Acta Cryst B
STRUCTURAL SCIENCE
CRYSTAL ENGINEERING
MATERIALS
ISSN 2052-5206

Exploring polymorphism of benzene and naphthalene with free energy based enhanced molecular dynamics

Elia Schneider,[a]‡ Leslie Vogt[a]‡ and Mark E. Tuckerman[a,b,c]*

[a]Department of Chemistry, New York University, New York, NY 10003, USA, [b]Courant Institute of Mathematical Science, New York University, New York, NY 10003, USA, and [c]NYU-ECNU Center for Computational Chemistry at NYU Shanghai, 3663 Zhongshan Road North, Shanghai 200062, China. *Correspondence e-mail: mark.tuckerman@nyu.edu

Received 16 February 2016
Accepted 13 May 2016

Edited by G. M. Day, University of Southampton, England

‡ These authors contributed equally.

Keywords: crystal structure prediction; enhanced sampling; molecular dynamics; benzene; naphthalene.

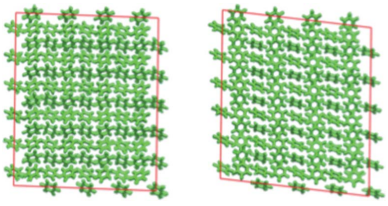
Supporting information: this article has supporting information at journals.iucr.org/b

Prediction and exploration of possible polymorphism in organic crystal compounds are of great importance for industries ranging from organic electronics to pharmaceuticals to high-energy materials. Here we apply our crystal structure prediction procedure and the enhanced molecular dynamics based sampling approach called the Crystal-Adiabatic Free Energy Dynamics (Crystal-AFED) method to benzene and naphthalene. Crystal-AFED allows the free energy landscape of structures to be explored efficiently at any desired temperature and pressure. For each system, we successfully predict the most stable crystal structures at atmospheric pressure and explore the relative Gibbs free energies of predicted polymorphs at high pressures. Using Crystal-AFED sampling, we find that mixed structures, which typically cannot be discovered by standard crystal structure prediction methods, are prevalent in the solid forms of these compounds at high pressure.

1. Introduction

Predicting the crystal structures of organic molecules is an immensely challenging task. Great progress has been made in crystal structure prediction (CSP) in the last decade, as shown by the increasing ability to correctly determine low-energy packed structures starting only from a molecular diagram (Lommerse *et al.*, 2000; Motherwell *et al.*, 2002; Day *et al.*, 2005, 2009; Bardwell *et al.*, 2011; Reilly *et al.*, 2016). However, it is still difficult to establish which of multiple polymorphs will be obtained under a given set of experimental conditions (Bernstein, 2007; Price, 2013). One of several significant bottlenecks is the problem of calculating the relative free energies of the crystal structures, for which specialized techniques based on statistical mechanics are needed (Dunitz & Scheraga, 2004; Yu & Tuckerman, 2011*b*; Raiteri *et al.*, 2005).

As simple yet pertinent examples, the high-pressure structures of well studied molecules such as benzene and naphthalene have been in question for decades. For solid benzene under standard conditions, the crystal structure (known as benzene I) is well known and has been accurately characterized using theoretical methods (Yang *et al.*, 2014). However, the structure at elevated temperatures and pressures is difficult to determine experimentally (Block *et al.*, 1970; Akella & Kennedy, 1971; Thiéry & Léger, 1988; Katrusiak *et al.*, 2010). Using computational methods based on enhanced sampling techniques (Yu & Tuckerman, 2011*b*; Yu *et al.*, 2014; Martoňák *et al.*, 2003), several pure structures previously proposed theoretically (Dzyabchenko, 1984; Rérat & Rérat, 1990; van Eijck *et al.*, 1998) were seen to be stable only at high pressures (Raiteri *et al.*, 2005). Further characterization of benzene at



2 GPa using enhanced molecular dynamics (MD) based free-energy sampling methodology showed that mixed structures with stacking faults had comparable free energies (Yu & Tuckerman, 2011*b*).

Given the suggestion that benzene is able to support mixed structures at high pressure, we were prompted to examine more closely the behavior of the second smallest member of the acene family, naphthalene, under the same conditions. Early experimental observations indicated that unlike benzene, the volume changes observed for naphthalene at high pressure (Bridgman, 1938) were not the result of a new polymorph (Block *et al.*, 1970). Recent characterization based on powder X-ray diffraction data indicates that there may be amorphous structures, as opposed to a new polymorph, at elevated temperatures and pressures in naphthalene (Likhacheva *et al.*, 2014). Predicting amorphous structures, mixed structures, stacking faults and other such motifs in molecular systems cannot be done using standard crystal structure prediction techniques, which only identify pure crystals. A set of techniques better suited to the discovery of these types of structures are derived from statistical mechanics and enhanced sampling. Not only can such methods identify purely ordered structures, but if amorphous, mixed or stacking-fault structures are lower in free energy and, consequently, more stable, such structures should appear in the sampling.

In this article we demonstrate that enhanced free-energy based molecular dynamics sampling techniques are able to correctly predict the relative stability of benzene polymorphs at high pressure. At intermediate pressures of ~ 1 GPa, several structures, including those with stacking faults, have similar free energies. When applied to naphthalene, the methodology generates stable structures that exhibit stacking faults or misaligned domains, which appear to be among the dominant structures under these experimental conditions.

2. Crystal structure prediction

2.1. Force field

Based on the prior success of Buckingham-type (exp-6-1) potentials to describe polymorphs of benzene (Dzyabchenko, 1984; Rérat & Rérat, 1990; van Eijck *et al.*, 1998), we used a modified Williams 99 intermolecular potential (Williams, 2001) with no bond foreshortening (referred to as mW99 in this work) to study high-pressure polymorphs of benzene and naphthalene. Atomic point charges were assigned by fitting the electrostatic potential (Bayly *et al.*, 1993) of each molecule optimized at the RHF/6-31G* level of *ab initio* electronic structure theory as implemented in *GAUSSIAN*09 (Frisch *et al.*, 2009). For the initial round of random structure generation, we used the OPLS van der Waals parameters (Jorgensen *et al.*, 1996). MD simulations were run using our mW99 parameters and intramolecular interaction parameters from the Generalized Amber force field (GAFF; Wang *et al.*, 2004), as assigned by the *antechamber* program (Wang *et al.*, 2001). The force-field parameters are reported in Tables S1, S2 and S3 of the supporting information.

2.2. Crystal structure generation

Sets of possible crystal structures for benzene and naphthalene were generated using the *UPACK* program suite (van Eijck & Kroon, 1999). To avoid short interatomic distances that are poorly described by the exponential form of the force field, random structure generation was performed using the OPLS van der Waals parameters (Jorgensen *et al.*, 1996). In this step, 5000 structures were generated for each of the 13 most common space groups for organic molecules ($P2_1/c$, $P\bar{1}$, $P2_12_12_1$, $P2_1$, $Pbca$, $C2/c$, $Pna2_1$, Cc, $Pca2_1$, $C2$, $P1$, $Pbcn$, Pc), with rigid molecules and an external pressure of 1 bar. Lattice energies were evaluated using a cutoff of 12 Å with an Ewald damping range of $\alpha = 3$ nm^{-1} and reciprocal space cutoff of 2 nm^{-1} for both Coulomb and dispersion terms. Structures within 30 kJ mol^{-1} of the lowest energy in each space group were subsequently optimized using our mW99 force field. Finally, these structures were clustered with the radial distribution function available in *UPACK*, using a cutoff of 7 Å and a tolerance of 0.25 Å to remove duplicates. We found 147 and 189 unique structures within 10 kJ mol^{-1} of the lowest energy for benzene and naphthalene, respectively.

2.3. Molecular dynamics refinement

We performed isothermal–isobaric MD simulations in a fully flexible simulation cell in order to evaluate the stability and relative energy ranking of candidate crystal structures. The 84 lowest energy unit cells from the *UPACK* ranking were used to build supercells containing 256 and 512 molecules for benzene and naphthalene, respectively, for which all sides were at least 25 Å in length. MD simulations were run with the *PINY_MD* package (Tuckerman *et al.*, 2000) using the mW99 intermolecular potential and the GAFF intramolecular parameters reported in §2.1. Energies were evaluated with an intermolecular cutoff of 12 Å and a Particle Mesh Ewald summation ($\alpha = 3.75$ nm^{-1}, $k_{\max} = 0.55$ nm^{-1}) for the periodic Coulomb interactions.

Each supercell was minimized prior to a 10 ps constant volume isothermal (NVT) simulation at $T = 100$ K. MD simulations used massive Nosé–Hoover chain (NHC) thermostats (Martyna *et al.*, 1992; length = 4, $\tau = 20$ fs, Suzuki–Yoshida order = 7, multiple time step = 4, one thermostat on each degree of freedom) and a time step of 1 fs. The r-RESPA multiple time step approach (Tuckerman *et al.*, 1992) was employed to compute rapidly varying forces n times per step (bond/bends $n = 6$, torsions $n = 4$). The atomic positions and velocities were then used as the starting point for a 40 ps flexible-cell isothermal–isobaric (NPT) MD simulation to equilibrate each crystal structure at $P = 1$ bar, using the MTK barostat (Martyna *et al.*, 1994, 1996) and reversible, measure-preserving integrator of Tuckerman and coworkers (Tuckerman *et al.*, 2006; Yu *et al.*, 2010; $\tau = 500$ fs) and an NHC thermostat ($\tau = 100$ fs) on the barostat. An additional 100 ps of flexible-cell NPT MD was run to evaluate the average potential energy of each structure. Unit-cell parameters were determined by averaging supercell snapshots every 1 ps along

the trajectory. For each predicted structure, the space group was then assigned using *PLATON* (Spek, 2015).

2.4. Crystal-Adiabatic Free Energy Dynamics

In addition to demonstrating the stability of a given polymorph, an MD simulation can, in principle, be used to determine the relative free energies between a set of conformations from their relative populations sampled during the trajectory. However, exploring polymorphism in organic molecular crystals requires sampling on a rough and complex energy landscape, not unlike the conformational space exploration of a protein (Dunitz & Scheraga, 2004). In practice, standard MD runs are impractical for determining the relative free energy of molecular crystal structures due to the computational time required to observe transitions over the large energy barriers that separate the polymorphs. To address this issue, we have employed the Crystal-Adiabatic Free Energy Dynamics (Crystal-AFED) enhanced sampling approach (Yu & Tuckerman, 2011*b*).

The Crystal-AFED method is derived from its canonical precursors, Adiabatic Free Energy Dynamics (AFED; Rosso *et al.*, 2002) and driven Adiabatic Free Energy Dynamics (d-AFED; Abrams & Tuckerman, 2008), both of which subject a set of collective variables (CVs), capable of distinguishing conformational states of a system, to enhanced sampling *via* temperature acceleration and adiabatic decoupling. Crystal-AFED is based on flexible-cell isothermal–isobaric MD in which the components of the supercell vectors, or elements of the **h**-matrix, are used as CVs. Targeting these system variables for the aforementioned enhanced sampling generates large fluctuations in the shape and size of the simulation cell and allows the MD run to explore a greater range of molecular packing interactions in the condensed phase environment. Each element of the **h**-matrix is coupled to a heat bath, here the generalized Gaussian moment thermostat of Liu & Tuckerman (2000), which is maintained at a high temperature $T_{\mathbf{h}} \gg T$, where T is the physical temperature. In practice, each diagonal element has an independent heat bath and one heat bath is coupled to all off-diagonal terms; for more details see Yu & Tuckerman (2011*a*). In order to ensure that these fluctuations are adiabatically decoupled from the atomic degrees of freedom, *i.e.* the algorithm directly explores the Gibbs Free Energy $G(\mathbf{h}, T)$, the effective mass of the barostat is much larger than it would be for a standard MD run. In *Appendix A*, we provide the equations of motion for additional clarification.

In the resulting MD, slow but large fluctuations of the cell vectors promote changes in the intermolecular packing arrangement. This method has been successfully applied to explore polymorphs of benzene at 2 GPa using a different force field (Yu & Tuckerman, 2011*b*). In this work, we extend the use of this enhanced sampling method to determine the relationship between benzene polymorphs over a range of applied pressures, and we apply it to explore the behavior of solid naphthalene at high pressures.

Given sufficiently long simulations, Crystal-AFED can be used to explore and generate the Gibbs free-energy surface $G(\mathbf{h}, T)$, from which relative Gibbs free energies for different crystal structures can be determined directly from the population of each structure (Yu & Tuckerman, 2011*b*). However, as was shown by Yu & Tuckerman (2011*b*), the accessible structures under a given set of conditions are generally visited at least once in a run much shorter than that needed to converge the full set of relative populations of these structures. Therefore, in this work, we employed relatively short simulations to explore possible polymorphism and look for transition paths between two ordered structures. Once a path has been identified from these short exploratory trajectories, thermodynamic integration (TI) is used to compute the free-energy difference between the two structures (see *Appendix B*). In order to characterize structural transitions, snapshots taken every 10 ps are annealed with 60 ps NVT MD runs to obtain the forces on the simulation cell at each step along the Crystal-AFED trajectory path. The structures from the free-energy minima along this path are subsequently equilibrated using a standard flexible-cell NPT MD run at the desired pressure. We note that the choice of CV is not limited to the **h**-matrix used in this work. Crystal-AFED was recently applied to predict polymorphs and stacking faults in high-pressure xenon (Yu *et al.*, 2014) and the melting transition of solid copper (Samanta *et al.*, 2014) using CVs characterizing atomistic packing arrangements, and molecular CVs describing relative separations and orientations is the subject of ongoing work.

3. Results

3.1. Ranked crystal structures

The average potential energy from the flexible isothermal–isobaric MD simulations is used to rank the candidate crystal structures for benzene and naphthalene. We find that at T = 100 K and P = 1 bar, the experimentally observed structure is the lowest energy for each molecule. To compare predicted and experimental structures, we use the Crystal Packing Similarity Tool (Chisholm & Motherwell, 2005) available in *Mercury*3.6 (Macrae *et al.*, 2008) to compare clusters of 20 molecules. The root mean-squared deviation (RMSD) is calculated for the 20 overlaid molecules if intermolecular distances and angles of all 20 molecules agree within 20% and 20°, respectively.

For benzene, the predicted low-energy *Pbca* structure matches the experimental form (I) (Nayak *et al.*, 2010) with an RMSD_{20} = 0.20 Å after equilibration at T = 100 K and RMSD_{20} = 0.24 Å at the reported crystal temperature T = 150 K. We also found the previously reported form (I′) (Thiéry & Léger, 1988; Raiteri *et al.*, 2005; Yu & Tuckerman, 2011*b*), with a space group of *Cmca*, approximately +5.6 kJ mol^{-1} higher in energy; however, this structure readily converted to form (I) during the 100 K flexible-cell MD simulations.

For naphthalene, we also obtain the experimentally observed $P2_1/c$ packing as the lowest energy, with an RMSD_{20}

= 0.29 Å for the structure reported at 100 K (Oddershede & Larsen, 2004) and $RMSD_{20}$ = 0.13 Å for the structure at room temperature (Capelli *et al.*, 2006). While the simulation temperature is not expected to match the physical temperature for the mW99 force field parameters, the fact that our $RMSD_{20}$ values are consistently less than 0.3 Å indicates that the force field is capable of capturing those intermolecular interactions that are most important for crystal structure stability in these molecules.

The distribution of potential energies for all ranked crystal structures is shown in Fig. 1 for both benzene and naphthalene. The reported densities are calculated from the equilibrated MD results and are smaller than experimental values for this set of force field parameters. Despite this systematic shift, the difference in the polymorph landscape between benzene and naphthalene is clear. Interestingly, there are a number of structures found at similarly low energies for benzene, including the high-pressure polymorph (II) (Katrusiak *et al.*, 2010) as the most densely packed form even at P = 1 bar. For naphthalene, the experimentally observed structure is more clearly separated from other molecular packings as the low-energy polymorph.

3.2. Pressure effects on benzene crystal structure stability

To evaluate the stability of high-pressure benzene crystal structures, we ran MD simulations for low-energy polymorphs

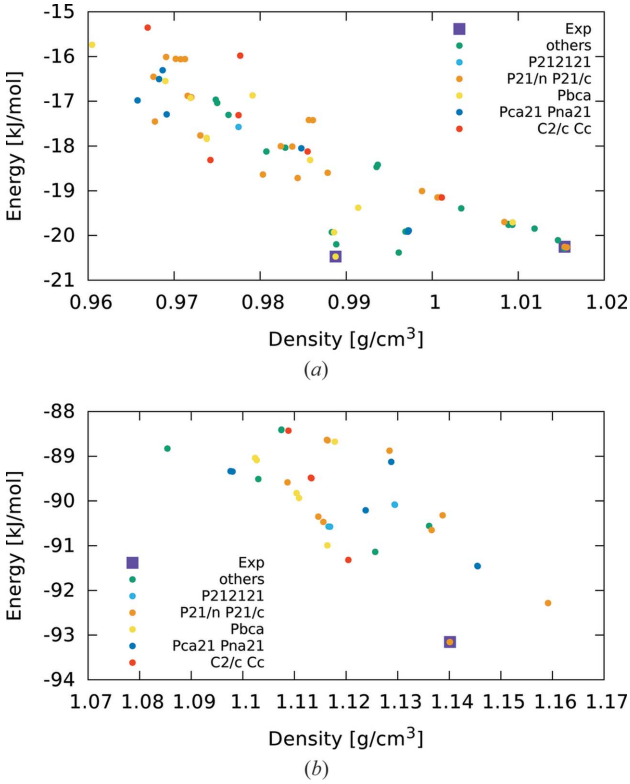


Figure 1
Potential energy per molecule *versus* predicted density for structures after MD refinement. Structures are shown for (*a*) benzene and (*b*) naphthalene, with purple squares indicating the experimental polymorphs after MD equilibration at T = 100 K and P = 1 bar.

Table 1
Predicted unit-cell parameters (in Å and °) for benzene structures from averaged flexible-cell NPT MD simulations run at T = 100 K with the mW99 force field.

SG is the space group for each structure, P is the simulation pressure and V_m is the molecular volume in Å^3

Polymorph	SG	P	a	b	c	β	V_m
(I)	*Pbca*	1 bar	7.28	7.63	9.43	90	131.0
(II)	$P2_1/c$	1 GPa	5.40	5.61	7.92	106.1	113.4
(V′)	$P2_1/c$	2 GPa	10.12	6.11	7.07	97.3	108.4

for a range of applied pressures from 1 bar (= 10^{-4} GPa) to 10 GPa. The results for benzene forms (I) and (II) are shown in Fig. 2. At ≥ 1.5 GPa, benzene (I) converts to a new structure during standard flexible-cell NPT simulations at T = 100 K. This new polymorph [named (I(HP))] is not predicted to be observed as it has a higher potential energy than form (II), which is stable at all simulated pressures, but (I(HP)) is the relevant structure for comparison of possible polymorphs at high pressures. Cell parameters are reported in Table 1 and in the supporting information Table S4.

We find that the mW99 force field is also adequate to describe the high-pressure form of benzene. While the alignment of the 0.97 GPa form (II) structure (Katrusiak *et al.*, 2010) with the MD structure at P = 1 bar results in $RMSD_{20}$ = 0.43 Å, the P = 1 GPa structure has $RMSD_{20}$ = 0.26 Å. Therefore, we proceeded to use the Crystal-AFED method to identify possible transition paths between the benzene structures at elevated pressures.

3.3. Crystal-AFED sampling of benzene polymorphs

To explore the relationship between benzene polymorphs, we ran a number of 500 ps Crystal-AFED simulations with T_h in the range (2.5–7.5) × 10^4 K and large effective mass (τ_h = 2 ps). Since the experimental transition from benzene (I) to

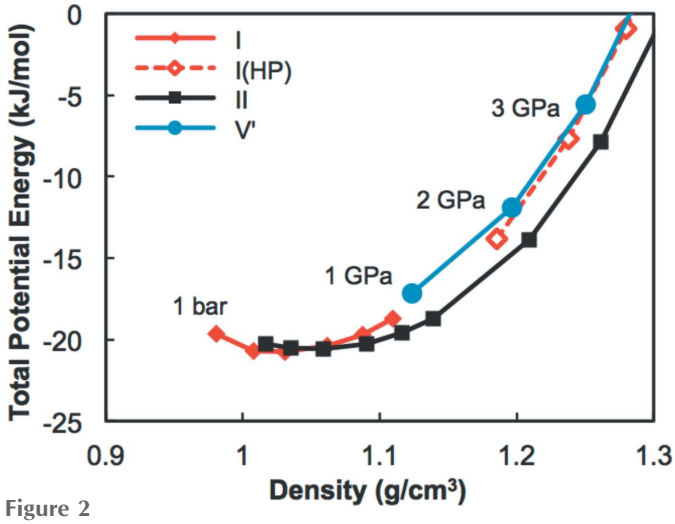


Figure 2
Potential energy per molecule for selected benzene structures at pressures ranging from 1 bar to 4 GPa. Structures with stacking faults [similar to III(d) from (Yu & Tuckerman, 2011*b*)] have a pressure-dependent behavior that is indistinguishable from pure form (II).

Table 2
Relative energy, enthalpy and Gibbs free energy differences (kJ mol^{-1}) between benzene polymorphs (II) and (I) (above 1 GPa, the new form (I(HP)).

P (GPa)	0.0001	0.5	1	1.5	2	3	5	10
ΔE	−0.2	−0.1	0.0	0.0	0.0	0.2	0.6	2.1
ΔH	−0.2	0.8	1.8	2.1	2.5	3.7	6.0	10.7
ΔG	−3.0	−1.7	−0.5	0.6	1.9	3.8	6.3	13.8

Table 3
Relative energy, enthalpy and Gibbs free-energy differences (kJ mol^{-1}) between benzene polymorphs (II) and (V′), which are not stable below 1 GPa.

P (GPa)	1	1.5	2	3	5	10
ΔE	1.6	1.81	2.0	2.3	2.8	4.0
ΔH	2.5	3.0	3.3	3.9	4.8	6.3
ΔG	0.0	0.7	1.4	2.6	4.6	9.1

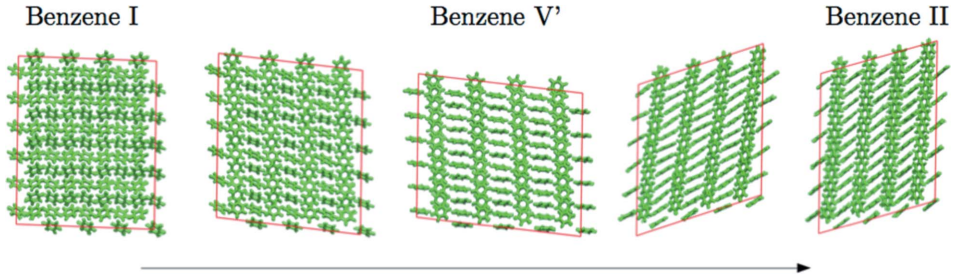


Figure 3
A smooth phase transition from a Crystal-AFED trajectory at T = 100 K and P = 2 GPa.

benzene (II) has been observed at ∼1.3 GPa at room temperature (Akella & Kennedy, 1971; Katrusiak *et al.*, 2010), we also ran Crystal-AFED with external pressure in the range P = 1 bar to P = 2 GPa.

During all Crystal-AFED simulations, benzene (I) transforms into other packing structures. For example, a commonly observed structure in the Crystal-AFED trajectories is similar to structure (V) reported previously (Raiteri *et al.*, 2005) and is denoted (V′) in this work. At low pressures, the large box fluctuations induce localized melting, even at an atomic temperature of T = 100 K. Simulations with larger external pressures, however, maintain crystalline order throughout the supercell and are used for our analysis of benzene polymorph transitions. As an example, a Crystal-AFED trajectory with $T_{\mathbf{h}}$ = 50 000 K and P = 2 GPa is analyzed in detail below.

The P = 2 GPa Crystal-AFED simulation samples structures with calculated densities ranging from 1.11 to 1.25 g cm^{-3}.

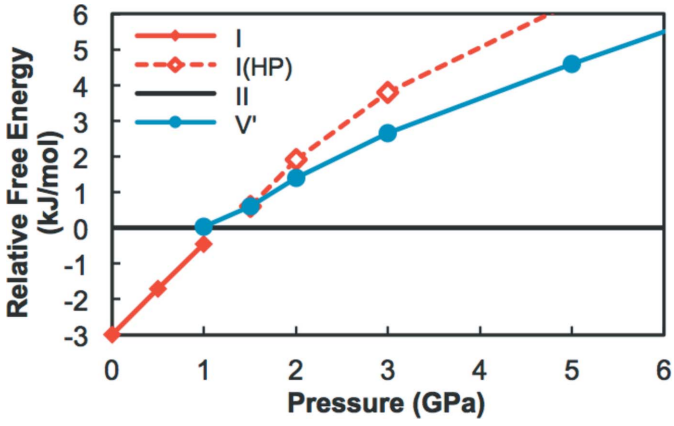


Figure 4
Calculated Gibbs free-energy difference relative to benzene (II) as a function of applied pressure. Lines connect the points determined for each distinct structure.

Over the course of the 500 ps MD simulation, several ordered structures were observed, with benzene (II) as the predominant form.

The typical transition proceeded from benzene (I) to (V′) to (II), with all the molecules in the supercell uniformly switching to the new structure within a few ps (see Fig. 3). During some simulations, a layer defect similar to the structure named III(d) in Yu & Tuckerman (2011*b*) was observed; these defects were stable even at low pressures and have a pressure-dependent behavior that is indistinguishable from the pure benzene (II) form.

Using the TI method, we were able to calculate relative Gibbs free energies of the benzene polymorphs at various pressures. As shown in Fig. 4, below 1 GPa the most stable polymorph is benzene (I). Furthermore, at this low pressure, we observed that benzene (V′) is not stable and converts to benzene (I) in the first 5 ps during standard flexible-cell NPT simulations. Between 1 and 1.5 GPa benzene (I), (II) and (V′) polymorphs are stable, and thus we can compute their relative free energies. In addition, at these pressures we observe that an interconversion between the polymorphs is thermally accessible (the difference in free energy being smaller than 1 kJ mol^{-1}). Above 1.5 GPa we can also determine the relative free energy between benzene (I(HP)) and the observed (II), and conclude that the most stable structure is benzene (II), as shown in Fig. 4.

In Tables 2 and 3 we report the comparison between the relative energy, enthalpy and Gibbs free-energy differences between benzene polymorphs; this comparison serves to underscore the importance of using free-energy differences to rank the polymorphs. At low pressure the energy difference between polymorphs (II) and (I) is only around −0.2 kJ mol^{-1}, while the free energy difference of −3.0 kJ mol^{-1} clearly indicates the stability of benzene (I). At high pressure the stability inverts, and benzene (II) clearly becomes the more stable structure based on the free energy difference. Above 1.5 GPa the (I(HP)) polymorph has free energy higher than (V′) despite having a lower potential energy. These examples demonstrate the relevance of the free energy evaluation in the study of crystal structure stability.

By comparing different structures along the Crystal-AFED trajectory that relax to the same polymorph, we estimate an error of ∼0.1 kJ mol^{-1} in the free energy calculation.

3.4. Pressure effects on naphthalene crystal structure stability

We ran high-pressure MD simulations for low-energy structures of naphthalene under the same conditions described above for benzene. Similar to benzene (II), the $P2_1/c$ experimental structure of naphthalene is found to be stable at simulation pressures up to 20 GPa (see Fig. 5). As we did not observe any crystal transitions during standard MD runs at T = 100 K, we used Crystal-AFED to determine structural transformations that could account for the experimentally observed volume changes at high pressures (Bridgman, 1938; Block *et al.*, 1970; Likhacheva *et al.*, 2014).

As for benzene, Crystal-AFED simulations were run for a range of external pressures and cell temperatures, $T_{\mathbf{h}}$. In all high-pressure Crystal-AFED trajectories, local defects were introduced in the otherwise ordered naphthalene packing structure. Some of these defects involved stacking faults similar to benzene (II); these defects were only slightly higher in energy than the known naphthalene structure. At high pressures, some trajectories produced a structure in which the naphthalene molecules were aligned in sheets (orange circles in Fig. 5), but each observed structure contained defects and was, on average, 13 kJ mol^{-1} higher in energy than the experimental polymorph. Another type of defect was also obtained during several independent Crystal-AFED simulations. In these MD runs, domains were introduced with different relative orientations along one crystal axis, as shown in Fig. 6. These domains persist even at low pressures during standard flexible-cell NPT simulations, with approximately 25% of the naphthalene molecules in the minority domain. We expect that with a larger supercell, a variety of domain sizes would be supported, similar to the results reported for mixed benzene structures (Yu & Tuckerman, 2011*b*; Raiteri *et al.*, 2005).

Using the Crystal-AFED method for benzene, we observed both crystal-to-crystal transitions and structures with lattice defects. Since all Crystal-AFED trajectories of naphthalene at high pressure produced supercells with stacking faults or misaligned domains, we suggest that formation of these partially disordered structures is responsible for the volume changes observed as naphthalene is compressed (Bridgman, 1938; Block *et al.*, 1970; Likhacheva *et al.*, 2014). The computational cost of larger supercells prohibits sampling the full entropic effect of disordered domains, but we note that previous work on naphthalene polymorphs based on single unit cells was inherently unable to identify this type of long-range disorder (Likhacheva *et al.*, 2014; Fedorov *et al.*, 2015). The Crystal-AFED approach is therefore a valuable tool to determine the likely polymorphs, and types of local disorder, found in molecular crystals.

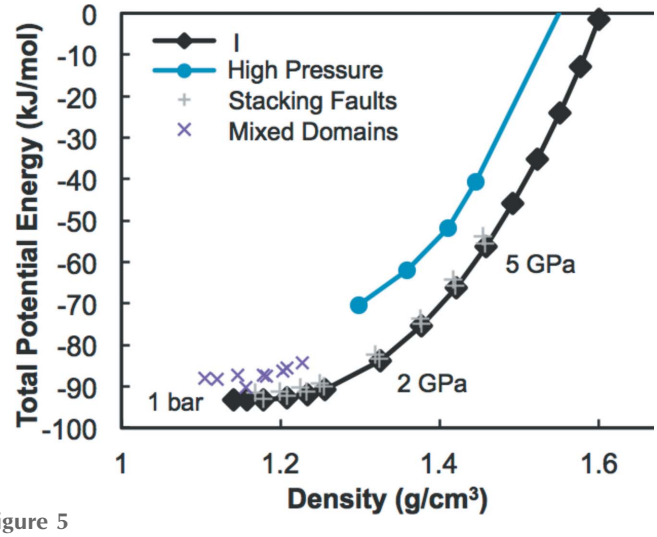


Figure 5
Potential energy per molecule for selected naphthalene structures at pressures ranging from 1 bar to 10 GPa. A set of similar high-pressure structures were stable above 2 GPa, but all contained some disorder. All higher-energy structures sampled by the Crystal-AFED method contain defects (stacking faults or mixed domains) that are stable during standard flexible-cell NPT simulations.

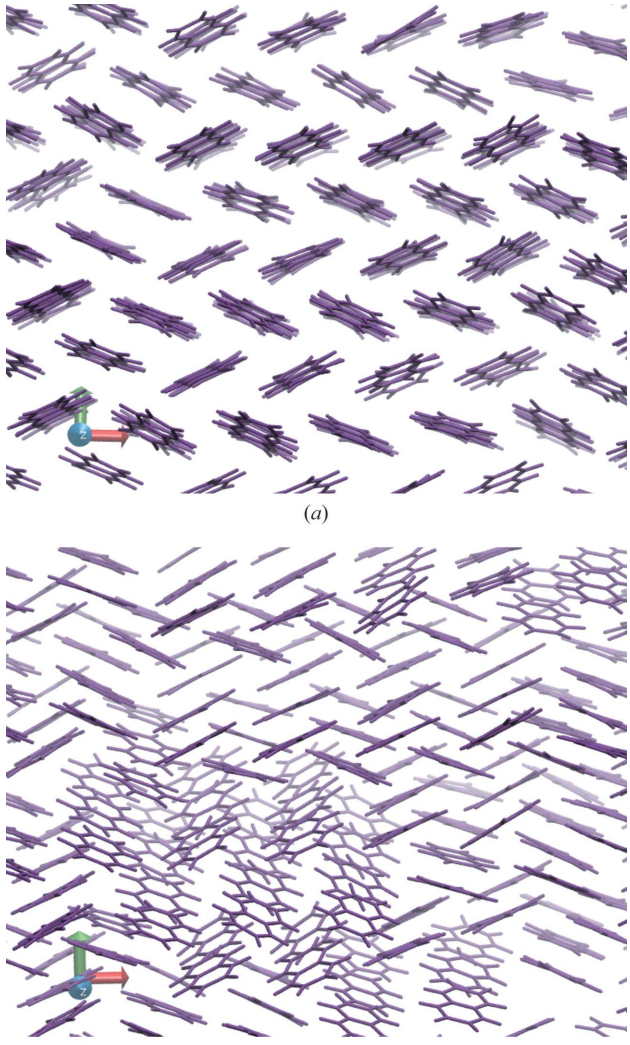


Figure 6
Naphthalene structure at 2 GPa with 25% of molecules aligned with a different orientation relative to the *a*-axis. Both figures show the same structure with the view along (*a*) the *c*-axis and (*b*) rotated 15° around the *a*-axis (indicated by the red arrow in the lower-left corner of each panel).

4. Conclusion

In this work, we have successfully predicted the low-energy structures of benzene and naphthalene using our CSP procedure and Crystal-AFED approach. We have also investigated the relationships between possible polymorphs of these molecules at high pressure.

A set of possible crystal structures was generated using *UPACK* and ranked based on the mW99 force field. Then, flexible-cell NPT MD runs were performed in order to investigate the stability of the predicted structures and to include thermal effects in our crystal structure prediction. Determination of the $RMSD_{20}$ values between our predicted structures and the experimental crystals reported in the literature shows that we are able to obtain good agreement with the reported structures of benzene and naphthalene.

During benzene Crystal-AFED simulations, we observed a transition between different crystalline polymorphs, which allowed us to estimate their relative Gibbs free-energy difference at different pressures using the TI technique. This study shows the importance of the free energy, rather than lattice energy, in obtaining a correct ranking of polymorphs, particularly under high-pressure conditions. Furthermore, our calculations demonstrated the robustness of the Crystal-AFED technique in obtaining results consistent with previous theoretical and experimental studies. In particular, in comparing our work to a previous Crystal-AFED study on benzene (Yu & Tuckerman, 2011*b*), we note that this approach is independent of the force field employed and could be used just as well within *ab initio* molecular dynamics when force fields are unavailable. In the present case, an improved description of benzene interactions, based on high-quality *ab initio* calculations (Podeszwa *et al.*, 2008; Yang *et al.*, 2014), would refine the estimate of free-energy differences.

Both benzene and naphthalene Crystal-AFED simulations predict the presence of stable mixed stacking structures at high pressure. In naphthalene, alternate defect states composed of mixed domains with commensurate lattice vectors are also predicted to be present. The prevalence of mixed structures could explain the experimental difficulties that have existed thus far in solving the high-pressure structures for both benzene and naphthalene.

APPENDIX *A*
Equations of Crystal-AFED

The equations of motion of Crystal-AFED are based on the barostat of Martyna *et al.* (1994, 1996), which generate an anisotropic isothermal–isobaric ensemble. Consider a molecular crystal consisting of N atoms with masses $m_1, ..., m_N$ at temperature T and pressure P. The atoms have coordinates $\mathbf{r}_1, ..., \mathbf{r}_N$ and momenta $\mathbf{p}_1, ..., \mathbf{p}_N$, and they interact *via* a potential $U(\mathbf{r}_1, ..., \mathbf{r}_N)$. The periodic simulation cell is described by vectors **a**, **b** and **c**, which are collected into the columns of the cell matrix denoted **h**, where the cell volume $V = \det(\mathbf{h})$. The equations of Crystal-AFED read

$$\begin{aligned}
\dot{\mathbf{r}}_i &= \frac{\mathbf{p}_i}{m_i} + \frac{\mathbf{p}_\mathbf{g}}{W}\mathbf{r}_i \\
\dot{\mathbf{p}}_i &= \mathbf{F}_i - \frac{\mathbf{p}_\mathbf{g}}{W}\mathbf{p}_i - \frac{1}{N_\mathrm{f}}\frac{\mathrm{Tr}[\mathbf{p}_\mathbf{g}]}{W}\mathbf{p}_i + \mathbf{H}_i(T) \\
\dot{\mathbf{h}} &= \frac{\mathbf{p}_\mathbf{g}\mathbf{h}}{W} \\
\dot{\mathbf{p}}_\mathbf{g} &= \det(\mathbf{h})\left[\mathbf{P}^{(\mathrm{int})} - P\mathbf{I}\right] + \frac{1}{N_\mathrm{f}}\sum_{i=1}^{N}\frac{\mathbf{p}_i^2}{m_i}\mathbf{I} + \mathbf{H}(T_\mathbf{h}).
\end{aligned} \tag{1}$$

Here, $\mathbf{H}_i(T)$ is a vector of heat baths that couple each degree of freedom to its own thermostat [the Nosé–Hoover chain algorithm (Martyna *et al.*, 1992) is used in this work], N_f is the number of degrees of freedom, $\mathbf{F}_i = -\partial U/\partial \mathbf{r}_i + \mathbf{f}_i^{(\mathrm{constr})}$ is the force on particle i, including any holonomic constraints that might be imposed, $\mathbf{I}$ is the 3×3 identity matrix, $\mathbf{p}_\mathbf{g}$ is a 3×3 barostat matrix that ensures the average of the internal pressure tensor $\mathbf{P}^{(\mathrm{int})} = V^{-1}\sum_{i=1}^{N}[\mathbf{p}_i\mathbf{p}_i + \mathbf{r}_i\mathbf{F}_i]$ is equal to the experimental pressure P, and $\mathbf{H}(T_\mathbf{h})$ is a matrix of heat baths on the barostat that ensure the barostat has a temperature $T_\mathbf{h} \gg T$, as indicated in the main text. In this work, the barostat thermostat is the generalized Gaussian moment thermostat (GGMT; Liu & Tuckerman, 2000). This high-temperature separation ensures that barriers to changes in the shape of the cell can be easily crossed. The parameter W determines the timescale of the barostat and is chosen so sufficiently large that the evolution of the cell is adiabatically decoupled from the evolution of the molecules in the cell. These equations can be expanded to include additional collective variables describing the internal orientations and packing arrangements, as was shown by Yu *et al.* (2014). In the work of Yu & Tuckerman (2011*b*), the above equations of motion were used to generate the Gibbs free energy surface $G(\mathbf{h}, T)$ as a function of the cell matrix from the adiabatic probability distribution $P_\mathrm{adb}(\mathbf{h}, T, T_\mathbf{h})$ generated by a long Crystal-AFED trajectory *via* $G(\mathbf{h}, T) = -k_\mathrm{B}T_\mathbf{h}\ln P_\mathrm{adb}(\mathbf{h}, T, T_\mathbf{h})$. In the present work, however, short Crystal-AFED trajectories are used to generate transitions between structures, and thermodynamic integration (described in *Appendix B*) is subsequently used to obtain the free-energy differences.

APPENDIX *B*
Thermodynamic Integration

The Crystal-AFED technique is first used to explore the configuration space to find a transition path that connects two crystal structures. We can then reconstruct the free-energy difference by calculating the free-energy gradient (*i.e.* mean forces) for each **h**-matrix visited along the simulation that connects the two different states (Maragliano & Vanden-Eijnden, 2008; Yu *et al.*, 2014)

$$f_{\mu\nu}(\mathbf{h}) = -\nabla_{\mathbf{h}} G(\mathbf{h}, T) = \det(\mathbf{h}) \sum_{\gamma} \langle h^{-1}_{\mu\nu} \left(P^{(\mathrm{int})}_{\gamma\nu} - P\delta_{\gamma\nu} \right) \rangle_{\mathbf{h}}, \quad (2)$$

where $P^{(\mathrm{int})}_{\gamma\nu}$ are the elements of the internal pressure introduced in *Appendix A*. By integrating the mean forces along the path we calculate the free-energy difference between the initial structure and alternate structures during the Crystal-AFED simulation using

$$\Delta G = -\sum_{\mu\nu} \int_{\mathrm{path}} f_{\mu\nu}(\mathbf{h}) dh_{\mu\nu}. \quad (3)$$

Acknowledgements

This work was supported, in part, by the US Army Research Laboratory and the US Army Research Office under contract/grant number W911NF-13-1-0387 (MET and LV) and, in part, by the Materials Research Science and Engineering Center (MRSEC) program of the National Science Foundation under Award Number DMR-1420073 (MET and ES).

References

Abrams, J. B. & Tuckerman, M. E. (2008). *J. Phys. Chem. B*, **112**, 15742–15757.

Akella, J. & Kennedy, G. C. (1971). *J. Chem. Phys.* **55**, 793–796.

Bardwell, D. A., Adjiman, C. S., Arnautova, Y. A., Bartashevich, E., Boerrigter, S. X. M., Braun, D. E., Cruz-Cabeza, A. J., Day, G. M., Della Valle, R. G., Desiraju, G. R., van Eijck, B. P., Facelli, J. C., Ferraro, M. B., Grillo, D., Habgood, M., Hofmann, D. W. M., Hofmann, F., Jose, K. V. J., Karamertzanis, P. G., Kazantsev, A. V., Kendrick, J., Kuleshova, L. N., Leusen, F. J. J., Maleev, A. V., Misquitta, A. J., Mohamed, S., Needs, R. J., Neumann, M. A., Nikylov, D., Orendt, A. M., Pal, R., Pantelides, C. C., Pickard, C. J., Price, L. S., Price, S. L., Scheraga, H. A., van de Streek, J., Thakur, T. S., Tiwari, S., Venuti, E. & Zhitkov, I. K. (2011). *Acta Cryst.* B**67**, 535–551.

Bayly, C. I., Cieplak, P., Cornell, W. & Kollman, P. A. (1993). *J. Phys. Chem.* **97**, 10269–10280.

Bernstein, J. (2007). *Polymorphism in Molecular Crystals*, Vol. 14. Oxford University Press.

Block, S., Weir, C. E. & Piermarini, G. J. (1970). *Science*, **169**, 586–587.

Bridgman, P. W. (1938). *Proc. Am. Acad. Arts Sci.* **72**, 227–268.

Capelli, S. C., Albinati, A., Mason, S. A. & Willis, B. T. M. (2006). *J. Phys. Chem. A*, **110**, 11695–11703.

Chisholm, J. A. & Motherwell, S. (2005). *J. Appl. Cryst.* **38**, 228–231.

Day, G. M., Cooper, T. G., Cruz-Cabeza, A. J., Hejczyk, K. E., Ammon, H. L., Boerrigter, S. X. M., Tan, J. S., Della Valle, R. G., Venuti, E., Jose, J., Gadre, S. R., Desiraju, G. R., Thakur, T. S., van Eijck, B. P., Facelli, J. C., Bazterra, V. E., Ferraro, M. B., Hofmann, D. W. M., Neumann, M. A., Leusen, F. J. J., Kendrick, J., Price, S. L., Misquitta, A. J., Karamertzanis, P. G., Welch, G. W. A., Scheraga, H. A., Arnautova, Y. A., Schmidt, M. U., van de Streek, J., Wolf, A. K. & Schweizer, B. (2009). *Acta Cryst.* B**65**, 107–125.

Day, G. M., Motherwell, W. D. S., Ammon, H. L., Boerrigter, S. X. M., Della Valle, R. G., Venuti, E., Dzyabchenko, A., Dunitz, J. D., Schweizer, B., van Eijck, B. P., Erk, P., Facelli, J. C., Bazterra, V. E., Ferraro, M. B., Hofmann, D. W. M., Leusen, F. J. J., Liang, C., Pantelides, C. C., Karamertzanis, P. G., Price, S. L., Lewis, T. C., Nowell, H., Torrisi, A., Scheraga, H. A., Arnautova, Y. A., Schmidt, M. U. & Verwer, P. (2005). *Acta Cryst.* B**61**, 511–527.

Dunitz, J. D. & Scheraga, H. A. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 14309–14311.

Dzyabchenko, A. (1984). *J. Struct. Chem.* **25**, 416–420.

Eijck, B. P. van & Kroon, J. (1999). *J. Comput. Chem.* **20**, 799–812.

Eijck, B. P. van, Spek, A. L., Mooij, W. T. M. & Kroon, J. (1998). *Acta Cryst.* B**54**, 291–299.

Fedorov, I., Marsusi, F., Fedorova, T. & Zhuravlev, Y. (2015). *J. Phys. Chem. Solids*, **83**, 24–31.

Frisch, M., *et al.* (2009). *GAUSSIAN*09. Gaussian Inc., Wallingford, CT, USA.

Jorgensen, W. L., Maxwell, D. S. & Tirado-Rives, J. (1996). *J. Am. Chem. Soc.* **118**, 11225–11236.

Katrusiak, A., Podsiadło, M. & Budzianowski, A. (2010). *Cryst. Growth Des.* **10**, 3461–3465.

Likhacheva, A. Y., Rashchenko, S. V. & Litasov, K. D. (2014). *J. Appl. Cryst.* **47**, 984–991.

Liu, Y. & Tuckerman, M. E. (2000). *J. Chem. Phys.* **112**, 1685–1700.

Lommerse, J. P. M., Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Mooij, W. T. M., Price, S. L., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams, D. E. (2000). *Acta Cryst.* B**56**, 697–714.

Macrae, C. F., Bruno, I. J., Chisholm, J. A., Edgington, P. R., McCabe, P., Pidcock, E., Rodriguez-Monge, L., Taylor, R., van de Streek, J. & Wood, P. A. (2008). *J. Appl. Cryst.* **41**, 466–470.

Maragliano, L. & Vanden-Eijnden, E. (2008). *J. Chem. Phys.* **128**, 184110.

Martoňák, R., Laio, A. & Parrinello, M. (2003). *Phys. Rev. Lett.* **90**, 075503.

Martyna, G. J., Klein, M. L. & Tuckerman, M. (1992). *J. Chem. Phys.* **97**, 2635–2643.

Martyna, G. J., Tobias, D. J. & Klein, M. L. (1994). *J. Chem. Phys.* **101**, 4177–4189.

Martyna, G. J., Tuckerman, M. E., Tobias, D. J. & Klein, M. L. (1996). *Mol. Phys.* **87**, 1117–1157.

Motherwell, W. D. S., Ammon, H. L., Dunitz, J. D., Dzyabchenko, A., Erk, P., Gavezzotti, A., Hofmann, D. W. M., Leusen, F. J. J., Lommerse, J. P. M., Mooij, W. T. M., Price, S. L., Scheraga, H., Schweizer, B., Schmidt, M. U., van Eijck, B. P., Verwer, P. & Williams, D. E. (2002). *Acta Cryst.* B**58**, 647–661.

Nayak, S. K., Sathishkumar, R. & Row, T. N. G. (2010). *CrystEngComm*, **12**, 3112–3118.

Oddershede, J. & Larsen, S. (2004). *J. Phys. Chem. A*, **108**, 1057–1063.

Podeszwa, R., Rice, B. M. & Szalewicz, K. (2008). *Phys. Rev. Lett.* **101**, 115503.

Price, S. L. (2013). *Acta Cryst.* B**69**, 313–328.

Raiteri, P., Martoňák, R. & Parrinello, M. (2005). *Angew. Chem. Int. Ed.* **44**, 3769–3773.

Reilly, A. M., Cooper, R. I., Adjiman, C. S., Bhattacharya, S., Boese, A. D., Brandenburg, J. G., Bygrave, P. J., Bylsma, R., Campbell, J. E., Car, R., Case, D. H., Chadha, R., Cole, J. C., Cosburn, K., Cuppen, H. M., Curtis, F., Day, G. M., DiStasio Jr, R. A., Dzyabchenko, A., van Eijck, B. P., Elking, D. M., van den Ende, J. A., Facelli, J. C., Ferraro, M. B., Fusti-Molnar, L., Gatsiou, C.-A., Gee, T. S., de Gelder, R., Ghiringhelli, L. M., Goto, H., Grimme, S., Guo, R., Hofmann, D. W. M., Hoja, J., Hylton, R. K., Iuzzolino, L., Jankiewicz, W., de Jong, D. T., Kendrick, J., de Klerk, N. J. J., Ko, H.-Y., Kuleshova, L. N., Li, X., Lohani, S., Leusen, F. J. J., Lund, A. M., Lv, J., Ma, Y., Marom, N., Masunov, A. E., McCabe, P., McMahon, D. P., Meekes, H., Metz, M. P., Misquitta, A. J., Mohamed, S., Monserrat, B., Needs, R. J., Neumann, M. A., Nyman, J., Obata, S., Oberhofer, H., Oganov, A. R., Orendt, A. M., Pagola, G. I., Pantelides, C. C., Pickard, C. J., Podeszwa, R. l., Price, L. S., Price, S. L., Pulido, A., Read, M. G., Reuter, K., Schneider, E., Schober, C., Shields, G. P., Singh, P., Sugden, I. J., Szalewicz, K., Taylor, C. R., Tkatchenko, A., Tuckerman, M. E., Vacarro, F., Vasileiadis, M., Vázquez-Mayagoitia, Á., Vogt, L., Wang, Y., Watson, R. E., de Wijs, G. A., Yang, J., Zhu, Q. & Groom, C. R. (2016). *Acta Cryst.* B**72**, 439–459.

Rérat, B. & Rérat, C. (1990). *J. Chim. Phys.* **87**, 2003–2015.

Rosso, L., Miháry, P., Zhu, Z. & Tuckerman, M. E. (2002). *J. Chem. Phys.* **116**, 4389–4402.

Samanta, A., Tuckerman, M. E. & Yu, T. Q. (2014). *Science*, **346**, 729–732.
Spek, A. L. (2015). *Acta Cryst.* C**71**, 9–18.
Thiéry, M. M. & Léger, J. M. (1988). *J. Chem. Phys.* **89**, 4255–4271.
Tuckerman, M. E., Alejandre, J., López-Rendón, R., Jochim, A. L. & Martyna, G. J. (2006). *J. Phys. A Math. Gen.* **39**, 5629–5651.
Tuckerman, M., Berne, B. J. & Martyna, G. J. (1992). *J. Chem. Phys.* **97**, 1990–2001.
Tuckerman, M. E., Yarne, D., Samuelson, S. O., Hughes, A. L. & Martyna, G. J. (2000). *Comput. Phys. Commun.* **128**, 333–376.
Wang, J., Wang, W., Kollman, P. A. & Case, D. A. (2001). *J. Am. Chem. Soc.* **222**, U403.
Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. (2004). *J. Comput. Chem.* **25**, 1157–1174.
Williams, D. E. (2001). *J. Comput. Chem.* **22**, 1154–1166.
Yang, J., Hu, W., Usvyat, D., Matthews, D., Schütz, M. & Chan, G. K.-L. (2014). *Science*, **345**, 640–643.
Yu, T.-Q., Alejandre, J., López-Rendón, R., Martyna, G. J. & Tuckerman, M. E. (2010). *Chem. Phys.* **370**, 294–305.
Yu, T.-Q., Chen, P.-Y., Chen, M., Samanta, A., Vanden-Eijnden, E. & Tuckerman, M. (2014). *J. Chem. Phys.* **140**, 214109.
Yu, T. Q. & Tuckerman, M. E. (2011*a*). *Eur. Phys. J.* **200**, 183–209.
Yu, T.-Q. & Tuckerman, M. E. (2011*b*). *Phys. Rev. Lett.* **107**, 015701.