NJC

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: A. Gavezzotti and L. Lo Presti, *New J. Chem.*, 2018, DOI: 10.1039/C8NJ05825C.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/njc

View Article Online
DOI: 10.1039/C8NJ05825C

Dynamic simulation of liquid molecular nanoclusters. Structure, stability and quantification of internal (pseudo)symmetries

Angelo Gavezzotti*[,a] and Leonardo Lo Presti[a,b,c]

Received (in XXX, XXX) XthXXXXXXXXX 200X, Accepted Xth XXXXXXXXX 200X
First published on the web Xth XXXXXXXXX 200X
DOI:

The atom–atom intermolecular force field AA–CLP is applied to the Molecular Dynamics simulation of liquid nanoclusters of benzene, chloroform, methanol and pyridine. Bulk liquids are also simulated for validation and comparison with experimental data. The applied software has been produced *de novo* to deal with the unusual analytical form of the intermolecular potential, and includes some novel features for control of net rotational momenta in isolated systems. The nanoclusters have been studied as a function of size (150–1000 molecules) in what concerns cohesion energies, rotational correlation, self–diffusion coefficients, and evaporation rates. Internal structure has been studied with traditional radial distribution functions, plus diagrams of the distribution of intermolecular vectors for flat compounds. In addition, a new algorithm for the quantification of pseudo– or near–symmetries between molecules in aggregates of any structure has been developed and tested, with reference to inversion, mirror plane and rotation axes symmetries proper of organic crystals, with possible import on the investigation of crystallization processes. The results confirm the reliability of the AA–CLP formulation for Molecular Dynamics simulation and throw some light on relationships between cluster and bulk properties. The computer codes are available for open–source download to ensure full reproducibility of all results.

Introduction

In parallel with an ever growing availability of microscopic analysis techniques, there is presently an ever growing interest in chemical objects of nanometer dimensions, both on the theoretical and applicative side. These objects have been around in the literature for a long time, since pioneering studies[1]that used supersonic flow experiments to generate liquid– and crystal–like molecular nanoparticles, characterizing them by diffraction techniques and molecular dynamics simulation. A recent overview[2] proves that the subject is still hot today, and although most studies refer to monoatomic gases or, typically, to water,[3] the experimental and theoretical apparatus is now up to the task of investigating the nucleation of larger molecules.

Nanocluster science is an essential part of the development of functional nanoparticles, mostly made of salts or metals, but also organicmolecular aggregates is important in many sectors of applied chemistry and physics as, to quote just an example, atmospheric pollution.[4]

Static optimization of the structure of molecular clusters has been pursued by quantum chemical methods.[5] Such studies involve the determination of temperature–less, static energy minima that may have little to do with the actual behaviour of these aggregates in real thermal conditions. Molecular dynamics (MD) and Monte Carlo (MC) are by now mature computational techniques of wide applicability. In the cluster arena most frequent applications are to model systems, likethe Lennard–Jones sphere,[6] which allows large particle numbers but lacks chemical significance, or, again, to the vapor–liquid interface of water droplets,[7]

Rather than in model systems, we are interested in the aggregation properties of polyatomic organic molecules. Currently available MC–MD software include an early but seminal item, GROMOS,[8] recently evolved into GROMACS.[9] However, these packages are mainly oriented to the simulation of biological systems and their setup is seldom adaptable to dealing with crystal symmetry or isolated clusters. Besides, they can only be used with the analytical form of the embedded force field. One aim of the present contribution is a validation in MD simulation of the AA–CLP force field,[10,11] an atom–atom recipe mainly developed for molecular crystals and liquids, up to now only applied in

* To whom correspondence should be addressed: angelo.gavezzotti@unimi.it
[a] *Department of Chemistry, University of Milano, viaGolgi 19, 20133 Milano, Italy*
[b] *Istituto di Scienze e Tecnologie Molecolari, Italian CNR, Via Golgi 19 I-20133 Milano (Italy)*
[c] *Centre for Materials Crystallography, Århus University, Langelandsgade 140, DK-8000 Århus C. (Denmark)*

† Electronic Supplementary Information (ESI) available: Force field parameters. Simulation outcomes: supplementary Figures. See DOI:

static or Monte Carlo simulation.

On these premises we present here a suite of Molecular Dynamics programs developed *ex novo* to treat molecular clusters of up to 1000 molecules of up to 20 atoms, living in a common PC environment, where the code can be easily modified to deal with new situations emerging on the fly, and significant results can be obtained in reasonable solar times without the use of supercomputing resources. Being cheap in the above sense allows a flexible investigation of many different compounds, temperatures and cluster sizes in unforeseen conditions. The formulation uses standard MD algebra but includes some new features specific to isolated clusters, like algorithms for disposal of net overall translational and rotational momentum, or for dealing with evaporation by detachment of boundary molecules from the cluster surface. Molecular clusters are interesting also as precursors of crystal nucleation and growth mechanisms. Therefore, a new algorithm for the quantification of molecular symmetries in systems of any structure between fully liquid and fully crystalline has been developed, with specific reference to symmetries of organic crystals, namely inversion, twofold rotation or screw axes and mirror or glide planes, plus pure translation.

Chemically very different test-case liquids, spanning apolar aromatics (benzene), strongly protic (methanol) and clorurated (chloroform) hydrocarbons, and aromatic bases (pyridine), have been considered. The results prove the reliability of the AA–CLP intermolecular potential scheme in molecular dynamic applications, yielding on the way a number of valuable insights into the structure and evolution of small liquid clusters as a function of size, chemical composition and temperature, also in comparison with bulk liquid properties.

Computational methods

Intermolecular force field: the AA–CLP scheme

The AA–CLP (Atom–Atom Coulomb–London–Pauli) intermolecular potential[10,11] is an atom–centered empirical scheme that consists of a Coulombic term over atomic point–charge parameters, plus terms in inverse powers of interatomic distance R to represent polarization, dispersion and repulsion energies. For the interaction between an i–j atom pair the potential reads:[10]

$$E(i,j) = \frac{q_i \cdot q_j}{R} - \frac{A}{R^4} - \frac{B}{R^6} + \frac{C}{R^{12}} \quad (1)$$
$$= E(\text{Coul}) + E(\text{Pol}) + E(\text{Dis}) + E(\text{Rep})$$

Charge parameters q come from Mulliken population analysis on a molecular orbital scheme. A, B and C factors are calculated from basic properties like local charge, ionization potential and polarizability for each atom pair in each molecule. More detail can be found in the Literature[12–16] and in the documentation available on line.[17] This scheme has been optimized to reproduce the properties of a hundreds of organic compounds in the liquid or crystalline state and owes its success to exhaustive parameter fitting. Rather than being set for pairs of atom types, as in most potential schemes, the interaction coefficients are molecule–dependent, which means that the potential cannot be adapted to other available MD packages. For this reason, capital among others, a new molecular dynamics code had to be developed, as described below.

Monte Carlo (MC)

The MC procedure, described[11] and tested[11,18] previously,considers any given molecule as either fully rigid (*e.g.* naphthalene) or partitioned into rigid fragments (like the two rings in biphenyl or the ring and the OH group in phenol) connected by rotatable but un–stretchable bonds. The intramolecular force field includes only bond bending for C–O–H fragments and torsional potentials over rotatable bonds. NPT–MC (constant number of particles, pressure and temperature) with periodic boundary conditions is used for liquids and crystalline states.

Molecular Dynamics (MD)

A description of the essential features of the newly developed MD procedure for clusters follows. More detail can be found in the deposited on line material.[19]

The only structural input to the MD modules is a set of Cartesian atomic position coordinates for the computational box. NPT–MD with periodic boundary conditions of 26 replicas of the cubic reference box, with virial pressure control over center–of–mass degrees of freedom,[20] is applied to simulate bulk liquids. Cluster simulations are by definition carried out *in vacuo*. The total intermolecular configurational energy is calculated over all atom–atom pairs in all molecule–molecule pairs below a cutoff distance between centers of mass, usually 18–20 Å. In this way the summations run always between neutral charge groups and there is no need of range control in Coulomb terms. Zero–step atomic velocities are assigned by an approximate Maxwellian distribution according to[9]

$$V = \sqrt{\frac{k_B T}{M}}\left[\sum_{i=1}^{12} R_{ni} - 6\right] \quad (2)$$

where k_B is the Boltzmann constant, T is the set temperature, M is the atomic mass, and R_{ni} is a random number between 0 and 1. To reduce translational or rotational biases in small isolated systems the components of any velocity vector V are further randomized as follows:

$$V = aV_x + bV_y + cV_z$$
$$b^2 + c^2 = 1 - a^2 \quad (3)$$

$$b^2 = a' \cdot (1 - a^2)$$
$$c^2 = (1 - a') \cdot (1 - a^2)$$

with a and a' random numbers between 0 and 1. Each component is assigned a plus or minus sign according to a random number being grater or smaller than 0.5.

Temperature has the standard equipartition expression and is constrained by rescaling all atomic velocities by the ratio between T(set) and T(current) usually every 500–1000 MD steps. In isolated clusters "temperature" is better regarded as

View Article Online
DOI: 10.1039/C8NJ05825C

just a measure of dynamic freedom with only limited reference to the corresponding thermodynamic quantity, and no pressure control applies.

The trajectory is integrated by a standard leap–frog algorithm with time (t), velocity (V), position (r), mass (M) and force (F):[9]

$$V(t + 1/2\Delta t) = V(t - 1/2\Delta t) + \frac{\Delta t}{M}F(t) \quad (4a)$$

$$r(t + \Delta t) = r(t) + \Delta t \cdot V(t + 1/2\Delta t) \quad (4b)$$

The intramolecular potentials and forces are taken as such from the GROMOS96[20] formulation:

$$E(\text{str}) = 1/2 \cdot k_{str}(R - R^0)^2 \quad (5a)$$

$$E(\text{bend}) = 1/2 \cdot k_{bend}(\cos\vartheta - \cos\vartheta^0)^2 \quad (5b)$$

$$E(\text{tors}) = k_{tors}[1 + f \cdot \cos(m\varphi)] \quad (5c)$$

where R is a bond distance, ϑ is a bond angle, φ is a (proper or improper) torsion angle and the k's are parametric force constants. f is a phase factor equal to ±1, and m is an integer equal to 1, 2 or 3. The applied force field includes one stretching potential for each bond, one bending potential for each bond triad, and one torsion potential along each chemical bond. Improper dihedrals help preserving planar centers (e.g. sp^2 C–atoms in aromatic rings). The force constants are assigned values from the literature[21]or (especially for torsions) from ab initio calibration; for nearly rigid systems, the choice of force constants is only moderately influential as high–frequency internal vibrations are not relevant to the problem at hand, namely intermolecular structure and phase change. Intramolecular potentials serve only as restraints to avoid undue molecular bond and angle distortions. At least for molecules without extensive internal flexibility degrees of freedom, the simple setup performs well with considerable time savings over more complex analytical constraint procedures. All numerical data are collected in ESI–Force field (Section S1 ESI).

Unwanted net momenta. In MD simulations of isolated clusters net translational and rotational momenta may develop. Empirical procedures have been devised to deal with of these adverse factors. The array of centers of mass at step zero is taken as a reference configuration. To dispose of Tr–drift, all coordinates are reset to the origin of the current center of mass (c.o.m) every 20–50 MD steps. To suppress Rot–momentum a rotation matrix is prepared as the product of three rotation matrices, one around each inertial axis, and an array of possible back–rotations of the whole cluster is explored by varying the three rotation angles from –15 to +15° in steps of 3°. The back–rotation actually performed is the one for which the sum of the distances between current centers of mass and the reference configuration at zero simulation time is minimum. These corrections avoid complex algebra and affect the whole set of atomic coordinates in a rigid manner without discontinuities in the trajectories. Their efficiency has been checked by tests on correlation functions and distributions of centers of mass with and without the corrections.

Evaporation. In MD runs on isolated clusters some surface molecules may drift away in the simulation analog of evaporation. Two procedures are introduced to deal with this problem. In the first, if the c.o.m. of a molecule is farther from the overall c.o.m. of the cluster by more than an assigned threshold, R_{ev}, the c.o.m. vector is reduced by a factor $0 < F_{ev} < 1$ so that the whole molecule is "tethered" back by a vector $1–F_{ev}$. Using $F_{ev} \approx 0.9$ and $R_{ev} \approx 1.5$ times the cluster radius a few molecules are thus kept in orbit, being then excluded from the estimation of rms displacements. In the second procedure when a c.o.m vector exceeds the diameter of the cluster the molecule is simply deleted from the simulation. This applies for larger clusters where deletion of a few molecules is not influential, or when evaporation rates are estimated.

Trajectory analysis. Periodic checks are carried out to make sure that molecular distortions do not exceed a limit of a few RT units. The intermolecular structure of the sample is described by radial distribution functions (RDF):

$$g(R_k) = \frac{N(R_k)}{4\pi R_k^2 \mathrm{d}R} \cdot \left(\frac{N}{V}\right)^{-1} \quad (6)$$

where $N(R_k)$ is the number of distances in the k–th distance bin and is divided by the volume of the bin of thickness dR (usually 0.1–0.2 Å). The term $(N/V)^{-1}$ normalizes the RDF to the number density. The volume of an isolated cluster is approximated by the sum of molecular volumes divided by a packing coefficient of 0.6.

Rotational correlation is estimated by the following expression:

$$C(\mathbf{u}) = 100 \cdot \langle \mathbf{u}(t) \cdot \mathbf{u}(0) \rangle \quad (7)$$

where $\mathbf{u}$ is a unit vector along any chosen molecular axis at time t with respect to a chosen reference. The brackets denote the average of dot products over the whole simulation sample. The rotational correlation time is estimated as the time for $C(\mathbf{u})$ to decay from 100 to about 20, and should be of the order of 5–20 ps for organic liquids. The translational diffusion coefficient D is given by the Einstein relation:

$$D = \frac{1}{6\Delta t} \cdot \langle \mathbf{d}(t) - \mathbf{d}(0) \rangle \quad (8)$$

where $\mathbf{d}(t)$ is the placement vector of the center of mass at time t and $\mathbf{d}(0)$ is the chosen reference. D is calculated as 1/6 of the slope of the quadratic displacement versus time and should be of the order of 10^{-8} m^2 s^{-1} (or Å^2 ps^{-1}).

Reproducibility. Fortran source modules, sample input and output data with documentation of the computational scheme are available for download from www.angelogavezzotti.it, CLP–dyn link. Together with deposited force field data (Section S1 ESI) this ensures complete reproducibility of all calculations.

Generation of starting aggregates

Molecular models of standard geometry (see ESI,Section S1) are used to prepare approximate starting computational boxes with molecules in random orientation. A sequence of energy

View Article Online
DOI: 10.1039/C8NJ05825C

minimization and NPT–MC runs produces equilibrated models of the liquids. The resulting final computational boxes are input as such to the NPT–MD simulations to check the correct reproduction of bulk liquid properties. Then, roughly globular, isolated nanoclusters (droplets) of variable size are prepared by deleting outer molecules. The clusters are input to starting MD runs of 60–100 ps followed by production runs of 100 ps, with a timestep of 2 fs. Such lengths are enough to characterize the main dynamic features, rotational correlation and self–diffusion. Longer runs are used to explore evaporation. C.o.m. and atom–atom radial distribution functions, eq. (6), are calculated on the final MD frames. C.o.m. diffusion coefficients, eq. (8), and rotational correlation functions, eq. (7), are calculated from the resulting MD trajectories.

Quantification of symmetries in a molecular array

Molecular clusters are interesting in themselves, but also in the perspective of a possible transition to crystalline structure along the nucleation and growth processes. Through phase changes, a crucial parameter is therefore the degree of crystallinity of the system. A simple but efficient way to obtain a quantitative measure of this parameter is described in the following.

In any pair of mlecules of a static computer model of an ordered molecular crystal or of a molecular liquid composed of rigid molecules, the pair members are related by an isometry, *i.e.* a distance–preserving transformation. In a perfect crystal the molecules are also related by perfect symmetry, for organic crystals by translation *T*, inversion *I*, twofold screw axis *S* and glide mirror plane *G*; less common are twofold axis *A* and mirror plane *M*. In what follows the translations associated with G and S operators can be neglected and their symmetry reduces to M or A respectively, making the analysis much easier.

Consider an ensemble of N_{mol} molecules each of N_{at} atoms, in which molecule *i* is described by the set of atomic coordinates $w_i°$ [x,y,z] with center of mass $w_{im}°$, in any external reference system. Firstly, all molecules are set in the reference system of their center of mass, $w'_i = w_i° - w_{im}°$. If the *i–j* pair is related by a perfect symmetry, the two sets of coordinates w_i' and w_j' are identical, except for some combination of sign inversions (first column of Table 1). In general for rigid or flexible molecules in any environment, a numerical index of deviation from symmetry can be constructed as follows. The overlap between the two sets of coordinates is maximized by transformation to the common inertial reference system. The six possible sums over the resulting final coordinates of the two partners, from $\Sigma_{(1)}$ to $\Sigma_{(6)}$, are then computed:

$$\Sigma_{(1)} = \frac{1}{N_{at}} \cdot \sum_{ij} |x_i + x_j|$$
$$\Sigma_{(2)} = \frac{1}{N_{at}} \cdot \sum_{ij} |x_i - x_j|$$
$$\Sigma_{(3)} = \frac{1}{N_{at}} \cdot \sum_{ij} |y_i + y_j| \quad (9)$$
$$\Sigma_{(4)} = \frac{1}{N_{at}} \cdot \sum_{ij} |y_i - y_j|$$
$$\Sigma_{(5)} = \frac{1}{N_{at}} \cdot \sum_{ij} |z_i + z_j|$$
$$\Sigma_{(6)} = \frac{1}{N_{at}} \cdot \sum_{ij} |z_i - z_j|$$

where each sum runs over all N_{at} pairs of corresponding atoms in the two molecules. If the molecule has internal point–group symmetry, the six sums are evaluated for each equivalent permutation of atom numberings generated by the point–group symmetry, and the lowest values are taken.

If the molecular pair is exactly symmetric three of these sums are exactly zero (Table 1). For any molecular pair not necessarily endowed with symmetry, the six sums are ranked in ascending order and it is assumed that the three lowest ones are the best approach to a possible symmetry relationship; their sum can be called an index of symmetry breaking, g_{ij}, in units of Å per atom (eq. 10). The average of g_{ij}'s over all pairs in the entire molecular array is the global asymmetry index of the array, *G* (eq. 11), representing an inverse order parameter:

$$g_{ij} = \sum_{i=1}^{3} \min(\Sigma_i) \quad (10)$$
$$G = \frac{1}{N_{pairs}} \cdot \sum_{i<j} g_{ij} \quad (11)$$

where $N_{pairs} = N_{mol}(N_{mol}-1)/2$. More revealing than just the global index G can be the distribution of g_{ij}.'s over the entire sample: a peak near $N(g_{ij}) = 0$ indicates approximate crystalline symmetry while a flat profile denotes an aggregate close to the liquid state.

Table 1. Values of the six sums for perfect symmetries. *T* = translation, *G,M* = glide or mirror plane, *S,A* = twofold screw or axis, *I* = inversion.

symmetry relationship		$\Sigma_{(i)}$					
	(*i*)	(1)	(2)	(3)	(4)	(5)	(6)
x, y, z	*T*	any	0	any	0	any	0
–x, y, z	*G,M*	0	any	any	0	any	0
x ,–y ,z	*G,M*	any	0	0	any	any	0
x, y, –z	*G,M*	any	0	any	0	0	any
–x ,–y ,z	*S,A*	0	any	0	any	any	0
–x , y ,–z	*S,A*	0	any	any	0	0	any
x, –y, –z	*S,A*	any	0	0	any	0	any
–x , –y, –z	*I*	0	any	0	any	0	any

Results

Bulk liquids *vs.* clusters

Cohesive energies and densities. For bulk liquids the results of the NPT–MC treatment (Table 2) reproduce the experimental conditions, as expected from previous calibration.[10,11] The NPT–MD results are also in good correspondence with experiment, in a preliminary validation of the whole setup and of the potential energy schemes for molecular dynamics application The comparison between cohesive energies and experimental vaporization enthalpies is only qualitative, since no account is taken of possible changes in intramolecular energies between liquid and vapor. The agreement between calculated and observed liquid densities is satisfactory, confirming that the potentials place their minima at the correct position, and parallels the acceptable reproduction of intermolecular energetics. Differences

View Article Online
DOI: 10.1039/C8NJ05825C

between MC and MD results may come from differences in box dimensions and treatment of intermolecular flexibility. MD performs even better than MC although smaller sample boxes are used.

Intermolecular orientation angles. For flat molecules like benzene or pyridine, one would think intuitively that ring stacking might be favored in condensed aggregates. This is not the case in the crystal structures of both compounds, where T–shaped arrangement predominates. This feature can be probed by calculating the distribution of angles between the normal to the ring in molecular pairs, δ. Results for the angles between ring normals in benzene are shown in Figure 1: the distribution shows a decrease in frequency of higher angles at short distances, as expected. By contrast, the distribution in a crystalline aggregate reveals sharp peaks for the symmetry–related molecular orientations (Figure S1 ESI). The distribution of angles between molecular C–H vectors in chloroformclusters is on the contrary perfectly uniform in distance (ESI, Figure S2).

Quantitative treatment of orientation probability densities in bulk systems[22–24] could be given for a more detailed discussion of the matter; this would be however outside the scope of the present work. We merely point out that the distribution of inter ring angles in liquid benzene clusters is identical to that of the bulk liquid (ESI, Figure S3).

Table 2. Results of NPT–periodic boundary MC and MD simulation for bulk liquids.

Substance	Method	Energies / kJ·mol^{-1}		Density / g·cm^{-3}		n molecules [c]
		CE[a]	exptl $\Delta_{vap}H$[b]	Calc	Exptl	
benzene	MC	31.6(2)	34	0.897(2)	0.87	686
	MD	32.5(5)		0.917(9)		250
chloroform	MC	26.8(1)	31	1.410(2)	1.47	1458
	MD	26.4(3)		1.451(10)		432
methanol	MC	37.8(1)	37	0.770(1)	0.79	1458
	MD	42.6(5)		0.799(10)		432
pyridine	MC	39.2(2)	40	1.047(5)	0.982	432
	MD	38.9(7)		1.040(13)		250

[a]Cohesive intermolecular energy. Rmsd in parentheses from the average in production runs.

[b]From https://webbook.nist.gov/chemistry/.

[c]Number of molecules in central box.

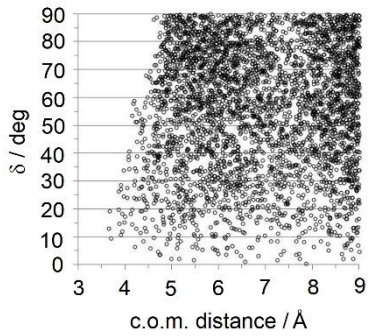


Figure 1. The distribution of inter–ring angles, δ, computed as the angle between ring normals, in the final frame of the MD simulation of a 475–molecule benzene cluster. At short c.o.m. distances small angles predominate, corresponding to approximate stacking. The distribution in the final frame of the bulk liquid simulation is nearly identical.

Nanocluster properties. In the MD runs for clusters the intramolecular force fields keep the molecules undistorted, with torsion angles oscillations < 10°, rms bond stretchings of 0.03 Å and rms bond angle bendings of 3.5°, and corresponding strain energies of 0.6 to 1.5 kJ·mol^{-1} per intramolecular degree of freedom. The few evaporating surface molecules are taken care of by the tethering or deletion strategies. All considered, MD production runs for nanoclusters are stable and robust.

Graphs for the correlation functions are given in Figure 2. Calculated dynamic properties in Table 3 generally show a reasonable agreement with experimental data and other simulation results[25–28] (which are also not free from oscillations, see footnotes to Table 3) and nicely reflect the different strength of intermolecular interactions, from weakly bound benzene or chloroform to tighter binding in pyridine and even more in the hydrogen–bonded alcohol. Self–diffusion coefficients and rotational correlation times in clusters are both overestimated with respect to bulk liquid, presumably due to an imbalance in boundless systems where diffusion (expansion) takes the best part of kinetic energy. As expected, the benzene rotational correlation time is higher for the rotation axis with the higher moment of inertia (perpendicular to the ring).

The results in Table 3 further confirm the reliability of the computational scheme. A somewhat unexpected, new result is that rotational correlation times are identical in the bulk liquid and in nanoclusters made of 400–1000 molecules (see also Figure S4 ESI).

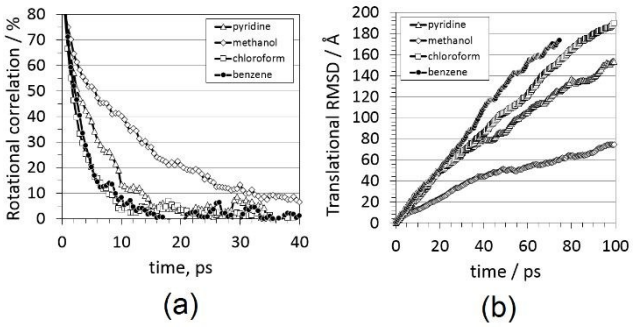


Figure 2. Time evolution of (a) rotational correlation function and (b) translational root mean squared c.o.m. displacement, as obtained from the 100 ps MD production runs for liquid clusters (see Table 3).

Table 3. Correlation results from MD. Self–diffusion coefficients from the slope of the rmsd *vs*. time graphs in Figure 2 from cluster calculations.

	Rotational correlation, ps		Self–diffusion Coefficient 10^{-9} m^2 s^{-1}		n (molecules)	
	Calc [a]	Exptl	Calc [b]	Exptl	Cluster	bulk
benzene	6.7(‖),4.5(⊥)[c]	4.7(‖), 3.0(⊥)[c]	4.2	2.3[c]	484	250
chloroform	4.9	3.1–5.3[d]	3.4	2.3[d]	1024	432
methanol	20	15[e]	1.4	1.3–2.3[e], 2.4[f]	958	432
pyridine	9.1	3.0[g]	2.8	1.88[g]	432	250

[a]Same results for clusters or bulk.

[b] Cluster.

[c]MD simulation and experiment, ref. 15; (‖) parallel, (⊥) perpendicular to the ring plane normal.

[d] As quoted in ref. 25 along with MD results.

[e] MD calculation, ref. 26.

[f] As quoted in ref. 27, with comparable DFT–MD results.

[g] As quoted in ref. 28, along with similar results from MD calculations.

View Article Online
DOI: 10.1039/C8NJ05825C

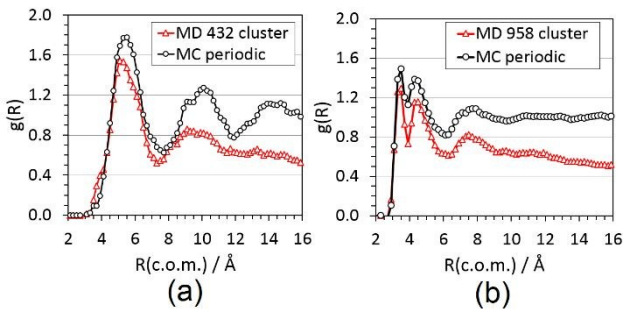


Fig. 3. Center–of–mass radial distribution functions, g(R), for bulk liquid from MC and for clusters from MD. (a) pyridine; (b) methanol.

Figure 3 shows a comparison of the center–of–mass radial distribution functions from bulk MC simulation and for clusters: they are in agreement with the results of accurate literature simulations on bulk liquids as well as with results from periodic-box simulation of bulk liquids. Similar results are obtained also for benzene and chloroform (see also Figure S5 ESI). The MD description of the hydrogen bond in methanol is adequate, with about one bond per molecule and RDF peaks at 1.8, 2.2 and 2.8 for the O···H, H···H and O···O distances, respectively (see Figure S6 ESI). The dynamic simulation correctly reproduces the effect of surface tension in changing a starting cubic slab into a spherical drop (Figure 4). Again remarkably, RDF's from large clusters are identical to those of bulk liquids.

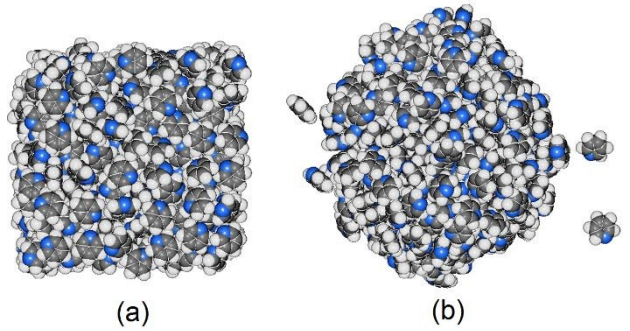


Figure 4. MD simulation of an isolated cluster of liquid pyridine. (a) Starting cubic slab; (b) final spherical droplet, showing a few incipient evaporated molecules.

Use of symmetry indices g_{ij} and G

During production runs of both MC and MD simulation for all considered liquids, the global *G*–index of eq. (11) is strictly constant. This is a valuable additional signal that the simulation has reached equilibration. Figure 5 shows the distribution of g_{ij} indices in the final frames of the cluster MD simulations. The trace for methanol shows a single peak more or less symmetric around the average, but the trace for pyridine is strongly skewed to the left, indicating that smaller values of g_{ij} predominate.According to this result the pyridine liquid counts a relatively large number of molecular pairs with some sort of approximate symmetry. Based on molecular shape, one can only guess that this is due to partial stacking of the aromatic rings. According to this result the pyridine liquid counts a relatively large number of molecular pairs with some sort of fractional symmetry, presumably of the translational or inversion type as proper of near bimolecular stacking.

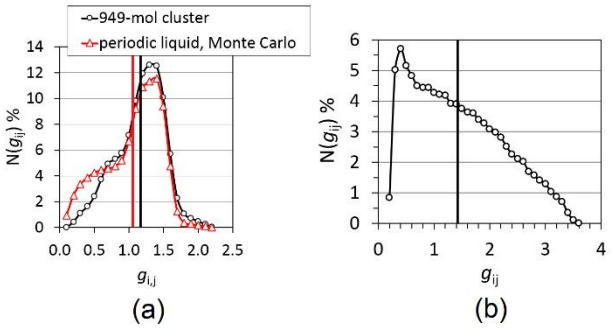


Figure 5. (a) The distribution of g_{ij} across the final frames of methanol. (b) Same as (a), for pyridine (427–mol cluster). Vertical bars indicate the corresponding average values.

The distribution of g_{ij} can also be represented as a function of distance between centers of mass: the result for benzene, and even more for pyridine (Figure 6) shows that the closest molecular pairs (which also have smaller interplanar angles, see Figure 1) tend to have a higher degree of symmetry. The same trend is observed in chloroform, but not in methanol. A possible interpretation is that closeness plays a role in inducing symmetry for systems whose aggregation is dictated by shape rather than by stronger interactions like a hydrogen bond.[29]These indications are by necessity somewhat qualitative. More experience on the use of this novel function is needed before clear-cut structural conclusions can be drawn.

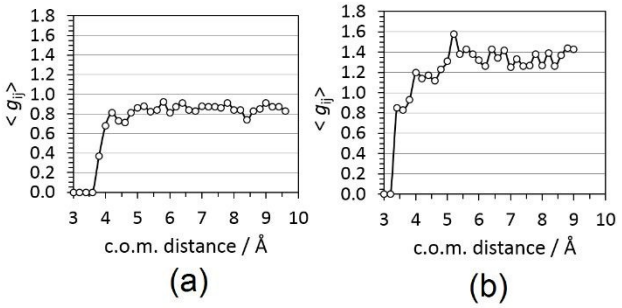


Fig. 6. Graphs for the average value of g_{ij} as a function of intermolecular distance between centers of mass. (a) Benzene cluster, 475 molecules at the last step. (b) Pyridine cluster, 427 molecules at the last step. Closer molecular pairs tend to have higher symmetry relationship.

Cluster size effects and evaporation rates

Simulations have been carried out for chloroform and for benzene with variable cluster sizes.Figure 7 and Table 4 show the results. For chloroform, smaller clusters have slightly smaller rotational correlation times; as the cluster becomes large enough, the correlation time becomes identical to that of the bulk liquid. Smaller clusters also have smaller cohesive energies and higher rms center–of–mass displacements, as expected. A cluster of about 1000 molecules is in all these respects quite similar to the bulk liquid. For benzene the results are a bit more erratic, and a 500–molecule cluster is still somewhat far away from the properties of bulk liquid. There seems to be an indication that cluster size is influential for benzene but not for chloroform, but the differences in rotational decay times could

View Article Online
DOI: 10.1039/C8NJ05825C

also be within the confidence intervals of the corresponding quantities.

Table 4. Results of variable–size cluster simulations for chloroform and benzene .

n (molecules)	radius, Å	$t(20)^a$	RMSD(10)[b]	$E_{coh}{}^c$ / kJ·mol^{-1}	Evaporation rate	
					mol/ps	10^5 · mol/(ps Å^2)
$CHCl_3$						
102	13	3.2	40	16	0.05	2.3
260	19	3.7	44	20	0.08	1.7
514	24	4.5	43	22	0.08	1.4
998	31	5.2	27	26	0.11	0.9
bulk	–	4.7	–	26	–	–
C_6H_6						
149	17	6.1	41	24	0.09	2.5
278	21	5.7	40	26	0.04	0.8
479	26	5.1	28	27	0.04	0.5
bulk	–	6.7	–	33	–	–

[a]Time (ps) for rotational correlation to decay from 100 to 20 % .
[b]Root–mean–square center–of–mass displacement after 10 ps simulation (starting diffusion; Å).
[c]Cohesive intermolecular energy (kJ/mol).

The simulations of evaporation were carried out by deleting the molecules for which the distance between the center of mass and the cluster center of mass exceeds the diameter of the cluster (see "Evaporation" in the Methods Section). Longer simulation times (up to 500 ps) were carried out in order to observe a sizeable number of evaporated molecules. The absolute evaporation rate, in molecules per picosecond, increases with increasing cluster surface, but the rate normalized to surface decreases due to increasing cohesion in larger clusters. An estimate of the lifetime of a 1000–molecule chloroform cluster or of a 500–molecule benzene cluster at room temperature, assuming a constant evaporation rate, is then about 10 ns.

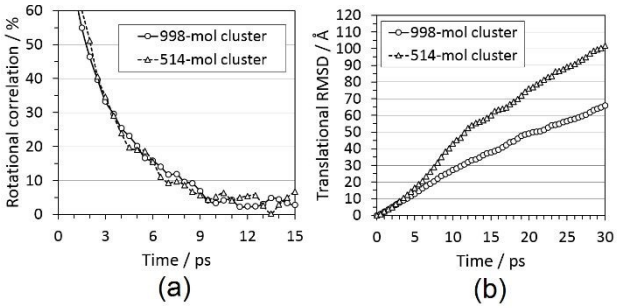


Figure 7. Correlation functions for chloroform*vs*. simulation time. (a) Rotational coefficient.(b) Translational root mean squared displacement. The results for benzene (see ESI, Figure S4) are quite similar.

Conclusions

This work describes a Molecular Dynamics (MD) formulation, in part using standard algorithms and in part explicitly devised for application to a study of the dynamic properties of isolated liquid nanoclusters, or droplets. Standard stretch-bend-torsion force fields are used to control intramolecular deformation without recourse to comlex constraint algorithms. The software includes simple, time-saving methods for controlling net rotational and translational momenta and for dealing with evaporation in isolated systems. The AA–CLP intermolecular force field is employed, originally parametrized to reproduce the bulk properties of molecular crystals and liquids. The setup is tested against nanoclusters of 150–1000 molecules of benzene, chloroform, methanol and pyridine, analyzing their cohesion energies, rotational correlation, self–diffusion coefficients, evaporation rates and internal structure. A newly developed algorithm is presented for the analysis of hidden or emerging molecule-molecule symmetry relationships, with reference to common space-group symmetry operations like inversion, screw or twofold rotation axes, and mirror or glide planes. On the other hand, this Molecular Dynamics setup includes a number of approximations, especially in the treatment of intramolecular flexibility, of temperature and pressure control, that prevent it from standing up against more refined and widely used MD packages. It should be taken rather as a tool and a help alongside the exploration of experimental occurrences in the realm of nanosized aggregates, not being intended as a device for the delivery of quantitative termophysical quantities by simulation. A more detailed summary of main results follows.

(a) Overall, appraisal of the applications confirms the reliability of the AA–CLP formulation for Molecular Dynamics simulations, complementing the already established reliability of the method in static lattice energy calculations and in Monte Carlo simulation.

(b) Simulations on nanoclusters for different chemistry (a hydrocarbon, a chloro- derivative, an azahydrocarbon, and a hydrogen bonded species) reveal the differences in dynamic behavior related to the relative strengths of intermolecular forces. Molecular mobility, in terms of rotational correlation times and self-diffusion coefficients, increases at smaller cluster sizes. Larger, 500– to 1000–molecule clusters are already quite similar to bulk liquids in internal structure (radial distribution functions), and thermochemical and dynamic properties (intermolecular correlation times).

(c) Evaporation rates are estimated by the number of molecules that leave an isolated cluster as a function of time and cluster surface area. For example, a 1000–molecule nanodroplet of a highly volatile substance like chloroform is predicted to have a lifetime of about 10 ns in isolation, while a cluster of hydrogen-bonded methanol molecules with similar dimensions is much more persistent.

(d) Using the above mentioned symmetry-control algorithm, no trace of persistent symmetric relationships are detected in bulk liquids or in clusters; however, a small but significant increase in symmetric arrangements appears for nearest neighbors of flat molecules like benzene and pyridine. These are presumably shape–driven due to incipient stacking, as demonstrated by the increase of small interplanar angles at short distance between molecular centers. This effect is absent in methanol and chloroform, where shape effects play a minor role. At least in principle, the algorithm shows promise for application to future studies of phase transitions (crystal nucleation and growth).

(e) Our simulations have been conducted for time lengths necessary and sufficient to characterize the main dynamic features of the sample. For bulk liquids, already heavily optimized by preliminary MC treatment, longer duration would have carried no new information. For isolated clusters, except for evaporation studies, simulation for a few hundred picoseconds was more than enough to allow the estimate of correlation times and to characterize structures. All calculations

View Article Online
DOI: 10.1039/C8NJ05825C

have been carried out in reasonable solar times of the order of a few hours on ordinary PCs without recourse to special computing resources; this approach might help transferring molecular simulation into the usual arsenal of general and experimental chemists, rather than being restricted to a small community of computer experts.

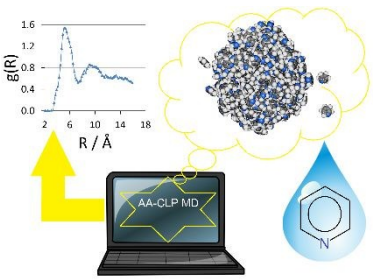
Finally, it is found that Molecular Dynamics supported by the AA-CLP potentials is able to reproduce key macroscopic thermodynamic quantities and to give a good picture of the dynamic behavior of organic liquids and droplets. As such it may provide a viable approach to the molecular–level study of the dynamics of nanoclusters in general, with potential application in non–equilibrium solution chemistry, exploring complex collective phenomena such as supersaturation / supercooling, the early stages of crystal nuclei formation and phase transitions, and spinodal decompositions. Such developments will be explored in forthcoming studies.

Acknowledgements. One author (LLP) wishes to thank Università degli Studi di Milano (Department of Chemistry) for partial funding through the Unimi Development Plan–Line 2, Action B, project NOVAQ, nº PSR2015-1716FDEMA_08.

Conflict of interest. There are no conflicts to declare.

References

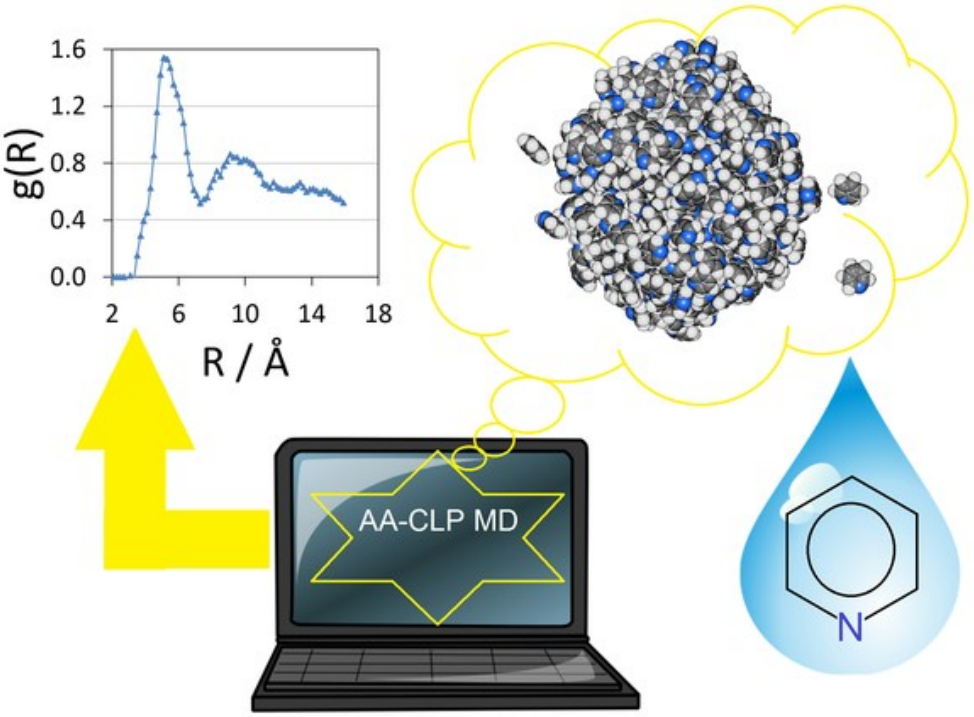
[1]L. S. Bartell, *Chem. Revs.* 1986, **86**, 491–505.
[2]B. E. Wyslouzil and J. Wölk, *J. Chem. Phys*. 2016, **145**, 211702–1,26.
[3]Y. J. Kim, B. E. Wyslouzil, G. Willemski, J. Wölk and R. Strey, *J. Phys. Chem.* 2004, **108**, 4365–4377.
[4]F. Yu and R. P. Turco, *J. Geophys. Res.* 2001, 106, 4797–4814.
[5]S.R. Gadre, S. D. Yeole and N. Sahu, Chem. Revs. 2014, **114**, 12132–12173.
[6]R. Angelil, J. Diemand, K. K. Tanaka and H. Tanaka, *J. Chem. Phys*. 2014, **140**, 074303.
[7]H. Matsubara, *J. Chem. Phys*. 2007, **127**, 214507
[8]W. F. van Gunsteren and H. J. C. Berendsen, A*ngew. Chem. Int. Ed.* 1990, **29**, 992–1023.
[9]See at http://www.gromacs.org
[10] A. Gavezzotti,*New J. Chem*. 2011, **35**, 1360–1368.
[11] A. Gavezzotti, *New J. Chem*. 2013, **37**, 2110–2119.
[12]A. Gavezzotti and L. Lo Presti, *Crystal Growth Des*. 2016, **16**, 2952–2962.
[13]A. Gavezzotti, V. Colombo and L. Lo Presti, *Crystal Growth Des.* 2016, **16**, 6095–6104.
[14]V. Colombo, L. Lo Presti and A. Gavezzotti, *CrystEngComm*, 2017,**19**, 2413-2423.
[15]L. Lo Presti, *CrystEngComm*, 2018, **20**, 5976-5989.
[16]A. Gavezzotti, S. Rizzato and L. Lo Presti, *Cryst. Growth Des*., 2018, **18**, 7219–7227.
[17] A. Gavezzotti, *The Coulomb – London – Pauli (CLP) model of intermolecular interaction, CLPdyn, Monte Carlo and Molecular Dynamics modules*, Description and user manual, www.angelogavezzotti.it (2018); follow the CLPdyn link, file CLPmanual.doc.
[18]A. Gavezzotti and L. Lo Presti, *Crystal Growth Des*. 2015, **15**, 3792–3803
[19]van Gunsteren, W. F., Billeter, S. R., Eising, A. A., Hunenberger, P. H., Kruger, P., Mark, A. E., Scott, W. R. P., Tironi, I. G. *Biomolecular Simulation: The GROMOS96 Manual and User Guide,* BIOMOS b.v., Zürich–Groningen, 1996.
[20] A. Gavezzotti, *New. J. Chem.* 2016, **40**, 6848–6853.
[21]I. Cacelli, G. Cinacchi, G.Prampolini and A. Tani, *J. Am. Chem. Soc.* 2004, **126**, 14278–14286
[22] P. M. Piaggi and M. Parrinello, *PNAS*, 2018, **115**, 10251–10256
[23] G. Gobbo, M. A. Bellucci, G. A. Tribello, G. Ciccotti, B. L. Trout, *J. Chem. Theory Comput*. 2018, **14**, 959−972
[24]I. Gimondi and M. Salvalaglio, *Mol. Syst. Des. Eng*. 2018, **3**, 243–252
[25]I.G.Tironi and W.F. van Gunsteren, *Mol. Phys.* 1994, **83**, 381–403.
[26]M.Haughney, M.Ferrario and R.MacDonald, *J. Phys. Chem.* 1987,**91**, 4934–4940
[27]J. W. Handgraaf, T.S. van Erp and E.J.Meijer, *Chem. Phys. Lett*. 2003, **367**, 617–624
[28]Z.Trumpakaj and B.Linde, *J. Mol. Struct*. 2015, **1085**, 268–275.
[29] R. Destro, E. Sartirana, L. Loconte, R. Soave, P. Colombo, C. Destro and L. Lo Presti, *Cryst. Growth Des*. 2013, **13**, 10, 4571-4582



For Table of Contents use only

Synopsis:*In a few hours on a standard laptop, AA–CLP MD correctly reproduces thermodynamic properties of bulk liquids and provides information on the nanoscale dynamics of liquid nanoclusters.*

View Article Online
DOI: 10.1039/C8NJ05825C



49x37mm (300 x 300 DPI)