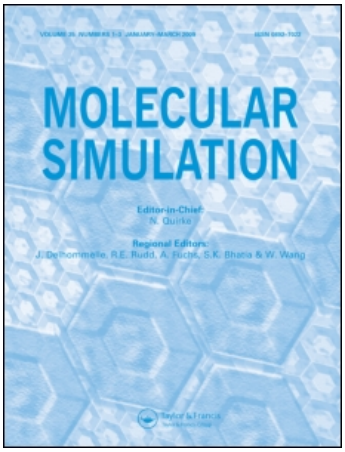

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

Mitosis method for directly calculating the interfacial free energy of nuclei

Nathan Duff[a]; Baron Peters[ab]
[a] Department of Chemical Engineering, University of California, Santa Barbara, CA, USA [b] Department of Chemistry and Biochemistry, University of California, Santa Barbara, CA, USA

Online publication date: 03 August 2010

To cite this Article Duff, Nathan and Peters, Baron(2010) 'Mitosis method for directly calculating the interfacial free energy of nuclei', Molecular Simulation, 36: 7, 498 — 504
To link to this Article: DOI: 10.1080/08927022.2010.483684
URL: http://dx.doi.org/10.1080/08927022.2010.483684

Molecular Simulation
Vol. 36, Nos. 7–8, June–July 2010, 498–504

Taylor & Francis
Taylor & Francis Group

Mitosis method for directly calculating the interfacial free energy of nuclei

Nathan Duff[a] and Baron Peters[ab]*

[a]*Department of Chemical Engineering, University of California, Santa Barbara, CA, USA;* [b]*Department of Chemistry and Biochemistry, University of California, Santa Barbara, CA, USA*

(*Received 21 December 2009; final version received 5 April 2010*)

Nucleation from solution is an important first step in many natural and industrial phase transformations. In contrast to nucleation in pure component systems, the nucleation of solutes from solution remains a major challenge for simulation. The challenge stems from difficulty in inserting solutes to maintain an excess solute chemical potential in solution as the pre-critical nucleus grows. We present a 'mitosis' method to estimate the effective interfacial free energy of nanoscale nuclei in solution. The new method circumvents the need for controlling chemical potential by particle insertion, and thus it can provide insight on real crystal nucleation processes in explicit solvents. To test the new mitosis method, we demonstrate it on the Potts lattice gas model of nucleation from solution. This lattice model enables an alternative method for computing the interfacial free energy that has been used in simulations of nucleation in pure component systems. The two methods for computing the interfacial free energy agree to within 6%. This confirms that the mitosis method is a viable approach for computing the interfacial free energy of solute nuclei in real solutions. The method should facilitate a molecular level understanding of how solvents and additives influence nucleation rates and polymorph selection.

Keywords: nucleation; interfacial free energy; lattice model; surface tension; polymorph selection

1. Introduction

Simulations of nucleation in pure component systems often compute the free energy as a function of nucleus size by importance of sampling along a nucleus size coordinate in canonical or isothermal–isobaric ensembles [1–23]. Unfortunately, this approach has been unsuccessful for nucleation from solution because the excess solute chemical potential (the driving force) cannot be maintained throughout the nucleation process. Alternatively, some simulations of nucleation identify the critical nucleus size by studying the trajectories that relax from a seeded configuration [19,24–30]. In the seeded simulation approach, the critical nucleus size is identified as the size from which nuclei have an equal tendency to grow or dissolve back into solution [12,30–32]. Again, seeded simulations face particular challenges for nucleation of solutes from an explicit solvent. When seeded trajectories are computed with a fixed number of solutes and solvents, growth and dissolution are artificially self-arrested. This occurs because the solute concentration in the surrounding solution is depleted or enriched as the nucleus grows or dissolves, respectively [33].

Simulations in the grand canonical ensemble preserve the chemical potential driving forces for nucleation from multicomponent vapours [20,21,29], but they are impractical for nucleation from explicit solvents. The impasse stems from the low acceptance probability for inserting a solute molecule into a dense liquid phase [34]. Several methods have been developed to improve the probability of solute insertion and removal including particle insertions into cavities [35], configurational bias moves [36–38] and staged fractional insertion methods [38–40]. The semigrand canonical ensemble of Kofke and Glandt [41] uses solute–solvent exchange moves instead of particle insertions. Despite these advanced methods, chemical potential remains extremely difficult to control in simulations of crystal nucleation from solution [15].

Classical nucleation theory (CNT) [42,43] has been widely used to study pure component nucleation. CNT relies on bulk thermodynamic values for the interfacial free energy and for the chemical potential driving force. The bulk thermodynamic values are intuitively inappropriate for nanoscale nuclei, and, indeed, using the thermodynamic values of the driving force and interfacial free energy gives nucleation rates that are many orders of magnitude different from experimental values [44–47]. Therefore, simulation methods that calculate the macroscopic surface tension [48–50] are not applicable for nucleation.

Nevertheless, simulations suggest that some aspects of CNT are qualitatively correct [3–7,10,11]. Specifically, a variety of systems show that the free energy as a function of nucleus size resembles the form $F(n) = \mu_{\text{eff}} n + \gamma_{\text{eff}} n^{2/3}$ as predicted by CNT, but with effective values for the driving force (μ_{eff}) and interfacial free energy (γ_{eff}) that are appropriate at the nanoscale.

*Corresponding author. Email: baronp@engineering.ucsb.edu

ISSN 0892-7022 print/ISSN 1029-0435 online

DOI: 10.1080/08927022.2010.483684
http://www.informaworld.com

We show how seeded simulations can be used to estimate the interfacial free energy for nanoscale nuclei. The interfacial free energy strongly influences the free energy barrier to nucleation. The mitosis method relates the interfacial free energy to the reversible work required to form two nuclei of size n from one nucleus of size $2n$. The composition of the surrounding solution does not change during the nucleus scission, so for points far from a critical concentration there is no need to control chemical potential with solute insertions. The method is therefore easy to implement using hybrid Monte Carlo (MC) and molecular dynamics algorithms in packages such as GROMACS [51], AMBER [52], NAMD [53], CHARMM [54] and LAMMPS [55].

Mitosis simulations may help engineer solvents or additives to enhance or inhibit nucleation by changing the interfacial free energy. Furthermore, the interfacial free energy may depend on the structure or polymorph of the growing nucleus [56]. By using the mitosis method while requiring the nucleus to maintain a particular polymorph structure, we can ask how the interfacial free energy for nanoscale nuclei of a particular polymorph depends on solvent properties and additives. Simulations of this type may thus provide insight on polymorph selection. Such insight could be gained by comparing how the interfacial free energy of competing polymorphs depends on solvents or additives. We note that computing the interfacial free energy for nuclei of a specific polymorph would require the use of polymorph-specific nucleus size coordinates as described by Peters [9] to preserve the desired structure during mitosis. Simulation results could also be compared to experiments using recently developed techniques by Kim et al. [57] to measure the interfacial free energy of nano-sized crystals. Currently, these techniques still require crystals that are larger than is computationally feasible for atomistic simulations.

We provide an example of the mitosis method in the Potts lattice gas (PLG) model of crystal nucleation from solution [10]. For this lattice model, it is also possible to estimate the interfacial free energy by the conventional approach of first computing the free energy as a function of nucleus size and then fitting the free energy profile to the CNT form. We demonstrate that the two methods give nearly the same estimate for the interfacial free energy. We emphasise that, of these two methods, only mitosis is currently feasible for simulating nucleation from a solution with explicit solutes and solvents.

2. Theory

In CNT, the free energy to form a nucleus of n monomers in an incompressible phase is given by

$$\Delta F(n) = n\Delta\mu + n^{2/3}\phi\sigma, \tag{1}$$

where $\Delta F(n)$ is the free energy, n is the number of particles in the nucleus, $\Delta\mu$ is the chemical potential difference between phase of the embryo and the bulk phase, σ is the bulk interfacial tension between the two phases and ϕ is a shape factor for the nucleus. Qualitative aspects of Equation (1) are supported by nucleation experiments. For example, the experimentally verified Lifshitz–Slyozov law for ripening in solution arises naturally from the $\phi\sigma n^{2/3}$ term in Equation (1) [58,59]. Simulations that compute the nucleation barriers as a function of nucleus size also suggest agreement with the general form of Equation (1), as shown in Figure 1 [3–7,10,11]. However, simulation results suggest that bulk thermodynamic properties $\Delta\mu$, σ and ϕ should be replaced by effective properties at the nanoscale [3,11,19]. After adopting effective values of $\Delta\mu$ and $\sigma\phi$, Equation (1) reduces to

$$\beta F_{\mathrm{CNT}}(n) = \mu_{\mathrm{eff}} n + \gamma_{\mathrm{eff}} n^{2/3}, \tag{2}$$

where μ_{eff} is the effective driving force, γ_{eff} is an effective surface energy density and both of these quantities have been normalised by $k_{\mathrm{B}}T$.

The mitosis method is motivated by the observation that, according to CNT, splitting a nucleus of size $2n$ into two smaller equally sized nuclei of size n requires only work against the interfacial tension:

$$\begin{aligned}\beta\Delta F &= 2\beta F_{\mathrm{CNT}}(n) - \beta F_{\mathrm{CNT}}(2n) \\ &= \gamma_{\mathrm{eff}}(2 - 2^{2/3})n^{2/3}.\end{aligned} \tag{3}$$

Thus, the free energy change from splitting a large nucleus into two smaller sub-nuclei of equal size would provide an estimate of the effective interfacial free energy for a nanoscale nucleus.

Equation (3) suggests a simple and feasible procedure for estimating the interfacial free energy. As the nucleus is split into two equal sub-clusters, it is advantageous to constrain these sub-clusters to a common axis for reasons explained in the Appendix. With r defined as the centre of mass distance between each sub-cluster, $F(r)$ can be

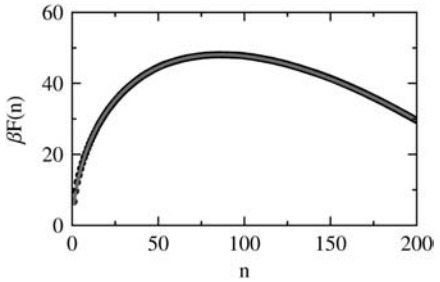

Figure 1. Free energy curve of PLG for $k_{\mathrm{B}}T = 0.8$, $f_1/f_0 = 2.25$, $A = 1.0$, $A' = 0.0$, $K = 25/24$, $K' = 1.0$ and $Q = 24$. The black points represent simulation results, and the grey curve is a least-squares fit of Equation (2). The interfacial free energy from the CNT fit is $\gamma_{\mathrm{eff}} = 7.4$.

computed as a potential of mean force along r. The sub-clusters will still form a single nucleus when r is small, as r grows the sub-clusters will separate and become independent nuclei. $F(r)$ is analogous to an activation barrier and can be directly related to the free energy difference in Equation (3):

$$\beta\Delta F = \frac{1}{L}\int_{\cup} \mathrm{e}^{-\beta F(r)}\,\mathrm{d}r, \tag{4}$$

where L is the unit of length, and the integral over $\cup$ is the configurations at low free energies where the two sub-clusters are conjoined. Combining Equations (3) and (4) gives a direct equation to calculate the interfacial free energy:

$$\gamma_{\mathrm{eff}} = [(2-2^{2/3})n^{2/3}]^{-1}\ln\left\{\frac{1}{L}\int_{\cup} \mathrm{e}^{-\beta F(r)}\,\mathrm{d}r\right\}. \tag{5}$$

Equation (5) can also be related to the effective surface tension (σ_{eff}). For example, if the nuclei are approximately spherical,

$$\gamma_{\mathrm{eff}} n^{2/3} = 4\pi R^2 \beta\sigma_{\mathrm{eff}}, \tag{6}$$

where R is the droplet radius.

A key assumption of the mitosis method is that γ_{eff} is constant with nucleus size. However, in the case of faceted nuclei, γ_{eff} may vary with nucleus size. In the case of faceted NaCl crystal nucleation from the melt studied by Zykova-Timan et al. [22], γ_{eff} did not vary significantly with nucleus size for nuclei larger than $\sim$25. Therefore, the method proposed here may also work for a faceted system if n is sufficiently large.

One challenge with the mitosis method is to *reversibly* split the nucleus so that the pair potential $F(r)$ can be computed accurately without any hysteresis effects. We propose the use of a minimum bisection coordinate to reversibly cut the nucleus into two parts. To find the minimum bisection, we divide the nucleus into two arbitrary initial sub-clusters of equal size. The bisection coordinate q is a continuous measure of the number of contacts between these two sub-clusters:

$$q = \sum_{i\neq j}(1-\delta_{c(i),c(j)})\min(1,(r_{ij}/r_c)^{-4}), \tag{7}$$

where the sum is over sites i in the first cluster and j in the second cluster, and where r_c is a typical distance for monomer contacts. In Equation (7), $c(i)=1$ if site i is in the first sub-cluster, and $c(i)=2$ if site i is in the second sub-cluster. The bisection coordinate q depends on the coordinates of all atoms in the cluster and also on how the cluster is partitioned into two sub-clusters. The bisection coordinate is large when the cluster is partitioned in a manner that gives a large interface between the two sub-clusters, and small when the interface between the sub-clusters is small. The cluster partition that minimises the interface between the two sub-clusters is found by an MC simulation using q as an effective energy to sample different partitions of a fixed cluster. The optimal partition for a given cluster configuration is identified by a minimal bisection coordinate that we denote as q_{min}. The optimal partition and q_{min} are found by swapping elements between sub-clusters using an MC simulation. The MC move set consists of randomly selecting a site in each of the two sub-clusters and attempting to swap their sub-cluster membership assignments. MC moves are accepted using the Metropolis-like acceptance criteria where q takes the place of energy:

$$P_{\mathrm{acc}}(q\rightarrow q') = \min(1,\exp[-(q'-q)]). \tag{8}$$

The MC simulation is run for 5000 MC sweeps with the minimum q configuration, q_{min} being taken as the optimal sub-cluster partition. Note that q_{min} continuously decays to zero as the nuclei completely separate and move apart into solution.

Splitting or recombining the nucleus by umbrella sampling along q_{min} avoids hysteresis errors that may occur if umbrella sampling [60,61] is performed along the coordinate r. To later obtain the potential of mean force $F(r)$, one should compute joint histograms of (q_{min}, r) while sampling along the q_{min}-coordinate. The coordinate q_{min} can then be integrated to recover an accurate result for $F(r)$.

3. PLG example

The PLG has been used previously to study nucleation from solution [10] and laser-induced nucleation [62]. Here, we apply the mitosis method to determine the surface free energy of a PLG system. The PLG has solvent and solute lattice sites like a lattice gas, with the addition of orientation-dependent interactions. Each lattice site has Q possible orientations, where $Q=24$ for a cubic lattice site with no symmetric interactions. The PLG Hamiltonian is defined as

$$\begin{aligned} H = &-\sum_{<i,j>} m(i)m(j)\{[(K-A/Q)]+\delta_{s(i),s(j)}A\} \\ &-\sum_{<i,j>}(1-m(i))(1-m(j))\{[(K'-A'/Q)] \\ &+\delta_{s(i),s(j)}A'\}, \end{aligned} \tag{9}$$

where $m(i)=1$ if lattice site i is a solute, and $m(i)=0$ if lattice site i is a solvent, and the orientation of lattice site i is given by $s(i)=1,2,3,\ldots$ or Q. The sum over $<i,j>$ is a sum over nearest neighbours. The parameters A and A' control the pure phase melting temperature of solutes and solvents, respectively, and K and K' control the solubility of solutes into solvents and vice versa.

MC simulations are carried out in the semigrand canonical ensemble [41]. Note that since the mitosis method maintains a constant cluster size, the supersaturation is approximately constant, making simulations in the canonical ensemble (NVT) with the same average supersaturation approximately equivalent. The MC move set consists of particle orientation changes and particle identity swaps. Each move type is attempted with equal probability. Orientation moves consist of the random selection of a lattice site, with the selection of a new orientation chosen at random from a uniform distribution of orientations. The move is then accepted or rejected using Metropolis acceptance criteria. Particle identity exchanges are performed in the semigrand canonical ensemble. The acceptance criterion for a solvent to solute exchange is

$$P_{\mathrm{ACC}}[(m(i)=0)\rightarrow(m(i)=1)] = \min\left(\frac{f_1}{f_0}\exp[-\beta(H_{m(i)=1}-H_{m(i)=0})],1\right), \quad (10)$$

where f_1/f_0 is the solvent to solute fugacity ratio and H is the PLG Hamiltonian from Equation (9) for lattice site i changed from a solvent to a solute.

We chose PLG parameters of $A=1.0$, $A'=0$, $K=25/24$, $K'=1.0$, $k_{\mathrm{B}}T=0.8$ and $f_1/f_0=2.25$ to correspond with conditions explored in our previous work which examined temperature effects on the PLG [10]. Note that the parameters reported in Ref. [10] are in error [63]. The nucleation barriers in that work actually correspond to the parameters listed above. This parameter set was found to produce post-critical liquid like solute nuclei [10]. We demonstrate the mitosis method by reversibly dividing a nucleus of size $2n=150$ into two nuclei of size $n=75$. Since nucleus structure remains liquid-like over the whole nucleation barrier, there is no need to constrain the nucleus to a particular structure [9].

We calculate the free energy of the PLG as a function of distance r between the sub-clusters and q_{min}. Umbrella sampling was performed using windows of width two units along the q_{min}-coordinate with hard walls at the window edges [60,61]. Sampling along the r-coordinate was unbiased. Each window was sampled using hybrid MC, where an MC trajectory is run for one sweep, and then the entire trajectory is accepted if the configuration lies within the simulation window. Each window is sampled for 5,000,000 trajectories, after being equilibrated for 100,000 trajectories. The centre of mass of each sub-cluster was constrained to a maximum displacement of one lattice site from the pairwise interaction axis.

Figure 2 shows the free energy surface for r and q_{min} for the PLG. As the interface between sub-clusters contracts, q_{min} decreases, r increases and the free energy increases. Figure 2 also shows the importance of sampling along q_{min} and not sampling along r directly. The nature

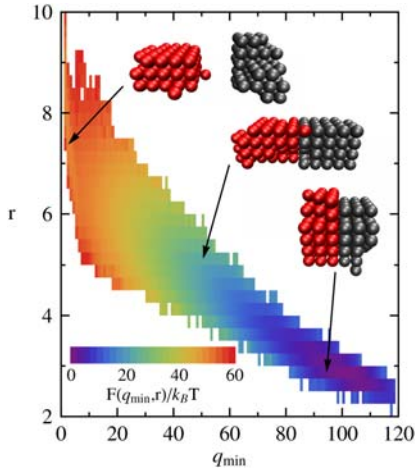

Figure 2. Free energy as a function of r and q_{min}. The inset nuclei are typical configurations at the indicated points. Spheres represent solute lattice sites, and the two colours (grey and red) distinguish the sub-clusters (colour online).

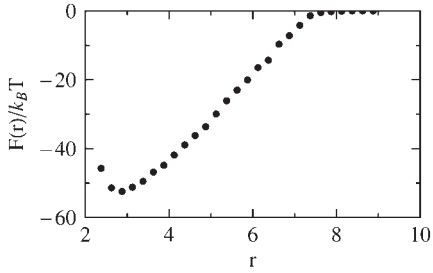

Figure 3. Free energy $F(r)$ vs. r, the centre of mass distance between the sub-clusters. The curve has been shifted such that $F(r)$ decays to 0 with increasing r.

of the free energy landscape near $q_{\mathrm{min}}=6$ would likely result in hysteresis if sampling was performed along the r-coordinate alone. Snapshots of typical nuclei along q_{min} are shown in Figure 2.

The free energy surface shown in Figure 2 can be projected onto a free energy as a function of r by direct integration [60]. Applying Equation (5) to the free energy curve in Figure 3 gives an interfacial free energy parameter of $\gamma_{\mathrm{eff}}=7.0$. For a comparison to the conventional approach for computing γ_{eff}, we fit Equation (2) from CNT to the results in Figure 1. The fit, shown in Figure 1, is effectively indistinguishable from the computed free energy as a function of nucleus size. The interfacial free energy obtained from the CNT fit to $F(n)$ is $\gamma_{\mathrm{eff}}=7.4$. These estimates of γ_{eff} differ by only 6%.

4. Conclusions

Importance of sampling methods are routinely used to compute free energies as a function of nucleus size for nucleation in pure component systems or from multi-

component vapours. Motivated by CNT, these free energy results have been used to estimate the effective driving force and interfacial free energy parameters that are appropriate at the length scale of critical nuclei [4]. However, such analyses have not been possible for crystal nucleation from solution with an explicit solvent. The need to maintain a constant chemical potential driving force and the difficulty of particle insertion into a condensed phase have precluded efforts to compute the free energy as a function of nucleus size for nucleation from solution.

This paper presented a new mitosis method to directly compute the interfacial free energy of nanoscale nuclei. The mitosis method computes the interfacial free energy by reversibly bisecting a seeded nucleus. The method was tested using the PLG model of nucleation from solution. The PLG model enabled a comparison to the conventional approach of first computing the free energy as a function of nucleus size. Fitting the free energy profile to a model based on CNT with empirical chemical potential and interfacial free energy parameters gave an interfacial free energy within 6% of the value obtained by the mitosis method.

Unlike the conventional method, the mitosis method can be used for off-lattice simulations of nucleation in multicomponent systems with an explicit solvent. Because the composition of the surrounding solution does not change during the bisection process, the mitosis method avoids the need for particle insertion. The mitosis method thus provides a practical simulation approach for understanding interfacial free energy effects on nucleation from solution. By using the new approach with polymorph-specific nucleus size coordinates, mitosis simulations may reveal how different solvents and additives influence the interfacial free energy of different polymorphs.

Acknowledgements

We thank Gregg Beckham, Brandon Knott, Mike Doherty, Allan Myerson, Dan Hooks and Daan Frenkel for stimulating discussions about this work. This work was supported by NSF CAREER Award No. 0955502.

References

[1] P.R. ten Wolde and D. Frenkel, *Computer simulation study of gas–liquid nucleation in a Lennard-Jones system*, J. Chem. Phys. 109 (1998), pp. 9901–9918.
[2] P.R. ten Wolde and D. Frenkel, *Enhancement of protein crystal nucleation by critical density fluctuations*, Science 277 (1997), pp. 1975–1978.
[3] S. Auer and D. Frenkel, *Prediction of absolute crystal-nucleation rate in hard-sphere colloids*, Nature 409 (2001), pp. 1020–1023.
[4] S. Auer and D. Frenkel, *Numerical prediction of absolute crystallization rates in hard-sphere colloids*, J. Chem. Phys. 120 (2004), pp. 3015–3029.
[5] C. Valeriani, E. Sanz, and D. Frenkel, *Rate of homogeneous crystal nucleation in molten NaCl*, J. Chem. Phys. 122 (2005), 194501.
[6] A.C. Pan and D. Chandler, *Dynamics of nucleation in the Ising model*, J. Phys. Chem. B 108 (2004), pp. 19681–19686.
[7] K. Brendel, G.T. Barkema, and H. van Beijeren, *Nucleation times in the two-dimensional Ising model*, Phys. Rev. E 71 (2005), 031601.
[8] B. Peters and B.L. Trout, *Obtaining reaction coordinates by likelihood maximization*, J. Chem. Phys. 125 (2006), 054108.
[9] B. Peters, *Competing nucleation pathways in a mixture of oppositely charged colloids: Out-of-equilibrium nucleation revisited*, J. Chem. Phys. 131 (2009), 244103.
[10] N. Duff and B. Peters, *Nucleation in a Potts lattice gas model of crystallization from solution*, J. Chem. Phys. 131 (2009), 184101.
[11] J.M. Leyssale, J. Delhommelle, and C. Millot, *Hit and miss of classical nucleation theory as revealed by a molecular simulation study of crystal nucleation in supercooled sulfur hexafluoride*, J. Chem. Phys. 127 (2007), 044504.
[12] D. Moroni, P.R. ten Wolde, and P.G. Bolhuis, *Interplay between structure and size in a critical crystal nucleus*, Phys. Rev. Lett. 94 (2005), 235703.
[13] S. Auer, C.M. Dobson, and M. Vendruscolo, *Characterization of the nucleation barriers for protein aggregation and amyloid formation*, HFSP J. 1 (2007), pp. 137–146.
[14] S. Auer, C.M. Dobson, M. Vendruscolo, and A. Maritan, *Self-templated nucleation in peptide and protein aggregation*, Phys. Rev. Lett. 101 (2008), 258101.
[15] S. Punnathanam and P.A. Monson, *Crystal nucleation in binary hard sphere mixtures: A Monte Carlo simulation study*, J. Chem. Phys. 125 (2006), 024508.
[16] B. Chen, J.I. Siepmann, K.J. Oh, and M.L. Klein, *Aggregation-volume-bias Monte Carlo simulations of vapor–liquid nucleation barriers for Lennard-Jonesium*, J. Chem. Phys. 115 (2001), 10903–10913.
[17] B. Chen, J.I. Siepmann, and M.L. Klein, *Simulating vapor–liquid nucleation of water: A combined histogram-reweighting and aggregation-volume-bias Monte Carlo investigation for fixed-charge and polarizable models*, J. Phys. Chem. A 109 (2005), pp. 1137–1145.
[18] B. Chen, J.I. Siepmann, K.J. Oh, and M.L. Klein, *Simulating vapor–liquid nucleation of n-alkanes*, J. Chem. Phys. 116 (2002), pp. 4317–4329.
[19] F. Trudu, D. Donadio, and M. Parrinello, *Freezing of a Lennard-Jones fluid: From nucleation to spinodal regime*, Phys. Rev. Lett. 97 (2006), 105701.
[20] B. Chen, H. Kim, S.J. Keasler, and R.B. Nellas, *An aggregation-volume-bias Monte Carlo investigation on the condensation of a Lennard-Jones vapor below the triple point and crystal nucleation in cluster systems: An in-depth evaluation of the classical nucleation theory*, J. Phys. Chem. B 112 (2008), pp. 4067–4078.
[21] B. Chen, J.I. Siepmann, and M.L. Klein, *Simulating the nucleation of water/ethanol and water/n-nonane mixtures: Mutual enhancement and two-pathway mechanism*, J. Am. Chem. Soc. 125 (2003), pp. 3113–3118.
[22] T. Zykova-Timan, C. Valeriani, E. Sanz, D. Frenkel, and E. Tosatti, *Irreducible finite-size effects in the surface free energy of NaCl crystals from crystal-nucleation data*, Phys. Rev. Lett. 100 (2008), 036103.
[23] L. Maibaum, *Phase transformation near the classical limit of stability*, Phys. Rev. Lett. 101 (2008), 256102.
[24] A.R. Browning, M.F. Doherty, and G.H. Fredrickson, *Nucleation and polymorph selection in a model colloidal fluid*, Phys. Rev. E 77 (2008), 041604.
[25] D. Moroni, P.R. ten Wolde, and P.G. Bolhuis, *Interplay between structure and size in a critical crystal nucleus*, Phys. Rev. Lett. 94 (2005), p. 235703.
[26] R. Radhakrishnan and B.L. Trout, *Nucleation of hexagonal ice (Ih) in liquid water*, J. Am. Chem. Soc. 125 (2003), pp. 7743–7747.
[27] S. Punnathanam and D.S. Corti, *Homogeneous bubble nucleation in stretched fluids: Cavity formation in the superheated Lennard-Jones liquid*, Ind. Eng. Chem. Res. 41 (2002), pp. 1113–1121.
[28] C. Desgranges and J. Delhommelle, *Molecular mechanism for the cross-nucleation between polymorphs*, J. Am. Chem. Soc. 128 (2006), pp. 10368–10369.
[29] S.M. Kathmann, G.K. Schenter, B.C. Garrett, B. Chen, and J.I. Siepmann, *Thermodynamics and kinetics of nanoclusters controlling gas-to-particle nucleation*, J. Phys. Chem. C 113 (2009), pp. 10354–10370.

[30] Z.-J. Wang, C. Valeriani, and D. Frenkel, *Homogeneous bubble nucleation driven by local hot spots: A molecular dynamics study*, J. Phys. Chem. B 113 (2009), pp. 3776–3784.
[31] R. Du, V.S. Pande, A.Y. Grosberg, T. Tanaka, and E.S. Shakhnovich, *On the transition coordinate for protein folding*, J. Chem. Phys. 108 (1998), pp. 334–350.
[32] P.G. Bolhuis, D. Chandler, C. Dellago, and P.L. Geissler, *Transition path sampling: Throwing ropes over rough mountain passes, in the dark, Annu*, Rev. Phys. Chem. 53 (2002), pp. 291–318.
[33] J. Wedekind, D. Reguera, and R. Strey, *Finite-size effects in simulations of nucleation*, J. Chem. Phys. 125 (2006), 214505.
[34] B. Widom, *Some topics in the theory of fluids*, J. Chem. Phys. 39 (1963), pp. 2808–2812.
[35] M. Mezei, *A cavity-biased (*T, V, μ*) Monte Carlo method for the computer simulation of fluids*, Mol. Phys. 40 (1980), p. 901.
[36] M. Laso, J.J. de Pablo, and U.W. Suter, *Simulation of phase equilibria for chain molecules*, J. Chem. Phys. 97 (1992), pp. 2817–2819.
[37] G.C.A.M. Mooij, D. Frenkel, and B. Smit, *Direct simulation of phase equilibria of chain molecules*, J. Phys.: Condens. Matter 4 (1992), pp. L255–L259.
[38] M.G. Martin and J.I. Siepmann, *Predicting multicomponent phase equilibria and free energies of transfer for alkanes by molecular simulation*, J. Am. Chem. Soc. 119 (1997), pp. 8921–8924.
[39] W. Shi and E.J. Maginn, *Continuous fractional component Monte Carlo: An adaptive biasing method for open system atomistic simulations*, J. Chem. Theory Comput. 3 (2007), pp. 1451–1463.
[40] J.J. De Pablo and J.M. Prausnitz, *Phase equilibria for fluid mixtures from Monte-Carlo simulation*, Fluid Phase Equilib. 53 (1989), pp. 177–189.
[41] D.A. Kofke and E.D. Glandt, *Monte-Carlo simulation of multicomponent equilibria in a semigrand canonical ensemble*, Mol. Phys. 64 (1988), pp. 1105–1131.
[42] P.G. Debenedetti, *Metastable liquids: Concepts and principles*, Princeton University Press, Princeton, NJ, 1996.
[43] D. Kashchiev, *Nucleation: Basic theory with applications*, Butterworth Heinemann, Oxford, 2000.
[44] G.W. Adams, J.L. Schmitt, and R.A. Zalabsky, *The homogeneous nucleation of nonane*, J. Chem. Phys. 81 (1984), pp. 5074–5078.
[45] J.L. Schmitt, R.A. Zalabsky, and G.W. Adams, *Homogeneous nucleation of toluene*, J. Chem. Phys. 79 (1983), pp. 4496–4501.
[46] M.A. Sharaf and R.A. Dobbins, *A comparison of measured nucleation rates with the predictions of several theories of homogeneous nucleation*, J. Chem. Phys. 77 (1982), pp. 1517–1526.
[47] R. Strey, P.E. Wagner, and T. Schmeling, *Homogeneous nucleation rates for n-alcohol vapors measured in a two-piston expansion chamber*, J. Chem. Phys. 84 (1986), pp. 2325–2335.
[48] J.K. Lee, J.A. Barker, and G.M. Pound, *Surface structure and surface tension: Perturbation theory and Monte Carlo calculation*, J. Chem. Phys. 60 (1974), pp. 1976–1980.
[49] T.S. Jain and J.J. de Pablo, *Calculation of interfacial tension from density of states*, J. Chem. Phys. 118 (2003), pp. 4226–4229.
[50] J.R. Errington and D.A. Kofke, *Calculation of surface tension via area sampling*, J. Chem. Phys. 127 (2007), 174709.
[51] E. Lindahl, B. Hess, and D. van der Spoel, *GROMACS 3.0: A package for molecular simulation and trajectory analysis*, J. Mol. Model. 7 (2001), pp. 306–317.
[52] P.K. Weiner and P.A. Kollman, *AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions*, J. Comput. Chem. 2 (1981), pp. 287–303.
[53] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé, and K. Schulten, *Scalable molecular dynamics with NAMD*, J. Comput. Chem. 26 (2005), pp. 1781–1802.
[54] B.R. Brooks, C.L. Brooks, III, A.D. Mackerell, Jr, L. Nilsson, R.J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A.R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R.W. Pastor, C.B. Post, J.Z. Pu, M. Schaefer, B. Tidor, R.M. Venable, H.L. Woodcock, X. Wu, W. Yang, D.M. York, and M. Karplus, *CHARMM: The biomolecular simulation program*, J. Comput. Chem. 30 (2009), pp. 1545–1614.
[55] S. Plimpton, *Fast parallel algorithms for short-range molecular dynamics*, J. Comput. Phys. 117 (1995), pp. 1–19.
[56] M.A. Lovette, A.R. Browning, D.W. Griffin, J.P. Sizemore, R.C. Snyder, and M.F. Doherty, *Crystal shape engineering*, Ind. Eng. Chem. Res. 47 (2008), pp. 9812–9833.
[57] K. Kim, I.S. Lee, A. Centrone, T.A. Hatton, and A.S. Myerson, *Formation of nanosized organic molecular crystals on engineered surfaces*, J. Am. Chem. Soc. 131 (2009), pp. 18212–18213.
[58] I.M. Lifshitz and V.V. Slyozov, *The kinetics of precipitation from supersaturated solid solutions*, J. Phys. Chem. Solids 19 (1961), pp. 35–50.
[59] O. Krichevsky and J. Stavans, *Ostwald ripening in a two-dimensional system: Correlation effects*, Phys. Rev. E 52 (1995), pp. 1818–1827.
[60] D. Frenkel and B. Smit, *Understanding molecular simulation: From algorithms to applications*, Academic Press, San Diego, 2002.
[61] G.M. Torrie and J.P. Valleau, *Non-physical sampling distributions in Monte-Carlo free-energy estimation – Umbrella sampling*, J. Comput. Phys. 23 (1977), pp. 187–199.
[62] B.C. Knott, N. Duff, M.F. Doherty, and B. Peters, *Estimating diffusivity along a reaction coordinate in the high friction limit: Insights on pulse times in laser-induced nucleation*, J. Chem. Phys. 131 (2009), 224112.
[63] N. Duff and B. Peters, *Erratum: "Nucleation in a Potts lattice gas model of crystallization from solution [J. Chem. Phys. 131 (2009) 184101]"*, J. Chem. Phys. 132 (2010), 129901.
[64] L. Maibaum, *Comment on "elucidating the mechanism of nucleation near the gas–liquid spinodal"*, Phys. Rev. Lett. 101 (2008), 019601.
[65] K.F. Riley, M.P. Hobson, and S.J. Bence, *Mathematical methods for physics and engineering*, Cambridge University Press, Cambridge, 2006.
[66] D. Chandler, *Introduction to modern statistical mechanics*, Oxford University Press, New York, 1987.

Appendix

This appendix shows that an analysis using Poisson statistics or an equilibrium constant with the relative population interpretation of CNT does not isolate γ_{eff} from μ_{eff}. The problem arises from a need to normalise the cluster population distribution in the metastable solution. This normalisation effectively reintroduces a factor that depends on μ_{eff} in the expression for γ_{eff}. The approach outlined in the main text is preferable, but we show this alternative result only to motivate computing $F(r)$ with the two sub-nuclei restricted to a particular line.

As in the discussion above, we assume that CNT is qualitatively correct in predicting the relative populations of cluster sizes, but that effective parameters are needed to replace the macroscopic chemical potential driving force and surface tension parameters [4]. In that case, the average number of clusters of size n in a system of volume V is [23,64]

$$\mathbf{N}(n) = A \exp[-\mu_{\text{eff}} n - \gamma_{\text{eff}} n^{2/3}], \tag{A1}$$

where A is a normalisation constant. If this expression remains true even for clusters as small as isolated monomers ($n = 1$), then we have the following equation for A:

$$\mathbf{N}(1) = A \exp[-\mu_{\text{eff}} - \gamma_{\text{eff}}]. \tag{A2}$$

For a small system, the probability of observing a large nucleus is vanishingly small. Indeed, this is why nucleation requires a long waiting time even at macroscopic system sizes [64]. The Poisson distribution [65] can be used to compute the ratio of the probabilities to observe a single rare nucleus of size $2n$ and the probabilities to observe two rare nuclei of size n:

$$\frac{P[\mathbf{N}(2n) = 1]}{P[\mathbf{N}(n) = 2]} = \frac{\mathbf{N}(2n)\mathrm{e}^{-\mathbf{N}(2n)}}{\mathbf{N}(n)^2\mathrm{e}^{-\mathbf{N}(n)}/2}. \tag{A3}$$

If the system size is small, then $\mathbf{N}(n)$ and $\mathbf{N}(2n)$ are vanishingly small and the exponentials approach unity. The equilibrium ratio of these two cluster sizes in a small system of volume V is approximately

$$\frac{P[\mathbf{N}(2n) = 1]}{P[\mathbf{N}(n) = 2]} \approx \frac{2\mathbf{N}(2n)}{\mathbf{N}(n)^2}. \tag{A4}$$

The ratio in Equation (A4) is a dimerisation equilibrium [66] that can be further simplified using the pair potential $F(r)$

$$\frac{2\mathbf{N}(2n)}{\mathbf{N}(n)^2} = \frac{1}{V}\int_0^{R_{\text{cut}}} \mathrm{e}^{-\beta F(r)} 4\pi r^2 \,\mathrm{d}r. \tag{A5}$$

Here, R_{cut} is the distance at which $F(r)$ becomes approximately zero. The calculation is not sensitive to R_{cut} as long as the minimum in $F(r)$ is many times $k_{\text{B}}T$ in depth. Equation (A1) can be combined with Equation (A5) to give

$$\frac{2}{A}\exp\left[\gamma_{\text{eff}}(2 - 2^{2/3})n^{2/3}\right] = \frac{1}{V}\int_0^{R_{\text{cut}}} \mathrm{e}^{-\beta F(r)} 4\pi r^2 \,\mathrm{d}r. \tag{A6}$$

Finally, inserting the definition of A and solving for γ_{eff} gives

$$\gamma_{\text{eff}} = \{(2 - 2^{2/3})n^{2/3} - 1\}^{-1} \ln\left[\frac{\mathbf{N}(1)\mathrm{e}^{\mu_{\text{eff}}}}{2V}\int_0^{R_{\text{cut}}} \mathrm{e}^{-\beta F(r)} 4\pi r^2 \,\mathrm{d}r\right]. \tag{A7}$$

It is important to note two problems with the calculation presented in this appendix. Equation (A2) which defines A is not likely to hold for very small cluster sizes such as monomers and dimers. Unfortunately, the normalisation constant A enters the final expression for γ_{eff}. This is one of the primary reasons that we recommend the alternative procedure described in the text. Furthermore, the effective interfacial free energy, tension parameter γ_{eff} is not isolated from the effective driving force parameter μ_{eff}.

In some cases, it may be justifiable to omit the term $\mu_{\text{eff}}\{(2 - 2^{2/3})n^{2/3} - 1\}^{-1}$ in comparison to the log of the equilibrium constant. However, the strategy in the text is generally preferable because it does not require this additional approximation and does not require Equation (A3) to hold at very small cluster sizes.