

Microscopic Mechanism and Kinetics of Ice Formation at Complex Interfaces: Zooming in on Kaolinite

SUPPORTING INFORMATION

Gabriele C. Sosso,^{,†} Tianshu Li,[‡] Davide Donadio,[¶] Gareth A. Tribello,[§] and
Angelos Michaelides[†]*

*[†]Thomas Young Centre, London Centre for Nanotechnology and Department of Physics and
Astronomy, University College London, Gower Street London WC1E 6BT, United Kingdom*

*[‡]Department of Civil and Environmental Engineering, George Washington University,
Washington, D.C. 20052, United States*

*[¶]Department of Chemistry, University of California Davis, One Shields Avenue, Davis,
CA 95616, USA*

*[§]Atomistic Simulation Centre, Department of Physics and Astronomy, Queen's University
Belfast, University Road Belfast BT7 1NN, United Kingdom*

E-mail: g.sosso@ucl.ac.uk

We provide supporting information on the calculation of the heterogeneous ice nucleation rate on the kaolinite (001) hydroxylated surface. The computational geometry is specified together with the details of the molecular dynamics simulations used in this work. Moreover, we discuss the choice of the order parameter we have employed within the forward flux sampling calculations, and we provide additional information about the implementation of

the algorithm and the results obtained at each stage of the latter. A brief discussion about heterogeneous classical nucleation theory is also presented together with the technical details of the topological criteria used to characterize the ice nuclei and a discussion about finite size effects.

Computational Geometry

The computational setup we have used is depicted in Fig. S1a. A single layer of kaolinite, cleaved along the (001) plane (perpendicular to the normal to the slab) was prepared by starting from the experimental cell parameters and lattice positions.¹ Specifically, a kaolinite bulk system made of two identical slabs was cleaved along the (001) plane. The triclinic symmetry of the system (space group $C1$) was modified by setting the α and γ angles (experimentally equal to 91.926 and 89.797 degrees respectively¹) to 90 degrees in order to make the cell orthorhombic. We explicitly verified that this modification does not introduce any structural change within the clay. The final slab has in-plane dimensions of 61.84 and 71.54 Å, corresponding to a 12 by 8 supercell. We positioned 6144 water molecules randomly atop this kaolinite slab at the density of the TIP4P/Ice model² at 300 K, and expanded the dimension of the simulation cell along the normal to the slab to 150 Å. This setup allows for a physically meaningful equilibration of the water at the density of interest at a given temperature, but suffers from two distinct drawbacks: i) the kaolinite slab possesses a net dipole moment which is not compensated throughout the simulation cell and ii) the presence of the water-vacuum interface can alter the structure and the dynamics of the liquid film. However, we have verified that compensating the dipole moment by means of a mirror slab does not affect our simulations, as we have been able to replicate the results of Ref. 3 independently of the computational geometry. Furthermore, the water film is thick enough to allow a bulk-like region to exist in terms of both structure and dynamics. The effect of the water-vacuum interface is therefore negligible. In Fig. S1b we highlight the layered

nature of the slab, while in Fig. S1c we zoom in on a portion of the (001) hydroxylated surface and show the hexagonal arrangement of the hydroxyl groups. This arrangement is important as the water can interact with the hydroxyls, so this arrangement is responsible for the templating effect of the clay which serves to promote ice nucleation. The amphoteric nature of the hydroxyl groups at the surface is depicted in Fig. S1d.

Molecular Dynamics Simulations

The CLAY_FF⁴ force field was used to model the kaolinite slab. We have not included the - optional - angular term (see Ref. 4), as we have verified that it does not affect the structure of the surface. In order to mimic the experimental conditions, we have constrained the system at the experimental lateral dimensions (see above), and have also restrained the positions of the silicon atoms at the bottom of the slab by means of a harmonic potential characterized by a spring constant of 1000 kJ/mol. All the other atoms within the kaolinite slab are unconstrained. We have verified that the thermal expansion of the clay at 230 K ($\sim 0.4\%$ with respect to each lateral dimension) does not alter the structure nor the dynamics of the water-kaolinite interface. This setup is thus as close as we can get to the realistic (001) hydroxylated surface within the CLAY_FF model. The interaction between the water molecules have been modeled using the TIP4P/Ice model,² so that our results are consistent with the homogeneous simulations of Ref. 5. The interaction parameters between the clay and the water were obtained using the standard Lorentz-Berthelot mixing rules.^{6,7} Extreme care must be taken in order to correctly reproduce the structure and the dynamics of the water-clay interface. The Forward Flux Sampling (FFS) simulations reported in this work rely on a massive collection of unbiased Molecular Dynamics (MD) runs, all of which have been performed using the GROMACS package, version 4.6.7. The code was compiled in single-precision, in order to alleviate the huge computational workload needed to converge the FFS algorithm and because we have taken advantage of GPU acceleration, which is not

available in the double-precision version. The equations of motions were integrated using a leap-frog integrator with a timestep of 2 fs. The van der Waals (non bonded) interactions were considered up to 10 Å, where a switching function was used to bring them to zero at 12 Å. Electrostatic interactions have been dealt with by means of an Ewald summation up to 14 Å. The NVT ensemble was sampled at 230 K using a stochastic velocity rescaling thermostat⁸ with a very weak coupling constant of 4 ps in order to avoid temperature gradients throughout the system. The geometry of the water molecules (TIP4P/Ice being a rigid model) was constrained using the SETTLE algorithm⁹ while the P_LINCS algorithm¹⁰ was used to constrain the O-H bonds within the clay. We have verified that these settings reproduce the dynamical properties of water reported in Ref.⁵ The system was equilibrated at 300 K for 10 ns, before being quenched to 230 K over 50 ns. This is the starting point for the calculation of the flux rate discussed in the next section.

Forward Flux Sampling Simulation

Order Parameter

The first step in setting up the FFS simulation involved choosing a suitable order parameter λ . We start by labeling as ice-like any water molecule whose oxygen atom displays a value of $lq^6 > 0.45$, where lq^6 is constructed as follows: we first select only those oxygens which are hydrogen-bonded to four other oxygens. For each of the i -th atoms of this subset S_{4HB} , we calculate the local order parameter:

$$lq_i^6 = \frac{\sum_{j=1}^{N_{S_{4HB}}} \sigma(\mathbf{r}_{ij}) \sum_{m=-6}^6 q_{i,m}^{6*} \cdot q_{j,m}^6}{\sum_{j=1}^{N_{S_{4HB}}} \sigma(\mathbf{r}_{ij})} \quad (1)$$

where $\sigma(\mathbf{r}_{ij})$ is a switching function tuned so that $\sigma(\mathbf{r}_{ij})=1$ when atom j lies within the first coordination shell of atom i and which is zero otherwise. $q_{i,m}^6$ is the Steinhardt vector¹¹

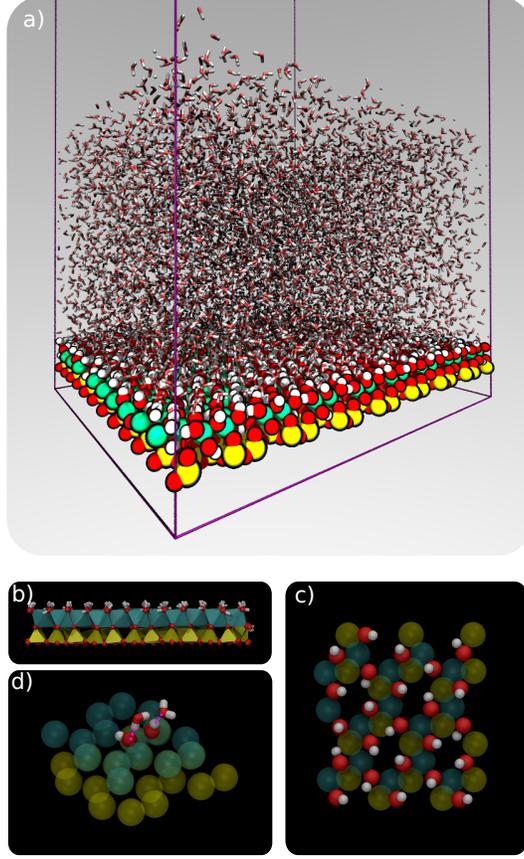


Fig. S1: a) The simulation cell used in this work. A film of liquid water about 40 Å thick is in contact with a single slab of kaolinite, cleaved along the (001) plane. This slab geometry is thus characterized by two interfaces: the water-kaolinite interface and the water-vacuum interface. The dimension of the simulation box along the normal to the slab is extended up to 150 Å. Water molecules are depicted as sticks, while atoms within the kaolinite slab as balls. Red, white, light blue and yellow atoms correspond to oxygen, hydrogen, aluminum and silicon atoms respectively. b) (side view) The layered structure of the kaolinite slab: yellow tetrahedra and light blue octahedra represent the tetrahedral silica sheet and the octahedral alumina sheet, terminated with hydroxyl groups, respectively. c) (top view) A small portion of the kaolinite slab depicting the hexagonal arrangement of the hydroxyl groups exposed. d) Sketch of the amphoteric character of the hydroxylated (001) face of kaolinite: the hydroxyl groups on top can either donate or accept an hydrogen bond from e.g. water molecules at the water-clay interface.

$$q_{i,m}^6 = \frac{\sum_{j=1}^{N_{SAHB}} \sigma(\mathbf{r}_{ij}) Y_{6m}(\mathbf{r}_{ij})}{\sum_{j=1}^{N_{SAHB}} \sigma(\mathbf{r}_{ij})}, \quad (2)$$

$Y_{6m}(\mathbf{r}_{ij})$ being one of the 6th order spherical harmonics. We have used 3.2 Å as the cutoff for $\sigma(\mathbf{r}_{ij})$ to be consistent with Ref. 5. Note that by selecting oxygen atoms within the

S_{AHB} subset exclusively we ensure that the hydrogen bond network within the ice nuclei is reasonable. Having identified a set of ice-like water molecules, we pinpoint all the connected clusters of oxygen atoms which: i) belong to the S_{AHB} subset; ii) have a value of $lq^6 > 0.45$ and; iii) are separated by a distance $\leq 3.2 \text{ \AA}$. We then select the largest of these clusters (i.e. the one containing the largest number of oxygen atoms or equivalently water molecules). The final step is to find all the *surface molecules* that are connected to this cluster, as this procedure allows us to account for the diffuse interface between the solid and the liquid. Surface molecules are defined as the water molecules that lie within 3.2 \AA from the molecules in the cluster. The final order parameter λ used in this work is thus the number of water molecules within the largest ice-like cluster plus the number of surface molecules. This approach allow us to include ice-like atoms sitting directly on top of the kaolinite surface, which are never labeled as ice-like (and which would thus never be included into the ice nuclei) because they are undercoordinated and because they display a different symmetry to the molecules within bulk water (which in turn leads to different values of lq^6). Note that the order parameter used in Ref. 5 differs with respect to our formulation in that i) a slightly stricter criterion has been used to label molecules as ice-like, namely $lq^6 > 0.5$ to be compared with our choice of $lq^6 > 0.45$; and ii) surface molecules are not included in the largest ice-like nucleus. This means that in order to compare quantitatively our results with those of Ref. 5 in terms of e.g. the size of the critical nucleus, the very same order parameter has to be used. The calculation of the order parameter is performed on the fly during our MD simulations thanks to the flexibility of the PLUMED plugin¹² (version 2.2). This code deals chiefly with metadynamics simulations, but can be adapted to a FFS simulation. Note that PLUMED benefits from a fully parallel implementation that flawlessly couples with the GPU-accelerated version of GROMACS, and thus provides a very fast tool for performing FFS simulations. Indeed, while several implementations of FFS are beginning to appear, the main issue preventing wider adoption remains the implementation of the order parameter, which can be as complex as the one used in this work. PLUMED allows a wide range of

order parameters to be exploited without the need to re-code them elsewhere.

Converging the Flux Rate and the Individual Crossing Probabilities

In order to calculate the flux rate Φ_0 we have performed a 1.5 ms long unbiased MD simulation, and subsequently built the probability density distribution for $P(\lambda)$ shown in Fig. S2a. We have thus delimited the liquid basin in terms of the order parameter as $0 < \lambda < \lambda_{Liq} = 32$, while setting the initial interface for the FFS $\lambda_0=75$, corresponding to a value of the cumulative distribution function of $P(\lambda)$ (also reported in Fig. S2a) of 0.99. The flux rate is then computed as the number of direct crossings of λ_0 (i.e. coming from $\lambda < \lambda_{Liq}$) divided by the total simulation time, and as such should flatten as a function of time. Meanwhile, the number of direct crossings should increase linearly with time. The value obtained for Φ_0 and the number of crossings as a function of time are reported in Fig. S2b. This figure demonstrates that, as previously noted in Ref. 13, long simulation times are needed in order to converge this quantity for inhomogeneous systems. The calculated value of Φ_0 is 0.00056359 ps⁻¹, which normalized by the average volume of the water film (189350.2980352 Å³) leads to the final value of $3.0 \cdot 10^{-9 \pm 1}$ ps⁻¹ Å⁻³. Note that we have chosen to normalize the flux rate by the average volume of the water film instead of by the surface area for the slab. While the latter choice could in principle be thought as more meaningful in the context of heterogeneous nucleation, our objective is to compare our numbers with the homogeneous case, which is why we choose the volume normalization rather than the surface area one. However, it should be noticed that the two different normalizations only introduce a difference of an order of magnitude in the nucleation rate. The number of starting configurations, one for each direct crossing of λ_0 , is of the order of eight hundred, providing a comprehensive sampling including ice-like clusters in the bulk of the water film as well as on top of the water surface (albeit the latter represent about 25%).

Converging the individual crossing probabilities $P(\lambda_i|\lambda_{i-1})$ required in our case as many as 10,000 trial MD runs for the first few interfaces. The initial velocities for each MD run

were randomly initialized consistent with the corresponding Maxwell-Boltzmann distribution at 230 K. In line with the coarse graining approach discussed in Ref. 5, we have decided to compute the value of λ on the fly every 4 ps, a frequency far smaller than the relaxation time of the liquid at this temperature (about 0.5 ns) which allows us to neglect meaningless fluctuation on very short timescales. The individual crossing probabilities, normalized by their value after 250 crossing events, are reported in Fig. S2c. Note that at the interfaces corresponding to critical/post-critical ice nuclei a much smaller number (about 500) of trial MD runs have been shot, as for large ice nuclei to get back to the liquid phase simulation times of the order of 10-40 ns are needed, dramatically increasing the computational cost - albeit more and more nuclei proceed to grow as λ increases leading to a faster convergence of the crossing probabilities. In fact, crossings for $n > 250$ are not reported in Fig. S2c as the crossing probabilities are already converged well before $n = 250$ within the last stages of the algorithm. The confidence intervals for each $P(\lambda_i | \lambda_{i-1})$ have been computed according to the binomial distribution of the number of successful trial runs collected at λ_i (see e.g. Ref. 14).

Heterogeneous Classical Nucleation Theory

Within the framework of classical nucleation theory, the homogeneous rate of nucleation J_{Homo} can be written as:^{15,16}

$$J_{Homo} = A_{Homo} \cdot e^{-\frac{\Delta G_{Homo}^*}{k_B T}} \quad (3)$$

where A_{Homo} is a kinetic prefactor, ΔG_{Homo}^* is the height of the free energy barrier for nucleation and k_B is the Boltzmann constant. On the other hand, the heterogeneous rate of nucleation J_{Hetero} can be written as:^{15,16}

$$J_{Hetero} = A_{Hetero} \cdot e^{-\frac{\mathcal{F}_S \cdot \Delta G_{Homo}^*}{k_B T}} \quad (4)$$

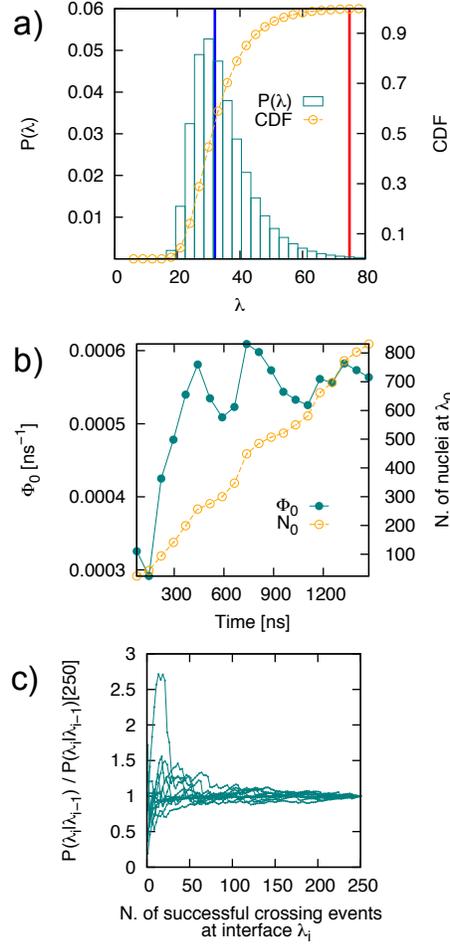


Fig. S2: a) Probability density distribution for the order parameter λ ($P(\lambda)$ left y-axis, boxes) and correspondent cumulative distribution function (CDF, right y-axis, empty circles). The blue and red vertical arrows mark the upper limit of the liquid basin λ_{Liq} and the position of λ_0 respectively. b) Flux rate (Φ_0 , left y-axis, filled circles) and number of direct crossing of the λ_0 interface (N_0 , right y-axis, empty circles) as a function of simulation time. c) Individual crossing probabilities $P(\lambda_i|\lambda_{i-1})$ (normalized by their value at $N=250$) as a function of the number of crossing events.

where A_{Hetero} is a kinetic prefactor which in principle can differ from A_{Homo} and \mathcal{F}_S is a shape factor, or potency factor, which embeds the effectiveness of the substrate to promote nucleation. The value of \mathcal{F}_S ranges from one (the surface does not contribute at all in lowering the free energy barrier for nucleation) to 0 (the nucleation proceeds in a barrierless fashion). By taking the ratio $\frac{J_{Hetero}}{J_{Homo}}$ and assuming that $A_{Hetero} = A_{Homo}$ (which is in many

cases a perfectly reasonable assumption, see e.g. Refs. 17–19), one can write the shape factor for heterogeneous nucleation as:

$$\mathcal{F}_S = 1 - \left[\frac{k_B T}{\Delta G_{Homo}^*} \cdot \ln \left(\frac{J_{Hetero}}{J_{Homo}} \right) \right] \quad (5)$$

The value of $\Delta G_{Homo}^* = \frac{1}{2} |\Delta \mu_{sl}| N_{Homo}^C$ is $86 \pm 5 k_B T$, obtained from Ref. 5 by using the definition of λ we have employed here (thus using a slightly different lq_i^6 cutoff and including surface molecules, see Eq. 1) - which accounts for an homogeneous critical nucleus size of 540 ± 30 water molecules and makes a direct comparison possible. Inserting this value into the expression above leads to a shape factor of 0.46 ± 0.09 .

Double-Diamond and Hexagonal Cages

Double-Diamond (DDC) and Hexagonal cages (HC) are the building blocks of cubic and hexagonal ice respectively. We have identified water molecules involved in DDC and/or HC within the largest ice nucleus in the system (defined according to the order parameter λ , see Eqs. 1 and 2) following the topological criteria detailed in Ref. 5. The first step in order to locate DDC and HC is the construction of the ring network of the oxygen atoms belonging to each water molecule. In this work, we have obtained all the six-atom rings needed to build DDC and HC using King’s shortest path criterion^{20,21} as implemented in the R.I.N.G.S. code.²² The same distance cutoff of 3.2 Å used for the construction of the order parameter λ has been employed to determine the nearest neighbors of each oxygen atom. The same algorithm described in Ref. 5 has subsequently been used to determine DDC and HC.

Asphericity Parameter

Many different choices are available to quantify the asphericity of clusters of molecules. We have considered the gyration radius as well as the α (Δ in Ref. 23) and S asphericity

parameters reported in Ref. 23. All of these quantities provided the same qualitative picture, so we have chosen to report the asphericity trends for α only, the latter being defined as:

$$\alpha = \frac{3}{2(\text{tr}\mathcal{T})^2} \sum_{i=1}^3 (\mu_i - \bar{\mu})^2 \quad (6)$$

where μ_i are the three eigenvalues of the inertia tensor \mathcal{T} for a given cluster, and $\bar{\mu} = \frac{\text{tr}\mathcal{T}}{3} = \frac{\sum_{i=1}^3 (\mu_i)}{3}$

Spatial extent Δz

The spatial extent Δz for a given ice nucleus has been calculated as the difference between the minimum and maximum values of the z- components of the position vector of all the oxygens belonging to the nucleus. As the direction normal to the kaolinite slab coincides to the z-axis of our simulation box, Δz provides a qualitative indication of the number of ice layers in the nuclei. Ice nuclei are defined to be on top of the kaolinite surface (*Surface*, see main text) if the minimum value of the z- components of the position vector of all the oxygens belonging to the nucleus is $< 15.0 \text{ \AA}$, which correspond to the position of the main peak in the density profile of the water film along the z-axis. If this is not the case, the ice nuclei are considered to sit in the bulk of the water film (*Bulk*, see main text).

Avoiding Finite Size Effects

Special care has to be taken when dealing with atomistic simulations of crystal nucleation from the liquid phase. Specifically, the presence of periodic boundary conditions can introduce significant finite effects, most notably spurious interactions between the crystalline nuclei and their periodic images. This artefact results in nonphysically large nucleation rates and/or crystal growth speeds. In this work we have considered simulation boxes with lateral dimensions of the order of 60 \AA , which is sufficient to ensure that finite size effects do not

affect our results. We also measured the distance between the ice nuclei and their periodic images using the average set-set distance $d(A, B)$, which is defined as:

$$d(A, B) = \inf_{x \in A, y \in B} \lim_{\lambda \rightarrow \infty} |\mathbf{x} - \mathbf{y}| \quad (7)$$

where x and y are the position vectors of each oxygen atoms belonging to the largest ice nucleus (defined according to the order parameter λ) A and its first periodic image B respectively. At the FFS interface closest to the critical nucleus size ($\lambda=225$), $d(A, B)=20\pm 6\text{\AA}$, and even at the last FFS interface we have considered ($\lambda=325$) the ice nuclei are still quite far away from their periodic images, $d(A, B)$ being $15\pm 7\text{\AA}$, which is of the order of 1/4 of the lateral dimension of the simulation box.

References

- (1) Bish, D. L. *Clays and Clay Minerals* **1993**, *41*, 738–744.
- (2) Abascal, J. L. F.; Sanz, E.; Fernandez, R. G.; Vega, C. *The Journal of Chemical Physics* **2005**, *122*, 234511.
- (3) Zielke, S. A.; Bertram, A. K.; Patey, G. N. *J. Phys. Chem. B* **2016**, *120*, 1726–1734.
- (4) Cygan, R. T.; Liang, J.-J.; Kalinichev, A. G. *The Journal of Physical Chemistry B* **2004**, *108*, 1255–1266.
- (5) Haji-Akbari, A.; Debenedetti, P. G. *Proceedings of the National Academy of Sciences* **2015**, 201509267.
- (6) Lorentz, H. A. *Annalen der Physik* **1881**, *248*, 127–136.
- (7) Berthelot, D.; Hebd, C. R. *Seances De L' Academie Des Sciences*. 1898.
- (8) Bussi, G.; Donadio, D.; Parrinello, M. *The Journal of Chemical Physics* **2007**, *126*, 014101.

- (9) Miyamoto, S.; Kollman, P. A. *Journal of Computational Chemistry* **1992**, *13*, 952–962.
- (10) Hess, B.; Bekker, H.; Berendsen, H. J. C.; Fraaije, J. G. E. M. *Journal of Computational Chemistry* **1997**, *18*, 1463–1472.
- (11) Steinhardt, P. J.; Nelson, D. R.; Ronchetti, M. *Physical Review B* **1983**, *28*, 784–805.
- (12) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. *Computer Physics Communications* **2014**, *185*, 604–613.
- (13) Bi, Y.; Li, T. *The Journal of Physical Chemistry B* **2014**, *118*, 13324–13332.
- (14) Allen, R. J.; Frenkel, D.; Wolde, P. R. t. *The Journal of Chemical Physics* **2006**, *124*, 194111.
- (15) Sear, R. P. *Journal of Physics: Condensed Matter* **2007**, *19*, 033101.
- (16) Kalikmanov, V. *Nucleation Theory*; Lecture Notes in Physics; Springer Netherlands: Dordrecht, 2013; Vol. 860.
- (17) Li, T.; Donadio, D.; Russo, G.; Galli, G. *Physical Chemistry Chemical Physics* **2011**, *13*, 19807–19813.
- (18) Li, T.; Donadio, D.; Galli, G. *Nature Communications* **2013**, *4*, 1887.
- (19) Gianetti, M. M.; Haji-Akbari, A.; Longinotti, M. P.; Debenedetti, P. G. *Physical Chemistry Chemical Physics* **2016**, *18*, 4102–4111.
- (20) King, S. V. *Nature* **1967**, *213*, 1112–1113.
- (21) Franzblau, D. S. *Physical Review B* **1991**, *44*, 4925–4930.
- (22) Le Roux, S.; Jund, P. *Computational Materials Science* **2010**, *49*, 70–83.
- (23) Rawat, N.; Biswas, P. *Physical Chemistry Chemical Physics* **2011**, *13*, 9632–9643.