AIP | The Journal of Chemical Physics

Free energy calculations for molecular solids using GROMACS

J. L. Aragones, E. G. Noya, C. Valeriani, and C. Vega

Citation: J. Chem. Phys. **139**, 034104 (2013); doi: 10.1063/1.4812362
View online: http://dx.doi.org/10.1063/1.4812362
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v139/i3
Published by the AIP Publishing LLC.

THE JOURNAL OF CHEMICAL PHYSICS **139**, 034104 (2013)

Free energy calculations for molecular solids using GROMACS

J. L. Aragones,[1,2,a)] E. G. Noya,[3,b)] C. Valeriani,[1,c)] and C. Vega[1,d)]
[1]*Departamento de Química Física, Facultad de Ciencias Químicas, Universidad Complutense de Madrid, 28040 Madrid, Spain*
[2]*Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[3]*Instituto de Química Física Rocasolano, CSIC, Serrano 119, E-28006 Madrid, Spain*

(Received 22 March 2013; accepted 5 June 2013; published online 16 July 2013)

In this work, we describe a procedure to evaluate the free energy of molecular solids with the GROMACS molecular dynamics package. The free energy is calculated using the Einstein molecule method that can be regarded as a small modification of the Einstein crystal method. Here, the position and orientation of the molecules is fixed by using an Einstein field that binds with harmonic springs at least three non-collinear atoms (or points of the molecule) to their reference positions. The validity of the Einstein field is tested by performing free-energy calculations of methanol, water (ice), and patchy colloids molecular solids. The free energies calculated with GROMACS show a very good agreement with those obtained using Monte Carlo and with previously published results. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4812362]

I. INTRODUCTION

The existence of a fluid-solid transition for hard spheres is now well established but it has been a matter of debate for long time after Alder and Wainwright[1] found evidences of its existence when performing Molecular Dynamic (MD) simulations. The debate ended when Hoover and Ree[2] in 1968 calculated the free energy of the fluid and solid phases showing that at a certain pressure the chemical potential of the two phases was the same (i.e., they found the coexistence conditions). To evaluate the free energy of the solid phase, Hoover and Ree[2] proposed a new method, the cell method, inspired in the successful cell theory of solids.[3] Since then the interest in determining the free energy of solids and/or the fluid-solid coexistence by using computer simulations has kept growing. In 1997, Bruce, Wilding, and Ackland[4–7] proposed a new method to determine the fluid-solid equilibria, the phase switch method. Later on, Grochola[8,9] proposed a method to compute the free energy difference between a fluid and a solid phase. The methodology was further developed by Eike, Brennecke, and Maginn.[10] More recently, other methods have been proposed to compute the free energy of a solid, as the self-referential methodology of Sweatman,[11] the density of states partitioning method by Do and Wheatley,[12] the modified constrained cell model by Orkoulas and Nayhouse,[13] and the new method recently proposed by Wilding and Sollich,[14] just to mention a few. In the last years, direct coexistence simulations techniques,[15–18] that require no free energy calculations, have become very popular to determine the fluid-solid equilibrium. However, direct coexistence techniques cannot be used to determine solid-solid equilibria and may become computationally expensive when the crystal growth becomes very slow.[19] Moreover, they do not provide direct information about the relative stability between the different polymorphs. It is becoming more and more evident that the relative stability between the different solid phases (i.e., the phase diagram) provides very interesting information on the orientational dependence of the interactions between molecules, and that phase diagram predictions can be very useful to test and further develop potential models (being the TIP4P/2005 water model a good example[20,21]). As anticipated by Bernal and Fowler,[22] Whalley,[23] Finney,[24] Monson,[25] and Morse and Rice[26] among others, there is a lot of information about intermolecular interactions in the low temperature region of the phase diagram (where one or more solid phases are thermodynamically stable). For this reason, free energy calculations of solids are likely to be needed still for some time.

One of the most popular methods to evaluate free energy of solids is the "Einstein crystal method" proposed by Frenkel and Ladd,[27–29] which consists of a thermodynamic integration scheme to compute the free energy of a solid using as a reference system an ideal Einstein crystal, i.e., a solid in which each particle is bound to its lattice position by a harmonic spring. The method was later extended to molecular solids including an angular-dependent orientational field.[30] Next, Vega and Noya[31,32] developed the "Einstein molecule (Em) method" which can be regarded as a small modification of the Einstein crystal method. While in the Einstein crystal method a quasi-divergence in the integration from the solid to the Einstein crystal was avoided by fixing the center of mass of the system, in the Einstein molecule approach this was achieved by fixing the position of the reference point of one molecule (but not its orientation). This simplifies considerably the derivation of the free energy of the reference system.

Even though several numerical techniques are nowadays available to compute the free-energy of solids, their calculation is still limited due to the need of having accessible *a la*

[a)] aragones@mit.edu
[b)] eva.noya@iqfr.csic.es
[c)] cvaleriani@quim.ucm.es
[d)] cvega@quim.ucm.es

carte Monte Carlo (MC) codes. It is becoming quite obvious that in the last decade the number of groups (specially the number of young groups) performing simulations using open source packages such as GROMACS[33] or LAMMPS[34] is steadily growing. However, free energy calculations for solids are not implemented in these open source packages and for this reason the number of groups determining free energy for solids still forms a small family.[35–56] Nevertheless as it will be shown here, it is possible to compute free energies for solid phases using these open source programs.

Free energy calculations have been recently reported for atomic systems such as Si,[57] alloys of copper and zirconium,[58] and a confined Lennard-Jones (LJ) system[59] in all cases using LAMMPS. Few months ago, some of us[60] showed how to carry out free-energy calculations (Einstein crystal and Einstein molecule) for *atomic solids* within open source molecular dynamics simulation programs, such as GROMACS and LAMMPS. Free energy calculations were performed for a LJ solid and for a NaCl solid, and the results were compared to previously published free energy calculations obtained using our home made Monte Carlo code. The agreement was found to be excellent so that the message was: for atomic systems it works! Besides, in Ref. 60, input files for both GROMACS and LAMMPS were provided as supplementary material (we provide again three files needed for the calculations with LAMMPS amended to remove misprints[61] that we have checked did not affect the results presented in Ref. 60). Therefore, one can now compute with GROMACS and/or LAMMPS the free energy of solids formed by noble gases, ionic solids, metals, and spherically interacting models. In spite of that, the situation is not fully satisfactory. With the methodology described so far it is not yet possible to compute the free energy for molecular solids.

In this work we will show that, by slightly modifying (from the traditional form) the expression of the field forcing the molecules to adopt the position and orientation of the solid lattice, it is possible to compute free energies for molecular solids using open source molecular dynamics packages such as GROMACS. The procedure will be applied to two interesting molecules: methanol and water for which free energy calculations have been previously reported, and a patchy colloidal solid. Quite often small molecules are described by rigid models (methanol,[62] water,[21] carbon dioxide[63]). Therefore, the procedure described in this work describes how to perform free energy calculations for rigid molecular systems.

II. METHODOLOGY

A. Einstein molecule approach

The Helmholtz free-energy $A_{sol}(T, V)$ computed with the *Em* method can be written as

$$A_{sol}(T, V) = A_0(T, V) + \Delta A_1(T, V) + \Delta A_2(T, V), \quad (1)$$

where A_0 is the free energy of the reference system including a correction for fixing the reference point of one molecule (i.e., molecule 1), ΔA_1 is the free-energy difference between the ideal Einstein crystal and an Einstein crystal in which particles interact through the Hamiltonian of the real solid ("interacting" Einstein crystal), and ΔA_2 is the free-energy difference between the interacting Einstein crystal and the real solid (see Ref. 50 for further details).

B. The choice of the Einstein field

In the most common approach, the external field used to calculate the free energy of molecular solids is obtained as the sum of two terms: a translational field (U_t, consisting on harmonic springs that bind the reference point of each molecule to its lattice position) and an angular-dependent orientational field (U_r, that forces the molecules to adopt the same orientation as in the reference crystal)[50]

$$U_{Em}^{t+r}(\mathbf{r_2}, \ldots, \mathbf{r_N}, \Omega_1, \ldots, \Omega_N) = \sum_{j=2}^{N} U_t(\mathbf{r_j}) + \sum_{j=1}^{N} U_r(\Omega_j), \quad (2)$$

where $\mathbf{r}_\mathrm{j}$ is the position of the reference point of molecule j, and Ω_j are the angular coordinates that specify the orientation of molecule j. Notice that since molecule 1 is fixed as a reference point, there is no need of including a translational spring for the reference point of this molecule. However, molecule 1 can still rotate around its reference point and for this reason it is necessary to include an orientational field for molecule 1 (for a more detailed discussion, see Ref. 32).

When choosing an angular-dependent orientational field, one usually chooses a functional form that is invariant under rotations exchanging equivalent atoms in the molecule. For instance, for N_2 it is useful to have an external field that is invariant to the exchange of the two nitrogen atoms, or in the case of water H_2O it is useful that the external field does not change when the two hydrogen atoms are exchanged. When the external field is invariant to the symmetry rotations of the molecule, there is no need of including in the Monte Carlo code moves that exchange equivalent atoms of the molecule. These moves look certainly "strange." Imagine an instantaneous configuration of ice. Now for, say, molecule 27, you exchange the position of the two hydrogen atoms. Obviously, the intermolecular energy of the system is invariant to this exchange. Why to include such a move? If your external orientational field is not invariant to this symmetry operation, then you should include this move (specially when the strength of the external field is small so that the energy penalty of the exchange is small and the move will likely be accepted within the Markov chain) to obtain the correct energy as was first illustrated by Schroer and Monson.[64] Since the inclusion of these "strange" moves when performing free energy calculations of solids is rather unusual, the solution is: "use an external field that is invariant to the symmetry rotations of the molecule and there is no need then of including the strange exchange moves." This is what everybody does when performing free energy calculations for solids. So far so good. This approach of imposing an orientational field that respects the symmetry of the molecule even though rather straightforward, has a disadvantage: it is not implemented in molecular dynamic packages. The situation may change and it may happen that, in the future, developers of these programs will include such kind of external fields. However, as it will be shown here, this is not strictly necessary. With the tools already available in these programs (and in this work we shall focus on GROMACS), it is now possible to compute the free energy of solids without introducing any single line of code.

A possible alternative to avoid using an angular-dependent orientational field is to fix the position and orientation of the molecules by binding each atom of the molecule to its lattice position by a harmonic spring

$$U_{Em}^{springs} = \sum_{j=2}^{N_{spr.}} \Lambda'_E\left(\mathbf{r}_{1,j} - \mathbf{r}^0_{1,j}\right)^2 + \sum_{i=2}^{N}\sum_{j=1}^{N_{spr.}} \Lambda'_E\left(\mathbf{r}_{i,j} - \mathbf{r}^0_{i,j}\right)^2, \tag{3}$$

where N is the number of molecules of the system, $N_{spr.}$ is the number of springs per molecule, $\mathbf{r}_{i,j}$ represents the instantaneous position of atom j of molecule i, and $\mathbf{r}^0_{i,j}$ is the position of the same atom in the reference lattice. Λ'_E is the current value of the coupling parameter. To be consistent with our previous work,[50] the maximum value of the external field will be denoted as Λ_E and we shall denote with the prime the values used in the integration leading to ΔA_2. Note that the position of the reference point of one of the particles (for example, atom 1 of molecule 1) is kept frozen to avoid a quasi-divergence in the evaluation of the term ΔA_2 as proposed in the Einstein molecule method.

How many springs per molecule ($N_{spr.}$) can be used? There are two answers to this question. The first one is that you can use a spring for each atom (or interaction site) of the molecule. That will always be correct. Thus for water, when using, for example, the SPC/E model[65] you should use three springs, one for each hydrogen atom and another one for the oxygen atom. For a united atom model of benzene (with 6 interaction centers), the field will consist of six springs, bounded to each interaction center. For a linear molecule, such as nitrogen you could use two springs, and for a three site interaction model of carbon dioxide[63] you could use three springs. Let us now reformulate the question. What is the minimum number of springs per molecule that could be used? In fact even though using a spring per atom is correct, the number of springs can be reduced by realizing that the position and orientation of a rigid molecule can be fully described by the position of three non-collinear points for a nonlinear molecule and two springs for a linear molecule, regardless of its number of atoms. Thus, you can use as many springs as atoms, or in case you want to use the minimum number of springs you should use three (non-collinear) for a nonlinear molecule and two for a linear molecule.

Now the reader is probably having the following question: is the external field given by Eq. (3) invariant to symmetry rotations of the molecule? The answer, unfortunately, is negative. When two nitrogen atoms are exchanged, the field given by Eq. (3) still links each nitrogen atom to its original position in the lattice and, therefore, the exchange of the two nitrogen atoms increases considerably the energy of the molecule with the external field. Programs such as GROMACS do indeed incorporate the possibility of having harmonic springs linking each atom of the system to the initial lattice position as described in Eq. (3). Therefore, the field given by Eq. (3) is available in GROMACS. The problem is that it is not invariant under symmetry rotations of the molecule. However, a rather simple solution to this problem exists. We shall return to this point later.

C. Evaluation of A_0

Having defined the Einstein field, we first need to evaluate $A_0(T, V)$. This term can be calculated from the partition function. The partition function of the ideal Einstein crystal molecule is given by[32]

$$Q_{Ein-id} = \frac{1}{N!}\frac{1}{\Lambda^{3N}}(q_r q_v q_e)^N \times \int e^{-\beta U_{Em}(\mathbf{r}_1,\ldots,\mathbf{r}_N,\Omega_1,\ldots,\Omega_N)} d\mathbf{r}_1,\ldots,d\mathbf{r}_N, d\Omega_1,\ldots,d\Omega_N, \tag{4}$$

where Λ is the thermal de Broglie wavelength (we shall set the thermal de Broglie wavelength to an arbitrary value since this will not affect coexistence conditions as far as the same choice is done for all phases); $(q_r q_v q_e)$ are the rotational, vibrational, and electronical partition function (set to $q_r q_v q_e = 1$, that is a correct procedure as long as we adopt the same value in all the competing phases). U_{Em} is the Einstein field that depends on the position and orientation of the N molecules of the system $(U_{Em}(\mathbf{r}_1, \Omega_1, \ldots, \mathbf{r}_N, \Omega_N))$. The purpose of this notation is to explicitly show the dependency of the Einstein field with the integration variables. At this point, the derivation is independent of the particular functional form, Eq. (2) or Eq. (3), of the Einstein field.

Equation (4) can be simplified by performing a change of variables so that all the positions are expressed with respect to that of the fixed molecule 1 ($\mathbf{r}_1, \mathbf{r}'_2 = \mathbf{r}_2 - \mathbf{r}_1, \ldots, \mathbf{r}'_N = \mathbf{r}_N - \mathbf{r}_1$), which allows to trivially integrate over $\mathbf{r}_1$

$$Q_{Ein-id} = \frac{1}{N\Lambda^{3N}} V \times \int_{1perm} e^{-\beta U_{Em}(\mathbf{r}'_2,\ldots,\mathbf{r}'_N,\Omega_1,\ldots,\Omega_N)} \times d\mathbf{r}'_2 \ldots d\mathbf{r}'_N d\Omega_1 \ldots d\Omega_N. \tag{5}$$

We have already taken into account that for a system of N indistinguishable particles where the position of molecule 1 is fixed, there are $(N-1)!$ possible permutations,[66] so that the integral in Eq. (5) is evaluated for one particular permutation (see Ref. 32).

Once the partition function has been evaluated, we compute the free energy A_0

$$\frac{\beta A_0}{N} = -\frac{1}{N}\ln(Q_{Ein-id}) = \frac{1}{N}\ln\left(\frac{N\Lambda^3}{V}\right) - \frac{1}{N}\ln\frac{1}{\Lambda^{3(N-1)}}\int e^{-\beta U_{Em}} d\mathbf{r}'_2 \ldots d\mathbf{r}'_N d\Omega_1 \ldots d\Omega_N. \tag{6}$$

Now, if one uses an Einstein field that depends on the position of at least three points of the molecule, such as $U_{Em}^{springs}$ defined in Eq. (3), the translational and orientational degrees of freedom are coupled and, therefore, the integral cannot be separated in two contributions. However, it is possible to separately integrate the contribution of each molecule

$$\begin{aligned}\frac{\beta A_0}{N} &= \frac{1}{N}\ln\left(\frac{N\Lambda^3}{V}\right)\\ &\quad -\frac{1}{N}\ln\int e^{-\beta u_{Em,1}(\Omega_1)}d\Omega_1\\ &\quad -\frac{(N-1)}{N}\ln\frac{1}{\Lambda^3}\int e^{-\beta u_{Em,2}(\mathbf{r}_2,\Omega_2)}d\mathbf{r}_2 d\Omega_2, \quad (7)\end{aligned}$$

where $u_{Em,\,1}$ is the Einstein field for the molecule with the reference atom fixed (i.e., molecule 1) given by

$$\sum_{j=2}^{N_{spr.}}\Lambda_E\left(\mathbf{r}_{1j}-\mathbf{r}_{1j}^0\right)^2 \quad (8)$$

and $u_{Em,\,2}$ is the Einstein field for molecule 2 given by

$$\sum_{j=1}^{N_{spr.}}\Lambda_E\left(\mathbf{r}_{2j}-\mathbf{r}_{2j}^0\right)^2. \quad (9)$$

The $(N-1)$ prefactor in the last term on the right hand side of Eq. (7) arises from the fact that there are $(N-1)$ molecules whose Einstein field is identical to that of molecule 2 so that it is sufficient to evaluate the integral for molecule 2 and to multiply it by $(N-1)$. Notice that the integrals of Eq. (7) must be computed for the maximum value of the external field Λ_E. For the sake of simplicity, we write Eq. (7) as

$$\begin{aligned}\frac{\beta A_0}{N} &= \frac{1}{N}\ln\left(\frac{N\Lambda^3}{V}\right)\\ &\quad -\frac{1}{N}\ln I_1 - \frac{(N-1)}{N}\ln I_2. \quad (10)\end{aligned}$$

Note that for the molecule whose reference point is kept fixed the integral is only performed over the orientational degrees of freedom (I_1), whereas for the remaining molecules the integral is performed over both the position and the orientational coordinates (I_2).

Therefore, when using the Einstein field defined in Eq. (3), one needs to evaluate the two integrals, I_1 and I_2 to obtain A_0. Those integrals can be estimated numerically, for example, using Monte Carlo integration. To compute I_2, we situate the molecule with its reference point located at the origin of coordinates and with a chosen orientation. For the evaluation of I_2, in each Monte Carlo step the following procedure is used: (1) a set of coordinates (r_x, r_y, r_z) and Euler angles $\Omega = (\theta, \phi, \chi)$ are chosen randomly, (2) the rotational matrix associated to these Euler angles is calculated, (3) the molecule is rotated from its reference state by applying the rotational matrix obtained in the previous step, (4) the molecule is translated from the origin to the position (r_x, r_y, r_z), and (5) the integrand of I_2 is evaluated and accumulated to calculate its average value. I$_1$ is obtained using a completely analogous procedure, but in this case since the integration is only performed over the rotational degrees of freedom, no translations are applied in the MC steps. The sampling in the MC integration can be considerably improved by noticing that the integrands $e^{-\beta u_{Em,1}}$ and $e^{-\beta u_{Em,2}}$ are rapidly decreasing functions with a maximum at $\mathbf{r} = 0$ and $\Omega = 0$ for $e^{-\beta u_{Em,2}}$ and $\Omega = 0$ for $e^{-\beta u_{Em,1}}$. Therefore, previous to the evaluation of the integrals, a short run can be performed to estimate the maximum displacement of the reference point of the molecule (r_x, r_y, r_z) and the maximum rotation (θ, ϕ, χ) for which the integrand adopts a value higher than a given threshold, for example, 10^{-16}. Thus, the integrals I_1 and I_2 can be calculated by sampling only the relevant regions of the configurational space. About 10^8–10^9 MC steps are needed to evaluate these integrals with enough accuracy. The evaluation of the integral typically takes about a few hours of CPU time. The value of A_0 as given by Eq. (7) depends on the geometry of the molecule, number of springs, and the value of Λ_E but it does not depend on the polymorph considered. A FORTRAN code to evaluate the term A_0 is provided as supplementary material.[61]

To compute the remaining free-energy terms (Eq. (1)), ΔA_1 and ΔA_2, molecular simulations are needed. In Sec. II D, we show how to implement their calculation in an open source molecular dynamic package such as GROMACS.

D. Evaluation of ΔA_1 and ΔA_2 using GROMACS

First of all, it is important to point out that the free energy calculations for solids are typically performed in the *NVT* ensemble. If one is interested in evaluating the free energy of a solid for a certain temperature and pressure, one should perform first an anisotropic *NpT* simulation, to determine the equilibrium density and shape of the simulation box. By anisotropic *NpT* simulation, we mean that the length of the three sides and the angle between the three vectors that define the simulation box are sampled and allowed to change independently. This is easily achieved by using a Parrinello-Rahman anisotropic barostat.[67] It should be emphasized that it is for the equilibrium simulation box that free energy calculations should be performed. The previous remark is general and it does not depend on whether you are using a MC home made program or GROMACS. Once the equilibrium density and simulation box is known at a certain *T* and *p*, one can start the *NVT* simulations leading to ΔA_1 and ΔA_2. In this work, the free energy calculations will be performed for the same *T* and *p* (i.e., density and shape of the equilibrium simulation box) as those previously reported using MC simulations.

ΔA_1 corresponds to the free-energy difference between the ideal Einstein crystal and the Einstein crystal in which molecules interact through the Hamiltonian of the real solid ("interacting" Einstein crystal). The free-energy difference between these two systems, in terms of the lattice energy, is given by

$$\beta\Delta A_1 = \beta U_{lattice} - \ln\langle\exp\left[-\beta(U_{sol}-U_{lattice})\right]\rangle_{Ein-id}\,, \quad (11)$$

where U_{sol} is the potential energy of the instantaneous configuration (evaluated using the intermolecular potential of interest), $U_{lattice}$ is the potential energy of the perfect lattice (which can be estimated running GROMACS for a perfect lattice using just one MD step and zero as the time step), and the angular brackets denote an ensemble average. This step allows us to choose a convenient value of the harmonic spring constant (Λ_E) that keeps the atoms attached to their lattice positions. As a rule of thumb, a good choice of Λ_E is the one that leads a difference between ΔA_1 and the lattice energy ($U_{lattice}$) of about 0.02–0.04 Nk_BT.

To compute ΔA_1 with GROMACS, we prepare an ideal Einstein crystal with the crystalline structure of the solid of interest, in which each atom is an ideal gas particle attached to its lattice position by a harmonic spring (*[position_restraints]*, see GROMACS manual[68]). Then, we carry out a *NVT* molecular dynamics simulation at the density and temperature of interest storing around 10^4–10^5 configurations of the trajectory to properly compute the ensemble average. This run must be done fixing the position of the reference point of one molecule, usually molecule 1 (*[freezegrps]*[68]) while letting all other particles vibrate through harmonic springs around their lattice positions; which entails letting the center-of-mass move freely (*[comm-mode = none]*). Then, we compute the Hamiltonian of the solid of interest U_{sol} for the stored configurations of the ideal Einstein crystal run, and calculate ΔA_1 from Eq. (11).

ΔA_2, i.e., the free-energy difference between the interacting Einstein crystal and the solid of interest, is obtained by a Hamiltonian thermodynamic integration in which the harmonic springs that tether the atoms to their initial lattice positions are gradually turned off. Thus, ΔA_2 can be computed from the integral of the mean square displacement of each particle from its lattice position

$$\Delta A_2 = -\int_0^{\Lambda_E} \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N_{springs}} \left(\mathbf{r}_{i,j} - \mathbf{r}^0_{i,j}\right)^2 \right\rangle_{N,V,T,\Lambda'_E} d\Lambda'_E. \tag{12}$$

In the previous equation, we have included for simplicity the case $i = 1$ (i.e., molecule 1), $j = 1$ (i.e., the reference atom of molecule 1) even though this atom remains fixed along the simulation at $r^0_{1,1}$ so that it does not contribute to the integral. We perform *NVT* MD simulations at different values of the harmonic spring constant (Λ'_E). We typically use 12–16 values of Λ'_E chosen according to the Gaussian quadrature method, and use the same method to compute the integral in Eq. (12). From the GROMACS output, we obtain the total harmonic energy ($U_{pos-rest}$)

$$U_{pos-rest} = \Lambda'_E \left\langle \sum_{i=1}^{N} \sum_{j=1}^{N_{springs}} \left(\mathbf{r}_{i,j} - \mathbf{r}^0_{i,j}\right)^2 \right\rangle_{N,V,T,\Lambda'_E} \tag{13}$$

from which the mean square displacement (which is the integrand of Eq. (12)) can be easily obtained by dividing by Λ'_E.

Four factors should be carefully taken into account when one performs free-energy calculations with a MD package

1. In order to implement the Einstein molecule method, we should allow the center-of-mass to freely move.[60]
2. When simulating truncated potentials, the discontinuity created by the truncation of the potential generates impulsive forces that cannot be handled within GROMACS. This effect can be significantly reduced by using longer cut-off distances.
3. It is important to choose the MD integration time step carefully since the period of an oscillation of an ideal harmonic spring is a function of the spring constant

$$\frac{1}{\nu} = 2\pi\sqrt{\frac{m}{2\Lambda'_E}}. \tag{14}$$

   Therefore, the integration time step should be chosen to allow to properly sample the vibrations for the highest Λ'_E values. In other words, smaller time steps are required for high values of Λ'_E.
4. A thermostat that treats properly harmonic vibrations should be chosen, as, for instance, the thermostat of Bussi *et al.*[69] On the contrary, one cannot use a Nose-Hoover thermostat[70] as it does not treat properly harmonic vibrations.

Sample input files to evaluate the terms ΔA_1 and ΔA_2 with GROMACS for the examples studied in this work are provided as supplementary material.[61]

E. Molecules with rotational symmetry

When dealing with rotationally asymmetric molecules, the integration to the Einstein field with at least three springs by the evaluation of Eq. (1) directly leads to the free energy of the solid. However, for molecules with rotational symmetry the free energy associated to the symmetry has to be explicitly taken into account.

As already mentioned, the orientational field is usually chosen to have the same symmetry as the molecule, so that the free energy associated with symmetry-preserving rotations is naturally taken into account.[50] Schroer and Monson[64] showed that for a field that does not preserve the symmetry of the molecule, the free energy can be calculated by including special Monte Carlo moves that perform rotations of the molecules consistent with the symmetry operations, so that the configurational space is correctly sampled at low values of the orientational field. Obviously, these moves cannot be incorporated trivially in a MD run.

However, as we will demonstrate in Sec. IV, it is possible to include the free energy associated to the symmetry operations by simply adding to the free energy as calculated with Eq. (1) the analytical value of

$$\frac{\beta A_{sym.}}{N} = -\ln(\Sigma_{rot}), \tag{15}$$

where Σ_{rot} is the number of proper rotations of the molecule. This number can be calculated by adding up the number of proper rotation operations of the corresponding point symmetry group plus the identity symmetry operation (see Table I). It is interesting to point out that Σ_{rot} is identical to the number used to correct the classical partition function of an asymmetric rotor (configurations that can be distinguished in classical statistical mechanics are indistinguishable in quantum statistical mechanics), which is usually denoted as σ in standard Statistical Thermodynamic book (see, for instance,

TABLE I. Point group symmetries for the molecules studied in this work.

		Symmetry operations				
Molecule	Point group symmetry	E	Proper rotations	Improper rotations	Reflection planes	Σ_{rot}
Methanol (OPLS)	C_1	1				1
Water	C_{2v}	1	C_2		σ_v,σ_s	2
Benzene	D_{6h}	1	$2C_6,2C_3,C_2,3C_2',3C_2''$	$2S_3,2S_6$	$\sigma_h,3\sigma_d,3\sigma_v$	12
Patchy model	O_h	1	$8C_3,6C_2,6C_4,3C_2$	$6S_4,8S_6$	$3\sigma_h,6\sigma_d$	24

McQuarrie[71]). Malanoski and Monson[39] used a similar correction when computing the free energy of hard diatomic molecule using an external field similar in form to that of Eq. (3). We should stress that the estimation of the free energy associated to the rotational symmetry using this analytical expression is correct only when molecules are not able to spontaneously flip between equivalent configurations. Otherwise the free energy associated to the rotational symmetry is overcounted. Therefore, it is very important to make sure that molecules do not flip from their initial orientation. In Sec. IV, we will discuss ways to check that no flip has occurred and how to calculate the free energy when flipping occurs.

III. MODELS AND SIMULATION DETAILS

In order to check the validity of the proposed method, we have evaluated the free energy for different systems for which the free energy was previously calculated.

A. Methanol: A rotationally asymmetric molecule

The first system we consider is methanol described using the OPLS model[62] (see Figure 1). Methanol is an asymmetric molecule, which in the OPLS model[62] is described by three centers of interaction. The complete phase diagram of this model was calculated by Gonzalez-Salgado and Vega[72] who found that the α and γ phases of methanol are predicted to be thermodynamically stable, in qualitative agreement with experiments, whereas OPLS was not able to reproduce the stability of the β phase.[72]

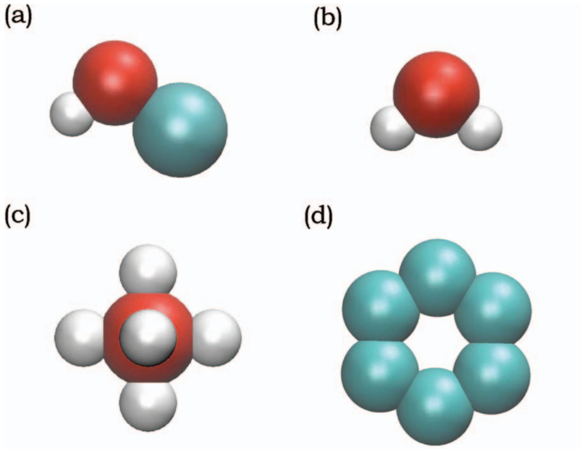

FIG. 1. Molecular models considered in this work: (a) OPLS methanol, (b) TIP4P/2005 water, (c) octahedral model, and (d) benzene toy model.

In what follows, we will calculate the free energy of the α solid phase. In this case, the Einstein field used is simply defined by binding each center of interaction to its lattice position with harmonic springs. Since the molecule is asymmetric, Eq. (1) leads to the correct free energy without the need of including any extra term.

B. Ice II and ice Ih: A rotationally symmetric molecule without and with proton disorder

The second system to test the procedure described is ice, simulated using TIP4P/2005[73] (see Figure 1). Different from methanol, water has C_{2v} symmetry, so the use of an Einstein field with three springs would not give us the correct free energy. As shown in Table I, the C_{2v} point group has one proper rotation symmetry operation (C_2) plus the identity E, so that in this case the number of proper rotations is $\Sigma_{rot} = 2$. Thus, we need to add the term $\beta A_{sym.}/N = -\ln(2)$ to the calculated free energy

$$\frac{\beta A_{sol}}{N} = \frac{\beta A_0}{N} + \frac{\beta \Delta A_1}{N} + \frac{\beta \Delta A_2}{N} - \ln(2). \tag{16}$$

For ice II, which is proton ordered (i.e., all water molecules are oriented in the same way in all the unit cells of the solid), Eq. (16) is sufficient to evaluate the free energy of the solid.

We have also computed the free energy for ice I_h. Ice I_h has an additional configurational entropy associated to proton disorder; considering that there are many possible hydrogen bond arrangements consistent with the ice rules.[22] Unless special moves are implemented, in the usual simulations time only one possible configuration of the protons is sampled.[74,75] The free energy associated to the proton disorder in ice I_h was estimated by Pauling[76] ($\beta A_{Pauling}/N = -\ln(3/2)$), and needs to be added to the total free energy of the solid. Therefore, for ice I_h, the total free energy is calculated using

$$\frac{\beta A_{sol}}{N} = \frac{\beta A_0}{N} + \frac{\beta \Delta A_1}{N} + \frac{\beta \Delta A_2}{N} - \ln(2) - \ln(3/2). \tag{17}$$

C. An example of complex molecules: Octahedral patchy model

The applicability of the proposed Einstein field (Eq. (3)) to molecules with more than three interaction sites was further tested by performing free energy calculations for a simple colloidal anisotropic model with six patches arranged in octahedral symmetry (see Figure 1). The interaction potential between two particles is given by a LJ like interaction multiplied by an orientational term. This type of potential

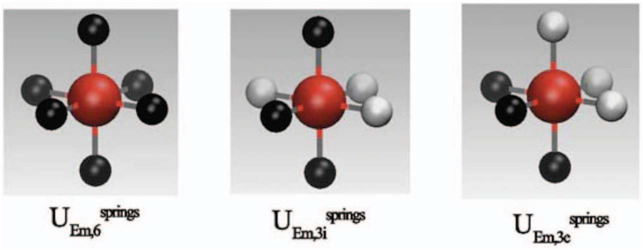

FIG. 2. Schematic representation of the three different Einstein fields used for the evaluation of the free energy of the octahedral patchy model. Black balls represent patches that are bound to their lattice positions by harmonic springs whereas patches shown in white are free to move. In the field on the left panel, all patches are bound to their lattice positions, whereas in the two other fields only three patches are bound to their lattice positions, located either in an isosceles triangle (middle panel) or in an equilateral triangle geometry (right panel).

was first proposed by Doye and co-workers.[77] The complete phase diagram of this model has already been computed using the Einstein crystal method.[78] This model exhibits several thermodynamic stable phases, namely, a simple cubic, a body centered cubic, and a face centered cubic crystal which is orientationally ordered at low temperatures but becomes a plastic crystal at high temperatures. This model is not implemented in GROMACS but we have chosen it because we had already initial configurations for the solid structures and the free energy was already calculated.

Moreover, once the validity of GROMACS was proven with the methanol and water examples, our purpose was to check how to define the Einstein field for more complex molecules, an issue which concerns the methodology rather than the particular computational code used. In this system, we have used a bespoken Monte Carlo code implementing the Einstein field given in Eq. (3) with several possibilities of the number and location of harmonic springs in the molecule (see Fig. 2): (1) six springs, one on each patch, $U_{Em,6}^{springs}$, (2) three springs located on three patches forming an isosceles triangle $U_{Em,3i}^{springs}$, and (3) three springs located on three patches forming an equilateral triangle, $U_{Em,3e}^{springs}$. The point group symmetry of the octahedral patchy model is O_h. The number of proper rotations in this case is $\Sigma_{rot} = 24$, which, as mentioned before, is calculated by adding the proper rotations symmetry operations plus identity, namely, 1(E) +8(C_3)+6(C_2)+6(C_4)+3(C_2) = 24 (see Table I). Therefore,

$$\frac{\beta A_{sol}}{N} = \frac{\beta A_0}{N} + \frac{\beta \Delta A_1}{N} + \frac{\beta \Delta A_2}{N} - \ln(24). \tag{18}$$

D. Simulation details

For methanol, ices Ih and II, the free energy was evaluated with the GROMACS code (version 4.5);[79] whereas for the patchy colloids, using a home-made code. We simulated a 300 methanol molecules system, a 432 water molecules system, and a 216 patchy colloids system. Ice I_h and II free energies were calculated at 1 bar and 200 K, whereas the α-methanol free energy was calculated at 1 bar and 150 K.

The term ΔA_2 was computed by performing *NVT* MD simulations for about 5 ns, and the integration time step was of 0.001 ps (i.e., 5×10^6 time steps). On the other hand, for the term ΔA_1 *NVT* MD simulations of an ideal gas were performed for about 1 ns, using a time step of 0.001 ps (1×10^6 time steps). We stored configurations every 100 MD steps, which correspond to approximately 10^5 independent configurations per simulation run. The temperature was kept constant using the velocity rescaled thermostat proposed by Bussi *et al.*[69] with a relaxation time of 2 ps. Constraints were used to fix the geometry of the molecules, both methanol and water, by using the algorithm LINCS,[80] which allows to use the domain decomposition of GROMACS. The LJ interactions were truncated at 8.5 Å (for water) and 10.0 Å (for methanol), and standard long range corrections were employed. Ewald sums were used to deal with Coulombic interactions; the real part was truncated at the same distance as the LJ interactions, and the reciprocal contribution was evaluated by using PME.[81]

For ices Ih and II, we also computed each term of the free energy using our Monte Carlo code, to check term by term the evaluation of the free energy with GROMACS. We carried out *NVT* simulations.[82] Coulombic interactions were calculated using the Ewald summation technique.[29] The LJ part of the potential and the real space contribution of the Coulombic interactions were truncated at the same distances as in the MD simulations for each system. Both dispersive and screened-Coulombic interactions up to the cut-off distance were calculated, and to accelerate the calculations we used a linked cell list.[83] Standard long range corrections to the Lennard-Jones part of the potential were added. The term ΔA_1 was calculated in a simulation consisting of 3×10^5 MC cycles for equilibration and 5×10^5 MC cycles for averages. The integrand of ΔA_2 was evaluated by performing a *NVT* MC simulation consisting of 3×10^5 MC cycles for equilibration and 7×10^5 MC cycles for taking averages at each value of the coupling parameter ($\Lambda_{E'}$).

The maximum value of the spring constant (Λ_E) was $\Lambda_E = 5000 k_B T/\text{Å}^2$ for ices I_h and II, and $\Lambda_E = 6666.67 k_B T/\text{Å}^2$ for methanol. We typically use 16 $\Lambda_{E'}$ values to compute the integral in Eq. (12) using the Gaussian quadrature method. However, we have also computed this integral with up to 60 $\Lambda_{E'}$ values for ice I_h.

For the patchy model, the free energy was calculated using our own Monte Carlo code. The simulation box contained $N = 216$ molecules arranged in a simple cubic solid. The energy was truncated and shifted at a cut-off distance of 2.5 σ_{LJ}. Typically 20 *NVT* simulations were used to evaluate the term ΔA_2, each simulation consisting of about 50 000 MC cycles for equilibration plus 200 000 MC cycles for taking averages (one Monte Carlo cycle is defined as N attempts to displace or rotate one molecule). Rather long simulations were used to evaluate the term ΔA_1 (about 500 000 MC cycles plus 50 000 for equilibration). The maximum value of the coupling parameter Λ_E was chosen for each considered Einstein field so that ΔA_1 was not more than $0.03 N k_B T$ units different from the lattice energy of the system. Calculations of the free energy were performed at two different temperatures $T^* = k_B T/\epsilon_{LJ} = 0.1$ and $T^* = 0.2$. The free energy was also evaluated using an angular-dependent orientational field that preserves the symmetry of the molecule (we used the Einstein field described in Ref. 78).

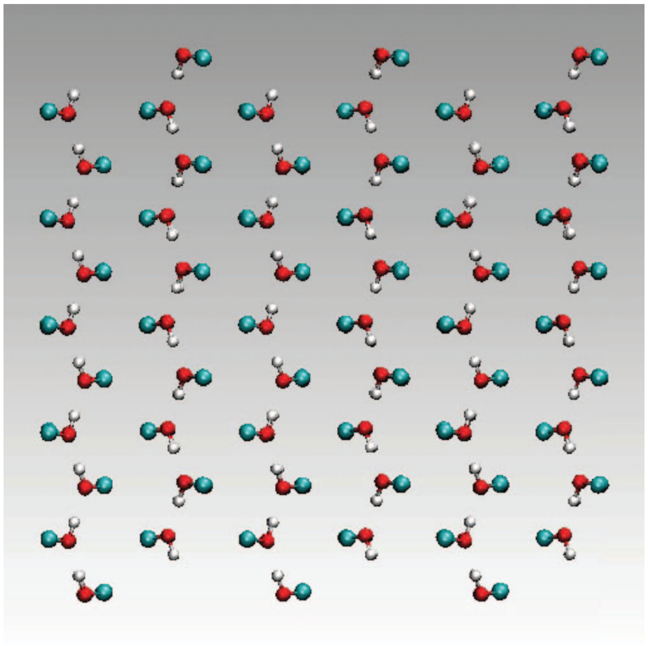

FIG. 3. Top view of an α-phase of methanol.

In this work to determine A_0 from Eq. (7), we have set the thermal wavelength Λ to 1 Å when computing the free energy of methanol and of ices Ih and II. For the octahedral patchy model, we have set Λ to σ_{LJ} which is the distance at which the LJ potential is zero.

IV. RESULTS

A. Methanol

To test the proposed methodology, we compute the free energy of the α-phase of methanol using GROMACS. The methanol molecule described by the OPLS force field represents the simplest example to test the Einstein field proposed here; it is an asymmetric molecule described by three interaction centers.[62] Therefore, the Einstein field is constructed by binding each center of interaction to its lattice position with a harmonic spring. The results are presented in Table II.

TABLE II. Free-energies for the methanol α-phase as obtained from Einstein Molecule (EM) method using MC[72] or GROMACS (G) at 1 bar and 150 K. Free energies are in Nk_BT units and number densities in Å^{-3}. The maximum value of the spring constant was $\Lambda_E = 6666.67 k_BT/\text{Å}^2$. The energy in bold is the free energy computed previously in Ref. 72.

		ρ(N/V)	A_0	ΔA_1	ΔA_2	A_{sol}
G	Methanol	0.01878	29.05	−41.27(1)	−17.33(3)	−29.55(4)
Ref. 72	Methanol	0.01878	…	…	…	**−29.55**

The free energy computed for the α-phase (see Fig. 3) is in agreement, within the statistical error bar, with previous results obtained for the Einstein molecule approach using an Einstein field with separate translational and orientational terms.[72] In this case, and since the methanol is an asymmetric molecule, there are no issues with the rotational symmetry: the free energy obtained from the integration with this Einstein field directly provides the free energy of the methanol solid.

B. Ice II and ice Ih

Next, we compute the free energy of two solid phases of water, one proton ordered phase (ice II), and another in which the protons are disordered (ice Ih) (Fig. 4).

The Einstein field for the water molecule is defined in a similar fashion as for the methanol molecule; tethering the three atoms of the molecule to their reference positions with harmonic springs, fix both the positions and orientations of the water molecules. We have used either GROMACS (G) or

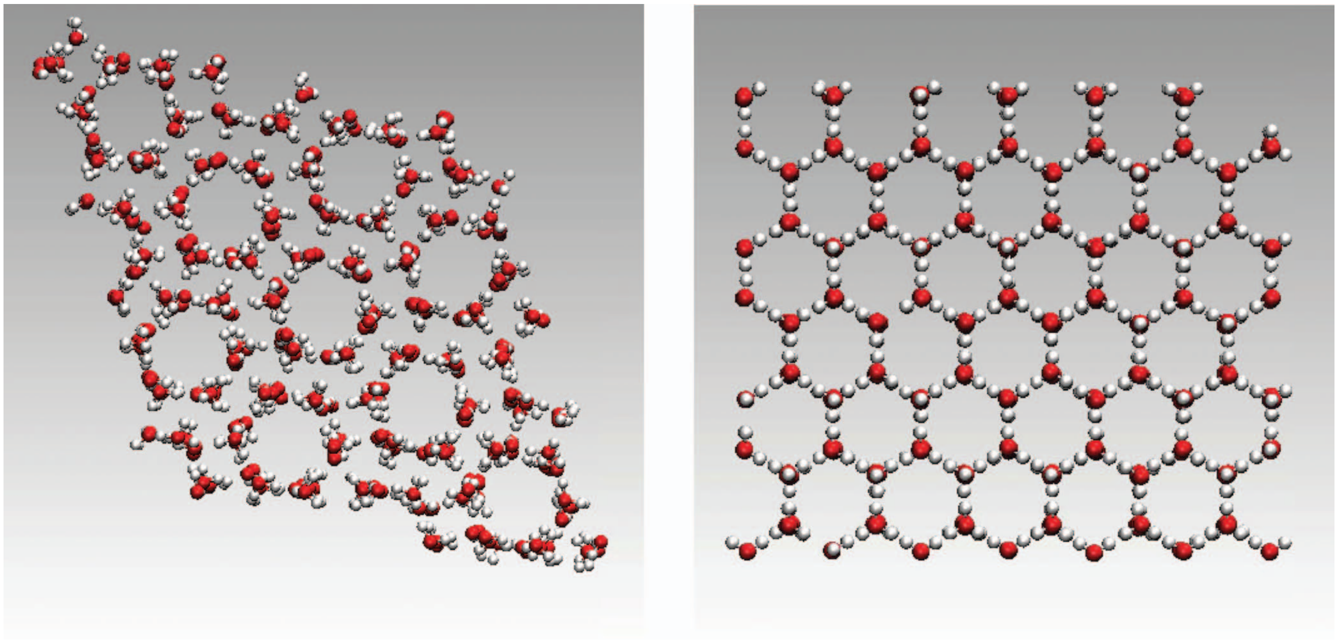

FIG. 4. Top view of an ice II (left-hand side) and ice Ih (right-hand side).

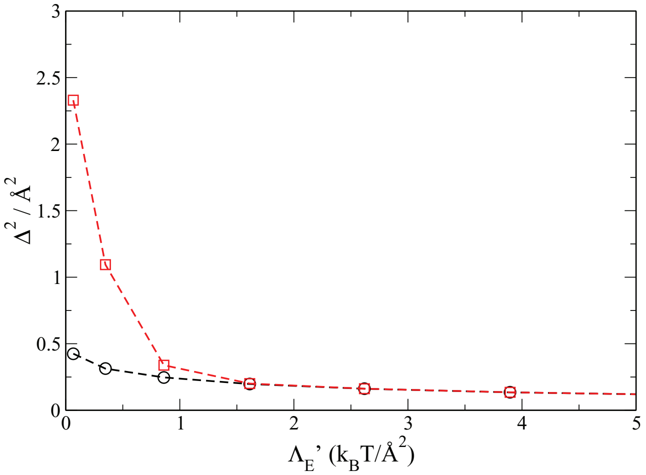

FIG. 5. Mean square displacement ($\Delta^2 = \sum_{j=1}^{N_{spr}} \langle (r_{i,j} - r_{i,j}^0)^2 \rangle$) as a function of the spring constant strength Λ'_E for ice Ih for the normal MC (black circles) and for the MC with rotationally symmetric moves (red squares).

our *home made* MC code to compute all free energy terms (Eq. (1)) of these solid phases.

Let us start by presenting the results for ice Ih at 200 K and 1bar. For this state, we have reported in previous work[84] a free energy of -26.25 in Nk_BT units. This is the benchmark value to be reproduced. Before implementing the calculations with GROMACS, we determined the free energy of the solid using the external field given by Eq. (3) using our home made Monte Carlo program. After adding ($A_0 + \Delta A_1 + \Delta A_2 + A^{Pauling}$), we obtained -25.56 Nk_BT (see second line of Table III). Thus, using the external field of Eq. (3) the free energy does not agree with our previous published value. This clearly points out that there is a problem with the form of the external field given by Eq. (3) when applied to water. The origin of the discrepancy comes from the free energy associated to the rotational symmetry of the water molecule. Different from the methanol, water has rotational symmetry (2-fold rotational symmetry along the axis of the molecule). How to solve this problem? There are two possible routes.

In the first route in order to explicitly take into account the free energy associated to the rotational symmetry of the molecule, we carried out a free energy calculation in which we sampled the symmetric rotations of the water molecules by including special MC movements exchanging the positions of the two H atoms, in order to correctly sample the configurational space at low values of the orientational field (similar to Schroer and Monson's[64] free-energy calculations for a model of benzene). Within the MC approach, we use 60 Λ'_E. In Fig. 5, the integrand of Eq. (12) is plotted in the two cases: for the normal MC and for the MC with rotationally symmetric moves.

As shown in Fig. 5, both curves coincide except for values smaller than 1 k_BT/Å^2. If the spring strength is small enough, the mean square displacement sharply grows when the symmetric rotations are sampled. The ΔA_2 values computed including or not flipping moves are $-15.75(2)$ and $-15.05(2)$ Nk_BT, respectively. The difference between them is about -0.7 Nk_BT, which corresponds to $-ln(\Sigma_{rot})$, where Σ_{rot} is the number of proper rotational operations of the molecule ($-ln(2)$ for the water molecule). The free energy obtained now (when including flipping moves at low strengths of the external field) is $-26.27Nk_BT$ (see third row of Table III) which is fully coincident with that obtained in our previous work. Thus, when using an external field of the form of Eq. (3) it is possible to recover the correct free energy by including flipping moves at low strengths of the external field. However, this is not the only possibility. In fact, one can summarize the situation by stating that the inclusion of flipping moves simply amounts to incorporating a contribution that can be determined analytically. Indeed, the inclusion of flipping moves provokes that ΔA_2 decreases by $-ln(2)$, and one could simply perform the calculations without including flipping moves and adding a $-ln(2)$ term at the end of the calculations. This is shown in the fourth row of Table III where we added a $-ln(2)$ term to the results obtained in the second row. The free energy obtained in this way is in complete agreement with the one previously reported. Although flipping moves can be easily implemented in a Monte Carlo code, this is not the case for molecular dynamics. For this reason, the suggested route is to compute the free energy without flipping and adding a $-ln(2)$ term. Using GROMACS, we have implemented the free energy using the external field of Eq. (3) (without including any type of flipping moves). At the end of the calculations, we simply added a $-ln(2)$ term. As it can be seen, we obtained term by term the same results that were obtained with our Monte Carlo program for the same external field, and that after adding $-ln(2)$ we recovered the free energy reported in our previous work. The free energy of a solid is unique (it

TABLE III. Free-energies of solids as obtained from Einstein Molecule (EM) method using MC or GROMACS (G). Ice I_h and II free energies were calculated at 1 bar and 200 K. Free energies are in Nk_BT units and densities in Å^{-3}. The maximum value of the spring constant was $\Lambda_E = 5000k_BT$/Å^2 for both ice phases. MC* = calculated including flipping movements.

		ρ(N/V)	A_0	ΔA_1	ΔA_2	$A_{sym.}$	$A^{Pauling}$	A_{sol}
Ref. 84	Ih	0.03103	…	…	…	…	…	**−26.252**
MC	Ih	0.03103	27.59	−37.70(1)	− 15.05(1)		−0.405	−25.56(2) (?)
MC*	Ih	0.03103	27.59	−37.70(1)	− 15.75(1)	-	−0.405	−26.27(2)
MC	Ih	0.03103	27.59	−37.70(1)	− 15.05(1)	−0.69	−0.405	−26.26(2)
G	Ih	0.03103	27.59	−37.70(1)	− 15.05(3)	−0.69	−0.405	−26.26(4)
G	II	0.03929	27.594	−37.08(1)	− 15.359(3)	−0.69	…	−25.54(4)
Ref. 84	II	0.03929	…	…	…	…	…	**−25.563**

does not depend on the procedure used to compute it) so that its value obtained from different types of external field should agree. This is indeed the case.

Therefore, the previous discussion proves that it is possible to include the free energy associated with the symmetry operations of the molecule by simply adding its analytical value (Eq. (15)) to the free energy calculated with Eq. (1). This will allow us to perform free-energy calculations of molecular solids with GROMACS, where we do not need to include movements to sample properly the configurational space. However, a word of caution is necessary. The analytical estimation of the free energy associated to the rotational symmetry of the molecule is correct *only* when no spontaneous flipping occurs in the simulation. Otherwise, we will be overcounting the rotational symmetry contribution to the free energy. In our simulations, we have checked that no molecule has flipped during the simulations, comparing the initial and final configurations.

Also with the aim to test the methodology for another solid phase of water (ice II), we have calculated the free energy of ice II, whose unit cell is rhombohedral. As we mentioned above, ice II is a proton ordered phase, therefore, the Pauling entropy has not been taken into account. The free energy calculated for the rhombohedral ice II is presented in Table III, and is also in agreement with the value previously published.

The configurations of the crystalline structures used in this work together with the GROMACS input files (*topol.top* and *grompp.mdp*) and a FORTRAN code to evaluate A_0 are provided as supplementary material.[61]

It has been shown in a recent work, that for certain potential models[20] ice II may be more stable than ice I_h at room pressure and at temperatures close to the melting point. With the tools provided in this section, it is possible now to compute the relative stability between ices I_h and II (at a certain T and p) for any potential model that can be simulated with GROMACS.

C. Benzene toy model

The two molecular models presented so far have in common that they both have only three interaction centers, which only leaves us one possible choice of where to locate the three springs of the Einstein field. However, we would also like to explore how to apply the proposed Einstein field to more complex molecules with more than three atoms or interaction sites. As mentioned before (Sec. II), one possible option would be to locate one spring on each atom of the molecule. However, using only three springs bound to three non-collinear points of the molecules must also be a possible route because the coordinates of three points fully specify the position and orientation of a rigid molecule.

In order to investigate theoretically how the proposed Einstein field is applied to those cases, we have designed a benzene *toy model*, consisting of six interaction sites located at the vertex of a hexagon. Since our aim is to investigate the behavior of the integrand of Eq. (12) for low values of ΔA_2, we shall perform the calculations for only one molecule. Moreover, the only possible Monte Carlo moves will be rotations exchanging the positions of the interactions sites (i.e., small vibrations around the lattice points are not allowed). The number of rotational operations of this molecule (whose point group symmetry is D_{6h}) is $\Sigma_{rot} = 12$ (which is obtained as the sum of $1(\mathrm{E})+2(\mathrm{C}_6)+2(\mathrm{C}_3)+1(\mathrm{C}_2)+3(\mathrm{C}_{2'})+3(\mathrm{C}_{2''}) = 12$, see Table I).

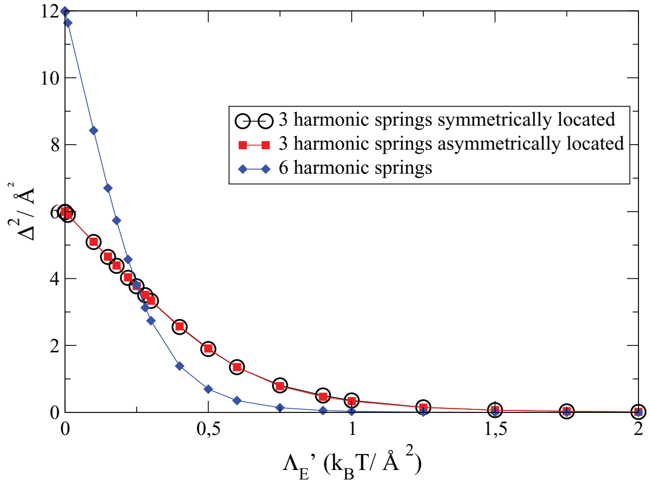

FIG. 6. $\Delta^2 = \sum_{j=1}^{N_{springs}} \langle (\mathbf{r}_{i,j} - \mathbf{r}^0_{i,j})^2 \rangle_{N,V,T,\Lambda'_E}$ versus Λ'_E for three different Einstein fields: in blue with 6 harmonic springs (one on each atom), in black 3 harmonic springs symmetrically located, and in red 3 harmonic springs asymmetrically located. Calculations are performed for just one molecule and the integral of the area always gives ln (12).

For this model, we considered different choices of the Einstein field, differing on the number of springs and on their location: 6 harmonic springs (one on each atom), 3 harmonic springs symmetrically located, and 3 harmonic springs asymmetrically located. Next, we run a Monte Carlo simulation sampling all 12 configurations. Special movements that perform the molecule symmetry operations were applied to evaluate the integrand at low values of the coupling parameter Λ'_E as it has been observed that virtually no molecular flips occurred from relatively low values of Λ'_E.

We then evaluated the integral ΔA_2 (Fig. 6) and obtained ln (12) in all three cases. Thus, one should always add $-ln(\Sigma_{rot})$ regardless of the number of springs used to fix the location of a certain molecule, and regardless of the particular choice of the location of the springs. The conclusion is that the correction $-ln(\Sigma_{rot})$ is related to the symmetry of the molecule and not to the symmetry of the external field.

D. Patchy model

The calculated free energies for the simple cubic solid formed by the octahedral patchy model (see Figure 7) are given in Table IV. Let us first look at the results of the free energy at $T^* = 0.1$.

It can be seen that the results obtained with the three different reference Einstein fields, namely, those with six springs ($U^{springs}_{Em,6}$), three springs in an equilateral triangle ($U^{springs}_{Em,3e}$), and three springs in an isosceles triangle ($U^{springs}_{Em,3i}$) (see Figure 2), all give the same results within statistical error, and the results are also in excellent agreement with those

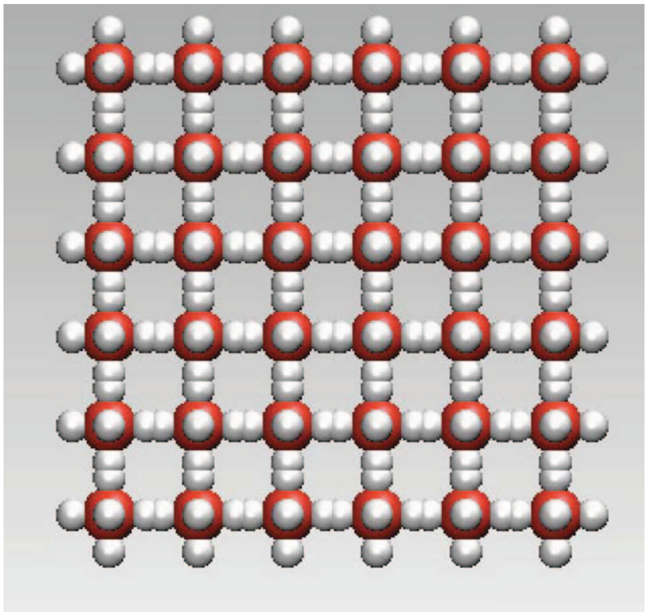

FIG. 7. Simple cubic structure for the octahedral patchy model.

obtained with the angular-dependent orientational field (U_{EM}^{t+r}). Note that, consistently with the results obtained for the benzene toy model, the correction that needs to be applied to Eq. (1) is $\beta A_{sym.}/N = -\ln(24)$ regardless of the position or the number of springs included in the field (Eq. (3)). This definitely proves that one can freely choose the number and distribution of harmonic springs used to define the Einstein field and that all choices lead to the correct free energy, as long as a minimum of three non-collinear springs are included for nonlinear molecules.

However, when the free energy was calculated at the same density but at a higher temperature, $T^* = 0.2$, we observed that the three different Einstein fields lead to slightly different values of the free energy, and neither of them was able to reproduce the value obtained with the angular-dependent orientational field that preserves the symmetry of the molecule. The origin of the discrepancy was clear when the results were analyzed in more detail. By looking at the final configurations of the simulations at the lowest values of Λ'_E, it became evident that some of the molecules had rotated according to the symmetry operations of the molecule (see Figure 8). As mentioned before, this leads to an overcounting of the free energy associated with the molecule symmetry operations. Because we were using our own Monte Carlo code, it was fairly easy for us to modify it to avoid the flipping by searching for the closest interaction sites to the spring lattice positions. Once the molecular flipping was removed, the free energy was computed by using Eq. (1) with the analytical symmetry correction. By doing this the results of the free energy are in complete agreement with those obtained with an orientational field that preserves the symmetry of the molecule (see Table IV).

TABLE IV. Free energy of the simple cubic solid for the octahedral patchy model. The simulation box contained 216 molecules and the energy was truncated and shifted at a distance $r^*_{cutoff} = 2.5$. All these calculations were performed at a density $\rho^* = 0.763$. Three different implementations of the Einstein field were tested, namely, six springs one on each patch ($U_{Em,6}^{springs}$), three springs located in three of the patches, one in each axis ($U_{Em,3e}^{springs}$), and finally three springs, two on two patches on the same axis and the other one in any of the remaining patches ($U_{Em,3i}^{springs}$). The coupling parameter Λ_E is given in k_BT/σ_{LJ}^2 and free energies in Nk_BT units. The free energy in bold is the free energy computed using a translational and an orientational field invariant to symmetry rotations of the molecule.

Eins. field	T^*	Λ_E	A_0	ΔA_1	ΔA_2	$A_{sym.}$	A_{sol}
$U_{Em,6}^{springs}$	0.1	30000	33.26	−29.31	−14.31	−ln(24)	−13.55
$U_{Em,3e}^{springs}$	0.1	60000	33.75	−29.31	−14.84	−ln(24)	−13.58
$U_{Em,3i}^{springs}$	0.1	60000	34.06	−29.31	−15.13	−ln(24)	−13.55
U_{Em}^{t+r}	0.1	20000	28.96	−29.27	−13.25	…	**−13.55**
$U_{Em,3e}^{springs}$	0.2	30000	31.67	−14.64	−15.11	−ln(24)	−1.25
U_{Em}^{t+r}	0.2	20000	27.75	−14.62	−14.35	…	**−1.22**

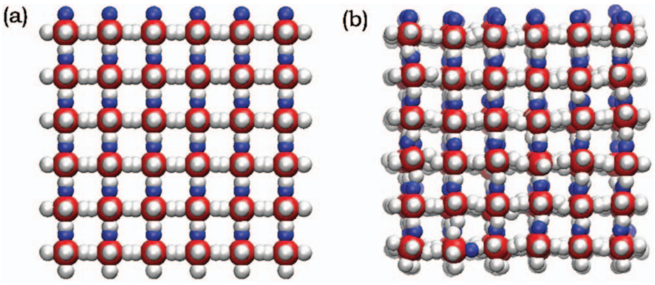

FIG. 8. (a) Initial and (b) final configuration at the lowest value of Λ'_E ($\Lambda'_E = 3.077 k_BT/\sigma_{LJ}^2$) used for the evaluation of the term ΔA_2 using a three springs Einstein field $U_{Em,3e}^{springs}$ for the tetrahedral patchy model at $T^* = 0.2$. One of the patches is colored in blue, so that it is possible to track any rotation of the particles. In the final configuration, it is evident that one of the molecules in the bottom row has rotated away from its initial orientation.

E. Avoiding molecular flipping

As we have just seen, it is possible that some molecular flipping occurs at the low values of the Einstein field, and if that happens, it invalidates the procedure to calculate the free energy described in this work. Therefore, it is necessary to discuss ways of how to avoid flipping using a molecular dynamics package such as GROMACS. As mentioned before, one quick way to know whether flipping has occurred is by looking at the final configurations and checking that neither of the molecules has flipped from the initial state. Although this method would most likely be a good indicator of flipping, the rigorous way of studying the flipping would be to calculate the angular distribution functions of the molecules with respect to their initial states. For the octahedral patchy model, we calculated the probability distribution function of the cosine of the angle ($\cos\theta$) formed between the instantaneous and the initial orientation of the axis that contains one of the patches, for example, patch 1. We observed that whereas for large values of the coupling parameter Λ'_E this function exhibits one single peak at angles close to $\cos\theta = 1$, it adopts a double peaked structure at low values of Λ'_E (see Figure 9). This can be done by writing a simple analysis code that reads in the output generated by GROMACS and calculates the angular histograms, which would have multiple peaks in

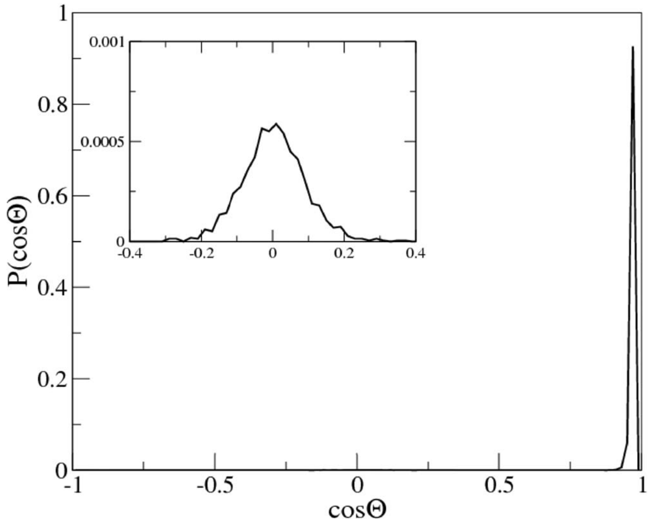

FIG. 9. Angular distribution function at the lowest value of Λ'_E ($\Lambda'_E = 3.077 k_B T/\sigma_{LJ}^2$) used for the evaluation of the term ΔA_2 with a three springs Einstein field $U_{Em,3e}^{springs}$ for the tetrahedral patchy model at $T^* = 0.2$. The distribution is double peaked, with a strong peak at angles close to $\cos\theta = 1$ that represents the molecules that are rotating about their orientation in the initial configuration, and a much smaller peak around $cos\,\theta = 0$ (an enlarged view of this smaller peak is shown in the inset) that evidences that some of the molecules have rotated according to the rotational operations of the molecule.

case of flipping and will be singled-peaked if no flipping has occurred.

If we observe flipping in the system, the most obvious advice would be to decrease the temperature of the system. At lower temperatures, the probability of having particles that are able to overcome the energy barrier for the molecular flipping would decrease considerably. We have seen that for the patchy model this is indeed a possible route. The free energy was not calculated correctly at $T^* = 0.2$ but it was possible to evaluate it at $T^* = 0.1$. In principle, this route must be valid for any studied solid. Then one could perform a thermodynamic integration to compute the free-energy of the solid at the desired temperature.

Another possible route would be to try to evaluate the integrand of ΔA_2 at low values of the coupling parameter by manually removing the contribution of those molecules that have undergone flipping. This can be done with the GROMACS package by saving the trajectory of the simulations and writing a code that calculates the mean square displacement (i.e., the integrand of ΔA_2) but considering only those molecules that have not undergone flipping. This is a valid route as long as most of the molecules remain in the initial orientation. We checked that for the patchy model this method leads to the same value of ΔA_2 as removing the flipping on the fly in the simulations (by searching for the closest spring site in the instantaneous configuration with respect to the reference spring positions).

V. SUMMARY AND CONCLUSIONS

In this work, we describe how the free energy of molecular solids can be calculated via the Einstein molecule method using MD packages, such as GROMACS. Given that the usual Einstein fields that include an angular-dependent term would involve making changes in the MD package, we propose to use an alternative field in which at least three non-collinear points of each molecule are bound to their lattice positions by harmonic springs, which can be implemented easily in the available MD packages. Note that by specifying the position of three non-collinear atoms, both the position and orientation of a rigid molecule are fully defined. Even though for molecules with rotational symmetry this Einstein field is not invariant under the symmetry operations of the molecule, we showed that in most of the cases the correct free energy is obtained by simply adding an analytical correction term at the end of the calculations. This route is valid as far as the molecules in the solid phase are not able to flip within the typical simulation length. A flipping operation corresponds to exchanging the position of two equivalent atoms. For instance, in the case of water, it would mean that two hydrogen atoms exchange their positions. In solids, flipping is a rare event as the flipping configurations are separated by rather large free energy barriers. Even when flipping occurs, there is a simple way of solving the problem: reducing the temperature so that flipping does no longer exist. After that you can always use thermodynamic integration to obtain the free energy at the wished temperature and pressure. Notice that, in computer simulations, molecular solids are usually mechanically stable up to very low temperatures even though they may not be the thermodynamically stable phase.[85] Another possibility when flipping occurs is to eliminate its effect by not including the contribution of the rotated molecules when computing the integrand of ΔA_2 at low values of Λ'_E.

Despite the importance of evaluating free energy of solids and the availability of methods that allow its calculation, there are not yet that many works in which such calculations are performed. We believe that one of the reasons for this relies on the fact that the methods to calculate free energies, for example, the Einstein crystal method, have not yet been implemented in commercial and freely distributed MD packages. We hope that the method described here (that shows how the Einstein molecule method can be easily implemented in such codes) will encourage other groups to calculate free energies of solid in the future. Although in this work we used the program GROMACS, it is likely that the procedure described here can be implemented in other open source MD packages (such as LAMMPS). The only requirements are: a good thermostat able to deal with harmonic vibrations as the one proposed by Bussi *et al.*,[69] the existence of harmonic restraints (i.e., harmonic springs) linking each atom to a certain reference lattice position, and the possibility of freezing the position of an atom of the system (frozen atom) which is needed when implementing the Einstein molecule methodology. These three requirements are in general all available in open source MD programs.

Finally, we would like to point out that it is likely that the procedure described in this work could be extended to flexible models. Free energy calculations for flexible models of n-alkanes[86,87] and water[88] have already being reported by using Monte Carlo simulations. The extension to flexible molecules of the procedure outlined in this work would be certainly an interesting topic for future studies.

ACKNOWLEDGMENTS

This work was funded by the grants FIS2010-16159, FIS2010-15502 (DGI), and MODELICO-P2009/ESP/1691 (CAM). C. Valeriani acknowledges financial support from a Juan de la Cierva Fellowship and from the Marie Curie Integration Grant PCIG-GA-2011-303941 ANISOKINEQ. We would like to thank Lunna Li for discovering a typographical error in two files supplied for LAMMPS in Ref. 60's Supplementary Materials.

[1] B. J. Alder and T. E. Wainwright, J. Chem. Phys. **27**, 1208 (1957).

[2] W. G. Hoover and F. H. Ree, J. Chem. Phys. **49**, 3609 (1968).

[3] J. E. Lennard-Jones and A. F. Devonshire, Proc. R. Soc. London, Ser. A **163**, 53 (1937).

[4] A. D. Bruce, N. B. Wilding, and G. J. Ackland, Phys. Rev. Lett. **79**, 3002 (1997).

[5] A. D. Bruce, A. N. Jackson, G. J. Ackland, and N. B. Wilding, Phys. Rev. E **61**, 906 (2000).

[6] G. C. McNeil-Watson and N. B. Wilding, J. Chem. Phys. **124**, 064504 (2006).

[7] J. R. Errington, J. Chem. Phys. **120**, 3130 (2004).

[8] G. Grochola, J. Chem. Phys. **120**, 2122 (2004).

[9] G. Grochola, J. Chem. Phys. **122**, 046101 (2005).

[10] D. M. Eike, J. F. Brennecke, and E. J. Maginn, J. Chem. Phys. **122**, 014115 (2005).

[11] M. B. Sweatman, Phys. Rev. E **72**, 016711 (2005).

[12] H. Do and R. J. Wheatley, J. Chem. Theory Comput. **9**, 165 (2013).

[13] G. Orkoulas and M. Nayhouse, J. Chem. Phys. **134**, 171104 (2011).

[14] N. B. Wilding and P. Sollich, EPL **101**, 10004 (2013).

[15] A. J. C. Ladd and L. V. Woodcock, Chem. Phys. Lett. **51**, 155 (1977).

[16] A. J. C. Ladd and L. V. Woodcock, Mol. Phys. **36**, 611 (1978).

[17] J. R. Morris and X. Song, J. Chem. Phys. **116**, 9352 (2002).

[18] E. G. Noya, C. Vega, and E. de Miguel, J. Chem. Phys. **128**, 154507 (2008).

[19] J. L. Aragones, E. Sanz, and C. Vega, J. Chem. Phys. **136**, 244508 (2012).

[20] C. Vega, J. L. F. Abascal, M. M. Conde, and J. L. Aragones, Faraday Discuss. **141**, 251 (2009).

[21] C. Vega and J. Abascal, Phys. Chem. Chem. Phys. **13**, 19663 (2011).

[22] J. D. Bernal and R. H. Fowler, J. Chem. Phys. **1**, 515 (1933).

[23] E. Whalley, J. Chem. Phys. **81**, 4087 (1984).

[24] J. L. Finney, J. Mol. Liq. **90**, 303 (2001).

[25] P. A. Monson, AIChE J. **54**, 1122 (2008).

[26] M. D. Morse and S. A. Rice, J. Chem. Phys. **76**, 650 (1982).

[27] D. Frenkel and A. J. C. Ladd, J. Chem. Phys. **81**, 3188 (1984).

[28] J. M. Polson, E. Trizac, S. Pronk, and D. Frenkel, J. Chem. Phys. **112**, 5339 (2000).

[29] D. Frenkel and B. Smit, *Understanding Molecular Simulation* (Academic Press, London, 2002).

[30] D. Frenkel and B. M. Mulder, Mol. Phys. **55**, 1171 (1985).

[31] C. Vega and E. G. Noya, J. Chem. Phys. **127**, 154113 (2007).

[32] E. G. Noya, M. M. Conde, and C. Vega, J. Chem. Phys. **129**, 104704 (2008).

[33] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. C. Berendsen, J. Comput. Chem. **26**, 1701 (2005).

[34] S. Plimpton, J. Comput. Phys. **117**, 1 (1995).

[35] B. B. Laird and A. D. J. Haymet, Mol. Phys. **75**, 71 (1992).

[36] C. Vega, E. P. A. Paras, and P. A. Monson, J. Chem. Phys. **96**, 9060 (1992).

[37] P. Bolhuis and D. Frenkel, J. Chem. Phys. **106**, 666 (1997).

[38] R. Agrawal and D. A. Kofke, Mol. Phys. **85**, 43 (1995).

[39] A. P. Malanoski and P. A. Monson, J. Chem. Phys. **107**, 6899 (1997).

[40] F. S. S. Prestipino and V. P. Giaquinta, Phys. Rev. E **71**, 050102 (2005).

[41] H. Nada and J. P. J. M. van der Eerden, J. Chem. Phys. **118**, 7401 (2003).

[42] G. T. Gao, X. C. Zeng, and H. Tanaka, J. Chem. Phys. **112**, 8534 (2000).

[43] L. M. Ghiringhelli, J. H. Los, E. J. Meijer, A. Fasolino, and D. Frenkel, Phys. Rev. Lett. **94**, 145701 (2005).

[44] M. A. Barroso and A. L. Ferreira, J. Chem. Phys. **116**, 7145 (2002).

[45] J. Anwar, D. Frenkel, and M. G. Noro, J. Chem. Phys. **118**, 728 (2003).

[46] A. P. Hynninen, M. E. Leunissen, A. van Blaaderen, and M. Dijkstra, Phys. Rev. Lett. **96**, 018303 (2006).

[47] M. Marechal and M. Dijkstra, Phys. Rev. E **77**, 061405 (2008).

[48] E. Sanz, C. Vega, J. L. F. Abascal, and L. G. MacDowell, Phys. Rev. Lett. **92**, 255701 (2004).

[49] D. M. Eike and E. J. Maginn, J. Chem. Phys. **124**, 164503 (2006).

[50] C. Vega, E. Sanz, J. L. F. Abascal, and E. G. Noya, J. Phys. Condens. Matter **20**, 153101 (2008).

[51] F. Moucka, M. Lisal, and W. R. Smith, J. Phys. Chem. B **116**, 5468 (2012).

[52] F. Romano, E. Sanz, and F. Sciortino, J. Chem. Phys. **132**, 184501 (2010).

[53] S. Ravipati and S. N. Punnathanam, Ind. Eng. Chem. Res. **51**, 9419 (2012).

[54] T. Schilling and F. Schmid, J. Chem. Phys. **131**, 231102 (2009).

[55] A. Alauddin and R. J. Sadus, J. Chem. Phys. **131**, 174504 (2009).

[56] L. Jensen, K. Thomsen, N. von Solms, S. Wierzchowski, M. R. Walsh, C. A. Koh, E. D. Sloan, D. T. Wu, and A. K. Sum, J. Phys. Chem. B **114**, 5775 (2010).

[57] H. Zhao and N. Aluru, Interact. Multiscale Mech. **1**, 303 (2008).

[58] C. Tang and P. Harrowell, J. Phys. Condens. Matter **24**, 245102 (2012).

[59] L. Wan, C. R. Iacovella, R. Christopher, T. D. Nguyen, H. Docherty, and P. T. Cummings, Phys. Rev. B **86**, 214105 (2012).

[60] J. L. Aragones, C. Valeriani, and C. Vega, J. Chem. Phys. **137**, 146101 (2012).

[61] See supplementary material at http://dx.doi.org/10.1063/1.4812362 for initial configurations of the studied systems, the GROMACS input files to evaluate the terms ΔA_1 and ΔA_2, and a FORTRAN code to evaluate the term A_0. We provide also three files to perform free energy calculations for atomic solids with LAMMPS following the procedure described in Ref. 60. There were three typographical errors in the Supplementary Material provided in Ref. 60 to perform calculations with LAMMPS. The three typographical errors in the LAMMPS input files are the following. When computing ΔA_1: in the *deltaa*1*.sh* script, a parenthesis was missing when expressing the lattice energy in $k_B T$ units; when computing the lattice energy Ulattice ("$lmp\_serial < input - real - EM.dat$" as in the $README\_LAMMPS.txt$), we now suggest to use another input file ("$lmp\_serial < input - real - LATTICE - EC$") that differs from $input - real - EM.dat$ only by the initial configuration read. When computing ΔA_2: in the *deltaa*2*.sh* script, there was a typo when expressing the values of Λ in $k_B T$ units. We have checked that the typos did not affect the results.

[62] W. L. Jorgensen, J. Chem. Phys. **90**, 1276 (1986).

[63] J. J. Potoff and J. I. Siepmann, AIChE J. **47**, 1676 (2001).

[64] J. W. Schroer and P. A. Monson, J. Chem. Phys. **112**, 8950 (2000).

[65] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, J. Phys. Chem. **91**, 6269 (1987).

[66] N. B. Wilding, Comput. Phys. Commun. **146**, 99 (2002).

[67] M. Parrinello and A. Rahman, J. Appl. Phys. **52**, 7182 (1981).

[68] S. Pronk *et al.*, Bioinformatics **29**, 845 (2013).

[69] G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).

[70] W. G. Hoover, Phys. Rev. A **31**, 1695 (1985).

[71] D. A. McQuarrie, *Statistical Mechanics* (Harper & Row, New York, 1976).

[72] D. Gonzalez-Salgado and C. Vega, J. Chem. Phys. **132**, 094505 (2010).

[73] J. L. F. Abascal and C. Vega, J. Chem. Phys. **123**, 234505 (2005).

[74] S. W. Rick and A. D. J. Haymet, J. Chem. Phys. **118**, 9291 (2003).

[75] J. L. Aragones, L. G. MacDowell, and C. Vega, J. Phys. Chem. A **115**, 5745 (2011).

[76] L. Pauling, J. Am. Chem. Soc. **57**, 2680 (1935).

[77] J. P. K. Doye, A. A. Louis, I.-C. Lin, L. R. Allen, E. G. Noya, A. W. Wilber, H. C. Kok, and R. Lyus, Phys. Chem. Chem. Phys. **9**, 2197 (2007).

[78] E. G. Noya, C. Vega, J. P. K. Doye, and A. A. Louis, J. Chem. Phys. **127**, 054501 (2007).

[79] B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, J. Chem. Theory Comput. **4**, 435 (2008).

[80] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, J. Comput. Chem. **18**, 1463 (1997).

[81] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, J. Chem. Phys. **103**, 8577 (1995).

[82] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, 1987).

[83] L. G. MacDowell, "Termodinámica estadística de moléculas flexibles: Teoría y simulación," Ph.D. thesis, Universidad Complutense de Madrid, 2000.

[84] M. M. Conde, C. Vega, G. A. Tribello, and B. Slater, J. Chem. Phys. **131**, 034510 (2009).

[85] J. L. Aragones, E. G. Noya, J. L. F. Abascal, and C. Vega, J. Chem. Phys. **127**, 154518 (2007).

[86] J. M. Polson and D. Frenkel, J. Chem. Phys. **111**, 1501 (1999).

[87] A. P. Malanoski and P. A. Monson, J. Chem. Phys. **110**, 664 (1999).

[88] S. Habershon and D. Manolopoulos, Phys. Chem. Chem. Phys. **13**, 19714 (2011).