PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Phys. Chem. Chem. Phys.*, 2021, **23**, 26843

Homogeneous nucleation of NaCl in supersaturated solutions

C. P. Lamas, [ab] J. R. Espinosa, [c] M. M. Conde, [d] J. Ramírez, [d] P. Montero de Hijes, [a] E. G. Noya, [b] C. Vega[a] and E. Sanz*[a]

Received 12th May 2021,
Accepted 2nd November 2021

DOI: 10.1039/d1cp02093e

rsc.li/pccp

The seeding method is an approximate approach to investigate nucleation that combines molecular dynamics simulations with classical nucleation theory. Recently, this technique has been successfully implemented in a broad range of nucleation studies. However, its accuracy is subject to the arbitrary choice of the order parameter threshold used to distinguish liquid-like from solid-like molecules. We revisit here the crystallization of NaCl from a supersaturated brine solution and show that consistency between seeding and rigorous methods, like Forward Flux Sampling (from previous work) or spontaneous crystallization (from this work), is achieved by following a mislabelling criterion to select such threshold (*i.e.* equaling the fraction of the mislabelled particles in the bulk parent and nucleating phases). This work supports the use of seeding to obtain fast and reasonably accurate nucleation rate estimates and the mislabelling criterion as one giving the relevant cluster size for classical nucleation theory in crystallization studies.

I. Introduction

Crystallization from a pure liquid is usually viewed as a two step process in which homogeneous nucleation precedes growth. The crystal nucleus that determines the end of nucleation and the beginning of growth is the critical one.[1,2] This state corresponds to the top of the free energy barrier between the supercooled liquid and the solid which appears due to the energetic penalty of forming a new interface. In the absence of surfaces or impurities, the free energy barrier is only overcome by thermal fluctuations, which makes this process a rare event. Moreover, the typical size and lifetime of a critical nucleus are in the nanometer and nanosecond scale, respectively. Thus, studying nucleation is like trying to find a needle in a haystack both in a spacial and in a temporal sense.[3]

Computer simulation has become a valuable tool for studying nucleation because it has access to the relevant length scale.[4–6] However, the time disparity between nucleus lifetime and nucleation occurrence makes it necessary to use special rare event methods. Some examples are transition path sampling (TPS),[7,8] forward flux sampling (FFS)[9,10] and metadynamics.[11–13] Although rigorous, these methods are computationally expensive.[14]

Seeding is a less expensive but approximate approach in which the simulation starts with a fluid in which a crystal nucleus has been inserted.[15–17] Seeding exploits the fact that the critical nucleus is a state of unstable equilibrium. Under isobaric–isothermal (*NpT*) conditions, a critical cluster in the top of the free energy barrier has fifty percent chances of either dissolving or growing towards the crystallization of the whole system. Thus, by inserting a crystal seed into a pre-equilibrated metastable liquid and then following the time evolution of the size of the crystal seed, one can estimate the critical size under the conditions of interest. The size of the critical cluster obtained in such simulations, alongside other properties of the bulk liquid and crystal phases, is plugged into the Classical Nucleation Theory (CNT) formalism[18–21] to get relatively inexpensive estimates of nucleation parameters such as the nucleation rate, the height of the nucleation free energy barrier or the interfacial free energy.

However, the main drawback of seeding is the arbitrariness with which the nucleus size is determined. It has been recently shown that different ways to obtain the number of particles belonging to the solid nucleus may lead to discrepancies of several orders of magnitude in the nucleation rate of NaCl from a supersaturated brine,[22] which is a quite discouraging result for current or potential seeding users.[23–40] In this work we revisit this problem and show that when the order parameter threshold to identify particles belonging to the crystal nucleus is tuned with the mislabelling criterion[41] (*i.e.* equal number of

[a] *Departamento de Química Física, Facultad de Ciencias Químicas, Universidad Complutense de Madrid, 28040 Madrid, Spain. E-mail: esa01@ucm.es*
[b] *Instituto de Química Física Rocasolano, Consejo Superior de Investigaciones Científicas, CSIC, Calle Serrano 119, 28006 Madrid, Spain*
[c] *Maxwell Centre, Cavendish Laboratory, Department of Physics, University of Cambridge, Cambridge CB3 0H3, UK*
[d] *Departamento de Ingeniería Química Industrial y Medio Ambiente, Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid, 28006, Madrid, Spain*

mislabelled particles in the liquid and the solid bulk phases), seeding can give nucleation rates consistent with those obtained by methods that do not depend on an order parameter, such as Forward Flux Sampling or spontaneous nucleation.[41] Thus, the nucleation of NaCl from supersaturated brine adds up to the list of systems for which crystal nucleation has been successfully investigated with seeding in conjunction with a mislabelling criterion: hard spheres,[41–43] Lennard-Jones,[41] NaCl from its melt,[44] ice from pure water,[17,45–48] ice from brine[49,50] and oppositely charged colloids.[51] Therefore, although seeding does strongly depend on the order parameter, it seems to work for a wide range of systems when the mislabelling criterion is employed to tune the order parameter.

II. Model and simulation details

We use the JC/SPC/E force field which uses +1 and −1 charges plus a Lennard-Jones potential for the ions and the nonpolarizable SPC/E water model.[52] The solubility of this model has been widely studied, reaching a consensual value of approximately 3.7 mol kg^{-1} at 1 bar and 298.15 K.[53–56] Moreover, there are nucleation rate estimates obtained *via* FFS to compare with.[10]

We shall compute the nucleation rates for solutions of molality 8, 10, and 12 mol kg^{-1} using seeding. Furthermore, we shall compute the nucleation rates for high concentrations *via* brute force runs of metastable NaCl solutions for molalities ranging from 14 to 15 mol kg^{-1}, thereby checking how our seeding results extrapolate to such regime.

Molecular Dynamics simulations are performed with GROMACS package 2016.4.[57] We use a Parrinhello–Rahman barostat[58] to maintain the system at 1 bar, with a relaxation time of 2 ps and a compressibility of 10^{-5} bar^{-1}. To control the temperature and keep it at 298.15 K we use a Nosé–Hoover thermostat[59] with a relaxation time of 2 ps. For the Verlet integration of the equations of motion, we use a time step of 2 fs. To deal with electrostatic interactions we use particle mesh Ewald summations.[60] The cut-off radius for the neighbour list, the dispersive and electrostatic interactions is 9 Å. LINCS algorithm is used to fix the geometry of the water molecules, with an order of 6 and 4 iterations.[61,62] Long range dispersion corrections are applied for energy and pressure.

For the seeding simulations, boxes are cubic and the edge length is approximately $L \approx 8.3$ nm and contain nearly 15 000 water molecules and the corresponding number of ions (between 4000 and 6600 ions depending on the concentration and the size of the inserted seed).

When using seeding, we need to compute the number of ions in the solid cluster (see the following section). To label ions as solid-like or solution-like we use a modified version of the $\bar{q}_4$ parameter proposed by Lechner and Dellago,[63] which we compute with a cut-off radius of 4 Å. Note that water is not included in the evaluation of the order parameter and cations and anions are treated as identical particles. The introduced modification is that we do not divide in eqn (6) of ref. 63 by $\tilde{N}_b$, the number of neighbors plus 1. As can be seen in Fig. 1(b), the usual normalized $\bar{q}_{4,n}$ is not able to distinguish solid-like from solution-like atoms, because the probability distributions of $\bar{q}_{4,n}$ in the solutions overlap completely with that of the solid, a problem that is solved by removing the normalization. The unnormalized $\bar{q}_4$, instead, well separates solid-like from solution-like atoms (Fig. 1(a)). This modification is recommended for problems regarding precipitation of salts from solution.

We use the mislabelling criterion[41] to set the threshold value, $\bar{q}_{4,t}$, above which a particle is considered to belong to the solid phase. Such criterion consists in finding the $\bar{q}_{4,t}$ value for which the fraction of the mislabelled particles in the bulk solid is equal to that of the bulk solution. To find this threshold, we first run a simulation of the bulk solid and gather 1000 configurations of 1000 ions each. With these configurations we work out, for a given $\bar{q}_{4,t}$ value, the fraction of ions that are not labelled as solid (mislabelled particles). This calculation is repeated for the same set of configurations and different values of $\bar{q}_{4,t}$ until we get the black curve shown in Fig. 1(c). This curve monotonously increases because the fraction of ions which are not detected as belonging to the solid increases as we increase $\bar{q}_{4,t}$. We do the same with the solution. We run a bulk solution simulation and gather 1000 independent configurations. Then, we compute the fraction of ions that are detected as solid-like in such configurations for a series of $\bar{q}_{4,t}$ values and we obtain the red, green and blue curves for 8, 10 and 12 mol kg^{-1} solutions, respectively (the ionic concentration is given in molality, or mol of NaCl per kg of water). When computing these curves we did not take into consideration isolated ions with no neighbors. We have checked, however, that, if we had considered these ions as belonging to the solution, the curves presented in the figure would have not changed much, and the crossing point between the solid and the solution mislabelling curves would remain almost unchanged. The solution mislabelling functions decay monotonously because the fraction of the particles labelled as solid-like in the solution decreases when $\bar{q}_{4,t}$ increases. As can be seen in Fig. 1(c), the solid and the three solution curves cross at 3.06 (the concentration barely affects the crossing point). Therefore, for the three concentrations, particles with $\bar{q}_4 > 3.06$ are considered solid-like and particles with $\bar{q}_4 < 3.06$ are considered to belong to the solution. Two solid-like particles are considered neighbors (thus belonging to the same cluster) if they are closer than 4 Å to each other.

Although one might think that by removing normalization we might be artificially considering as solid-like solution ions with many neighbours belonging to disordered aggregates, the plot of $\bar{q}_4$ *versus* the number of neighbours shown in Fig. 1(d) indicates that atoms in the solution with three or more neighbours most often adopt low values, rarely reaching values above the threshold. This graph clearly shows that by choosing not to normalize the order parameter, only particles with at least three neighbours can have larger values than the chosen threshold $\bar{q}_{4,t} = 3.06$, *i.e.*, isolated dimers and trimers are always considered solution-like with this order parameter. Note also that atoms with three or more neighbours can adopt values larger than $\bar{q}_{4,t}$, but, as shown in panel (c), the probability of mislabelled

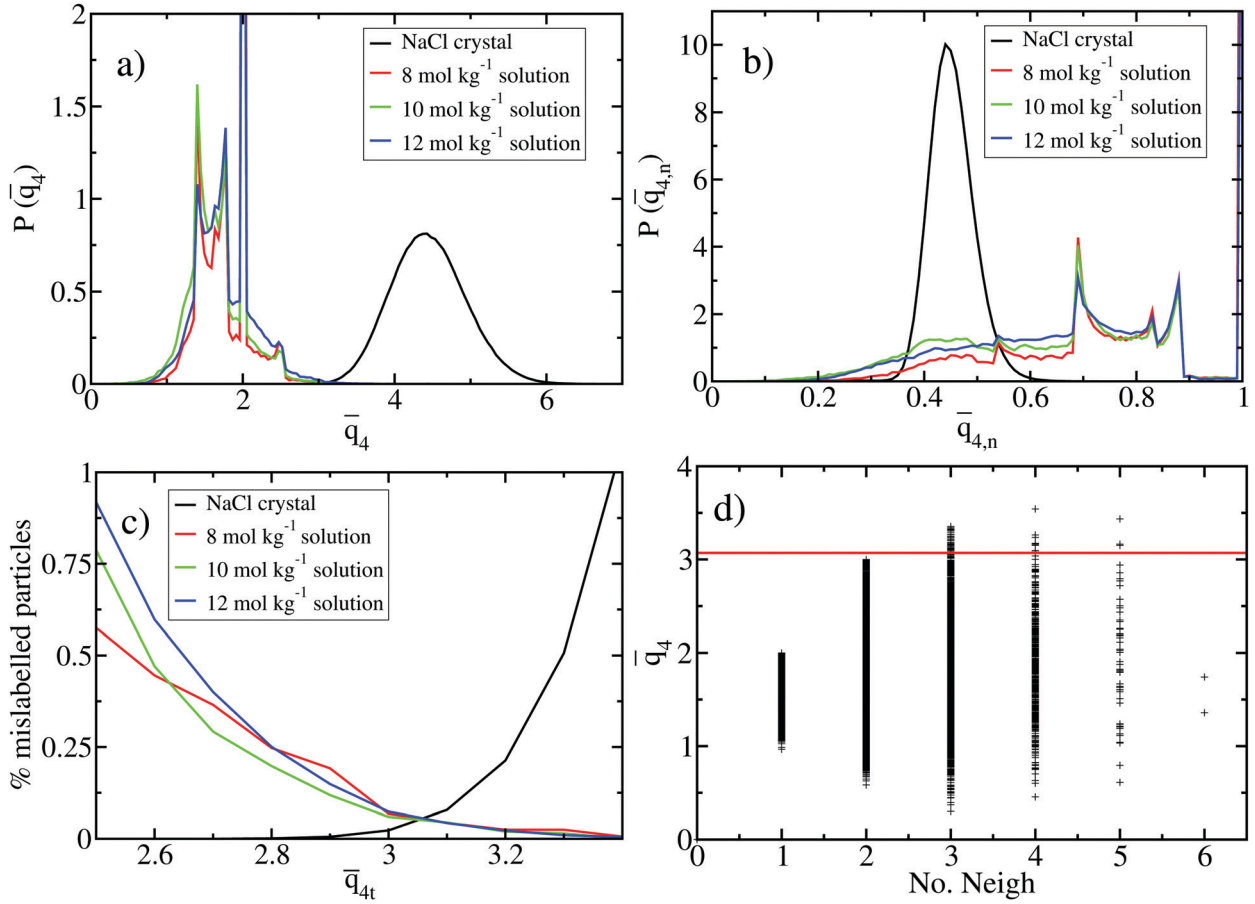

Fig. 1 Probability distributions of the order parameter $\bar{q}_4$ for the NaCl crystals (black curve), and 8 mol kg^{-1} (red curve), 10 mol kg^{-1} (green curve) and 12 mol kg^{-1} (blue curve) NaCl solutions at 1 bar and 298.15 K, both in (a) the unnormalized ($\bar{q}_4$, the one used in this work) and (b) the normalized ($\bar{q}_{4,n}$) versions of the order parameter. Note that the probability distribution functions of the normalized version completely overlap with those of the solid, making this order parameter unsuitable to distinguish atoms within solid environments from atoms within solution environments. This problem is solved if the order parameter is not normalized, as in (a). (c) Percentage of mislabelled particles for the NaCl crystals, and 8 mol kg^{-1}, 10 mol kg^{-1} and 12 mol kg^{-1} NaCl solutions as a function of the chosen $\bar{q}_4$ threshold to discriminate between solid and solution-like ions. (d) Correlation between $\bar{q}_4$ and the number of neighbours measured in the 8 mol kg^{-1} NaCl solution.

particles for the chosen $\bar{q}_{4,t}$ is only about 0.1%. In summary, particles need to have at least three neighbours that also need to be oriented properly to attain larger values of $\bar{q}_4$ than $\bar{q}_{4,t}$, so that the probability that the order parameter classifies disordered aggregates as solid is very low, as seen in panel (c).

III. Results

A. Nucleation rate *via* seeding

Here, we combine molecular dynamics simulations with CNT to get an approximation of the nucleation rate of NaCl crystals in supersaturated aqueous solutions. According to CNT the free energy barrier height is given by:

$$\Delta G = N_c|\Delta\mu|/2 \quad (1)$$

where N_c is the number of ions in the critical cluster and $|\Delta\mu|$ is the chemical potential difference per ion between the solid and the solution at the temperature and pressure under which the study is carried out. We use the $|\Delta\mu|$ values presented by Zimmermann *et al.*[22] (same ones employed by Jiang *et al.*[10]).

To estimate N_c, we first build systems which have a NaCl crystal seed by inserting a spherical cluster cut from the equilibrium NaCl crystal phase in the corresponding solution. We investigate systems with the following solution concentrations: 8 mol kg^{-1}, 10 mol kg^{-1}, and 12 mol kg^{-1}. For each concentration we generate systems with different cluster sizes. We then run simulations and monitor the size of the inserted seed (see Fig. 2). As can be seen in Fig. 2, the lower the concentration, the longer it takes for the cluster to either grow or redissolve.

To quantitatively determine the critical cluster size, N_c, we count the number of times the trajectories shown in Fig. 2 experience a cross-over respect to a given cluster size N_{ions}. We refer to such number of crosses as $N_{recrossing}(N_{ions})$. We expect that the trajectories can go back and forth most easily across N_c because the free energy slope is zero at such cluster size. Accordingly, we identify N_c with the position of the

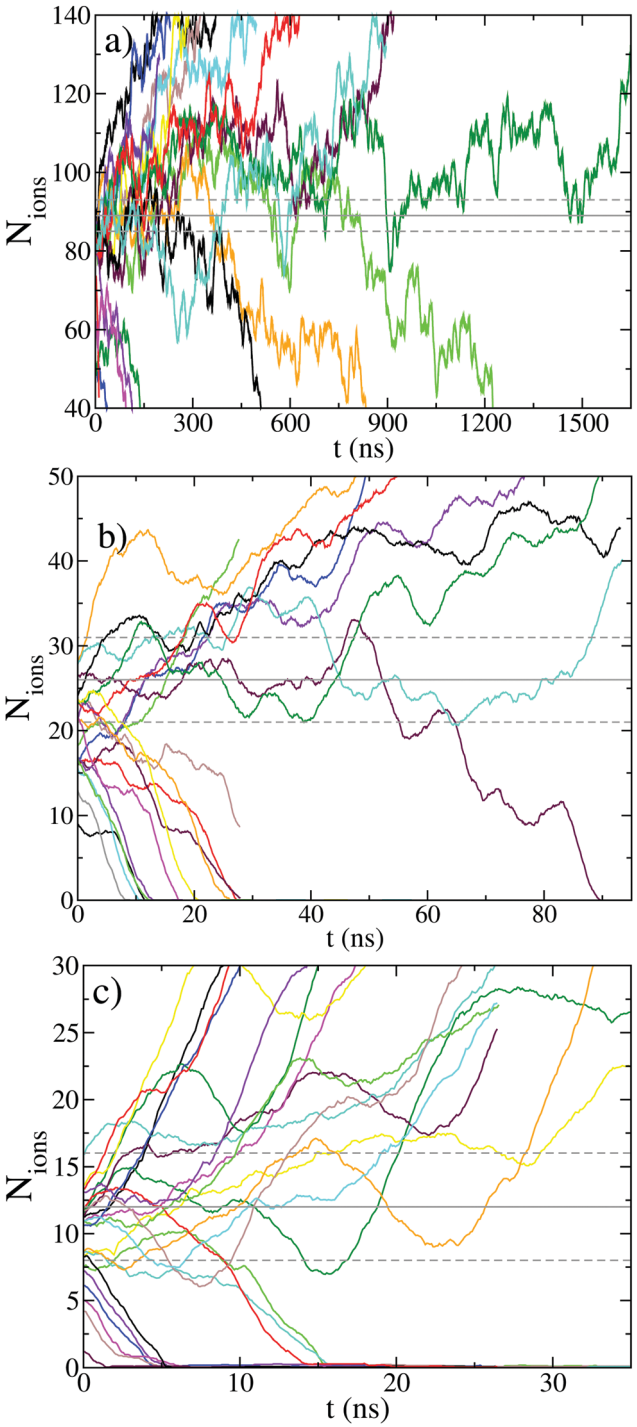

Fig. 2 The number of particles in the NaCl crystal cluster *versus* time for different starting configurations in which a NaCl crystal seed was inserted in NaCl aqueous solutions (at (a) 8 mol kg^{-1}, (b) 10 mol kg^{-1}, and (c) 12 mol kg^{-1} at T = 298.15 K and p = 1 bar).

maximum at $N_{recrossing}(N_{ions})$. To count $N_{recrossing}$ we do not take into account the first time a trajectory crosses a given N_{ions}. Therefore, the trajectories where the cluster monotonously grows or shrinks are not counted in our $N_{recrossing}$ statistics. Fig. 3 shows $N_{recrossing}$ *vs.* N_{ions} for the three studied concentrations. From the maxima of these plots we conclude that the

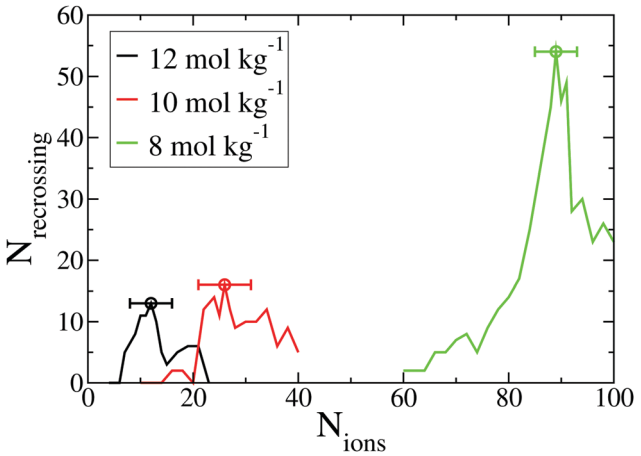

Fig. 3 Number of recrossing *vs.* number of particles in the cluster in NaCl solutions with concentrations of 8 mol kg^{-1}, 10 mol kg^{-1}, and 12 mol kg^{-1}. These graphs have been obtained by analysing the trajectories shown in Fig. 2.

critical clusters contain around 89, 26, and 12 ions for the 8 mol kg^{-1}, 10 mol kg^{-1}, and 12 mol kg^{-1} solutions respectively. It is evident from Fig. 3 that the larger the N_c the higher the number of recrossing. This is consistent with the fact that a wider barrier top is expected for higher free energy barriers (or larger N_c).

The CNT equation that we used to calculate the nucleation rate is

$$J = K\exp(-\Delta G/(k_B T)), \quad (2)$$

where the kinetic constant K is obtained using the following expression:

$$K = \rho_{ions} D Z, \quad (3)$$

where ρ_{ions} is the density of ions in solution and D is the attachment frequency, which we estimate in two different ways. On the one hand, D has been shown to be desolvation limited and to obey the following expression:[6,64]

$$D_{des} = 4\pi R^2 k_s \sigma_s \rho_{ions}, \quad (4)$$

the product of k_s (a second order rate constant) and σ_s (the surface concentration of attachment sites) is only weakly dependent on the nucleus size and is estimated to be 0.066 m s^{-1} by Zimmermann *et al.*[64] This value, alongside eqn (4), can be used to estimate D_{des} reported in Table 1. On the other hand, we use the Auer and Frenkel approach to estimate the attachment frequency as follows:[65]

$$D_{AF} = \langle (N_{ions}(t) - N_{ions}(0))^2 \rangle / (2t), \quad (5)$$

which requires evaluating the slope of the mean squared displacement of the cluster size. In Fig. 4, we show in black such mean squared displacement averaged over three of the trajectories in Fig. 2(a) while the cluster fluctuates around its critical size. The slope of the linear fit shown in red in Fig. 4 divided by two gives an estimate of D, reported as D_{AF} in Table 1. We acknowledge that the slope in Fig. 4 is quite noisy.

Table 1 Seeding simulation data leading to the calculation of the nucleation rate. $\Delta\mu$, given per mol of ions in the table, has been taken from ref. 53, 55 and 56. See the main text for the meaning of each variable

m/(mol kg^{-1})	N_c	ρ_{ions}/nm^{-3}	R/nm	D_{AF}/s^{-1}	D_{des}/s^{-1}	$\Delta\mu$/(kJ mol^{-1})	Z	K/(m^{-3}s^{-1})	$\Delta G/(k_BT)$	$\log_{10}$[J/(m^{-3} s^{-1})]
8	89(4)	8.2074	0.6352	$0.4(3) \times 10^9$	2.7×10^9	4.55	0.0661	2.2×10^{35}	81(4)	0(3)
10	26(5)	9.7791	0.4215	$0.4(3) \times 10^9$	1.4×10^9	6.30	0.1440	5.6×10^{35}	33(6)	21(4)
12	12(4)	11.203	0.3257	$0.9(4) \times 10^9$	0.98×10^9	7.65	0.2336	2.4×10^{36}	18(6)	28(4)

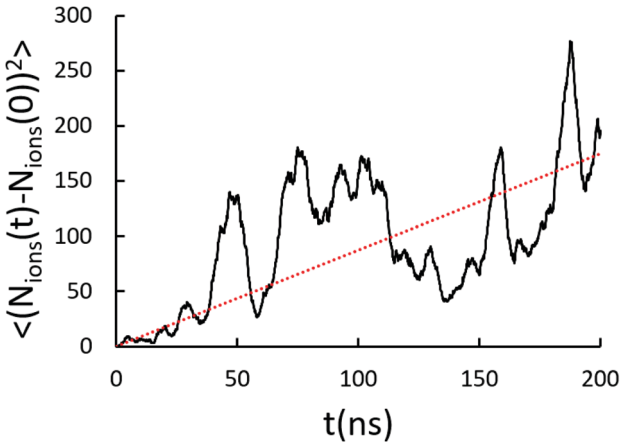

Fig. 4 Mean squared displacement of the cluster size averaged over three trajectories at 8 mol kg^{-1} where the cluster fluctuates around its critical size.

However, for all three studied concentrations we get values of the attachment rate of the order of 10^8–10^9 s^{-1}, which gives us confidence that we are capturing the right order of magnitude in our calculations. Moreover, also for all three concentrations, D_{des} and D_{AF} are within the same order of magnitude, which is quite satisfactory when dealing with nucleation rates. The agreement between theory and simulation gives us further confidence that we are capturing the correct order of magnitude of the attachment rate. We use the D_{AF} values to obtain the nucleation rate given that the Auer and Frenkel approach is a more direct simulation – as opposed to theoretical – estimate of the attachment rate.

R is the radius of the critical nucleus:

$$R = \sqrt[3]{3N_c/(\rho_{solid}4\pi)}, \quad (6)$$

where ρ_{solid} is the density of ions in the NaCl solid at 1 bar and 298.15 K: ρ_{solid} = 41.445 nm^{-3}. The assumption that the critical clusters are spherical has been supported by Espinosa *et al.*, who compared cubic and spherical NaCl clusters in a seeding study.[44] In any case, our results are not much affected by such an assumption given that: (i) eqn (1), that gives ΔG_c, is valid both for spherical and cubic shapes; and (ii) the kinetic prefactor does not even change by an order of magnitude from one shape to another.

Finally, Z is the Zeldovich factor[1,17] and is given by:

$$Z = \sqrt{|\Delta\mu|/(6\pi k_B T N_c)}. \quad (7)$$

In Table 1 the nucleation rates inferred from our seeding simulations, and all variables leading to their calculation, are reported. The rate strongly depends on salt concentration and increases by almost 30 orders of magnitude when going from 8 to 12 mol kg^{-1} concentration. In Section IV we present the comparison of the nucleation rate data obtained by seeding with those obtained by spontaneous crystallization (see next section) and those calculated with FFS in a previous work.[10]

B. Nucleation rate *via* brute force

At high concentrations, where ΔG is low, spontaneous nucleation of NaCl crystals can be seen in brute force molecular dynamics simulations. In Fig. 5 we show the number of particles in the largest NaCl cluster *versus* time for different trajectories started from a 15 mol kg^{-1} solution. We see a cluster growing in all cases after some induction period that varies from run to run due to the stochastic nature of nucleation. The nucleation rate can be easily estimated from these simulations as:

$$J = \frac{1}{\langle t \rangle V}, \quad (8)$$

where $\langle t \rangle$ is the average time for observing spontaneous crystallization within the average simulation volume, V. To obtain these average properties we run 10 independent simulations starting from the same initial configuration and with different initial particle velocities (an example of such trajectories is shown in Fig. 5). The crystallization time for a given trajectory is identified as that at which the number of particles in the cluster irreversibly shoots up (see Fig. 5). This criterion is subject to some arbitrariness which could introduce an error in the average nucleation time. However, this error is small because the time that it takes to detect irreversible growth is much smaller than the nucleation time, as suggested by the abrupt increase in the slope of the curves in Fig. 5 before and after irreversible growth.

In Table 2 we report the system sizes employed in our brute force simulations and the resulting nucleation rates, along with $\langle t \rangle$ and V. We present nucleation rates for concentrations ranging from 14 to 15 mol kg^{-1}. We also tried with 16 mol kg^{-1} (with 14 985 molecules of water and 4320 of NaCl) but we could not compute J since ions quickly aggregate and we observed the growth of a crystal nucleus in spinodal fashion without any induction period[66]. This is in agreement with ref. 66 where a spinodal aggregation regime was found beyond 15 mol kg^{-1}. Therefore, the selected concentrations for the study of spontaneous crystallization are close to the onset of spinodal aggregation but still lie within the metastable equilibrium regime.

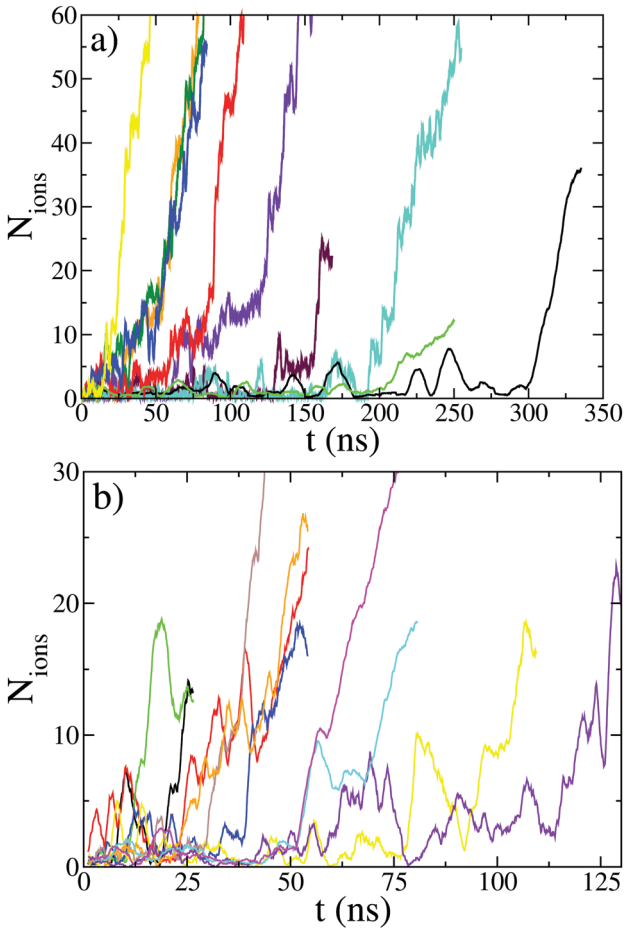

Fig. 5 The number of particles in the largest NaCl crystal cluster along time for 15 mol kg^{-1} systems. Each color represents one of the 10 independent simulations performed. Panels correspond to (a) *NVT* and (b) *NpT* ensembles at T = 298.15 K and p = 1 bar (NVT simulations correspond to the system density obtained at p = 1 bar).

Table 2 Salt nucleation data leading to the calculation of J *via* brute force simulations in both *NpT* (top) and *NVT* (bottom) ensembles

NpT

m/(mol kg^{-1})	N_{NaCl}	N_{water}	$\langle t\rangle$/ns	$\langle V\rangle$/nm^3	p/bar	J/(s^{-1} m^{-3})
15.00	4800	17 760	47(24)	732.43	1	3(1) × 10^{31}
14.71	37 632	142 080	16(8)	5822.18	1	11(4) × 10^{30}
14.50	37 120	142 080	26(13)	5797.41	1	7(2) × 10^{30}

NVT

m/(mol kg^{-1})	N_{NaCl}	N_{water}	$\langle t\rangle$/ns	V/nm^3	$\langle p\rangle$/bar	J/(s^{-1} m^{-3})
15.0	4800	17 760	110(55)	731.19	93	12(4) × 10^{30}
14.5	37 120	142 080	17(9)	5774.81	190	10(3) × 10^{30}
14.0	35 840	142 080	31(16)	5715.07	189	6(2) × 10^{30}

Coming back to the brute force simulations, if some runs do not show crystallization after a long time for a given concentration, we estimate J through the half-life time $t_{1/2}$, which in our case is the crystallization time for the 5th run to crystallize – recall that we launch ten in total – in increasing order of crystallization times. Then, we use the typical expression that relates the meanlife $\langle t\rangle$ and $t_{1/2}$ in a first order kinetic process to get an estimate of the former: $\langle t\rangle \approx t_{1/2}/\ln(2)$. The usefulness of this approach is that one does not need to wait for ten trajectories to crystallize, which sometimes can take a very long time. For instance, a system at 15 mol kg^{-1} (with 17 760 molecules of water, 4800 of NaCl and 731.19 nm^3 volume) that was simulated over 10 independent runs in the *NVT* ensemble led to a mean nucleation time of $\langle t\rangle$ = 110.35 ns (we show these trajectories in Fig. 5(a). Hence, we estimate the nucleation rate to be J = 1.2 × 10^{31} s^{-1} m^{-3}. The 5th run to crystallize does it at 59.65 ns, which leads $\langle t\rangle \approx t_{1/2}/\ln(2) \approx$ 86 ns and to J = 1.6× 10^{31} s^{-1} m^{-1}, which is in good agreement with the direct measurement of runs that completed crystallization. Thus, this calculation gives a good approximation of the nucleation rate.

We have used both *NpT* and *NVT* ensembles with no significant discrepancies, as it can be seen in Table 2 or in Fig. 5. Coming back to the previous example, we study the same 15 mol kg^{-1} system in the *NpT* ensemble. The average volume of the system in this case is 732.43 nm^3 and the mean nucleation time equals $\langle t\rangle$ = 47(7) ns. Then, the nucleation rate is J = 29(4) × 10^{30} s^{-1} m^{-3}. In this case, as can be seen in Fig. 5(b), the time evolution of the number of crystal-like ions suggests a critical cluster of around 9 ions. Due to the agreement of J obtained by either ensemble we can take the nucleation rate as the average: J = 2 × 10^{31} s^{-1} m^{-3}. This value is in excellent agreement with J = 4 × 10^{31} s^{-1} m^{-3} reported in ref. 66 for the same model and at the same concentration. Such agreement gives us great confidence in our brute force simulations.

Another interesting information that can be inferred from the spontaneous nucleation runs is the size distribution of the solid-like clusters that spontaneously appear in the system. In Fig. 6(a) we show the number of clusters containing a certain number of solid-like ions, n_{Nions}, normalised by the total number of ions in the system, $2N_{NaCl}$. This distribution has been obtained by analysing the configurations generated in our 14.5 mol kg^{-1} runs before nucleation takes place. The solid clusters have been detected as described in Section II. By taking −ln of $n_{Nions}/(2N_{NaCl})$ we get an estimate of the nucleation barrier,[65] as shown in Fig. 6(b). The appearance probability of clusters larger than 9 ions is very small and the statistics are quite poor. To acknowledge such lack of statistics, data for $n_{Nions} > 9$ are shown with a dashed line in Fig. 6(b). The barrier height for the largest cluster that we can reliably sample (N_{ions} = 9) is $\sim 15k_BT$, which is consistent with the $18.5k_BT$ barrier height estimated by seeding at 12 mol kg^{-1} (see Table 1). The barrier seems to flatten beyond 9 ions (clusters with 10, 11 and 12 ions roughly have the same appearance probability), also consistent with the 12-ion critical cluster seeding estimate at 12 mol kg^{-1}. In any case, the dashed part of the curve shown in Fig. 6(b) should be just taken as a very rough estimate of the nucleation barrier given that in spontaneous crystallization one does not sample an equilibrium but rather a stationary state distribution of critical clusters. It is nonetheless worth mentioning that in ref. 66 a nucleation barrier for the same model

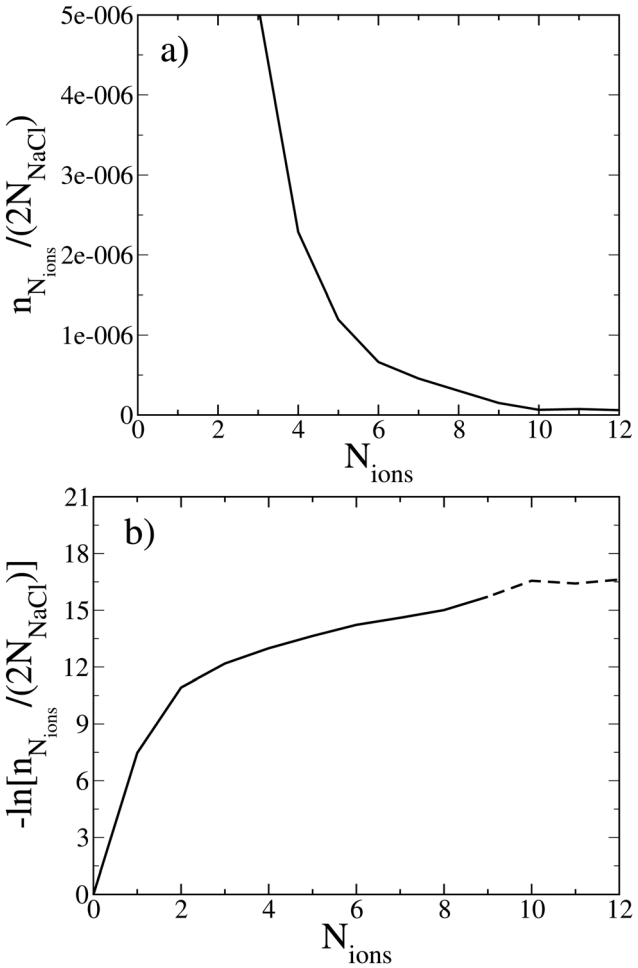

Fig. 6 (a) The number of clusters containing a given number of solid-like ions, $n_{N_{ions}}$, divided by the total number of ions for a 14.5 mol kg^{-1} solution; (b) free energy barrier estimated as $-\ln$ of such a ratio. Data in (b) beyond $N_{ions} > 9$ are shown with a dashed line because the probability of appearance of such clusters is very low and the statistics is poor.

and a close concentration (15 *vs.* 14.5 mol kg^{-1}) were rigorously computed using two different methods and the reported barrier (Fig. 2 in ref. 66) looks very similar to the one shown in Fig. 6(b). Our simulations suggest that spontaneous nucleation is expected if clusters larger than ~10 ions – detected with the order parameter presented in Section II – are observed in the simulations.

Another property we can obtain from the simulated ionic solutions is the number of contact ion pairs, n^{CIP}:

$$n^{CIP} = \rho_{ions}/2\int_0^{r_{min}} 4\pi r^2 g_{Na-Cl}(r)dr \quad (9)$$

Where r_{min} is the first minimum of the Na^+–Cl^- radial distribution function, g_{Na-Cl}. We obtain 1.0 and 1.5 ion pairs for the 12 and 14.5 mol kg^{-1} solutions respectively. For NaCl, a number of CIP larger than 1.00 is a sign of warning for imminent spontaneous precipitation of the salt (as it often occurs in many force fields of NaCl which have quite small value of solubility as compared to the experimental one).[67,68]

IV. Discussion

In Fig. 7 we show the decimal logarithm of the nucleation rate as a function of supersaturation, S, of the salty aqueous solution, defined as the ratio between the concentration of interest and the solubility. The black circles represent our estimation of the nucleation rate by seeding calculations for concentrations of 8 mol kg^{-1}, 10 mol kg^{-1} and 12 mol kg^{-1}. The nucleation rate trend provided by seeding is very similar to that of the FFS calculations by Jiang *et al.*[10] However, seeding always gives higher rates, the discrepancy being slightly larger for lower nucleation rates. It has been reported that FFS calculations are very demanding from a computational point of view and often yield underestimates of the nucleation rate.[14,51,69] This problem becomes more accentuated for state points where the rate is low (low concentrations in our case) because tiny probabilities to reach the critical nucleus need to be sampled. This could explain the fact that the discrepancy between FFS and seeding is larger for low concentrations (Fig. 7). Also in the case of hard spheres it seems that FFS gives lower rates than other techniques like seeding or Umbrella Sampling in the low driving force regime (*i.e.* low pressures in this case).[41,70] In any case, the agreement between seeding and FFS is quite reasonable. Moreover, seeding is consistent with Brute Force estimates of the nucleation rate, which are represented using red circles (this work), orange diamonds (ref. 10), and a magenta star (ref. 71). We cannot extrapolate the seeding results, which are based on CNT, to a higher supersaturation

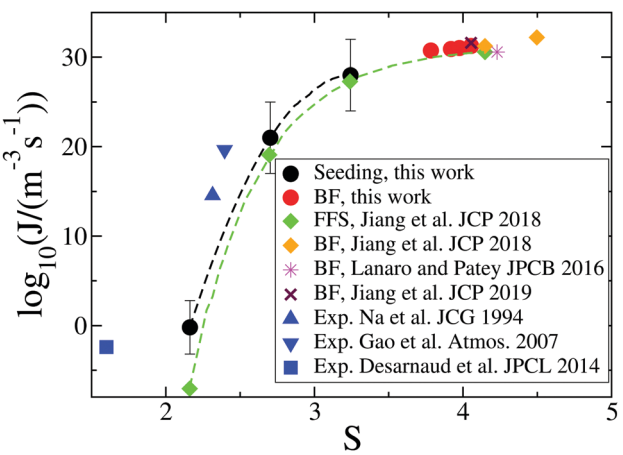

Fig. 7 Decimal logarithm of the nucleation rate of NaCl in aqueous solution as a function of supersaturation, S, defined as the ratio between the concentration of interest and the solubility concentration (both in molality units). Black circles correspond to the nucleation rates estimated from seeding and red circles to rates estimated from Brute Force (BF), both obtained in this work. The black dashed curve is a guide for the eyes for our seeding data. Diamonds are the nucleation rates reported by Jiang *et al.*,[10] green corresponding to the forward flux sampling method and orange to BF. Magenta star is the BF nucleation rate estimated from the induction times reported by Lanaro and Patey.[71] Maroon cross is the BF nucleation rate estimated with MFPT reported by Jiang *et al.*[66] Filled blue symbols are the values of the nucleation rate estimated in ref. 64 from the experiments by Na *et al.*[77] (up triangle), Gao *et al.*[76] (down triangle) and Desarnaud *et al.*[78] (square).

(beyond 15 m [66]) where crystallization is expected to take place in a 2-step spinodal fashion as opposed to a classical 1-step mechanism.[66,72] In general, the seeding approach employed in this work is only valid in principle for classical 1-step nucleation pathways that can be described within the CNT framework. There are several instances where solute precipitation is thought to be a multiple-step process[6,73–75] in which case the approach presented in this paper would not be suitable.

In this paper we revisit a study performed a few years ago where it was concluded that seeding does not provide accurate values of the rate because the results strongly depend (by tens of orders of magnitude) on the choice of the order parameter used to detect the cluster size.[22] Although we subscribe to such statement, here we show that when the mislabelling criterion is employed to tune the order parameter, seeding yields reasonable results in agreement with independent techniques that do not have such dependence on the order parameter. With such mislabelling criterion, seeding has also been consistent with independent methods for a wide variety of systems ranging from hard spheres to water.[41] Therefore, this work provides further support to the use of a mislabelling criterion as a means to identify the cluster size relevant for CNT in crystal nucleation studies.

Finally we compare the model predictions for the nucleation rate with experimental data. Please note that the comparison between model and experiments is not straightforward since the model solubility, 3.7 mol kg^{-1}, is different from the experimental one, 6.1 mol kg^{-1}. To enable a sensible comparison, model results and experimental data are plotted against supersaturation in Fig. 7. It is also worth noting that experimental data are not direct measurements of the nucleation rate, but estimates made on the supplementary material of ref. 64 based on efflorescence,[76] electrodynamic balance drop levitation,[77] and capillary crystallization[78] experiments. For a given supersaturation, the experimental data lie above the model. This indicates that the model requires higher supersaturations to achieve the same crystallization rate as in the experiments. It would be highly desirable to have a model that reproduces both the experimental solubility of NaCl and its crystal nucleation rate, but this is not the case of the JC/SPC/E model used here, that gives reasonable predictions but with a wide margin for improvement. Hopefully, the big effort currently being devoted to the development of force fields to simulate NaCl–H_2O mixtures will give rise to such a good model.[67,79–81]

V. Conclusions

In this work we have obtained the nucleation rates of NaCl in supersaturated aqueous solutions using the JC/SPC/E model. To do so, we have applied both the seeding technique and brute force simulations. Seeding results extrapolate very well toward brute force results. Moreover, the agreement of seeding and previous studies using rigorous techniques like FFS is quite reasonable. Thus, this work suggests that as long as a good order parameter is implemented in the seeding scheme, this can be a very powerful approach to study nucleation even in complex media as ionic solutions.

The determination of the critical cluster size is indeed the most delicate issue in the seeding method. Different criteria may lead to substantial differences in the rate estimates, as concluded in ref. 22. Motivated by ref. 22, and by the fact that the mislabelling criterion has proven successful in the past for hard spheres,[41] Lennard Jones,[41] water[82] and, more recently, oppositely charged colloids,[51] we decided to use the mislabelling criterion in the case of NaCl precipitation from a brine solution. We do not have a proof that the mislabelling criterion must necessarily work. The fact that the criterion works for a seeding study of NaCl crystallization from the brine is an empirical finding of this paper. In principle the mislabelling criterion seems a very reasonable approach because, by construction, it equally favours or penalises both bulk phases involved in the phase transition. However, just being "'reasonable'' is not a guarantee of success and one has to test it for every specific case, which is precisely what we do in this work for NaCl nucleation from solution. The list of systems for which the mislabelling criterion combined with seeding gives satisfactory estimates of the nucleation rate is quite wide now, making a strong case for the use of seeding to study nucleation.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This project has been funded by grants FIS2016-78117-P and PID2019-105898GB-C21 of MEC. E. G. N. thanks Agencia Estatal de Investigación and Fondo Europeo de Desarrollo Regional (FEDER), Grant No FIS2017-89361-C3-2-P. C. P. L. thanks Ministerio de Educacion y Formación Profesional for a predoctoral Formacion Profesorado Universitario Grant No. FPU18/03326 and also Ayuntamiento de Madrid for a Residencia de Estudiantes grant. The authors acknowledge the computer resources and technical assistance provided by RES. PMdH acknowledges financial support from the FPI grant no. BES-6712017-080074. J. R. E. acknowledges funding from the Oppenheimer Research fellowship and the Roger Ekins Research Fellowship of Emmanuel College. M. M. C. thanks financial support from PID2019-105898GA-C22 of the MICINN and CAM and UPM through the CavitieS project No. APOYO-JOVENES-01HQ1S-129-B5E4MM from ''Acción financiada por la Comunidad de Madrid en el marco del Convenio Plurianual con la Universidad Politecnica de Madrid en la linea de actuacion estimulo a la investigacion de jovenes doctores''. The authors gratefully acknowledge Universidad Politecnica de Madrid (www.upm.es) for providing computing resources on Magerit Supercomputer.

References

1 K. F. Kelton, *Solid State Physscis*, Elsevier, 1991, vol. 45, pp. 75–177.
2 P. G. Debenedetti and F. H. Stillinger, *Nature*, 2001, **410**, 259–267.
3 T. Nakamuro, M. Sakakibara, H. Nada, K. Harano and E. Nakamura, *J. Am. Chem. Soc.*, 2021, **143**, 1763–1767.
4 G. C. Sosso, J. Chen, S. J. Cox, M. Fitzner, P. Pedevilla, A. Zen and A. Michaelides, *Chem. Rev.*, 2016, **116**, 7078–7116.
5 R. P. Sear, *J. Phys.: Condens. Matter*, 2007, **19**, 033101.
6 V. Agarwal and B. Peters, *Adv. Chem. Phys.*, 2014, **155**, 97–160.
7 P. G. Bolhuis, D. Chandler, C. Dellago and P. L. Geissler, *Annu. Rev. Phys. Chem.*, 2002, **53**, 291–318.
8 D. Moroni, P. R. ten Wolde and P. G. Bolhuis, *Phys. Rev. Lett.*, 2005, **94**, 235703.
9 R. J. Allen, P. B. Warren and P. R. ten Wolde, *Phys. Rev. Lett.*, 2005, **94**, 018104.
10 H. Jiang, A. Haji-Akbari, P. G. Debenedetti and A. Z. Panagiotopoulos, *J. Chem. Phys.*, 2018, **148**, 044505.
11 A. Laio and M. Parrinello, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 12562.
12 M. M. Matteo Salvalaglio, Thomas Vetter and M. Parrinello, *Angew. Chem., Int. Ed.*, 2013, **52**, 13369.
13 T. Karmakar, P. M. Piaggi and M. Parrinello, *J. Chem. Theory Comput.*, 2019, **15**, 6923–6930.
14 A. Haji-Akbari and P. G. Debenedetti, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 10582–10588.
15 X.-M. Bai and M. Li, *J. Chem. Phys.*, 2006, **124**, 124707.
16 B. C. Knott, V. Molinero, M. F. Doherty and B. Peters, *J. Am. Chem. Soc.*, 2012, **134**, 19544–19547.
17 E. Sanz, C. Vega, J. Espinosa, R. Caballero-Bernal, J. Abascal and C. Valeriani, *J. Am. Chem. Soc.*, 2013, **135**, 15008–15017.
18 J. W. Gibbs, *Trans. Conn. Acad. Arts Sci.*, 1876, **3**, 108–248.
19 J. W. Gibbs, *Trans. Conn. Acad. Arts Sci.*, 1878, **16**, 343–524.
20 M. Volmer and A. Weber, *Z. Phys. Chem.*, 1926, **119**, 277–301.
21 R. Becker and W. Döring, *Ann. Phys.*, 1935, **24**, 752.
22 N. E. Zimmermann, B. Vorselaars, J. R. Espinosa, D. Quigley, W. R. Smith, E. Sanz, C. Vega and B. Peters, *J. Chem. Phys.*, 2018, **148**, 222838.
23 T. Dasgupta, G. M. Coli and M. Dijkstra, *ACS Nano*, 2020, **14**, 3957–3968.
24 P. I. Hurtado, J. Marro and P. L. Garrido, *J. Stat. Phys.*, 2008, **133**, 29–58.
25 Y. Lifanov, B. Vorselaars and D. Quigley, *J. Chem. Phys.*, 2016, **145**, 211912.
26 P. Pedevilla, M. Fitzner, G. C. Sosso and A. Michaelides, *J. Chem. Phys.*, 2018, **149**, 072327.
27 A. O. Tipeev, E. D. Zanotto and J. P. Rino, *J. Phys. Chem. C*, 2018, **122**, 28884–28894.
28 A. R. Browning, M. F. Doherty and G. H. Fredrickson, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2008, **77**, 041604.
29 A. O. Tipeev, E. D. Zanotto and J. P. Rino, *J. Phys. Chem. B*, 2020, **124**, 7979–7988.
30 L. Separdar, J. P. Rino and E. D. Zanotto, *Comput. Mater. Sci.*, 2021, **187**, 110124.
31 C. Parks, A. Koswara, F. DeVilbiss, H.-H. Tung, N. K. Nere, S. Bordawekar, Z. K. Nagy and D. Ramkrishna, *Phys. Chem. Chem. Phys.*, 2017, **19**, 5285–5295.
32 Y. Sun, F. Zhang, H. Song, M. I. Mendelev, C.-Z. Wang and K.-M. Ho, *J. Chem. Phys.*, 2018, **149**, 174501.
33 T. Mandal and R. G. Larson, *J. Chem. Phys.*, 2017, **146**, 134501.
34 T. Mizuguchi, Y. Higeta and T. Odagaki, *Phys. Rev. E*, 2017, **95**, 042804.
35 C. Parks, A. Koswara, H.-H. Tung, N. Nere, S. Bordawekar, Z. K. Nagy and D. Ramkrishna, *Cryst. Growth Des.*, 2017, **17**, 3751–3765.
36 L. Zepeda-Ruiz, B. Sadigh, A. Chernov, T. Haxhimali, A. Samanta, T. Oppelstrup, S. Hamel, L. Benedict and J. Belof, *J. Chem. Phys.*, 2017, **147**, 194704.
37 Y. Sun, F. Zhang, H. Song, M. I. Mendelev, C.-Z. Wang and K.-M. Ho, *J. Phys. Chem. C*, 2019, **123**, 6685–6692.
38 Y. Sun, H. Song, F. Zhang, L. Yang, Z. Ye, M. I. Mendelev, C.-Z. Wang and K.-M. Ho, *Phys. Rev. Lett.*, 2018, **120**, 085703.
39 F. Leoni, R. Shi, H. Tanaka and J. Russo, *J. Chem. Phys.*, 2019, **151**, 044505.
40 I. Zubieta, M. V. Del Saz, P. Llombart, C. Vega and E. G. Noya, *Phys. Chem. Chem. Phys.*, 2019, **21**, 1656–1670.
41 J. R. Espinosa, C. Vega, C. Valeriani and E. Sanz, *J. Chem. Phys.*, 2016, **144**, 034501.
42 P. Montero de Hijes, J. R. Espinosa, E. Sanz and C. Vega, *J. Chem. Phys.*, 2019, **151**, 144501.
43 P. Montero de Hijes, J. R. Espinosa, V. Bianco, E. Sanz and C. Vega, *J. Phys. Chem. C*, 2020, **124**, 8795–8805.
44 J. R. Espinosa, C. Vega, C. Valeriani and E. Sanz, *J. Chem. Phys.*, 2015, **142**, 194709.
45 J. Espinosa, E. Sanz, C. Valeriani and C. Vega, *J. Chem. Phys.*, 2014, **141**, 18C529.
46 J. R. Espinosa, A. Zaragoza, P. Rosales-Pelaez, C. Navarro, C. Valeriani, C. Vega and E. Sanz, *Phys. Rev. Lett.*, 2016, **117**, 135702.
47 A. Zaragoza, M. M. Conde, J. R. Espinosa, C. Valeriani, C. Vega and E. Sanz, *J. Chem. Phys.*, 2015, **143**, 134504.
48 V. Bianco, P. M. de Hijes, C. P. Lamas, E. Sanz and C. Vega, *Phys. Rev. Lett.*, 2021, **126**, 015704.
49 G. D. Soria, J. R. Espinosa, J. Ramirez, C. Valeriani, C. Vega and E. Sanz, *J. Chem. Phys.*, 2018, **148**, 222811.
50 J. R. Espinosa, G. D. Soria, J. Ramirez, C. Valeriani, C. Vega and E. Sanz, *J. Phys. Chem. Lett.*, 2017, **8**, 4486.
51 I. Sanchez-Burgos, A. Garaizar, C. Vega, E. Sanz and J. R. Espinosa, *Soft Matter*, 2021, **17**, 489–505.
52 I. S. Joung and T. E. Cheatham III, *J. Phys. Chem. B*, 2008, **112**, 9020–9041.
53 A. Benavides, J. Aragones and C. Vega, *J. Chem. Phys.*, 2016, **144**, 124504.
54 F. Moučka, I. Nezbeda and W. R. Smith, *J. Chem. Phys.*, 2013, **138**, 154102.
55 Z. Mester and A. Z. Panagiotopoulos, *J. Chem. Phys.*, 2015, **143**, 044505.
56 Z. Mester and A. Z. Panagiotopoulos, *J. Chem. Phys.*, 2015, **142**, 044507.

57 B. Hess, C. Kutzner, D. Van Der Spoel and E. Lindahl, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.
58 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.
59 G. J. Martyna, M. L. Klein and M. Tuckerman, *J. Chem. Phys.*, 1992, **97**, 2635–2643.
60 D. R. Wheeler and J. Newman, *Chem. Phys. Lett.*, 2002, **366**, 537–543.
61 B. Hess, H. Bekker, H. J. Berendsen and J. G. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.
62 B. Hess, *J. Chem. Theory Comput.*, 2008, **4**, 116–122.
63 W. Lechner and C. Dellago, *J. Chem. Phys.*, 2008, **129**, 114707.
64 N. E. Zimmermann, B. Vorselaars, D. Quigley and B. Peters, *J. Am. Chem. Soc.*, 2015, **137**, 13352–13361.
65 S. Auer and D. Frenkel, *J. Chem. Phys.*, 2004, **120**, 3015–3029.
66 H. Jiang, P. G. Debenedetti and A. Z. Panagiotopoulos, *J. Chem. Phys.*, 2019, **150**, 124502.
67 W. R. Smith, I. Nezbeda, J. Kolafa and F. Moučka, *Fluid Phase Equilib.*, 2018, **466**, 19–30.
68 A. Z. Panagiotopoulos, *J. Chem. Phys.*, 2020, **153**, 010903.
69 A. Haji-Akbari, *J. Chem. Phys.*, 2018, **149**, 072303.
70 L. Filion, M. Hermes, R. Ni and M. Dijkstra, *J. Chem. Phys.*, 2010, **133**, 244115.
71 G. Lanaro and G. Patey, *J. Phys. Chem. B*, 2016, **120**, 9076–9087.
72 D. Chakraborty and G. Patey, *Chem. Phys. Lett.*, 2013, **587**, 25–29.
73 P. G. Vekilov, *Nanoscale*, 2010, **2**, 2346–2357.
74 D. Erdemir, A. Y. Lee and A. S. Myerson, *Acc. Chem. Res.*, 2009, **42**, 621–629.
75 D. Gebauer, M. Kellermeier, J. D. Gale, L. Bergstrom and H. Colfen, *Chem. Soc. Rev.*, 2014, **43**, 2348–2371.
76 Y. Gao, L. E. Yu and S. B. Chen, *J. Phys. Chem. A*, 2007, **111**, 10660–10666.
77 H.-S. Na, S. Arnold and A. S. Myerson, *J. Cryst. Growth*, 1994, **139**, 104–112.
78 J. Desarnaud, H. Derluyn, J. Carmeliet, D. Bonn and N. Shahidzadeh, *J. Phys. Chem. Lett.*, 2014, **5**, 890–895.
79 H. Jiang, Z. Mester, O. A. Moultos, I. G. Economou and A. Z. Panagiotopoulos, *J. Chem. Theory Comput.*, 2015, **11**, 3802–3810.
80 A. Benavides, M. Portillo, V. Chamorro, J. Espinosa, J. Abascal and C. Vega, *J. Chem. Phys.*, 2017, **147**, 104501.
81 J. Kolafa, *J. Chem. Phys.*, 2016, **145**, 204509.
82 J. R. Espinosa, C. Vega and E. Sanz, *J. Chem. Phys.*, 2018, **122**, 22892–22896.

