

Molecular Simulation of Nucleation and Dissolution of
Alkali Halides

by

Gabriele Lanaro

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Chemistry)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

March 2017

© Gabriele Lanaro 2017

Abstract

The process of crystal nucleation, despite being so fundamental and ubiquitous in industrial and natural processes, is still not fully understood because of its stochastic nature, and the high spatial and temporal resolution needed to observe it through experiments. This thesis investigates several aspects of nucleation through the use of molecular dynamics, a computational technique that is able to simulate systems up to $\sim 10^{12}$ atoms (as of today's computational power).

The projects in this thesis focus on the nucleation from aqueous solution of alkali halide salts, with supplementary studies on the related processes of dissolution in water, and crystallization from the melt.

The mechanism of NaCl nucleation from solution is examined in Chapter 3 by direct simulation. The NaCl supersaturated solution was found to contain many small ionic clusters that continuously form and disappear from solution until one (or more) of them nucleates and grows irreversibly. An original method was developed to detect and follow clusters in time, producing results useful in the study of their characteristics and lifetimes. Most importantly, it was found that the lifetime of transient clusters is ~ 1 ns, and that both the cluster lifetime and nucleation probability are significantly higher if the cluster is more geometrically ordered. The dissolution of NaCl crystals was also investigated. The process was found to happen in stages, is characterized by an activation barrier, and can be

described by a simple rate law.

The crystal nucleation of LiF from supersaturated solution was observed, in our simulations, only at high pressure and temperature. The growth rate for an already nucleated crystal was found to have a temperature dependence that follows the Arrhenius law, and further evidence suggests that the reason for such behavior is the high activation energy required to dehydrate the ions.

The crystallization from the melt of the Joung-Cheatham and Tosi-Fumi models for lithium halides was also investigated. We found that, for the Tosi-Fumi model, all lithium halides crystallize as wurtzite. For the Joung-Cheatham model, LiF and LiCl crystallize as rock salt, while LiBr and LiI crystallize as wurtzite.

Preface

The Chapters in this thesis are based on work that has been, or will be, published by G. Lanaro and G. N. Patey.

A version of Chapter 3 has been published by G. Lanaro and G. N. Patey, The Birth of NaCl Nanocrystals: Insights from Molecular Simulation, The Journal of Physical Chemistry B **120**, 34 (2016), and was inspired by previous work by D. Chakraborty and G. N. Patey, How Crystals Nucleate and Grow in Aqueous NaCl Solution, The Journal of Physical Chemistry Letters **4**, 4 (2013).

Versions of Chapters 4 and 5 are currently being prepared for publication.

Chapter 6 is based on a publication by G. Lanaro and G. N. Patey, Molecular Dynamics Simulation of NaCl Dissolution, The Journal of Physical Chemistry B **119**, 11 (2015).

In all projects my contribution consisted of performing all of the simulations and data analysis, writing codes for visualization and analysis, formulating hypothesis, designing the projects along with my supervisor, and producing initial drafts for the manuscripts. Editing, reviews and production of final drafts were carried out together with my supervisor.

Table of Contents

Abstract	ii
Preface	iv
Table of Contents	v
List of Tables	viii
List of Figures	x
List of Symbols	xviii
Dedication	xix
1 Introduction	1
1.1 Phase Transitions and Nucleation	1
1.2 Classical Nucleation Theory	3
1.3 Computer Simulations	8
2 Models and Methods	12
2.1 Overview	12
2.2 Molecular Dynamics	12
2.3 Interaction Potentials	14
2.3.1 Van der Waals Forces	14
2.3.2 Electrostatic Forces	16
2.3.3 Water and Ion Models	16
2.3.4 Infinite Systems	17
2.4 System Initialization	23

2.4.1	Liquid Systems	23
2.4.2	Crystals	24
2.4.3	Crystal-liquid Interfaces	24
2.5	Time Evolution	25
2.5.1	Constraints	26
2.6	Temperature and Pressure Control	28
2.6.1	Velocity Rescaling Thermostat	29
2.6.2	Berendsen Barostat	31
2.6.3	Parrinello-Rahman Barostat	32
3	Nucleation of NaCl from solution	34
3.1	Introduction	34
3.2	Models and Methods	37
3.2.1	Simulation Details	37
3.2.2	Cluster Detection	39
3.3	Results and Discussion	47
3.3.1	Cluster Lifetimes	50
3.3.2	Crystallinity and Nucleation	55
3.4	Summary and Conclusions	67
4	Crystallization of Lithium Halides from Molecular Simulation.	69
4.1	Introduction	69
4.2	Models and Methods	71
4.3	Results and Discussion	75
4.3.1	Lattice Energies	75
4.3.2	Infinitely Periodic System Simulations	82
4.3.3	Finite Size Cluster Simulations	83
4.3.4	Temperature Effects	89
4.4	Summary and Conclusions	90
5	Crystallization of Lithium Fluoride in Aqueous solution from Molecular Simulation.	91
5.1	Overview	91

5.2	Introduction	92
5.3	Models and Methods	93
5.4	Results and Discussion	97
5.4.1	Diffusion	98
5.4.2	Growth Barrier	99
5.4.3	Solution Structure and Dynamics	101
5.4.4	Solubility	106
5.5	Summary and Conclusions	109
6	Dissolution of NaCl nanocrystals	111
6.1	Overview	111
6.2	Introduction	112
6.3	Models and Methods	114
6.4	Results and Discussion	118
6.4.1	Stages of Dissolution	118
6.4.2	Concentration Effects	132
6.4.3	Temperature Dependence and Activation Energy	134
6.5	Summary and Conclusions	136
7	Conclusions	138
7.1	Summary	138
7.2	Future Directions	141
	Bibliography	144

Appendices

A	Distribution of Various Properties for Nucleated and Failed Clusters.	157
----------	--	------------

List of Tables

3.1	Joung-Cheatham Lennard-Jones parameters for NaCl. ¹	37
3.2	Summary of the simulations performed. The number of water molecules (Water), the number of ions (Ions), the mole fraction x_{NaCl} , the run length (Length), and the number of clusters that achieved nucleation (Nuclei) are indicated. At $x_{NaCl} = 0.3$, multiple nucleations occurred resulting in a polycrystal. The five simulations at $x_{NaCl} = 0.22$ listed at the bottom of the table were used for the statistical analysis given in the text.	39
4.1	Lattice energies for the JC model for the wurtzite (E_{wurtz}) and rock salt (E_{rock}) structures. The difference $E_{rock} - E_{wurtz}$ is reported as ΔE . The NaCl lattice energies are added for comparison.	80
4.2	Lattice energies (TF model) for the wurtzite (E_{wurtz}) and rock salt (E_{rock}) structures. The difference $E_{rock} - E_{wurtz}$ is reported as ΔE . The NaCl lattice energies are added for comparison.	80
4.3	Final structures obtained by melting and freezing periodic crystals. The final temperature is 300 K, and the periodic box contained 1000 ions. The results are consistent with the lattice energy calculations (Tables 4.1 and 4.2).	82
4.4	Final structures obtained by melting and freezing finite size clusters using the JC and TF models, the final temperature is 300 K. A wurtzite-like structure corresponds to a structure that is not perfectly crystalline but shows hexagonal motifs similar to that shown in the left panel of Figure 4.5.	84
4.5	LiY lattice energies at 0 K and 300 K. The value E^w was obtained by simulating a box to initialized from the perfect wurtzite structure, while $E^{w\ddagger}$ is the energy of the crystal obtained from freezing of the molten salt.	89
5.1	Joung-Cheatham Lennard-Jones parameters for LiF ¹ adapted for the SPC/E water model. ²	93

5.2	Summary of the simulations performed to study LiF nucleation and growth. The temperature at which a simulation was conducted is indicated by T, the length of the simulation is indicated by Length, the concentration in mole fraction is abbreviated as m.f., and the time to nucleation (ttn) is the time at which the first nucleus was observed. LiF and H ₂ O indicate the number of molecules of each type used in the simulation. The pressure values for the crystal growth simulations are reported as a range for some simulations because, when substantial crystal growth occurs, the pressure drifts to higher values.	94
5.3	Diffusion constants for Li ⁺ (D ₊) and F ⁻ (D ₋) at different temperatures.	99
5.4	Growth rates at different temperatures	101
5.5	Average number of clusters (n _c) per time frame at different temperatures, for LiF and NaCl supersaturated solutions (see Chapter 3 for details). n _c is normalized by the number of ion pairs in the simulations to allow comparisons between simulations with different numbers of ions. The concentration in mole fraction is reported in the m.f. column.	106
6.1	Joung-Cheatnam NaCl parameters ¹ for the Lennard-Jones potential.	114
6.2	Summary of crystal shape, the number of ion pairs, the number of water molecules, and the temperatures used in the simulations. . . .	116
6.3	Goodness of fit parameters R ² obtained for different rate laws. The best fits are indicated in bold.	130

List of Figures

1.1	Kinetic model of CNT. Droplets can grow only by single molecule attachments and shrink by single molecule detachments. The forward rate $a(n) = A(n)\beta$ is proportional to the surface area of the nucleus, $A(n)$, and the rate of arrival of single molecules on the droplet, β . The backward rate $b(n) = A(n)\alpha$ is proportional to the rate of detachment per unit area, α , and the surface area.	4
2.1	The LJ potential with parameters $\sigma = 1$ and $\epsilon = 1$. The σ correspond to the x-intercept of the potential, while ϵ corresponds to the well depth.	15
2.2	Positions of the charges in the models employed in our simulations. The SPC/E water model on the left represents the atoms as point charges, with a O-H distance of ~ 0.1 nm and a HOH angle of ~ 109.47 deg. The alkali and halide ions are modeled as single point charges. The short-range interaction parameters are not shown in the picture. For the water model, the masses of the oxygen and hydrogen atoms were employed for the negative and positive charged sites, respectively. Similarly, the mass for the corresponding element was chosen for the alkali halide models.	17
2.3	Depiction of PBC. The central system is repeated in all directions. A particle exits from the central box on the right side (black arrow) and reenters from the opposite side.	19
2.4	The short-range and long-range contribution to the electrostatic potential for $\kappa = 1$	21
3.1	Density of q_8 values in a sample crystal (~ 2000 ions)(blue) and a supersaturated solution ($x_{NaCl} = 0.20$) (green). The two distributions show excellent separation. The vertical orange line indicates the selected threshold ($q_8 > 0.35$).	41

3.2	Schematic diagram illustrating the DBSCAN algorithm for $\eta = 4$. Top left panel: an ion is chosen at random and ions within ϵ are counted; the value is 4 so the ion is marked as <i>core</i> (red). Its neighbours are labeled as <i>border</i> (pink). Top right panel: the procedure is repeated for one of the border ions which becomes a core ion. Bottom panel: all ions in the cluster have been processed, and no more ions are reachable from the core ions. The isolated ion is labeled as noise (blue).	43
3.3	Schematic diagram illustrating the cluster time resolution algorithm. At $t = 0$ the clusters are labeled employing the DBSCAN algorithm. A cluster is compared with clusters in the following frame, and a connection is made if the clusters are sufficiently similar, judged by the Jaccard index as described in the text. The procedure is repeated recursively for the following frames. By performing this procedure for all clusters, it is possible to obtain a connectivity graph (bottom) where each connected component represents the time evolution of a single cluster.	45
3.4	Size (top left) and crystallinity (top right) histograms for clusters at first detection. The bottom panel shows the joint distribution of size and crystallinity. Note that very few clusters begin with both high size and high crystallinity, suggesting that this feature develops as clusters evolve in time. The red points indicate the thirteen clusters that achieve nucleation, and we note that these show no obvious preference for any region of the joint distribution.	49
3.5	Examples of clusters of different sizes and crystallinities (given in parenthesis below each cluster). Na^+ and Cl^- ions are colored purple and green, respectively. At inception larger clusters tend to be elongated and of lower crystallinity (top left), whereas larger high crystallinity clusters (not observed at early stages) tend to be more compact (top right). For small clusters structural features associated with low (bottom left) and high crystallinity (bottom right) are not so apparent.	51
3.6	Histograms of “failed” cluster lifetimes (left) and the overall survival function (right) for all clusters detected in the simulations.	52
3.7	Survival curves for clusters of size 10 grouped by high (> 0.40) and low (≤ 0.40) crystallinity.	54
3.8	Ratio of median lifetimes for high versus low crystallinity clusters of different size. The high crystallinity clusters always have the higher median lifetime.	55

3.9	Trajectory shown in (size, crystallinity) space for a nucleated (blue) and a failed (red) cluster as they evolve in time. In early stages both clusters oscillate in size up to ~ 30 ions, but the crystallinity of the cluster that eventually achieves nucleation reaches higher values than the failed case. This suggests that it is a combined effect of both size and crystallinity that promotes nucleation.	56
3.10	Comparison of growth and crystallinity profiles for nucleated (colored) and failed (gray) clusters. After a short period, the nucleated cluster manages to reach quite a high crystallinity and maintain its size. In contrast, the failed cluster, while maintaining its size, experiences a steady decrease in crystallinity.	58
3.11	Size (blue) and crystallinity (orange) profiles for a nucleated cluster. The fluctuation in crystallinity is fairly high, and, as the crystal increases in size, the crystallinity reaches a plateau at ~ 0.47 . The snapshots represent the cluster at 0 ns (left), 40 ns (center) and 80 ns (right).	59
3.12	Crystallinity distributions (blue histogram) of failed clusters of different size (number of ions). Clusters that achieved nucleation are indicated by single orange lines. Note that clusters that achieve nucleation come preferentially from the upper part of the crystallinity distribution.	60
3.13	Probabilities of high ($\bar{q}_8 > 0.40$) and low ($\bar{q}_8 < 0.40$) crystallinities achieving nucleation. The error bars (vertical black lines) represent one standard deviation. The lower panel shows the results for small clusters on an expanded scale. Note that for all cluster sizes $P(N high)$ is always larger than $P(N low)$, and the high/low ratio is shown as a gray line in the upper panel.	62
3.14	Joint distribution of crystallinity and cluster energy (kJ mol^{-1}) for clusters of size 6 (left) and size 10 (right). Values corresponding to clusters that eventually nucleated are displayed in orange. The correlation coefficient is negative and very low at size 6, and moderately low at size 10, indicating that crystallinity carries different information than energetic stability.	64

3.15	Running coordination numbers for: high q_8 Na^+ with all Cl^- (blue); high q_8 Cl^- with all Na^+ (green); and all Na^+ and Cl^- (red). In the first coordination shell (~ 0.35 nm) both high q_8 ions are surrounded by more counterions (3-4) than average (1). Also, high q_8 Na^+ are surrounded by more counterions than high q_8 Cl^- , suggesting that on average they occupy positions deeper within the clusters. The difference in the first shell coordination number decreases at longer times, as larger crystals develop in the simulation, and the internal preference of Na^+ becomes less noticeable.	65
3.16	Δq_8 as a function of cluster size. The shaded area indicates one standard deviation. The average value is slightly positive indicating that Na^+ ions tend to occupy ordered environments than Cl^- ions.	66
4.1	Unit cells used for the rock salt (right) and wurtzite (left) crystal structures. The configurations are dependent on a single cell parameter a , all other parameters being constrained. In the rock salt structure, a is the length of any edge, while for the wurtzite structure a is the length of any of the two short edges.	73
4.2	Radial distribution functions for an infinite LiCl crystal for the TF model (left panel) and JC model (right panel). The initial rock salt configuration and the liquid phase are colored in blue and red, respectively, and do not show noticeable differences between models. The rdf corresponding to the crystallized structures are displayed in green and correspond to wurtzite on the left and rock salt on the right. The inset plots show the rdf for ideal wurtzite on the left and ideal rock salt on the right (black lines) superimposed onto the simulation result (green).	76
4.3	TF and JC potentials for the pairs Li^+-Cl^- (top panel), Li^+-Li^+ (middle panel) and Cl^--Cl^- (bottom panel). The TF model has a shorter contact radius for the Cl^- ion, and is more slowly varying function at short range. Notice also how at very short range (~ 0.1 nm), the TF model becomes attractive (but this does not influence the simulations).	78
4.4	Lattice energies as a function of the cell length a for the JC (left column) and TF (right column) potentials. The solid red lines indicate the energy of the wurtzite structure, while the dashed red lines indicate the rock salt structure. The energy minima are annotated below each curve. Blue and green lines represent the contributions of the positive and negative ion sites, respectively.	81

4.5	Spontaneously crystallized LiCl/TF (left panel) and LiCl/JC (right panel) clusters of 216 ions. For the cluster on the left, the structure has hexagonal motifs characteristic of the wurtzite structure, but is mostly hollow inside. On the right, the cluster crystallizes into a rock salt structure, without substantial surface deformations. In this figure, Li^+ is represented in purple and Cl^- is in green.	85
4.6	LiI/JC rdfs at 300 K for the perfect wurtzite crystal (red), and for the interior of a finite size crystal of 1000 ions (blue). The peaks are well defined for the perfect wurtzite crystal, while for the finite size crystal there is wide broadening and merging of adjacent peaks. . . .	86
4.7	Example of surface reconstruction for a LiI crystal (Li^+ and I^- ions are colored in purple and grey, respectively). On the left, surface charges are present on the wurtzite crystal face (0001). After simulating the crystal in vacuum at 300 K, the resulting surface (on the right) rearranges to reduce the dipole moment. The dipole moment per ion is reduced from ~ 6.3 D for the structure on the left, to ~ 0.7 D for the structure on the right.	87
4.8	Spontaneously crystallized LiBr/JC cluster (1000 ions) using the LiBr/JC model (Li^+ in purple, and Br^- in brown). One notices both hexagonal motifs, characteristic of the wurtzite structure, as well as square motifs characteristic of the rock salt structure.	88
5.1	Clustering based on neighbor distance. The dashed lines represent distances lower than a fixed threshold, and the colors represent different clusters. Free particles form a cluster of their own.	97
5.2	Schematic diagram of the free energy landscape at different temperatures for the crystal growth process. If the driving force (which is the difference in free energy of initial and final stages) increases and the barrier is weakly affected by the increase in temperature, the activation barrier (which is the height of the curve between initial and final stages) of the process decreases.	98
5.3	Linear regression to determine the activation energy for the diffusion process for Li^+ ions.	100
5.4	Arrhenius fit to the growth rates	101
5.5	Fraction of ions in connectivity based clusters of a certain size at different temperatures.	102
5.6	Li^+F^- radial distribution functions at different temperatures. The contact peak, characteristic of ion pairs increases with temperature, while the second peak (solvent-separated pairs) decreases.	103

5.7	Diagram of surviving pairs. Initially, the {1, 2, 3, 4} cluster contains 6 unique pairs, after the cluster splits, 2 of the pairs are still associated.	104
5.8	Surviving pairs at a time delay of t ns for a range of different temperatures. As the temperature increases, there are more associated pairs that decay at a faster rate.	105
5.9	Solution concentration profiles obtained by dissolving a crystal at 300 K (blue), by growing a crystal at 500 K (red, top line), and dissolving a crystal at 500 K (red, bottom line). The estimated solubilities are ~ 0.025 mole fraction at 300 K and ~ 0.01 at 500 K.	108
5.10	Possible free energy landscape for LiF nucleation from solution. A first barrier is required to go from the free ions (solution) to the cluster aggregates. The cluster aggregates are metastable, continuously form and disappear from solution, and are more easily found at high temperature. The second barrier reflects the fact that most clusters do not nucleate but dissolve back into solution, depending on cluster related properties such as size, crystallinity, and surface tension. In absence of nucleation rate measurements, it is not possible to estimate the relative magnitude of the two barriers, and two possible options are displayed as solid and dashed lines.	110
6.1	A representation of the crystals simulated. The number of NaCl ion pairs is given below each illustration.	115
6.2	An example of the order parameter applied to a surface ion. The central atom (yellow) is surrounded by 14 neighbors (red). The 0.6 nm cutoff applied to the central atom is highlighted with a gray transparent sphere.	118
6.3	Dissolution profiles for the cubical and spherical crystals displayed as number of ions in the crystal vs time. The starting and final sections of the profiles are detailed in the zoomed-in plots.	120
6.4	Snapshots corresponding to different points in the dissolution profile of the cubic crystal. After 20 ns, only ions located at edges and corners are removed. The edges and corners gradually get consumed (100 ns) and the crystal becomes roughly spherical after about 185 ns. After that point the shape doesn't change until the final stage of dissolution.	121

6.5	Comparison of the dissolution profiles for different shapes expressed in terms of fraction of crystal dissolved $f_{\text{cry}}(t)$, the derivatives df_{cry}/dt (on the order of -10^{-3} ns^{-1}) are depicted in the inset. The color coding is: dark blue for the sphere; green for the cube; red for the tablet; light blue for the rod. The rod-shaped and tablet-shaped crystals show a higher overall dissolution rate.	122
6.6	Snapshots along the rod-shaped crystal trajectory. Ions are detached mainly from the ends of the rod.	122
6.7	Snapshots along the tablet-shaped crystal trajectory. The top and bottom surfaces are never attacked by water molecules.	123
6.8	Three quantities obtained for the spherically shaped crystal as functions of the number of neighbor in the first coordination shell (within 0.37 nm), Left panel: The relative probability of detachment (within 1.6 ns) of a surface ion, given the number of neighbors. Upper right panel: The total number of ions on the surface at $t = 0$. Lower right panel: The total number of ions detached over 1.6 ns. All quantities are averaged over 50 time slices between 20 and 40 ns of the dissolution simulation.	125
6.9	Radial ion density profile starting from the center of a spherical crystal at 300 ns. The dashed horizontal line in the inset indicates the density at saturation. The calculated radius of the crystal is ~ 2 nm, and a concentration gradient (that could be related to a diffusion layer) is not observed.	131
6.10	Snapshots of the last stage of dissolution of the spherical crystal. When the crystal reaches ~ 64 ions, the structure begins to disintegrate. In the final snapshot, water quickly penetrates the structure and breaks apart the remaining crystal nucleus.	132
6.11	Concentration effects on rate. Curves (a), (b), and (c) are for NaCl mole fractions (calculated assuming complete dissolution of the entire crystal) of 0.0254, 0.0163, and 0.0127. Note that curve (a) shows a sharp decrease in rate at ~ 120 ns, while there is no substantial difference between the profiles (b) and (c).	133
6.12	Detachment and reattachment events for the spherical crystal. During the course of dissolution, the detachment rate systematically slows while the attachment rate remains substantially constant.	134
6.13	Fits to the cube root law for the spherical crystal at temperatures of 300, 320 and 340 K.	135

6.14	Fits to the Arrhenius equation of the rate constants obtained for the spherical crystal at 300, 320, and 340 K. Error bars represent one standard deviation of uncertainty.	135
A.1	Distributions of various properties for failed clusters (blue histogram) of size 10. Clusters that achieved nucleation are indicated by single orange lines. Note that the property values of the clusters that achieve nucleation are found around the mean of the distribution.	159
A.2	Distributions of various properties for failed clusters (blue histogram) of size 30. Clusters that achieved nucleation are indicated by single orange lines. At this size, some preference for low radius of gyration, low surface area and higher sphericity can be observed.	160
A.3	Two dimensional representation of the algorithm used to estimate the volume. The cluster is encased in a grid, and the grid points that lie within the cluster are represented in red.	161

List of Symbols

k_B	Boltzmann constant
\mathcal{K}	Kinetic energy
\mathcal{V}	Potential energy
$\text{Tr}(A)$	Trace of matrix A
bcc	Body centered cubic
CNT	Classical Nucleation Theory
fcc	Face centered cubic
JC	Joung-Cheatham
LJ	Lennard-Jones
PBC	Periodic boundary conditions
PME	Particle-mesh Ewald
rdf	radial distribution function
TF	Tosi-Fumi

To my parents

Chapter 1

Introduction

1.1 Phase Transitions and Nucleation

Physical matter can exist in different phases, and phase transitions are the processes that describe how matter can transform from one phase to another, after a change in one or more thermodynamic parameters. In equilibrium thermodynamics, the stable phase of a system at constant temperature T and pressure p is the one with the lowest chemical potential μ . In general, as the temperature and pressure of the system are changed, the chemical potential of each phase will vary by different amounts, defining the regions of stability in the phase diagram.

While thermodynamics describes systems at equilibrium, it gives no information on the time evolution, nor on the atomistic details of phase transitions. The dynamical aspects of phase transitions are, however, extremely important as the time scales involved can be very large, and their understanding is crucial in many applications.

A system that is stable to small fluctuations in thermodynamic parameters, but can transition to a more stable state given a large enough perturbation, is termed

metastable. Metastable states are due to the fact that the system is in a local free energy minimum, and to reach the global minimum it is necessary to overcome a free energy barrier. An extreme example is the graphite-diamond transition. At 298 K and 1 bar, graphite is the stable phase of carbon, with a free energy that is $\sim 2.90 \text{ kJ mol}^{-1}$ lower than that of diamond.³ However, the time to convert diamond to graphite at room temperature and pressure is extremely long, making diamond remarkably metastable.

The dynamics of phase transitions are usually explained through a process termed nucleation. In order to form a new phase, particles in the metastable phase spontaneously arrange to form the initial seed of the new phase, that will grow as the transition progresses. The formation of the initial seed can happen directly in the metastable phase (homogeneous nucleation), or on a surface, such as the wall of the container, or can be initiated by impurities (heterogeneous nucleation).

Nucleation is at the core of many physical and biological processes. For example, some organisms in cold environments have developed strategies to prevent freezing of bodily fluids through the use of antifreeze proteins that inhibit ice nucleation.^{4,5} Another example are high clouds (clouds that occur between heights of 5 to 13 km above the earth's surface), where water droplets exist in a supercooled state, and freezing can be triggered⁶ at about $-15 \text{ }^\circ\text{C}$, catalyzed by other particles, such as clay particles, carried into the atmosphere by the wind.^{7,8} Once water droplets freeze they can grow and coalesce into snowflakes or hail that precipitate to the ground.⁹

In the remainder of this Chapter, we will describe the classical kinetic theory of nucleation, the experimental and theoretical attempts to validate it, and how computer simulations have proven to be a valuable tool in the study of nucleation.

1.2 Classical Nucleation Theory

The kinetic treatment of nucleation dates back to 1926, when Volmer and Weber¹⁰ postulated that the nucleation rate J (number of nuclei formed per unit time and volume) depends on the negative exponential of the free energy required to form a new nucleus. Later developments and refinements^{11–13} gave rise to the standard treatment for nucleation, often called classical nucleation theory (CNT).

In the original treatment of CNT, applied to a homogeneous gas-liquid transition, the metastable vapor is characterized to be a mixture of single molecules and liquid droplets of various sizes. The system is assumed to be in an equilibrium¹⁴ characterized by a droplet distribution $N(n)$, that is proportional to the negative exponential of the free energy ΔG needed to form a droplet of size n

$$N(n) = N_0 \exp\left(-\frac{\Delta G(n)}{k_B T}\right), \quad (1.1)$$

where N_0 is the number density of the nonassociated molecules, and k_B is the Boltzmann constant. The equilibrium distribution $N(n)$ describes the state of the vapor before the phase transition, and to characterize the kinetics of the process it is necessary to consider the time evolution of the actual, instantaneous, droplet distribution $f(n, t)$.

To recover this dynamical quantity, CNT assumes that the interchange of molecules between the liquid droplets and the vapor happens by single molecule attachment or detachment, each collision between a molecule and a droplet resulting in an attachment, and every attachment or detachment event is uncorrelated with the previous ones (Figure 1.1). If this is true, the droplet formation process described above can reach a steady state and the nucleation rate, J , can be found solely from

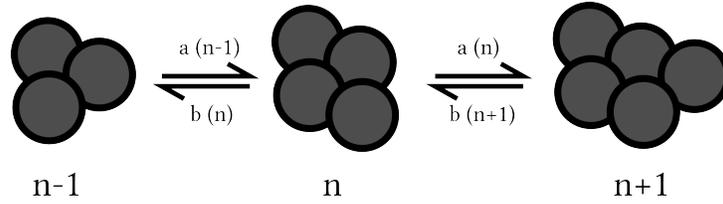


Figure 1.1: Kinetic model of CNT. Droplets can grow only by single molecule attachments and shrink by single molecule detachments. The forward rate $a(n) = A(n)\beta$ is proportional to the surface area of the nucleus, $A(n)$, and the rate of arrival of single molecules on the droplet, β . The backward rate $b(n) = A(n)\alpha$ is proportional to the rate of detachment per unit area, α , and the surface area.

time independent quantities, such that⁹

$$J = \frac{1}{\sum_{n=1}^{\infty} \frac{1}{\beta A(n) N(n)}}, \quad (1.2)$$

where $A(n)$ is the surface area of a droplet of size n , and β is the rate of arrival of single molecules at the droplet per unit area.

Other assumptions are necessary to simplify Equation (1.2). According to CNT, droplets are postulated as large, homogeneous, spherical, and incompressible objects. The inside of the droplet is made of bulk liquid, and the surface free energy of the cluster can be described by the product of a size-independent surface tension, γ , and the cluster surface area $A(n)$. This set of assumptions is usually referred to as the *capillary approximation*.

If the capillary approximation applies, the free energy of formation for a droplet of size n will depend on the free energy of formation of the new *bulk* phase, as well as on the free energy required to form a new surface,

$$\Delta G(n) = -n\Delta\mu + \gamma A(n), \quad (1.3)$$

where $\Delta\mu$ is the difference between the chemical potential of the liquid and the vapor, and γ is the surface tension of the liquid. Even though CNT was originally formulated for vapor to liquid transitions, the theory can be applied to other transitions, such as crystal nucleation, by using the chemical potentials and interfacial free energies of the phases of interest.

When the cluster is very small, the free energy required to create the new surface will be higher than the free energy gained by the formation of the more stable phase, and as the size of the droplet grows, ΔG will go through a maximum, then decrease to the point of being negative. The size n^* at which the maximum is found, also called *critical size*, is

$$n^* = \frac{32\pi v'^2 \gamma^3}{3 \Delta\mu^3}, \quad (1.4)$$

and the maximum $\Delta G(n^*)$, also termed the *nucleation barrier*, is

$$\Delta G(n^*) = \frac{16\pi v'^2 \gamma^3}{3 \Delta\mu^2}, \quad (1.5)$$

where v' is the volume per particle in the condensed phase.

Given the above assumptions, the sum in Equation (1.2) can be expressed as a solvable integral⁹ and the rate J is found to depend on $\Delta G(n^*)$ and a prefactor J_0

$$J = J_0 \exp\left(-\frac{\Delta G(n^*)}{k_B T}\right). \quad (1.6)$$

The factor J_0 can be expressed as

$$J_0 = Z\beta A(n^*)N_0, \quad (1.7)$$

$$Z = \sqrt{\frac{-\Delta G''(n^*)}{2\pi kT}}, \quad (1.8)$$

where Z , also termed *Zeldovich factor*,⁹ physically represents the fact that there is a probability that critical nuclei can revert to vapor, while the term $A(n)\beta$ is the frequency of arrival of single particles on a nucleus of size n . The term $\Delta G''(n^*)$ is the second derivative with respect to n of ΔG , evaluated at n^* . The main parameters affecting the nucleation rate, J , are the temperature, the difference in the chemical potential of the two phases, and the surface tension.

The temperature dependence is expressed through the prefactor J_0 , the exponential dependency on $1/k_B T$, as well as on the degree of undercooling that will affect the chemical potential $\Delta\mu$. While lower temperatures tend to increase the rate because of the higher degree of undercooling, the rate of arrival β will substantially decrease because of decreased mobility. The chemical potential dependency is expressed through $\Delta G(n^*)$. As $\Delta\mu$ increases, the barrier $\Delta G(n^*)$ will decrease, causing an increase in nucleation rate J . A higher surface tension, γ , leads to an increase in nucleation barrier and a decrease in nucleation rate.

Early experiments using expansion chambers,^{15–18} showed that classical nucleation theory predicts the correct dependence of nucleation rate on supersaturation, but fails to reproduce the temperature dependence, a fact that has been attributed to an inadequate description of the prefactor J_0 .¹⁹

Auer and Frenkel²⁰ were able to characterize the crystal nucleation rate of the hard-sphere model using molecular simulation, and compared their results with

experiments on colloids.^{21,22} They showed that, albeit CNT adequately describes the functional form of $\Delta G(n^*)$, the nucleation rates are more strongly dependent on temperature in models than in experiments, suggesting that the problem is, again, in the description of the temperature dependent prefactor. Further experiments on colloidal systems also showed issues in considering the surface tension as concentration independent.²³

Very interesting developments on the microscopic mechanism of nucleation were evidenced by experiments on CaCO_3 crystal nucleation.²⁴⁻²⁶ In this system, crystal nucleation proceeds quite differently from the molecule-by-molecule growth process described by CNT. It was observed²⁴ that in the saturated solution, amorphous calcium carbonate clusters are first formed, these subsequently undergo transitions to vaterite and only then, convert to the final structure calcite. This indicates a nucleation process that happens in stages, similar to a mechanism that was first postulated by Ostwald²⁷ and termed *rule of stages*.

In contrast with what was found for CaCO_3 , the microscopic process described in CNT was shown to apply to the description of certain systems. For example, Sleutel et al., performed direct observation (through atomic force microscopy) of nucleation in two dimensions of the protein glucose isomerase and showed that subcritical nuclei have the same structure as the bulk crystalline phase.²⁸ Measurements of the nucleation rate of NaCl crystals also seem to be in good agreement with estimates from CNT.²⁹ Additionally, calculations of the critical size for glycine and NaCl aqueous solutions were found to be consistent with estimates of the average cluster size.³⁰

Attempts to describe the first stages of nucleation have proven hard to perform experimentally, because of the intrinsic small size and the short lifetimes of

most nuclei.³¹ Additionally, estimates of the homogeneous nucleation rates are subject to significant uncertainties, and often complicated by unwanted heterogeneous nucleation.³²

Computer simulations are a great way to avoid many of the shortcomings that affect experiments. As discussed above, the foundation of CNT relies upon assumptions about the microscopic nature of the metastable and nascent phases, and through molecular simulation it is possible to observe nucleation in its very early stages for a variety of model systems and physical conditions.

1.3 Computer Simulations

Computer simulations have long been used to study phase transitions of simple atomic systems,^{33–36} and early studies of nucleation were performed on Lennard-Jones (LJ) liquids.^{37–40} It was found that, despite the fact that the stable solid structure for the LJ system is face centered cubic (fcc), the LJ liquid nucleates at moderate supercooling with a body centered cubic (bcc) structure and, starting from the interior, gradually converts to the fcc structure, suggesting that the nucleation process, even for this very simple system, happens in stages.³⁹

Other early work includes simulations of hard spheres that were used to compare crystal nucleation rates with experimental data for colloids.^{20,41} More recently, thanks to a dramatic increase in parallel computational power, simulations are being used to study homogeneous and heterogeneous nucleation molecular systems,^{42–45} including aqueous solutions.^{46–50}

One of the most studied processes is perhaps the nucleation of ice in liquid water.³² Homogeneous nucleation measurements for this system vary by orders of

magnitude between theory and experiments and also between different studies.³² Attempts to study homogeneous nucleation through unbiased simulations have also proven to be problematic³² given the very long time scales involved in ice nucleation, but has been observed when an electric field is applied to the system.^{51–53} Coarse grained simulations have suggested that ice nucleation involves both cubic and hexagonal variants of ice,⁴² also observed in unbiased simulations of heterogeneous ice nucleation,^{54–56} and in experiments.⁵⁷

A considerable body of work, including some of the projects in this thesis, has been dedicated to the NaCl/water system. Early work on NaCl crystal nucleation^{46,47} used small systems to study the early stages, and showed that the Na⁺ ion tend to be positioned near the center of the cluster aggregates present in NaCl solutions.⁴⁷ The mechanism of cluster formation was further explored by Hassan,^{58,59} showing that the supersaturated NaCl solution is characterized by a combination free ions and ionic clusters with relatively long lifetimes. Later work on larger systems^{48–50} gave evidence of a two-stage nucleation process.^{49,50} The formation of the NaCl nucleus was characterized by a density fluctuation followed by an ordering transition, where the cluster rearranges to form a more regular crystal structure, gradually dehydrating as the crystal grows.

In order to accurately describe highly concentrated alkali halide solutions, Joung and Cheatham¹ developed models tuned to reproduce solid properties, such as lattice energies and lattice constants, as well as solution properties (such as the free energy of solvation). Joung-Cheatham (JC) parameter sets have been extensively used to investigate solubility and crystal nucleation from solution.^{60–66}

The main focus of this thesis is the study of crystallization of alkali halides from solution. In addition to using more recent models,¹ we also employ advanced

data analysis techniques to observe and study the morphology and dynamical behavior of ionic clusters in solution.

In Chapter 2, we describe the molecular dynamics techniques and algorithms used for the simulations performed in the subsequent chapters.

In Chapter 3 we employ large-scale molecular simulations to investigate the nucleation of NaCl in water using the JC parameter set, and develop computational methodology to detect and follow in time a large number of pre- and post-nucleation clusters. Statistical analysis shows that a very large number of clusters are formed, and that the cluster lifetime and nucleation probability is strongly affected by the geometrical arrangement of the ions contained in the cluster.

In Chapter 4, we evaluate pair potentials for pure lithium halides by calculating lattice energies, and by melting and freezing finite size clusters to assess their possible role in nucleation studies. The results indicate that for most lithium halide models, the stable structure is not rock salt but wurtzite, and that, in certain cases, the finite size structure can differ from that of the bulk.

In Chapter 5, we examine the crystallization of LiF from solution. We perform molecular simulations of supersaturated LiF solutions at temperatures ranging from 300 to 500 K and, while nucleation was observed only at a temperature of 500 K, we were able to measure the growth rate of an already nucleated crystal at lower temperatures. We found that the growth process is activated and that the barrier to growth is much larger than that of diffusion. We suggest that the barriers to growth and nucleation are likely related to ion dehydration.

In Chapter 6, we focus on the NaCl dissolution process. The study of dissolution provides insights into the dynamics of ion attachment and detachment, and

into the stability limit of small ionic clusters in dilute solutions. By simulating the dissolution of NaCl nanocrystals of different sizes, shapes, and at different temperatures, we find that the dissolution process is activated, and that the activation energy is affected by the position of ions on the crystal surfaces.

A global summary, further remarks, and future directions are presented in Chapter 7.

Chapter 2

Models and Methods

2.1 Overview

Molecular dynamics is a simulation method widely used to study model systems at the microscopic level. To perform a simulation, it is necessary to specify a system in terms of its initial state and interaction forces that apply. Numerical integration of the equations of motion is then performed to obtain a trajectory. In this chapter we explain the main algorithms used to perform molecular dynamics simulations, and how to include other effects such as periodic boundary conditions, constraints, temperature, and pressure controls.

2.2 Molecular Dynamics

Molecular dynamics is a computational technique used to simulate the evolution of a molecular system given the initial conditions and a description of the interatomic forces. The result of a molecular simulation is a time trajectory that

can be used to calculate the thermodynamic, structural and dynamic properties of the system under study.

Atoms in molecular systems are subject to intra and intermolecular interactions that originate from chemical bonds, electrostatic charges, short-range repulsions, and dispersion forces. In molecular simulations, these interactions are usually approximated employing simple potentials that can be used in the framework of classical mechanics. Even though molecular interactions are best described by quantum mechanics, this approximation has proven quite successful in describing many systems of physical and biological interest, especially since molecular dynamics simulations are typically concerned with nuclear motion.

The time evolution of a system, where the initial interaction site positions $\mathbf{r}_i(t = 0)$, velocities $\dot{\mathbf{r}}_i(t = 0)$, and forces \mathbf{f}_i are specified, can be described by Newton's equations of motion

$$\mathbf{f}_i = m_i \ddot{\mathbf{r}}_i, \quad (2.1)$$

where m_i is the mass associated with the interaction site, and $\ddot{\mathbf{r}}_i$ is the acceleration. To obtain a trajectory $\mathbf{r}_i(t)$, the equations of motion are integrated numerically.

In this chapter we discuss the functional forms of the molecular interactions used in the following chapters, and some of the techniques used to initialize the atomic positions in our systems of interest. We also describe how numerical integration is performed, and how to impose further conditions such as periodic boundaries, bond constraints, temperature and pressure controls.

2.3 Interaction Potentials

The first step in performing molecular dynamics is the definition of the models that represent the interactions between the atoms and molecules of interest. The simulations performed in this thesis involve water and alkali halide ions which are subject to van der Waals and electrostatic forces. The chemical bonds between oxygen and hydrogen are not explicitly modeled, rather the bond distances and angles are kept fixed by using constraints (See Section 2.5.1 for details.)

2.3.1 Van der Waals Forces

Van der Waals forces are commonly expressed through the use of Lennard-Jones (LJ) type potentials. In this model the interaction between a pair of atoms is characterized by a strong repulsive core and a soft attractive term. The functional form of the LJ potential requires the specification of two parameters that depend on the pair of atoms considered, σ_{ij} and ϵ_{ij} ,

$$u_{\text{LJ}}(r_{ij}) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right], \quad (2.2)$$

where $r_{ij} = |\mathbf{r}_{ij}| = |\mathbf{r}_j - \mathbf{r}_i|$ is the interatomic distance. The σ_{ij} and ϵ_{ij} parameters are usually termed the size and energy parameters, respectively, and they correspond to the x-intercept and well depth of the potential as shown in Figure 2.1.

Quite often, σ_{ij} and ϵ_{ij} are obtained by combining parameters that depend on the individual particles $\sigma_i, \sigma_j, \epsilon_i, \epsilon_j$, effectively reducing the number of parameters needed from $2\binom{N}{2}$ to $2(2N)$, where N is the number of atom (interaction site) types.

The combination rules depend on the specific parameter set and, in this

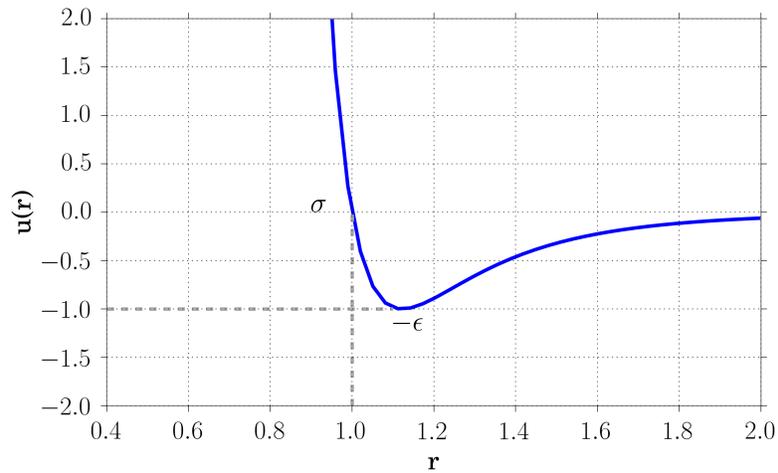


Figure 2.1: The LJ potential with parameters $\sigma = 1$ and $\epsilon = 1$. The σ correspond to the x-intercept of the potential, while ϵ corresponds to the well depth.

thesis, the Lorentz-Berthelot rules are used. According to these rules, σ_i and σ_j are combined using the arithmetic mean, while the geometric mean is used to combine ϵ_i and ϵ_j

$$\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2}, \quad (2.3)$$

$$\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}. \quad (2.4)$$

The LJ and other short-range potentials aimed at describing van der Waals forces decay rapidly with distance, and are typically calculated by truncating the interactions after a certain distance, also termed the cutoff radius.

2.3.2 Electrostatic Forces

The electrostatic interactions are modeled using Coulombic potentials, that depends on the interatomic distance, r_{ij} , and the partial charges, q_i and q_j ,

$$u_C(r_{ij}) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}, \quad (2.5)$$

where ϵ_0 is the dielectric constant in vacuum. The electrostatic potential does not decay rapidly with distance, and cannot be truncated by imposing a cutoff. Strategies have been developed to deal with electrostatic interactions, as explained in Section 2.3.4 below.

2.3.3 Water and Ion Models

The modelling of atoms and molecules requires the parameterization of the potential functions. Parameters can be obtained by fitting molecular properties obtained from *ab initio* calculations,⁶⁷⁻⁶⁹ or by fitting the parameters to reproduce known experimental properties^{2,61} (e.g. hydrogen bond lengths, densities, local structures, hydration energies).

To model water, a simple approach is to use single point charges on the oxygen and hydrogen sites, and to represent the molecule as a sphere centered on the oxygen atom. In the SPC/E (extended simple point charge) water model² used in our simulations, partial charges are positioned on the oxygen and hydrogen atoms according to the geometry shown in Figure 2.2. The masses for the three interaction sites correspond to the masses of the oxygen and hydrogen atoms. The hydrogen atoms lack LJ interaction terms ($\sigma_H = 0$), while the oxygen $\sigma_O = 0.3166$ nm. The

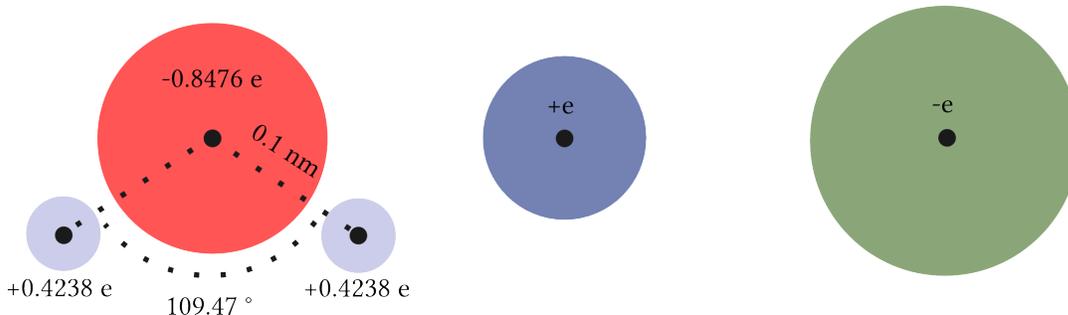


Figure 2.2: Positions of the charges in the models employed in our simulations. The SPC/E water model on the left represents the atoms as point charges, with a O-H distance of ~ 0.1 nm and a HOH angle of ~ 109.47 deg. The alkali and halide ions are modeled as single point charges. The short-range interaction parameters are not shown in the picture. For the water model, the masses of the oxygen and hydrogen atoms were employed for the negative and positive charged sites, respectively. Similarly, the mass for the corresponding element was chosen for the alkali halide models.

parameters are obtained to reproduce experimental properties such as the radial distribution function, diffusion constant, and density at 300K.²

The alkali halide ions used in our simulations are modeled as point charges plus LJ interaction terms.

2.3.4 Infinite Systems

Currently, typical system sizes used in molecular simulation are on the order of $\sim 10^5 - 10^7$ atoms with a record holding simulation of $\sim 10^{12}$ atoms.⁷⁰ As macroscopic systems are much larger (a mole of substance is $\sim 10^{23}$ particles) than molecular simulations can handle, special techniques are usually employed as described below.

A large system can be approximated as a periodic repetition of one of its parts. Using this construction, the simulation can be performed on a small system

that interacts with identical copies that extend in all directions, a construction termed periodic boundary conditions (PBC).

A depiction of a two dimensional cell under PBC is shown in Figure 2.3. The central cell is surrounded by identical copies and, as the system evolves in time, particles that exit from one side of the box are reinserted from the opposite side. As there are an infinite number of copies of the central cell, to calculate the distance between particles it is necessary to specify a particular periodic image. In molecular simulation, the image is chosen so that the distance between the atoms in consideration is minimum. This choice is termed the *minimum image convention* and is implemented by calculating each component of the displacement vector $\mathbf{r}_{ij} = \{x_{ij}, y_{ij}, z_{ij}\}$ as follows

$$x_{ij} = \begin{cases} x_j - x_i - L_x & \text{if } x_j - x_i > L_x/2, \\ x_j - x_i + L_x & \text{if } x_j - x_i < -L_x/2, \\ x_j - x_i & \text{otherwise,} \end{cases} \quad (2.6)$$

where L_x is the box length in the x direction. The calculation is performed similarly for the y_{ij} and z_{ij} components.

The treatment of interactions is simple for the short-range (van der Waals) terms. As already mentioned, since they become negligible at a relatively short distance (e.g. ~ 1 nm), they can be calculated by truncating all interactions at a distance larger than a given cutoff radius.

For the long-range electrostatic interactions, the expression of the potential

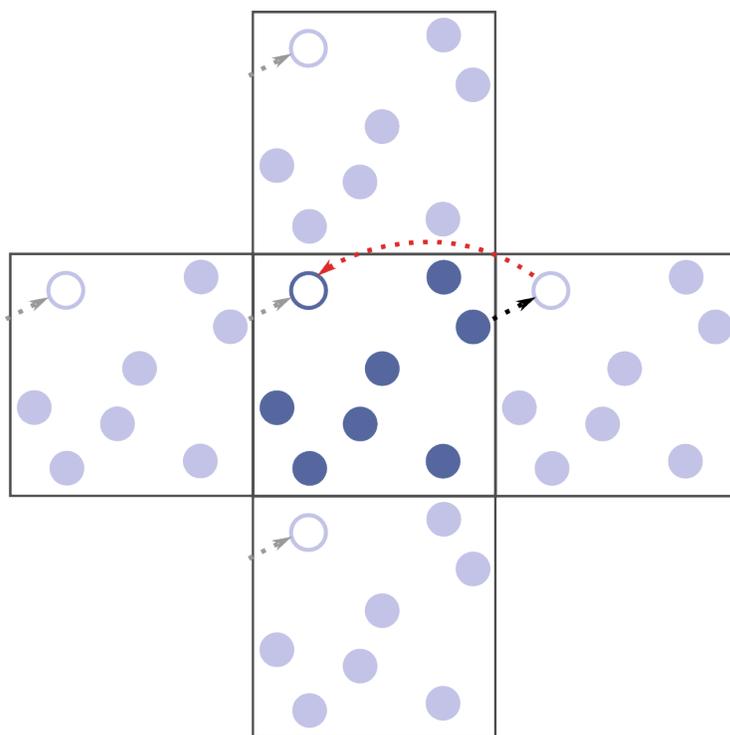


Figure 2.3: Depiction of PBC. The central system is repeated in all directions. A particle exits from the central box on the right side (black arrow) and reenters from the opposite side.

energy felt by a charge in the central cell can be written as

$$u_i = \sum_{\mathbf{n}=-\infty}^{\infty} \sum_j \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{n}|} \quad i \neq j, \quad \text{when } \mathbf{n} = \{0\}, \quad (2.7)$$

where $\mathbf{n} = \{n_a, n_b, n_c\}$ is the displacement of the periodic cell, and \mathbf{r}_{ij} is the displacement vector between the i^{th} and j^{th} atoms in the central cell. Notice that the term $|\mathbf{r}_{ij} + \mathbf{n}|$ corresponds to the position of the charge in the periodic cell defined by \mathbf{n} . The infinite series expressed in Equation (2.7) is conditionally convergent (i.e. its result depends on the way the terms are summed), and special methods have been developed for its evaluation.⁷¹

One of the methods used to calculate electrostatic energies is the Ewald summation,⁷² that was developed to calculate Madelung constants for ionic crystals, and can be used to calculate the electrostatic interactions in a simulation subjected to PBC. It was later shown⁷¹ that the evaluation of the sum of Equation (2.7) gives different results depending on the medium surrounding the infinite system and, in the case of the original Ewald summation,⁷² the medium is a perfect conductor (its relative dielectric constant $\epsilon_r = \infty$.)

In the Ewald summation method, the contribution of the Coulombic potential is split into short-range and long-range parts by adding and subtracting an $\text{erf}(\kappa r)/r$ term

$$\frac{1}{r} = \frac{1 + \text{erf}(\kappa r) - \text{erf}(\kappa r)}{r} = \frac{\text{erfc}(\kappa r)}{r} + \frac{\text{erf}(\kappa r)}{r}, \quad (2.8)$$

$$\text{erf}(r) = \frac{2}{\sqrt{\pi}} \int_r^{\infty} e^{-z^2} dz, \quad (2.9)$$

$$\text{erfc}(r) = 1 - \text{erf}(r). \quad (2.10)$$

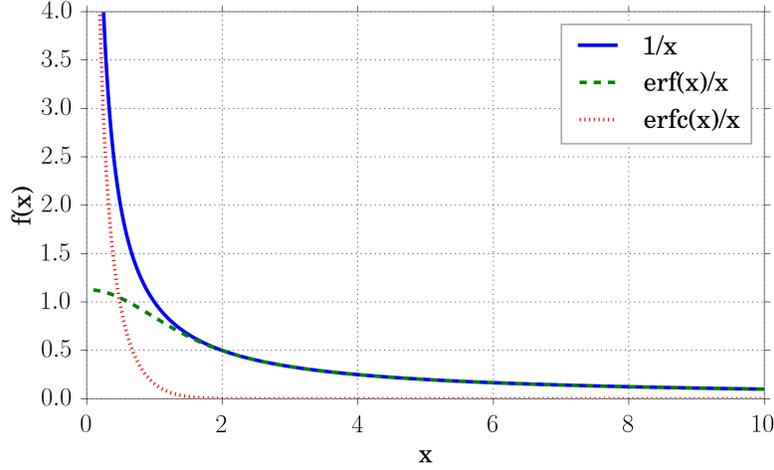


Figure 2.4: The short-range and long-range contribution to the electrostatic potential for $\kappa = 1$.

The first term on the right of Equation (2.8) decays rapidly with r (Figure 2.4) and, similar to the other short-range interactions, its contribution to the sum can be calculated directly by using a cutoff, while the second term can be Fourier transformed and converges rapidly in reciprocal space.⁷³ The parameter κ can be used to modulate the convergence of the two terms as explained below.

The expression of the Ewald summation for the potential energy felt by an charged site i in a cell of volume V can be expressed as follows⁷³

$$u_i = u_i^r + u_i^k + u_i^{\text{self}}, \quad (2.11)$$

$$u_i^r = \sum_{\mathbf{n}} \sum_j \frac{q_i q_j \text{erfc}(\kappa |\mathbf{r}_{ij} + \mathbf{n}|)}{|\mathbf{r}_{ij} + \mathbf{n}|} \quad i \neq j \text{ when } \mathbf{n} = \{0\}, \quad (2.12)$$

$$u_i^k = \frac{1}{\pi V} \sum_{\mathbf{k} \neq 0} \sum_j q_i q_j \left(\frac{4\pi^2}{|\mathbf{k}|^2} \right) \exp\left(-\frac{|\mathbf{k}|^2}{4\kappa^2}\right) \cos(\mathbf{k} \cdot \mathbf{r}_{ij}), \quad (2.13)$$

$$u_i^{\text{self}} = -\frac{\kappa}{\pi^{1/2}} q_i^2. \quad (2.14)$$

The first term in Equation (2.11) u_i^r is simply the direct sum, similar to Equation

(2.7), of the rapidly decaying part, and \mathbf{n} is truncated to include charges up to a certain cutoff (often the truncation will only include the central cell.) The second term u_i^k is the sum of the slowly varying part of the potential over the reciprocal lattice vectors \mathbf{k} , and u_i^{self} is a correction needed to remove the interaction of a particle with itself, introduced by the term u_i^k . The parameter κ can be tuned to modulate the convergence rate of either of the sums and is often set to $5/L$, where L is the box length. Given the value $\kappa = 5/L$, to achieve convergence in reciprocal space, ~ 200 reciprocal vectors \mathbf{k} are required.⁷³

To retrieve the total electrostatic potential of the cell, it is sufficient to sum over all particles i in the following way

$$U = \frac{1}{2} \sum_i (u_i^r + u_i^k) + \sum_i u_i^{\text{self}}. \quad (2.15)$$

Additionally, if the system is assumed to be surrounded by a medium with a certain dielectric constant ϵ_r , a correction that depends on the dipole of the central cell must be added to the total

$$U_{\text{dip}} = \frac{2\pi}{(1 + 2\epsilon_r)V} \left| \sum_i q_i \mathbf{r}_i \right|^2. \quad (2.16)$$

Notice that, in the case of ionic systems (and in the original Ewald formulation), this term is neglected, corresponding to a system surrounded by a perfect conductor ($\epsilon_r = \infty$).

In terms of computational time, the Ewald summation method (when a constant cutoff on the short-range part is applied) scales as $O(N^2)$, where N is the number of charges. Most simulation programs, however, implement other variants such as particle-mesh Ewald⁷⁴ (PME) that evaluates the reciprocal term (i.e., the

Fourier transform of the charge density) on a grid using the fast Fourier transform algorithm.⁷⁵ This reduces the time required to evaluate the electrostatic potential to $O(N \log(N))$.

2.4 System Initialization

Generating the initial configuration is one of the most important aspect of molecular simulation, where the computational scientist sets up “experiments” targeted towards the validation of a specific hypothesis. In this section we explain some of the techniques used to initialize atomic positions, as velocities are usually randomly sampled from a Maxwell-Boltzmann distribution.

2.4.1 Liquid Systems

Liquid systems can be initialized by placing the molecules on an evenly spaced grid. This method can be effective when the objects to be placed are of similar size so that they are able to fit onto a grid without overlapping with each other.

Another option is to insert particles at random and to perform a neighbor distance test to avoid particle collision. However, as the box becomes filled, random insertions will likely cause collisions and, in high density systems, the many attempts required can make this method impractically slow. Despite this drawback, this method is very general as it doesn’t require specification of a particular grid size, and for this reason it was used to initialize the liquid phase simulations in this thesis.

2.4.2 Crystals

Crystals can be generated given information about the unit cell. Typically, the parameters required to specify the initial structures are the unit cell lengths a , b and c , the unit cell angles, α , β and γ , the fractional coordinates of the atoms, and the space group.

This information can be usually found electronically as Crystallographic Information Files,⁷⁶ that contain the data required to generate the unit cell. The unit cell can then be repeated and translated along the three crystal axes to yield arbitrarily large crystals. Spherical crystals can be obtained by creating a large crystal and discarding all molecules or ions beyond a certain cutoff radius. Crystals can also be extracted from molecular simulation after nucleating and growing a crystal, however, as such crystals are often imperfect, care must be taken to include the same number of positive and negative ions necessary to maintain electroneutrality.

2.4.3 Crystal-liquid Interfaces

Crystal-liquid interfaces can be generated by placing a crystal in a simulation box and then positioning liquid particles in the empty spaces using the techniques described above. Notice that, when using the crystals generated using the procedure mentioned above, the results may vary greatly depending on the crystal face exposed and/or the presence of defects.

2.5 Time Evolution

The equations of motion described by Equation (2.1) can be integrated numerically to yield a trajectory. The method used in the simulations performed in this thesis is the leap-frog method, termed such because the velocities are updated at a different time step than the positions

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t + \Delta t/2), \quad (2.17)$$

$$\mathbf{v}_i(t + \Delta t/2) = \mathbf{v}_i(t - \Delta t/2) + \Delta t \mathbf{a}_i(t), \quad (2.18)$$

$$\mathbf{a}_i(t) = \frac{\mathbf{f}_i(t)}{m_i}, \quad (2.19)$$

$$\mathbf{f}_i(t) = -\nabla u_i(\mathbf{r}_i). \quad (2.20)$$

This method is particularly effective in molecular dynamics because it is time reversible (integrating forward n steps and backward n steps brings the system to the same initial position) and is able to keep the total energy constant to a good approximation. Care should be taken in choosing a time step Δt that is small compared to the motion in the system (a typical value is $\sim 1-2$ femtoseconds).

One drawback of the leapfrog integrator is that it doesn't calculate velocities at integer time steps, and these need to be approximated from the velocities at half time steps as follows

$$\mathbf{v}_i(t) = \frac{\mathbf{v}_i(t + \Delta t/2) + \mathbf{v}_i(t - \Delta t/2)}{2}. \quad (2.21)$$

2.5.1 Constraints

The vibration of chemical bonds and bond angles can be approximated in molecular simulation using a harmonic potential. However, as the vibrational frequency is often high compared to the intermolecular motion, in certain models such as the SPC/E water model, bonds and angles are kept at a fixed value (the molecule is modeled as a rigid body). In simulation programs, rigid bonds and angles are usually kept fixed using a constraining algorithm.

A constraint is a function that introduces a dependency between coordinates, for example, a constraint that imposes a distance d between two sites l and m can be expressed as follows

$$g(\mathbf{r}, t) = |\mathbf{r}_l - \mathbf{r}_m| - d = 0 \quad (2.22)$$

$$\mathbf{r} = \{\mathbf{r}_1 \dots \mathbf{r}_N\}. \quad (2.23)$$

If the constraints are holonomic (i.e., they depend only on the coordinates \mathbf{r} and time t), the equations of motion for the constrained system can be obtained from the Lagrangian

$$\mathcal{L} = \mathcal{K} - \mathcal{V} + \boldsymbol{\lambda} \cdot \mathbf{g}, \quad (2.24)$$

$$\mathcal{K} = \frac{1}{2} \mathbf{M} \mathbf{v} \cdot \mathbf{v}, \quad (2.25)$$

$$\mathcal{V} = \sum_i \sum_{j>i} u(r_{ij}), \quad (2.26)$$

where $\boldsymbol{\lambda} = \{\lambda_0, \dots, \lambda_K\}$ are the K Lagrange multipliers associated with the K constraints $\mathbf{g} = \{g_0, \dots, g_K\}$, \mathcal{K} and \mathcal{V} are, respectively, the kinetic and potential energy of the system. \mathbf{M} is a diagonal $(3N, 3N)$ matrix containing the mass of each particle

repeated three times (one for each of the x, y, z coordinates)

$$\mathbf{M} = \text{diag}(m_1, m_1, m_1, m_2, m_2, m_2, \dots, m_N, m_N, m_N). \quad (2.27)$$

The resulting equations of motion can be expressed in matrix notation⁷⁷

$$\ddot{\mathbf{r}} = (\mathbf{I} - \mathbf{TB})\mathbf{M}^{-1}\mathbf{f} - \mathbf{TB}\dot{\mathbf{r}}, \quad (2.28)$$

$$\mathbf{T} = \mathbf{M}^{-1}\mathbf{B}^T(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}, \quad (2.29)$$

where \mathbf{B} is the $(K, 3N)$ matrix containing the partial derivatives with respect to the particle coordinates of the constraints

$$B_{ki} = \frac{\partial g_k(\mathbf{r}, t)}{\partial x_i}. \quad (2.30)$$

The matrix $\mathbf{I} - \mathbf{TB}$ projects the acceleration vector $\mathbf{M}^{-1}\mathbf{f}$, and the last term of Equation (2.28) represents centripetal forces due to rotation.

In our simulations we adopted the LINC algorithm⁷⁷, which is a method to efficiently integrate the constrained equations of motion discussed above. The integration of Equation (2.28) using the leap-frog algorithm can be performed by first conducting an unconstrained step to obtain the new coordinates and velocities (termed \mathbf{r}^u and \mathbf{v}^u), and then by projecting them using the matrix $\mathbf{I} - \mathbf{TB}$. In computer implementations, an additional term $\mathbf{T}(\mathbf{B}\mathbf{r} - \mathbf{d})/\Delta t$ is necessary to prevent accumulation of numerical errors⁷⁷

$$\mathbf{v}(t + \Delta t/2) = (\mathbf{I} - \mathbf{TB})\mathbf{v}^u(t + \Delta t/2) - \frac{\mathbf{T}(\mathbf{B}\mathbf{r}(t) - \mathbf{d})}{\Delta t}, \quad (2.31)$$

$$\mathbf{r}(t + \Delta t) = (\mathbf{I} - \mathbf{TB})\mathbf{r}^u(t + \Delta t) + \mathbf{T}\mathbf{d}. \quad (2.32)$$

The slow step in this algorithm is the calculation of the matrix \mathbf{T} , as it contains the inversion of $(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1}$. In the LINCS method, the inversion is performed efficiently by rearranging the matrix $\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T$ into a $\mathbf{I} - \mathbf{A}$ form and approximating its inverse using the series expansion

$$(\mathbf{I} - \mathbf{A})^{-1} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \dots \quad (2.33)$$

The matrix $\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T$ can be transformed into the $\mathbf{I} - \mathbf{A}$ form by using the diagonal matrix \mathbf{S} , defined as the inverse square root of the diagonal elements of $\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T$ as follows

$$(\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T)^{-1} = \mathbf{S}(\mathbf{S}\mathbf{B}\mathbf{M}^{-1}\mathbf{B}^T\mathbf{S})^{-1}\mathbf{S} = \mathbf{S}(\mathbf{I} - \mathbf{A})^{-1}\mathbf{S}. \quad (2.34)$$

The advantages of the series expansion in Equation (2.33) are a high storage efficiency (since \mathbf{A} is usually sparse) and a faster computation time, compared to standard matrix inversion. Nevertheless, Equation (2.33) is only valid when the absolute value of all the eigenvalues is less than 1, a condition that is not verified when there is a high degree of bond connectivity. The structure of matrix \mathbf{A} also allows for easy parallelization, making the algorithm scalable to very large simulations.

2.6 Temperature and Pressure Control

A system described by the equations of motion presented in Equation (2.1), will conserve energy and, if we enclose the system in a container or apply periodic boundary conditions, will conserve its volume (or density). Many systems of interest, like a laboratory environment, are best described by the conditions of constant temperature and pressure. In the subsections below we will describe some of the

algorithms used to keep temperature and pressure to a fixed value in molecular dynamics simulations.

2.6.1 Velocity Rescaling Thermostat

The instantaneous temperature in a molecular simulation can be calculated from the kinetic energy of the system \mathcal{K}

$$T = \frac{2\mathcal{K}}{k_{\text{B}}N_f}, \quad (2.35)$$

where N_f is the number of degrees of freedom which, for a system subject to constraints, is equal to the number of coordinates minus the number of constraints.

A very simple approach, called the Berendsen thermostat,⁷⁸ consists of fixing the temperature by rescaling the velocities so that the kinetic energy assumes a target value \mathcal{K}' consistent with the target temperature T' . The scaling factor α , that is applied to the velocities is defined as

$$\alpha = \sqrt{\frac{\mathcal{K}'}{\mathcal{K}}}, \quad (2.36)$$

$$\mathcal{K}' = \frac{1}{2}k_{\text{B}}N_fT'. \quad (2.37)$$

The factor α is usually applied with a predetermined frequency. The procedure outlined above will keep the kinetic energy constant but will not reproduce fluctuations characteristic of the canonical ensemble, it was, however, used in some of the calculations in this thesis to relax the systems to our target temperatures before running the simulations for data collection.

In order to retrieve the correct sampling for the canonical ensemble, a dif-

ferent method, named the *velocity rescaling thermostat*,⁷⁹ was used. In this method, the rescaling factor can be chosen to obtain a target kinetic energy that is sampled from the canonical ensemble distribution

$$P(\mathcal{K}_t) \propto \mathcal{K}_t^{\frac{N_f}{2}-1} e^{-\frac{\mathcal{K}_t}{k_B T}}. \quad (2.38)$$

While, in principle, the kinetic energy could be rescaled to match a value sampled from this distribution at each time step, in practice this causes abrupt changes in velocity. In the velocity rescaling thermostat,⁷⁹ this value is computed by evolving the kinetic energy according to an auxiliary stochastic process. The kinetic energy is rescaled to the target value gradually by updating the previous kinetic energy value according to the equation

$$\mathcal{K}(t + \Delta t) = \mathcal{K}(t) + \Delta\mathcal{K} \quad (2.39)$$

$$\Delta\mathcal{K} = (\mathcal{K}' - \mathcal{K}(t)) \frac{\Delta t}{\tau} + 2 \sqrt{\frac{\mathcal{K}' \mathcal{K}(t)}{N_f}} \frac{dW}{\sqrt{\tau}}, \quad (2.40)$$

where dW is a Wiener noise (a random displacement) and τ is a parameter, with the dimensions of time, that can regulate how fast the system equilibrates to the target kinetic energy.

The velocity rescaling thermostat was used to control the temperature in all of the simulations performed in this thesis.

2.6.2 Berendsen Barostat

For a system consisting of N particles, volume V , and at an instantaneous temperature T , the microscopic pressure can be calculated as

$$p = k_{\text{B}}T \frac{N}{V} + \frac{1}{3V} \sum_i \mathbf{f}_i \cdot \mathbf{r}_i, \quad (2.41)$$

where \mathbf{f}_i is the force acting on the i^{th} particle at position \mathbf{r}_i . Control of the pressure can be applied, in a manner similar to the velocity-rescaling technique, by rescaling the particle positions and volume of the system by a factor α

$$\dot{V} = 3\alpha V, \quad (2.42)$$

$$\dot{\mathbf{x}} = \mathbf{v} + \alpha \mathbf{x}. \quad (2.43)$$

In the case of the Berendsen barostat,⁷⁸ in order to bring the system pressure to a value p' , the coefficient can be calculated as

$$\alpha = c_{\beta}(p' - p)/(3\tau_p), \quad (2.44)$$

where c_{β} is the isothermal compressibility of the system and the τ_p is a parameter with the units of time. As c_{β} and τ_p appear as a ratio, only their relative magnitude is important, and they affect how fast the system will equilibrate to the target pressure.

The Berendsen barostat does not produce the correct energy fluctuations for the isobaric-isoenthalpic ensemble, but is often used for its simplicity and to quickly bring the system to a target pressure. In this thesis, this barostat was used

when fluctuations generated by other methods would make the system unstable (more details are provided in Chapter 4), and was also used in other instances, after testing that it did not produce different results than more advanced methods such as the Parrinello-Rahman barostat described below.

2.6.3 Parrinello-Rahman Barostat

The Parrinello-Rahman barostat⁸⁰ is another method to control pressure in a molecular dynamics simulation. The equations of motion are modified such as to allow the simulation cell vectors to evolve in time, in a way that produces the correct statistical properties of the isobaric-isoenthalpic ensemble.

If we identify the cell vectors with \mathbf{a} , \mathbf{b} and \mathbf{c} , we can construct a matrix \mathbf{h} by stacking the vectors as its rows

$$\mathbf{h} = \{\mathbf{a}, \mathbf{b}, \mathbf{c}\}. \quad (2.45)$$

The equations of motion for the system subject to an isotropic pressure p can be obtained from the Lagrangian

$$\mathcal{L} = \mathcal{K} - \mathcal{V} + \frac{1}{2}M\text{Tr}(\dot{\mathbf{h}}^T\dot{\mathbf{h}}) - pV, \quad (2.46)$$

where the additional terms can be interpreted as the kinetic and potential energy related to a piston. The term $\frac{1}{2}M\text{Tr}(\dot{\mathbf{h}}^T\dot{\mathbf{h}})$ represents the kinetic energy of a piston with mass M that is compressing the cell, while pV is the potential energy driving the compression.

The resulting equations of motion can be expressed as a function of the scaled

coordinates of the particles, $\mathbf{s}_i = \mathbf{h}^{-1}\mathbf{r}_i$,

$$\ddot{\mathbf{s}}_i = - \sum_{i \neq j} \frac{du}{dr_{ij}} \frac{1}{m_i r_{ij}} (\mathbf{s}_i - \mathbf{s}_j) - \mathbf{G}^{-1} \dot{\mathbf{G}} \dot{\mathbf{s}}_i, \quad (2.47)$$

$$\ddot{\mathbf{h}} = \frac{1}{M} (\boldsymbol{\pi} - p) \boldsymbol{\sigma}. \quad (2.48)$$

The other quantities defined in the above equations are $\boldsymbol{\sigma}$, a matrix containing the direction of the reciprocal vectors, \mathbf{G} also termed the metric tensor, and $\boldsymbol{\pi}$, the stress tensor defined as follows

$$\boldsymbol{\sigma} = \{\mathbf{b} \times \mathbf{c}, \mathbf{c} \times \mathbf{a}, \mathbf{a} \times \mathbf{b}\}, \quad (2.49)$$

$$V \boldsymbol{\pi} = \sum_i m_i \mathbf{v}_i \mathbf{v}_i^T - \sum_i \sum_{j>i} \frac{du}{dr_{ij}} \frac{1}{r_{ij}} \mathbf{r}_{ij} \mathbf{r}_{ji}^T, \quad (2.50)$$

$$\mathbf{G} = \mathbf{h}^T \mathbf{h}. \quad (2.51)$$

If the system is far from equilibrium the Parrinello-Rahman barostat can generate large volume fluctuations and, in this case, the Berendsen barostat can be used to quickly bring the system to equilibrium before switching to the Parrinello-Rahman barostat. The Parrinello-Rahman barostat was used to control pressure in some of the simulations in this thesis. As will be discussed in Chapter 4, the Berendsen barostat was preferred for simulations where the fluctuations produced by the Parrinello-Rahman barostat caused the system to become unstable.

Chapter 3

The Birth of NaCl Crystals: Insights from Molecular Simulations

3.1 Introduction

As stated in Chapter 1, the initial process by which one phase transitions to another is termed nucleation. Nucleation has a central role in a wide range of physically interesting processes such as crystal and rock formation,^{81,82} drug development,^{83,84} and the formation of atmospheric aerosols.^{85,86} There have been a number of interesting experimental studies of crystal nucleation,^{28,29,87-89} but there remain significant challenges for current experimental methods, particularly for the analysis of early prenucleation events. Therefore, to augment experimental studies, molecular simulations are being increasingly used to gain physical insight into crystal nucleation and growth.^{43-45,47,49,50,66,90}

The usual approach to nucleation, classical nucleation theory (CNT), is a subject of current debate.^{26,91,92} In particular, the assumption that early stage potential nuclei can be characterized by properties of the nascent bulk phase (plus a surface term) is being challenged by both experiments and molecular simulations. One such example is provided by CaCO_3 , where experiments suggest that prenucleation clusters serve as precursors to nucleation.²⁵ There is also evidence that crystallization can occur via a two-step mechanism consisting of an initial density “transition”, followed by a slower ordering transition whereby the dense liquid region becomes geometrically ordered.^{49,93,94}

In the present Chapter we use direct molecular dynamics simulations to investigate the nucleation of NaCl nanocrystals. Supersaturated aqueous NaCl solutions are a good choice for detailed study because relatively accurate models exist,¹ and previous work has shown that spontaneous nucleation occurs on simulation timescales.^{47,49,50,66} Earlier direct simulation studies of NaCl crystallization from supersaturated solution suggest a process by which less ordered and more hydrated NaCl clusters evolve with time into a largely anhydrous crystalline arrangement.^{49,59} There is also some evidence that the two-step mechanism applies with a large density (concentration) fluctuation preceding any spatial ordering.⁴⁹

Nucleation and crystal growth have also been considered employing indirect simulation techniques. In an early investigation Zahn⁴⁷ approached the problem using trajectory sampling methods. He noticed that small NaCl clusters tended to prefer a Na^+ ion at the center, and suggested that such clusters might be important in nucleation. A recent study reported by Zimmerman et al.⁶⁶ used a seeded trajectory approach to determine the sizes of critical nuclei, ion attachment frequencies, and nucleation rates for NaCl in supersaturated solutions.

However, despite earlier efforts, the factors that influence the formation of critical nuclei have not been determined, and this is particularly true of the very early stages or the birth of the crystal. For example, it is interesting to ask if size alone determines the relative stability of small NaCl clusters, or if other factors such as the cluster geometry are important. This is one of the questions addressed in the present Chapter.

Spontaneous nucleation events for NaCl are typically rare (on simulation timescales) except at very high solute concentration, and multiple nucleation events must be observed in order to draw any meaningful conclusions. Additionally, it is necessary to define, detect, and follow in time a great many NaCl clusters that do not achieve nucleation. Here we carry out multiple direct simulations sufficient to generate a number of nucleation events. We develop methodology to define and detect crystal-like NaCl clusters, and to follow them in time from very early stages (~ 6 ions) until nucleation is achieved or, much more frequently, the cluster dissolves back into solution. An important conclusion of our analysis is that cluster size is not the only factor influencing cluster lifetime and the probability of nucleation. By introducing a new parameter called the cluster “crystallinity” we show that cluster geometry is also a very significant factor influencing cluster lifetime and nucleation probability; moreover, this is true for clusters as small as six ions.

The remainder of this Chapter is divided into three parts. The models and simulation methods are described in Section 3.2, results are presented in Section 3.3, and our conclusions are summarized in Section 3.4.

	σ (nm)	ϵ (kJ mol ⁻¹)
Na ⁺	0.2160	1.4754533
Cl ⁻	0.4830	0.0534924

Table 3.1: Joung-Cheatham Lennard-Jones parameters for NaCl.¹

3.2 Models and Methods

3.2.1 Simulation Details

In our simulations we adopt the Joung-Cheatham parameter set¹ for the Na⁺ and Cl⁻ ions paired with the SPC/E water model.² For this parameter set, the saturation mole fraction of NaCl, x_{NaCl} , at 298 K was calculated^{60,64,95} to be ~ 0.06 using the chemical potential route, while results from direct coexistence methods^{1,96-98} report a value of $\sim 0.09-0.11$ (experimentally,⁹⁹ the value is ~ 0.10). With this force field all nonbonded interactions consist of Lennard-Jones (LJ) plus electrostatic terms such that the site-site pair potentials have the form

$$u(r_{ij}) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right] + \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}, \quad (3.1)$$

where ϵ_0 is the permittivity of free space, q_i and q_j are the partial charges on sites i and j , and σ_{ij} and ϵ_{ij} are the usual LJ length and energy parameters, respectively. The Joung-Cheatham parameters for Na⁺ and Cl⁻ are given in Table 3.1, and σ_{ij} and ϵ_{ij} are obtained from these and the SPC/E parameters² using the Lorentz-Berthelot combining rules, $\sigma_{ij} = (\sigma_i + \sigma_j)/2$, and $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$.

Simulations were carried out under *NPT* conditions employing the GRO-MACS¹⁰⁰ molecular dynamics package, version 4.5.4. The temperature was controlled using the velocity-rescale thermostat⁷⁹ with a relaxation time of 0.1 ps, and

the pressure was kept constant at 1 bar using a Berendsen barostat⁷⁸ with a compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$ and a relaxation time of 1.0 ps. A timestep of 2 fs was used in all simulations. A spherical cutoff of 0.9 nm was applied to the pair potentials, and the long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) method.⁷⁴

Details of all simulations performed are summarized in Table 3.2. Initial simulations were carried out for x_{NaCl} ranging from 0.20 to 0.30, and the lowest concentration where a nucleation event was observed within 200 ns at 300 K was $x_{\text{NaCl}} = 0.22$. This time frame is feasible for an investigation employing direct molecular dynamics simulations, so we focus on this concentration. More nucleation events occur at higher concentrations, but the occurrence of many simultaneous nucleations can complicate the analysis. Note that $x_{\text{NaCl}} = 0.22$ is, based on the estimates provided above, between two and four times the saturation value at 298 K. The atomic coordinates for the system were collected at intervals of 0.04 nanoseconds, and the coordinate sets collected with this time interval are referred to as time frames elsewhere in the paper.

The analysis given in this paper is based on five “replica” simulations at $x_{\text{NaCl}} = 0.22$ carried out as follows. Initially, 7040 NaCl pairs and 24960 water molecules were distributed randomly on a lattice within the simulation cell. The system was then “equilibrated” at 400 K and five configurations were extracted at intervals of ~ 10 ns. Five simulations, initiated with these configurations, were cooled to 300 K, and allowed to evolve for ~ 500 ns, as summarized in Table 3.2. The number of nucleation events observed in each simulation is also given in Table 3.2.

Water	Ions	x_{NaCl}	Length (ns)	Nuclei
25600	12800	0.20	157	0
24960	14080	0.22	187	3
24000	16000	0.25	515	5
22400	19200	0.30	202	n.a.
24960	14080	0.22	523	1
			547	2
			532	3
			530	4
			550	3

Table 3.2: Summary of the simulations performed. The number of water molecules (Water), the number of ions (Ions), the mole fraction x_{NaCl} , the run length (Length), and the number of clusters that achieved nucleation (Nuclei) are indicated. At $x_{NaCl} = 0.3$, multiple nucleations occurred resulting in a polycrystal. The five simulations at $x_{NaCl} = 0.22$ listed at the bottom of the table were used for the statistical analysis given in the text.

3.2.2 Cluster Detection

The aim of the present work is to detect NaCl clusters and follow their evolution in time in order to understand the factors influencing (or not) the probability of achieving crystal nucleation. From the perspective of classical nucleation theory (CNT), these clusters can be viewed as subcritical nuclei unless and until nucleation is achieved. Since we are trying to observe the very early stages of a phase transition, it is necessary to find some measure to distinguish local liquid-like and solid-like structures, and to determine which ions are part of the same solid-like cluster. Generally, such measures are based on local environments, and several approaches to this problem have been employed. These include methods based on ion connectivity,^{58,59} local ion density,¹⁰¹ local solvent density,⁴⁷ and local bond order parameters.^{102,103} The last approach has been used extensively in studies of ice nucleation.^{104–106}

In the present paper, the computational procedures used to detect (define) ion clusters and follow their time evolution consists of three steps: filtering, clustering, and entity resolution, which are described in detail below.

Filtering

The purpose of this step is to identify for each configuration along a trajectory the ions that are most likely to be part of a local solid-like structure. To do this we follow the general bond-orientational order parameter approach of Steinhardt et. al.¹⁰⁷ This method defines potential order parameters of the form

$$q_l = \left[\frac{4\pi}{2l+1} \sum_{m=-l}^l |q_{lm}|^2 \right]^{1/2},$$

where

$$q_{lm} = \frac{1}{N} \sum_{r_i=r_1}^{r_N} Y_l^m(\theta(\mathbf{r}_i), \phi(\mathbf{r}_i)).$$

In the above equation, Y_l^m is a spherical harmonic, \mathbf{r}_i is the position vector of the i^{th} neighbor with respect to a central ion, and $\theta(\mathbf{r}_i)$, $\phi(\mathbf{r}_i)$ are, respectively, the polar and azimuthal angles with respect to an arbitrary frame of reference. The actual choice of reference frame is irrelevant as the order parameter q_l is rotationally invariant.

The objective is to find the q_l that best distinguishes between ions with liquid-like and solid-like local environments. To do this we examined q_l distributions for $l = 2, 4, 6, 8$, and 10 , using a spherical crystal of ~ 2000 ions to represent the solid phase, and a supersaturated solution at $x_{NaCl} = 0.20$ (that did not nucleate) as representative of the liquid phase. Both reference systems were held at 300 K. Based on this investigation, the order parameter q_8 provides the best separation

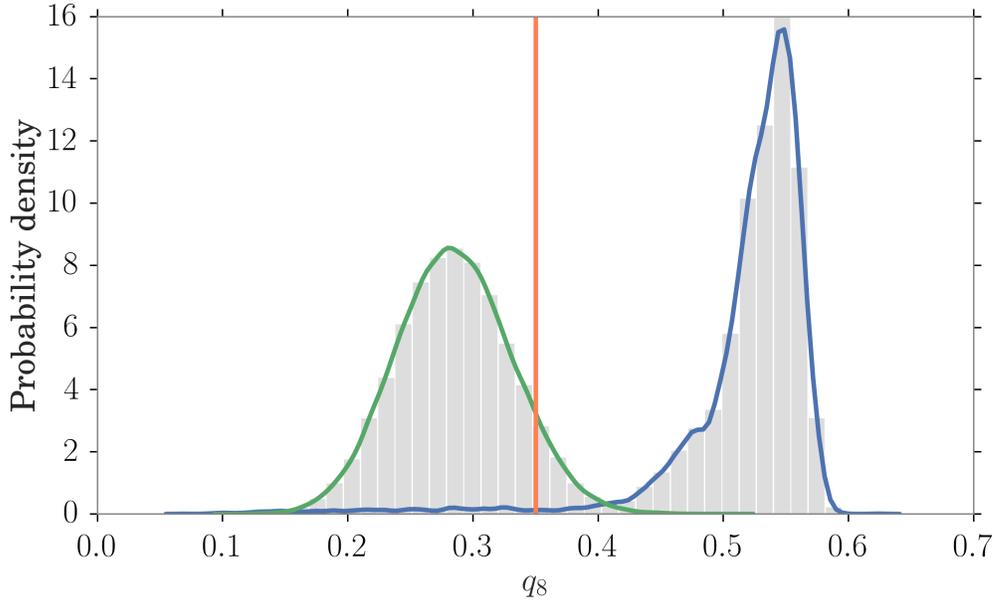


Figure 3.1: Density of q_8 values in a sample crystal (~ 2000 ions)(blue) and a supersaturated solution ($x_{NaCl} = 0.20$) (green). The two distributions show excellent separation. The vertical orange line indicates the selected threshold ($q_8 > 0.35$).

between liquid and solid phase distributions (Figure 3.1). Also, by testing a reference solution of lower concentration ($x_{NaCl} \approx 0.025$), we found that the q_8 distribution is largely independent of concentration (unlike order parameters based on local ion densities¹⁰¹). In the remainder of this chapter we refer to q_8 as the bond order parameter.

The order parameter q_8 is used to filter out ions that are part of liquid-like disordered structures. From Figure 3.1, we see that the best separation point between liquid-like and solid-like structures occurs at $q_8 \approx 0.4$. However, we select $q_8 = 0.35$ as the threshold for our analysis to capture structures from the upper end of the solution distribution. As clusters that achieve nucleation form spontaneously from solution, selecting a lower threshold value allows us to trace nucleation processes from the very initial stages. Selecting a very low threshold would include too

many transient, short-lived structures, while a very high threshold would filter out interesting prenucleation stages. The value of 0.35 selected proved to be a good compromise between these two factors. We verified that our results and conclusions are not affected by small variations in the selected threshold value. Note that with the 0.35 threshold only $\sim 10\%$ of the ions are classified as solid-like for each configuration. It is also important to emphasize that in our simulations all clusters that eventually achieved nucleation originated within the set of ions identified as solid-like; thus our filtering process does not eliminate any interesting events.

Cluster Identification

While order parameter filtering efficiently detects ions in solid-like local environments, another algorithm is necessary to identify which of the selected (solid-like) ions are part of the same cluster or aggregate. For this purpose we adopted the density-based cluster algorithm DBSCAN¹⁰⁸ as implemented in the Python library scikit-learn.¹⁰⁹ In the DBSCAN algorithm, clusters are identified by means of two parameters, that in the present context are a distance (ϵ) and the number of neighboring ions (η) within that distance. To find ions connected to each other, an ion is selected at random, and the number of ions within ϵ are counted. If the number of ions within ϵ is equal to or greater than η , then the central ion is labelled as *core* and its neighbors are labelled as *border*, otherwise, the ion is labelled as *noise* (Figure 3.2). The procedure is repeated by picking an ion from the *border*. If that ion is *core*, it will generate a new *border* and the procedure is repeated recursively. Eventually, no more points are reachable and we label all of the collected *core* and *border* ions as part of the same cluster. The procedure is then repeated by choosing another ion at random (that might belong to a new cluster) until all ions are processed.

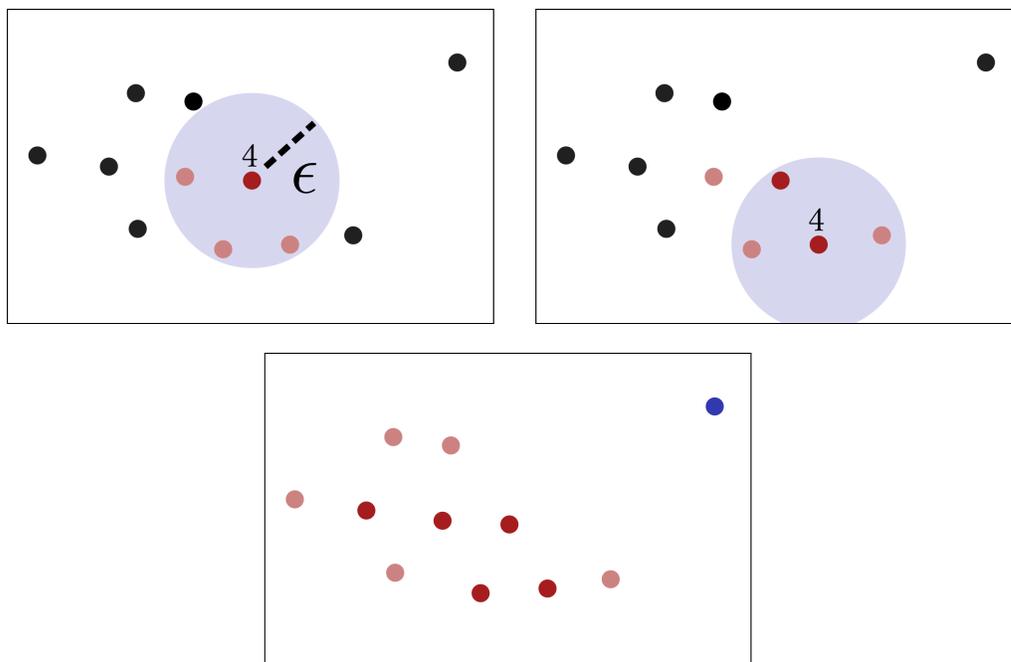


Figure 3.2: Schematic diagram illustrating the DBSCAN algorithm for $\eta = 4$. Top left panel: an ion is chosen at random and ions within ϵ are counted; the value is 4 so the ion is marked as *core* (red). Its neighbours are labeled as *border* (pink). Top right panel: the procedure is repeated for one of the border ions which becomes a core ion. Bottom panel: all ions in the cluster have been processed, and no more ions are reachable from the core ions. The isolated ion is labeled as noise (blue).

The value of ϵ is the minimum distance for which two ions, are considered part of the same aggregate. A high value of ϵ causes some clusters that are close but not connected to each other to be aggregated into a single cluster. On the other hand, a small value of ϵ would fail to connect ions that are effectively bonded (every ion would be detected as a single cluster). The value of η acts as a filter for very small clusters, as every cluster that has fewer than η ions is discarded. A very low value of η detects too many short-lived fluctuations that are difficult to follow and analyse, while a high value would detect only large clusters providing no opportunity to investigate early stages of cluster formation. In the present analysis the values $\epsilon = 0.3$ nm and $\eta = 6$ ions are used. The value 0.3 nm is approximately the NaCl bond length, and $\eta = 6$ ions was chosen to be the minimum size cluster. These values proved to work well in practice, avoiding the problems noted above.

Time Resolution of Clusters

We wish to follow the time evolution of all clusters identified as described above. Therefore, it is necessary to determine how clusters are connected in time. This would be a simple question if we had only a single cluster that grew or dissolved as time progressed. However, in the present case we have many clusters that continually change (divide, grow, dissolve) as the simulation advances in time. Therefore, we need to clearly define how we identify and follow clusters in time. Our general approach is to find “similar” clusters in successive configurations (using an appropriate similarity index), and build a time connectivity graph using the following procedure.

First we initialize an empty graph where each cluster detected appears as a node labelled $c_i(s)$, where i indicates the cluster and s the time frame (configuration).

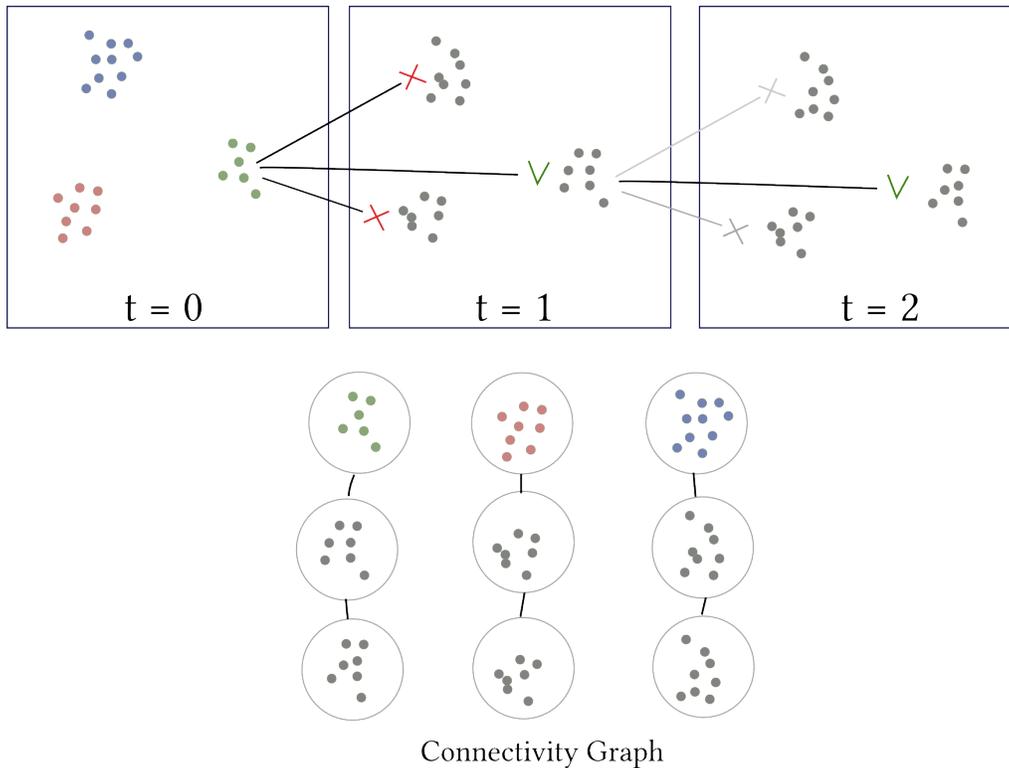


Figure 3.3: Schematic diagram illustrating the cluster time resolution algorithm. At $t = 0$ the clusters are labeled employing the DBSCAN algorithm. A cluster is compared with clusters in the following frame, and a connection is made if the clusters are sufficiently similar, judged by the Jaccard index as described in the text. The procedure is repeated recursively for the following frames. By performing this procedure for all clusters, it is possible to obtain a connectivity graph (bottom) where each connected component represents the time evolution of a single cluster.

Then for all consecutive frames pairwise similarities, $d_{ij}(s) = sim(c_i(s), c_j(s + 1))$, are calculated between clusters using the similarity measure defined below. If two clusters in different frames are detected to be sufficiently similar (satisfy an appropriate threshold) they are deemed to be the same cluster at different points in time, and the nodes representing these clusters are connected through an edge. The outcome of this procedure is a graph that connects clusters in different time frames, such that every connected component of this graph is the trajectory of a particular cluster as illustrated in Figure 3.3.

Several similarity measures could be used to look for matching clusters, and here we employ the Jaccard index,¹¹⁰ defined as the size of the intersection divided by the size of the union of two sets,

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}.$$

Here A and B represent two clusters in consecutive time frames, and clearly $J(A, B)$ is one, if both clusters contain exactly the same ions, and zero if they do not share any ions. The threshold value of $J(A, B)$ must be selected with care to avoid ambiguities in cluster identity. For example, if two clusters, a and b , detected in frame one merge into a single cluster in frame two, it is not clear if the merged cluster is a , b , or an entirely new cluster c . Similar ambiguities can arise from other possible cluster evolution scenarios. In order to assign a unique well-defined time line to a cluster, it is necessary to remove all ambiguities. A natural way to do this is to connect clusters in consecutive frames only if they satisfy the threshold $J(A, B) > 0.5$. By definition of the Jaccard index, this threshold ensures that a cluster will match at most one cluster in the following frame, giving a set of unique cluster trajectories.

One potential issue with this method of obtaining cluster trajectories is how to take account of splits and merges that happen over multiple time frames. For example, if we have a large cluster (e.g., 500 ions) and a piece (e.g., 100 ions) becomes temporarily detached, the connectivity with the original cluster would be lost and the detached piece would appear as a newborn cluster of 100 ions. This problem is handled by detecting abnormally large newborn clusters and merging them with the original cluster. Fortunately, such events are rare (it happened once in our five simulations) and the post processing step was sufficient to solve the problem.

Another possible issue is that small clusters can partially redissolve (in the sense that some of their ions become less ordered), fall below the 6 ion threshold, and go undetected for a few time frames. To solve this problem we modified the algorithm to look for a match in 10 consecutive frames following the initial cluster detection. Newborn clusters that had no match in the following 10 frames were assumed to be short-term fluctuations, and discarded from the analysis.

3.3 Results and Discussion

Ionic clusters form and disintegrate continuously in the supersaturated solution. These clusters are detected and followed in time as described above. We are interested in understanding what features (if any) of ionic clusters influence their ability to survive and eventually achieve nucleation. We considered several cluster attributes or properties, but only two had a significant easily observable influence on the nucleation probability of a cluster. One of those, cluster size (calculated as the number of ions in the cluster), is of rather obvious importance, and is in fact the crucial parameter in the usual application of CNT, which assumes spherical nuclei that become critical at a certain radius. The other cluster property that proved very influential, especially so in smaller clusters, we call \bar{q}_8 , or the cluster *crystallinity*. This property is simply the average value of the q_8 bond order parameter defined above, taken over all ions in a cluster.

Other cluster properties considered include volume, surface area, sphericity, average neighbor count, hydration, and radius of gyration. These properties are precisely defined and discussed in Appendix A. For smaller clusters (e.g., 10 ions), none of these properties showed a significant correlation with the probability of nu-

cleation. For larger clusters (e.g., 30 ions), we do find noticeable correlations for three properties, with smaller surface areas, higher sphericities, and smaller radii of gyration all favoring nucleation. These attributes are all measures of cluster compactness, and it is not surprising that they are correlated with nucleation probability for larger clusters.

We note that all results reported below are based on the five simulations of 523 – 550 ns at $x_{\text{NaCl}} = 0.22$, as listed in Table 3.2. In our five simulations a total of 13 nucleation events were observed. In these simulations we noticed that clusters with a lifetime of at least 30 ns that reached a size greater than ~ 50 ions never redissolved, and continued to grow during the simulation. These clusters were deemed to have achieved nucleation. Note that we would expect this observational “definition” of nucleation to be concentration dependent.

At inception (defined as the earliest time a cluster is detected) the ionic clusters span a rather wide range of size and crystallinity. The size and crystallinity distributions of all clusters detected are shown in Figure 3.4, the top left and top right panels, respectively. Note that on average our algorithm detects ~ 2 new clusters per time frame, giving a total of ~ 80000 clusters detected over the course of the five simulations. It is also worth noting that the filtering step of the algorithm removes $\sim 90\%$ of the ions, indicating that only $\sim 10\%$ of the ions present in solution belong to a cluster of any sort.

From the histogram of initial sizes (Figure 3.4, top left panel) we see that the peak occurs at 6 ions (the smallest cluster detected by our algorithm), and that the frequency of larger sizes follows an exponential-like decay. Although very rare, clusters of more than 30 ions are found but, as noted below, these correspond to elongated, amorphous-looking structures. The distribution of crystallinities (Figure

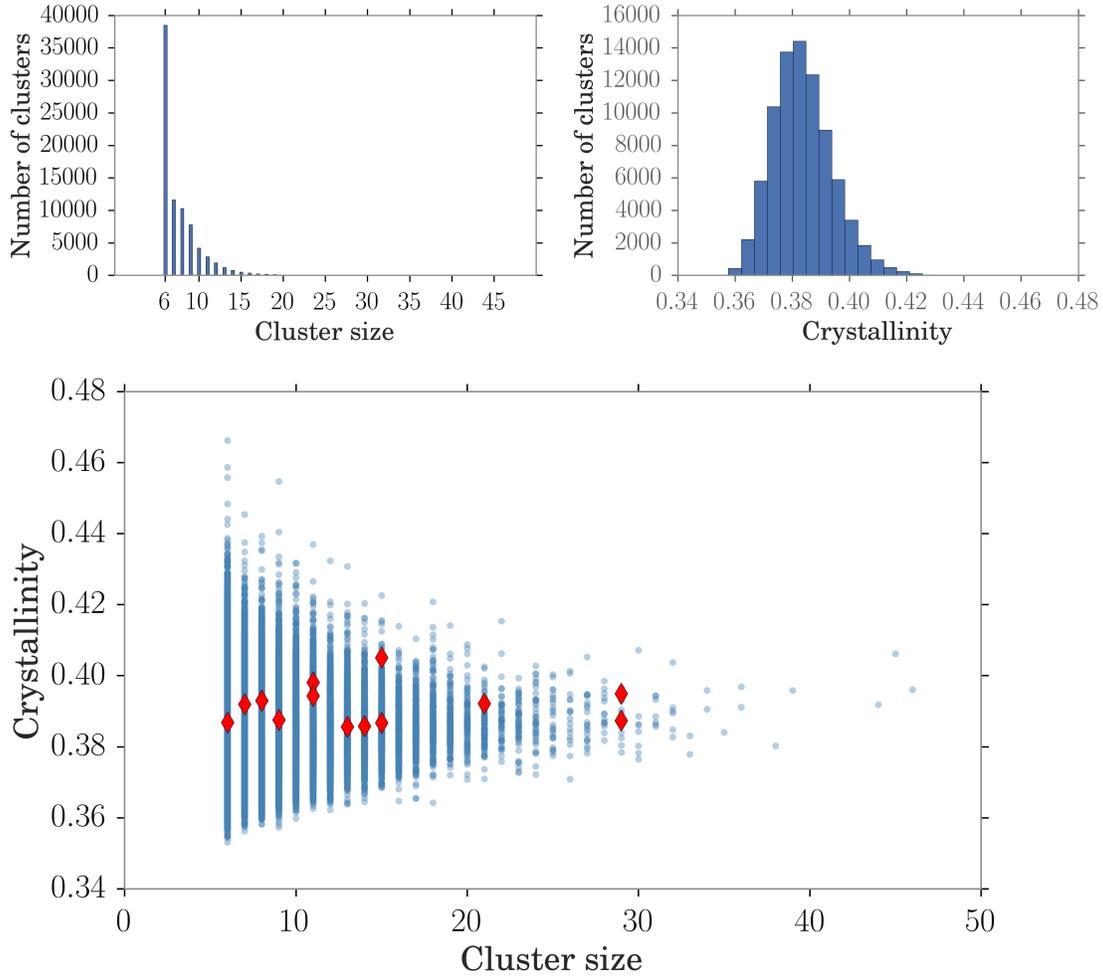


Figure 3.4: Size (top left) and crystallinity (top right) histograms for clusters at first detection. The bottom panel shows the joint distribution of size and crystallinity. Note that very few clusters begin with both high size and high crystallinity, suggesting that this feature develops as clusters evolve in time. The red points indicate the thirteen clusters that achieve nucleation, and we note that these show no obvious preference for any region of the joint distribution.

3.4, top right panel) is skewed to the right, but one does not observe the high values characteristic of a bulk crystal (Figure 3.1). A scatter plot of cluster crystallinity versus size is shown in the bottom panel of Figure 3.4. It is apparent that at cluster inception the correlation between these variables is modest at most. For smaller clusters, there are large fluctuations of crystallinity for a given cluster size, with large and small values occurring with high frequency. For larger cluster sizes, the crystallinity fluctuations decrease markedly, which is not surprising, since statistically variance is expected to decrease with increasing sample size, which here is the number of ions in the cluster.

Ionic clusters exhibit a range of crystallinities and the examples given in Figure 3.5 provide an idea of how cluster structure and crystallinity are related. Newly detected small clusters can have low (bottom left panel) or high (bottom right panel) crystallinity, with high values being associated with more regular structures. For larger clusters (more than 30 ions) we do not find any very high crystallinities at the point of initial detection, as such clusters tend to have elongated shapes (Figure 3.5, top left panel.)

3.3.1 Cluster Lifetimes

As discussed above, clusters are born with certain characteristics, and, after some time, they will either achieve nucleation (very rare) and continue to grow as a crystal or, much more commonly, dissolve back into solution. Given that one would expect a connection between cluster lifetime and nucleation, it is of interest to investigate which cluster characteristics influence their lifetime, defined here as the total time a cluster is detected (by the algorithm described above) as a distinct entity in solution. As, by definition, clusters that nucleate have infinite lifetimes,

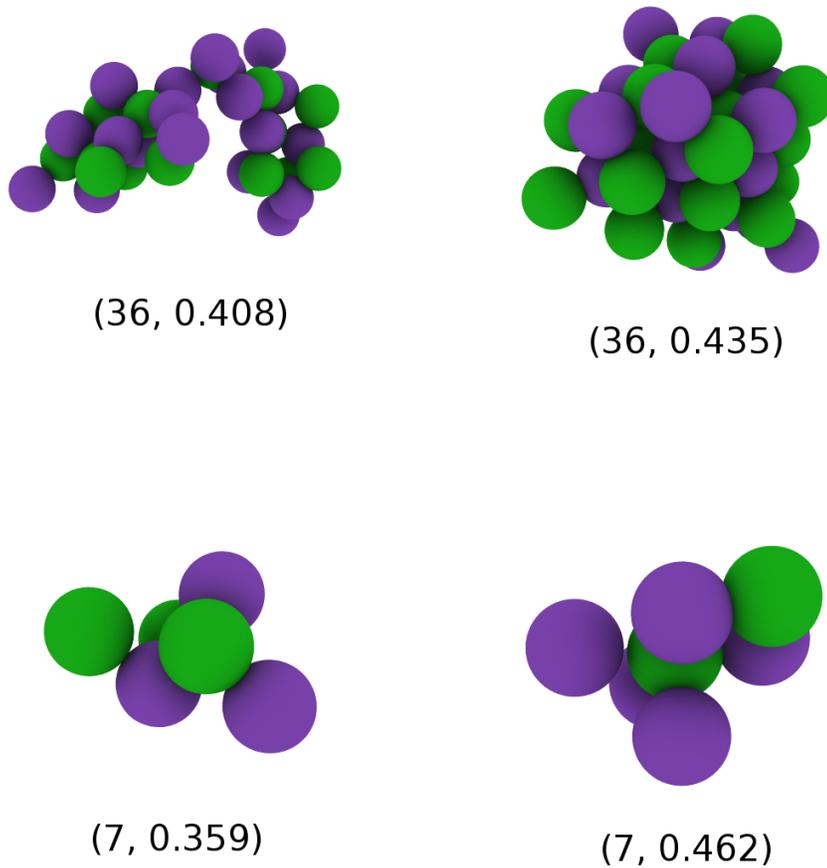


Figure 3.5: Examples of clusters of different sizes and crystallinities (given in parenthesis below each cluster). Na^+ and Cl^- ions are colored purple and green, respectively. At inception larger clusters tend to be elongated and of lower crystallinity (top left), whereas larger high crystallinity clusters (not observed at early stages) tend to be more compact (top right). For small clusters structural features associated with low (bottom left) and high crystallinity (bottom right) are not so apparent.

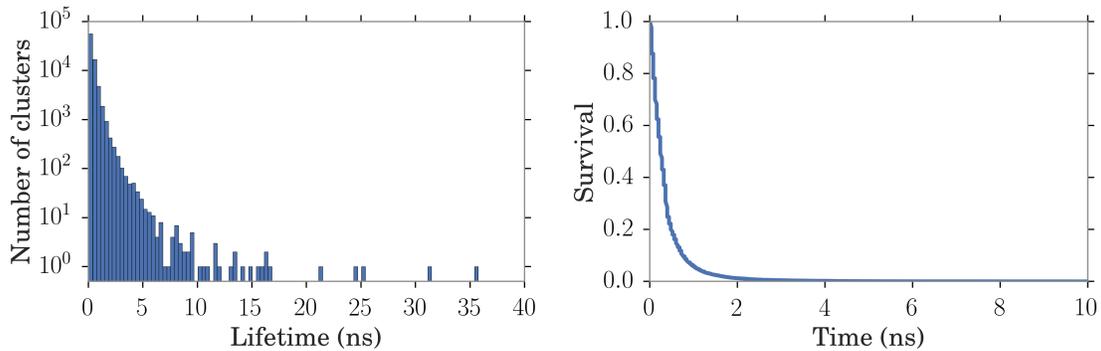


Figure 3.6: Histograms of “failed” cluster lifetimes (left) and the overall survival function (right) for all clusters detected in the simulations.

we excluded these from our statistical analysis. Note that in the five simulations (Table 3.2) included in the analysis only 13 clusters achieved nucleation, whereas ~ 80000 failed to nucleate. Focusing on the large number of failures provides us with excellent statistics, and clearly, the influence of the very few that did nucleate would have a negligible effect on the lifetime and survival distributions discussed below.

The overall lifetime distribution of failed clusters (uncontrolled for size or crystallinity) is shown in the left panel of Figure 3.6. This plot shows that the great majority of clusters are very short lived (lifetimes less than a nanosecond), but there are a significant number of clusters that live for more than 10 ns.

A more instructive way to represent lifetime information is to construct survival functions. The survival function, $S(t)$, is defined as the probability that a cluster has a lifetime, τ , larger than t , or

$$S(t) = Pr(\tau > t).$$

The survival function starts at 1 (all clusters have a non-zero lifetime) and

decreases with t , depending on the surviving fraction of the population. The survival function can be written as the cumulative product

$$S(t) = R(1)R(2) \dots R(t) ,$$

with

$$R(t) = \frac{\text{Population}(t)}{\text{Population}(t-1)} ,$$

where $\text{Population}(t-1)$ represents the population at frame $t-1$, and $\text{Population}(t)$ is the surviving population at the next frame t . $R(t)$ is the probability of surviving for one time interval at t . We note that this method of calculating $S(t)$ is commonly used, and is known as the Kaplan-Meier estimator.¹¹¹ The overall survival function including all clusters is shown in the right panel of Figure 3.6, and we note that over 90% of all clusters survive less than 2 ns.

Survival functions can be used to investigate if and how particular cluster characteristics influence their probability of survival. To isolate the influence of crystallinity, we divide clusters of fixed size into groups of high (> 0.40) and low (≤ 0.40) crystallinity, and compare the survival functions. Note that the survival functions are constructed using the size and crystallinity of a cluster at first detection. Obviously, neither of these cluster characteristics remains fixed as the cluster evolves in time. Results for clusters of 10 ions are shown in Figure 3.7, and it is obvious that crystallinity has a large effect on the survival probability. Clusters in the higher crystallinity group have a substantially higher survival probability than clusters in the lower crystallinity group. This is true for all cluster sizes, remarkably, even in the 6 ion case, which is the smallest cluster size we consider. This is illustrated in Figure 3.8, where we plot the ratio (high/low) of the median lifetimes for

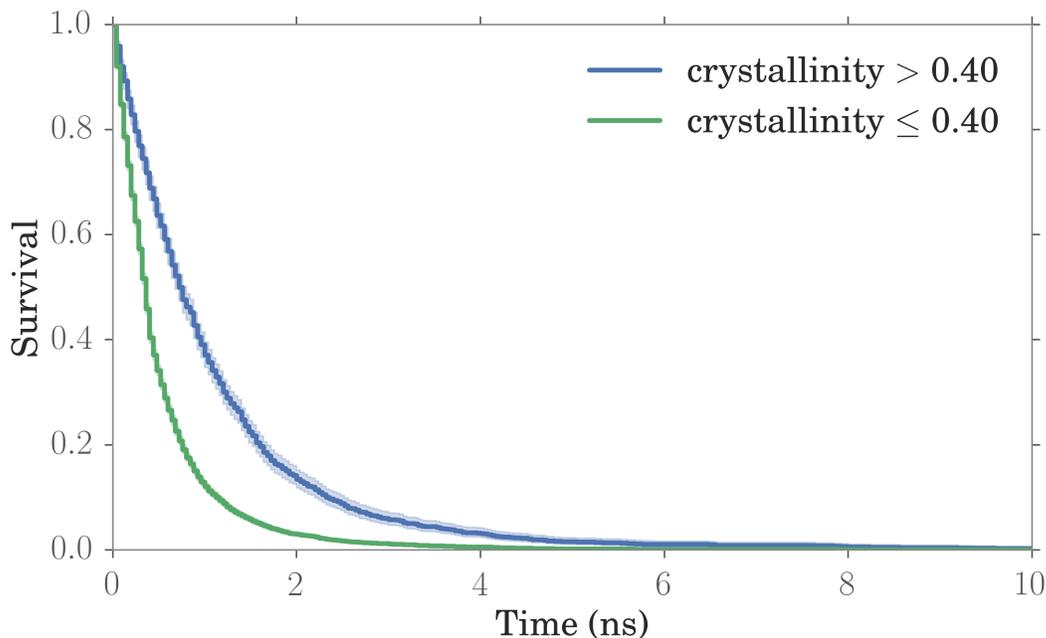


Figure 3.7: Survival curves for clusters of size 10 grouped by high (> 0.40) and low (≤ 0.40) crystallinity.

clusters of high and low crystallinity as a function of cluster size. We see that the ratio is significant (~ 1.5) for clusters of six ions, and rises rapidly with increasing cluster size, reaches a plateau at clusters of ~ 10 ions, then rises again for clusters containing more than ~ 18 ions. Results for clusters larger than 30 ions are not included because for the larger clusters the statistical analysis is less reliable due to small sample sizes. In an analogous manner, one can fix crystallinity and calculate how survival functions vary with size. This is less interesting, simply showing that, as expected, larger clusters have a larger survival probability.

While cluster survival is not a direct measure of nucleation probability, one would expect both phenomena to be related to cluster stability. Therefore, based on the above analysis, we would expect cluster crystallinity to be an important factor influencing the probability that a cluster will achieve nucleation. Specifically, for

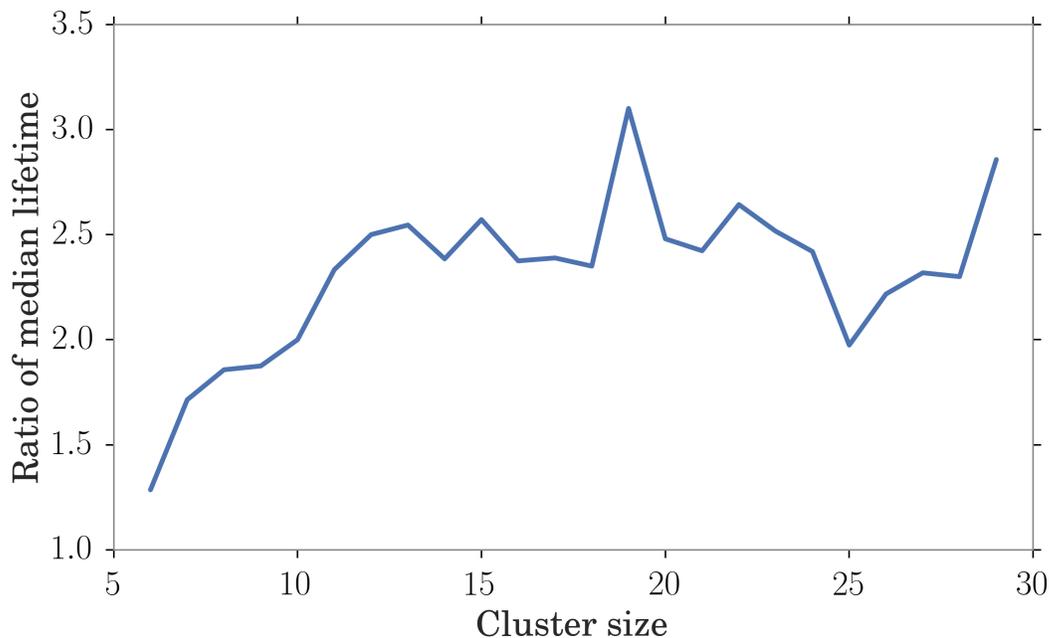


Figure 3.8: Ratio of median lifetimes for high versus low crystallinity clusters of different size. The high crystallinity clusters always have the higher median lifetime.

fixed cluster size higher crystallinities should favor higher nucleation probabilities. The relationship between cluster crystallinity and nucleation is treated directly in the following section, and we show that the behavior is indeed as we would anticipate based on the survival analysis. We remark that the survival analysis has the great advantage that the large number of clusters (events) included provides convincing statistics. On the other hand, we observe only a few nucleation events, so a statistical analysis based on nucleation alone would be less convincing.

3.3.2 Crystallinity and Nucleation

To get a qualitative idea of the influence of cluster size and crystallinity on nucleation, it is useful to plot cluster trajectories in (size, crystallinity) space. This is done in Figure 3.9 for a cluster that nucleates (blue) and one that does not (red).

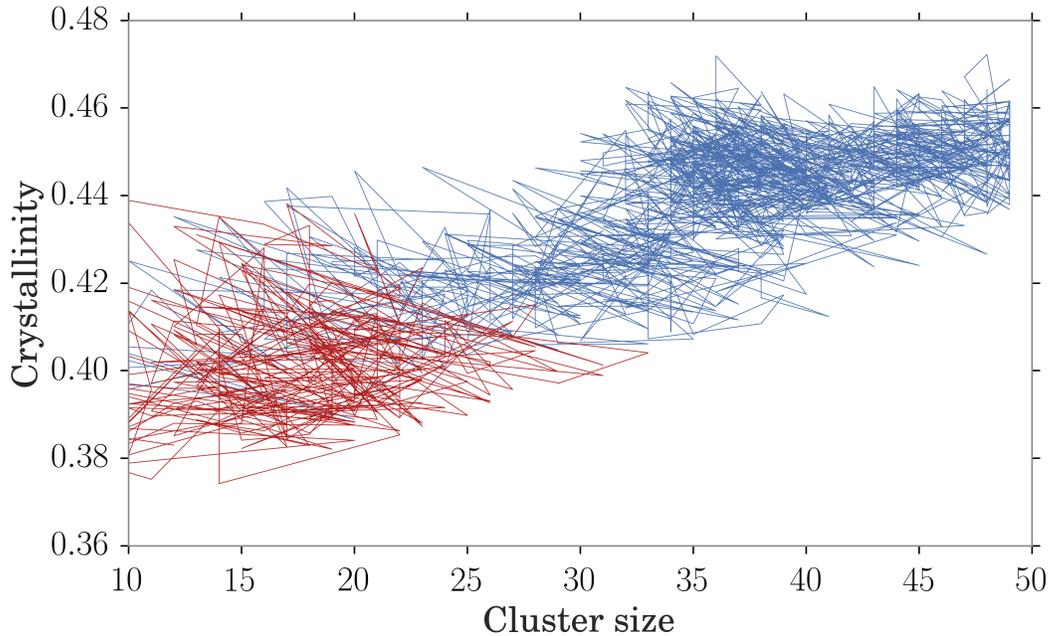


Figure 3.9: Trajectory shown in (size, crystallinity) space for a nucleated (blue) and a failed (red) cluster as they evolve in time. In early stages both clusters oscillate in size up to ~ 30 ions, but the crystallinity of the cluster that eventually achieves nucleation reaches higher values than the failed case. This suggests that it is a combined effect of both size and crystallinity that promotes nucleation.

It is easy to see that after a brief residency at low sizes and low crystallinities, the cluster that nucleates appears to surpass a “critical region” in the (size, crystallinity) space and never falls back. This suggests that both size and crystallinity influence the probability of nucleation.

This can also be seen in the early stage time evolution profiles shown in Fig. 3.10 for the nucleating and failing clusters. We see that while relatively long lived (~ 12 ns), and sometimes exceeding 30 ions in size, for most of the trajectory the crystallinity of the failing cluster lies below that of the nucleating case. Snapshots (not shown) suggest that the failing cluster grows in an elongated manner with a resulting lower crystallinity. The growth profile for a nucleating cluster is shown

in Fig. 3.11. It is interesting to note that the growth rate increases with size, attributable to the growing surface area of the nucleus. The crystallinity profile initially grows but plateaus as the cluster becomes larger in size, and the crystallinity approaches that of a perfect crystal.

The influence of cluster size and crystallinity on nucleation probability can be explored more quantitatively by comparing statistically these properties for clusters that nucleate with those that fail. If cluster size were the only factor influencing nucleation, then two clusters of the same size would have the same probability of nucleation. Therefore, if we fix the cluster size and observe a difference in crystallinity between nucleations and failures, we can isolate the effect of crystallinity.

One way to fix cluster size is to follow the trajectory of each cluster in time until it reaches a particular size, at which point the cluster crystallinity is recorded. Values obtained in this way can be reasonably assumed to be independent because the clusters exist at different points in time or in different simulations, and we register only a single crystallinity value for each cluster. Crystallinity distributions obtained as described are shown in Figure 3.12 for clusters of different size. The crystallinities of the clusters that achieved nucleation are also indicated in the plots, and we see that for all cluster sizes, the crystallinities of the nucleated clusters fall mainly in the upper half of the crystallinity distribution. In other words, the clusters that eventually nucleate have on average higher crystallinities than clusters that fail, and it is remarkable that this distinction exists even for clusters as small as six ions. For comparison, similar plots for other properties are reported in Appendix A.

The influence of cluster crystallinity on nucleation is further demonstrated in Figure 3.13, where the probability of nucleation, $P(N)$, for clusters of N ions is shown for clusters of “high” and “low” crystallinities. If the cluster crystallinity is

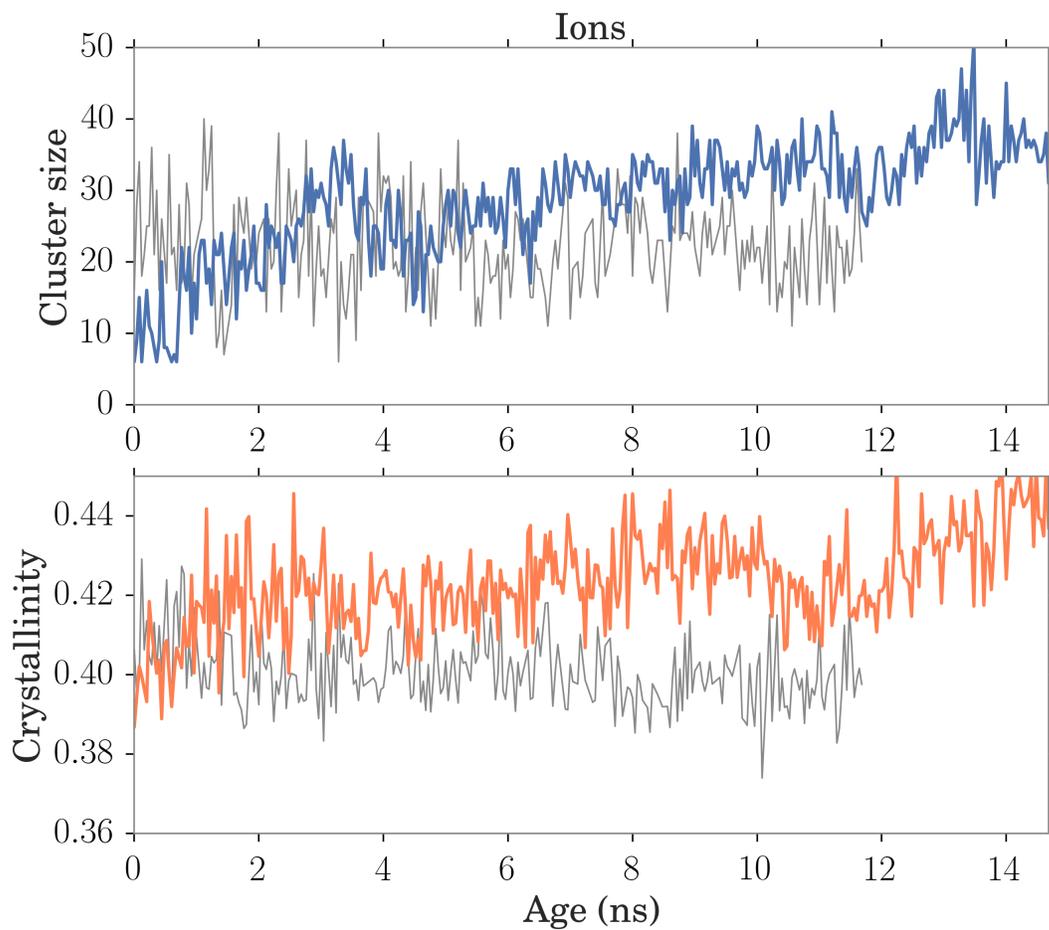


Figure 3.10: Comparison of growth and crystallinity profiles for nucleated (colored) and failed (gray) clusters. After a short period, the nucleated cluster manages to reach quite a high crystallinity and maintain its size. In contrast, the failed cluster, while maintaining its size, experiences a steady decrease in crystallinity.

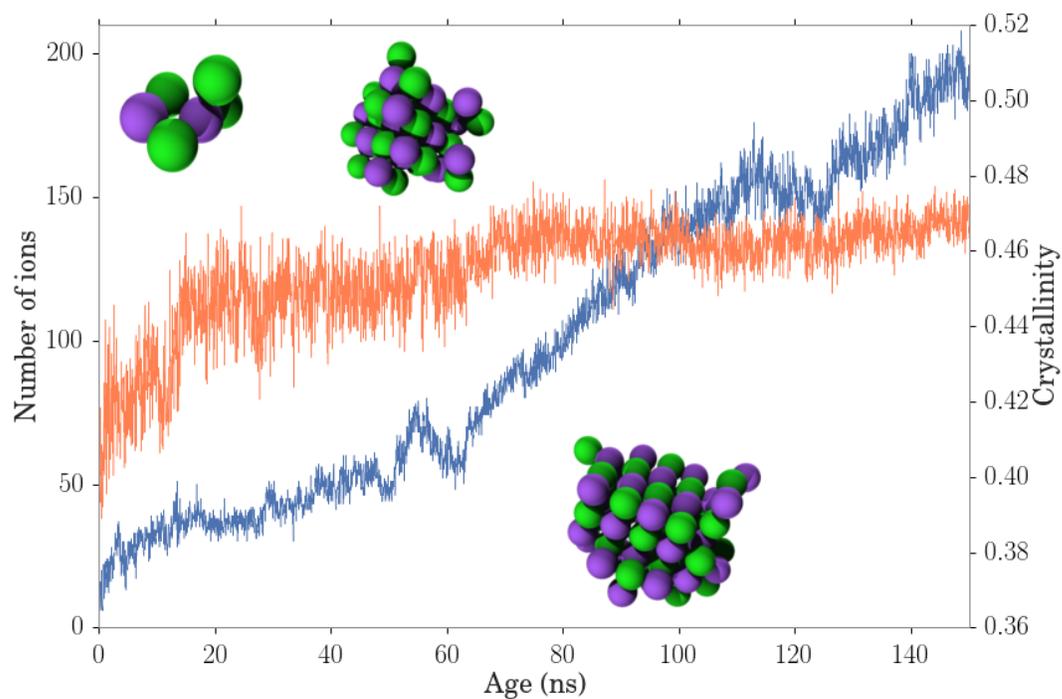


Figure 3.11: Size (blue) and crystallinity (orange) profiles for a nucleated cluster. The fluctuation in crystallinity is fairly high, and, as the crystal increases in size, the crystallinity reaches a plateau at ~ 0.47 . The snapshots represent the cluster at 0 ns (left), 40 ns (center) and 80 ns (right).

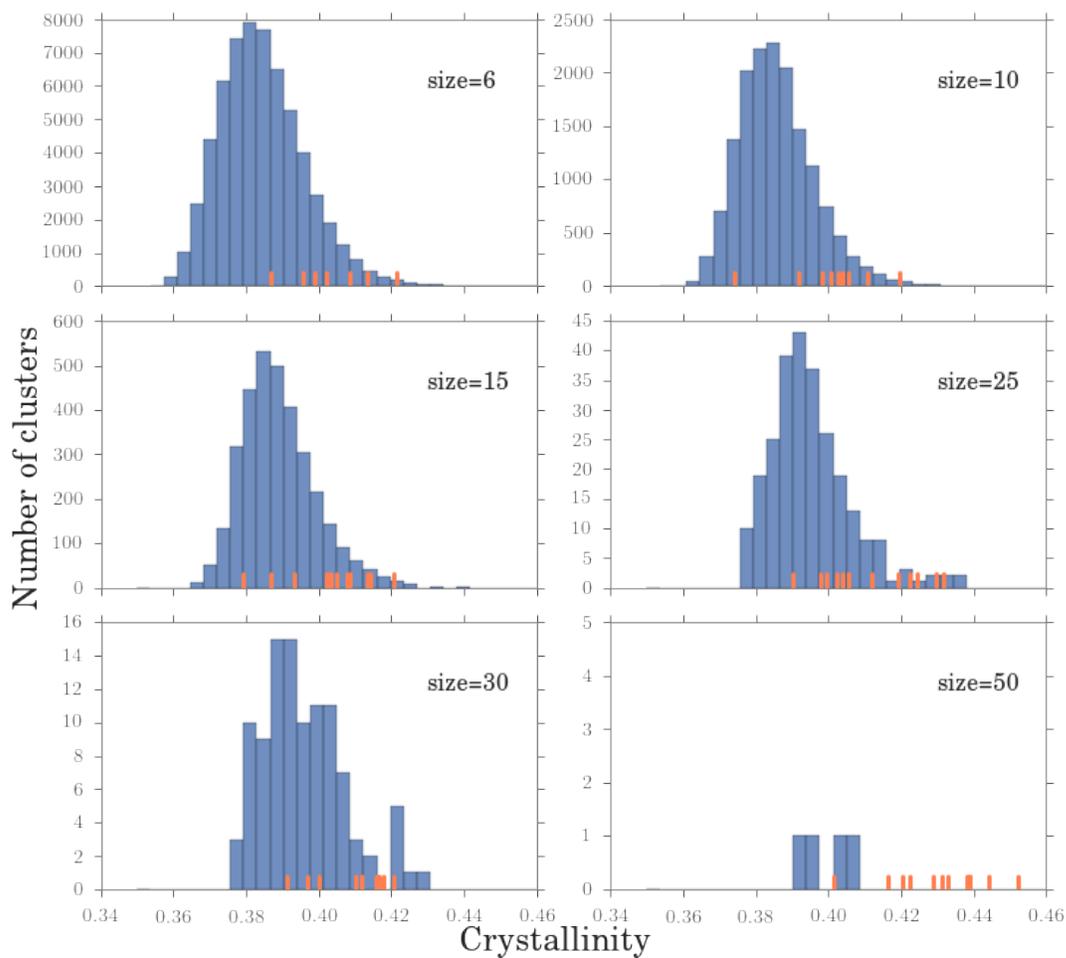


Figure 3.12: Crystallinity distributions (blue histogram) of failed clusters of different size (number of ions). Clusters that achieved nucleation are indicated by single orange lines. Note that clusters that achieve nucleation come preferentially from the upper part of the crystallinity distribution.

> 0.40 , it is labelled high, otherwise it is labelled low. The nucleation probabilities $P(N|\text{high})$ and $P(N|\text{low})$ are shown in Figure 3.13 for N ranging from 6 to 35 ions. Results for larger clusters are not shown because the sample size is too small to give good estimates of the probabilities. We note that $P(N|\text{high})$ is much higher than $P(N|\text{low})$ for all cluster sizes. The effect is most pronounced for small clusters (which have a very low total probability of nucleation) where $P(N|\text{high})$ is up to eight times larger than $P(N|\text{low})$. The ratio $P(N|\text{high})/P(N|\text{low})$ generally decreases with increasing N , but remains at ~ 3 for $N = 35$.

This analysis highlights the fact that nucleation is substantially influenced by factors other than cluster size. For a given cluster size, geometrical order, as captured by the crystallinity parameter, increases cluster lifetime and the probability of achieving nucleation.

Cluster Binding Energy

It is interesting to ask why high crystallinity leads to increased cluster lifetimes and influences the probability that a cluster will achieve nucleation. One possibility that immediately comes to mind is that in high crystallinity clusters the direct ion-ion interactions lead to lower cluster energies and hence increased stability. To investigate this possibility we calculated the energies

$$U_k = \frac{1}{2} \sum_i^{N_k} \sum_j^{N_k} u(r_{ij}) , \quad (3.2)$$

where N_k is the total number of ions in cluster k , and $u(r_{ij})$ is the ion-ion interaction as defined in Section 3.2. One potential problem with this analysis is that large fluctuations can occur in the apparent cluster energy depending on whether or not

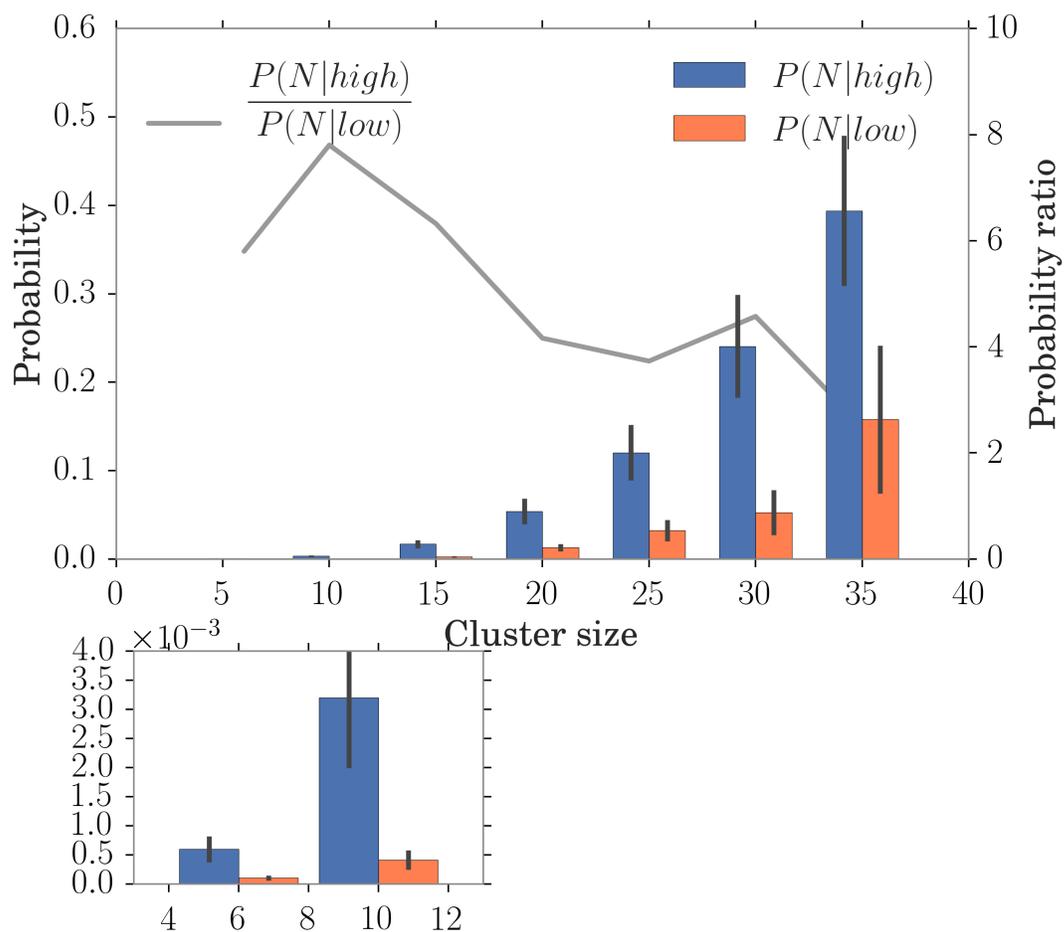


Figure 3.13: Probabilities of high ($\bar{q}_8 > 0.40$) and low ($\bar{q}_8 < 0.40$) crystallinities achieving nucleation. The error bars (vertical black lines) represent one standard deviation. The lower panel shows the results for small clusters on an expanded scale. Note that for all cluster sizes $P(N|high)$ is always larger than $P(N|low)$, and the high/low ratio is shown as a gray line in the upper panel.

a cluster has a net charge. These fluctuations can be created by a single ion moving in or out across the defined cluster boundary, and hence can be rather arbitrary and artificial. Therefore, to avoid this problem we include only electrically neutral clusters in this analysis.

Joint distributions of crystallinity and cluster energy for clusters of six and ten ions are shown in Figure 3.14. For clusters of six ions the correlation coefficient $r = -0.074$ indicating only a very weak correlation between energy and crystallinity. For clusters of ten ions, $r = -0.22$ showing that the correlation increases with cluster size, but still remains rather weak. Moreover, we see from Figure 3.14, that the clusters that eventually nucleate come mainly from the high crystallinity tail of the crystallinity distribution, but are nearly evenly spread over the energy distribution. These observations show that crystallinity is not a measure of increased cluster stability due to lower cluster energies coming through the direct ion-ion interactions. Clearly, other factors must be involved.

The Role of Na^+ and Cl^- Ions in Small Crystals

It is also of interest to more closely examine the structural nature of small NaCl clusters, since these clusters represent a fundamental step in the nucleation process. To do this we divide all ions in the system into two classes, high q_8 (≥ 0.4) and low q_8 (< 0.4), and calculate running counterion coordination numbers for Na^+ and Cl^- .

Results for single frames at 40 and 240 ns are shown in Figure 3.15. For the high q_8 class we see that both Na^+ and Cl^- have more counterions (3-4) in the first coordination shell (~ 0.35 nm) than the solution average (~ 1). This is not surprising because the high q_8 ions generally belong to ionic clusters. However, from

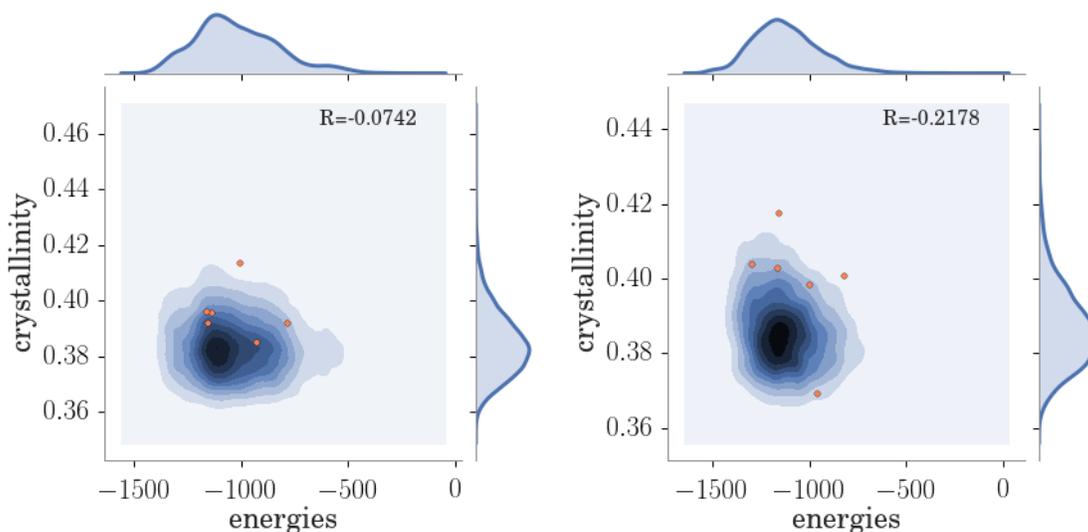


Figure 3.14: Joint distribution of crystallinity and cluster energy (kJ mol^{-1}) for clusters of size 6 (left) and size 10 (right). Values corresponding to clusters that eventually nucleated are displayed in orange. The correlation coefficient is negative and very low at size 6, and moderately low at size 10, indicating that crystallinity carries different information than energetic stability.

Figure 3.15 we also notice that high q_8 Na^+ ions have more counterion neighbors than high q_8 Cl^- ions, suggesting that Na^+ ions tend to lie more within the “interior” of small clusters than Cl^- ions, consistent with an earlier observation of Zahn.⁴⁷ A possible reason for this is that the smaller Na^+ ions interact more strongly with water molecules, and hence more first shell counterions are needed to compensate for water molecules lost from the first hydration shell when ionic clusters are formed. Note that the gap between the coordination numbers of high q_8 Na^+ and Cl^- decreases in later frames as larger clusters develop in the simulation.

To obtain additional evidence that Na^+ and Cl^- do not contribute symmetrically to the structure of small clusters, for each cluster we calculate the parameter

$$\Delta q_8 = \langle q_8^+ \rangle - \langle q_8^- \rangle,$$

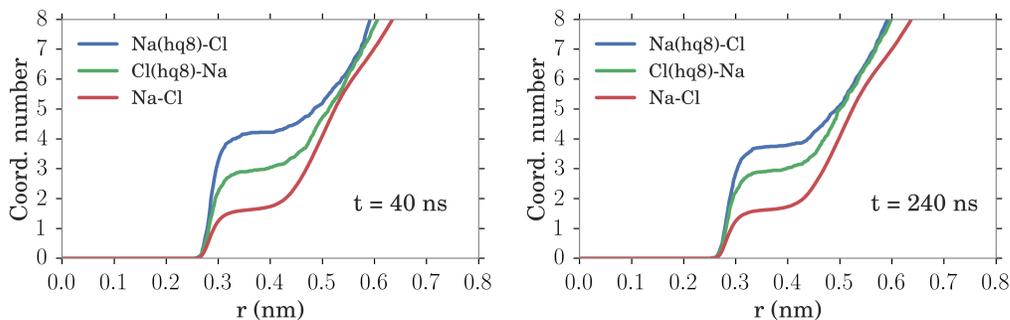


Figure 3.15: Running coordination numbers for: high q_8 Na^+ with all Cl^- (blue); high q_8 Cl^- with all Na^+ (green); and all Na^+ and Cl^- (red). In the first coordination shell (~ 0.35 nm) both high q_8 ions are surrounded by more counterions (3-4) than average (1). Also, high q_8 Na^+ are surrounded by more counterions than high q_8 Cl^- , suggesting that on average they occupy positions deeper within the clusters. The difference in the first shell coordination number decreases at longer times, as larger crystals develop in the simulation, and the internal preference of Na^+ becomes less noticeable.

where $\langle q_8^+ \rangle$ and $\langle q_8^- \rangle$ are the average values q_8 for the Na^+ and Cl^- ions in the cluster.

Figure 3.16 shows Δq_8 as a function of cluster size. These results are from the second replica simulation (Table 3.2). We note that for small clusters Δq_8 is positive on average indicating that Na^+ tends to lie in more ordered environments than Cl^- . As we would expect, the distinction between Na^+ and Cl^- grows smaller as clusters grow in size, and Δq_8 approaches zero for large clusters. Survival functions and nucleation probabilities were analysed separately for positive and negative values of Δq_8 , but no significant differences were found. Thus, while there is a structural distinction for small clusters, this effect does not have a significant influence on cluster survival or the probability of nucleation.

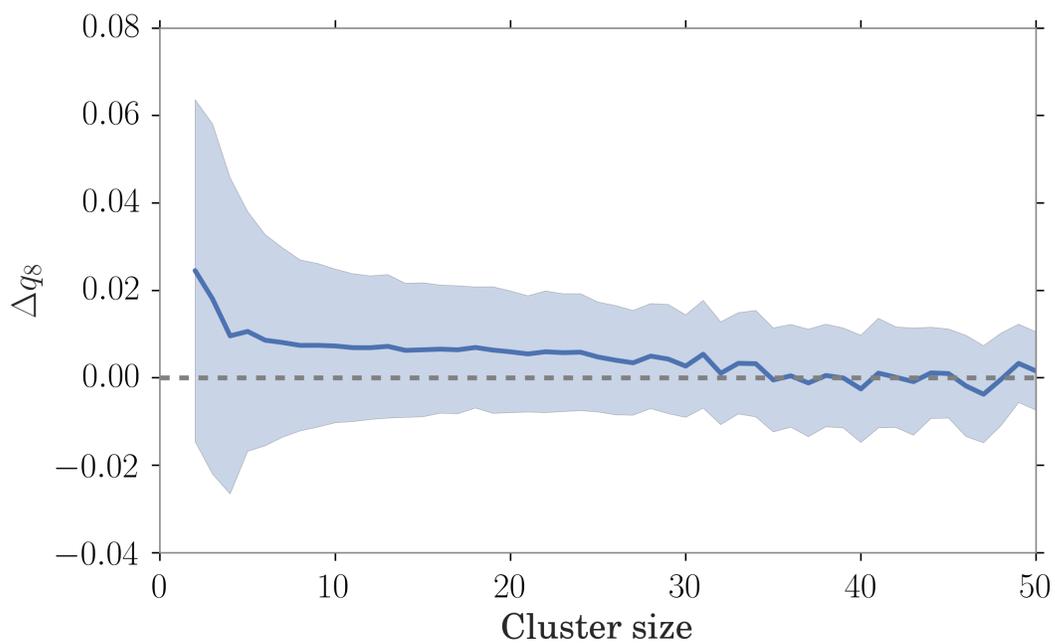


Figure 3.16: Δq_8 as a function of cluster size. The shaded area indicates one standard deviation. The average value is slightly positive indicating that Na^+ ions tend to occupy ordered environments than Cl^- ions.

3.4 Summary and Conclusions

Direct molecular dynamics simulations have been employed to identify and investigate factors that influence the nucleation of NaCl crystals in model supersaturated aqueous solutions. We develop methods that allow potential nuclei to be detected as small clusters (~ 6 ions), and followed in time until nucleation is achieved, which occurs very rarely, or the cluster dissolves back into solution.

Our analysis clearly demonstrates that cluster size is not the only property that has an important influence on the expected lifetime and nucleation probability of a particular cluster. We show that the geometric arrangement of the ions in the cluster, as measured by a single parameter which we call the cluster “crystallinity” is also very influential. For example, for clusters of ten ions the median lifetime for clusters of high crystallinity is double that of those of low crystallinity, and their probability of achieving nucleation is ~ 8 times greater. Similarly, smaller and larger clusters also have significantly longer lifetimes and greater probabilities of nucleation.

Physically, it is not entirely clear why crystallinity (as measured by our parameter) has such a large influence on cluster lifetime and nucleation probability, and this is especially true for small clusters of six or ten ions. One possibility is that there is a connection with the binding energy of a cluster. However, we did not find a strong correlation between binding energy and our crystallinity parameter, indicating that the crystallinity parameter is not merely a proxy for energetic stability. For small clusters, we did find that Na^+ ions had some preference for a “central” position, in accordance with Zahn’s observation,⁴⁷ but such cluster arrangements had no influence on the probability of nucleation.

Given that there is no obvious energetic explanation for the increased stability of small clusters of higher crystallinity, it appears that the advantage must lie in the microscopic dynamics of cluster growth and/or disintegration. One possibility is that ions in more ordered environments are less exposed to water molecules, and hence high crystallinity clusters are less susceptible to disintegration.

Chapter 4

Crystallization of Lithium Halides from Molecular Simulation.

4.1 Introduction

Alkali halides are inorganic compounds with high melting points that are usually found in nature as crystalline solids. Due to their importance in industrial and scientific applications such as corrosion,¹¹² desiccants, nanotube preparation,¹¹³ but also because of their simple structure, they have been widely studied for many years both in experiments and theory.^{114,115} More recently, sodium chloride has been used as a model compound in nucleation from solution,^{47,49} providing important insights in the nucleation process. However, not much has been done regarding other compounds, partly because their phase behavior in simulation has not been thoroughly examined.

In nature, the stable crystal structures for alkali halides are rock salt (FCC) and CsCl (BCC), however both experimental¹¹⁶⁻¹¹⁸ and theoretical studies show

the presence of wurtzite-type crystal structures,^{119–123} especially in lithium halides. Using ab-initio methods, the wurtzite-type structure was shown to be metastable at standard pressure for most of the alkali halides,¹²² and its existence was confirmed through low temperature deposition experiments (aimed at capturing metastable structures) for LiCl,¹¹⁶ LiBr^{117,118} and LiI.¹¹⁸ Four coordinated structures were also hypothesized by Pauling,¹²⁴ who derived the radius ratio rules using a simple hard sphere model.

Classical simulations of LiCl crystals also showed that, under the Tosi-Fumi¹²⁵ potential, the preferred structures achieved by melting and freezing small clusters were described as having hexagonal motifs.¹²¹ These structures were found to be energetically favorable (with respect to the rock salt structure) for sizes up to 64 ions.¹²¹ The competition between the hexagonal and cubic structure of LiCl was later found to be also present in larger systems.¹¹⁹

Despite the evidence of the existence of a metastable wurtzite structure in alkali halides, a systematic review of the most commonly used classical potentials with respect to this structure has not been performed. Such a study is important for the determination of solid and solution properties (e.g. solubilities) which could greatly affect the nucleation mechanism from the melt and/or from aqueous solutions.

Through molecular simulation we study the role of the wurtzite structure for lithium halides in both infinite crystals and finite size clusters. By comparing two established pair potentials used^{62–65,126–130} in classical simulations, the Tosi-Fumi¹²⁵ (TF) and Joung-Cheatham¹ (JC) parameter sets, we find that the wurtzite structure is the lowest potential energy minimum for some of the lithium compounds. This suggests that current models need to be adjusted to reflect the higher stability of the rock salt structure, in order to agree with experiments.

4.2 Models and Methods

As noted above, in this Chapter we consider two commonly used classical models for alkali halides. The TF potential^{125,131} is of the Born-Mayer-Huggins,^{132–134} consisting of a short-range repulsion combined with electrostatic interactions, and attractive dispersion terms of dipole-dipole and dipole-quadrupole order. The interaction $u_{ij}(r)$ between ions i and j can be expressed as

$$u_{ij}(r) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r} + B_{ij} e^{-\alpha_{ij} r} - \frac{C_{ij}}{r^6} - \frac{D_{ij}}{r^8}. \quad (4.1)$$

The parameters required are the charges q_i and q_j , B_{ij} , C_{ij} , and D_{ij} , which depend on the particular pair interaction (with different sets for different salts), and α_{ij} which is a constant depending only on the salt. The TF potential was obtained by fitting experimental data (e.g. equation of state and volume derivatives) for pure salts. Because of its focus on pure salts, this model has been mainly used for crystalline and molten salts,^{102,126–130} but it has also been adapted for use in aqueous solution.⁹⁸

The JC model combines electrostatic and Lennard-Jones (LJ) interactions and has the form

$$u_{ij}(r) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r} + 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r} \right)^{12} - \left(\frac{\sigma_{ij}}{r} \right)^6 \right], \quad (4.2)$$

where σ_{ij} and ϵ_{ij} are the usual LJ length and energy parameters. Unlike the TF potentials, the JC interactions were tuned to reproduce aqueous solution properties (e.g. solvation free energies), as well as solid state properties (e.g. lattice constants and energies). However, we note that all solid state calculations assumed the rock

salt structure, and no attempt was made to determine if a more energetically stable crystal structure existed. The parameter sets given by JC for a particular salt vary a little depending on the water model employed in calculations of the solution properties. Here we use the set given for the SPC/E water model, which is the model we employ throughout this thesis. We note that the rock salt lattice energies obtained for these parameters are in good agreement with experiments.¹ JC parameter sets have been used extensively to study solubility and solid-liquid transitions in solution.^{60–65}

Lattice energies for LiX ($X = \text{F}^-$, Cl^- , I^- , Br^- and I^-) salts are obtained for both rock salt and wurtzite crystal structures. The unit cells (see Figure 4.1) for the two crystal types were generated by calculating the unit cell parameters (a, b, c), angles (α, β, γ), and atomic fractional positions given the parameter a (the unit cell configuration for both the wurtzite and rock salt structures is completely determined by a , as the other parameters are constrained by symmetry.) The crystal structures, as well as the final crystals, were generated using the chemlab python library.¹³⁵ For the models considered, the lattice energy can be divided into short-range (van der Waals) and Coulombic contributions, denoted E^w and E^c , respectively, such that the total energy $E = E^w + E^c$.

The short-range interaction felt by ion i is given by

$$E_i^w = \sum_{\mathbf{n}=-\mathbf{n}_{\max}(r_c)}^{\mathbf{n}_{\max}(r_c)} \sum_j^N u_{ij}^{sr}(|\mathbf{s}_{ij} + \mathbf{n}|) \quad \text{with } i \neq j, \quad \text{when } \mathbf{n} = \{\mathbf{0}\}, \quad (4.3)$$

where u_{ij}^{sr} is the short-range part of the pair interaction, \mathbf{s}_{ij} is the difference between the position vectors of ions i and j in the unit cell, and \mathbf{n} is the displacement of the periodic cell. The first sum in Eq. (4.3) is over periodic repetitions of the unit

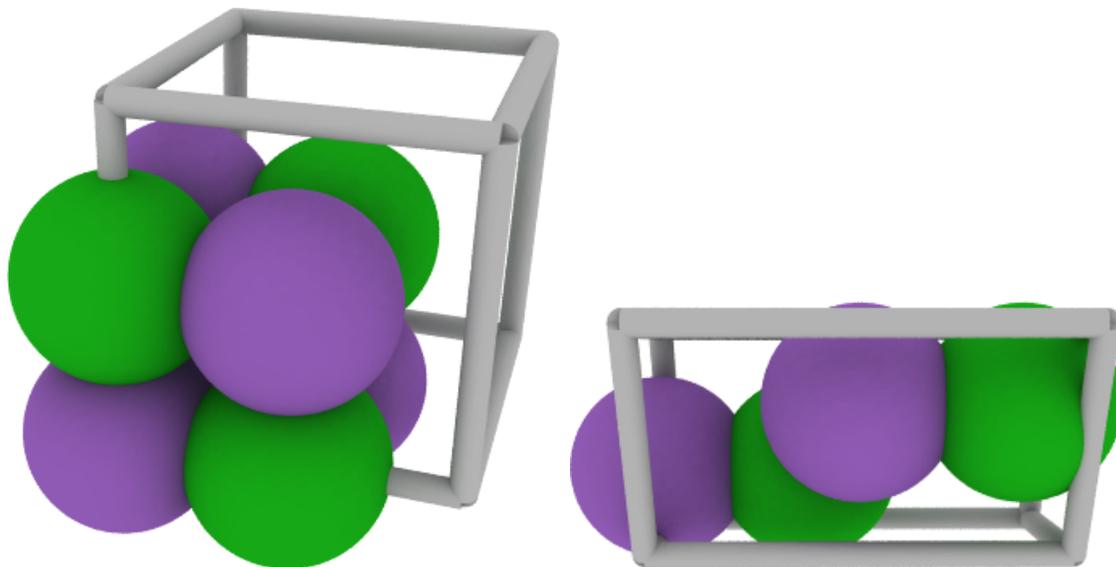


Figure 4.1: Unit cells used for the rock salt (right) and wurtzite (left) crystal structures. The configurations are dependent on a single cell parameter a , all other parameters being constrained. In the rock salt structure, a is the length of any edge, while for the wurtzite structure a is the length of any of the two short edges.

cell, and \mathbf{n}_{\max} is the minimum cell displacement needed to include all ions within a cutoff radius, r_c , of the central ion. In the present calculations, r_c is taken to be 2.4 nm.

The electrostatic interactions were calculated using the Ewald summation method⁷³ with conducting boundary conditions. Note that conducting and vacuum boundary conditions will give the same lattice energy for rock salt because, for that crystal structure, the unit cell does not have a dipole moment, but this is not true for wurtzite where a dipole moment is present. The Ewald parameter κ (see Ref. 73) was set at $\sim 5/a$ (a is the cell parameter), and 1000 wave vectors were used in the Fourier space sum. Short-range interactions in the real space part of the Ewald sum were truncated at a . The lattice energies obtained were in excellent agreement (within ~ 0.001 kJ mol⁻¹) with values reported in the literature.¹³⁶

Since in both rock salt and wurtzite crystals there are only two types of lattice

site (positive and negative ion sites), the total lattice energy can be written as the sum of two terms $E = E_+ + E_-$, where E_+ and E_- are contributions from the ions on different sites interacting with the remainder of the lattice. This decomposition of E gives some useful insight into the relative stabilities of rock salt and wurtzite structures for various salts (see Section 4.3).

Molecular dynamics simulations were carried out for both finite clusters and infinitely periodic systems. Starting configurations were generated by repeating the rock salt unit cell along the 3 crystal axes, 3, 4 and 5 times for finite clusters (corresponding to 256, 512, and 1000 ions, respectively), and 5 times for infinitely periodic systems.

Simulations were conducted using the GROMACS molecular dynamics package version 4.6.¹⁰⁰ Finite clusters were simulated using a cubic simulation cell of length L that was ~ 4 -5 times larger than the cluster, such that interactions with its periodic images were negligibly small. All interactions were spherically truncated at $L/2$, and the cell volume was kept constant during the simulation. The temperature was regulated using the velocity-rescale thermostat.⁷⁹ These simulation conditions effectively correspond to a finite size crystal at zero external pressure.

Periodically infinite systems (both liquid and solid) were simulated by applying periodic boundary conditions in the usual manner.⁷³ All short-range interactions were spherically truncated at 0.9 nm, and the long-range electrostatic interactions were calculated using the particle mesh Ewald (PME) method.⁷⁴ The pressure was fixed at 1 bar using the Berendsen barostat,⁷⁸ and the temperature was regulated using the velocity rescale thermostat.⁷⁹ The Berendsen barostat was chosen because it produced less drastic fluctuations during convergence, making the simulation less likely to become unstable (especially with the TF potential). Some tests were per-

formed with the Parrinello-Rahman barostat¹³⁷ to verify that the choice of barostat did not influence the results.

We are mainly interested in determining crystal structures at finite temperature by spontaneous nucleation and growth. Therefore, the simulations were performed in the following manner. The simulations were started at 300 K in a rock salt configuration and equilibrated for $\sim 0.1 - 0.5$ ns. The temperature was then raised to $\sim 1200 - 2000$ K (depending on the melting point of the crystal) over a time period of ~ 1 ns, causing the salt to melt. The molten salt was then frozen by cooling the system to $\sim 500 - 600$ K over a time interval of ~ 5 ns. The solid was then maintained at the freezing temperature for $\sim 5 - 20$ ns to allow the system to anneal (the difference in annealing times is because small crystals require less times to anneal). Finally, the temperature was lowered to 300 K to collect properties for comparison with the initial rock salt crystal equilibrated at the same temperature.

After this procedure, the final structure was determined by visual examination, and by comparing radial distribution functions of the initial and final frames of the trajectory, as illustrated in Figure 4.2. A summary of the simulations performed, along with the resulting crystal structures, is reported in Table 4.4 and 4.3.

4.3 Results and Discussion

4.3.1 Lattice Energies

To illustrate how the TF and JC potentials differ, we compare the pair interactions for LiCl in Figure 4.3. One notices (see the top and bottom panels of Figure 4.3) that the Cl^- ion is slightly “larger” for the JC model. Also, the repulsive part

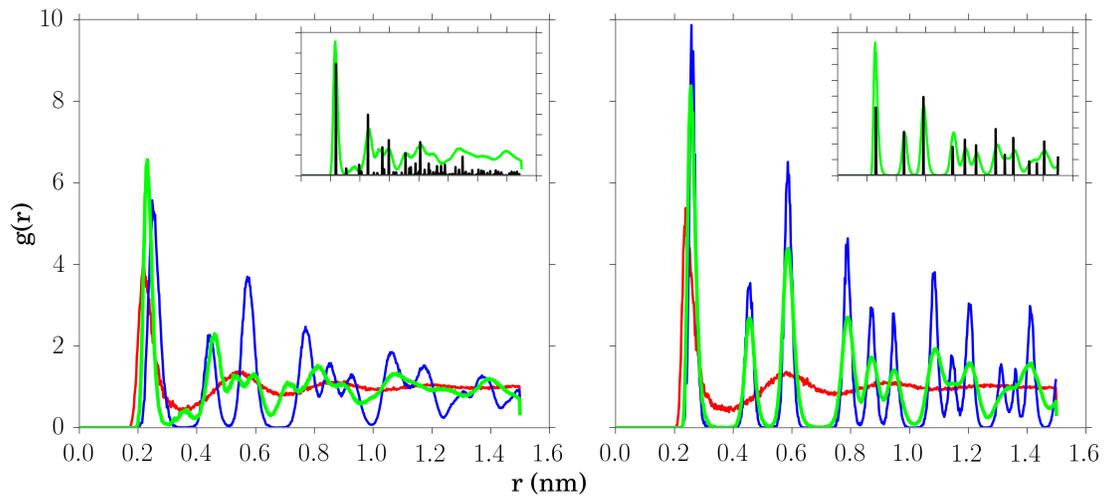


Figure 4.2: Radial distribution functions for an infinite LiCl crystal for the TF model (left panel) and JC model (right panel). The initial rock salt configuration and the liquid phase are colored in blue and red, respectively, and do not show noticeable differences between models. The rdf corresponding to the crystallized structures are displayed in green and correspond to wurtzite on the left and rock salt on the right. The inset plots show the rdf for ideal wurtzite on the left and ideal rock salt on the right (black lines) superimposed onto the simulation result (green).

of the potential, which dominates at short range, grows less sharply for the TF than for the JC model (this is due to the different functional form, $\sim e^{-r}$ vs $\sim r^{-12}$, employed in the potentials). For the attractive cation-anion interaction (middle panel of Figure 4.3), there is a difference in the well depth and in the position of the minimum, with the TF interaction more attractive than JC. At longer range the potentials are equally dominated by the electrostatic interaction (r^{-1}).

Of course simply comparing pair potentials in this way does not provide much information on interactions in particular crystal structures, therefore we now turn to the lattice energies. For lithium halides occupying rock salt or wurtzite lattices, the electrostatic (Coulombic) interaction is the same for both positive and negative ion sites such that $2E_+^c = 2E_-^c = E^c$, where E^c is the total electrostatic contribution to the lattice energy. For the crystal structures we consider

$$E^c = \frac{1}{4\pi\epsilon_0} \frac{M_{\pm}}{r_0}, \quad (4.4)$$

where r_0 is the distance of an ion from its nearest neighbor, and M_{\pm} is the Madelung constant, which depends only on the crystal structure. Since $M_{\pm}^{\text{rocksalt}} > M_{\pm}^{\text{wurtzite}}$, the electrostatic contribution to the lattice energy, for fixed r_0 , is always lower for the rock salt structure (a consequence of the higher coordination number). The wurtzite structure can only become electrostatically more stable than rock salt by allowing the ions to pack more tightly, thus increasing the electrostatic interaction by reducing r_0 .

Equilibrium ion-ion distances and the minimum lattice energies are significantly influenced by the short-range repulsive part of the pair potential. In the case of oppositely charged ions, the short-range repulsion counteracts the Coulombic

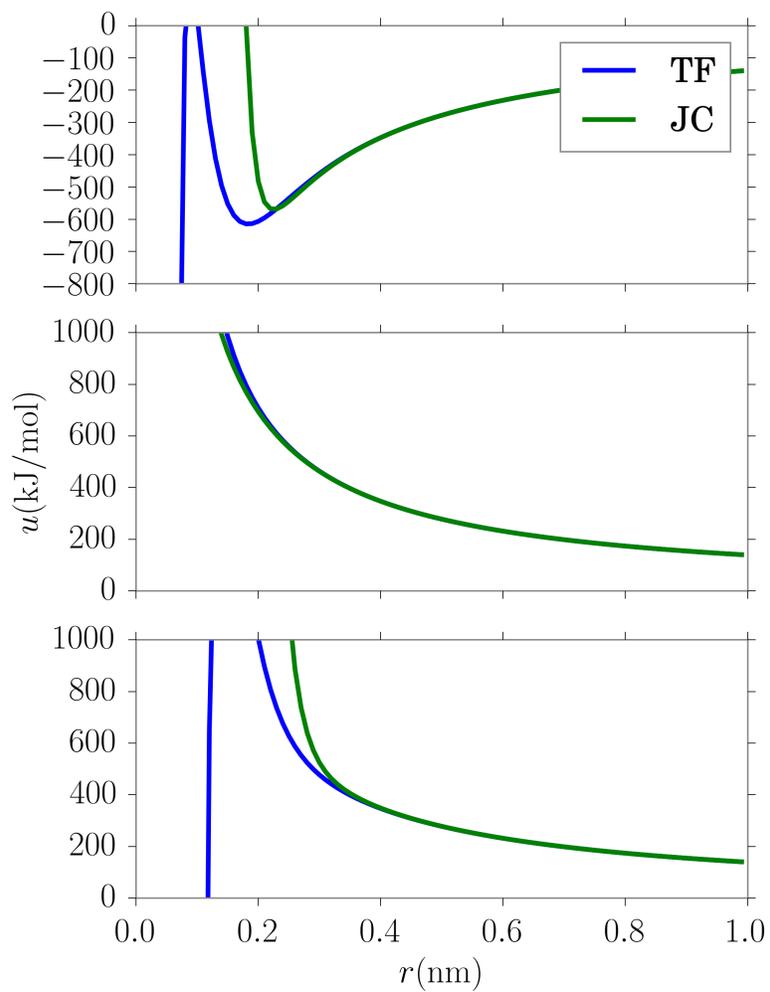


Figure 4.3: TF and JC potentials for the pairs Li⁺-Cl⁻ (top panel), Li⁺-Li⁺ (middle panel) and Cl⁻-Cl⁻ (bottom panel). The TF model has a shorter contact radius for the Cl⁻ ion, and is more slowly varying function at short range. Notice also how at very short range (~ 0.1 nm), the TF model becomes attractive (but this does not influence the simulations).

attractive forces creating a well defined “contact” distance (Figure 4.3, top panel). The interaction between ions of like charge is always repulsive, with a sharp increase at short range that strongly penalizes close contact (Figure 4.3, bottom panel). The balance between the two competing interactions determines the nearneighbor equilibrium distance and the coordination number.

In general the lattice-positive-ion interaction, E_+ , will be different from the lattice-negative-ion interaction, E_- , because the short-range, like-charge interactions are different. To obtain the lattice energy it is necessary to find the structure that minimizes the sum of those two contributions.

Lattice energies for perfect lithium halide crystals at 0 K are given in Tables 4.1 and 4.2 for the JC and TF models, respectively. We see from that for the JC model (Table 4.1) the rock salt structure has the lower lattice energy for LiF and LiCl, but that the wurtzite energy is lower for LiBr and LiI. For the TF model, (Table 4.2) the wurtzite lattice energy is lower than that of rock salt for all four lithium halide salts. Thus, only the JC results for LiF and LiCl are in accord with the rock salt structure experimentally observed for all lithium halides.¹³⁸ Lattice energies for NaCl are included in Tables 4.1 and 4.2 for comparison purposes, and we note that in this case the rock salt structure has the lower lattice energy for both the JC and TF models.

One can gain some insight into this behavior from Figure 4.4, where the lattice energies for lithium halides are plotted as functions of the cell parameter a . The lattice-positive-ion and lattice-negative-ion contributions, respectively, are also shown. Note that in the figure the actual plots represent $2E_+$ and $2E_-$; this preserves the same vertical scale such that features of all three curves (e.g. the positions of the minima) can be easily compared. For the JC model (left column, Figure 4.4) we see

Compound	$E_{\text{rock}}(\text{kJ mol}^{-1})$	$E_{\text{wurtz}}(\text{kJ mol}^{-1})$	$\Delta E (\text{kJ mol}^{-1})$
LiF	-1052.56	-1030.42	-22.14
LiCl	-870.31	-863.49	-6.82
LiBr	-826.04	-826.67	0.63
LiI	-761.83	-770.08	8.25
NaCl	-793.15	-769.67	-23.48

Table 4.1: Lattice energies for the JC model for the wurtzite (E_{wurtz}) and rock salt (E_{rock}) structures. The difference $E_{\text{rock}} - E_{\text{wurtz}}$ is reported as ΔE . The NaCl lattice energies are added for comparison.

Compound	$E_{\text{rock}} (\text{kJ mol}^{-1})$	$E_{\text{wurtz}} (\text{kJ mol}^{-1})$	$\Delta E (\text{kJ mol}^{-1})$
LiF	-1043.04	-1051.83	8.78
LiCl	-844.65	-846.66	2.02
LiBr	-795.60	-796.48	0.87
LiI	-720.01	-729.20	9.18
NaCl	-777.73	-762.51	-15.22

Table 4.2: Lattice energies (TF model) for the wurtzite (E_{wurtz}) and rock salt (E_{rock}) structures. The difference $E_{\text{rock}} - E_{\text{wurtz}}$ is reported as ΔE . The NaCl lattice energies are added for comparison.

that as the anion becomes larger the repulsive interactions become more important, and the minima become less negative for both rock salt and wurtzite structures. The effect is larger for rock salt, hence the crossover to the wurtzite structure at LiBr. The apparent reason for this is the fact that for rock salt the minima for E_+ (blue lines) and E_- (green lines) are found at very different a values, or, in other words, the most favorable cell length for the lattice-positive-ion interaction is different from that of the lattice-negative-ion interaction. A similar effect occurs, but is much less pronounced for the wurtzite structure, hence wurtzite becomes more stable than rock salt as the ion size asymmetry becomes larger. For the TF model (right panel, Figure 4.4), the wurtzite is the more stable structure for all four salts.

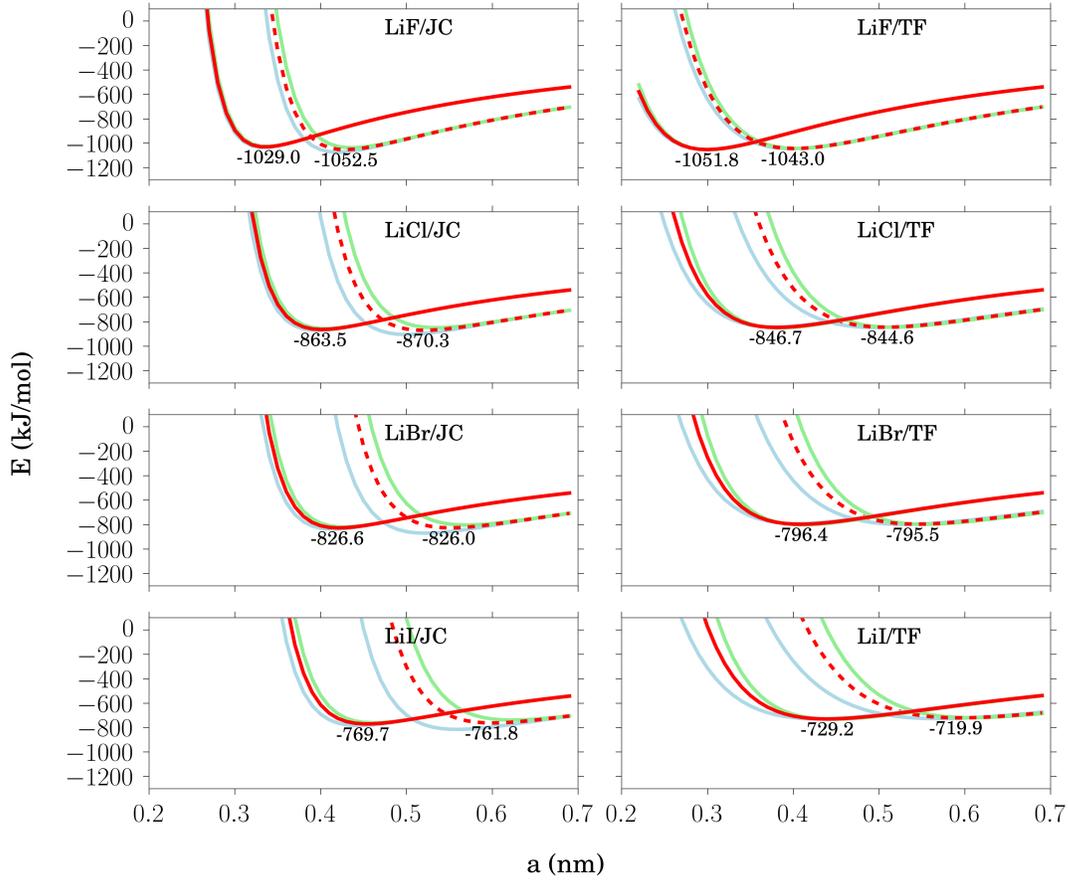


Figure 4.4: Lattice energies as a function of the cell length a for the JC (left column) and TF (right column) potentials. The solid red lines indicate the energy of the wurtzite structure, while the dashed red lines indicate the rock salt structure. The energy minima are annotated below each curve. Blue and green lines represent the contributions of the positive and negative ion sites, respectively.

Compound	JC	TF
LiF	rock salt	wurtzite
LiCl	rock salt	wurtzite
LiBr	wurtzite	wurtzite
LiI	wurtzite	wurtzite

Table 4.3: Final structures obtained by melting and freezing periodic crystals. The final temperature is 300 K, and the periodic box contained 1000 ions. The results are consistent with the lattice energy calculations (Tables 4.1 and 4.2).

4.3.2 Infinitely Periodic System Simulations

The lattice energy calculations given above determine the stable crystal structures at $T = 0$ K. It remains interesting to ask if these are in fact the structures that crystallize spontaneously from the melt at finite temperature. To answer this question we carried out MD simulations, first heating a rock salt crystal to a temperature above the melting point, then cooling the molten salt until spontaneous crystallization occurs, following the protocol described in Section 4.2. The results obtained for both JC and TF models are summarized in Table 4.3, and we see that in all cases the salt crystallized into the structure predicted by the lattice energy calculations.

The lattice structures obtained upon spontaneous crystallization can be identified by inspecting configurational snapshots, or, more quantitatively, by comparing radial distribution functions with expected results for rock salt and wurtzite crystal structures. As an example, various rdfs for LiCl/JC and LiCl/TF are shown in Figure 4.2. Comparing the rdfs, it is apparent that LiCl/JC crystallizes as rock salt and LiCl/TF as wurtzite, consistent with the lattice energy predictions.

In addition to identifying crystal structures, one can make some other observations based on the rdfs in Figure 4.2. For the initial rock salt structure both

models give qualitatively similar rdfs (blue curves), as expected, but the peaks tend to be lower and broader for the TF model. We attribute this to the softer short-range repulsion of the TF potential, which allows the ions to have larger oscillations around their equilibrium positions. We see also that the molten salts give very similar rdfs for both models.

4.3.3 Finite Size Cluster Simulations

We have also carried out melting/freezing simulations on finite size clusters of 216, 512, and 1000 ions. This is of interest in the context of crystal nucleation, for example, from aqueous solution. It is important to know if one can expect small clusters to nucleate in the stable structure of the bulk phase, or in some other size-dependent structure. The final structures obtained for finite clusters are summarized in Table 4.4.

We see from Table 4.4 that for salts where the rock salt and wurtzite lattice energies differ by $\sim 2 \text{ kJ mol}^{-1}$ or more, the finite cluster structures are consistent with the infinite system results discussed above, and with predictions based on the lattice energies. One important difference is that for cases where rock salt is the stable structure, such as LiCl/JC and LiF/JC, even small clusters achieve well-defined rock salt structures, whereas this is not true of salts that prefer wurtzite. Rather, in the wurtzite case, one obtains structures that have a hexagonal motif characteristic of wurtzite, but are by no means perfect wurtzite structures, and in the small cluster case can be mostly hollow inside. We refer to these structures as wurtzite-like.

An example structure obtained for LiCl/TF with 216 ions is shown in Fig-

Compound	Ions	JC	TF
LiF	216	rock salt	wurtzite-like
	512	rock salt	wurtzite-like
	1000	rock salt	wurtzite-like
LiCl	216	rock salt	wurtzite-like
	512	rock salt	wurtzite-like
	1000	rock salt	wurtzite-like
LiBr	216	rock salt	wurtzite-like
	512	mixed	wurtzite-like
	1000	wurtzite-like	wurtzite-like
LiI	216	wurtzite-like	wurtzite-like
	512	wurtzite-like	wurtzite-like
	1000	wurtzite-like	wurtzite-like

Table 4.4: Final structures obtained by melting and freezing finite size clusters using the JC and TF models, the final temperature is 300 K. A wurtzite-like structure corresponds to a structure that is not perfectly crystalline but shows hexagonal motifs similar to that shown in the left panel of Figure 4.5.

ure 4.5 (left panel). We note that these hexagonal structures are similar to those reported in previous cluster simulations of LiCl employing the TF potential.¹²¹ In fact, based on considering finite and infinite hexagonal structures, Rodrigues and Silva Fernandes¹¹⁹ suggested that the TF potential does not produce the experimental rock salt structure for LiCl, and our lattice energy and simulations confirm this conclusion.

It is also interesting to compare the rdf of a perfect wurtzite and a finite size wurtzite-like crystal. In Figure 4.6, the Li-I rdf for the JC model is shown for both a perfect, infinite, wurtzite crystal, and a finite size crystal of 1000 ions, both at 300 K. To avoid the influence of surface ions, the rdf is computed by using only the Li⁺ ion closest to the cluster’s geometric centre. One notices that the rdf peaks for the finite size structure are broad and, with the exception of the contact peak, by no means a perfect match for the perfect wurtzite structure.

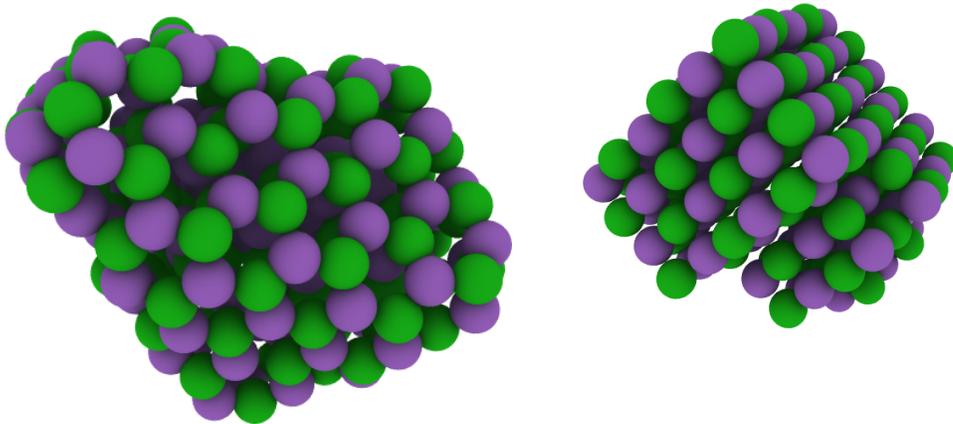


Figure 4.5: Spontaneously crystallized LiCl/TF (left panel) and LiCl/JC (right panel) clusters of 216 ions. For the cluster on the left, the structure has hexagonal motifs characteristic of the wurtzite structure, but is mostly hollow inside. On the right, the cluster crystallizes into a rock salt structure, without substantial surface deformations. In this figure, Li^+ is represented in purple and Cl^- is in green.

The reason why small wurtzite crystals are not stable is because the wurtzite unit cell has a non-zero dipole moment directed along the c axis. For a unit cell with $a = 0.45$ the dipole moment is ~ 26 D, therefore a perfect wurtzite crystal will have a surface dipole with charges of the same sign occurring at the surface. This means that the surface configuration of a finite wurtzite crystal is energetically unfavorable, and will undergo surface reconstruction. For small clusters, one observes rather hollow structures (most of the cluster is surface) with hexagonal patterns on the surface. This effect is confirmed by noticing that the dipole moments per ion of the finite 216 ion crystals obtained in our simulations are $\lesssim 0.1$ D, which is much smaller than the value ~ 6.5 D expected for a perfect wurtzite crystal.

Surface reconstruction of a larger (2048 ion) wurtzite crystal of LiI/JC is illustrated in Figure 4.7. Simulated in vacuum at 300 K, the reconstruction of the (0001) face is apparent in the before and after snapshots shown in Figure 4.7. We

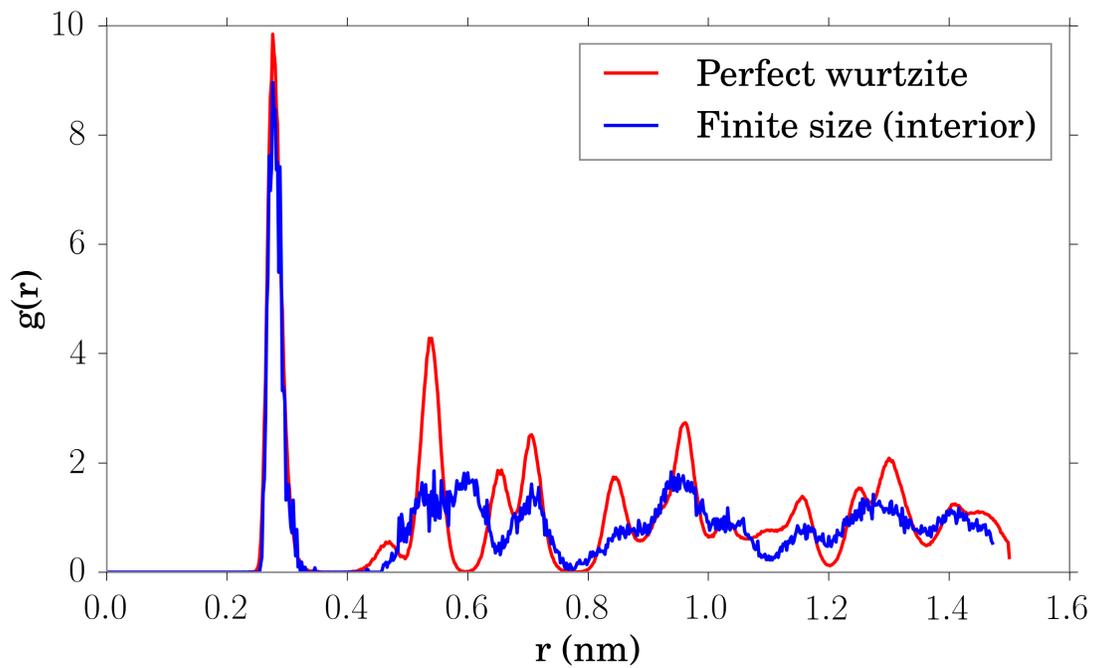


Figure 4.6: LiI/JC rdfs at 300 K for the perfect wurtzite crystal (red), and for the interior of a finite size crystal of 1000 ions (blue). The peaks are well defined for the perfect wurtzite crystal, while for the finite size crystal there is wide broadening and merging of adjacent peaks.

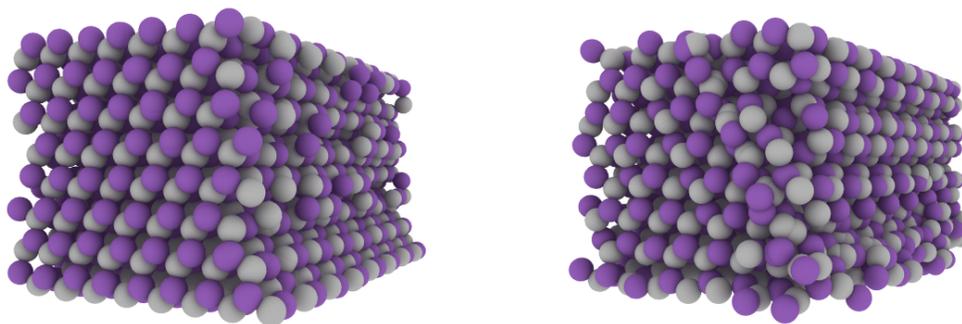


Figure 4.7: Example of surface reconstruction for a LiI crystal (Li^+ and I^- ions are colored in purple and grey, respectively). On the left, surface charges are present on the wurtzite crystal face (0001). After simulating the crystal in vacuum at 300 K, the resulting surface (on the right) rearranges to reduce the dipole moment. The dipole moment per ion is reduced from ~ 6.3 D for the structure on the left, to ~ 0.7 D for the structure on the right.

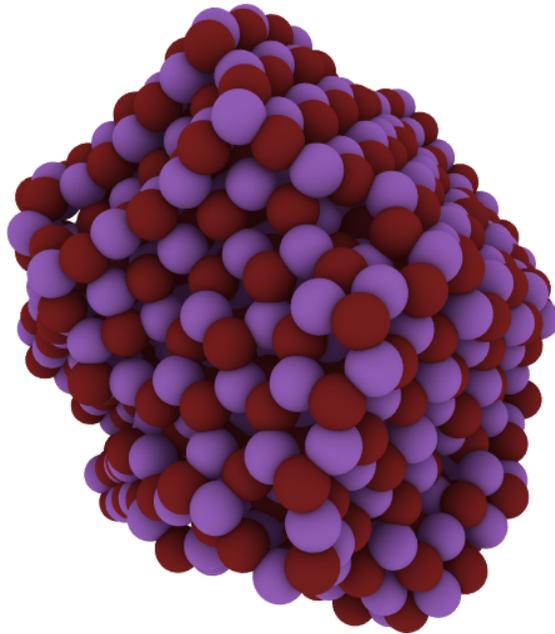


Figure 4.8: Spontaneously crystallized LiBr/JC cluster (1000 ions) using the LiBr/JC model (Li^+ in purple, and Br^- in brown). One notices both hexagonal motifs, characteristic of the wurtzite structure, as well as square motifs characteristic of the rock salt structure.

note that the reconstruction reduces the dipole moment of the crystal from ~ 6.3 D per ion initially to ~ 0.7 D after simulated reconstruction at 300 K.

An interesting situation occurs when the rock salt and wurtzite lattice energies are very close together, such as LiBr/JC. The rock salt structure crystallizes spontaneously at small sizes but, as size increases, it crosses over to wurtzite, giving rise to mixed crystal structures, as shown in Figure 4.8. This confirms that the rock salt structure has a lower surface tension (due to the lack of surface charges) than the wurtzite structure and crystallizes more easily. Also, since the standard procedure we used involved ~ 20 ns annealing time at high temperature, (see Section 4.2 for details) we can exclude that this is due to dynamical effects.

Temp (K)	E^r	E^w	$E^{w\dagger}$	ΔE (kJ mol ⁻¹)
0	-894.53	-889.66	n.a.	-4.86
300	-886.34	-881.70	-878.71	-4.65

Table 4.5: LiY lattice energies at 0 K and 300 K. The value E^w was obtained by simulating a box to initialized from the perfect wurtzite structure, while $E^{w\dagger}$ is the energy of the crystal obtained from freezing of the molten salt.

4.3.4 Temperature Effects

By constructing a fictitious system, we studied how temperature favors one or the other crystal structure. In our lithium halide set, we obtained cases where the wurtzite and rock salt lattice energies are spaced well apart and, for LiBr/JC, where the wurtzite structure is marginally more stable than the rock salt structure, but no system was found to have a marginally stable rock salt structure (the compound closest to this condition is LiCl/JC, where rock salt is ~ 6.82 kJ mol⁻¹ more stable than wurtzite).

A fictitious LiY/JC system, based on the LiBr/JC parameter set with a smaller σ_{Br} (in our case, reduced by 0.04 nm) was constructed to have a rock salt crystal structure that is slightly more stable than the wurtzite structure. According to lattice energy calculations, the minimum energies for the rock salt and wurtzite configurations for the LiY/JC were found to be, respectively, -894.53 and -889.66 kJ mol⁻¹, with the rock salt structure being more stable by ~ 4.87 kJ mol⁻¹.

However, the simulations of crystallization from the melt at 300 K show that, despite its higher energy (Table 4.5), the preferred structure is wurtzite, suggesting that the wurtzite structure is favored by entropy.

4.4 Summary and Conclusions

In this Chapter, we investigated the relative stability of the wurtzite and rock salt crystal structures for a series of lithium halides using the popular Joung-Cheatham¹ and Tosi-Fumi¹²⁵ potentials.

While wurtzite crystals are metastable in real systems, the TF model was found to crystallize preferentially into the wurtzite structure. In contrast, the JC model correctly predicts the rock salt crystal structure for LiCl and LiF but not for LiBr and LiI. The origin of this effect can be attributed to the short-range part of the pair potential and the excessive destabilization of the rock salt structure as ion size asymmetry increases.

Our simulations of finite size clusters in vacuum suggest that the wurtzite structure is hard to form at small sizes because of the surface rearrangements needed to avoid surface charge formation. We also found that if the lattice energies of the rock salt and wurtzite structures are close to each other, like in the case of the LiBr/JC model, the system may crystallize as rock salt at small sizes, and as wurtzite at larger sizes.

The behavior of molten salts is interesting in its own right, and an understanding of finite size cluster structures can be related to homogeneous nucleation from aqueous solution. Further work is required to show if the finite size crystal structure could prevent the stable bulk structure from forming or if there are structure rearrangements during the nucleation process.

Chapter 5

Crystallization of Lithium

Fluoride in Aqueous solution from Molecular Simulation.

5.1 Overview

Molecular simulation is employed to study nucleation and growth of lithium fluoride crystals in aqueous solution. It was possible to observe nucleation on simulation timescales only at a temperature of 500 K, and a fixed density of ~ 1.24 g/cm³ to prevent water evaporation. We find that the growth rate is temperature dependent, and follows Arrhenius' law, with an activation energy of ~ 50 kJ mol⁻¹. Since the activation energy for the diffusion of Li⁺ and F⁻ ions in solution was found to be much smaller (~ 20 kJ mol⁻¹) the growth barrier is likely related to the high energy required to remove water from the first solvation shell of the ions. The solubility for the model salt was estimated to be ~ 0.03 mole fraction at 300 K, which

is much higher than the experimental value (~ 0.001 mole fraction). Additionally, a decrease of solubility with temperature was observed, likely due to a decrease in the dielectric constant of the solvent.

5.2 Introduction

The study of alkali halide phase transitions with molecular simulation has been an intense subject of study because of the simplicity of the compounds, and the availability of models designed to reproduce both properties of their crystals and aqueous solutions.¹ Particular attention has been devoted to NaCl models,^{1,125,139} for which exist nucleation studies,^{47,49,66,140} as well as estimates of their solubility.⁶⁰ Lithium halides crystallization from the melt has also been examined,^{121,127} but the study of their nucleation from solution is very limited.¹⁴¹

In this chapter we investigate LiF nucleation from aqueous solution using molecular simulation. LiF, with an experimental solubility¹⁴² of ~ 0.135 g/100 mL (~ 0.001 mole fraction) at 298 K and 1 bar, is the least soluble of the lithium halides, and molecular dynamics simulations of its nucleation process provides insights into how the nucleation mechanism may differ from that of a more soluble salt such as NaCl. Also, LiF does not introduce the complication of stable hydrate structures that are common for compounds such as LiCl.¹⁴³

We find that, within simulation timescales, nucleation from solution can only be obtained by heating the solution to ~ 500 K, and that, for LiF, crystal growth is an activated process with an activation energy of ~ 50 kJ mol⁻¹. In the remainder of this Chapter we explore the origin of this temperature dependence by examining properties of the LiF solution such as diffusion, solubility, as well as the distribution

	σ (nm)	ϵ (kJ mol ⁻¹)
Li ⁺	0.1409	1.4088967
F ⁻	0.4022	0.0309637

Table 5.1: Joung-Cheatham Lennard-Jones parameters for LiF¹ adapted for the SPC/E water model.²

and lifetime of LiF clusters in solution.

5.3 Models and Methods

The model used to simulate LiF aqueous solutions are the Joung-Cheatham¹ parameters for the ions, combined with the SPC/E² parameters for water. In this model the pairwise interactions are represented by a Lennard-Jones term plus an electrostatic contribution (cf. Equation (3.1)). The Joung-Cheatham parameters for Li⁺ and F⁻ are given in Table 5.1.

All simulations employed the GROMACS version 4.5.5 software package.¹⁰⁰ To keep the temperature fixed in our simulation we used the velocity rescaling thermostat,⁷⁹ and to keep pressure fixed we used the Parrinello-Rahman barostat.¹³⁷ The time step used in our simulations was 2 fs. The short-range interactions were calculate up to a cutoff of 0.9 nm while the electrostatic interactions were calculated using the particle mesh Ewald method.⁷⁴

A series of LiF solutions at different concentrations were initialized by randomly placing water molecules, together with Li⁺ and F⁻ ions, in a cubic simulation cell of length ~ 4 nm. The system was relaxed at a fixed temperature of 300 K and a pressure of 1 bar until its density, the Li⁺-F⁻ radial distribution function, and the ion cluster distribution (see below for explanation) were stabilized, which required ~ 20 ns.

Nucleation

T (K)	Length (ns)	ttn (ns)	LiF	H ₂ O	m.f.	Pressure (bar)
300	60.60	n/a	419	7977	0.05	1
300	134.12	n/a	878	7903	0.10	1
300	147.20	n/a	1677	7640	0.18	1
300	200.00	n/a	2095	7430	0.22	1
300	158.56	n/a	1369	7761	0.15	1
350	199.74	n/a				1500
400	102.12	n/a				3100
450	136.82	n/a				4800
500	132.84	15				6500
500	80.00	5				
500	80.00	3				

Growth

T (K)	Length (ns)	LiF	H ₂ O	m.f.	Pressure (bar)
300	80.00	1369	7761	0.15	1
350	80.00				1500
400	80.00				3100
450	249.30				4800-6000
500	24.64				6500-7200

Table 5.2: Summary of the simulations performed to study LiF nucleation and growth. The temperature at which a simulation was conducted is indicated by T, the length of the simulation is indicated by Length, the concentration in mole fraction is abbreviated as m.f., and the time to nucleation (ttn) is the time at which the first nucleus was observed. LiF and H₂O indicate the number of molecules of each type used in the simulation. The pressure values for the crystal growth simulations are reported as a range for some simulations because, when substantial crystal growth occurs, the pressure drifts to higher values.

Starting from the ~ 0.15 mole fraction LiF equilibrated solution (which corresponds to ~ 5 to 15 times the saturation concentration for our model, depending on the temperature), we simulated the system over a range of temperatures (between 300 and 500 K) and observed crystal nucleation at a temperature of 500 K. In all simulations the volume was kept constant (with density ~ 1.24 g/cm³) to maintain the solution in the liquid state. Three replica simulations at 500 K were also performed to have an estimate of the time necessary to achieve the first nucleus, the replicas showed a substantial variability resulting in a time to first nucleation ranging from 3 to 15 ns. Lower concentrations were also tested but we were not able to obtain nucleation at any temperatures (between 300 and 500 K), concentrations higher than ~ 0.15 mole fraction immediately yielded multiple nucleations at a temperature of 500 K. A summary of the simulations performed is given in the left panel of Table 5.2. It is useful to notice that the experimental solubility for LiF at 300 K is ~ 0.001 mole fraction,¹⁴² which is an order of magnitude lower than our estimate for the model employed (see next section).

To study crystal growth of a LiF crystal in solution, a configuration containing a partially grown crystal made of ~ 300 ions was extracted from one of the simulations at 500 K and 0.15 mole fraction. Beginning with this configuration, *NVT* simulations were performed at the temperatures 300, 350, 400, 450 and 500 K (cf. Table 5.2).

The amount of crystal phase present in the growth simulations was estimated by calculating the Steinhardt-Nelson bond order parameter¹⁰⁷ q_8 for each ion and by counting the number of ions with $q_8 > 0.40$ (cf. Chapter 3 for further details).

The order parameter q_l is defined as

$$q_l = \left[\frac{4\pi}{2l+1} \sum_{m=-l}^l |q_{lm}|^2 \right]^{1/2},$$

where

$$q_{lm} = \frac{1}{N} \sum_{r_i=r_1}^{r_N} Y_l^m(\theta(\mathbf{r}_i), \phi(\mathbf{r}_i)),$$

where $\{r_1 \dots r_N\}$ are the positions of the first N neighboring ions relative to the central ion, Y_l^m are spherical harmonics, and θ and ϕ are the polar and azimuthal angles corresponding to the vector \mathbf{r} . The method and parameters used are the same as those described in Chapter 3 (which correspond to parameters $N = 12$ and $l = 8$).

Another clustering method, equivalent to that employed by Hassan,⁵⁸ was used to study dissolution and the properties of the solution, including information about pairs and other ion aggregates that do not arrange with an octahedral geometry. We will refer to this method as connectivity based clustering.

In this method, given a system configuration, all the pairwise distances between positive and negative ions are first calculated. Every pair of oppositely charged ions that are closer than a certain threshold are deemed to be in contact and placed in the same cluster. The natural threshold for this procedure is a number slightly larger than the contact radius (which can be found from the first peak of the Li-F radial distribution function), and it was chosen to be 0.3 nm for LiF. The final result of the above procedure is a set of clusters where every ion can be reached from every other ion through a path of consecutive connections. Notice also that free ions are considered to be in a cluster of their own. See Figure 5.1 for a diagram.

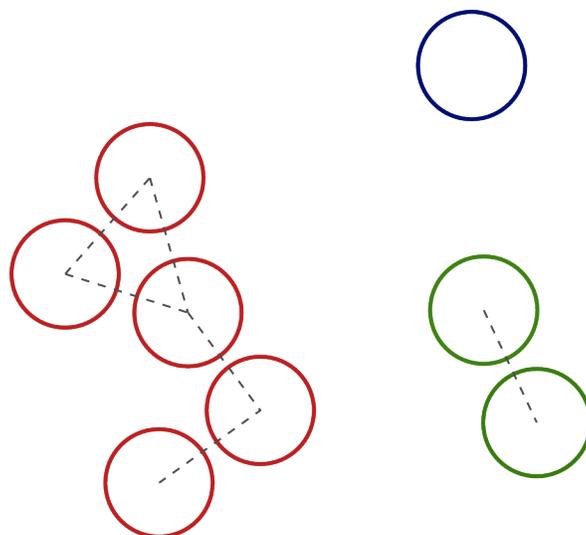


Figure 5.1: Clustering based on neighbor distance. The dashed lines represent distances lower than a fixed threshold, and the colors represent different clusters. Free particles form a cluster of their own.

5.4 Results and Discussion

A series of LiF solutions at different concentrations and temperatures were simulated at constant volume and it was found that spontaneous, homogeneous nucleation within ~ 3 to 15 ns can be obtained at a high temperature (500 K) and at a concentration ~ 0.15 mole fraction and higher. Simulations of crystal growth rate also show an increase with temperature and, as we will show below, the growth rate dependency on temperature follows an Arrhenius law.

The Arrhenian temperature dependency of the crystal growth simulations suggests that there is an activation barrier to crystal growth. The origin of the barrier could be due to the very slow diffusion of ions towards the surface of the crystal. Another hypothesis is that solvation shells are strongly attached to the ions, and there is a high barrier to displace water molecules for ion attachment to occur. Lastly, a temperature-dependent crystal growth rate could also be explained by an

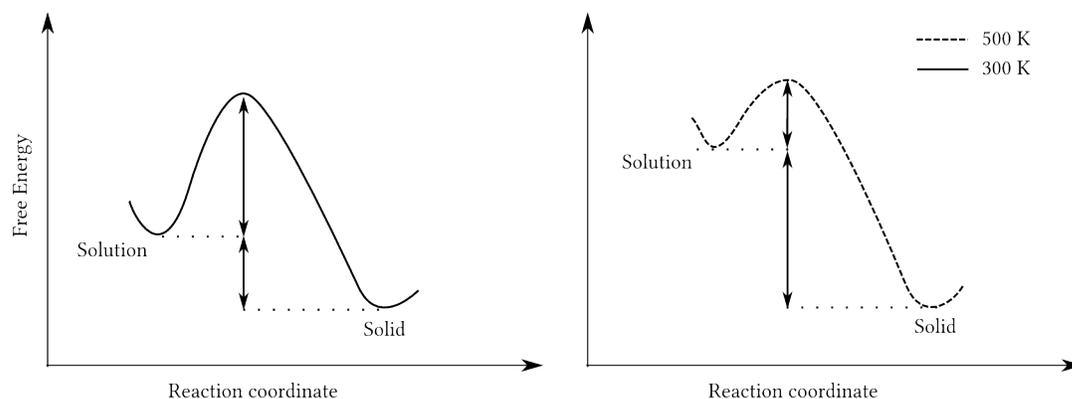


Figure 5.2: Schematic diagram of the free energy landscape at different temperatures for the crystal growth process. If the driving force (which is the difference in free energy of initial and final stages) increases and the barrier is weakly affected by the increase in temperature, the activation barrier (which is the height of the curve between initial and final stages) of the process decreases.

increase in the driving force at high temperature. If the solution is less stable at high temperature, while the energy of the “transition” state may be weakly affected, the activation barrier might be reduced (see Figure 5.2 for a diagram).

Since we did not observe crystal nucleation at different temperatures, similar statements are not directly applicable to the nucleation process. However, it is reasonable to assume that nucleation is affected by growth in the sense that nucleation cannot be observed without crystal growth. In this sense, we would expect the activation energy for growth to constitute a lower bound for the crystal nucleation activation energy.

5.4.1 Diffusion

One might expect diffusion to influence the speed at which crystal growth occurs by decreasing the overall mobility of the ions in solution. The diffusion constants for the Li^+ and F^- ions were estimated by calculating the mean square

T (K)	D ₊ (10 ⁻⁵ cm ² /s)	D ₋ (10 ⁻⁵ cm ² /s)
300	0.0190 (± 0.0007)	0.0205 (± 0.0002)
350	0.1348 (± 0.0050)	0.1350 (± 0.0106)
400	0.3842 (± 0.0174)	0.4216 (± 0.0133)
450	0.7918 (± 0.0313)	0.8329 (± 0.0115)
500	1.3514 (± 0.0239)	1.4928 (± 0.0089)

Table 5.3: Diffusion constants for Li⁺ (D₊) and F⁻ (D₋) at different temperatures.

displacement between 1 and 9 ns of our simulations, while making sure that no nucleation is taking place, and the results are reported in Table 5.3.

We calculated the diffusion coefficient at different temperatures to estimate the activation energy of the process using the Arrhenius equation:

$$\ln(D) = -\frac{E_a}{RT} + \ln(A), \quad (5.1)$$

where D is the diffusion coefficient, E_a is the activation energy, R is the gas constant, and A is the preexponential factor. The activation energy for diffusion can be estimated by performing a linear regression of $\ln(D)$ versus $\frac{1}{T}$ (see Figure 5.3 for an example), the resulting E_a for both Li⁺ and F⁻ diffusion was found to be ~ 26 kJ mol⁻¹. Interestingly, this value is less than half the activation energy for crystal growth (see below), suggesting that diffusion is not a controlling factor in the growth process.

5.4.2 Growth Barrier

By measuring the growth rate at different temperatures it is possible to estimate an activation energy for the growth process, similarly to the estimation of the activation energy for diffusion.

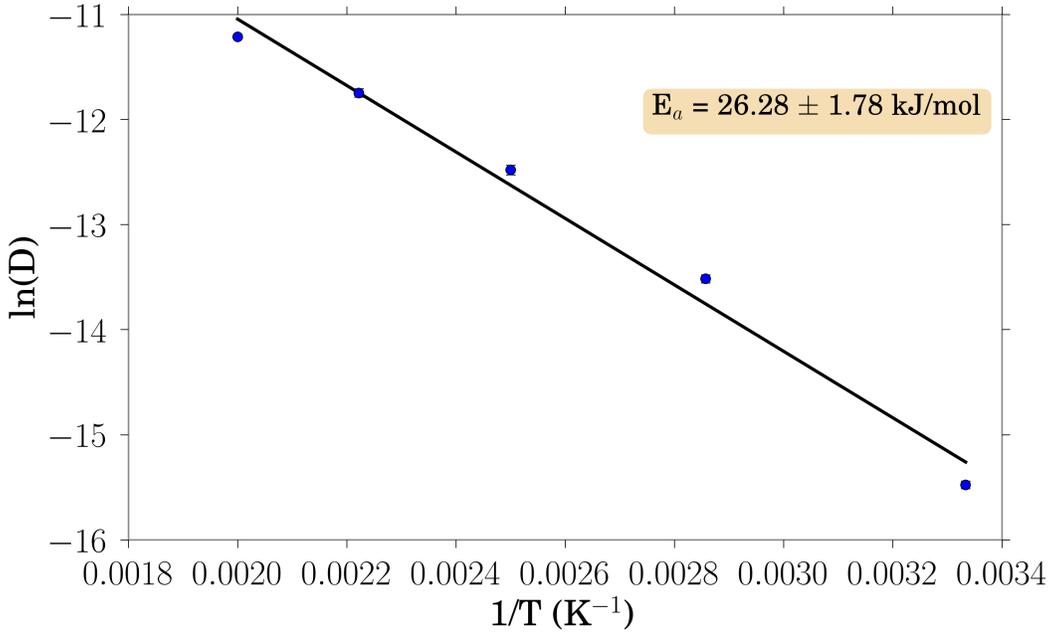


Figure 5.3: Linear regression to determine the activation energy for the diffusion process for Li^+ ions.

The growth simulations summarized in Table 5.2, were initialized by extracting a configuration containing a small cluster of ~ 300 ions from the simulation at 500 K and ~ 0.15 mole fraction. This seed was surrounded by the supersaturated solution, simulations were carried out at the temperatures 300, 350, 400, 450 and 500 K, and the growth profiles were monitored using a threshold on the q_8 order parameter as described in Chapter 5.3. The initial growth rate, r_g (over ~ 1 ns) was obtained for each temperature (Table 5.4) and fitted to the Arrhenius equation (Figure 5.4). The activation energy was estimated to be $\sim 50 \text{ kJ mol}^{-1}$.

If we compare the activation energy for growth with the activation energy for diffusion, it is easy to see that the activation energy for growth is much slower than that of diffusion, suggesting that the growth rate is not diffusion limited. The origin of such high activation barrier probably lies in the solvent-ion interaction, for

T(K)	r_g (ion ns ⁻¹)
300	-0.05
350	0.23
400	1.66
450	11.41
500	32.40

Table 5.4: Growth rates at different temperatures

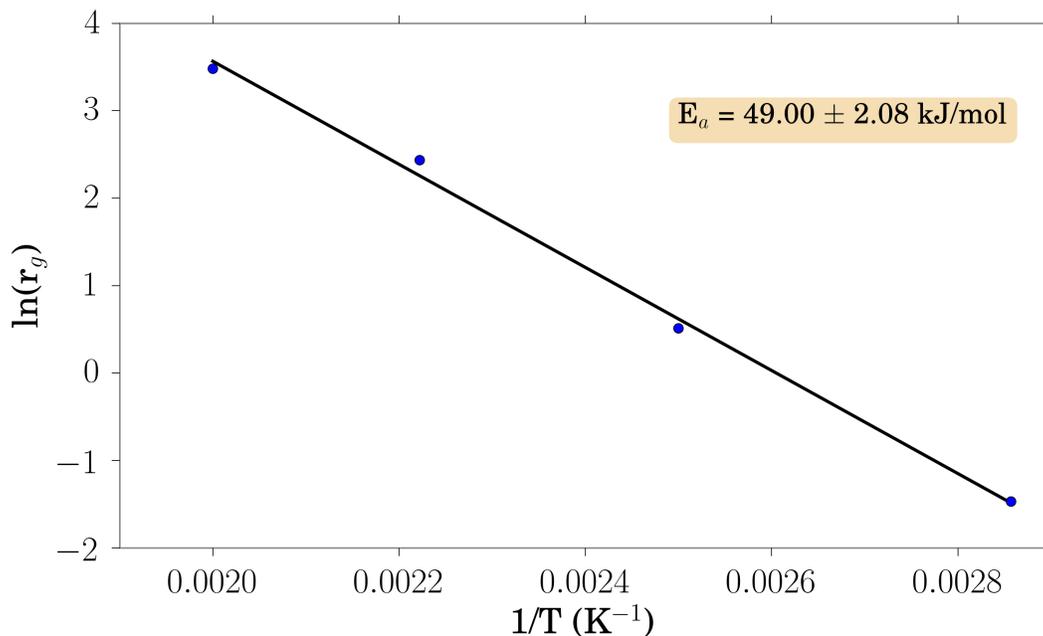


Figure 5.4: Arrhenius fit to the growth rates

this reason we further explored the structure and dynamics of the supersaturated solution.

5.4.3 Solution Structure and Dynamics

One important property of the solution is the degree of association of the ions. To study ion association, we performed connectivity based clustering, as described in Chapter 5.3. Using this method, every ion is assigned to a cluster, and it is

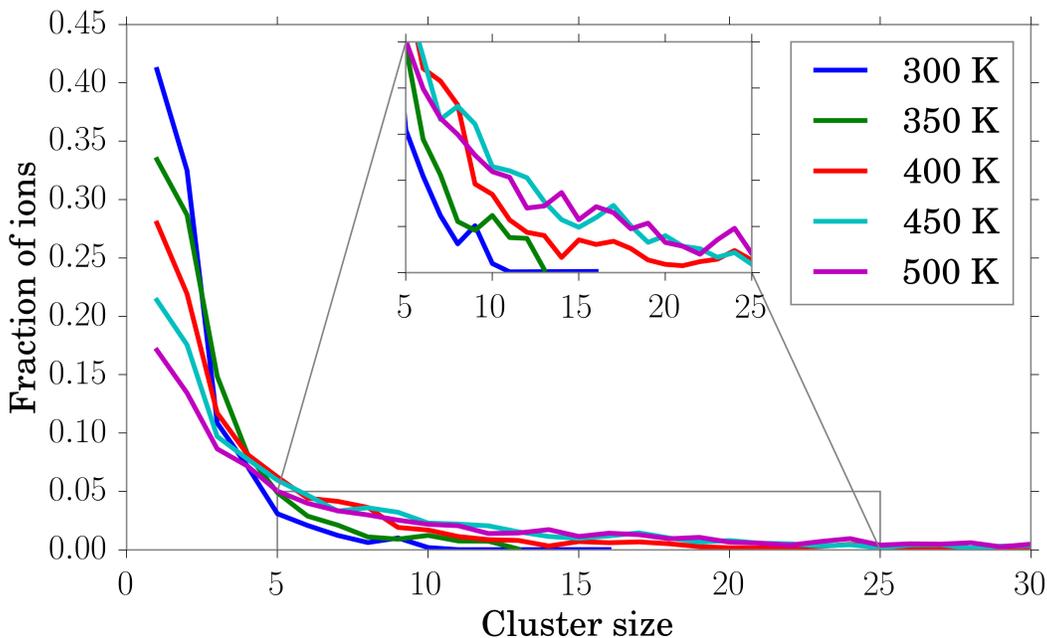


Figure 5.5: Fraction of ions in connectivity based clusters of a certain size at different temperatures.

possible to study how the distribution of free ions, ion pairs, ion triples, and higher order aggregates varies at different temperatures.

The fraction of ions that are part of a cluster of a certain size at different temperatures is shown in Figure 5.5. At 300 K, about 40 % of the ions are not associated (in clusters of size 1), about 35% of the ions are part of clusters of size 2, and the remaining 15% are part of larger clusters up to ~ 16 ions. As the temperature increases, the distribution gradually shifts towards clusters of higher size and, at 500 K only $\sim 15\%$ of the ions are not associated. This shows that as temperature increases the ions have a tendency to associate into larger aggregates, thus favoring nucleation.

Further evidence can be found from the Li^+F^- radial distribution functions of the solution at different temperatures shown in Figure 5.6. At higher temperatures,

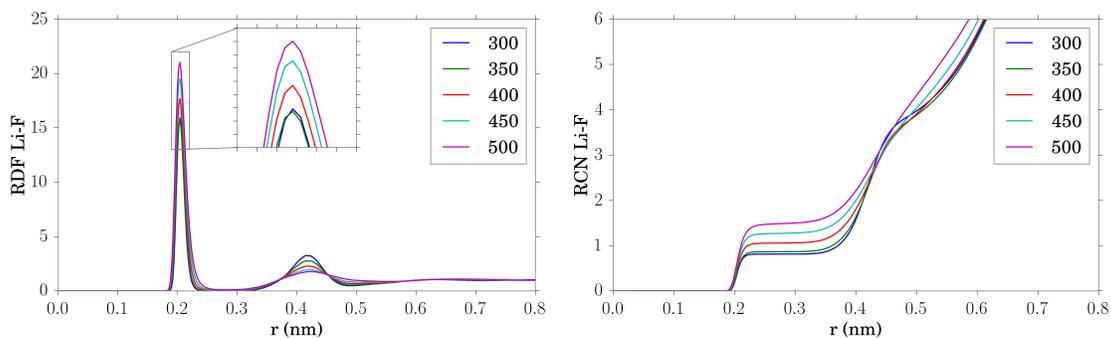


Figure 5.6: Li^+F^- radial distribution functions at different temperatures. The contact peak, characteristic of ion pairs increases with temperature, while the second peak (solvent-separated pairs) decreases.

the LiF contact peak increases substantially, while the second peak, corresponding to solvent-separated pairs, decreases in intensity. This result is in agreement with the ion association plots (Figure 5.5) as more ion association increases the number of contact pairs.

An important point is that the ion association distributions shown in Figure 5.5 describe the metastable “equilibrium” of the solution at different temperatures. This means that, before nucleation happens, the system spontaneously relaxes to have a certain ion association pattern. This effect, as well as the change in solubility, is likely to be due to the changing dielectric constant of water at different temperatures and pressures.

Ion association distributions, however, do not describe the dynamics of the solution. To assess how clusters exchange ions, it is useful to introduce a new measure that we will refer to as *surviving pairs*.

At a given time frame, the ions are associated in a number of clusters. Two different frames may (and usually) have a different number of clusters and, in order to define a lifetime, it is necessary to match the clusters in two different frames

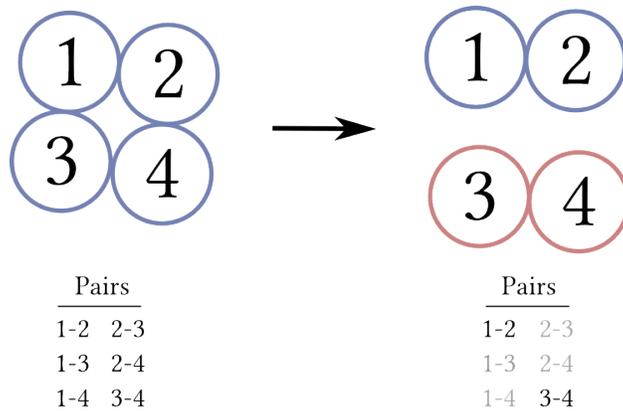


Figure 5.7: Diagram of surviving pairs. Initially, the $\{1, 2, 3, 4\}$ cluster contains 6 unique pairs, after the cluster splits, 2 of the pairs are still associated.

(this process is usually called *entity resolution* and is described in detail in Chapter 3). The matching can be quite problematic when there is substantial rearranging between the frames, as there is a large redistribution of ions between the clusters.

Another approach to the problem is to count the number of pairs that are associated in a certain time frame, and count the number of pairs that are still associated at a different time frame. As shown in Figure 5.7, the *surviving pairs* is perfectly defined even when clusters split and, while it doesn't provide information about the lifetime of individual clusters, can be used as a global measure of how long lived are ion clusters in solution.

By plotting the surviving pairs as a function of time delay between frames it is possible to obtain a plot that indicates how many pairs are still associated up to time t (Figure 5.8). As temperature increases, the initial number of pairs increases (corresponding to time delay $t = 0$) as there is a higher degree of association, however, their decay with t is much more pronounced. The 300 K surviving pairs curve shows that, at this temperature, the associated structures are very stable, and decay slowly over the course of $\sim 20 - 30$ nanoseconds (the range is too wide to be

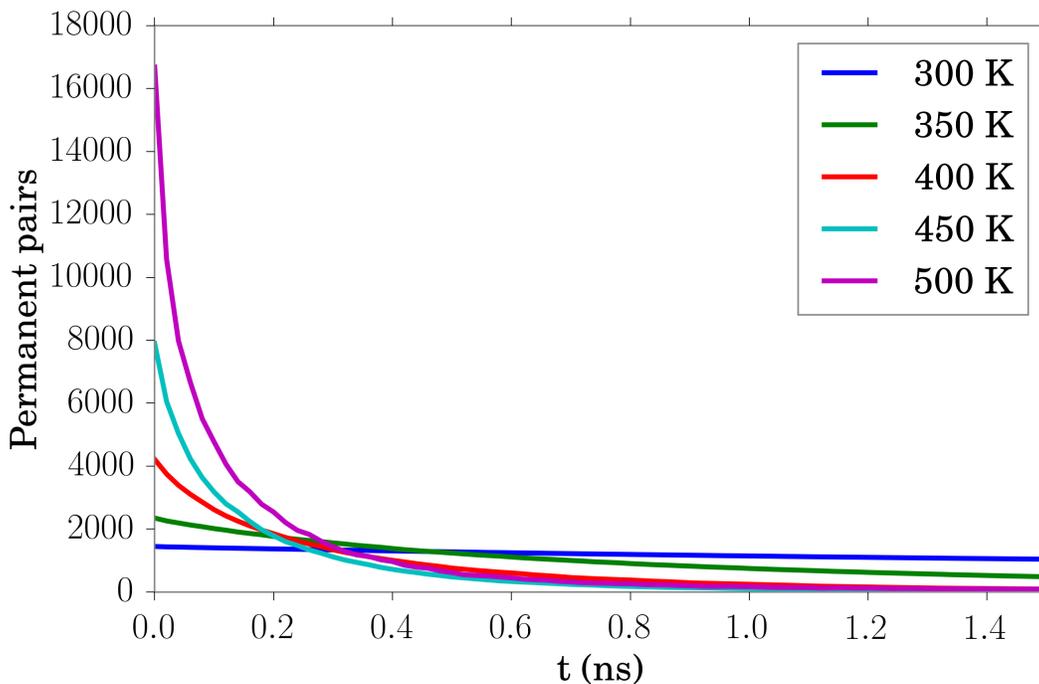


Figure 5.8: Surviving pairs at a time delay of t ns for a range of different temperatures. As the temperature increases, there are more associated pairs that decay at a faster rate.

shown on the plot). In contrast, temperatures higher than 400 K show that after ~ 1 ns most of the pairs have decayed, suggesting fast rearrangements of the ion aggregates present in solution.

A cluster analysis similar to the one presented in Chapter 3 was performed for the LiF supersaturated solution. The number of clusters detected per frame, using the methodology described in Chapter 3, is reported in Table 5.5. An important finding is that, at low temperatures, no clusters are detected using this method. The main reason for this effect is that the aggregates that form in the LiF supersaturated solution do not possess high enough crystallinity to pass through the filtering step (cf. Chapter 3). At higher temperatures, there is a sharp increase in number of clusters that, presumably, are a prerequisite for crystal nucleation. In contrast, the

Compound	T (K)	$n_c \times 10^5$	m.f.
LiF	300	0	0.15
	350	0	
	400	0.15	
	450	1.33	
	500	27.65	
NaCl	300	142.33	0.22

Table 5.5: Average number of clusters (n_c) per time frame at different temperatures, for LiF and NaCl supersaturated solutions (see Chapter 3 for details). n_c is normalized by the number of ion pairs in the simulations to allow comparisons between simulations with different numbers of ions. The concentration in mole fraction is reported in the m.f. column.

NaCl solution, as discussed in Chapter 3, is composed of a large quantity of clusters with longer lifetimes.

Lifetime analysis was not performed on the cluster data produced from the LiF simulations because at high temperatures the time interval at which configurations were sampled (0.02 ns) is large compared to the time it takes a typical cluster to redissolve, hence it was not possible to follow the time evolution of most clusters.

5.4.4 Solubility

If the solubility decreases substantially at high temperature, one would expect nucleation and growth rates to increase due to a higher driving force (since our simulations are all at the same concentration, the solubility determines the degree of supersaturation). Experimentally, the solubility of LiF increases with temperature (at least in the range ~ 300 to 350 K at 1 atm).¹⁴² However, it is still possible for our LiF model to show a decrease in solubility with temperature due to changes in the solvent, for example the dielectric constant.

We estimated the solubility of LiF at 300 K by dissolving a crystal until dissolution stops, and at 500 K by both dissolving a crystal and, by growing a nucleus in a supersaturated solution until growth stops. More specifically, the dissolution and growth were assumed to be completed when there was no change in the solution concentration which occurs after ~ 100 to 200 ns.

The solution concentration was calculated by detecting connectivity based clusters (see Chapter 5.3) and counting the number of ions that are part of clusters with size less than 100 ions. As the simulation contains a single cluster of size ~ 2000 ions, and transient clusters are of size definitely smaller than 100 ions (see Figure 5.5), these numbers represent a good estimate of the ions present in solution. The dissolution and growth profiles are shown in Figure 5.9.

It is useful to notice that the starting point of the dissolution at 300 K and at 500 K is larger than 0 ns. This is due to the fact that, since the initial crystal was extracted from a particular configuration, it possessed a small net charge due to ions fluctuating in proximity of the interface. To avoid net charges, extra ions were added to the solution to retain the electroneutrality. Geometry optimization was also required and produced slightly different starting condition for the two temperatures. The fluctuations observed for the dissolution simulation at 500 K show that the initial conditions are quite close to equilibrium; in fact, after an initial rapid dissolution, the cluster grows again until the solution stabilizes around a concentration of ~ 0.01 mole fraction.

The solubilities estimated using these methods yielded values in the range of ~ 0.01 to 0.03 mole fraction, which is much larger than the experimental solubility of ~ 0.001 mole fraction. Interestingly, at high temperature there is a decrease in solubility and this could be attributed to a change in solvent nature. To test this

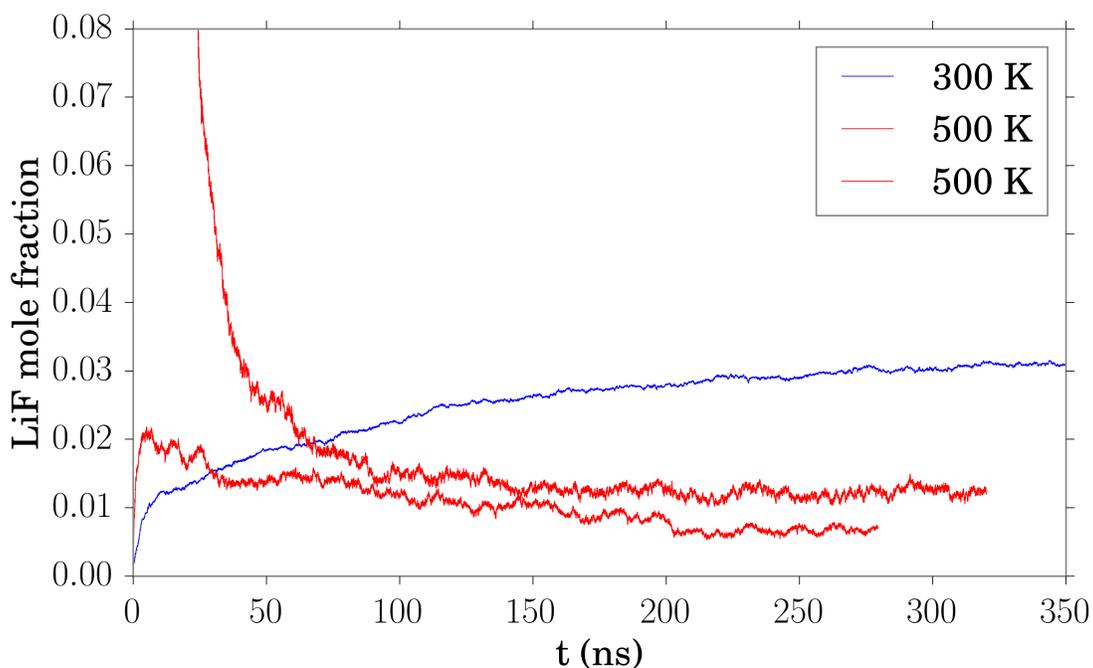


Figure 5.9: Solution concentration profiles obtained by dissolving a crystal at 300 K (blue), by growing a crystal at 500 K (red, top line), and dissolving a crystal at 500 K (red, bottom line). The estimated solubilities are ~ 0.025 mole fraction at 300 K and ~ 0.01 at 500 K.

hypothesis, we measured the dielectric constant of SPC/E water to be ~ 71 at 300 K and ~ 40 at 500 K. The decrease in dielectric constant of the solvent is in line with the decrease in solubility at high temperature.

The degree of supersaturation can certainly influence crystal growth rate as it is the driving force of the process. However, as it was shown above, the crystal growth rate shows an Arrhenian dependency on temperature, and is not likely due to a change in solubility, because we would expect the associated change in activation energy to yield a nonlinear Arrhenius plot.

Similar effects can also be expected for the nucleation rate; in fact, increasing the degree of supersaturation by pushing the concentration to ~ 22 mole fraction

didn't result in nucleation at 300 K, showing that an increase in driving force is not enough to trigger nucleation at that temperature.

5.5 Summary and Conclusions

In this Chapter we focused on the nucleation and growth of LiF crystals from solution. Our results suggest that ions in the supersaturated solution are strongly solvated and increasing the temperature helps ions to overcome the barrier to dehydration, favoring growth.

While we could not directly measure nucleation rate at different temperatures, we were able to measure growth rate of an already nucleated crystal, and we found that the temperature dependence of the growth rate follows the Arrhenius law, indicating that the process is activated and has an activation energy of ~ 50 kJ mol⁻¹. Such a high activation energy can not be explained by ion diffusion, which has a much smaller activation energy (~ 20 kJ mol⁻¹).

At high temperature, the LiF solution contains a large quantity of ion clusters that form and dissolve very quickly but, despite the fast decay of these structures, the solution is able to achieve nucleation. This situation is in contrast with what was observed in Chapter 3, where the NaCl solution was found to have many long lived clusters that failed to achieve the size and crystallinity for nucleation to occur.

A mechanism consistent with our results for LiF is a two-step nucleation process.¹⁴⁴ The first step involves cluster generation from solution and, in the case on LiF, clusters appear at high temperature and pressure, where they are easier to form because dehydration is fast enough to achieve kinetically, and ion clusters are also favored because the dielectric constant of the water decreases. Once a sufficient

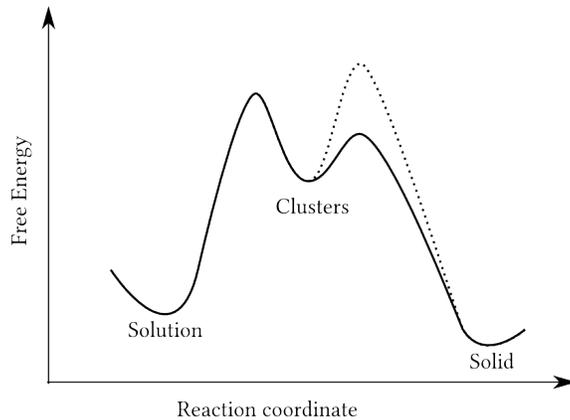


Figure 5.10: Possible free energy landscape for LiF nucleation from solution. A first barrier is required to go from the free ions (solution) to the cluster aggregates. The cluster aggregates are metastable, continuously form and disappear from solution, and are more easily found at high temperature. The second barrier reflects the fact that most clusters do not nucleate but dissolve back into solution, depending on cluster related properties such as size, crystallinity, and surface tension. In absence of nucleation rate measurements, it is not possible to estimate the relative magnitude of the two barriers, and two possible options are displayed as solid and dashed lines.

number of clusters are generated from the saturated solution, there is another barrier, similar to that postulated by CNT, that is likely related to cluster properties such as size and crystallinity as we found for NaCl (Chapter 3). Unfortunately, the data we collected is not sufficient to determine the magnitude of the second barrier, as it requires precise measurement of nucleation rates sampled at different temperatures. A diagram of possible free energy landscapes of the LiF nucleation process is shown in Figure 5.10.

The solubility of LiF was estimated to be ~ 0.03 mole fraction at 300 K, and was found to decrease at high temperature to reach a value of ~ 0.01 . The variation is likely due to the decreasing dielectric constant of the solvent. The calculated value of the solubility is substantially different from the experimental solubility of ~ 0.001 mole fraction, which likely indicates that, in this respect, the JC model is not accurate for LiF, and further model improvements are needed.

Chapter 6

Molecular Dynamics Simulation of NaCl Dissolution.

6.1 Overview

Molecular dynamics simulations are used to investigate the dissolution of NaCl nanocrystals (containing ~ 2400 ions) in water. We focus on systems under sink conditions at 300 K, but the influences of concentration and temperature are also investigated. Cubical, spherical, tablet-shaped, and rod-shaped nanocrystals are considered, and it is shown that the initial shape can influence the dissolution process. Dissolution is observed to occur in three stages: an initial period where the most exposed ions are removed from the crystal surface, and the crystal takes on a solution-annealed shape which persists throughout the second stage of dissolution; a second long intermediate stage where dissolution roughly follows a fixed rate law; and a final stage where the small residual crystal ($\lesssim 200$ ions) dissolves at an ever increasing rate until it disappears. The second stage of dissolution which applies for

most of the dissolution process is well described by classical rate equations which simply assume that the dissolution rate is proportional to an active surface area from which ions are most easily detached from the crystal. The active area depends on the initial crystal shape. We show that for our model NaCl nanocrystals the rate determining step for dissolution under sink conditions is ion detachment from the crystal, and that diffusion layers do not exist for these systems.

6.2 Introduction

Solid dissolution is an important process in many physical systems and situations, with one obvious example being in the area of drug development.^{145,146} Many drugs are administered in solid form and dissolve in the gastrointestinal tract before being absorbed by the body. As the dissolution properties are related to the drug bioavailability,^{147,148} there is considerable current interest in dissolution processes.¹⁴⁹

During the past century, several dissolution models based on different assumptions have been suggested^{147,149,150} to describe and interpret dissolution rates, and identify the factors that affect them. The original model was put forward by Noyes and Whitney,¹⁵¹ and further generalized by Brunner and Tolloczko¹⁵² to account for varying solid surface area during particle dissolution. For spherical solid samples, the integrated form of the Brunner-Tolloczko equation is the well-known cube root law, whereby the cube root of the particle weight decreases linearly in time during dissolution, first derived by Hixson and Crowell.¹⁵³ Another early model, based on the assumption that the dissolution rate is controlled by solute diffusion from a concentrated layer of solution adjacent to the solid surface, was suggested by

Nernst¹⁵⁴ and Brunner.¹⁵⁵ Under some conditions the Nernst-Brunner equation also leads to the cube root law. Additionally, square root¹⁵⁶ and two-thirds root^{156,157} dissolution laws have been obtained and used in dissolution research. These models are all based on theoretical assumptions or empirical observation, depending on the substance being dissolved and its environment; however, little is known from a microscopic, mechanistic point of view. Furthermore, the classical models do not take any account of shape and finite size effects (other than surface area) that one might expect to be of some importance in nanoscale crystals.

Thanks to large increases in computational power, molecular simulation has started to gain traction as a tool in dissolution research. Studies have been conducted to uncover the mechanics of the first steps of dissolution and growth of small crystalline compounds.^{158–160} Gao and Olsen¹⁵⁹ have reported simulations of the drug acetaminophen in water, giving insights into the initial stages of the dissolution process. Several studies have also been done to analyze the interactions that take place at the interface between water and NaCl.^{161–165}

In the present paper, we investigate the complete dissolution of NaCl nanoparticles in water employing molecular dynamics simulations. NaCl was chosen because of its simple structure and because of the availability of force fields developed to reproduce both solid and solution phase properties.¹ Our results provide an atomistic view of the dissolution process at every stage, from initiation, through a fixed rate law regime, to the final disintegration of the residual crystal. Additionally, we examine the influence of particle shape on dissolution, and compare our results with suitable, shape-adapted, classical dissolution models. The influences of solution concentration and temperature on the dissolution rate are considered. The temperature dependence allows us to estimate the activation energy associated with ion

	σ (nm)	ϵ (kJ mol ⁻¹)
Na ⁺	0.2160	1.4754533
Cl ⁻	0.4830	0.0534924

Table 6.1: Joung-Cheatnam NaCl parameters¹ for the Lennard-Jones potential.

detachment, which for NaCl crystals controls the dissolution rate.

The remainder of this Chapter is divided into three parts. The models and methods are given in Section 6.3, our results are described and discussed in Section 6.4, and our conclusions are summarized in Section 6.5.

6.3 Models and Methods

Several simulations of NaCl nanocrystals dissolving in water are reported. In all simulations, the nonbonded, site-site interactions $u(r_{ij})$ consist of Lennard-Jones (LJ) plus electrostatic terms such that

$$u(r_{ij}) = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}, \quad (6.1)$$

where q_j and q_i are the charges on sites i and j , r_{ij} is the site-site separation, σ_{ij} and ϵ_{ij} are the LJ length and energy parameters, and ϵ_0 is the permittivity of free space.

To model Na⁺ and Cl⁻ ions we adopt the parameters of Joung and Cheatnam¹ (Table 6.1), used together with the SPC/E water model.² The Joung and Cheatnam parameters were developed to reproduce both solid and solution properties, such as the solubility.¹ The usual Lorentz-Berthelot combining rules, $\sigma_{ij} = (\sigma_i + \sigma_j)/2$, $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$, were used to calculate the LJ interactions.

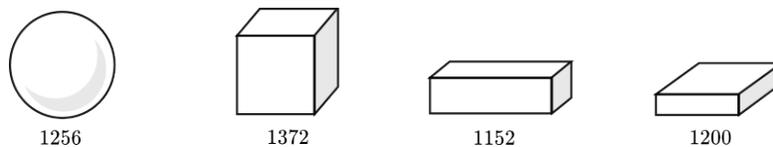


Figure 6.1: A representation of the crystals simulated. The number of NaCl ion pairs is given below each illustration.

In order to investigate the possible influence of particle shape (number of crystal layers, surface area, etc.) on dissolution rate, NaCl crystals of different shapes were generated by repeating the NaCl unit cell along the three crystal axes directions, to produce the following four structures (Fig. 6.1):

Cubic Crystal (1372 ion pairs) generated by repeating the unit cell 7 times in all crystal axes directions.

Spherical Crystal (1256 ion pairs) generated by taking a cubic crystal of ~ 2000 ion pairs and removing all ions pairwise (such as to maintain a charge neutral crystal) whose distance from the center is greater than 2.28 nm.

Rod-Shaped Crystal (1152 ion pairs) generated by repeating the unit cell 18, 4, and 4 times along the crystal axes.

Tablet-Shaped crystal (1200 ion pairs) generated by repeating the unit cell 10, 10 and 3 times along the crystal axes.

All dissolution simulations were carried out under *NPT* conditions with the pressure fixed at 1 bar. The numbers of NaCl and water molecules in each simulation, together with the temperatures considered are summarized in Table 6.2.

For the present model the saturation concentration estimates reported by Aragonés *et al.*⁹⁸ correspond to NaCl mole fractions in the range 0.08 – 0.09. Most of our simulations were carried out at NaCl mole fractions $\lesssim 0.015$, or at $\sim 1/6$ of the

	NaCl	H ₂ O	T (K)
Sphere	1256	97956	300
	1256	97956	320
	1256	97956	340
	1256	75917	300
	1256	48238	300
Cube	1372	92838	300
Rod	1152	77952	300
Tablet	1200	81200	300

Table 6.2: Summary of crystal shape, the number of ion pairs, the number of water molecules, and the temperatures used in the simulations.

saturation concentration. This meets the so-called “sink condition” which refers to solutions sufficiently dilute that ion reattachments do not significantly influence the dissolution rate.¹⁶⁶ Note that sink conditions are usually attained when the volume of solvent is 3 – 10 times greater than the saturation volume.¹⁶⁶ We tested that our systems did obey sink conditions by directly monitoring ion attachment and detachment events, as well as by running simulations at different concentrations, as described below (Section 6.4).

All simulations were performed employing the GROMACS¹⁰⁰ molecular dynamics package version 4.5.4. The temperature was controlled using the velocity-rescale thermostat⁷⁹ with a relaxation time of 0.1 ps. The pressure was kept constant by applying the Berendsen barostat⁷⁸ with a compressibility of $4.5 \times 10^{-5} \text{ bar}^{-1}$ and a relaxation time of 1.0 ps. The time step chosen for all the simulations was 2 fs. Periodic boundary conditions with the usual minimum image convention were applied, and all short-range interactions were spherically truncated at 0.9 nm. The long-range electrostatic interactions were accounted for using the particle mesh Ewald (PME) method⁷⁴ with a Fourier spacing of 0.20 nm. Constraints on all bonds were

maintained using the LINCS algorithm.⁷⁷

Initial configurations were generated by placing the NaCl crystal at the center of the simulation cell surrounded by water molecules randomly placed on a grid spaced by 0.3 nm. The system was equilibrated for 1 to 2 ns to obtain a stable pressure, temperature, and potential energy during which time no substantial part of the crystal dissolved. The system was then evolved in time until complete dissolution of the crystal was achieved. For the systems considered here, this required times ranging from ~ 100 to ~ 700 ns, depending on the nature of the crystal, the amount of water present, and the temperature. Note that all simulations were at 300 K, except for the spherical crystal, where two additional simulations were performed at 320 K and 340 K, and used to estimate an activation energy for the dissolution process.

An order parameter based on the number of neighboring ions is used to classify the ions as being part of the crystal or part of the solution. Ions belonging to the crystal have a substantially higher number of neighbors than those in solution, and this difference is sufficient to define a relatively simple order parameter. The order parameter for a particular ion is calculated by counting the number of neighboring ions (both positive and negative) within a specified radius. The actual value of the neighbor search radius did not greatly affect the result as long as there is a substantial difference between the number of neighbors around ions in the crystal and in the solution. In the present calculations we selected a search radius of 0.6 nm. For this radius, ions associated with the crystal have on average 20 neighbors, while those in solution have 6. Therefore, an intermediate value of 11 was chosen as the classification threshold; ions having 11 or more neighbors were classified as being part of the crystal, otherwise they were classified as being part of the solution.

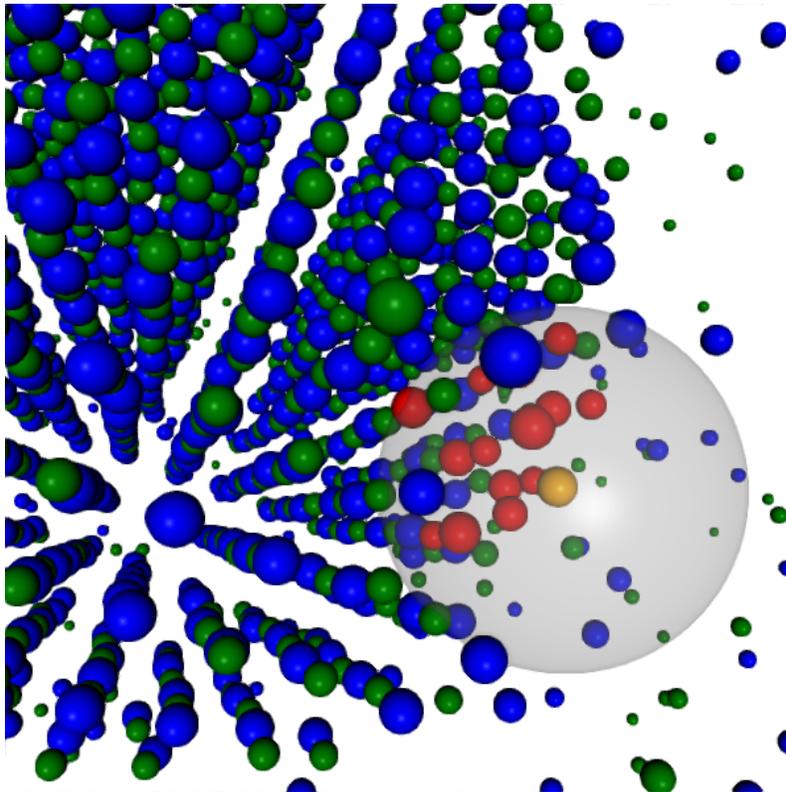


Figure 6.2: An example of the order parameter applied to a surface ion. The central atom (yellow) is surrounded by 14 neighbors (red). The 0.6 nm cutoff applied to the central atom is highlighted with a gray transparent sphere.

An illustration of how this order parameter detects an ion at the crystal surface for a particular configuration is given in Fig. 6.2. Using this order parameter we could closely follow the dissolution of each crystal independent of its shape.

6.4 Results and Discussion

6.4.1 Stages of Dissolution

We first examine in detail the stages of the dissolution process. Simulations were used to obtain dissolution profiles for several model NaCl crystals. A dissolu-

tion profile is a plot of the crystal weight, or, equivalently, the number of ions in the crystal, as a function of time. The time derivative of this profile is the dissolution rate, or the change in crystal weight (or ion number) per unit time. The dissolution profile can be influenced by the solution concentration and structural properties of the crystal, such as size and shape. Qualitatively, we observed that the dissolution process occurs in three stages that can be roughly described as follows: Initial Stage: detachment of the most exposed ions located on any sharp edges. Fixed Rate Regime: crystal dissolution after the edges are removed appears to closely follow a fixed rate law. Rapid Dissolution: the crystal reaches a “stability limit” and quickly disappears.

Initial Stage

During the first few nanoseconds, water interacts with the most exposed ions and these move readily from crystal to solution. The detachment of these ions leads to defects on the surface of a geometrically shaped crystal. In the initial stage, the dissolution rate depends on the crystal shape, as this influences the arrangement of the most exposed ions.

Dissolution profiles for initially cubic and spherical crystals are plotted in Fig. 6.3, and snapshots of the cubic crystal at various points on the dissolution curve are shown in Fig. 6.4. For the cubic crystal, one observes (see the magnification of the short-time region in Fig. 6.3) an initial steep slope between 0 and ~ 20 ns; dissolution then quickly slows down, and enters a nearly linear regime (the actual time dependence is discussed below). During the rapid initial stage ~ 200 ions leave the crystal, and as can be seen from Fig. 6.4 (snapshot at 20 ns), these come mainly from the corners and edges of the crystal. From Fig. 6.4 we see that after the

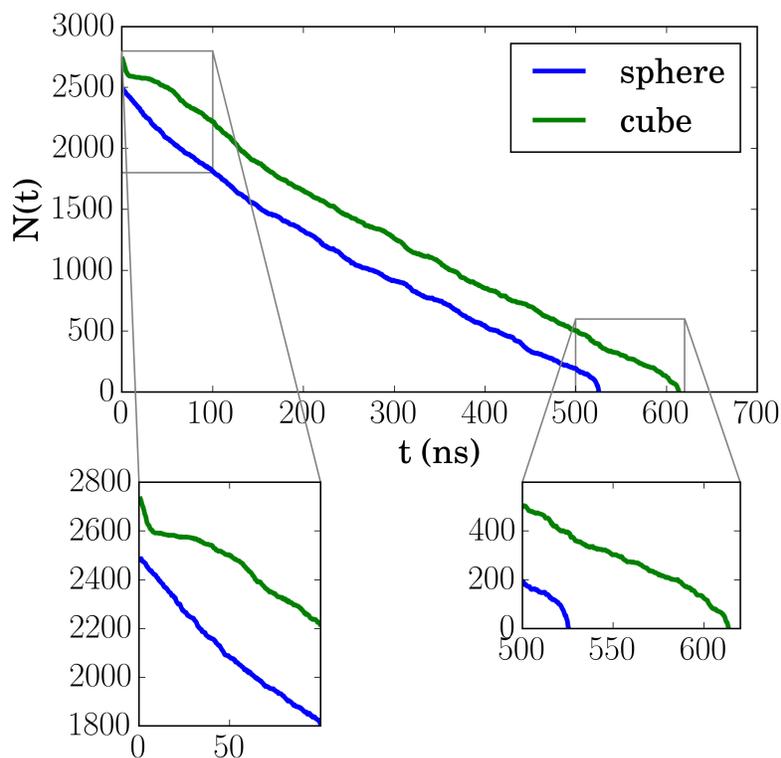


Figure 6.3: Dissolution profiles for the cubical and spherical crystals displayed as number of ions in the crystal vs time. The starting and final sections of the profiles are detailed in the zoomed-in plots.

corners and edges of the initial crystal are gone, more ions are detached from the sites generated by their removal, and water continues to gradually consume crystal edges until they are completely rounded off. At 185 ns (Fig. 6.4), the dissolving crystal is almost spherical in shape, and its structure and dissolution rate are very similar to those of the initially spherical crystal. Note from Fig. 6.3 that, after the initial stage, the dissolution profiles of the cubic and spherical crystals are essentially parallel to each other. Note also, that the initial rate is much faster for the cube than for the sphere, which has no corners and edges. Nevertheless, for the sphere as well, dissolution is faster at the beginning, as the more weakly attached ions are swept from the surface.

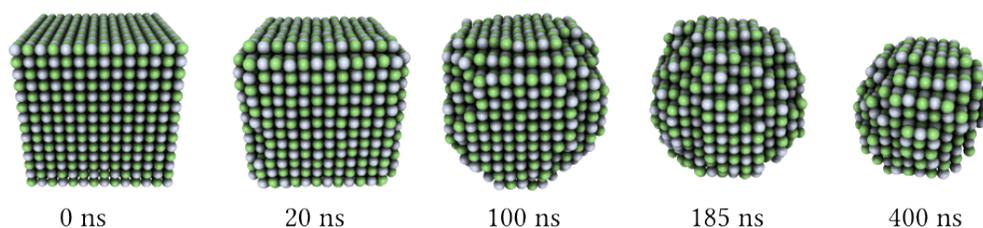


Figure 6.4: Snapshots corresponding to different points in the dissolution profile of the cubic crystal. After 20 ns, only ions located at edges and corners are removed. The edges and corners gradually get consumed (100 ns) and the crystal becomes roughly spherical after about 185 ns. After that point the shape doesn't change until the final stage of dissolution.

The dissolution profiles of rod-shaped and tablet-shaped crystals also show fast initial stages followed by slower nearly linear behavior (see Fig. 6.5, and the discussion below). The overall qualitative picture can be seen from snapshots taken from rod and tablet dissolution trajectories shown in Figs. 6.6 and 6.7, respectively. For the rod (Fig. 6.6), ions are preferentially removed from the ends, and for the tablet (Fig. 6.7), water peels layers from the thin sides, while none are removed from the interior of the larger flat surfaces. In general, as the dissolution continues, the detachments progressively round the corners of the crystals, giving both the tablet-shaped and rod-shaped crystals a more cylindrical appearance. After the initial stage, for the rod and tablet crystals we are left with structures that resemble long and flat cylinders, respectively. From that point on, the dissolution takes place from the side in the case of the flat cylinder (originally a tablet) and from the ends in the case of the long cylinder (originally a rod). These dissolution induced shape changes are included in the dissolution rate laws discussed below.

Physically, it is not difficult to understand the detachment patterns discussed above. We would expect ion detachment to be an activated process (see below for confirmation) and the activation energy of a particular ion will depend on its location

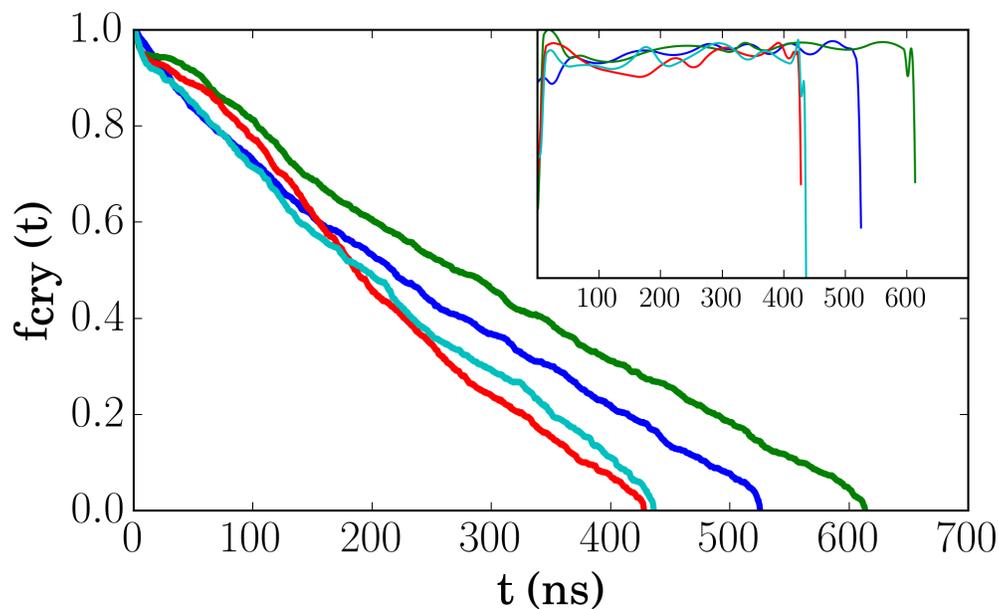


Figure 6.5: Comparison of the dissolution profiles for different shapes expressed in terms of fraction of crystal dissolved $f_{\text{cry}}(t)$, the derivatives df_{cry}/dt (on the order of -10^{-3} ns^{-1}) are depicted in the inset. The color coding is: dark blue for the sphere; green for the cube; red for the tablet; light blue for the rod. The rod-shaped and tablet-shaped crystals show a higher overall dissolution rate.

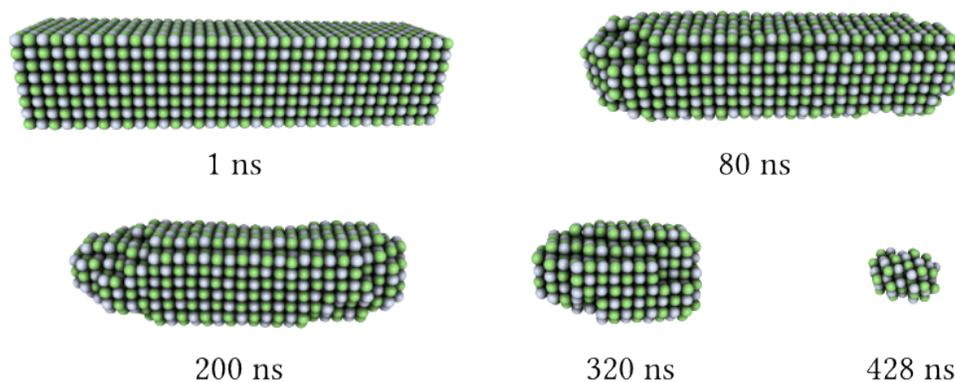


Figure 6.6: Snapshots along the rod-shaped crystal trajectory. Ions are detached mainly from the ends of the rod.

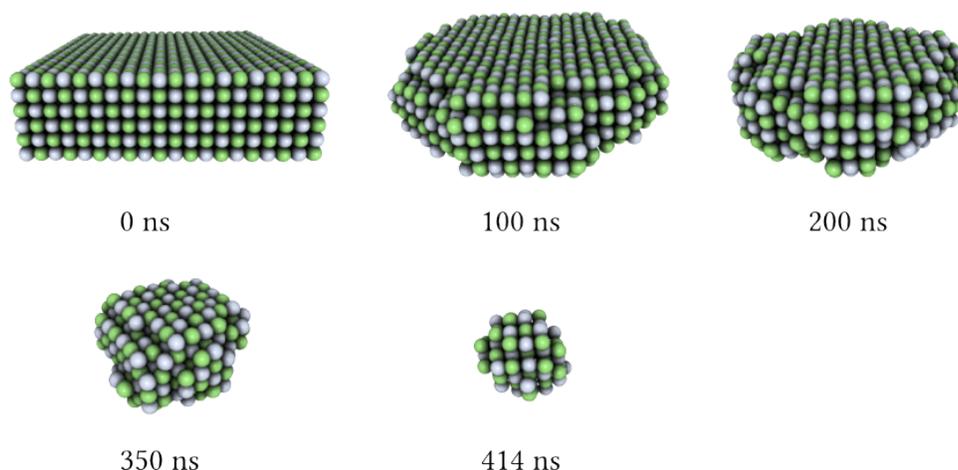


Figure 6.7: Snapshots along the tablet-shaped crystal trajectory. The top and bottom surfaces are never attacked by water molecules.

in the crystal. We would expect ions at corners and edges to be more weakly bound to the crystal and hence to have a lower activation energy. Moreover, the edges and corners are more exposed to water molecules, which should aid in getting over activation barriers inherent in the crystal.

To further examine this reasoning, we simulated a system with a perfect NaCl (001) crystal face (no defects, corners, or edges) in contact with water. The crystal face was constructed by repeating the unit cell 7, 7, and 4 times (for a total of 1578 ions) along the x , y , and z axes, and placing the resulting crystal wall in a simulation cell periodic in the x and y directions. The crystal wall was in contact with 2855 water molecules, with all conditions, parameters, and potentials identical to our other dissolution simulations at 300 K. During an ~ 200 ns simulation, not a single ion left the crystal wall, demonstrating a much lower probability of detachment for ions located in large flat crystal surfaces, and highlighting again the importance of corners, edges etc., in the dissolution process.

The picture that emerges is that the detachment of ions depends on the local

environment on the surface. To quantify this effect we performed a calculation of the probability of detachment (within a certain time frame) given the number of ions present (0 to 6) in the first coordination shell.

The ions detached from the surface of the spherical crystal within 1.6 ns was recorded and averaged in a time frame between 20 and 40 ns. The distribution of the order parameter showed that the detached ions assume a value no larger than 22 neighbors. Given that the detached ions come from the surface, this threshold value was used to classify ions as part of the surface.

The ions on the surface were further classified based on the number of immediate neighbors (within a radius of 0.37 nm), corresponding to the first coordination shell. This step was necessary to analyse the relationship between the local environment and the probability of detachment.

The total number of ions on the surface, given the local environment is reported in Fig. 6.8, top right corner. The number of detached ions is reported in the bottom right corner, by calculating the fraction of detached ions for each local environment, it is possible to obtain the probability of detachment (reported on the left).

The probabilities of detachment show that, as the coordination number increases (the ion is more internal), the probability of detachment greatly decreases. However, the total number of ions detached shows that most of the ions dissolved come from an environment containing 3 neighbors. This is due to the fact that, even if the probability of detachment is very low (~ 0.08), the larger supply of those ions on the surface makes them, on the 1.6 ns time frame, the most common kind of ion to detach.

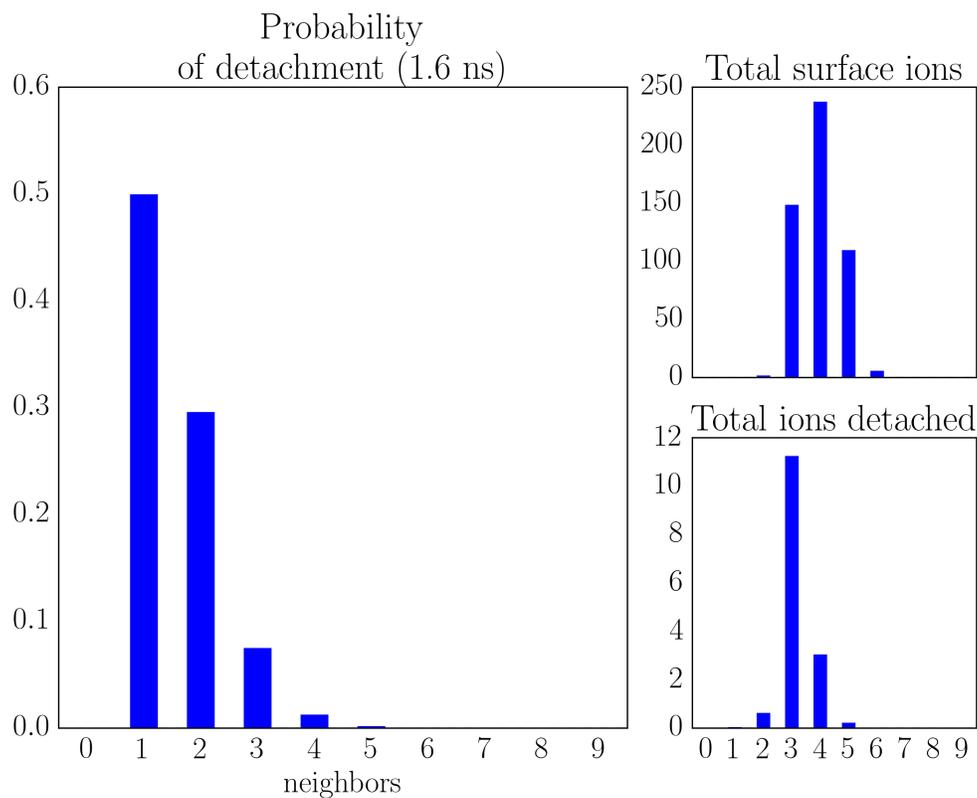


Figure 6.8: Three quantities obtained for the spherically shaped crystal as functions of the number of neighbor in the first coordination shell (within 0.37 nm), Left panel: The relative probability of detachment (within 1.6 ns) of a surface ion, given the number of neighbors. Upper right panel: The total number of ions on the surface at $t = 0$. Lower right panel: The total number of ions detached over 1.6 ns. All quantities are averaged over 50 time slices between 20 and 40 ns of the dissolution simulation.

Fixed Rate Regime

After the initial stage, the dissolution of all crystals follows a similar route. The dissolution profile is almost linear (the actual behavior is discussed in more detail below), until the remaining crystal (~ 200 ions) becomes very unstable and there is a sudden and continuing increase in dissolution rate as shown in Figure 6.3.

During this stage of dissolution, the crystal surfaces are well annealed by the water, and dissolution takes place evenly over reasonably well defined “active” areas of the crystal surface. Roughly, the active areas are spherical surfaces for the cube and sphere, the cylindrical bases for the rod, and the cylindrical side for the tablet. For these active surfaces that develop in the partially dissolved crystals, the probability of detachment appears essentially uniform such that dissolution follows a fixed rate law until the final rapid dissolution stage.

We next consider more closely the influence of crystal shape, and compare our dissolution curves with classical models of the type briefly discussed in Section 6.2.

Influence of the crystal shape. The crystals we consider differ not only in shape, but also a little in size. Therefore, in order to compare dissolution profiles in Fig. 6.5 we plot the fraction of the crystal dissolved, $f_{\text{cry}}(t) = (N_0 - N(t))/N_0$, where $N(t)$ is the number of ions in the crystal at time t , and N_0 is the number present initially. It is apparent that for the crystals we consider the tablet and rod dissolve substantially faster (~ 6 ions ns^{-1}) than the cube and sphere (~ 4 ions ns^{-1}). For the present examples, this is likely explained by the fact that ions on the active surfaces of the tablet and rod interact more weakly with the bulk crystal than those on the surface of the water-annealed cube or sphere. Note that in their

smallest dimension the tablet and rod have only 8 and 6 crystal layers, respectively, whereas the cube has 14 layers initially. Thus the faster rates we observe for the tablet-shaped and rod-shaped crystals, while true for the present examples, are not only functions of shape, and likely will not apply as general rules.

Comparison with dissolution models. As discussed in Section 6.2, dissolution processes have been studied for well over a century, and a number of models have been put forward in the form of differential and integrated rate laws. Recent reviews of these rate laws and their history are given in refs. 147 and 149. Here we find that simple rate laws give a reasonably good description of our results in the fixed rate law regime. If we assume that sink conditions apply, and that detachment from the surface is the rate-determining step, we would expect the dissolution rate to be given by

$$\frac{dN(t)}{dt} = -kS_{\text{active}}(t) , \quad (6.2)$$

where $N(t)$ is the number of ions remaining in the crystal at time t , $S_{\text{active}}(t)$, is the “active surface area” where ion detachments occur, and k is a constant. This rate law that takes account of the changing surface area was first suggested by Brunner and Tolloczko,¹⁵² as an extension of the original Noyes-Whitney equation¹⁵¹ (*i.e.*, $dN(t)/dt = -k$) which applies to experimental situations where the surface area is kept fixed during dissolution. The full Brunner-Tolloczko and Noyes-Whitney equations also apply if sink conditions are not obeyed, but here we consider only the sink condition limit. The applicability of the sink condition to our simulations is confirmed below (Section 6.4).

We also remark that another model of dissolution proposed by Nernst¹⁵⁴ and

Brunner¹⁵⁵ assumes that solute diffusion from a surface layer where the concentration approaches the value at saturation, C_s , is the rate determining step in the dissolution process.¹⁴⁹ If we consider a spherical diffusion layer of surface area $S(t)$, set $S(t) = S_{\text{active}}(t)$, and $k = DC_s/\delta$, where D is a solute diffusion coefficient, and δ is the thickness of the layer, then Eq. (6.2) becomes the Nernst-Brunner equation. In principle, both D and δ could be time dependent, but if these quantities are regarded as constant, then at least for spherical crystals, Eq. (6.2) and the Nernst-Brunner equation are formally identical, even if the underlying physical mechanisms and rate constants are different. Below, we show explicitly that the diffusion layer model does not apply to the dissolution of NaCl.

Integrated rate laws can be obtained by specifying particular forms for $S_{\text{active}}(t)$. Above, we argued that in the fixed rate law regime, the active surface areas are approximately spherical for the sphere and the cube, the area of a nearly cylindrical wall for the tablet, and mainly the area of the circular base for the rod which anneals into a roughly cylindrical shape. Assuming these active surface areas, and using appropriate area to volume relationships, it is easy to deduce that for the sphere and the cube $S_{\text{active}}(t) \propto N^{2/3}(t)$, for the tablet $S_{\text{active}}(t) \propto N^{1/2}(t)$, and for the rod S_{active} is a constant independent of t . From Eq. (6.2), we then obtain the integrated rate laws

$$\sqrt[3]{N(t)} = \sqrt[3]{N_0} - kt \quad (\text{Sphere, Cube}) , \quad (6.3)$$

$$\sqrt{N(t)} = \sqrt{N_0} - kt \quad (\text{Tablet}) , \quad (6.4)$$

$$N(t) = N_0 - kt \quad (\text{Rod}) , \quad (6.5)$$

where of course k represents a different constant in each rate law. The cube root law applicable to spherical crystals is just the well-known Hixson-Crowell result¹⁵³ originally obtained by integrating the Brunner-Tolloczko equation (or, equivalently, the Nernst-Brunner equation with DC_s/δ held fixed) assuming a uniform spherical surface that decreases with time in proportion to $w^{2/3}$, where w is the weight of the crystal. We note that a square root law has been obtained empirically by Niebergall and Goyan,¹⁵⁶ but this is unlikely related to the tablet expression obtained here. Diffusion controlled assumptions have also been used to obtain two-thirds^{157,167} and square root¹⁶⁷ laws.

We tested all suggested rate laws by attempting linear least squares fits to the dissolution profiles. The initial and final portions of the dissolution profiles were removed from the fit as they constitute a deviation from the laws. The goodness of fit parameters R^2 are given in Table 6.3. Results for a two-thirds root law are also included since this model is sometimes used in analysis of experimental data. From Table 6.3 we see that even though for some crystal shapes (sphere and tablet) the best fits correspond to the rate laws derived above, in general the differences amongst the goodness of fit parameters are not large enough to draw any firm conclusions. Essentially, different power laws can fit a particular dissolution profile nearly equally well, and profiles on much longer time scales would be necessary to make meaningful distinctions.

We can also easily check to see if a diffusion layer can reasonably explain the observed dissolution rates. By calculating the $\Delta N/\Delta t$ in for the spherical crystal in the fixed rate regime (at ≈ 200 ns), we estimate the dissolution rate $dN/dt \approx -2.36 \times 10^{-3}$ ions ps⁻¹. If a diffusion layer is rate determining, then $dN/dt = -D * S * C_s/\delta$, and for our model $C_s \approx 3.29$ ions nm⁻³, S at ≈ 200 ns (assuming

R ² values for the rate laws		
	SPHERE	CUBE
linear	0.98873	0.99846
cube root	0.99738	0.99291
square root	0.99598	0.99635
two thirds root	0.99407	0.99838
	ROD	TABLET
linear	0.98962	0.98798
cube root	0.99775	0.99805
square root	0.99808	0.99885
two thirds root	0.99679	0.99733

Table 6.3: Goodness of fit parameters R² obtained for different rate laws. The best fits are indicated in bold.

spherical geometry) is 58.08 nm², and at saturation we obtain $D_+ \approx 0.36 \times 10^{-3}$ and $D_- \approx 0.42 \times 10^{-3}$ nm² ps⁻¹. This means that the thickness of the diffusion layer δ would need to be ~ 29 nm in order to explain the observed dissolution rate. This is obviously much too thick to be physically relevant, since it greatly exceeds the size of our crystal (radius ~ 2 nm) and system. The ion density as a function of distance from the crystal center at ~ 300 ns is shown in Fig. 6.9, and it is obvious that the ion density in solution is nearly uniform and considerably less than that implied by C_s , except possibly for a very narrow region near the crystal surface. Therefore, the diffusion layer model clearly does not apply to NaCl dissolution.

Rapid Dissolution

Dissolution profiles (Fig. 6.3) show that in the final stage of crystal dissolution there is a sharp increase in the rate at which ions are lost to solution. The crystal becomes very unstable and disappears rapidly once it has been sufficiently reduced in size, in all likelihood due to the decreased lattice energy of the small crys-

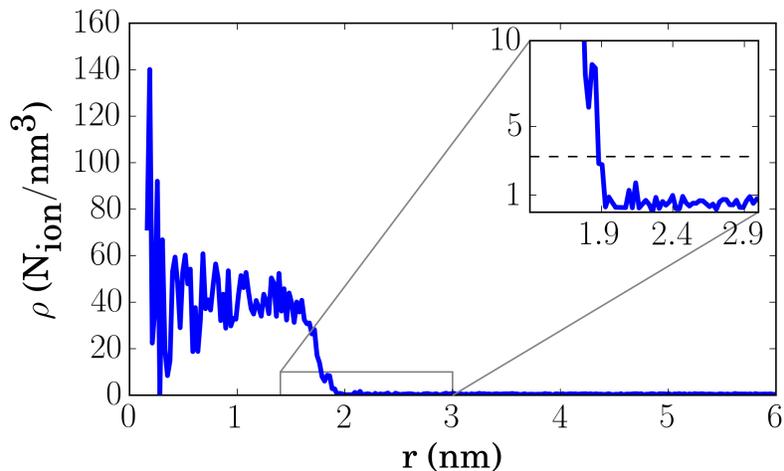


Figure 6.9: Radial ion density profile starting from the center of a spherical crystal at 300 ns. The dashed horizontal line in the inset indicates the density at saturation. The calculated radius of the crystal is ~ 2 nm, and a concentration gradient (that could be related to a diffusion layer) is not observed.

tal. The dissolution profiles indicate that the onset of increased instability occurs when the crystal contains ~ 200 ions.

A closer view of the final stage of dissolution of the spherical crystal is provided by the snapshots shown in Fig. 6.10. We see that the crystal holds its structure down to ~ 64 ions, then deforms and disintegrates completely when only ~ 12 ions remain. As one would expect, the onset of the final stage marked by increasing instability is independent of the initial crystal size. For example, spherical crystals initially containing 532 and 240 ion pairs also enter an instability region at ~ 200 ion. A similar picture pertains for crystals of other shapes.

It is worth mentioning a possible connection to the opposite process of crystal nucleation from supersaturated solution. For model NaCl solution, the critical nucleus just above saturation is estimated to contain ~ 75 ions,⁵⁰ which is reasonably close to the size where the residual NaCl crystal becomes very unstable in the dissolution process.

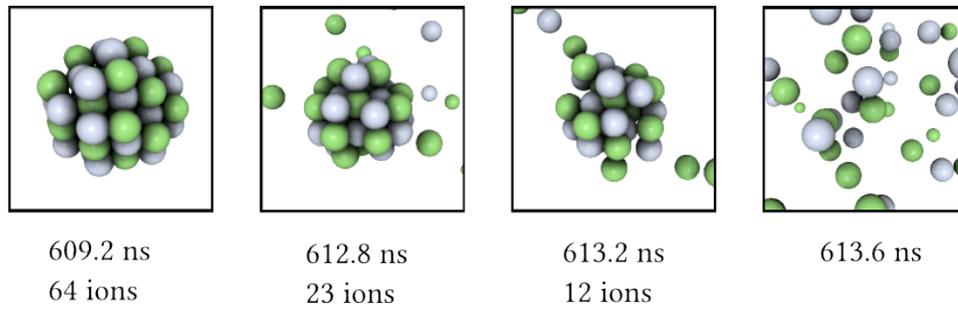


Figure 6.10: Snapshots of the last stage of dissolution of the spherical crystal. When the crystal reaches ~ 64 ions, the structure begins to disintegrate. In the final snapshot, water quickly penetrates the structure and breaks apart the remaining crystal nucleus.

6.4.2 Concentration Effects

In order to confirm that the dissolution profiles discussed above correspond to sink conditions and are not strongly influenced by increasing salt concentration, dissolution simulations of the spherical crystal were carried out at three salt mole fractions (calculated at complete dissolution), 0.0254, 0.0163, and 0.0127, keeping all other conditions fixed. The dissolution profiles are shown in Fig. 6.11, and we note that the results are very similar for 0.0163, and 0.0127, indicating that any salt concentration effects are small, and we are indeed in the sink regime in these systems. However, at the highest concentration considered, 0.0254, the dissolution rate clearly slows down after the initial ~ 120 ns. There are several possible explanations for the observed concentration effect at the highest salt mole fraction, and it is interesting to examine some possibilities in more detail.

One possible reason is ion reattachments. When the solution surrounding the crystal reaches a certain concentration, one might expect some ions to reattach to the crystal, hence slowing down the net dissolution rate. This possibility can be eliminated by counting the number of detachments and attachments over time.

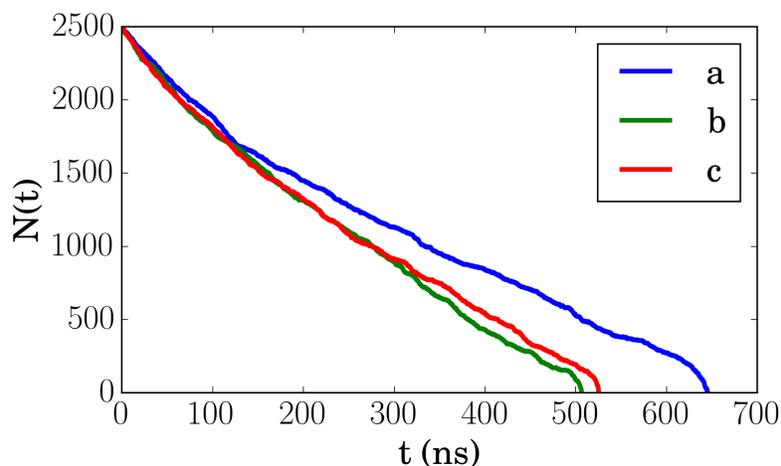


Figure 6.11: Concentration effects on rate. Curves (a), (b), and (c) are for NaCl mole fractions (calculated assuming complete dissolution of the entire crystal) of 0.0254, 0.0163, and 0.0127. Note that curve (a) shows a sharp decrease in rate at ~ 120 ns, while there is no substantial difference between the profiles (b) and (c).

Figure 6.12 shows that the reattachments are approximately constant at a rate of 2 events/ns, while detachments show a sharp rate change at about 120 ns, corresponding to the change in the dissolution profile noted above (Fig. 6.11). Since there is no substantial variation in the attachment rate, reattachments are not responsible for the change in the dissolution rate. We further remark that since the reattachment rate oscillates but does not increase with time (concentration), most of the reattachments counted are likely due to ions at the crystal-solution interface moving back and forth across the artificially sharp boundary introduced by our order parameter. Another possibility is that the ions in solution change the properties of the solvent, for example, by effectively “sequestering” a significant fraction of the water molecules, hence reducing the dissolution rate. To isolate the effect of the ions in solution on the dissolution rate, we took the system at 120 ns (before the crossover) and simply removed the dissolved ions from the solution, while the partially dissolved crystal was left in the system. This procedure completely removed

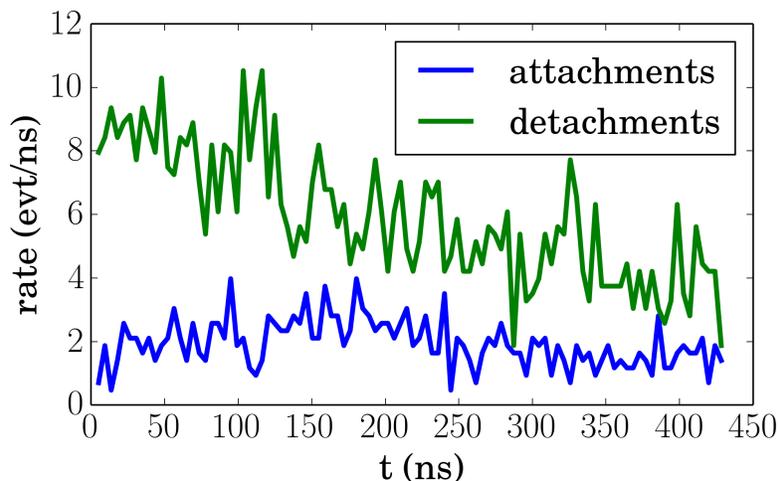


Figure 6.12: Detachment and reattachment events for the spherical crystal. During the course of dissolution, the detachment rate systematically slows while the attachment rate remains substantially constant.

the crossover in the dissolution rate, confirming that the drop in rate is caused by the presence of ions in solution, and not by any change in the crystal.

6.4.3 Temperature Dependence and Activation Energy

We would expect ion detachment to be an activated process, and in order to estimate the activation energy dissolution simulations of spherical crystals (~ 1256 ion pairs) were carried out at three temperatures, 300, 320, and 340 K. Rate constants are found by fitting the cube root law [Eq. (6.2)] in the fixed rate law regime (*i.e.*, eliminating the initial and final regions of the dissolution profiles). Excellent linear fits are obtained as shown in Fig. 6.13. The rate constants obtained follow the Arrhenius equation (Fig. 6.14)

$$\ln(k) = \ln(A) - E_a/RT, \quad (6.6)$$

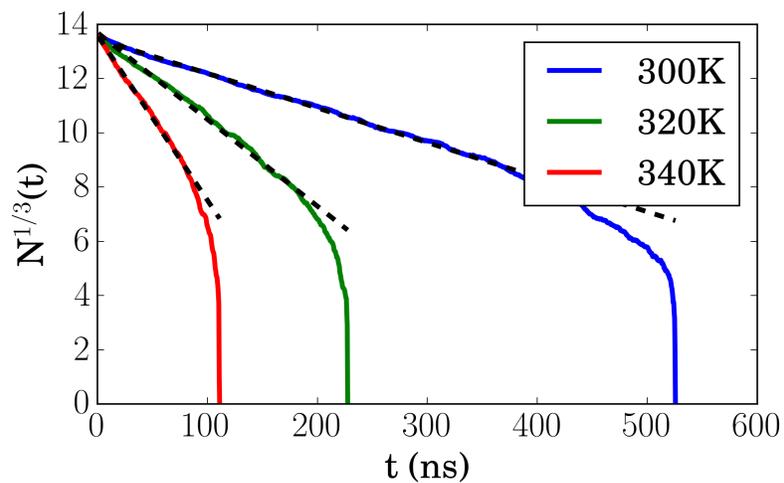


Figure 6.13: Fits to the cube root law for the spherical crystal at temperatures of 300, 320 and 340 K.

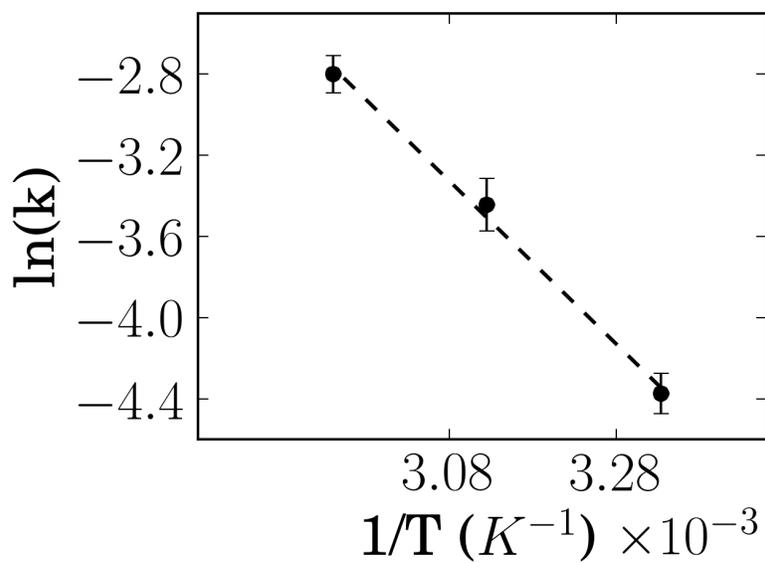


Figure 6.14: Fits to the Arrhenius equation of the rate constants obtained for the spherical crystal at 300, 320, and 340 K. Error bars represent one standard deviation of uncertainty.

giving an activation energy of 33.46 ± 1.46 kJ mol⁻¹ (SD).

The standard deviation in the activation energy quoted above and those included as error bars in Figure 6.14) were estimated as follows. The trajectories obtained for the three temperatures considered were divided into four equal parts (time slices), and rate constants for each part were obtained by fitting to the cube root law (eq. 6.2). The four different estimates were then used to obtain the error bars shown in Figure 6.14. Similarly, by fitting the four different sets of rate constants to the Arrhenius equation, four estimates of the activation energy were obtained and used to estimate the standard deviation in the activation energy.

6.5 Summary and Conclusions

Molecular dynamics simulations were employed to examine in detail the dissolution of NaCl nanocrystals of different shape. Specifically, cubic, spherical, tablet-like and rod-like crystals were considered with the dissolution carried out under so-called sink conditions, where the solution is always very dilute compared to saturation. In all cases, dissolution was found to occur in three distinct stages. Initially, the more exposed and/or most weakly bound ions are quickly detached from the crystal. The initial crystal shape greatly affects the first stage as it mainly involves ions at edges and corners. After the rapid initial period, the crystals take on solvent-annealed shapes (cubic crystals become nearly spherical and rods and tablets roughly cylindrical) that persist until the final stage of dissolution. During the long intermediate stage the dissolution appears to closely follow a fixed rate law. In the final stage ($\lesssim 200$ ions) the crystal becomes very unstable and dissolves at an ever increasing rate until it disappears.

In the intermediate fixed rate law regime, the dissolution process is well described by assuming that the rate is proportional to an active surface area from which the ions are preferentially detached. For the solvent-annealed cubic and spherical crystals the active area is the entire spherical surface, and we obtain the classical cube root law often employed in dissolution studies of macroscopic crystals, assuming a spherical shape. We show that the cube root law also provides a good description of the nanoscopic crystals considered here in the intermediate stage. We also show that ion detachment from the surface, and not the existence of a Nernst-Brunner diffusion layer,^{149,154,155} is the rate determining step in the dissolution process for NaCl. For the spherical crystal, simulations were done at three temperatures and the rate constants, determined by fitting to the cube root law, closely follow the Arrhenius equation, giving an activation energy of ~ 33.5 kJ mol⁻¹.

After the initial stage, both the tablet-shaped and rod-shaped crystals have acquired roughly cylindrical shapes, and we observed that ions did not leave uniformly over the entire surface area for these geometries. Rather, for the annealed tablet and rod we identified the active surface areas to be the cylinder walls, and the cylinder base, respectively, giving a square root law for the tablet and a linear law for the rod. These rate laws give good fits to the dissolution profiles, but the total dissolution times for the nanocrystals considered here are not long enough to make clear, unambiguous distinctions amongst the different rate laws. Nevertheless, our simulations show that, apart from the initial and final stages of dissolution, models of the classical type give a good description of the dissolution of NaCl nanocrystals.

Chapter 7

Summary and Conclusions

7.1 Summary

The microscopic mechanism of nucleation is still largely unknown because the high temporal and spatial resolution needed to observe the process is difficult to reach using current experimental techniques.³² A considerable body of research has been conducted using molecular simulation that, thanks to the availability of parallel computational power, can be used to model nucleation for a range of compounds of scientific and industrial interest.³² In this thesis, we investigated different aspects of nucleation using alkali halide salts as a model, and designed data analysis techniques able to better exploit the results produced by molecular dynamics simulations.

In Chapter 2, the key algorithms and methods of molecular dynamic simulations are introduced, as well as the models used to represent alkali halides in aqueous solutions.

Chapter 3 describes an investigation of nucleation of NaCl from aqueous solution. We found that NaCl nucleation is observable within simulation timescales

(10 to 200 ns), at 300 K and 1 bar, starting from a solution at about twice the saturation concentration for the model. A major challenge to the investigation of the early stages of nucleation was the lack of a reliable method to detect small clusters in solution. To address this issue, we developed a novel three-step method to detect cluster formation, and follow cluster evolution over time. Thanks to the new methodology, it was possible to collect and monitor several properties of the large number of clusters that continuously form and disappear from solution.

By analyzing a variety of cluster properties we showed, for the first time by direct measurement, that the lifetime and probability of nucleation of a cluster depend not only on size, but also on the specific geometrical arrangement of the cluster. The importance of this result lies in its contrast with the common CNT assumption that the probability of nucleation depends only on cluster size. A way to extend CNT to include structural effects could be the addition of a structure dependent free energy in the exponential term of the nucleation rate. It is also useful to notice that factors such as geometrical arrangement are not merely surrogates of surface tension, as they do not scale with surface area.

In Chapter 4 we investigated crystallization from the melt of the TF and JC models for lithium halides, partially to assess their probable behavior in simulations of crystallization from aqueous solutions. We found that the TF model incorrectly predicts the wurtzite crystal structure as the stable structure for lithium salts, while the more recent JC model reproduces the correct (rock salt) stable crystal structure for LiF and LiCl, but not for LiBr and LiI. Furthermore, we found that the wurtzite crystal structure is highly irregular in finite size crystals, and that this can be likely attributed to surface rearrangements that are necessary to avoid the formation of a surface dipole moment.

In Chapter 5 we further explore crystal nucleation from solution focusing on lithium fluoride. For the JC model, which has the correct rock salt crystal structure, nucleation and growth can be observed only at high temperature and a high degree of supersaturation. By measuring the growth rate at different temperatures we found that growth requires a high activation energy, which is likely associated with the high barrier required to remove water molecules from the first solvation shell of an ion. While we couldn't measure an activation energy for nucleation, it is reasonable to assume that the activation energy for nucleation is equal to, or greater, than the activation energy for growth. We also found that the structure of the metastable solution is quite different from that observed for NaCl. In LiF solutions, the metastable ionic clusters tend to be much shorter lived and less regular in shape.

Based on our observations, in supersaturated LiF aqueous solutions, a crucial step to nucleation is the ability of an ion to lose at least part of its solvation shell and join a cluster. In this system, the applicability of CNT is in question because the mechanism seems to invalidate some of the foundational assumptions of CNT. For example, the LiF supersaturated solution is characterized by the presence of metastable ionic clusters, that, compared to those found in NaCl solutions, are small and do not resemble the bulk crystal. Interestingly, these metastable species increase in number at higher temperatures and, their formation likely constitutes a necessary step to nucleation.

In Chapter 6, the dissolution of NaCl was investigated to gain insight into the stability of finite size ionic crystals in water, and the dynamics of ion attachment and detachment. It was found that NaCl dissolution can be modeled as a three-stage process where the crystal initially loses ions around the edges, then follows an essentially fixed dissolution rate law until it reaches a certain size ($\lesssim 200$ ions),

where it becomes very unstable and quickly dissolves into solution. By measuring the dissolution rate of NaCl at different temperatures, it was found that the process is activated, with a barrier of $\sim 33 \text{ kJ mol}^{-1}$. Additionally, we found that it is possible to obtain rate laws from a simple expression involving the surface area of the crystal.

7.2 Future Directions

Our results suggest that, in general, the mechanism of homogeneous nucleation commonly assumed in CNT does not correctly model the species that precede the formation of a critical nucleus. In the case of NaCl nucleation, we found that size is not the only factor affecting nucleation, and geometric arrangement also influences the probability of nucleation, as well as the survival of prenucleation clusters. Despite the influence of geometric regularity, the process observed resembles the CNT process in that there is a distribution of clusters that stochastically form or disappear in solution, even though the clusters do not have the same properties as the bulk phase, since they are mostly made of interfacial ions. The presence of small, metastable clusters can be also seen as evidence of a multi-step nucleation process, where the prenucleation clusters are intermediate species that convert to the bulk phase as the cluster grows.

A mechanism that explains the data obtained from LiF simulations of Chapter 5 is a process with two kinetic barriers. The first, is a barrier to cluster formation. In order to form metastable clusters, it is necessary to surpass a barrier that appears to be related to the solvent-solute interactions. The second barrier, is a nucleation barrier that likely depends on cluster related properties (e.g. size, surface tension

and crystallinity). Characterization of nucleation rates at different temperatures (through simulation) could help probe the nature of the second barrier, and lead to comparison, validation, and extension of multi-step nucleation theories.^{144,168}

Further speculating, a compound where the barrier to generating prenucleation clusters is very weak could have a behavior similar to that exhibited by CaCO_3 ,²⁵ where there is aggregation into an amorphous phase, followed by rearrangement into a crystal structure. Investigation on more limiting cases would be useful in formalizing a theory that encompasses multiple nucleation mechanisms. In this respect, a kinetic formulation involving multiple steps such as diffusion, dehydration, and aggregation would possibly constitute a more realistic model for nucleation.

Model development for solute and solvent interactions is also of crucial importance. Models, in order to give realistic results in simulations of phase transitions, should reproduce both solution and solid properties such as solubilities and lattice energies. This is especially important for compounds that form hydrates (e.g. LiCl) where a model that at least supports the correct stable structure at the pressure and temperature of interest is necessary. Additionally, increasing attention should be devoted to the study of finite-size structures since, as we have shown throughout this thesis, crystallization nuclei can be very small and have different properties than the bulk crystal phase.

As we have shown, by using medium- and large-scale molecular dynamics simulations combined with advanced data analysis techniques, it is possible to directly observe individual nucleation events, degree of ion association, cluster lifetimes and other properties of interest. More experiments are definitely needed to validate the models and, while some work has been done for NaCl ,²⁹ a better coverage of simple

compounds would be effective in evaluating the results obtained through molecular simulation.

A combination of realistic models, experiments, and data analysis techniques will ultimately lead to a better understanding of crystal nucleation and a refinement of classical theories.

Bibliography

- [1] I. S. Joung and T. E. Cheatham, *J. Phys. Chem. B* **112**, 9020 (2008).
- [2] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma, *J. Phys. Chem.* **91**, 6269 (1987).
- [3] J. Angus, *Structure and Thermochemistry of Diamond*, The Physics of Diamond, OS Press, Amsterdam, 1997.
- [4] A. L. DeVries and D. E. Wohlschlag, *Science* **163**, 1073 (1969).
- [5] F. Franks, *Biophysics and biochemistry at low temperatures*, Cambridge University Press, 1985.
- [6] M. K. Yau and R. Rogers, *A short course in cloud physics*, Elsevier, 1996.
- [7] M. Kumai, *J. Meteor.* **18**, 139 (1961).
- [8] M. Kumai and K. E. Francis, *J. Atmos. Sci.* **19**, 474 (1962).
- [9] P. G. Debenedetti, *Metastable Liquids: Concepts and Principles*, Princeton University Press, 1996.
- [10] M. Volmer and A. Weber, *Z. Phys. Chem.* **119**, 277 (1926).
- [11] L. Farkas, *Z. Phys. Chem.* **125** (1927).
- [12] R. Becker and W. Döring, *Ann. Phys. (Berlin)* **416**, 719 (1935).

- [13] J. B. Zeldovich, *Acta physicochim. URSS* **18**, 1 (1943).
- [14] J. Frenkel, *Kinetic Theory of Liquids*, Dover Publications, Dover, 1955.
- [15] J. L. Schmitt, G. W. Adams, and R. A. Zalabsky, *J. Chem. Phys.* **77**, 2089 (1982).
- [16] A. Kacker and R. H. Heist, *J. Chem. Phys.* **82**, 2734 (1985).
- [17] R. Strey, T. Schmeling, and P. E. Wagner, *J. Chem. Phys.* **85**, 6192 (1986).
- [18] C. Hung, M. J. Krasnopoler, and J. L. Katz, *J. Chem. Phys.* **90**, 1856 (1989).
- [19] J. Wlk, , and R. Strey*, *J. Phys. Chem. B* **105**, 11683 (2001).
- [20] S. Auer and D. Frenkel, *Nature* **409**, 1020 (2001).
- [21] J. L. Harland and W. van Meegen, *Phys. Rev. E* **55**, 3054 (1997).
- [22] A. Heymann, A. Stipp, C. Sinn, and T. Palberg, *J. Colloid Interface Sci.* **207**, 119 (1998).
- [23] P. Wette and H. J. Schöpe, *Phys. Rev. E* **75**, 051405 (2007).
- [24] J. D. Rodriguez-Blanco, S. Shaw, and L. G. Benning, *Nanoscale* **3**, 265 (2011).
- [25] D. Gebauer, A. Völkel, and H. Cölfen, *Science* **322**, 1819 (2008).
- [26] D. Gebauer, M. Kellermeier, J. D. Gale, L. Bergström, and H. Cölfen, *Chem. Soc. Rev.* **43**, 2348 (2014).
- [27] W. Ostwald, *Z. Phys. Chem.* **22**, 289 (1897).
- [28] M. Sleutel, J. Lutsko, A. E. Van Driessche, M. A. Durán-Olivencia, and D. Maes, *Nat. Commun.* **5** (2014).

- [29] J. Desarnaud, H. Derluyn, J. Carmeliet, D. Bonn, and N. Shahidzadeh, *J. Phys. Chem. Lett.* **5**, 890 (2014).
- [30] H.-S. Na, S. Arnold, and A. S. Myerson, *J. Cryst. Growth* **139**, 104 (1994).
- [31] R. J. Davey, S. L. M. Schroeder, and J. H. terHorst, *Angew. Chem. Int. Ed.* **52**, 2166 (2013).
- [32] G. C. Sosso, J. Chen, S. J. Cox, M. Fitzner, P. Pedevilla, A. Zen, and A. Michaelides, *Chem. Rev.* (2016).
- [33] J.-P. Hansen and L. Verlet, *Phys. Rev.* **184**, 151 (1969).
- [34] J. Q. Broughton, G. H. Gilmer, and K. A. Jackson, *Phys. Rev. Lett.* **49**, 1496 (1982).
- [35] R. O. Rosenberg and D. Thirumalai, *Phys. Rev. A* **36**, 5690 (1987).
- [36] D. S. Corti and P. G. Debenedetti, *Chem. Eng. Sci.* **49**, 2717 (1994).
- [37] F. de Wette, R. Allen, D. Hughes, and A. Rahman, *Phys. Lett. A* **29**, 548 (1969).
- [38] M. J. Mandell, J. P. McTague, and A. Rahman, *J. Chem. Phys.* **64**, 3699 (1976).
- [39] P. Rein ten Wolde, M. J. RuizMontero, and D. Frenkel, *J. Chem. Phys.* **104**, 9932 (1996).
- [40] P. R. ten Wolde, M. J. Ruiz-Montero, and D. Frenkel, *J. Chem. Phys.* **110**, 1591 (1999).
- [41] B. O'Malley and I. Snook, *Phys. Rev. Lett.* **90**, 085702 (2003).

- [42] E. B. Moore and V. Molinero, *Phys. Chem. Chem. Phys.* **13**, 20008 (2011).
- [43] J. Anwar and D. Zahn, *Angew. Chem. Int. Ed.* **50**, 1996 (2011).
- [44] M. Salvalaglio, C. Perego, F. Giberti, M. Mazzotti, and M. Parrinello, *Proc. Natl. Acad. Sci.* **112**, E6 (2015).
- [45] F. Jimenez-Angeles and A. Firoozabadi, *J. Phys. Chem. C* **118**, 11310 (2014).
- [46] H. Ohtaki and N. Fukushima, *Pure Appl. Chem.* **63** (1991).
- [47] D. Zahn, *Phys. Rev. Lett.* **92**, 040801 (2004).
- [48] F. Giberti, G. A. Tribello, and M. Parrinello, *J. Chem. Theory Comput.* **9**, 2526 (2013).
- [49] D. Chakraborty and G. N. Patey, *Chem. Phys. Lett.* **587**, 25 (2013).
- [50] D. Chakraborty and G. N. Patey, *J. Phys. Chem. Lett.* **4**, 573 (2013).
- [51] J. Yan and G. N. Patey, *J. Phys. Chem. Lett.* **2**, 2555 (2011).
- [52] J. Yan and G. N. Patey, *J. Phys. Chem. A* **116**, 7057 (2012).
- [53] J. Yan, S. Overduin, and G. N. Patey, *J. Chem. Phys.* **141**, 074501 (2014).
- [54] S. A. Zielke, A. K. Bertram, and G. N. Patey, *J. Phys. Chem. B* **119**, 9049 (2015).
- [55] S. A. Zielke, A. K. Bertram, and G. N. Patey, *J. Phys. Chem. B* **120**, 1726 (2015).
- [56] S. A. Zielke, A. K. Bertram, and G. N. Patey, *J. Phys. Chem. B* **120**, 2291 (2016).

- [57] D. Murphy, *Geophys. Res. Lett.* **30** (2003).
- [58] S. A. Hassan, *Phys. Rev. E* **77**, 031501 (2008).
- [59] S. A. Hassan, *J. Chem. Phys.* **134**, 114508 (2011).
- [60] A. Benavides, J. Aragonés, and C. Vega, *J. Chem. Phys.* **144**, 124504 (2016).
- [61] I. S. Joung and T. E. Cheatham, *J. Phys. Chem. B* **113**, 13279 (2009).
- [62] G. A. Orozco, O. A. Moulton, H. Jiang, I. G. Economou, and A. Z. Panagiotopoulos, *J. Chem. Phys.* **141**, 234507 (2014).
- [63] J. L. Aragonés, M. Rovere, C. Vega, and P. Gallo, *J. Phys. Chem. B* **118**, 7680 (2014).
- [64] F. Moučka, I. Nezbeda, and W. R. Smith, *J. Chem. Phys.* **138**, 154102 (2013).
- [65] J. Aragonés, E. Sanz, C. Valeriani, and C. Vega, *J. Chem. Phys.* **137**, 104507 (2012).
- [66] N. E. Zimmermann, B. Vorselaars, D. Quigley, and B. Peters, *J. Am. Chem. Soc.* **137**, 13352 (2015).
- [67] T. A. Halgren, *J. Comput. Chem.* **17**, 490 (1996).
- [68] P.-O. Åstrand, P. Linse, and G. Karlström, *Chem. Phys.* **191**, 195 (1995).
- [69] H. Sun, S. J. Mumby, J. R. Maple, and A. T. Hagler, *J. Am. Chem. Soc.* **116**, 2978 (1994).
- [70] W. Eckhardt, A. Heinecke, R. Bader, M. Brehm, N. Hammer, H. Huber, H.-G. Kleinhenz, J. Vrabc, H. Hasse, M. Horsch, M. Bernreuther, C. W. Glass,

- C. Niethammer, A. Bode, and H.-J. Bungartz, *591 TFLOPS Multi-trillion Particles Simulation on SuperMUC*, pages 1–12, Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [71] S. W. de Leeuw, J. W. Perram, and E. R. Smith, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **373**, 27 (1980).
- [72] P. P. Ewald, Ann. Phys. (Berlin) **369**, 253 (1921).
- [73] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*, Clarendon Press: Oxford, United Kingdom, New York, NY, USA, 1989.
- [74] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).
- [75] J. W. Cooley and J. W. Tukey, Math. Comp. **19**, 297 (1965).
- [76] S. R. Hall, F. H. Allen, and I. D. Brown, Acta Crystallogr. Sect. A **47**, 655 (1991).
- [77] B. Hess et al., J. Comput. Chem. **18**, 1463 (1997).
- [78] H. J. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. Di Nola, and J. Haak, J. Chem. Phys. **81**, 3684 (1984).
- [79] G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).
- [80] M. Parrinello and A. Rahman, J Appl Phys **52**, 7182 (1981).
- [81] H. Blatt, R. Tracy, and B. Owens, *Petrology: Igneous, Sedimentary, and Metamorphic*, Macmillan, 2006.

- [82] K. L. Webber, A. U. Falster, W. B. Simmons, and E. E. Foord, *J. Petrol.* **38**, 1777 (1997).
- [83] P. Arosio, M. Vendruscolo, C. M. Dobson, and T. P. Knowles, *Trends Pharmacol. Sci.* **35**, 127 (2014).
- [84] B. E. Rabinow, *Nat. Rev. Drug Discovery* **3**, 785 (2004).
- [85] A. Laaksonen, V. Talanquer, and D. W. Oxtoby, *Annu. Rev. Phys. Chem.* **46**, 489 (1995).
- [86] M. Kulmala, T. Petäjä, M. Ehn, J. Thornton, M. Sipilä, D. Worsnop, and V.-M. Kerminen, *Annu. Rev. Phys. Chem.* **65**, 21 (2014).
- [87] D. Aastuen, N. Clark, L. Cotter, and B. J. Ackerson, *Phys. Rev. Lett.* **57**, 1733 (1986).
- [88] Z. Wang, F. Wang, Y. Peng, and Y. Han, *Nat. Commun.* **6** (2015).
- [89] N. Pienack and W. Bensch, *Angew. Chem. Int. Ed.* **50**, 2014 (2011).
- [90] M. Matsumoto, S. Saito, and I. Ohmine, *Nature* **416**, 409 (2002).
- [91] M. Kellermeier, A. Picker, A. Kempter, H. Cölfen, and D. Gebauer, *Adv. Mater.* **26**, 752 (2014).
- [92] J. Merikanto, E. Zapadinsky, A. Lauri, and H. Vehkamäki, *Phys. Rev. Lett.* **98**, 145702 (2007).
- [93] P. G. Vekilov, *Cryst. Growth Des.* **4**, 671 (2004).
- [94] V. J. Anderson and H. N. Lekkerkerker, *Nature* **416**, 811 (2002).

- [95] Z. Mester and A. Z. Panagiotopoulos, *J. Chem. Phys.* **143**, 044505 (2015).
- [96] H. M. Manzanilla-Granados, H. Saint-Martín, R. Fuentes-Azcatl, and J. Alexandre, *J. Phys. Chem. B* **119**, 8389 (2015).
- [97] K. Kobayashi, Y. Liang, T. Sakka, and T. Matsuoka, *J. Chem. Phys.* **140**, 144705 (2014).
- [98] J. Aragoes, E. Sanz, and C. Vega, *J. Chem. Phys.* **136**, 244508 (2012).
- [99] Merck index 14th ed. whitehouse station, nj: Merck and co, 2006.
- [100] B. Hess, C. Kutzner, D. Van der Spoel, and E. Lindahl, *J. Chem. Theory Comp.* **4**, 435 (2008).
- [101] G. Lanaro and G. N. Patey, *J. Phys. Chem. B* **119**, 4275 (2015).
- [102] J. R. Espinosa, C. Vega, C. Valeriani, and E. Sanz, *J. Chem. Phys.* **142**, 194709 (2015).
- [103] W. Lechner and C. Dellago, *J. Chem. Phys.* **129**, 114707 (2008).
- [104] R. Radhakrishnan and B. L. Trout, *J. Am. Chem. Soc.* **125**, 7743 (2003).
- [105] A. Reinhardt, J. P. Doye, E. G. Noya, and C. Vega, *J. Chem. Phys.* **137**, 194504 (2012).
- [106] T. Li, D. Donadio, G. Russo, and G. Galli, *Phys. Chem. Chem. Phys.* **13**, 19807 (2011).
- [107] P. J. Steinhardt, D. R. Nelson, and M. Ronchetti, *Phys. Rev. B* **28**, 784 (1983).

- [108] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise., in *Kdd*, volume 96, pages 226–231, 1996.
- [109] F. Pedregosa et al., *J. Mach. Learn. Res.* **12**, 2825 (2011).
- [110] P. Jaccard, *New Phytol* **11**, 37 (1912).
- [111] E. L. Kaplan and P. Meier, *J. Am. Stat. Assoc.* **53**, 457 (1958).
- [112] C. P. O’Hagan, B. J. O’Brien, F. Griffin, B. Hooper, S. B. Leen, and R. F. D. Monaghan, *Energy & Fuels* **29**, 3082 (2015).
- [113] A. R. Kamali and D. J. Fray, *Carbon* **77**, 835 (2014).
- [114] J. R. Hardy and A. M. Karo, *The Lattice Dynamics and Static of Alkali Halide Crystals*, Springer Science + Business Media, 1979.
- [115] T. Martin, *Phys. Rep* **95**, 167 (1983).
- [116] A. Bach, D. Fischer, and M. Jansen, *Z. Anorg. Allg. Chem.* (2009).
- [117] Y. Liebold-Ribeiro, D. Fischer, and M. Jansen, *Angew. Chem.* **120**, 4500 (2008).
- [118] D. Fischer, A. Mueller, and M. Jansen, *ChemInform* **36** (2005).
- [119] P. C. Rodrigues and F. M. S. Fernandes, *EPJ D* **44**, 109 (2007).
- [120] P. C. Rodrigues and F. M. S. Fernandes, *J Mol Struc-THEOCHEM* **946**, 94 (2010).
- [121] T. Croteau and G. N. Patey, *J. Chem. Phys.* **124**, 244506 (2006).

- [122] Ž. P. Čančarević, J. C. Schoen, and M. Jansen, *Chem Asian J* **3**, 561 (2008).
- [123] A. Aguado, A. Ayuela, J. M. López, and J. A. Alonso, *Phys. Rev. B* **56**, 15353 (1997).
- [124] L. Pauling, *The Nature of the Chemical Bond and the Structure of Molecules and Crystals: an Introduction to Modern Structural Chemistry*, volume 18, Cornell university press, 1960.
- [125] M. Tosi and F. Fumi, *J. Phys. Chem. Solids* **25**, 45 (1964).
- [126] J. R. Espinosa, C. Vega, C. Valeriani, and E. Sanz, *J. Chem. Phys.* **144**, 034501 (2016).
- [127] J. Wang, J. Wu, Z. Sun, G. Lu, and J. Yu, *J. Mol. Liq.* **209**, 498 (2015).
- [128] C. Valeriani, E. Sanz, and D. Frenkel, *J. Chem. Phys.* **122**, 194501 (2005).
- [129] T. Zykova-Timan, D. Ceresoli, U. Tartaglino, and E. Tosatti, *J. Chem. Phys.* **123**, 164701 (2005).
- [130] A. B. Belonoshko, R. Ahuja, and B. Johansson, *Phys. Rev. B* **61**, 11928 (2000).
- [131] F. Fumi and M. Tosi, *J. Phys. Chem. Solids* **25**, 31 (1964).
- [132] M. Born and J. E. Mayer, *Z. Phys.* **75**, 1 (1932).
- [133] M. L. Huggins and J. E. Mayer, *J. Chem. Phys.* **1**, 643 (1933).
- [134] J. E. Mayer, *J. Chem. Phys.* **1**, 270 (1933).
- [135] G. Lanaro, chemlab 0.4, <https://chemlab.readthedocs.io>, 2015.

- [136] Y. Sakamoto, J. Chem. Phys. **28**, 164 (1958).
- [137] M. Parrinello and A. Rahman, Phys. Rev. Lett. **45**, 1196 (1980).
- [138] D. B. Sirdeshmukh, L. Sirdeshmukh, and K. Subhadra, *Alkali Halides: A Handbook of Physical Properties*, volume 49, Springer Science & Business Media, 2013.
- [139] D. E. Smith and L. X. Dang, J. Chem. Phys. **100**, 3757 (1994).
- [140] G. Lanaro and G. N. Patey, J. Phys. Chem. B **120**, 9076 (2016).
- [141] L. Le and V. Molinero, J. Phys. Chem. A **115**, 5900 (2010).
- [142] C. B. Stubblefield and R. O. Bach, J. Chem. Eng. Data **17**, 491 (1972).
- [143] W. A. Hart, O. Beumel, and T. P. Whaley, *The Chemistry of Lithium, Sodium, Potassium, Rubidium, Cesium and Francium: Pergamon Texts in Inorganic Chemistry*, volume 7, Elsevier, 2013.
- [144] D. Erdemir, A. Y. Lee, and A. S. Myerson, Acc. Chem. Res. **42**, 621 (2009), PMID: 19402623.
- [145] J. B. Dressman, G. L. Amidon, C. Reppas, and V. P. Shah, Pharm. Res. **15**, 11 (1998).
- [146] O. Anand, X. Y. Lawrence, D. P. Conner, and B. M. Davit, AAPS J. **13**, 328 (2011).
- [147] A. Dokoumetzidis and P. Macheras, Int. J. Pharm. **321**, 1 (2006).
- [148] S. S. Jambhekar and P. J. Breen, Drug Disc. Today **18**, 1173 (2013).

- [149] J. Siepmann and F. Siepmann, *Int. J. Pharm.* (2013).
- [150] P. Costa and J. M. Sousa Lobo, *Eur. J. Pharm. Sci.* **13**, 123 (2001).
- [151] A. A. Noyes and W. R. Whitney, *J. Am. Chem. Soc.* **19**, 930 (1897).
- [152] L. Brunner and S. Tolloczko, *Z. Phys. Chem* , 283 (1900).
- [153] A. Hixson and J. Crowell, *Industrial & Engineering Chemistry* **23**, 923 (1931).
- [154] W. Nernst, *Z. Phys. Chem* **47**, 52 (1904).
- [155] E. Brunner, *Z. Phys. Chem* **43**, 56 (1904).
- [156] P. Niebergall, G. Milosovich, and J. Goyan, *J. Pharm. Sci.* **52**, 236 (1963).
- [157] W. Higuchi and E. Hiestand, *J. Pharm. Sci.* **52**, 67 (1963).
- [158] S. Piana and J. D. Gale, *J. Am. Chem. Soc.* **127**, 1975 (2005).
- [159] Y. Gao and K. W. Olsen, *Mol. Pharm.* **10**, 905 (2013).
- [160] N. H. De Leeuw, S. C. Parker, and J. H. Harding, *Phys. Rev. B* **60**, 13792 (1999).
- [161] H. Shinto, T. Sakakibara, and K. Higashitani, *J. Phys. Chem. B* **102**, 1974 (1998).
- [162] R. Bahadur, L. M. Russell, S. Alavi, S. T. Martin, and P. R. Buseck, *J. Chem. Phys.* **124**, 154713 (2006).
- [163] O. Engkvist and A. J. Stone, *J. Chem. Phys.* **112**, 6827 (2000).
- [164] J. Klime, D. R. Bowler, and A. Michaelides, *J. Chem. Phys.* **139**, (2013).

- [165] N. Holmberg, J. C. Chen, A. S. Foster, and K. Laasonen, *Phys. Chem. Chem. Phys.* (2014).
- [166] C. M. Riley and T. W. Rosanske, *Development and Validation of Analytical Methods*, volume 3, Elsevier, 1996.
- [167] J. Wang and D. R. Flanagan, *J. Pharm. Sci.* **88**, 731 (1999).
- [168] W. Pan, A. B. Kolomeisky, and P. G. Vekilov, *J. Chem. Phys.* **122**, 174905 (2005).
- [169] T. J. Richmond, *J. Mol. Biol.* **178**, 63 (1984).
- [170] W. E. Lorensen and H. E. Cline, Marching cubes: A high resolution 3d surface construction algorithm, in *ACM siggraph computer graphics*, volume 21, pages 163–169, ACM, 1987.
- [171] D. Xu and Y. Zhang, *PloS one* **4**, e8140 (2009).

Appendix A

Distribution of Various Properties for Nucleated and Failed Clusters.

In Figures A.1 and A.2, distributions of various properties for nucleated and failed clusters of size 10 and 30 are reported. A description of such properties is given below:

- **Average neighbor count** is calculated by counting for each ion of the cluster the number of neighbors within 0.6 nm. The counts are then averaged for every ion in the cluster.
- **Volume** is approximated by encasing a cluster in a cuboid box, subdivided into $32 \times 32 \times 32$ grid points, and counting the number of grid points that are inside of the cluster, the volume is obtained by multiplying this number by the volume of the cell (Figure A.3). The radius that each ion covers is chosen to be the ionic radius plus the solvent radius, according to the definition of excluded volume.¹⁶⁹
- **Surface area** is approximated by extracting the surface from the volumetric representation described above using the marching cube algorithm,¹⁷⁰ the procedure is inspired by the work of Xu et al.¹⁷¹

- **Sphericity** is the ratio of the surface area of a sphere with equivalent volume and the surface area of the cluster, and is defined as

$$\pi^{1/3} \frac{(6V)^{2/3}}{A},$$

where V is volume and A is surface area. Its value is 1 when the object is perfectly spherical and assumes lower values for less spherical objects. It is often used to represent the compactness of a three dimensional object.

- **Hydration** is a measure of water content in the cluster and is calculated by counting for each ion the number of water molecules within 0.3 nm, the counts are then averaged over all the ions in the clusters.
- **Radius of gyration** is defined as

$$R_g^2 = \frac{1}{N} \sum_a (\mathbf{x}_a - \bar{\mathbf{x}}_a)^2,$$

where the sum is over the N ions in the cluster, \mathbf{x}_a is the position of the ion, and $\bar{\mathbf{x}}_a$ is the geometric center of the cluster. It is commonly used as a measure of compactness.

At size 30, the results are consistent with the intuition that more compact clusters are favored for nucleation. Surface area, sphericity, and the radius of gyration of nucleated cluster tend to assume values compatible with compact shapes. The effect, however, is not noticeable at size 10, suggesting that those measures are not sensitive enough to describe the fine structure of smaller clusters, and that these aspects do not influence the early stages of nucleation.

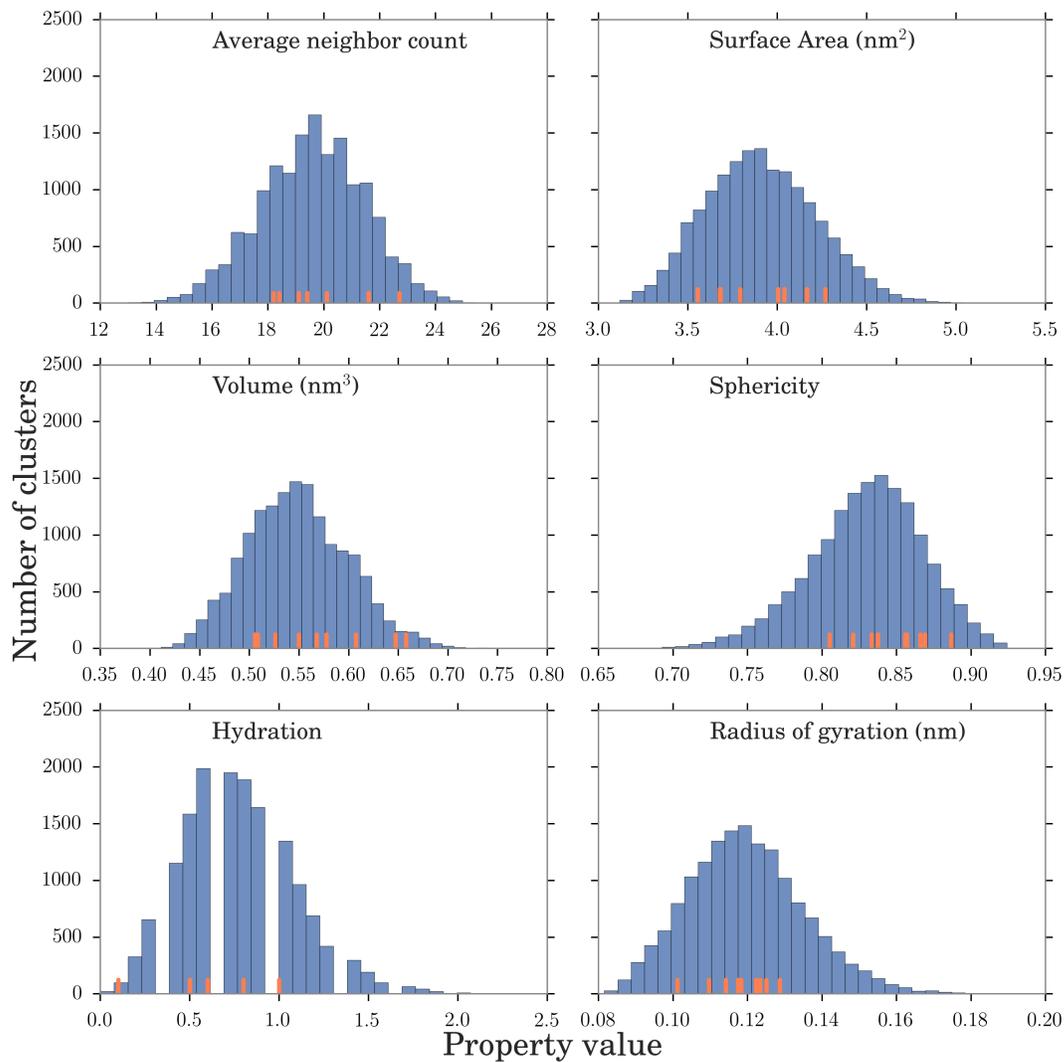


Figure A.1: Distributions of various properties for failed clusters (blue histogram) of size 10. Clusters that achieved nucleation are indicated by single orange lines. Note that the property values of the clusters that achieve nucleation are found around the mean of the distribution.

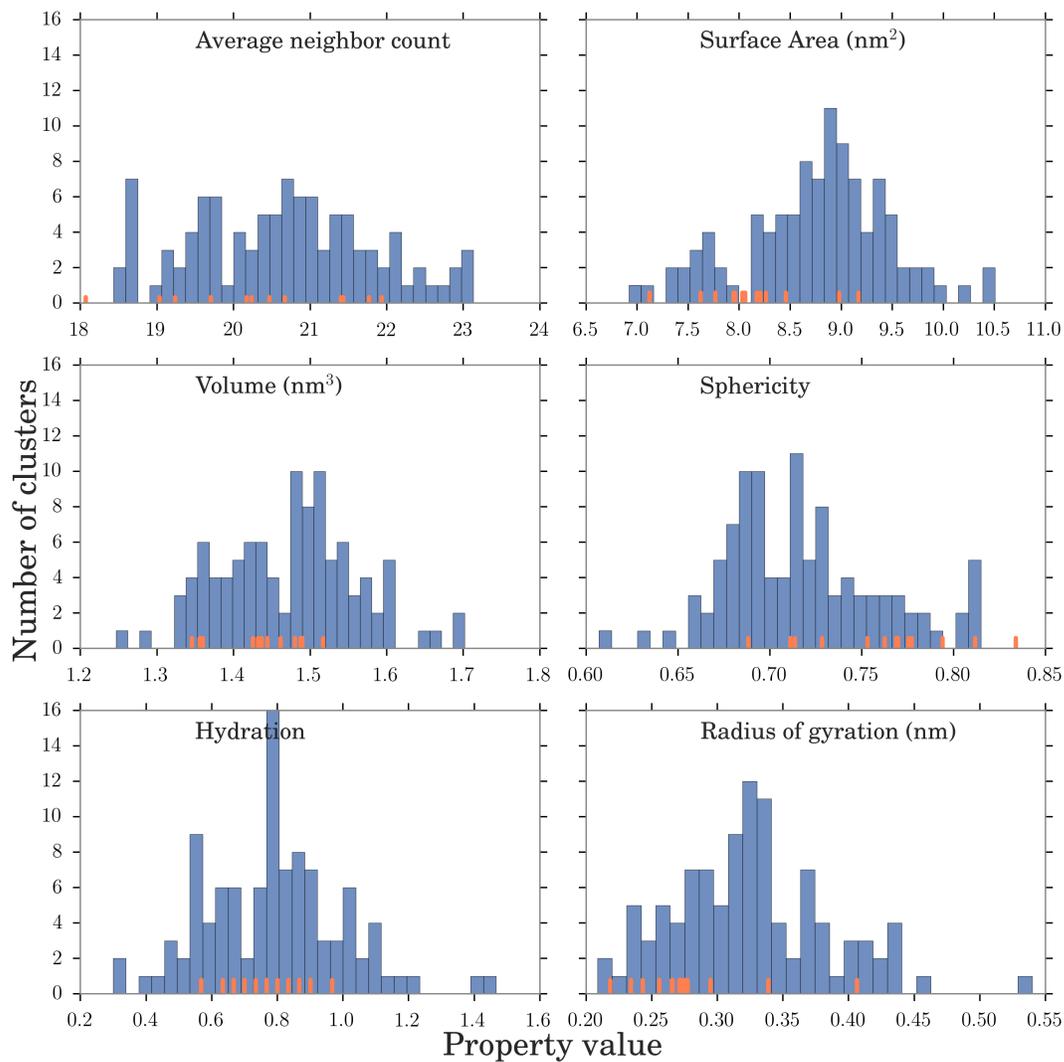


Figure A.2: Distributions of various properties for failed clusters (blue histogram) of size 30. Clusters that achieved nucleation are indicated by single orange lines. At this size, some preference for low radius of gyration, low surface area and higher sphericity can be observed.

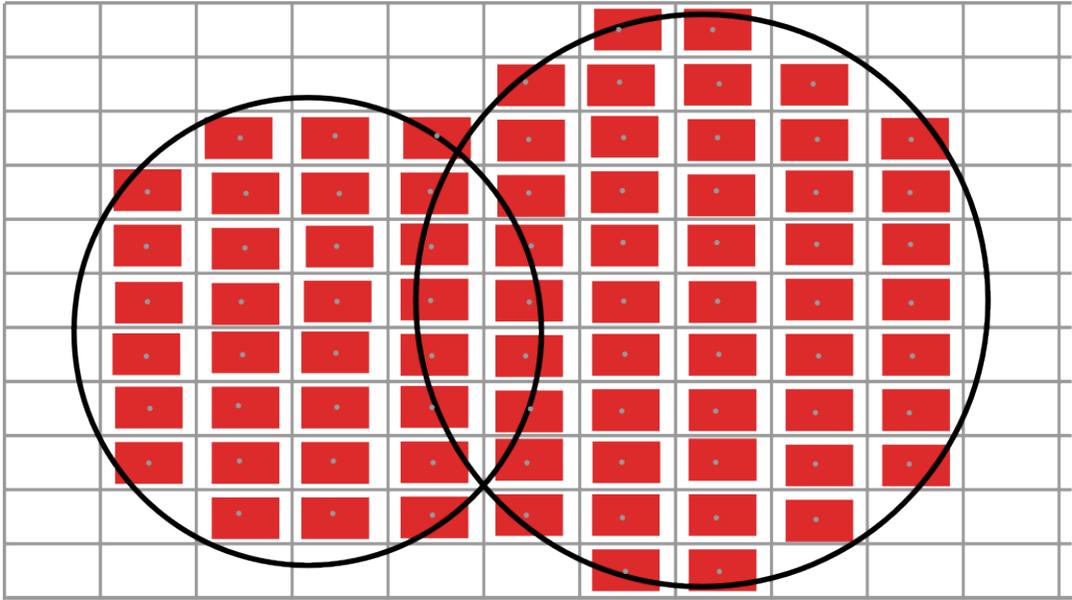


Figure A.3: Two dimensional representation of the algorithm used to estimate the volume. The cluster is encased in a grid, and the grid points that lie within the cluster are represented in red.