

Water droplet excess free energy determined by cluster mitosis using guided molecular dynamics

Gabriel V. Lau¹, Patricia A. Hunt², Erich A. Müller¹, George Jackson¹ and Ian J. Ford³

¹*Department of Chemical Engineering, Imperial College London,
South Kensington Campus, London SW7 2AZ, U.K.*

²*Department of Chemistry, Imperial College London,
South Kensington Campus, London SW7 2AZ, U.K. and*

³*Department of Physics and Astronomy and London Centre for Nanotechnology,
Gower Street, University College London, London WC1E 6BT, U.K.*

Atmospheric aerosols play a vital role in affecting climate by influencing the properties and lifetimes of clouds and precipitation. Understanding the underlying microscopic mechanisms involved in the nucleation of aerosol droplets from the vapour phase is therefore of great interest. One key thermodynamic quantity in nucleation is the excess free energy of cluster formation relative to that of the saturated vapour. In our current study, the excess free energy is extracted for clusters of pure water modelled with the TIP4P/2005 potential using a method based on nonequilibrium molecular dynamics and the Jarzynski relation. The free energy change associated with the “mitosis” or division of a cluster of N water molecules into two $N/2$ sub-clusters is evaluated. **This methodology is an extension of the disassembly procedure used recently to calculate the excess free energy of argon clusters [Phys. Rev. E 91, 023308 (2015)].** Our results are compared to the corresponding excess free energies obtained from classical nucleation theory (CNT) as well as internally consistent classical theory (ICCT). The values of the excess free energy that we obtain with the mitosis method are consistent with CNT for large cluster sizes but for the smallest clusters, the results tend towards ICCT. For intermediate sized clusters, we obtain values between the ICCT and CNT predictions. Furthermore, the curvature-dependent surface tension is obtained by regarding the clusters as spherical droplets of bulk density and found to be a monotonically increasing function of cluster size for the studied range. The results are compared to other values reported in the literature, agreeing qualitatively with some but disagreeing with the values determined by Joswiak *et al.* [J. Phys. Chem. Lett. 4, 4267 (2013)] using a biased mitosis approach; an assessment of the differences is the main motivation for our current study.

I. INTRODUCTION

Atmospheric aerosols play a key role in climate, weather, pollution and human health. These substances can influence the optical properties of clouds as well as their lifetimes and consequently have an effect on the distribution of precipitation and the global radiation budget (the difference between solar energy accumulated by the earth and the energy radiated into space). Aerosols can give rise to urban pollution episodes [1, 2] and are thought to be responsible for significant morbidity in vulnerable populations [3–6]. There is therefore strong motivation to understand the underlying molecular mechanisms involved in atmospheric nucleation, namely the formation of condensed phase molecular clusters from the ambient metastable vapour phase [7–10]. Important progress has been made in recent years in identifying the species responsible for the initial stages of aerosol formation in the atmosphere [11–15] but the task of properly characterising the formation mechanism in detail still remains, particularly understanding the population dynamics and thermodynamics of the molecular clusters in question.

The most important quantity in nucleation modelling within a traditional kinetic/thermodynamic framework is the *excess free energy* which is often interpreted as the cost in free energy associated with the formation of the

interface between the two phases. In classical nucleation theory (CNT) of a liquid cluster forming from a super-saturated vapour, the excess free energy F_s is one of the two terms that comprise the work of cluster formation ϕ :

$$\phi = F_s - N\Delta\mu, \quad (1)$$

where N is the number of molecules in the cluster and $\Delta\mu$ is the difference in chemical potentials (which is related to the vapour supersaturation) between the condensed and vapour phases. The second term is the “bulk” contribution corresponding to the thermodynamic driving force $\Delta\mu$. In traditional CNT, the excess free energy is represented by the capillarity approximation, namely $F_s^{\text{CNT}} = \gamma_\infty A$, where γ_∞ is the vapour-liquid interfacial tension and A is the surface area of a cluster regarded as a spherical liquid drop with a constant density matching that of the bulk condensed phase. A more general framework extending CNT can be devised from underlying microscopic models, such as the introduction of a size-dependent surface tension [16], allowing for models which do not assume the vapour to be ideal [17, 18] and mean-field kinetic theories [19].

The CNT framework leads to a relationship for the nucleation rate J which has a sensitive dependence on the surface tension:

$$J = K \exp \left(-\frac{16\pi\gamma_\infty^3 v_l^2}{3(\Delta\mu)^2 k_B T} \right), \quad (2)$$

where K is the kinetic prefactor that includes the aggregation rate of particles to the cluster, k_B is the Boltzmann constant, T is the temperature, and v_l is the volume per particle in the condensed phase.

In fact, the surface tension of curved interfaces is known not to be constant, and its size-dependence has been the focus of recent attention [20, 21]. A variety of different methods has been employed for this purpose, ranging from evaluation of the components of the pressure tensor [22–24] and thermodynamic perturbation methods [25–27] to techniques involving a finite-size scaling analysis in the grand canonical ensemble [28–30]. Consistency in the representation of the interfacial properties has not yet emerged in all cases. In some studies of water droplets, for example, the tension is found to be a monotonically increasing function of the droplet size until the planar limit is reached, but even in cases where the qualitative trend is the same, the cluster size where the tension essentially reaches the planar value differs from study to study [31–34]. It should be noted, however, that different models of water have been employed, and furthermore the various studies were conducted under slightly different conditions, so that a direct comparison is not always possible. In stark contrast, Joswiak *et al.* [35] have predicted that the surface free energy of clusters of water molecules for the popular TIP4P/2005 model [36] will increase monotonically from the planar limit with decreasing cluster size. These authors employ the “mitosis method” using a hybrid molecular dynamics (MD)/Monte Carlo (MC) technique involving umbrella sampling to evaluate the free energy associated with splitting a water cluster into two equal sized sub-clusters. This free energy is then related to the vapour-liquid interfacial tension of the clusters. It is this recent investigation which motivates our current study.

Our aim is to re-assess the unusual nature of cluster-size dependence obtained by Joswiak *et al.* [35] using the same TIP4P/2005 molecular model, a similar mitosis procedure, but a different fundamental approach to computing the thermodynamic properties. In our current work, we use a combination of nonequilibrium molecular dynamics (NEMD) and the Jarzynski equality [37] to extract the excess free energies. The Jarzynski equality is a simple but powerful formula relating an equilibrium change in the free energy to the mechanical work performed during the nonequilibrium simulations. We carry out mitosis by dynamically driving the separation process and relating the extracted difference in free energy to the excess free energy of the initial and final clusters. The vapour-liquid interfacial tension can then be estimated by dividing the excess free energy by an area commensurate with the representation of the clusters as spherical drops with a constant density equal to that of the bulk condensed phase.

In Sections II A to C, we describe the mitosis method and how the work calculated from the simulations can be related to the free energy difference through the Jarzynski equality. An analysis relating the free energy of mitosis extracted from the simulations to the excess free energy of the free cluster taking into account the perturbation due to the presence of the mechanical driving forces is made in Sections II D and E. Further details of the analysis are provided in Appendix A. In Section III, we outline the procedure employed for the NEMD simulations. The results and discussion are presented in Section IV, and the conclusions of our work are given in Section V.

II. CLUSTER SURFACE FREE ENERGIES FROM GUIDED MOLECULAR DYNAMICS

A. Guided Molecular Dynamics

The method we employ for evaluating cluster free energies involves the disassembly of a cluster via nonequilibrium molecular dynamics (MD) simulations, originally developed for studies of argon clusters [38]. The technique involves tethering every cluster molecule to an artificial “guide” particle using a harmonic potential. The guide particles are then arranged to move apart at constant speeds, driving cluster disassembly until a desired final state is reached. In our current study, the cluster is separated into two sub-clusters of equal size, corresponding to a mitosis process (cf. Figure 1). The strength of the harmonic potentials can be varied throughout the course of the simulation. This allows the cluster to be initially weakly affected by their presence, ensuring minimal perturbation of the initial cluster structure. Towards the end stages of the simulation, the springs can be strengthened to ensure that the mitosis is completed. The work done during the mitosis process can then be related to the change in free energy.

It is important to mention that the computation of the change in free energy associated with the separation of a bulk liquid into two (or more) liquid slabs separated by vapour has a long history. In the early days of the development of molecular simulation techniques, Mijazaki *et al.* [39] determined the excess surface free energy of a Lennard-Jones (LJ) fluid with a thermodynamic procedure involving the separation of a homogeneous liquid state into two slabs separated by two vapour-liquid interfaces; a step-wise procedure was employed by introducing an external hard-wall potential to decouple the liquid slabs which were then allowed to relax to form the free interfaces. The vapour-liquid surface tension could then be estimated from the cost in free energy associated with the formation of the interfaces.

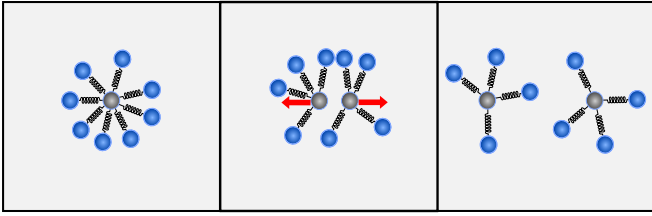


Figure 1. The cluster mitosis process. Guide particles and the cluster molecules are indicated in grey and blue, respectively. Initially, the cluster molecules are weakly tethered to the guide particles placed at the origin. The guide particles are then made to drift apart until the cluster is separated into two distinct sub-clusters.

B. Work of Mitosis

In general, when the Hamiltonian H is a function of an externally controlled parameter λ , the work done W on a system over some sampling time interval τ is

$$W = \int_0^\tau \frac{d\lambda}{dt} \frac{\partial H(\lambda)}{\partial \lambda} dt. \quad (3)$$

For a cluster of N molecules tethered to their corresponding guide particles, the Hamiltonian consists of intermolecular interactions between the cluster molecules as well as the harmonic interactions between the molecules and their guides. The total Hamiltonian is therefore a function of the time-dependent spring constant $\kappa(t)$, the molecular positions $\{\mathbf{x}_i(t)\}$ and the guide positions $\{\mathbf{X}_i(t)\}$. The total work can be expressed as

$$\begin{aligned} W &= \int_0^\tau \frac{d\kappa(t)}{dt} \frac{\partial H}{\partial \kappa} dt + \sum_{i=1}^N \int_0^\tau \frac{d\mathbf{X}_i(t)}{dt} \frac{\partial H}{\partial \mathbf{X}_i} dt \\ &= \frac{1}{2} \int_0^\tau \frac{d\kappa(t)}{dt} \sum_{i=1}^N (\mathbf{x}_i(t) - \mathbf{X}_i(t))^2 dt \\ &\quad - \int_0^\tau \kappa(t) \sum_{i=1}^N (\mathbf{x}_i(t) - \mathbf{X}_i(t)) \cdot \frac{d\mathbf{X}_i}{dt} dt, \end{aligned} \quad (4)$$

where $d\mathbf{X}_i/dt$ is the velocity of the guide particle tethered to molecule i . The first term corresponds to the work done in tightening or loosening the spring constant over the course of the simulation, and the second term is the work done on the molecules by the moving guide particles through the springs.

C. Jarzynski Equality

The work performed in the simulations needs to be related to the free energy change between the initial and final states. According to the first and second laws of

thermodynamics, the work done on a system as its Hamiltonian evolves is greater than or equal to the corresponding change in Helmholtz free energy: $W \geq \Delta F$. Only in the case of an infinitesimally slow and reversible (quasistatic) process will the work be equal to the free energy difference between the initial and final states. However, running simulations at the quasistatic limit in molecular dynamics is not practicable. On the other hand, one can perform many fast process simulations and estimate the free energy change from the average of the work performed. Nevertheless, as a consequence of the second law, the outcome can only be at best an upper bound to the free energy difference: $\langle W \rangle \geq \Delta F$.

One solution to this problem is provided by the approach of Jarzynski [37], who derived a nonequilibrium relationship, which is commonly known as the Jarzynski equality:

$$\Delta F = -k_B T \ln \left[\left\langle \exp \left(-\frac{W}{k_B T} \right) \right\rangle \right], \quad (5)$$

where the free energy difference between the initial and final states can be related *exactly* to an ensemble average of the Boltzmann factor of the work. The initial state needs to be represented in the canonical ensemble, but no conditions are placed on the statistics of the final state after the work process. Since the Jarzynski equality can be used to relate the change in free energy to the work performed in nonequilibrium processes (far from the quasistatic limit), it is particularly suitable for use in conjunction with MD simulations. It is important to note, however, that because of limited sampling, the computed ΔF may depend on the rate of processing: slower processes will give rise to a narrower work distribution and hence an estimate of the free energy change with lower statistical error than that derived from faster processes with a broader distribution of work.

D. Contributions to the Excess Free Energy

The free energy difference extracted from the MD simulations correspond to that associated with the mitosis of *tethered* clusters. This quantity can be related to the excess free energy difference for the mitosis of a free (undistorted) cluster using a statistical thermodynamic analysis (details are given in Appendix A). We find that the difference in excess free energy between two $N/2$ sub-clusters and the parent N cluster (with N even) can be written as a sum of four terms:

$$2F_s(N/2) - F_s(N) = \sum_{i=1}^4 f_s^i = f_s^{\text{tot}}, \quad (6)$$

with

$$\begin{aligned}
f_s^1 &= \Delta F_{\text{mit}} \\
f_s^2 &= -k_B T \ln \left[\frac{N!}{[(N/2)!]^2} \right] \\
f_s^3 &= k_B T \ln \left[\frac{(4\pi)^{3/2} \rho_{vs} v_{\text{HO}}}{(N/2)^{3/2}} \right] \\
f_s^4 &= (N/2)^{5/3} \kappa_i \frac{3}{5} \left(\frac{3v_l}{4\pi} \right)^{2/3} (2^{2/3} - 1),
\end{aligned} \tag{7}$$

where ρ_{vs} is the saturated vapour density, κ_i is the initial spring constant, $v_{\text{HO}} = (k_B T / \kappa_i)^{3/2}$, and v_l is the volume per particle in the condensed phase. The first term f_s^1 corresponds to the change in free energy associated with the mitosis of a tethered cluster, as obtained from MD simulations. The second term f_s^2 accounts for the correction that relates calculations derived from an MD simulation with distinguishable particles to a system of indistinguishable particles. The third term f_s^3 is an entropic contribution associated with correcting for the confinement of the centre of mass of the cluster in the initial state, and the centres of mass of the two clusters in the final state, through the tethering process. The final term f_s^4 corrects for the perturbation in the cluster energy, before and after mitosis, caused by the presence of the tethers.

E. Surface Tension from Mitosis

The excess free energy difference obtained from the mitosis procedure can be related to the surface tension. A similar analysis has been previously performed by Joswiak *et al.* [35]. For a system with a planar interface, the thermodynamic surface tension is defined as the excess free energy per unit area $\gamma = F_s / A$ for an appropriate choice of dividing surface. For small liquid clusters, linking the excess free energy associated with the formation of the surface to the tension requires a similar analysis but has the further issue of specifying an appropriate density of the condensed phase. In this study, we introduce a size-dependent surface free energy $\gamma(N)$ while making the assumption that clusters are spherical (with a well defined radius and constant density corresponding to that of the bulk liquid phase), such that we can write the work of formation of a cluster of N particles from the vapour as the sum of a surface term and a volume term:

$$\phi(N) = 4\pi[R(N)]^2 \gamma(N) - \frac{4}{3}\pi[R(N)]^3 \rho_l \Delta\mu, \tag{8}$$

where $R(N)$ is the radius of the cluster and $\rho_l = 1/v_l$ is the bulk density of the condensed phase. The density corresponding to a bulk liquid phase at the same chemical potential as a coexisting vapour phase would be a more rigorous choice but it is more difficult to implement such a procedure; this would be consistent with the

formal thermodynamic approaches of Gibbs [40] and Tolman [16]. We take this simple approach to make qualitative comparisons of the size-dependent tension with other approaches in the literature but we are aware of the difficulties in assigning the macroscopic concept of a surface tension to small systems. To calculate the surface tension, Joswiak *et al.* used the expression of Tolman [16] relating the surface free energy to the surface curvature of large droplets. The Tolman expression can be useful for larger clusters but it is expected to break down at smaller length scales and we opt not to employ it in our current work.

With the excess free energy of a N molecule cluster written in the form of a surface term $F_s = 4\pi[R(N)]^2 \gamma(N)$ and employing the incompressibility condition $R(N) = 2^{1/3} R(N/2)$, then the excess free energy difference extracted from the mitosis procedure can be expressed as

$$\begin{aligned}
2F_s(N/2) - F_s(N) &= \\
&= 8\pi[R(N/2)]^2 \gamma(N/2) - 4\pi[R(N)]^2 \gamma(N) \\
&= 8\pi[R(N/2)]^2 \gamma(N/2) - 4\pi 2^{2/3} [R(N/2)]^2 \gamma(N) \\
&= 8\pi[R(N/2)]^2 [\gamma(N/2) - 2^{-1/3} \gamma(N)].
\end{aligned} \tag{9}$$

In order to progress from here, we assume the planar limit is valid for the largest cluster studied (of size N_{max}), allowing us to insert a value for $\gamma(N_{\text{max}})$ in Eq. (9). We check the validity of this assumption by ensuring that the calculated surface tensions of the N_{max} cluster and the $N_{\text{max}}/2$ cluster differ negligibly. The surface tensions of the smaller clusters can then be calculated using Eq. (9) from the excess free energy obtained through simulation.

III. METHODOLOGY

Molecular dynamics simulations are carried out for the mitosis of clusters for $N = 6, 12, 24, 48, 76, 96$ and 128 TIP4P/2005 [36] water molecules. Simulations are performed in the canonical ensemble in a nonperiodic box at a temperature of 300 K maintained using a Langevin thermostat [41] with a friction constant of 1 ps^{-1} . For each cluster, 1000 initial configurations are sampled from an equilibrium trajectory of duration 10 ns with the guide particles fixed at the origin. An initial spring constant $\kappa_i = 0.05 \text{ kJ mol}^{-1}$ is used, deemed to be minimally perturbing since the mean tether energy of a molecule for all cluster sizes is then less than the typical thermal energy $k_B T$. The radii of the clusters are calculated based on the bulk liquid density (of the TIP4P/2005 model) at the prevailing temperature.

Starting from the initial configurations, half the guide particles are made to drift at a constant speed in the positive direction parallel to one of the Cartesian axes while the other half drift in the negative direction. The spring constant is time-dependent, smoothly varying from the

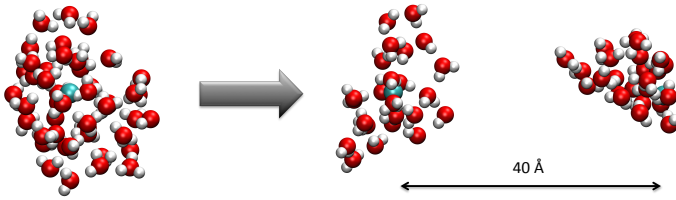


Figure 2. Illustration of the mitosis of a $N = 48$ water cluster into two $N/2 = 24$ clusters over a separation time of 10.24 ns. The guides are indicated by the turquoise coloured particles.

initial value $\kappa_i = 0.05 \text{ kJ mol}^{-1}$ to a final value of $\kappa_f = 0.5 \text{ kJ mol}^{-1}$. The time-dependent form is

$$\begin{aligned} \kappa(t) &= \kappa_i & t \leq t_i \\ \kappa(t) &= \kappa_i + \frac{\kappa_f - \kappa_i}{2} \left[1 - \cos \left(\pi \frac{t - t_i}{t_f - t_i} \right) \right] & t_i \leq t \leq t_f \\ \kappa(t) &= \kappa_f & t \geq t_f \end{aligned}$$

where t_i is the time when the spring constant starts changing and t_f is the time when the spring constant reaches its final value. In our simulations, the values t_i and t_f are pre-set as the following fractions of the simulation time: $t_i = t_{\text{sim}}/5$ and $t_f = 4t_{\text{sim}}/5$ and we use a total simulation time of $t_{\text{sim}} = 10.24 \text{ ns}$. **Further details on the justification of our particular choices of κ_i and t_{sim} are given in Appendix B.**

The guide particles drift until the two sets of guide particles are separated by a distance of 40 Å as illustrated in Figure 2. At this point, the guide particle velocities are reduced to zero and the sub-clusters are allowed to relax while the tether spring constants are loosened back to their initial value κ_i over a further period of t_{sim} .

The differences in free energy obtained from simulation using the Jarzynski relation are converted to excess free energy following the analysis presented in Section II. The saturated vapour density ρ_{vs} and the molecular volume v_l of TIP4P/2005 water at 300 K are required as inputs. The corresponding values of $\rho_{vs} = 5.642 \times 10^{-6} \text{ g cm}^{-3}$ and v_l calculated from the density $\rho_l = 0.9965 \text{ g cm}^{-3}$ are taken from an extensive study of the properties of TIP4P-like models by Vega *et al.* [42, 43].

IV. RESULTS

An example of the distribution of work obtained for the mitosis of the $N = 48$ cluster is depicted in Figure 3. The distributions for all clusters exhibit a smooth, approximately Gaussian form suggesting that the statistics are good enough to extract accurate free energy differences. Values for the excess free energy for different cluster sizes obtained from a sum of the four contributions f_s^i are shown in Table I. The vapour-liquid surface tensions derived from the procedure described in Section II

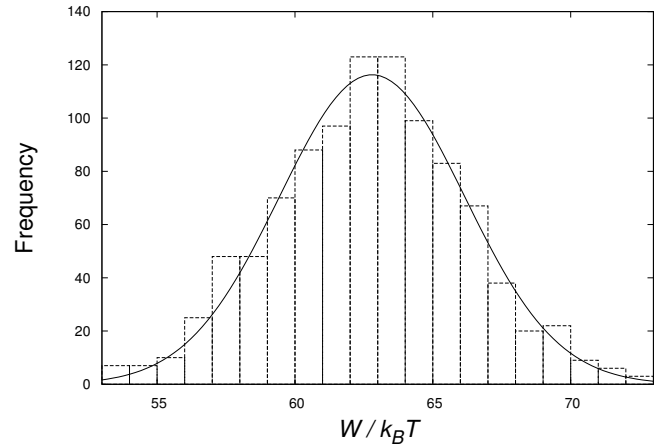


Figure 3. The distribution of work (1000 trajectories) for the mitosis of a $N = 48$ cluster into two $N/2 = 24$ clusters at 300 K with a separation time of 10.24 ns.

N	f_s^1	f_s^2	f_s^3	f_s^4	f_s^{tot}	$\gamma/\text{mN m}^{-1}$
3						41.5
6	10.41	-3.00	-7.47	0.16	0.11	52.0
12	17.86	-6.83	-8.51	0.52	3.04	60.3
24	33.05	-14.81	-9.55	1.65	10.34	65.0
48	57.08	-31.10	-10.59	5.24	20.63	68.0
76	81.75	-50.28	-11.28	11.27	31.46	69.2
96	97.75	-64.03	-11.63	16.64	38.73	69.3
128	119.07	-86.07	-12.06	26.88	47.82	69.7

Table I. Cluster size, the four contributions to the excess free energy in Eq. (6), the total excess free energy, all given in units of $k_B T$, and the estimated surface tension γ .

are also presented in Table I. The surface tension for the $N = 96$ cluster is set to that for the planar vapour-liquid interface ($\gamma_\infty = 69.3 \text{ mN m}^{-1}$ for this model including long range corrections [43]) and serves as an initial input in Eq. (9) for the surface tensions of the smaller clusters. The results for $N = 76$ and $N = 128$ are obtained using interpolated values for γ at $N = 38$ and $N = 64$ (derived from the results for N in the range 24 to 96) and using them as inputs in Eq. (9) to obtain γ at $N = 76$ and $N = 128$, since these sizes were those studied by Joswiak *et al.* [35]. The interpolated values for the tension are obtained by fitting a cubic spline between the data points at $N = 24, 48$ and 96.

A plot of the surface tension normalised by $\gamma(N = 96)$ is depicted in Figure 4. It can be seen that this ratio is close to unity for $N = 48$ (there is a small difference of $\sim 2\%$ in the value between $N = 48$ and $N = 96$), suggesting that the tension has almost reached the planar limit for this cluster size range. This justifies our choice of setting the tension at $N = 96$ to the planar value although we are aware that in principle the planar value can only be reached in the limit

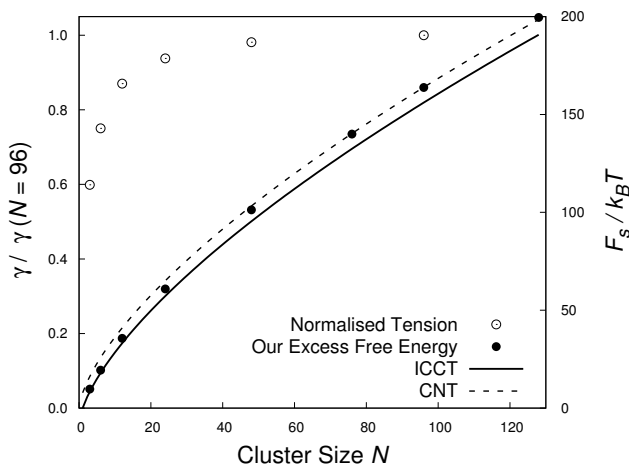


Figure 4. The normalised surface tensions $\gamma(N)/\gamma(N = 96)$ (○) and the excess free energies F_s (●) as a function of cluster size N at 300 K. The dashed and continuous curves represent the corresponding excess free energy obtained from classical nucleation theory (CNT) and internally consistent classical nucleation theory (ICCT), respectively.

$N \rightarrow \infty$. The excess free energy F_s as a function of cluster size N is also illustrated in Figure 4. In addition, we present curves of the excess free energy obtained from classical nucleation theory (CNT) and internally consistent classical theory (ICCT) [44], $F_s^{\text{CNT}} = \gamma_\infty A(N)$ and $F_s^{\text{ICCT}} = \gamma_\infty [A(N) - A(1)]$, respectively. In these cases, the tension is assumed to be that of a planar vapour-liquid interface of the TIP4P/2005 system at 300 K (i.e. 69.3 mN m^{-1}). Our values for the excess free energy are therefore consistent with CNT for the larger cluster sizes. For clusters of $N = 3$ and $N = 6$ particles our results tend towards the ICCT curve. CNT is known to break down at small cluster sizes since it fails to predict a zero excess free energy at $N = 1$. ICCT addresses this inconsistency by shifting the CNT free energy curve down by a constant so that $F_s = 0$ at $N = 1$. For $N = 12$ and $N = 24$, our values lie between the CNT and ICCT curves, indicating that there is an intermediate length scale corresponding to a transition from ICCT to CNT.

The vapour-liquid surface tension estimated for the drops of TIP4P/2005 water as a function of cluster size is plotted again in Figure 5 where it is compared with results obtained in other studies. We observe the tension to be a monotonically increasing function of size that approaches the planar limit. The capillarity approximation, namely the assumption that the clusters all have a tension equal to the planar tension, is surprisingly accurate at this temperature, at least down to cluster sizes of about $N = 40$ molecules. After this point, the tension drops off rapidly to about 57% of the planar value for our smallest system (three water molecules). Our results do not agree with those obtained by Joswiak *et al.* [35] who find that the tension increases from the planar limit upon decreasing the size of the cluster. However, we

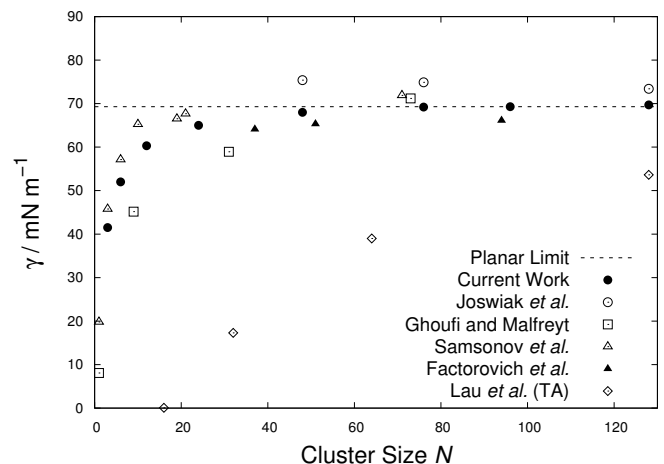


Figure 5. The surface tension of water clusters as a function of cluster size at 300 K. The data obtained in our current guided MD mitosis study (●) for the TIP4P/2005 model are compared with: the free energy of excision simulations of Samsonov *et al.* [32] for the Stockmayer model at 300 K (△); the **volume perturbation simulations** of Ghoufi and Malfreyt [31] for the DPD model at 298 K (□); the vapour pressure simulations of Factorovich *et al.* [33] at 298 K using the mW coarse-grained model and the Kelvin relation to obtain the tension (▲); the test-area (TA) simulations for the TIP4P/2005 model at 293 K [34] (◇); and the mitosis simulations of Joswiak *et al.* [35] for the TIP4P/2005 model at 300 K (○). The dashed line corresponds to the planar limit of the interfacial tension for the TIP4P/2005 model at 300 K [43].

find that our results are in near-quantitative agreement with the values obtained independently by Samsonov *et al.* [32], by Ghoufi and Malfreyt [31], and by Factorovich *et al.* [33]. Samsonov *et al.* used the Stockmayer potential to represent water and equated the change in free energy of formation of a drop to the change in internal energy due to excision from the bulk liquid phase at 300 K. Ghoufi and Malfreyt performed mesoscale simulations of water at 298 K using a dissipative particle dynamics (DPD) model of water and employed a method to calculate the local normal and transverse components of the pressure tensor using volume perturbation. The curvature dependent surface tension was then computed from the pressure components. Factorovich *et al.* took a different approach by analysing the saturated vapour pressure over different sized drops of water represented with the mW coarse-grained model at 298 K. The latter authors found that the macroscopic Kelvin relation held down to clusters of about 40 molecules, suggesting that the tension has little dependence on curvature until the dimensions become very small; the values of the tension reported for Factorovich *et al.* in Figure 5 are deduced from a rearrangement of the Kelvin equation.

In a recent study [34] we applied the test-area method to TIP4P/2005 water clusters consisting of between $N = 16$ and $N = 1000$ molecules at 293 K. In contrast to the other studies discussed thus far, including our cur-

rent work, the tension obtained with the test-area approach was found to depart from the planar limit at significantly larger cluster sizes. Nevertheless, the trend of a decreasing tension at small N is still in qualitative agreement with the other studies. At least at a temperature of $T \approx 300$ K, there is qualitative agreement between most of the existing literature and our current work. The general consensus appears to indicate that the capillarity approximation only starts to break down for clusters with fewer than about 40 particles, at least at 300 K. Care should however be employed in the analysis of macroscopic thermodynamic properties such as the surface tension (or for that matter the average liquid density) for such small clusters.

The differences between the size dependence of the vapour-liquid interfacial tension for water obtained using the test-area (TA) and mitosis methodologies may appear to be rather puzzling. We should, however, acknowledge that the two approaches characterise the free energy associated with a (nominal) surface area of the cluster in rather different ways.

In a macroscopic thermodynamic treatment [46], the interfacial tension for a system in the canonical ensemble can be expressed as the derivative of the Helmholtz free energy with respect to the surface area, $(\partial F/\partial A)_{N,V,T}$. In the case of a system with a curved interface the relation still holds but one has to be careful with the definition of the dividing surface. The surface of tension, useful for a mechanical analysis employing the differential surface tension, is distinct from the equimolar Gibbs dividing surface that most conveniently characterises the surface tension as a specific surface free energy per unit area. If the clusters are large enough, the radii defining these surfaces are similar (the Tolman length is small [21]), and the definition of the surface area is not problematic with the differential and specific surface tensions not differing significantly.

However, for small clusters, an interfacial tension defined using the excess Helmholtz free energy F_s (derived from a microscopic statistical mechanical analysis) will generally differ from the differential surface tension. In the macroscopic thermodynamic analysis [47], the counterpart to F_s is the surface grand potential $\Omega_\sigma = \gamma_E A_E$ where A_E is the equimolar surface area and γ_E is a specific surface free energy defined with respect to this area. Such an analysis tells us that the statistical mechanical approach should produce results consistent with the macroscopic surface tension for large enough clusters, though it unfortunately has little to say about the behaviour of the excess free energy at small sizes.

In the case of the TA method we determine a quantity associated with deforming the area of a mechanically stable (long lived) cluster and relate it to a differential definition of the surface tension. It is a statistical mechanical approach based on an evaluation of $(\partial F/\partial A)_{N,V,T}$ using a distortion of the metric of space as a proxy for differentiation with respect to a surface area [34]. We again expect the results to tend towards the macroscopic sur-

face tension for large clusters, but for smaller sizes, we expect differences to arise. Furthermore, we do not expect coincidence between the resulting differential surface tension and the excess free energy divided by the area of equimolar surface. We properly regard these as distinct cluster properties.

The methods are most appropriate for different ranges of cluster size. Difficulties associated with the TA approach applied to very small drops include the lack of mechanical stability, stabilization effects due to keeping particles within a finite-size simulation cell (potentially distorting the statistics), high uncertainties arising from the proportionality of the free energy of deformation to second order in the distortion parameter (large numerical fluctuations are seen in Figure 11 of reference [34]), and also the non-sphericity of small clusters which can make the method impracticable. This may mean that it is not advisable to employ the methodology for very small clusters.

In the case of the mitosis approach we determine a well defined excess Helmholtz free energy of cluster formation F_s , free of any continuum assumptions, even for small clusters of arbitrary geometry. The only concession to a continuum treatment is that in relating F_s to a surface free energy per unit area, we impose a model $F_s(N) = 4\pi[R(N)]^2\gamma(N)$, and make an assumption of incompressibility of the condensed phase in order to define the equimolar radius $R(N)$. Such a conversion is not necessary in an assessment of cluster stability, and is only done for convenience of presentation. It is noted that the density at the interior of small LJ clusters can be found to be considerably larger than that of the corresponding bulk system [20], indicating some degree of compressibility, though it is not clear that this will be an issue for water at the conditions investigated.

So while we can reasonably assume that the TA and mitosis methods will provide an equivalent estimate of the thermodynamic surface tension for large clusters, for small clusters we are clearly computing different thermodynamic quantities. For example, it is straightforward to show that the TA “surface tension” of a harmonically bound dimer is zero, while the excess free energy of the same cluster is not. Differences will persist for more complex interactions or larger clusters. We regard the TA surface tension as the quantity most analogous to the thermodynamic surface tension, and the excess free energy per unit area as the quantity most relevant for an accurate characterisation of cluster stability.

We collect together results for the size dependence of various formulations of the interfacial tension, normalised by the planar value, in Figure 6 for both LJ and TIP4P/2005 systems. We acknowledge that the different studies are carried out at a variety of temperatures and cutoffs of the potential. Results obtained with the TA method for drops of LJ [27] and TIP4P/2005 [34] water particles appear to follow the same trend (although they are not expected to in general); the water drops are also found to be stable for smaller clusters as might be ex-

pected for a system with stronger intermolecular forces. The behaviour obtained with the disassembly and mitosis methods for the LJ [38] and TIP4P/2005 clusters, respectively, also display close correspondence. There are clear differences between the surface tensions derived by the two approaches for both systems when cluster size is decreased.

Interestingly, results obtained from an FMT-DFT approach [20], based on the excess grand potential, is in line with the results from the mitosis and disassembly simulations, as might be expected of a method based on a computation of specific surface excess quantities. The dependence obtained by Schrader *et al.* [29] for LJ drops, involving a finite-size scaling analysis within the grand canonical ensemble, is somewhere between those obtained with the TA and mitosis approaches. The various methodologies tend to agree for large systems. Unfortunately, we are not able to rationalise the apparent increase in the interfacial tension observed by Joswiak *et al.* although their results also tend towards the planar value in the large R limit [35].

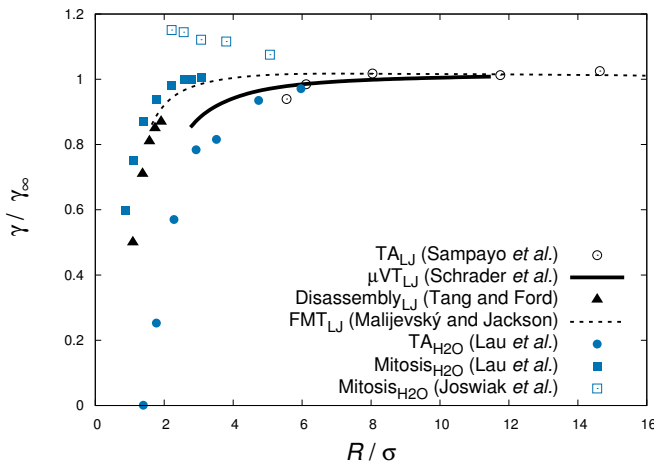


Figure 6. The curvature dependence of the interfacial tension relative to the planar tension for LJ (black symbols and lines) and TIP4P/2005 water clusters (blue symbols and lines). For the LJ systems, we compare: the TA simulations of Sampayo *et al.* [27] at $T^* = 0.8$ ($\odot$); the grand canonical ensemble simulations of Schrader *et al.* [29] at $T^* = 0.68$ (continuous black curve); the non-equilibrium disassembly simulations of Tang and Ford [38] at $T^* = 0.5$ ($\blacktriangle$); and the canonical mean-field FMT-DFT calculations of Malijevský and Jackson [20] at $T^* = 0.8$ (dashed black curve). For TIP4P/2005 water clusters, we compare: our current guided MD mitosis study at 300 K ($\blacksquare$); the TA simulations of Lau *et al.* [34] at $T = 293$ K ($\bullet$); and the MD/MC mitosis simulations of Joswiak *et al.* [35] at $T = 300$ K ($\square$).

V. CONCLUSIONS

We have developed a method based on mitosis by guided molecular dynamics and the use of the Jarzynski equality to extract the excess free energy and estimate the vapour-liquid surface tension of water clusters. Simulations have been carried out for clusters of $N = 6$ to $N = 128$ TIP4P/2005 water molecules at a temperature of 300 K. The size-dependence of the excess free energy has been extracted and compared to the CNT and ICCT models. Our results are consistent with CNT for larger clusters ($N \geq 40$) and with ICCT for the smallest clusters ($N = 3$ and $N = 6$). For intermediate sized clusters ($N = 12$ and $N = 24$), our values transition between the ICCT and CNT curves. The surface tension is observed to be a monotonically increasing function of cluster size, converging to the planar limit for clusters with $N \geq 40$ molecules but reducing significantly in magnitude at the lower end of the size scale. Our results are found not to agree with those obtained by Joswiak *et al.* [35], who carried out a similar mitosis procedure using umbrella sampling; by contrast, these authors found that the tension increased from the planar limit with decreasing cluster size. On the other hand, our results are qualitatively in agreement with other studies [31–34]. We conclude that mitosis simulation in nonequilibrium molecular dynamics appears to be a powerful and efficient method for the determination of thermodynamic properties of small clusters.

ACKNOWLEDGMENTS

Gabriel Lau was supported through a studentship in the Centre for Doctoral Training on Theory and Simulation of Materials managed by the Department of Physics at Imperial College, funded by the Engineering and Physical Sciences Research Council (EPSRC) of the UK under grant number EP/G036888. EAM and GJ acknowledge additional funding from the EPSRC to the Molecular Systems Engineering group through grants EP/E016340 and EP/J014958.

[1] N. Kim Oanh, N. Upadhyay, Y. H. Zhuang, Z. P. Hao, D. Murthy, P. Lestari, J. Villarin, K. Chengchua, H. Co, N. Dung, *et al.*, *Atmos. Environ.* **40**, 3367 (2006).

[2] P. K. Hopke, D. D. Cohen, B. A. Begum, S. K. Biswas, B. Ni, G. G. Pandit, M. Santoso, Y. S. Chung, P. Davy, A. Markwitz, *et al.*, *Sci. Total Environ.* **404**, 103 (2008).

- [3] A. Seaton, D. Godden, W. MacNee, and K. Donaldson, *The Lancet* **345**, 176 (1995).
- [4] J. Sunyer, C. Spix, P. Quenel, A. Ponce-de Leon, A. Pönka, T. Barumandzadeh, G. Touloumi, L. Bacharova, B. Wojtyniak, J. Vonk, *et al.*, *Thorax* **52**, 760 (1997).
- [5] F. Ballester, S. Medina, E. Boldo, P. Goodman, M. Neuberger, C. Iñiguez, and N. Künzli, *J. Epidemiol. Commun. H.* **62**, 98 (2008).
- [6] C. A. Pope III, M. Ezzati, and D. W. Dockery, *New Eng. J. Med.* **360**, 376 (2009).
- [7] D. Kashchiev, *Nucleation* (Butterworth-Heinemann, Oxford, U.K., 2000).
- [8] I. J. Ford, *Proc. Instn Mech. Engrs Part C: J. Mech. Eng. Sci.* **218**, 883 (2004).
- [9] H. Vehkamäki, *Classical nucleation theory in multicomponent systems* (Springer, Berlin, 2006).
- [10] V. Kalikmanov, *Nucleation Theory* (Springer, Dordrecht, 2013).
- [11] J. Duplissy, M. B. Enghoff, K. L. Aplin, F. Arnold, H. Aufmhoff, M. E. Avngaard, U. Baltensperger, T. T. E. Bondo, R. Bingham, K. Carslaw, *et al.*, *Atmos. Chem. Phys.* **10**, 1635 (2010).
- [12] J. Kirkby, J. Curtius, J. Almeida, E. Dunne, J. Duplissy, S. Ehrhart, A. Franchin, S. Gagné, L. Ickes, A. Kürten, *et al.*, *Nature* **476**, 429 (2011).
- [13] M. Noppel, H. Vehkamäki, and M. Kulmala, *J. Chem. Phys.* **116**, 218 (2002).
- [14] J. Merikanto, I. Napari, H. Vehkamäki, T. Anttila, and M. Kulmala, *J. Geophys. Res.-Atmos.* **112** (2007).
- [15] T. Kurtén, V. Loukonen, H. Vehkamäki, and M. Kulmala, *Atmos. Chem. Phys.* **8**, 4095 (2008).
- [16] R. C. Tolman, *J. Chem. Phys.* **17**, 333 (1949).
- [17] A. Laaksonen, I. J. Ford, and M. Kulmala, *Phys. Rev. E* **49**, 5517 (1994).
- [18] J. K. Lee, J. Barker, and F. F. Abraham, *J. Chem. Phys.* **58**, 3166 (1973).
- [19] V. Kalikmanov, *J. Chem. Phys.* **124**, 124505 (2006).
- [20] A. Malijevský and G. Jackson, *J. Phys. Condens. Matter* **24**, 464121 (2012).
- [21] M. Horsch, H. Hasse, A. K. Shchekin, A. Agarwal, S. Eckelsbach, J. Vrabec, E. A. Müller, and G. Jackson, *Phys. Rev. E* **85**, 031605 (2012).
- [22] J. G. Kirkwood and F. P. Buff, *J. Chem. Phys.* **17**, 338 (1949).
- [23] A. Trokhymchuk and J. Alejandre, *J. Chem. Phys.* **111**, 8510 (1999).
- [24] J. Vrabec, G. K. Kedia, G. Fuchs, and H. Hasse, *Mol. Phys.* **104**, 1509 (2006).
- [25] G. J. Gloor, G. Jackson, F. J. Blas, and E. de Miguel, *J. Chem. Phys.* **123**, 134703 (2005).
- [26] J. R. Errington and D. A. Kofke, *J. Chem. Phys.* **127**, 174709 (2007).
- [27] J. G. Sampaio, A. Malijevský, E. A. Müller, E. de Miguel, and G. Jackson, *J. Chem. Phys.* **132**, 141101 (2010).
- [28] K. Binder, *Phys. Rev. A* **25**, 1699 (1982).
- [29] M. Schrader, P. Virnau, and K. Binder, *Phys. Rev. E* **79**, 061104 (2009).
- [30] A. Tröster, M. Oettel, B. Block, P. Virnau, and K. Binder, *J. Chem. Phys.* **136**, 064709 (2012).
- [31] A. Ghoufi and P. Malfreyt, *J. Chem. Phys.* **135**, 104105 (2011).
- [32] V. Samsonov, A. Bazulev, and N. Y. Sdobnyakov, *Dokl. Phys. Chem.*, **389**, 83 (2003).
- [33] M. H. Factorovich, V. Molinero, and D. A. Scherlis, *J. Am. Chem. Soc.* **136**, 4508 (2014).
- [34] G. V. Lau, I. J. Ford, P. A. Hunt, E. A. Müller, and G. Jackson, *J. Chem. Phys.* **142**, 114701 (2015).
- [35] M. N. Joswiak, N. Duff, M. F. Doherty, and B. Peters, *J. Phys. Chem. Lett.* **4**, 4267 (2013).
- [36] J. L. Abascal and C. Vega, *J. Chem. Phys.* **123**, 234505 (2005).
- [37] C. Jarzynski, *Phys. Rev. Lett.* **78**, 2690 (1997).
- [38] H. Y. Tang and I. J. Ford, *Phys. Rev. E* **91**, 023308 (2015).
- [39] J. Miyazaki, J. Barker, and G. Pound, *J. Chem. Phys.* **64**, 3364 (1976).
- [40] J. W. Gibbs, *Am. J. Sci.* **16**, 441 (1878).
- [41] M. P. Allen and D. J. Tildesley, *Computer simulation of liquids* (Oxford university press, 1989).
- [42] C. Vega, J. Abascal, and I. Nezbeda, *J. Chem. Phys.* **125**, 034503 (2006).
- [43] C. Vega and E. De Miguel, *J. Chem. Phys.* **126**, 154707 (2007).
- [44] S. L. Girshick and C. P. Chiu, *J. Chem. Phys.* **93**, 1273 (1990).
- [45] I. J. Ford, *Phys. Rev. E* **56**, 5615 (1997).
- [46] J. S. Rowlinson and B. Widom, *Molecular Theory of Capillarity* (Clarendon Press, Oxford, 1982).
- [47] I. J. Ford, *J. Phys. Chem. B* **105**, 11649 (2001).

APPENDIX A: THERMODYNAMICS OF TETHERED CLUSTER MITOSIS

The change in free energy extracted from the MD simulations is associated with the mitosis of *tethered* clusters. This quantity needs to be related to the excess free energy difference between an untethered (free) cluster of N molecules and two untethered sub-clusters of size $N/2$. We follow a procedure similar to that used recently for cluster disassembly [38].

We start by expressing the free energy of an untethered cluster of N indistinguishable particles as $F_F = -k_B T \ln Z_F$ with

$$Z_F = \frac{1}{N!h^{3N}} \int \prod_{j=1}^N d\mathbf{x}_j d\mathbf{p}_j \exp \left[-\frac{H(\{\mathbf{x}_j\})}{k_B T} \right], \quad (10)$$

where h is Planck's constant, and $H(\{\mathbf{x}_j\})$ is the Hamiltonian of the system, a function of the molecular positions $\mathbf{x}_j$ and momenta $\mathbf{p}_j$. We can write $Z_F = V Z_F^c$ where Z_F^c is the partition function of a free cluster with its centre of mass fixed at the origin, and V is the system volume. This is done by inserting $1 = \int \delta \left(\sum_{i=1}^N (\mathbf{x}_j - \mathbf{x}_c) / N \right)$ into Eq. (10), and recasting the equation in terms of particle coordinates with respect to the centre of mass $\mathbf{x}_c$. The partition function becomes

$$\begin{aligned}
Z_F &= \frac{1}{N!h^{3N}} \int \prod_{j=1}^N d\mathbf{x}'_j d\mathbf{p}_j d\mathbf{x}_c \exp[-\beta H(\{\mathbf{x}'_j\})] \delta \left(\frac{1}{N} \sum_{j=1}^N \mathbf{x}'_j \right) \\
&= \frac{V}{N!h^{3N}} \int \prod_{j=1}^N d\mathbf{x}'_j d\mathbf{p}_j \exp[-\beta H(\{\mathbf{x}'_j\})] \delta \left(\frac{1}{N} \sum_{j=1}^N \mathbf{x}'_j \right) \\
&= V Z_F^c,
\end{aligned} \tag{11}$$

where $\mathbf{x}'_j = \mathbf{x}_j - \mathbf{x}_c$ are the transformed particle coordinates.

The free energy of a tethered cluster is $F_T = -k_B T \ln Z_T$ with

$$\begin{aligned}
Z_T &= \frac{1}{N!h^{3N}} \int \prod_{j=1}^N d\mathbf{x}_j d\mathbf{p}_j \\
&\times \exp \left[-\frac{1}{k_B T} \left(H(\{\mathbf{x}_j\}) + \sum_{j=1}^N \frac{1}{2} \kappa_i x_j^2 \right) \right]. \tag{12}
\end{aligned}$$

In a similar fashion to the case of the free cluster, the partition function of the tethered cluster can be written as $Z_T = [2\pi k_B T / (N\kappa_i)] Z_T^c$ where Z_T^c is the partition function of a cluster with the particles tethered to the origin as well as being constrained to have its centre of mass at the origin:

$$\begin{aligned}
Z_T^c &= \frac{1}{N!h^{3N}} \int \prod_{j=1}^N d\mathbf{x}'_j d\mathbf{p}_j \exp \left[-\frac{H(\{\mathbf{x}'_j\})}{k_B T} \right] \\
&\times \exp \left[-\sum_{j=1}^N \frac{\kappa_i x_j'^2}{2k_B T} \right] \delta \left(\frac{1}{N} \sum_{j=1}^N \mathbf{x}'_j \right). \tag{13}
\end{aligned}$$

Assuming that the mean tethering energy is less than the typical thermal energy scale $k_B T$, the difference in free energy between the free and tethered cluster can be expressed as [38]:

$$F_F(N) - F_T(N) \approx -k_B T \ln[\rho_c^N(0)V] - \frac{N\kappa_i}{2} \int \rho_0^N(\mathbf{y}) y^2 d\mathbf{y}, \tag{14}$$

where $\rho_c^N(0) = (N\kappa_i/2\pi k_B T)^{3/2}$ is the probability density function, evaluated at the origin, of the centre of mass of the N tethered particles in the cluster. $\rho_0^N(\mathbf{y})$ represents the spatial density profile of a single particle in a free cluster of N particles with its centre of mass constrained to the origin.

The cluster excess free energy can be defined as

$$\begin{aligned}
F_s(N) &= \phi(N) - \phi(1) + (N-1)\Delta\mu \\
&= F_F(N) - F(1) - (N-1)\mu_s,
\end{aligned} \tag{15}$$

where $\phi(N)$ is the work of formation of an N -cluster, $\Delta\mu$ is the chemical potential difference between the condensed and vapour phases at a given vapour pressure, $F_F(N)$ is the Helmholtz free energy of an N -cluster, and μ_s is the coexistence chemical potential. **Eq. (15) is subtly different from Eq. (1) in that the the work of formation of a cluster minus that of a monomer is the appropriate quantity that emerges from careful development of nucleation theory [45]. Eq. (1) is commonly taken as an approximation in CNT.** If the vapour phase is considered ideal then the monomer Helmholtz free energy is $F(1) = -k_B T \ln(V/\Lambda^3)$ and the coexistence chemical potential is $\mu_s = k_B T \ln(\rho_{vs}\Lambda^3)$ where ρ_{vs} is the saturated vapour density, and Λ is the thermal de Broglie wavelength. Combining Eqs. (15) and (14), the excess free energy of an N -cluster can now be written as

$$\begin{aligned}
F_s(N) &= F_F(N) + k_B T \ln(V/\Lambda^3) - (N-1)k_B T \ln(\rho_{vs}\Lambda^3) \\
&= F_T(N) - k_B T \ln[\rho_c^N(0)V] - \frac{N\kappa_i}{2} \int \rho_0^N(\mathbf{y}) y^2 d\mathbf{y} \\
&\quad + k_B T \ln(V/\Lambda^3) - (N-1)k_B T \ln(\rho_{vs}\Lambda^3). \tag{16}
\end{aligned}$$

It should be recognized that the difference in free energy obtained from guided molecular dynamics simulations corresponds to a process involving distinguishable molecules. However, we require the free energy difference between a cluster consisting of N indistinguishable molecules and two sub-clusters which are distinguishable from each other but each consist of $N/2$ indistinguishable molecules. To relate the free energy change obtained from simulation to the desired quantity, **we need** to apply the usual classical corrections for indistinguishability. The partition function corresponding to an initial cluster of N indistinguishable molecules is $Z_N = 1/(N!) Z_N^{\text{dist}} = \exp(-F_F(N)/k_B T)$ where the superscript “dist” denotes a system of distinguishable particles. The partition function corresponding to a final state of two distinguishable clusters of $N/2$ indistinguishable molecules is $1/((N/2)!)^2 (Z_{N/2}^{\text{dist}})^2$. The free energy difference is thus $\Delta F = 2F_F(N/2) - F_F(N) = -k_B T \ln\{N!/(N/2)!^2 [(Z_{N/2}^{\text{dist}})^2/Z_N^{\text{dist}}]\}$. However, the difference extracted from molecular dynamics simulations is $\Delta F_{\text{mit}} = -k_B T \ln\{(Z_{N/2}^{\text{dist}})^2/Z_N^{\text{dist}}\}$, and thus $\Delta F = \Delta F_{\text{mit}} + k_B T \ln[(N/2)!^2/N!]$.

We now turn our attention to the excess free energy difference between a N cluster and two $N/2$ subclusters:

$$\begin{aligned}
2F_s(N/2) - F_s(N) = & \\
2F_T(N/2) - 2k_B T \ln[\rho_c^{N/2}(0)V] - \frac{N\kappa_i}{2} \int \rho_0^{N/2}(\mathbf{y}) y^2 d\mathbf{y} & \\
+ 2k_B T \ln(V/\Lambda^3) - 2 \left(\frac{N}{2} - 1 \right) k_B T \ln(\rho_{vs}\Lambda^3) & \\
- \left[F_T(N) - k_B T \ln[\rho_c^N(0)V] - \frac{N\kappa_i}{2} \int \rho_0^N(\mathbf{y}) y^2 d\mathbf{y} \right. & \\
\left. + k_B T \ln(V/\Lambda^3) - (N-1)k_B T \ln(\rho_{vs}\Lambda^3) \right], & \quad (17)
\end{aligned}$$

which after grouping terms reduces to the following four contributions:

$$\begin{aligned}
2F_s(N/2) - F_s(N) = & \\
\Delta F_{\text{mit}} - k_B T \ln \left[\frac{N!}{((N/2)!)^2} \right] & \\
+ k_B T \ln \left[\frac{\rho_c^N(0)\rho_{vs}}{(\rho_c^{N/2}(0))^2} \right] & \quad (18) \\
- \frac{N}{2} \kappa_i \int [\rho_0^{N/2}(\mathbf{y}) - \rho_0^N(\mathbf{y})] y^2 d\mathbf{y}.
\end{aligned}$$

We note that there is a cancellation of the terms involving the system volume V , as required in an expression representing excess free energies. The third term can be written as

$$k_B T \ln \left[\frac{\rho_c^N(0)\rho_{vs}}{(\rho_c^{N/2}(0))^2} \right] = k_B T \ln \left[\frac{(4\pi)^{3/2} \rho_{vs} v_{\text{HO}}}{(N/2)^{3/2}} \right], \quad (19)$$

where $v_{\text{HO}} = (k_B T / \kappa_i)^{3/2}$ is the volume scale characterising the confinement of particles in the final harmonic potentials.

For the fourth term, we represent the single particle density profile as a step function $\rho_0(\mathbf{y}) = \rho_l / N$ for $0 < y < r_{\text{max}}$. In this approximation the cluster is imagined to be spherical with a constant molecular density ρ_l and a radius r_{max} . Inserting this into the fourth term gives

$$\begin{aligned}
& - \frac{N\kappa_i}{2} \int [\rho_0^{N/2}(\mathbf{y}) - \rho_0^N(\mathbf{y})] y^2 d\mathbf{y} \\
& = - \frac{N\kappa_i}{2} \left[\frac{3}{5} \left(\frac{3(N/2)v_l}{4\pi} \right)^{2/3} - \frac{3}{5} \left(\frac{3Nv_l}{4\pi} \right)^{2/3} \right] \quad (20) \\
& = - \frac{N\kappa_i}{2} \frac{3}{5} \left(\frac{3(N/2)v_l}{4\pi} \right)^{2/3} (1 - 2^{2/3}),
\end{aligned}$$

where v_l is the volume per molecule in the bulk condensed phase. Inserting Eqs. (19) and (20) into Eq. (18) leads to the following sum of four terms:

$$\begin{aligned}
2F_s(N/2) - F_s(N) = & \\
\Delta F_{\text{mit}} - k_B T \ln \left[\frac{N!}{((N/2)!)^2} \right] & \\
+ k_B T \ln \left[\frac{(4\pi)^{3/2} \rho_{vs} v_{\text{HO}}}{(N/2)^{3/2}} \right] & \quad (21) \\
+ \left(\frac{N}{2} \right)^{5/3} \kappa_i \frac{3}{5} \left(\frac{3v_l}{4\pi} \right)^{2/3} (2^{2/3} - 1) \\
= f_s^1 + f_s^2 + f_s^3 + f_s^4,
\end{aligned}$$

which is the form given in Eq. (7) in the main text.

APPENDIX B: CONVERGENCE OF THE FREE ENERGY

An analysis of the convergence of the change in free energy associated with the mitosis of a tethered cluster of TIP4P/2005 water obtained from our MD simulations for the 6, 12 and 24 particle systems as a function of the guided separation time t_{sim} is depicted in Figure 7. A value of $t_{\text{sim}} = 10.24$ ns is seen to be sufficient to obtain a converged change in free energy without being too computationally expensive. The time scales are in line with what was found for the disassembly of argon clusters [38], though one should note that in such a process longer times are expected than for the mitosis method.

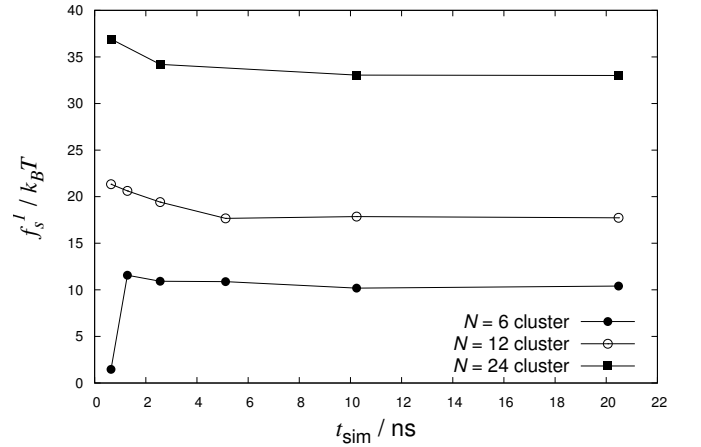


Figure 7. The change in free energy associated with the mitosis of $N = 6, 12$ and 24 tethered water clusters obtained from MD simulations for a range of separation times.

Another possible issue with the calculation of the excess free energy with the mitosis methodology is sensitivity of the results to the value of the spring constant employed to tether the clusters. The excess free energy difference obtained to split a cluster of six TIP4P/2005 water molecules (into two clusters of 3 molecules) as a function of the initial value of the spring constant is illustrated in Figure 8; the largest variation between the

values is found to be significantly smaller than the thermal energy $k_B T$ ($\sim 0.1 k_B T$). The corresponding value of the excess free energy per unit area for the $N = 3$ cluster system is depicted in Figure 9 and found to be quite insensitive to the value of the initial spring constant with small deviations of $\sim 0.2 \text{ mN m}^{-1}$. These results suggest $\kappa_i = 0.05 \text{ kJ mol}^{-1}$ is a suitable (and sufficient) choice for the value of initial spring constant. It is also noted that for lower spring constants, the clusters are not sufficiently long lived in some trajectories within the duration of the simulation, so those are not included in the computation of the change in free energy as shown in the Table II.

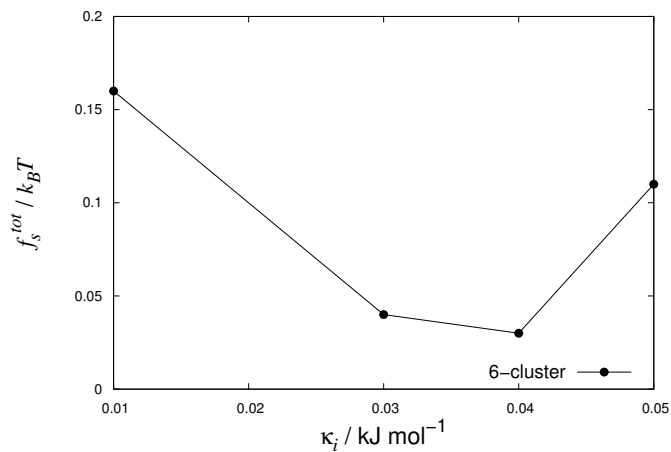


Figure 8. The extracted excess free energy difference for the mitosis of the $N = 6$ cluster as a function of the initial spring constant.

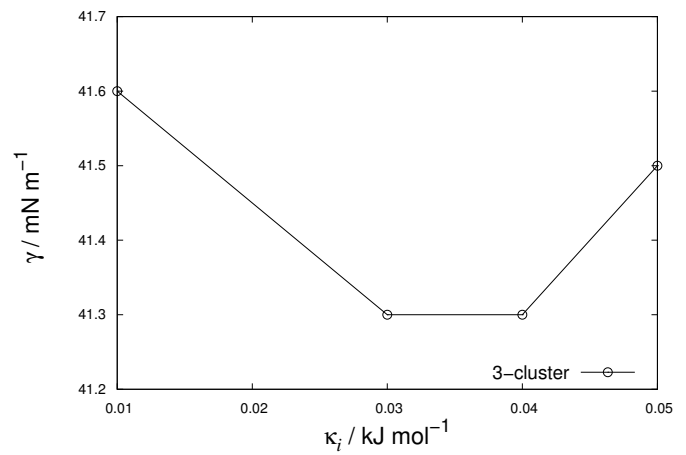


Figure 9. The corresponding values (given the excess free energy difference values for the mitosis of the $N = 6$ cluster in Figure 8) of the surface tension for the $N = 3$ cluster as a function of the initial spring constant.

κ_i / kJmol ⁻¹	Successful trajectories
0.05	1000
0.04	996
0.03	999
0.01	979

Table II. The number of trajectories with systems that remain bound (clustered) for the duration of the simulation at different values of the initial spring constant.