-C bonds in gaseous and liquid aniline,^{51,52} solid local anaesthetics procaine and benzocaine.53

As mentioned earlier, molecules simulated with the MM3 force field exhibited considerable motion – methyl and amino groups in alanine rotate even at 60 K, imidazole and benzene molecules wobble significantly around their planes, benzene at 218 K loses even short-range order. Therefore, we were unable to analyse dynamics in some MM3 simulations. Given the poor performance of this force field in reproducing structural parameters of our model compounds it was excluded completely from our investigation of molecular motion.

We define reorientation as an event where a molecule or a molecular fragment rotates by a certain angle and spends no less than 1 ps (1000 timesteps of the simulation) near that new angle. The residence time of 1 ps corresponds to the experimental limit of determining correlation times from spin–lattice relaxation studies. An example of reorientations of a benzene molecule throughout the trajectory is given in Fig. 10. The rate constants of such reorientations were calculated as their frequency, which is the average number of reorientational events per unit of time per molecule or per group in a molecule:

$$k = \nu = \frac{N_{\text{reorient}}}{t \times n_{\text{mol}} \times n_{\text{equiv.}gr}},$$
(11)

where t is the length of the simulation in seconds, and $n_{\text{equiv},gr}$ is the number of groups exhibiting reorientation in one molecule, *e.g.* for reorientation of hydrogens in a methyl group $n_{\text{equiv},gr} = 1$ in alanine and $n_{\text{equiv},gr} = 2$ in n-decane and isopropylcyclohexane. To obtain reliable values of these frequencies longer simulations (6 to 60 ns) were conducted to obtain better statistics of such dynamic processes. Results are given in Table 7.

Examining the table we can see that the frequencies of reorientations are overestimated for deuterated benzene and benzene at 218 K, but slightly underestimated for benzene at 138 K and for alanine at 295 K. Simulated values for n-decane and isopropylcyclohexane are probably underestimated by an order of magnitude or so.

Arrhenius activation energies (ΔH^{\dagger}) for reorientation of benzene and rotation of the methyl group in alanine are also available from the literature. To obtain these values from our simulations, we plotted the logarithm of reorientation frequency *versus* inverse temperature for benzene (the Arrhenius plots), which showed straight lines (Fig. S1 of ESI). We used data from protonated and deuterated benzene in the same plot since at low temperature the difference in their frequencies of reorientation is not significant. Given the Arrhenius equation in the transition state theory form, one can obtain the enthalpy, entropy and free energy of activation ($\Delta H^{\ddagger}, \Delta S^{\ddagger}, \Delta G^{\ddagger}$) from such a plot (see section S.2 of ESI for details of calculations). The values obtained for benzene from this



Fig. 10 Angle of reorientation of a benzene molecule at 218 K as simulated by GAFF, CHARMM and OPLS. The angle in the plots varies only between 0 and 180° , not 360° , due to inversion symmetry of the benzene molecule in the crystal, so angles of reorientation of 60° and 300° are both recorded as 60° , and angles of 120° and 240° are both recorded as 120° .

Table 7	Frequencies of	different types of mol	ecular motion as s	simulated by GA	AFF, CHARMM	and OPLS. All values are in s ⁻¹
		<i></i>			· · · · · · · · · · · · · · · · · · ·	

Type of molecular motion	GAFF	CHARMM	OPLS	Experiment
6-fold reorientations of d ₆ -benzene at 123 K	2.3×10^{6}	10.4×10^{6}	9.2×10^{6}	$\sim 2.0 \times 10^{6}$ ref. 45
6-fold reorientations of benzene 138 K	2.6×10^{7}	5.5×10^{7}	6.4×10^{7}	7.2×10^7 ref. 45
6-fold reorientations of benzene 218 K	7.5×10^{9}	12.3×10^{9}	11.6×10^{9}	1.8×10^9 ref. 45
CH ₃ rotation in alanine 295 K	0.8×10^8	0.2×10^{8}	0.7×10^8	2.8×10^8 ref. 50
CH ₃ rotation in decane 150 K	2.6×10^{7}	11.3×10^{7}	10.1×10^{7}	$\sim 10^{9 a}$
CH ₃ rotation in isopropylcyclohexane 150 K	2.8×10^{6}	5.8×10^{6}	5.6×10^{6}	$\sim 10^{9 a}$
NH ₂ inversion and rotation in diaminoheptane 213 K	0	7.8×10^6	0	
^{<i>a</i>} Value is for heptacosane ($C_{27}H_{56}$) at 150 K. ⁵⁴				

analysis are shown in Table 8. In the absence of multitemperature simulations for methyl reorientation we cannot construct a similar Arrhenius plot for alanine. Instead we obtained a free energy profile using umbrella sampling (Fig. 11b) from which we can obtain a value of the free energy barrier for methyl rotation. As all angles are sampled well, we can construct a similar free energy plot for benzene directly from the probability distribution of angles of reorientation from equilibrium simulations (Fig. 11a).

Data obtained as described above from free energy profiles (Fig. 11) and from Arrhenius plots using simulation data (Fig. S1 of ESI) are compared to experimental data for benzene⁴⁵ and alanine⁵⁰ in Table 8. Although the free energy barrier for rotation for benzene determined from the simulation *via* the Arrhenius plot and from the free energy surface are not identical, they do have similar magnitude and reproduce the same ordering of the force fields. Thus in our simulations the barrier to reorientation is underestimated for benzene, which explains the overestimated frequency; and overestimated for alanine, which is consistent with underestimated frequency of reorientation.

Correlation between different types of results

Having simulated a large range of data on a range of structural and dynamic properties of organic molecular crystals it is worth checking that the various results we have found are consistent with each other. The different types of results do make sense when considered together. For example, molecular structure is consistent with crystal structure – a shorter diaminoheptane molecule is consistent with the shrinking of the diaminoheptane unit cell.

Fig. 12 has differences between simulated and experimental ADPs plotted against the differences between simulated and experimental cell constants. Fig. 12 shows a general trend for the difference between simulated and observed ADPs to be most negative if the difference between simulated and observed unit cell distances is very negative. This indicates that the minimum intermolecular interaction energy (composed of the corresponding intermolecular potentials) occurs at distances that are too short and the effective potential itself is too narrow. Benzene is the only compound for which ADPs are overestimated (at 138 and at 218 K) with all force fields.

Table 8 Thermodynamic data on benzene and alanine from simulations using GAFF, CHARMM, OPLS and experiment. Subscript "A.plot" refers to values obtained from Arrhenius plots (experimental and constructed using data from simulations), and subscript "free" refers to values obtained from free energy surfaces. All values except ΔS^{\ddagger} are in kJ mol⁻¹. ΔS^{\ddagger} is in J mol⁻¹ K⁻¹

GAFF	CHARMM	OPLS	Experiment
18.7	16.6	16.4	$15.5 \pm 0.8,^{45} 17.4^{55}$
33.1 11.5	10.7	25.9 10.8	8.2^{a} 13.7 ^a
10.4 ± 0.1	8.3 ± 0.1	8.8 ± 0.1	50 57
			$22.6, {}^{30}22\pm1{}^{36}$ 33.6^{b}
28.6 ± 0.5	32.2 ± 0.3	31.2 ± 0.3	12.7°
	$\begin{array}{c} & & & & \\ & & & & \\ & & & & \\ & & & & $	$\begin{array}{c cccc} & & & & & & & \\ \hline & & & & & & \\ & & & &$	OATE CHARMM OPLS 18.7 16.6 16.4 33.1 27.0 25.9 11.5 10.7 10.8 10.4 \pm 0.1 8.3 \pm 0.1 8.8 \pm 0.1 pt 28.6 \pm 0.5 32.2 \pm 0.3 31.2 \pm 0.3



Fig. 11 Free energy (in kJ mol⁻¹) of reorientation (angle in degrees) of benzene molecules at 218 K and hydrogens in the alanine methyl group at 295 K as simulated by GAFF, CHARMM and OPLS.

The effect is largest for the highest temperature and weakly correlates with an overestimation of cell constants. This indicates that the minimum of the corresponding effective intermolecular potentials is too long and the potential itself to shallow, in qualitative agreement with the trend noted above for negative differences.

A similar but less pronounced correlation is found for reorientation frequencies. On average simulated frequencies tend to be higher than experimental ones for benzene, but lower for alanine and the saturated hydrocarbon molecules. This behaviour is consistent with the corresponding over- and underestimated ADPs and unit cell dimensions.

Conclusions

In this study we have taken a structural chemist's look at classical MD simulations of organic molecular crystals. Most simulation results agree qualitatively with experiment, but not quantitatively. In some situations simple adjusting of key parameters, as done in ref. 13, or finding a way to scale the results might give the necessary more quantitative level of agreement with experiment for a given physical property, however, this may not work for multiple properties simultaneously.

The force fields tested are not generally applicable, each of them having its domain among organic molecules where it performs best. Not surprisingly, the GAFF, CHARMM and OPLS force fields, which share the functional form and some of the parameters, perform very similarly, although CHARMM seems to do best in the case of polar molecules. The MM3 force field is significantly different in its functional form and was designed for gas phase calculations. A conclusion of this study is that it should not be used in condensed phase simulations.



Fig. 12 Correlation of differences between simulated and experimental ADPs and differences between simulated and experimental cell constants. Data points for GAFF are coloured red, for CHARMM in green and for OPLS in blue.

The magnitude of thermal motion produced in a classical MD simulation depends on the size of the simulation box, but is close to the converged value for a system of the size allowing NAMD's standard cutoff of 12 Å. It is generally distinctly underestimated by the GAFF, CHARMM22 and OPLS force fields. To some extent this deficiency of the simulations may be attributed to the classical treatment of atomic motion in conventional MD, *i.e.* a lack of quantum effects, which has the biggest influence at low temperature and on light atoms. At higher temperatures the deficiencies of the simulations in the force fields.

Depending on the aim of the study there are ways to overcome the intrinsic limitations of classical molecular dynamics as described in this paper. These may include using a polarizable force field, like AMOEBA,⁵⁷ instead of a force field based on point charges; employing *ab initio* calculations to study crystal properties with codes like CRYSTAL06⁵⁸; introducing quantum effects into MD simulations with Car-Parinello MD⁵⁹ combined with Feynman path integral calculations.⁶⁰

In view of the many possible applications of MD simulations to organic molecular crystals listed in the introduction, we recommend to compare as many experimental quantities as available to the simulation results and to determine simulation conditions (force field, box size, temperature, *etc.*) that produce results as close as possible to experiment before venturing into the study of crystal properties not amenable to experiment.

Further work exploring aspects of dynamics and disorder in organic molecular crystals using molecular dynamics is underway in our laboratory.

Acknowledgements

The authors thank Associate Professor Dylan Jayatilaka for fruitful discussions. HBB thanks the Australian Research Council for a Linkage International Fellowship. AN acknowledges the Australian Government for the International Postgraduate Research Scholarship and the University of Western Australia for the University Postgraduate Award.

Computer time was allocated under the merit allocation scheme of the National Computational Infrastructure (Australia).

References

- 1 A. T. Brunger, J. Kuriyan and M. Karplus, *Science*, 1987, 235, 458–460.
- 2 A. T. Brunger, Annu. Rev. Phys. Chem., 1991, 42, 197-223.
- 3 A. T. Brunger and P. D. Adams, Acc. Chem. Res., 2002, 35, 404-412.
- 4 J. Kua, R. C. Daly, K. M. Tomlin, A. C. T. van Duin, T. B. Brill, R. W. Beal and A. L. Rheingold, *J. Phys. Chem. A*, 2009, **113**, 11443–11453.
- 5 S. Piana and J. D. Gale, J. Am. Chem. Soc., 2005, 127, 1975-1982.
- 6 A. Gavezzotti, G. Filippini, J. Kroon, B. P. van Eijck and P. Klewinghaus, *Chem.-Eur. J.*, 1997, **3**, 893–899.
- 7 A. Gavezzotti, Molecular aggregation, structure analysis and molecular simulation of crystals and liquids, Oxford University Press Inc., New York, 2007, p. 273.
- 8 A. Gavezzotti, J. Am. Chem. Soc., 2000, 122, 10724-10725.

- 9 J. van Eerden, S. Harkema and D. Feil, Acta Crystallogr., Sect. B: Struct. Sci., 1990, 46, 222–229.
- 10 N. M. Glykos, Acta Crystallographica Section D, 2007, 63, 705–713.
- 11 A. M. Reilly, D. A. Wann, C. A. Morrison and D. W. H. Rankin, *Chem. Phys. Lett.*, 2007, 448, 61–64.
- 12 M. Rueda, C. Ferrer-Costa, T. Meyer, A. Perez, J. Camps, A. Hospital, J. L. Gelpi and M. Orozco, *Proc. Natl. Acad. Sci.* U. S. A., 2007, **104**, 796–801.
- 13 D. C. Chatfield and S. E. Wong, J. Phys. Chem. B, 2000, 104, 11342-11348.
- 14 D. C. Chatfield, A. Augsten, C. D'Cunha and S. E. Wong, J. Comput. Chem., 2003, 24, 1052–1058.
- 15 B. Dittrich, J. E. Warren, F. P. A. Fabbiani, W. Morgenroth and B. Corry, *Phys. Chem. Chem. Phys.*, 2009, 11, 2601–2609.
- 16 G. A. Jeffrey, J. R. Ruble, R. K. McMullan and J. A. Pople, Proc. R. Soc. London, Ser. A, 1987, 414, 47–57.
- 17 G. E. Bacon, N. A. Curry and S. A. Wilson, Proc. R. Soc. London, Ser. A, 1964, 279, 98–110.
- 18 D. R. Allan, P. A. McGregor, L. J. McLachlan, S. Parsons and C. R. Pulham, *Acta Crystallogr.*, Sect. E: Struct. Rep. Online, 2002, 58, 0652–0655.
- 19 A. D. Bond and J. E. Davies, Acta Crystallogr., Sect. E: Struct. Rep. Online, 2002, 58, 0196–0197.
- 20 R. K. McMullan, J. Epstein, J. R. Ruble and B. M. Craven, Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem., 1979, 35, 688–691.
- 21 J. Epstein, J. R. Ruble and B. M. Craven, Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem., 1982, 38, 140–149.
- 22 V. R. Thalladi, R. Boese and H. C. Weiss, Angew. Chem., Int. Ed., 2000, 39, 918–922.
- 23 R. Gotthardt, J. H. Fuhrhop, J. Buschmann and P. Luger, Acta Crystallogr., Sect. C: Cryst. Struct. Commun., 1997, 53, 1715–1717.
- 24 C. C. Wilson, D. Myles, M. Ghosh, L. N. Johnson and W. G. Wang, New J. Chem., 2005, 29, 1318–1322.
- 25 G. S. Padiyar and T. P. Seshadri, Acta Crystallogr., Sect. C: Cryst. Struct. Commun., 1996, 52, 1693–1695.
- 26 J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, J. Comput. Chem., 2004, 25, 1157–1174.
- 27 A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus, J. Phys. Chem. B, 1998, **102**, 3586–3616.
- 28 W. L. Jorgensen and J. Tirado-Rives, J. Am. Chem. Soc., 1988, 110, 1657–1666.
- 29 N. L. Allinger, Y. H. Yuh and J. H. Lii, J. Am. Chem. Soc., 1989, 111, 8551–8566.
- 30 J. H. Lii and N. L. Allinger, J. Am. Chem. Soc., 1989, 111, 8566–8575.
- 31 J. H. Lii and N. L. Allinger, J. Am. Chem. Soc., 1989, 111, 8576–8582.
- 32 X. Gu, L. Zhang, X. Gong, W. M. Lau and Z. F. Liu, J. Phys. Chem. B, 2008, 112, 14851–14856.
- 33 L. Kale, R. Skeel, M. Bhandarkar, R. Brunner, A. Gursoy, N. Krawetz, J. Phillips, A. Shinozaki, K. Varadarajan and K. Schulten, J. Comput. Phys., 1999, 151, 283–312.
- 34 J. Ponder, Tinker-Software Tools for Molecular Design; http:// dasher.wustl.edu/tinker/.
- 35 W. Humphrey, A. Dalke and K. Schulten, J. Mol. Graphics, 1996, 14, 33–38.
- 36 M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia and M. Parrinello, *Comput. Phys. Commun.*, 2009, **180**, 1961–1972.
- 37 S. Kumar, J. Rosenberg, D. Bouzida, R. Swendsen and P. Kollman, J. Comput. Chem., 1995, 16, 1339–1350.
- 38 B. Roux, Comput. Phys. Commun., 1995, 91, 275-282.
- 39 A. Grossfield, An implementation of WHAM: the Weighted Histogram Analysis Method http://membrane.urmc.rochester.edu/ Software/WHAM/WHAM.html.
- 40 S. L. Price, B. Patel, P. Prithanani-Jethani and A. Torrisi, ACA Transactions, 2004, 39, 2–13.

- 41 R. A. Bonham, L. S. Bartell and D. A. Kohl, J. Am. Chem. Soc., 1959, 81, 4765–4769.
- 42 H. B. Burgi and S. C. Capelli, Helv. Chim. Acta, 2003, 86, 1625–1640.
- 43 H. B. Burgi, S. C. Capelli, A. E. Goeta, J. A. K. Howard, M. A. Spackman and D. S. Yufit, *Chem.-Eur. J.*, 2002, 8, 3512–3521.
- 44 E. G. Cox, D. W. J. Cruickshank and J. A. S. Smith, Proc. R. Soc. London, Ser. A, 1958, 247, 1–21.
- 45 E.R. Andrew and R. G. Eades, *Proc. R. Soc. London, Ser. A*, 1953, **218**, 537–552.
- 46 P. A. Beckmann, L. Happersett, A. V. Herzog and W. M. Tong, J. Chem. Phys., 1991, 95, 828–835.
- 47 A. M. I. Ahmed and R. G. Eades, J. Chem. Soc., Faraday Trans. 2, 1972, 68, 1337.
- 48 L. Latanowicz, Concepts Magn. Reson., 2005, 27a, 38-53.
- 49 P. Schiebel, R. J. Papoular, W. Paulus, H. Zimmermann, A. Detken, U. Haeberlen and W. Prandl, *Phys. Rev. Lett.*, 1999, 83, 975–978.
- 50 L. S. Batchelder, C. H. Niu and D. A. Torchia, J. Am. Chem. Soc., 1983, 105, 2228–2231.
- 51 N. W. Larsen, E. L. Hansen and F. M. Nicolaisen, Chem. Phys. Lett., 1976, 43, 584–586.

- 52 G. Fronza, R. Mondelli, F. Lelj, E. W. Randall and C. A. Vveracini, *Journal of Magnetic Resonance*, 1980, **37**, 275–284.
- 53 M. A. Palafox, Spectrosc. Lett., 1997, 30, 1495–1514.
- 54 R. Kitamaru, F. Horii, M. Nakagawa, K. Takamizawa, Y. Urabe and Y. Ogawa, J. Mol. Struct., 1995, 355, 95–104.
- 55 W. E. Sanford and R. K. Boyd, Can. J. Chem., 1976, 54, 2773–2782.
- 56 M. Keniry, A. Kintanar, R. Smith, H. Gutowsky and E. Oldfield, *Biochemistry*, 1984, 23, 288–298.
- 57 J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, Jr., M. Head-Gordon, G. N. I. Clark, M. E. Johnson and T. Head-Gordon, J. Phys. Chem. B, 2010, 114, 2549–2564.
- 58 R. Dovesi, V. Saunders, R. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N. M. Harrison, I. J. Bush, D. P. and M. Llunell, *CRYSTAL06*, CRYSTAL06 User's Manual.
- 59 R. Car and M. Parrinello, Phys. Rev. Lett., 1985, 55, 2471-2474.
- 60 R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York, 1965.