Nucleation and Dehydration of Calcium Carbonate

by

Aaron R. Finney

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Chemistry

2014

THE UNIVERSITY OF

WARWICK

Contents

List of Tables

List of Figures

Acknowledgments

My thanks and appreciation go to my supervisor, Mark Rodger, for his guidance, advice and patience throughout my degree. His seemingly limitless knowledge of the field, enthusiasm for new ideas, and his encouragement for me to fulfil my own research goals have helped and enabled me to develop into the competent researcher I am today. The opportunities that Mark has given me at Warwick, the wider UK and beyond have certainly broadened my horizons. I also wish to thank Mark for his personal support.

My gratitude is extended to John Harding and all the members of the MIB programme for financial support under an EPSRC grant, and for all the opportunities the consortium has provided. The scientific discussions and workshops we have held together have been extremely useful and have motivated me greatly. The successful collaborations that I have made during my PhD were only possible with the connections made through my association with MIB.

My progress over the course of my degree has been aided by insightful discussions and support from David Quigley, Matt Bano, and Yuriy Bushuev. I would like to thank all the members of the CSC, and in particular wish to single out Salvatore Cosseddu, Louise Wright, Julia Choe and Jasmine Desmond for time well spent having what seemed like endless and inconclusive debates about the qualities of metadynamics (often over a much needed coffee).

Finally, but by no means least, I wish to thank my family and friends, both past and present, for their continued and unfaltering support and encouragement. Any success I achieve is possible thanks to the support of my parents, Debbie and Robert, and my brother and his partner, Leevan and Claire (who often let me squat at their house during the writing of this thesis). Particular thanks go to WURAS for the occasional sanity breaks (aka trips to the pub).

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

The work presented (including data generated and data analysis) was carried out by the author except in the cases outlined below:

- The force field for Aspartate was provided by David J. Sparkes and Dr. Colin L. Freeman of the *Department of Materials Science and Engineering, University of Sheffield.*

- Simulations of amorphous calcium carbonate and subsequent analyses of pair correlation functions in Chapter 7 were generated by Dr. Yuriy G. Bushuev of the *Centre for Scientific Computing and Department of Chemistry, University of Warwick.*

- Experimental studies of calcium carbonate precipitation in Chapter 6 were conducted by Paul J. Smeets and colleagues under the supervision of dr. Nico A. J. M. Sommerdijk of the *Department of Chemical Engineering and Chemistry, Eindhoven University of Technology.*

A selection of the material in Chapters 3 and 4 has been published in,
A. R. Finney and P. M. Rodger, Probing the structure and stability of prenucleation clusters of calcium carbonate. Faraday Discussions, 159:47–60, 2012.

A selection of the material in Chapter 6 has been written into the submitted manuscript,
P. J. M. Smeets, A. R. Finney, W. J. E. M. Habraken, F. Nudelman, H. Friedrich, J. Laven, P. M. Roger, N. A. J. M. Sommerdijk. A classical view on non-classical nucleation.

A selection of the material in Chapter 7 has been written into the submitted manuscripts:
Y. G. Bushuev, A. R. Finney and P. M. Rodger. Stability, structure and dehydration of hydrated amorphous calcium carbonate.

A. R. Finney, Y. G. Bushuev, and P. M. Rodger. On the mobility of water in amorphous calcium carbonate.

Abstract

Great challenges remain in our understanding of biomineralisation which impede the design and production of a new class of materials. Recent studies have suggested that calcium carbonate clusters are stable in solution before nucleation, and that these participate in the formation of amorphous calcium carbonate (ACC). The structure and stability of these clusters has not been fully determined. Furthermore, the dehydration of stable ACC before crystallisation remains ambiguous.

Exhaustive computational searches have been carried out to provide a sample of clusters up to the sizes suggested for particles persisting before nucleation. A large sample of clusters were simulated at high concentration in water using molecular dynamics (MD). The results suggest that cluster stability is a balance between ionic coordination and ion hydration. At high concentration clusters are generally dynamic in the lower limit of stability, forming chains to which ions frequently aggregate and dissolve.

Free energy calculations showed a transition in the favoured coordination levels with cluster size. One dimensional Umbrella Sampling calculations showed that at small sizes a collection of clusters with low average cation–anion coordination were thermodynamically stable. For systems containing sixty ions and above, more compact clusters with internal water, close to the stoichiometries identified for stable hydrated ACC *in vivo*, were lower in free energy.

From MD simulations at experimental concentrations and pH, while dynamic ordering was found, ion pairs dominated in solution and the largest clusters observed contained no more than four ions. These findings, combined with the data at high concentration, allow for a re-evaluation of the proposed nucleation mechanisms for calcium carbonate from solution.

ACC simulations identified water–filled channels within the ionic framework. Percolating clusters were found when $H_2O/CaCO_3$ was greater than 0.75–0.8. The ACC system fitted well with the percolation theory on a simple cubic site lattice for water, and critical exponents showed a good fit to the theoretical values. Non–standard diffusion was found for water, with a "jump" mechanism of diffusion observed and a cascade of molecule displacements within channels. This original result allows new light to be shed on the dehydration mechanism of ACC.

Abbreviations

ABMD	Adiabatic Biased Molecular Dynamics
ACC	Amorphous Calcium Carbonate
ASP	Aspartate
AUC	Analytical Ultracentrifugation
CG	Conjugate Gradient
CNT	Classical Nucleation Theory
cryo-TEM	Cryotransmission Electron Microscopy
CSD	Cluster Size Distribution
CV	Collective Variable
DLNP	Dense Liquid Nanoparticle
DOLLOP	Dynamically Ordered Liquid-like Oxyanion Polymer
DSC	Differential Scanning Calorimetry
FE	Free Energy
FES	Free Energy Surface
FTIR	Fourier Transform Infrared
HRTEM	High-resolution Transmission Electron Microscopy

MC   Monte Carlo

MD   Molecular Dynamics

MHC   Monohydrocalcite

NMR   Nuclear Magnetic Resonance

OP   Order Parameter

PEEM   Photoelectron Emission Microscopy

PES   Potential Energy Surface

PILP   Polymer-induced Liquid Precursor

PNC   Prenucleation Cluster

RDF   Radial Distribution Function

REMD   Replica Exchange Molecular Dynamics

RMC   Reverse Monte Carlo

RSS   Random Structure Searching

SAM   Self Assembled Monolayer

SEM   Scanning Electron Microscope

TGA   Thermogravimetric Analysis

US   Umbrella Sampling

WHAM   Weighted Histogram Analysis Method

XANES   X-ray Absorption Near Edge Spectroscopy

Chapter 1

Introduction

Minerals are a key feature of the structural and defensive mechanisms of organisms in nature. Minerals found in biology – biominerals – are composite, formed from "hard" crystalline phases with inclusive, "soft" organic molecules such as proteins and peptides [Sigel *et al.*, 2008]. Calcium based minerals are found throughout the animal kingdom. Calcium carbonate is particularly common in marine organisms; for instance, the mollusc shell is formed from a number of calcium carbonate polymorphs, and is rich with organic components including glycoproteins and β-chitin [Addadi *et al.*, 2003]. The inclusion of organic molecules lends biominerals special mechanical properties. The human skeleton is a biomineral composed of hydroxyapatite with chemical formula $Ca_{10}(PO_4)_6(OH)_2$, deposited onto collagen fibrils [Martin and Ishida, 1989]. The composite material provides strong and hard tissue to support the body and protect vital organs, yet bone is both lightweight and porous, allowing for fluid flow. The incorporation of collagen into bone tissue provides a certain amount of elasticity to the material, and an increased resistance to stress [Martin and Boardman, 1993].

Organic molecules not only tailor the mechanical properties and functionality of biominerals, but also direct the growth of particular crystalline polymorphs with high selectivity [Cölfen, 2003]. The level of intricacy and selectivity displayed in biomineral formation is currently inaccessible to materials engineers in the laboratory. It is no wonder that scientists have endeavoured to mimic biomineral formation mechanisms in the pursuit of novel materials. However, to do this, a fundamental understanding of the stages involved in biomineral formation – biomineralisation – must first be sought.

In the following Chapter, the properties and features of calcium carbonate, the most abundant biomineral in nature, are presented. Insight into the biomin-

eralisation of calcium carbonate from both experiment and computational studies are discussed, highlighting the remaining challenges to understand these processes. Finally, the aims of the work in this thesis are outlined, in order to address a number of the challenges.

1.1   Calcium Carbonate

Calcium carbonate is a ubiquitous mineral in nature. As the major constituent of limescale, it is also a mineral which can be found in abundance in industry and the home. It can be found in three crystalline anhydrous forms: vaterite, aragonite and calcite in order of increasing thermodynamic stability [Radha *et al.*, 2010]. There are a further two hydrated crystalline forms known as monohydrocalcite (MHC) and ikaite with stoichiometries $CaCO_3 \cdot H_2O$ and $CaCO_3 \cdot 6\,H_2O$, respectively. Aragonite and calcite can be found in abundance in nature, both in biological organisms and geological samples, while vaterite is less common in biology [Meldrum, 2003]. Ikaite is unstable under standard conditions, and decomposes at temperatures just above $0\,°C$ [Marland, 1975]. In the following subsections, each of the structures and examples of crystalline phases in nature are briefly discussed. Representations of the unit cells for anhydrous and hydrated crystals are shown in Figures 1.1 and 1.2, respectively, while Table 1.1 provides the unit cell parameters and spacegroups for crystals. Crystal structures have been obtained using the *United Kingdom Chemical Database Service* [Fletcher *et al.*, 1996].

1.1.1   Aragonite

Aragonite has an orthorhombic crystal lattice. As can be seen in Figure 1.1, the crystal structure is made from consecutive layers of calcium and carbonate ions perpendicular to the *c* axis. Carbonates are co-aligned within layers, but adjacent layers of carbonate have alternating orientation. Hexagonal packing is found for calcium ions and also for carbonates between two adjacent layers in the *ab* plane, hence crystals of aragonite tend to have a hexagonal shape.

The most studied example of aragonite in nature is within the nacreous layer of the mollusc shell. Below a thin prismatic layer, the main constituent of the shell is nacre which is formed from layers of hexagonal (or rectangular) aragonite tablets, typically around 10 μm in diameter and 50 nm deep [Addadi and Weiner, 1997; Nudelman *et al.*, 2006]. Extracellular organic matrices can be found both between adjacent tablets and between parallel layers of aragonite. Levi-Kalisman showed [Levi-Kalisman *et al.*, 2001] that the intralamellar layer of organic material is
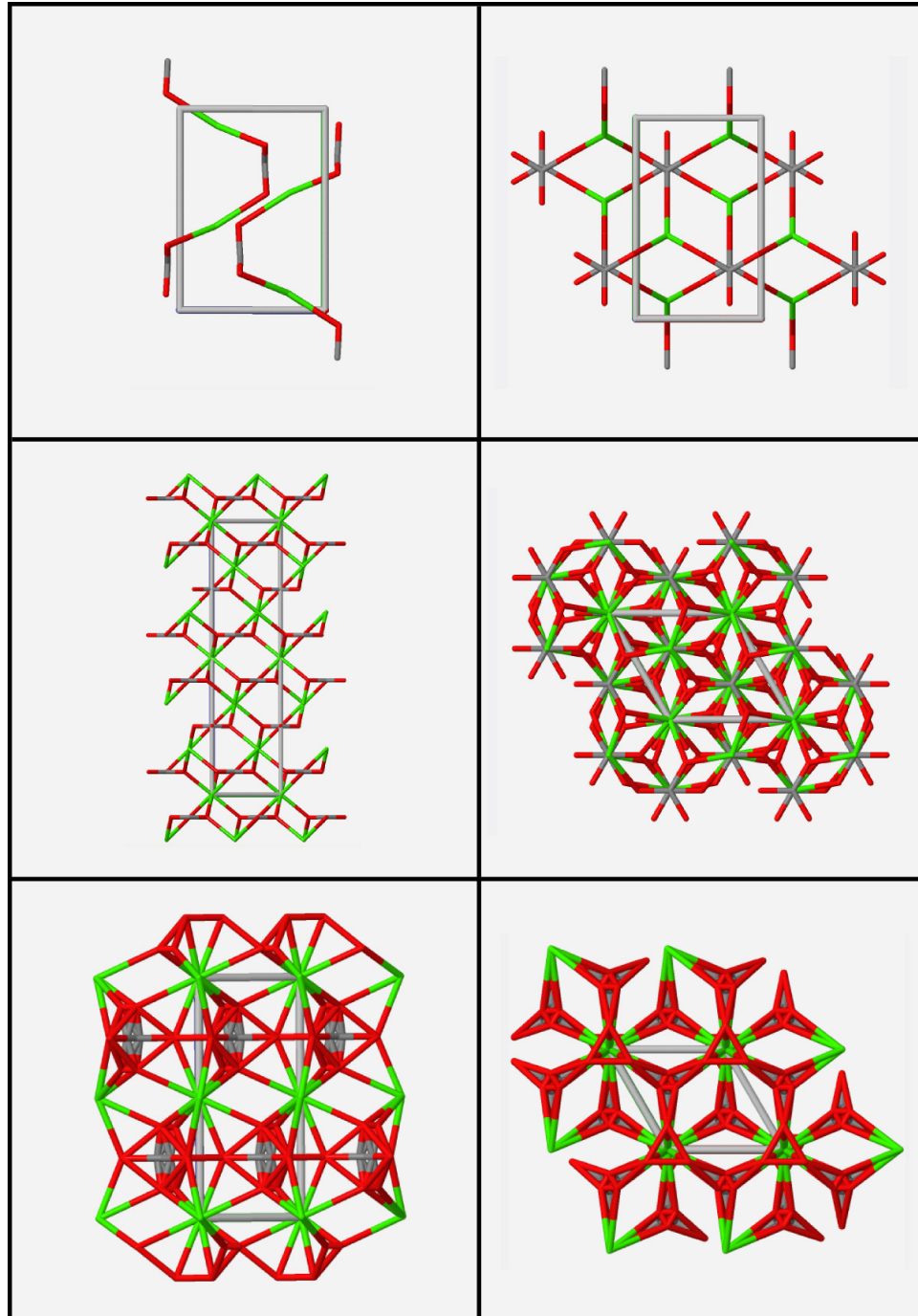
Figure 1.1: Unit cell projections along cell a (left) and c (right) axes for aragonite (top), calcite (middle) and vaterite (bottom). The occupancy of carbonates in the vaterite unit cell is one third, hence not all of those shown are present in the lattice. The unit cell can be seen in grey, and atoms are coloured green, dark grey and red for calcium, carbon and oxygen, respectively. The source for crystal structures are provided in Table 1.1.
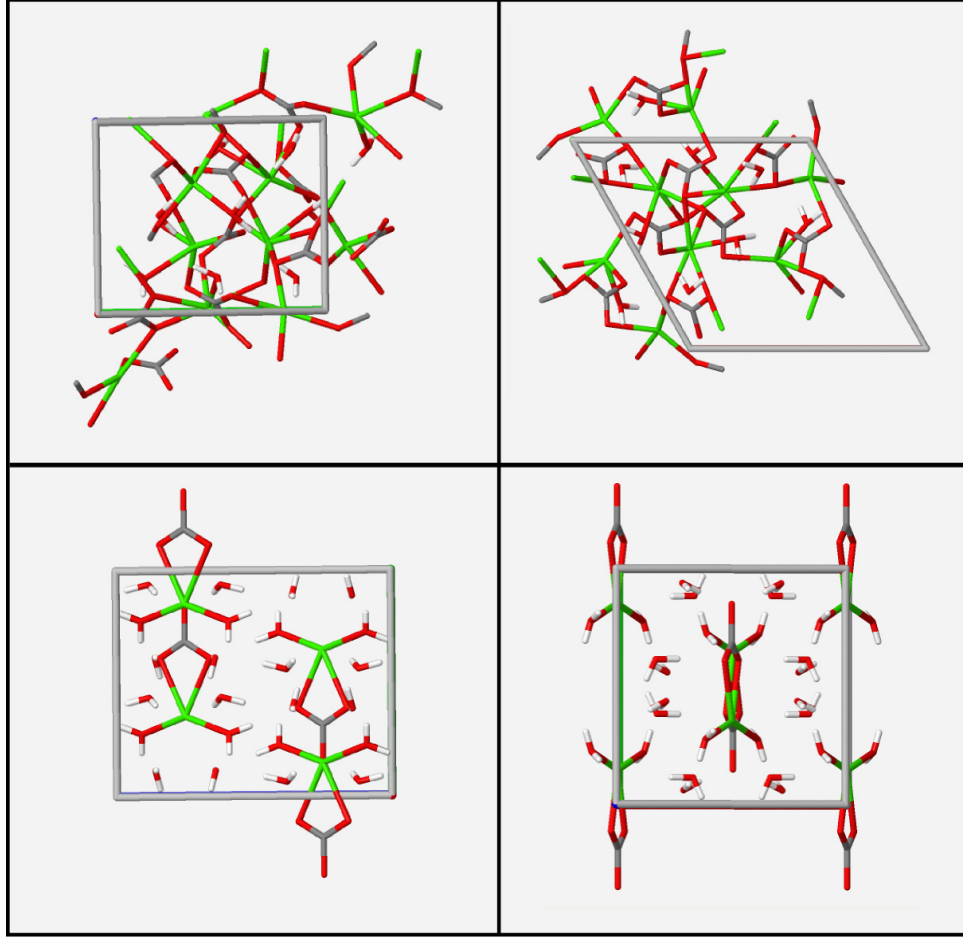
Figure 1.2: Unit cell projections along cell a (left) and c (right) axes for monohydro-calcite (top) and ikaite (bottom). The unit cell can be seen in grey, and atoms are coloured green, dark grey, red and white for calcium, carbon, oxygen and hydrogen, respectively. The source for crystal structures are provided in Table 1.1.

Table 1.1: Spacegroup and unit cell information for crystalline phases of calcium carbonate. Data was taken from: Antao and Hassan [Antao and Hassan, 2009] for aragonite; Antao and Hassan [Antao and Hassan, 2010] for calcite; Kamhi [Kamhi, 1963] for vaterite; Swainson [Swainson, 2008] for MHC; and Lennie *et al.* [Lennie *et al.*, 2004] for ikaite.

	Spacegroup	Minimum cell length (Å)	Cell length ratio ($a : b : c$)	Unit cell angles (α, β, γ /°)
Aragonite	$Pmcn$	4.96	$1 : 1.61 : 1.16$	$90, 90, 90$
Calcite	$R\bar{3}c$	4.98	$1 : 1 : 3.6$	$90, 90, 120$
Vaterite	$P6_3/mmc$	7.15	$1 : 1 : 2.37$	$90, 90, 120$
MHC	$P3_1$	7.57	$1.39 : 1.39 : 1$	$90, 90, 120$
Ikaite	$C2/c$	8.31	$1.06 : 1 : 1.33$	$90, 110.6, 90$

mostly formed from ordered β-chitin fibrils which direct the crystal growth of aragonite along the a axis. However, Nudelman and colleagues showed [Nudelman *et al.*, 2006], using a combination of fluorescence and electron microscope techniques, that the nucleation is driven by the presence of organic compounds rich in carboxylates and sulphates, located in specific zones. Evolution has therefore led to a sophisticated system comprising many components to grow the mollusc shell. The organic layers not only control the growth of aragonite, but also provide the material with its special mechanical properties. As highlighted by Meldrum [Meldrum, 2003], organics typically compose around one percent of nacre, yet the distribution of "soft" molecules mean that aragonite is much stronger and tougher than most other composite materials. Cracks in nacre propagate through the organic layer, making nacre orders of magnitude more resistant to fracture than pure aragonite [Meldrum, 2003].

Aragonite is present in Scleractinian corals, also known as stony corals. Organic molecules are again present in these corals, but the organic component is more acidic than in the mollusk shell. The organics aid in the deposition of aragonite, and combined with magnesium impurities, control the growth of aragonite at multiple scales to give the coral its distinctive shape [Falini *et al.*, 2013]. Microstructural units of coral are grown via the deposition of densely packed micron sized crystals (which can be randomly oriented), surrounded by an organic matrix, from which aragonite skeletal fibres radiate [Cuif and Dauphin, 1998]. Coral growth has received much attention lately from environmental scientists, as changes in ocean temperature and acidity (associated with climate change) can diminish corals ability to calcify; by studying historical fluctuations in coral populations, current climatic events may be better understood [Pandolfi and Kiessling, 2014].

1.1.2 Calcite

The most stable phase of calcium carbonate is calcite under standard conditions, and hence the presence of calcite in nature is widespread. Calcite has trigonal crystal symmetry; layers of planar carbonate oriented perpendicular to the c axis separate layers of calcium ions, with spacings $\approx c/12$. Adjacent layers of carbonate have alternating orientation along the cell a axis. Six-fold coordination is found for calcium between carbonates in adjacent layers.

Perhaps the most common calcite biomineral in the home is avian egg shell. Egg shell is constructed especially for its purpose as a secure environment for a developing embryo, with both water and air regulation, and which provides a resistive layer to contamination from external microscopic organisms. During formation, calcium carbonate is secreted onto the outer shell membrane from which calcite

columns grow radially, with pores interspersed to allow chemical and temperature regulation [Dieckert *et al.*, 1989]. Avian egg shell contains organic material providing mechanical strength and structural control. As highlighted by Chien and co-workers [Chien *et al.*, 2008], typically, four percent of chicken egg is formed from a number of proteins; the acidic glycoprotein, osteopontin, which is thought to regulate mineralisation, is a major component of the shell. Another prominent protein, Ovocleidin–17, has been shown to promote crystallisation of calcium carbonate nanoparticles to calcite [Freeman *et al.*, 2010].

Coccoliths are an exquisite example of single crystal assemblies of calcite which reside at the surface of coccolithophores, as shown in Figure 1.3. The coccolithophore is a single cell eukaryotic phytoplankton, upon which disc shaped coccoliths form an exoskeleton and presumably impart defensive support. However, the exact function of the coccoliths is unknown. Young *et al.* have provided a detailed study of coccolith formation [Young *et al.*, 1999]. Coccoliths nucleate within vesicles upon a preformed organic scale which has a microfibrill structure. Calcite nucleates at the rim of this scale, and subsequent growth upward and outward results in the formation of the coccolith disc. Once formed, the vesicle and coccolith are separated from the cell to the surface, and an organic layer surrounds the calcite structure.
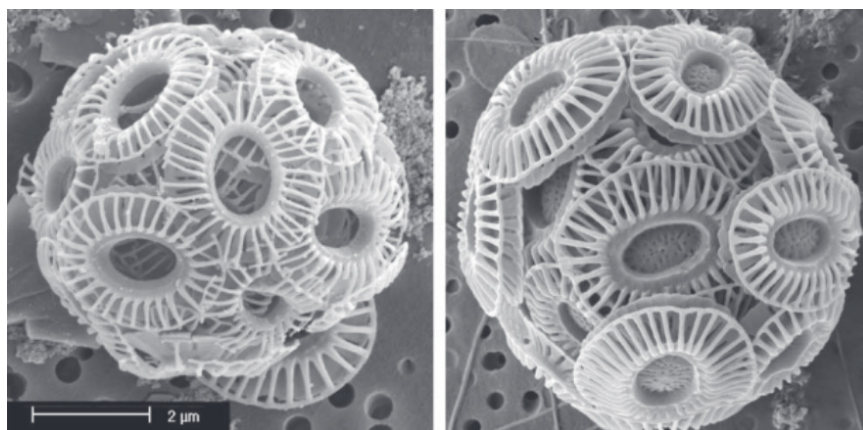
Figure 1.3: Optical microscope images of coccoliths at the surface of coccolithophores, taken from two samples of the species *Emiliana huxleyi*. Images were taken from Ridgwell *et al.* [Ridgwell *et al.*, 2009].

1.1.3 Vaterite

The crystal structure for vaterite has often been misunderstood. Over decades, many different crystal structures have been proposed, some of which significantly contrast with each other [Meyer, 1969; Kamhi, 1963; Wang and Becker, 2009; Mugnaioli *et al.*, 2012]. The commonalities between suggested crystal structures are that calcium forms a hexagonal lattice with six-fold coordination. Calcium and carbonate layers are found in the *ab* plane; however, carbonate orientation tends to be close to perpendicular to the plane. The oxygen sub-lattice is disputed between the suggested crystal structures. The unit cell provided in Figure 1.1 was one of the earliest identified by Kamhi [Kamhi, 1963], which shows partial carbonate occupancy. Clusters of three overlapping carbonates can be seen which are slightly offset from each other, but the occupancy for each is one third, hence only one site in the lattice is fully occupied.

Recently, Kabalah-Amitai and co-workers have elucidated the vaterite structure [Kabalah-Amitai *et al.*, 2013]. Using high-resolution synchrotron powder diffraction and high-resolution transmission electron microscopy (HRTEM), the authors were able to analyse ascidian spicules. These particular samples were used as the single crystals obtained were relatively large compared with those from other biological sources or synthetically prepared vaterite. Using HRTEM on small volumes of the samples (the beam width being ~ 5 nm), two coexisting structures were found. The major component showed a very good fit to the Kamhi structure (see Figure 1.1), while the minor component was not able to be resolved. The distribution of the two crystal lattices within the samples was random. This supports the fact that the Kamhi structure has become the most widely endorsed structure for vaterite, but the identification of minor crystalline domains inconsistent with the Kamhi structure were, for the first time, observed thanks to the high resolution analysis. The authors highlight that the single crystal scattering for vaterite is dominated by calcium, and suggest that if the lattice sites for the cation in both major and minor components are uniform, the "single crystals" will appear to show a single diffraction pattern.

Vaterite has been observed as the product of calcium carbonate growth in experimental studies, often being formed in the presence of organic templates [Naka *et al.*, 2006; Pouget *et al.*, 2010]. The crystal is, however, relatively rare in biogenic settings compared to the more stable aragonite and calcite. There are some examples, such as in egg shells of soft-shelled turtles, where the insoluble peptide, pelovaterin, has been shown to promote the nucleation and growth of vaterite as well as stabilising the phase [Lakshminarayanan *et al.*, 2005]. Glycoproteins were found to help nucleate and stabilise vaterite in freshwater pearls [Natoli *et al.*, 2010].

1.1.4 Monohydrocalcite

MHC ($CaCO_3 \cdot H_2O$) has a trigonal crystal symmetry with a rhombohedral unit cell. Layers of calcium can be observed which cut through cell a, b and c axes, but these are not perfectly planar. Planar carbonates and water molecules occupy sites within these layers as well as above and below. There are three types of calcium (calcium has coordination numbers 3, 4 and 5), carbonate and water according to their symmetries. Hydrogen-bonding is observed in the crystal, with three networks identified in the unit cell; a water molecule is H-bonded to one carbonate molecule with a linear H-bond to one oxygen of carbonate, while bifurcated H-bonds are formed to the remaining two oxygen atoms and water ($D \cdots A = 2.68 - 3.16$ Å). MHC is not found biogenically, and under ambient conditions it is metastable with respect to anhydrous phases. MHC is predominantly found in marine sediments where there are high concentrations of magnesium present [Neumann and Epple, 2007].

1.1.5 Ikaite

The crystal system of ikaite ($CaCO_3 \cdot 6\,H_2O$) is monoclinic. Planes of calcium and carbon can be found parallel to the cell a axis, with one carbonate binding in a bidentate fashion to one calcium ion. The carbonates are planar, with the OOO plane forming an angle of approximately 11.5° with the bc plane. The coordination number of calcium is four, and so two water molecules coordinate directly with the cation as well as the carbonate. Two lattice sites can be identified for water. Water molecule geometries deviate from those found in solid water: O–H distances were measured at −25°C to be $0.73 - 0.83$ Å, and $\theta_{HOH} = 107.3 - 111.4°$ [Hesse *et al.*, 1983]. Water molecules participate in hydrogen bonding with adjacent carbonate molecules, with $D \cdots A = 2.74 - 2.91$ Å, and $\theta_{DHA} \approx 164 - 177°$ [Hesse *et al.*, 1983]. First-principles calculations have recently shown that H-bonding between waters in ikaite leads to the formation of extended, complex networks, with each water molecule also strongly binding to carbonate [Demichelis *et al.*, 2014].

Ikaite is not found in biominerals and is generally unstable under ambient conditions (where dehydration to MHC can occur). However, a recent study has shown that in confinement, micrometre sized ikaite precipitates can be stabilised at room temperature, without the use of additives [Rodrguez-Ruiz *et al.*, 2014]. Natural ikaite is found below sea level (where conditions favour its stability), and in fact ikaite is named after the Ika Fjord in Greenland where it was first discovered.

1.2 Biomineralisation

1.2.1 *In vivo*

A general scheme for the formation of biominerals *in vivo* is provided in Figure 1.4. A great number of researchers have provided insight into the mechanism of formation, with a vast literature dedicated to the structure and stabilities of phases identified at each stage [Sigel *et al.*, 2008; Weiner, 2008; Meldrum and Colfen, 2008; Weiner and Addadi, 2011]. The crystallisation of calcium carbonate is believed to be "non-classical": crystallisation is preceded by the formation of an amorphous calcium carbonate (ACC) solid, as opposed to direct crystal formation from free ions in solution. It is conceivable that crystalline phases could be formed via a multi-step classical pathway which includes the classical nucleation (where the likelihood of solid phase growth is governed by bulk and surface energies) of amorphous phases, and so the "non-classical" description may be incorrect. Biogenic ACC was identified as a precursor phase in sea urchin spicules by Weiner and co-workers [Beniash *et al.*, 1997] (although amorphous precursors to biogenic minerals were identified much earlier [Towe and Lowenstam, 1967]), and have since been observed in other organisms prior to crystallisation [Weiner and Addadi, 2011]. Both stable and transient ACC have been observed as the first solid deposits during calcification. Stable ACC contains structural water: about one mole of water is found, on average, per mole of calcium carbonate [Levi-Kalisman *et al.*, 2000; Addadi *et al.*, 2003]. Depositing an amorphous phase can be beneficial as this can provide a store for calcium and carbonate, from which crystallisation can occur at a more appropriate time for the organism [Raz *et al.*, 2002; Sigel *et al.*, 2008]. In some organisms, stable ACC has been found to provide structural support or to impart special mechanical properties in composites [Aizenberg *et al.*, 1996; Nassif *et al.*, 2005].

Aizenberg and colleagues provided one of the first studies into the mechanism of calcium carbonate biomineralisation in samples of ascidian invertebrate skeletal tissue [Aizenberg *et al.*, 2002]. Body and tunic spicules were extracted from the ascidian *Pyura pachydermatina*. X–ray spectroscopy confirmed that antler spicules were amorphous, while tunic spicule surfaces were composed of crystalline calcite. However, the core of tunic spicules was amorphous, with crystalline and amorphous phases separated by an organic sheath in the core. Both ACC and calcite in tunic spicules contained magnesium, with ACC containing higher levels of the impurity (1.7 and 5.9 mol % in the calcite and amorphous layers, respectively). Organic molecules were extracted from both phases; the calcite layer contained proteins rich in aspartate, while ACC contained relatively large amounts of glutamic and hydrox-
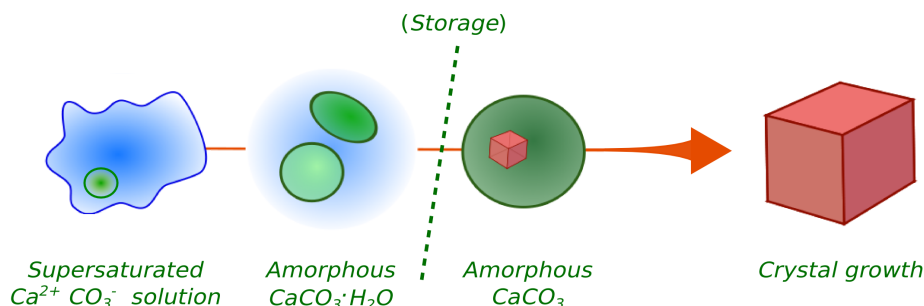
Figure 1.4: A general scheme for biomineralisation. The orange arrow shows a path from a supersaturated ionic solution in which solid nucleates (shown by the green phase). The first phase precipitated is hydrated amorphous calcium carbonate (ACC) which grows to be nanometres in size. Functional hydrated ACC has been found, but typically, hydrated ACC is transient and dehydrates to amorphous $CaCO_3$ (perhaps after a period of storage marked by the dashed green line). Crystallisation via solid-state transformation takes place in dry ACC, indicated by the emergence of the red cube. Further growth proceeds via re-precipitation of the surrounding amorphous material to produce bulk crystal [Gong *et al.*, 2012].

yamino acids. In solution the calcite proteins were shown to increase crystallisation rates, while those from ACC inhibited crystal growth. This is evidence that the organic molecules directly control the mechanism of calcification in organisms. The organic sheath layer prevents propagation of crystal into ACC, with the composite structure providing mechanical properties necessary for structural support.

Sea urchins have proved to be invaluable in the study of biomineralisation. Sea urchin larval spicules have been studied extensively, with a great deal of insight into the mechanism of growth provided by Gilbert and co-workers [Politi *et al.*, 2008; Radha *et al.*, 2010; Gong *et al.*, 2012] Initially, a single crystal of calcite is deposited inside a syncytium (produced by fusion of a number of cells) which is shown in Figure 1.5. Ion regulation in vesicles via ATP transporters concentrates Ca^{2+} and HCO_3^- in the intracellular region, leading to the precipitation of hydrated ACC particles up to approximately 120 nm in size. Confinement within the vesicles may lead to stabilization of ACC [Stephens *et al.*, 2010]. Upon fusion of the vesicle and synctinial walls (see Figure 1.5), the solid is deposited onto the spicule which forms radially following the crystallographic orientations of the seed crystal [Politi *et al.*, 2008]. ACC subsequently dehydrates as shown by X–ray absorption near edge
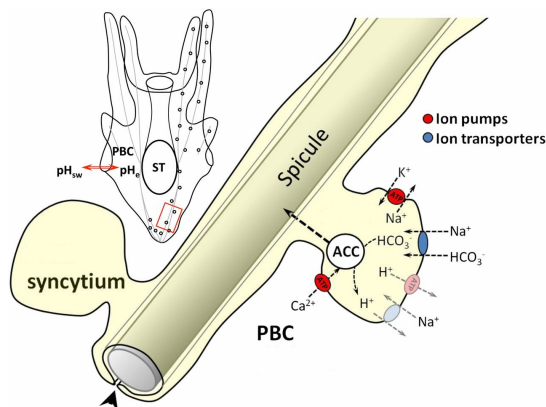
Figure 1.5: The formation of spicules in the sea urchin primary body cavity (PBC). The image on the top left is the juvenile sea urchin with the stomach (ST) highlighted (there is a noticeable difference in the internal pH, pH_e, and that of seawater, pH_{SW}). The red box highlights a spicule section which is expanded in the foreground. The syncytium, where calcite is first formed, is shown. Growth occurs via deposition of ACC from vesicles which fuse with the syncytial membrane. The image was taken and modified from Stumpp *et al.* [Stumpp *et al.*, 2012].

spectroscopy (XANES), before crystallisation to calcite. This mechanism allows for fast growth of the skeleton. By combining XANES and photoelectron emission microscopy (PEEM), Gong *et al.* were able to identify the structural phases on the surface of spicule cross-sections from a number of biological samples [Gong *et al.*, 2012]. This showed that after 36, 48 and 72 hours from fertilisation, spicules grew in size, with the major phase at each stage being hydrated ACC, ACC and calcite, respectively, confirming stepwise transformations. After 72 hours, the spicules were formed almost entirely from crystalline calcite, but small regions of hydrated ACC were found to remain on the spicule surface, and these were hypothesised to be stabilised by organic inhibitors to crystallisation. The protein SM50 was extracted from the spicules, and this was shown to stabilise hydrated ACC *in vitro* [Gong *et al.*, 2012].

Radha *et al.* analysed the enthalpies of biogenic ACC relative to calcite [Radha *et al.*, 2010]. By employing a number of thermochemical cycles, the enthalpy of dissolution was measured to be -42.5 ± 0.5 kJ mol^{-1}, and enthalpy of crystallisation was -12.86 ± 0.76 kJ mol^{-1}. ACC was shown to be metastable with respect to all anhydrous crystalline phases of calcium carbonate, and therefore a possible precursor to the formation of all $CaCO_3$ crystals in biology. A small amount of water is present in biogenic ACC and this was lost during crystallisation. Also, the samples were shown to contain 0.05 : 0.95 Mg:Ca, but the presence of magne-

sium did not significantly change relative enthalpies when compared to synthetically produced ACC in the absence of ionic impurities.

1.2.2 *Ex vivo* Model Studies

Elucidating biomineralisation mechanisms directly from biological samples can be impractical, and the complex environments in which calcification takes place can make it difficult to understand the effects of variables to nucleation and growth. *In vitro* and *in silico* experiments are essential to studying biomineralisation. The literature is vast and cannot be summarised easily; nonetheless a number of examples will be provided here.

Idealised solution experiments are perhaps the most commonly employed model to study biominerlisation. Bolze and co-workers showed that the initial solid produced in pure solutions is ACC which is found in a colloidal form [Bolze *et al.*, 2002]. Particles up to around 270 nm (in terms of colloidal diameter) were found with low polydispersity and a mass density lower than that of crystalline polymorphs. Later studies showed that ACC particles were stable up to around a size of 100 nm in diameter [Nudelman *et al.*, 2010].

Wang *et al.* used turbidity measurements to study the precipitation of calcium carbonate, which confirmed the limiting size to ACC stability of around 100 nm [Wang *et al.*, 2012]. They showed that further growth above this size took place via aggregation of ACC, but that crystallisation to calcite in aggregates occurred. Vaterite was also found, and this transformed in time to calcite [Meldrum and Colfen, 2008]. The transformation presumably takes place via Ostwald's step rule, where the emerging phase is the one nearest in stability to the original; a number of phases (with increasing stability) may be visited in a multi-step, energetically downhill path to the thermodynamically favoured one. The mechanism of transformation between states is not defined by Ostwald's rule, and it is not true that this multi-step pathway to the most stable phase occurs for all materials [Threlfall, 2003]. Later experiments [Wang *et al.*, 2013; Aziz *et al.*, 2011] showed that there is a certain degree of kinetic control to calcium carbonate precipitation, with changes in stirring rate, pH and temperature affecting particle size and the product of crystallisation. Understanding the co-action between kinetic and thermodynamic control could be key to exploiting biomineralisation processes.

Nanoparticle computer simulations have provided a much needed understanding of solid phases at length and time scales difficult to probe with experimental techniques. Molecular dynamics (MD) simulations by Cooke and Elliot have shown that calcite crystals containing less than forty formula units of calcium car-

bonate amorphise in solution, indicated by a change in pair distribution functions and carbonate orientation [Cooke and Elliott, 2007]. Calcite crystalline order was found to persist in solution for clusters containing more than forty formula units [Cooke and Elliott, 2007]. Tribello *et al.* measured no energy barrier to the growth of ACC from free energy calculations of ion addition to nanoparticles, highlighting that calcification by the initial deposition of an amorphous phase is faster than classical crystallisation [Tribello *et al.*, 2009]. Quigley *et al.* have shown using free energy calculations that calcite particles are thermodynamically stable at sizes of 2 nm, with the amorphous phase only lower in energy due to confinement effects [Quigley *et al.*, 2011]. Unfortunately, the force fields employed in these simulations do not reproduce the solvation enthalpies of calcium carbonate, and therefore cannot predict the stabilities of phases in solution accurately [Raiteri *et al.*, 2010]. Later work by Raiteri and Gale showed that hydrated ACC particles up to around 5 nm in diameter are more stable than crystalline phases [Raiteri and Gale, 2010]. These studies seem to contradict the experimental findings. However, it is plausible that metastable ACC up to $\sim$ 100 nm persists in solution before crystallisation.

Additive experiments allow scientists to understand the influence of certain functional groups on the mechanism of biomineralisation. Xiao and Yang studied the effect of hydroxyl, carboxyl and amine–rich organic molecules on the polymorph control of calcium carbonate when formed from constituent ions in solution [Xiao and Yang, 2011]. Amorphous particle intermediates ($\sim$200 nm in diameter) all contained up to around 13 wt% water and organic molecules. While carboxyl groups showed a preference for calcite formation, hydroxyl and amine–rich molecules led to the formation of aragonite. The polymorph control was observed during the formation of ACC, and the inclusion of organics did not affect crystallisation during the ACC$\longrightarrow$crystalline transition [Xiao and Yang, 2011].

Some organic molecules have been used to stabilise ACC in solution. Examples include polydopamine [Wang and Xu, 2013], (poly)aspartate and DNA [Gower, 2008], and these additives generally work by kinetic control; coating ACC particles in organics can limit the extent to which they can grow, and so inhibit crystallisation [Nudelman *et al.*, 2010]. It is true that organic additives can influence many stages in biomineralisation, and Gebauer and co-workers have suggested a classification scheme according to how additives control crystal formation [Gebauer *et al.*, 2009]. For example, polyaspartate was shown to bind to calcium, inhibit nucleation and to influence the local structure of the bound phase, making it a type I, III and V additive [Verch *et al.*, 2011].

An often employed model system to study biomineralisation is precipitation

in the presence of organic monolayers [Harding *et al.*, 2014]. This system mimics the deposition of solid phases onto an insoluble organic matrix, which takes place in living organisms. Experiments have shown [Han and Aizenberg, 2003, 2008] that self assembled monolayer (SAM) chain length can control the crystallisation front, and head groups can stabilise particular polymorphs. Amorphous calcium carbonate has been found to be initially deposited onto acidic monolayers, with subsequent crystal orientation being directed by the organic template [Fricke and Volkmer, 2007]. Indeed, the geometry of the the SAM head groups can be highly selective to the formation of particular crystallographic planes [Aizenberg *et al.*, 1999b]. Simulations have shown that monolayer surface charge density is crucial to crystallisation; the lattice structure which closely matches SAM surface charge density is likely to be low in energy, and crystal propagation tends to follow this orientation [Freeman *et al.*, 2008]. Aizenberg and colleagues showed [Aizenberg *et al.*, 1999a] that calcite growth can be controlled on patterned SAMs, with crystal nucleation rates affected by the polarity of monolayer head groups.

1.3 Amorphous Calcium Carbonate

The amorphous phase of calcium carbonate deserves further discussion. The small sizes and often transient nature of ACC means that investigating the structure and properties of this material is difficult. Furthermore, the structural differences between stable "wet" ACC (with an approximate composition $CaCO_3 \cdot H_2O$ [Addadi *et al.*, 2003]) and transient "dry" ACC are difficult to determine from experiment. Nonetheless, a number of experiments have been conducted which help us to understand this phase. For an in depth insight into ACC and its role in biomineralisation see the review by Addadi, Raz and Weiner [Addadi *et al.*, 2003].

1.3.1 ACC Structure

Amorphous materials show no Bragg reflections in elastic scattering measurements, indicating that no long range atomic order is present. However, amorphous solids can still be described in terms of short range atomic order. Energetic studies have shown [Radha *et al.*, 2010] that synthetic ACC prepared from similar methods can have different stabilities relative to calcite; Radha *et al.* attributed this difference to the level of short range atomic ordering in the samples.

Short range atomic order which resembles crystalline phases has been found in biogenic ACC; analyses of sea urchin embryos have shown [Raz *et al.*, 2003; Politi *et al.*, 2006] that ACC which transforms to calcite has short range calcite–

like ordering. During transformation to crystal, the coordination sphere transforms to resemble that of calcite, before long–range rearrangement of ions takes place to produce the calcite lattice [Politi *et al.*, 2006]. Epple and co-workers have shown in growing embryos of snail that ACC short range atomic order becomes aragonite–like as the embryo develops [Hasse *et al.*, 2000; Marxen *et al.*, 2003]. Stable hydrated ACC from a number of biological samples have also shown atomic ordering consistent with MHC in the first 3–4 coordination shells [Neumann and Epple, 2007; Addadi *et al.*, 2003]. Intriguingly, it is not always the case that the crystal–like short range order observed in ACC evolves to the same respective crystalline lattice on crystallisation [Cartwright *et al.*, 2012].

The short range ordering in biogenic ACC could be a reason for its increased stability compared with synthetically prepared samples. This could possibly be induced by the presence of macromolecules, which have been known to be present in crystalline phases after the transition from ACC [Meldrum and Colfen, 2008]. However, it has been shown that ACC prepared synthetically in the absence of additives does display calcite–like short range order [Günther *et al.*, 2005]. Moderate supersaturation is required to produce ACC with distinct short range order, in contrast to ACC samples prepared at high supersaturation [Cartwright *et al.*, 2012]. Additives can be used to tailor the type of ACC and the crystal product [Lam *et al.*, 2007]. Gebauer and co-workers termed calcite–like and vaterite–like ACC as proto–calcite and proto–vaterite, respectively, to reflect the amorphous phases which precede the formation of crystal [Gebauer *et al.*, 2010].

The structure of ACC has been investigated by computer simulation. From MD simulatons Tribello *et al.* suggested that ACC is formed from sub-domains of vaterite and aragonite [Tribello *et al.*, 2009]. However, the force field used to conduct the simulations has since been shown [Raiteri *et al.*, 2010] to be unreliable. Goodwin *et al.* used reverse Monte Carlo (RMC) techniques to fit a computational model with chemical composition $1620(CaCO_3 \cdot H_2O)$ to experimentally measured scattering data (further constraints had to be implemented to avoid close ion contact) [Goodwin *et al.*, 2010]. The resulting structure had a wide distribution of coordination, with average Ca–O coordination of 5.3–5.8. The coordination number of biogenic ACC can vary widely, but the average distance of the first coordination shell (measured up to 2.8 Å) found from RMC was slightly larger than those measured in biogenic samples [Addadi *et al.*, 2003; Neumann and Epple, 2007]. Most interestingly, charge separation was found, with carbonate and water–rich channels within a calcium–rich framework. The carbonate and water formed hydrogen bonds, and so two types of water and carbonate molecules were identified: those in chan-

nels and in the framework. The observation of a heterogeneous structure could have important implications for the stability of ACC.

Subsequent investigations by Singer *et al.* measured coordination distances much closer to those found for biogenic ACC and other synthetically prepared samples [Singer *et al.*, 2012]. From the RMC result of Goodwin *et al.*, MD simulations were performed with the force field of Raiteri *et al.* [Raiteri and Gale, 2010]. Structural reorganisation was observed over a 2 ns trajectory, with channels decreasing in size. The result suggests that with the adopted force field, the RMC structure is unstable. In a subsequent article [Saharay *et al.*, 2013], the structure of ACC was characterised: average coordination numbers for calcium were larger than the results of RMC suggested, with a maximum value of 7–8 and a density of 2.59 g cm^{-3}, slightly higher than the density of MHC.

Raiteri and Gale produced ACC from annealed melts of calcite [Raiteri and Gale, 2010]. ACC particles containing 18 to 864 CaCO$_3$ units were prepared with hydration levels of H$_2$O/CaCO$_3$=0–3, producing particles up to around 5.5 nm in diameter. The surface of particles was found to be rough and fluctuating, distinct from crystalline faces. The hydration level did not significantly affect the pair distribution functions for calcium and carbonate, while the peak positions and shape of the Ca–Ca radial distribution function did not change. The relative heights between the first and second peak were found to correlate with the hydration level. Heterogeneity or crystal–like short range order was not identified. However, particles were prepared by random addition of internal water, and restructuring of the framework at 300 K over a one nanosecond relaxation window may be hindered due to strong electrostatic interactions.

Tomono *et al.* performed MD simulations of ACC for a range of hydration levels, and which included Mg^{2+} impurities [Tomono *et al.*, 2013]. The force field of Raiteri *et al.* [Raiteri *et al.*, 2010] was adopted with added potential functions for the incorporation of magnesium. They showed that a local ordering of ions similar to the vaterite crystal structure was apparent in anhydrous ACC, and this was reduced by the presence of magnesium. The addition of water did not affect the structure significantly, but large amounts of water led to a local ordering which was comparable to MHC.

1.3.2 Water and Dehydration

Although a great deal of understanding about the dehydration and crystallisation of ACC has been gained over the past decade, there remains little evidence to the mechanisms through which these processes take place. Studies [Aizenberg *et al.*,

2002; Politi *et al.*, 2008; Gong *et al.*, 2012; Lam *et al.*, 2007] have highlighted that the removal of water from ACC precedes the onset of crystallisation. It is therefore possible that water acts to kinetically stabilise ACC and once removed, reconstruction of the ionic framework results in crystal nucleation [Cartwright *et al.*, 2012].

Raiteri and Gale studied the influence of water loss on the energetics of ACC using MD by preparing a range of particle sizes with varying hydration levels, as discussed above [Raiteri and Gale, 2010]. As particle size increased up to a maximum of 864 calcium carbonate formula units, the enthalpy of the system per unit increased, and the enthalpy decreased as a function of increasing hydration for fixed particle size. When plotted against the enthalpies for calcite nanoparticles of equal size (in terms of the number of calcium carbonate units), the data shows that ACC is always lower in energy, except in the case of the largest particles, and this fits with experimental observations that for large clusters, anhydrous ACC is transient. The authors note that the data are consistent with a negative dissolution enthalpy for calcium carbonate. While the data are noisy for small cluster sizes, at around 20 formula units all clusters are shown to be anhydrous. Unfortunately, no discussion of these results is provided.

By using the experimentally reported and calculated entropies of water of inclusion into MHC, Raiteri and Gale were able to estimate the free energies of clusters [Raiteri and Gale, 2010]. The free energies for clusters of the same size showed a minimum as a function of hydration level. By considering these values as a function of size there is a clear trend that for the largest clusters the preferred water content was around $H_2O/CaCO_3=2$, while anhydrous particles were most stable at small sizes. The results indicate that thermodynamically, the hydration level will increase as a function of size. Although heterogeneity was not found in the structure of particles, the authors suggest that if growth rates exceed those for diffusion of structural water, then it is likely that the hydration levels will increase for larger particles.

Saharay and colleagues have also used MD to study ACC dehydration [Saharay *et al.*, 2013; Saharay and James Kirkpatrick, 2014]. By simulating bulk ACC with varying levels of hydration from $CaCO_3 \cdot H_2O$ to anhydrous ACC, they found that the mass density of ACC decreases with hydration level from a value that was comparable to calcite. The density for the highest hydration level was 7 % higher than that for MHC. Loss of water led to changes in ionic coordination shells as calcium carbonate coordination increased due to a reduction in calcium–water binding. The enthalpy of dehydration per water molecule, ΔE_H, was found to decrease as the hydration level decreased. For $CaCO_3 \cdot H_2O$ and anhydrous ACC, $\Delta E_H \approx -66$ and

-85 kJ mol^{-1}, respectively. The authors note that this is lower than the reported ΔE_H for complete dehydration of ACC from experimental measurements [Radha *et al.*, 2010]. However, the values do not account for the possibility that some water is retained within the ionic framework and is only lost upon crystallisation.

Ihli and co-workers have provided a significant contribution to the understanding of ACC dehydration using experiments which mimic biological systems [Ihli *et al.*, 2014]. The transient amorphous $CaCO_3 \cdot H_2O$ in sea urchin spicules is formed in the absence of bulk solution. To mimic this natural system, Ihli *et al.* prepared three types of ACC particles around 100 nm in size, and these were either uncoated or coated (typically a 5–10 nm shell) with porous silica or a phosphatidylcholine-dihexadecyl phosphate (DHP) membrane.

ACC particles were subjected to thermogravimetric analysis (TGA) and differential scanning calorimetry (DSC), with heating up to 800°C in the case of coated particles and 400°C for uncoated particles in incremental 5°C steps. All types of ACC showed systematic dehydration with application of heating cycles and annealing (which can be seen in Figure 1.6 (a)). From a number of TGA experiments, three dehydration regions were identified. For uncoated ACC these were: below 40°C the surface adsorbed water layers were lost; at $40 - 85$°C internal water was lost and the particle contracted in size; at $140 - 260$°C the remaining water was lost, and with continued heating crystal was formed. Figure 1.6 (b) shows the activation energies associated with removing water from uncoated ACC as a function of fraction of water molecules. In general there is an increase in activation energy for removing water as the hydration level decreases. The activation energy increases up to $H_2O/CaCO_3 \approx 1$, following this there is a plateau in the data, suggesting that the removal of a large volume of water in ACC ($H_2O/CaCO_3 \approx 1 \longrightarrow 0.55$) does not require any additional energy. Finally there is a steep increase in energy to 250 kJ mol^{-1}, associated with removing the remaining water necessary for crystallisation.

NMR spectroscopy has been applied to study ACC. In separate studies, Nebel *et al.* [Nebel *et al.*, 2008] and Michel *et al.* [Michel *et al.*, 2008] used ^{1}H NMR to investigate the type and mobility of hydrogen in ACC. Both studies found hydroxide present, and Michel *et al.* suggested that this formed 7 % of the total amount of hydrogen. Very little bicarbonate was found in the sampled materials. Two types of water were identified: rigid water and mobile water, which showed restricted motion over 0.2 ms [Michel *et al.*, 2008]. Quantifying the motion of water further is difficult with the available techniques. Nonetheless, the presence of mobile water could have important consequences for the dehydration of ACC.
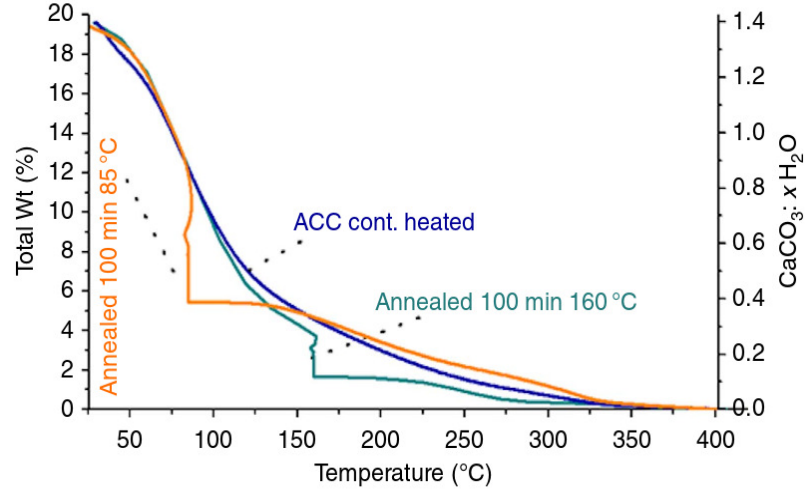
(a)
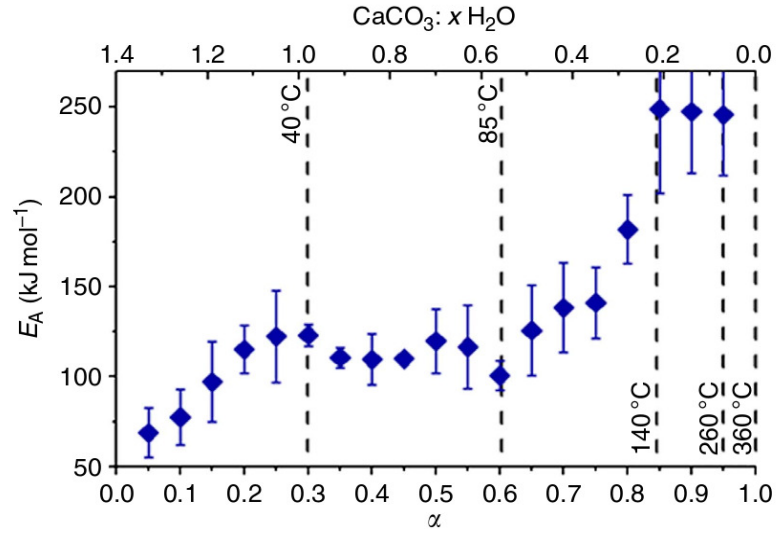
(b)

Figure 1.6: (a) Total weight percent of water in ACC nanoparticles during dehydration as measured from thermogravimetric analysis. (b) Activation energies of water removal from ACC as a function of water content calculated from an Arrhenius relation, $f(\alpha, t) = ln[Af(\alpha)] - \frac{E_a}{RT}$ where $\alpha = \frac{(W_{max} - Wt)}{(W_{max} - W_{min})}$ and W_{max}, W_{min} and W_t are the fractions of water at initial, final and times of measurement. Reprinted from [Ihli *et al.*, 2014].

1.4 Nucleation from Solution

The emergence of solid phases – either crystalline or amorphous – from solution will always involve a phase transition where particles form from constituent atoms/ions/ion pairs/clusters. Over the course of the past decade a surge of research activity [Gebauer and Cölfen, 2011; Gebauer *et al.*, 2014] has been focussed on understanding the species emerging from solution when ions first associate, and any subsequent phase separations. Prior to a discussion of these findings, a brief overview of the physics of phase separation, with emphasis on mineral formation, is provided. The discussion generally follows that provided by Gebauer *et al.* [Gebauer *et al.*, 2014], which should be consulted for further information.

1.4.1 Phase Diagrams

To facilitate the discussion of nucleation, it is helpful to review the concept of phase diagrams. The phase diagram identifies thermodynamically stable phases found at equilibrium as a function of the physical properties of a system. Two dimensional phase diagrams are common, but higher dimension diagrams are also found. A phase boundary marks the conditions where more than one phase can exist in equilibrium (that is, where the free energies of more than one phase are equal), and is sometimes referred to as a coexistence curve or binodal. Phase boundaries are represented as lines separating two phases on a two dimensional phase diagram (*e.g.* the phase diagram for water in pressure–temperature space contains curves separating gaseous, liquid and solid regions). Moving across the phase boundary results in a change in the thermodynamically favoured phase, and ultimately a phase transition. The triple point of substances is a common feature of $p - T$ phase diagrams which identifies the conditions where three phases may coexist at equilibrium, and is marked by the intercept of two phase boundaries. Another common feature are critical points, which mark the conditions where there ceases to be a distinction between phases.

An example of a $p - T$ phase diagram for calcium carbonate at relatively modest conditions is shown in Figure 1.7. The plot shows that calcite is the most stable phase under standard conditions. Only at high pressure does aragonite become more stable than calcite. At low temperature and modest pressures vaterite was identified by Albright as the most stable phase [Albright, 1971]. The triple point on the diagram is highlighted, but this is an estimate as the coexistence curves have been extrapolated. The phase diagram also marks regions of metastability, where the system may spend long time periods in a state higher in energy than the most stable one; perturbing the system will lead to transition to the thermodynamically
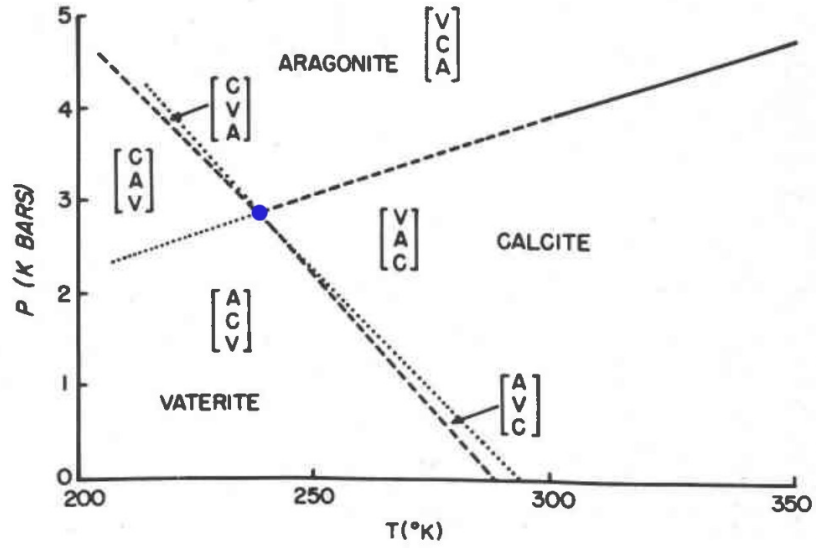
Figure 1.7: $p - T$ phase diagram suggested by Albright for calcium carbonate [Albright, 1971]. The data has been extrapolated. Dashed lines show phase boundaries, while dotted lines highlight metastable equilibria. The relative orders of thermodynamic stability for aragonite (A), vaterite (V) and calcite (C) phases are shown in square brackets, and the blue circle shows the estimated triple point.

favoured state.

Often, nucleation studies are performed in solution at constant pressure, where the amount of solute is increased over time. It is therefore common to exhibit the stabilities of phases as a function of temperature and reaction coordinate, where this could be the amount of solute or ion activity product. Transitions which occur at constant temperature, as the amount of solute in solution increases, can then be followed by crossing lines on the diagram, passing through stability regions (see the following section for examples).

1.4.2 Nucleation Models

Classical Nucleation Theory

If the amount of a solute in solution increases, it is likely that solid will precipitate once the conditions are such that solid phases are thermodynamically more favourable than solution. To link with the phase diagram discussion in the previous section, this is equivalent to passing a coexistence curve on a phase diagram defined by temperature and solution concentration/activity. From the earliest studies, classical nucleation theory (CNT) has been applied to understand the emergence of

solid phases from solution. Becker and Döring provided the first quantitative theory for classical nucleation as we know it today [Becker and Döring, 1935].

CNT stipulates that in a supersaturated solution, the growth of spherical particles is determined by,

$$\Delta G(r) = \frac{4}{3}\pi r^3 \rho \Delta \mu + 4\pi r^2 \gamma, \tag{1.1}$$

where the change in free energy, ΔG, is a function of r, the particle radius. Figure 1.8 shows a general scheme for CNT. The first term in equation 1.1 is the volume bulk free energy, and is shown by the green curve in Figure 1.8, where ρ is the density of the emerging phase and $\Delta \mu$ is the chemical potential difference between the solution and solid phases. The second term is the interfacial energy contribution, given by the red curve in Figure 1.8, and this is a function of γ: the surface tension of the emerging interface. At small r, the interfacial term dominates, but as the particle size becomes large, bulk free energy governs particle growth, resulting in the typical blue curve (Figure 1.8). Density fluctuations in solution lead to the creation of particles, but these will probably re-dissolve unless $r > r_{crit}$, where r_{crit} is the critical size for spontaneous particle growth, and defines the radius where the bulk energy compensates the unfavourable interfacial energy.
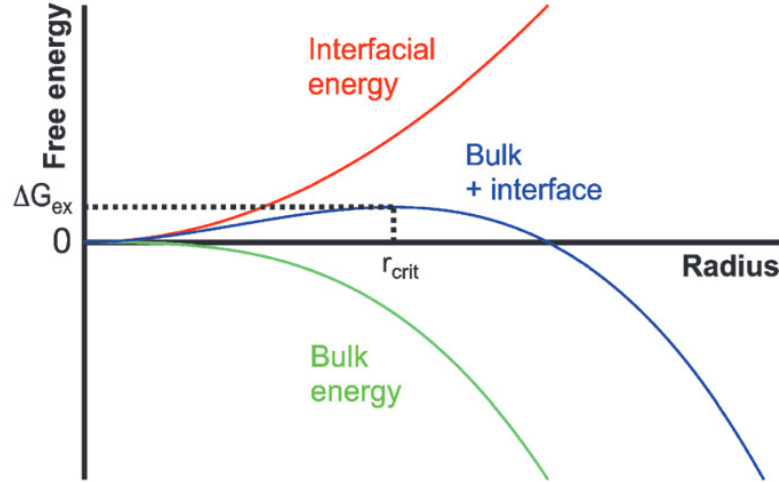
Figure 1.8: Schematic diagram for classical nucleation theory; ΔG_{ex} is the free energy barrier to nucleation and r_{crit} is the nuclei critical radius (if $r > r_{crit}$ then growth is favourable). Taken and adapted from Gebauer *et al.* [Gebauer *et al.*, 2014].

The bulk free energy can be determined by the affinity term, ϕ, provided in equation 1.2. Here, the Boltzmann constant, temperature, ionic activity product and bulk solubility are given by k_{B}, T, IP and K_{sp}, respectively. Crucially, at $IP/K_{sp} = 1$ the solution is saturated. Further addition of solute will reduce the energy barrier to nucleation, ΔG_{ex}, and the value of r_{crit} ($r_{crit} \sim \gamma/\phi$).

$$\phi = k_{\mathrm{B}}T\ln\left(\frac{IP}{K_{sp}}\right) \tag{1.2}$$

At the heart of CNT is the capillary assumption which states that the emerging phase from solution has the same properties and structure as the bulk phase. The values of ρ, $\Delta\mu$, γ and IP/K_{sp} in the CNT equations are therefore often treated as parameters. The Gibbs–Thomson effect (which takes into account changes in interfacial energy with particle shape) can be incorporated into CNT to minimise error associated with spherical particles as opposed to bulk phases which often have well defined planar interfaces. However, at the smallest particle sizes, the shape and structure of particles is often unknown, which makes it difficult to determine these variables. For calcium carbonate, nanoparticles have been found to be amorphous in many experiments [Addadi *et al.*, 2003; Gebauer and Cölfen, 2011], for which K_{sp} is difficult to measure. This in turn can make it difficult to define calcium carbonate nucleation using CNT.

Phase Separation and Spinodal Decomposition

If the IP of solution is increased rapidly, then at a critical level, solid will crash out of solution. The limit of stability is known as the spinodal, and is preceded by an IP range where phase separation is likely to occur given enough time. To best explain this phenomenon Figure 1.9 shows the phase diagram for a generic two phase system. An arrow at constant temperature shows the evolution of the system in the direction of increasing solute concentration. Point A marks the region of composition where a stable homogeneous solution can be found. With increasing IP, the system reaches point B and the phase boundary (binodal), where the chemical potential of solution and solid phases are equal. On passing B, phase separation becomes likely. The metastable region labelled in Figure 1.9 represents a supersaturated solution, and is where CNT can take place. On reaching point C, the solution becomes unstable and no free energy barrier to phase separation exists (which is markedly different from CNT). The system instantly phase separates into solid and liquid. This process is known as spinodal decomposition and is marked by the red spinodal curve on the phase diagram (where $\partial G/\partial IP = 0$).
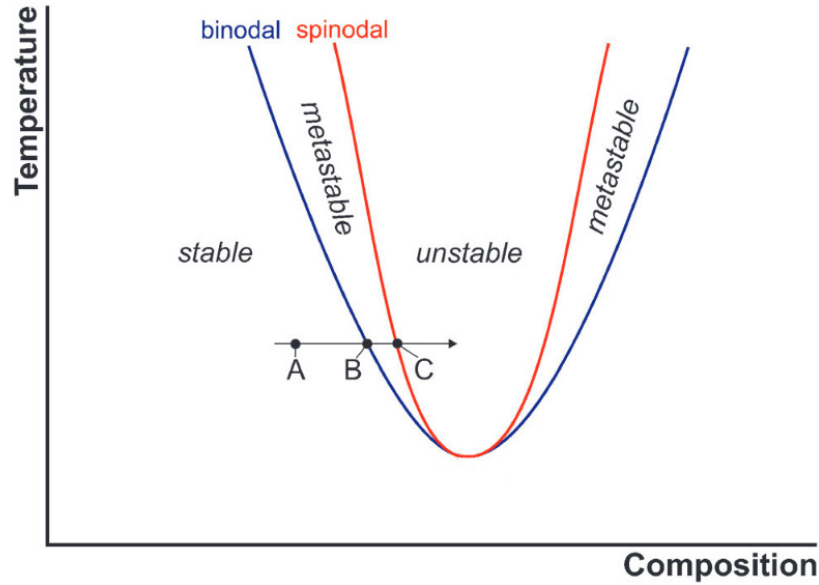
Figure 1.9: A schematic for binodal demixing and spinodal decomposition for a generic system which separates into two phases. Taken and adapted from Gebauer *et al.* [Gebauer *et al.*, 2014].

In the case of calcium carbonate ions in solution, at the binodal curve an equilibrium exists between ions and solid mineral. However, the emerging solid may be of a metastable phase which goes through subsequent transformations (by Ostwald ripening) to the most stable phase. Multiple local equilibria may therefore exist, in which case further curves representing the coexistence of different solid phases can be added to the region of metastability in Figure 1.9. Phase separation in the bionodal region is due to microscopic but large fluctuations in solution composition (*cf.* CNT). However spinodal decomposition occurs because of system–wide, infinitesimal fluctuations in composition. With no free energy barrier, phase separation is diffusion dependent; Ostwald ripening can subsequently lead to reductions in interfacial energy.

As well as liquid–solid phase separation, it is also possible that the homogeneous solution phase could separate into two liquids with high and low solute concentration. With reference to Figure 1.9, a binodal curve on a phase diagram for liquid–liquid separation refers to the coexistence of two liquids. Phase separation may occur by binodal demixing, as was the case in solid–liquid phase separation. The compositions of the two phases will be determined by the binodal curve for a given temperature (presuming temperature is constant). This is also true for systems which reach the spinodal curve when the dense and lean solutions equilibrate.

1.4.3 Experimental Evidence

Prenucleation Clusters

A fundamental observation of calcium carbonate precipitation to this work was made by Gebauer, Völkel and Cölfen during titration experiments of calcium chloride solution into carbonate buffer at constant pH in the range pH=9.0–10.0 [Gebauer *et al.*, 2008]. The authors found a constant difference between the amount of added Ca^{2+} and that which was free in solution before nucleation of ACC, suggesting that calcium was bound in clusters in solution. Importantly, Ca^{2+} was found at all pH (with the amount of binding increasing as a function of pH) and in both under– and super–saturated solution. This observation means that the clusters which form must be thermodynamically stable, and as they form prior to nucleation of solid ACC, were termed prenucleation clusters (PNCs).

Analytical ultracentrifugation (AUC) experiments were carried out to measure the size of clusters. A peak in the data was observed at ~ 2 nm hydrodynamic diameter, d, and the authors state this equates to approximately 70 ions bound in a cluster on average. Larger clusters with $d > 4$ nm were found, and this led to the hypothesis that PNCs aggregate in solution to form ACC. The fact that ACC phases produced at different pH subsequently crystallised to calcite (pH 9.0–9.5) and predominantly vaterite (pH=9.75–10.0) led the authors to suggest that structural preformation takes place at the PNC stage.

Later work by Kellermeier *et al.* employed silica to coat PNCs formed in solution [Kellermeier *et al.*, 2012]. Using cryotransmission electron microscopy (cryo–TEM), clusters in the size range 0.5–3 nm were detected, with the distribution of sizes peaking at 0.6 nm. Dynamic aggregation of PNCs took place as the silica stabilised a colloidal solution. The effect of silica on PNCs and the system in general is largely unknown as the authors acknowledge [Kellermeier *et al.*, 2012].

Pouget *et al.* carried out precipitation experiments in the presence of a stearic acid monolayer [Pouget *et al.*, 2009]. Using cryo–TEM, particles of size 0.6 to 1.1 nm were visualised in freshly prepared solutions. A small proportion of particles < 4 nm in diameter were also found, and later analysis of the solution showed particles were around 30 nm in diameter. This suggests that the particles in solution were PNCs which aggregated. The larger particles were found to persist after nucleation in contrast to the findings of Gebauer *et al.*, and were confirmed to have an amorphous structure, the same as the nucleated solid phase. These particles were in coexistence with larger particles (70–120 nm). Larger particles still were found at the monolayer surface which were suggested to grow by a dissolution–reprecipitation mechanism,

fed by the smaller particles in solution.

Combined, these works create a compelling case for the formation of solid calcium carbonate by PNC aggregation and a non–classical mechanism of growth. The possibility that calcium binding in solution could be in the form of ion–pairs was highlighted by Gebauer and Cölfen in review of the literature [Gebauer and Cölfen, 2011], but they state that structural preformation cannot be encoded into hydrated dimers, and so growth of crystalline–like ACC is unlikely to occur by ion pair addition to solid nuclei.

Liquid Precursors

Liquid–liquid phase separation in calcium carbonate solutions with polyaspartate additive was shown by Gower and Odom [Gower and Odom, 2000]. As the polyelectrolyte led to the phase separation, the mechanism was termed the polymer–induced liquid–precursor (PILP) process. Making use of a glass coverslip, the authors showed that on supersaturation, 2–5 μm liquid droplets phase separated from solution and coated the coverslip before coalescing. Crystal nucleation took place on the film and propagation led to the creation of thin crystals on the surface. A patchwork of crystals with different orientations was found due to multiple nucleation sites and the deposition of liquid droplets which partially merged with the remaining film. PILP was later shown [DiMasi *et al.*, 2002; Jee *et al.*, 2010] for organic molecules native to biology. The composition of the polymer is key to its function, and polymers rich in aspartic acid have proved successful to PILP formation [Schenk *et al.*, 2012].

Recently, Bewernitz *et al.* showed that close to neutral pH and in the absence of additives, a liquid precursor phase to crystallisation is formed in calcium carbonate solution [Bewernitz *et al.*, 2012]. Liquid droplets, hundreds of nanometres in diameter were stable in solution. The liquid was stable at relatively low pH (pH=8.5), leading the authors to state that stabilisation by bicarbonate is taking place. The authors argue that a bicarbonate–rich liquid would consume Ca^{2+} from solution and render any PNC nucleation mechanism less likely. Subsequent nucleation and growth of solid takes place within this liquid phase.

Faatz *et al.* used liquid–liquid phase separation to describe the formation of ACC nanoparticles [Faatz *et al.*, 2004]. The experiment involved a 10 mM calcium chloride solution in which carbon dioxide was released (via hydrolysis of dialkyl carbonate). Initially, a homogeneous solution was stable, but on further release of carbon dioxide, the solution phase separates into ion–rich and –poor solutions. This was verified by the fact that a homogeneous solution was found at very high concentrations (see the phase diagram in Figure 1.9). The authors suggest that

in the high concentration liquid phase, gelation takes place in which water is lost and amorphous solid is formed, and further state that these can act as nuclei for calcification.

Classical Nucleation and Multiple Pathways

A recent report by Hu *et al.* showed the nucleation of calcite in the presence of carboxyl– and hydroxyl–terminated SAMs [Hu *et al.*, 2013]. Nucleation rates were measured using optical microscopy, from which interfacial energies were found. The precipitation of calcite on carboxylated SAMs showed a good fit to a classical pathway for nucleation, with the effect of the SAM reducing the free energy barrier to nucleation. Computational investigations showed that the chain length of the SAM affects the nucleation kinetics. Interestingly, ACC was found in solution in experiments with both types of SAM; ACC particles with sizes of hundreds of nanometres were found before crystal formation on hydroxyl–terminated SAMs and post–nucleation on carboxyl–terminated SAMs, but these played no part in the nucleation of crystal.

Wang *et al.* also found the concomitant precipitation of calcite and ACC (which transformed to vaterite) in solution [Wang *et al.*, 2012]. It is possible that the formation of calcite was classical in this sense, with the additional formation of ACC. The authors suggested that multiple pathways to nucleation may take place; homogeneous nucleation of calcite and secondary nucleation of vaterite from ACC was suggested.

The possibility of multiple pathways to crystal formation has been shown for other biominerals and proposed for calcium carbonate [Baumgartner *et al.*, 2013; De Yoreo, 2013]. According to a number of variables (pH, temperature, solution mixing, *etc.*) it seems possible that there may be several paths to nucleation and growth between solution and crystalline phases. This has led some researchers to try to fit precursor and intermediate phases to these two limits into a phase diagram or free energy path [De Yoreo, 2013; Bewernitz *et al.*, 2012; Gebauer *et al.*, 2014].

1.4.4 Computational Evidence

Force Fields

Any atomistic level study of calcium carbonate nucleation requires the use of a force field which accurately captures the structure of emerging phases, and also the correct energetics of solvation and ion binding equilibria. A number of force fields accurately reproduce the cell parameters of stable anhydrous phases, such as those

of Dove *et al.* [Dove *et al.*, 1992] and Pavese *et al.* [Pavese *et al.*, 1992]. Both of these force fields were shown to match well with experimental measurements for calcite elastic constants and lattice frequencies. Carbonates were treated as a single molecule with angle bending and improper dihedrals.

Unfortunately, these early models failed to accurately reproduce interfacial structures and energetics of crystalline phases with water, which is crucial for studies of biomineralisation. Parker and co–workers [Leeuw and Parker, 2001; Kerisit and Parker, 2004] used a shell model for water, combined with elements of the Pavese potential to model the free energies and structure of water adsorption onto the surface of calcite. Freeman *et al.* took the most successful elements from the Pavese potential and that of de Leeuw and Parker [Leeuw and Parker, 2001], combined with a TIP3P water model [Jorgensen *et al.*, 1983] to derive a force field specifically for biomineralisation studies [Freeman *et al.*, 2007]. They developed a generic method to produce force fields which combine off-the-shelf potentials for organic molecules with pre-existing mineral potentials, in order to enable efficient studies of organic–mineral interfaces.

Raiteri *et al.* highlighted that while previous force fields perform well when modelling crystalline phases, they fail to reproduce the thermodynamics of component ions in solution as well as the free energies differences between crystalline phases of calcium carbonate [Raiteri *et al.*, 2010]. They, therefore, developed a new force field based on rigid carbonates with a number of Buckingham interatomic potentials, and with the rigid four–site TIP4P–Ew water model [Horn *et al.*, 2004]. As well as reproducing the cell parameters and properties of crystalline phases, the force field accurately modelled the solvation enthalpies of ions in solution. The use of rigid carbonates may not be ideal when investigating amorphous phases, and so the force field was later adapted [Raiteri and Gale, 2010] to include flexible carbonate and water, making use of the SPC/Fw potential for water [Wu *et al.*, 2006]. Further adaptation was made by Demichelis *et al.* when bicarbonate ions were included and cross–terms were added to molecular anions in an effort to improve vibrational spectra modelling [Demichelis *et al.*, 2011].

DOLLOP

Demichelis *et al.* performed large free ion solution simulations at concentrations of 0.06, 0.26 and 0.5 M in the pH range 8.5–11.5 [Demichelis *et al.*, 2011]. The pH was defined by the ratio of bicarbonate to carbonate ions in solution, and simulations were conducted for up to 70 ns. Spontaneous aggregation of ions was observed, beyond the expected ion pairs. Large numbers of ions clustered to form chains, rings

and a multitude of cluster shapes. A typical example of a cluster is shown inset in Figure 1.10. At high pH, where carbonate dominates the equilibrium, clusters grew to large sizes and resembled a branched polymer with average Ca–C coordination of two. At lower pH, smaller clusters formed as bicarbonate tended to bind to one cation only, resulting in termination of chains. Cluster size distributions suggest that maximum cluster sizes were ≈ 25 and 220 ions at the extremes of low and high pH. Clusters at all pH levels showed a large amount of conformational freedom and were liquid–like. Furthermore, the cluster sizes both decreased and increased continually throughout the simulations as ions dissolved and aggregated, suggesting that a dynamic (dis)ordering was taking place (where order in this description refers to short range atomic order). The authors therefore described clusters as dynamically ordered liquid–like oxyanion polymers (DOLLOPs).
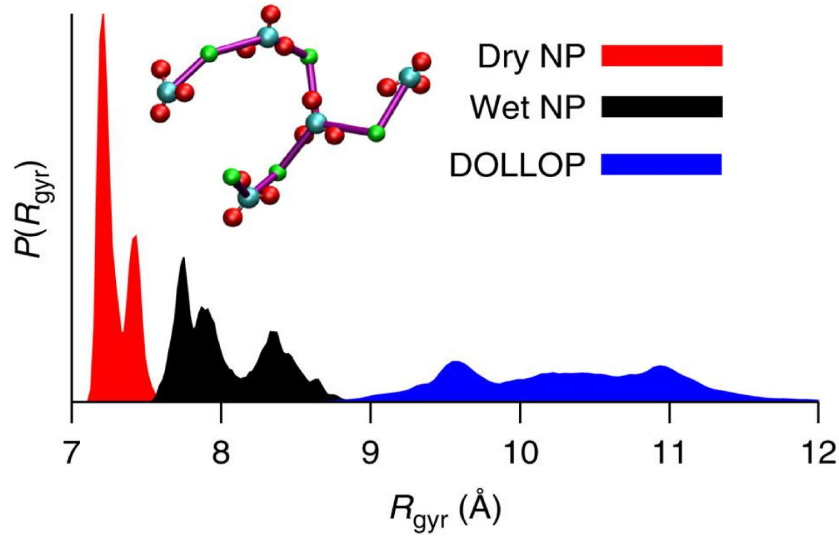
Figure 1.10: Probability distribution for a 36 formula unit calcium carbonate cluster in water, highlighting radius of gyration, R_{gyr}, regions for three states. Inset is an example of calcium carbonate DOLLOP, and here cyan, red and green atoms show carbon, oxygen and calcium, respectively, while purple lines highlight connections between ions. Images were taken from Demichelis *et al.* [Demichelis *et al.*, 2011].

Free energy sampling of a six formula unit DOLLOP as a function of radius of gyration showed that a dense, high coordination cluster was unstable compared with the highly disordered DOLLOP, for which there was no energy barrier to changing conformation or even partially dissolving. The probability distribution for a 36 formula unit cluster as a function of radius of gyration is shown in Figure 1.10. Three regions representing an anhydrous (Dry NP) and hydrous nanoparticle (Wet NP) as well as DOLLOP are clearly identifiable in the distribution. DOLLOP has a much wider distribution of accessible gyration radius, suggesting that the entropy of the calcium carbonate in this state is higher than in solid nanoparticles.

A speciation model was fitted to the data to estimate equilibrium constants and free energies of ion binding in solution. Using the model, the authors were able to show the fraction of calcium bound in ion pairs and DOLLOP ($\sim$70%) was close to those found by Gebauer *et al.* [Gebauer *et al.*, 2008], leading the authors to conclude that PNCs are DOLLOP. The speciation model was fitted to the simulation data at concentrations hundreds of times larger than those of experiment, so to verify the stability of the liquid phase, four small DOLLOPs containing four ion pairs were simulated at 0.4 mM for one nanosecond. Partial disassembly was observed, but as the clusters remained largely unbroken on a time scale much longer than the time for water exchange in the calcium coordination sphere ($\sim$ 80 ps), the stability of DOLLOP was confirmed.

Liquid–Liquid Separation

A separate study of the nucleation of calcium carbonate from constituent ions in solution was made by Wallace *et al.* [Wallace *et al.*, 2013]. As opposed to previous simulations, the authors ensured the concentration of ions was comparable to experiment, with $[Ca^{2+}]$=15 mM. Calcium carbonate cluster growth was simulated using the method of Kawska and Zahn [Kawska *et al.*, 2006]. From an initial solvated cluster (water molecules within 4.5 Å of ions were considered to be solvating), an additional solvated ion pair was placed on a surrounding sphere. With the cluster fixed, relaxation of the solvated ion pair was performed. The new cluster coordinates were fixed in a pre-equilibrated water box to relax solvent molecules for 50 ps. Finally, all atoms were mobilised and replica exchange molecular dynamics (REMD) was performed; eight replicas were simulated in the temperature range 300–400 K for 0.5 ns. The 300 K cluster was extracted, from which further growth was initiated.

Clusters were grown up to a size of around 2 nm and comprised of a maximum of 40 ions. For small clusters, low density configurations were observed, with coordination numbers comparable to those found for DOLLOP. As cluster size

increased, higher densities were observed and compact clusters were found. The diffusion of ions decreased as a function of size, approaching a bulk liquid phase, and the diffusion remained much higher than the diffusion of ions in ACC. The average coordination was found to change from two to three at a size of ~ 26 ions. The free energy of clusters decreased monotonically as a function of increasing size. Wallace *et al.* note that this is indicative of a solution undergoing a spontaneous phase separation by spinodal decomposition, and enables the formation of crystal either by ion–ion or cluster mediated pathways. Making use of an Ising lattice gas model of phase separation, they show that in the dense liquid phase ~ 100 nm clusters are likely to form, while smaller clusters are likely to evolve in undersaturated solutions. The coexistence of small and large clusters seems to fit well with experimental observations [Pouget *et al.*, 2009].

Wallace *et al.* proposed the phase diagram shown in Figure 1.11 [Wallace *et al.*, 2013]. In the diagram, the solubility of all solid polymorphs is shown by the solubility line, SL, and the blue region represents undersaturated solution. Following the addition of ions at constant temperature (green line), a stable homogeneous system crosses the binodal line (L–L) with increasing IP and becomes metastable. On further increase in IP, the system crosses the spinodal curve (SP) and phase separation is spontaneous. T_c marks the lower critical solution temperature, below which no liquid–liquid phase separation occurs.
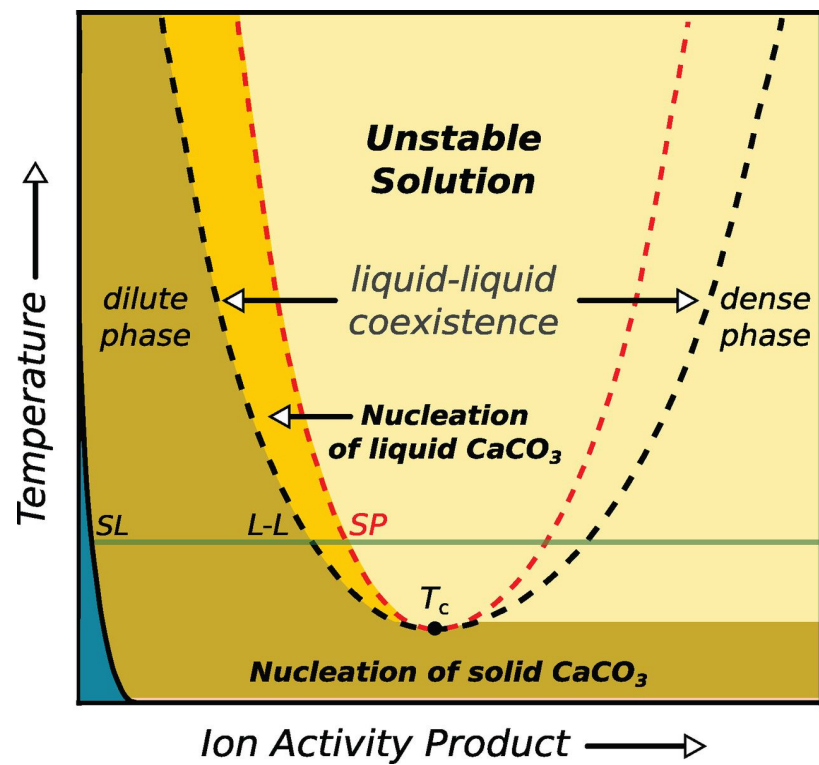
Figure 1.11: A schematic of a suggested phase diagram for hydrated calcium carbonate, as suggested by Wallace *et al.* [Wallace *et al.*, 2013]. For a full description see the text.

1.5 Summary and Outlook

Many of the stages in biomineralisation are yet to be fully understood. The nucleation mechanism is one which is particularly perplexing.

The stability of PNCs has been argued eloquently [Gebauer *et al.*, 2014], but the nature of this phase is hardly understood. The limiting size to stability has been suggested to be on the order of a few nanometres, yet Pouget *et al.* find particles up to 30 nm both before and after nucleation [Pouget *et al.*, 2009]. The term prenucleation therefore seems to be inappropriate from these results. Furthermore, there is a lack of convincing evidence to suggest that PNCs are directly involved in the formation of solid ACC. ACC formation has been described both in terms of PNC aggregation and liquid–liquid phase separation, which further makes the precipitation mechanism unclear.

DOLLOP has been proposed as the structure for PNCs. No barrier to the growth of DOLLOP was found, with further addition of ions being diffusion limited only. At low pH bicarbonate will inhibit the growth of DOLLOP, and the cluster size distributions provided by Demichelis *et al.* [Demichelis *et al.*, 2011] match well with those of PNCs from Kellermeier *et al.* [Kellermeier *et al.*, 2012], but these experiments were conducted at different pH and concentrations. Raiteri and Gale showed [Raiteri and Gale, 2010] that for small clusters anhydrous particles are most stable, yet DOLLOP is lower in energy than ACC in the same size regime, which seems contradictory (admittedly a thermodynamic link between DOLLOP and a dense hydrated ACC phase was not provided).

Computational [Wallace *et al.*, 2013] and experimental [Bewernitz *et al.*, 2012] studies have shown that liquid–liquid phase separation is likely to occur at modest supersaturations. Classical nucleation has been observed [Hu *et al.*, 2013] in the presence of dense phases, and so it is possible that the phase diagram provided by Wallace *et al.* [Wallace *et al.*, 2013] accurately represents the true picture for $CaCO_3$ growth.

This thesis aims to shed new light on the structures and stabilities of ion aggregates emerging from solution. The relationships between free ions, ion pairs, DOLLOP, dense liquid phases and hydrated and anhydrous ACC will be discussed. Possible mechanisms for the transformation between these phases will also be presented.

Ihli and co–workers have provided compelling evidence for a number of different types of water molecules in hydrated ACC, with associated changes in dehydration energies [Ihli *et al.*, 2014]. However, the internal structure of ACC remains

ambiguous, and a mechanism of dehydration on the molecular scale is unknown. Computational studies of the structure and mobility of water in hydrated ACC in this thesis will help to explain the mechanisms by which different types of water are expelled from the amorphous phase.

Chapter 2

Methods

2.1 Introduction

Observations of nucleation and material growth processes at the molecular scale can be difficult using experimental techniques, owing to the length scales over which they take place. Molecular models which accurately describe the interactions between atoms can provide information to elucidate these mechanisms [Anwar and Zahn, 2011]. Sampling nucleation in simulation does pose many challenges, due to the time-scales on which nucleation events occur, although this has been done for some systems with standard techniques [Matsumoto *et al.*, 2002; Walsh *et al.*, 2009]. Nonetheless, understanding the dynamics and stabilities of phases emerging from solution can go some way to interpreting the initial formation of material from solution [Xu *et al.*, 2014; Ma, 2014]. Furthermore, sophisticated techniques are available to enhance the sampling of phases separated by high energy barriers, and to collect information on the relative stabilities between them [Wales, 2003].

There are two main families commonly used in atomistic level simulations: molecular dynamics (MD) and Monte Carlo. In MD, time averages are found by evolving atomic positions according to the equations of motion, while in Monte Carlo, equilibrium averages are assessed according to the probability of accepting randomly generated configurations in a Markovian sequence. In this Chapter the focus is on MD, but many of the concepts discussed are true for both types of atomistic simulations.

The following sections provide a brief introduction into atomistic simulations and the statistical and thermodynamic theories which underpin them. For a more detailed discussion see: *Statistical Mechanics* by McQuarrie, [McQuarrie, 2000] *Understanding Molecular Simulation* by Frenkel and Smit, [Frenkel and Smit, 2001]

and *Computer Simulation of Liquids* by Allen and Tildesley [Allen and Tildesley, 1989]. The DL_POLY suite of codes [Todorov *et al.*, 2006] developed by the *Computational Science and Engineering Department* at the UK *Science and Technology Facilities Council* have been used throughout to run MD simulations.

2.2   Statistical Foundation for Molecular Simulation

Statistical mechanics provides a link between macroscopic variables and an ensemble of microstates. A microstate is a point in a $6N$ dimensional phase space, $\Gamma(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N, \mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N)$, where $\mathbf{r}$ and $\mathbf{p}$ are particle positions and momenta. The phase space probability density, $f(\mathbf{r}^N \mathbf{p}^N)$, satisfies,

$$\int \int \mathrm{d}\mathbf{r}^N \mathrm{d}\mathbf{p}^N f(\mathbf{r}^N \mathbf{p}^N) = 1. \tag{2.1}$$

Each configuration of a molecular simulation containing N atoms is a microstate from which the energy of the system can be calculated via the Hamiltonian, $\mathcal{H}$,

$$\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + \sum_{i=1}^{N} \sum_{j>i}^{N} U(\mathbf{r}_{ij}) + \phi(\mathbf{r}^N), \tag{2.2}$$

where the first term is the kinetic energy, T, (where m_i is the mass of atom i), the second term is the potential energy, U, and the third term represents the influence of an external field if present. The potential energy in molecular simulation is composed of a number of functions which capture the chemistry in the system, as discussed in section 2.3.

Observable properties (*i.e.* $\mathcal{B}$) are averages measured from microstates within the ensemble,

$$\langle \mathcal{B} \rangle = \int \int \mathrm{d}\mathbf{r}^N \mathrm{d}\mathbf{p}^N f(\mathbf{r}^N, \mathbf{p}^N) \mathcal{B}(\mathbf{r}^N, \mathbf{p}^N) \tag{2.3}$$

$$= \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \mathrm{d}\tau \mathcal{B}(\mathbf{r}^N, \mathbf{p}^N). \tag{2.4}$$

In the limit of infinite sampling, ensemble averages are equal to time averages, and so the results from Monte Carlo and MD simulations are equivalent. This is known as the ergodic hypothesis and is a fundamental result in statistical mechanics.

2.2.1 Ensembles

Sampling of the ensemble is performed with a number of constrained thermodynamic variables, which define the type of thermodynamic system. Ensembles include: the microcanonical ensemble (N, V, E), where the number of particles, N, volume, V, and energy, E, are fixed; the canonical ensemble (N, V, T), where particle numbers, volume and temperature, T, are fixed; the isothermal-isobaric ensemble (N, P, T), in which particle numbers, pressure, P, and temperature are fixed; and the grand canonical ensemble (μ, V, T) where the chemical potential, μ, volume and temperature remain constant. Ensembles with constant numbers of atoms work well with MD, while Monte Carlo simulations are often performed to sample within μ, V, T.

N, V, E Ensemble

The microcanonical ensemble relates to an ideal, isolated thermodynamic system. The number of microstates, $\Omega(N, V, E)$, in this ensemble is,

$$\Omega(N, V, E) = \int \int d\mathbf{r}^N d\mathbf{p}^N \delta \left[\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) - E \right], \tag{2.5}$$

and therefore, the probability density of states is,

$$f(\mathbf{r}^N, \mathbf{p}^N) = \frac{1}{\Omega(N, V, E)} \delta \left[\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) - E \right]. \tag{2.6}$$

$\Omega(N, V, E)$ is a normalisation constant and is known as the partition function. This is a particularly important function in statistical mechanics as it offers a complete description of the thermodynamic system. All thermodynamic state functions can be calculated if the partition function is known (*e.g.* entropy, $S = k_{\mathrm{B}} \ln \Omega(N, V, E)$). Unfortunately, the partition function cannot be measured using MD, and so quantities such as entropy and free energy need to be evaluated by other means.

N, V, T Ensemble

The canonical ensemble is a thermodynamically closed system in which heat can be exchanged with an external reservoir. The partition function, $Q(N, V, T)$, is as follows,

$$Q(N, V, T) = \int \int d\mathbf{r}^N d\mathbf{p}^N \exp \left(\frac{-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)}{k_{\mathrm{B}} T} \right). \tag{2.7}$$

$Q(N, V, T)$ can be factorised into ideal and excess parts, where the ideal function is essentially the partition function for an ideal gas and the excess part cannot be solved analytically, but is approximated by a sum over the configurational energies for sampled microstates. The probability density of states in N, V, T is,

$$f(\mathbf{r}^N, \mathbf{p}^N, V) = \frac{1}{Q(N, V, T)} \exp\left(\frac{-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)}{k_{\mathrm{B}}T}\right). \tag{2.8}$$

N, P, T Ensemble

In the isothermal-isobaric ensemble, which has been mostly sampled in simulations conducted in this thesis, the system is coupled to a heat reservoir and homogeneous volume fluctuations are allowed in order to conserve pressure. The ensemble is a superposition of N, V, T with changing V. The partition function is therefore,

$$Q(N, P, T) = \int_0^\infty \mathrm{d}V \int \int \mathrm{d}\mathbf{r}^N \mathrm{d}\mathbf{p}^N \exp\left(\frac{-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) + PV}{k_{\mathrm{B}}T}\right), \tag{2.9}$$

and the probability density of states is,

$$f(\mathbf{r}^N, \mathbf{p}^N, V) = \frac{1}{Q(N, P, T)} \exp\left(\frac{-\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) + PV}{k_{\mathrm{B}}T}\right). \tag{2.10}$$

2.3 The Force Field

While a full electronic description of simulated atoms and molecules is desirable, the computational expense associated with this severely limits the size and time length of MD simulations. Instead, by making use of the Born–Oppenheimer approximation and by removing electrons from the calculation, a much simpler treatment of atoms as charged point masses turns out to be a reasonable approximation to explore molecular scale processes. Interactions between atoms can be modelled using a number of classical physical potential energy functions which are incorporated into the atomistic force field (or model potential). Potential energy functions are generally chosen to accurately model the chemistry, but also to ensure efficient computational sampling of the ensemble. The fitting of the interaction functions can be informed from higher levels of theory and experimental data. The potential energy, U, is calculated from the force (described by the force field) on atom i, $\mathbf{f}_i$, as,

$$\mathbf{f}_i = m_i \ddot{\mathbf{r}}_i = -\nabla_i U(\mathbf{r}_1, \mathbf{r}_1, \ldots, \mathbf{r}_N), \tag{2.11}$$

where $\mathbf{r}$ are the Cartesian coordinates of atoms. The force field can be defined as a sum of the potential energies arising from intra- and inter-molecular interactions and it is helpful to discuss these in turn in the following sections.

There are several off-the-shelf force field packages available, such as CHARMM [MacKerell *et al.*, 1998] and AMBER [Wang *et al.*, 2004], which are commonly used for organic molecule simulations. However, for calcium carbonate studies a number of force fields have been developed specifically to simulate this material [Harding *et al.*, 2008]. Throughout this work the force fields of Gale and co-workers have been used. These potential models are based on one with rigid carbonates [Raiteri *et al.*, 2010], but later adaptations, which have been adopted, combine flexible carbonates and water molecules [Raiteri and Gale, 2010; Demichelis *et al.*, 2011]. These models have been used as they accurately predict the solvation energies of calcium carbonate as discussed in Chapter 1 (section 1.4.4), as well as reproducing the properties of crystalline phases. The force field of Demichelis *et al.* [Demichelis *et al.*, 2011] includes bicarbonate molecules, and so this has been used in simulations where both protonated and deprotonated anionic species were simulated. The energies, structures and dynamics of calcium carbonate systems were rigorously tested with both potential models which did not highlight any differences in assessed properties. The energies in the force fields of Gale and co-workers were defined in units of electronvolts, and so this energy unit was adopted in the current work (1 eV = 96.49 kJ mol^{-1} = 23.06 kcal mol^{-1}).

2.3.1 Intramolecular Potentials

Common molecular features can be treated with two, three or four body potentials which describe, for example, bonds, bond angles and dihedral angles, respectively. Bond potential energies, U_{bnd}, are often modelled using a harmonic potential of the form,

$$U_{bnd} = \frac{1}{2} k_{ij}^b \left(r_{ij} - r_{ij}^0 \right)^2 . \tag{2.12}$$

Here k_{ij}^b is the force constant, while r_{ij} and r_0 are the distance between bonded atoms i and j, and the equilibrium bond distance, respectively. The function is a reasonable approximation as bond stretching close to equilibrium is dominated by the first term from a Taylor expansion of bond distortion. The Morse potential, shown in equation 2.13, is a more accurate model for bond stretching with zero point energy, E_0.

$$U_{bnd} = E_0 \left[\left(1 - \exp\left(-k_{ij} \left(r_{ij} - r_{ij}^0 \right) \right) \right)^2 - 1 \right] \tag{2.13}$$

This potential better models the asymmetries associated bond stretching, particularly far from equilibrium, but requires more computational effort than the harmonic potential.

Bond angles and dihedral angles can be further treated with potentials similar to equation 2.12, where a harmonic spring models the deviation of bond angles from some equilibrium constant. However, it is also common to find cosine potentials, such as in equation 2.14, to model the periodic energy changes associated with rotation between 'cis' and 'trans' conformations. For four connected atoms, $i - j - k - l$, the dihedral angle, ϕ_{ijkl}, is defined as the angle between planes formed from atoms ijk and jkl, which deviates from some equilibrium angle, ϕ_{ijkl}^0.

$$U_{dih} = k_{ijkl}^d \left[1 + cos \left(\phi_{ijkl} - \phi_{ijkl}^0 \right) \right] \tag{2.14}$$

In simulations of calcium carbonate, a further potential function has been introduced by Raiteri and Gale to model the out-of-plane distortion of the carbonate ion [Raiteri and Gale, 2010]. The four body potential takes the form,

$$U_{inv} = Au^2 + Bu^4, \tag{2.15}$$

where u is the displacement of a central atom (C) from a plane of bonded atoms (O–O–O), and the parameters A and B control the stiffness of the molecular plane.

Some atomistic force fields adopt cross-term potentials, which combine molecular features such as bond and angle stretching into a single model potential. Coupling these motions results in a model which better captures the natural dynamics of the system being simulated. Such a term was introduced into one of the force fields used in this work in order to reproduce the phonon spectrum of calcite, and is known as the compass potential,

$$U_{jik} = A \left(r_{ij} - r_{ij}^0 \right) \left(r_{ik} - r_{ik}^0 \right) + \left(\theta_{jik} - \theta_{jik}^0 \right) \left(B \left(r_{ij} - r_{ij}^0 \right) + C \left(r_{ik} - r_{ik}^0 \right) \right). \tag{2.16}$$

The potential combines stretching between atoms $i - j$ and $i - k$ with the associated angular displacement, θ_{jik}; A, B and C are spring constants, and r^0 and θ^0 are equilibrium bond lengths and angles, respectively.

2.3.2 Intermolecular Potentials

The short range forces between non-bonded atoms are due to interacting electron densities. Pauli repulsion dominates at small distances where electron orbitals overlap, yet as the distance between atoms increases, van der Waals forces dominate.

The Lennard-Jones potential is a two-body potential which is an efficient model for short range forces:

$$U_{LJ} = 4\varepsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right]. \tag{2.17}$$

Here, ε is the strength of the attraction between atoms i and j, and σ is the atom separation distance where the potential energy is zero. The r^{-12} term prevents atom overlap, and the attractive r^{-6} dispersive interaction approaches zero as r increases.

While there is added computational efficiency associated with the Lennard-Jones potential, a more accurate description of Pauli repulsion is often desirable, particularly for dense material simulations. The Buckingham potential,

$$U_{Buck} = A\exp\left(-\frac{r_{ij}}{\rho} \right) - \frac{C}{r_{ij}^{6}}, \tag{2.18}$$

offers this, with an exponential term modelling the decay in electronic density surrounding nuclei. In the below equation, A and C are constants which incorporate ionisation potentials and polarisabilities, while ρ is an ionic distance parameter. One has to be careful when using this potential as there is a sharp decrease in energy at small r_{ij}, which can lead to fusion of nuclei if the repulsive part of the potential is not sufficiently large.

Electrostatic interactions between charged atoms are the long range forces affecting non-bonded atoms. Coulomb's law provides a potential function between pairs of point charges q_i and q_j, where the permittivity of free space and the medium are given by ε_0 and ε_m:

$$U_{Coul} = \frac{q_i q_j}{4\pi\varepsilon_0\varepsilon_m r_{ij}}. \tag{2.19}$$

While it is not modelled in the force fields adopted in this study, polarisability is a common feature of atomistic simulations, particularly those involving ions. The polarisability of an atom in an electric field, $\mathbf{E}$, is defined by $\alpha = \boldsymbol{\mu}/\mathbf{E}$, where $\boldsymbol{\mu}$ is the induced dipole moment. DL_POLY treats polarisability using the shell model [Mitchell and Fincham, 1993]. Here, the total charge on an atom is split between a point mass representing the atomic nucleus and a massless shell which are connected by a spring. Any intermolecular interactions are performed using the shell, while the atomic core and shell do not interact electrostatically. Within an electric field, the atom core and shell are separated and $\alpha = q_s^2/k$, where q_s is the shell charge and k is the spring constant.

Computational Considerations

Often in atomistic simulations, the user is interested in studying extended systems such as a bulk liquid or solid, or perhaps, processes occurring at the surface of a crystal. In order to do this, the simulation cell configuration can be repeated in each of the three Euclidean dimensions, and in this work, three dimensional periodicity of simulation cells was used. One caveat to doing this is that atoms may interact with their periodic images, and correlated atomic motion can arise in simulation following the calculation of intermolecular forces.

Calculation of short range forces only needs to be conducted up to atomic distances of around 3σ. By truncating the potential around this distance, computational resources are not wasted. Furthermore, by cutting off the interactions at a distance $< L/2$ (where L is the smallest simulation cell length) calculation of atoms with periodic images is avoided. However, to circumvent unfavourable discontinuities, in this work a tapering function, as detailed in the force field [Raiteri *et al.*, 2010], has been utilised to scale the potential energy between the upper and lower limits, r_c and r_t, respectively:

$$f(x) = \left(1 + 3x + 6x^2\right)\left(1 - x\right)^3 ; \qquad x = \frac{r_{ij} - r_t}{r_c - r_t}. \tag{2.20}$$

The long range interaction of point charges in a full, periodic, extended system must be calculated in order to obtain the potential energy (the Coulombic potential decays as r^{-1}). As the Coulombic sum is conditionally convergent, truncating the potential would introduce non-physical results. Ewald summation offers an efficient alternative by separating the potential into short and long range parts. The technique involves surrounding point charges with a diffuse Gaussian shaped cloud of equal but opposite charge. At long range charges are neutralised, and short range forces due to electrostatics are calculated up to a truncation distance; this sum is convergent in real space. Additional Gaussian-shaped charges are superimposed to nullify the additional charge in the system. The potential from these compensatory charges can be calculated in reciprocal space using a Fourier series. Equations 2.21 and 2.22 show the real- and reciprocal-space (running over **k** wavevectors) terms to the sum. An error function must be added to the real-space part which removes the effect of compensatory charges (α is the shielding width). In this work, a smooth particle-mesh version of the sum has been used which interpolates the contribution of charges to a regular grid and greatly speeds up the calculation [Essmann *et al.*, 1995].

$$U_{r-space} = \sum_{i=1} \sum_{j>1} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} f(\alpha r_{ij}) \tag{2.21}$$

$$U_{k-space} = \frac{1}{2V\varepsilon_0} \sum_{\mathbf{k}} \frac{e^{-k^2/4\alpha^2}}{k^2} \left| \sum_{j} q_j e^{-i\mathbf{k}\cdot\mathbf{r}_j} \right|^2 \tag{2.22}$$

2.4 Molecular Dynamics

2.4.1 Integrating the Equations of Motion

Through equation 2.11, the force field provides the forces on atoms from which the velocity can be calculated. The trajectory of atoms can be integrated iteratively by making use of a finite timestep, τ, and evolving the positions of atoms by a multiple number of steps. Verlet integration is the most common method to evolve atomic positions in time, which results from a Taylor expansion of $\mathbf{r}(t + \tau)$:

$$\mathbf{r}(t + \tau) = 2\mathbf{r}(t) - \mathbf{r}(t - \tau) + \tau^2 \ddot{\mathbf{r}}(t) + O\tau^4. \tag{2.23}$$

Verlet integration is advantageous as it requires a single evaluation of the forces per iteration: this is the most expensive part of any MD simulation. Furthermore, the algorithm is time reversible and quite simple to implement. The uncertainty in the new positions is on the order τ^4. One caveat in the standard Verlet algorithm is that evaluating atomic velocities across two timesteps can introduce numerical errors, and so alternative algorithms have been put forward. The two most common are velocity Verlet and the Verlet leapfrog algorithms, the latter of which has been used in simulations in this work. During leapfrog Verlet, the velocities at half timesteps are calculated (see equation 2.24). The half timestep advancement of velocities are then used in equation 2.25 to evolve the atomic coordinates by a full timestep. The velocity at the current step can be evaluated from equation 2.26. Both the potential and kinetic energy can be calculated during each iteration of the timestep.

$$\mathbf{v}\left(t + \frac{\tau}{2}\right) = \mathbf{v}\left(t - \frac{\tau}{2}\right) + \ddot{\mathbf{r}}(t)\tau \qquad \mathbf{v}\left(t - \frac{\tau}{2}\right) = \frac{\mathbf{r}(t) - \mathbf{r}(t - \tau)}{\tau} \tag{2.24}$$

$$\mathbf{r}(t + \tau) = \mathbf{r}(t) + \mathbf{v}\left(t + \frac{\tau}{2}\right)\tau \tag{2.25}$$

43

$$\mathbf{v}(t) = \frac{1}{2}\left[\mathbf{v}\left(t - \frac{\tau}{2}\right) + \mathbf{v}\left(t + \frac{\tau}{2}\right)\right] \tag{2.26}$$

The choice of timestep deserves some consideration. Ideally a large timestep would be used to generate trajectories which overlap with experiment. However, the timestep must be small enough to accurately evaluate the fastest motions in the simulation, and thereby avoid large forces which could arise from the close contact of nuclei. Bond vibration periods are on the order of femtoseconds, and a 1 fs timestep is commonly used in MD. With such a small timestep, MD simulations of 10^6 atoms typically run for tens of nanoseconds; this should be sufficient time for most systems to establish an equilibrium from which averages can be calculated. It is possible to use multiple timesteps in a simulation. At $t = 0$, the initial velocities on atoms are calculated from the kinetic energy which is a function of the input temperature; velocities are distributed to atoms according to the Boltzmann distribution.

2.4.2 Thermostats and Barostats

Experiments are often conducted at constant temperature and pressure. In order to do this in simulations, thermostats and barostats must be employed. A number of techniques are available to control temperature and pressure, and in this work, Nosé-Hoover control has been utilised.

Nosé-Hoover thermostats are available in most MD packages which offer an extended Hamiltonian approach to conserving temperature in a simulation [Hoover, 1985]. The velocities of particles, $\mathbf{v}(t)$, from the Nosé-Hoover equations are,

$$\frac{\mathrm{d}\mathbf{v}(t)}{\mathrm{d}(t)} = \frac{\mathbf{f}(t)}{m} - \chi(t)\mathbf{v}(t), \tag{2.27}$$

where $\mathbf{f}(t)$ are the forces on atoms and $\chi(t)$ is a friction coefficient,

$$\frac{\mathrm{d}\chi(t)}{\mathrm{d}(t)} = \frac{N_d k_{\mathrm{B}}(T(t) - T_c)}{Z}. \tag{2.28}$$

where N_d are the number of degrees of freedom. The input parameter for temperature is T_c, while Z provides the magnitude of the coupling of the system to a thermal reservoir, s,

$$Z = N_d k_{\mathrm{B}} T_c \tau_T^2, \tag{2.29}$$

where τ_T is the relaxation time for the coupling. The algorithm is implemented iteratively as the velocities are first required to obtain $T(t)$. It is therefore more expensive

than simulations conducted at constant energy, but unlike other thermostats, Nosé-Hoover dynamics implemented with constant volume do lead to average properties in the canonical ensemble.

Pressure conservation is particularly important when trying to map the phase diagram of a material, as this often requires simulating with both constant temperature and constant pressure, making use of NPT sampling. In order to ensure a constant value for pressure, the simulation cell volume must be allowed to fluctuate. The Hoover barostat coupled with a Nosé-Hoover thermostat provides a technique to regulate volume fluctuation for a set pressure, P_c [Melchionna *et al.*, 1993]. Several versions of the algorithm are available according to the type of periodic boundaries used in simulation. In this work, three dimensional boundaries have been used with cubic simulation cells, and so isotropic volume fluctuations are discussed. The new equations of motion, as implemented in DL_POLY are:

$$\frac{d\mathbf{r}(t)}{d(t)} = \mathbf{v}(t) + \eta(\mathbf{r}(t) - R_0); \tag{2.30}$$

$$\frac{d\mathbf{v}(t)}{d(t)} = \frac{\mathbf{f}(t)}{m} - [\chi(t) + \eta(t)]\,\mathbf{v}(t); \tag{2.31}$$

$$\frac{d\chi(t)}{d(t)} = \frac{1}{Z}\left[N_d k_{\mathrm{B}}(T(t) - T_c) + (M\eta(t)^2 - k_{\mathrm{B}}T_c)\right]; \tag{2.32}$$

$$\frac{d\eta(t)}{d(t)} = \frac{3V(t)\left[P(t) - P_c\right]}{M} - \chi(t)\eta(t); \tag{2.33}$$

$$\frac{dV(t)}{d(t)} = 3\eta(t)V(t). \tag{2.34}$$

In the above equations, the magnitude of the coupling of atoms to the barostat is provided by $M = N_d k_{\mathrm{B}}T_c^2$ (where τ_P is the relaxation time for pressure coupling), in a similar way that coupling to the thermostat is given by Z. $P(t)$ and P_c are the instantaneous and desired pressures, η is the dynamical friction coefficient for the barostat and R_0 is the centre of mass for atoms in the cell. The extended Hamiltonian for the system in NPT simulations is,

$$\mathcal{H} = T + U + P_c V(t) + \frac{1}{2}Z\chi(t)^2 + \frac{1}{2}M\eta(t)^2 + \int_0^t ds\left(\frac{Z}{\tau_T^2}\chi(s) + k_{\mathrm{B}}T_c\right). \tag{2.35}$$

2.4.3 Equilibrium Averages and Experiment

The time required to establish an equilibrium in simulation depends completely on the system of interest, and slow correlated modes of motion, as are often found in protein simulations, might mean that insufficient MD configurations can be sampled to reach equilibrium.

For state functions, establishing whether equilibrium has been reached can simply be a case of monitoring the time dependence of the property of interest; a plateau in the time dependence of a state function indicates that equilibrium has been reached. Once the simulation has achieved equilibrium, averages of other properties can be measured. However, it is not always clear which portions of trajectories should be used to measure averages. Correlation functions can be used to measure the correlation of a property, $\mathcal{B}$, at different times in the trajectory. From a time origin, t_0, the time auto-correlation function, $C_{\mathcal{B}}(t)$,

$$C_{\mathcal{B}}(t) = \langle \mathcal{B}(t_0)\mathcal{B}(t_0 + t) \rangle . \tag{2.36}$$

Equation 2.37 shows how the correlation time, $\tau_{\mathcal{B}}$, can be calculated. To ensure good statistical sampling, the time window used for averaging a property must be much longer than the correlation time, $\tau_{\mathcal{B}}$. The statistical uncertainty of $\mathcal{B}$ is the standard deviation of the property at equilibrium. However, by making use of multiple trajectories or uncorrelated "blocks" of long simulation trajectories, the uncertainty can be reduced to the standard error of the mean: $\bar{\sigma} = \sigma/\sqrt{N}$ (where N is the number of simulations or trajectory blocks).

$$\tau_{\mathcal{B}} = \int_0^\infty \frac{C_{\mathcal{B}}(t)}{C_{\mathcal{B}}(0)} \, \mathrm{d}t \tag{2.37}$$

Correlation functions can also be used to measure time-dependent properties of a system to compare to experiment. For instance, correlation times relate to physical quantities such as bond lifetimes. The velocity auto-correlation function can be used to measure diffusion coefficients, D, and transformation of this function from time to frequency domains provides vibrational data which can be compared to spectroscopic measurements. In this work, D have been measured according to the mean squared displacement of atoms and applying Einstein's relation (see equation 2.38).

$$D = \frac{1}{6} \lim_{t \to \infty} \frac{\mathrm{d}}{\mathrm{d}t} \left\langle |\mathbf{r}(t) - \mathbf{r}(0)|^2 \right\rangle \tag{2.38}$$

Pair distribution functions are often measured in MD simulations as they capture the relative distances between atoms in equilibrium simulations. The radial distribution function (RDF) is,

$$g_{\alpha\beta}(r) = \sum_{i \in \alpha} \sum_{j \in \beta} \frac{V \delta(r - r_{ij})}{4\pi N_\alpha N_\beta r^2 \delta r}, \qquad (2.39)$$

where the function runs over atom types α and β, the total number of which is N_α and N_β respectively. $g_{\alpha\beta}(r)$ is the probability that an atom will be found at a distance r from an atom at the origin. The calculation is performed by running over all atom pairs and calculating the distance between them, building up a histogram of distances which is then normalised according to the density of atoms in radial shells (where the shell volume, $\delta V \simeq 4\pi r^2 \delta r$ with shell radius width δr). The function is important as it provides structural information about the system which is directly related to elastic scattering measurements: structure factors (functions describing the phase and amplitude of waves diffracted by atoms in a condensed phase) can be produced via Fourier Transform of $g_{\alpha\beta}(r)$. Furthermore, integrating the function gives coordination numbers as a function of distance. As the RDF is related to the probability of atomic distributions, thermodynamic properties such as enthalpy and pressure can be calculated using the function, although this is not usually done in MD.

2.5 Advanced Simulation Techniques

2.5.1 Collective Variables

In simulations it is often desirable to promote a system to a particular state. For example, in protein simulations, advancing the conformation of the organic molecule may be necessary, or as is the case in this thesis, promoting a coordination number for ionic associates may be required. In standard equilibrium MD simulations this can be computationally demanding if high energy barriers exist between the equilibrium state and that of the desired one. Advanced simulation techniques offer ways to overcome energy barriers in simulation. There are many techniques available to simulators, and those discussed here are just the tip of the iceberg. In this work, collective variable (CV) techniques have been exploited to evolve systems to a target state. The CV is an order parameter which defines regions of phase space, and so the distribution of the CV for different phases must be well separated. In biomolecule simulations, common CVs are molecule dihedral angles, and radius of

gyration is often biased in conformational sampling of polymers. CVs must be continuous functions as the force on atoms is calculated from the ordinary differential of the potential.

In simulations of condensed phases, the relative distances and distribution around a central atom are often used as CVs. Steinhardt and co-workers provided a set of bond-orientational order parameters which are able to measure these properties in a single CV [Steinhardt $et~al.$, 1983]. Quigley and Rodger have provided a continuous version of the Steinhardt CV, $Q_l^{\alpha\beta}$, for molecular simulations [Quigley and Rodger, 2009]. This has been shown to define liquid and solid phases, and has proved successful to characterise different crystalline polymorphs [Quigley and Rodger, 2008b,a].

Equation 2.40 gives the continuous function for $Q_l^{\alpha\beta}$, where the function runs over "bonds", b, between atom types α and β. l in DL_POLY is either 4 or 6, providing a measure for tetrahedral or cubic ordering. $Y_{lm}(\theta_b, \phi_b)$ are the spherical harmonic functions which define the spatial distributions of β surrounding α. The number of total bonds is N_C, while the total number of α atoms is N_α. $f_c(r_b)$ is a continuous function where r_b is the distance between bonded atoms, and is provided in equation 2.41. This function defines the depth of the coordination sphere around a central atom α. r_1 and r_2 are user defined parameters which take the function smoothly from one to zero.

$$Q_l^{\alpha\beta} = \left[\frac{4\pi}{2l+1} \sum_{m=-l}^{l} \left| \frac{1}{N_C N_\alpha} \sum_{b=1}^{N_b} f_c(r_b) Y_{lm}(\theta_b, \phi_b) \right|^2 \right]^{\frac{1}{2}} \tag{2.40}$$

$$f_c(r_b) = \begin{cases} 1 & r_b \leq r_1 \\ \frac{1}{2} \left\{ cos \left[\frac{\pi(r_b - r_1)}{(r_2 - r_1)} \right] + 1 \right\} & r_1 < r_b \leq r_2 \\ 0 & r_b > r_2 \end{cases} \tag{2.41}$$

2.5.2 Adiabatic Biased Molecular Dynamics

Steered MD offers a way to drag a system, often using a harmonic potential, to the desired value of the CV(s). However, one has to be careful to set the velocity of the CV efficiently to ensure relaxation of the system during the pulling. A more appropriate method for systems which do not have large energy barriers between phases is adiabatic biased molecular dynamics (ABMD). In ABMD, the CV, S, evolves by promoting sampling in new regions of CV space; the evolution of the CV into previously visited regions is biased. No work is performed on the system as no

bias is applied in the direction of the target value of the CV [Marchi and Ballone, 1999]. The bias potential is,

$$
U(t) = \begin{cases} \frac{\alpha}{2}\left[(S(t) - S_0)^2 - S_m(t)\right]^2 & (S(t) - S_0)^2 > S_m(t) \\ 0 & (S(t) - S_0)^2 \le S_m(t), \end{cases} \tag{2.42}
$$

where the target value of the CV is S_0, and $S_m(t)$ is the minimum value of S in the previously visited configurations. ABMD has been used in this thesis to generate configurations of the dissolution of calcium carbonate for Umbrella Sampling calculations.

2.5.3 Umbrella Sampling

As discussed in earlier sections, MD fails to sample low-probability regions of phase space, and so the free energies of phases cannot be calculated; entropy must be taken into account in the calculation of free energy and this is dependent on the probability density of states. Umbrella Sampling (US) allows constraining of the sampling to regions of the CV (known as windows), and so improved sampling can be performed for high energy states [Torrie and Valleau, 1977].

In an unconstrained system, the Landau free energy is given by,

$$
\mathcal{F}(q) = -k_\mathrm{B}T\ln P(q), \tag{2.43}
$$

where q is a general coordinate of the system, and in this case the collective variable(s), and $P(q)$ is the probability distribution. In US, the system is often constrained to a window i using a harmonic potential bias, typically of the form, $W(s^{(i)}, q) = \frac{k_W}{2}(s(q) - s^{(i)})^2$, where $s(q)$ and $s^{(i)}$ are the instantaneous and minimum values of the CV, and k_W is the spring constant. The Hamiltonian for the system is perturbed by the addition of the bias potential $W(s^{(i)}, q)$. By reweighting, the unbiased probability distribution in window i, $P^{(i)}(s^{(i)}, q)$, can be obtained from the biased one, $\mathcal{P}^{(i)}(s^{(i)}, q)$:

$$
P^{(i)}(s^{(i)}, q) = \left\langle \exp(-\beta W(s^{(i)}, q) \right\rangle \mathcal{P}^{(i)}(s^{(i)}, q)\exp(-\beta W(s^{(i)}, q)), \tag{2.44}
$$

where $\beta = 1/k_\mathrm{B}T$.

By constraining sampling to a number of windows along the CV, where the sampled CV distributions overlap between windows, relative probabilities can be

calculated, and therefore free energies as a function of q can be found from post-processing. The weighted histogram analysis method (WHAM) allows for stitching together of probability distributions from windows to produce a continuous free energy surface [Kumar *et al.*, 1992]. In WHAM, a linear combination of $P^{(i)}(s^{(i)}, q)$ is performed,

$$P(q) = \sum_i c_i(q) P^{(i)}(s^{(i)}, q), \tag{2.45}$$

where,

$$c_i(q) = \frac{n_i \exp\left(-\beta W(s^{(i)}, q)\right)}{\left\langle \exp(-\beta W(s^{(i)}, q)) \right\rangle}. \tag{2.46}$$

In equation 2.46, n_i are the number of configurations in window i. The algorithm is iterative, with successive calculation of $\mathcal{F}(q)$ and $P(q)$ until self-consistency is achieved.

2.5.4   Metadynamics

Metadynamics is a collective variable method developed by Laio and Parrinello in which the system is pushed away from regions of sampled CV space [Laio and Parrinello, 2002]. The system is biased to explore new regions of space by overcoming large energy barriers. Several flavours of metadynamics exist, but standard metadynamics has been performed in this thesis. The method works by building a history dependent bias potential which perturbs the Hamiltonian:

$$\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} + \sum_{i=1}^{N} \sum_{j>i}^{N} U(\mathbf{r}_{ij}) + V\left[\mathbf{s}(\mathbf{r}^N), t\right]. \tag{2.47}$$

The bias potential, $V\left[\mathbf{s}^M(\mathbf{r}^N), t\right]$, is a function of the instantaneous values of the collective variable(s), $\mathbf{s}$, and takes the form of a number of Gaussian hills as follows,

$$V\left[\mathbf{s}(\mathbf{r}^N), t\right] = w \sum_{k}^{N_G} \exp\left[\frac{-|\mathbf{s}(k\tau) - \mathbf{s}(t)|^2}{2\delta h^2}\right]. \tag{2.48}$$

The Gaussian weight and width are given by w and h, which are deposited every τ steps and $N_G = int(t/\tau)$. Conceptually, the algorithm works by adding bias to the free energy basin where the system resides, pushing it to higher energy states. When enough bias has been added, the system is able to escape a minimum and explore new regions of CV space. Eventually, the whole of phase space will have been

sampled and the built up bias results in a random walk of the system as a function of the CV(s). The deposition time must be chosen carefully to allow relaxation of the system after the addition of bias. The dynamics of atoms in the system will be adiabatically separated from $\mathbf{s}$ if the CV evolves over a much longer time-scale than atomic spatial motions. The free energy of the system can be calculated from the bias potential:

$$F(\mathbf{s}) = -\lim_{t \to \infty} V\left[\mathbf{s}(\mathbf{r}^N), t\right].$$

(2.49)

Chapter 3

Generating Calcium Carbonate Clusters

Declaration

A selection of the material in the following Chapter has been published in the article: A. R. Finney and P. M. Rodger, Probing the structure and stability of prenucleation clusters of calcium carbonate. Faraday Discussions, 159:47–60, 2012.

3.1   Introduction

While Demichelis *et al.* have established a possible state for liquid-like clusters on the nanoscale, dynamically ordered liquid-like oxyanion polymer (DOLLOP) was formed from the spontaneous aggregation of dispersed ions in solution, and its observation does not prove that there are no, more compact clusters of greater stability [Demichelis *et al.*, 2011]. It is therefore of interest to undertake a comprehensive search of possible pre-nuclear cluster structures to determine the range of (meta)stability that can be exhibited.

Estimates for the size of calcium carbonate prenucleation clusters (PNCs) place them in the limit of 1.1–2 nm [Gebauer *et al.*, 2008] or 0.6–1.1 nm [Pouget *et al.*, 2009] hydrodynamic diameter, according to the measurement technique. The upper limit of this estimate is thought to correspond to a 70 calcium and carbonate ion spherical cluster with a mass density comparable to ACC [Gebauer *et al.*, 2008]. While the hydration level and mass density of PNCs are not yet established, it is sensible to expect PNCs to have a lower mass density than ACC; therefore one could expect a cluster containing $35CaCO_3$ to be in the limit of stability for PNCs.

Later studies showed that a wide distribution of PNC sizes form in solution, with the upper limit being around 3 nm [Kellermeier *et al.*, 2012].

The following Chapter discusses the techniques used to search for possible calcium carbonate clusters up to a size of $40\,CaCO_3$. Clusters have been generated which contain a range of mass densities. The remainder of the work in this thesis makes use of the candidate cluster structures discussed herein.

3.2   Methods

Cluster searching is a well established field within molecular simulation, with a great deal of literature particularly dedicated to Lennard-Jones cluster searching [Wales, 2003]. Exploration of the potential energy surface (PES) for clusters of low energy can provide a route to stable cluster configurations which may not be sampled during a normal molecular dynamics simulation, due to the existence of insurmountable barriers between different states. Surveying of minima within the PES can therefore provide a range of states which can be subsequently investigated for stability. This is particularly true for calcium carbonate clusters in the PNC size range; searching the PES for large clusters can be computationally intensive as the number of minima in the PES increases exponentially with the number of atoms.

3.2.1   Cluster Searching

There are several commonly employed techniques for cluster searching. Simulated annealing is one which is analogous to physical annealing in metals [Woodley and Catlow, 2008]. Starting from a candidate structure, annealing can be carried out using either Monte Carlo (MC) or molecular dynamics (MD). In both cases, re-configuration takes place at a higher temperature, chosen such that energy barriers are easily surpassed. New configurations are accepted if lower in energy, and with probability $e^{-(E_{new}-E_{old})T}$ (where E_{new} and E_{old} are the current and previous configurational energies and T is temperature) if higher in energy for the case of Metropolis MC. The temperature is then slowly reduced to localise the searches in the low energy regions of the PES, and with careful cooling, to the global minimum. However, the method is susceptible to trapping of the system in basins, and the final result can be sensitive to the choice of initial temperature and/or MC move.

Genetic algorithms, a subset of evolutionary algorithms, are a sophisticated way to search for the global minimum. Beginning with a sample of candidate structures, each described in terms of a unique representative bit array, or gene (which could be generated according to the topological features of the cluster), both

"crossover" and "mutations" are applied to a sample of the population configurations. Crossover involves combining the bit arrays of two clusters to produce an "offspring" cluster, while mutations may involve MC moves to perturb the bit array. The process is iterative with energy relaxation steps, and according to the selection criteria (*i.e.* cluster energy), a refinement of the bit arrays is performed until the lowest energy clusters are found. For an in depth discussion of genetic algorithms see the review by Johnston [Johnston, 2003].

Random structure searching (RSS) is perhaps the most simple method to generate candidate clusters of calcium carbonate. It involves optimising the potential energy of a sample of randomly arranged atoms in space to a minimum in the PES. In terms of clusters, long-range attraction and short-rang repulsion means that a large volume of the PES will be for high energy configurations, and so this technique is an efficient route to energy basins. While this method of cluster searching may seem like a naïve route to the global minimum, the method offers advantages in the current exercise.

RSS provides the widest coverage of the PES and can identify states which are far from equilibrium. Furthermore, as highlighted by Pickard and Needs, low energy minima have large volumes, and the fact that the distribution of these decreases as one approaches basin minima, means that RSS can quickly locate low energy structures [Pickard and Needs, 2011]. While not necessarily true for glasses, a general feature of the PES is that low energy basins are expected to be grouped, which in turn are widely separated. RSS can therefore be an efficient route to these subsets of basins in the PES. Making use of the chemistry of the system of interest can greatly decrease the computational cycles required to locate a minimum. For instance, in the case of calcium carbonate it is highly probable that one carbon and three oxygen atoms will be optimised to the configuration of the carbonate ion. Fixing chemical units at the onset of the search is therefore sensible. Finally, in terms of computational cost, the algorithm is simple and fast, and can be executed in an embarrassingly parallel fashion.

3.2.2 Conjugate Gradient Optimisation

It is helpful to consider the optimisation method for RSS. A simple approach would be that of steepest descent. This involves calculating the downhill gradient of the PES from the derivatives of the forces on atoms, and minimising along this line by a distance which could be proportional to the gradient. The algorithm is iterative and further steps of descent are taken until the stationary point is reached, determined by a convergence criterion. However, this can be slow, as at each iteration,

the new direction of the gradient will always be orthogonal to the previous search direction. Further, the minimisation path can be affected by the initial direction of optimisation.

The conjugate gradient (CG) algorithm circumvents these issues by using conjugate search directions at each iteration, providing a more direct route to the minimum of the basin. Schlick provides an introduction to CG and other optimisation schemes [Schlick, 2010]. CG is robust and only requires calculation of the first derivatives of the potential. The basics of the algorithm are as follows. For an initial position, in this case atomic coordinates, r^0, the initial search direction is that of the force direction, and is a steepest descent one: $h^0 = f^0 = -\nabla U(r^0)$. The scheme is then iterated as follows,

$$r^{(i)} = r^{(i-1)} + \alpha h^{(i-1)}. \tag{3.1}$$

$$h^{(i)} = f^{(i)} + \gamma^{(i)} h^{(i-1)} \; ; \qquad \gamma^{(i)} = \frac{|f^{(i)}|^2}{|f^{(i-1)}|^2} \tag{3.2}$$

A move of step size $\alpha h^{(i-1)}$ is made along the search direction (equation 3.1), as one would do in a steepest descent. This is followed by calculation of the new search direction, $h^{(i)}$, and as this is dependent on the force, $f^{(i)}$, it is sensible to make two moves during a single loop in the algorithm. The step size is chosen to be around 10^{-5} Å, and the convergence is reached when the norm of the gradient is sufficiently small.

3.2.3 Producing Clusters

To generate candidate structures for possible prenucleation clusters up to 40 CaCO$_3$ units in size, random structure searches were conducted. Ten thousand initial structures were condensed from random configurations of nCa^{2+} and nCO$_3^{2-}$ in the gas phase using CG optimisation, for each value of n formula units in the range 1–40. All clusters generated were therefore neutral; while it is probable that some clusters will carry a net charge, it is difficult to perform such a large variety of cation–anion combinations. Ions were positioned randomly in a sphere, with a density of 0.04 atoms Å^{-3}. This number density was chosen pragmatically to ensure reasonable optimisation time, but did not limit condensation to spherical clusters: a wide range of aspect ratios was observed amongst the clusters generated for any given range of n. The distance between ions was not restricted: close contact in the random structure was permitted, and this provided negligible extra computational

cost, as high energy configurations were neglected early on in the optimisation by way of an energy check.

To investigate the effect of water, optimisation of ten thousand random configurations was also conducted for hydrated clusters with the composition $n(CaCO_3 \cdot H_2O)$ for $n = 4$, 6, 10, and 15. A 1:1 $CaCO_3$:H_2O stoichiometry was chosen, as previous studies have indicated this to be the composition of stable hydrated ACC [Levi-Kalisman et al., 2002; Neumann and Epple, 2007]. These particular cluster sizes were chosen as they spanned the size range suggested for PNCs [Pouget et al., 2009].

The Raiteri and Gale force field [Raiteri and Gale, 2010] was employed in these studies which includes improper dihedral distortion of the CO_3^{2-} plane and flexible water described by a modified SPC/Fw potential [Wu et al., 2006]. As the optimisations were conducted in vacuum, electrostatics were treated with a Coulombic potential, and all intermolecular interactions were cut off at 70 Å; in the largest optimised clusters, ion–ion separation distances were found up to around 30 Å.

3.3   Results

When describing clusters throughout this Chapter, and indeed, this thesis, unless otherwise stated, density shall refer to the mass density of ionic constituents. Mass density can be inferred from the cation–anion coordination numbers, which gives an indication of the relative amount of ion packing in clusters. Low and high energy will be used to describe the energy of clusters relative to the minimum and maximum of the energy distributions sampled. For example, when referring to clusters optimised in vacuum, the lowest energy cluster will be the one which has the most negative value of energy.

3.3.1   Fingerprinting Structures

Categorising amorphous structures and different structural phases by spectroscopic methods is a common technique in experiment. Hohl et al. used a combination of diffraction techniques and structure prediction software to compare spectroscopic fingerprints of short range atomic order, in an effort to determine the amorphous structure of SiO_x [Hohl et al., 2003]. Bates and co-workers used PDF fingerprints of short- and long-range atomic order, measured from X-ray spectroscopy and computational modelling, to determine defects in the crystalline/amorphous structure of organic compounds during dehydration [Bates et al., 2007]. Short-range atomic

order can also be measured in the optimised calcium carbonate clusters, and so fingerprinting has been performed in the current study.

From the random structure searches, a distribution of cluster energies per formula unit of calcium carbonate was found. In all cases, low energy clusters had high density and small surface area, while the opposite was true for high energy clusters. In order to characterise cluster types, fingerprinting was performed by means of a structural analysis. Radial distribution functions (RDFs) for carbon and calcium atoms, $g(r)_{(C-Ca)}$, were calculated for each cluster. Gaussian smoothing of the functions was performed, the degree of which was informed from atom vibrations at 300 K and 1 atm (as measured from the Fourier transform of velocity autocorrelation functions).

A comparison of structures was carried out by calculating the χ^2 statistic to quantify the similarity between RDF fingerprints, as a way to locate distinct minima in the PES. The form of the function was,

$$\chi^2_{ij} = \sum_k \frac{(g(r)^j - g(r)^i)^2}{g(r)^i}, \qquad (3.3)$$

here, $g(r)^i$ and $g(r)^j$ are the RDFs for Ca–C in cluster configurations i and j, where i is a reference structure, and the sum is calculated over all RDF histogram bins, k. In the algorithm, χ^2 is calculated over all i and j cluster pairs, and subsequent analysis is performed to match clusters if their χ^2 values fall within a specified tolerance, $\chi^2{}_{crit}$. Unfortunately, there is no general choice of the tolerance: too large a value would lead to matching of configurations from very different regions of the PES, whilst an insufficient tolerance would give a large number of clusters. A wide range of tolerances were considered, and the similarity between matches was subsequently tested making use of order parameters (OPs) defining features of the clusters. OPs included local calcium density, carbonate out-of-plane distortion, cluster size along the principal axis, carbonate orientation and particle moment of inertia, I.

To emphasise this point, the number of clusters originating from 4,000 of the lowest energy minimised configurations containing $20\,CaCO_3$ have been calculated for a range of $\chi^2{}_{crit}$ and are shown in Figure 3.1 (a). A pragmatic choice of the tolerance would be 5.75, as shown by the red line. At this value, most of the configurations have been matched (into clusters containing at least two structures), ensuring an accurate representative set of structures from reliable statistical sampling of all minimised configurations. To test the adequacy of the chosen tolerance, Figure 3.1 (b) gives a scatter plot of I as a function of energy, E, for configurations constituting five of the total 18 clusters identified from the χ^2 comparison, with

(a)

(b)
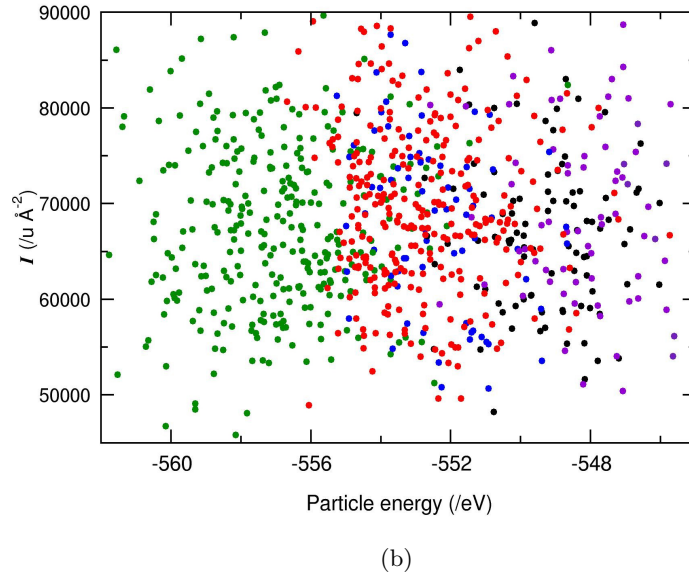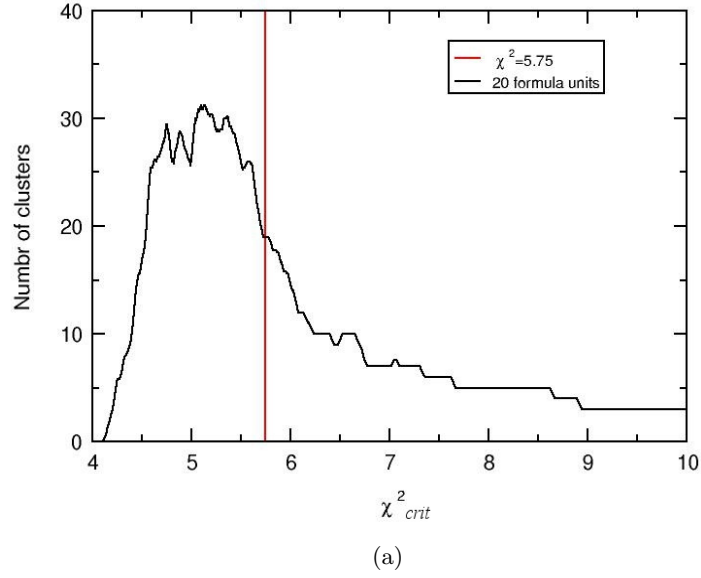
Figure 3.1: (a) The number of clusters of calcium carbonate configurations as a function of χ^2_{crit} (a critical value of χ^2). χ^2 (of $g(r)_{(C-Ca)}$) has been calculated for 4,000 low energy configurations of $20\,CaCO_3$, and structures are matched if the χ^2 value falls below a specified χ^2_{crit}. A cluster here must contain at least two structures. (b) Five cluster types plotted as a function of their moment of inertia, I, and total energy when $\chi^2_{crit} = 5.75$. Colours highlight configurations in the same cluster.

colours used to indicate the members of each cluster. Ideally, the plot should identify distinct regions of the same colour, showing that there are distinct minima in the PES. Instead, a wide distribution is produced, with overlap between the configurations in different clusters, and this was found for all tolerances and sets of OPs tested.

While the fingerprinting technique did not lead to distinct configurations from RSS, it suggests that the PES of calcium carbonate in vacuum up to 40 formula units is reminiscent of a glass, with few well defined 'funnels' that drive the formation of structurally quite different particles. From these studies, it therefore seems that the corresponding free energy surface from the PES described for this particular force field would lack well defined structural phases. This is congruent with the expected energy landscape associated with amorphous phases [Wales, 2003].

3.3.2 Candidate Clusters

In order to classify configurations, global shape OPs were measured for each cluster along with cluster energy, U, and average Ca–C coordination number in the first coordination shell ($r_{Ca-C} < 3.825$ Å), n_{Ca}. Shape OPs included: cluster asphericity, $asph$; cluster acylindricity, $acyl$; cluster radius of gyration, R_g; and cluster relative shape anisotropy, κ^2. These shape OPs have been successfully applied to analyse dendrimers and polymers in computational studies [Bosko et $al.$, 2006]. Functions describing how shape OPs were calculated from the tensor of gyration or clusters are provided in Appendix B.

As indicated in Figure 3.2 (and Figure B.1), a combination of U, n_{Ca} and one shape descriptor (of $asph$, R_g or κ^2) provided a successful set of classes to guide the selection of candidate cluster types. Five candidates were extracted where these were minimum energy structures taken for particular values of n_{Ca} at regular intervals of the total distribution of n_{Ca}. This ensured simultaneous sampling of all three distributions. For many structure sets, the cluster with the highest value of n_{Ca} was also the minimum energy cluster, but for instances where this was not the case, the cluster with lowest energy was sampled.

Figure 3.3 shows a selection of minimised structures for $7\,CaCO_3$ from the ten thousand structures generated, and provides an example of sampled configurations. As might be expected, configurations with lowest n_{Ca} are those of highest energy in the gas phase, while the opposite is true for structures of highest coordination, and in this case density. The lowest energy structure found for this particular composition of cluster was -194.223 eV; the energies are relative to constituent ions in $vacuo$, for which energies will be approaching zero. Between the bounds of lowest to highest
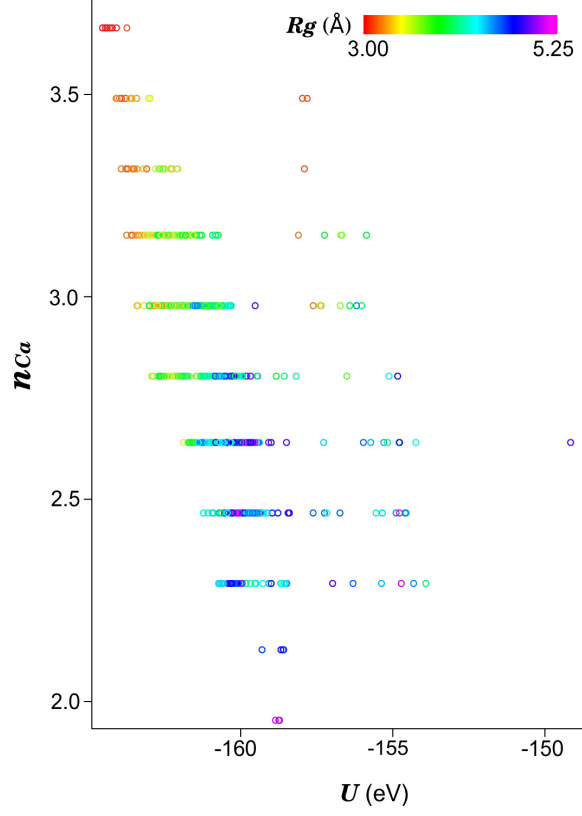
Figure 3.2: A plot of average Ca–C coordination number (n_{Ca}) as a function of potential energy (U) for ten thousand minimised structures of $6\,CaCO_3$ from ions in the gas phase. The points are coloured according to a gradient representing cluster radius of gyration (R_g). The minimum energy data points for each discrete level of n_{Ca} span the n_{Ca}, U and R_g distributions.
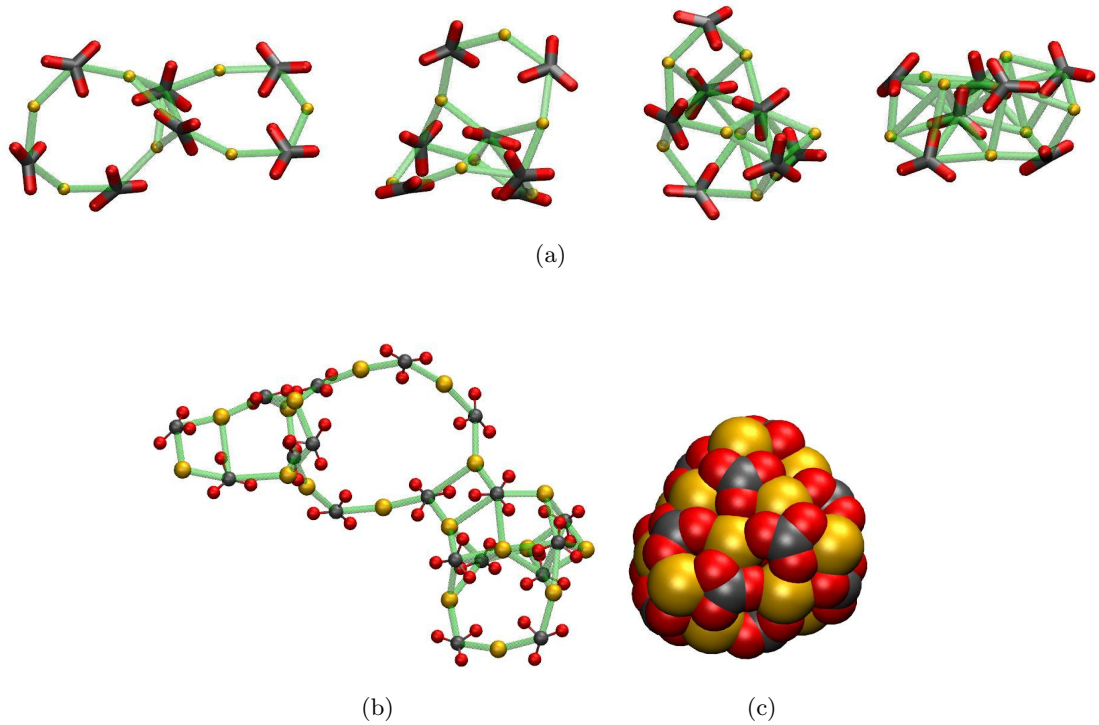
Figure 3.3: Representative configurations from random structure minimisations of (a) $7\,CaCO_3$, and (b,c) $18\,CaCO_3$ from ions in the gas phase. Red, black and yellow represent oxygen, carbon and calcium atoms, respectively, while green lines highlight distances between carbon and calcium atoms below 3.825 Å, indicating ions connected in the first coordination shell. (a) Shows structures with potential energies between $\sim$ -185 to -194 eV which decrease from left to right, while n_{Ca} increases from 2.0 to 4.0. (b) Shows an open $18\,CaCO_3$ cluster with average n_{Ca} of 2.6, and (c) shows a compact $18\,CaCO_3$ cluster with a spacefill representation, and n_{Ca} of 4.6.

energy, the coordination of ions decreased, and this led to structures with average coordination ranging from that shown for DOLLOP in solution [Demichelis *et al.*, 2011] to an anhydrous, amorphous form of calcium carbonate, which are described here as "open" and "compact" clusters, respectively.

Biogenic ACC has Ca–C coordination numbers in the second coordination shell of 2–4 in the range 3.36–3.47 Å, but interestingly, in the samples presented by Addadi and co-workers in their review [Addadi *et al.*, 2003] (lobster carapace, plant cystoliths and ascidian spicule), Ca–Ca coordination ranged from 2–4 in the third shell with distances around 3.8 Å. Michel *et al.* measured the coordination number of Ca–O in the first coordination sphere of synthetic ACC to be 6.1–6.7 [Michel *et al.*, 2008]. The coordination measurements in the first and second coordination

shells for compact clusters in this study corroborate with those measured for both synthetic and biogenic ACC, but there is a shorter Ca–Ca distance in biogenic ACC in the third shell. The Ca–Ca distance in calcite is approximately 4 Å, while in monohydrocalcite this is around 3.8 Å, so it may be the case that internal water leads to shorter Ca–Ca distances [Michel *et al.*, 2008].

Similar coordination motifs were observed in all minimised structures of $n\mathrm{CaCO_3}$. For small n, many minima for low coordination are visited during the random structure optimisation, giving rise to ring and offset–chain type motifs (see Figure B.2). The rings which can be seen in the higher energy clusters of Figure 3.3 (a), were also observed for clusters with larger n, as shown in Figure 3.3 (b). In larger clusters, the average coordination did increase, which is simply an effect of adding more ions. Nonetheless, open clusters were still sampled as n increased. While compact clusters (as in Figure 3.3 (c)) are more favourable in vacuum, kinetic hindrance during optimisation results in sampling of a wide variety of states, and further exemplifies the usefulness of the method adopted.

Water Effects

In the biomineralisation process, ions of $\mathrm{Ca^{2+}}$ and $\mathrm{CO_3^{2-}}$ associate in solution. During this process, water could become kinetically trapped during transformation to a more dense $\mathrm{CaCO_3}$ phase [Tribello *et al.*, 2009]. Indeed, the presence of internal water may stabilise the emerging phase from solution [Raiteri and Gale, 2010]. With this in mind, random structure searches have been conducted for $n(\mathrm{CaCO_3 \cdot H_2O})$ where $n = 4$, 6, 10 and 15.

For these sizes, the structures found for hydrated clusters were similar to those found for anhydrous clusters. There was little variation in the average coordination numbers for any of the cluster sizes. In fact, the minimum energy structures found for both four and six formula unit hydrated clusters were essentially identical to those found in the anhydrous optimisations, and this is shown in Figure 3.4. In these systems, the role of water appeared to be to solvate the particles found from the anhydrous optimisations.

As the cluster size increased, the amount of structures with inclusive water also increased. Water was seen to occupy voids in the ionic framework as well as to solvate ions, as shown in Figure B.3. This is expected, as water will become trapped during minimisation. However, the structures containing large numbers of internal water molecules were generally found to have higher potential energies. For instance, the water molecules in the minimum energy structure found for $10(\mathrm{CaCO_3 \cdot H_2O})$ were located at the surface of the cluster, forming the first solvation shell, as was

(a)   (b)
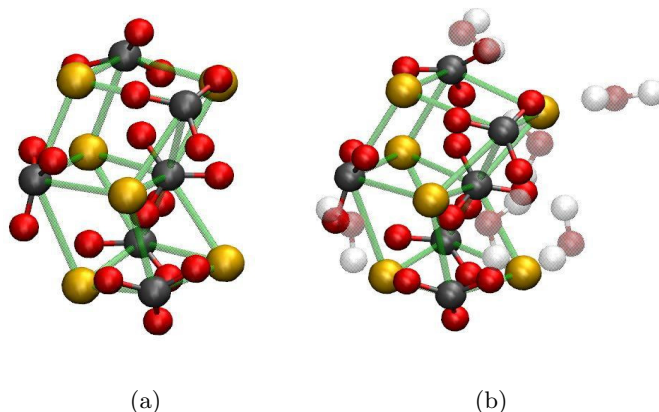
Figure 3.4: Minimum energy clusters found from random structure searches of (a) $6\,CaCO_3$ and (b) $6\,(\,CaCO_3 \cdot H_2O\,)$. For a description of atom colours see Figure 3.3. Water molecules are shown as transparent with oxygen and hydrogen coloured red and white respectively.

the case for four and six unit clusters. It is only for 15 formula units that inclusive water was found within the minimum energy structure. In this system, two water molecules were found embedded in the cluster, completely surrounded by Ca^{2+} and CO_3^{2-} ions; the remaining water molecules again formed the first solvation shell.

For these hydrated systems, the energy of clusters was observed to correlate with n_{Ca}, as shown in Table B.1; hence, water favourably solvated dense clusters, allowing for the maximum coordination of ions below the cluster surface. The result is a structure sample set which is very similar to those found from anhydrous optimisations. It is possible that for the largest cluster sizes (*i.e.* approaching $40\,(\,CaCO_3 \cdot H_2O\,)$) significant numbers of water molecules reside below the surface of low energy clusters. While only 500 optimisations were performed, the lowest energy cluster for $30\,(\,CaCO_3 \cdot H_2O\,)$ did contain three water molecules below the cluster surface, and these were found to form a chain with oxygen atoms separated by around 3.8 Å. Detailed investigation of the statistical likelihood of finding internal water at these cluster sizes becomes difficult, due to the large increase in computational cost associated with increasing system size.

3.3.3   Surface Charge Bias

Clusters minimised from the gas phase, in their native state, have an interesting property in that the distribution of charges within the clusters is non-uniform in many cases (see Figure 3.5). An excess concentration of anions was usually observed

63
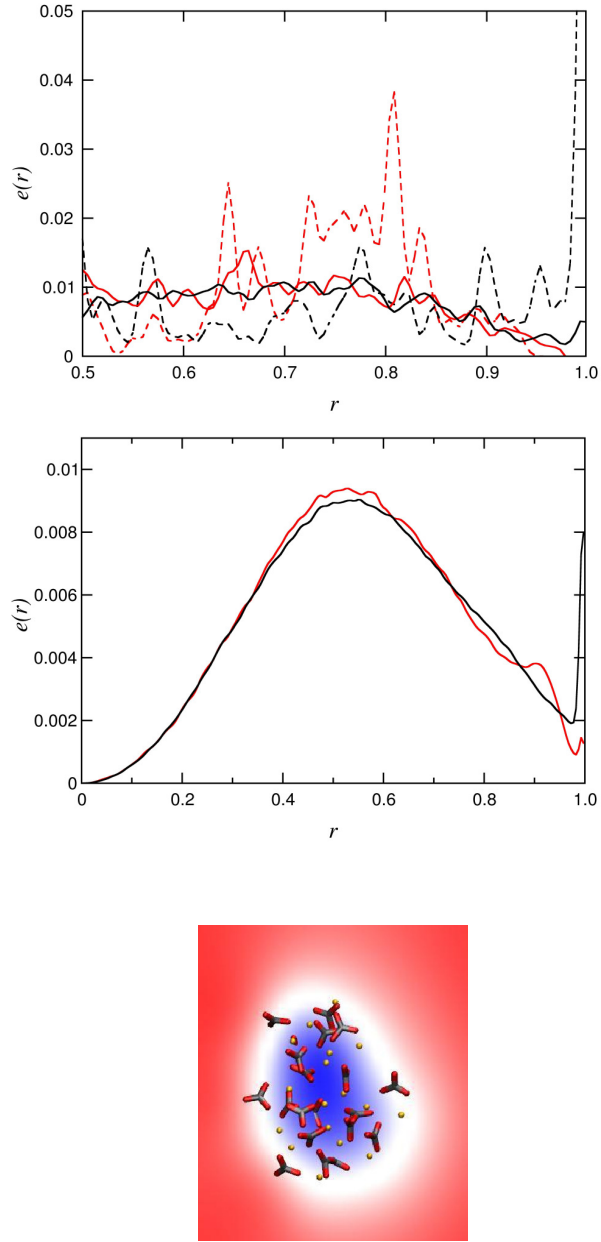
Figure 3.5: Top and middle: ionic charge distributions for minimised anhydrous $n\text{CaCO}_3$ clusters; red lines indicate cationic charge distributions from cluster centres of mass to the surface (normalised to $r = 1$), while black lines are the equivalent anionic distributions. Top: mean ionic charge distributions for minimum energy clusters of $n = 1$–10 (dotted lines) and 31–40 (solid lines) for $n\text{CaCO}_3$. Middle: mean charge distributions for $20\,\text{CaCO}_3$, taken from ten thousand minimised structures. Bottom shows an electrostatic potential map for $27\,\text{CaCO}_3$ taken half way through the simulation cell z axis [Aksimentiev and Schulten, 2005]. Shading represents negative (red) and positive (blue) regions, ranging from -100 – $280\,\frac{k_B T}{e}$.

64

at the surface of clusters, and this was seen in both minimum energy clusters (which usually showed the highest density) and for clusters with lower coordination (albeit to a much lesser extent for low coordination clusters of less than ten $CaCO_3$ units). This can be explained by considering the packing of ions in anhydrous clusters. To achieve a higher density for the cluster, and therefore increased ionic coordination, smaller cationic species, Ca^{2+}, will favourably occupy sites below the cluster surface, with larger anionic species surrounding cations, reducing the potential energy of the system via electrostatic attraction. This will result in carbonate ions more often residing at the surface of the cluster, and hence, introduces the negative surface charge. The data in Figure 3.5 represents structures minimised in vacuum; this will accentuate the surface charge effect as carbonate ions tend to orient with the carbonate and surface planes aligned, in order to maximise coordination.

To further examine the significance of the charge bias, two compact clusters of $9\,CaCO_3$ and $20\,CaCO_3$ were immersed into water with a distance of 15 Å separating cluster centres of mass. Water was relaxed with the cluster atoms frozen before all atoms were mobilised. A series of simulations was performed at 300 K and 1 atm with all other MD input parameters taken as suggested by Raiteri and Gale [Raiteri and Gale, 2010]. A harmonic potential was introduced such that the equilibrium distance between cluster centres was decreased iteratively at a rate of 1 Å per 0.1 ns. This was done quickly to ensure minimal relaxation of the clusters. Figure B.4 shows initial and final states of the cluster from the simulation series. When the cluster separation was around 7–8 Å, deformation of both clusters was observed. n_{Ca} in the smaller cluster decreased, and an open structure resulted, while loss of ions was observed at the surface of the larger cluster. The result suggests that aggregation of dense particles in solution is unfavourable.

3.4   Summary and Conclusions

A rigorous search has been conducted for possible calcium carbonate clusters in the size range expected for prenucleation clusters, which were suggested as the first ion association species in the precipitation of $CaCO_3$ from aqueous solution [Gebauer *et al.*, 2008]. Structures were generated from random starting configurations of Ca^{2+}, CO_3^{2-}, and H_2O in the gas phase, and were condensed to produce configurations for anhydrous and hydrous clusters of calcium carbonate. From the thousands of possible structures available per formula unit of calcium carbonate, cluster analyses have been performed and OPs have been identified which can successfully classify structures, and sample a representative subset of candidate clusters.

It is informative to consider whether any of the clusters show features of crystalline calcium carbonate. While a rigorous comparison of coordination environments has not been performed, the RDFs, as measured from vacuum optimisations, did not show that crystallinity was apparent even in the most compact clusters. There was some periodicity of ion centres of mass which could be studied further to investigate for crystalline-like structure. Rigorous studies by Bano indicated that at these cluster sizes, amorphous clusters are likely to form in vacuum over crystalline ones, and while OPs to classify the difference between crystalline and amorphous states were difficult to pinpoint, clusters with bond ordering indistinguishable to crystalline phases were not found [Bano, 2012]. Fernández-González *et al.* investigated the atomic structure of relaxed configurations of a range of cluster sizes up to $2000CaCO_3$ in vacuum, generated from initial aragonite and calcite crystalline lattices [Fernández-González *et al.*, 2011]. They found that for small sizes (< 14 formula units), only amorphous clusters were found, and for clusters between 15 and 160 formula units, amorphisation of the lattice occurred with some features of the initial crystalline lattice evident in diffraction patterns [Fernández-González *et al.*, 2011].

For all clusters studied, the majority of the more dense clusters were found to have a bias in the charge distribution from cluster centres of mass, due to a large number of anions at the surface of the clusters. This was seen both in minimised structures in the gas phase and for systems that had been relaxed in water. This effect could be a contributing factor to the (meta)stability of small clusters of calcium carbonate in solution, if relatively dense structures (of the size range considered here) form. The effect has been shown for other nanoparticles, such as TiO_2, where surface charge (either net negative or positive according to the solution chemistry) can affect the dispersion of particles in solution, as well as agglomeration and aggregation [Jiang *et al.*, 2009]. Furthermore, surface bound negatively charged ions can lead to a kinetically stabilised suspension of nanoparticles in solution [Özkar and Finke, 2002].

Chapter 4

Simulating Clusters in Aqueous Solution

Declaration

A selection of the material in the following Chapter has been published in the article:
A. R. Finney and P. M. Rodger, Probing the structure and stability of prenucleation clusters of calcium carbonate. Faraday Discussions, 159:47–60, 2012.

In the following Chapter, the force field for aspartate was provided by David J. Sparkes and Dr. Colin L. Freeman working in the group of Professor John H. Harding at the *Department of Materials Science and Engineering, University of Sheffield, UK*.

4.1 Introduction

The following Chapter describes molecular dynamics (MD) investigations into the dynamics and energetics of calcium carbonate clusters in water. A sample of clusters generated, as discussed in Chapter 3, have been immersed into water and simulated with MD. The (meta)stability of clusters is discussed and implications for the nucleation of calcium carbonate are considered.

Proteins and polymers have been found to direct the growth of particular crystalline polymorphs of calcium carbonate [Freeman *et al.*, 2010; Meldrum and Colfen, 2008; Wolf *et al.*, 2011]. Observing the effect of organics on $CaCO_3$ clusters in the size regime of prenucleation clusters (PNCs) may allow explanation of nucleation mechanisms. Of particular interest is the role of aspartate (ASP), which can be found in calcite promoting proteins *in vivo* [Aizenberg *et al.*, 2002; Weiner, 2008]. The

dynamics of ASP inclusive clusters have consequently been studied for comparison to $CaCO_3$ clusters of equivalent size.

4.2 Methods

As discussed in section 3.3.2, five clusters were generated per formula unit of calcium carbonate, n, in the range $n = 1 - 40$, and for $n(CaCO_3 \cdot H_2O)$ where $n = 4$, 6, 10, and 15. In the current Chapter, the effect of ASP on the stability of clusters was considered, and therefore random structure searching was repeated for ten thousand hydrated clusters of the same $H_2O/CaCO_3$ composition, but with one ASP molecule also included for $n = 4$, 6, 10, and 15. The same cluster classification and sampling procedure was adopted as for anhydrous and hydrous clusters. The initial conformation of the amino acid was its minimum energy conformation in water (see Figure C.2). In solution, the pKa of $HCO_3^- \rightleftharpoons CO_3^{2-} + H^+$ is 10.3, which suggests that a highly basic solution is necessary for pure carbonate clusters to form. As such, ASP was chosen to be fully deprotonated (pKa values: α-COOH = 2.1, β-COOH = 3.9, and α-NH$_3^+$ = 9.8). The protonation state of ASP is unknown when bound in solid calcium carbonate, but the selected one is sensible for this study.

As for cluster optimisations, the adopted force field was that of Raiteri and Gale with flexible water [Raiteri and Gale, 2010]. The choice of potential at the mineral – organic interface is crucial to the accuracy of the simulation. As highlighted by Freeman *et al.*, combining off-the-shelf organic potentials with mineral force fields using standard mixing rules may lead to inaccurate atomic interactions at the interface [Freeman *et al.*, 2007]. Ideally, a widely available and reputable organic potential would be used with cross-terms refitted to the Raiteri and Gale potential. However, this is an unnecessarily costly exercise for the current study.

Freeman *et al.* have developed a generalised method for fitting organic potentials to mineral ones, and have demonstrated this to be an acceptable potential development tool in biomineral simulation studies [Freeman *et al.*, 2007]. The method has been adopted in the current study and, in short, involves refitting of the AMBER force field [Wang *et al.*, 2004] with that of Raiteri and Gale [Raiteri and Gale, 2010], such that the Coulombic interaction between equivalent atoms at the mineral–organic interface is consistent. Although this re-parametrisation was conducted with a TIP3P water model [Jorgensen *et al.*, 1983], studies have indicated little structural and energetic changes in peptides when modelled with TIP3P and SPC/Fw water [Desmond *et al.*, 2013].

The handful of candidate structures extracted from cluster analyses were

subsequently simulated with MD, in enough water to ensure solution behaviour (see Appendix C for details). The number of water molecules was kept constant for simulations of the same nCaCO$_3$. The concentration of solutions was on the order of 10 g dm^{-3}. This is much higher than the solubility of calcium carbonate ($\sim$ 0.015 g dm^{-3}), but results in speeding up of dynamical processes to extract accurate information on solution systems with reasonable computational effort.

All systems were simulated in the *NPT* ensemble at 298 K and 1 atm, employing a Nosé–Hoover thermostat and barostat with 0.1 ps and 1.0 ps relaxation times. A short, 20 ps simulation was initially carried out to enable relaxation of water whilst keeping solute atoms fixed in space. A 1 fs timestep was used throughout. Trajectories of 10 ns for minimum energy structures, and 5 ns for other candidate clusters were generated, with all atoms mobile. 50 ns simulations were carried out for anhydrous clusters containing $n = 10, 20, 30$ and 40 (for nCaCO$_3$) for the two clusters at highest and lowest values of Ca–C coordination, n_{Ca}. The cut-off for short range interactions was set to 9.0 Å, with a tapering function applied to the Ca, C and O calcium carbonate atom respective pairwise interactions from 6.0–9.0 Å, as recommended in the chosen force field. Electrostatics were treated using the smooth particle mesh Ewald method [Essmann *et al.*, 1995]. The simulation package used to run the MD simulations was DL_POLY Classic [Smith *et al.*, 2002].

Further simulations were carried out for 20 CaCO$_3$ which had the lowest n_{Ca} at time zero, to observe the effect of temperature on clusters. This particular configuration was chosen as it was the lowest energy cluster (in solution), and of size which fell in the ranges expected for PNCs from experimental measurements [Gebauer *et al.*, 2008; Pouget *et al.*, 2009]. 20 ns trajectories were collected at temperatures of 320 – 500 K in 20 K increments. Aside from temperature, input parameters were taken as for other simulations in this Chapter.

4.3   Results

4.3.1   Compact Clusters in Water

The minimum energy and most compact structures of nCaCO$_3$ clusters (*i.e.* those with initial highest Ca–C coordination) for each n in the range 1–40 were simulated for 10 ns in water, and the mean potential energy, U, taken from the latter regions of the trajectories are shown in Figure 4.1. The underlying trends in the data show a rapid monotonic decrease in potential energy with increasing cluster size for clusters up to around nine formula units. For larger clusters, Figure 4.1 shows some scatter superimposed on a general increase of U with cluster size, up to a plateau around
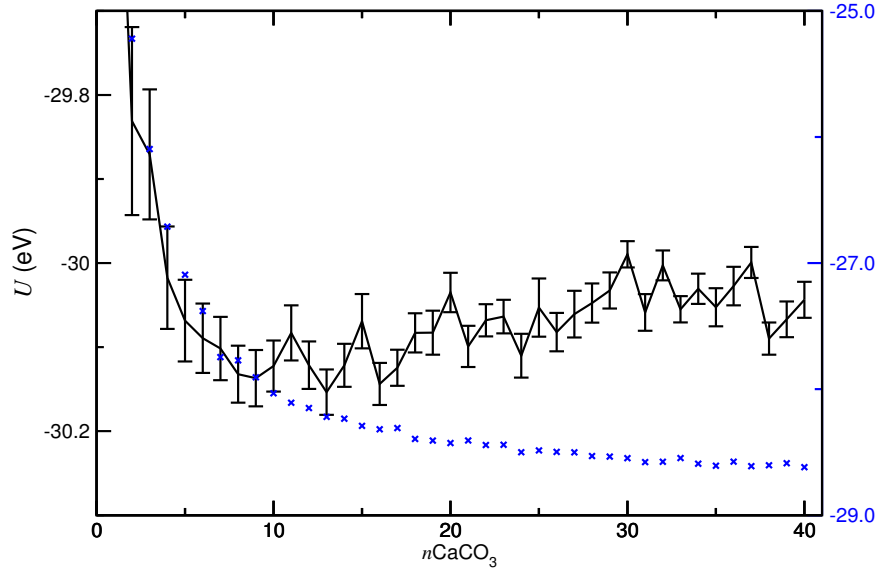
Figure 4.1: Average potential energy, U, per formula unit of calcium carbonate calculated for clusters of initial $n\text{CaCO}_3$. Average values were calculated from the final 2 ns of a 10 ns MD trajectory of clusters simulated in water. The energy values are exclusive of bulk water energies in respective systems. Error bars indicate uncertainties relating to one standard error of the mean. The potential energy of the clusters in the gas phase are also provided and shown by the blue data points, with the shifted scale (in blue) on the right of the graph applicable to these data points only.

$n = 30$. The difference in U between the minimum around $n = 9$ and the largest cluster sizes is on the order of ~ 0.1 eV per CaCO_3 ion pair, which is comparable in strength to a moderately strong hydrogen bond. The general trends observed here contrast with those seen from cluster optimisations in vacuum (shown by the blue curve in Figure 4.1), where extra stability was found from increasing ionic coordination in dense amorphous clusters, up to the largest cluster size considered in this study (~ 2 nm in diameter).

The increase in cluster energy as a function of size can be ascribed to ion hydration in solution, and can be explained by considering the time evolution of n_{Ca} for the various system sizes. Figure 4.2 shows an example of the initial and final states of large and small clusters at the beginning and end of 10 ns simulations. In small clusters ($1 \leq n \leq 10$ for $n\text{CaCO}_3$) nearly every ion was in contact with the solvent, even for particles with maximum n_{Ca}. The presence of the cluster–solvent boundary introduced an energy penalty (*i.e.* the interfacial energy), and the clusters were too small to be stabilised by a favourable bulk energy contribution. Energetically, it was more favourable for the clusters to partially dissolve (see Figure
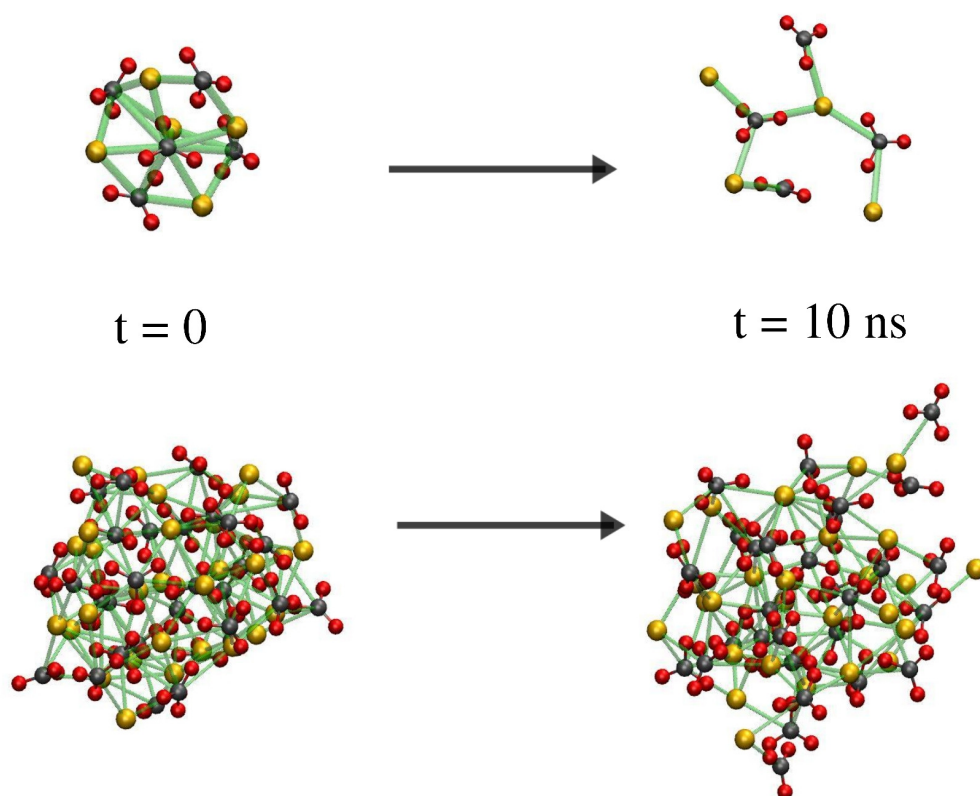
Figure 4.2: Initial and final snapshots of minimum energy calcium carbonate clusters (from vacuum optimisations) simulated in water for 10 ns. Top shows $5\,CaCO_3$ at the beginning of the simulation which partially dissolved to give the $4\,CaCO_3$, low density cluster shown on the right, plus one ion pair which has been omitted. Bottom shows a $32\,CaCO_3$ cluster, for which partial loss of surface ions occurred over the 10 ns simulation, but for which a well defined ionic core was retained. Ca, C and O atoms are shown as yellow, black and red, respectively, while green lines show distances between carbon and calcium below 3.825 Å, highlighting ions connected in the first coordination shell. Water has been omitted for clarity.

4.2 for $5\,CaCO_3$), allowing retention of limited ionic coordination, but also greatly increasing ion stabilisation through solvation. In the case of $9\,CaCO_3$, by $t = 10$ ns, the cluster dissociated and small clusters of ions dynamically dissolved and aggregated during the simulation. This dynamic (dis)ordering was found for all small clusters of initial nCaCO3, with cluster size distributions (CSDs) indicating a wide range of possible cluster sizes after relaxation in water. A particularly interesting case was observed for a six formula unit cluster; this dissolved into solution to form smaller clusters, which underwent dynamic aggregation and dissolution, before a six formula unit cluster was reformed, and a considerable volume of configuration space was explored on the time-scale of the experiment.

As cluster size increased, the stabilisation from high ionic coordination within the cluster was sufficient to avoid complete cluster dissolution over a 10 ns trajectory, as shown in Figure 4.2 for the *in vacuo* minimum energy $32\,CaCO_3$ cluster (CSDs showed little change over the course of the simulation for large clusters). The internal core sub-lattice structure of ions in the cluster over 10 ns was retained for clusters of initial high density with n larger than ~ 20. Surface ions in larger particles coordinated strongly with solvent molecules, and this binding was sufficient to disassemble surface ions, creating dendritic arms and rings which protruded into solution and, on occasion, dissociated into small clusters. Ion pairs observed to dissociate from the cluster surface, yet this was insufficient to "break" the cluster on the time-scale of the simulations. This result appears to contrast from the findings of earlier simulation studies, reporting the structural stability of highly ordered nanoparticles of $CaCO_3$ in water, where calcite nanoparticles (of comparable size to the larger dense clusters in this study) were found to be stabilised by the effect of strongly coordinating surface water [Kerisit and Parker, 2004; Cooke and Elliott, 2007]. Although we have not considered calcite stability at this scale, the general differing result can most probably be ascribed to the choice of force field.

As small clusters partially dissolved in solution, n_{Ca} was observed to decrease over time, resulting in structures with coordination motifs that were found in higher energy clusters from gas phase optimisations. Oligomers of ions formed with solvent stabilisation of ions at the chain ends of dendritic arms. The coordination between monomeric ions in the chain was observed to be dynamic, with ion–ion separation reaching considerable distances, and ion pair loss and recombination prevalent over long time-scales. This behaviour is very much analogous to the that of Ca^{2+}, CO_3^{2-} and HCO_3^- ions in aqueous solution, and is further evidence for the stability of dynamically ordered liquid-like oxyanion polymer (DOLLOP) over more ordered structures in solution [Demichelis *et al.*, 2011].

The radial distribution functions (RDFs) for large clusters which retained high coordination during simulation are provided in Appendix A. These show that compact clusters were amorphous with no long range atomic order. It is not expected that at the sizes studied there is likely to be crystalline order in clusters, but a study to investigate the bond ordering in the core of amorphous clusters would indicate whether atomic coordination environments were comparable to the crystalline phases of calcium carbonate. The RDFs show that for $n > 10$, atomic density is retained throughout the simulation at large r, which is indicative that clusters remained particulate over long simulation times (*i.e.* up to 50 ns). The doublet in the first peak in the Ca–C RDF (see Figure A.1) shows that both monodentate and bidentate binding of carbonate to calcium was apparent in the clusters. Bidentate binding tended to be more probable in compact clusters, which is not surprising as this maximises the electrostatic interaction between ions.

4.3.2   Comparing Open and Compact Clusters

To investigate whether larger cluster sizes are likely to become dynamically ordered structures when immersed into water, clusters of $n\text{CaCO}_3$, but with lower initial n_{Ca}, have also been simulated. As discussed in section 3.3.2, as n_{Ca} decreased over the five sampled clusters (from gas phase optimisations) per formula unit, more open structures were found, and the dissipation of a well defined ionic core was observed. Figure 4.3 shows the potential energy of clusters containing 20–29 formula units, after relaxation in water. It is clear that lower initial n_{Ca} does lead to a solvated cluster with lower potential energy, with the difference being ~ 0.1 eV per CaCO_3 unit. This difference is comparable to the potential energy barrier evident in Figure 4.1 for compact clusters.

The mobility of ions around the centre of mass of clusters with low coordination is higher than for equivalently sized clusters with more dense packing. The final configurations for a 28CaCO_3 cluster have been included in Figure 4.3; these show a much less ordered structure after 5 ns of simulation for a cluster with low initial n_{Ca}, than for the corresponding high coordination structure (simulated from the minimum energy structure in vacuum) after 10 ns. The average U of all five samples considered appears to converge over the range of cluster sizes considered.

Longer simulations (50 ns) were conducted for the most open and compact clusters that were sampled for $n =$10, 20, 30 and 40. The potential energies measured across this range from the final 5 ns of simulation (see Figure C.3), showed the same energy difference between open and compact cluster states. For $n = 10$, the energy of clusters in the limits of low and high initial coordination converged to -32.5 eV.
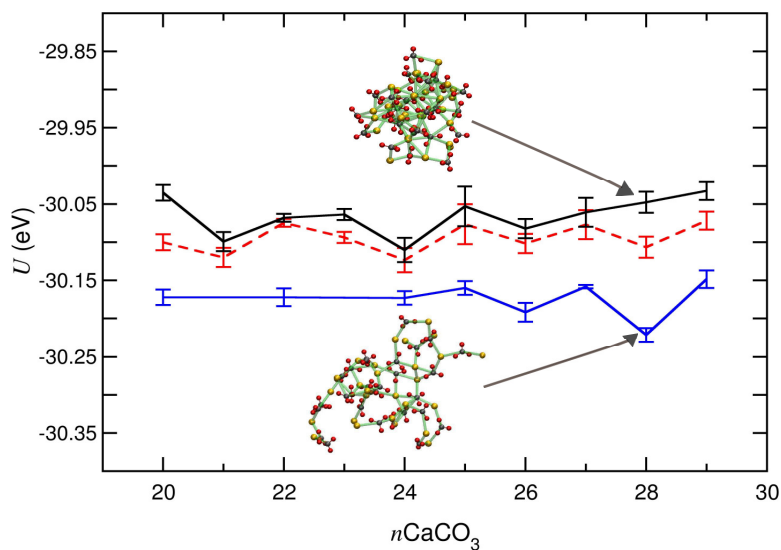
Figure 4.3: Potential energy, U, per formula unit, n, of clusters in aqueous solution for (20–29)$CaCO_3$ with maximum (black) and minimum (blue) initial n_{Ca}. The red–dashed line gives the average U for the five samples simulated, where these samples span the n_{Ca} distribution, as described in Chapter 3. Data points were calculated from the final 1 ns of respective 10 or 5 ns trajectories, with uncertainty of one standard error of the mean indicated. Final configurations for clusters of $28CaCO_3$ are provided. Atom and bond colours are as for Figure 4.2. Bulk water energies have been subtracted.

There does appear to be a slight positive gradient in the potential energy as a function of cluster size for the most compact clusters, and this is likely to be due to increased levels of ionic coordination, and therefore a concomitant decrease in ionic solvation.

Figure C.4 shows representative structures for clusters with initial low (open) and high (compact) coordination after relaxation in water. As expected from simulations of compact clusters, n_{Ca} (Figure 4.4) and CSDs for both open and compact clusters converged very quickly in the simulations. However, for larger clusters, while there was an evident but slight decrease in the average n_{Ca}, as shown in Figure 4.4, there was a consistent difference in the level of n_{Ca} as a function of time.

The average coordination of larger compact clusters by the end of long simulations was consistent with ACC, and that for open clusters was closer to DOLLOP $(1.9 \pm 0.2 - 2.5 \pm 0.1)$, albeit with a slight difference in the pattern of bond ordering. As shown by Demichelis *et al.*, when DOLLOP was formed with clusters containing large amounts of carbonate, the probability of calcium binding to two carbons

74
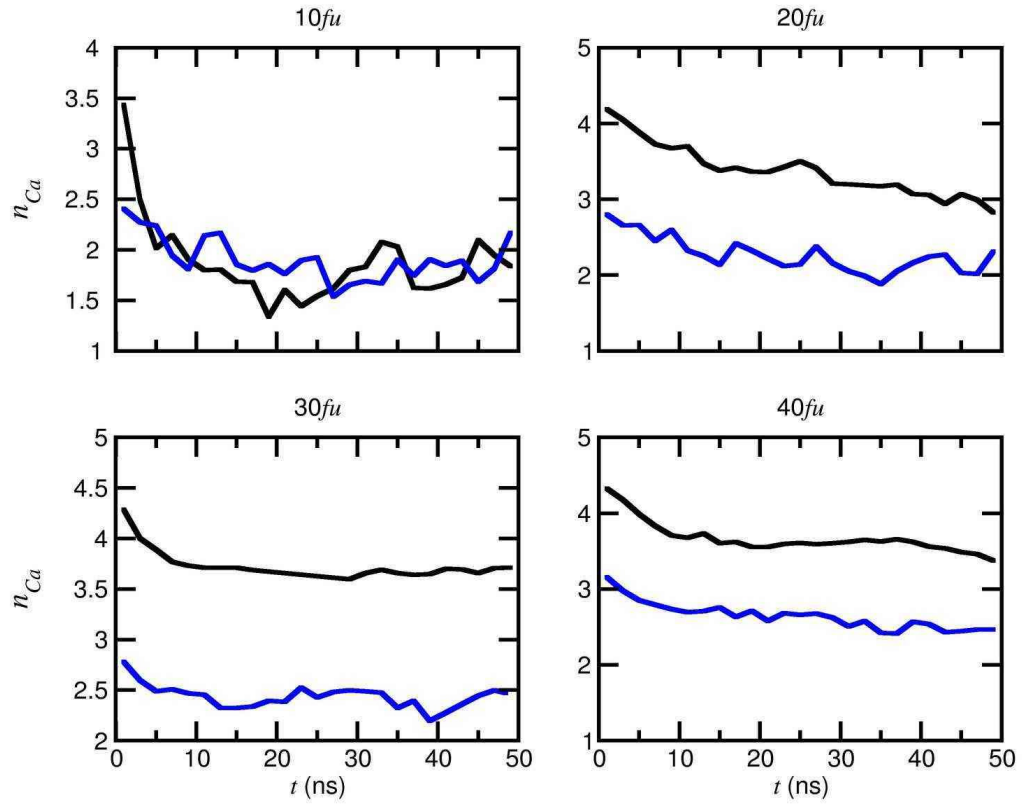
Figure 4.4: Calcium carbonate coordination number, n_{Ca}, as a function of time for simulations of $n\text{CaCO}_3$ where $n =10$, 20, 30, and 40, in water. Blue and black curves represent clusters with initial low and high density respectively.

was around 0.3, with binding to three and four carbons being ~ 0.2 and 0.05 respectively. [Demichelis *et al.*, 2011] Even for $10\,CaCO_3$, the binding probabilities measured, as shown in Table 4.1, suggest an increase in the mass density of clusters due to an increased amount of two-fold coordination and a decrease in the level of calcium binding to a single anion. For larger open clusters, a similar binding pattern was found, albeit with a decrease in the probability of Ca–C, as Ca–4C and Ca–5C increased. It is not clear how the coordination in DOLLOP will evolve over long simulation times, and whether DOLLOP with increased "branching" will become more likely. The difference in bond ordering could be due to increased concentrations in these simulations compared with those of Demichelis *et al.*. For the most compact clusters, the binding probabilities showed a greater likelihood of binding to more than three carbons (see Table 4.1). This is further suggestive that in large compact clusters, the packing of ions in the core remained high, and that these clusters were not indicative of a DOLLOP structure.

Table 4.1: Coordination probabilities for calcium binding to N carbons, measured as a mean average from the final 10 ns of a 50 ns simulation of open and closed clusters containing 10 and 40 $CaCO_3$ units in water.

N	$10\,CaCO_3$ open	$40\,CaCO_3$ open	$40\,CaCO_3$ compact
1	0.30 ± 0.11	0.15 ± 0.04	0.10 ± 0.03
2	0.52 ± 0.06	0.45 ± 0.09	0.18 ± 0.02
3	0.16 ± 0.06	0.24 ± 0.05	0.25 ± 0.02
4	0.01 ± 0.02	0.10 ± 0.02	0.19 ± 0.01
5	0	0.06 ± 0.01	0.18 ± 0.01
6	0	0	0.10 ± 0.01

The behaviour of the larger, open clusters in water were more comparable with those of smaller sizes, in that they displayed dynamic coordination of ionic species throughout the duration of the trajectories. Calcium–carbonate bond lifetime probability densities were measured for clusters over the final 10 ns of a 50 ns simulation. A bond was defined by a distance criteria of $r_{Ca-C} < 4$ Å. As the RDFs of Ca–C (Figure A.1) and Ca-O_{wat} (Figure A.4) show, this distance includes the first Ca–C coordination shell, but is also large enough such that solvation of calcium is likely if ionic coordination is broken. This ensured that large spatial fluctuations in the coordinating ions did not lead to inaccurate estimation of short bond lifetimes. In larger clusters, a fraction of the connections were unbroken over the 10 ns sampling window: 2.4% and 1.8% of connections were fixed in compact clusters containing 30 and 40 units, while 0.6% of connections were fixed in the low density $40\,CaCO_3$ cluster.

Figure 4.5 shows the bond lifetime probability densities for simulations containing ten formula units of calcium carbonate, as well as those for open and compact clusters of the largest size studied. The plot shows that ionic connections are retained for long times in both open and compact clusters. Compact clusters showed much longer Ca–C binding, with probability density up to around 4 ns. The most open cluster, that of $10\,CaCO_3$, had the highest probability density for lifetimes < 0.5 ns, highlighting that this cluster showed the most dynamic (dis)ordering. The dynamic ordering in the largest open cluster was on a similar time-scale to that of $10\,CaCO_3$, but the probability density of bond ordering at short times is slightly reduced. Figure C.5 provides the bond lifetime probability densities for open clusters across the cluster size range studied, which further shows that as cluster size was increased, calcium carbonate connections persisted for longer times. This can be explained by considering the dynamics of the breaking and reforming ionic connections within the clusters. For very small clusters, when a connection is broken, there is a limited, small number of possible connections that can be made to reproduce a cluster of equal size to the initial state. However, for larger clusters, there are many possible ionic connections apparent upon cluster dissociation. Furthermore, the lifetime of an insufficiently coordinated ion in larger clusters is reduced, compared with smaller clusters of comparable density, as ionic concentration in the local proximity of cluster constituents increases.
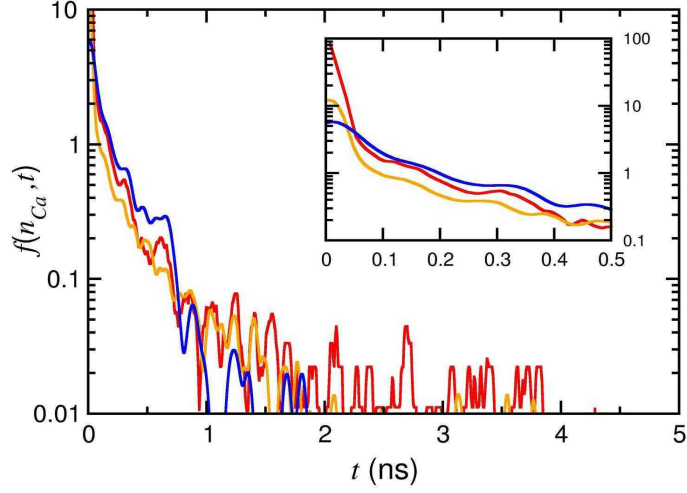
Figure 4.5: Calcium carbonate bond lifetime probability densities measured from the final 10 ns of 50 ns simulations of $40\,CaCO_3$ with low (orange) and high (red) density, and $10\,CaCO_3$ with low density (blue) in water. Bonds were defined by a distance criteria between calcium and carbon atoms of < 4 Å. The average lifetime of the bond is given on the x axis. Inset is the probability densities for bonds with a lifetime under 0.5 ns.

4.3.3 Temperature Effects

If one is to construct a phase diagram for calcium carbonate, then the effect of temperature on low density clusters can help to resolve the features of the diagram in the prenucleation regime. To this effect, a low density $20\,CaCO_3$ cluster was simulated in water for 20 ns in the temperature range 320–500 K. As temperature increased, the potential energies of the cluster per formula unit increased linearly (as shown in Figure C.6) from $\sim$ -30 – -18 eV. The linear scaling of energy with temperature is expected for a homogeneous bulk phase.

Figure 4.6 (a) shows the change in n_{Ca} and average cluster size as a function of temperature. Both of these quantities increased as T increased, with the coordination increasing from a value comparable to that of DOLLOP, to one which approached the values expected for the most dense cluster when relaxed in water (see Figure 4.4 for $20\,CaCO_3$). The change in coordination was approximately linear, consistent with the change in potential energy. The average cluster size at 500 K is close to $N = 100$, which is the maximum possible value, suggesting that the cluster condensed in solution as the temperature increased.

The change in coordination could be ascribed to just aggregation of smaller

(a)

(b)
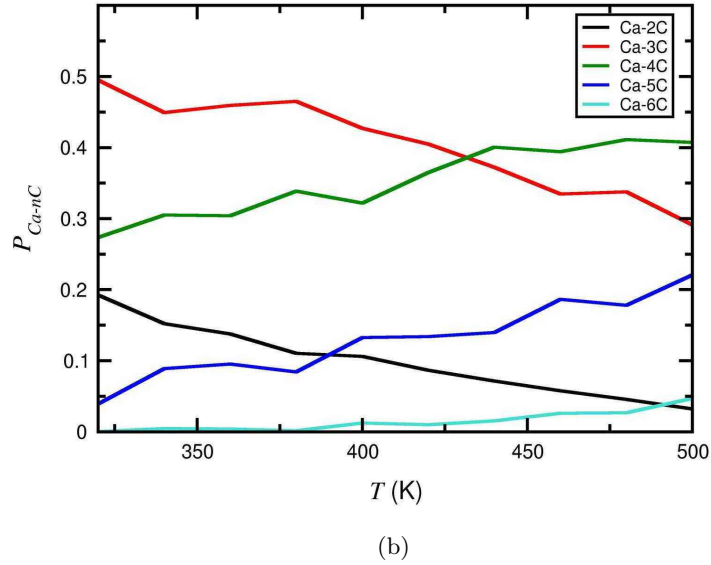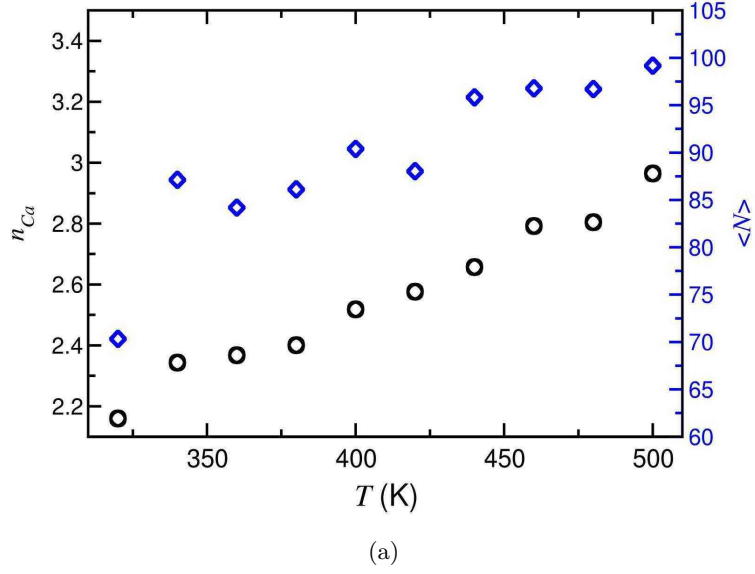
Figure 4.6: (a) Coordination of carbon to calcium, n_{Ca}, and average cluster size, $< N >$ (as a measure of number of cluster atoms), for simulations of a low density $20\,CaCO_3$ in water. Simulations were performed for 20 ns, with measurements taken from the final 10 ns of trajectories. Black circles show n_{Ca} and blue diamonds show $< N >$, with y axis scales shown on the left and right, respectively.

(b) Coordination probabilities of n carbons to a single calcium as a function of temperature, for the simulations described in (a). Coordination was measured using a distance criteria of $r_{Ca-C} < 3.825$ Å.

clusters in solution. However, through consideration of the change in bonding probabilities, P_{Ca-nC}, as a function of temperature, that was not the case. Coordination of calcium to two and three carbons decreased as T increased, as shown in Figure 4.6 (b), and this was associated with an increase in calcium coordination to $4-6$ carbons. At $T = 460$ K, the probability of calcium binding to four carbons was higher than P_{Ca-3C}, and there was also a small probability of six–fold coordination ($P_{Ca-6C} \sim 0.05$ at $T = 500$ K). These results suggest that as temperature increased, clusters of calcium carbonate aggregate in solution, and at the same time, ions pack more closely. There was a relatively large increase in the average lifetime of bonds at the higher temperatures. Probability densities of 0.1 were found for bond lifetimes at 0.45 ns (measured over the final 5 ns of simulation) for $T = 500$ K; similar probability densities were found at a maximum bond lifetime of 0.1 ns for $T = 320$ K. While these lifetimes are smaller than compact clusters at 300 K, they do indicate that dynamic ordering decreased as temperature increased, and that the cluster was behaving more like ACC than a dense liquid at the highest temperature. Higher temperature should favour high entropy states, which does not, at first sight, appear to be the case for calcium carbonate. However, increasing ion coordination results in the release of highly constrained water from ion solvation shells. The disruption of water order around ions will lead to a favourable increase in water entropy. This is likely to be the driving force for condensation at high temperature.

4.3.4 Surface Charge Bias

It is interesting to consider the effects of solvation on the strength of the charge bias of clusters which was found during optimisations, and is discussed in Chapter 3 (section 3.3.3). The charge bias which was observed for clusters (particularly those which were compact) in the gas phase was reduced when clusters were immersed and relaxed in water. This can be attributed to ion hydration and partial dissolution of clusters. However, the nature of a charge distribution bias did persist in these systems, as shown in Figure 4.7, which provides the final configurations for an open and compact cluster of $27CaCO_3$ when simulated in water. The cluster with maximum initial n_{Ca} remained relatively dense by the end of a 5 ns simulation, and the majority of positively charged species resided below the surface of the cluster. In contrast, ions in the sampled cluster with lowest initial n_{Ca}, became further exposed to solvent molecules over the course of the simulation, and density continued to reduce as the structure disassembled, resulting in charge distributions approaching uniformity. Although data are presented here for one size of cluster, similar behaviour was commonly found for all of the larger cluster sizes studied.
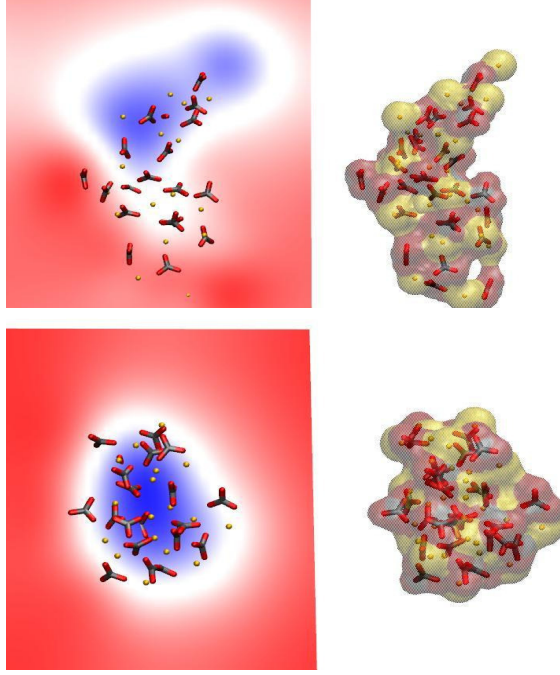
Figure 4.7: Configurations for clusters of $27CaCO_3$ after 5 ns of simulation in water. Top and bottom show clusters with n_{Ca} =2.5 (t_0, $n_{Ca} = 3.1$) and n_{Ca} =3.8 (t_0, $n_{Ca} = 4.7$); the explicit structure is shown on the right (see the caption in Figure 4.2 for atom identities), while the image on the left is an electrostatic potential map, indicating the electrostatic charge distribution half way through the simulation cell z axis for Ca, C and O atoms only [Aksimentiev and Schulten, 2005]. Here shading represents negative (red) and positive (blue) regions, ranging from -100 – 280 $\frac{k_\mathrm{B}T}{e}$.

4.3.5   Clusters Containing Water and Additives

Clusters of equivalent numbers of ions, n =4, 6, 10 and 15, were sampled for $n\text{CaCO}_3$ (anhydrous), $n(\text{CaCO}_3 \cdot \text{H}_2\text{O})$ (hydrous) and $n(\text{CaCO}_3 \cdot \text{H}_2\text{O}) \cdot \text{ASP}$ (hydrous+ASP) optimised clusters. Therefore, direct comparisons of the dynamics and stabilities of clusters inclusive of water and water and ASP can be made.

Effect of ASP on Optimised Cluster Configurations

Table 4.2: Average n_{Ca} values for anhydrous, hydrous and hydrous+ASP optimised clusters, from optimisations of random structures in vacuum.

n	Anhydrous	Hydrous	Hydrous+ASP
4	2.38	2.39	1.80
6	2.55	2.62	2.18
10	2.90	2.93	2.61
15	3.21	3.17	2.91

Before discussing the dynamics of clusters, it is first helpful to consider the effect of ASP on the optimised structures of clusters in vacuum. As shown in Table 4.2, while the presence of water during optimisations served to generally solvate the structures found for anhydrous optimisations, ASP did perturb these configurations (see Figure C.7). The carboxylate oxygens and amine hydrogens of ASP bound to the Ca^{2+} ions and oxygens of CO_3^{2-} in the cluster, respectively. This resulted in a more distorted network amongst the remaining calcium carbonate, and clusters with greater asphericity were found than was observed in the anhydrous and hydrous clusters. ASP displaced ions within the cluster as functional groups of the organic species competed to coordinate with ions. While n_{Ca} values shown in Table 4.2 suggest a less dense structure for clusters containing ASP, these data do not include the coordination between Ca^{2+} and CO_3^{2-} ions with the functional groups of ASP.

Hydrous and Hydrous+ASP Cluster Simulations

The hydrous clusters, with and without ASP, were immersed into water and MD simulations were conducted. The behaviour of the smaller solvated hydrous clusters was very similar to that observed for anhydrous clusters. Small hydrous clusters were found to dissociate into smaller clusters and ion pairs over the course of the simulation; these subsequently aggregated and re-dissolved, displaying dynamic ordering. Larger, compact hydrous clusters tended not to disassemble in solution, and remained particulate throughout the trajectory, albeit with some reconfiguration

apparent inside, and at the surface of, the cluster. The case of the minimum energy *in vacuo* $15(CaCO_3 \cdot H_2O)$ cluster is particularly interesting as this contained two embedded water molecules at $t = 0$. Figure 4.8 shows snapshots of the cluster at the beginning and end of a 5 ns simulation in water. After approximately 80 ps of simulation, a third water molecule was absorbed from the surface and became integrated internally within the cluster (see Figure 4.8 (b)). All three water molecules remained embedded within the cluster for the remainder of the trajectory. The rate of disassembly in this cluster, as indicated by the fraction of maximum n_{Ca} (see Figure 4.9 (a)), was comparable to that of the anhydrous cluster with the same number of ions. A reduction in coordination was primarily seen at the cluster–solvent interface as ion solvation was maximised, which was observed in the case of the compact $15\,CaCO_3$ cluster.

(a)   (b)
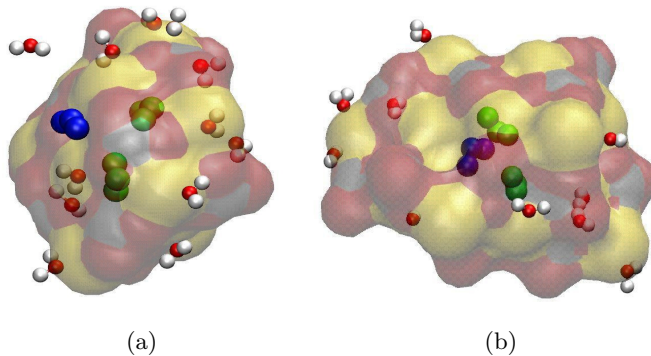
Figure 4.8: (a) Initial and (b) final structures of a $15\,(\,CaCO_3 \cdot H_2O)$ cluster from a 5 ns simulation in water. The ionic surface is shown and carbon, oxygen, calcium and hydrogen are coloured black, red, yellow and white, respectively. Highlighted are two water molecules (green) which resided below the cluster surface throughout the simulation, and one water molecule (blue) which initially resided at the surface of the cluster, but was absorbed below the surface at ~ 80 ps.
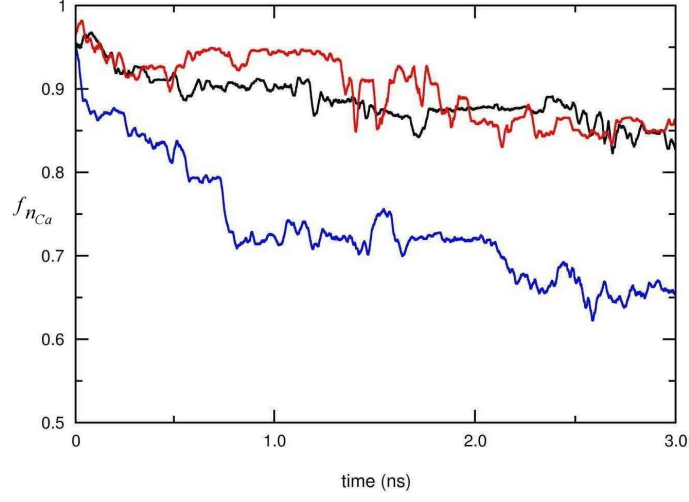
(a)
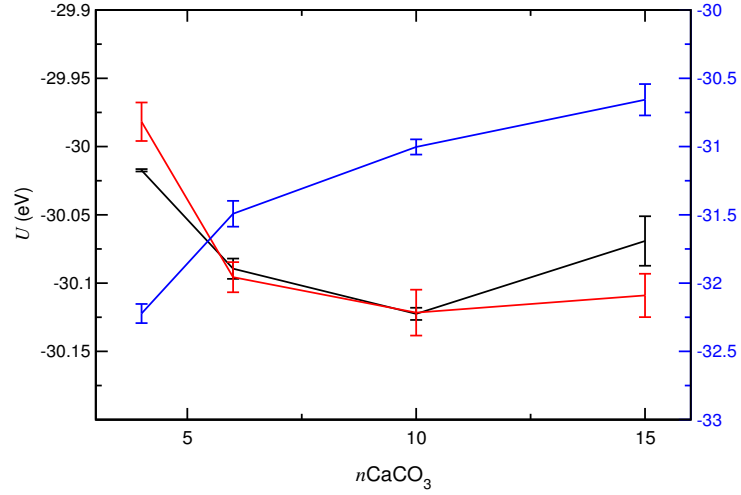
(b)

Figure 4.9: (a) Fraction of maximum coordination number ($f_{n_{Ca}}$), as a function of time for minimum energy (*in vacuo*) clusters containing 15 $CaCO_3$ units simulated in water, and (b), average U per $nCaCO_3$ for clusters of size $n =$4, 6, 10 and 15, simulated in water. Anhydrous (red), hydrous (black) and hydrous+ASP (blue) clusters have been simulated. The black data in (b) is taken from Figure 4.1 for anhydrous clusters, and the shifted scale on the right of the plot in blue relates to the data for hydrous+ASP only (due to the energy associated with ASP inclusion). Mean values were taken from the final 1 ns of 10 ns trajectories with bulk water energy subtracted, and uncertainty in the data representing one standard error of the mean are provided.

Figure 4.9 (b) shows the mean potential energy, U, for hydrous cluster systems over the final portions of a 10 ns trajectory. The data are similar to most for anhydrous clusters, and there was little difference between the structures and energetics found between anhydrous and hydrous clusters in this size regime. From this data, we can expect the number of ions within a cluster to define the limit of cluster stability, rather than the level of hydration for dense cluster systems, especially when one considers the nature of the dynamics seen for clusters in the size range studied here.

The potential energy data for clusters containing one ASP molecule did not follow the trends found for other, solvated cluster systems. Alternatively, U increased as the number of ions in the cluster increased. This can be explained by considering the interaction of ASP functional groups to ions of Ca^{2+} and CO_3^{2-} in solvated clusters. The binding between ASP and ions in the cluster was sufficiently strong to limit the extent to which the clusters dissociated. Although disassembly was observed to occur for ASP–bound clusters (of initial maximum n_{Ca}) at a faster rate than in similar anhydrous and hydrous clusters, as shown in Figure 4.9 (a), with n_{Ca} decreasing and R_g increasing at early times during the trajectory, cluster separation into smaller species was less likely in the presence of ASP. This was consistent over the whole size range of clusters studied, as shown in Figure C.8, and may be due to the fact that ASP inclusion leads to a cluster which is less compact than equivalent anhydrous clusters (simulation of anhydrous clusters indicated that with initial lower coordination, cluster disassembly to open states is more rapid.)

The CSDs shown in Figure 4.10 highlight the possible cluster sizes accessible during the final 3 ns of a 10 ns trajectory for clusters containing ASP, and those for the equivalent control experiments, for which ASP was absent. For small systems, with all ions initially in contact with solution, distributions indicated many possible cluster sizes, and the influence of the organic species is unclear. However, as cluster size increased, the effect of ASP was to reduce the amount of cluster dissociation to smaller species. For instance, when simulating clusters containing $15\,(\,CaCO_3 \cdot H_2O) \cdot ASP$, distributions indicated dissolution into approximately two smaller clusters when ASP was bound in the initial state; on the other hand, a wider range of accessible cluster sizes was displayed in the control system, over the timescales studied here.
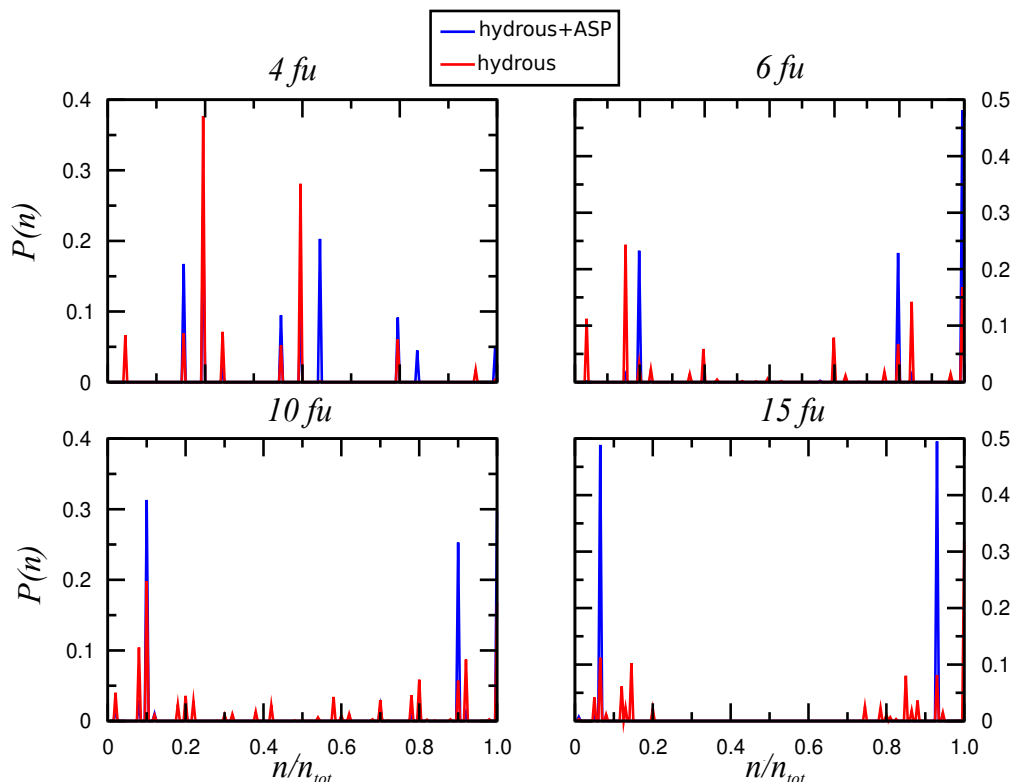
Figure 4.10: Cluster size distributions for simulations of clusters containing $n(\text{CaCO}_3 \cdot \text{H}_2\text{O})$ (red) and $n(\text{CaCO}_3 \cdot \text{H}_2\text{O}) \cdot \text{ASP}$ (blue) for $n = 4, 6, 10$, and 15. Distributions have been sampled from the final 3 ns of a 10 ns trajectory, and have been normalised against maximum cluster size for ease of comparison.

Initial and relaxed structures for clusters containing 15 CaCO_3 units, with and without ASP, are depicted in Figure 4.11. Note that although the cluster with ASP was optimised with a 1:1 ratio of CaCO_3:H_2O, water was not found below the cluster surface, and instead this water formed a solvation shell. The images show that after 10 ns of simulation, the cluster without ASP had disassembled and dissolution was apparent, with dynamic ordering also observed. This results in the open cluster species described in section 4.3.1. However, for the cluster which included ASP, while coordination decreased (see Figure C.8), leading to a more open structure, the local concentration of ions remained high. During the course of a 10 ns simulation, the strength of the binding between ASP and ions in the cluster was such that the organic molecule did not readily desorb from the cluster. Furthermore, because at $t = 0$ the cluster configuration had maximum n_{Ca}, and that ASP was coordinated to ions across the surface of a cluster, the organic molecule
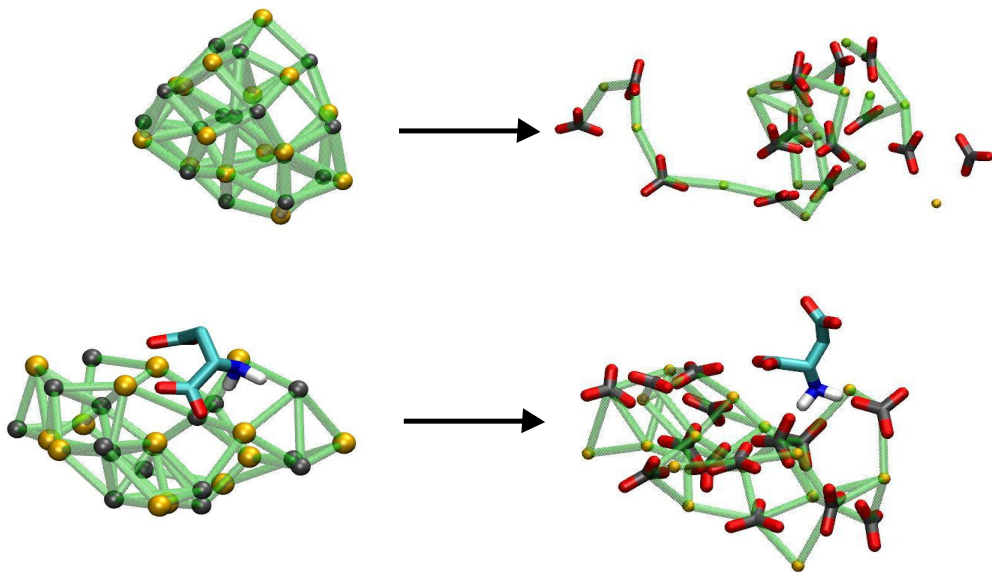
Figure 4.11: Configurations are provided for clusters containing 15 formula units of $CaCO_3$ (top), and $CaCO_3$ with one molecule of ASP (bottom). Calcium carbonate atoms are coloured as described for previous Figures (see Figure 4.2 caption), while nitrogen, oxygen, carbon and hydrogen of ASP are coloured: blue, red, cyan and white, respectively. The structures on the left indicate the configuration of clusters at $t = 0$ and do not show carbonate oxygen atoms for clarity (top: $n_{Ca} = 4.5$, bottom: $n_{Ca} = 3.9$). Structures on the right show the systems at $t = 10$ ns (top: $n_{Ca} = 2.4$, bottom: $n_{Ca} = 3.0$.

reduced the possibility for the cluster to dissociate into smaller species.

4.4 Summary and Conclusions

(Meta)stable Clusters

Simulations of clusters up to the size of $40\,CaCO_3$ in water indicate that calcium–carbon coordination can aid in understanding the stability of structures. The balance between ionic coordination and ion solvation was reflected in the potential energies of solvated clusters, which can be used to infer cluster stability. While this is not strictly a measure of thermodynamic stability, due to neglected entropic effects, monotonically decreasing free energy has been found for amorphous particles with increasing numbers of ion pairs of calcium carbonate, indicating that larger amorphous clusters are thermodynamically more stable than smaller ones [Raiteri and Gale, 2010]. Samples with lower levels of coordination, and therefore lower

density, tended to disassemble more readily in solution, with regular breaking of the cation–anion network at chain ends of dendritic clusters. These low coordination clusters were found to have a lower potential energy than structures with maximum coordination, and this has been ascribed to ion solvation, which is maximised in clusters with a large surface area.

Stabilisation of ions in solution was greatest for small dynamically ordered clusters, which is in agreement to the studies of Demichelis *et al.* [Demichelis *et al.*, 2011], suggesting that DOLLOP is most stable in solution. However, the definition of DOLLOP is not so clear. The features of DOLLOP which are laid out by Demichelis *et al.* are that dynamic ordering is present, yet this was found in all clusters covering a range of densities, even those with coordination numbers comparable to stable ACC. Without a reference time-scale for dynamic ordering, it is difficult to say exactly what classifies as DOLLOP. Further, the coordination profile for DOLLOP suggests that the most probable bonding arrangement is Ca–2C (with some Ca–C and Ca–3C). However, open clusters in this study tended to display higher coordination levels, even for those found in the minimum of Figure 4.1. As the cluster size increased, the coordination profiles of clusters with lowest energy diverged further from that suggested by Demichelis *et al.* for DOLLOP. One key parameter in the description of DOLLOP is R_g. Probability distributions of the measured R_g for a dry, wet and DOLLOP particle of $36\,CaCO_3$ provided by Demichelis *et al.*, showed distinct accessible regions for each state. The accessibility of large volumes of conformation space, as indicated by R_g distributions, is used to quantitatively show the liquid-like character of DOLLOP. Average R_g for low density clusters of 68–73 ions (see Figure C.9) showed that the open clusters simulated here do fall in the region of DOLLOP, but at the interface with the "wet" nanoparticle (*i.e.* hydrated ACC).

Wallace *et al.* studied the growth of dense liquid particles by addition of solvated ion pairs to a growing cluster of calcium carbonate followed by replica exchange MD simulations [Wallace *et al.*, 2013]. The coordination numbers suggest that for clusters with more than thirty ions, binding of three anions to calcium dominates, with a significant amount of four-fold coordination observed. This is congruent with the clusters examined in this study. Furthermore, it was shown that this change in coordination was complementary to a decrease in the mobility of ions in the clusters. Wallace *et al.* describe these clusters as "*more compact dense states*" to DOLLOP [Wallace *et al.*, 2013]. It is not entirely clear how the two states are thermodynamically linked, and if there is a smooth transition from a low coordination DOLLOP cluster to a higher coordinated, but highly mobile, dense

liquid phase.

While highly solvated, dynamically ordered calcium carbonate particles were found to be lowest in potential energy (and likely free energy at 300 K – due to the entropic contribution coming from a wide distribution of accessible configurations), it is not true to say that open clusters over the size range of prenucleation clusters studied are DOLLOP. The clusters will instead be described as dense liquid nanoparticles (DLNPs) for the remainder of this thesis. It is not clear whether DLNPs will have a limiting size to stability, and therefore whether these are PNCs. Wallace *et al.* showed no free energy barrier to the growth of dense liquid clusters up to a size of 45 ions [Wallace *et al.*, 2013], and growth of calcium carbonate DOLLOP by addition of ion pairs (with activation energies on the order of k_BT) only ceased when the supply of ion pairs diminished [Demichelis *et al.*, 2011].

Although solvent-induced deformation was seen at the surface of the largest, most dense clusters in solution, these clusters remained highly coordinated by the end of a 10 ns or 50 ns trajectory, with the internal structure of the clusters appearing to be quite stable. Coordination remained high in the largest, most compact clusters throughout the simulations, and the observation of high coordination environments, along with long bond lifetimes, suggests that these clusters will remain particulate over long times. This is not surprising as in larger clusters, the electrostatic interactions for high local ionic concentrations will lead to strong binding. These clusters are ACC in the core, which is stabilised by the diffuse interface created by a dynamically ordered layer of solvated ions. ACC in computer simulations is the most stable phase for a $75 \, CaCO_3$ cluster (albeit with a different force field to the one used here), and particles of ACC < 100 nm in diameter can be stable in solution [Quigley and Rodger, 2008a; Nudelman *et al.*, 2010]. However, as these are higher in energy, it is unlikely they will be found in solution over DLNPs in the size range investigated here.

Following these results, questions remain about the nature of species which are thought to exist before nucleation. To gain further insight, the equilibrium between ions in solution, DOLLOP and DLNPs needs to be further examined. Also, whether DLNPs are PNCs which aggregate, as suggested in a stepwise nucleation mechanism following, ion solution $\longrightarrow$ PNCs $\longrightarrow$ ACC, or if DLNPs are involved in liquid–liquid phase separation and spinodal decomposition, the relative stability compared with more dense states (*i.e.* ACC) needs to be established.

Temperature Effects

When DLNPs were simulated over a range of temperatures, condensation towards compact clusters was observed as T increased. It is known that carbonates condense as temperature is increased, and so this result is not so surprising. As an increase in T reduces the size of free energy barriers between states, the accessibility of compact clusters under standard conditions is not clear without suitable free energy calculations. Nonetheless, the appearance of condensed states allows one to understand the changing thermodynamic stability of calcium carbonate and to consider the phase diagram.

de Yoreo and co-workers have shown that multiple pathways to calcium carbonate growth may occur, and that nucleation via PNCs is not always necessary for precipitation [De Yoreo, 2013; Hu *et al.*, 2013; Wallace *et al.*, 2013]. Furthermore, in the paper of Wallace *et al.*, from simulations of liquid clusters the authors were able to propose a phase diagram based on liquid–liquid phase separation (see Chapter 1, section 1.4.4 for a more detailed discussion) [Wallace *et al.*, 2013]. This was informed from the lack of energy barrier to cluster association and the monotonic decrease of free energy with cluster size, suggesting that the solution was undergoing a spontaneous phase separation [Wallace *et al.*, 2013].

It is possible that the DLNPs observed in this study are a liquid phase coexisting with a dilute calcium carbonate solution. If that is the case, then according to the phase diagram suggested by Wallace *et al.*, solid can nucleate either from a classical mechanism in the lean solution, or perhaps by a non–classical route from the dense liquid phase [Wallace *et al.*, 2013]. Simulations here showed that as temperature was increased, the dense liquid nanoparticle equilibrated to a state which was more comparable in density and coordination levels to an amorphous solid phase. This suggests that nucleation of solid could possibly occur by reconfiguration within a dense liquid phase. The transition in the stability of the liquid to solid phase with temperature is indicative of spinodal decomposition. Therefore, while the phase diagram of Wallace *et al.* only includes a binodal for liquid–liquid separation, then according to the results in this Chapter, a second binodal should be present for the coexistence of dense liquid and solid phases.

Aspartate

While the simulations for hydrous clusters, containing a 1:1 ratio of $CaCO_3$:H_2O, were analogous to the simulations for anhydrous clusters of the same size, solvated clusters which also included ASP did not show the same trends in solution as other

clusters of equivalent size. The potential energy of these clusters (per formula unit) increased as the number of ions in the cluster increased. It was shown that ASP reduced the amount of ion dissolution and subsequent aggregation in solution, but did not hinder the reduction of coordination in clusters and any associated increase in ion solvation. Coordination was observed to decrease more rapidly and to a greater extent in clusters containing ASP than was found in equivalent anhydrous systems. ASP was observed to form strong bonds to ions and limited the change in CSDs during simulations.

Our studies do not fully explain the role of ASP on DLNPs; this is due to the fact that we have investigated clusters with bound ASP in the initial state. In reality, ASP is more likely to associate with an emerging cluster from a dilute ionic solution. In this environment, the influence of ASP on cluster stability may be different from that found here. When ASP is bound in clusters, the decreasing dissolution of calcium carbonate may suggest a limiting of the growth of prenucleation species, due to diminished ion aggregation. Conversely, the multiple binding sites associated with ASP could lead to the organic species bridging clusters in solution and leading to more rapid cluster–cluster aggregation, and therefore promoting cluster growth. By measuring the free energy of binding of ASP to ion pairs and ACC combined with experimental investigations, Raiteri *et al.* were able to conclude that ASP acts to inhibit the nucleation of calcium carbonate, which suggests a stabilising of any PNCs in solution and limited cluster aggregation [Raiteri *et al.*, 2013].

Polyaspartate has been shown to promote liquid-liquid separation in the PILP (Polymer Induced Liquid Precursor) process, and ASP rich matrices are often used in the stabilisation of PILP [Evans, 2013; Gower, 2008]. Considering the behaviour of ASP on calcium carbonate here, it is understandable how ASP-rich proteins achieve this. The ability for ASP to retain high local concentration of ions with low coordination in the form of DLNPs demonstrated here means that extended solutions of polymer and interconnecting ionic networks may be possible.

Chapter 5

Cluster Free Energies

5.1 Introduction

During the precipitation of calcium carbonate, ions associate in solution. According to simulations conducted in Chapter 4, the first product of that association at 300 K and 1 atm, is a dense liquid nanoparticle (DLNP) phase. But the identification of low (potential) energy DLNPs does not preclude the emergence of the more dense nanoparticles which were observed to be mechanically stable (in the simulation times investigated), particularly for clusters containing more than 20 formula units of calcium carbonate (see Chapter 4). The energetic measurement to infer stability in the dense phases was potential energy; however, to understand the relative thermodynamic stabilities of free ions, DLNPs and more dense structures, the entropic contribution to the energy must also be included. The following Chapter presents the findings from a number of free energy calculations of calcium carbonate solutions.

Collective variable (CV) techniques, namely, metadynamics and Umbrella Sampling (US), have been used in order to traverse a reversible path between the possible states available for 20–80 ions in solution. A number of CVs were tested and a comparison is made on the efficiencies and adequacy of each. From the results, the likely stable species to emerge from solution are identified and their significance to the biomineralisation mechanism are outlined.

5.2 Methods

Free energy calculations were initiated on cluster configurations which were prepared as described in Chapter 3. Configurations for 10, 15, 20, 30 and 40 calcium carbon-

ate unit (low energy) clusters which were sampled with high and low coordination of carbon to calcium were extracted and immersed in water; 1200, 2753, 1600, 2000 and 2400 water molecules were used to solvate 10, 15, 20, 30 and 40 formula unit clusters respectively. Across the size range, the concentration of systems also increased. Short 50–200 ps molecular dynamics (MD) simulations were conducted to equilibrate the volume of the simulation cell, while the spatial coordinates of calcium and carbonate atoms remained fixed. Following this, all atoms were mobilised for the free energy calculations.

Unless otherwise stated, simulations were conducted in the NpT ensemble with the application of cubic periodic boundaries. The temperature and pressure was 298 K and 1 atm which was regulated by a Nosé-Hoover thermostat and barostat with 0.1 and 1.0 ps relaxation times. The force field employed was that of Raiteri and Gale which included flexible water [Raiteri and Gale, 2010]. Intermolecular interactions were truncated at 9 Å, with the addition of a tapering function between Buckingham short-range potentials between 6 and 9 Å, as suggested in the force field. Electrostatics were treated using the smooth particle mesh Ewald method [Essmann *et al.*, 1995]. A 1 fs timestep was used throughout. Simulations were conducted with the DL_POLY suite of MD simulation packages [Todorov *et al.*, 2006]. Metadynamics calculations were performed with a version of DL_POLY 4.05, and DL_POLY 2.2 with the Plumed Plugin of Bonomi and co-workers [Bonomi *et al.*, 2009]. US was performed with DL_POLY 2.2 with the Plumed Plugin. Analysis of free energy surfaces was performed using the sum_hills utility of Plumed for metadynamics, and the weighted histogram analysis method (WHAM) as packaged by Grossfield for US [Grossfield, 2014].

The CVs evaluated were radius of gyration, R_g, ion solvation and ion coordination. R_g was evaluated from the centre of mass of a group of defined cluster atoms (see equation 5.1; r_i and m_i are atom i coordinates and mass), and the closest periodic image of atoms was used initially to build the cluster configuration (thereby removing discontinuities in R_g as a cluster extended across a boundary). Both ion solvation and coordination were defined by a coordination CV between two groups of atoms, G_1 and G_2. Two definitions of coordination were employed. In the Plumed plugin, the total coordination was defined by a double sum over the coordination, n_{ij}, between i and j atoms (with coordinates r_i and r_j) in these two groups as shown in equation 5.2. A continuous function was used to define n_{ij} as shown in equation 5.3, where $r_{ij} = | r_i - r_j | -d_0$. The parameters n, m, r_0 and d_0 allow for fine tuning of the definition of coordination. The Plumed manual suggests that r_0 and d_0 can initially be taken as the position and full width at half maximum

of the first peak in the radial distribution function (RDF) for G_1 and G_2. n and m parameters can be tuned so that n_{ij} is approximately zero at the position of the first minimum in the RDF. DL_POLY allows for biasing a continuous version of Steinhardt bond-orientational CVs, based on the spherical harmonics and distances between atoms in G_1 and G_2. A Q_0 CV was implemented into DL_POLY 4.05 to define coordination. The function for $Q_0^{G_1 G_2}$ is shown in equation 5.4, which is dependent upon the number of Q_0 sites, N_S, the number of atoms in G_1, N_{G_1}, and the continuous function $f_c(r_b)$, as shown in equation 5.5. $f_c(r_b)$ takes the coordination between two atoms smoothly to zero when the distance between them, r_b, is between the parameters r_1 and r_2. Sensible choices for r_1 and r_2 would be the position of the first peak and the first minimum of the RDF between atoms in G_1 and G_2 [Quigley and Rodger, 2009]. The use of two different definitions for coordination should not lead to inconsistencies in results, so long as the parameters are chosen carefully.

A number of CV combinations for different system sizes have been used. The specific details of CVs and parameters adopted are provided in the following section.

$$R_g^2 = \frac{\sum_i^n (r_i - r_0)^2}{\sum_i^n m_i}; \qquad r_0 = \frac{\sum_i^n r_i m_i}{\sum_i^n m_i} \tag{5.1}$$

$$N_{ij} = \sum_{i \in G_1} \sum_{j \in G_2} n_{ij} \tag{5.2}$$

$$n_{ij} = \begin{cases} 1 & r_{ij} \leq 0 \\ \frac{1 - \left(\frac{r_{ij}}{r_0}\right)^n}{1 - \left(\frac{r_{ij}}{r_0}\right)^m} & r_{ij} > 0 \end{cases} \tag{5.3}$$

$$Q_0^{G_1 G_2} = \left[4\pi \left| \frac{1}{N_S N_{G_1}} \sum_{b=1}^{N_b} \frac{1}{2\sqrt{\pi}} f_c(r_b) \right|^2 \right]^{\frac{1}{2}} \tag{5.4}$$

$$f_c(r_b) = \begin{cases} 1 & r_b \leq r_1 \\ \frac{1}{2} \left\{ cos \left[\frac{\pi(r_b - r_1)}{(r_2 - r_1)} \right] + 1 \right\} & r_1 < r_b \leq r_2 \\ 0 & r_b > r_2 \end{cases} \tag{5.5}$$

5.3 Results

5.3.1 Biasing R_g and N_{Ca-C}

A crucial feature of collective variables is that all possible phases must be described by a unique set of CV values, such that the distributions of the CVs for different

phases are well separated. In this way, complex systems with high dimensionality can be simplified and described in terms of a small number of parameters. In Chapter 3, shape descriptors such as R_g and Ca-C coordination were used in combination to analyse the similarity/difference between clusters with varying densities. It was further shown in Chapter 4 that these parameters can define the different states accessible to calcium carbonate in solution.

Both R_g and global Ca-C coordination, N_{Ca-C}, ($n = 4$; $m = 16$; $d_0 = 3.3$; $r_0 = 0.7$) were biased in a metadynamics simulation of a calcium carbonate cluster with composition $15 \, CaCO_3$. Bias was accumulated in the form of Gaussian hills which were deposited every 400 MD steps with a height of 10 $k_B T$ and with σ parameters (ascertained from the standard deviation of CVs during MD simulations) of 0.03 Å and 0.8 for R_g and N_{Ca-C}, respectively. The time evolution of CVs is provided in Figure D.1. Initially, the cluster had a relatively low density with $R_g = 11.4$ Å and $N_{Ca-C} = 29$. During the course of the 30 ns simulation, a wide distribution of both parameters was found: both R_g and global N_{Ca-C} fluctuated throughout the simulation, highlighting the fact that many states were visited during sampling. The relatively large height of the Gaussian hills assisted in the fast sampling of a wide distribution of states.

The free energy surface of the system as a function of the adopted CVs is shown in Figure 5.1. The surface is very noisy, and a large basin can be observed with the minimum centred at $R_g \approx 15$ Å and $N_{Ca-C} \approx 18$. This minimum relates to a collection of small clusters each with low coordination, an example snapshot of which is shown in Figure 5.1. Low density clusters with high levels of solvation were observed to have low enthalpy in Chapter 4, and as these clusters have the highest degree of dynamic coordination, and therefore entropy (at 300 K), from relatively short bond lifetimes, it is not surprising that this state is the one with the lowest free energy. The region of the free energy surface (FES) which is associated with a single cluster ($R_g \lesssim 12$ Å) is relatively high in energy, with a negative gradient of free energy (FE) in this region leading to the minimum shown. A cluster with average Ca–C coordination of approximately 2.6 is shown in Figure 5.1, and the energy of this cluster was approximately 0.2 eV per formula unit higher than the minimum energy state. Anhydrous clusters with maximum ion packing were not observed in the calculation, which suggests that while a wide region of phase space had been visited, the FE surface may not have been fully sampled. The most dense cluster was found with $R_g \approx 6$ Å and $N_{Ca-C} \approx 55$; two water molecules were observed below the ionic surface of this cluster. A minimum in FE was observed for this state. However, while the barrier height between the high density cluster and

95
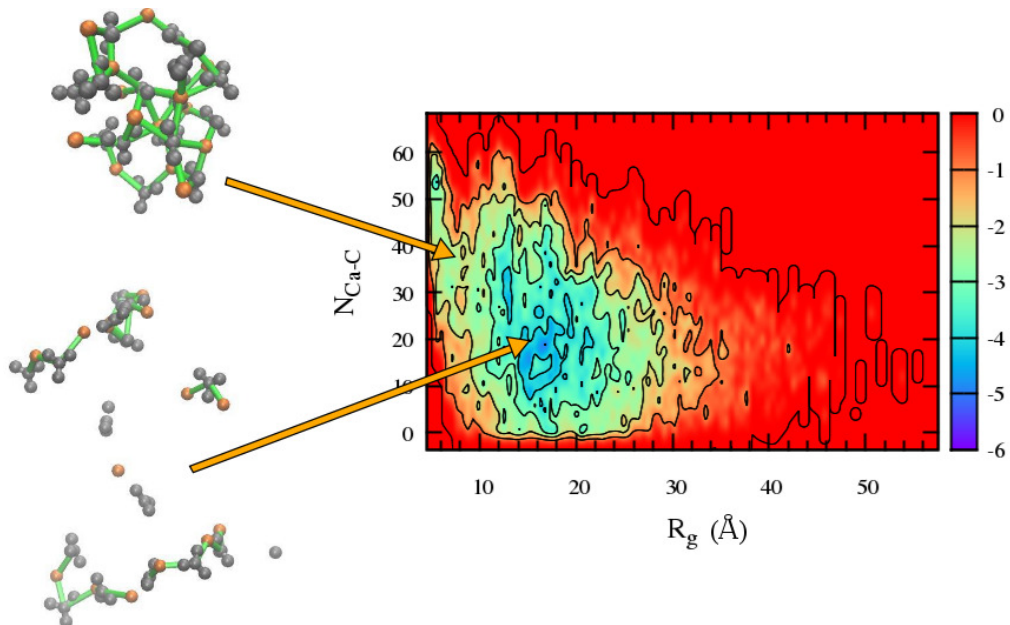
Figure 5.1: The free energy surface for $15\,CaCO_3$ defined by R_g and N_{Ca-C} collective variables. The scale on the right shows the free energy in units of electronvolts. Two representative configurations are shown on the left with arrows pointing to the general region of the free energy surface where they emanated from. Calcium is shown as orange with carbonate in black, while water has been omitted for clarity; green lines highlight distances between calcium and carbon where $r < 3.825$ Å.

the energy minimum can be approximated to be around 2 eV, the noise in the data makes interpreting this difficult.

While N_{Ca-C} is able to accurately define the different states efficiently, R_g is calculated for all cluster atoms in solution, and so a wide range of possible values exist for dissociated clusters (*e.g* a wide distribution of R_g would be observed for a solution of ions), hence the large basin in Figure 5.1. This means that the calculation is somewhat inefficient, and insufficient sampling of R_g is likely to be the cause for the noise in the data.

5.3.2 Biasing Ion Solvation

Water molecules must be removed from the solvation shell of ions in order for cations and anions to bind as calcium carbonate precipitates; ion solvation parameters offer further CVs for biased sampling calculations. A metadynamics calculation was performed on a high density $40CaCO_3$ cluster (as shown in Figure D.2) which was the minimum energy cluster of this size in vacuum. Two CVs were biased in combination: N_{CO_3-Owat} ($n = 4$; $m = 10$; $d_0 = 2.65$; $r_0 = 0.4$) and $N_{Ca-Owat}$

96

($n = 6$; $m = 10$; $d_0 = 2.35$; $r_0 = 0.2$) with $\sigma = 3.5$ and 0.8, respectively. The coordination parameters allowed for biasing of the first solvation shell around ions. The large σ parameter for N_{CO_3-Owat} highlights the large fluctuations in the CV during standard MD simulations (total solvation is taken over four atoms in the anion). Gaussian hills with height 15 $k_B T$ were added every 500 MD steps and the calculation was run for 44 ns.

Figure D.3 provides the values of CVs as a function of simulation time. Correlation of the CVs was found during the sampling, which suggests that these may not be independent; an increase in calcium solvation will likely lead to the release of anions into solution, and so this is not so surprising. The time-dependent CV plots show that there were approximately two regions of sampling with lower values of CVs visited at 20–30 ns; however, there was overlap of the distribution of CVs for both regions. Very large fluctuations were observed in $N_{Ca-Owat}$, highlighting that large changes in the calcium carbonate configuration were found during sampling.

The FES is shown in Figure 5.2 (a). From the initial configuration where $N_{CO_3-Owat} = 146$ and $N_{Ca-Owat} = 49$, the value of both CVs increased rapidly as the coordination of ions decreased. A wide basin can be seen with the minimum in FE at $N_{CO_3-Owat} \approx 550$ and $N_{Ca-Owat} \approx 200$, this equates to, on average, five water molecules surrounding a calcium ion, which is slightly lower than would be expected for the solvation of free ions. Figure 5.2 (c) provides a snapshot of a representative configuration from the minimum of the FES. A large branched cluster can be observed with predominantly two-fold coordination between calcium and carbonate. Surrounding the large low dimensional cluster are a number of smaller clusters and free ions, which again show low coordination and high ion solvation. The configurations in the minimum are indicative of low coordination DLNPs. The difference between this result and that for the 15 formula unit cluster when biasing R_g and N_{Ca-C} could be due to a difference in concentration. The concentration according to the total number of calcium ions in the 40 formula unit system was three times higher than that for the 15 formula unit system (15 $CaCO_3$: [Ca^{2+}]=0.3 M; 40 $CaCO_3$:[Ca^{2+}]=0.9 M), and the increased number of counterions locally may lead to large ionic liquid clusters as opposed to small liquid clusters and ion pairs.

An additional minimum in the FES can be seen at lower values of the CVs. By projecting the FES onto a single CV (see Figure 5.2 (b)), the minimum can be seen at $N_{Ca-Owat} \approx 145$, and this is ~ 2 $k_B T$ per $CaCO_3$ formula unit higher in free energy than the lowest energy state. The barrier to the lowest energy basin is approximately 10 $k_B T$. However, the uncertainty in the FES is on this order, and

<center>(a)</center>

<center>(b)</center>

<center>(c)</center>

<center>(d)</center>
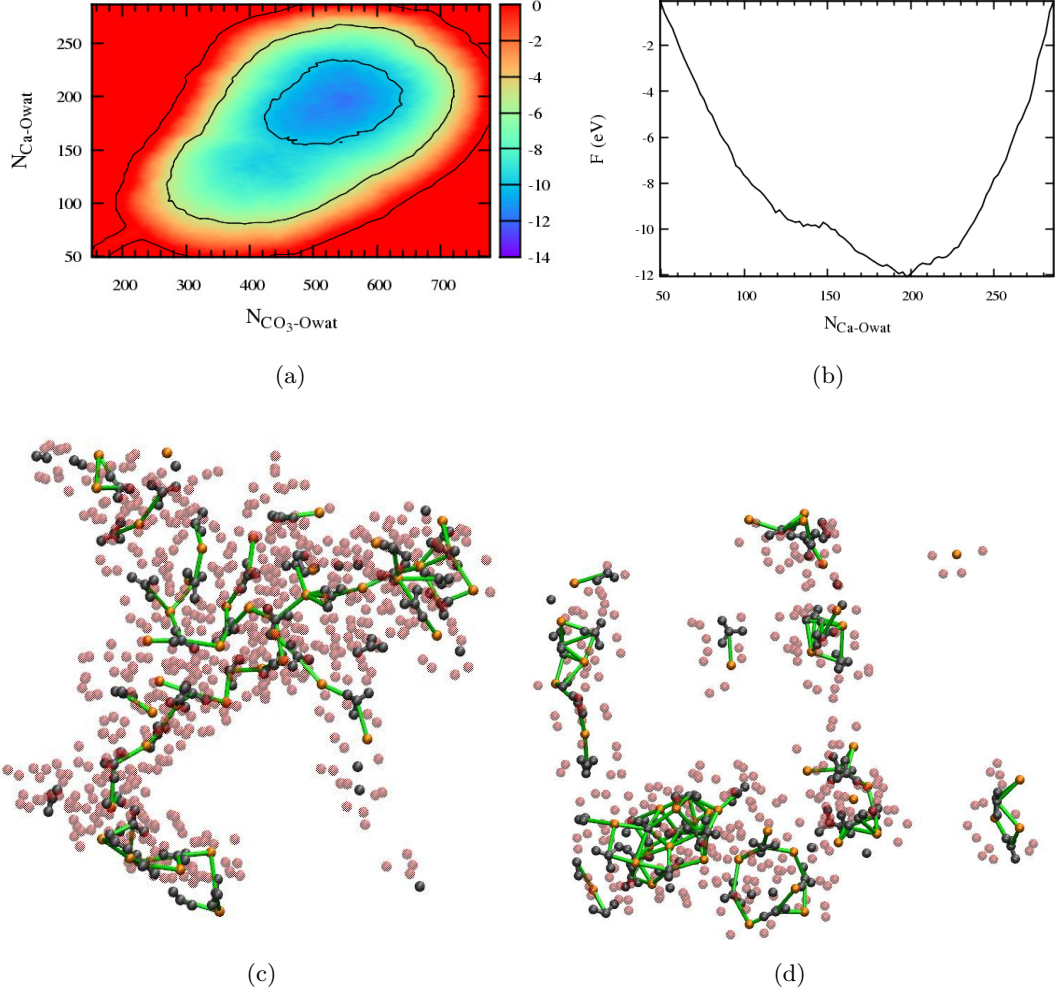
Figure 5.2: Metadynamics simulations of a $40\,CaCO_3$ cluster in water. (a) Free energy for the system as a function of N_{CO_3-Owat} and $N_{Ca-Owat}$ with the energy scale in units of eV. (b) The free energy projected onto $N_{Ca-Owat}$. Representative snapshots of the calcium carbonate species found in the minimum (c) and high energy basin (d) of FE, where the oxygen of water molecules in the first solvation shell of ions are shown in red. For a description of atom and bond colours see Figure 5.1.

so it is not clear whether the two states are separated by an energy barrier or not. A typical configuration from the higher energy minimum is shown in Figure 5.2 (d). The snapshot shows a collection of small clusters with a higher average coordination to the state found in the minimum of the FES. A collection of small clusters which have relatively high coordination still allows for high solvation of ions. This was evident in the case of $6\,CaCO_3$, as discussed in Chapter 4, where clusters with high coordination were found to form and dissolve throughout the trajectories. No minimum for a single cluster of a dense phase, such as hydrated or dry amorphous calcium carbonate (ACC) was found. This may signify that those phases are unstable at the system sizes studied, or could be due to insufficient sampling of the FES; while it was found that the FES had converged locally (because the relative difference in energy between the two identified minima remained unchanged), it is possible that the energy barrier to higher density states exceeds the depth of the FES.

5.3.3 Sampling Ion Coordination

Metadynamics

A metadynamics calculation has been performed on a low density cluster containing ten calcium carbonate units. A single coordination CV between calcium and oxygen of carbonate, N_{Ca-OC}, was used with parameters: $n = 2$; $m = 10$; $d_0 = 2.55$; $r_0 = 0.3$. Ca–OC connectivity is a more appropriate definition of coordination than the previously used Ca–C, as this differentiates between monodentate and bidentate binding of the anion to calcium. Gaussian hills with height 10 $k_B T$ and width $\sigma = 1.0$ were deposited every 400 MD steps. 6.5 ns of trajectory was collected at which point the calculation failed due to large forces on atoms resulting in high energy bonds. Figure D.4 shows that around 6.5 ns, N_{Ca-OC} was being sampled at the high limit of the CV before the calculation failed. The average coordination at this point was Ca–6C, and so this represents a high density configuration. The reason for the calculation failing is not entirely clear, but at the natural limits of the CV there will exist valleys with steep walls in the FE. In an effort to sample more of CV space, it is possible that atoms overlap, and the resulting large forces lead to expansion of the system to dissipate any excess energy. While the forces in the system should be in balance, at these natural limits, deposited Gaussians with relatively large height and weight could lead to this behaviour. Indeed the calculation was observed to fail during the following MD step after bias was added.

The FES from the metadynamics calculation is presented in Figure 5.3. Ini-

tially, most sampling was done in the CV range for free ions in solution, but with the addition of more bias a basin emerged in the range $N_{Ca-OC} = 20 - 35$ (where this is the global value of coordination). Following this, higher values of CV were sampled and a third basin emerged which extended up to the natural limit of the CV. The average coordination in the global minimum of the FES is Ca–(2.2)OC with a $\sim$ 10 eV barrier to Ca–(2.8)OC. This is indicative of a low dimensional cluster (the average coordination in a one dimensional calcium carbonate chain will range between 2–3 for Ca–OC according to the number of bidentate bonds). A representative configuration of a cluster occupying the global minimum is shown in the basin. The cluster configurations were of the type found for this size of cluster in Chapter 4 and is confirmation that at the concentrations studied, DLNPs are likely to form in solution when ions first associate at the concentration studied. The energy barrier from free ions to the dense liquid is minimal. However, once DLNPs form in solution, it is unlikely that a homogeneous solution of free ions will be recovered due to the high energy barrier to dissolution. The basin with a minimum at $N_{Ca-OC} = 50$ is for high density clusters with no internal water; essentially this is a dry amorphous calcium carbonate phase. The energy difference between DLNP(s) and the high density cluster is approximately 10 eV per formula unit, and so dry ACC particles of this size are unlikely to form over a dense liquid in solution. However, if high density clusters are formed, they will represent a long-lived state, as the energy barrier to DLNP(s) is $\sim$ 5 eV per formula unit. No minimum for dense particles with internal water was found, and so it can be concluded that larger clusters are necessary to produce hydrated or anhydrous ACC.
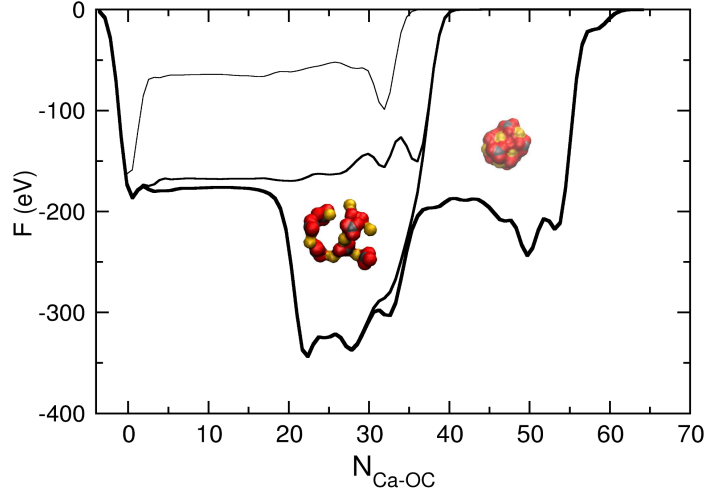
Figure 5.3: The free energy surface for $10\,CaCO_3$ as a function of N_{Ca-C}. Black curves show the time evolution of the FE as more bias was added, with the thickness of lines increasing as a function of time (in increments of $\sim$ 1.5 ns). Inset are representative configurations for the two low energy minima. The surface of calcium, carbon and oxygen is shown as yellow, black and red respectively, while water has been omitted for clarity.

Umbrella Sampling

To investigate whether clusters with densities expected for ACC can be found in the size range suggested for prenucleation clusters, a series of US simulations were carried out on solutions containing a total of 10, 20, 30 and 40 $CaCO_3$ units. To generate the starting configurations for US windows, one dimensional adiabatic biased molecular dynamics (ABMD) simulations were carried out: from the most dense particles, ABMD was performed where the target value was $N_{Ca-OC} = 0$. Figure D.5 shows the time evolution of N_{Ca-OC} during ABMD simulations. The spring constant was chosen to be 0.05 eV, which was increased where necessary if the CV became stationary within noise for long periods. From the trajectory of the dissolution, configurations were extracted such that $\delta = 3.5$ (where δ is the difference in N_{Ca-OC} for successive US window configurations). The large majority of extracted configurations matched the target value; some configurations at high N_{Ca-OC} diverged from the target, but this was never more than $N_{Ca-OC} = 1$. A total of 22, 36, 46 and 56 US starting configurations were extracted for 10, 20, 30 and 40 formula unit systems, respectively.

US simulations were run for approximately 7, 3, 15 and 4 ns for 10, 20, 30

101

and 40 formula unit simulations, respectively. Figures D.6 and D.7 show that the sampling in windows was well performed: significant overlap of N_{Ca-OC} distributions was found between window centres. Subsequently, free energy surfaces were computed using WHAM. A FES was considered to be converged if the free energy at US window centres did not change by more than $k_{\mathrm{B}}T$ in successive 1 ns simulations. Figure D.8 shows the time evolution of the free energies over the final 1 ns of simulation, which highlight that the FE for 10 and 30 formula unit US simulations had converged.

The free energy profiles for the converged calculations are shown in Figure 5.4. For the smallest cluster studied, the FES shows a wide basin with the minimum at $N_{Ca-OC} \approx 20 - 27$ (an average coordination of 2–2.7 per calcium ion). At low values of N_{Ca-OC}, the FES shows a series of successive minima associated with increasing Ca–OC coordination, but at large N_{Ca-OC} this becomes smooth. Inset is a snapshot of a representative structure of a configuration from the minimum with $N_{Ca-OC} = 27.6$. The low energy state is that which was found in Chapter 4, with a number of free ions, ion pairs and small clusters. This further corroborates the observation that at smaller cluster sizes, and at high concentration, DLNPs are likely to be found in solution along with free ions and small associates. The high energy dry ACC state shown in Figure 5.3 was not found, which suggests that the anhydrous ACC from which ABMD was initiated is unstable; structural rearrangement of ions within the anhydrous phase will be slow on the time-scale of the simulations.

For 30 formula units, again a wide basin in a relatively smooth free energy surface was observed. Two minima at relatively low and high N_{Ca-OC} were found, separated by an energy barrier of $\sim 3k_{\mathrm{B}}T$. At low N_{Ca-OC} a number of low coordination clusters, free ions and ion pairs were observed, and this is comparable with the minimum energy state for smaller clusters. Average coordination was around 2.5, which is consistent with the minimum for 10 formula units and for low density clusters of this size simulated with standard MD. The configuration at high N_{Ca-OC} is that of a dense state with a well defined ionic core. At this value of the CV ($N_{Ca-OC} = 127.6$), average coordination per formula unit was approximately 4.3, and analysis of the radial distribution function shows that average Ca–C coordination was 3.7. While the average is relatively small compared with crystalline or ACC phases, this is due to system size effects giving a large surface:volume ratio, and high coordination of ions in the core of the particle were found: visual inspection showed Ca–C coordination numbers of 7–8. This is entirely consistent with high density clusters simulated with standard MD. However, unlike dense dry
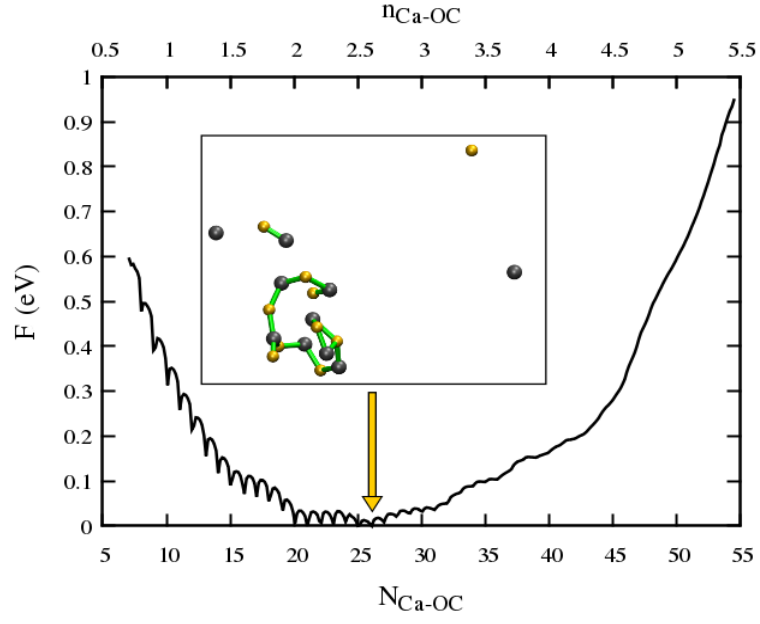
(a)
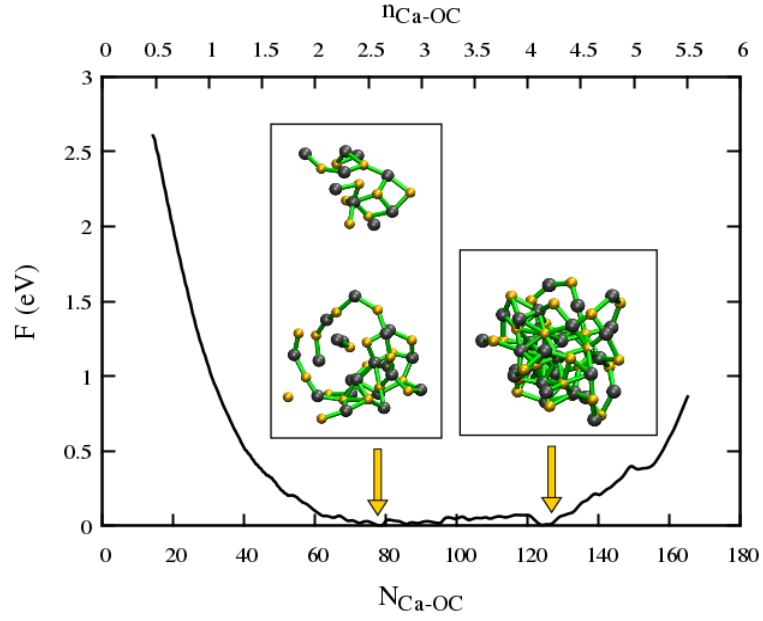
(b)

Figure 5.4: Free energies as a function of N_{Ca-OC} for 10 and 30 $CaCO_3$ formula unit clusters in water, measured from umbrella sampling simulations. Average Ca–OC coordination per formula unit (n_{Ca-OC}) is shown on the x axis. Inset are representative structures for the values of N_{Ca-OC} designated by the arrows. Black and yellow atoms are carbon and calcium respectively. Green lines highlight connections between cation and anion centres where $r < 3.825$ Å.

states, the high density configuration in Figure 5.4 (b) contained a large number of internal water molecules. 27 water molecules were found to be contained within the particle, as shown in Figure 5.5, which gives a stoichiometry of $CaCO_3 \cdot (0.9)H_2O$, and is consistent with that of stable biogenic ACC [Levi-Kalisman *et al.*, 2002]. The size of the particle was measured to be ~ 2 nm in diameter which is congruent with the size of prenucleation clusters measured by analytical ultracentrifuge experiments (although the size measured in experiment was in terms of hydrodynamic diameter) [Gebauer *et al.*, 2008]. The results show that hydrated ACC particles are likely to be found in equilibrium with DLNPs, and, at high concentration, within the size range of prenucleation clusters. While the FES for 40 formula units did not converge at small N_{Ca-OC}, further sampling of windows in Figure D.8 (d) show that the FES has a minimum at relatively large N_{Ca-OC}, with the energy of low density states increasing. The cluster with high coordination was also found to contain a number of water molecules, with the stoichiometry of the particle being approximately equal to that observed in the high coordination minimum for the 30 formula unit FES. It is therefore likely that for particles of $CaCO_3$ containing ≥ 30 formula units, a dense amorphous structure becomes favourable in solution.

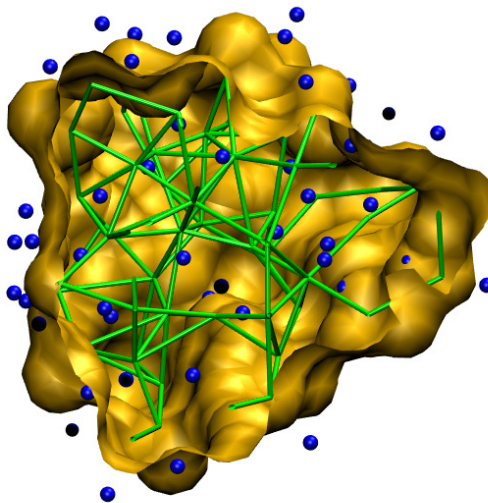Figure 5.5: A snapshot of a $30CaCO_3 \cdot (0.9)H_2O$ particle identified to be stable from Umbrella Sampling with the N_{Ca-OC} CV. The surface of the particle is shown in yellow, with a cutting plane in the x dimension allowing the internal structure to be seen. Green lines show connections between ion centres where $r < 3.825$ Å, and blue spheres are water oxygens which are within 3.2 Å of ions.

5.3.4 Biasing Ion Coordination and Solvation

At the concentrations studied, free ion solutions were found to be higher in energy than DLNPs or high density hydrous phases. However, on close inspection of the configurations where $N_{Ca-OC} = 0$ or $N_{Ca-C} = 0$, a number of ion pairs were observed in the solutions. While the distance between ions was larger than the CV definition, these associates cannot be truly considered as dissociated, as free ion proximities suggest that not all ions were solvent separated. To gain a fuller understanding of the relative free energies between associated ions and free ions in solution, a series of metadynamics simulations have been carried out where Q_0^{Ca-OC} and $Q_0^{Ca-O_{wat}}$ have been biased. r_1 values were 2.65 while r_2 were set to 2.95 and 3.05 for Q_0^{Ca-OC} and $Q_0^{Ca-O_{wat}}$, respectively. Gaussians were deposited every 400 steps with height $10k_BT$ and width $\sigma = 0.015$ (this is equivalent to 0.18 when scaled by N_S). Simulations were run in the NVT ensemble for around 15, 11, 8 and 10 ns for 10, 20, 30 and 40 formula unit simulations, respectively.

Figures D.9 and D.10 show the time evolution of Q_0^{Ca-OC} and $Q_0^{Ca-O_{wat}}$ for all the simulations. These plots highlight that a wide range of states were sampled, with sampling of Q_0^{Ca-OC} in the range 0–5 and $Q_0^{Ca-O_{wat}}$ in the range 0–15. While the distribution of Q_0^{Ca-OC} is expected, and signifies the sampling of high and low density states, the maximum number of water molecules in the solvation shell of calcium was very high. In equilibrated solvent, the number of water molecules found to solvate calcium was up to approximately eight. The RDFs for Ca–OC and Ca–O$_{\text{wat}}$ are shown in Figure 5.6, along with those measured from equilibrium simulations. Extra peaks were observed in both RDFs centred at $r = 3.1$ and 3.2 for Ca–OC and Ca–O$_{\text{wat}}$, which are just outside the values of r_2 specified to define water coordination. Visual inspection of trajectories showed that in a number of configurations free cations in solution were coordinated by water at distances greater than expected for solvation at equilibrium. Furthermore, some configurations showed excessive numbers of water molecules surrounding ions within a distance of 3.0 Å. The results suggest that the forces arising in the metadynamics simulation led to unnatural solvation of ions.

Free energy surfaces measured from 10 and 30 formula unit simulations are provided in Figure 5.7. For the smallest system, three low energy minima were identified at scaled values of $Q_0^{Ca-OC} = 0$, 1.5 and 2.9, representing disassociated ions and two DLNP states, with the lower coordination state approximately 5 eV lower in energy than at $Q_0^{Ca-OC} = 2.9$. The relative free energies for the DLNP with low coordination and disassociated ions is equivalent, and the energy barrier between the states is around 30 eV per unit. No stable high density phase was
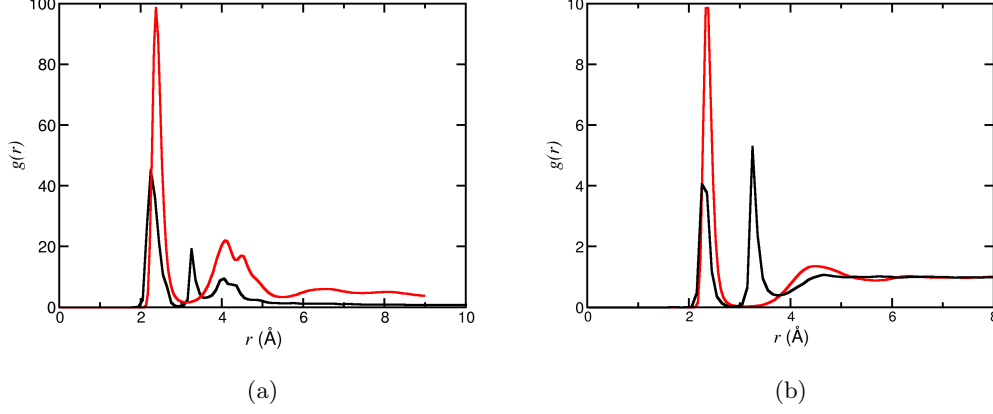
Figure 5.6: Radial distribution functions of Ca–OC and Ca–O$_{\text{wat}}$ measured during metadynamics simulations of 20 CaCO$_3$ in water where Q_0^{Ca-OC} and $Q_0^{Ca-O_{wat}}$ were biased are shown in black. Red curves show the RDFs measured during equilibrium simulations of DLNP clusters taken from Figures A.2 and A.4 for twenty formula unit simulations respectively.
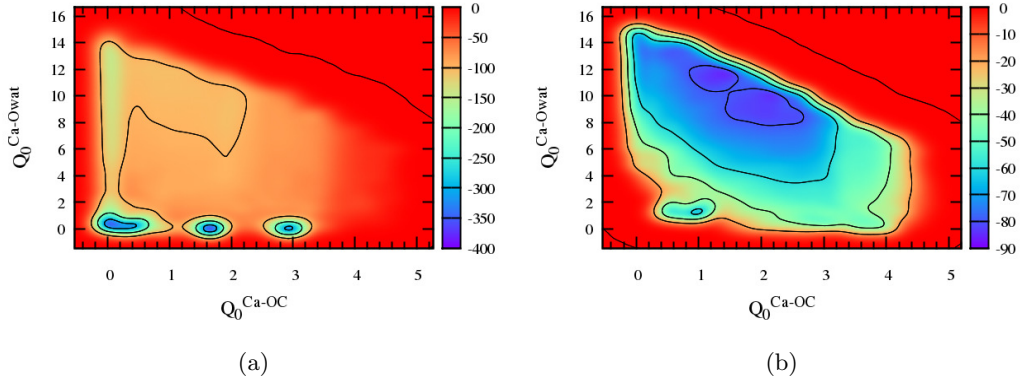
Figure 5.7: Free energy of (a) 10 and (b) 30 calcium carbonate formula units in water as a function of Q_0^{Ca-OC} and $Q_0^{Ca-O_{wat}}$ collective variables from metadynamics simulations. CVs have been scaled by N_S to give the equivalent coordination numbers. The free energy colour scale is shown on the right of plots with units of eV.

identified, unlike the case when N_{Ca-C} was biased. All of the minima are centred at $Q_0^{Ca-O_{wat}} = 0$, and this is due to the artificial solvation of ions caused by the introduction of the bias. For both free ions and DLNPs, a considerable amount of solvent water should surround ions, and so this result cannot be considered as an accurate representation of the FES. For 30 formula units, Figure 5.7 (b) shows a very large basin at values of relatively low Q_0^{Ca-OC} and high $Q_0^{Ca-O_{wat}}$. The free energy surfaces for 20 and 40 formula unit systems were very similar to that of 30 formula units. In Figure 5.7 (b), two minima with energies of ~ -85 eV were found in the basin at $Q_0^{Ca-OC} = 1.25$ and 2.1, and $Q_0^{Ca-O_{wat}} \approx 11$ and 10, respectively. At $Q_0^{Ca-OC} = 1.25$, predominantly ion pairs were found in solution with a small number of small ion associates. An energy barrier on the order of $k_B T$ separated this minima and the one at $Q_0^{Ca-OC} = 2.1$, where DLNP clusters outnumbered ion pairs, and these were, in general, up to ten ions in size and consisted of chains of ions with some branching. A minimum was observed at $Q_0^{Ca-OC} \approx 1$ and $Q_0^{Ca-O_{wat}} \approx 1$. However, visual inspection of the configurations showed a collection of DLNP clusters with relatively large ionic coordination, but with distances between ions larger than was observed at equilibrium. This is further evidence that forces in the simulation led to unnatural coordination environments as the system endeavoured to sample large ranges of CV space.

5.4 Summary and Conclusions

From a number of computational experiments, ion coordination CVs proved to be the most efficient to sampling calcium carbonate associates in solution. The most reliable results from FE calculations are those from US and metadynamics with Ca–C and Ca–OC CVs. Furthermore, defining the coordination by a distance criteria, rather than the distance between atoms and a spherical harmonic function, gave the most credible free energy surfaces. Refinements in the CV definitions may improve the accuracy of free energy calculations. To understand the stability of phases, the results in sections 5.3.3 and 5.3.2 should be considered.

At the concentrations investigated, for small numbers of ions the most likely species to be found in solution are small clusters of associated ions along with free ions in solution. These findings corroborate the findings of Chapter 4 which showed that for less than ~ 15 calcium carbonate formula units, the small ion associates were low in energy and clusters with initially high coordination dissolved to the same collection of states. Metadynamics simulations were performed on 15 and 10 unit clusters where R_g and N_{Ca-C}, as well as only N_{Ca-C} were biased, respectively.

Small ionic clusters were found to be low in energy and in the simulation where one collective variable was biased, the energy difference between this state and a high density anhydrous (in terms of the internal stoichiometry) cluster was approximately 10 eV per formula unit of calcium carbonate.

On increasing cluster size, and also concentration, the free energy differences between low and high coordination phases decreased. From US simulations, a high coordination particle, which contained water molecules with a hydration level close to stable hydrated ACC, was found to be as thermodynamically stable as the small ion associate clusters, and the energy barrier between the states was on the order of $k_{\mathrm{B}}T$. For the largest clusters studied, a wide basin was found with a single minimum at $N_{Ca-C} \approx 150$ (*i.e.* four carbonates per calcium on average), and while the FES did not converge at small N_{Ca-C} during the simulation, the position of the minimum at large N_{Ca-C} did not change. The cluster configurations found at the minimum were again particulate with a well defined core and rough surface. As for the 30 formula unit system, a number of water molecules were observed below the particle surface. Interestingly, for these dense hydrated particles, clusters were aspherical: oblate particles have a large surface area and therefore increased solvation of ions. As clusters increase in size this suggests that there is a transition between low coordination dynamic species and dense hydrated particles. The effect of concentration is also worth considering. As the size of clusters increased, so too did the concentration, and so the transition to dense solid phases may be associated with changing concentration. Further simulations to validate the cause of the change are necessary; calculation of the FES for $30\,CaCO_3$ at the same concentration as $10\,CaCO_3$ would indicate whether system size or concentration led to changes in FE.

When metadynamics simulations were performed with biasing of ion solvation, a dense particle was not observed. However, the local FE converged to identify two distinct states: that of a large DLNP with additional small ion associates which were low in energy, and a slightly higher energy state with a number of clusters with increased density; the energy difference being $\sim 2k_{\mathrm{B}}T$ per formula unit. In the high density clusters, coordination numbers of individual ions reached levels found in meta(stable) amorphous particles (see Chapter 4). The nature of the high density clusters needs to be ascertained further by analysis of ion mobility within clusters, and by measurement of calcium carbon "bond" lifetimes in comparison with the clusters simulated in Chapter 4. However, in the high energy minimum, DLNPs are found in solution alongside what appear to be small amorphous clusters, and this result may be important to biomineralisation where additives may remove energy

barriers to high density particles. Investigations of the effect of organic molecules found in biogenic calcium carbonate might shed light on the significance of the higher energy state to calcification in biology.

Finding amorphous calcium carbonate clusters to be stable in solution is not so surprising, as the amorphous phase has been identified as the first deposited solid product by experiments [Addadi *et al.*, 2003]. Computational free energy studies have identified ACC to be more stable than crystalline phases [Freeman *et al.*, 2010; Quigley *et al.*, 2011; Raiteri and Gale, 2010]. Hydrated ACC has been shown to form in biogenic samples before dehydration and subsequent crystallisation to calcite, and so the results found here for large particles are consistent with this observation [Gong *et al.*, 2012]. What is surprising is that the size of these dense particles is within the range expected for prenucleation clusters. Gebauer and co-workers identified 2 nm particles in solution which were suggested to aggregate to form an amorphous phase [Gebauer *et al.*, 2008], and so it is possible that these clusters are hydrated ACC. Tribello and colleagues [Tribello *et al.*, 2009] provided evidence to show how water can become kinetically trapped during the growth of ACC, and highlighted that hydrated ACC wins out to anhydrous ACC in solution because of this. However, in the results discussed here, anhydrous particles were found to be much higher in energy than ones containing water, meaning that at the system sizes and concentrations investigated, hydrated ACC is more stable. The stability of anhydrous ACC has been shown for bulk systems, and so the particle size or presence of impurities may be the trigger for the phase transition [Politi *et al.*, 2008].

The results here suggest that the first product of ion association will be ion-pairs, which have no barrier to formation. The resulting neutral associates will then agglomerate, and this is likely to be diffusion limited as the dipole–dipole interactions between ion-pairs will be screened by solvent. DLNPs will form from a small number of ions, but as further ions are added to the cluster and the concentration is increased, hydrated ACC phases will become more stable within the size range identified for prenucleation clusters. Further growth of hydrated ACC particles cannot be speculated upon from the data here, but other studies have shown no energy barrier to the addition of ion pairs to ACC [Raiteri and Gale, 2010].

Chapter 6

Solution and Speciation Effects

Declaration

A selection of the material in the following Chapter has been written into the submitted manuscript: P. J. M. Smeets, A. R. Finney, W. J. E. M. Habraken, F. Nudelman, H. Friedrich, J. Laven, P. M. Roger, N. A. J. M. Sommerdijk. A classical view on non-classical nucleation.

In the following Chapter, experiments, presented in section 6.4, were performed by Paul J. Smeets and colleagues working under the supervision of dr. Nico A.J.M. Sommerdijk from *The Department of Chemical Engineering and Chemistry, Eindhoven University of Technology, The Netherlands.*

6.1   Introduction

The equilibrium between dense liquid nanoparticles (DLNPs) and constituent ions at experimental concentrations has not yet been determined. Furthermore, ions which are found in solution in experiment have been omitted from the simulations in previous Chapters. This does not detract from the usefulness of such calculations. It is common practice to simulate molecular systems at relatively high concentrations to save computational resources, as solution equilibria can be achieved more quickly.

Combining classical simulation and experimental studies offers a powerful approach to elucidating molecular processes [Mayor *et al.*, 2003; Morris *et al.*, 2010; Lai *et al.*, 2010]. This can be applied in the pursuit of calcium carbonate nucleation mechanisms from constituent ions in solution. Success in this, from a simulation perspective, relies on the proper inclusion of contributing species to precipitation, and a suitable force field (and algorithm) to model the chemistry. With this in mind,

a series of computational experiments have been conducted and the results, combined with those of *in vitro* studies made by collaborators, allow for a reassessment of the stabilities of species emerging from calcium and carbonate solutions.

In the following Chapter a series of simulations investigating ion aggregation in solution at millimolar concentrations and a range of pH are reported. Cluster simulations were also performed to assess the extent of the equilibrium of free ions with DLNPs, and to investigate concentration effects. Comparison of the data with experimental studies is performed, and the implications for the nucleation mechanism of calcium carbonate are discussed.

Experiments by collaborators were based on the ground-breaking study by Gebauer and co–workers [Gebauer *et al.*, 2008] which involved a series of titrations: 10^{-3} mol dm^{-3} calcium chloride solution was titrated into a buffer solution of 10^{-3} mol dm^{-3} sodium carbonate in the range pH 9–10. Simulations under basic conditions and millimolar concentrations will therefore allow a more direct comparison to experiment. Treatment of pH in simulation can be difficult, so it is useful to discuss this in more detail.

6.1.1 Modelling pH

An ideal and full electronic treatment of solution species would allow for accurate modelling of speciation at equilibrium, but the computational cost involved makes this unrealisable to study ion by ion growth at experimental concentrations. In a recent study employing Born-Oppenheimer Molecular Dynamics, the structure and dynamics of solvated anions was investigated, but calculations containing a single ion and 124 water molecules were limited to 205 ps of total simulation due to the computational effort involved [Zhao *et al.*, 2013].

Alternatively, bond order force fields can be employed. ReaxFF is arguably the most common of these, which makes use of the initial assumption that all atom pairs are non–bonded [van Duin *et al.*, 2001]. Through calculation of van der Waals and Coulombic interactions, bond ordering can be estimated according to atom distances and pairwise potential energies. Importantly, sigma and pi bond ordering can be incorporated into the force field. While significantly cheaper than *ab initio* treatment of atomic configurations, ReaxFF is typically an order of magnitude slower than simulations with empirical atomistic force fields.

Demichelis *et al.* made use of an atomistic force field with fixed bonding, but approximated the pH of solution by using ratios of the total numbers of protonated and deprotonated carbonate ions at the beginning of MD simulations [Demichelis *et al.*, 2011]. From the experimental acid dissociation constant for bicarbonate in

water ($pK_a = 10.329$), pH for the initial configuration can be estimated. By varying $[CO_3^{2-}]$:$[HCO_3^-]$ over a number of simulations, the effects of pH can be inferred. One caveat to this method is that the change of pK_a for carbonate in different phases of calcium carbonate cannot be accounted for. However, the practicalities of simulating millimolar solutions for nanoseconds using MD means this approach is a reasonable one, and as such has been employed in this Chapter.

6.2 Methods

6.2.1 Simulation

Ionic solutions were prepared as shown in Table 6.1, by randomly populating a cubic simulation cell with calcium, carbonate and bicarbonate ions, such that the distance between two ions was > 5 Å. The box was filled with a number of pre-equilibrated water molecules which conserved target concentrations according to molar ratios. The total number of carbonate and bicarbonate ions in the simulation cell was used to calculate the total carbon concentration, $[C]$. The number of bicarbonate and carbonate ions remained fixed during the simulations. The carbonate:bicarbonate ratio was calculated from the target pH using the Henderson–Hasselbach equation,

$$pH = pK_a + \log\frac{\left[CO_3^{2-}\right]}{\left[HCO_3^-\right]}, \tag{6.1}$$

and the pK_a of carbonate under standard conditions: $pK_a = 10.329$.

Table 6.1: Details of simulations of free ions in aqueous solution. The number of carbonate, bicarbonate and calcium ions are given by n_{Carb}, n_{Bicarb} and n_{Ca}, while the number of water molecules is labelled n_{Wat}. The system label gives the type of system (I representing free ions), number of anions and number of cations in solution (e.g. system I_{22-14} is an ionic solution containing 22 anions and 14 cations).

System	n_{Carb}	n_{Bicarb}	n_{Ca}	n_{Wat}	$[C]$/mM	$[Ca^{2+}]$/mM	pH
I_{22-14}	6	16	14	24424	50.5	32.1	9.90
I_{46-29}	12	34	29	83132	31.0	19.6	9.88
I_{64-47}	30	34	47	100000	35.9	26.4	10.27
$I_{100-100}$	100	0	100	10000	566	566	–

A number of cluster solution simulations were also performed. Configurations for clusters were taken where possible from earlier cluster optimisation studies as discussed in Chapter 3. Where this was not possible, random structure searching was performed. Approximately 1,000 structures were optimised and subjected to

the previously employed classification criteria (energy, coordination and shape order parameters). In all cluster simulations, low density clusters of lowest energy were sampled and solvated. Low coordination clusters were chosen as these were found to initially form from the association of free ions [Demichelis *et al.*, 2011]. The set-up for cluster solution simulations are detailed in Table 6.2. As all ions were bound in clusters, the ionic concentration is effectively zero at the beginning of simulations. However, the carbon and calcium concentrations, as well as pH can still be defined by the total number of ions in the simulation cell.

Table 6.2: Details of simulations of clusters in aqueous solution. Column labels are the same as in Table 6.1. The system label gives the type of system (C representing clusters), number of anions and number of cations in solution (*e.g.* system C_{22-14} is an ionic solution containing 22 anions and 14 cations)

System	n_{Carb}	n_{Bicarb}	n_{Ca}	n_{Wat}	[C]/mM	[Ca^{2+}]/mM	pH
C_{22-14}	6	16	14	24424	50.5	32.1	9.90
C_{20-20}	20	0	20	55510	20.2	20.2	–
C_{20-20}	20	0	20	37006	30.4	30.4	–
C_{20-20}	20	0	20	27755	40.5	40.5	–
C_{20-20}	20	0	20	22204	50.6	50.6	–
C_{30-30}	30	0	30	83264	20.2	20.2	–
C_{40-40}	40	0	40	111019	20.2	20.2	–

The force field employed was that of Demichelis *et al.* [Demichelis *et al.*, 2011] with truncated intermolecular potentials at 9.0 Å. Simulations were performed in the NpT ensemble at 298 K and 1 atm with a Nosé–Hoover thermostat and barostat (with 0.1 and 1.0 ps relaxation times), and electrostatics were treated using the smooth particle mesh Ewald method [Essmann *et al.*, 1995]. The timestep was taken to be 1 fs and simulations were run for 15 ns in the case of ion solutions, and up to 50 ns for cluster simulations. Owing to the large system sizes, calculations were performed using DL_POLY 4.05 on local and national computing facilities. Equilibrium was taken as the time after which cluster size distributions (CSDs) and coordination probabilities, as well as energetic and volume averages, remained unchanged within statistical noise. Equilibrium averages were therefore taken from the final portions of trajectories. The coordination between ions was defined by a distance criterion of $r_{Ca-C} < 3.825$ Å, where r_{Ca-C} is the distance between calcium and carbon atoms. This distance was the minimum after the first peak in Ca–C radial distribution functions for DLNPs and dense amorphous clusters in high concentration simulations (see appendix A).

6.2.2 Experiment

Solutions were freshly prepared for experiments using Milli-Q water. Sodium hydroxide pellets were dissolved into water to produce a 10 mM soldium hydroxide solution. Calcium chloride dihydrate was dissolved into CO_2-purified water to give a 10 mM calcium chloride solution. Carbonate buffer was prepared by mixing 10 mM sodium carbonate and sodium bicarbonate to obtain the desired pH. 10mM hydrochloric acid was prepared by diluting HCl in water.

Polyvinyl chloride membrane calcium and sodium ion-selective electrodes (ISE) were used in the titration set-up. These were calibrated using constant pH solutions (measured using a calibrated pH meter) in the concentration range that was expected for the experiment. For example, the Ca^{2+}-ISE was calibrated by addition of pH=9.75 calcium chloride solution at 10 μL min^{-1} into 25 mL of nitrogen-bubbled Millipore water. pH was maintained by titrating 10 mM sodium hydroxide and 10 mM hydrochloric acid into the solution. The calibration for the Na^+-ISE was similar to that for calcium, but 1 M sodium chloride was used (as opposed to a 10 mM calcium chloride solution), as this better matched the experimental concentrations.

A 10 mM $CaCl_2$ solution was titrated at 10 μL min^{-1} into 25 mL of 10 mM carbonate buffer solution under standard conditions. pH was held constant at 9.75 by addition of 10 mM sodium hydroxide solution, and monitored using a pre-calibrated pH electrode. The calcium and sodium potentials were measured using the ISEs to determine solution concentrations.

Fourier transform infra-red (FTIR) spectroscopy was performed on the samples using a Varian 670 FTIR Spectrometer. Samples of 1 mL were extracted and analysed under standard conditions. The buffer spectrum was created using nitrogen saturated Millipore water.

A scanning electron microscope (SEM) was used to image objects in solution. High magnification SEM images were recorded using a Quanta 3D FEG (FEI, The Netherlands) with a field emission electron gun as the electron source at an acceleration voltage of 5 kV and a current of 1.48–2.96 pA. Samples were taken from solution and put onto carbon tape for 24 hours. Once dry, the samples were sputter coated with a layer of gold, applying a 20 mA current for a duration of three minutes.

Cryotransmission electron microscopy (cryo-TEM) was used to analyse samples throughout the titration. 3 μL samples were extracted from solution and vitrified by inserting onto a grid inside an automated vitrification robot (FEI Vitrobot^{TM} Mark III) chamber at 100% relative humidity and 293 K; after blotting to remove excess solution the sample was plunged into liquid ethane. High resolution trans-

mission electron microscopy (HRTEM) was performed using a TU/e CryoTitan equipped with a post column Gatan Energy Filter (GIF), and a field emission gun operating at 300 kV was used for higher resolution imaging. The images were then recorded with the use of a post-GIF 2k by 2k pixel MultiScan^{TM} Gatan CCD camera. For HRTEM with selected area electron diffraction (SAED) a Titan Krios^{TM} transmission electron microscope (FEI) was used, equipped with a field emission gun and operating at 300 kV. Image recording was performed using a 4k by 4k Eagle CCD camera.

6.3 Results

6.3.1 Systems at Basic pH

One cluster and three free ion solutions were simulated where both bicarbonate and carbonate were present in solution. The initial concentration and pH was taken to be as close to the experimental values of 10 mM and pH=9.75 as computationally feasible with the resources available. However, the actual concentration of ions in solution can only be measured at equilibrium. As the number of free ions in solution changes, so does the pH, and so the solution concentrations and pH were measured from the final 5 ns of trajectories and are reported in Table 6.3.

Table 6.3: Concentrations and pH for systems simulated under basic conditions, measured after equilibration in water.

System	[C]/mM	[Ca^{2+}]/mM	pH
I$_{22-14}$	26.4 ± 1.3	5.1 ± 2.4	7.8 ± 1.2
C$_{22-14}$	23.8 ± 1.4	3.3 ± 1.2	8.4 ± 0.7
I$_{46-29}$	16.4 ± 0.7	4.5 ± 0.8	7.4 ± 1.5
I$_{64-47}$	15.5 ± 0.6	2.9 ± 0.4	8.3 ± 0.8

In all free ion solutions, ions were observed to associate very early on in the simulations. Figure E.1 (a) shows the most common species found in solution once equilibrium had been achieved in system I$_{22-14}$. Ion pairs CaCO$_3^0$ and CaHCO$_3^+$ quickly emerged, and the relative concentration of these species remained high throughout the simulations. Occasional aggregation of ion pairs and free ions was found, and this led to larger ion associates in solution, however the lifetime and concentration of these species tended to be lower than ion pairs.

Figure 6.1 shows the cluster composition probabilities and CSD calculated over the final 5 ns of simulation for system I$_{22-14}$. The most probable cluster composition supports the visual observation that ion pairs dominated in solution,
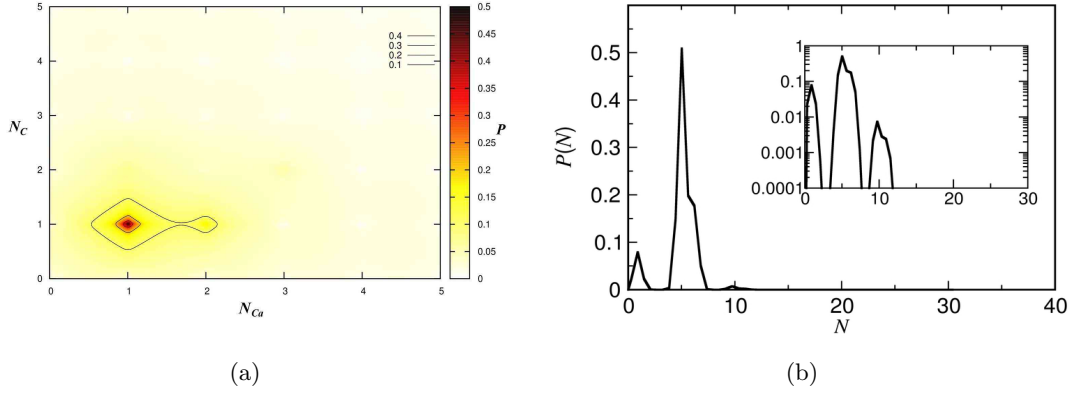
Figure 6.1: (a) Cluster composition probabilities for clusters containing N_{Ca} cations and N_C anions, and (b) a cluster size distribution for clusters containing N atoms. The data was taken during the final 5 ns of a 15 ns simulation of system I_{22-14}. Inset of (b) is the same probability distribution on a logarithmic scale.

and inspection of the CSD suggests that these were more probable than free ions at a concentration of 26.4 ± 1.3 mM. A second peak in the composition probability shows that larger clusters were predominantly ion trimers containing a single anion and two cations. The largest clusters, as identified from the CSD contained 10–11 atoms, which could be due to the presence of clusters containing two ion pairs. The probability of finding the cluster was very low (~ 0.01), as shown on the plot inset of Figure 6.1 (b).

Convergence of the CSD after ~ 3 ns of simulation suggests that larger ionic networks/clusters are not stable in solution at the concentration studied. Demichelis *et al.* found the spontaneous formation and growth of dynamically ordered liquid-like oxyanion polymer (DOLLOP) [Demichelis *et al.*, 2011]. However, their simulations were at higher concentrations (at least 60 mM). It is possible that liquid clusters are more stable than ion pairs in system I_{22-14}, but that longer simulation times are required to sample such states. The dynamics of the cluster (shown in Figure E.1 (b)) in system C22-12 showed that this was not the case. The cluster was observed to dissolve rapidly in the same volume of water as for system I_{22-14}; the time evolution of nC coordinating to Ca is provided in Figure 6.2 (a). Single (bi)carbonate binding to calcium increased rapidly at the beginning of the simulation, reaching a probability of 0.9 around 7.5 ns. This coincided with a large reduction in the probability for calcium to bind to two and three carbonates. By 10 ns, the maximum coordination number for calcium was two. CSDs provided in
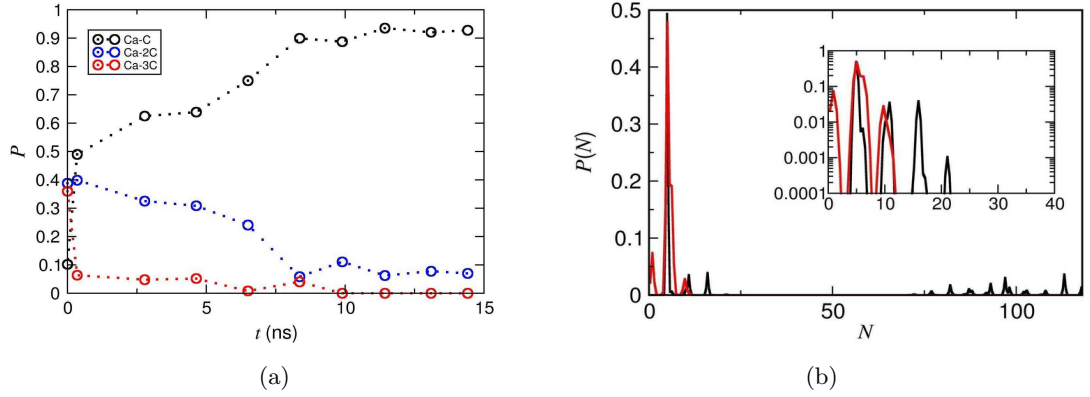
Figure 6.2: (a) Coordination probabilities of Ca coordination to nC as a function of time, and (b) a cluster size distribution for clusters containing N atoms, measured from the initial 0.8 ns (black) and final 2.5 ns (red) of a 15 ns simulation of system C_{22-14}. Inset of (b) is the probability distribution on a logarithmic scale.

Figure 6.2 (b) had converged to those measured for system I_{22-14} at equilibrium after 10 ns. The species found in solution matched those of the ionic solution system, and composition probabilities highlighted an increased probability of finding ion pairs over larger associates.

A lack of large clusters in systems I_{22-14} and C_{22-14} could be a result of the limited number of ions in solution due to the system size: only six carbonate ions were present in both simulation cells. Carbonate offers higher bond ordering in solution than bicarbonate and leads to larger, more stable clusters at high concentration [Demichelis *et al.*, 2011]. An excess of bicarbonate will therefore limit the growth of clusters emerging from free ions in solution. The carbonates found in solution in these systems was preferentially bound to calcium leading to an excess number of free bicarbonate ions, as was indicated by a reduction in the pH at equilibrium (see Table 6.3). The pH therefore deviated from the target value.

Systems I_{46-29} and I_{64-47} were simulated to address these concerns. I_{46-29} contained 12 carbonate ions, keeping the target pH the same as in the smaller systems (pH=9.9), while I_{64-47} contained a higher ratio of CO_3^{2-}:HCO_3^-, and had a target pH of 10.3. Figure 6.3 provides the coordination probabilities for all systems shown in Table 6.3. The coordination probabilities for I_{22-14} showed a preference for $CaCO_3^0$: the probability for calcium to bind to a carbon of carbonate was approximately 0.6, compared to 0.3 for carbon in bicarbonate, and 66% of ion pairs were of this type. For the cluster simulation (system C_{22-14}), the probability of

117
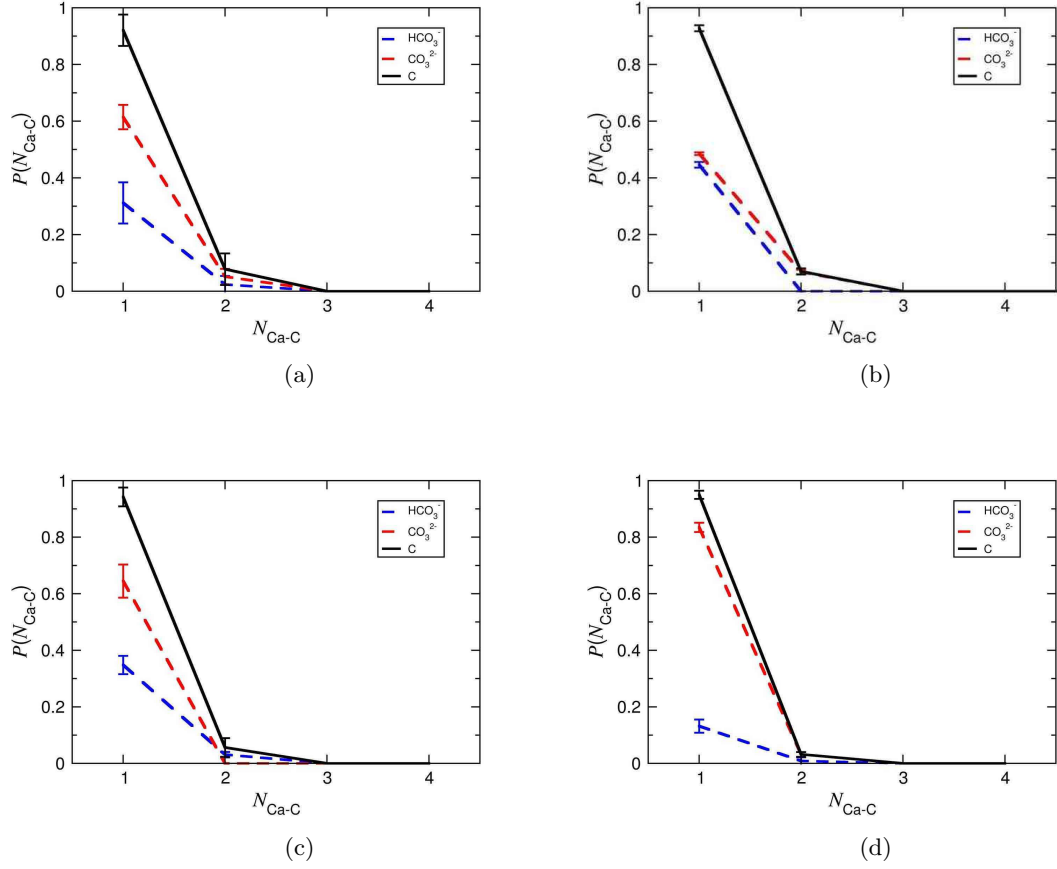
Figure 6.3: Coordination probabilities calculated from the final 5 ns of 15 ns simulations of systems (a) I_{22-14}, (b) C_{22-14}, (c) I_{46-29} and (d) I_{64-47}. Uncertainties of one σ are shown.

calcium binding to a single carbonate or bicarbonate was approximately equal, and there was a reduction in the amount of calcium coordinating to two bicarbonates compared to system I_{22-14}. The data highlights that diffusion-limited exchange of ions at these concentrations is slow, and further sampling may lead to the convergence of plots in Figures 6.3 (a) and (b). In both I_{22-14} and C_{22-14}, around 7% of bound calcium was coordinated to two carbons.

As the number of carbonate ions in solution increased (Figure 6.3 (c)), the coordination probabilities and CSD (see Figure E.2 (a)) remained largely unchanged. The cluster composition probabilities for system I_{46-29} showed a single peak for calcium bound to a single anion. The system size therefore played no role in the limiting size of ion associates. An increase in the amount of carbonate in solution

in system I_{64-47} did lead to an increase in the final pH compared with other ionic solution systems; however, the pH measured at equilibrium (pH=8.3 ± 0.8) was still considerably lower than the target pH. The probability for calcium to bind to carbonate increased compared to systems at lower initial pH. 87% of calcium bound to a single anion coordinated with carbonate. The presence of increased proportions of carbonate did not lead to larger clusters in solution, and there was a reduction in the probability of calcium coordinating to two anions. The maximum peak in the CSD for this system was centred at 11 atoms as shown in Figure E.2 (b), and the composition probabilities suggested that the clusters formed from the largest number of ions accounted for 0.15% of bound calcium.

It is sensible to conclude that over the pH range studied (between pH 7–9), and at concentrations close to experiment, extended networks of coordinated ions and DLNPs are not expected to be present in solution. If these species are formed, it is likely they will dissolve. It is sensible to speculate that the lifetime of larger clusters will increase as a function of increasing pH and concentration, where there is an increased amount of carbonate in solution.

This is in stark contrast to the conclusions of Demichelis et $al.$ who note the increasing size of DOLLOP during simulation is evidence for its stability [Demichelis et $al.$, 2014]. Over the whole pH range (8.5–11.5), DOLLOP was found in solution; however, close inspection of the data suggests that this was not the case for all concentrations. At 0.5 M, ion pairs were the most probable species containing bound calcium, and for larger clusters, calcium preferentially bound to two anions, with bicarbonate and carbonate forming the majority of the polymer chain at pH=9.5 and 10.0 respectively. However, at 0.06 M (the lowest concentration where ion aggregation was studied), while ion pairs again dominated in solution, the fraction of calcium bound to two anions was significantly reduced ($P_{Ca-2C} \approx 0.1$ for pH=9.5 and $P_{Ca-2C} \approx 0.05$ for pH=10.0; see [Demichelis et $al.$, 2011] Supplementary Figure S4). The conditional probabilities (where ion pairs are omitted in the calculation) show that all of the carbon species for Ca-2C in clusters at pH=9.5 are bicarbonate (see [Demichelis et $al.$, 2011] Supplementary Figure S5). The probability that bicarbonate binds to calcium in these clusters is ≈ 0.05. Carbonate did bind in clusters to two calcium ions, but these must have been at the "chain end" as calcium did not bind to two carbonates. At pH=10.0, the probability for calcium to bind to carbonate or bicarbonate was approximately 0.05. Only carbonate was observed to bind to two cations with significant probability at this pH. This all points to the fact that at 0.06 M and the pH values discussed, the vast majority of clusters in solution were ion pairs and small ion associates (such as ion trimers), and that DOLLOP

was not likely to be found in significant concentrations.

6.3.2   Ion Solutions at High Concentration

To link between the dense liquid clusters in Chapter 4 where DLNPs were found to be (meta)stable in simulations over multiple nanoseconds, a simulation starting from a random distribution of solvated ions at high concentration was performed. A solution of calcium and carbonate ions was prepared at 0.57 M and MD simulations were conducted. 40 ns of trajectory was collected in order to establish an equilibrium before averages were collected from a subsequent 2 ns window.

Ion association was observed very early on in the trajectory and clusters grew and aggregated. By 40 ns, a large dense liquid nanoparticle (DLNP) was evident in solution along with a number of smaller clusters (see Figure 6.4 (d)) with sizes and compositions comparable to those found at the lowest concentrations investigated in this Chapter. The cluster size distribution in Figure 6.4 (a) highlights the presence of a large cluster with a wide size distribution, ranging between 300 and 400 atoms (equivalent to approximately 60–80 ions), along with a high probability of finding clusters containing < 80 atoms. The wide distribution in cluster sizes is due to the dynamic nature of coordination. To examine this further, the distribution of bond lifetimes was measured from the 2 ns window, and for this the distance criteria for coordination was increased to 4.2 Å in an effort to reduce the noise in the data. Figure 6.4 (b) gives the probability densities for the lifetimes of bonds which were broken during simulation (9% of bonds remained fixed over the 2 ns window). Bond breaking was prolific during simulation, with connections between ions most likely lasting for < 50 ps. This region is dominated by the bond breaking in small clusters and at the periphery of the large cluster, where ions with small coordination numbers are found and bond breaking is more prevalent. At the centre of the largest cluster, where the local ionic concentration is large, changes in bond orders are minimised by the strong coordination of ions. A smooth distribution is found up to the maximum bond lifetime of 0.28 ns.

Figure 6.4 (c) shows the coordination probabilities for associated ions in solution, the data differs to that presented by Demichelis *et al.* for=pH 10 (as shown in [Demichelis *et al.*, 2011] Supplementary Figure S2), where the coordination probabilities of calcium to carbon were $P_{Ca-C} \sim 0.4$; $P_{Ca-2C} \sim 0.45$; $P_{Ca-3C} \sim 0.05$; $P_{Ca-4C} = 0$ (in the absence of ion pairs). The difference could be due to the presence of bicarbonate in previous studies which will limit the extent of bond ordering in liquid clusters (bicarbonate terminates chains of ions). However, the data does show that at high concentration, dense liquids are likely to form in solution,

(a)

(b)

(c)

(d)
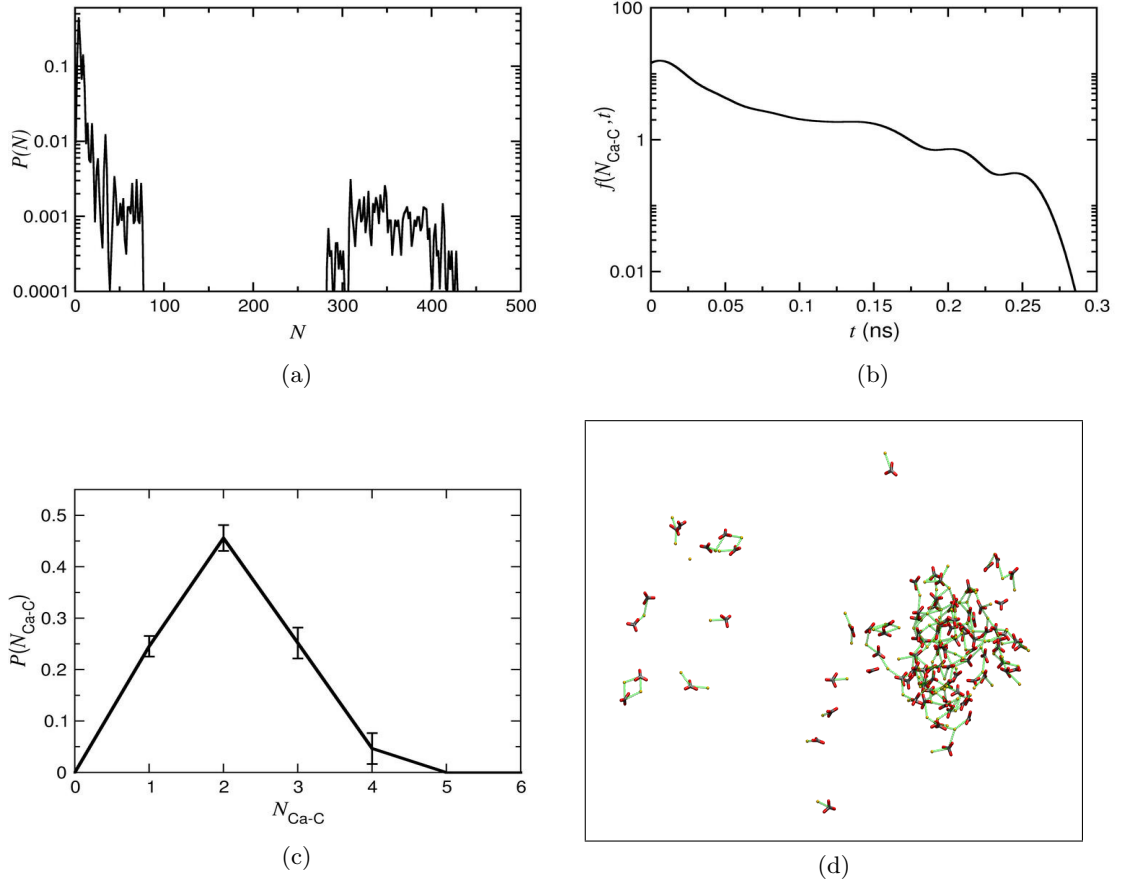
Figure 6.4: Simulation of 100 Ca^{2+} and 100 CO$_3^{2-}$ in water at 0.57 M. Averages were taken from a 2 ns window following 40 ns of simulation. (a): Cluster size probability distribution for clusters containing N atoms. (b): Probability densities for bond lifetimes between calcium and carbon defined by a distance criteria of $r_{Ca-C} < 4.2$ Å. (c): Coordination probabilities for N carbon atoms binding to calcium. (d): A snapshot of the system after 42 ns of simulation; calcium, carbon and oxygen are coloured yellow, black and red respectively, while green lines show connections between calcium and carbon ($r_{Ca-C} < 4.2$ Å).

and these can reach considerable sizes if enough ions are present locally.

6.3.3 Cluster Simulations at the Limit of High pH

To understand the equilibrium between ion pairs and a dense liquid phase, a number of simulations were conducted in the concentration range 20–51 mM (see Table 6.2). Liquid clusters have been shown to be prevalent in solution close to the upper limit of this concentration range at high pH [Demichelis *et al.*, 2011]. The C_{20-20} system was set up for this purpose. In this system, no bicarbonate ions were included, as carbonates lead to a more stable, long-lived cluster, and so simulations in the absence of bicarbonate should overestimate the equilibrium sizes and lifetimes of any dense liquid clusters. Furthermore, as the formation of a dense liquid phase is diffusion limited, simulations initiated where there is a high local concentration of ions should allow an equilibrium to be established relatively quickly.

While a low density cluster was sampled for the current study, this had been optimised in vacuum (see Chapter 3), and so the coordination of ions was relatively large compared with those suggested for dense liquid clusters [Demichelis *et al.*, 2011; Wallace *et al.*, 2013]. The cluster was therefore relaxed at 20 mM for 15 ns. After this period, the coordination probabilities showed (see Figure E.3) a preference for Ca–2C binding, along with coordination of calcium to three and four anions. Partial dissociation was observed with a small cluster $(2\,CaCO_3)$ dissolving into solution. The largest cluster contained 36 ions, for which a snapshot is provided in Figure E.3. Subsequently, water was removed from the final configuration to produce systems with concentrations 20-51 mM. A short 300 ps simulation was performed to relax water with ion coordinates frozen. This was taken to be $t = 0$, from which 27–50 ns simulations, with all atoms mobile, were performed.

The mass weighted average cluster size, S, was calculated: $S = \sum sW_s / \sum W_s$ where s and W_s are the number of atoms and weight fraction of clusters, and the sum is over all clusters in each configuration. The time dependence of ln$S$ as a function of concentration is presented in Figure 6.5 (a). It is clear from the plot that S decreases exponentially with time at all concentrations. The simulations did not reach an equilibrium by the end of up to 50 ns simulations, and further dissolution of associates is likely to be observed in longer simulations. The general trends in the data show that as the concentration increased, the dissolution rate for clusters decreased. Dissolution at [C]=30 mM was observed to deviate from the general trend. It was observed that for this simulation, there was a retention of relatively high probability of calcium binding to three anions late into the trajectory. Dissolution rates are readily calculated from the slope of the semi–log plots, giving
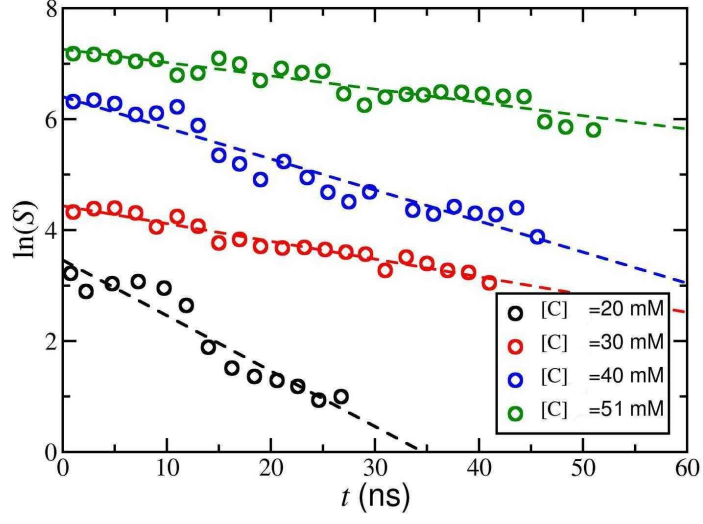
(a)
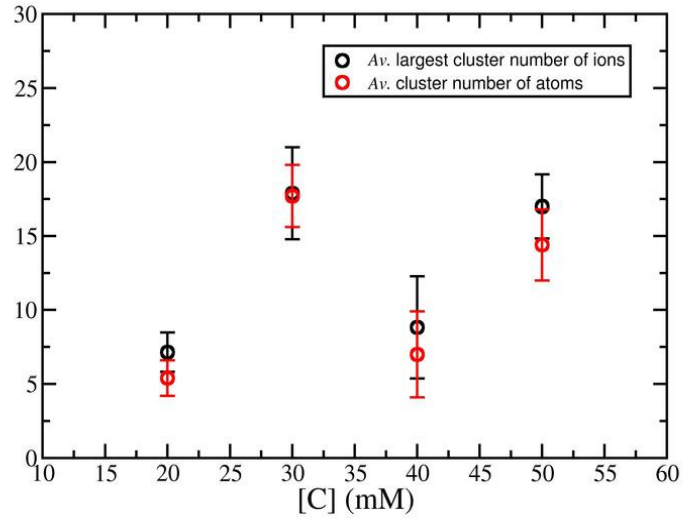
(b)

Figure 6.5: Simulations of system C_{20-20} at [C]=20, 30, 40 and 51 mM. (a) $\ln(S)$ as a function of time; the data has been shifted for clarity as follows, $\ln(S) - 1$, $\ln(S)$, $\ln(S) + 2$, $\ln(S) + 3$ for [C]=20, 30, 40 and 51 mM, respectively. (b) The average cluster size in units of number of atoms, and average largest cluster size in units of number of ions, for all concentrations studied for system C_{20-20} by the end of simulation. Error bars show uncertainties of one standard deviation in the data from the final 2 ns of simulation (as this time window was fixed there may be some bias in the data for larger concentrations as cluster dissolution is slower).

values shown in Table 6.4.

Table 6.4: Dissolution rates measured as $\ln(S)/t$ (see Figure 6.5 (a)), of a $20\,CaCO_3$ cluster at varying concentrations.

[C]/mM	$\ln(S)/t$ /ns^{-1}
20	0.100 ± 0.010
30	0.032 ± 0.003
40	0.056 ± 0.004
51	0.029 ± 0.004

Average and maximum cluster sizes at the end of simulation are shown in Figure 6.5 (b). At 20 and 40 mM these values are, within statistical uncertainties, consistent with ion pairs. Larger clusters are seen at 30 and 51 mM, though we note that from the $\ln(S(t))$ plots that dissolution is still in progress at this stage, and so these are overestimates of the equilibrium cluster sizes. It is important to note that all of these concentrations are higher than experimental ones, which are typically performed at 10 mM. While extrapolation to this concentration is not possible, due to the fact that simulations did not fully equilibrate, it is highly likely that ion pairs and small ion associates would also be found in solution at this value. It is sensible to conclude that even in the limit of high pH, dense liquids are unlikely to exist at experimental concentrations.

Closer inspection of the $S(t)$ curves shows a strong stochastic element to the dissolution process, and this was seen most clearly at 20 mM. To investigate this further, Figure 6.6 (a) provides S as a function of time, and is extended (compared with Figure 6.5 (a)) to include early dissolution. This shows a sudden and rapid decrease of S between 5–15 ns, superimposed on a slower but consistent decrease through the rest of the simulation. This change in rate can be explained by considering coordination probabilities; Figure 6.6 (b) shows the carbon to calcium binding probabilities as a function of time during the simulation. Around 10 ns, the probability of calcium coordinating to three anions decreased, and this was associated with an increase of calcium binding to a single anion. Conceptually, this is a change from a relatively compact cluster (where calcium binds to three and four carbons), to a system where there is predominantly ion pairs, and small ion associates in solution.Visual observation of the cluster at this stage showed the cluster "opening" to maximise ion solvation.

The cluster size distributions over the final 6 ns of simulation are provided in Figure 6.6 (c). At the end of the trajectory clusters continued to decrease in size and, ideally, further simulation would provide information on the extent to
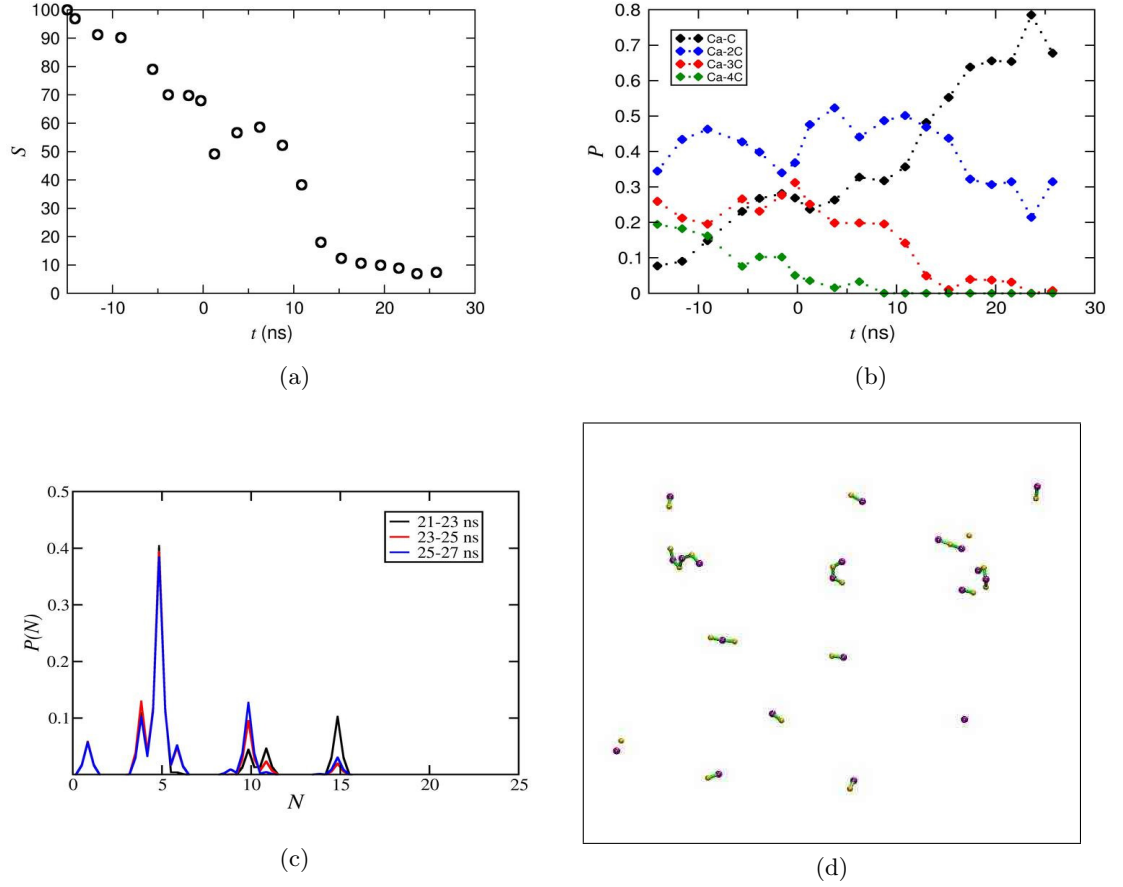
Figure 6.6: 42 ns simulation of $20\,CaCO_3$ in water at a concentration of 20 mM. (a): Average cluster size as a function of time; (b): average coordination probabilities of calcium to N carbons over time; (c): cluster size probability distributions calculated in 2 ns windows from the final 6 ns of simulation; and (d), a snapshot (with oxygen and hydrogen omitted) from the final 2 ns of simulation. Calcium is shown in yellow while carbonate and bicarbonate ion centres are shown in purple and blue. Green lines show coordinating ions where the distance between calcium and carbon is < 3.825 Å.

which equilibrium had been achieved. Ion pairs dominated in solution showed by the peak at $N = 5$; however, contrary to simulations at moderately basic pH, peaks are observed at $N \approx 15$. This suggests clusters reached a maximum size of around three formula units, in fact the average size of the largest cluster during the final 2 ns of simulation was measured as 5.5 ± 0.8 ions. From the calculated composition probabilities of associated species, ion pairs were most common ($P = 0.63$), followed by binding of two calcium and carbonate ions ($P = 0.21$). The probability of finding three calcium and three carbonate ions in a cluster was 0.05. The concentration of these species is likely to be much lower at equilibrium and a concentration of 10 mM. A snapshot in Figure 6.6 (d) shows the types of clusters which were found after 40 ns of simulation.

At the highest concentration, 51 mM, a large cluster was found at the end of the simulation (see Figure E.4 (b)). This was formed of 16 ions, with an average coordination of two between ion centres, and can be considered DLNP. The extent to which this cluster will further dissolve is unknown, as the system was still relatively far from equilibrium by 52 ns of simulation. The cluster size distribution provided in Figure E.4 (a) from the end of simulation shows a large concentration of ion pairs and a smaller concentration of larger associates. Analysis of the cluster compositions showed that 81% of associates were ion pairs, 9% were ion trimers and the remaining 10% were larger clusters.

6.3.4 System Size Effects at Low Concentration

To understand the effect of system size on cluster dissolution, clusters containing 30 and 40 calcium carbonate units with low density were generated. Clusters were immersed in water to produce solution concentrations of 20 mM (in terms of total numbers of ions) and simulated for 15 ns with the same input parameters as for the $20\,CaCO_3$ cluster at 20 mM. Clusters were sampled such that the average coordination and coordination profiles (as shown in Figure 6.7 (a)) were as close to the 20 unit cluster as possible. However, this became increasingly difficult as cluster size increased. The 30 unit cluster showed a reasonable match with 20 units: average coordination numbers at $t = 0$ were equivalent at 2.8, although as shown in Figure 6.7 (a), an increase in Ca-4C was observed for 30 units as Ca-3C decreased. $40\,CaCO_3$ had a higher average coordination of 3.5, and this was due to increased binding of calcium to four and five carbons, and the presence of dense ionic regions.

After 15 ns of simulation, the coordination profiles (Figure 6.7 (b)) for 30 and 20 units had converged (and were equivalent to the coordination profile for dense liquid phases at high concentration); however, the 40 unit cluster retained high
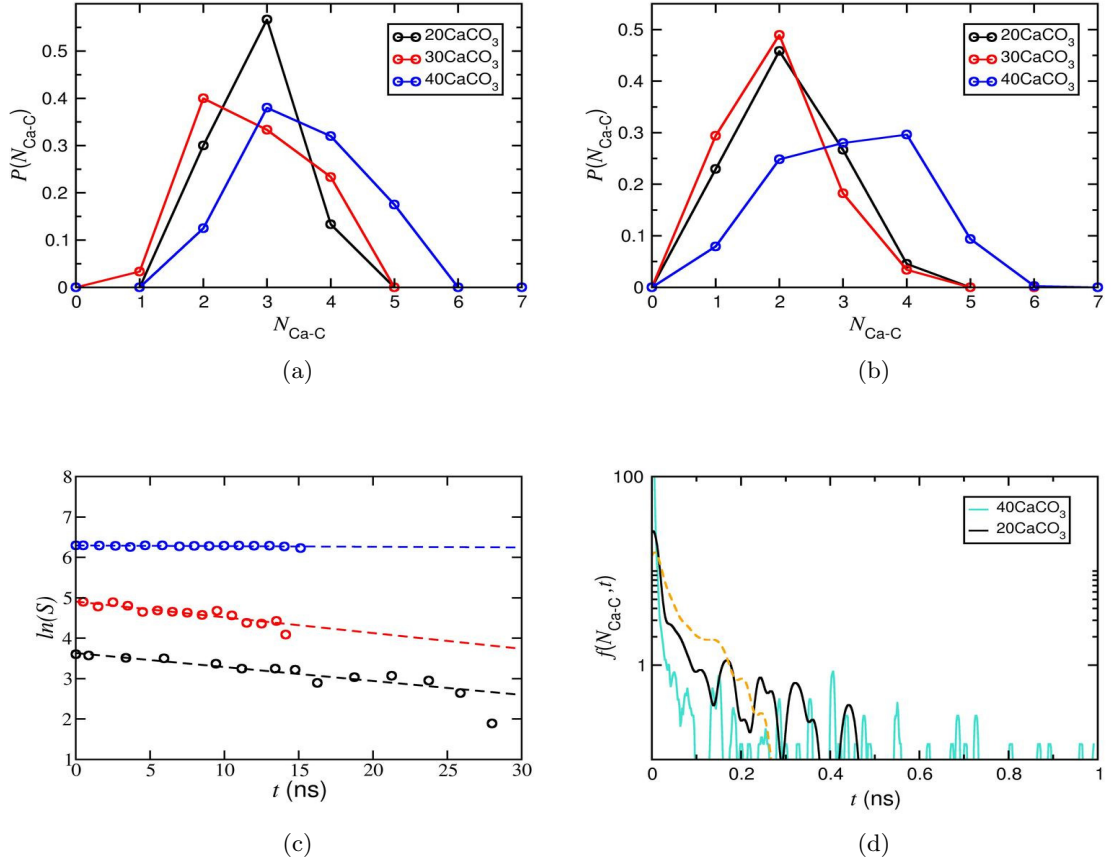
Figure 6.7: Simulations of clusters containing 20, 30 and 40 calcium carbonate units in water at 20 mM. (a): The coordination probabilities at $t = 0$. (b): The coordination probabilities at $t = 15 - 17$ ns for $20\,\mathrm{CaCO_3}$ and $t = 13 - 15$ ns for 30 and 40 unit clusters. (c): The natural logarithm of average cluster size as a function of time; the data has been shifted for clarity as follows, $\ln(S) - 1$, $\ln(S)$, $\ln(S) + 1$, for n=20 (black), 30 (red), and 40 (blue) ($n\mathrm{CaCO_3}$). (d): Probability densities for bond lifetimes between calcium and carbon defined by a distance criteria of $r_{Ca-C} < 4.2$ Å, measured from the time windows described for (b); the data from 6.4 (b) is shown as the orange dashed line. The probability densities for $30\,\mathrm{CaCO_3}$ are not shown for clarity, however these were very similar to the data for 40 units.

coordination levels, albeit with a reduction in Ca-5C as Ca-C increased. Figure 6.7 (c) shows the change in the logarithm of average cluster size over time. Values of $\ln(S)/t$ were -0.033 ± 0.003 and $-0.037 \pm .004$ ns^{-1} for the dissolution of 20 and 30 unit clusters respectively, up to 15 ns. The equivalent rates of dissolution can be attributed to the similarity in the coordination of ions at the beginning of the simulation. The comparable dissolution rates suggest that the simulations will eventually evolve to the same collection of ion pairs with a low concentration of small ion associates. The largest cluster remained particulate throughout the 15 ns trajectory, except for the loss of a single ion pair, and the average cluster size barely changed, as shown in Figure 6.7 (c). The lack of disassembly may suggest a (meta)stability to large clusters at this concentration, but analysis of the bond lifetimes identified that bonds were broken up to one nanosecond as shown in Figure 6.7 (d), and this was consistent with the data for 30 units. The fraction of unbroken bonds in the 2 ns windows were 9%, 4% and 6% for 20, 30 and 40 formula unit clusters, respectively. The bond lifetimes for 20 units converged to those for dense liquid nanoparticles at high concentration, with an extra two peaks at increasing t up to 0.45 ns, and so it is possible that concentration plays a role in the dynamic nature of coordination.

In the study of Wallace *et al.*, increasing the size of ions in liquid calcium carbonate clusters resulted in a reduction in the free energy [Wallace *et al.*, 2013]. In the current study, clusters of different sizes showed the same rates of dissolution. While the concentration in Wallace *et al.* was close to experiment ($[Ca^{2+}]=[CO_3^{2-}]=15$ mM), the sampling of a range of optimised cluster configurations (using replica exchange molecular dynamics) was carried out for 0.6 ns, from which the free energies were subsequently estimated. This is a completely satisfactory way to measure the free energy of growing liquid clusters, but sampling is insufficient to show the relative free energies of all possible states, and the study does not show that liquid clusters are more stable than free ions or ion pairs in solution at this concentration.

6.4 Experimental Studies of Precipitation

During the controlled titration of 10 mM CaCl$_2$ solution into a 10 mM carbonate buffer solution at pH=9.75, precipitation of calcium carbonate was found and a highly reproducible LaMer diagram (shown in Figure 6.8) was obtained. In the LaMer diagram the concentration of precursor solute increases until the critical concentration for nucleation is reached; the time before this point is known as the
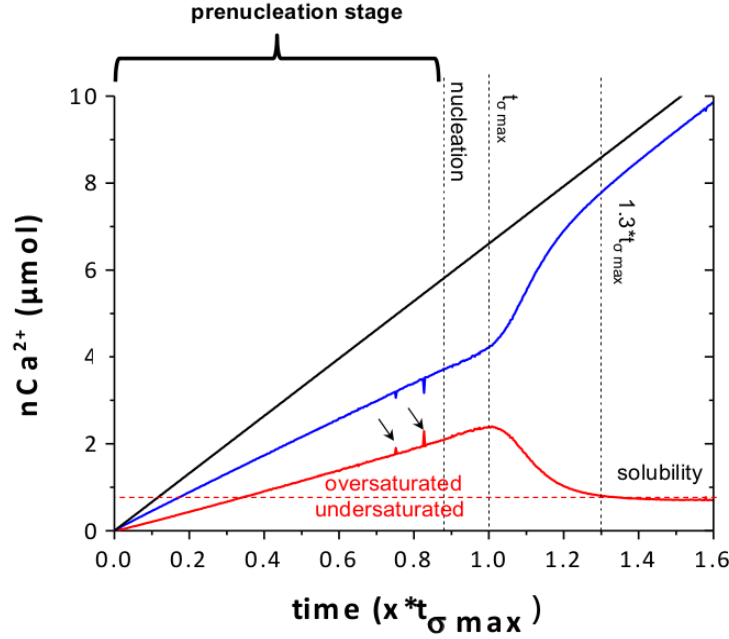
Figure 6.8: Titration curve showing the development of free Ca^{2+} ions (red line) compared to the total amount of dosed calcium (black) at pH=9.75, as a function of time normalised with respect to $t_{\sigma max}$ (the time at which the red curve reached a maximum). Different stages of the experiment are defined in the plot ($t = 1.3t_{\sigma max}$ is the time when solubility is reached).

prenucleation stage. Nucleation occurs at the critical concentration, but the concentration of free solute continues to increase up to a maximum, where the nucleation rate is balanced by the consumption rate of solute. The subsequent decrease in the concentration of solute occurs as the consumption rate is higher than the nucleation rate. Once the solute concentration falls below the critical concentration to nucleation, any nucleation ceases and the postnucleation stage is entered, where solute is consumed by growing nuclei. Finally, the solute concentration reaches the solubility limit for the bulk phase of growing solid.

Prenucleation Stage

The time axes in the LaMer diagrams from multiple experiments were normalised according to the maximum ($t_{\sigma max}$) in the free Ca^{2+} concentration (see Figure 6.8), and the critical concentration was $0.88t_{\sigma max}$. 62.9 ± 1.0 mol% of dosed calcium was measured to be bound in solution in the prenucleation stage, corroborating with the findings of Gebauer *et al.* [Gebauer *et al.*, 2008]. Constant pH was maintained by the constant linear addition of sodium hydroxide solution.

The free and bound calcium in solution before nucleation can be considered to be in equilibrium, in which case an equilibrium constant can be defined as follows,

$$K_{eq} = \frac{\left[\text{Ca}_x(\text{CO}_3)_y^{x-y} \right]}{\left[\text{Ca}_{free}^{2+} \right]^x \left[\text{CO}_{3free}^{2-} \right]^y},$$

(6.2)

with a similar function for bicarbonate binding. As the free (bi)carbonate concentration is excessive in the prenucleation stage, this can be considered as constant. The function reduces to,

$$K_{eq} = \frac{\left[\text{Ca}_x(\text{CO}_3)_y^{x-y} \right]}{\left[\text{Ca}_{free}^{2+} \right]^x}.$$

(6.3)

As the ratio of bound to free calcium was found to be constant, it follows that $x = 1$ and that one calcium ion is incorporated into prenucleation complexes. This fits with classical ion pair formation in solution before nucleation.

As the pH is constant before nucleation, the ratio of bicarbonate to carbonate must also be constant. 7.24 mM of bicarbonate is present in 10 mM buffer solution. On anion binding to calcium, this ratio will be re-established as the equilibrium of free anions ($\text{HCO}_3^- \rightleftharpoons \text{CO}_3^{2-} + \text{H}^+$) shifts to replenish the relative amounts of (de)protonated species. 1 mmol of carbonate binding to 1 mmol of calcium results in the release of 0.724 mmol of H^+, while 1 mmol of bicarbonate binding to 1 mmol of calcium requires consumption of 0.276 mmol of H^+. If these processes were taking place, then 0.724 mmol of OH^- or 0.276 mmol of H^+ would need to be added per mmol of added Ca^{2+}, respectively, to keep the pH constant. The amount of added NaOH across the prenucleation stage can, therefore, be used to calculate the ratio of bicarbonate to carbonate binding: $\text{NaOH}/\text{Ca}_{bound}^{2+} = 0.724$ for only carbonate binding and $\text{NaOH}/\text{Ca}_{bound}^{2+} = -0.276$ for only bicarbonate binding.

As the experiment is conducted under atmospheric conditions, the effect of CO_2 indiffusion must also be accounted for. When CO_2 dissolves into solution H_2CO_3 is generated, from which bicarbonate and carbonate is produced by loss of protons. NaOH is added to account for the release of 1.31H^+ (measured from the relative activities of bicarbonate and carbonate) per molecule of CO_2.

Accounting for the indiffusion of CO_2 (which is distinctly different to the studies of Gebauer et $al.$), during the prenucleation stage $\text{NaOH}/\text{Ca}_{bound}^{2+} = 0.31 \pm 0.06$. Using all of the titration data in a speciation program (Visual MINTEQ), the concentration of species in solution at $t = 0, 0.2, 0.4, 0.8$ and $1.0 \times t_{\sigma max}$ were estimated. The software determined that 99 % of bound calcium before nucleation

130

was present as $CaCO_3^0$ and $CaHCO_3^+$, with calcium carbonate ion pairs dominating (97% of ion complexes were of this type). The titration profiles were reproduced with high accuracy, providing further evidence that bound calcium in solution before nucleation can be descried in terms of simple ion pairs.

Further support for the lack of prenucleation clusters was provided by cryo–TEM. The TEM images which were obtained following analysis of samples extracted from solution at $t = 0.86t_{\sigma max}$ were subjected to an analysis procedure. Firstly, a smoothing filter was applied to a TEM image which resulted in the smoothing of each pixel according to the intensity of 250 surrounding pixels. The standard deviation in pixel intensity was 0.004% of the mean. Objects were identified according to the pixel intensity once background noise (from the filtered image) was extracted, so long as a minimum of four-way pixel connectivity was found. A minimum of three pixels were, therefore, required to define the diameter of an object. The threshold, below which an object cannot be distinguished from noise, was 7–8 pixels (approximately 0.9 nm). An equivalent diameter of objects was measured from the maximum length of connected pixels. No object in the prenucelation stage was measured to be greater than the threshold of 0.9 nm. That is to say, no prenucleation clusters can be identified from cryo–TEM images.

Nucleation and Post–nucleation stages

cryo–TEM was used to analyse the species emerging at nucleation. At $t = 0.88 - 1.0t_{\sigma max}$ TEM images showed the presence of 200-400 nm objects which had a very low contrast for their size, as shown in Figure 6.9. The low contrast could be due to high degrees of hydration in the objects, similar to those observed for polymer induced liquid precursor phases [Cantaert *et al.*, 2012]. No long range atomic order was found in the objects from low dose electron diffraction analysis, and so these objects were proposed to be a dense liquid phase formed via liquid–liquid separation.

At $t = 0.96t_{\sigma max}$ objects of size 2 nm were identified in solution using cryo–TEM; these were much higher contrast than the dense liquid phase. These objects were too small for diffraction analysis. However, SEM revealed that the high contrast objects had reached micrometer sizes with spherical morphology by $t = 1.03t_{\sigma max}$ (see Figure 6.9). *In-situ* FTIR identified characteristic vibrations for vaterite (875, 1072 and 1087 cm^{-1}) in these objects. As in the prenucleation stage, $NaOH/Ca_{bound}^{2+}$ in the post-nucleation stage was found to be constant. The change in the concentration of sodium hydroxide, $\Delta[NaOH]$ and bound calcium, $\Delta[Ca_{bound}^{2+}]$, in solution during the growth stage can be used to determine the number of protons and calcium involved in the nucleation and growth of solid. $\Delta[NaOH]/\Delta[Ca_{bound}^{2+}]$ was
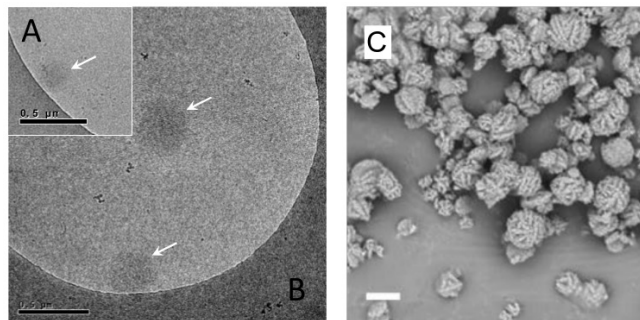
Figure 6.9: Images of (sub)micron sized objects during the nucleation and growth stages. cryo-TEM images were taken at (A) $t = 0.92t_{\sigma max}$, and (B) $t = 0.96t_{\sigma max}$ showing low contrast objects of 215 nm and 320–350 nm, respectively (indicated by arrows; scale bar is 0.5 μm). C shows a SEM image taken at $t = 1.03t_{\sigma max}$ showing spherical–framboidal objects typical of vaterite; scale bar is 2 μm.

measured to be 1.47 ± 0.14. Growth and nucleation of solid from free ions would give $\Delta[\text{NaOH}]/\Delta[\text{Ca}^{2+}_{bound}] = 1.38$ (1/0.724 H$^+$ released per Ca$^{2+}_{bound}$). The data suggest that vaterite is grown by addition of free ions to growing nuclei, and this fits with a classical mechanism of nucleation and growth.

6.5 Conclusions

From MD simulations at equilibrium, ion pairs and free ions were found to dominate species probability distributions at pH and concentration levels close to experiment. No calcium carbonate clusters in the size ranges expected for prenucleation clusters (PNCs) were found. Only when the concentration of ions was much higher than experiment (51 mM) were large calcium carbonate clusters observed. DOLLOP has been suggested as the structural form of PNCs, but at the concentrations investigated, only a small concentration of ion associates containing up to three to four ions are likely to be found in solution at experimental concentrations and pH. Even at the limit of high pH, and at a concentration of 20 mM, clusters containing a maximum of six ions were found.

The findings are in contrast with the proposed thermodynamic equilibria of ion association in solution [Demichelis *et al.*, 2011]. However, speciation data for the highest concentrations studied in the work of Demichelis *et al.* was used to fit a speciation model at the experimental conditions of 10 mM buffer solution and a calcium concentration of 0.4 mM [Demichelis *et al.*, 2011]. The fraction of calcium bound in CaCO$_3^0$, CaHCO$_3^+$ and DOLLOP was suggested to be 0.2, 0.03 and 0.73, respectively. The authors note that this shows good agreement with the fraction of

bound calcium ($f_{Ca-bound}$) in experiment: at pH 9.75$\pm$0.05, $f_{Ca-bound} = 0.69\pm0.07$, and at pH 10.0 ± 0.05, $f_{Ca-bound} = 0.76 \pm 0.08$ [Demichelis *et al.*, 2011]. However, at pH 9.5-10 and 0.06 M, the fraction of Ca^{2+} bound in DOLLOP is likely to be extremely smaller than this. At the pH range and concentrations considered in this study, the fractions of calcium bound in solution are provided in Table 6.5. Reasonable agreement to the experimental values is found, especially for ionic solutions simulated at an initial pH of 9.9, but none of the bound calcium was present as DLNPs. While Demichelis *et al.* note that bond orders at low concentration are system size limited, this was not the case in the current study: at the same pH, coordination probabilities for different system sizes were equivalent within statistical noise.

Table 6.5: Fraction of bound calcium, $f_{Ca-bound}$, in solution at equilibrium.

System	$f_{Ca-bound}$
I$_{22-14}$	0.84 ± 0.07
C$_{22-14}$	0.90 ± 0.04
I$_{46-29}$	0.77 ± 0.04
I$_{64-47}$	0.89 ± 0.02

Simulations at higher concentrations were consistent with recent computational studies [Demichelis *et al.*, 2011; Wallace *et al.*, 2013] showing that dense liquid phases are stable in solution. The possibility of dense liquid phases forming locally in solution is relevant. During experiments, concentration gradients are likely to be found in solution, which may lead to the association of ions and a liquid–liquid phase separation. As these studies have shown, dense liquid clusters which form are likely to be long lived and could possibly be detected by experimental analysis. Furthermore, as others have pointed out [Faatz *et al.*, 2004; Wallace *et al.*, 2013], the formation of solid phases in the dense liquid may offer a low energy route to the precipitation of calcium carbonate.

The results of computer simulations are in good agreement with the experimental findings presented in section 6.4. Only free ions and ion pairs CaCO$_3^0$ and CaHCO$_3^+$ were found in solution in the prenucleation stage of titrations. No PNCs in the size ranges suggested by Gebauer *et al.* [Gebauer *et al.*, 2008] and Pouget *et al.* [Pouget *et al.*, 2009] were found in TEM images. To reconcile differences in this study and previous experiments, it is useful to consider the experimental methods. In the analytical ultracentrifugation experiments used to determine the size distribution of PNCs [Gebauer *et al.*, 2008], it is likely that large concentration gradients exist in solution. At relatively high concentrations, clusters may be

metastable, as shown from simulations here. In the experiments of Pouget *et al.* the Kitano method for nucleation was adopted in which local high supersaturations at the gas–liquid interface will be found, and so direct comparison to these experiments is difficult [Pouget *et al.*, 2009]. However, at the interface it is likely that metastable clusters would form in these high concentration regions, and vitrification for TEM analysis would capture clusters which would otherwise dissolve in an equilibrated homogeneous solution.

The fact that only free ions and ion pairs are found at low concentration and moderate pH levels in both experiment and simulation is evidence that PNCs are unstable. Instead, a classical mechanism of growth of crystalline phases from free ions or ion pairs seems more likely. Indeed, analyses of objects formed during nucleation in experiment support this. From the simulation side, a thermodynamic link between ion pairs and nanocrystalline particles may help to understand the likelihood of this mechanism. Nonetheless, the data presented here does compare reasonably well with the findings of Hu *et al.* where classical nucleation of calcite on self assembled monolayers was observed [Hu *et al.*, 2013].

The observation of low contrast spherical objects (from cryo–TEM) which reach considerable sizes is consistent with the formation of a dense liquid phase. The formation of a dense liquid via binodal demixing was suggested by Wallace *et al.* [Wallace *et al.*, 2013], and is further indication of the emergence of liquid (precursor) phases. It is possible that the DLNP found at high concentration is the structural form of the dense liquid phase in experiment. At very high supersaturation a homogeneous solution of free ions and ion pairs will be unstable and so phase separation becomes likely. However, nucleation of crystal (as found in the experiment) will not be seen on the time-scales of the simulation. Liquid–liquid phase separation, where there is a low energy barrier due to a low surface tension of dense liquid phases, is, on the other hand, more likely to be observed. Further analysis must be performed in order to compare the dense liquid phases found in experiment and DLNPs from simulation.

Chapter 7

Postnucleation Phases and Dehydration of Amorphous Calcium Carbonate

Declaration

A selection of the material in this Chapter has been written into two submitted manuscripts which are listed. 1) Y. G. Bushuev, A. R. Finney and P. M. Rodger. Stability, structure and dehydration of hydrated amorphous calcium carbonate. 2) A. R. Finney, Y. G. Bushuev, and P. M. Rodger. On the mobility of water in amorphous calcium carbonate.

   In the following Chapter, my analysis of a series of simulations of ACC will be presented. The simulation trajectories were provided by Dr. Yuriy G. Bushuev (YGB) from the *Department of Chemistry* and *Centre for Scientific Computing, University of Warwick*. A number of the simulations were repeated to extend the data set provided by YGB. By way of background and to make the significance of the work clear, atomistic spatial correlation functions were calculated by YGB and are presented in section 7.3.1 (*i.e.* Figure 7.1). All other work in this Chapter is my own.

7.1   Introduction

Amorphous calcium carbonate (ACC) has been identified as a precursor phase to crystalline forms of the mineral, with both stable and transient ACC identified *in vivo* [Beniash *et al.*, 1997; Weiner *et al.*, 2005; Politi *et al.*, 2008]. The purpose of

ACC may be to stabilise the size and shape of emerging crystalline phases, without the necessity of directing organic templates [Nassif *et al.*, 2005; Bots *et al.*, 2012]. Both re-precipitation and solid state transformation have been identified in ACC at the onset of crystal growth, and so it is possible that both of these processes take place during the nucleation and growth of biominerals [Pouget *et al.*, 2010; Rodriguez-Blanco *et al.*, 2011].

Hydrated and anhydrous forms of ACC are found *in vivo* [Addadi *et al.*, 2003]. Gilbert and co-workers have recently shown that the crystallisation pathway in sea urchin spicules follows hydrated ACC $\longrightarrow$ dry ACC $\longrightarrow$ calcite [Gong *et al.*, 2012], with the stabilities of each phase (relative to calcite) decreasing in the same order [Radha *et al.*, 2010]. The ratio of water to calcium carbonate in hydrated ACC has been identified as $1 : 1$ H_2O:$CaCO_3$ [Levi-Kalisman *et al.*, 2002]. Further studies have shown that in transient ACC, water is released from the amorphous phase on crystallisation, and that atomistic scale features of crystalline phases can be identified before crystal growth [Cartwright *et al.*, 2012]. The dehydration mechanism is therefore an important component in the biomineralisation process which remains unresolved.

In the following Chapter, molecular dynamics (MD) simulations of ACC with varying levels of hydration are outlined. The structure of ACC from a number of preparation methods is discussed and compared with experimental and other computational findings. The role of water in hydrated ACC is studied in detail, and a comparison of ACC with different amounts of water is made. Following this, a dehydration mechanism is proposed.

7.2 Methods

7.2.1 Preparation and Simulation

Two types of ACC with varying hydration levels were generated: bulk ACC and mesoporous ACC.

Bulk ACC was prepared from initial random particle distribution configurations with a composition of $CaCO_3 \cdot nH_2O$, where $n =$ 0, 0.55, 0.75, 0.8, 1, 2, and 3. For this purpose, 2880 $CaCO_3$ units and 8640 water molecules were placed in a cubic simulation cell with an edge length of 70.5 Å. A 1 ns *NVT* MD simulation was then performed at 750 K using the force field of Raiteri and Gale [Raiteri and Gale, 2010] in which all electrostatic interactions were switched off. A Ca–Ca Lennard-Jones interaction, which does not feature in the $CaCO_3$ force field, was added with parameters $\varepsilon = 0.00674$ eV, $\sigma = 5.17$ Å, to avoid cation overlap. The published in-

termolecular potential, including electrostatic interactions, was then reinstated and the system relaxed in NpT simulations at 300 K for 3 ns. Other stoichiometries were obtained by removing further water molecules selected at random, and repeating the relaxation simulations.

A series of simulations were carried out to investigate specially prepared extended ACC systems, termed mesoporous ACC. In the first stage of generation, ACC nanoparticles were created by immersing a nanocrystal of ikaite (containing 2880 $CaCO_3$ units) in 30,000 water molecules. NVT simulations were then conducted for 100–500 ps at 1500 K to amorphise the carbonate phase. Subsequently, solvent water and some randomly selected internal water were removed from final nanoparticle configurations to give systems with $H_2O/CaCO_3$ of 1 and 3. NpT simulations were then performed at 1000 K and 0.1 MPa for 300–500 ps, followed by simulations at 300 K and 0.1 MPa.

MD simulations were implemented with DL_POLY 4.03. All systems were simulated in the NpT ensemble at 300 K and 0.1 MPa, employing a Nosé–Hoover thermostat and barostat with 0.1 ps and 1.0 ps relaxation times, respectively. Cubic periodic boundaries were used throughout. The cut off for short range interactions was set to 9.0 Å, and electrostatic interactions were treated using the smooth particle mesh Ewald method [Essmann *et al.*, 1995]. A 2 fs timestep was used throughout, and trajectories of approximately 5 ns were generated.

## 7.3 Results

7.3.1 ACC Atomistic Structure

In order to compare the atomistic structure of bulk ACC with varying degrees of hydration to experimentally analysed ACC, total scattering structure factors, $S(q) - 1$, and total radial distribution functions, $G^x(r)$, are presented in Figure 7.1. Data from Fernandez-Martinez *et al.* [Fernandez-Martinez *et al.*, 2013] was used to compare calculated pair correlation functions to those obtained from experimental X-ray diffraction patterns. In the experiment, ACC was found to contain 0.5 moles of water for every mole of calcium carbonate. In general, there is good agreement between experimental and simulated ACC correlation functions, with calculated structure factor peaks in Figure 7.1 (a) matching on to those found for synthetic ACC from experiment.

The effect of water on calculated $G^x(r)$ can be seen in Figure 7.1 (b); as hydration levels in ACC were increased, peaks in $G^x(r)$ shifted towards larger r. This was coupled with a decrease in the strength of oscillations in radial distribu-
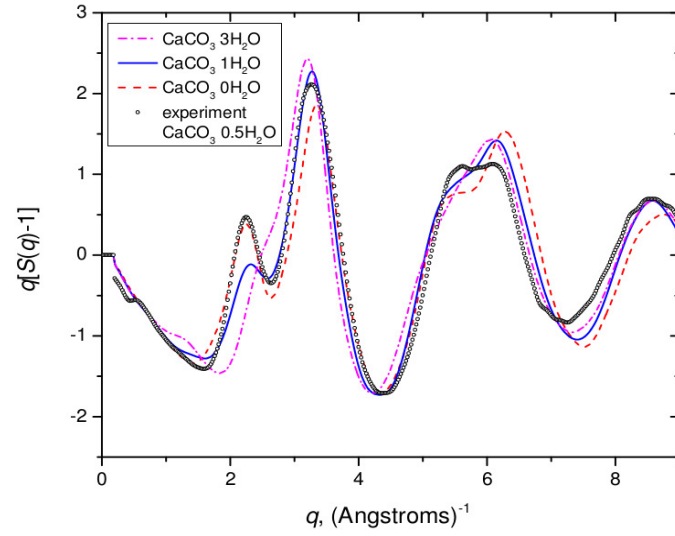
(a)
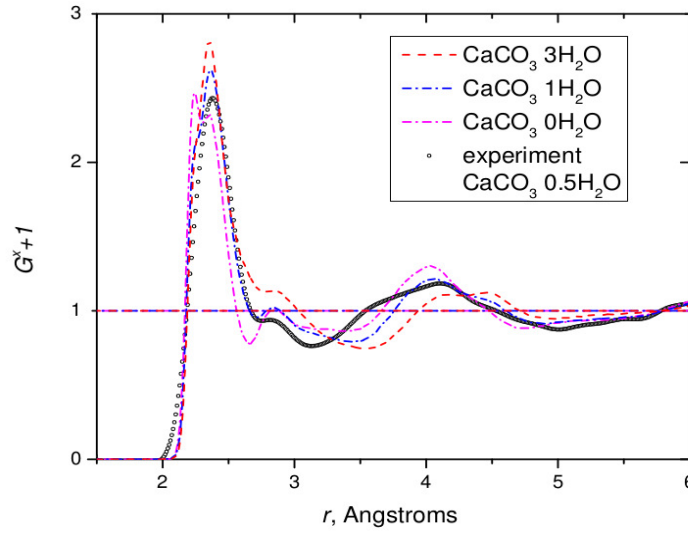
(b)

Figure 7.1: (a) Total scattering structure factors, and (b), total radial distribution functions of bulk ACC with composition $n = 0$, 1 and 3 for $CaCO_3 \cdot nH_2O$. Experimental data was taken from Fernandez-Martinez *et al.* [Fernandez-Martinez *et al.*, 2013]. Correlation functions were measured from time averaged spatial atom pair distributions from the final portions of MD simulations after equilibration. See Appendix F for the functional form of calculated correlation functions. Printed with the permission of YGB.

tion functions, suggesting a more amorphous material in $CaCO_3 \cdot 3\,H_2O$ than for ACC with smaller amounts of water present. The experimental curve matched most closely to simulated anhydrous ACC, but for both $n = 0$ and 1 (for $CaCO_3 \cdot nH_2O$), the largest discrepancy between simulated and experimental ACC was found in the region $r = 2.5 - 3.5$ Å. To understand this, G^x was decomposed into individual contributions from $g_{ij}(r)$. It was shown that at $2.6 - 3.6$ Å, $G^x(r)$ is dominated by O–O, Ca–C and O–O_{wat} spatial correlations, while at $3.6 - 4.5$ Å, Ca–O and Ca–Ca atom pairwise spatial correlations mainly contribute to $G^x(r)$. At 3.2 Å, a large density was found, which was not observed in the corresponding experimental distribution function, suggesting that there is slight inaccuracy in the spatial distributions of carbonate and water, which has been highlighted in recent studies of the calcite – water interface [Fenter *et al.*, 2013].

A comparison of ACC with initial composition $CaCO_3 \cdot nH_2O$, where $n = 0$ and 3, was made with four crystalline forms of calcium carbonate: calcite, vaterite, aragonite and monohydrocalcite (MHC). ACC was prepared from an initial nanoparticle with $3 : 1$ H_2O:$CaCO_3$, and after relaxation in water, this was calculated as $2.8 : 1$ H_2O:$CaCO_3$. Nanoparticles of crystalline phases were immersed into water and allowed to relax, before spatial correlation functions were calculated. Some disordering at the surface of nanoparticles was found, but the large size of particles meant that this did not lead to perturbations in the calculated $g(r)$. Figure F.1 shows $g_{Ca-Ca}(r)$, $g_{Ca-C}(r)$ and $g_{C-C}(r)$ for ACC and the four crystalline polymorphs. Ca–Ca spatial distributions matched well with vaterite and MHC. Ca–C spatial correlations showed that there was good agreement between ACC and both aragonite and MHC at small r, but the similarities in peak positions became unclear between aragonite and ACC as r increased. From C–C spatial correlations, it was not clear that ACC compared well to any anhydrous crystalline polymorphs (C–C close contact where $r < 3.5$ Å was observed in ACC, MHC and aragonite), but reasonable overlap of peaks was seen when compared with MHC. In general, it is clear that the radial distribution functions of anions in ACC compare well with MHC.

7.3.2 Heterogeneities in ACC

In characterising ACC it is important to understand the material and to compare with crystalline polymorphs. As well as features on the order of Angstroms, such as those provided in section 7.3.1, analysis of the supramolecular structure of ACC can shed light on a material which can be difficult to study by experiment because of its often transient nature. Figure 7.2 shows snapshots of typical configurations
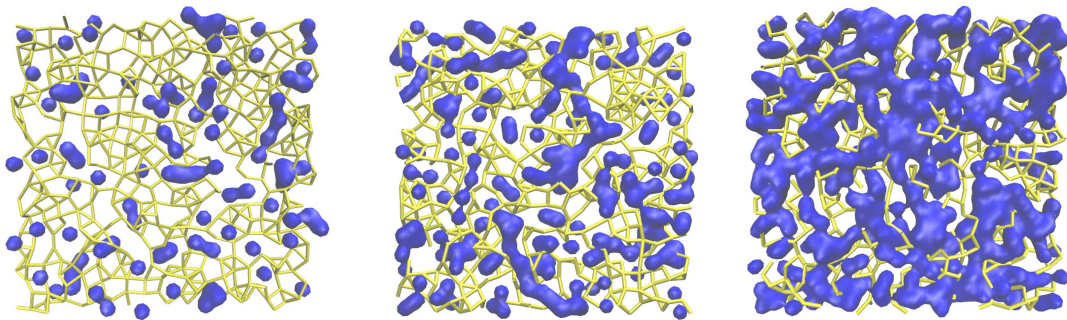
Figure 7.2: Snapshots of a 4 Å slice through the centre of the simulation z axis of equilibrated configurations of $n = 0.55$, 1 and 2 (from left to right respectively) for $CaCO_3 \cdot nH_2O$. Calcium and carbon atoms are connected by yellow bonds if $r_{Ca-C} < 3.825$ Å. Water is shown as a blue surface, where the probe radius is 1σ of oxygen atoms. Hydrogen and carbonate oxygens have been omitted for clarity.

for systems containing $n = 0.55$, 1 and 2 water molecules per $CaCO_3$ unit. A heterogeneous distribution of water can be seen, with water clusters tending to occupy channels and cavities within an ionic framework formed by calcium and carbonate. The distribution of ions locally was equivalent for all hydration levels, but there was a clear difference between the local concentration of water with respect to ions. At low levels of hydration (*i.e.* $n = 0.55$), water tended to form small clusters which were isolated in cavities, but increased water content led to structures with very few isolated water in cavities and large interconnected water clusters that formed within channels in the ionic framework. The clusters extended throughout the framework, and the hydration level taken from different regions of time averaged configurations was consistent with the average stoichiometry for the whole system, as shown in Figure F.2 for ACC for the highest water content.

Goodwin *et al.* have used Reverse Monte Carlo (RMC) simulations to generate a proposed structure of hydrated ACC based on experimental measurements [Goodwin *et al.*, 2010]. They showed a similar heterogeneity within ACC, identifying water-rich regions that had excess carbonate, within a porous, calcium-rich ionic framework. This is further support for heterogeneous water distributions in ACC; however, it is unlikely that cation- and anion-rich regions will form in ACC as this would lead to a high energy state, as was shown by Singer *et al.* [Singer *et al.*, 2012].

The distribution of water in mesoporous ACC was not consistent throughout the simulation cells. Two systems were simulated, with overall average stoichiometries of $CaCO \cdot H_2O$ and $CaCO_3 \cdot 3\,H_2O$. However, heat treatment led to partial phase separation, and introduced a concentration gradient to both systems. While

the local concentrations of ions was again consistent, as for bulk systems, there was a depletion of water around the centre of the simulation cell, leading to an increase in the amount of water close to the simulation cell boundaries, exemplified by Figure F.3 (a). This led to a mesoporous system where repeating periodic units of calcium carbonate were encapsulated within channels of water.

Analysis of the water content in mesoporous systems is provided in Figure F.3 (b). For both low and high hydration, a plateau region is found in H_2O/Ca^{2+} up to around 23 Å (from the simulation cell origin), after which there is a sharp increase in water content to the edge of the simulation cell. This cut-off signifies a partition in the systems from a core with low hydration, to increasingly hydrated calcium carbonate, whereupon the concentration of ions gradually decreases. The surface of ACC is ill-defined and is both rough and fluctuating, in agreement with other studies [Raiteri and Gale, 2010]. The chemical composition in the core of overall low and high hydration mesoporous ACC was measured as $n = 0.16 \pm 0.01$ and $n = 0.64 \pm 0.01$, respectively, where n is the number of water molecules per calcium ion. Within the core, water was mainly isolated in cavities within the ionic framework, with larger clusters observed in channels at the periphery of the ionic surface as the hydration level increased. In the mesoporous system with total average composition of $CaCO_3 \cdot 3\,H_2O$, no ions were found at the outermost regions of the simulation cell, and so this water can be considered to be contained in mesoporous channels.

In this work, it has been helpful to define three types of water molecule which can be found in ACC: occluded, channel and mesoporous channel water. The distance between water molecules was defined by the relative positions of oxygen atoms, and the average maximum distance between nearest neighbour water molecules (as measured from the minimum following the first peak in $g_{O-O}(r)$ for $CaCO_3 \cdot H_2O$) was ~ 3.8 Å. Water molecules which maintained $r_{O-O} > 3.8$ Å over the course of the trajectory were deemed to be "occluded" in the ionic framework and were labelled H_2O_{Occ}. Water molecules for which $r_{O-O} < 3.8$ Å at any stage in the trajectory were described as "clustered" water. For clustered water, it was further helpful to differentiate between water molecules found within nanometre sized channels in the ionic framework of ACC, and those which were contained within mesoporous channels (where a relatively low concentration of ions was found). Water in channels within the ACC framework and in water-rich regions in mesoporous ACC were labelled H_2O_{Chan} and H_2O_{Mes}, respectively. Water clusters can be occluded in the ACC framework, and so the definition of occluded water is a conservative one. A snapshot showing H_2O_{Occ} and H_2O_{Chan} in mesoporous ACC is provided in Figure

F.4.

In the case of bulk ACC, H_2O_{Occ} and H_2O_{Chan} were present, and in mesoporous systems, additional H_2O_{Mes} was found. On analysis of bulk systems, 4.76 (0.25) % of water molecules were measured to be H_2O_{Occ} in $CaCO \cdot H_2O$. The probability of finding H_2O_{Occ} in all types of hydrated ACC, measured as a function of hydration, was observed to decrease exponentially as shown in Figure 7.3, with almost all water in $CaCO_3 \cdot 3\,H_2O$ present in clusters, while 80.76 (2.01) % of water in $CaCO_3 \cdot (0.16)H_2O$ was H_2O_{Occ}. The results suggest that the size of water-rich channels in ACC decreases as the water content in the system is reduced, leading to a reduction in the amount of H_2O_{Chan} for bulk ACC and mesoporous systems.

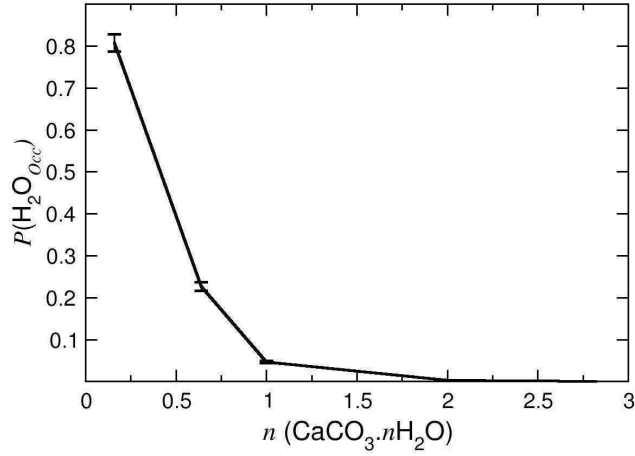Figure 7.3: Probabilities of finding occluded water, H_2O_{Occ}, in hydrated ACC. Water types were analysed from bulk ACC and the core of mesoporous ACC systems. H_2O_{Occ} was defined as water molecules which maintained a distance of 3.8 Å to all other water throughout simulation trajectories.

7.3.3 Water Clusters

It would be helpful to obtain further information about the cavities and channels in ACC. Measuring accessible internal surface area and the size and shape of voids in the ionic framework are ways to do this. However, as these spaces in ACC are occupied by water, characterising the water networks can provide information on the supramolecular features of the amorphous ionic material. This approach is pragmatic, as any voids in the material which do not contain water will be omitted from any measurements; nonetheless, it is a sensible approach. In the current section, water clusters have been analysed in bulk ACC for the compositions $n =0$, 0.55, 0.75, 0.8, 1 and 2 in $CaCO_3 \cdot nH_2O$.

Using the same definition to label water types, waters were considered clustered if $r_{O-O} < 3.8$ Å at any stage in the analysed trajectories. Adjacent water molecules, for which $r_{O-O} < 3.8$ Å, were therefore described as being "bonded" – a bond here is equivalent to an edge in the terminology of graph theory. This distance was taken from $g_{O-O}(r)$, and is larger than the typical distances measured for liquid or solid bulk water phases. At this distance, calcium carbonate mediated nearest neighbour waters are found rather than directly interacting water molecules. The definition of bonded water is therefore appropriate for the current exercise and is distinct from hydrogen-bonding. The fraction of water molecules participating in H-bonds was measured and is shown in Table 7.1. The data shows that the fraction of water molecules H-bonding with other water molecules is relatively small, even for the highest hydration levels. When measurements were repeated with carbonates included in the H-bond calculations, there was a significant increase in the number of H-bonds per water molecule. The data suggests that water in ACC is structural and that H-bonding increases as the hydration level increases.

The bonding (defined by $r_{O-O} < 3.8$ Å) probabilities for a range of hydration levels are provided in Figure 7.4. As the hydration level increased, a greater affinity for increased bond orders was observed, with the maximum in the bond probability

Table 7.1: Percentage of total water molecules in bulk ACC which engaged in hydrogen-bonding directly to other water molecules ("water H-bonds"), or either a water or a carbonate oxygen (simply "H-bonds"). H-bonds were defined by $r_{O-O} < 3.4$ Å and $\theta_{H-D-A} < 40°$.

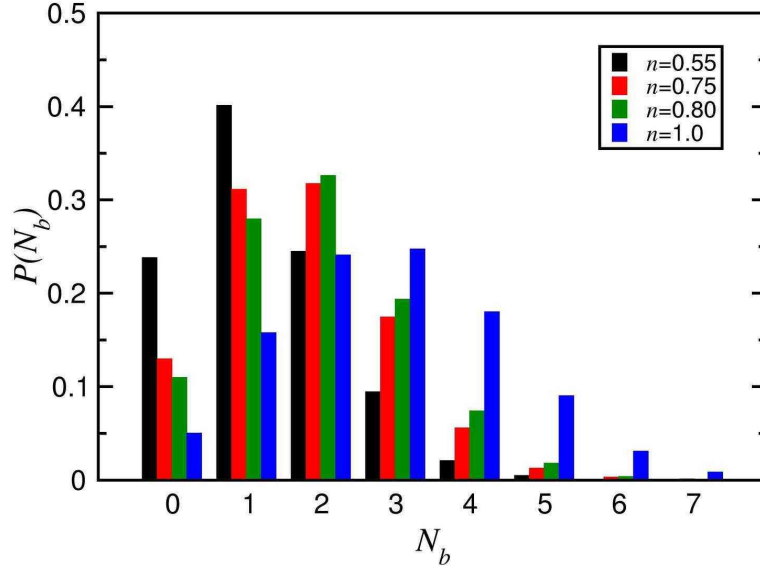| $H_2O/CaCO_3$	Water H-bonds	H-bonds
0.55	5.4 ± 2.2	89.1 ± 1.4
1	14.0 ± 0.7	97.3 ± 1.2

Figure 7.4: Probability of finding N_b bonds between water molecules in ACC for a range of hydration levels, n, in $CaCO_3 \cdot nH_2O$. A bond was defined by a distance criteria of $r_{O-O} < 3.8$ Å, and data was measured from the final 0.5-1.0 ns of trajectories after equilibration.

distributions shifting from $N_b = 1$ to $N_b = 3$ for $n = 0.55$ and 1, respectively. At the intermediate hydration levels a maximum of $N_b = 2$ was found. The distributions were also observed to widen with increasing n, and this was due to a greater variety of accessible cluster topologies and increasing size of clusters. By enumerating all the distinct clusters observed during a simulation, it was possible to calculate the average water coordination number as a function of cluster size, s (in terms of number of water molecules), and the results are presented in Figure F.5. There appeared a limit to average N_b for different hydration levels. As cluster size increased, an increase in average N_b was observed, but this reached a plateau for clusters where $s > 60$. Large clusters are rarely seen for $n = 0.55$, and so the data for this system spans a limited range of cluster sizes; the average N_b was around that for $n = 0.75$ and 0.8, which was measured to be 2.4 ± 0.1 (for $n = 0.8$). For $n = 1$, the maximum average N_b was 2.6 ± 0.2. The data suggests that low coordination water clusters in the channels of ACC were present..

Cluster size distributions (CSDs) are provided in Figure F.6. These showed that as the water content in ACC increased, there was an associated increase in the maximum cluster size. For $n = 0.55$ and 1, the maximum peak in CSDs was

around $s = 37$ and 2600 (the fraction of water molecules in the largest clusters was 0.023 and 0.903), respectively, and as discussed earlier, this was due to an increase in the amount of H_2O_{Chan}. The data for $n = 0.75$ and 0.8 was noisy at the limit of maximum s, with a wide range of cluster sizes at these hydration levels. However, for $n = 1$, the data shows a high probability of finding clusters up to around $s = 100$, but no larger clusters were found until the maximum of $s = 2600$. The largest cluster in the system was found to be periodically infinitely large, with connections found across at least two periodic boundaries of the simulation cell, and with an infinite number of cycles possible throughout the network defined by molecule spatial coordinates.

To further investigate cluster shape, the average radius of gyration, R_g (measured from the cluster centre of mass), was calculated for the largest finite clusters in the same range of hydration levels, with infinite clusters omitted from the calculation. The data are provided in Figure 7.5. As the amount of water in ACC increased, the size and R_g of average finite clusters also increased. The average size of the largest finite cluster, N_{lrg}, was calculated by averaging the number of atoms in the largest finite cluster over successive frames in simulation trajectories. N_{lrg} reached a maximum of around 145 when $H_2O/CaCO_3$ was 0.8, and decreased for $n = 1$ to the levels observed for ACC with the lowest hydration (see Figure 7.5). The maximum in R_g was in the range $n = 0.75 - 0.8$. This non-monotonicity is due to the emergence of infinite clusters, which by $n = 1$, composed a significant proportion of the total number of waters in the system.

The eigenvalues of the moment of inertia tensor were calculated for finite clusters in $CaCO_3 \cdot (0.8)H_2O$. Ratios of eigenvalues of decreasing size, λ_3, λ_2 and λ_1, were calculated and can be found in Figure F.7. Above a cluster size of around $s = 20$, ratios of $\frac{\lambda_2}{\lambda_1}$, $\frac{\lambda_3}{\lambda_1}$ and $\frac{\lambda_3}{\lambda_2}$ reached plateaus, albeit with significant noise in the data. $\frac{\lambda_3}{\lambda_2}$ was consistent across the range of cluster sizes, signifying that the shape of clusters did not change. The ratios suggest clusters were oblate on average, and considering the coordination profile, this is consistent with a backbone chain of water molecules, with a large amount of branching from the chain. To verify this, the average fractal dimension, d_f, of clusters was measured using the cluster growing method. In this method, the radial distance distribution between nodes in the network is calculated; d_f is measured from the resulting double logarithmic plot as $d_f = \ln N(r)/\ln r$, where N is the number of nodes at distance r. The fractal dimension data for finite clusters in simulations with different numbers of water molecules are provided in Table 7.2. A d_f of 2 is consistent with a branched cluster system with average coordination of 2. Combined, all the data shows that ACC

channels were fractal within the ionic framework, and increasing hydration led to connections between nanoporous channels to form the fractal system.

Table 7.2: Mean fractal dimension, d_f, of finite water clusters found in bulk ACC ($CaCO_3 \cdot nH_2O$) simulations with varying levels of hydration, n.

n	d_f
0.55	1.7
0.75	2.1
0.8	2.1
1.0	1.9
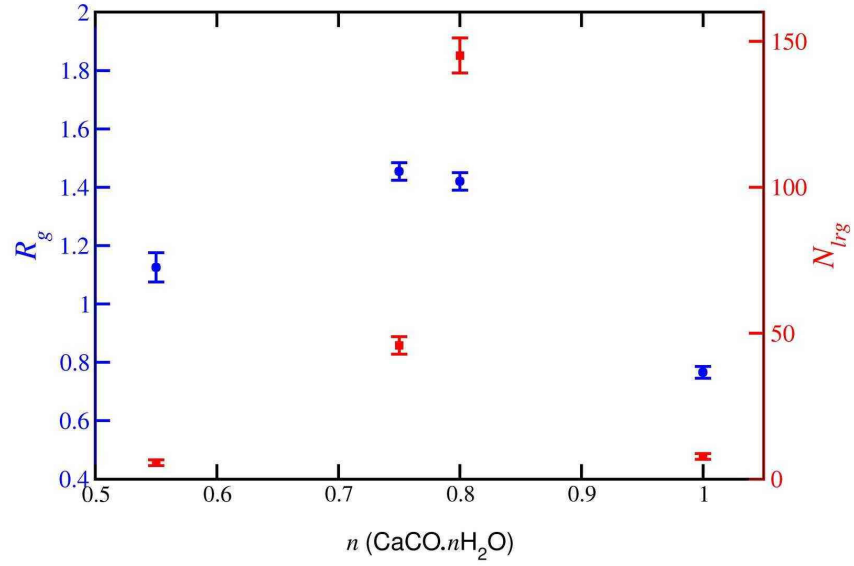
Figure 7.5: Average radius of gyration, R_g (blue), and the average largest cluster size, N_{lrg} (red), as a function of composition of bulk hydrated ACC. The y axis scales in blue and red relate to R_g and N_{lrg} respectively. Averages were measured from the final 0.5-1.0 ns of a 5 ns simulation, and error bars show uncertainties of one standard deviation.

146

7.3.4 ACC and Percolation Theory

The fact that as hydration in ACC increases, water clusters increase in size, both in terms of mass and R_g, along with the fact that at a certain threshold, infinite clusters are found in the material, suggests that percolation theory may be applicable to hydrated ACC. In a stochastic lattice system, such as a random graph, sites can be occupied with probability p or unoccupied with probability $1 - p$. As the probability of site occupancy is increased from zero, stochastic clusters increase in size until an infinite cluster is formed, and the system is termed to be percolating. The percolation threshold, p_c, determines the probability at which a continuous path can be formed through the lattice, and in the current case, the concentration of water in ACC which is required for percolation. The value of p_c is governed by the topology and Euclidean dimension of the lattice.

At the percolation threshold, the infinite cluster which is formed is a fractal, with replicated patterning at every scale. Therefore, while the cell length of the systems studied is on the order of tens of Angstroms, the stochastic nature of cluster formation means that system sizes should not affect the measure of p_c, and analysis of the supramolecular structures arising from simulation will provide all necessary information to understand the percolation in macroscopic systems [Stauffer and Aharony, 1994].

Infinite clusters were only found for ACC hydration levels where $H_2O/CaCO_3$ ≥ 0.8. Only a sample of the total number of configurations at $n = 0.8$ were observed to contain percolating clusters. The fraction of water molecules in infinite clusters at these levels was $20 - 30\%$ and $85 - 90\%$ for $n = 0.8$ and 1, respectively, suggesting that the value of $n = 0.8$ is very close to p_c. The data here are taken as an ensemble average; hence, while there are insufficient statistics to pinpoint the true value of p_c, it is clear that for hydrated ACC, the threshold fulfils $0.75 < p_c < 0.8$. For hydration levels above the percolation threshold, the remaining water molecules still formed smaller finite clusters, hence the decrease in measurements such as R_g and average cluster size for finite clusters as n increased above 0.8 (shown in Figure 7.5).

It is worth considering the sensitivity of these conclusions to the maximum distance used to define a bond, r_{bond}. Throughout, the definition of r_{bond} is $r_{O-O} < 3.8$ Å; this is physically significant, but a sensitivity analysis on this cut-off can help to validate the conclusions. At a value of $r_{bond} < 3.6$ Å, infinite clusters were still only found when $n \geq 0.8$, however the fraction of water molecules within the infinite cluster decreased (68% for $n = 1$). With a slight increase of r_{bond} to 3.85 Å, percolation was found at $n = 0.75$, and percolation was only observed for a fraction of the total number of configurations analysed: infinite clusters were

formed due to fluctuations in the spatial coordinates of oxygen atoms. This is further indicative that the region of n identified for p_c is sensible. With a value of $r_{bond} < 4.4$ Å all hydration levels contained infinite clusters (with 76% of water molecules in percolating clusters at $n = 0.55$). However, at this distance waters in the second coordination sphere are included in the calculation, and so this is not a realistic indication of extended water networks.

Figure 7.6 shows the probability of water–water bonding in ACC. Stochastically formed clusters in percolation theory exhibit a random distribution of sites, connections between which follow a binomial distribution. To verify that site occupation is random, and therefore that percolation theory is applicable, the coordination probability distributions were plotted alongside binomial distributions of the form,

$$f(k, N_b, p) = \binom{k}{N_b} p^k (1-p)^{k-N_b}; \qquad \binom{k}{N_b} = \frac{k!}{N_b!(k-N_b)!} \tag{7.1}$$

where p is the probability of site occupation, N_b is the number of coordinated sites (which is dependent on the distance criteria for bond formation) and k is the lattice topology constant. A number of functions with different k were plotted alongside the measured data from simulation. As shown in Figure 7.6, the best fit with the theoretical data was found when $k = 6$ ($k = 8$ also showed a good fit, but $k = 6$ was the simplest model which fit the data), and while only one hydration level is shown, this was found to be the case for all levels studied. This fit indicates that the lattice has six-fold coordination, and may correspond to a simple cubic topology. This differs from the topology of water clusters in bulk liquid, where a four-site model is used to describe the lattice, and shows that structural water is present in ACC. p_c for site percolation in the simple cubic lattice is 0.312. The probability of site occupation for $k = 6$ gave $p = 0.294$ and 0.318 for $n = 0.75$ and 0.8, which is further support that water in ACC can be represented by a simple cubic lattice model.

The percolation model allows for a description of systems which have microscopically very different structure, but display the same critical point behaviour [Stauffer and Aharony, 1994]. Critical exponents allow for definition of the percolation behaviour which are invariant to the microscopic structure of the system, these neither depend on the type of lattice or p_c, and are measured according to the features of finite clusters in the system. One is D_f, the fractal dimension of the stochastic clusters, which can be measured by investigating the change in R_g of finite clusters as a function of size. Figure 7.7 (a) shows the data for $n = 0.75$ and 0.8. Close to the percolation threshold $R_g(s) \sim s^{\frac{1}{D_f}}$ and for a 3D system,
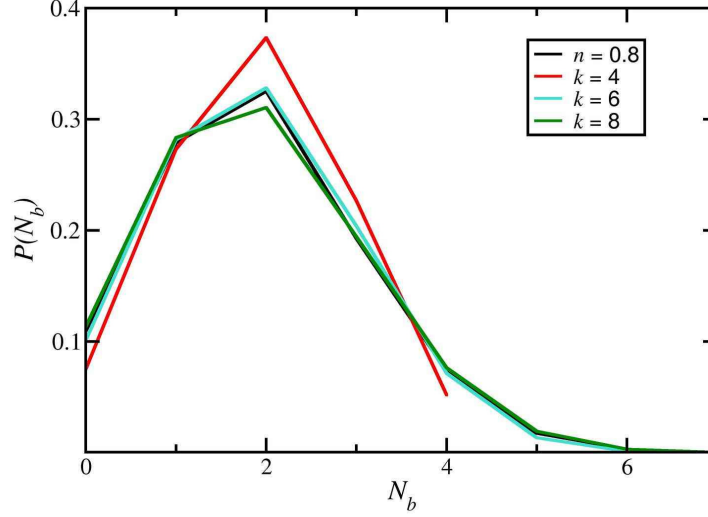
148

Figure 7.6: Probability of finding N_b bonds to other water molecules in amorphous CaCO$_3\cdot$(0.8)H$_2$O (shown in black). A bond was defined by a distance criteria of $r_{O-O} < 3.8$ Å. Also plotted are the theoretical predictions from binomial distributions with k values of 4, 6 and 8. See equation 7.1 for the functional form of the distribution.

$\frac{1}{D_f} = 0.396$. The straight lines fitted through the data show that for both $n = 0.75$ and 0.8, D_f is close to the predicted value ($D_f = 0.431$ and 0.412 for $n = 0.75$ and 0.8, respectively). Another critical exponent is τ, known as the Fisher exponent, which characterises cluster size distributions. Figure 7.7 (b) provides a double logarithm plot of the weight fraction, W_s, of clusters as a function of size s. Through the relationship $W_s = cs^{-(\tau-1)}$, the value of τ was 2.15 and 2.27 for $n = 0.75$ and 0.8, respectively. These values are very close to the theoretical prediction of τ for a 3D system, where $\tau = 2.18$. Overall, hydrated ACC fits very well to the percolation model, and so percolation theory can be used to predict the properties of water in ACC, without the need for a detailed understand of the atomistic features of the material.
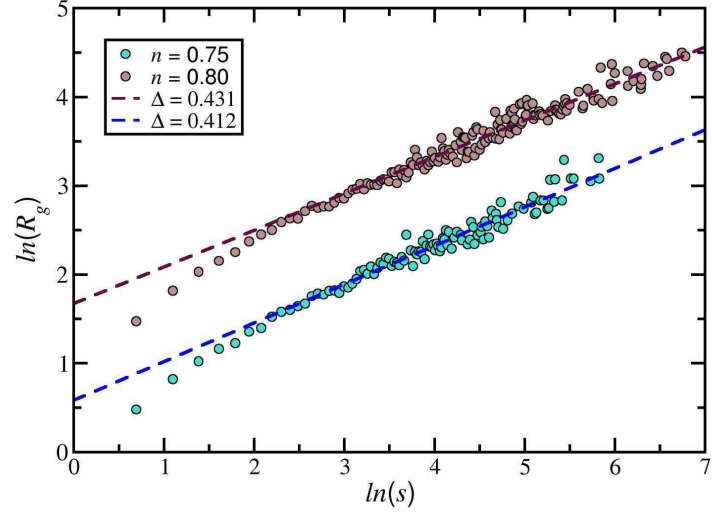
(a)
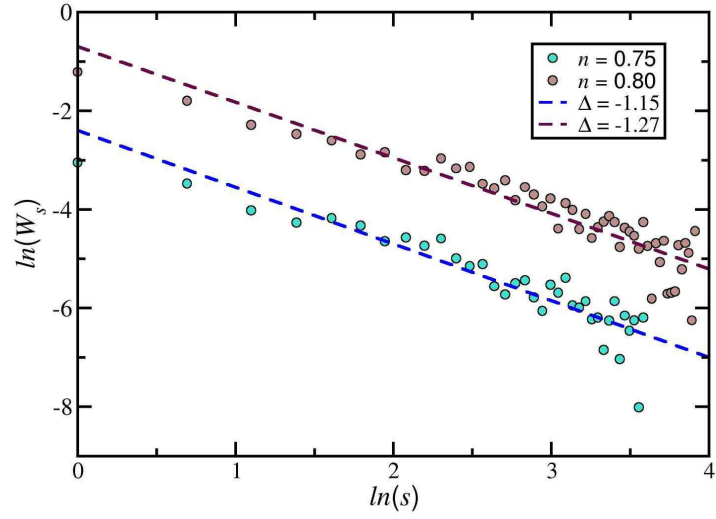
(b)

Figure 7.7: Double logarithm plots of (a) the radius of gyration, R_g, and (b) the weight fraction, W_s, of clusters of size s for ACC with composition $CaCO_3 \cdot nH_2O$. Data have been shifted: $\ln(y) - 1$ and $\ln(y) + 1$ for $n = 0.75$ and 0.8, respectively, to make the graphs clear. Linear regression fits are shown by dashed lines, the gradients of which are provided in legends. Different regions of s have been used in calculation to ensure good statistics, and measurements were taken from relaxed configurations.

7.3.5 Water and Ion Mobility

Movement of water was shown in the last section to lead to fluctuations in the number of molecules forming infinite clusters. To gain a full understanding of the nature of ACC, it is helpful to discuss dynamical features and the significance of these to the stability of ACC. By considering the mobilities of water and ions, information can be gained to help elucidate mechanisms of biomineralisation, and so diffusion analysis has been carried out on bulk and mesoporous ACC systems.

The diffusion coefficients for ions and water were calculated from the time dependence of the mean squared displacement of atoms, and are presented in Figure 7.8. In the case of mesoporous systems, which exhibited a water concentration gradient, a general increase was found in the diffusion as a function of radius from the simulation cell origin (see Figure 7.8). Diffusion in the core ($r_O < 23$ Å) of both systems was uniform (*cf.* regions at the edge of the simulation cell). Ionic diffusion was measured to be on the order of 10^{-14} m^2 s^{-1}. Outside of the core region, D increased through several orders of magnitude, correlating with increased hydration levels up to $n \sim 4$, where the effects of water-filled mesopores became apparent.

Diffusion of ions in bulk ACC systems was also found to increase with hydration, and was on the order of 10^{-13} m^2 s^{-1}. For the hydration level most closely related to that found in biogenic samples, $n = 1$, the diffusion of calcium and carbonate was 1.32 ± 0.01 and 1.23 ± 0.02 $\times 10^{-13}$ m^2 s^{-1}, respectively. Diffusion coefficients for ionic species in calcite have been measured in the range $10^{-23} - 10^{-22}$ m^2 s^{-1}, therefore these measurements suggest a much higher mobility of ions than in crystalline phases [Lahav and Bolt, 1964; Anderson, 1969; Fisler and Cygan, 1999]. The magnitude of diffusion coefficients for ions in ACC is consistent with that of glasses [Brady, 1995].

Water diffusion was consistently faster than that of ions for all systems investigated, and on average, was around four times larger. The intermolecular interactions between ions will be stronger than that for ions with water, and so this is understandable. Diffusion coefficients for water were of the order 10^{-13} m^2 s^{-1} (5.48 ± 0.05 $\times 10^{-13}$ m^2 s^{-1} for $n = 1$). For the mesoporous system with $< n >= 3$, the water content in mesoporous channels was sufficiently large to obtain an estimate of diffusion for H$_2$O$_{Mes}$: $4.15 \pm 0.4 \times 10^{-10}$m^2 s^{-1}. However, in this region a mixture of water environments was present – these included occluded, surface bound and mesopore channel water – and so the average D of *ca.* 10^{-10} m^2 s^{-1} in these regions is an underestimate for the diffusion of H$_2$O$_{Mes}$. This was much greater than diffusion coefficients for internal water, but was still an order of magnitude smaller than bulk diffusivites of liquid water: $D_{bulk} = 2.5 \times 10^{-9}$ m^2 s^{-1} [Holz *et al.*, 2000].

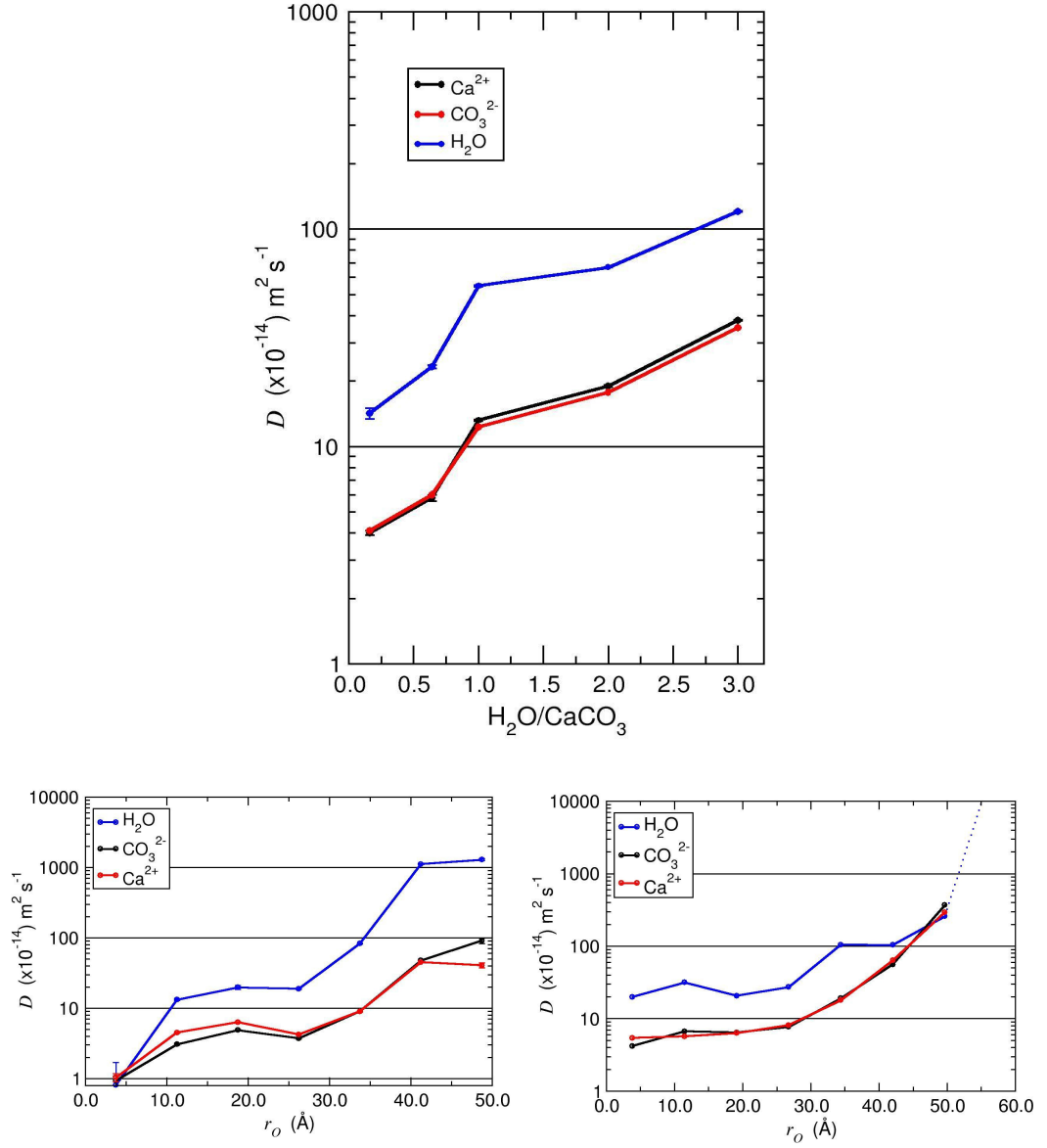Figure 7.8: Diffusion coefficients for hydrated ACC. Top shows data calculated for five levels of hydration. D for $H_2O/CaCO_3$=1–3 was measured from bulk ACC simulations. The data for $n \leq 0.64$ was taken from the carbonate rich region of mesoporous systems ($r_O < 23$ Å) where a plateau was seen in the diffusion coefficients as a function of radius from the simulation origin, as shown (bottom) for $< n >= 1$ (left) and 3 (right).

The effect of percolation can be observed in the measured diffusion coefficients: a step change in the gradient of D as a function of n (see Figure 7.8) is clear between sub- and super-critical ACC (the terms from percolation theory refer to systems below and above the percolation threshold). The presence of continuous clusters allows for a greater degree of translational motion of water, and the increased hydration levels reduce the regions of ACC with a high local concentration of ions, therefore in the range $0.64 < n < 1.0$, the average self diffusion of all water increased.

To further examine the effect of channels on the mobility of water, the diffusion coefficients of different water types have been analysed. Statistics were sufficient to measure diffusion coefficients for H_2O_{Occ} and H_2O_{Chan} in systems with hydration levels of $n = 0.64$ and 1, and these are presented in table F.1. The diffusion coefficient for H_2O_{Occ} was 91% and 42% of that for H_2O_{Chan} for the different hydration levels, respectively. The difference can be ascribed to the fact that at low hydration levels, a significant fraction of water molecules form dimers and trimers, where diffusion is limited similarly to H_2O_{Occ}. In general, the diffusion of water was found to follow, $H_2O_{Occ} < H_2O_{Chan} < H_2O_{Mes}$.

Close inspection of the molecular dynamics trajectories showed that water in the channels could undergo infrequent large hops, during which the water molecules moved several Angstroms within a few hundred femtoseconds, as shown in Figure F.8. Accordingly, it was possible to describe water molecules as "fast" or "slow" depending on whether or not a "jump", defined by displacement of > 3.5 Å, was observed in the multiple nanosecond trajectories. The concentration of fast water was low: typically 1% of species displayed such motion. To quantify this motion, van Hove correlation functions were calculated for the fast and slow water molecules. The van Hove self correlation function takes the form,

$$G_s(r,t) = \frac{1}{N} \left\langle \sum_{i=1}^{N} \delta\left(r - (r_i(t) - r_i(0))\right) \right\rangle \qquad (7.2)$$

which runs over all N water oxygen atom coordinates, r, selected for fast or slow water. The functions are shown in Figure 7.9 (a). The results confirm the significance of the fast/slow distinction for the mechanism of water diffusion through ACC. While the shape of the curve for the slow water is consistent with the Gaussian shape expected for normal diffusion (*i.e.* diffusion as described by the Langevin equation, with frequent random forces affecting particle velocities), the curve for fast waters shows a pronounced broadening in the range 1–3 Å and, for longer time delays, additional peaks occurring up to 4.5 Å. This broadening is present for all time de-
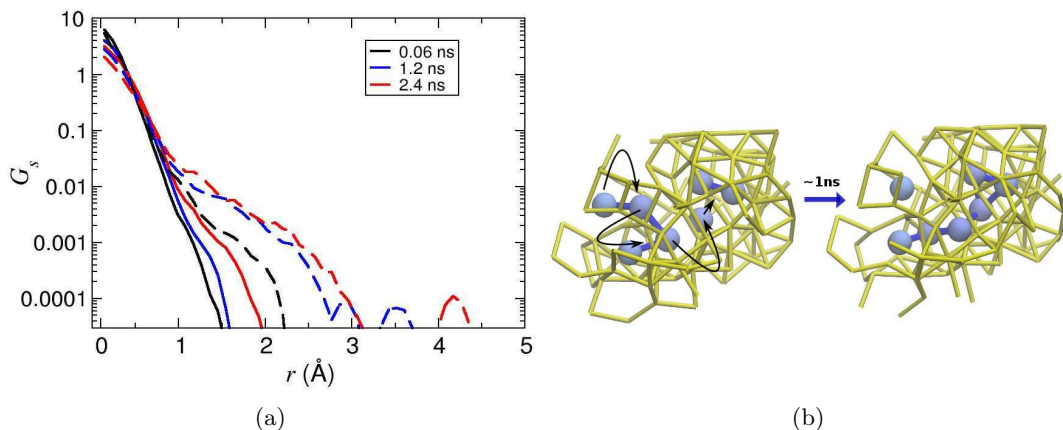
Figure 7.9: Correlated jumps of water molecules in amorphous $CaCO_3 \cdot (0.64)H_2O$. (a): van Hove self correlation functions for slow (solid lines) and fast (dashed lines) water molecules measured at three time delays. (b): A cascade of correlated hops of water molecules (blue) moving through channels in the ionic framework (yellow); arrows show the progression of displacements, each of which occurred < 1 ps, although the concerted motion took ~ 1 ns.

lays shown, but is substantially more pronounced for time delays greater than 1 ns. This is indicative of a jump diffusion mechanism in which there is a range of jump distances, and multiple jumps occurring on the time-scale of 1 ns.

Inspection of the trajectories revealed cascade-type motions of rotations and translations, with correlated hops of adjacent molecules, as the system sampled configuration space. Figure 7.9 (b) provides a typical example of the concerted motion that was observed. The arrows in the Figure show that as one water molecule was displaced, this generated further translational motion to an adjacent chain of water molecules in a channel. The result was the formation of two new water channel clusters and a correlated restructuring of the surrounding ionic framework. Each of the consecutive jumps occurred on the order of femtoseconds, but the collective local reconfiguration took approximately one nanosecond. Cascade translational motion was observed at all hydration levels, but these became more prevalent as the water level increased. Experimental measurements suggest that a large amount of water in ACC undergoes rotations on the time scale of milliseconds [Nebel *et al.*, 2008]. The time scales on which rotations and translations take place here is much smaller than this, and is inaccessible by the methods used for the experimental estimate.

7.4 Summary and Conclusions

Using the force field of Raiteri and Gale [Raiteri and Gale, 2010] bulk hydrated ACC, $CaCO_3 \cdot nH_2O$, was prepared for a range of hydration levels, $n =$ 0, 0.55, 0.75, 0.8, 1, 2, and 3. Specially prepared ACC was also generated, and this was composed of configurations with a concentration gradient of ions and water, comprising an ion-rich core, where the hydration level was $n = 0.16$ and 0.64 up to around 23 Å from the origin. These systems were described as mesoporous due to the emergence of water-rich and ion-rich domains. The average $H_2O/CaCO_3$ was 1 and 3 in the two mesoporous systems constructed. Analysis of atomistic spatial correlation functions showed there was little change in the atomistic structure of ACC as the hydration level changed, and that these matched well to experimental measurements. The adopted force field was therefore a suitable one to study ACC across length scales.

Analysis of the supramolecular structure of ACC at all hydration levels highlighted the existence of water-rich cavities and channels which penetrated into the ionic framework. Water molecules did not display a significant amount of hydrogen bonding within the channel clusters; instead, H-bonding between water and carbonate ions was more likely, as has been observed in other simulation studies [Saharay et al., 2013]. Clustering of water molecules within the ACC was characterised using a distance criterion derived from the first minimum in the water–water RDF: $r_{O-O} < 3.8$ Å. From analysis of water clusters, three types of water were identified: occluded, channel and – within the mesoporous systems – mesoporous water. Experimental studies [Radha et al., 2010] have shown the existence of at least two different water types, namely, structural and physisorbed water, but the present study suggests that there are at least two types of structural water. In bulk ACC with $n = 1$, 5% of water was found to be occluded in the framework cavities, with the remaining water existing in channels. At the lowest hydration level, $n = 0.16$, 80% of water was occluded. Diffusion coefficients confirmed the existence of the three types of water identified, with the diffusion coefficients for occluded water being a small fraction of those for channel water, and mesoporous water having diffusion coefficients three orders of magnitude higher than other water types.

Water molecules in clusters were found to have largely two and three-fold coordination up to $n = 1$. Through visual inspection, along with knowledge of the radius of gyration distribution and fractal dimensions, clusters were demonstrated to take the form of branched chains with a random network of molecules. Cluster size distributions showed a wide range of cluster sizes, with the maximum cluster size increasing up to $n = 0.8$. At this hydration level, percolating clusters emerged

which spanned the simulation cell boundaries and formed continuous networks. The percolation model for $3D$ systems was applied to bulk ACC with $n = 0.55 - 1$. Water site occupation was found to be random with connections between sites modelled by a simple cubic lattice. The percolation threshold was found to fall close to the higher limit in the range $n = 0.75 - 0.8$. Analysis of critical exponents further justified the applicability of the percolation model to ACC. This is a significant step in the understanding of this material as it demonstrates that percolation theory can be a very useful tool for predicting the hydration of ACC.

Following these results, it is possible to speculate on the implications to biomineralisation, and to consider the mechanism of dehydration. The self diffusion of water in channels was on the order expected for glasses, and so, even in percolating systems, ACC with composition $CaCO_3 \cdot H_2O$ will not readily dehydrate under standard conditions by normal diffusion alone. However, correlated jumps in the spatial coordinates of water molecules, along with coordinated reconstruction of the ionic framework, was observed on the order of nanoseconds. With 3 Å jumps occurring roughly every nanosecond, it is possible to displace a water molecule by a distance of 10^{-7} m within a microsecond, and considering ACC typically forms nanoparticles, this introduces a possible mass transport mechanism for the expulsion of water from channels in ACC on experimental time-scales. Moreover, an internal pressure, perhaps arising form the evolution of a crystallisation front, would increase the rate of these processes, and so the percolating channels in ACC offer a route to the removal of water from the mineral. Nonetheless, a fraction of occluded water will remain in ACC, saving significant reconfiguration of the ionic framework, and therefore would contribute to the energy barrier associated with crystallisation via solid-state transformation.

Ihli and co-workers showed stepwise ACC dehydration (as discussed in Chapter 1), with three dehydration regions identified [Ihli *et al.*, 2014]. An increase in activation energy to removing water was found as the water content in ACC decreased. These findings are completely complementary to the results of the work in this Chapter. As demonstrated here, water in mesoporous channels has relatively high diffusion. In these channels there is no calcium carbonate and so removal of this water would not affect the average hydration level of ions. Agglomeration of particles was observed in experiment, and so the water lost in the first dehydration stage is likely to be water trapped in layers between particles. While the mesoporous water studied here and water trapped between ACC in experiment is not the same, it is possible that both of these types of water have low activation energies to desorption. The second dehydration stage corresponds well with the presence of

percolating channels, where a combination of increased diffusion and increasing stability of dry ACC (from the application of heat) would result in the expulsion of water from ACC, and a reduction in system size as density increases to maximise ion packing. From cluster analyses in this work, a fraction of water would remain occluded in the ionic framework, and removal of this fits well with a final dehydration stage, where a high activation barrier would exist as significant restructuring of the ionic framework is required to form a route to water expulsion. The fractions of water in each dehydration stage are also comparable to the water types identified here, with the majority of water excluded from ACC at stage two (65%; see Figure 1.6). 15% of water remained and was lost at the final stage in experiment, which is higher than the amount of occluded water identified here when $n = 1$. However, as the hydration level is reduced and restructuring of the ionic framework takes place, it is likely that further channel water will become occluded, and it is worth reiterating that the definition for occluded water here is a conservative one. In this study, approximately 20% of water was occluded at $n = 0.64$ and this fits well with the experimental measures.

In nature, hydrated ACC with composition $CaCO_3 \cdot H_2O$ is stable, and can be stored *en route* to crystalline phases [Gong *et al.*, 2012]. As ACC *in vivo* contains additives, such as ion impurities and organic molecules, the mechanism of stabilisation is an enticing one to speculate upon. The most common ionic impurity in ACC is Mg^{2+}, for which water in the solvation shell is more strongly bound than in Ca^{2+}. It is possible then, that strongly coordinating water around a small fraction of magnesium ions form blockages in channels and reduce the concentration and frequency of water molecules undergoing jumps. Organic molecules may act to block channels completely and hinder percolation, therefore eliminating any water expulsion pathway. Repeating the analyses presented in this Chapter for ACC which contains additives and impurities would go a long way to testing this conjecture. Understanding these stabilisation processes could be crucial to explaining biomineralisation.

Chapter 8

Conclusions

8.1 Summary

Using a simple and efficient algorithm, a wide range of states for calcium carbonate clusters have been sampled, as discussed in Chapter 3. Thousands of optimisations allowed sampling of low energy clusters with a range of mass densities and coordination in the size range suggested for prenucleation clusters (PNCs) [Gebauer and Cölfen, 2011], and containing a maximum of 80 ions. Clusters were optimised in vacuum and showed calcium-carbon coordination which ranged from an average of around two to five. Optimisations in the presence of water where $H_2O/CaCO_3=1$ were performed for clusters of 4, 6, 10 and 15 formula units of calcium carbonate. For small systems, water solvated low energy clusters which were found in anhydrous optimisations, with internal water found only in the largest clusters studied. A negative surface charge was found for clusters with a relatively high mass density.

When simulated in water at relatively high concentration (see Chapter 4), small clusters were found to be dynamically ordered with continual breaking and reforming of bonds between ions, and a wide distribution of clusters sizes was found. A large number of cluster conformations were sampled, and ions had high mobility within clusters, consistent with dynamically ordered liquid–like oxyanion polymer (DOLLOP) which has been suggested as the structural form of PNCs [Demichelis *et al.*, 2011]. Large clusters – particularly those containing more than sixty ions – of low and high density were found to be (meta)stable on the time-scales of the simulations (*i.e.* up to 50 ns). Ions at the surface of high density clusters partially dissolved, retaining limited binding to the cluster core but maximising hydration; the core of the largest clusters remained anhydrous. High coordination clusters were found to have a higher potential energy than low coordination clusters of the same

size. Low density clusters expanded in solution to maximise the hydration of ions, but generally ions remained within a single cluster throughout the simulation. Bond lifetimes were considerably longer than for small clusters of DOLLOP and average coordination was more akin to those measured for liquid droplets described by Wallace *et al.* [Wallace *et al.*, 2013]. However, the differences between the structural phases of DOLLOP and dense liquids modelled by computation is ambiguous. The higher coordination levels for dense liquid proposed by Wallace *et al.* may arise over time in DOLLOP. Clusters were therefore described as dense liquid nanoparticles (DLNPs).

Free energy calculations were performed to assess the stability of clusters in Chapter 5. Using a combination of adiabatic biased molecular dynamics (ABMD) and Umbrella Sampling (US), the free energy profile for clusters as a function of coordination was measured. At small cluster sizes a wide minimum was found in the free energy surface (FES) with small ion associates and DLNPs being the stable state. The energy required to change the average calcium–oxygen (of carbonate) coordination number between two and three was within $k_{\mathrm{B}}T$. For larger systems, which also had higher concentrations, a transition in the stability of different phases was observed. For $30\,\mathrm{CaCO_3}$ a wide basin was found; there appears to be a small energy barrier ($\sim 3k_{\mathrm{B}}T$) separating two minima centred at $N_{Ca-OC} = 2.5$ and 4.3. The minimum at lower values of the collective variable covers coordination levels for DLNPs, while that at $N_{Ca-OC} = 4.3$ represents a high density state. High coordination numbers were observed internally within cluster configurations sampled around the minimum of $N_{Ca-OC} = 4.3$, and a large number of water molecules were found below the cluster surface. The average number of water molecules was roughly equal to the number of calcium ions in the cluster, and the stoichiometry matched that of stable hydrated amorphous calcium carbonate (ACC) found in biogenic samples [Addadi *et al.*, 2003]. At the largest cluster size the FES showed a single minimum at high coordination that was consistent with the high density 30 formula unit cluster (although we note that the FES had not fully converged for small coordination numbers). Raiteri and Gale found hydrated ACC more stable than crystalline phases at these sizes [Raiteri and Gale, 2010], and so at the concentrations studied, hydrated ACC is the most probable solid phase to emerge from solution.

The effect of temperature and additives on cluster dynamics was also considered in Chapter 4. Increasing the temperature led to higher coordination levels in low density clusters of $20\,\mathrm{CaCO_3}$. This is consistent with the phase diagram for carbonates which precipitate readily at high temperature. The change is indicative of a low entropy to solvation: maximising the entropy of water requires a

reduction of water-order in calcium solvation shells, and increased calcium carbonate coordination results. The effect of aspartate (ASP) on clusters was to reduce dissociation of clusters into smaller species. The binding between functional groups of ASP and cluster ions was strong, and while dynamic ordering of liquid clusters was found, ASP retained large cluster sizes. This corroborates with experimental findings [Gower and Odom, 2000; Schenk *et al.*, 2012] that (poly)aspartate stabilises dense liquid phases of calcium carbonate. The energy of clusters increased in ASP–bound clusters, although this approached a limiting value which was approximately $8\ k_{\mathrm{B}}T$ per formula unit higher than ASP–free low density clusters.

Simulations at experimental conditions in Chapters 6 showed that concentration is an important consideration in simulations of calcium carbonate. Making use of the Henderson–Hasselbach equation, the pH in cluster and free ion solutions was defined by the total unbound HCO_3^-:CO_3^{2-} ratio at equilibrium. Clusters dissociated close to the experimental pH and concentrations where PNCs are suggested to be stable [Gebauer *et al.*, 2008], converging to the cluster size distributions found for random ionic solution systems. The most dominant species in solution were ion pairs of $CaCO_3^0$ and $CaHCO_3^+$. While some larger clusters were observed, the maximum size was no more than three to four ions, and these were found with low probability. The system size was not found to change the results.

Cluster simulations were conducted for low density $20\ CaCO_3$ at 20–51 mM. At the lowest concentration, cluster dissolution led to the formation of a large number of ion pairs and a smaller concentration of larger associates which contained a maximum of ~ 6 ions. At 51 mM large clusters were found which contained up to around 20 ions and were clearly DLNPs. The formation of DLNPs is therefore not only dependent upon pH but also concentration. Analysis of the simulation showed that the thermodynamics of speciation at high and low concentrations are inconsistent, and fitting to a speciation model at low concentration, as in the study of Demichelis *et al.* [Demichelis *et al.*, 2011] is incorrect. At experimental conditions ion pairs will preferentially form over larger clusters even at the highest levels of pH. Simulating larger clusters at 20 mM suggested that cluster dissolution is very slow and metastable clusters will persist in solution certainly on the nanosecond time-scale. Higher concentrations will further reduce the dissolution rate of large clusters.

The first solid calcium carbonate phases in many organisms is amorphous calcium carbonate (ACC). Hydrated ACC with composition $CaCO_3 \cdot H_2O$ is stable in biology, and transient ACC dehydrates before crystallisation takes place. Analysis of the structure and dynamics of ACC within this thesis was performed in Chapter

7, which showed that the material is heterogeneous. Water channels penetrate the ionic framework, and typically water molecules form chains and branched clusters when a distance criteria for cluster formation of $r_{O-O} < 3.8$ Å (corresponding to the first minimum in the water O–O radial distribution function) was invoked. The size of channels increased as a function of hydration, with a higher probability of finding occluded water molecules in the framework when the hydration level was low. We were able to define three water types: occluded water, water in nanometre sized channels, and water in mesoporous channels. The diffusion of water increased as a function of hydration, and the mobility of channel water was larger than for occluded water molecules. Water molecules showed correlated translational motion, with large displacements of water molecules in channels on just a few femtoseconds leading to the displacement of adjacent water molecules, a few fs later, in a "cascade" mechanism of diffusion. The different types of water match well with the finding that several stages of dehydration take place within ACC upon heat treatment, as molecules with different activation energies to dehydration are lost [Ihli *et al.*, 2014].

A site percolation model on a simple cubic lattice was applied to ACC. Critical exponents fitted well with a three dimensional percolation model. The validity of the percolation model indicates that the formation of channels is stochastic and that fractal clusters emerge as the hydration level increased. The percolation threshold was found to be in the range 0.75–0.8 $H_2O/CaCO_3$. Combined with the type of water diffusion discussed above, the findings show that if a strong driving force to dehydration is present, water channels offer a route to the expulsion of water from ACC. At the lowest hydration levels ($H_2O/CaCO_3=0.16$) around 80% of water molecules were found to be monomeric and occluded in the framework; this water is likely to remain in ACC unless significant rearrangement of the framework occurs. The removal of water concomitant with crystallisation in ACC has been seen in experiment [Ihli *et al.*, 2014].

8.2 Nucleation

From the results found in this thesis, it is possible to relate the findings to other studies of calcium carbonate nucleation and to propose a mechanism for this process.

At experimental conditions, clusters were found to be unstable and dissolved into ion pairs. Emerging solids are likely to form from these smallest associates in a homogeneous bulk phase. Experimental studies presented here (performed by Smeets *et al.*) and from Hu *et al.* [Hu *et al.*, 2013] have indicated that crystal nucleation can be described in terms of classical nucleation theory (CNT). The

simulation results fit with this description as no PNC or dense liquid phase was found at conditions which compare well with experiment (where often pH≈ 10.0 and buffer concentration is around 10 mM). This contrasts with the results of relatively recent computational studies of nucleation by Demichelis *et al.* [Demichelis *et al.*, 2011]. However, while this study carried out simulations at experimental concentrations, the sampling was insufficient. The simulations in Chapter 6 highlighted that longer simulation times, than the maximum of one nanosecond performed in the other computational studies, are required to establish an equilibrium.

It is not so easy to dismiss the existence of PNCs in solution. Gebauer *et al.* measured the size of PNCs to be 2 nm by centrifuge experiments [Gebauer *et al.*, 2008]. One can question the effect of this type of experiment on the stability of species in solution in light of kinetic studies on amorphous phase stability [Wang *et al.*, 2013; Aziz *et al.*, 2011]. Nonetheless, electron microscopy studies have visualised particles of similar sizes in solution before nucleation [Pouget *et al.*, 2009; Kellermeier *et al.*, 2012]. The general consensus is that these clusters contain large amounts of water [Kellermeier *et al.*, 2012; Gebauer and Cölfen, 2011] and are liquid–like [Demichelis *et al.*, 2011; Gebauer *et al.*, 2014]. Gebauer *et al.* earlier suggested [Gebauer and Cölfen, 2011] that ACC with crystalline–like domains cannot be formed from ion pairs as these units cannot encode structural information. However, this is also true in the case of a dynamically ordered liquid such as DOLLOP. Computational studies [Demichelis *et al.*, 2011; Wallace *et al.*, 2013] have indicated no barrier to the growth of liquid clusters, with the free energy of clusters decreasing as a function of increasing size. While bicarbonate may limit cluster size, it is not clear from the computational studies why DOLLOP should have a size distribution that tails around 3 nm [Kellermeier *et al.*, 2012].

In this work, it was only at concentrations much higher than typical precipitation experiments ($\gtrsim 40$ mM) where clusters of significant sizes were (meta)stable in solution. While the experiment can be described by an average concentration, undoubtedly, microscopic regions of relatively low and high supersaturation can be present. For instance, in the titration experiments, as solutions initially mix the local concentration of ions and supersaturation level in microscopic volumes will be high. This may lead to the formation of the PNCs which have been visualised and measured in experiment. The implications of the supersaturation on nucleation were discussed by Hu *et al.* [Hu *et al.*, 2013]. As shown here, any clusters which form will be long lived, and so detection by experimental analyses is possible. Ultimately though, during equilibration within the bulk homogeneous solution phase, liquid clusters up to 80 ions in size will dissociate in the absence of additives, unless

transformation to another phase occurs.

A dense liquid phase was found in solution at high supersaturations by Smeets *et al.*, which emerged in the nucleation region of the LaMer diagram from titrations (see Figure 6.8). It is possible that the DLNPs found in Chapters 4 and 6 are the structural form of this phase, which is consistent with the conclusions of Wallace *et al.* [Wallace *et al.*, 2013]. Further analyses must be carried out to compare experiment with simulation (comparison of TEM images from experiment and those calculated for DLNPs *in silico* will be performed); however, it is the view of the author that DLNPs are precursors to an ion–rich bulk liquid phase, which remain particulate in simulation due to system size effects. If the supply of ions from the lean solution in equilibrium with DLNPs in simulations were not limited, then further growth of liquid–like clusters would occur, as demonstrated by Demichelis *et al.* for high pH [Demichelis *et al.*, 2011]. While solid nucleation may take place within the dense liquid bulk phase, the results of this study do not contradict with a classical description of solid nucleation and growth at modest supersaturations. Indeed, experimental measurements discussed in Chapter 6, highlight the growth of vaterite by addition of free ions to growing nuclei.

Free energy calculations (Chapter 5) showed that DLNPs around 1.5 nm in diameter (and containing around 60 ions) could transform to hydrated ACC with no energetic cost. Clusters above this size were only stable in a dense, hydrated amorphous form. While improvement in the characterisation of these clusters (for instance, by measuring ion mobilities) should be performed, they are indicative of hydrated ACC. This is further suggestive that at high supersaturations, it is not unlikely that dense liquid and amorphous solid phases may be found in solution alongside crystal nuclei upon nucleation.

With stable ion pairs dominating in solution, it is possible that amorphous phases could be nucleated from ion pairs via a classical route. Amorphous phases will likely have a lower interfacial energy, while the bulk free energy of crystalline phases will outweigh any amorphous phase. Growth of ACC by ion pair addition at high concentration was shown to carry no free energy penalty and offers a fast route to precipitation; crystalline phases are only suggested to be more stable than ACC at sizes of around 4 nm [Raiteri and Gale, 2010]. The presence of both ACC and crystalline phases in solution is therefore possible, with the amorphous phase consumed in the growth of crystal in accordance with experimental observations [Pouget *et al.*, 2009].

Taking all of the information into consideration, then from this work it is possible to hypothesise a mechanism for the nucleation of calcium carbonate. It

is likely that crystallisation takes place via a classical pathway from ion pairs in a bulk homogeneous solution; this work provides strong evidence that PNCs are unstable in homogeneous solution. The general consensus that calcium carbonate nucleation is non–classical should be treated with caution. Other simulations at high concentration suggest that dense liquids will form, and on further increase in the concentration (this could be the microscopic local concentration), hydrated ACC phases will precipitate. This is indicative of spinodal decomposition between a liquid and solid phase. The phase diagram of Wallace *et al.* as discussed in Chapter 1 is reproduced here in Figure 8.1. The simulations in this study generally fit well with the phase diagram; however, the finding that ACC becomes stable with a relatively high concentration suggests that the true phase diagram is more complicated than the one presented. Our studies suggest that as well as liquid–liquid spinodal decomposition taking place, liquid–solid decomposition can also happen. This is because free energy calculations showed a single minimum at high concentration relating to hydrated ACC, and the fact that heating calcium carbonate liquid clusters at high concentration leads to the formation of solid. Solid precipitation in the metastable region (marked by the yellow region within black and red curves of Figure 8.1) is likely to occur if the local supersaturation levels are high. It is not clear whether this solid decomposition takes place directly as a function of ion activity (and therefore fits in the two dimensional phase diagram) or whether higher dimensions are required to explain the phase behaviour.
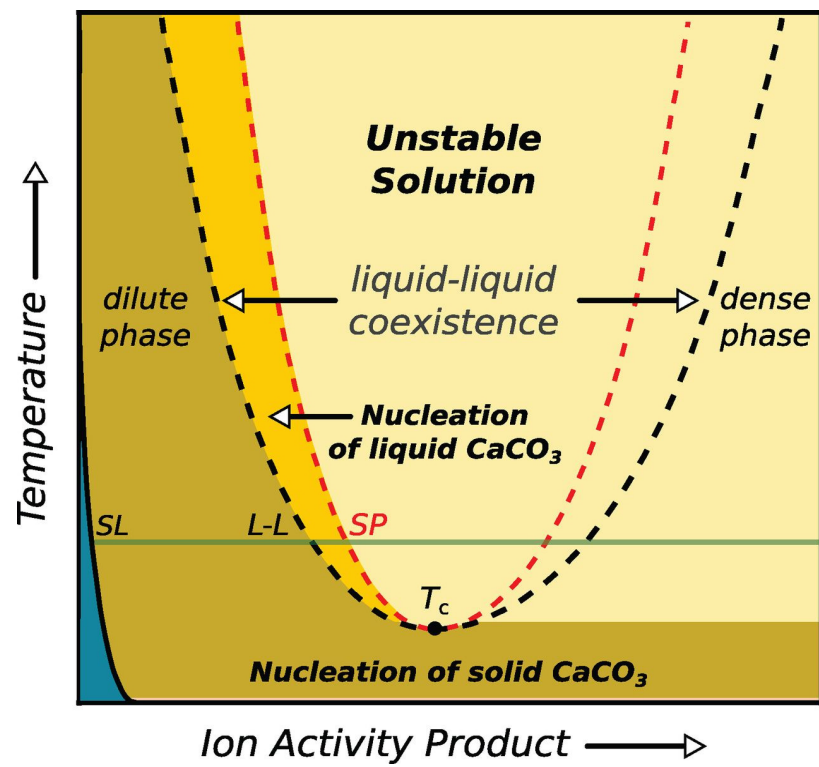
Figure 8.1: A schematic of phase diagram for calcium carbonate, as suggested by Wallace *et al.* [Wallace *et al.*, 2013].

8.3   Outlook

While the results presented here do make a valuable contribution to the under-standing of calcification *in vivo* and *in vitro*, the conclusions are speculative to some extent. Furthermore, questions remain which can be addressed following this work.

As discussed previously, the concentration adopted in simulations is crucial to estimating the correct equilibria between species in solution. To validate the proposed phase changes which take place on nucleation from ion pairs, establishing equilibria between ion pairs, dense liquid phases and amorphous calcium carbonate needs to be performed at experimental concentrations. To do this efficiently at an atomistic level requires methods which remove the computational burden associated with simulating large amounts of water to achieve desired concentrations. Gibbs ensemble molecular dynamics (GEMD) has been identified as a suitable method to conduct such experiments on the computer [Hentschke *et al.*, 1996]. The algorithm involves simulation of multiple periodic systems, in which atoms/molecules can move between simulation cells in an effort to establish an equilibrium. The method has proven to be successful at finding equilibria in small organic molecule systems such as hexane [Kotelyanskii and Hentschke, 1996], but will need to be adapted in or-der to simulate calcium carbonate. For instance, moving charged species between simulation cells requires careful thought on the treatment of electrostatic potential energies.

While the free energies of liquid phases and ACC have been established at high concentrations, in order to validate the hypothesised thermodynamic links between phases, Umbrella Sampling calculations at experimental conditions should be performed. These will be expensive for large systems; however, the effect of changing concentration on the smallest clusters might give insight into how the relative stabilities of phases change as a function of concentration.

The dehydration mechanism for ACC seems to fit well with experimental findings. However, the reason for stability in hydrated ACC, where the hydra-tion level exceeds the percolation threshold, is unclear. Comparable simulations and analyses, such as those presented in Chapter 7, can be performed for systems which contain impurities and additives. The effect of aspartate and magnesium on the percolation threshold and the nature of water channels in ACC is particularly intriguing.

Finally, modifications in the model potentials used to simulate calcium car-bonate will lead to a fuller understanding of the stability of species which have been identified to precede crystallisation. The force fields developed by Gale an

166

co-workers accurately reproduce key structural and thermodynamic features of calcium carbonate phases [Raiteri *et al.*, 2010; Raiteri and Gale, 2010; Demichelis *et al.*, 2011]. However, these models do not include hydroxide or magnesium ions which are known to be present *in vivo*. Including impurities and polarisation in calcium carbonate simulations are an important, and perhaps, a crucial next step to elucidating molecular mechanisms to biomineralisation in simulation.

Appendix A

Calcium Carbonate RDFs
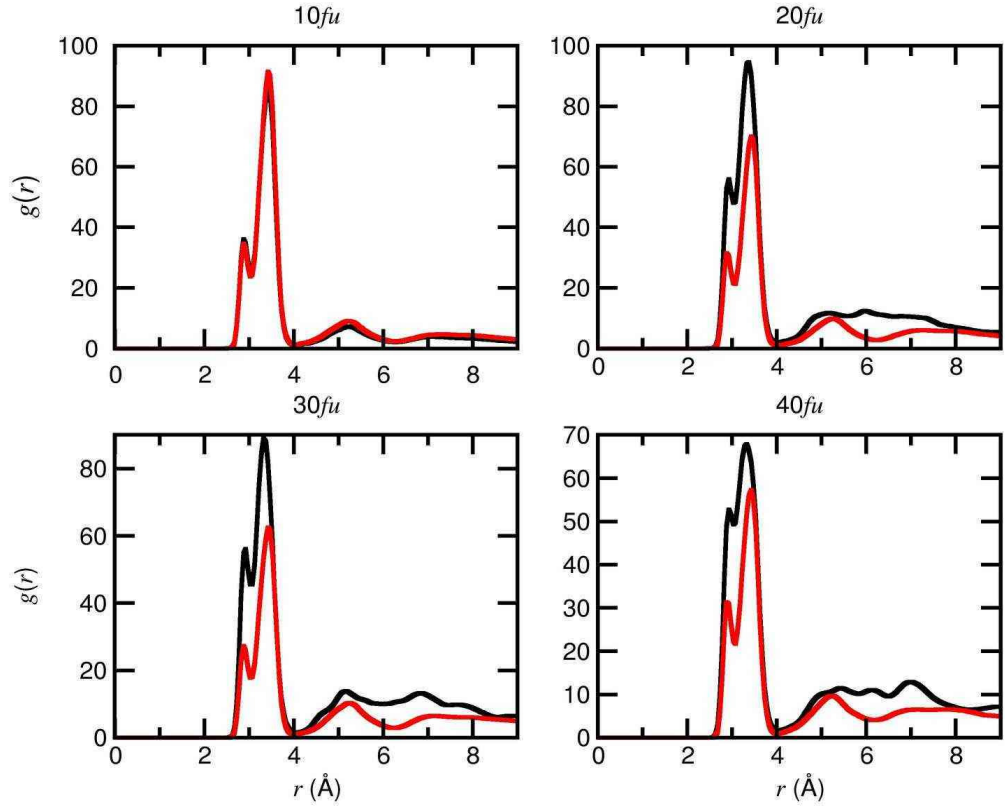
Figure A.1: Radial distribution functions for calcium and carbon atoms measured from a 50 ns simulation of clusters in water. Clusters of size $n\mathrm{CaCO_3}$, where $n =$10, 20, 30, and 40, which had open (red) and compact (black) configurations were simulated . The two curves for each cluster size therefore highlight the coordination spheres for the least and most dense clusters after relaxation in water.
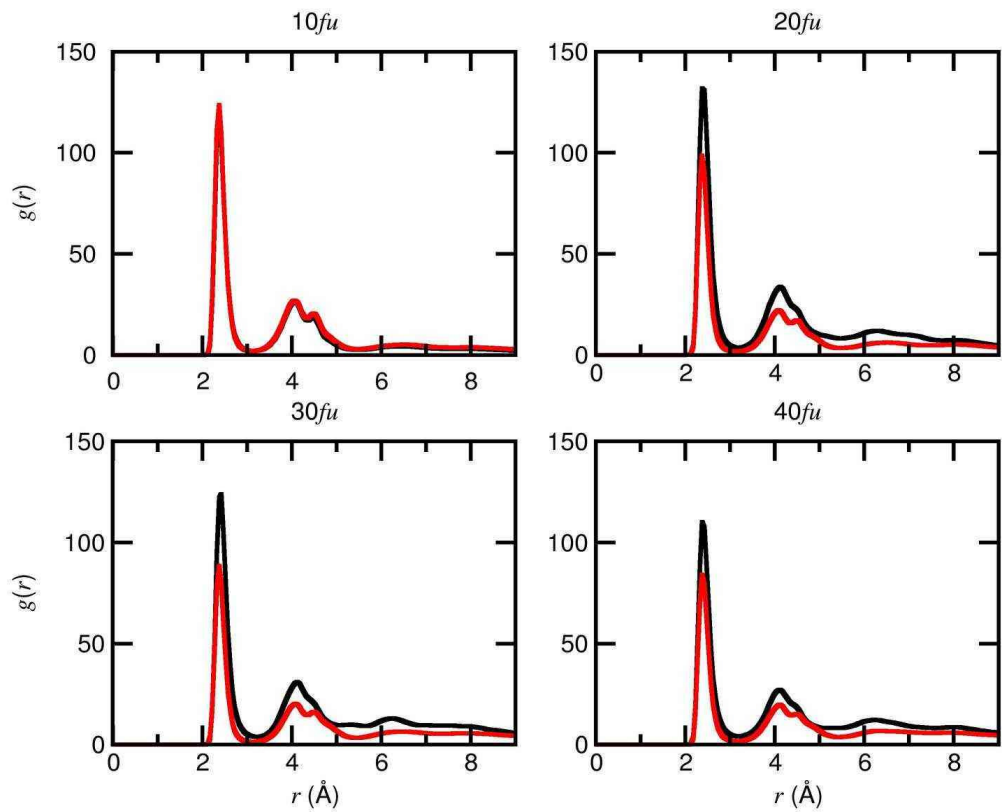
Figure A.2: Radial distribution functions for calcium and oxygen of carbonate atoms measured from a 50 ns simulation of clusters in water. Clusters of size $n\text{CaCO}_3$, where $n =$10, 20, 30, and 40, which had open (red) and compact (black) configurations were simulated . The two curves for each cluster size therefore highlight the coordination spheres for the least and most dense clusters after relaxation in water.
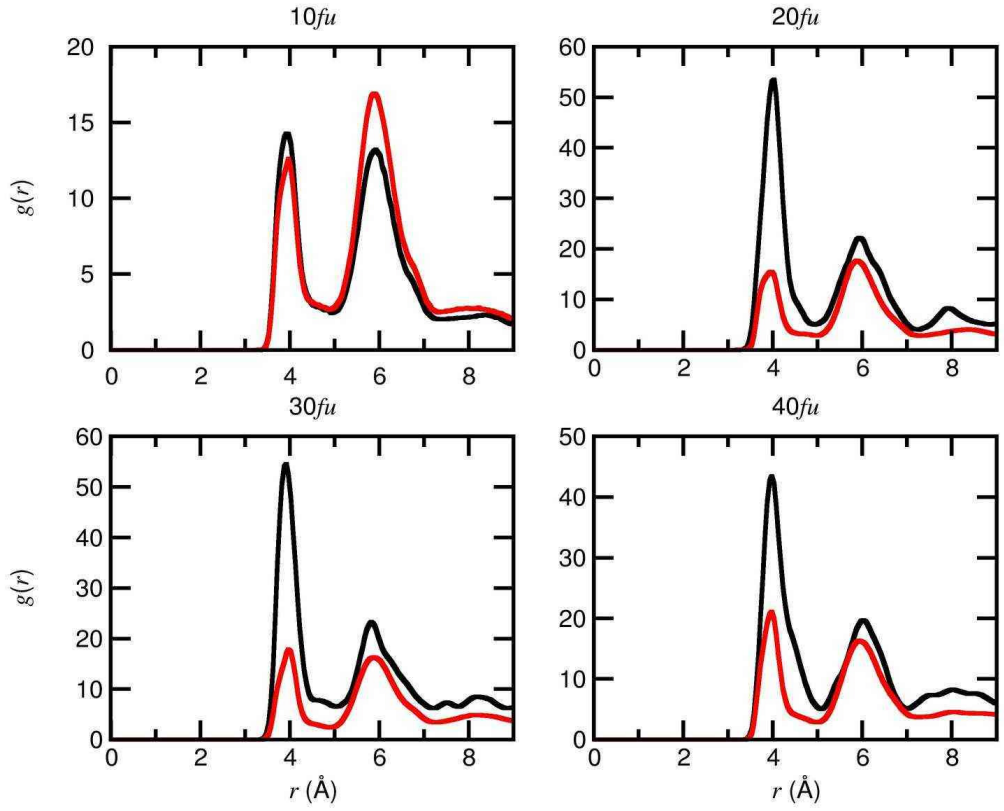
Figure A.3: Radial distribution functions for calcium atoms measured from a 50 ns simulation of clusters in water. Clusters of size $n\mathrm{CaCO_3}$, where $n =10$, 20, 30, and 40, which had open (red) and compact (black) configurations were simulated . The two curves for each cluster size therefore highlight the coordination spheres for the least and most dense clusters after relaxation in water.
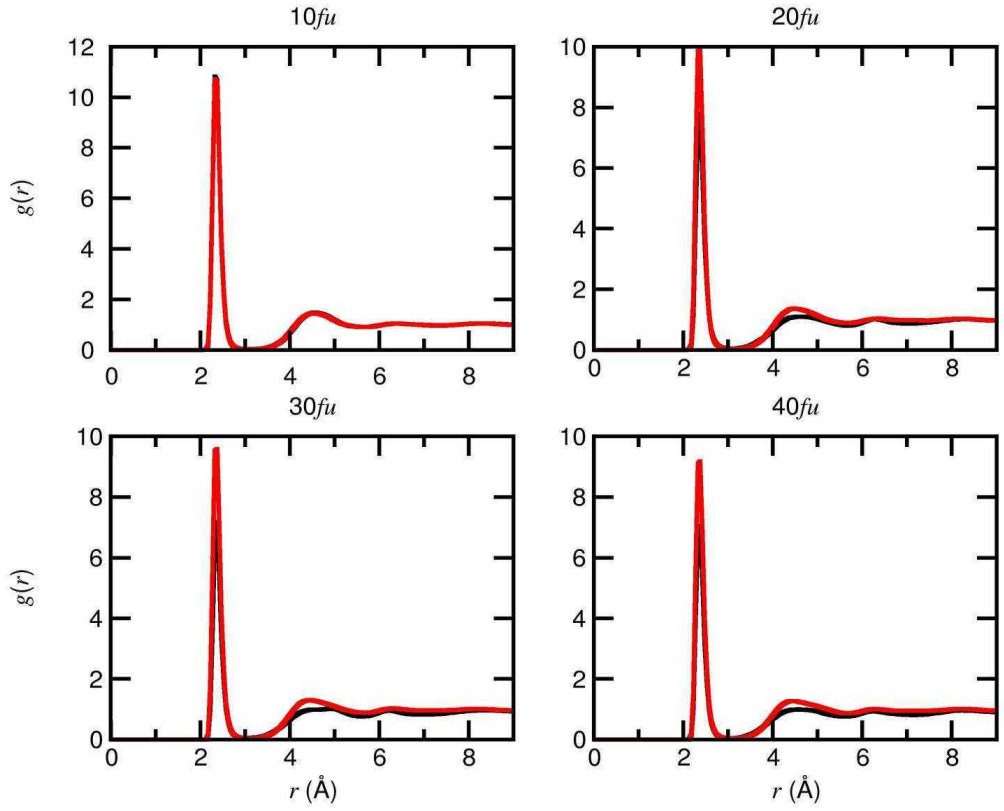
Figure A.4: Radial distribution functions for calcium and water oxygen atoms measured from a 50 ns simulation of clusters in water. Clusters of size $n\text{CaCO}_3$, where $n =10$, 20, 30, and 40, which had open (red) and compact (black) configurations were simulated . The two curves for each cluster size therefore highlight the coordination spheres for the least and most dense clusters after relaxation in water.

Appendix B

Generating Calcium Carbonate Clusters

Calculating Shape Order Parameters

To classify a calcium carbonate cluster with N atoms, a number of shape order parameters were calculated from an initial calculation of the gyration tensor,

$$\mathbf{S} = \frac{1}{N}\sum_{i}^{N}(\mathbf{r}_i - \mathbf{r}_{CM})^2, \tag{B.1}$$

where $\mathbf{r}_i$ and $\mathbf{r}_{CM}$ are the positions of atom i in the cluster and the cluster centre of mass, respectively. Diagonalising $\mathbf{S}$ provides the eigenvalues of the tensor: $\lambda_1^2 \leq \lambda_2^2 \leq \lambda_3^2$.

The squared radius of gyration is calculated as,

$$R_g^2 = \mathrm{Tr}(\mathbf{S}). \tag{B.2}$$

Asphericity is given by,

$$asph = \lambda_3^2 - \frac{\lambda_1^2 + \lambda_2^2}{2}, \tag{B.3}$$

while acylindricity is defined by,

$$acyl = \lambda_2^2 - \lambda_1^2 \tag{B.4}$$

Finally, relative shape anisotropy is given by,

$$\kappa^2 = \frac{asph^2 + \frac{3}{4}acyl^2}{R_g^4}. \tag{B.5}$$

Table B.1: Average Ca–C coordination (within 3.8 Å), n_{Ca}, and configurational energy per formula unit, U, for clusters of $n(CaCO_3 \cdot H_2O)$ optimised in vacuum.

n	n_{Ca}	U/eV
4	3.0	-28.18
	2.8	-27.97
	2.3	-27.74
	2.0	-27.67
6	3.3	-28.61
	3.0	-28.53
	2.5	-28.16
	2.2	-27.97
10	4.2	-29.04
	3.6	-29.03
	3.0	-28.78
	2.5	-28.38
15	4.2	-29.26
	3.7	-29.25
	3.1	-28.92
	2.7	-28.54
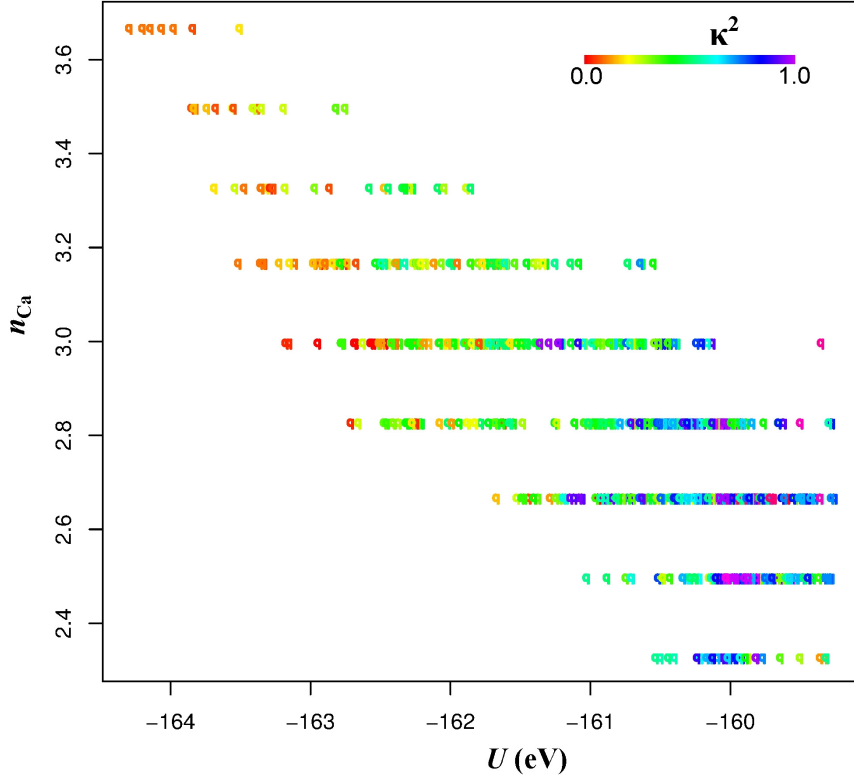
Figure B.1: A plot of average Ca–C coordination number (n_{Ca}) as a function of potential energy (U) for ten thousand minimised structures of $6\,CaCO_3$ from ions in the gas phase. The points are coloured according to a gradient representing cluster shape anisotropy, $\kappa^2 = \frac{\lambda_b^2 + \frac{3}{4}\lambda_c^2}{R_g^4}$, where λ_b and λ_c are the second and third largest eigenvalues from the diagonalisation of the gyration tensor matrix, and R_g is the radius of gyration. The minimum energy data points for each discrete level of n_{Ca}, span the n_{Ca}, U and κ^2 distributions.
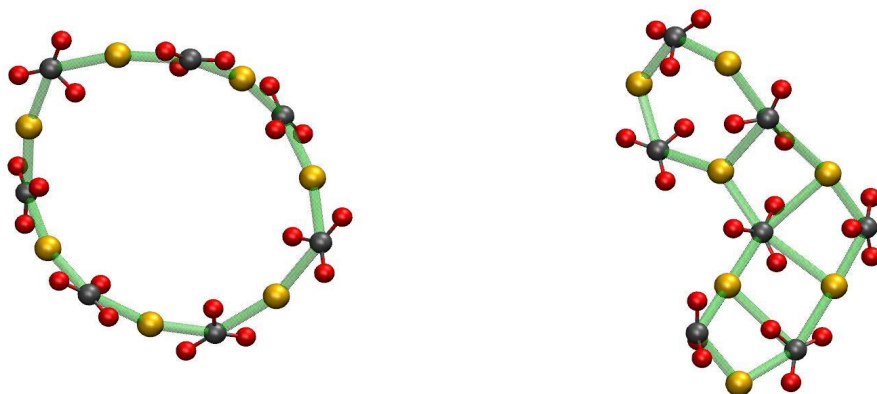
Figure B.2: Common structural motifs found during optimisation of random structures of calcium carbonate in vacuum are the ring (left) and off-set dual chains (right) which are exemplified by configurations of $7\,CaCO_3$. Yellow, black and red atoms represent calcium, carbon and oxygen, respectively. Green lines highlight $r_{Ca-C} < 3.825$ Å.

(a)   (b)
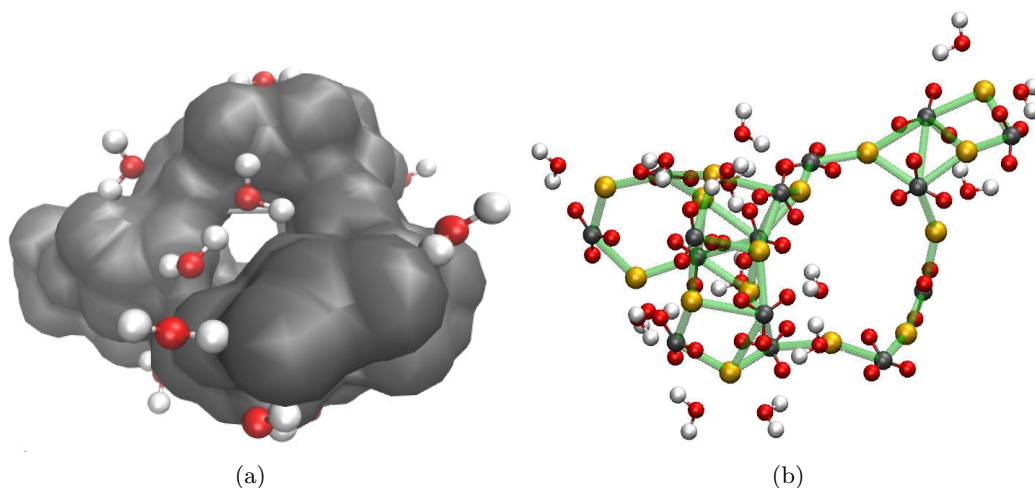
Figure B.3: Optimised clusters of $15\,(\,CaCO_3 \cdot H_2O)$ from random structure searches in vacuum. (a) shows the ionic surface of a cluster (in black) with water penetrating a hole in the cluster, while (b) shows an open cluster with low coordination structural motifs, as found for smaller clusters. Yellow, black, red and white atoms show calcium, carbon, oxygen and hydrogen, respectively.

177
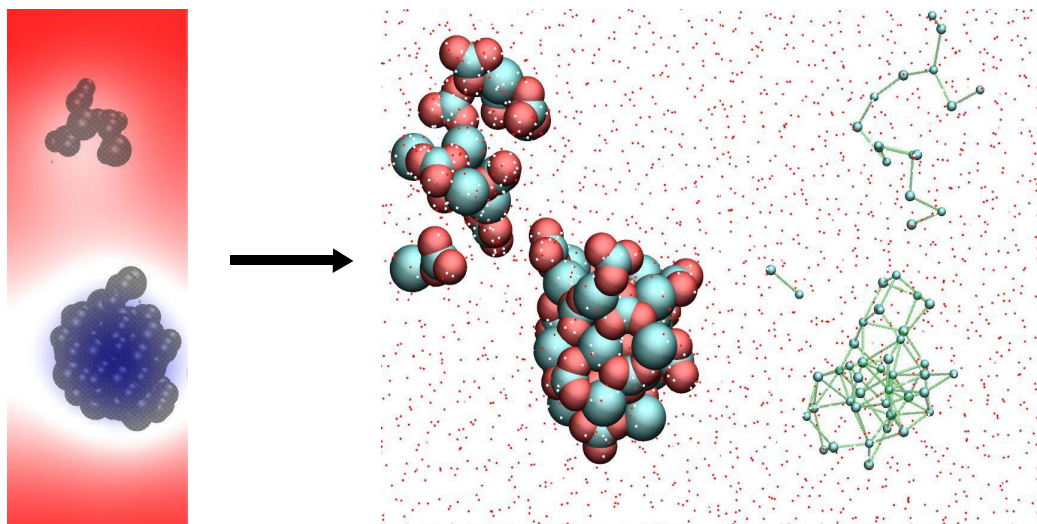
Figure B.4: Configurations taken from a simulation of high density $20\,CaCO_3$ and $9\,CaCO_3$ clusters in water. After initial relaxation, the distance between cluster centres was incrementally reduced using an applied force in the form of a harmonic potential. Left shows the initial configuration with clusters shown in spacefill representation and water removed. The electrostatic potential, as measured half way through the z axis of the simulation cell is also shown, with red and blue shading representing regions of negative and positive charge, respectively. Right of the arrow shows the final configuration after deformation of clusters. The initial representation shows atom van de Wall radii, with carbon and oxygen atoms in cyan and red. Far right shows the same configuration from an alternative angle with only carbon and calcium atoms shown as spheres connected by green lines where $r_{Ca-C} < 3.825$ Å. Solvent water is shown on the right as red points. Both clusters deformed as they were brought together, and the smallest cluster disassembled to form an open, low coordination cluster.

178

Appendix C

Simulating Calcium Carbonate Clusters in Aqueous Solution

Solvating Clusters

To be computationally efficient, one would like to simulate cluster systems with as little water as possible to ensure true solution behaviour. The energy of the system as a whole, E_T, can be treated as two interacting energetic systems: the cluster and its solvation shell(s), E_c, and the surrounding bulk solution, E_b, *i.e.* $E_T = E_c + E_b$. Once the cluster solvation shell(s) is complete, any additional water can be thought of as non-interacting with the cluster, and so water sufficiently far from the cluster should have the energy of bulk water, E_{H_2O}.

$$\frac{E_m - E_n}{m - n} \simeq (m - n)E_{H_2O} \tag{C.1}$$

As shown in the equation, if a system contains nH$_2$O molecules and water is added to give a system with mH$_2$O molecules, then the difference in energy of the system on addition of $m - n$ water should converge to the equivalent amount of bulk water, as shown in the example in Figure C.1.

For cluster systems of different size and density, this test must be repeated; however, the computational effort of doing this does not outweigh that of running unnecessarily large aqueous solutions. Tests were carried out for $n =$1, 5, 10, 20, 30 and 40, for nCaCO$_3$ clusters with maximum and minimum Ca–C coordination.
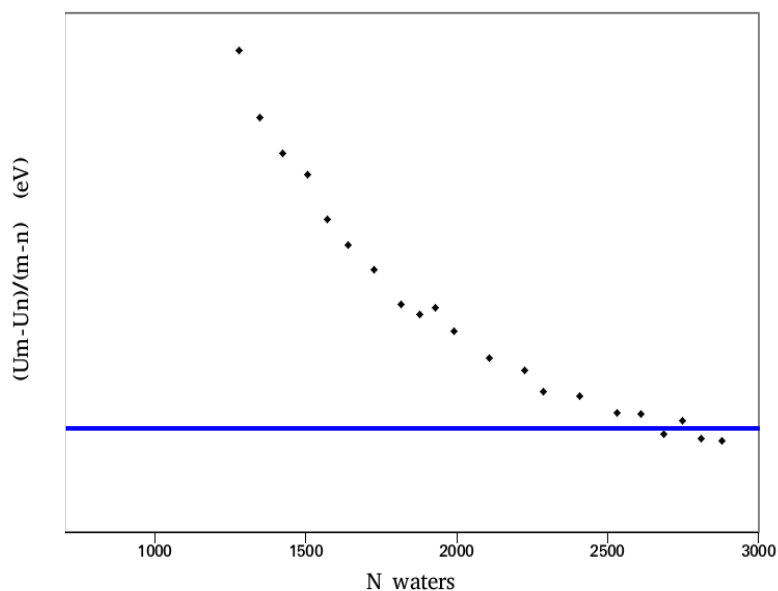
Figure C.1: Change in potential energy, U, of a cluster system as a function of water added $(m - n)$, against the total number of waters in the system containing a low density $40\,\mathrm{CaCO_3}$ cluster. The y axis ranges from -0.44 to -0.431 eV and the blue line indicates the energy of a bulk water molecule of SPC/Fw as estimated from a bulk water simulation.
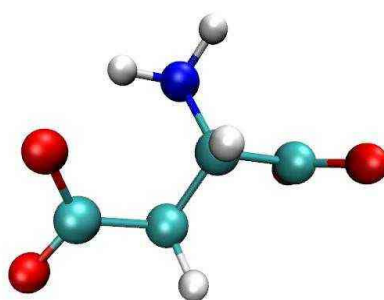
Figure C.2: Optimised conformation of aspartate (ASP) in vacuum, with cyan, red, blue and white atoms representing carbon, oxygen, nitrogen and hydrogen, respectively.
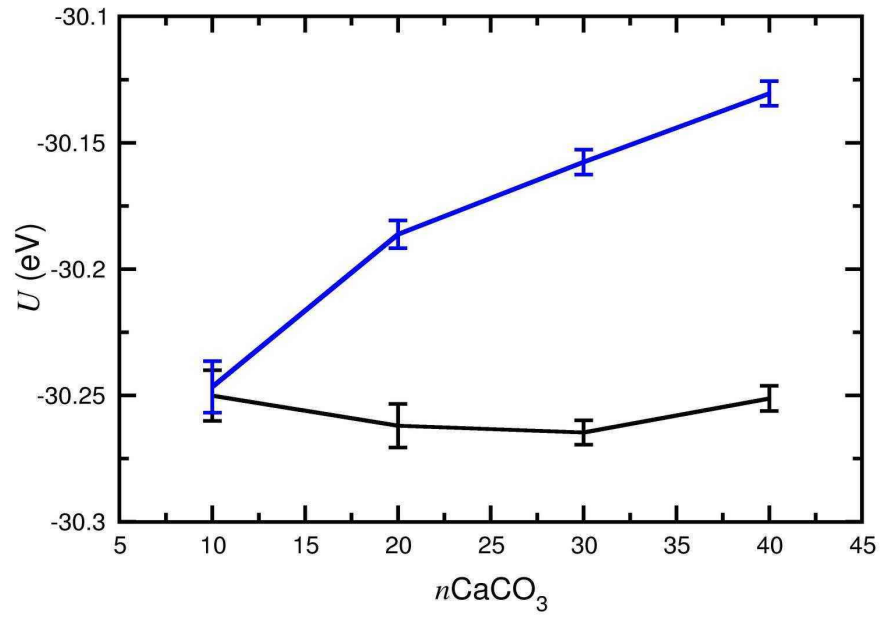
Figure C.3: Potential energies, U, per formula unit of calcium carbonate measured from the final 5 ns of a 50 ns simulation of open (black) and compact (blue) clusters in water. Error bars show uncertainties of one standard error of the mean. Bulk water energies have been subtracted.
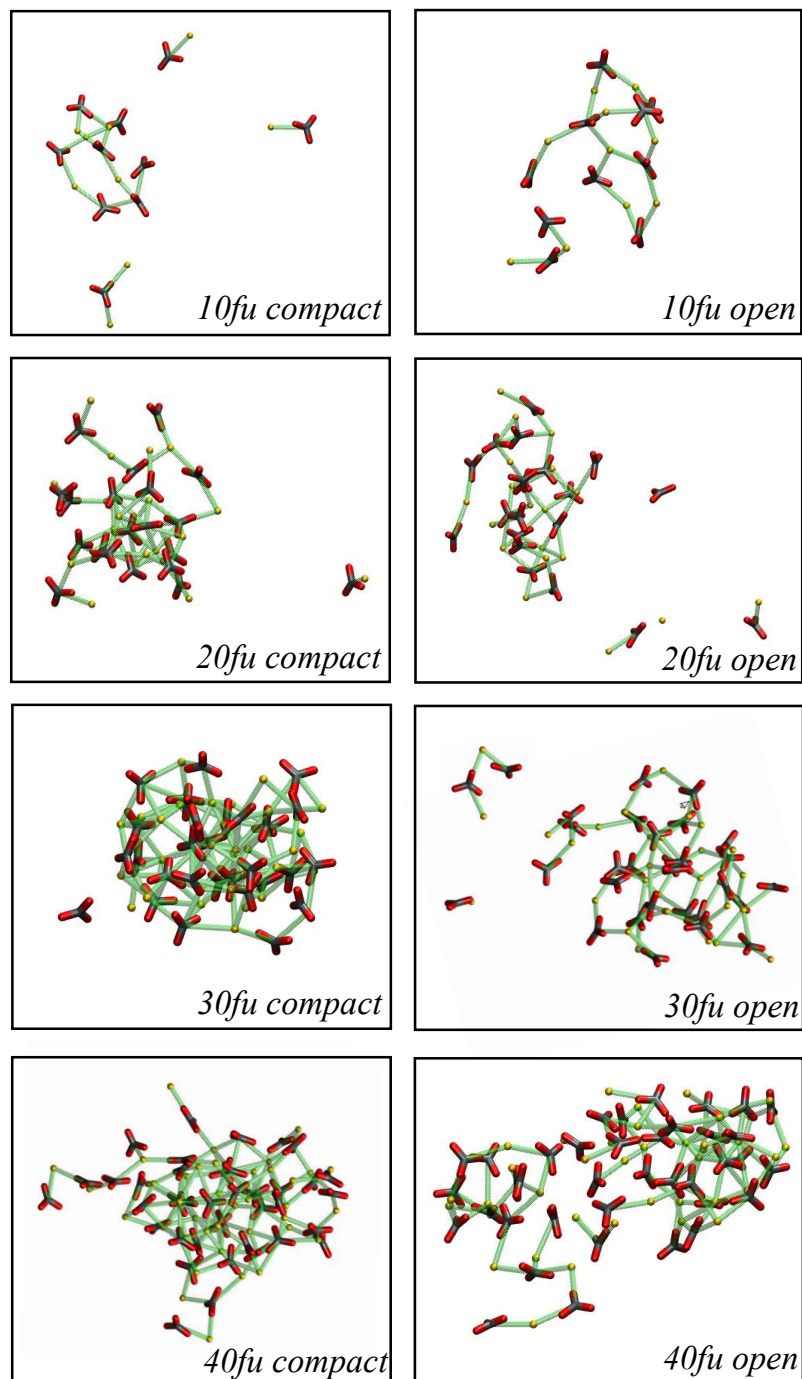
Figure C.4: Final snapshots of open and compact cluster configurations after 50 ns of simulation in water. Clusters of $n\text{CaCO}_3$ were simulated where n =10, 20, 30, and 40. Yellow, red and black atoms represent calcium, oxygen and carbon atoms, respectively, while green lines highlight connected calcium and carbon atoms where $r < 3.825$ Å.
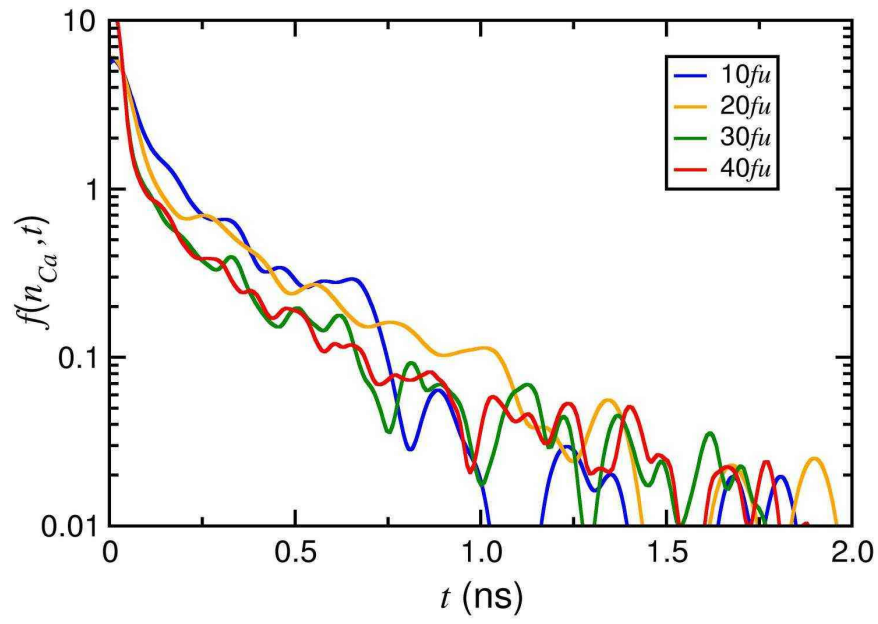
Figure C.5: Calcium carbonate bond lifetime probability densities measured from the final 10 ns of simulation of low density clusters, in the range 10–40 formula units, in water. Bonds were defined by a distance criteria between calcium and carbon atoms of < 4 Å. The average lifetime of the bond is given on the x axis.
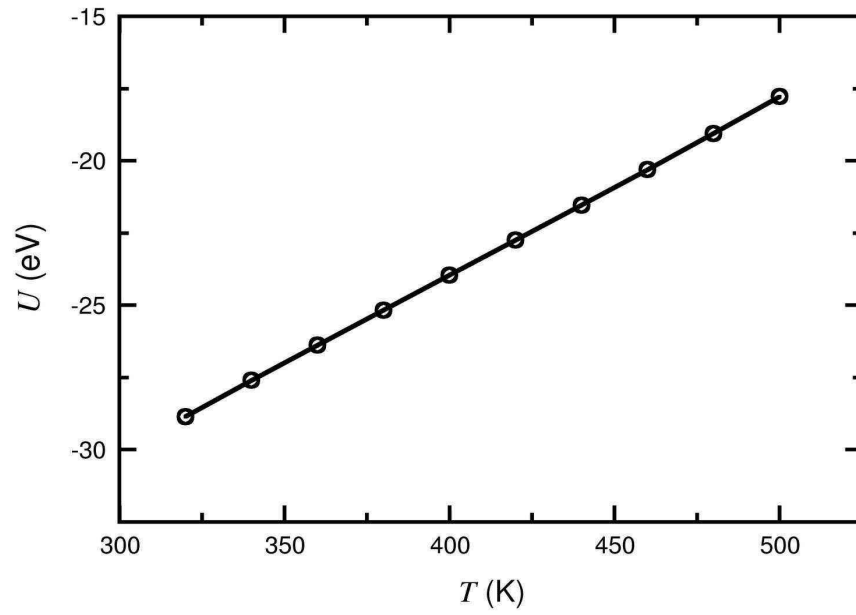
Figure C.6: Potential energies per calcium carbonate unit, U, as a function of temperature, T, measured from the final 5 ns of a 20 ns simulation of a low density $20\,CaCO_3$ cluster in water. Errors of one standard deviation of the mean are smaller than the data point markers at this scale.
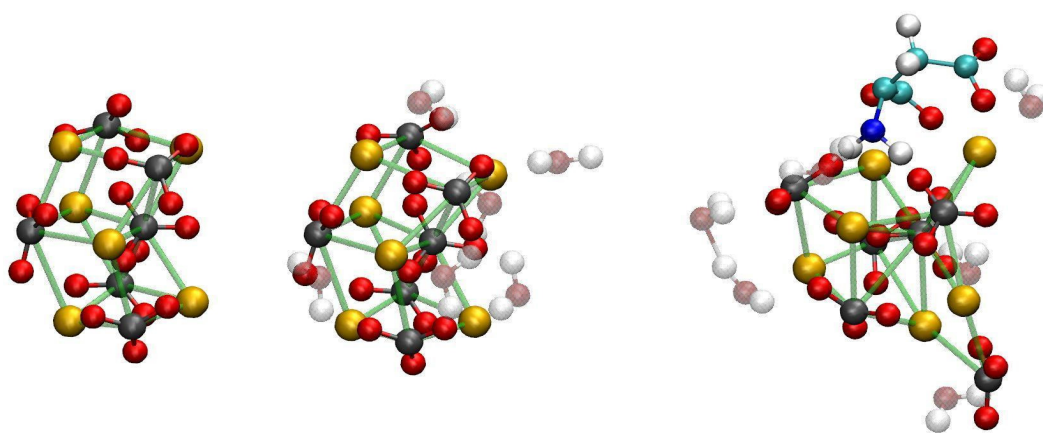
Figure C.7: Lowest energy configurations of $6\,CaCO_3$, $6\,(\,CaCO_3 \cdot H_2O)$ and $(6\,CaCO_3 \cdot H_2O) \cdot ASP$ from random structure searches in vacuum. Carbon of carbonate is shown in black and that of ASP in cyan. Red, blue, yellow and white atoms show oxygen, nitrogen, calcium and hydrogen, respectively. Green liness show connections between carbon of carbonate and calcium of < 3.825 Å. Water molecules are shown as transparent for clarity. While water serves to solvate the minimum energy configuration found from anhydrous optimisations, ASP perturbs the structure due to competitive binding from functional groups to ions.
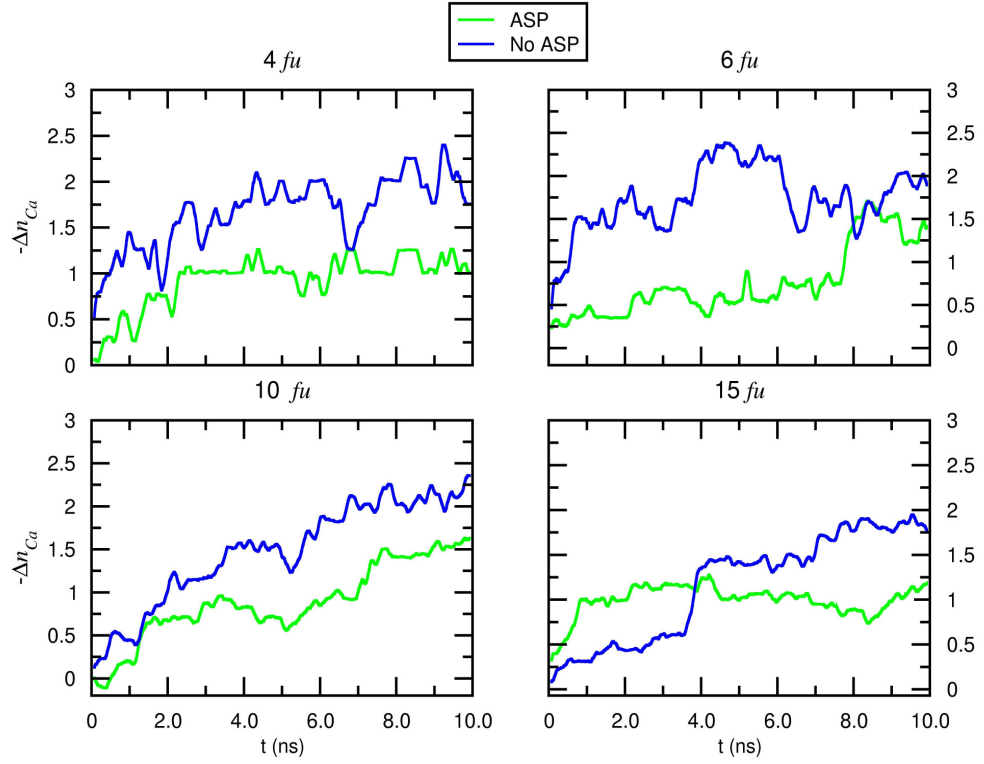
Figure C.8: Time evolution of the change in the absolute value of Ca–C coordination, Δn_{Ca}, for clusters containing 4, 6, 10 and 15 $CaCO_3$ units minimised in vacuum, with and without ASP, and simulated in water with MD for 10 ns. ASP inclusive clusters disassemble more rapidly, and to a greater degree than equivalent sized clusters which did not contain ASP.
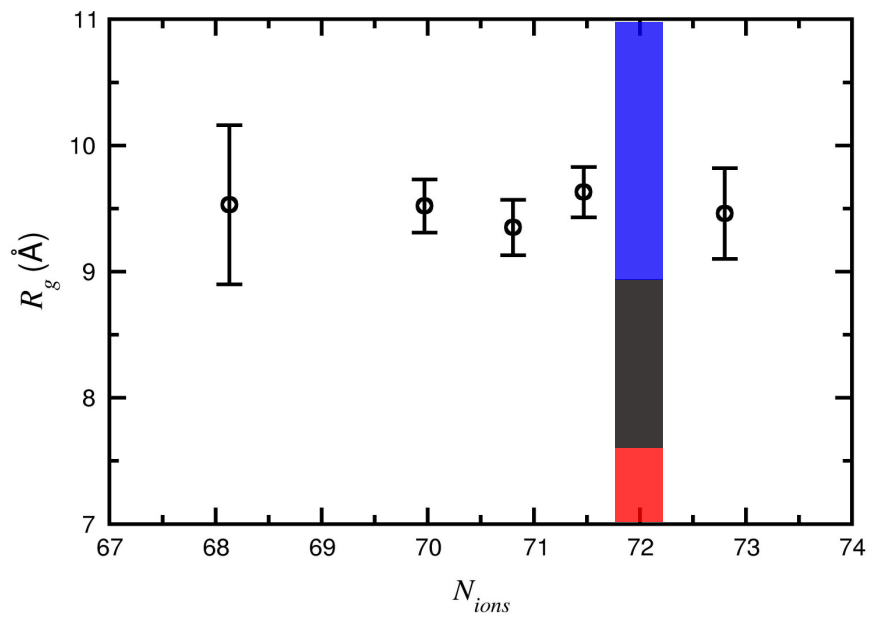
Figure C.9: Radius of gyration, R_g, as a function of cluster size, taken from the final 10 ns of simulation of a low density $40\,CaCO_3$ cluster in water. R_g and N_{ions} for clusters were calculated over 2 ns windows. The plot shows that R_g had converged. Overlaid are the coloured bars which represent distinct regions of accessible R_g for dry (red), wet (black) and DOLLOP (blue) states for a $CaCO_3$ particle of calcium carbonate as provided by Demichelis *et al.* [Demichelis *et al.*, 2011].

Appendix D

Cluster Free Energies
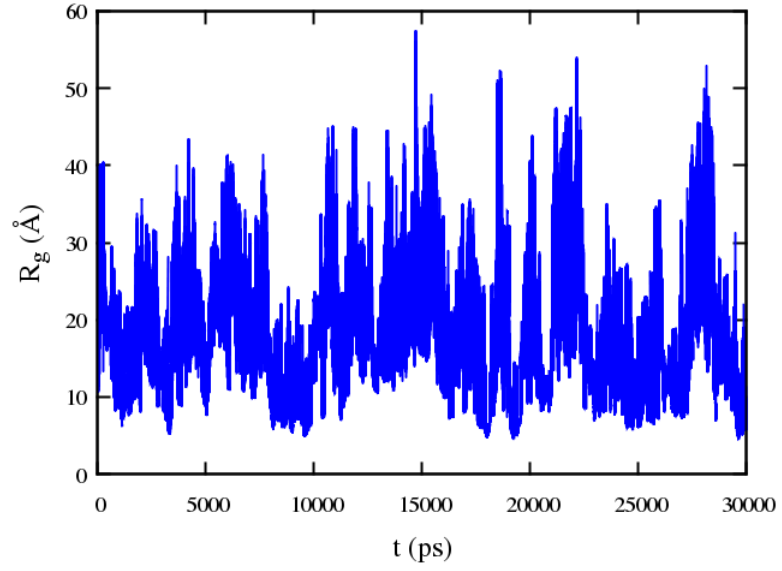
(a)
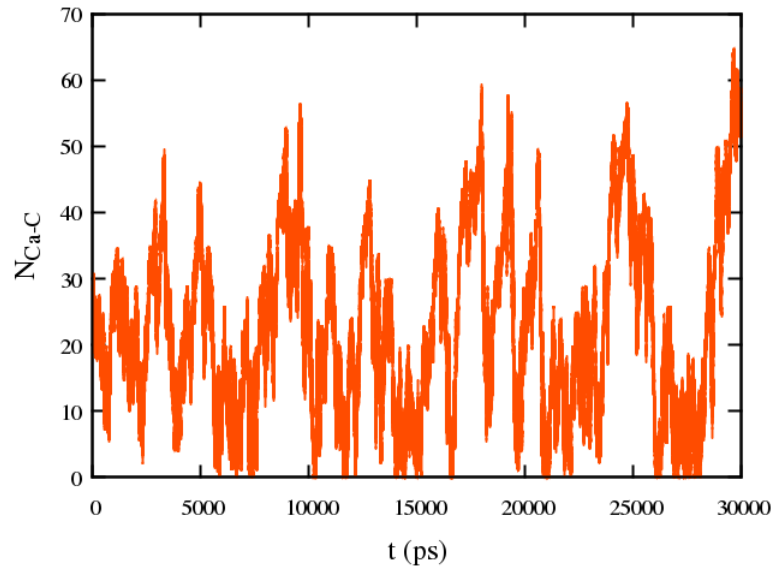
(b)

Figure D.1: N_{Ca-C} and R_g collective variables measured as a function of time for a metadynamics calculation of a cluster with composition $15\,CaCO_3$ in water.

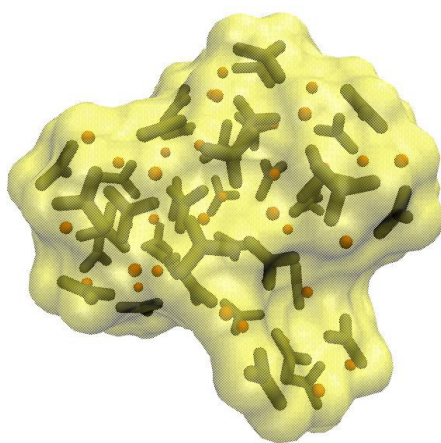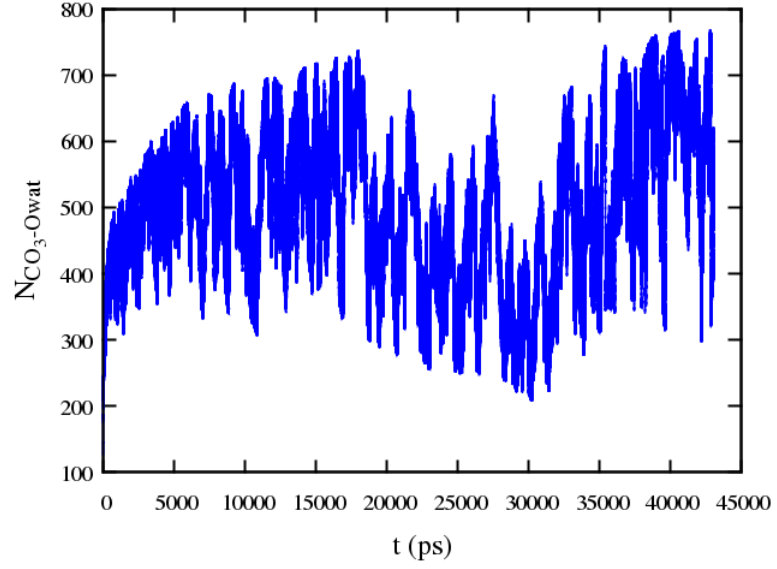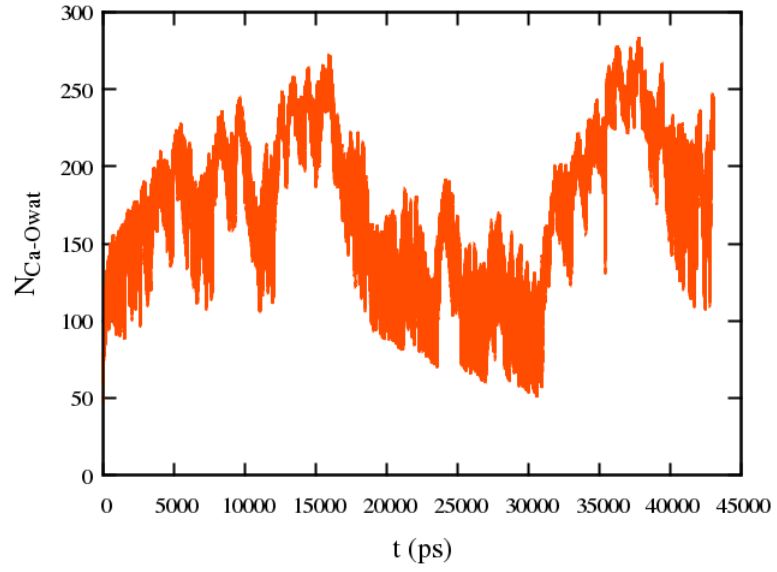Figure D.2: Initial configuration for a high density cluster of $40\,\mathrm{CaCO_3}$. Black molecules are carbonate while calcium is shown in orange. The ionic surface of the cluster is shown in yellow which was calculated with a 1σ distance from atom centres.

(a)

(b)

Figure D.3: N_{CO_3-Owat} and $N_{Ca-Owat}$ collective variables measured as a function of time for a metadynamics calculation of a high density cluster with composition $40 \, CaCO_3$ in water.
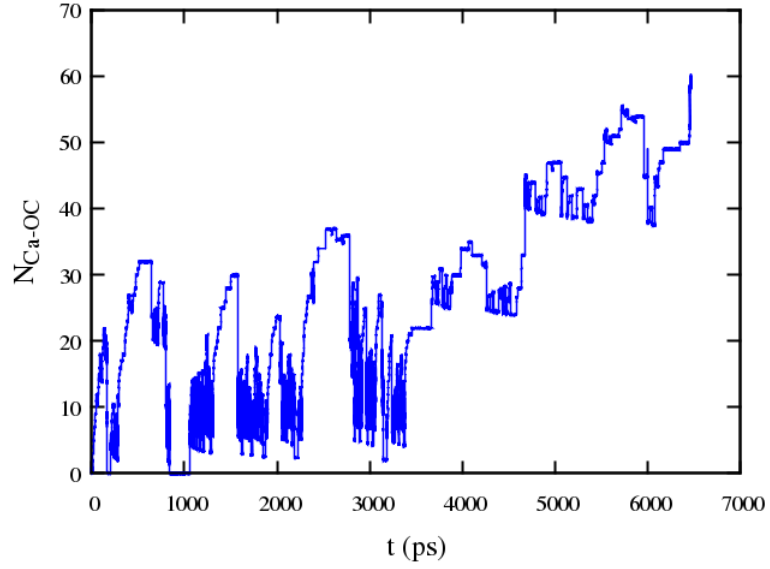
Figure D.4: N_{Ca-OC} collective variable measured as a function of time for a meta-dynamics calculation of a cluster with composition $10\,CaCO_3$ in water.
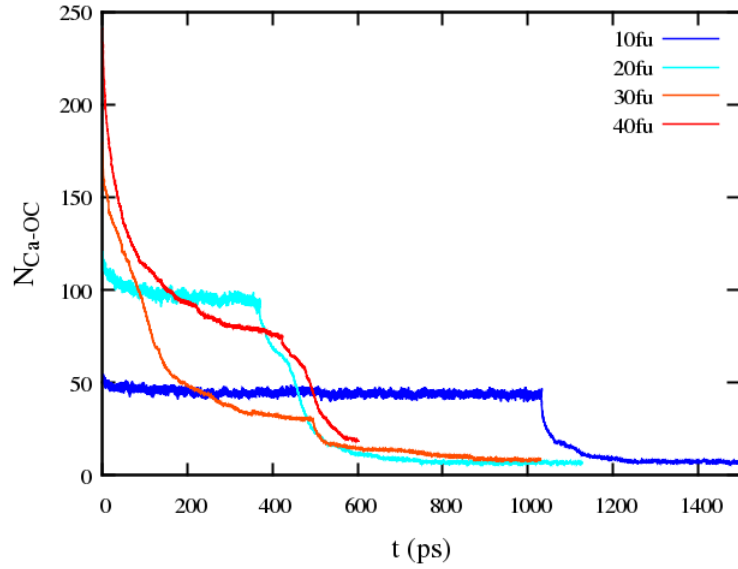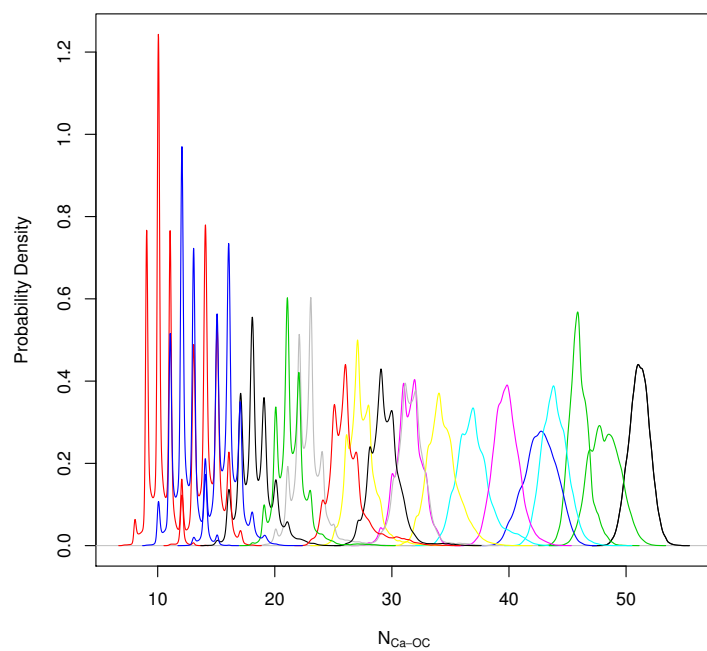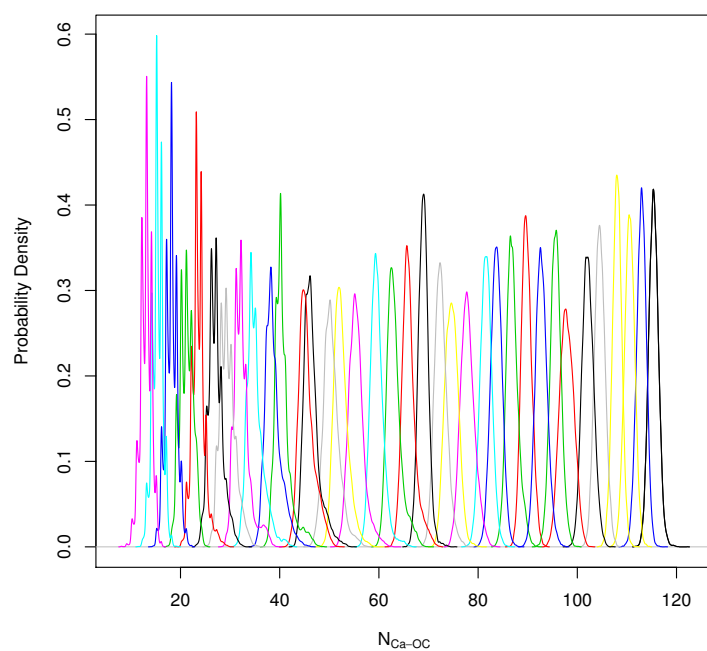

Figure D.5: N_{Ca-OC} as a function of time from adiabatic biased molecular dynamics simulations of high density 10, 20, 30 and 40 formula unit clusters in water, where the target for the order parameter was $N_{Ca-OC} = 0$.
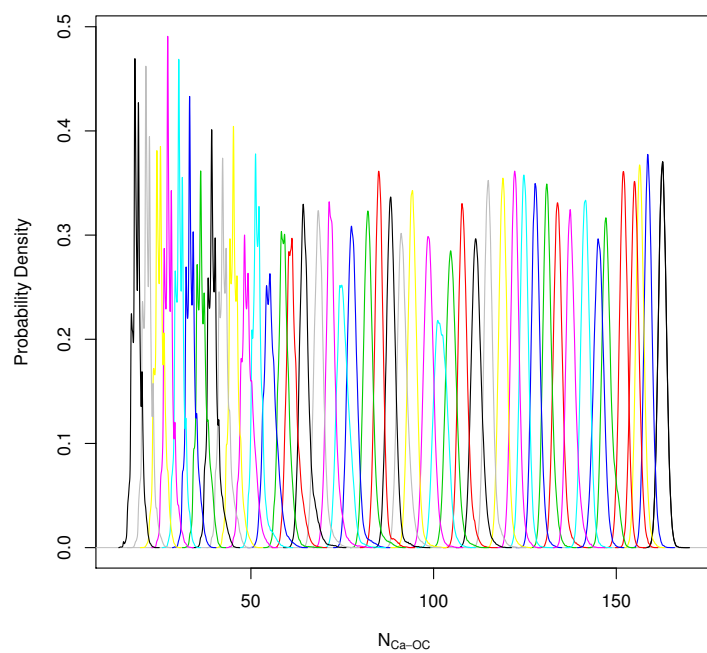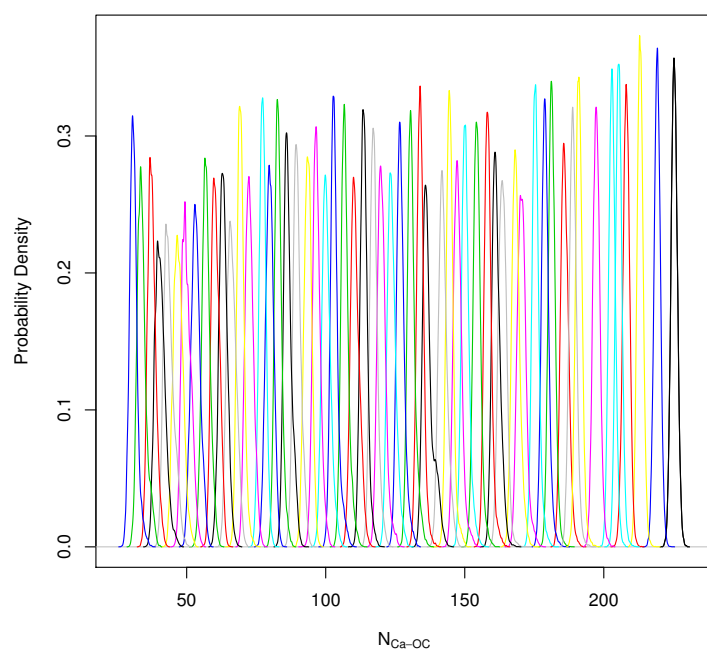
(a)

(b)

Figure D.6: Probability densities of N_{Ca-OC} from Umbrella Sampling simulations of 10 and 20 $CaCO_3$ in water.

(a)

(b)

Figure D.7: Probability densities of N_{Ca-OC} from Umbrella Sampling simulations of 30 and 40 $CaCO_3$ in water.

195
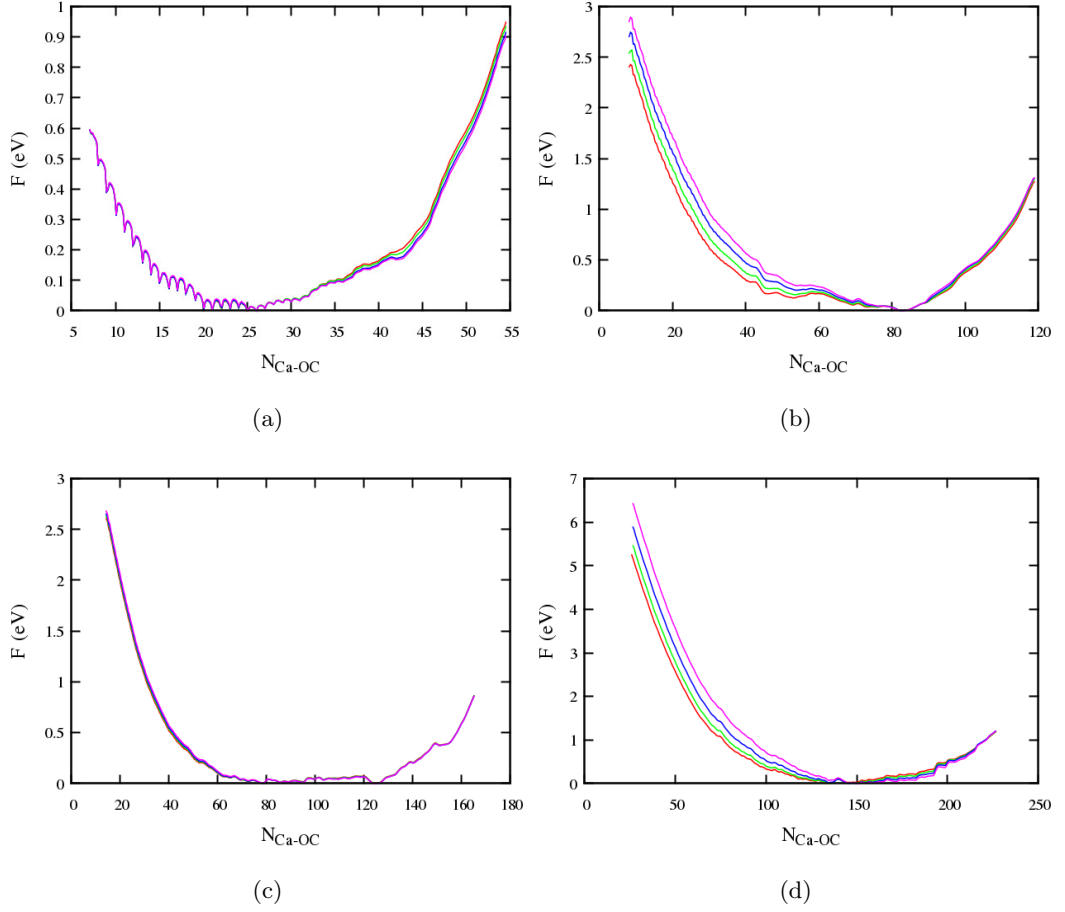
Figure D.8: Free energies for (a) 10, (b) 20, (c) 30 and (d) 40 formula unit clusters in water from Umbrella Sampling simulations (in which N_{Ca-OC} was constrained) and post-processing with WHAM. Red, green, blue and magenta curves show the free energy at successive 0.25 ns simulation times after approximately 6, 2, 14 and 3 ns for 10, 20, 30 and 40 formula unit simulations, respectively.

Figure D.9: Q_0^{Ca-OC} as a function of time during simulations of (a) 10, (b) 20, (c) 30 and (d) 40 calcium carbonate formula units in water where Q_0^{Ca-OC} and $Q_0^{Ca-O_{wat}}$ were biased. Q_0^{Ca-OC} has been scaled by N_S to give the equivalent coordination numbers.
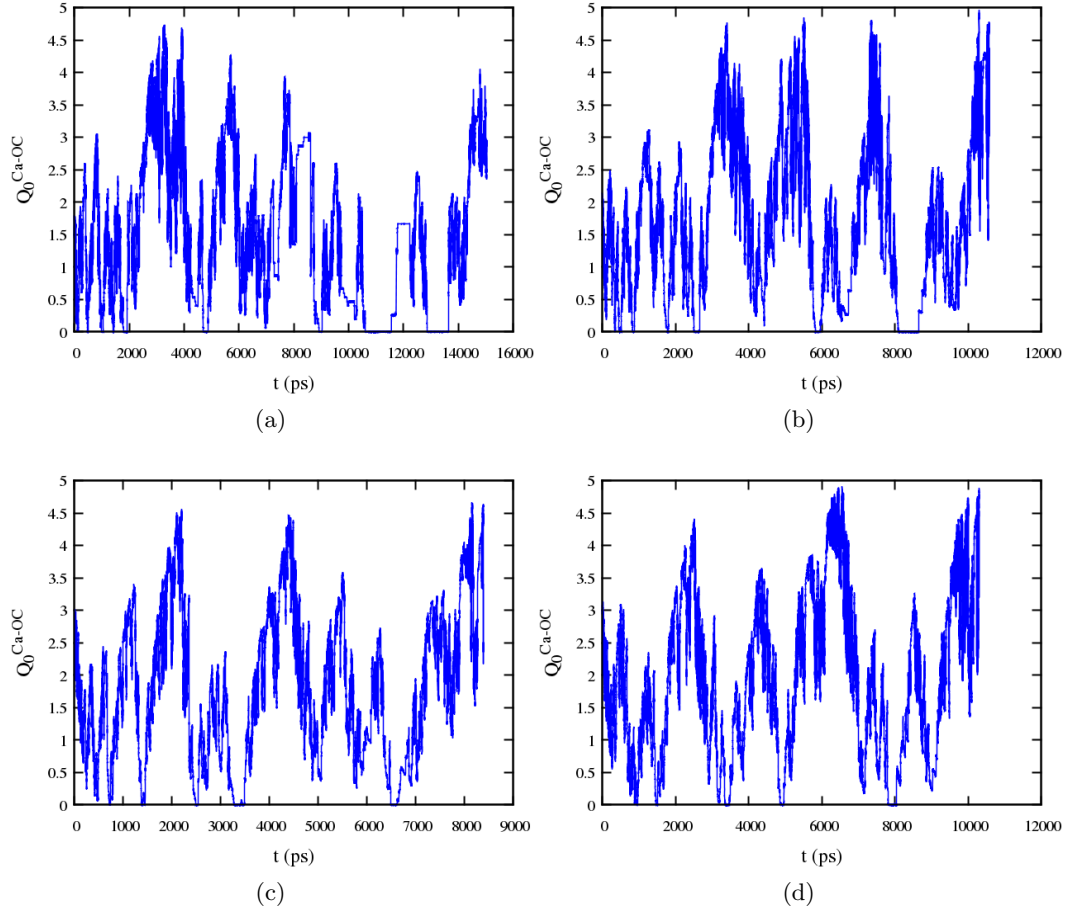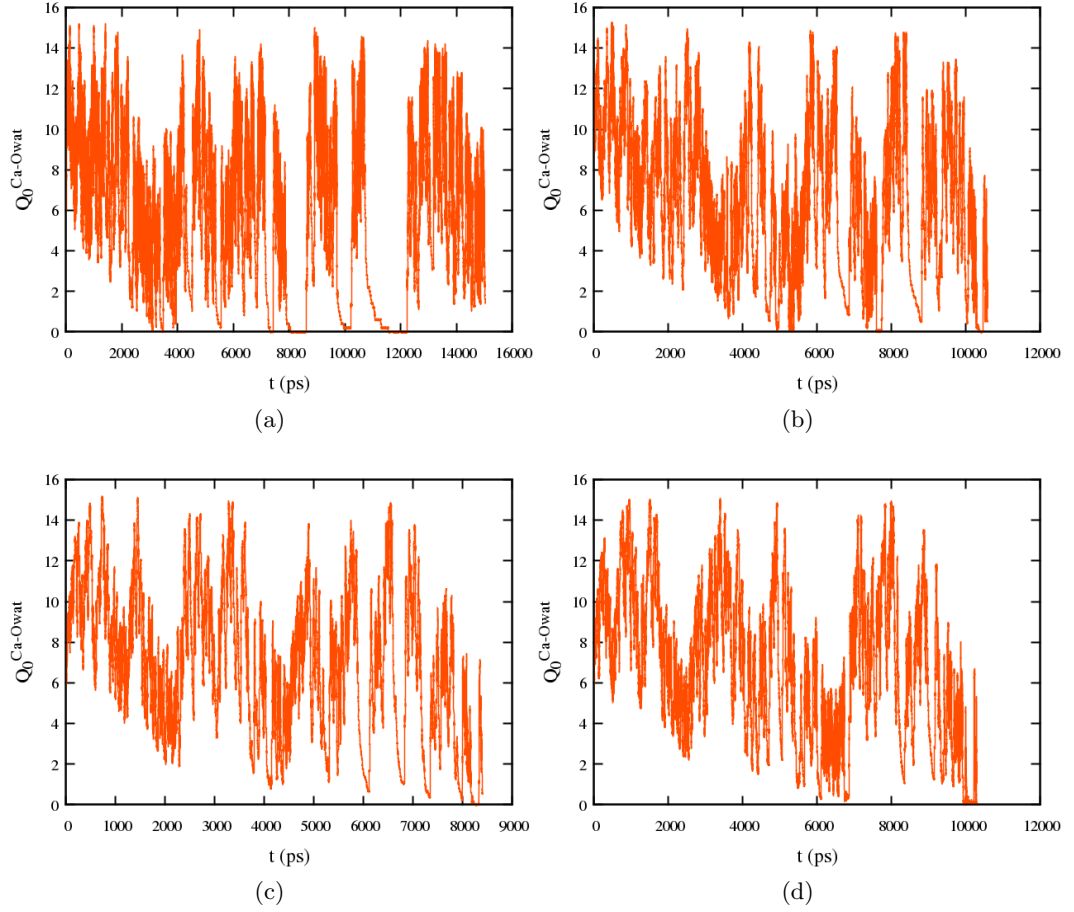
Figure D.10: $Q_0^{Ca-O_{wat}}$ as a function of time during simulations of (a) 10, (b) 20, (c) 30 and (d) 40 calcium carbonate formula units in water where Q_0^{Ca-OC} and $Q_0^{Ca-O_{wat}}$ were biased. $Q_0^{Ca-O_{wat}}$ has been scaled by N_S to give the equivalent coordination numbers.

Appendix E

Solution and Speciation Effects
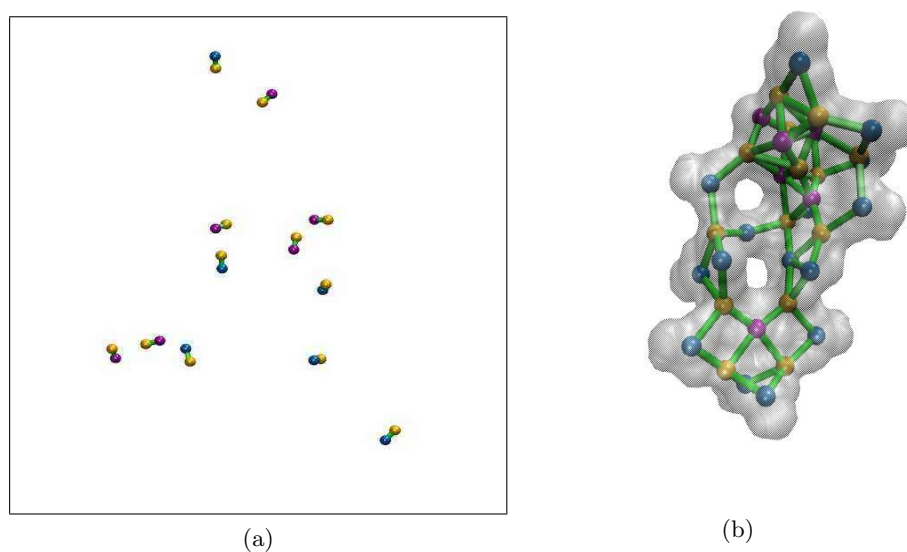
Figure E.1: Snapshots of (a) the final step of a 15 ns trajectory of system I_{22-14} (free ions and water are omitted for clarity), and (b) the initial cluster configuration for system C_{22-14}. Calcium is shown in yellow while carbonate and bicarbonate ion centres are shown in purple and blue. Green lines show coordinating ions where the distance between calcium and carbon is < 3.825 Å. Note that in (b), carbonate has a higher coordination than bicarbonate.
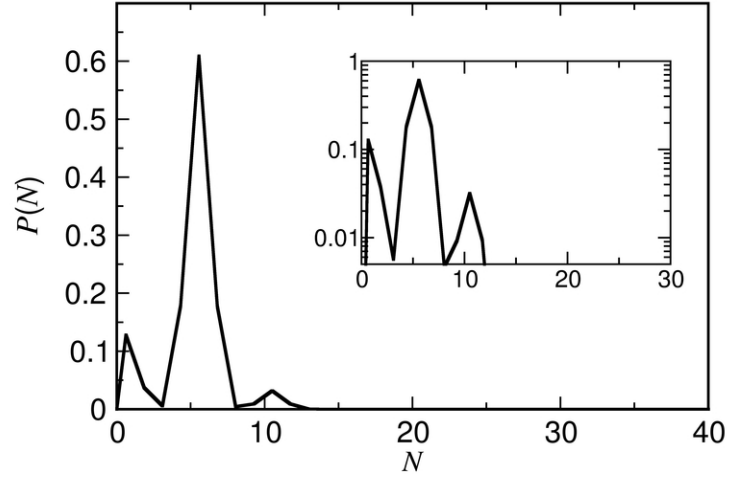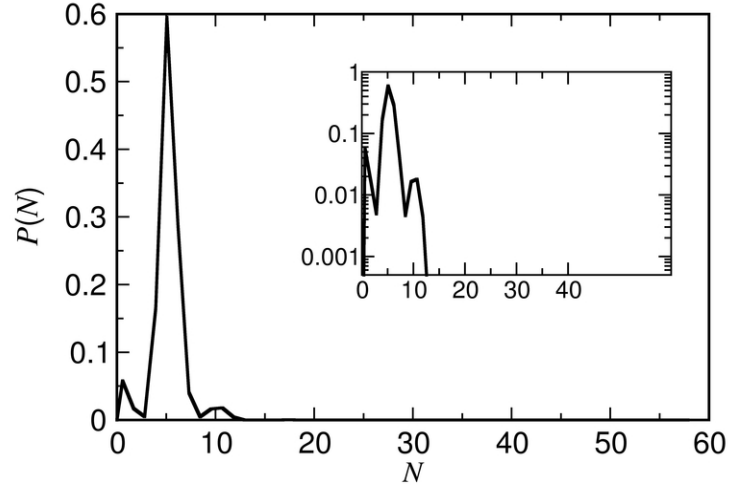
Figure E.2: Cluster size distributions for systems (a) I_{46-29}, and (b) I_{64-47} after equilibration in water. Data was taken from the final 5 ns of a 15 ns simulation. Inset are the probability distributions plotted with a logarithmic y scale.
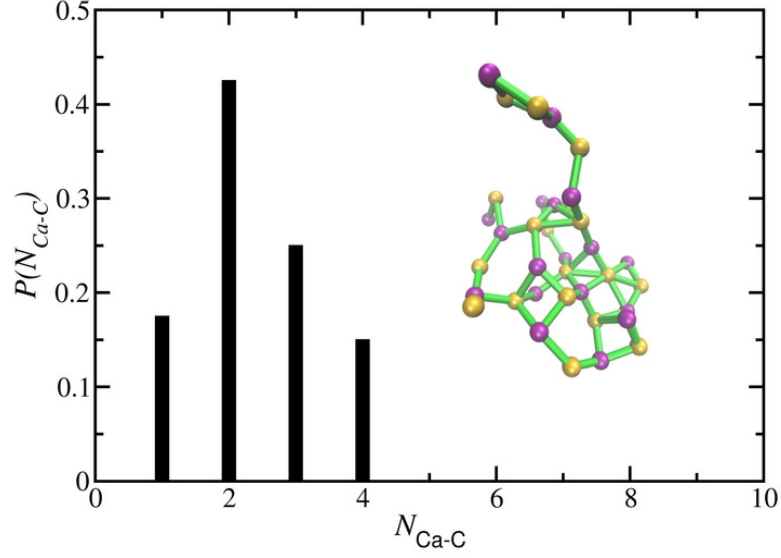
Figure E.3: Coordination probabilities for calcium binding to N carbon atoms at the beginning of simulations at 20–51 mM. Inset is a snapshot of a neutral cluster containing 36 ions which was present in solution, where calcium and carbon are shown as yellow, and magenta, respectively (oxygen and hydrogen are omitted for clarity). Green lines show distances between ions where $r < 3.825$ Å.

(a)	(b)
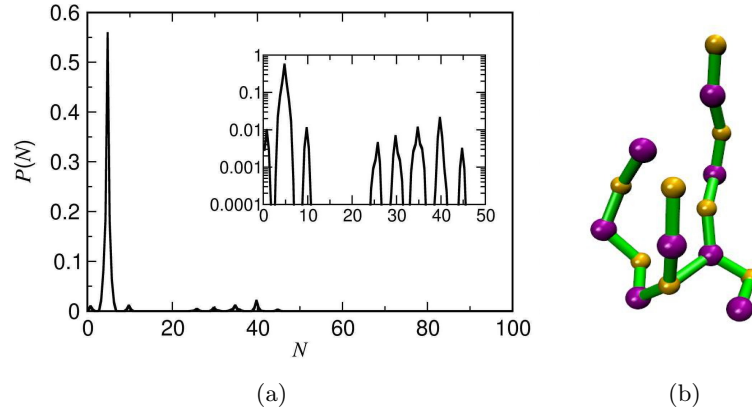
Figure E.4: (a) Cluster size ditribution, and (b) a cluster configuration from the final 2 ns of simulation of system C_{20-20} at [C]=51 mM. Inset of (a) is a logarithmic scale plot of the probability distribution. For a description of the representations in (b) see Figure E.3.

202

Appendix F

Postnucleation Phases and Dehydration of Amorphous Calcium Carbonate

Calculating Atomistic Spatial Correlation Functions

Time averaged spatial atomic distributions were measured, from which standard correlation functions, such as total radial distribution function, $G^x(r)$, and total scattering structure factor, $S(q)$ (which are related through the differential correlation function, $D(r)$) can be calculated. For a full description see Keen [Keen, 2001]. In the case of calculated correlation functions from atomistic simulation, atomic coordinates are used instead of electronic densities and truncated functions are used. The functional forms of $G^x(r)$, $S(q) - 1$ and $D(r)$ are,

$$G^x(r) = \sum_{i,j=1}^{n} c_i c_j \frac{Z_i Z_j}{(\sum_{k=1}^{n} c_k Z_k)^2} \left[g_{ij}(r) - 1 \right] \tag{F.1}$$

$$S(q) - 1 = 4\pi\rho_0 \int_0^{r_{max}} r^2 G^x(r) \frac{sin(qr)}{qr} \, dr \tag{F.2}$$

$$D(r) = 4\pi\rho_0 r G^x(r) = \frac{2}{\pi} \int_0^{q_{max}} q \left[S(q) - 1 \right] sin(qr) \, dq \tag{F.3}$$

In the above equations: Z_x and c_x are the atomic number and proportion of atoms, x, in the simulated material; $g_{ij}(r)$ is the radial distribution function for atoms i and j; ρ_0 is the average number density of atoms in the material; r_{max} and q_{max} are the truncation distances for correlation functions.

Table F.1: Diffusion coefficients for water occluded in cavities, H_2O_{Occ}, and channels, H_2O_{Chan}, in hydrated ACC with composition $CaCO_3 n H_2O$.

n	H_2O_{Occ} $(10^{-13}$ m^2 s$^{-1})$	H_2O_{Chan} $(10^{-13}$ m^2 s$^{-1})$
0.64	1.21 ± 0.02	2.98 ± 0.05
1	4.98 ± 0.05	5.39 ± 0.05
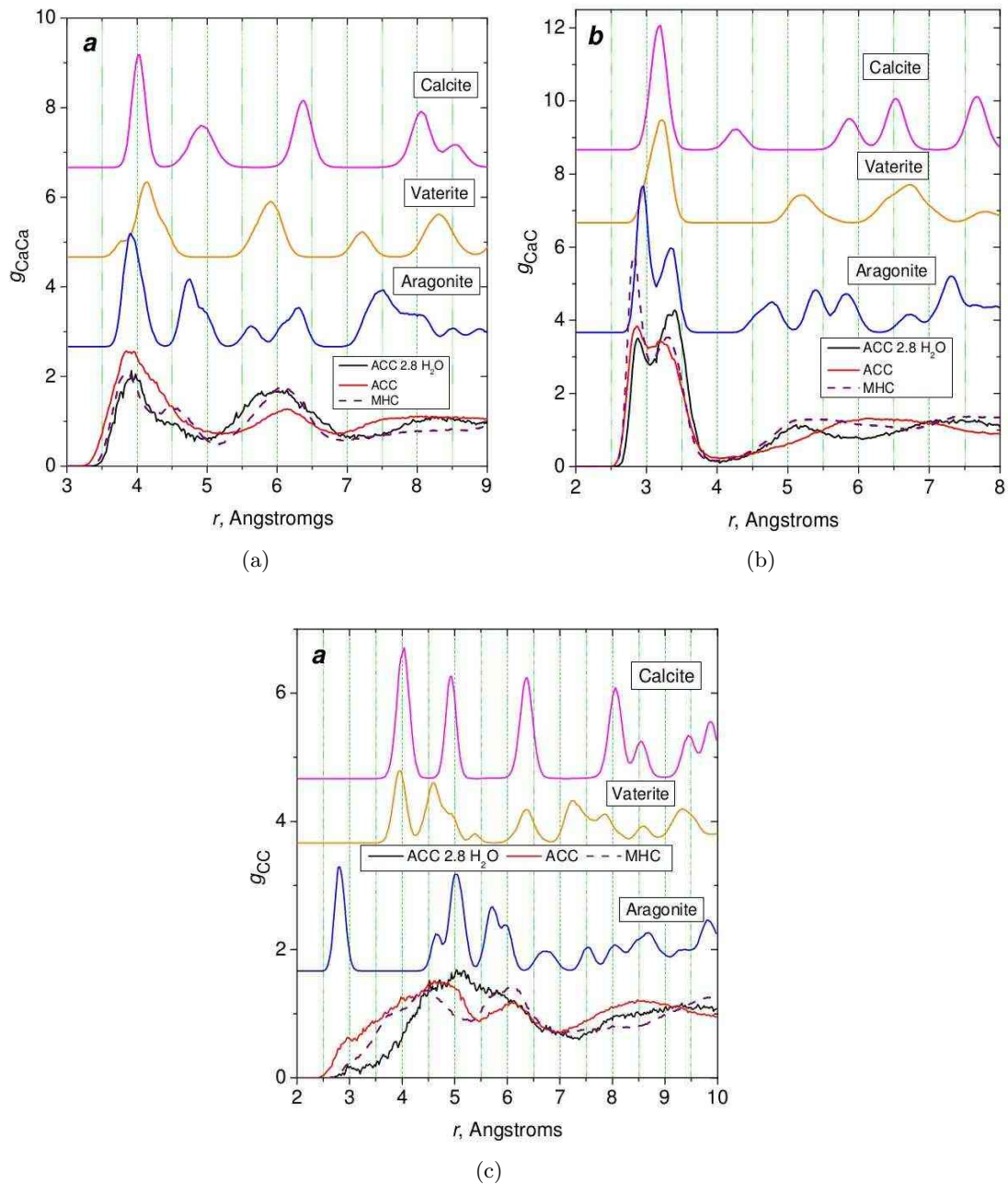
Figure F.1: Radial distribution functions for Ca–Ca (a), Ca–C (b) and C–C (c) measured from MD simulations of nanoparticles in water. Data was calculated as time averages from final portions of MD trajectories after relaxation of particles in water. Functions for anhydrous polymorphs have been shifted on the y-axis for clarity. Printed with permission of Dr. Yuriy G. Bushuev, *University of Warwick.*
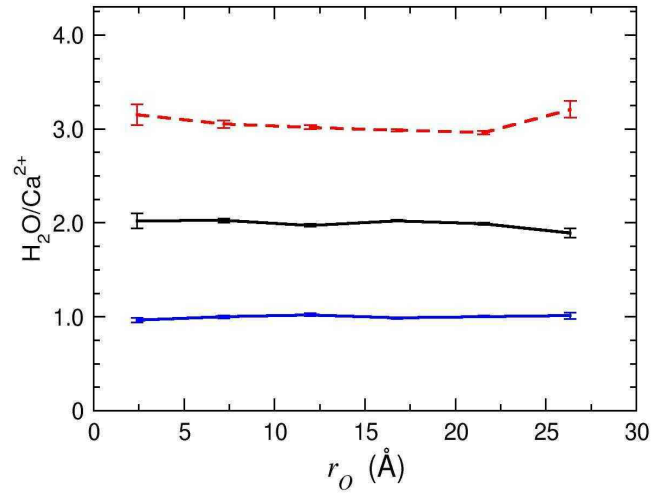
Figure F.2: Average number of water molecules per calcium ion in bulk ACC with composition $CaCO_3 \cdot nH_2O$, where $n = 1$, 2 and 3. Averages were taken from concentric spheres from the simulation cell origin, and during the final 200 ps of 5 ns simulations. Error bars show uncertainties of one standard deviation.
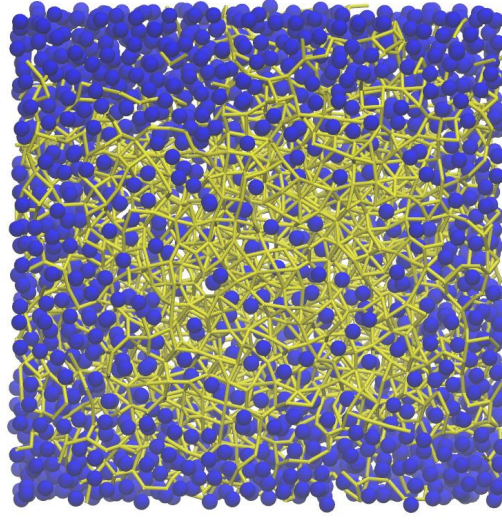
(a)
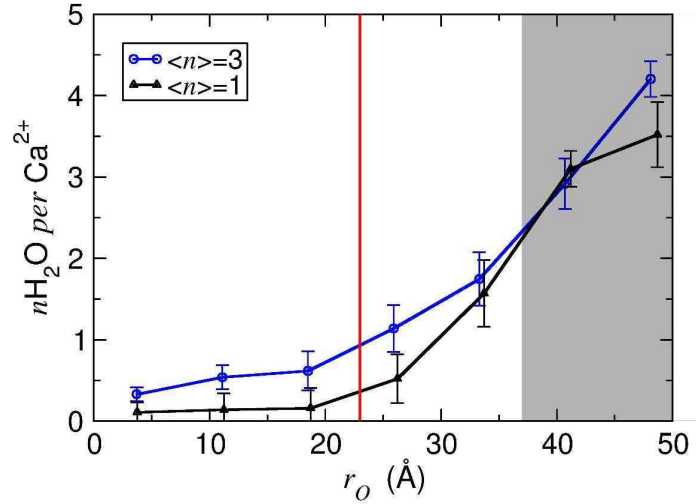
(b)

Figure F.3: (a) A snapshot of a 10 Å slice through the centre of the z axis of a simulation of mesoporous ACC with average composition $CaCO \cdot 3\,H_2O$. Calcium and carbon atoms are connected by yellow bonds if $r_{Ca-C} < 3.825$ Å, and water oxygens are shown as blue. (b) Average number of water molecules per calcium ion in mesoporous ACC with total compositions $n = 1$ and 3 for $CaCO_3 \cdot nH_2O$. Averages were taken from concentric spheres from the simulation cell origin, and during the final 200 ps of 5 ns simulations. Error bars show uncertainties of one standard deviation. The red line indicates the outer extremity of the core of the systems, where the average $H_2O{:}Ca^{2+}$ remains relatively constant, while the grey region highlights the diffuse ionic surface.
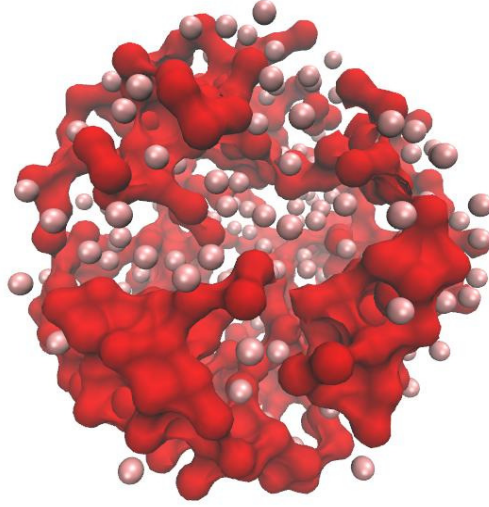
Figure F.4: The core of a mesoporous ACC system with composition $CaCO_3 \cdot (0.64)H_2O$ showing only water. Water occluded in the ionic framework is shown in pink, while water forming clusters (where coordinated water molecules are considered to be less than 3.5 Å apart) are shown in red.
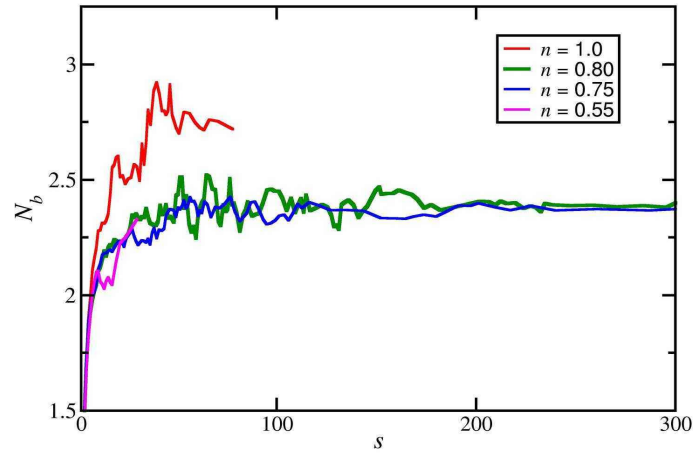
Figure F.5: Number of bonds, N_b, between water molecules in bulk ACC as a function of cluster size, s, for a range of hydration levels, n, in $CaCO_3 \cdot nH_2O$. A bond was defined by a distance criteria of $r_{O-O} < 3.8$ Å, and data was measured from the final 0.5-1.0 ns of trajectories after equilibration.

209
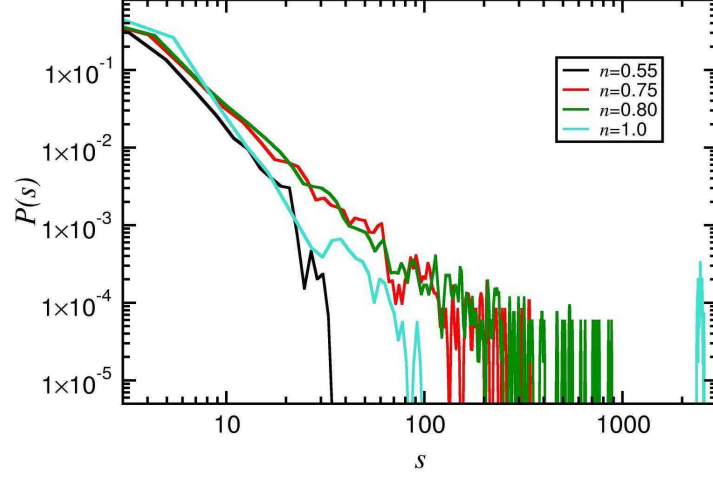
Figure F.6: Cluster size probability distributions for water clusters of size s in bulk hydrated ACC with composition $CaCO_3 \cdot n H_2O$. A logarithmic y scale is shown for clarity of peaks at large s.
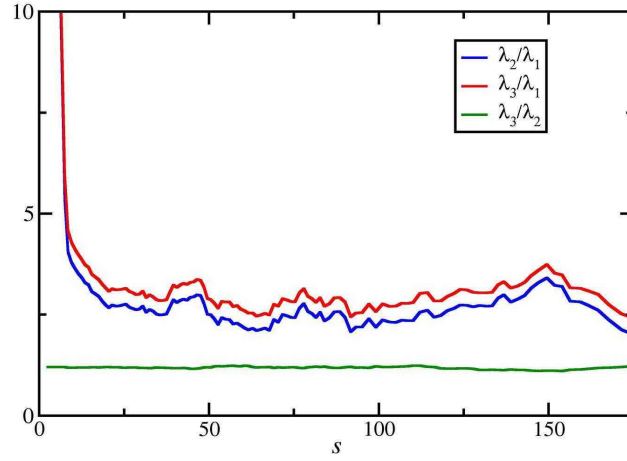
Figure F.7: Ratios of eigenvalues of the moment of inertia tensor, for finite water clusters of size s in $CaCO_3 \cdot (0.8)H_2O$. The largest eigenvalue is λ_3 and the smallest is λ_1.
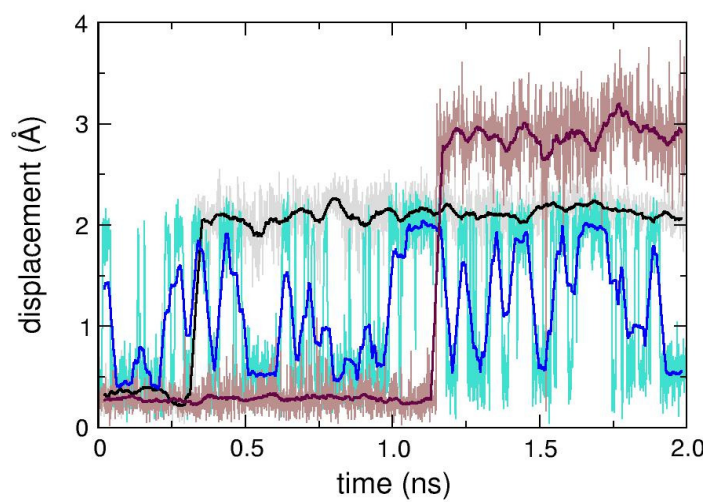
Figure F.8: Displacement of selected water oxygen atoms in the core of mesoporous ACC with composition $CaCO_3 \cdot (0.64)H_2O$. Occasional large displacement (2–3 Å) was observed in very short times ($\sim 10^2$ fs). Raw data is shown in light colours with a running average (taken over 20 ps) overlaid in dark colours.

Bibliography

L. Addadi and S. Weiner. Biomineralization: A pavement of pearl. *Nature*, 389 (6654):912–915, 1997.

L. Addadi, S. Raz, and S. Weiner. Taking advantage of disorder: Amorphous calcium carbonate and its roles in biomineralization. *Advanced Materials*, 15(12):959–970, 2003.

J. Aizenberg, L. Addadi, S. Weiner, and G. Lambert. Stabilization of amorphous calcium carbonate by specialized macromolecules in biological and synthetic precipitates. *Advanced Materials*, 8(3):222–226, 1996.

J. Aizenberg, A. J. Black, and G. M. Whitesides. Control of crystal nucleation by patterned self-assembled monolayers. *Nature*, 398(6727):495–498, 1999a.

J. Aizenberg, A. J. Black, and G. M. Whitesides. Oriented growth of calcite controlled by self-assembled monolayers of functionalized alkanethiols supported on gold and silver. *Journal of the American Chemical Society*, 121(18):4500–4509, 1999b.

J. Aizenberg, G. Lambert, S. Weiner, and L. Addadi. Factors involved in the formation of amorphous and crystalline calcium carbonate: a study of an ascidian skeleton. *Journal of the American Chemical Society*, 124(1):32–39, 2002.

A. Aksimentiev and K. Schulten. Imaging -hemolysin with molecular dynamics: Ionic conductance, osmotic permeability, and the electrostatic potential map. *Biophysical Journal*, 88(6):3745–3761, 2005.

J. N. Albright. Vaterite stability. *The American Mineralogist*, 56:620–624, 1971.

M. P. Allen and D. J. Tildesley. *Computer simulation of liquids.* Clarendon Press ; Oxford University Press, 1989. ISBN 0198556454 9780198556459 0198553757 9780198553755.

T. F. Anderson. Self-diffusion of carbon and oxygen in calcite by isotope exchange with carbon dioxide. *Journal of Geophysical Research*, 74(15):3918–3932, 1969.

S. M. Antao and I. Hassan. The orthorhombic structure of $CaCO_3$, $SrCO_3$, $PbCO_3$ and $BaCO_3$: Linear structural trends. *The Canadian Mineralogist*, 47(5):1245–1255, 2009.

S. M. Antao and I. Hassan. Temperature dependence of the structural parameters in the transformation of aragonite to calcite, as determined from in situ synchrotron powder x-ray diffraction data. *The Canadian Mineralogist*, 48(5):1225–1236, 2010.

J. Anwar and D. Zahn. Uncovering molecular processes in crystal nucleation and growth by using molecular simulation. *Angewandte Chemie International Edition*, 50(9):1996–2013, 2011.

B. Aziz, D. Gebauer, and N. Hedin. Kinetic control of particle-mediated calcium carbonate crystallization. *CrystEngComm*, 13(14):4641, 2011.

A. M. Bano. Polymorphism in biomineral nanoparticles, 2012.

S. Bates, R. C. Kelly, I. Ivanisevic, P. Schields, G. Zografi, and A. W. Newman. Assessment of defects and amorphous structure produced in raffinose pentahydrate upon dehydration. *Journal of Pharmaceutical Sciences*, 96(5):1418–1433, 2007.

J. Baumgartner, A. Dey, P. H. H. Bomans, C. Le Coadou, P. Fratzl, N. A. J. M. Sommerdijk, and D. Faivre. Nucleation and growth of magnetite from solution. *Nature Materials*, 12(4):310–314, 2013.

R. Becker and W. Döring. Kinetische behandlung der keimbildung in übersättigten dämpfen. *Annalen der Physik*, 416(8):719–752, 1935.

E. Beniash, J. Aizenberg, L. Addadi, and S. Weiner. Amorphous calcium carbonate transforms into calcite during sea urchin larval spicule growth. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 264(1380):461–465, 1997.

M. A. Bewernitz, D. Gebauer, J. Long, H. Coelfen, and L. B. Gower. A metastable liquid precursor phase of calcium carbonate and its interactions with polyaspartate. *Faraday Discussions*, 159:291–312, 2012.

J. Bolze, B. Peng, N. Dingenouts, P. Panine, T. Narayanan, and M. Ballauff. Formation and growth of amorphous colloidal $CaCO_3$ precursor particles as detected by time-resolved SAXS. *Langmuir*, 18(22):8364–8369, 2002.

M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, and M. Parrinello. PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications*, 180(10):1961–1972, 2009.

J. T. Bosko, B. D. Todd, and R. J. Sadus. Analysis of the shape of dendrimers under shear. *The Journal of Chemical Physics*, 124(4):044910, 2006.

P. Bots, L. G. Benning, J.-D. Rodriguez-Blanco, T. Roncal-Herrero, and S. Shaw. Mechanistic insights into the crystallization of amorphous calcium carbonate (ACC). *Crystal Growth and Design*, 12(7):3806–3814, 2012.

J. B. Brady. Diffusion data for silicate minerals, glasses, and liquids. In T. J. Ahrens, editor, *Mineral Physics and Crystallography: A Handbook of Physical Constants*, pages 269–290. American Geophysical Union, 1995. ISBN 9781118668191.

B. Cantaert, Y.-Y. Kim, H. Ludwig, F. Nudelman, N. A. J. M. Sommerdijk, and F. C. Meldrum. Think positive: Phase separation enables a positively charged additive to induce dramatic changes in calcium carbonate morphology. *Advanced Functional Materials*, 22(5):907–915, 2012.

J. H. E. Cartwright, A. G. Checa, J. D. Gale, D. Gebauer, and C. I. Sainz-Daz. Calcium carbonate polyamorphism and its role in biomineralization: How many amorphous calcium carbonates are there? *Angewandte Chemie International Edition*, 51(48):11960–11970, 2012.

Y. C. Chien, M. T. Hincke, H. Vali, and M. D. McKee. Ultrastructural matrixmineral relationships in avian eggshell, and effects of osteopontin on calcite growth in vitro. *Journal of Structural Biology*, 163(1):84–99, 2008.

H. Cölfen. Precipitation of carbonates: recent progress in controlled production of complex shapes. *Current Opinion in Colloid and Interface Science*, 8(1):23–31, 2003.

D. J. Cooke and J. A. Elliott. Atomistic simulations of calcite nanoparticles and their interaction with water. *The Journal of Chemical Physics*, 127(10):104706, 2007.

J.-P. Cuif and Y. Dauphin. Microstructural and physico-chemical characterization of centers of calcification in septa of some recent scleractinian corals. *Paläontologische Zeitschrift*, 72(3):257–269, 1998.

J. De Yoreo. Crystal nucleation: More than one pathway. *Nature Materials*, 12(4): 284–285, 2013.

R. Demichelis, P. Raiteri, J. D. Gale, D. Quigley, and D. Gebauer. Stable prenucleation mineral clusters are liquid-like ionic polymers. *Nature Communications*, 2: 590, 2011.

R. Demichelis, P. Raiteri, and J. D. Gale. Structure of hydrated calcium carbonates: A first-principles study. *Journal of Crystal Growth*, 401:33–37, 2014.

J. L. Desmond, P. M. Rodger, and T. R. Walsh. Testing the inter-operability of the CHARMM and SPC/Fw force fields for conformational sampling. *Molecular Simulation*, 0(0):1–10, 2013.

J. W. Dieckert, M. C. Dieckert, and C. R. Creger. Calcium reserve assembly: a basic structural unit of the calcium reserve system of the hen egg shell. *Poultry Science*, 68(11):1569–1584, 1989.

E. DiMasi, V. M. Patel, M. Sivakumar, M. J. Olszta, Y. P. Yang, and L. B. Gower. Polymer-controlled growth rate of an amorphous mineral film nucleated at a fatty acid monolayer. *Langmuir*, 18(23):8902–8909, 2002.

M. T. Dove, B. Winkler, M. Leslie, M. J. Harris, and E. K. H. Salje. A new interatomic potential model for calcite : applications to lattice dynamics studies, phase transition, and isotope fractionation. *The American mineralogist*, 77(3): 244–250, 1992.

U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen. A smooth particle mesh ewald method. *The Journal of Chemical Physics*, 103(19): 8577–8593, 1995.

J. S. Evans. "liquid-like" biomineralization protein assemblies: a key to the regulation of non-classical nucleation. *Crystengcomm*, 15(42):8388–8394, 2013.

M. Faatz, F. Gröhn, and G. Wegner. Amorphous calcium carbonate: Synthesis and potential intermediate in biomineralization. *Advanced Materials*, 16(12):996–1000, 2004.

G. Falini, M. Reggi, S. Fermani, F. Sparla, S. Goffredo, Z. Dubinsky, O. Levi, Y. Dauphin, and J.-P. Cuif. Control of aragonite deposition in colonial corals by intra-skeletal macromolecules. *Journal of Structural Biology*, 183(2):226–238, 2013.

P. Fenter, S. Kerisit, P. Raiteri, and J. D. Gale. Is the calcitewater interface understood? direct comparisons of molecular dynamics simulations with specular x-ray reflectivity data. *The Journal of Physical Chemistry C*, 117(10):5028–5042, 2013.

A. Fernández-González, L. Fernández-Seivane, M. P. Rubio, and J. Ferrer. A molecular study of CaCO$_3$ cluster configurations. *arXiv:1105.5371 [cond-mat, physics:physics]*, 2011.

A. Fernandez-Martinez, B. Kalkan, S. M. Clark, and G. A. Waychunas. Pressure-induced polyamorphism and formation of aragonitic amorphous calcium carbonate. *Angewandte Chemie International Edition*, 52(32):8354–8357, 2013.

D. K. Fisler and R. T. Cygan. Diffusion of Ca and Mg in calcite. *American Mineralogist*, 84(9):1392–1399, 1999.

D. A. Fletcher, R. F. McMeeking, and D. Parkin. The united kingdom chemical database service. *Journal of Chemical Information and Computer Sciences*, 36 (4):746–749, 1996.

C. L. Freeman, J. H. Harding, D. J. Cooke, J. A. Elliott, J. S. Lardge, and D. M. Duffy. New forcefields for modeling biomineralization processes. *The Journal of Physical Chemistry C*, 111(32):11943–11951, 2007.

C. L. Freeman, J. H. Harding, and D. M. Duffy. Simulations of calcite crystallization on self-assembled monolayers. *Langmuir*, 24(17):9607–9615, 2008.

C. Freeman, J. Harding, D. Quigley, and P. Rodger. Structural control of crystal nuclei by an eggshell protein. *Angewandte Chemie International Edition*, 49(30): 5135–5137, 2010.

D. Frenkel and B. Smit. *Understanding Molecular Simulation, Second Edition: From Algorithms to Applications*. Academic Press, 2 edition edition, 2001. ISBN 9780122673511.

M. Fricke and D. Volkmer. Crystallization of calcium carbonate beneath insoluble monolayers: Suitable models of mineralmatrix interactions in biomineralization? In K. Naka, editor, *Biomineralization I*, number 270 in Topics in Current Chemistry, pages 1–41. Springer Berlin Heidelberg, 2007. ISBN 978-3-540-46379-5, 978-3-540-46380-1.

D. Gebauer and H. Cölfen. Prenucleation clusters and non-classical nucleation. *Nano Today*, 6(6):564–584, 2011.

D. Gebauer, A. Völkel, and H. Cölfen. Stable prenucleation calcium carbonate clusters. *Science*, 322(5909):1819–1822, 2008.

D. Gebauer, H. Cölfen, A. Verch, and M. Antonietti. The multiple roles of additives in CaCO$_3$ crystallization: A quantitative case study. *Advanced Materials*, 21(4): 435–439, 2009.

D. Gebauer, P. N. Gunawidjaja, J. Y. P. Ko, Z. Bacsik, B. Aziz, L. Liu, Y. Hu, L. Bergström, C.-W. Tai, T.-K. Sham, M. Edn, and N. Hedin. Proto-calcite and proto-vaterite in amorphous calcium carbonates. *Angewandte Chemie International Edition*, 49(47):8889–8891, 2010.

D. Gebauer, M. Kellermeier, J. D. Gale, L. Bergstrom, and H. Cölfen. Pre-nucleation clusters as solute precursors in crystallisation. *Chemical Society Reviews*, 43(7): 2348–2371, 2014.

Y. U. T. Gong, C. E. Killian, I. C. Olson, N. P. Appathurai, A. L. Amasino, M. C. Martin, L. J. Holt, F. H. Wilt, and P. U. P. A. Gilbert. Phase transitions in biogenic amorphous calcium carbonate. *Proceedings of the National Academy of Sciences*, 109(16):6088–6093, 2012.

A. L. Goodwin, F. M. Michel, B. L. Phillips, D. A. Keen, M. T. Dove, and R. J. Reeder. Nanoporous structure and medium-range order in synthetic amorphous calcium carbonate. *Chemistry of Materials*, 22(10):3197–3205, 2010.

L. B. Gower. Biomimetic model systems for investigating the amorphous precursor pathway and its role in biomineralization. *Chemical Reviews*, 108(11):4551–4627, 2008.

L. B. Gower and D. J. Odom. Deposition of calcium carbonate films by a polymer-induced liquid-precursor (PILP) process. *Journal of Crystal Growth*, 210(4):719–734, 2000.

A. Grossfield. WHAM: the weighted histogram analysis method, version 2.0.9, 2014.

C. Günther, A. Becker, G. Wolf, and M. Epple. In vitro synthesis and structural characterization of amorphous calcium carbonate. *Zeitschrift für anorganische und allgemeine Chemie*, 631(13):2830–2835, 2005.

T. Y.-J. Han and J. Aizenberg. Calcium carbonate storage in amorphous form and its template-induced crystallization. *Chemistry of Materials*, 20(3):1064–1068, 2008.

Y.-J. Han and J. Aizenberg. Face-selective nucleation of calcite on self-assembled monolayers of alkanethiols: Effect of the parity of the alkyl chain. *Angewandte Chemie International Edition*, 42(31):3668–3670, 2003.

J. H. Harding, C. L. Freeman, and D. M. Duffy. Oriented crystal growth on organic monolayers. *CrystEngComm*, 16(8):1430, 2014.

J. H. Harding, D. M. Duffy, M. L. Sushko, P. M. Rodger, D. Quigley, and J. A. Elliott. Computational techniques at the organicinorganic interface in biomineralization. *Chemical Reviews*, 108(11):4823–4854, 2008.

B. Hasse, H. Ehrenberg, J. C. Marxen, W. Becker, and M. Epple. Calcium carbonate modifications in the mineralized shell of the freshwater snail biomphalaria glabrata. *Chemistry A European Journal*, 6(20):3679–3685, 2000.

R. Hentschke, T. Bast, E. Aydt, and M. Kotelyanskii. Gibbs-ensemble molecular dynamics: A new method for simulations involving particle exchange. *Molecular modeling annual*, 2(9):319–326, 1996.

K.-F. Hesse, H. Küppers, and E. Suess. Refinement of the structure of ikaite, $CaCO_3 6 H_2O$. *Zeitschrift für Kristallographie*, 163(3):227–231, 1983.

A. Hohl, T. Wieder, P. A. van Aken, T. E. Weirich, G. Denninger, M. Vidal, S. Oswald, C. Deneke, J. Mayer, and H. Fuess. An interface clusters mixture model for the structure of amorphous silicon monoxide (SiO). *Journal of Non-Crystalline Solids*, 320(1):255–280, 2003.

M. Holz, S. R. Heil, and A. Sacco. Temperature-dependent self-diffusion coefficients of water and six selected molecular liquids for calibration in accurate ^{1}H NMR PFG measurements. *Physical Chemistry Chemical Physics*, 2(20):4740–4742, 2000.

W. G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31(3):1695–1697, 1985.

H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, and T. Head-Gordon. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *The Journal of Chemical Physics*, 120(20): 9665–9678, 2004.

Q. Hu, M. H. Nielsen, C. L. Freeman, L. M. Hamm, J. Tao, J. R. I. Lee, T. Y. J. Han, U. Becker, J. H. Harding, P. M. Dove, and J. J. D. Yoreo. The thermodynamics

of calcite nucleation at organic interfaces: Classical vs. non-classical pathways. *Faraday Discussions*, 159(0):509–523, 2013.

J. Ihli, W. C. Wong, E. H. Noel, Y.-Y. Kim, A. N. Kulak, H. K. Christenson, M. J. Duer, and F. C. Meldrum. Dehydration and crystallization of amorphous calcium carbonate in solution and in air. *Nature Communications*, 5, 2014.

S. S. Jee, L. Culver, Y. Li, E. P. Douglas, and L. B. Gower. Biomimetic mineralization of collagen via an enzyme-aided PILP process. *Journal of Crystal Growth*, 312(8):1249–1256, 2010.

J. Jiang, G. Oberdörster, and P. Biswas. Characterization of size, surface charge, and agglomeration state of nanoparticle dispersions for toxicological studies. *Journal of Nanoparticle Research*, 11(1):77–89, 2009.

R. L. Johnston. Evolving better nanoparticles: Genetic algorithms for optimising cluster geometries. *Dalton Transactions*, (22):4193, 2003.

W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, 1983.

L. Kabalah-Amitai, B. Mayzel, Y. Kauffmann, A. N. Fitch, L. Bloch, P. U. P. A. Gilbert, and B. Pokroy. Vaterite crystals contain two interspersed crystal structures. *Science*, 340(6131):454–457, 2013.

S. R. Kamhi. On the structure of vaterite $CaCO_3$. *Acta Crystallographica*, 16(8): 770–772, 1963.

A. Kawska, J. Brickmann, R. Kniep, O. Hochrein, and D. Zahn. An atomistic simulation scheme for modeling crystal formation from solution. *The Journal of Chemical Physics*, 124(2):024513, 2006.

D. A. Keen. A comparison of various commonly used correlation functions for describing total scattering. *Journal of Applied Crystallography*, 34(2):172–177, 2001.

M. Kellermeier, D. Gebauer, E. Melero-Garcia, M. Drechsler, Y. Talmon, L. Kienle, H. Coelfen, J. Manuel Garcia-Ruiz, and W. Kunz. Colloidal stabilization of calcium carbonate prenucleation clusters with silica. *Advanced Functional Materials*, 22(20):4301–4311, 2012.

S. Kerisit and S. C. Parker. Free energy of adsorption of water and metal ions on the 1014 calcite surface. *Journal of the American Chemical Society*, 126(32): 10152–10161, 2004.

M. J. Kotelyanskii and R. Hentschke. Gibbs-ensemble molecular dynamics: Liquid-gas equilibria for lennard-jones spheres and n-hexane. *Molecular Simulation*, 17 (2):95–112, 1996.

S. Kumar, J. Rosenberg, D. Bouzida, R. Swendsen, and P. Kollman. The weighted histogram analysis method for free-energy calculations on biomolecules. i. the method. *J. Comput. Chem.*, 13(8):1011–1021, 1992.

N. Lahav and H. Bolt. Self–diffusion of ca^{45} into certain carbonates. : Soil science, 1964.

C.-L. Lai, K. E. Landgraf, G. A. Voth, and J. J. Falke. Membrane docking geometry and target lipid stoichiometry of membrane-bound PKC c2 domain: A combined molecular dynamics and experimental study. *Journal of Molecular Biology*, 402 (2):301–310, 2010.

A. Laio and M. Parrinello. Escaping free-energy minima. *Proceedings of the National Academy of Sciences*, 99(20):12562–12566, 2002.

R. Lakshminarayanan, E. O. Chi-Jin, X. J. Loh, R. M. Kini, and S. Valiyaveettil. Purification and characterization of a vaterite-inducing peptide, pelovaterin, from the eggshells of pelodiscus sinensis (chinese soft-shelled turtle). *Biomacromolecules*, 6(3):1429–1437, 2005.

R. S. K. Lam, J. M. Charnock, A. Lennie, and F. C. Meldrum. Synthesis-dependant structural variations in amorphous calcium carbonate. *CrystEngComm*, 9(12): 1226–1236, 2007.

N. H. d. Leeuw and S. C. Parker. Surfacewater interactions in the dolomite problem. *Physical Chemistry Chemical Physics*, 3(15):3217–3221, 2001.

A. R. Lennie, C. C. Tang, and S. P. Thompson. The structure and thermal expansion behaviour of ikaite, $CaCO_3 6 H_2O$, from T= 114 to T= 293 K. *Mineralogical Magazine*, 68(1):135–146, 2004.

Y. Levi-Kalisman, S. Raz, S. Weiner, L. Addadi, and I. Sagi. Structural differences between biogenic amorphous calcium carbonate phases using x-ray absorption spectroscopy. *Advanced Functional Materials*, 12(1):43, 2002.

Y. Levi-Kalisman, S. Raz, S. Weiner, L. Addadi, and I. Sagi. X-ray absorption spectroscopy studies on the structure of a biogenic amorphous calcium carbonate phase. *Journal of the Chemical Society, Dalton Transactions*, (21):3977–3982, 2000.

Y. Levi-Kalisman, G. Falini, L. Addadi, and S. Weiner. Structure of the nacreous organic matrix of a bivalve mollusk shell examined in the hydrated state using cryo-TEM. *Journal of Structural Biology*, 135(1):8–17, 2001.

J. Ma. A molecular dynamics study on the nucleation of calcium phosphate regulated by collagen. *Journal of Materials Science*, 49(8):3099–3106, 2014.

A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *Journal of Physical Chemistry B*, 102(18):3586–3616, 1998.

M. Marchi and P. Ballone. Adiabatic bias molecular dynamics: A method to navigate the conformational space of complex molecular systems. *The Journal of Chemical Physics*, 110(8):3697–3702, 1999.

G. Marland. The stability of $CaCO_3 6 H_2O$ (ikaite). *Geochimica et Cosmochimica Acta*, 39(1):83–91, 1975.

R. B. Martin and D. L. Boardman. The effects of collagen fiber orientation, porosity, density, and mineralization on bovine cortical bone bending properties. *Journal of Biomechanics*, 26(9):1047–1054, 1993.

R. B. Martin and J. Ishida. The relative effects of collagen fiber orientation, porosity, density, and mineralization on bone strength. *Journal of Biomechanics*, 22(5): 419–426, 1989.

J. C. Marxen, W. Becker, D. Finke, B. Hasse, and M. Epple. Early mineralization in biomphalaria glabrata: Microscopic and structural results. *Journal of Molluscan Studies*, 69(2):113–121, 2003.

M. Matsumoto, S. Saito, and I. Ohmine. Molecular dynamics simulation of the ice nucleation and growth process leading to water freezing. *Nature*, 416(6879): 409–413, 2002.

U. Mayor, N. R. Guydosh, C. M. Johnson, J. G. Grossmann, S. Sato, G. S. Jas, S. M. V. Freund, D. O. V. Alonso, V. Daggett, and A. R. Fersht. The complete folding pathway of a protein from nanoseconds to microseconds. *Nature*, 421 (6925):863–867, 2003.

D. A. McQuarrie. *Statistical Mechanics*. University Science Books, 2000. ISBN 9781891389153.

S. Melchionna, G. Ciccotti, and B. Lee Holian. Hoover NPT dynamics for systems varying in shape and size. *Molecular Physics*, 78(3):533–544, 1993.

F. C. Meldrum. Calcium carbonate in biomineralisation and biomimetic chemistry. *International Materials Reviews*, 48(3):187–224, 2003.

F. C. Meldrum and H. Colfen. Controlling mineral morphologies and structures in biological and synthetic systems. *Chemical Reviews*, 108(11):4332–4432, 2008.

H. J. Meyer. Struktur und fehlordnung des vaterits. *Zeitschrift für Kristallographie*, 128(3):183–212, 1969.

F. M. Michel, J. MacDonald, J. Feng, B. L. Phillips, L. Ehm, C. Tarabrella, J. B. Parise, and R. J. Reeder. Structural characteristics of synthetic amorphous calcium carbonate. *Chemistry of Materials*, 20(14):4720–4728, 2008.

P. J. Mitchell and D. Fincham. Shell model simulations by adiabatic dynamics. *Journal of Physics: Condensed Matter*, 5(8):1031, 1993.

W. Morris, B. Leung, H. Furukawa, O. K. Yaghi, N. He, H. Hayashi, Y. Houndonougbo, M. Asta, B. B. Laird, and O. M. Yaghi. A combined experimental–computational investigation of carbon dioxide capture in a series of isoreticular zeolitic imidazolate frameworks. *Journal of the American Chemical Society*, 132(32):11006–11008, 2010.

E. Mugnaioli, I. Andrusenko, T. Schüler, N. Loges, R. E. Dinnebier, M. Panthöfer, W. Tremel, and U. Kolb. Abinitio structure determination of vaterite by automated electron diffraction. *Angewandte Chemie International Edition*, 51(28): 7041–7045, 2012.

K. Naka, S.-C. Huang, and Y. Chujo. Formation of stable vaterite with poly(acrylic acid) by the delayed addition method. *Langmuir*, 22(18):7760–7767, 2006.

N. Nassif, N. Pinna, N. Gehrke, M. Antonietti, C. Jäger, and H. Cölfen. Amorphous layer around aragonite platelets in nacre. *Proceedings of the National Academy of Sciences of the United States of America*, 102(36):12653–12655, 2005.

A. Natoli, M. Wiens, H.-C. Schröder, M. Stifanic, R. Batel, A. L. Soldati, D. E. Jacob, and W. E. Müller. Bio-vaterite formation by glycoproteins from freshwater pearls. *Micron*, 41(4):359–366, 2010.

H. Nebel, M. Neumann, C. Mayer, and M. Epple. On the structure of amorphous calcium carbonatea detailed study by solid-state NMR spectroscopy. *Inorganic Chemistry*, 47(17):7874–7879, 2008.

M. Neumann and M. Epple. Monohydrocalcite and its relationship to hydrated amorphous calcium carbonate in biominerals. *European Journal of Inorganic Chemistry*, 2007(14):1953–1957, 2007.

F. Nudelman, B. A. Gotliv, L. Addadi, and S. Weiner. Mollusk shell formation: Mapping the distribution of organic matrix components underlying a single aragonitic tablet in nacre. *Journal of Structural Biology*, 153(2):176–187, 2006.

F. Nudelman, E. Sonmezler, P. H. H. Bomans, G. d. With, and N. A. J. M. Sommerdijk. Stabilization of amorphous calcium carbonate by controlling its particle size. *Nanoscale*, 2(11):2436–2439, 2010.

S. Özkar and R. G. Finke. Nanocluster formation and stabilization fundamental studies: ranking commonly employed anionic stabilizers via the development, then application, of five comparative criteria. *Journal of the American Chemical Society*, 124(20):5796–5810, 2002.

J. M. Pandolfi and W. Kiessling. Gaining insights from past reefs to inform understanding of coral reef response to global climate change. *Current Opinion in Environmental Sustainability*, 7:52–58, 2014.

A. Pavese, M. Catti, G. D. Price, and R. A. Jackson. Interatomic potentials for $CaCO_3$ polymorphs (calcite and aragonite), fitted to elastic and vibrational data. *Physics and Chemistry of Minerals*, 19(2):80–87, 1992.

C. J. Pickard and R. J. Needs. Ab initio random structure searching. *Journal of Physics: Condensed Matter*, 23(5):053201, 2011.

Y. Politi, Y. Levi-Kalisman, S. Raz, F. Wilt, L. Addadi, S. Weiner, and I. Sagi. Structural characterization of the transient amorphous calcium carbonate precur-

sor phase in sea urchin embryos. *Advanced Functional Materials*, 16(10):1289–1298, 2006.

Y. Politi, R. A. Metzler, M. Abrecht, B. Gilbert, F. H. Wilt, I. Sagi, L. Addadi, S. Weiner, and P. U. P. A. Gilbert. Transformation mechanism of amorphous calcium carbonate into calcite in the sea urchin larval spicule. *Proceedings of the National Academy of Sciences*, 105(45):17362–17366, 2008.

E. M. Pouget, P. H. H. Bomans, J. A. C. M. Goos, P. M. Frederik, G. d. With, and N. A. J. M. Sommerdijk. The initial stages of template-controlled $CaCO_3$ formation revealed by cryo-TEM. *Science*, 323(5920):1455–1458, 2009.

E. M. Pouget, P. H. H. Bomans, A. Dey, P. M. Frederik, G. de With, and N. A. J. M. Sommerdijk. The development of morphology and structure in hexagonal vaterite. *Journal of the American Chemical Society*, 132(33):11560–11565, 2010.

D. Quigley and P. M. Rodger. Free energy and structure of calcium carbonate nanoparticles during early stages of crystallization. *The Journal of Chemical Physics*, 128(22):221101, 2008a.

D. Quigley and P. M. Rodger. Metadynamics simulations of ice nucleation and growth. *The Journal of Chemical Physics*, 128(15):154518, 2008b.

D. Quigley and P. Rodger. A metadynamics-based approach to sampling crystallisation events. *Molecular Simulation*, 35(7):613–623, 2009.

D. Quigley, C. L. Freeman, J. H. Harding, and P. M. Rodger. Sampling the structure of calcium carbonate nanoparticles with metadynamics. *The Journal of Chemical Physics*, 134(4):044703, 2011.

A. V. Radha, T. Z. Forbes, C. E. Killian, P. U. P. A. Gilbert, and A. Navrotsky. Transformation and crystallization energetics of synthetic and biogenic amorphous calcium carbonate. *Proceedings of the National Academy of Sciences*, 107(38): 16438–16443, 2010.

P. Raiteri and J. D. Gale. Water is the key to nonclassical nucleation of amorphous calcium carbonate. *Journal of the American Chemical Society*, 132(49):17623–17634, 2010.

P. Raiteri, J. D. Gale, D. Quigley, and P. M. Rodger. Derivation of an accurate force-field for simulating the growth of calcium carbonate from aqueous solution: A new model for the calcitewater interface. *The Journal of Physical Chemistry C*, 114(13):5997–6010, 2010.

P. Raiteri, R. Demichelis, J. D. Gale, M. Kellermeier, D. Gebauer, D. Quigley, L. B. Wright, and T. R. Walsh. Exploring the influence of organic species on pre- and post-nucleation calcium carbonate. *Faraday Discussions*, 159(0):61–85, 2013.

S. Raz, P. Hamilton, F. Wilt, S. Weiner, and L. Addadi. The transient phase of amorphous calcium carbonate in sea urchin larval spicules: The involvement of proteins and magnesium ions in its formation and stabilization. *Advanced Functional Materials*, 13(6):480–486, 2003.

S. Raz, O. Testeniere, A. Hecker, S. Weiner, and G. Luquet. Stable amorphous calcium carbonate is the main component of the calcium storage structures of the crustacean orchestia cavimana. *The Biological Bulletin*, 203(3):269–274, 2002.

A. Ridgwell, D. N. Schmidt, C. Turley, C. Brownlee, M. T. Maldonado, P. Tortell, and J. R. Young. From laboratory manipulations to earth system models: scaling calcification impacts of ocean acidification. *Biogeosciences*, 6(11):2611–2623, 2009.

J. D. Rodriguez-Blanco, S. Shaw, and L. G. Benning. The kinetics and mechanisms of amorphous calcium carbonate (ACC) crystallization to calcite, via vaterite. *Nanoscale*, 3(1):265, 2011.

I. Rodrguez-Ruiz, S. Veesler, J. Gmez-Morales, J. M. Delgado-Lpez, O. Grauby, Z. Hammadi, N. Candoni, and J. M. Garca-Ruiz. Transient calcium carbonate hexahydrate (ikaite) nucleated and stabilized in confined nano- and picovolumes. *Crystal Growth and Design*, 14(2):792–802, 2014.

M. Saharay and R. James Kirkpatrick. Onset of orientational order in amorphous calcium carbonate (ACC) upon dehydration. *Chemical Physics Letters*, 591:287–291, 2014.

M. Saharay, A. O. Yazaydin, and R. J. Kirkpatrick. Dehydration-induced amorphous phases of calcium carbonate. *The Journal of Physical Chemistry B*, 117(12):3328–3336, 2013.

A. S. Schenk, H. Zope, Y.-Y. Kim, A. Kros, N. A. J. M. Sommerdijk, and F. C. Meldrum. Polymer-induced liquid precursor (PILP) phases of calcium carbonate formed in the presence of synthetic acidic polypeptidesrelevance to biomineralization. *Faraday Discussions*, 159:327, 2012.

T. Schlick. *Molecular modeling and simulation: an interdisciplinary guide*. Number v. 21 in Interdisciplinary applied mathematics. Springer, 2nd ed edition, 2010. ISBN 9781441963505.

A. Sigel, H. Sigel, and R. K. O. Sigel. *Biomineralization: From Nature to Application*. John Wiley and Sons, 2008. ISBN 9780470986318.

J. W. Singer, A. Ö. Yazaydin, R. J. Kirkpatrick, and G. M. Bowers. Structure and transformation of amorphous calcium carbonate: A solid-state ^{43}Ca NMR and computational molecular dynamics investigation. *Chemistry of Materials*, 24(10): 1828–1836, 2012.

W. Smith, C. Yong, and P. Rodger. DL_POLY: Application to molecular simulation. *Molecular Simulation*, 28(5):385–471, 2002.

D. Stauffer and A. Aharony. *Introduction To Percolation Theory*. CRC Press, 1994. ISBN 9781420074796.

P. J. Steinhardt, D. R. Nelson, and M. Ronchetti. Bond-orientational order in liquids and glasses. *Physical Review B*, 28(2):784–805, 1983.

C. J. Stephens, S. F. Ladden, F. C. Meldrum, and H. K. Christenson. Amorphous calcium carbonate is stabilized in confinement. *Advanced Functional Materials*, 20(13):2108–2115, 2010.

M. Stumpp, M. Y. Hu, F. Melzner, M. A. Gutowska, N. Dorey, N. Himmerkus, W. C. Holtmann, S. T. Dupont, M. C. Thorndyke, and M. Bleich. Acidified seawater impacts sea urchin larvae pH regulatory systems relevant for calcification. *Proceedings of the National Academy of Sciences*, 109(44):18192–18197, 2012.

I. P. Swainson. The structure of monohydrocalcite and the phase composition of the beachrock deposits of lake butler and lake fellmongery, South Australia. *American Mineralogist*, 93(7):1014–1018, 2008.

T. Threlfall. Structural and thermodynamic explanations of ostwald's rule. *Organic Process Research and Development*, 7(6):1017–1027, 2003.

I. T. Todorov, W. Smith, K. Trachenko, and M. T. Dove. DL_POLY_3: new dimensions in molecular dynamics simulations via massive parallelism. *Journal of Materials Chemistry*, 16(20):1911–1918, 2006.

H. Tomono, H. Nada, F. Zhu, T. Sakamoto, T. Nishimura, and T. Kato. Effects of magnesium ions and water molecules on the structure of amorphous calcium carbonate: A molecular dynamics study. *The Journal of Physical Chemistry B*, 117(47):14849–14856, 2013.

G. M. Torrie and J. P. Valleau. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, 23(2):187–199, 1977.

K. M. Towe and H. A. Lowenstam. Ultrastructure and development of iron mineralization in the radular teeth of cryptochiton stelleri (mollusca). *Journal of Ultrastructure Research*, 17(1):1–13, 1967.

G. A. Tribello, F. Bruneval, C. Liew, and M. Parrinello. A molecular dynamics study of the early stages of calcium carbonate growth. *The Journal of Physical Chemistry B*, 113(34):11680–11687, 2009.

A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard. ReaxFF: a reactive force field for hydrocarbons. *The Journal of Physical Chemistry A*, 105(41):9396–9409, 2001.

A. Verch, D. Gebauer, M. Antonietti, and H. Cölfen. How to control the scaling of $CaCO_3$: a fingerprinting technique to classify additives. *Physical Chemistry Chemical Physics*, 13(37):16811, 2011.

D. J. Wales. *Energy landscapes.* Cambridge molecular science. Cambridge University Press, 2003. ISBN 0521814154.

A. F. Wallace, L. O. Hedges, A. Fernandez-Martinez, P. Raiteri, J. D. Gale, G. A. Waychunas, S. Whitelam, J. F. Banfield, and J. J. D. Yoreo. Microscopic evidence for liquid-liquid separation in supersaturated $CaCO_3$ solutions. *Science*, 341(6148):885–889, 2013.

M. R. Walsh, C. A. Koh, E. D. Sloan, A. K. Sum, and D. T. Wu. Microsecond simulations of spontaneous methane hydrate nucleation and growth. *Science*, 326 (5956):1095–1098, 2009.

D. Wang, L. M. Hamm, A. J. Giuffre, T. Echigo, J. D. Rimstidt, J. J. D. Yoreo, J. Grotzinger, and P. M. Dove. Revisiting geochemical controls on patterns of carbonate deposition through the lens of multiple pathways to mineralization. *Faraday Discussions*, 159(0):371–386, 2013.

J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9):1157–1174, 2004.

J. Wang and U. Becker. Structure and carbonate orientation of vaterite ($CaCO_3$). *American Mineralogist*, 94(2):380–386, 2009.

S.-S. Wang and A.-W. Xu. Amorphous calcium carbonate stabilized by a flexible biomimetic polymer inspired by marine mussels. *Crystal Growth and Design*, 13 (5):1937–1942, 2013.

Y.-W. Wang, Y.-Y. Kim, C. J. Stephens, F. C. Meldrum, and H. K. Christenson. In situ study of the precipitation and crystallization of amorphous calcium carbonate (ACC). *Crystal Growth and Design*, 12(3):1212–1217, 2012.

S. Weiner, I. Sagi, and L. Addadi. Choosing the crystallization path less traveled. *Science*, 309(5737):1027–1028, 2005.

S. Weiner. Biomineralization: A structural perspective. *Journal of Structural Biology*, 163(3):229–234, 2008.

S. Weiner and L. Addadi. Crystallization pathways in biomineralization. *Annual Review of Materials Research*, 41(1):21–40, 2011.

S. E. Wolf, J. Leiterer, V. Pipich, R. Barrea, F. Emmerling, and W. Tremel. Strong stabilization of amorphous calcium carbonate emulsion by ovalbumin: Gaining insight into the mechanism of polymer-induced liquid precursor processes. *Journal of the American Chemical Society*, 133(32):12642–12649, 2011.

S. M. Woodley and R. Catlow. Crystal structure prediction from first principles. *Nature Materials*, 7(12):937–946, 2008.

Y. Wu, H. L. Tepper, and G. A. Voth. Flexible simple point-charge water model with improved liquid-state properties. *The Journal of Chemical Physics*, 124(2): 024503, 2006.

J. Xiao and S. Yang. Unveiling the critical process in which organic molecules control the polymorphism of magnesium-containing calcium carbonate: the early nucleation of amorphous precursors or the subsequent amorphous to crystalline transformations? *CrystEngComm*, 13(20):6223, 2011.

X. Xu, C. L. Ting, I. Kusaka, and Z.-G. Wang. Nucleation in polymers and soft matter. *Annual Review of Physical Chemistry*, 65(1):449–475, 2014.

J. R. Young, S. A. Davis, P. R. Bown, and S. Mann. Coccolith ultrastructure and biomineralisation. *Journal of Structural Biology*, 126(3):195–215, 1999.

Y. Zhao, H. Li, and X. C. Zeng. First-principles molecular dynamics simulation of atmospherically relevant anion solvation in supercooled water droplet. *Journal of the American Chemical Society*, 135(41):15549–15558, 2013.