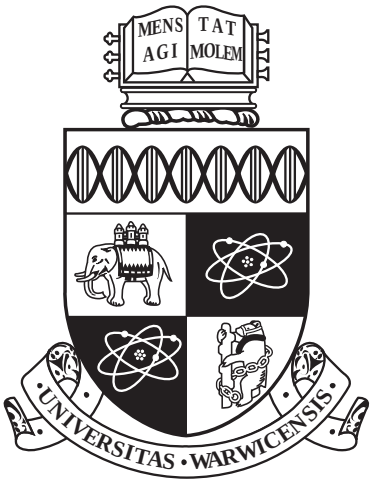

Polymorphism in Biomineral Nanoparticles

by

Anthony Matthew Bano

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

MOAC Doctoral Training Centre

July 2012

THE UNIVERSITY OF
WARWICK

Contents

Dedication

This work is dedicated to the memory of Carl Blakey, much missed by his friends.

Acknowledgments

I would like to thank, first and foremost, my supervisors, Prof. Mark Rodger and Dr. David Quigley, for their unfailing support and assistance. I have benefited from their guidance a great deal, and not just on topics related to this work.

I must also thank Prof. Alison Rodger for all manner of help throughout the past four years. Even aside from other considerations, I will forever be in her debt just for letting me join the MOAC Doctoral Training Centre in the first place, as this has been a life changing experience. I have also received much support from the wider MOAC and Warwick System biology communities over these years, both from my cohort of fellow students, and administrative and support staff past and present.

At an early stage in my studies Prof. Sascha Ott, Dr. Keith Vance, and Dr. Sach Mukherjee were extremely helpful when I was in difficulties, and were important in steering me towards the project detailed in this work.

Over the course of the project I have benefited from conversations with many academics in presentations, group meetings and the like. I would particularly like to single out and thank Prof. Mike Allen, Prof. Baron Peters, and Prof. Julian Gale.

My friends and family have been vital to any success I have had in my studies, and have been endlessly patient through good times and bad. I must particularly single out housemates past and present, Emma Argles, Dr. Sian Taylor-Phillips, and Iain Wallace. My climbing partner Philip Griffiths has also been key to helping me keep my sanity, particularly in the closing stages of the project.

Declarations

This thesis is all my own work except where explicitly stated in the text. No part of it has been submitted for examination for any other higher degree, and to date none of this work has been published elsewhere.

Abstract

Biomineralisation is the process by which living things produce hard mineral tissues with unique physical properties. The study of this process can help us produce biomimetic materials, reproducing such properties, with the study of nucleation and crystallisation of the materials being particularly important. I have used molecular simulation techniques to help gain a greater understanding of these processes, focussing particularly on identifying the conformations and solid phases available to nanoparticles of two biomineral compounds.

The bones and teeth of mammals are made largely of calcium phosphates. I have used metadynamics to study nanoparticles of tricalcium phosphate (TCP) and have identified high and lower order configurations. To facilitate this work I reviewed the extant empirical potentials for calcium phosphate systems, selecting the most appropriate for TCP.

Calcium carbonate, found in examples throughout the animal kingdom, has three crystalline polymorphs relevant to biomineralisation: calcite, aragonite and vaterite. While nanoparticles of calcite have been extensively studied the other polymorphs have been neglected to date. In this work I present a technique for predicting crystalline morphologies for all three polymorphs across a range of sizes, and compare the energetic ordering. In water the energetic ordering of the nanoparticles is heavily dependent on nanoparticle size.

Furthermore, I present work calculating the surface enthalpies of a variety of calcium carbonate surfaces, many of which are negative. It appears that entropic penalty of ordered water is key to understanding the stability of nanocrystals.

Also presented is an application of the nudged elastic band method to study transitions between nanoparticle crystal conformations. Between all three crystal polymorphs the nanoparticles passed through an amorphous region of phase space. These results have also been used to evaluate order parameters for use in metadynamics simulations

Abbreviations

ACC	Amorphous Calcium Carbonate
BFGS	Broyden-Fletcher-Goldfarb-Shanno
CBMC	Configuration Bias Monte Carlo
CV	Collective Variable
CNT	Classical Nucleation Theory
DFT	Density Functional Theory
FES	Free Energy Surface
GULP	General Unility Lattice Program
ISMC	Ion-site Monte Carlo
MC	Monte Carlo
MD	Molecular Dynamics
NEB	Nudged Elastic Band
OP	Order Parameter
PILP	Polymer-Induced Liquid-Precursor
RDF	Radial Distribution Function

RMSD	Root-Mean-square Deviation
RSS	Random Structure Seaching
SEM	Scanning Electron Microscopy
ss-NMR	solid-state Nuclear Magnetic Resonance
TCP	Tricalcium Phosphate
TEM	Transmission Electron Microscopy
WAXS	Wide Angle X-ray Scattering

Chapter 1

Introduction

Biomineralistion is the process by which living things produce hard materials. As discussed in sections 1.1 and 1.2, these materials can have unique physical properties. The ability to recreate such properties in novel advanced materials through biomimetic means would be extremely useful. In this thesis I present work using molecular simulation to gain new understanding of the nucleation and growth of the two most important classes of biominerals. Molecular simulation is ideal for such a study, as it provides atomistic resolution of structures and mechanisms impossible to gain through other means.

In this chapter I describe the properties of the bulk crystals of these materials, and give short reviews of the experimental evidence on the topic of biomaterial nucleation and crystallisation, as well as the results of prior simulation studies.

1.1 Calcium Carbonates

Calcium carbonates are perhaps the best studied class of biominerals. While comparatively rare in the mammalia, they feature extensively in many classes of living organisms, from bird egg-shells [Jeon et al., 1993] to crustacean skeletons [Addadi et al., 2003], to the crusts of bacterial colonies [Qiao et al., 2008].

Pure, bulk calcium carbonate ($CaCO_3$) has three main crystalline polymorphs: these are calcite, aragonite and vaterite. Here each is considered in turn, and the crystalline hydrated forms and amorphous phases are also reviewed.

1.1.1 Calcite

Calcite is the most stable polymorph of calcium carbonate. Found extensively in biological processes, and partly as a result of this, also in sedimentary rock types such as limestone and marble. Large single crystals can occur, with the variant Iceland spar exhibiting the unusual phenomenon of double refraction.

Crystal Structure

The crystal lattice is hexagonal with space group $R3\bar{c}$. As shown in figure 1.1 the carbonate groups form flat planar layers perpendicular to the c-axis, interleaved between layers of Ca^{2+} ions. Both layer types are arranged in a hexagonal lattice. Considering each carbonate as a single ion, both ion types are twelve-fold coordinated by their own kind of ion (six on the same layer, three each two layers above and below), and six fold coordinated by the opposite kind of ion (three each on the layers above and below). All the carbonate ions on a layer are in the same orientation, with either all ions with one C—O bond pointing along the **a** direction, or all ions with a bond pointing along the $-\mathbf{a}$ direction. The reference calcite structure used in this work is that of Maslen et al. [1993].

Role in Biomineralisation

As the most stable polymorph, calcite is a major component of many forms of biomineral calcium carbonate. For example, chicken egg shells are 96% calcite by weight with the remaining 4% being composed of biological matrix. Beyond simply protecting the developing embryo with its mechanical strength, the egg shell must also allow water and gas exchanges, protect against temperature fluctuations, and provide a source of calcium for the embryo [Chien et al., 2008]. Figure 1.2 shows the structure of the egg shell.

Another example of calcite as a biomineral are the coccoliths certain species of algae construct around themselves as shells (coccospheres) [Henriksen et al., 2004]. Coccolith growth begins inside liquid vesicles where the supersaturation of calcium carbonate can be controlled. Once nucleated, the crystal morphology and growth is tightly controlled, with a precision that is simply not possible to achieve in a laboratory.

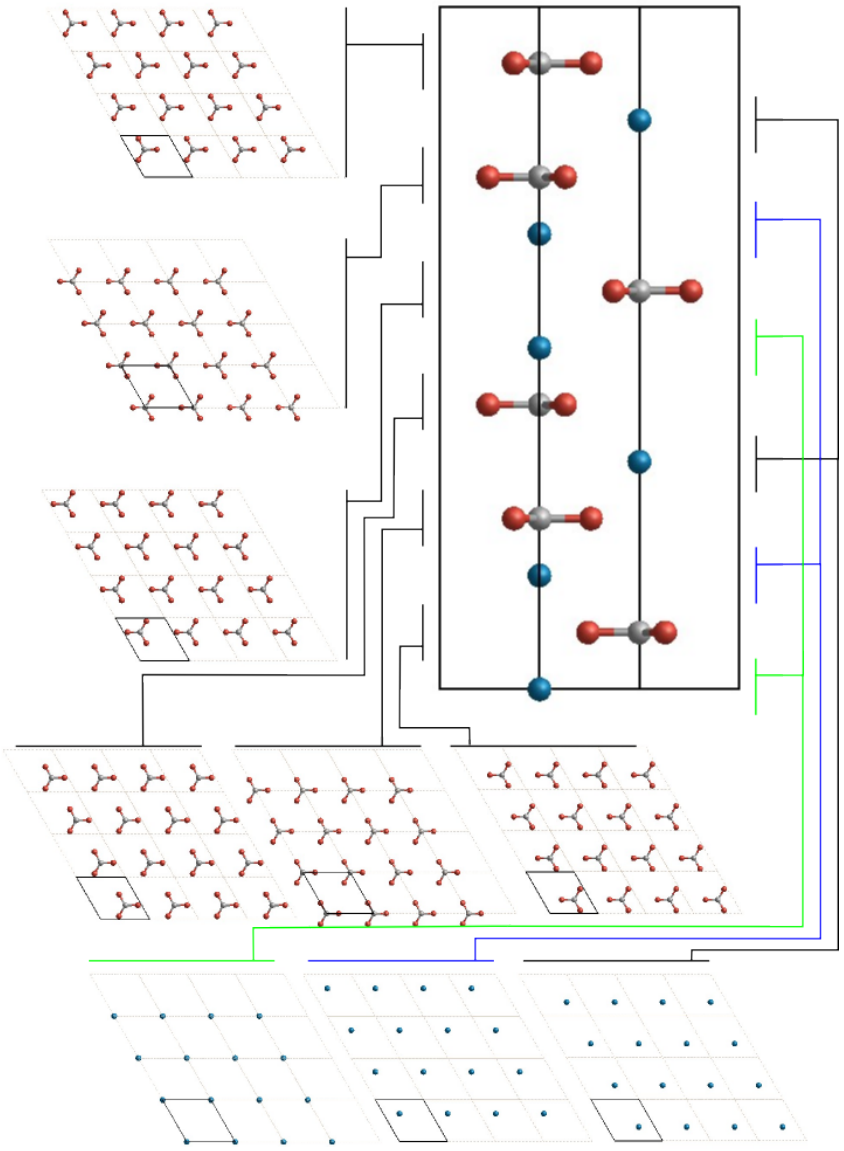

Figure 1.1: Calcite unit cell of Maslen et al. [1993] with Ca^{2+} and carbonate layers. Both layer types form a hexagonal sub-lattice. Layers are in the **ab** plane.

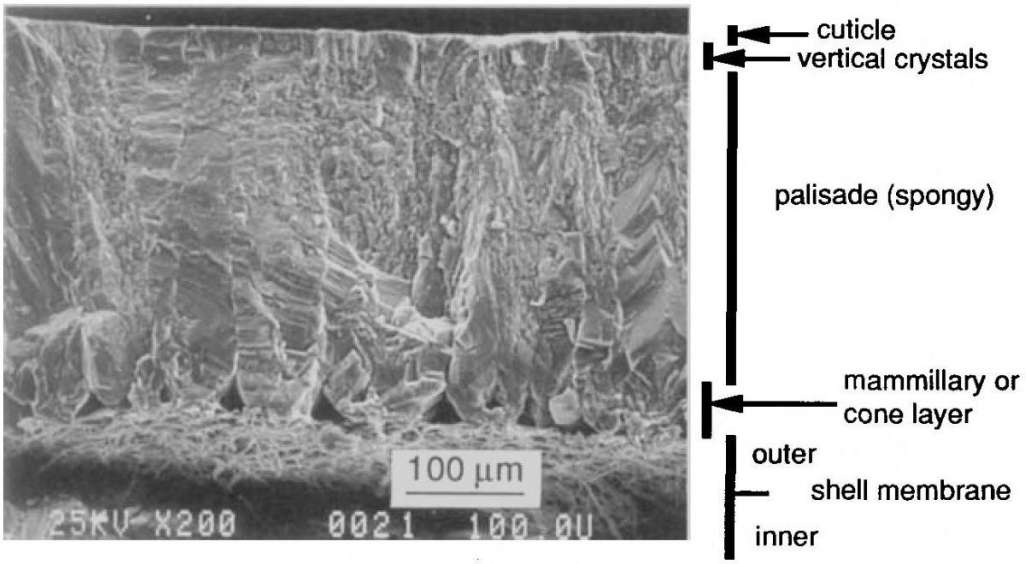

Figure 1.2: Egg shell cross-section, showing distinct layers. Figure is from Jeon et al. [1993].

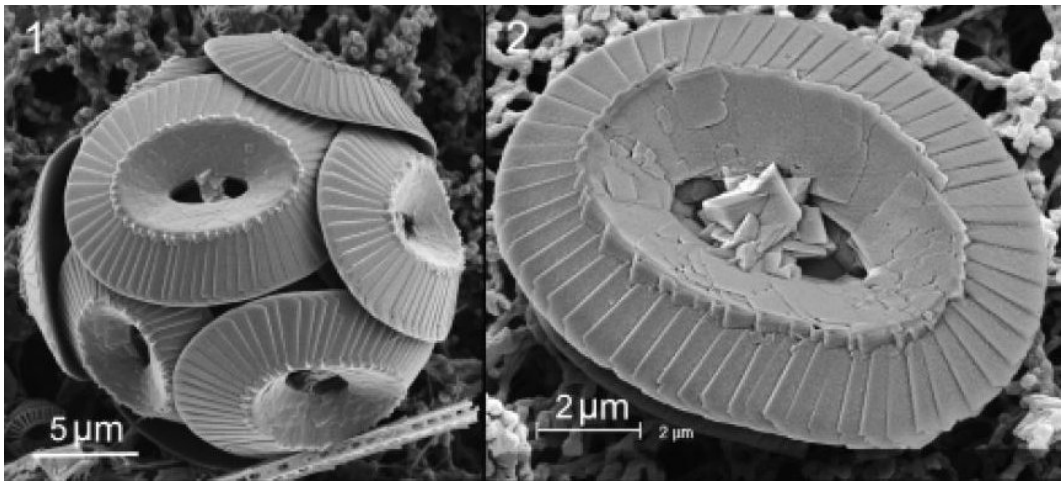

Figure 1.3: SEM (Scanning Electron Microscopy) pictures of a (1) coccosphere and (2) coccolith from the species *Coccolithus pelagicus*. From Henriksen et al. [2004]

1.1.2 Aragonite

Aragonite, named for the Aragon region of Spain, is the next most stable bulk polymorph of pure calcium carbonate. Also occurring (rarely) in large single crystals it is again an important biomineral.

Crystal Structure

The crystal lattice is orthorhombic, with space group *Pmcn*. The carbonate ions are organised into planar layers perpendicular to the **a** axis. In this case however, there are two carbonate layers for each Ca^{2+} layer, with the Ca^{2+} layer twice as densely occupied. All the carbonate ions on a layer are in the same orientation, with either all ions with one C—O bonds pointing along the **c** direction, or all ions with a bond pointing along the $-\mathbf{c}$ direction. If the first of these is denoted Cb+ and the latter Cb- then moving along the **a**-axis the layers are [Cb+][Ca][Cb+][Cb-][Ca][Cb-] with the sequence then repeating. Each carbonate layer forms a rectangular lattice, with the projection of two adjoining carbonate layers forming a hexagonal lattice. Each Ca^{2+} layer forms a hexagonal lattice. The Ca^{2+} ions are coordinated by their six neighbours on the same layer, and three Ca^{2+} ions in the next Ca^{2+} layers. Each Ca^{2+} ion is six fold-coordinated by carbonate groups by three ions each from the layer-pairs above and below. Each carbonate ion is twelve fold coordinated by ions of the same kind — four from the same layer, two from the adjoining layer, and three each from each of the layers on the far side of the nearest Ca^{2+} layers. The reference aragonite structure used for this work is that of de Villiers [1971].

Role in Biomineralisation

Mollusc shells are made from biogenic aragonite. These shells with an inner layer, made of an aragonite nacre, are composed of flat slabs around 400 nm thick, with an organic matrix between layers. While only a small proportion by weight of the shell, this organic matrix allows the shell to bend, not break under stress, vastly increasing the strength over pure aragonite, while maintaining hardness. The 400 nm gap explains the iridescence of many such shells, as this is close to the visual wavelength of light — most pearls are also made of aragonite nacre. The layers are themselves made of subcrystals, which on average are organised with the **c**-axis perpendicular to the plain of each layer, though it has recently been shown that for each subunit there is considerable error in this alignment [Metzler et al., 2007].

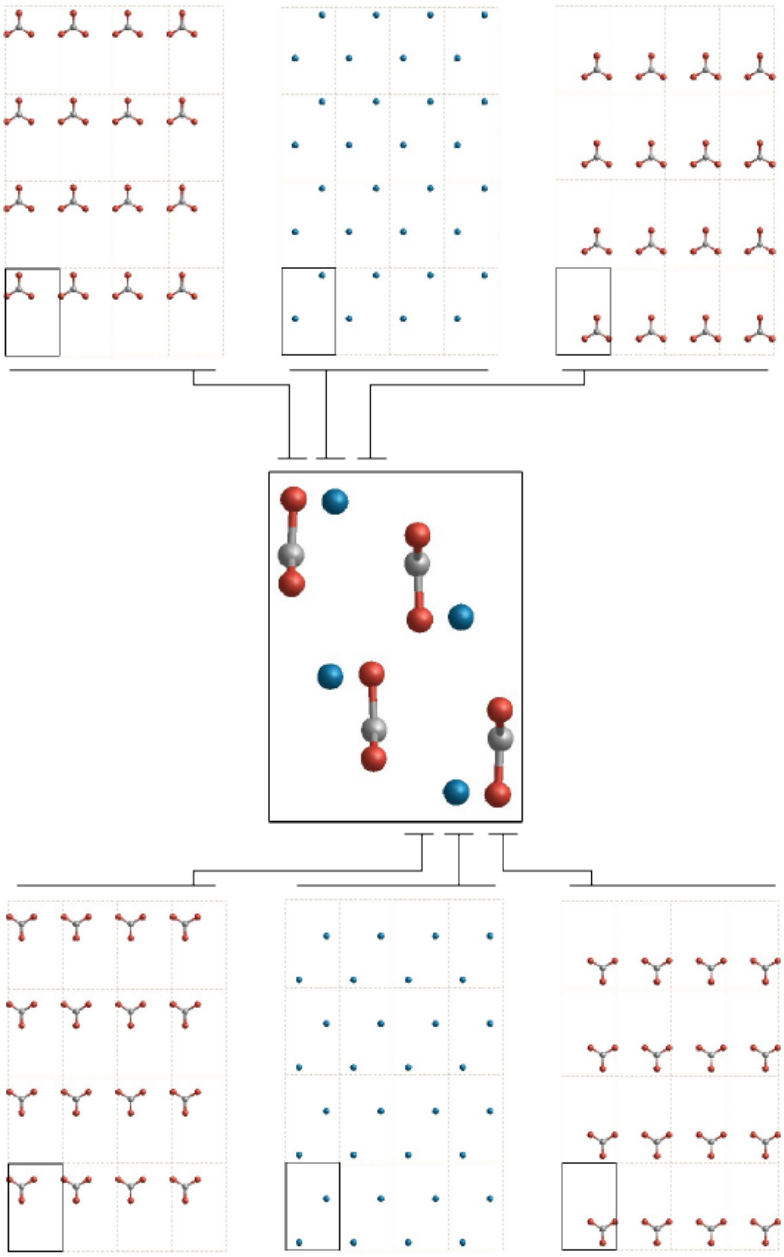

Figure 1.4: Aragonite unit cell of de Villiers [1971] with Ca^{2+} and carbonate layers. The projection of of two adjacent carbonate layers forms a hexagonal sub-lattice. Layers are in the **bc** plane.

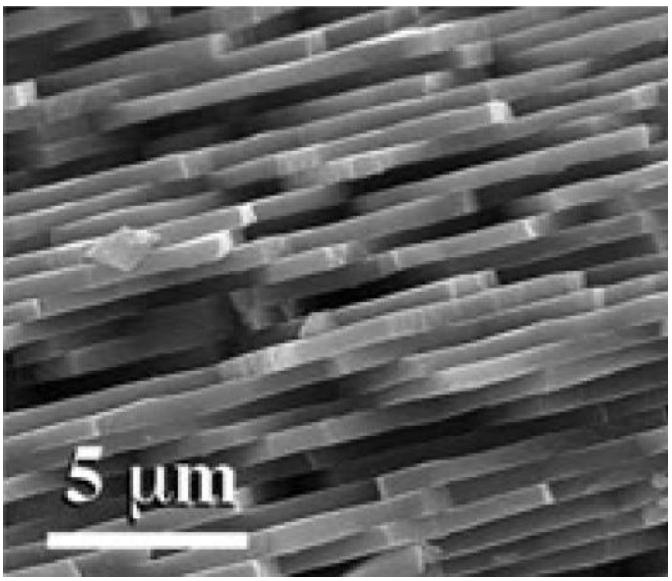

Figure 1.5: SEM picture of aragonite nacre (of species *U. pictorum*), showing the flat tablets arranged into layers. From Marie et al. [2007]

1.1.3 Vaterite

Unlike the other two bulk polymorphs vaterite does not form large single crystals and analysis is usually of the powder. Nevertheless, it is still important to the process of biomineralisation, both as a biomaterial in its own right and perhaps as in intermediate stage in aragonite or calcite growth through an Ostwald step process.

Crystal Structure

For aragonite and calcite the crystal structures are uncontroversial — refinements may be published from time to time but the overall details are well established. The final structure of vaterite is far from being a solved problem, though the following details are known:

- As with aragonite and calcite the Ca^{2+} ions form a hexagonal sublattice. With vaterite these layers are perpendicular to the c-axis, and unlike calcite or aragonite, they overlay one another. Each Ca^{2+} ion is therefore 8-fold coordinated by other Ca^{2+} ions.
- Between these layers are single carbonate layers. Rather than lying flat in

the plane of the layer these carbonates are at right angles to it. Ignoring the oxygen atoms the carbon atoms of these carbonates form rough hexagonal layers. These layers are offset from the layers below. Thus each carbonate is six fold coordinated by the carbonates on the same layer, and coordinated by three carbonates on each of the layers above and below.

- Each Ca^{2+} ion is six fold coordinated by carbonates — three each from the layers above and below. Similarly each carbonate is six fold coordinated by Ca^{2+} ions, again with three from each adjacent layer.

The source of the controversy is in the orientation of the carbonate groups. Meyer [1960] have all the carbonates lying in the **ac** plane, with a rhombohedral unit cell with space group *Pbmn* (figure 1.6A). In this structure all the oxygen sites are fully occupied, which is inconsistent with experimental results. Kamhi [1963] produced a structure with partially occupancy (figure 1.6B). The unit cell is hexagonal with space group $P6_3/mmc$. There are a total of six carbonate sites in the unit cell each with occupancy of one third. There are two clusters, each of three overlapping sites, with one and only one site on each cluster fully occupied. The subsites in each cluster each have one bond of the carbonate pointing towards the line joining two Ca ions in the adjoining layers. The carbon atoms in each subsite are at slightly different locations, though their van der Waals radii certainly overlap. the C—O on two carbonates next to each other cannot point towards the same Ca ion pair, introducing some local order, analogous to the ice rules that govern ice polymorphs such as ice I_H. Meyer [1969] produced a largely similar structure.

Medeiros et al. [2007] have produced a density functional theory (DFT) study of the orthorhombic structure of Meyer [1960], finding the local minimum. This structure again had 100 % occupancy of the carbonate oxygen sites, with the authors arguing that the experimental evidence is inconclusive. This structure does not have all the carbonates lying in the **ac** plane, with the carbonates having relaxed considerably.

More recently Wang and Becker [2009] have produced another hexagonal structure (figure 1.6C). They started with repeated units of the structure of Kamhi [1963] and, through a combination of molecular dynamics and simulated annealing, identified a low energy supercell, where the carbonates are arranged so as to minimise overall energy, with ordering not only within carbonate layers but between layers. Their unit cell is three times larger in the **c** axis, and twice as large in the **a** and

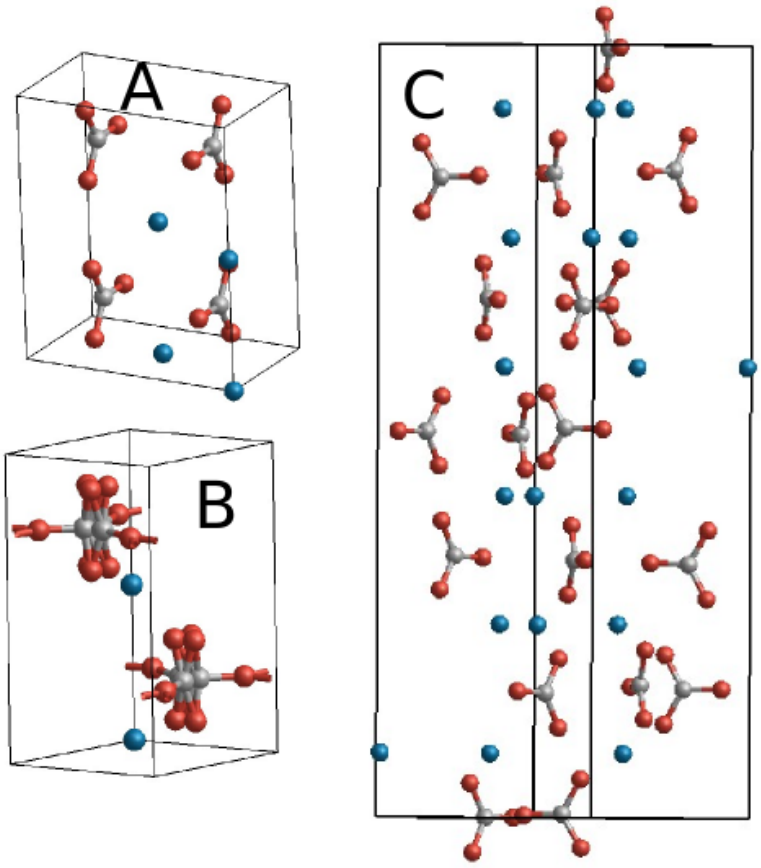

Figure 1.6: Proposed Vaterite Structures. **A.** Orthorhombic structure of Meyer [1960]. **B.** Hexagonal structure of Kamhi [1963], showing carbonate disorder — carbonates can point in any of three directions. **C.** Hexagonal structure of Wang and Becker [2009], a fully ordered supercell based on **B**.

b axes than that of Kamhi [1963]. The Ca^{2+} ions relax to positions slightly off the lattice sites due to their local environment. They argue that their supercell represents the low energy ideal and that freshly made experimental vaterite may be more disordered as some of energy penalties involved are very small. Perhaps surprisingly they do not evaluate this supercell using *ab initio* techniques such as DFT, as they did use these techniques to evaluate and dismiss the structure of Meyer [1960] as unsatisfactory. Figure 1.6 shows three of the proposed vaterite structures.

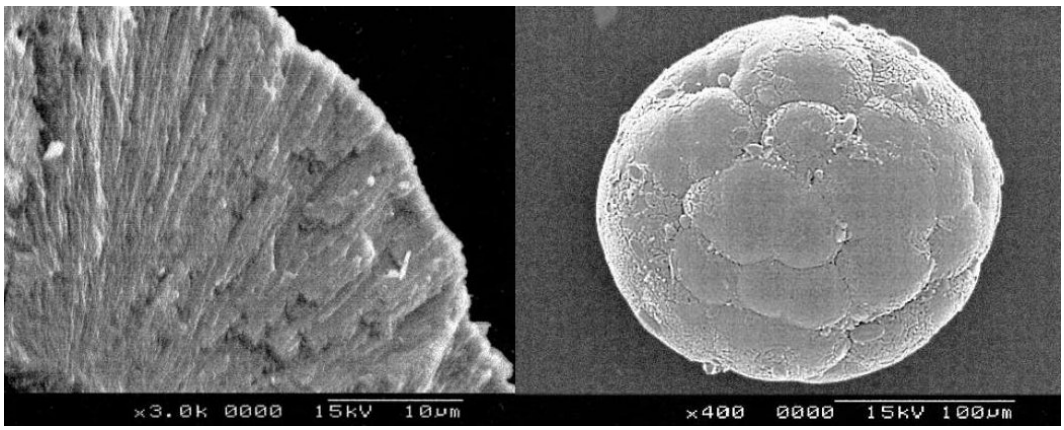

Figure 1.7: SEM pictures of microbial biscuits chiefly comprised of vaterite. Fibrous vaterite crystals (l) and an example biscuit (r). From Giralt et al. [2001]

Role in Biomineralisation

As the least stable bulk phase vaterite is less common than the other polymorphs in natural systems. Nevertheless, there are several biological examples of occurrence of this polymorph. Qiao et al. [2008] studied the vaterite freshwater pearls found in lakes in southern China. They observed that the powder of such pearls was very much more stable than synthetic vaterite. Giralt et al. [2001] found and studied 'microbial biscuits' of vaterite in a lake in Kyrgyzstan. These structures, which are spherical and around 150 μm across (figure 1.7) are composed of a fibrous vaterite coating of a bacterial colony. The authors believe the unusual chemistry of the lake, with high levels of phosphorus, are partially responsible for the stabilisation of vaterite with respect to the other phases. Vaterite is also found in systems such as gallstones, kidney stones and turtle eggshells [Qiao et al., 2008].

1.1.4 Hydrated Crystalline Phases

Monohydrocalcite ($CaCO_3.H_2O$) is mostly found in lake beds as spherules or as crusts. The crystal structure has a hexagonal unit cell with space group $P3_121$ and contains eight formula units.

Ikaite ($CaCO_3.6H_2O$) is found only in cold water and is only metastable — in water warmer than 8 °C it disintegrates into water and calcite. It has a monoclinic unit cell with space group $C2/c$, with each Ca^{2+}—$(CO_3)^{2-}$ ion pair surrounded by a cage of water molecules. There is no evidence in the literature that this material

is found in biological systems.

1.1.5 Amorphous Calcium Carbonates

Though often referred to as a single material, amorphous calcium carbonate (ACC) is in reality a whole continuum of materials with factors such as water content, particle size, local solute composition and pH, age, and contaminants all affecting stability. For the purposes of this thesis, concerned with nucleation of the crystal state from solution, the most important distinction is between wet, hydrated ACC and dry, anhydrous ACC. As extensively reviewed by Addadi et al. [2003], wet ACC (approximate formula $CaCO_3.H_2O$) is more stable than dry ACC, and in biogenic materials may be metastable over biological timescales. It is also noted in this paper that despite the name, stable ACC may possesses significant short-range order as evidenced by order in the IR spectra. Furthermore, from x-ray absorbtion experiments it has been shown that stable ACC from different organisms has significantly different local order despite similar stoichiometry.

As well as a bulk scale biomineral ACC is important as a precursor for the crystallisation of the ordered phases. It is known for example that for nacre tablets aragonite crystallisation begins with ACC nanoparticles [Metzler et al., 2007]. This is also observed in *in vitro* experiments as discussed in section 1.4.

A further phase, distinct from the solid ACC discussed above is the polymer-induced liquid-precursor (PILP) phase [Dai et al., 2008]. Where ACC is deposited from a supersaturation, the addition of polyaspartic acid has been observed to induce a liquid precursor phase of very small transient particles, which are deformable and soft. These particles are by their nature difficult to study, particularly as the drying required to study them using many techniques is sufficient to induce a phase transition to a more solid ACC form. Nevertheless, it is thought that PILP could form an important precursor phase to many biogenic materials.

1.2 Calcium Phosphates

Though more relevant for human health studies (mammalian bones and teeth are made mostly of calcium phosphates), the nucleation and growth of these compounds have not been as extensively studied as that of calcium carbonates, perhaps because the systems concerned represent a more complicated mixture of compounds. As extensively reviewed by Dorozhkin [2007], numerous compounds in various polymorphs

are composed of the basic species of Ca^{2+} ions, phosphate groups, and water before one considers protonated species. Here the apatites and tricalcium phosphate are considered, as the most biologically important and simplest systems respectively.

1.2.1 Apatites

The apatites have chemical formula $Ca_{10}(PO_4)_6X_2$ where X might be any of a number of ions with charge -1 e such as fluoride (fluorapatite) or hydroxyl groups (hydrxoyapatite: sometimes hydroxylapatite). Bones are made largely of hydroxyapatite, as is dental enamel and dentin. Synthetic hydroxyapatite is biocompatible with bone.

The fluorapatite unit cell (fig. 1.8) is hexagonal, with space group $P6_3/M$. The fluoride ions form channels in the c axis. Hydroxyapatite has a very similar unit cell with the hydroxyl groups lined up along the same channels. There is partial disorder in the hydroxyl group positions.

1.2.2 Tricalcium Phosphate

With only two species tricalcium phosphate (TCP — $Ca_4(PO_4)_2$) can be considered to be a simpler system to study than the apatites. There are two main polymorphs: the β polymorph is the more stable of the two, though α-TCP is metastable at standard temperature and pressure for long periods. There is also a γ polymorph, though this is only stable at high temperature and high pressure, and is not considered further in this thesis.

β-TCP occurs only in natural systems with a substantial proportion of the Ca^{2+} ions substituted by Mg^{2+}, in a mineral called Whitlockite. This is found in dental calculi and urinary stones. β-TCP is widely used in bone substitution bioceramics [Dorozhkin, 2007]. The unit cell of the pure substance (figure 1.9) is complicated compared to the calcium carbonate polymorphs, with total occupancy equivalent to 21 formula units. The cell is hexagonal, with space group $R3c$. There is disorder within the cell, as six of the Ca^{2+} ion sites have only 50% occupancy. The reference structure used in this work is that of Yashima et al. [2003].

α-TCP, while not found in biominerals in a pure form, is nonetheless important as a component of calcium phosphate cements. The substance may also be stabilised by doping with silicon. The only published structure of the unit cell is that of Mathew et al. [1977] (figure 1.10), which is a monoclinic structure of 24

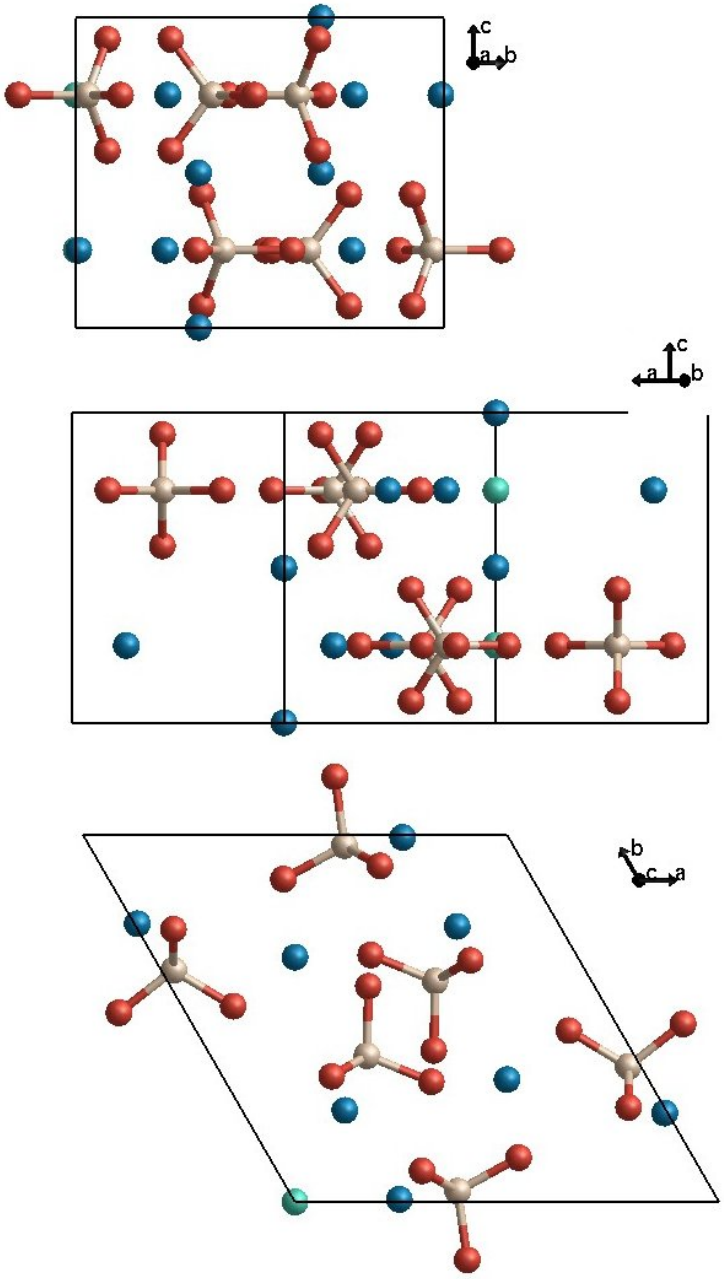

Figure 1.8: projections of fluorapatite unit cell along x,y and z axes as shown. Structure is that of [Hughes et al., 1989].

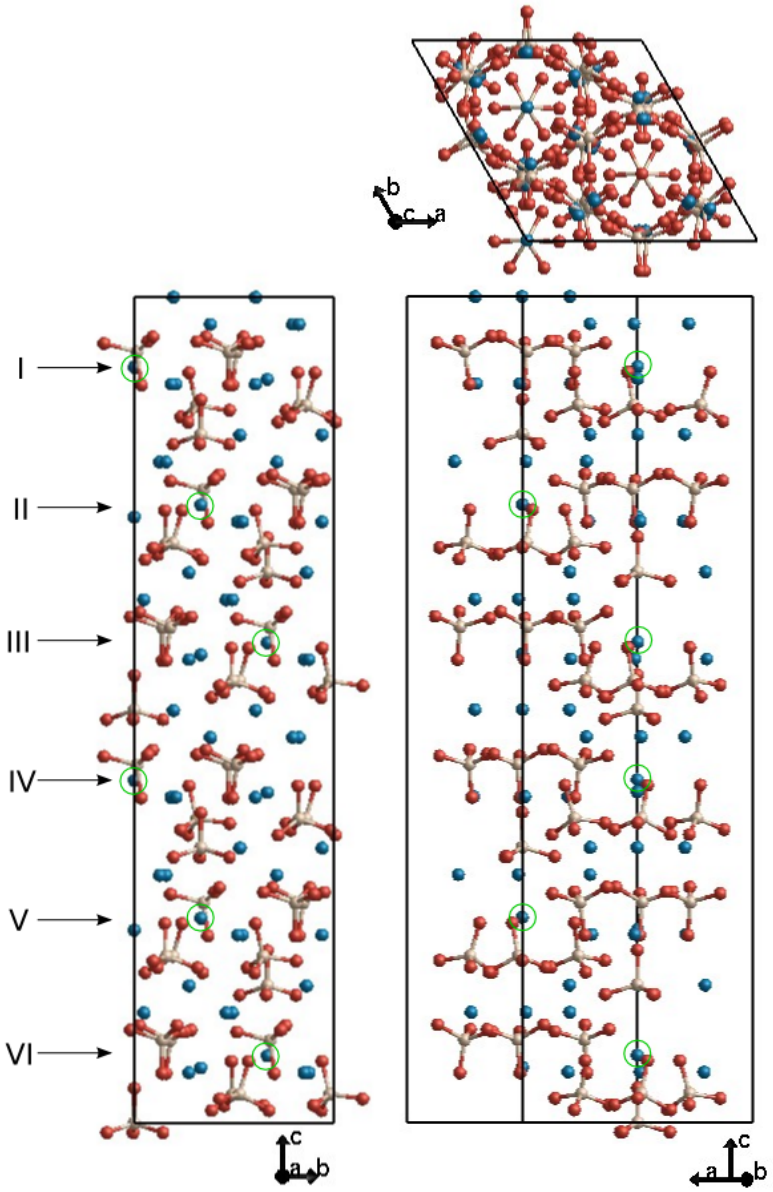

Figure 1.9: Projection of the unit cell of β-TCP unit cell of Yashima et al. [2003]. Partially occupied Ca^{2+} sites are highlighted, are numbered following Liang et al. [2010].

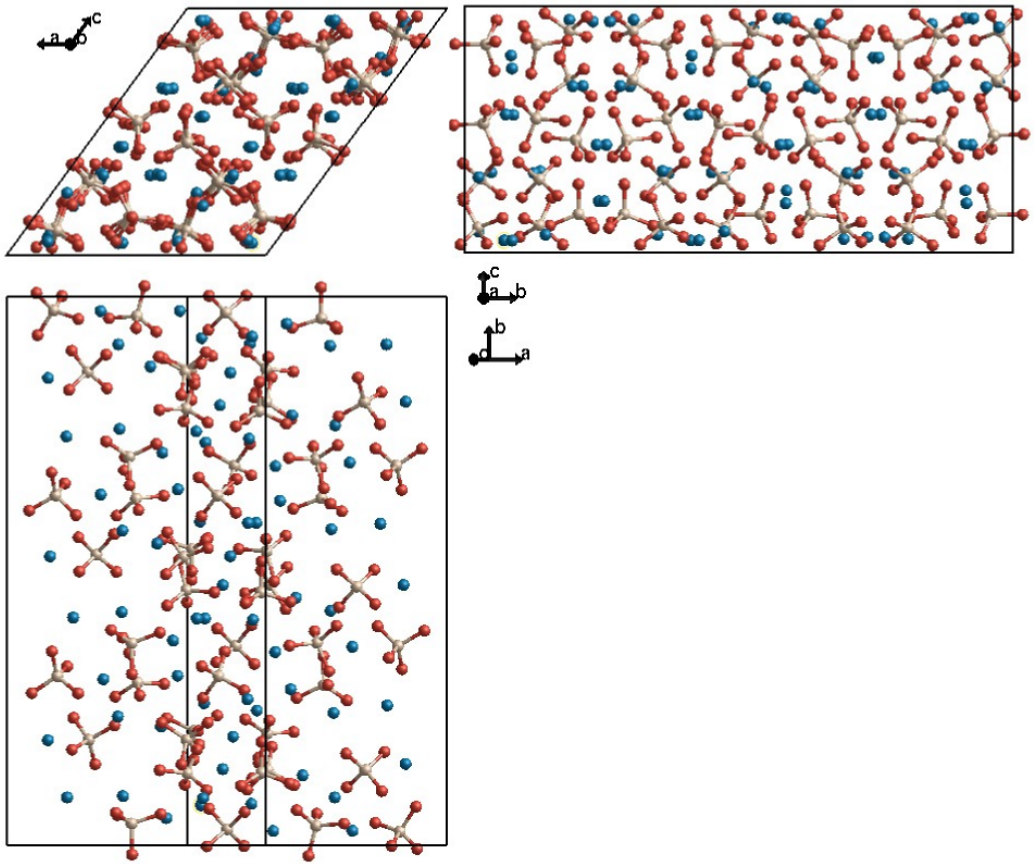

Figure 1.10: Projections of the unit cell of α-TCP unit cell of Mathew et al. [1977].

formula units, with space group $P2_1/a$.

1.3 Classical Nucleation Theory and the alternatives

Nucleation and growth of crystalline materials from solution can be a highly complicated process, and for many systems the process is poorly understood. Introduced here are some of the ideas and theories which are informing present thinking on the growth of biomaterials.

Classical nucleation theory (CNT) assumes a model of single unit growth in which particles transiently associate and disassociate in a supersaturated solution. Forming an interface between solvent and solute causes an energetic penalty scaling with the surface area, and hence square of the radius, of the particle. There is

an energetic gain in growth however, scaling with the volume of the particle, and hence the cube of the radius. Thus there is a critical radius at which these two terms balance — it is energetically favourable (and hence more probable) for a particle smaller than this critical radius to shrink, while a particle larger than this critical radius is thermodynamically likely to grow. On atomistic length scales (or in appropriate circumstances, macroscopic length scales) nucleation is a rare event — crossing the energy barrier to growth requires a particle to grow by chance to a size beyond the critical radius.

CNT as outlined above makes a great deal of assumptions. Growth is assumed to be homogeneous around the particle, with no dependence on particle size in the interaction between the particle and the growth unit. Importantly for biomineral systems, polymorphism is not considered within this framework — while calcite is the most stable bulk calcium carbonate phase any of the other polymorphs might be more stable for small particle systems due to surface and edge effects differing between structures. This might lead to the particles transforming through a series of phases as they grow.

Furthermore, in many systems (including calcium carbonates as discussed in section 1.4) pre-critical nuclei have been observed. These particles, smaller than the predicted CNT critical radius, are observed to be metastable for long periods and may well be crucial to the nucleation of crystalline phases.

Figure 1.11 outlines some of the possible mechanisms. Ions may crystallise out of solution through CNT, either directly to the crystalline phase or through an amorphous precursor stage. Alternatively crystallisation may occur through pre-critical clusters. These themselves may be amorphous or crystalline, and may assemble through aggregation. If crystalline, this aggregation requires some degree of orientation in the attachment process. Another possibility is mesocrystal formation. A mesocrystal is a system of orientated sub-crystals linked together by layers of another material, such as a polymer, or even the amorphous phase of the same substance [Niederberger and Colfen, 2006]. It can be difficult to distinguish between mesocrystal and true single crystals using experimental techniques, and it has been speculated that many biocrystals may actually be mesocrystals.

Any number of factors such as the pH of the solvent, biological agents such as proteins, or the confinement circumstances of the solution may effect nucleation and growth, as well as polymorph selection. It is the latter that is the focus of this thesis: for calcium carbonates I have attempted to discover which polymorphs

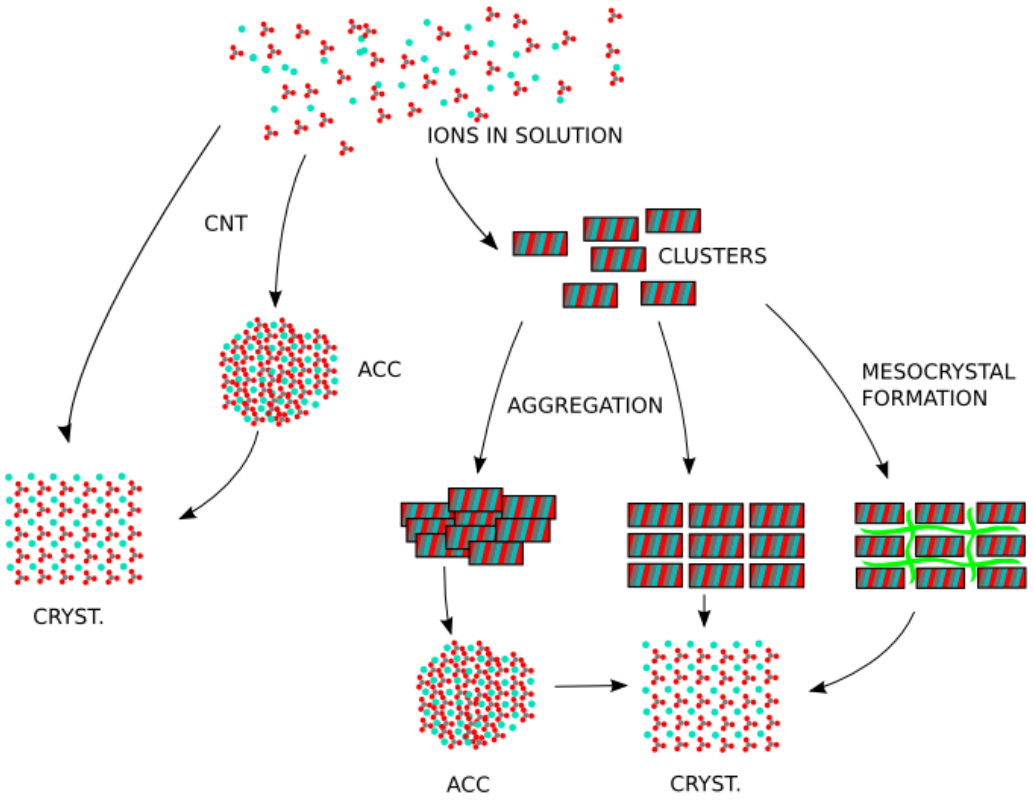

Figure 1.11: Proposed nucleation mechanisms for biominerals. Calcium carbonate is shown, though the principles also apply to other materials such as calcium phosphates.

are the most energetically favoured and which are stable across a range of sizes, and hence, should crystallisation take place at these sizes, which polymorph may be favoured. I have also attempted to understand the phases available to small nanoparticles of tricalcium phosphate.

1.4 Experimental results for calcium carbonate nucleation

Before considering the full complexity of the nucleation and growth of biominerals it is important to understand the simplest possible systems. Here some of the recent *in vitro* experimental evidence is reviewed, focussing on calcium carbonate systems as the more widely-studied.

The work of Gebauer et al. [2008] has proved to be an important milestone in the quest to understand the earliest stages of calcium carbonate nucleation. In this experimental study, Ca^{2+} ions were titrated into a carbonate solution where the pH was maintained at a constant value. By measuring the concentration of free calcium ions and comparing to the known quantity added, the authors were able to show that even at concentrations far below the point of saturation, a constant percentage of the present Ca^{2+} ions were no longer free in solution. This percentage varied as a function of pH and hence with the carbonate/bicarbonate ratio. Higher pH, and hence more carbonate, was associated with greater percentage of bound calcium, suggesting calcium carbonate clusters were forming. Analysis using ultra-centrifugation suggests these clusters were small, at around 2 nm diameter, or 70 formula units in size, although this is an average, and the size distribution is not known. A Ca^{2+} ion selective electrode was used to measure the free ion concentration.

As well as studying the pre-nucleation clusters, the authors also studied the nucleation product as calcium concentration continued to rise. They argue that as they could not detect clusters following nucleation, and as the cluster size was observed to increase with supersaturation, that nucleation proceeds through cluster aggregation rather than the classical model of single ion growth.

A considerable degree of supersaturation was necessary to induce nucleation, but once begun it proceeded rapidly. A further insight gained was that the product also varied with pH, with two distinct amorphous phases precipitating with different solubility levels. At relatively low pH (between 9.0 and 9.5) ACC1 was found, which was also found to be relatively stable in comparison to ACC2, which precipitated at pH 9.75 and 10.0. Both precipitates eventually transform to the crystalline phases, with the dominant product at pH 9.0 being calcite, while at pH 10.0 the dominant product is vaterite, with some calcite — aragonite was not observed. This suggests that ACC1 is a calcite precursor while ACC2 is a vaterite precursor, though some vaterite was also found in the product of experiments ran at pH 9.4.

The exact nature of ACC1 and ACC2 was later studied by Gebauer et al. [2010] in a paper where the nanoparticles were studied using a wide range of experimental techniques. The particles were found to be around 20 nm in diameter, with a relatively narrow distribution around this size range. While measures of long range order confirmed the generally amorphous nature of the nanoparticles, other techniques showed that the particles has significant short range order, with ACC1

having similar structure to calcite and ACC2 being associated with vaterite. The match is by no means exact, which is partly due to the large amounts of water in the amorphous phases; both phases have approximate stoichiometry $CaCO_3.H_2O$. Despite the strong association between the amorphous particles and bulk phases, it was possible to drive the crystallisation of the phases to either bulk polymorph. Heating to 350 °C induces both phases to crystallise to calcite, while intriguingly, the rapid spinning involved in solid-state nuclear magnetic resonance (ss-NMR) was found to induce both phases to crystallise to vaterite.

It might be argued that the rapid mixing inherent in a technique such as this, combined with the presence of solvent/surface interactions, might influence the nucleation mechanism. Wolf et al. [2008] studied calcium carbonate crystallisation in acoustically levitated water droplets, where the only interface in the system was the water/air boundary. An aqueous solution of calcium bicarbonate of about 4 μl was suspended, and as water and carbon dioxide evaporated the degree of supersaturation in the solution gradually rose until eventually only a calcium carbonate precipitate remained. Thus the rapid mixing of other techniques is avoided. The precipitation (and as it turned out, nucleation and crystallisation) could be studied in situ using wide angle x-ray scattering (WAXS), and at different stages the droplet was studied using electron microscopy techniques. They were able to show that at neutral pH, and in conditions with low mixing conditions, amorphous particles of up to 300 nm were observed, and when wet were found to be stable for periods up to several days. Calcite nucleation eventually occurred at multiple sites at the air/ACC interface, with much of the calcium carbonate remaining amorphous.

The work of Pouget et al. [2009] also involved the study of calcium carbonate clusters, in this case on a stearic acid template. Using cryo-TEM (transmission electron microscopy), they were able to directly visualise both the clusters and larger crystalline nanoparticles which nucleated at the interface (figure 1.12.

Together, these papers suggest a coherent picture of additive-free calcium carbonate crystallisation and growth. Calcium and carbonate ions will associate in solution into liquid clusters even at very low concentrations. Nucleation occurs from these clusters into hydrated amorphous particles, which may have significant local structure relating to a bulk polymorph. Crystallisation is likely to proceed to the corresponding bulk polymorph, though this is by no means certain.

In a biological context there is scope for manipulation of this process at every step. There is potential for the confinement and unique surfaces of biological sys-

tems, combined with the properties of proteins and other agents, to alter the nature of the clusters, their mechanism of aggregation, the composition and structure of the resulting ACC nanoparticles and the method and timing of transition to the crystalline phases.

1.5 Simulation results for calcium carbonate nucleation

The use of molecular simulation to study the nucleation of biomaterials has a long history, with most effort once again focusing on calcium carbonates. Here the work related to nanoparticles and clusters in solution is reviewed, though there has also been a wealth of work conducted studying nucleation on surfaces, such as that of Freeman et al. [2008] and Quigley et al. [2009].

de Leeuw and Parker [1998] used molecular simulation to calculate the surface energies of calcite, aragonite and vaterite, both for fully hydrated and vacuum systems, with the vaterite using a fully-ordered orthorhombic structure. They used these data to make Wulff construction [Wulff, 1901] estimates of nanoparticle morphology, though they did not make the final step of actually constructing nanoparticles, which, as explored in chapter 6, can be problematic on the nano-scale. For calcite the $(10\bar{1}4)$ surface was found to dominate both in water and vacuum, in good agreement with experiment. For aragonite and vaterite however, the agreement with experiment was much less satisfactory.

The first simulation of calcium carbonate nanoparticles was that of Kerisit et al. [2005]. While chiefly focussing on surface/water interactions, the authors also studied a calcite particle in water. At just 24 formula units, the particle rapidly distorted, though it did not entirely lose its ordered character over the 0.4 ns course of the simulation.

Cooke and Elliott [2007] took this work further, constructing larger calcite nanoparticles based on Wulff constructions at larger sizes and testing their stability both in vacuum and water. It was found that for this work, water stabilised the order in the nanoparticles, with particles as small as 75 units retaining their shape and order over 1 ns.

Quigley and Rodger [2008a] used metadynamics (Laio and Parrinello [2002] — see also section 2.7) to explore the conformations and phases available to a nanoparticle of such a 75 unit particle in water. This technique, which biases a set of collective variables describing the state of the system, recovers the free energy

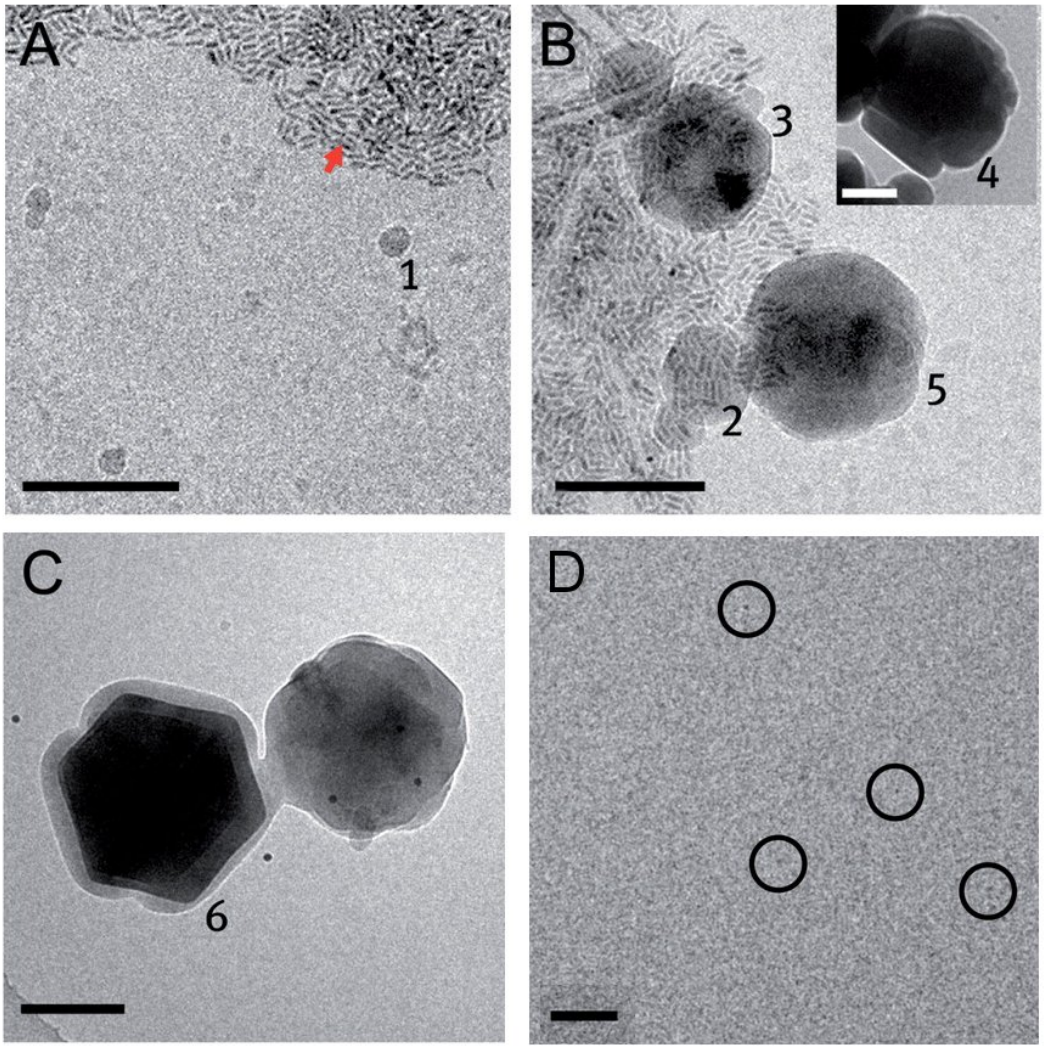

Figure 1.12: **A-C**. Cryo-TEM pictures of calcium carbonate super-solution after 6, 11 45 minutes respectively, from Pouget et al. [2009]. Highlighted particles 1 and 2 are amorphous, 3, 4, and 5 are polycrystalline, and 6 is a single vaterite crystal. Scale bar 200 nm. **D** Fresh solution from the same work with clusters, near detection limit of apparatus, highlighted. Scale bar 20 nm.

surface with respect to those collective variables. In this case the collective variables used were the Steinhardt parameters (Steinhardt et al. [1983], section 2.7), derived from spherical harmonics. Later work [Quigley et al., 2011] repeated these simulations for a variety of nanoparticle sizes, sampling from a more relevant (N, P, T) ensemble (see section 2.2), so the results of this later paper are described here. For both 75 and 192 formula unit nanoparticles, in pure water the free energy surface had three minima: metastable amorphous and vaterite, and also a clear minimum corresponding to calcite, exposing the $10\bar{1}4$ surfaces of that phase (figure 1.13). At the larger particle size the calcite minimum was separated by several hundred k_BT from the other states, while for 75 formula units calcite was around 80 k_BT lower than the amorphous state. For a 300 formula unit nanoparticle the recovered free energy surface did not possess the vaterite-like metastable minimum, with only amorphous and calcite configurations present.

A different study by the same authors [Freeman et al., 2010] extended this work also studied the action of the egg-shell protein ovocleidin-17 on calcium carbonate nanoparticles. For a 192 formula unit nanoparticle, the free energy surface changed dramatically, when attached to the protein, with calcite become considerably more energetically favourable. This suggests that by binding ACC nanoparticles, the protein catalyses the calcite/ACC transformation. Furthermore the authors discovered that the protein would not stay bound to a larger 300 formula unit nanoparticle, and postulated a catalytic cycle. This would involve the protein attaching to a small ACC nanoparticle, which then transforms to calcite. This particle then grows to a larger size, where calcite is more stable, before the protein detaches. A cycle such as this can explain how a small amount of protein is able to modify the nucleation and growth of vastly more calcium carbonate.

Tribello et al. [2009] studied the behaviour of supersaturated free ions in solution. They found that even over 20 ns, clusters would rapidly appear, initially driven by electrostatic forces, then simple diffusion. They found that their clusters did exhibit some vaterite like local order. They also conducted umbrella sampling looking at single ion growth of calcite and amorphous particles.

The accuracy of force-fields is crucial to the accuracy of any results acquired. The development of force-fields for the study of hard/soft interfaces is a particular challenge, and developments have ran in parallel with the simulations described above. For the hard, mineral terms, most work considered above has used the potential of Pavese et al. [1992] including later modifications (Pavese et al. [1996]).

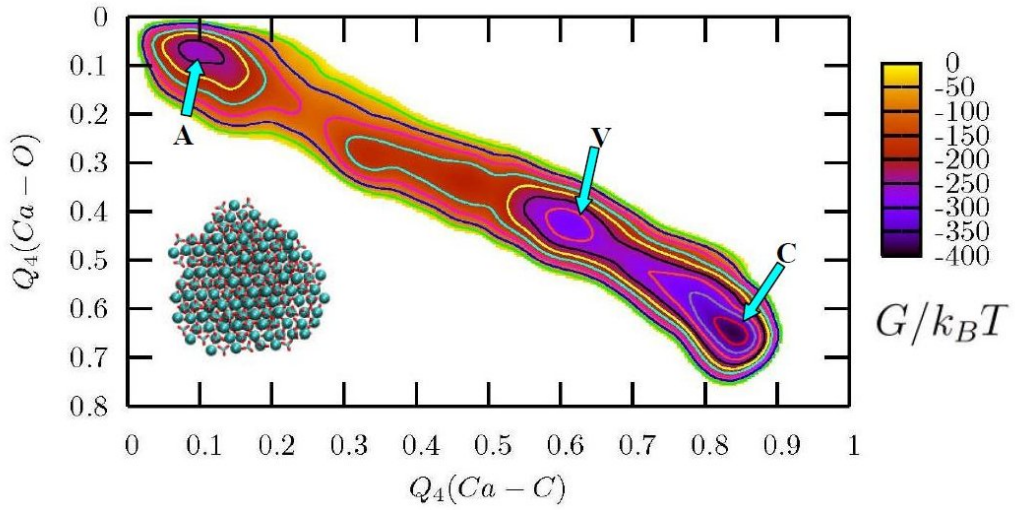

Figure 1.13: Gibbs free energy surface recovered for a 192 unit nanoparticle by Freeman et al. [2010]. The three minima highlighted are dry amorphous (A), vaterite-like (V) and calcite (C).

The water model and cross terms have evolved over time, with work during and after 2007 using the force-field of Freeman et al. [2007]. Tribello et al. [2009] however used a force-field produced by the authors [Bruneval et al., 2007].

The work of Raiteri et al. [2010], was significant for an in-depth study of these calcium carbonate force-fields used and their limitations. For the mineral terms, most prior work had fitted force fields to unit cell measurement, mechanical properties, and phonon modes. They persuasively argued that where phase behaviour is being studied, it is thermodynamic properties of a force-field that are most important. They therefore fit the mineral terms to the enthalpy difference between calcite and aragonite as well as the usual cell and mechanical parameters. By using a lattice-switching Monte Carlo algorithm to recover the phase diagram, they were able to show that this potential gives much better agreement with experiment than force fields such as that of Freeman et al. [2007].

Furthermore, they fitted mineral/water interactions by fitting the hydration free energy of the Ca^{2+} and carbonate ions. This was done using free energy perturbation, which involves gradually switching on an ion in a water box. There is no requirement for the configurations involved along the pathway to be physically realistic in this method. By fitting the hydration free energy accurately, they were able

to recover the experimental dissolution enthalpy exactly, though the exact agreement is likely coincidental. This is in contrast to the other force-fields tested, which get this property wrong by an order of magnitude, and a sign.

The extent to which prior work needs to be repeated in light of these results remains controversial: results such as the rapid aggregation on simulation timescales seen by Tribello et al. [2009] would be unlikely to occur to the same extent if water molecules bind the mineral ions so much more strongly, and as discussed in chapter (7), the stabilisation of mineral nanoparticles by water seen by Cooke and Elliott again appears to be a result of the potential used. The absence of aragonite in the Quigley and Rodger metadynamics simulations may also be explained by the potential used, which fails to model even bulk aragonite accurately.

In this paper they do not use their new force-field to study nanoparticles, instead considering the calcite $(10\bar{1}4)$ surface, and the structure of the associated water together with individual ion dissolution and adsorption free energies.

In later work however [Raiteri and Gale, 2010] they do consider nanoparticles. The first paper used rigid carbonate ions and water molecules (TIP4P-Ew — Horn et al. [2004]), which while giving excellent agreement for calcite and aragonite was not able to accurately recover the vaterite unit cell of Wang and Becker [2009] due to the carbonate distortions in that structure. The new work, while using the same inter-molecular terms, used harmonic bond and springs, together with an unusual out-of-plane potential, to model the carbonate ion, while using the SPC-Ew model of water [Wu et al., 2006], and gives good agreement with vaterite properties. It is this flexible force-field that has been selected as the calcium carbonate potential for this work.

Beyond the introduction of the new flexible force-field this paper details many nanoparticle simulations. As well as considering calcite, they also consider ACC nanoparticles with varying water content. As found in experiment, wet ACC was found to be energetically stable in comparison to dry ACC. The most stable proportion of water content in the nanoparticles actually grew with particle size.

More surprising was that even for quite large sizes, wet ACC was more stable than even idealized calcite nanoparticles. From free energy calculations they estimate that calcite is only more stable than ACC for nanoparticles bigger than around 3.5 nm or 500 formula units. Moreover, by performing umbrella sampling calculations, they were able to show that single formula unit growth onto a perfect calcite $(10\bar{1}4)$ surface is energetically unfavourable, restricting growth to kinks, edges and

corners. They noted that for all particle sizes and hydration states, for amorphous particles single formula unit growth was always energetically favourable. Thus even if calcite does occur at small sizes further growth is hindered, and where kinetic factors are important, calcite would be out-competed by wet amorphous particles.

The picture of nucleation produced in this paper, is in conflict with the experimental evidence. As detailed in section 1.4 the experimental evidence suggests the existence of liquid pre-nucleation clusters, which have to cross a significant energy barrier to nucleate into larger, solid ACC nanoparticles. The simulation study however is clear on there being no energy barrier to the growth of compact ACC nanoparticles direct from solution, with no region of particular stability for the pre-critical clusters to inhabit. From private communication with Professor Gale and Dr Quigley, who have conducted more work recently on this issue, it appears that the role of bicarbonate is vital. While most published simulation studies are conducted at pH that is effectively infinite, if bicarbonate is explicitly included as a separate species it acts to stabilise chain-like liquid clusters, similar to PILP, which might then nucleate into compact ACC particles.

Nevertheless, even if one accepts that the earliest stages of nucleation are understood, there is much scope for further study even of pure water systems with no biological agents present. Issues such as how hydrous ACC, where much of the water has very limited mobility, transforms to an anhydrous crystalline phase, remain poorly understood.

In this work, the focus on calcium carbonates has been on aragonite and vaterite. I am not aware of any prior work undertaken to simulate nanoparticles of these polymorphs, with most effort focusing on calcite and ACC as discussed above. The energetic differences between the bulk phases are small, and the differences between them might be very different on the nano-scale. Thus the study of the idealized morphology of aragonite and vaterite nanoparticles has been undertaken. The thermodynamic stability of these nanoparticles both in vacuum and solvent has also been studied. I have also studied transitions between the three crystalline phases, both as an end in itself and as a means of evaluating order parameters. New methods of characterising the output of nanoparticle simulations have also been developed, and extended to produce new continuous order parameters for metadynamics.

Calcium phosphates have also been studied, repeating the metadynamics work on calcium carbonate [Quigley and Rodger, 2008a] on these perhaps neglected compounds. Focusing on tricalcium phosphate as a simpler subsystem, the extant

force fields have been reviewed, with the most appropriate selected for use in metadynamics to try and understand the thermodynamic phases available to nanoparticles of this substance.

Chapter 2

Molecular Simulation

2.1 Introduction

Molecular simulation has become an important tool in the study of matter in all its forms, from studies of water at standard temperature and pressure [Wu et al., 2006] to the exotic materials at the surfaces of neutron stars [Horowitz et al., 2005]. This chapter gives a brief overview of several aspects of the field, particularly those techniques employed in this work. It begins with a brief summary of the statistical mechanics which underpin molecular simulation. The classical potentials employed to model atomic interactions are then reviewed, before a consideration of the molecular dynamics (MD) and Monte-Carlo (MC) techniques. After a short discussion of energy minimisation the accelerated sampling technique meta-dynamics is reviewed.

2.2 Classical Statistical Mechanics

In this section I briefly summarise some basic statistical thermodynamics results in order to motivate and clarify what follows. This treatment is largely based on that given in Frenkel and Smit [2002].

2.2.1 Ensembles

For a set of N particles, let the $3N$-vectors of the positions and momenta of all particles be denoted as $\mathbf{r}^N$, and $\mathbf{p}^N$ respectively. Together $\mathbf{r}^N$, and $\mathbf{p}^N$ describe every aspect of the state of the system, and the attainable values describe a *phase space*.

Each unique point on the phase space is defined as a *microstate*. The *Hamiltonian* $\mathcal{H}$, is a function of the microstate that evaluates the total energy of the system.

An *ensemble* can be defined by a function f on phase space giving a probability density to each microstate. Hence:

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

The following section briefly reviews some of the more common ensembles.

The microcanonical (N, V, E) ensemble

This is the set of all microstates with constant number of particles N, volume V, and energy E, and corresponds to a closed, thermally isolated system. All such microstates are equally likely, and hence:

$$f(\mathbf{r}^N, \mathbf{p}^N) = C\delta[\mathcal{H}(\mathbf{r}^N, \mathbf{p}^N) - E] \tag{2.2}$$

The normalisation constant C must have this value to ensure that equation 2.1 holds:

$$\frac{1}{C} = \Omega(N, V, E) = \int \int \delta[\mathcal{H}(\mathbf{r}, \mathbf{p}) - E] \mathrm{d}\mathbf{r}^N \mathrm{d}\mathbf{p}^N \tag{2.3}$$

This quantity $\Omega(N, V, E)$ is known as the partition function.

The canonical (N, V, T) ensemble

The canonical ensemble corresponds to the microstates of a system which can exchange heat with its environment. The system therefore has constant number of particles N, volume V, and temperature T. It can be derived that the probability density function is given by:

$$f(\mathbf{r}^N, \mathbf{p}^N) = \frac{1}{Q} \exp\left(\beta \mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)\right) \tag{2.4}$$

Here common notation is followed, with $\beta = \frac{1}{k_B T}$ where k_B is Boltzmann's constant. Again a normalisation constant $Q(N, V, T)$ is necessary, and again is known as the partition function for this ensemble. It is given by:

$$Q(N, V, E) = \int \int \exp\left(\beta \mathcal{H}(\mathbf{r}^N, \mathbf{p}^N)\right) \mathrm{d}\mathbf{r}^N \mathrm{d}\mathbf{p}^N \tag{2.5}$$

Other ensembles

The isothermal-isobaric (N, P, T) ensemble has constant N, pressure P and temperature T and represents a system connected by a freely moving piston to its environment. The grand canonical ensemble (μ, P, T) has constant chemical potential μ, volume V, and temperature T, and represents an open system, exchanging not just heat or volume with its environment, but particles.

2.2.2 The Ergodic Hypothesis

The ensemble average of a quantity can give useful information about the state of the full system. For a function $B(\mathbf{r}^N, \mathbf{p}^N)$ of the microstate it is defined as:

$$\langle B \rangle_{ens} = \int \int B(\mathbf{r}^N, \mathbf{p}^N) f(\mathbf{r}^N, \mathbf{p}^N) \mathrm{d}\mathbf{r}^N \mathrm{d}\mathbf{p}^N \tag{2.6}$$

An example of such a function B might be the kinetic energy temperature estimator, or a pressure estimator.

The time average of such a quantity is defined as:

$$\langle B \rangle_t = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau B(\mathbf{r}^N(t), \mathbf{p}^N(t)) \mathrm{d}t \tag{2.7}$$

The ergodic hypothesis is that these two quantities are equal: that a sufficiently long trajectory will sample all states in proportion to their weight within the ensemble. It is this hypothesis that underpins the entire field of molecular dynamics (see section 2.4).

2.3 Classical Potentials

The forces that govern how atoms interact are fundamentally grounded in quantum mechanics and the Schrödinger equation. Techniques such as density functional theory (DFT) are used to perform molecular simulation at the quantum level, and the use of such *ab initio* methods is certainly desirable where computationally practical, though they still entail the use of approximations and simplifications. The use of classical, empirically derived potentials is vastly less computationally expensive and can still give good agreement with experimental results. Such potentials have a number of parameters which may be fitted to *ab initio* calculations, experimental

results, or a combination of the two. In this thesis published empirical force fields for calcium phosphates are reviewed in chapter 3. For calcium carbonates the force-field of Raiteri and Gale [2010] has been chosen due to its accurate reproduction of thermodynamic quantities. In this section I review some of the terms used in these potentials.

2.3.1 Inter-molecular forces

Van der Waals Potentials

There are a number of standard models of van der Waals forces which are commonly used. Two of the more common are the Lennard-Jones potential (equation 2.8) and the Buckingham potential (equation 2.9).

$$U(r_{ij}) = \frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^{6}} = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^{6} \right] \tag{2.8}$$

$$U(r_{ij}) = A \exp\left(-\frac{r_{ij}}{\rho} \right) - \frac{C}{r_{ij}^{6}} \tag{2.9}$$

For both of these potentials, at longer distances the attractive r^{-6} term dominates, while at very short distances there is strong repulsion due to the other term. Both of these potentials, together with most of the other common van der Waals potentials, tend rapidly towards, but never actually reach, zero with increasing r. Rather than calculate extremely small forces for atoms a long distance apart, it is usual to apply a cut-off to such potentials so that for atoms above a certain distance, the potential is not calculated. This can create an undesirable discontinuity around the cutoff. To compensate for this several different tapering strategies can be employed, such as shifting or scaling the potential function, or applying some scaling only in the range around the cut-off, as in the Buckingham potentials of Raiteri and Gale [2010].

Electrostatics

The Coulomb potential between two charges is given by:

$$U(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \tag{2.10}$$

Where q_i, q_j are the charges on the particles and ϵ_0 is the permittivity of free space. This interaction decays over much longer distances than the Van der Waals terms described above, and cannot typically be truncated without introducing significant errors. For non-periodic systems, such as the nanoparticles in vacuum simulations that form a large part of this thesis, the explicit sum over all pairs of atoms must be calculated if these errors are to be avoided.

For periodic systems the standard method to overcome this difficulty is the Ewald summation [Ewald, 1921]. For a simple cubic periodic system of length L containing N atoms, the full Coulomb energy is given by:

$$E = \frac{1}{4\pi\epsilon_0}\frac{1}{2}\sum_{\mathbf{n}}\sum_{i=1}^{N}\sum_{j=1}^{N}{}' \frac{q_i q_j}{\|\mathbf{r}_{ij} + \mathbf{n}L\|} \tag{2.11}$$

Here the summation over n should be understood to refer to all possible values of $n_1\mathbf{i}+n_2\mathbf{j}+n_3\mathbf{k}$ for integer n_a i.e. the summation over all images of the cell. The $'$ indicates that the summation excludes the case where $i = j$ and $\mathbf{n} = 0$, i.e. the interaction of each atom with itself. Equation 2.11 converges extremely slowly and the result actually depends on the order on which the terms are evaluated. In the Ewald sum procedure the infinite sum over all atom pairs is separated into two separate series, both of which can be readily truncated. The treatment given here is partially derived from those given in Hark and Cai [2009] and also Frenkel and Smit [2002], and can be readily extended to non-cubic periodic systems.

As in figure 2.1B, the 'short range' series is composed of the simple point charges in addition to a Gaussian cloud of charge acting to screen each point charge. Starting at any point and calculating the electric field generated by this combined field, as the distance from the start point increases the net contribution of each point charge/Gaussian pair tends rapidly to zero. The short range interactions thus converge rapidly as a direct sum. More formally, it can be shown that the sum of the short range interactions is given by:

$$E_{SR} = \frac{1}{4\pi\epsilon_0}\frac{1}{2}\sum_{\mathbf{n}}\sum_{i=1}^{N}\sum_{j=1}^{N}{}' \frac{q_i q_j}{\|\mathbf{r}_{ij} + \mathbf{n}L\|}\,\mathrm{erfc}\left(\frac{\|\mathbf{r}_{ij} + \mathbf{n}L\|}{\sqrt{2}\sigma}\right) \tag{2.12}$$

Here erfc is the standard complementary error function given by:

$$\mathrm{erfc}(z) = \frac{1}{2\pi}\int_z^\infty \exp(-t^2)\mathrm{d}t \tag{2.13}$$

Note the similarity between equations 2.11 and 2.12. The distinction between the two is that the error function term in the latter rapidly tends to zero with increasing $\|\mathbf{n}L\|$.

The contribution of the Gaussian clouds (figure 2.1C) must then be considered in the 'long range' series. The problem of calculating the sum of an infinite series of point charge pairs has been transformed in the problem of calculating the infinite series of Gaussian cloud charge pairs. The difference between the two cases is that, as a smoothly varying function, the latter is readily computable in Fourier space. While the derivation runs to several pages, by transforming into this reciprocal space a sum is produced as below:

$$E_{LR} = \frac{1}{2L^3\epsilon_0} \sum_{\mathbf{k}\neq 0} \frac{\exp\left(\frac{-\sigma^2\|\mathbf{k}\|^2}{2}\right)}{\|\mathbf{k}\|^2} \|S(\mathbf{k})\| \tag{2.14}$$

Here $\mathbf{k}$ is analogous to $\mathbf{n}$ with the summation over all possible integer combinations of the reciprocal lattice vectors. $S(\mathbf{k})$ is a term referred to as the structure factor, and is given by:

$$S(\mathbf{k}) = \sum_{j=1}^{N} q_j \exp(i\mathbf{k}.\mathbf{r}_j) \tag{2.15}$$

The terms of the sum in equation 2.14 will tend rapidly to zero with increasing $\|\mathbf{k}\|$ due to the presence of $-\|\mathbf{k}\|^2$ in the exponential term, and hence this summation can again be rapidly truncated.

One minor complication is that implicit in the long range summation was the interaction of each atom with itself. Correcting for this self interaction reduces to calculating:

$$E_{SELF} = \frac{1}{4\pi\epsilon_0} \frac{1}{\sqrt{2\pi}\sigma} \sum_{i=1}^{N} q_i^2 \tag{2.16}$$

The full summation can now be given by $E = E_{SR} + E_{LF} - E_{SELF}$.

Assigning charges

The charges on atoms in a forcefield are parameters to be fitted, particularly for the atoms of molecules where each atom may have a differing partial charge.

To assign all the charge of an atom to a single point is an approximation to the reality of a nucleus surrounded by a fluctuating electron cloud. One method of

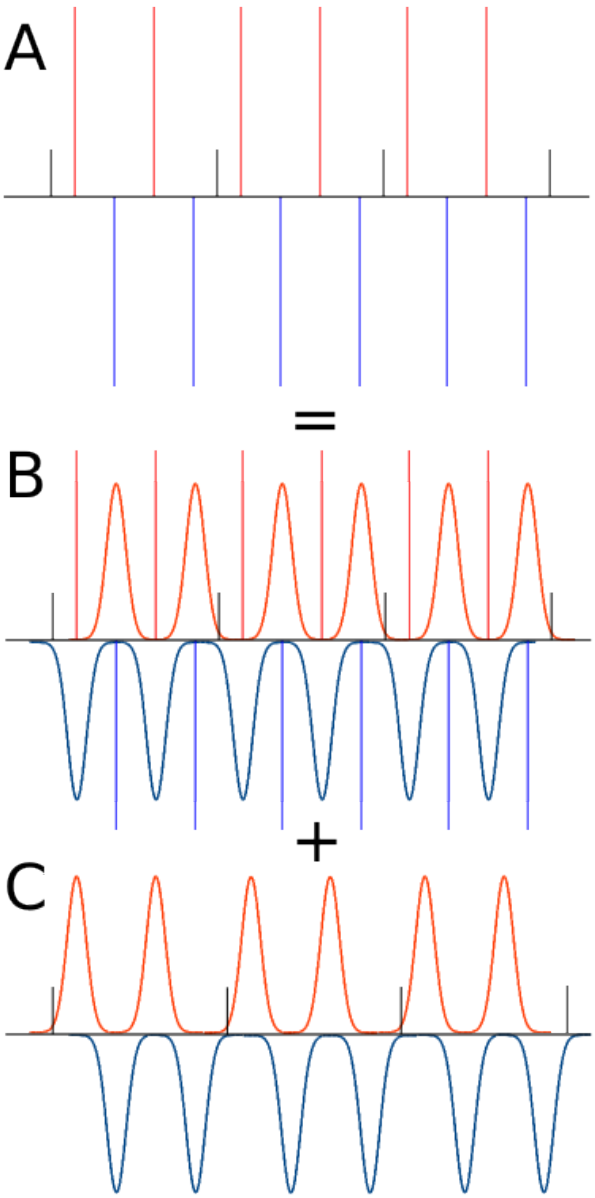

Figure 2.1: Charge distributions on a 1D lattice. The point charge distribution (**A**) is equal to the sum of a short range charge distribution (**B**) where the point charges are shielded by Gaussian clouds of charge, and the long range distribution (**C**), the negative image of those Gaussian charge clouds. Strictly speaking the point charges should be infinite in height as they have zero width, but the same total area as each Gaussian.

incorporating some polarisability is to use a core/shell model of the atom. Two common core/shell models exist for use in molecular dynamics, though both represent a significant increase in computational cost over single-body potentials:

- The relaxed shell model [Dick and Overhauser, 1958] treats the atom as a massive core (nucleus), and a massless shell, which are linked by a simple harmonic spring. The position of the shell is optimised every timestep by balancing the harmonic spring and the local electrostatic force. This requires additional potential evaluations per timestep.

- The dynamic shell model [Lindan and Gillan, 1993] again has a massive core, however in this case the shell is also assigned a small mass, chosen to ensure the natural frequency of the spring will be far in excess of the vibration frequency of the atom in the larger system. This requires the use of a smaller timestep than a one-body system, and hence again more potential evaluations than a single-body system.

2.3.2 Intra-molecular forces

With some exceptions such as the ReaxFF methodology [Chenoweth et al., 2008], employed for calcium carbonates in Gale et al. [2011], atomistic molecular simulation does not usually consider the formation and breakage of bonds which are mostly covalent in nature. Simulations end with the same bonds with which they began, and these are specified explicitly in the input files. The intra-molecular interactions within the molecule may or not include the van der Waals and electrostatic interactions discussed above, and this can introduce difficulty in reproducing force-fields from the literature, as shown in chapter 3.

Two-body terms

Two-body terms generally maintain the inter-atom distance between bonded atoms near a value, while allowing some flexibility. Often a simple harmonic spring is used:

$$U(r_{ij}) = \frac{1}{2}k(r_{ij} - r_0)^2 \tag{2.17}$$

An alternative is the Morse potential:

$$U(r_{ij}) = E_0[1 - \exp(-k(r_{ij} - r_0))^2 - 1] \tag{2.18}$$

This potential requires less energy to stretch the bond than shrink it, allowing a bond distance to smoothly increase with temperature. Similarly, the Lennard-Jones and Buckingham forces discussed in section 2.3.1 can also be used to model bond lengths, although with very different parameters for when they are used for van der Waals interactions.

Three-body terms

Three-body terms typically regulate not inter-atomic distances, but angles, such as the H–O–H angle in a water molecule. Once again, a simple harmonic spring is often used:

$$U(\theta_{ijk}) = \frac{1}{2}k(\theta_{ijk} - \theta_0)^2 \tag{2.19}$$

Four-body terms

The most common four-body potentials again drive inter-atomic angles rather than distances, typically a dihedral angle or improper torsion angle. In the latter case, one set of three atoms are considered to form a plane, and a different set of three atoms form a different plane. The improper torsion angle is the angle between the two planes. Harmonic springs are again often used to model such improper torsion potentials:

$$U(\psi_{ijkl}) = \frac{1}{2}k(\psi_{ijkl} - \psi_0)^2 \tag{2.20}$$

Four-body potentials which act on distances rather than angles are used. The calcium carbonate potential used in this thesis [Raiteri and Gale, 2010] includes an out-of-plane potential. Here the three oxygen atoms of the carbonate are considered to form a plane, and the distance of the central carbon atom from this plane is calculated. This is used as below to drive the carbonate group to maintain its planer form:

$$U(r_{i-jkl}) = k_1 r_{i-jkl}^2 + k_2 r_{i-jkl}^4 \tag{2.21}$$

2.4 Molecular Dynamics

As detailed in section 2.2, the Ergodic hypothesis is at the heart of molecular dynamics. The assumption is made that an average taken over a sufficiently long trajectory will be equal to the full ensemble average.

2.4.1 Integration Algorithms

Analytical solutions for the equations of motion using typical potentials are unobtainable for all but the simplest of cases. Instead, trajectories must be calculated using a numerical integration method derived from the Taylor series expansion of the position vector $\mathbf{r}^N$ in time, using a small discrete timestep, Δt. The algorithm of choice is almost always the Verlet algorithm, which is time reversible and computationally inexpensive compared to higher order methods. The standard algorithm does not require or generate the velocities $\mathbf{v}$, and so in molecular dynamics simulations it is usual to use either the leap-frog or velocity Verlet variations. Both of these similar algorithms give the same performance as the base algorithm at the cost of additional storage requirements for the velocities. The integration over each timestep is a four stage process. This is the velocity Verlet algorithm:

$$\mathbf{v}\left(t + \frac{\Delta t}{2}\right) = \mathbf{v}(t) + \frac{1}{2}\mathbf{a}(t)\Delta t \tag{2.22}$$

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}\left(t + \frac{\Delta t}{2}\right)\Delta t \tag{2.23}$$

$$\mathbf{a}(t + \Delta t) = -\frac{\nabla U(\mathbf{r}(t + \Delta t))}{\mathbf{m}} \tag{2.24}$$

$$\mathbf{v}(t + \Delta t) = \mathbf{v}\left(t + \frac{\Delta t}{2}\right) + \frac{1}{2}\mathbf{a}(t + \Delta t)\Delta t \tag{2.25}$$

Here the division by mass at the third stage should be understood to correspond to term by term division of the force on each particle by its own corresponding mass.

2.4.2 Molecular Dynamics in different ensembles

In the microcanonical or (N, V, E) ensemble the velocity Verlet algorithm proceeds exactly as above, though the velocities are usually scaled at the start of the simulation to give approximately the correct initial temperature. If the timestep is sufficiently short energy will be conserved over the time course of the simulation.

In a canonical (N, V, T) simulation however the velocities of the particles must be modified in order that the time-average of the temperature is correct. It is not sufficient to simply scale all the velocities so that the temperature is exactly accurate — temperatures fluctuate stochastically around the average with known variance, and this behaviour must be produced. Typically the velocities are coupled to a thermostat, requiring subtle alterations to the integration scheme. The Nosé-

Hoover thermostat is a commonly used example [Nosé, 1984][Hoover, 1985].

In an isothermal-isobaric (N, P, T) ensemble these techniques are extended to also allow variation in the volume based on the current estimate of pressure, to ensure that over a reasonable time scale both the average pressure, and the variance in this quantity, are at the correct value.

MD simulations in the grand canonical (μ, P, T) ensemble are very rare. It is difficult to maintain continuous average forces while the number of particles varies discretely.

2.5 Monte Carlo

Rather than calculating trajectories, Monte Carlo methods sample an ensemble directly, with the number of times an area of phase space visited in proportion to its probability in the underlying ensemble. This is done as a stochastic (hence Monte Carlo) process, with a Markov chain of configurations, with each link generated by a perturbation of the preceding configuration.

2.5.1 Move Selection Criteria in the canonical ensemble

Monte Carlo schemes are most straightforward in the (N, V, T) ensemble — it is this ensemble for which the earliest and still most common techniques were developed [Metropolis et al., 1953]. In this section I provide an overview.

A Markov chain is defined by the transition probabilities between each state. In the continuous case these can be specified by a probability density function. A suitable choice of probability density function for the transitions ensures the existence of a stationary probability density function for the states — a single function describing how likely the chain is to be in a region of phase space at an arbitrary point along the chain. A sufficient condition for the existence of such a stationary probability density function π is detailed balance. If, for any two states $\mathbf{r}_i$, $\mathbf{r}_j$, with transition probabilities $p_M(i \to j)$, $p_M(j \to i)$, detailed balance holds if a valid π can be constructed such that:

$$\pi_i * p_M(i \to j) = \pi_j * p_M(j \to i) \tag{2.26}$$

The goal is to ensure that the stationary probability density corresponds to the probability density function of the (N, V, T) ensemble. Neglecting the contribution

of the momenta, that density function is given by:

$$f(\mathbf{r}_i) = \frac{\exp(-\beta U(\mathbf{r}_i))}{Q(N,V,T)} \tag{2.27}$$

By dividing the transition probability into the product of $p_G(i \to j)$, the probability of generating the move from $\mathbf{r}_i$ to $\mathbf{r}_j$, and $p_A(i \to j)$, the probability of accepting the move then becomes:

$$\frac{\exp(-\beta U(\mathbf{r}_i))}{Q(N,V,T)} * (p_G(i \to j)p_A(i \to j)) = \frac{\exp(-\beta U(\mathbf{r}_j))}{Q(N,V,T)} * (p_G(j \to i)p_A(j \to i)) \tag{2.28}$$

The key insight of the Monte Carlo algorithm is that when transitioning between states it is not necessary to fully evaluate the partition function — it is only necessary to calculate the ratio between the probability density function of two states. Rearranging gives:

$$\frac{p_A(i \to j)}{p_A(j \to i)} = \frac{p_G(j \to i)}{p_G(i \to j)} \exp\left(-\beta(U(\mathbf{r}_j) - U(\mathbf{r}_i))\right) \tag{2.29}$$

This equation can be further simplified by ensuring that the move generation probabilities are chosen such that $p_G(i \to j)$ is equal to $p_G(j \to i)$. The only remaining task is to choose acceptance probabilities that balance the equation. While other choices are possible, the most common form is:

$$p_A(i \to j) = \begin{cases} 1 & \text{if } U(\mathbf{r}_j) - U(\mathbf{r}_i) < 0 \\ \exp\left(-\beta(U(\mathbf{r}_j) - U(\mathbf{r}_i))\right) & \text{if } U(\mathbf{r}_j) - U(\mathbf{r}_i) > 0 \end{cases} \tag{2.30}$$

Thus MC simulations proceed by generating a move, evaluating the energy difference of that move, definitely accepting the move if it lowers the energy, and accepting it with finite probability if it increases the energy.

Often situations arise where the probability of accepting a move is very low as most of the moves entail a large increase in energy. Sampling can sometimes be accelerated during a MC simulation by applying a configuration bias. Rather than choosing moves so that the probability of selecting the move is the same as for the backwards move, one chooses moves such that the probability of generating the forwards and backwards move can be calculated explicitly. These quantities can then be substituted into equation 2.30, and detailed balance preserved. Such a configuration bias MC (CBMC) scheme is used in the Ion Site Monte Carlo (ISMC)

code described in chapter 6.

2.5.2 Monte Carlo Sampling in other ensembles

The same principles are followed when sampling other ensembles: the full partition function is not evaluated, as it is only necessary to calculate the ratio of the probabilities of two states. Unlike molecular dynamics, the grand canonical ensemble can be sampled by Monte Carlo techniques, using the additional move type of particle creation or deletion.

2.5.3 Parallel Tempering

In recent years the adaptation of molecular simulation techniques to parallel computer architectures has been vital to their continued development, enabling the study of ever larger systems. For MD codes this has chiefly precoded by dividing up the calculation of the forces among cores, and this can be done reasonably efficiently. In MC codes the moves inherently proceed one at a time, and the energy change resulting from a single move cannot typically be sped up significantly through parallelisation. An alternative technique is parallel tempering (reviewed extensively by Earl and Deem [2005]). Here different processes run the same MC simulation at different temperatures. At random intervals, an additional move type of replica exchange is attempted. This entails two processes swapping the entire state of the system, with finite probability:

$$p = \min\left(1, \exp\left[\left(\frac{1}{k_B T_i} - \frac{1}{k_B T_j}\right)(E_i - E_j)\right]\right) \tag{2.31}$$

Thus if the higher temperature system has lower energy the move is always accepted, else the move is accepted with finite probability. Simulations at higher temperature will sample configurations more rapidly as more moves will be accepted, crossing energy barriers which are in practice insurmountable to lower temperature simulations. By allowing periodic swaps between the low and high temperature simulations, the low temperature simulation can gain access to the other side of these barriers, accelerating the sampling. Parallel tempering can also be employed in MD simulations through replica exchange MD (REMD) [Sugita and Okamoto, 1999]. As with the CBMC technique, parallel tempering is employed in this work in the ISMC technique described in chapter 6.

2.6 Local energy minimisation methods

Many tasks in molecular simulation involve relaxing an atomic configuration, or finding the local minimum in the potential energy landscape. For example the crystal structures of TCP are relaxed using various force-fields in chapter 3, and in chapter 6 the use of local minimisation in random structure searching is discussed.

For non-periodic systems such as nanoparticles in vacuum, I have used the conjugate gradient method, requiring only the first derivatives of the potential function (i.e.. the forces) to perform energy minimisations. This was based on a program written by David Quigley, using his interface to the MD program DL_POLY [Smith et al.]. The program GULP (General Utility Lattice Program)[Gale, 1997] has been used to find local minima in periodic systems. By default the program uses the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method, which uses an approximation to the Hessian of the potential field [Press et al., 2007].

2.7 Metadynamics

A central limitation of the basic molecular simulation techniques discussed above is one of time scale. MD simulations of the order of milliseconds have been performed, but nanosecond simulations are much more common. Events such as nucleation, which may appear to occur rapidly from the standpoint of unaided human perception, are rare events on such a time scale, particularly as the size of a simulation is so much smaller than many real problems of interest. The Ergodic hypothesis only applies in the limit of long sampling, which is often completely impractical

First introduced by Laio and Parrinello [2002], and extensively reviewed by Laio and Gervasio [2008], metadynamics is a widely used method of accelerating the sampling of rare events. This method proceeds through the use of a history dependent biasing potential, driving the system to explore new configurations. From the amount of bias applied to drive transitions it is possible to produce an estimate of the free energy difference between configurations.

The cornerstone of the technique is the description of the system by a small number, d, of collective variables (CVs). These are numerical descriptors of the state of the system as a function of the atomic positions. I denote the i-th collective variable $S_i(\mathbf{r}^N(t))$. As well as the standard potential $U(\mathbf{r}^N)$, molecular dynamics simulations are carried out with an additional time dependant potential $V_G(\mathbf{r}_N, t)$

given by:

$$V_G(\mathbf{r}_N, t) = w \sum_{[t'=\tau_G, 2\tau_G, ...(t'<t)]} \prod_{i=1}^{d} \exp\left(-\frac{S_i((\mathbf{r}^N(t)) - S_i(\mathbf{r}^N(t'))^2)}{2h_i^2}\right) \tag{2.32}$$

Essentially, the potential is built up by depositing Gaussians of height w at intervals of τ_G, with width parameter h_i for the i-th order parameter. This bias potential gradually drives the system away from configurations it has already visited towards new areas. Furthermore, it can be shown that, where F is the free energy surface (FES) in terms of the CVs:

$$\lim_{t\to\infty} V_G(S_1, ..., S_d, t) \approx -F(S_1, ..., S_d) \tag{2.33}$$

If the simulation could be run for an infinitely long time, there would be no need to run metadynamics at all, and the equation also only holds in the limit of long τ_G, but metadynamics can nevertheless provide a good estimate of the FES under appropriate circumstances.

Collective variables

Choice of CVs is crucial to any metadynamics simulation. The abstraction of a full $3N$ dimensional phase space into a much smaller system is what gives metadynamics its power, but is also its central limitation. If an important 'slow' process occurs orthogonally to the CV phase space, then this process will not be observed. The CVs must also distinguish not merely between the endpoints of a transition, but the set of states along the pathway between them. The CVs must also be smooth and continuous, as the evaluation of the action of the MD integrator requires the partial derivatives of the potential in all the $3N$ dimensions of phase space. Common choices include the coordination number, bond angles in protein folding simulations, and the potential energy, as reviewed by Laio and Gervasio [2008]. Much of the current work is based around the development and evaluation of order parameters: in chapter 5 I describe order parameter derived from graph theory, while chapter 8 describes an approach for evaluating combinations of order parameters without having to produce a full metadynamics implementation. In this work, predominantly concerned with driving polymorph selection and crystallisation, I refer to collective variables as order parameters (OPs) interchangeably, except where the distinction

between the two is particularly important.

It is instructive to review in depth the Steinhardt OPs as example CVs which have already been shown to drive phase transitions in nanoparticles [Quigley and Rodger, 2008a], and are used in the present work to study tricalcium phosphate (chapter 4). First introduced by Steinhardt et al. [1983], beginning with the study of ice freezing [Quigley and Rodger, 2008b] Quigley and Rodger have produced a series of papers with their own continuous version of these quantities. Q_l, where l is usually four or six, is given by:

$$Q_l = \left[\frac{4\pi}{2l+1} \sum_{m=-l}^{l} \left| \frac{1}{4N_n} \sum_{b=1}^{N_b} f_C(r_b) Y_{lm}(\theta_b, \phi_b) \right|^2 \right]^{\frac{1}{2}} \tag{2.34}$$

.The terms of this equation are as follows:

- Index b is of the N_b atom pairs involved in the order parameter. In their work on calcium carbonate Quigley and Rodger have used five distinct order parameter based on Ca—Ca, Ca—C, Ca—O, C—C, and C—O pairs.

- N_n is the number of nearest-neighbours to include in the calculation, typically set to a bulk phase coordination number.

- Y_{lm} is the spherical harmonic function for l and m, with polar coordinates θ_b and ϕ_b set according to arbitrary axes.

- The function f_C is the function chosen to ensure the contribution of each atom pair grows or shrinks smoothly as the distance between them varies. It is given by:

$$f_c(r) = \begin{cases} 1 & \text{if } r \leq r_c \\ \frac{1}{2}\left[\cos\left(\frac{r-r_c}{r_o-r_c}\pi\right) + 1\right] & \text{if } r_c < r < r_o \\ 0 & \text{if } r \geq r_o \end{cases} \tag{2.35}$$

  The inner and outer cutoffs r_c, r_o can be distinct for each OP, and for the calcium carbonates were chosen to fall in the region of low occupancy between the first and second coordination shells of the corresponding radial distribution functions (RDFs) for the bulk phases.

Well-tempered metadynamics

One problem with the accuracy of the recovered FES is that it will only ever be accurate to the same order as the Gaussian height parameter w. Even if by chance, the applied bias V_G matched the FES exactly at some time, as new Gaussians were deposited this would change, producing a FES with unphysical undulations of height of order w. A relatively simple remedy for this problem is the use of well tempered metadynamics [Barducci et al., 2008] where the height of each deposited Gaussian is adjusted according to following formula:

$$w = \omega \exp\left(-\frac{V_G(S_1, ..., S_d, t)}{V_{\max}}\right) \tag{2.36}$$

This restricts metadynamics to cross only those basins of a height less than $V_{\max}$, with the initial Gaussians deposited with height ω. This gives a much smoother convergence of the FES to this height. Typically, as in chapter 4, a standard metadynamics trajectory is performed and the important minima identified, and then a well-tempered metadynamics trajectory will be performed with $V_{\max}$ selected at a suitable value to cross the important energy barriers.

Data analysis

When analysing the results of a metadynamics simulation the challenge is to turn a list of deposited Gaussians and their positions in CV-space into a recovered estimate of the free energy surface. In the graphs produced for this thesis, this has been performed by a MATLAB program written especially for the task, working for up to six CVs.

Where the number of CVs is larger than the dimension of the desired FES, it is also necessary to integrate out the action of some of the collective variables. This is done by transforming the FES into a probability density function, integrating out one or more CV, then transforming back to a FES. More formally, for a FES $F(S_1, ..., S_d)$, in units of $k_B T$, it can be shown that:

$$p(S_1, ..., S_d) = \exp\left(-F(S_1, ..., S_d)\right) \tag{2.37}$$

$$p(S_1, ..., S_{d-1}) = \int p(S_1, ..., S_d) \mathrm{d}S_d \tag{2.38}$$

$$F(S_1, ..., S_{d-1}) = -\ln p(S_1, ..., S_{d-1}) \tag{2.39}$$

This has been implemented locally using MATLAB, with the integration carried out numerically. Two-dimensional FES plots produced in this work are produced using a contour plot.

Chapter 3

Potentials for tricalcium phosphate

3.1 Introduction

Any fully atomistic simulation is only as good as the potential model used. These potential models are typically created by fitting parameters to any of a number of standard functions, some of which are discussed in the previous chapter. For the study of calcium phosphates, a literature search has revealed five such potentials models, all focused on apatites ($Ca_{10}(PO_4)_6X_2$ where X might be any ion with charge -1 such as fluorine or a hydroxyl group). Each of these potential models has been fitted either to experimental values such as crystallographic parameters and elastic constants, results of *ab initio* calculations, or a mixture of the two.

This chapter reviews these potential models, giving full details and applying them to the published structures of the α and β polymorphs of bulk tricalcium phosphate [Mathew et al., 1977][Yashima et al., 2003]. I have evaluated the force-fields on their ability to accurately model the unit cell parameters as well as mechanical properties. For the latter I have used the values derived by Liang et al. [2010] through an *ab initio* study. As extensively discussed in that paper, the extant experimental studies of TCP have used porous samples with mixed compositions, and hence the DFT simulations presented represent a more accurate picture of the behaviour of a hypothetical perfect TCP crystal of either phase. The aim is to select the most appropriate potential for work studying the properties of tricalcium phosphate nanoparticles.

Potentials have been implemented in the program GULP 3.1 [Gale and Rohl, 2003]. Usefully, this program calculates predictions of properties such as the bulk modulus as well as the phonon modes, from which an indication of structural stability can be obtained. β-TCP has six Ca sites within the unit cell which have 50% partial occupancy. These sites are equally spaced along the c-axis. Following Yin et al. [2003], Pedone et al. [2007] and Liang et al. [2010] I used 100% occupancy at sites 2, 4 and 6 (figure 1.9, following the nomenclature of Yin et al. [2003]) and 0% occupancy at the other sites, thus evenly spreading the charge through the unit cell. As a check on this procedure, I also tested each combination of sites using the potential of Mkhonto and de Leeuw [2002] and found that this combination gave the joint lowest potential energy (100% occupancy for sites 1, 3 and 5 appears to be equivalent).

3.2 Meis et al. [2000] Potential Model

3.2.1 Details of Potential

In 2000 Meis et al. [2000] published the first potential model for fluorapatite, performing the fitting using the program GULP, written by one of the authors [Gale, 1997]. The purpose of this research was to investigate the suitability of fluorapatite for use in nuclear waste disposal. Plutonium and caesium ions would be incorporated into neodymium-fluorapatite. The potential primarily uses short range two-body Buckingham potentials together with the coulombic interactions. Non-electrostatic cation-cation interactions are set to zero, with only anion-cation, and anion-anion interactions having a Buckingham potential. In addition, to represent the tetrahedral configuration of the PO_4^{3-} group, a three body potential of the form $E_{ijk} = \frac{1}{2}k_{ijk}(\theta - \theta_0)^2$ is used. For the oxygen ions a core-shell model is used (see section 2.3.1). This allows the polarisation of these ions, governed by a spring constant K_P. All non-coulombic forces act on the oxygen ions acted on the shell, and hence only indirectly on the core.

The various parameters of the force field were fitted to give agreement with experimentally derived lattice parameters of synthetic fluorapatite [Boyer et al., 1997], as well as the elastic matrix elements and the bulk modulus [Hearmon, 1956]. Fitting was performed under the additional constraint that some of the two body interaction parameters were taken from other force fields, specifically the values

fixing the attractive, exponential section of the Buckingham potential (the A_{ij} and ρ_{ij} values) for some interactions. The oxygen-oxygen and oxygen-fluorine values come from D'Arco and Islam [1997], while the fluorine-fluorine values come from Catlow [1977]. Table 3.1 lists the parameters of the force field in full.

Buckingham Potential Parameters				
Interaction	A (eV)	ρ ($\AA$)	C (eV $\AA^6$)	Cut-off (Å)
O_{shell} — O_{shell}	22764	0.1490	32.58	12
O_{shell} — P	836	0.3513	0	10
O_{shell} — Ca	1288	0.3316	0	10
O_{shell} — F	198	0.1110	0	10
F — P	697	0.2389	0	10
F — Ca	2031	0.2711	0	10
F — F	1128	0.2753	0	10
Charges				
Entity	Charge (e)			
O_{core}	0.860			
O_{shell}	-2.860			
F_{core}	-1.000			
P_{core}	5.000			
Ca_{core}	2.000			
Three-body harmonic				
Interaction	k (eV rad^{-2})		θ_0 (deg.)	
O_{shell} — P_{core} — O_{shell}	2.840		109.5	
Core-shell spring constant				
O_{core} — O_{shell}	98.67 eV $\AA^{-2}$			

Table 3.1: Potential parameters for fluorapatite in Meis et al. [2000]

3.2.2 Implementation and Validation

Table 3.2 lists the experimental values, used to fit the force-field together with both the calculated values listed in the paper, and the local implementation in GULP. There is a small deviation from the published values in the local implementation, typically in the third or fourth significant figure. In the local implementation the force field was applied to a fluorapatite structure [Hendricks et al., 1932] from naturally occurring fluorapatite, while the published figures were based on the lattice

structure of a synthetic fluorapatite, the precise coordinates of which have only been published in a French PhD thesis [Boyer, 1998]. While applying the same force field to these two structures, and applying GULP's energy minimisation routines will likely reduce the differences between them, it is not surprising small differences remain. Furthermore, the authors do not state which version of GULP they use, but several updated versions have been released since 2000, likely including some changes to the energy minimisation routines. This means that even if using the same structure, some small differences between the two minimisations might be expected. While it cannot be stated with absolute certainty that the potential has been implemented correctly locally, it certainly seems probable, particularly as with the relative simplicity of the potential model, implemented in a version of the same software in the same units, there are few sources of ambiguity where errors could creep in.

3.2.3 Application of Force Field to Tricalcium Phosphates

When applied to TCP in GULP the force field proved to be incompatible with experimental results. Applying the GULP structure minimisation routines to the α-TCP structure using the force field, the resultant structure had negative phonon frequencies, indicating instability. When applied to the more stable β-TCP polymorph (specifically the structure of Yashima et al. [2003]), the force field minimisation stretches a subset of the P-O bonds to unrealistic lengths, effectively breaking the bonds. That finding the local minimum in the energy landscape produces such unphysical results indicates that this forcefield is unsuitable for use with this structure, and hence for further study of TCP.

3.3 Lee et al. [2000] Potential Model

3.3.1 Details of Potential

This potential was developed in order to study surface relaxations in hydroxyapatite, although details for fluorapatite and chlorapatite are also given. Released shortly after Meis et al. [2000], This potential references the results of that work, although any difference in the applicability of the two force fields is not discussed. The two force fields have much in common, with both fitted using GULP and both fitted to short-range two body Buckingham potentials to study van der Waals forces.

Property	Experiment	Calculated (published)	Calculated (local)
$a(\mathring{A})$	9.374	9.3688	9.3670
$c(\mathring{A})$	6.888	6.8694	6.8697
$V(\mathring{A}^3)$	524.2	522.2	522.0
c_{11} (GPa)	166.7	164.6	164.9
c_{13}	65.5	60.3	60.3
c_{33}	139.6	145.5	145.4
c_{55}	66.3	54.7	54.7
Bulk modulus (GPa)	98	91.8	91.8

Table 3.2: Fluorapatite properties from experiment and both published, and locally implemented, calculated values using the potential of Meis et al. [2000]

Both also use a three-body potential on phosphate groups, and the spring constant connecting the core and shell of phosphate oxygen ions. The experimental results used for the fitting were the crystal structure parameters of Hughes et al. [1989], and the vibrational modes of [Farmer, 1974]. They also fit to the elastic constants of Bhimasenachar [1945], which they state are "the only measurements of elastic constants we could find", apparently unaware of the work of Hearmon [1956] or Sha et al. [1994]. Table 3.3 gives all the published details of the potential.

Two features mark this potential out as unique compared to the others the others considered in this chapter:

- It is not fully atomistic. When applied to hydroxyapatite the OH^- group is considered as a single "averaged ion", with net charge -1 e. The mass assigned to this ion is not stated but is presumably the simple sum of the O and H atom masses. It seems curious to make such an approximation while also having a computationally expensive core-shell model for the oxygens on the phosphate, and no justification is given in the text.

- The charges applied to ions are not thedisa formal charges. For the other force fields the charge assigned to the Ca ion is always $+2$ e, while the total sum of the charges of the PO_4 group is always -3 e. For the purposes of applying these potentials to TCP this is extremely convenient — the charge on the $Ca_3(PO_4)_2$ formula units always sums to zero, as is required for straightforward implementation in programs such as GULP. If the X in an apatite $Ca_{10}(PO_4)_6X_2$ formula unit has charge -1 e (as it does in this and all the other force fields considered), these charges are mathematically the only val-

ues that can balance the charges of both apatites and TCP. In this force field however, the charge on the Ca ion is exactly +2.150 e while the sum of the charges on the phosphate group is exactly −3.247 . These values actually fail to balance either apatite (net charge +0.018 e per unit cell) or a TCP formula unit (−0.044 e per formula unit). It may be assumed that the published values have been rounded at some stage.

Buckingham Potential Parameters			
Interaction	A (eV)	ρ (Å)	C (eV Å^6)
P — O_{shell}	983.45	0.3260	
O_{shell} — O_{shell}	712.36	0.0660	25.8
OH — O_{shell}	4.417E+005	0.1990	27.9
OH — OH	23400.1	0.1250	184.3
Ca — O_{shell}	521.46	0.3790	
OH — Ca	10433.9	0.2420	
F — O_{shell}	9.451E+006	0.1630	
F — F	91.19	0.0550	
F — Ca	10433.9	0.2420	
Cl — O_{shell}	10125.4	0.1030	
Cl — Cl	3447.3	0.1350	
Cl — Ca	9.853E+007	0.1430	
Charges			
Entity	Charge (e)		
Ca	2.150		
P	4.345		
O_{core}	1.126		
O_{shell}	-3.024		
OH	-1.000		
F	-1.000		
Cl	-1.000		
Three-body harmonic			
Interaction	k (eV rad^{-2})	θ_0 (deg.)	
O_{shell} — P — O_{shell}	1.686	109.47	
Core-shell spring constant			
O_{core} — O_{shell}	89.34 eV $\mathring{A}^{-2}$		

Table 3.3: Potential parameters for fluorapatite in Lee et al. [2000]

3.3.2 Implementation

Implementation focused on fluorapatite. Implementation of this force field was not entirely straightforward, as important parameters were not published. The cut-off for the Buckingham potentials are not stated, and erring on the side of caution, the large value of $30\mathring{A}$ was used in the local implementation. The exact charges are also unknown as discussed above, although if the differences here are from a rounding error they can only vary within a narrow range. In the local implementation the $0.018e$ excess charge was subtracted from among the 40 Ca, P and O ions of the fluorapatite unit cell evenly. This change of 0.00045 e per ion is consistent with a rounding error in the published charges.

Even allowing for these problems, the deviation between the published and locally calculated values of the cell parameters and phosphate modes is larger than might be wished (around 1%), particularly as the correct structure [Hughes et al., 1989] is being optimised. The differences between the values produced by the local implementation and those published are however dwarfed by the differences between the experimental and published values. This implies that if this force field has not been implemented correctly, a new force field which gives similar agreement with experiment has been implemented, and should give similar performance when applied to TCP.

Property	Experiment	Calculated (published)	Calculated (local)
Unit cell $a(\mathring{A})$	9.3973	9.5153	9.5277
Unit cell $c(\mathring{A})$	6.8782	6.6457	6.6542
c_{11} (GPa)	166.7	152.8	151.1
c_{33}	139.6	134.7	134.1
c_{12}	13.96	56.9	56.1
c_{13}	66.3	60.5	60.2
c_{55}	66.3	33.2	33.1

Table 3.4: Fluorapatite properties from experiment and both published, and locally implemented, calculated values using the potential of Lee et al. [2000]

3.3.3 Application to TCP

As described above, no possible values of the charges, when rounded to the published values, could balance the formula of both the apatites and TCP. In the local

implementation new charges must be derived for application to this compound. I have done this by dividing the excess charge of 0.044 e per formula unit evenly among the 10 P and O ions and subtracting the resulting value from the published charges.

Applying this potential to β-TCP gave a structure with negative phonon frequencies, despite the stability of β-TCP at low temperatures, indicating that this force field is unsuitable for use in the study of β-TCP, and hence unsuitable for the purposes of this work.

3.4 Mkhonto and de Leeuw [2002] force field

3.4.1 Details of Potential

Mkhonto and de Leeuw [2002] published a third apatite model, focusing initially on fluorapatite. In the published paper this model is used to examine the effect of water on the structure and morphology of bulk fluorapatite. This force field, which was fitted using GULP, is part of a larger family of compatible force fields. Earlier force fields were fitted to calcite [Pavese et al., 1996] and fluorite [Catlow et al., 1977], while later extension have covered hydroxyapatite [de Leeuw, 2004] and other apatites [Rabone and de Leeuw, 2006], producing a system that covers a wide variety of compounds. At this stage I focus on the fluorapatite terms, as these contain all the interactions relevant for the study of bulk TCP, and the TCP terms are not modified in later papers. The force field was fitted to the experimental structure obtained by Hendricks et al. [1932], together with the elastic constants of Sha et al. [1994], and the phosphate vibrational modes of Taylor et al. [2001]. The interactions between the PO_4, Ca and F ions are fitted to a mixture of Coulombic interactions and Buckingham potentials, while the geometry of the phosphate groups is governed by a 3 body potential on the O-P-O angles, and a Morse potential for the O-P bond (formula $D((1 - \exp -a(r - r_0))^2 - 1)$). A core-shell model of charge distribution is applied to both the phosphate oxygens and the fluorine ions, although for the phosphate oxygens the spring constant linking the core and shell (507.4 eV$\mathring{A}^{-2}$) is around five times the values used in the other potentials (e.g. 98.67 eV$\mathring{A}^{-2}$ for Meis et al. [2000]), and the difference in charge between the core and shell is lower. The full details of the potential, as applied to fluorapatite, can be found in table 3.5.

Buckingham Potential Parameters			
Interaction	A (eV)	ρ ($\AA$)	C (eV $\AA^6$)
Ca — O_{shell}	1550.0	0.29700	0.00
Ca — F	1272.8	0.29970	0.00
O_{shell} — O_{shell}	16372.0	0.21300	3.47
O_{shell} — F	583833.7	0.21163	7.68
F — F	99731834.0	0.12013	17.02423
Short range cut-off 20 $\AA$			
Charges			
Species	Charge (e)		
Ca	2.000		
P	1.180		
F_{Core}	1.380		
F_{Shell}	-2.380		
O_{core}	1.250		
O_{shell}	-2.050		
Three-body harmonic			
Interaction	k (eV rad^{-2})		θ_0 (deg.)
O_{core} — P — O_{core}	1.322626		109.47
Core-shell spring constant			
O_{core} — O_{shell}	507.4000 eV $\AA^{-2}$		
F_{core} — F_{shell}	101.2000 eV $\AA^{-2}$		
Morse Potential			
Interaction	D (eV)	a ($\AA^{-1}$)	r_0 ($\AA$)
P — O_{Core}	3.47	2.03	1.60

Table 3.5: Potential parameters for fluorapatite in Mkhonto and de Leeuw [2002]

3.4.2 Implementation

Table 3.6 compares the properties of fluorapatite obtained by experiment with both published and locally implemented calculated values. As with the other potentials, there is a small difference between the calculated values published by the authors, and those calculated in the local implementation. The difference could be due to either the rounding errors in the published details of the potential or differences in the version of GULP used. The difference is minimal — in the third or fourth significant figure — and it appears probable that the potential has been implemented correctly.

		Calculated	
Property	**Experimental**	**Published**	**Local**
Unit cell a length ($\mathring{A}$)	9.3700	9.3750	9.3751
c length ($\mathring{A}$)	6.8800	6.8712	6.8712
c_{11} (GPa)	152	150.6	150.8
c_{33}	185.7	176.6	176.8
c_{44}	42.8	53.2	53.2
c_{12}	50	62.8	62.9
c_{13}	63.1	73.6	73.8
Bulk modulus (GPa)	93.4	98.8	87.8

Table 3.6: Fluorapatite properties from experiment and both published, and locally implemented, calculated values using the potential of Mkhonto and de Leeuw [2002]

3.4.3 Application to TCP

Both α and β-TCP are stable at 0 K using the force-field, with positive phonon modes. Table 3.7 shows how the unit cell parameters are reproduced using this potential. Agreement is generally good, with a small amount of shrinkage (about 2%) in each unit cell direction, leading to an overall drop in volume of around 4% for both the α and β-TCP polymorph unit cells. Potential energy calculations show that the β polymorph is 5.35 kJ/Mol more stable than the α-TCP polymorph, consistent with β-TCP being the more stable arrangement.

	α-TCP		β-TCP	
	Experiment	**Calculated**	**Experiment**	**Calculated**
a (Å)	12.887	12.673	10.435	10.265
b (Å)	27.280	27.202	10.435	10.265
c (Å)	15.219	14.916	37.403	37.457
α (deg)	90	90	90	90
β (deg)	126.2	126.3	90	90
γ (deg)	90	90	120	120
Volume (Å^3)	4317.52	4145.24	3527.3	3418.0

Table 3.7: TCP unit cell properties from experiment and simulation

3.5 Simplified Mkhonto and de Leeuw [2002] force field

3.5.1 Details of Potential

While as described in section 3.4 the forcefield of Mkhonto and de Leeuw [2002] gives excellent agreement with the properties of TCP, it has two features which lead to undesirably expensive MD simulations:

- The short-range cut off for the Buckingham interactions between non-bonded atoms is set at the high value of 20 Å. For simulations of periodic systems such as bulk TCP with the interactions calculated under the standard minimum image convention, this means unit cells must be at least 40 Å across in any direction, which leads to systems of around 7500 atoms, much larger than would otherwise be needed.

- The core-shell model of the phosphate oxygens and the fluorine ions of fluorapatite are also computationally expensive to simulate (as described in section 2.3.1).

Computationally expensive potentials do not merely represent an inconvenience — where the intention is to push other aspects of a simulation, such as the long time-scale methods employed, as far as possible, a speed-up factor of 2 could feasibly be the difference between the impractical and the merely long-winded.

It is for these reason that I have investigated the simplification of the Mkhonto and de Leeuw [2002] force field, both through the reduction of the short range cut-off for the Buckingham parameters, and a removal of the core-shell model. The latter

was accomplished by summing of the charges of the core and shell parts of the ion, and directing all forces to act on the core of the combined entity

3.5.2 Reducing the Short Range Cut-Off

The effect of lowering the short range cut-off of the potential through a range of values was investigated, to see whether it would significantly affect the ability of the force field to recreate the experimental data. Table (3.8) shows the effect of reducing the cut-off for fluorapatite. The effect of cutting off the potential at 10 Å rather than 20 is barely measurable up to to 4 significant figures, and the differences are particularly small in comparison with the difference between values calculated in the full force field and experiment. Using a value as low as 8 Å for the cut-off may be justified by these results, however to err on the side of caution a value of 10 Å has been selected. For simulations carried out by using minimum image convention, this reduces the minimum cell volume necessary by a factor of 8. The effect of reducing the cut-off on the TCP unit cells is similarly insignificant.

The other factor to consider when reducing such a cut-off is the discontinuity in the energy introduced when a particle pair move across the cut-off. As shown in section 10.1.1 over the course of standard MD simulations energy is still conserved to at least 5 significant figures when using this shorter cut-off, even for time-steps as long as 4 nanoseconds.

	Calculated values				
Parameter	**20 Å**	**12 Å**	**10 Å**	**8 Å**	**Experiment**
a (Å)	9.375	9.375	9.375	9.375	9.37
b (Å)	9.375	9.375	9.375	9.375	9.37
c (Å)	6.871	6.871	6.871	6.871	6.88
c_{11} (10^{10} Pa)	150.8	150.8	150.8	150.8	152.0
c_{33} (10^{10} Pa)	176.8	176.8	176.8	176.7	185.7
Bulk modulus (GPa)	98.82	98.81	98.81	98.79	93.4

Table 3.8: Effect of lowering the short-range cut-off in the Mkhonto potential on Fluorapatite Properties

3.5.3 Removing the core-shell Model

As described above, the core-shell spring used in the model of the oxygen atom is particularly stiff. A core-shell model gives an atom a larger effective radius, so ideally

treating each atom as a single unit would entail refitting the potential. However, the average oxygen core-shell distance in relaxed TCP polymorphs was measured at just 0.0055 Å, with the largest value observed just 0.007 Å. In such circumstances treating the oxygen atom as a single entity, with the combined charge of the core and shell, and identical values for the Van der Waals interactions, would lead to a minimal additional error. Table 3.9 shows that the error is indeed small — here the full Mkhonto potential, with long cut-off and core-shell model is compared to the short cut-off, simplified model. The differences are small, and as before are dwarfed by the distance between simulation and experiment. The energy difference between the two polymorphs is again similar to the full potential, with β-TCP being 5.41 kJ/mol more than stable, compared to 5.35 kJ/mol. The simplified model also gives positive phonon modes, and gives similar performance to the full potential.

3.6 Hauptmann et al. [2003] Force Field

3.6.1 Details of potential

The fourth potential considered was that of Hauptmann et al. [2003], which was fitted to both experimental results, and the results of quantum-chemical calculations. Unlike the three earlier potentials, many of the parameters of the model were fitted using MD simulations, rather than the energy minimisation employed by GULP. Furthermore, these simulations were conducted at standard temperature and pressure, rather than the effective zero values utilised by default in GULP. Two versions of the force field are presented, with one treating the phosphate ion as a rigid unit, while the other allows flexibility in P-O distances and O-P-O angles. The authors do not discuss the other pre-existing apatite potentials. Unlike the other potential models, they also test the ability of their potential to model the thermal expansion of the apatites.

3.6.2 Implementation

Due to several ambiguities in the paper, implementation proved difficult. For example, the paper does not specify if the intra-molecular terms included the electrostatic force between the atoms. Furthermore, many of the units are specified in kJ/mol, but the molar unit (formula unit or molecule) is not specified. Focussing on the flexible forcefield as the more portable between codes (GULP does not work with

Fluorapatite			
Property	**Experiment**	**Calc (full)**	**Calc (simp)**
a (Å)	9.3700	9.3751	9.3884
c (Å)	6.8800	6.8712	6.8668
c_{11} (GPa)	152.0	150.8	152.6
c_{33}	185.7	176.8	177.3
c_{44}	42.8	53.2	53.3
c_{12}	50.0	62.9	59.0
c_{13}	63.1	73.8	74.7
Bulk modulus (GPa)	93.4	87.8	98.6
α-TCP			
Property	**Experiment**	**Calc (full)**	**Calc (simp)**
a (Å)	12.887	12.673	12.691
b (Å)	27.280	27.202	27.215
c (Å)	15.219	14.916	14.943
α (deg)	90	90	90
β (deg)	126.2	126.3	126.29
γ (deg)	90	90	90
Volume (Å^3)	4317.52	4145.24	4160.12
c_{11} (GPa)	127.4	113.16	114.81
c_{33}	123.0	110.3	111.89
β-TCP			
Property	**Experiment**	**Calc (full)**	**Calc (simp)**
a (Å)	10.435	10.265	10.261
b (Å)	10.435	10.265	10.277
c (Å)	37.403	37.457	37.383
α (deg)	90	90	89.31
β (deg)	90	90	90.4
γ (deg)	120	120	119.98
Volume (Å^3)	3527.3	3418.0	3414.34
c_{11} (GPa)	140.5	150.03	150.73
c_{33}	132.3	136.02	143.58

Table 3.9: Comparison of behaviour of full and simplified Mkhonto potentials

rigid molecules) it was, after some experimentation, deemed most straightforward to exhaustively test every possible variant of the four sources of ambiguity in combination. Overall, 144 combinations were tested. Of these only a minority would converge to any structure without excessive stretching of the phosphate units. The variant that most closely matched the fluorapatite unit cell under zero Kelvin minimisation, was deemed to be the correct implementation. It is this variant that is described in table 3.10, with the units converted to eV for consistency with the other potentials. The inclusion of coulombic interaction between bonded phosphorus and oxygen, gives significantly improved performance compared to the other permutations tested. The Urey-Bradley equation, a 1—3 intra-molecular term, is of the form given in equation 3.1.

$$U_{ik} = \frac{k_{UB}}{2}(r_{ik} - r_0)^2 \tag{3.1}$$

Table 3.11 shows how the local implementation recreates the experimental structure as well as some mechanical properties. While the unit cell properties are adequately recovered, the mechanical properties are wrong by nearly a factor of two, perhaps because they formed no part of the fitting procedure.

3.6.3 Application to TCP

As for fluorapatite, when applied to both α and β-TCP, the potential models the properties of the unit cell reasonably well (table 3.12), and both polymorphs are stable with positive phonon modes. However, once again, the mechanical properties are wrong to a considerable extent. the potential energy of β-TCP is 71.95 kJ/mol lower in energy than α-TCP — a much larger figure than the Mkhonto or Pedone potentials.

3.7 Pedone et al. [2007] force field

3.7.1 Details of potential

Fitted entirely to the results of *ab initio* calculations, the force field of Pedone et al. [2007] is the most recent attempt to produce a classical atomistic forcefield for calcium phosphates of which I am aware. Fitting to a hydroxyapatite structure, inter-atomic forces are once again modelled by Buckingham and coulombic interactions. This potential is unusual in that no intra-atomic forces are explicitly included

Buckingham Potential Parameters			
Interaction	A (eV)	ρ (Å)	C (eV $\mathring{A}^6$)
Ca — Ca	1596695.04	0.1580	11.66
P — P	6231.51	0.3500	360
O — O	181775.15	0.1840	32.4
F — F	98..51	0.4180	63.5
P — Ca	32503.52	0.2540	64.8
O — Ca	494596.29	0.1710	19.44
F — Ca	1265.72	0.2880	27.22
O — P	18968.43	0.2670	108
F — P	649.51	0.3840	151.2
O — F	907.9	0.3010	45.36
Short range cut-off 13 $\mathring{A}$			
Charges			
Entity	Charge (e)		
Ca	2.000		
P	2.600		
O	-1.400		
F	-1.000		
Three-body harmonic			
Interaction	k (eV rad^{-2})	θ_0 (deg.)	
O — P — O	3.540	129.75	
Urey-Bradley			
Interaction	k (eV Å^{-2})	r_0 Å	
O — P — O	3.310	3.54	
Two-body harmonic			
Interaction	k (eV Å^{-2})	r_0 Å	
O — P	26.360	1.27	
Intramolecular coulombic O—P interactions are included. Intramolecular O—O coulombic interactions are not.			

Table 3.10: Parameters of the Hauptmann potential

Property	**Experiment**	**Calculated**
Unit cell a length (Å)	9.3700	9.3412
c length (Å)	6.8800	6.8829
Volume (Å^3)	523.12	519.49
c_{11} (GPa)	152	272.89
c_{33}	185.7	373.77
c_{44}	42.8	75.38
c_{12}	50	106.15
c_{13}	63.1	105.14
Bulk modulus (GPa)	93.4	170.6

Table 3.11: Fluorapatite properties using potential of Hauptmann et al. [2003]

α-TCP		
Property	**Experiment**	**Calculated**
a (Å)	12.887	12.891
b (Å)	27.280	27.423
c (Å)	15.219	15.295
Volume (Å^3)	4317.52	4365.71
Bulk modulus (GPa)	75.9	115.82
c_{11} (GPa)	127.4	219.41
c_{33}	123	179.56
β-TCP		
Property	**Experiment**	**Calculated**
a (Å)	10.435	10.608
b (Å)	10.435	10.608
c (Å)	37.403	37.227
Volume (Å^3)	3527.3	3628.2
Bulk modulus (GPa)	82	153.8
c_{11} (GPa)	140.5	221.63
c_{33}	132.3	324.77

Table 3.12: TCP properties using potential of Hauptmann et al. [2003]

— the phosphate groups are held together only by the coulombic and Van der Waals interactions, and at hot enough temperatures would break apart. A core shell model is used for the oxygen atoms, and fitting was performed using GULP. Table 3.13 gives the full details of the potential.

Buckingham Potential Parameters			
Interaction	A (eV)	ρ ($\AA$)	C (eV $\AA^6$)
P_{core}—$O_{shell}{}^{P}$	1099.3	0.3429	0
$Ca_{core}{}^{I}$—$O_{shell}{}^{P}$	3074.4	0.2882	0
$Ca_{core}{}^{II}$—$O_{shell}{}^{P}$	2899.5	0.2886	0
$Ca_{core}{}^{II}$—$O_{shell}{}^{H}$	948.0	0.3440	0
H_{core} — $O_{shell}{}^{H}$	463.3	0.2058	0
$O_{shell}{}^{P}$—$O_{shell}{}^{P}$	122152.7	0.1529	35.87
$O_{shell}{}^{P}$—$O_{shell}{}^{H}$	846.5	0.2133	0.09
$O_{shell}{}^{H}$—$O_{shell}{}^{H}$	54718.7	0.2533	62.67
Short range interaction cut-off of 15 Å			
Charges			
Entity	Charge (e)		
P	5.000		
Ca	2.000		
H	1.000		
$O_{core}{}^{P}$	1.561		
$O_{shell}{}^{P}$	-3.561		
$O_{core}{}^{H}$	0.494		
$O_{shell}{}^{H}$	-2.494		
Core-shell spring constants			
$O_{shell}{}^{P}$ — $O_{core}{}^{P}$	145.77 eV $\AA^{-2}$		
$O_{shell}{}^{H}$ — $O_{core}{}^{H}$	4389 eV $\AA^{-2}$		
Ca^{I} ions are not linked to OH groups			
Ca^{II} ions are linked to OH groups			
O^{P} ions are part of phosphate groups			
O^{H} ions are part of OH groups			

Table 3.13: Potential parameters for hydroxyapatite in Pedone et al. [2007]

3.7.2 Implementation

Implementation of the forcefield was trivial, as GULP input files were included in the supplementary information of the paper in which the force-field was published.

That there were very small differences between the published results and those locally calculated (table 3.14) supports the hypothesis that changes in the GULP program may have led to the small differences seen in the local implementations of the other potentials.

		Calculated	
Property	**Experiment**	**Published**	**Local**
Unit cell a (Å)	9.425	9.327	9.327
Unit cell c (Å)	6.884	6.951	6.951
Volume (Å^3)	530	524	524
c_{11} (GPa)	140	174	174.1
c_{33}	180	175	174.1
c_{44}	36.2	40	40.0
c_{12}	13	68	68.4
c_{13}	69	71	71.0

Table 3.14: Hydroxyapatite properties from experiment and both published, and locally implemented, calculated values using the potential of Pedone et al. [2007]. Published values are to same number of significant figures as paper.

3.7.3 Application to TCP

The authors apply their potential to both α and β-TCP within the paper, and comment on the excellent agreement with experiment. Naturally, they could not compare the mechanical properties to those calculated three years later by Liang et al. [2010], although as shown in table 3.15, the agreement is fair. Both polymorphs also have positive phonon modes. Using this potential β-TCP is 2.91 kJ/mol lower in energy than α-TCP.

3.8 Simplified Pedone et al. [2007] force field

As with the Mkhonto and de Leeuw [2002] force field the Pedone et al. [2007] force field gives good agreement with the properties of the polymorphs of TCP, yet features a computationally expensive core-shell model. With a full refitting of all parameters outside the scope of this work, I again created a variant of the force field with the oxygen atoms each treated as a single entity with the combined charge of the core and shell. This simplified forcefield gave fair agreement with the unit

α-**TCP**		
Property	**Experiment**	**Calculated**
a (Å)	12.887	12.727
b(Å)	27.280	27.529
c (Å)	15.219	15.304
Volume (Å^3)	4317.5	4289.9
Bulk modulus (GPa)	75.9	81.9
c_{11} (GPa)	127.4	121.7
c_{33}	123	120.7
β-**TCP**		
Property	**Experiment**	**Calculated**
a (Å)	10.435	10.284
b(Å)	10.435	10.284
c (Å)	37.403	38.613
Volume (Å^3)	3527.3	3536.9
Bulk modulus (GPa)	82	94.0
c_{11} (GPa)	140.5	162.48
c_{33}	132.3	115.02

Table 3.15: TCP unit cell properties from experiment and simulation

cell and mechanical properties of the polymorphs of TCP (table 3.16), though the difference is much larger than for the equivalent simplification of the Mkhonto potential, both from experimental values and from the parent potential. While for the full potential, β-TCP was 2.91 kJ/mol lower in energy than α-TCP, for the simplified variant this quantity swelled to 52.1 kJ/mol, underlying how this version of the potential has very different properties. This is likely due to the much weaker spring constant in the core-shell model leading to the atom having a much greater polarisability, particularly as the charges applied to the core and shell are much larger.

α-TCP			
Property	**Experiment**	**Calc (full)**	**Calc (simp)**
a (Å)	12.887	12.727	13.332
b(Å)	27.280	27.529	26.927
c (Å)	15.219	15.304	15.237
Volume (Å^3)	4317.52	4289.9	4495.4
Bulk modulus (GPa)	75.9	81.9	53.42
c_{11} (GPa)	127.4	121.7	76.5
c_{33}	123	120.7	82.93
β-TCP			
Property	**Experiment**	**Calc (full)**	**Calc (simp)**
a (Å)	10.435	10.284	10.818
b(Å)	10.435	10.284	10.818
c (Å)	37.403	38.613	33.955
Volume (Å^3)	3527.3	3536.9	3542.99
Bulk modulus (GPa)	82	94.0	103.87
c_{11} (GPa)	140.5	162.48	149.7
c_{33}	132.3	115.02	211.3

Table 3.16: Comparison of behaviour of full and simplified Mkhonto potentials

3.9 Conclusion

Five calcium phosphate force fields have been implemented, and evaluated for their ability to recreate the experimental properties of TCP. Those of Meis et al. [2000] and Lee et al. [2000] are unable to accurately simulate the known experimental structures. That of Hauptmann et al. [2003] does recreate the unit cell proper-

ties reasonably accurately, but fails to reproduce mechanical properties well. The remaining force-fields both give good agreement with unit cell and mechanical properties, but both feature an undesirably computationally expensive core-shell model for their oxygen atoms. When this is removed in the simplest possible way, only the Mkhonto and de Leeuw [2002] potential continues to give good agreement with the known properties of TCP. It is therefore this simplified potential that has been chosen for use in further work.

Chapter 4

Metadynamics simulation of tricalcium phosphate

Having selected an appropriate force-field for TCP I wished to repeat the work Quigley et al. [2009] had conducted on calcium carbonate nanoparticles, exploring the crystalline and amorphous conformations available to the system. As discussed in section 1.3 it is not clear at which stage of growth from solution crystallisation occurs. Through the use of metadynamics (see section 2.7) I have sought to locate the free energy minima in the phase space of nanoparticle conformations and investigate the extent to which they are crystalline, and hence begin to gain an understanding of at what size nanoparticle the amorphous to crystalline transition takes place.

4.1 Metadynamics

Using metadynamics (see section 2.7) I attempted to explore the free energy surface (FES) of the nanoparticle in terms of three separate sets of collective variables. As this is preliminary work, simulations in vacuum were performed due to the substantial saving in computational time gained — hydrated simulations could be performed later if appropriate CVs were identified. Full details of the simulations are given in section 10.1. Order parameters derived from graph theory are discussed in chapter 5 after the ideas behind these parameters are introduced. Here I present the results from the use of the principal moments of inertia, and Steinhardt order parameters.

4.1.1 Principal moments of inertia

For an object comprised of point masses, such as an atomistic simulation of a nanoparticle, the inertia tensor I with respect to the origin is given by:

$$I = \sum_{k=1}^{N} m_k \begin{pmatrix} y_k^2 + z_k^2 & -x_k y_k & -x_k z_k \\ -x_k y_k & x_k^2 + z_k^2 & -y_k z_k \\ -x_k z_k & -y_k z_k & x_k^2 + y_k^2 \end{pmatrix} \tag{4.1}$$

Here there are N atoms with atom k having Cartesian coordinates (x_k, y_k, z_k), and mass m_k. As the sum of symmetric matrices I is itself symmetric, and hence can be transformed to a diagonal matrix with the eigenvalues on the diagonal. These three eigenvalues are the principal moments of inertia, and are independent of particle orientation providing the centre of mass is at the origin.

In work inspired by Bealing et al. [2010], it is these quantities, roughly quantifying the morphology of the nanoparticle, that have been used as order parameters to drive the bias underlying metadynamics. By driving morphology it was hoped that internal structure would also alter and crystalline forms would appear.

Simulation results

An unexpected issue was encountered during early simulations. These CVs will eventually drive the nanoparticle into two or more pieces which rapidly move apart when sufficient bias has been applied. The entropic gain of a nanoparticle vaporising in an infinite vacuum is itself infinite, and the whole purpose of metadynamics is to explore free energy minima, so this is not an error. Unlike CVs such as the Steinhardt parameters considered next these CVs can describe pathways of separation events, though the estimate of the free energy surface cannot of course converge to an infinitely deep minima, so this behaviour cannot be fully explored. The FES presented here are therefore not fully converged, and may be approaching the point of separation.

The simulation presented here explored a wide variety of morphologies, from roughly spherical shapes to long tubes. Trajectories show the bias is sufficient to cause large changes in morphology, reminiscent in appearance to the kneading of bread dough. A full set of the 2D and 1D free energy surfaces has been produced (figures 4.1, 4.2 and 4.3). All the deepest minima were found along the first principal moment of inertia, with the deepest (A in figure 4.4) the point where all three

moments are approximately equal, and the particle is roughly spherical. As this CV increases the particle is compressed into more oblate round shapes, with a series of shallower minima such as configurations B and C in figure 4.4.

Configurations from the minima do not appear to be crystalline in nature. Furthermore, from visual inspection of the trajectories it does not appear that the nanoparticle is exploring crystalline configurations at any stage. It is possible order has been missed, but with no other minima to explore, these CVs appear unable to distinguish between crystalline and amorphous states. It would seem intuitively reasonable that crystalline and amorphous states could have similar densities and morphologies, and the ability to distinguish not only between the two, but along a pathway between the two, is vital if metadynamics is to explore crystalline states. Combined with the issue of particle separation detailed above, it appears that these order parameters are inappropriate for studying the phase behaviour of this system.

4.1.2 Steinhardt order parameters

As an existing set of order parameters which have been shown to drive crystallisation in nanoparticles of calcium carbonate, the Steinhardt order parameters (described in more detail in section 2.7) are an obvious choice for the application to TCP.

Simulation Results

Figures 4.5, 4.6 and 4.7 shows the recovered estimate of the free energy surface with respect to the order parameters in an assortment of combinations. Two minima are identified, separated by about 20 k_BT, which I denote as state A and state B. Example configurations corresponding to the states are shown in figure 4.8. From visual inspection it is difficult to distinguish between these two states, still less identify them with bulk crystal phases. The local structure methodology of Quigley and Rodger is based on the high symmetry within the calcium carbonate polymorph unit cells, and is dependent on there only being one site of each kind for each polymorph. As discussed in chapter 5 this is not the case for either TCP polymorph, and this methodology would not be useful without significant adaptation. More fundamentally, it was perhaps naïve to believe that a particle of a size of the same order as the polymorph unit cell could crystallise to that polymorph, when there are likely to be substantial surface effects counteracting the energetic benefits of order over disorder.

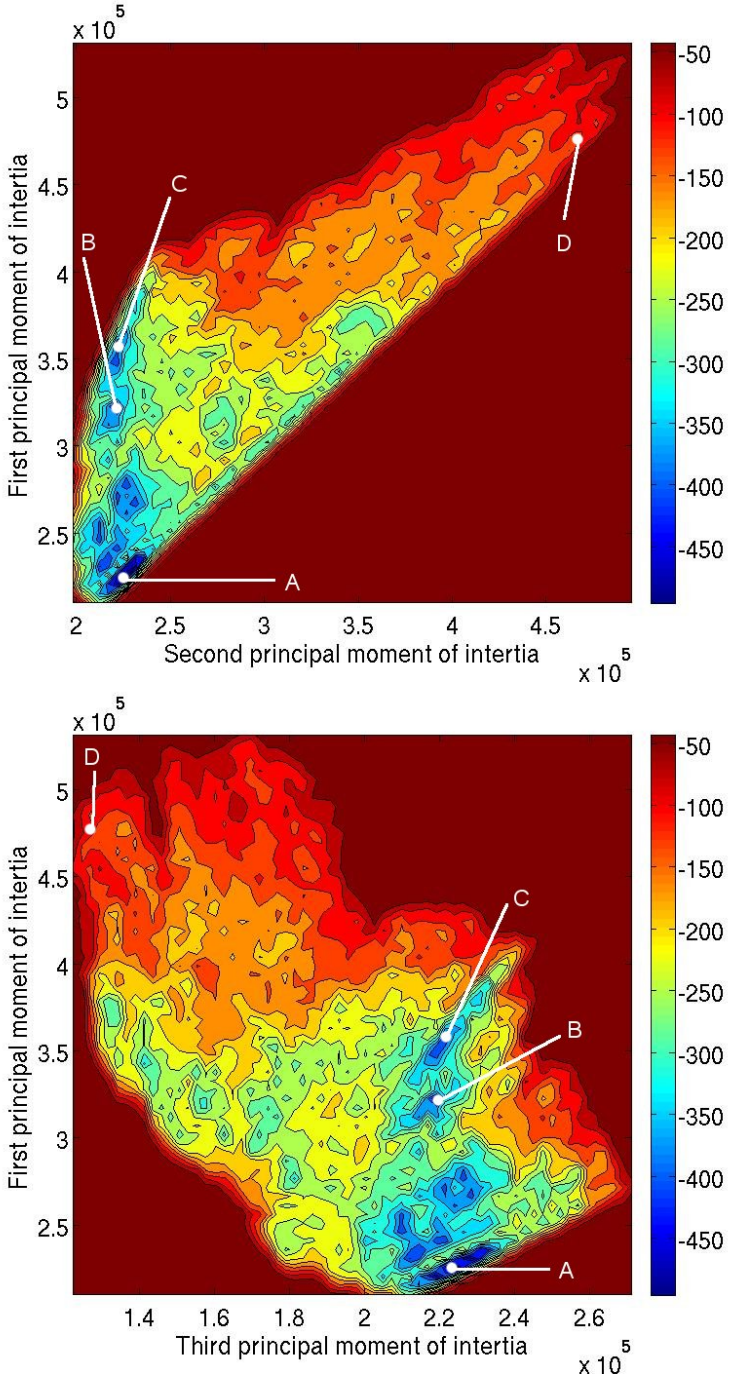

Figure 4.1: Approximations to the FES for the inertia tensor eigenvalues for a 273 atom nanoparticle of TCP. Units on the contour plot are k_BT. Letters A, B, C, D refer to figure 4.4.

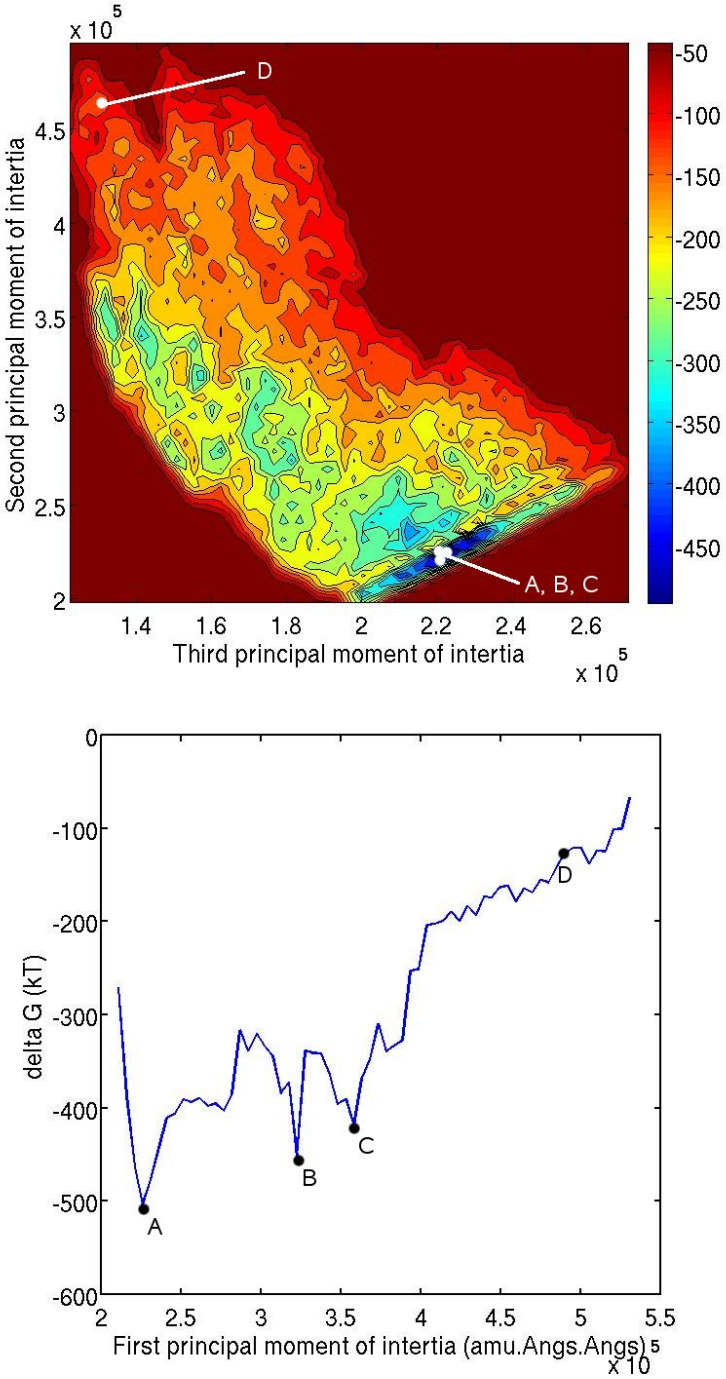

Figure 4.2: Approximations to the FES for the inertia tensor eigenvalues for a 273 atom nanoparticle of TCP. Units on the contour plot are k_BT. Letters A, B, C, D refer to figure 4.4.

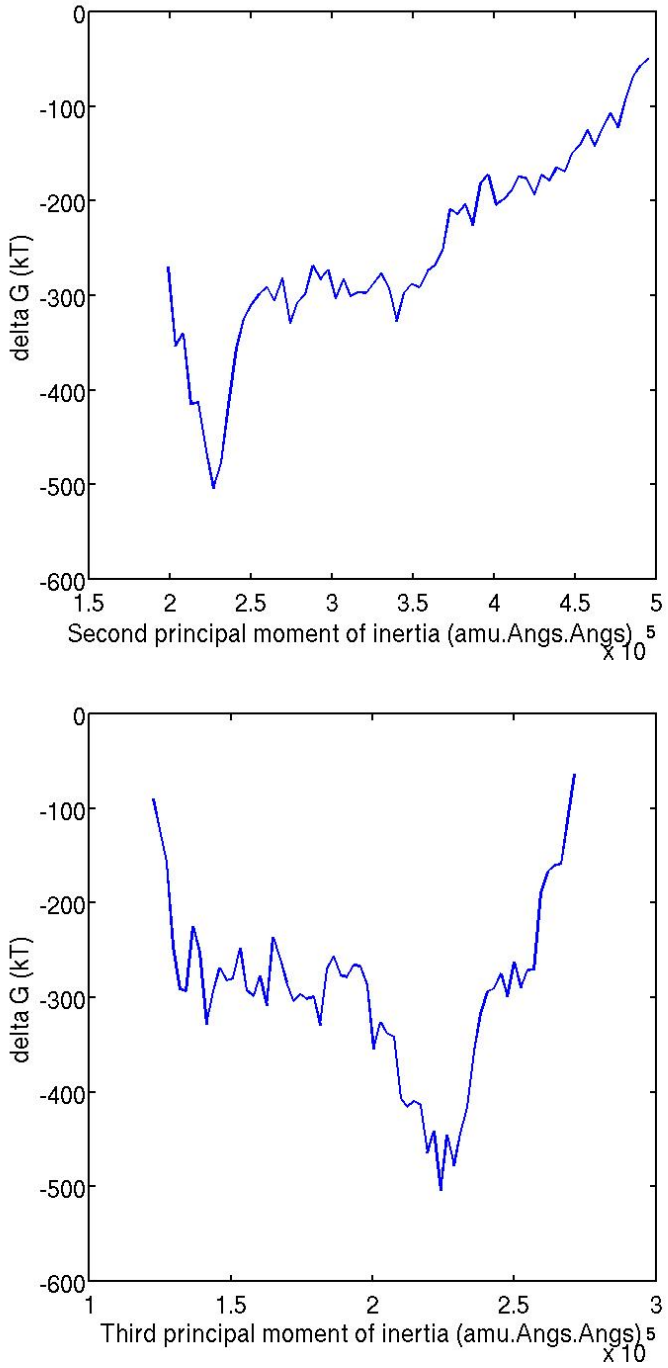

Figure 4.3: Approximations to the FES for the inertia tensor eigenvalues for a 273 atom nanoparticle of TCP

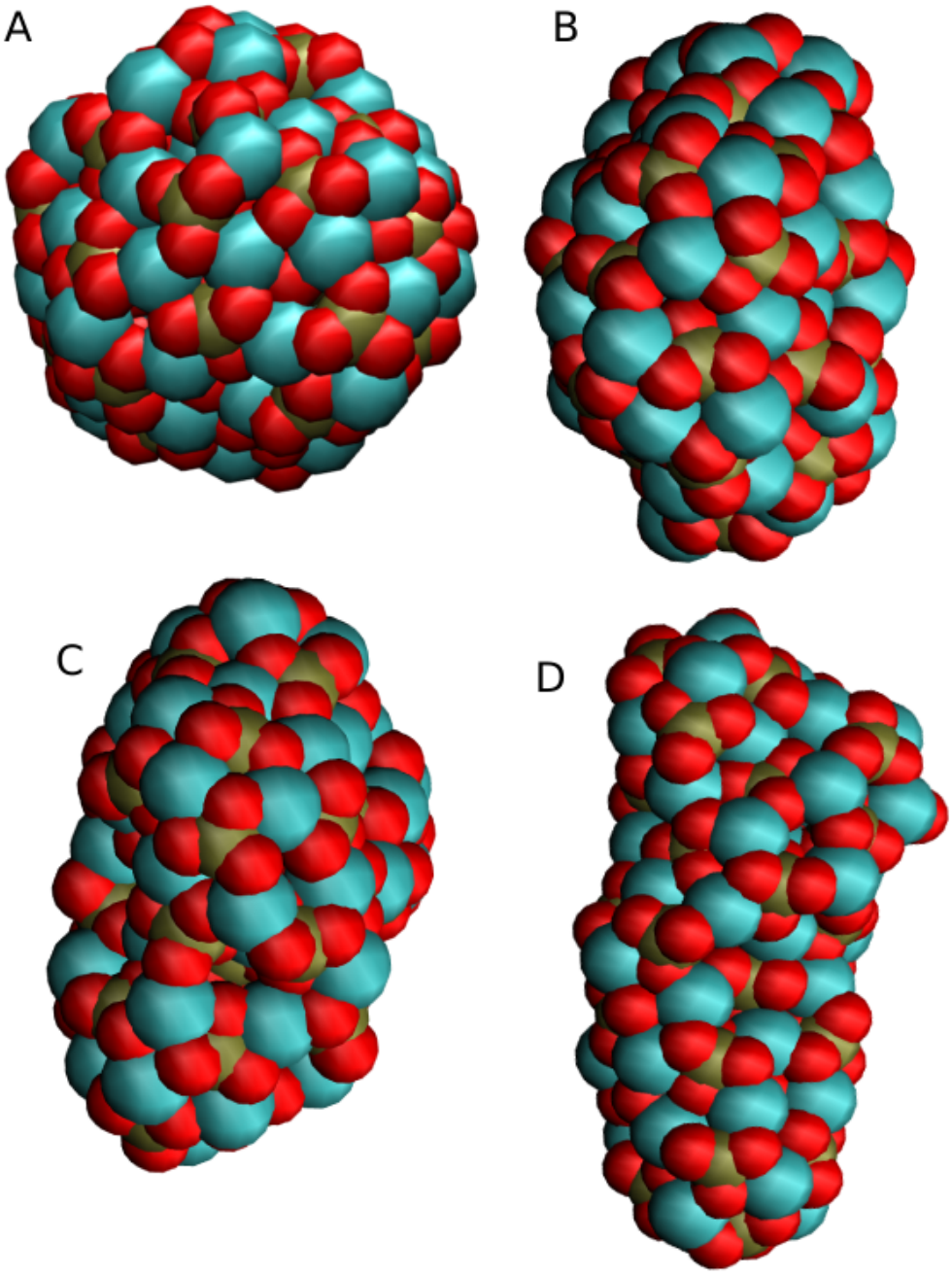

Figure 4.4: Configurations along the biased trajectory. A, B, and C are minima while D is a high energy configuration close to splitting in two.

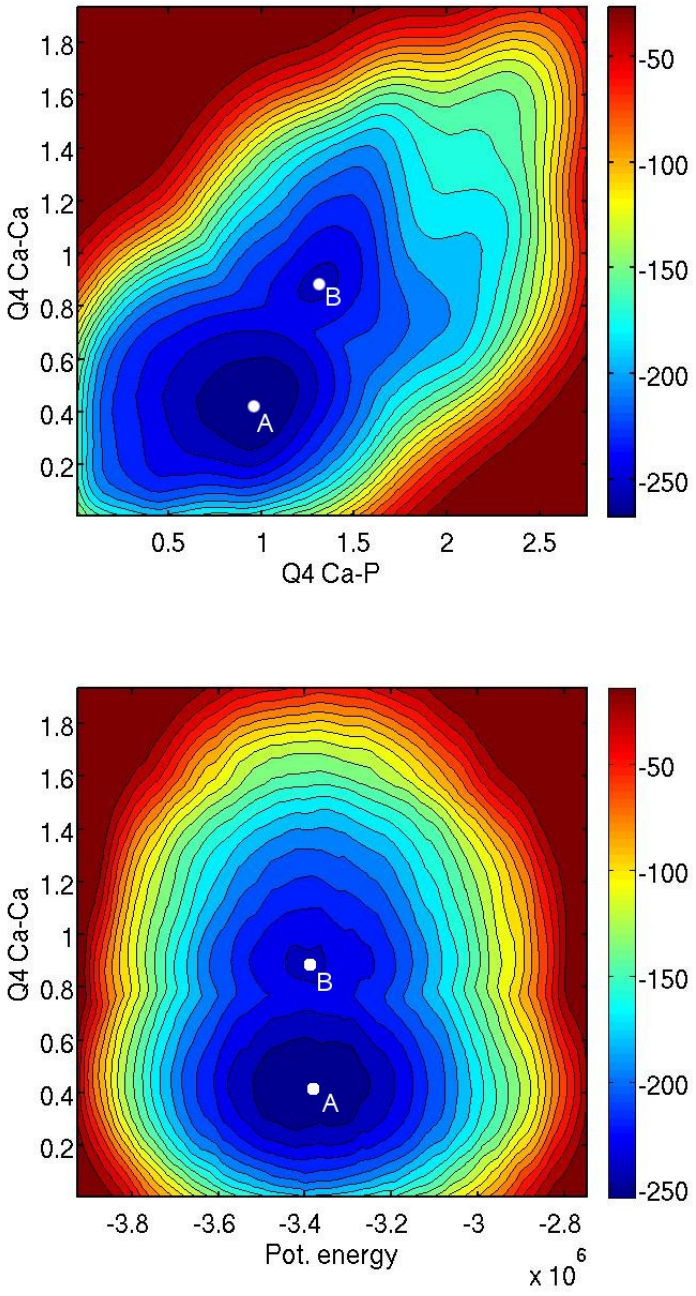

Figure 4.5: Approximations to the FES for the given CVs. The energy scale for the 2D contour plot is in units of k_BT. Points A and B correspond the states in figure 4.8.

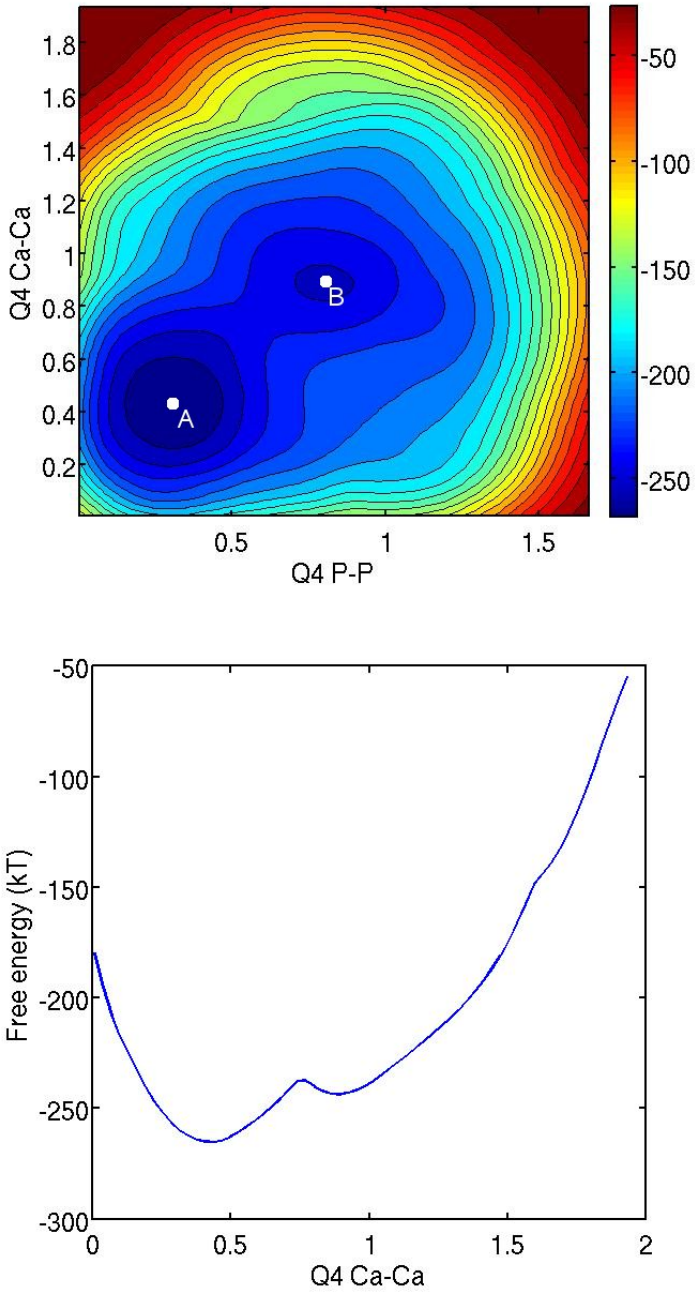

Figure 4.6: Approximations to the FES for the given CVs. The energy scale for the 2D contour plot is in units of k_BT. Points A and B correspond the states in figure 4.8.

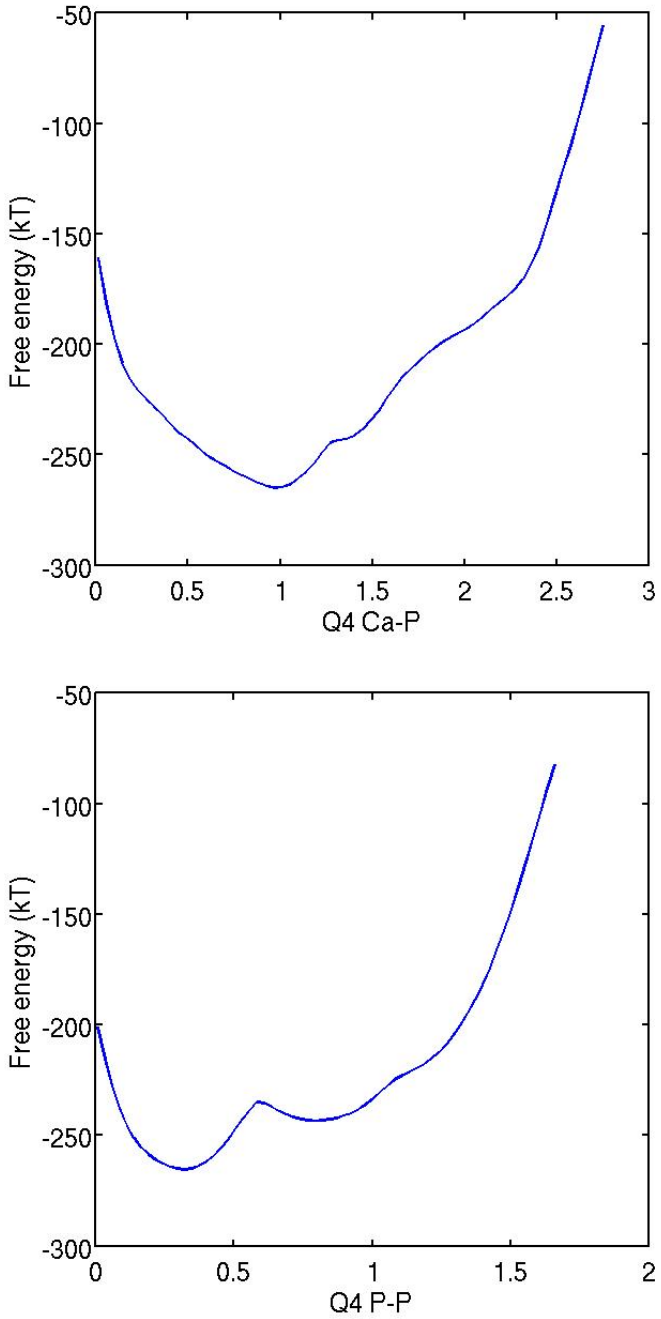

Figure 4.7: Approximations to the FES for the given CVs.

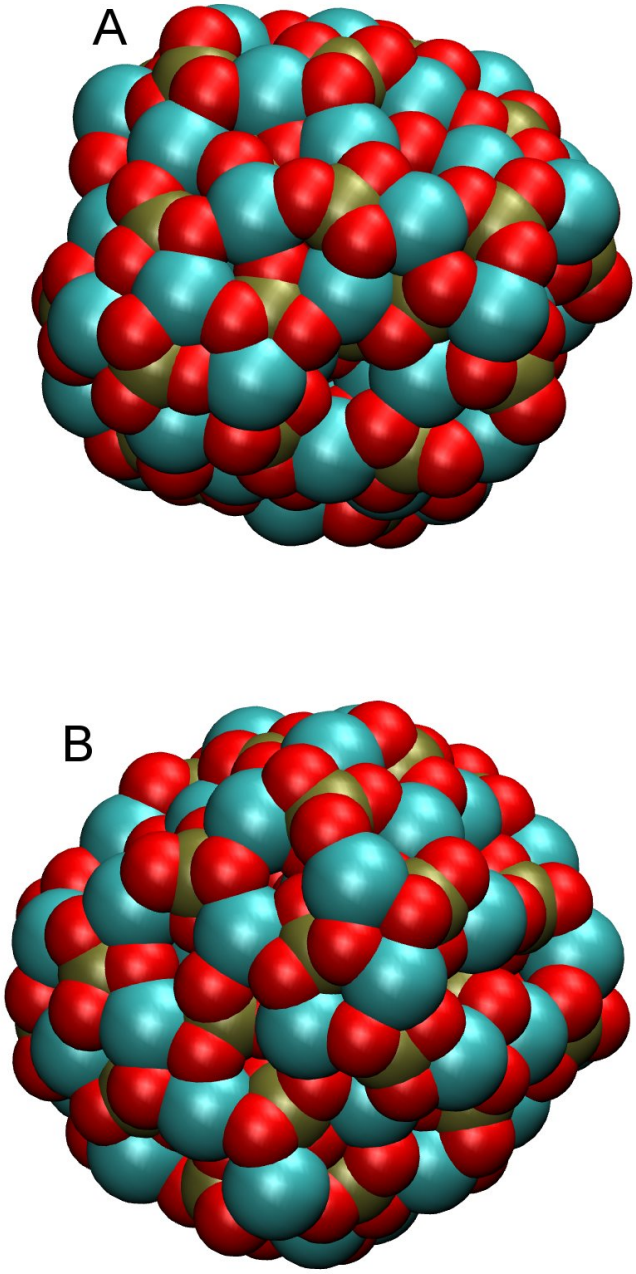

Figure 4.8: Configurations representative of the free energy minima found during metadynamics simulation.

Nevertheless, as both states have very similar potential energy yet state B is much higher in free energy, state B must also be higher in entropy, and hence arguably less disordered. This hypothesis is supported by state B having higher numerical values in the Steinhardt order parameters, and hence more order.

4.2 Conclusion

I have been able to use Steinhardt order parameters to identify two distinct minima in the free energy landscape of nanoparticles of TCP in vacuum. Beyond characterising them as high and low order however I am unable to identify them with a bulk phase.

The complications within the α and β-TCP unit cell, with their large number of distinct sites, makes local structure analysis extremely difficult, though even if considered as truly amorphous, the existence of two distinct classes of amorphous TCP is a novel result. Further work on TCP could involve studying larger particle sizes to see if the relative stability of the high and low ordered states change, though it seems probable that nanoparticles of sufficient size to identify with bulk polymorphs would be computational impractical to simulate. This is likely to present a significant barrier to the study of TCP nucleation and growth in the years ahead.

Chapter 5

Application of graph theory to molecular simulation

5.1 Use of topology to study local order

5.1.1 Introduction

Characterising output is a key consideration for any study of phase transitions. In the context of this project, the ability to distinguish between crystalline polymorphs and amorphous phases of the biomineral compounds with the greatest possible resolution and accuracy is vital.

It should be noted that assigning a polymorph based on local structure is inherently subjective. Unless one does not allow any deviation from the relaxed crystal unit cell, there will always be a choice in how much deviation to allow before no longer considering an area as being, for example, aragonite-like. A suitable scheme will be self consistent in that the same ion cannot be assigned to two phases, and flexible enough to allow some distortion in the structure, without erroneously categorising amorphous structure as crystalline.

Several approaches to local structure analysis have appeared in the literature. For their metadynamics work on calcium carbonate nanoparticles (beginning with Quigley and Rodger [2008a] and continuing through to Quigley et al. [2011]) have used local versions of the Steinhardt parameters used to drive the metadynamics bias. Peters [2009] has extended the work of Shetty et al. [2002] to study nucleation in oppositely charged colloids, developing local order parameters to do so. More recently, Santiso and Trout [2011] have presented a more general set of

order parameters for molecular crystals. What all of these approaches share (apart from a highly complicated derivation and implementation) is a dependence on the angles between the ions in the coordination shell of the ion of interest.

An alternative to such approaches is to consider not the precise angles between the ions but merely their presence in the correct approximate locations. The mathematical field of graph theory, where the precise orientation of units is regarded as irrelevant providing the overall network of links is conserved, provides a framework to formalise this idea and produce order parameters.

Each atom (or small molecule) is set as a node on a graph (figure 5.1). If two atoms are sufficiently close in space then there is considered to be an edge linking the two nodes. A *ring* is defined as any closed loop of edges. More restrictively, a *primitive ring* is a ring where the shortest path between two nodes on the ring is equal in length to the shortest path between them around the ring. A *local cluster* of an atom is defined as the part of the graph that includes all the nodes that form part of one or more primitive rings that include that atom (figure 5.2). For crystalline systems the local cluster can be considered analogous to the unit cell, as a method of summarising the full lattice into a smaller unit.

There is some history of using the local cluster and related ideas to determine local order as extensively reviewed by Linn W. Hobbs in a chapter of Novoa et al. [2008]. I have written a Fortran 90 program based on the algorithm of Hobbs et al. [1998] (see appendix A) to analyse the local clusters of a DL_POLY coordinate file, counting the number of atoms in each local cluster and the number and length of each primitive ring. In the following sections the application of this code to the biomineral systems under consideration is discussed.

5.1.2 Calcium carbonates

There are a number of parameters to select when applying the ring statistics methodology to a new system. The choice of which atoms are to be considered as nodes, and which node pairs are linked on the graph, is crucial to the ability to distinguish between polymorphs. In this implementation I have excluded the oxygen atoms on the carbonate groups from the local clusters. Excluding these atoms is advantageous as it means that only the carbon and calcium local clusters need to be considered. Apart from small deviations in the vaterite structure due to carbonate disorder, there is also only one distinct carbon or calcium site in each polymorph, repeated through symmetry throughout the unit cell. Furthermore the implementation only

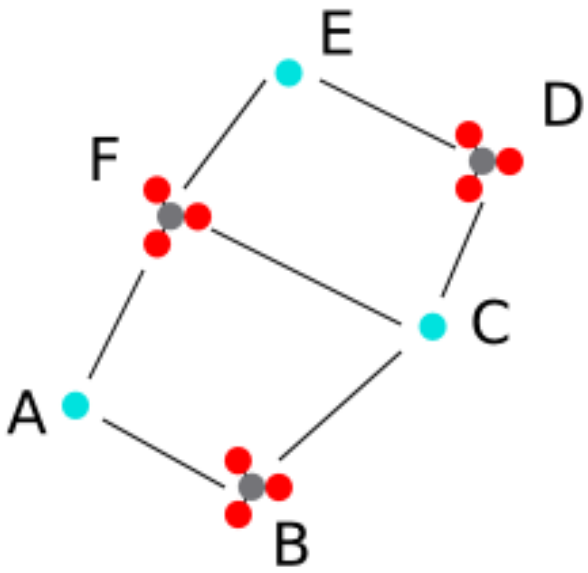

Figure 5.1: Ring theory illustration. The ions A, B, C, D, E and F are nodes on a graph linked by edges. The rings ABCF and CDEF are primitive, while the ring ABCDEF is not due the CF edge

includes Ca—C edges on the graph and not Ca—Ca or C—C edges. This has the advantage that only rings with an even number of edges need to be considered. An appropriate cut-off distance for two adjacent ions to be included on the graph was 3.7 Å. Figure 5.3 shows the local clusters for calcite, aragonite and vaterite, while table 5.1 presents the number of atoms and primitive rings in the local clusters. While the Ca vaterite local cluster has the same ring numbers as the C aragonite structures and vice versa, this does not present a difficulty. The code was able to recognise all the vaterite structures discussed in section 1.1.3 as vaterite. A copy of local structure code, implemented for calcium carbonates, is included in appendix B.

The ring statistics algorithm is clearly able to distinguish between the three polymorphs, but it is also important to check that this choice of ring statistics is unlikely to mis-categorise amorphous structures as crystalline. To this end a 175 formula unit calcite nanoparticle was simulated at 2000 K for 5 ns, before being annealed at 300 K for a further 5 ns. As argued above, local structure analysis is ultimately inherently subjective, but from visual inspection, the nanoparticle does appear to lack any order (figure 5.4), and this is supported by a consideration of

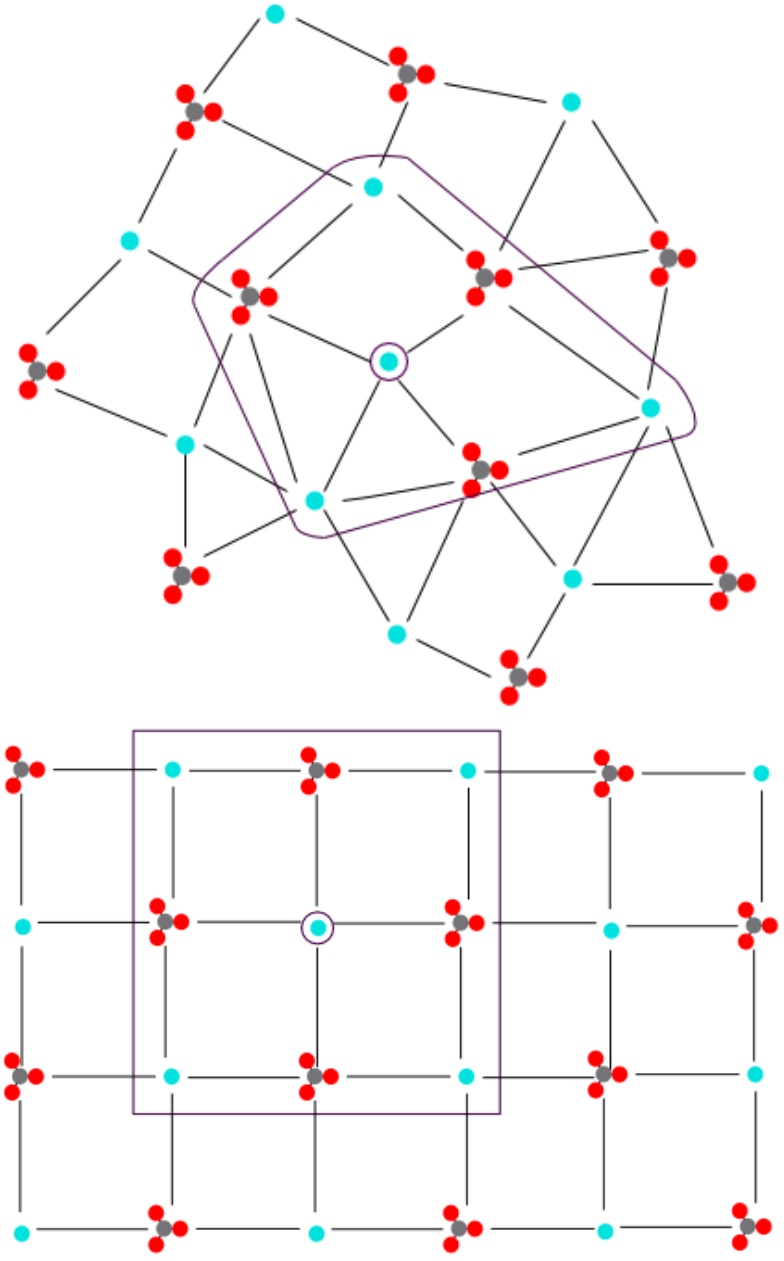

Figure 5.2: Local clusters for highlighted atoms, for ordered and disordered systems.

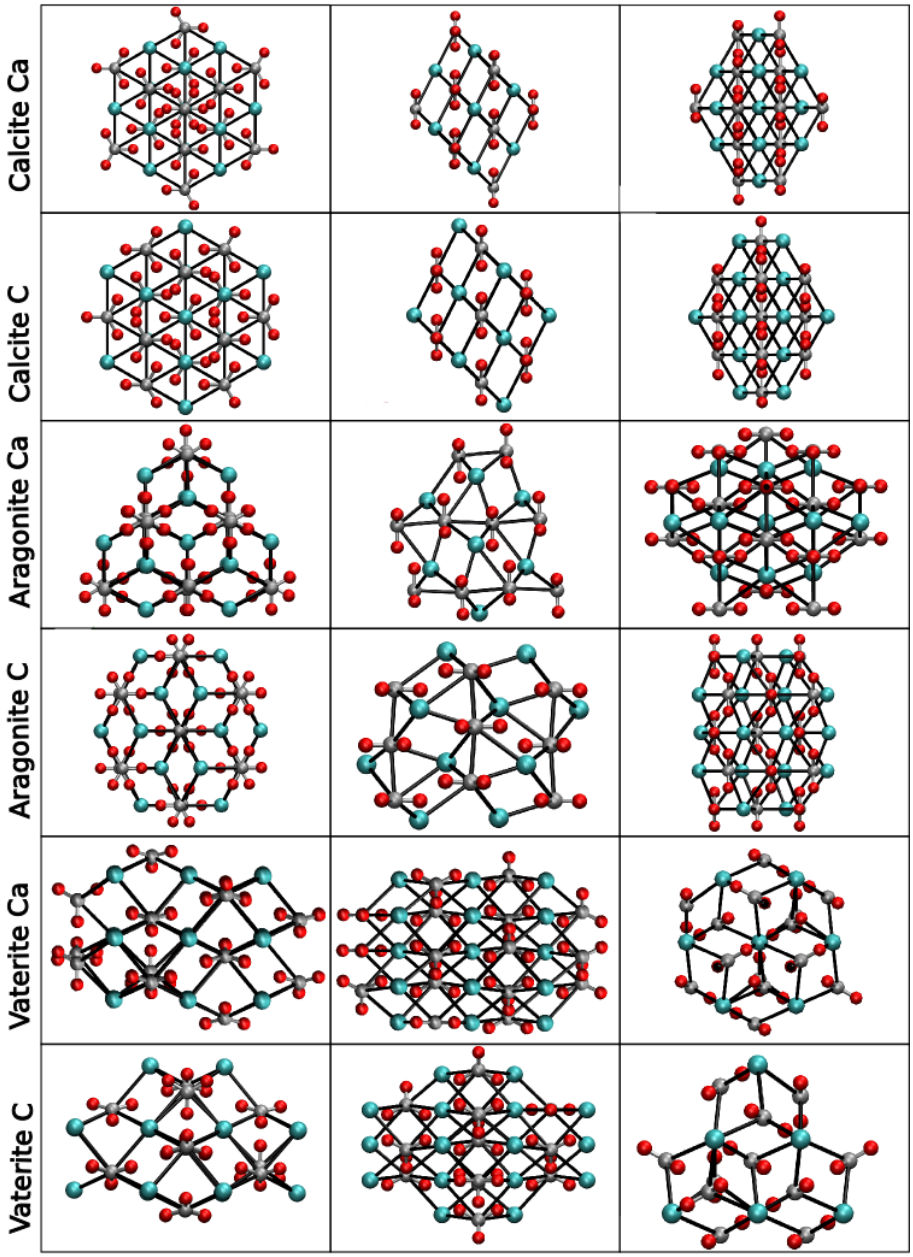

Figure 5.3: Three projections of each of the local clusters for the polymorphs of calcium carbonate. The technique is insensitive to the three possible rotations of each vaterite carbonate ion — here they are randomly populated.

Local cluster	Nodes	4-rings	6-rings
Calcite Ca	27	12	24
Calcite C	27	12	24
Aragonite Ca	31	12	42
Aragonite C	39	12	42
Vaterite Ca	39	12	42
Vaterite C	31	12	42

Table 5.1: Local cluster statistics for calcium carbonate polymorphs

the RDFs of the nanoparticle (figure 5.5). In comparison to the RDFs of the bulk polymorphs (figure 5.6), some of the peaks show some similar features to vaterite, but there are also clear differences and in no case can any secondary peaks be seen. If one accepts that this nanoparticle is sufficiently amorphous, it is supportive of the rings statistics methodology that there were no matches to any of the local clusters of the bulk polymorphs within this nanoparticle.

One drawback of the scheme is that the size of the local cluster varies between phases. This creates two difficulties:

- There is no possibility of exactly matching the local clusters of the bulk phase for surface ions. While for the calcite $(10\bar{1}4)$ surface the local cluster is sufficiently small that this scheme can recognise the next layer below as calcite-like, for the other polymorph the larger clusters mean that it would never be possible to unambiguously identify atoms so close to the surface as belonging to the crystalline phase. For smaller nanoparticles, this restricts local cluster analysis to the core for aragonite and vaterite.

- The larger cluster size means a greater area must be crystalline for the scheme to identify an ion as crystalline. 39 nodes on the graph (as for the aragonite carbon and vaterite calcium clusters) equates to nearly 20 formula units in the approximate arrangement of the bulk crystalline phase, in comparison to nearly 14 for calcite, a much weaker test.

When applying this algorithm, local structure must therefore be considered in terms of what it is possible to measure. Of the 294 ions in an unrelaxed 147 formula unit calcite nanoparticle (of the kind generated by the ISMC technique — see chapter 6) 100 of these ions can be identified as being calcite-like, as this many ions are below the surface. For an aragonite nanoparticle of the same size only 54 ions

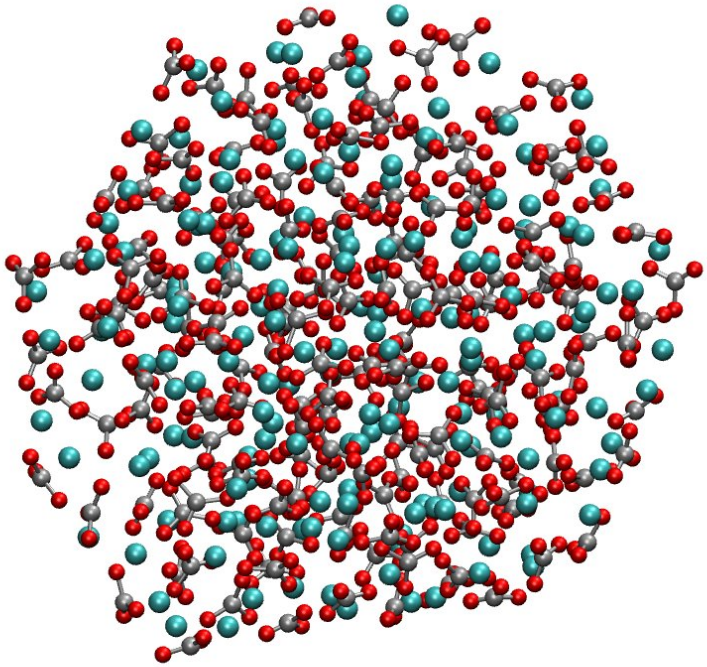

Figure 5.4: Amorphous $CaCO_3$ nanoparticle of 175 formula units.

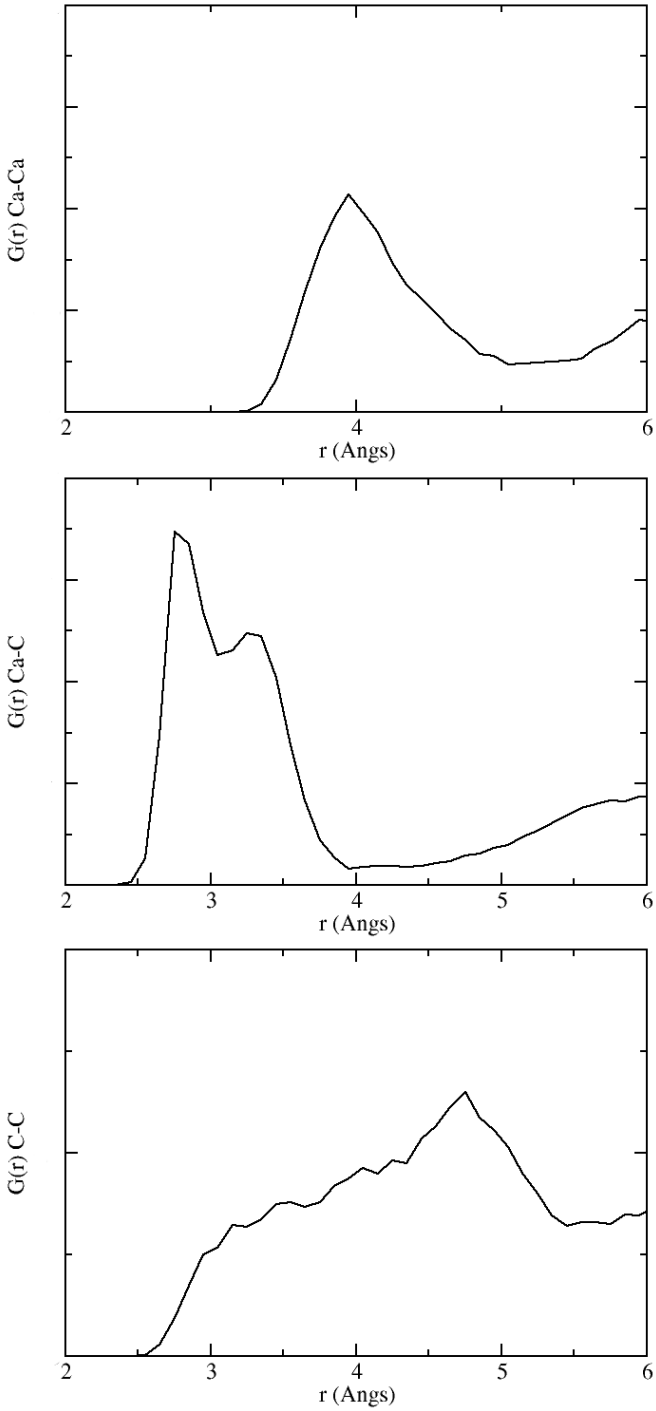

Figure 5.5: Radial distribution functions for calcium carbonate nanoparticle. Since the volume term in the RDF formulation is ill-defined for a nanoparticle in an infinite vacuum, the y-axis values have been removed.

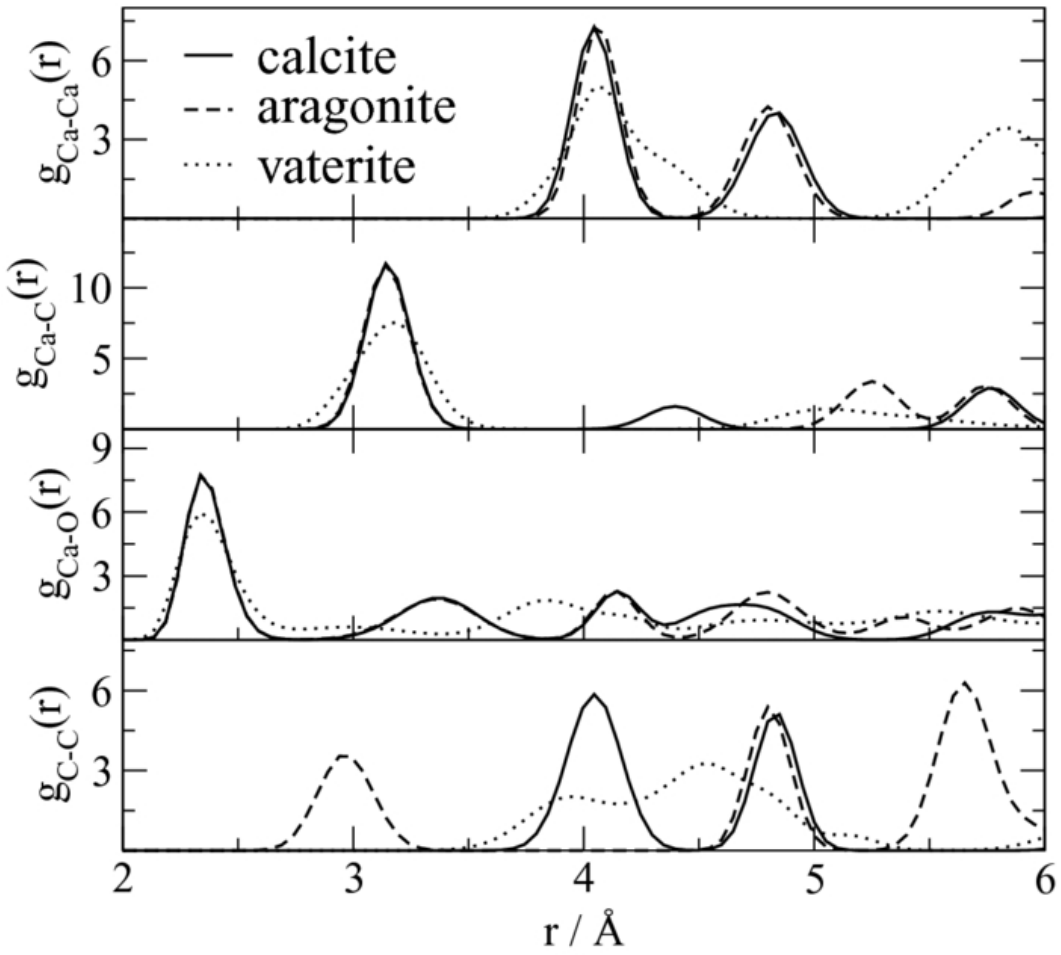

Figure 5.6: Radial distribution functions for calcium carbonate bulk phases.From Quigley et al. [2011].

match the local cluster perfectly. It seems clear that for a less idealised nanoparticle (from a metadynamics trajectory for example), a lower threshold for categorising a nanoparticle as 'aragonite-like' is warranted than for 'calcite-like' nanoparticles.

It is also reasonable to allow some deviation from the perfect aragonite and vaterite local clusters. If n_c, n_4, and n_6 are the number of cluster nodes, 4-membered rings and 6-membered rings in an bulk polymorph cluster, and m_c, m_4, and m_6 are the same quantities for an ion in a system undergoing local cluster analysis I define the quantity d as:

$$d = \sqrt{(n_c - m_c)^2 + (n_4 - m_4)^2 + (n_6 - m_6)^2} \tag{5.1}$$

This parameter gives an estimation of how similar the local cluster considered is to the bulk polymorph. For aragonite and vaterite I have experimented with this value and found that a threshold value of $d < 4.1$ appears to give a good compromise between allowing some deviation from the bulk local cluster, while preventing false positives in nanoparticles of other polymorphs. I thus identify in this work any ion with sufficiently low parameter d as being aragonite-like or vaterite-like as applicable. An ion is only identified as calcite-like if $d = 0$ for the calcite local clusters. Using this scheme, the amorphous nanoparticle of figure 5.4 still has no ions with local structure matching a crystalline phase.

5.1.3 Calcium phosphates

The fundamental difference between the calcium carbonate and TCP systems from the point of view of local structure is that while in the carbonate crystalline polymorphs all sites are identical (with the qualified exception of vaterite), for both α and β TCP there are many distinct sites for both the calcium and phosphate groups. Any local structure detection mechanism would need to be able to recognise all these sites and distinguish between them. I have experimented with a similar graph abstraction scheme as used for the carbonates, considering only the Ca and P atoms as nodes, and only the Ca—P pairs as links. From the radial distribution functions (see figure 10.3) a cut-off of 4.2 Å is adequate for the first coordination shell to be included.

Table 5.2 shows the different clusters for α-TCP. There are 30 distinct local clusters in this scheme, each appearing twice in the unit cell (the unit cell is made of two units identical up to symmetry). Similarly for β-TCP table 5.3 gives the

details of the 14 distinct clusters in the unit cell. Here there is more repetition of the clusters — 18 of the 63 Ca^{2+} ions have the same local cluster, but there is still considerable variation. For both polymorphs these local clusters are also very large, with the majority much larger than even the larger aragonite and vaterite clusters discussed above.

When confronted with a TCP system for local structure analysis it would appear insufficient to identify some of the atoms as having the same local cluster as some of the atoms of a bulk polymorph, when the scheme fails to recognise that each local cluster should only link in specific ways with others. There would also be a question of false positives when there are so many candidate clusters an ion could match, though this is mitigated by the difficulty of matching exactly such large clusters by chance. The large clusters also mean that only at the core of a small nanoparticle would local cluster analysis be possible. I applied the scheme to the A and B structures identified through metadynamics in chapter 4 and in each case only three or four ions had clusters of the same size as those identified for the bulk polymorphs, with the others all smaller, presumably because of surface effects. None of the identified local clusters matched any bulk cluster exactly.

As discussed in that chapter, it is unrealistic to expect crystallinity in a particle the same size or only a few times larger than the unit cell of the bulk polymorphs, to which difficulty we must add the problem of identifying any output as matching a crystalline phase.

5.2 Continuous order parameters

I have shown the ring statistics order parameters can successfully differentiate between the three polymorphs of calcium carbonate, and by exclusion, the amorphous state. A natural extension of the work is to try to use such concepts not just to detect crystallinity, but to drive phase transitions through metadynamics. There has been some prior work that has attempted this: Donadio et al. [2005] successfully used graph theory derived parameters to drive melting of water ice, though this work fails to recover the full free energy surface and drive the crystallisation of the liquid back to the solid. Much more recently Pietrucci and Andreoni [2011] have performed metadynamics on order parameters derived directly from the list of edges of the graph, and not the full ring statistics.

I have derived order parameters for metadynamics from the ring statistics

Local cluster	Nodes	4-rings	6-rings	Freq.
Ca	35	7	31	2
Ca	37	19	19	2
Ca	38	13	41	2
Ca	39	17	21	2
Ca	39	17	28	2
Ca	39	15	35	2
Ca	41	19	18	2
Ca	41	16	27	2
Ca	42	13	45	2
Ca	43	12	49	2
Ca	44	20	36	2
Ca	46	12	53	2
Ca	48	15	60	2
Ca	48	19	56	2
Ca	49	17	54	2
Ca	49	17	56	2
Ca	49	20	55	2
Ca	49	18	63	2
P	43	26	60	2
P	43	26	61	2
P	44	17	54	2
P	44	25	55	2
P	44	23	59	2
P	44	24	67	2
P	45	26	59	2
P	45	21	72	2
P	46	24	55	2
P	46	21	69	2
P	47	32	59	2
P	47	21	77	2

Table 5.2: Local clusters for α-TCP. Frequency is number of times each cluster appears in the unit cell

Local cluster	Nodes	4-rings	6-rings	Freq.
Ca	38	9	36	3
Ca	38	16	32	9
Ca	39	17	38	9
Ca	43	24	45	3
Ca	45	22	48	9
Ca	46	24	51	3
Ca	48	22	44	9
Ca	50	21	53	18
P	49	27	60	9
P	50	28	59	9
P	50	27	69	9
P	51	33	73	9
P	52	33	63	3
P	56	36	90	3

Table 5.3: Local clusters for β-TCP. Frequency is number of times each cluster appears in the unit cell

concepts above and applied them to a variety of systems. The key difference between my work and that of Donadio et al. [2005] is the method of converting the inherently binary properties of the definitions applied above — an edge either exists or it does not, a ring is either primitive or it is not — into the smooth, continuous collective variables necessary for the metadynamics algorithm.

5.2.1 Definitions

The key distinction between the discrete scheme described above and the continuous case is that an edge, rather than being a distinct binary property, moves smoothly in and out of existence as the two atoms it links move closer together or further apart. For these order parameters, each edge on the graph is assigned a value I denote as the 'f-score' between zero and one based on the distance r between the two nodes in the full atomistic description, as in equation 5.2.

$$f(r) = \begin{cases} 1 & \text{if } r \leq r_c \\ 1 - 3\left(\frac{r-r_c}{r_o-r_c}\right)^2 + 2\left(\frac{r-r_c}{r_o-r_c}\right)^4 & \text{if } r_c < r < r_o \\ 0 & \text{if } r \geq r_o \end{cases} \tag{5.2}$$

The function in the mid part of the range is chosen to ensure smoothness at the cross over points. In practice r_o is the cut-off chosen for which an edge ceases to exist in the graph structure.

A ring is in a sense 'better' or at least more notable, if it is composed of shorter, stronger, edges. To quantify this, I define a quality score q for each ring (equation 5.3).

$$q = \prod_{\text{Ring edges, } i} f_i \tag{5.3}$$

The concept of whether a ring is primitive must also be adapted to a continuous property. I assign a 'primitivity' score p as in equation 5.4, based on the lengths of any edge which shortcuts the ring. Where no shortcuts exist $p = 1$ by definition.

$$p = \prod_{\text{Edges crossing ring, } i} (1 - f_i) \tag{5.4}$$

Using this framework we define the order parameter for the number of n-rings as in equation 5.5. Primitive rings composed from short links count more than rings composed of long links with short-cuts across them. Figure 5.7 illustrates some of these concepts.

$$n_a = \sum_{\text{All rings of length } a, \, i} p_i q_i \tag{5.5}$$

I have have also found it helpful to include an order parameter, n_C, based on coordination number. This again needs to be converted to a continuously variable quantity, and the f-score provides the most convenient method in this context. The simple sum of the f-score for all edges is, up to division by a scalar, the same as the average coordination number (equation 5.6).

$$n_C = \sum_{\text{All graph edges, } i} (f_i) \tag{5.6}$$

Numerous systems were used for the evaluation and testing of these order parameters, including carbon fullerenes and diamond, xenon, iron, and silicon carbide. Presented here are the results for calcium carbonate, as the archetypal biomineral system, and cadmium selenide, as a system for which the technique has shown the most promise. Full details of the parameters of these simulations can be found in section 10.3.

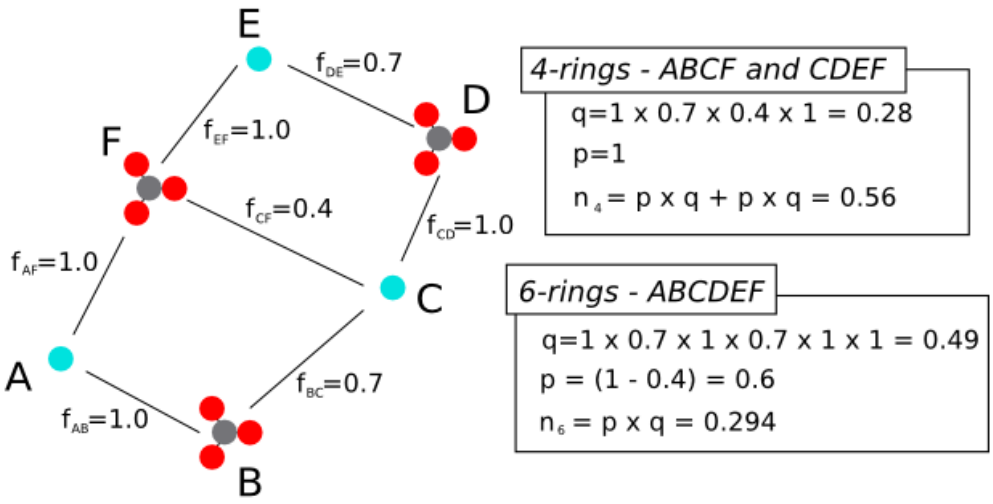

Figure 5.7: Illustration of continuous ring statistic order parameters

Calcium carbonates

The trajectory of the simulation shows rapid change in the position of ions within the nanoparticle. The particle quickly transforms from the calcite rhombohedron to a roughly spherical shape. The order parameters are thus clearly successfully driving change in the nanoparticle shape, with rapid change both internally and at the surface. Figures 5.8, 5.9 and 5.10 show the approximation to the FES produced from this simulation. The 1D energy surfaces each show several minima, though the n_6 surface is significantly more complicated than the other two. Points A and B on the free energy surfaces are the deepest identified minima, and representative configurations are shown in figure 5.11. Applying local structure analysis, neither appear to represent the bulk polymorphs to any degree whatsoever. Point C on the free energy surface is example of the calcite configurations from the beginning of the simulation and is not part of a free energy minimum in terms of any pair of collective variables. Furthermore, this is only one such configuration — many other configurations which were still recognisably calcite-like (both by eye and through local structure analysis) have very different continuous ring statistics and coordination number and would appear over an extremely wide area of these surfaces, including in the amorphous minima.

The 2D surfaces again show numerous shallow minima, though the n_4—n_6

FES is significantly less undulated — it appears that the number of four-membered primitive rings is correlated with the coordination number The entire energy surface has been explored to only a very shallow depth (around 30 k_BT), though a vast number of configurations have been explored. None of the minima are particularly distinct when considered in light of the Gaussian height of 5 k_BT, and all could be considered indistinguishable from noise.

It appears that these CVs are not appropriate for studying crystalline polymorph selection for this system. The free energy surface may be best considered as one vast, largely flat basin, within which both amorphous and crystalline configurations are scattered widely and are indistinguishable through consideration of the CVs alone. Two basic requirements of a set of CVs for the purposes of this work are that they can distinguish amorphous from crystalline states, and that all crystalline states of one polymorph should be confined to a relatively small area of phase space. These CVs fulfil neither criteria for this system. In order to give cut-off distances large enough to allow the bias to drive conformational change, the resolution necessary to distinguish between polymorphs is lost. It may also be the case that while the precise atomic angular orientation is not necessary to distinguish between polymorphs, this incorporation of this information is necessary when encapsulating important transition states into CVs.

Cadmium Selenide

Studies of nanoparticle conformation are not just restricted to biomineral systems. Cadmium selenide has been studied as a test system for the technique of transition path sampling [Grünwald and Dellago, 2009], and has the advantage of having two clear, simple, crystalline polymorphs, being stable in both rocksalt and wurtzite structures. As such it is an ideal system on which to experiment with advanced nanoparticle sampling techniques.

The trajectory shows rapid change in the structure, with the nanoparticle exhibiting both wurtzite and rocksalt structures. The transition between these states follows the same mechanism as that found by Grünwald and Dellago [2009], beginning at one edge and working gradually through the structure. After these configurations have been explored the system moves through a wide variety of amorphous structures, before eventually returning to the crystalline phases.

The free energy surfaces (figures 5.12, 5.13 and 5.14) for this substance are much smoother than those for calcium carbonate, perhaps due to the much larger

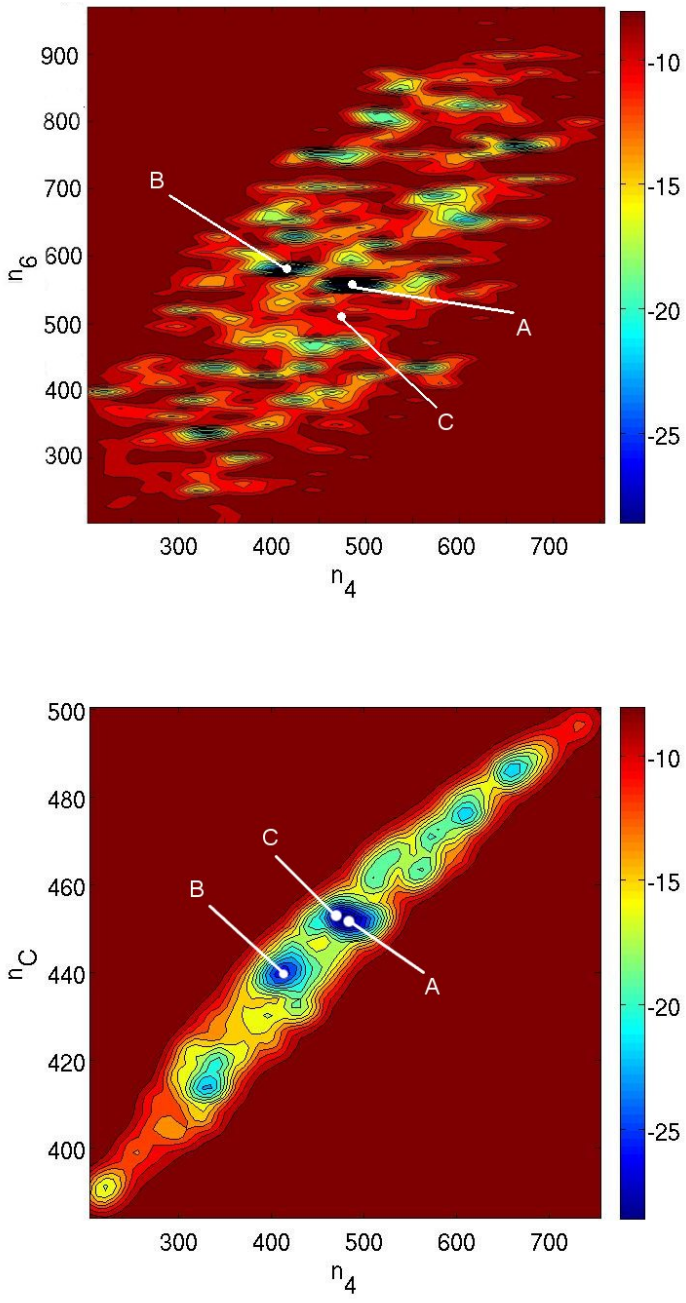

Figure 5.8: Approximations to the FES for the given CVs for a calcium carbonate nanoparticle. The energy scale for the 2D contour plot is in units of k_BT. Points A, B and C correspond to configurations shown in figure 5.11.

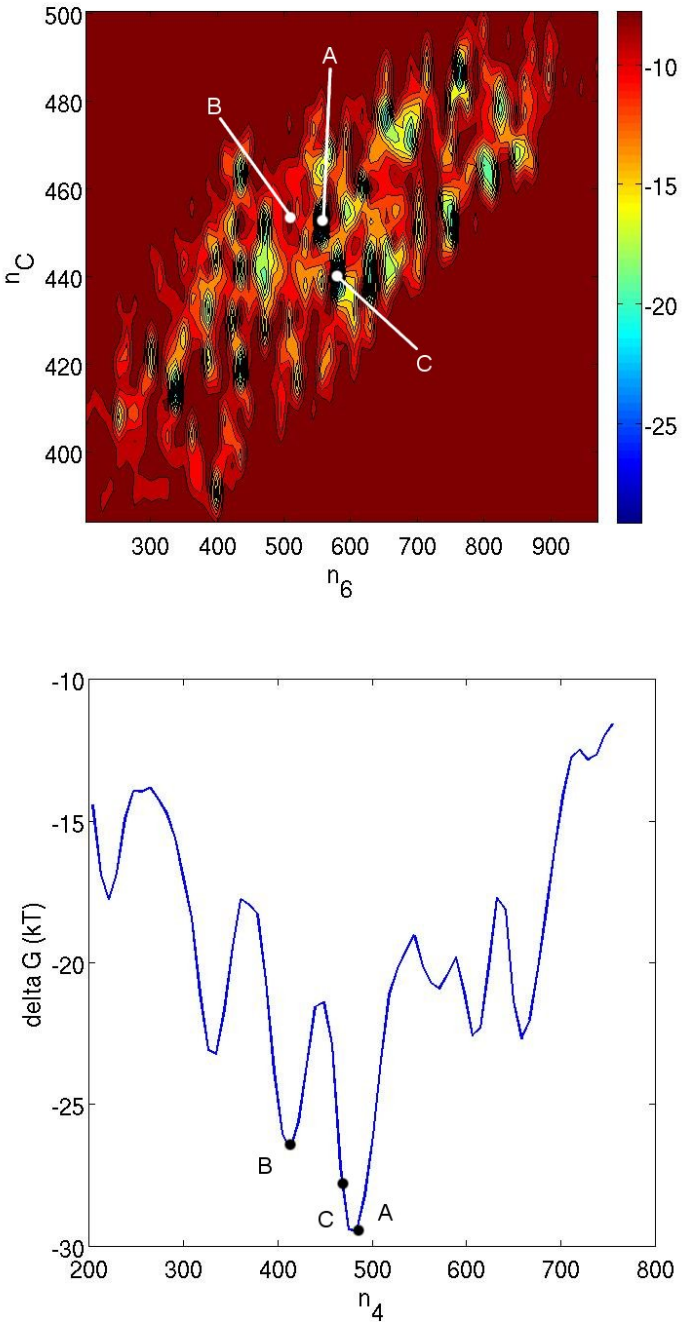

Figure 5.9: Approximations to the FES for the given CVs for a calcium carbonate nanoparticle. The energy scale for the 2D contour plot is in units of k_BT. Points A, B and C correspond to configurations in figure 5.11.

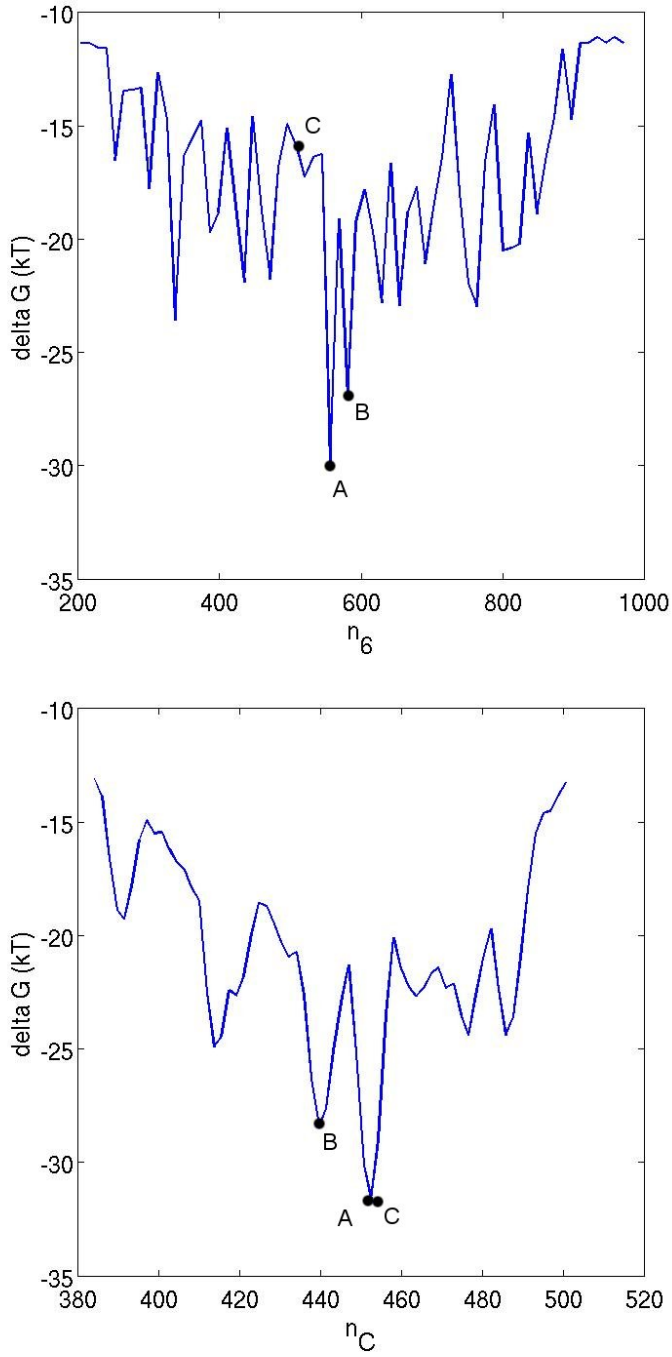

Figure 5.10: Approximations to the FES for the given CVs for a calcium carbonate nanoparticle. Points A, B and C correspond to configurations in figure 5.11.

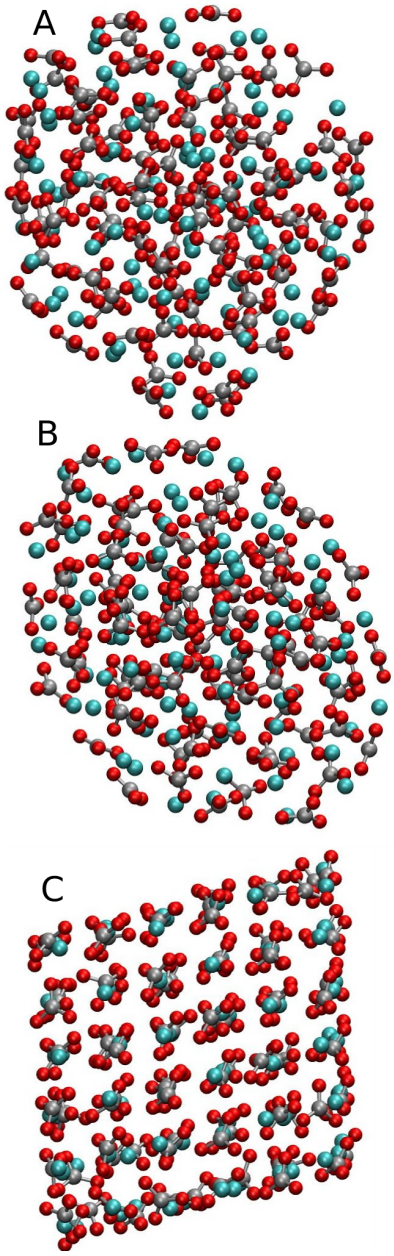

Figure 5.11: Calcium carbonate configurations A, B and C from the free energy surfaces. A and B are from the identified minima while C is representative of the the initial calcite configuration.

width of the Gaussians, and are explored to a much greater depth. Nevertheless important features can be extracted as for the 1D n_6 surface, which shows two minima. The 2D surfaces which include n_6 also show these minima. Representative configurations have been extracted from the trajectory and are shown in figure 5.15, corresponding to the highlighted points in the free energy surfaces. The deepest minimum (point A) corresponds to the wurtzite structure, while point B corresponds to the shallower basin, separated by about 100 k_BT from the wurtzite minimum. It appears to be highly amorphous. Points C and D correspond to differing rocksalt configurations, which while they have the same crystal structure, have different shapes. None of the free energy surfaces show a minima for rocksalt, which from the trajectories was unexpected: from visual inspection the nanoparticle spends significant time in that polymorph. It appears that as with the calcite in the calcium carbonate example the range of phase space available for this polymorph is wide, and poorly distinguished from amorphous states. Another factor might be the pressure in the simulation, since Grünwald and Dellago [2009] applied barostat using ideal gas particles to apply significant pressure to their nanoparticle, and in this way helped stabilise the rocksalt structure. At zero pressure wurtzite is the favoured structure.

5.2.2 Conclusion

The failure of these CVs when applied to calcium carbonate systems and the dubious utility of the recovered free energy surfaces for cadmium selenide is disappointing. The order parameters are by design only capable of directly driving ion-ion distances and not relative orientations. For calcium carbonate systems at least it appears some consideration of the orientations is essential to driving phase transitions.

The order parameters do show some promise however. They have successfully been shown to drive ordered-ordered phase transitions in nanoparticle systems. Furthermore, while in isolation they have been shown to be unsatisfactory for biomineral systems, in chapter 8 I discuss the possibility of combining one or more of these parameters with others to create a hybrid phase space in which more slow processes are encapsulated than would be by any one technique alone.

Future work could include longer, slower, simulations of CdSe system, using well-tempered metadynamics to get much better resolution on the free energy surfaces. For calcium carbonate, some attempts could be made to include carbonate orientation, perhaps by driving Ca—O bonds rather than Ca—C bonds. it would

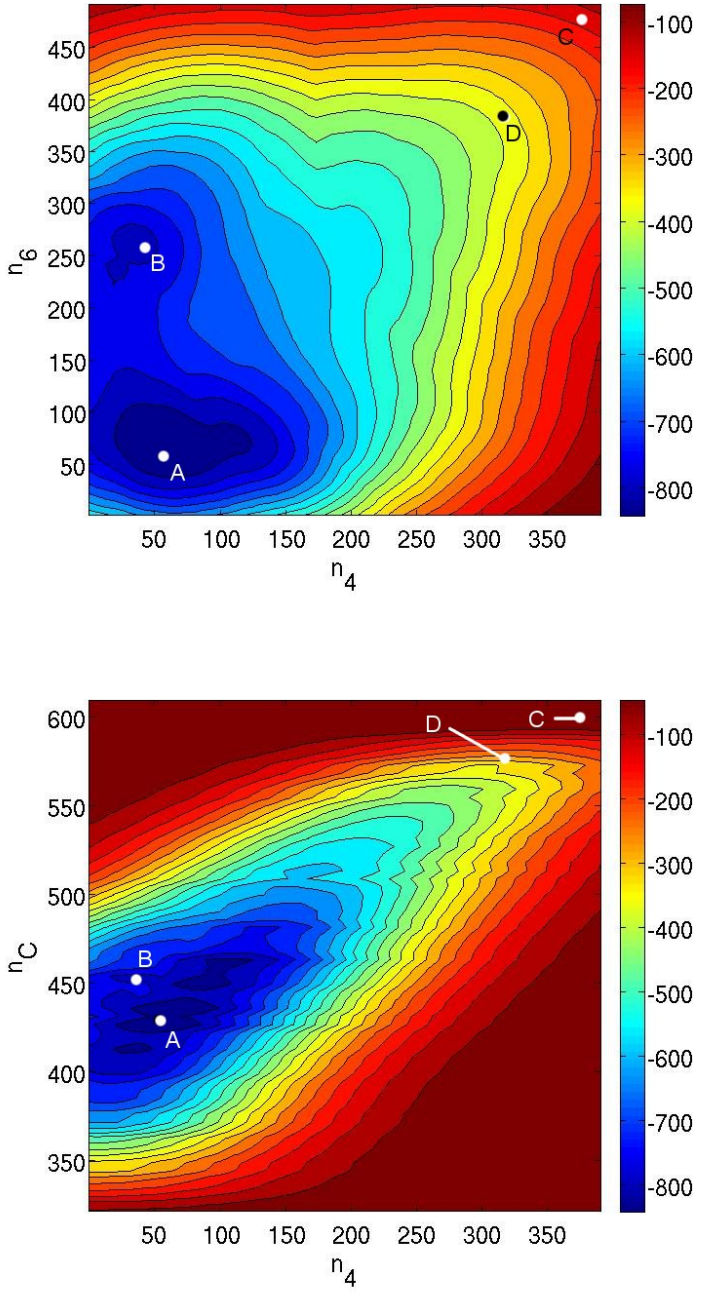

Figure 5.12: Approximations to the FES for the given CVs for a nanoparticle of CdSe. The energy scale for the 2D contour plot is in units of k_BT. Points A, B, C and D correspond to configurations in figure 5.15.

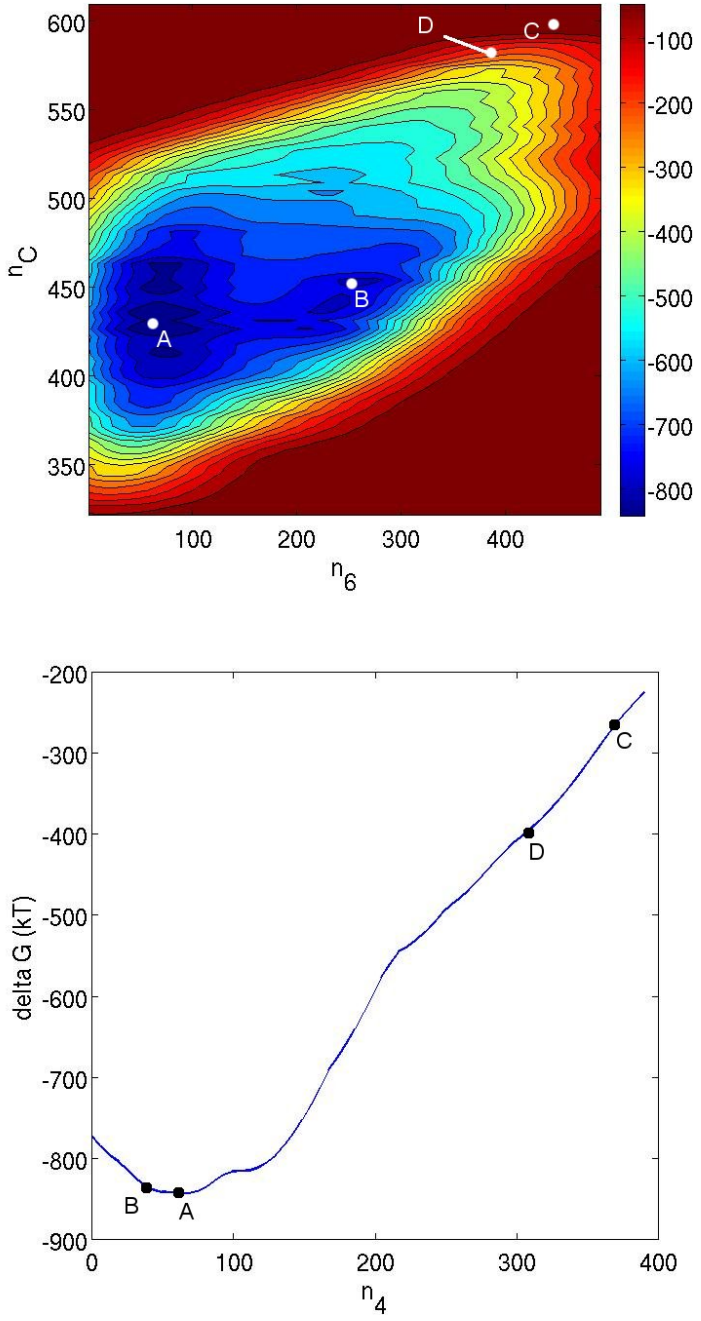

Figure 5.13: Approximations to the FES for the given CVs for a nanoparticle of CdSe. The energy scale for the 2D contour plot is in units of k_BT. Points A, B, C and D correspond to configurations in figure 5.15.

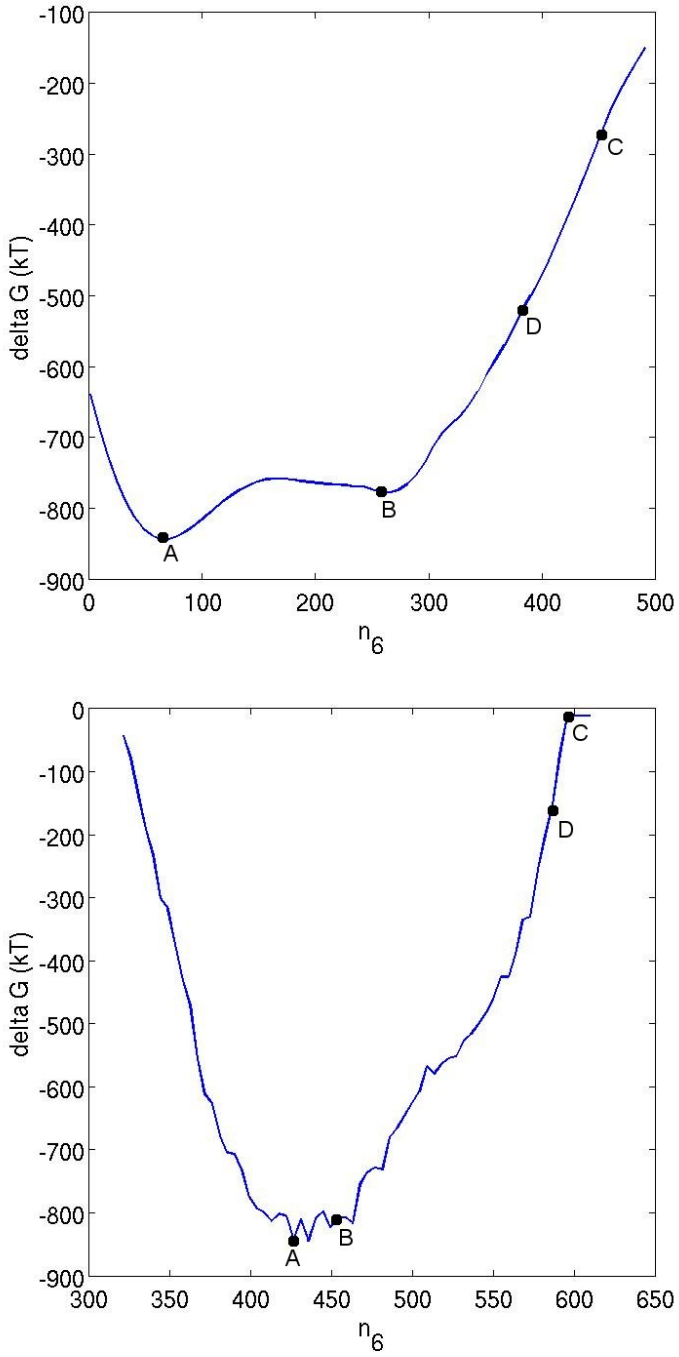

Figure 5.14: Approximations to the FES for the given CVs for a nanoparticle of CdSe. Points A, B, C and D correspond to configurations in figure 5.15.

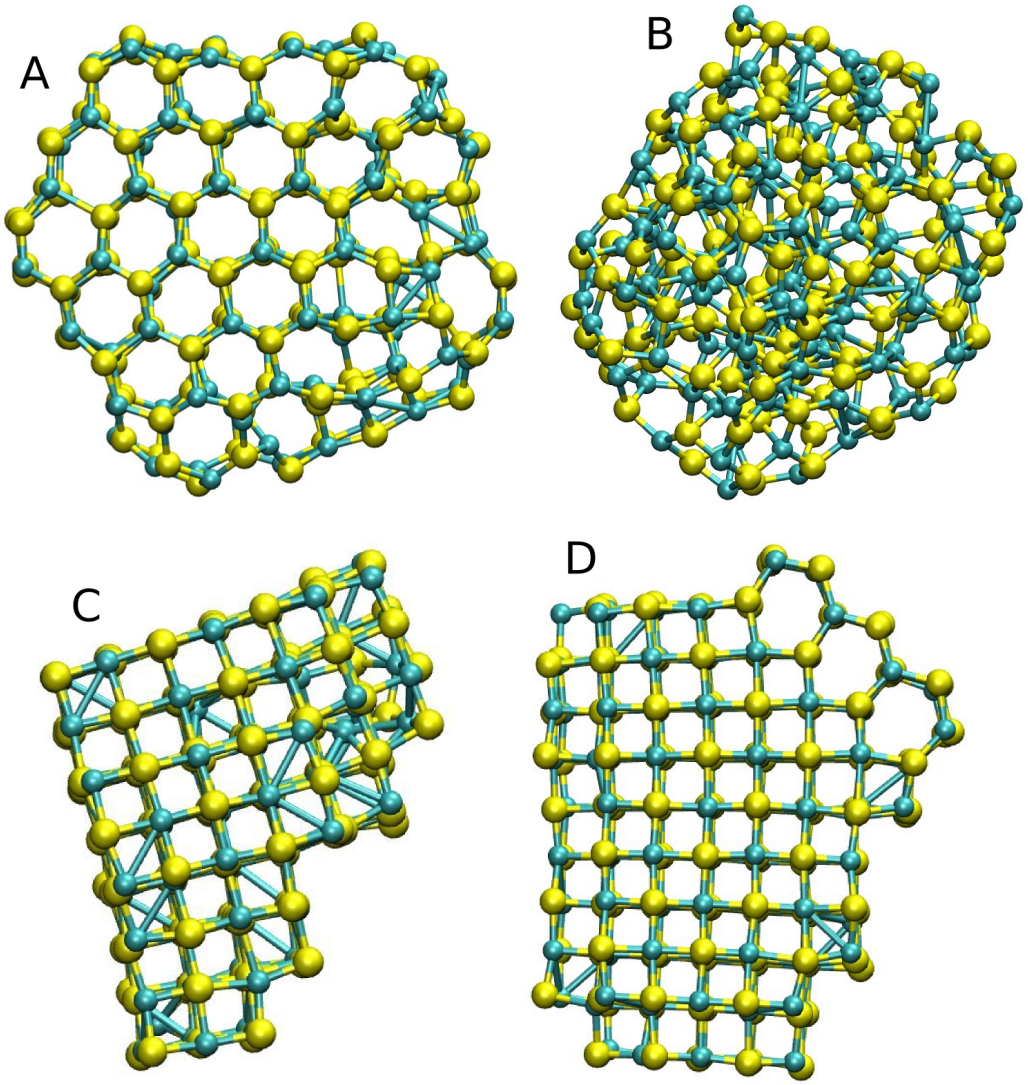

Figure 5.15: CdSe configurations A, B, C and D from the free energy surfaces. A and B are from the identified minima, with A clearly Wurtzite. C and D are example rocksalt configurations.

also be interesting to compare this methodology with that of Pietrucci and Andreoni [2011] on the same systems.

Chapter 6

Nanoparticle Generation Theory

6.1 Introduction

A barrier to understanding polymorph selection in calcium carbonate is a lack of knowledge of which conformations nanoparticle crystals of aragonite or vaterite would adopt. Producing such nanocrystals would facilitate a study of their stability, and is also a useful step towards understanding transitions between such nanocrystals. Although de Leeuw and Parker [1998] have used the Wulff construction methodology to predict the general shape of nanocrystals of these polymorphs, they did not generate structures, and, as extensively discussed in section 1.5, the force-field used does not accurately model the mineral/water interactions.

In this chapter I review common methods of generating low energy nanoparticle conformations that I have implemented, and also describe a novel Monte Carlo method which may have applications well beyond calcium carbonates.

6.2 Wulff Constructions

Wulff constructions (reviewed by Cerf and Picard [2006]) are based on the principle that the lowest energy conformation for a nanocrystal is that which minimises the total surface energy. If edge effects are neglected, it has been proven that this can be done by cleaving an infinite lattice of unit cells in the following manner:

- The surface energies are first calculated. Ideally these would be surface free

energies, however these quantities are not easily accessible. Instead in vacuum I have calculated the potential energy after the surfaces have been relaxed at zero Kelvin, and for hydrated surfaces I have calculated the surface free enthalpy. The methods used are described in section 10.4.1.

- From a start site lines are drawn normal to each surface of length proportional to the energy of that surface (figure 6.1A). At the end point of each line is a plane normal to the line, which represents the ideal location at which to cleave the lattice. The constant of proportionality, or scaling factor, determines the size of the resulting nanoparticle.

There are however several difficulties when this methodology is applied on the scale of interest:

- In reality cleavage planes can only be be defined with atomic resolution as atoms must be either included or excluded, as illustrated in figure 6.1B. In this example, this causes the [10] surface to be be excluded altogether from the resulting structure.
- Edge and corner effects have been neglected — some edges are less energetically favourable than others, and this is not accounted for by this methodology.
- For ionic crystals, there is no method of ensuring the resulting nanoparticle is charge neutral.
- The exposed surfaces are dependent on the choice of start site within the central unit cell.

All these factors become less important with increasing particle size, as small defects or undesirable edges become energetically insignificant.

I have generated Wulff construction morphologies for aragonite for vacuum surface energies, and have attempted to do so for hydrated surfaces. Full details of the surface energy calculations and the Wulff methodology implementation can be found in section 10.4.1.

6.3 Random Structure Searching

Appealingly parsimonious, random structure searching (RSS) essentially entails distributing the atoms of a structure at random, and then finding the nearest local

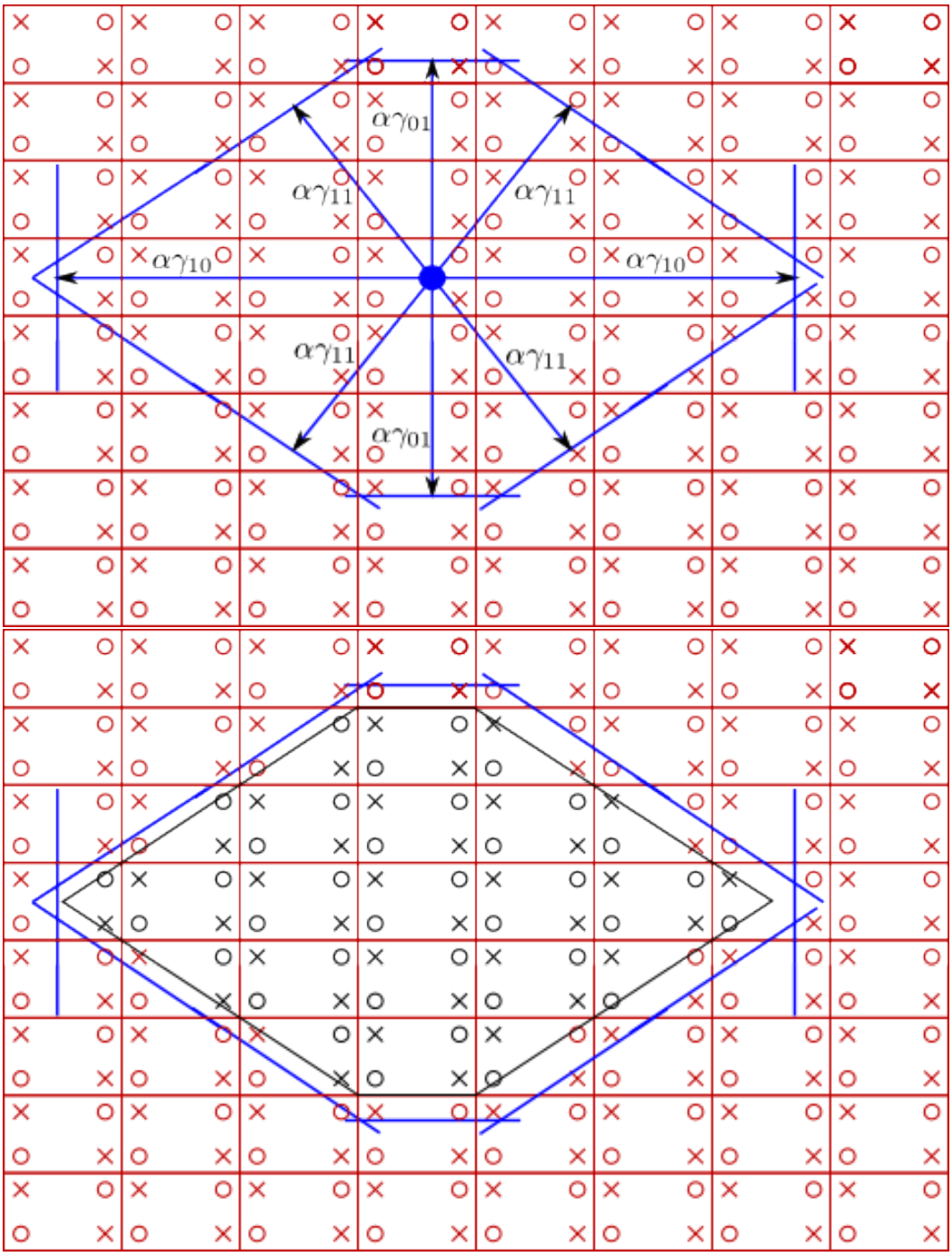

Figure 6.1: **A** Idealised Wulff construction for a 2D lattice with 3 relevant edges, before rounding. **B** Applying rounding to recover nearest applicable edges of A.

energy minimum. Any one structure is unlikely to be close to the ground state for a system, but for small systems it is perfectly possible to try very large numbers of configurations, and be reasonably confident of finding representative minima. The technique has been used both for periodic cells and for nanoparticle clusters [Pickard and Needs, 2011].

For any implementation a key consideration is how 'random' to make the random structures. Introducing biases can drastically reduce the length of the search at the cost of limiting the size of the search space. This can be as basic as whether to treat a CO_3^{2-} ion as a single entity, and position its atoms in their equilibrium positions at a random site, or as 4 separate atoms positioned independently. The latter case is perfectly possible if using *ab initio* methods, and might enable the study of high pressure phases where the carbonate ion has broken down, whereas the former is much more appropriate if using an atomistic force-field where the single carbonate ion is the only species of interest. There is also a choice of whether to set a minimum distance between atoms in the initial configuration. Excluding severely unphysical configurations from the search space would improve the rate of convergence of the energy minimisation tool, but again, limits the search space.

I have attempted some random structure searching to find low energy calcium carbonate nanoparticle conformations: full details can be found in section 10.5.

6.4 The Ion Site Monte Carlo method

Although often appropriate and useful, the methods described in section 6.2 and 6.3 are limited in utility for many purposes.

The proofs that guarantee the Wulff construction is the low energy conformation of a crystal hold only in the limit where edge effects are negligible, which is not the case on the nanoscale. Furthermore, particles can only be produced with certain numbers of formula units — growing the cleavage radius results in the addition of planes of atoms. Many combinations of cleavage radius and origin produce nanoparticles which are not charge neutral. Wulff constructions may also be unsuitable where there is partial disorder within the crystal unit cell, as with vaterite.

With RSS the number of attempts required to obtain a reasonable approximation to the low energy structure scales exponentially with the number of atoms. RSS ignores the information that it is a particular crystalline polymorph that is of interest. As discussed in chapter 7, it is currently completely impractical for the

particle numbers of interest — perhaps 100 atoms is a current feasible limit, though this is currently being investigated within the research group.

Here I present the Ion Site Monte Carlo (ISMC) method suited for the intermediate scale between these two regimes, designed to produce nanoscale crystalline conformations.

The method is essentially very simple. Consider a large 3D grid of unit cells of some crystalline material. Each unit cell has a number of ion, atom or small molecule sites. Only a small number of these sites are populated with actual matter — the rest are left vacant. Monte Carlo moves can be made attempting to move a unit from one location to a vacant site (figure 6.2). The system thus explores a highly constrained ensemble, and this massive reduction in both the number and extent of the degrees of freedom enables the efficient exploration of low energy configurations.

I have implemented this method, combining parallel tempering, configuration bias Monte Carlo, and tabulated potentials to further increase the speed of sampling and hence the discovery of the low energy state. As well as calcite and aragonite versions, I have also produced a vaterite version. This has an additional move type to allow for the partial disorder of the carbonate ions within this polymorph, with the differences in this version described in section 6.4.3. Details of the code and the simulations conducted can be found in section 10.6.

There are disadvantages to the method. The resulting nanoparticles are inherently optimised for vacuum, as the incorporation of realistic solvent would be difficult to achieve. It also does not allow for relaxation of the nanoparticle: it is quite possible that two candidate low energy structures might shift in energetic ordering after relaxation, and this is presently ignored. Nevertheless as shown in chapter 7, the resulting nanoparticles still appear to be lower in energy than Wulff constructions in this size range. It is also possible that when hydrated the second disadvantage mitigates the first: water atoms will help to stabilise under-coordinated mineral ions, limiting the amount of relaxation that occurs.

6.4.1 Configuration Bias MC Scheme

The most straightforward method of choosing moves would be to pick both the ion to move, and the destination, randomly without bias. This would ensure detailed balance (see section 2.5, equation 2.29) was maintained as the generation of the forward and backward moves would be equally probable. From figure 6.2 the limitations of such a scheme can be seen: the move labelled ‘B’ is enormously unlikely

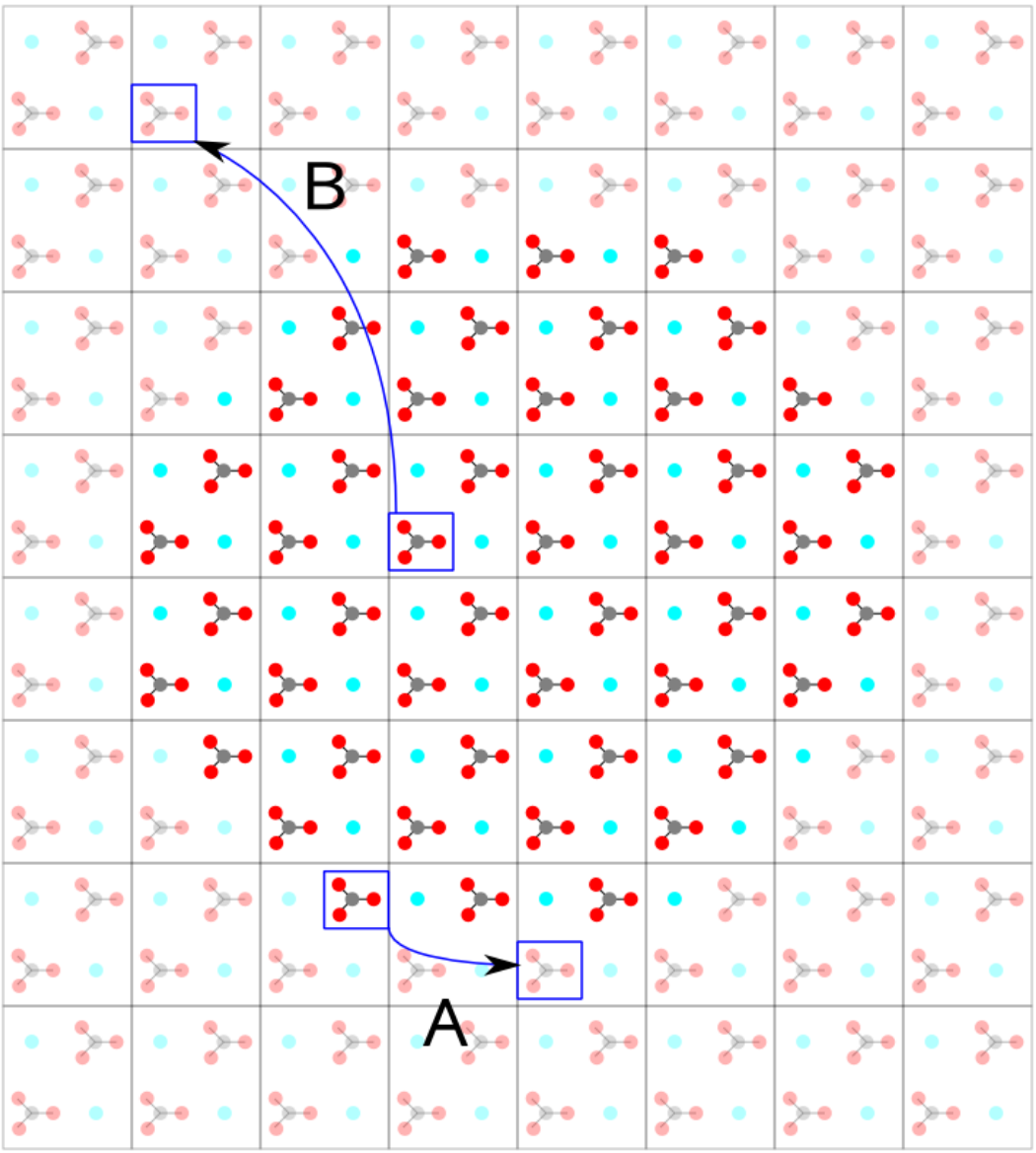

Figure 6.2: Representation of the ISMC lattice with two example trial moves.

to succeed. This is both because of the choice of ion, and the choice of destination. The ion is at the core of the nanoparticle and moving it would create a void, likely to be a high energy configuration. The destination is not part of the main nanoparticle and would result in breaking the nanoparticle into two section, which again is likely to be energetically unfavourable. In contrast the move labelled 'A' involves moving an ion from the the edge of the nanoparticle to another site on the edge of the nanoparticle, and is likely to be much more energetically favourable. Since most of the ions are below the surface of the nanoparticle and most of the possible destinations do not adjoin it, the system is much more likely to generate moves of type 'B' than type 'A', and hence waste computational effort.

The use of a configuration bias scheme can vastly improve the efficiency of sampling. Recall from section 2.5 that when implementing a CBMC scheme and preserving detailed balance, it is necessarily to know exactly both $p_G(o \to n)$, and the $p_G(n \to o)$, the probabilities of generating both the forward and backward moves. I improve on the random scheme in both the choice of ion to move and the choice of destination, and hence separate p_G, into the product of p_{CI}, and p_{CD}, the probability of choosing ion and destination respectively.

For both of these improvements it is necessary to track the number of neighbours of each site. In bulk calcite and aragonite, both the Ca^{2+} and CO_3^{2-} ions are twelve-fold coordinated - that is have 12 surrounding ions of the same kind, as described in more detail in section 1.1. For every site (occupied or void) the number of these 12 neighbouring sites that are occupied is stored, and updated when a move is attempted. In what follows N denotes the total number of formula units spread over the lattice.

Choice of ion to Move

For most contiguous nanoparticles, only a proportion of the occupied sites will be exposed at the surface of the structure. Before the program is started an unnormalised histogram $h(m),\ m \in \{0, \dots, 12\}$ is built of how many occupied sites have each number of neighbours (figure 6.3A). A discrete probability distribution $p(X = n),\ n \in \{1, \dots, N\}$ is defined with a skew towards the low end of the distribution. It is this skew that defines the extent of the bias in the sampling. A suitable choice might be a cosine curve on the range $(0, \frac{\pi}{2})$, suitably stretched and normalised, or a simple linear curve such as $p(X = n) = 3N - 2n$, again suitably normalised. In order to simplify implementation the choice should specify a function

with $p(X = n) > 0$ for all n. The curve is then converted to a cumulative distribution curve $F(n) = p(X \leq n)$. The probability of picking an ion with m neighbours is then set at:

$$p_{CI}(Y = m) = F\left(\sum_{i=1}^{m} h(i)\right) - F\left(\sum_{i=1}^{m-1} h(i)\right) \tag{6.1}$$

This is implemented by picking a uniform random number $x_0 \in (0, 1)$ and finding the lowest m such that $F\left(\sum_{i=1}^{m+1} h(i)\right) > x_0$ (figure 6.3B), and then picking sites at random without bias until one with the correct number of neighbours is found. The p_{CS} for the backwards move is evaluated from equation 6.1 using the number of neighbours of the new site.

Choice of destination

The 3D grid of unit cells must be large to ensure the nanoparticle conformation is not constrained. Picking a destination site in the grid at random without bias is then very inefficient, as most moves would take a unit from a contiguous nanoparticle to a far removed location, in what tends to be a very high energy, and hence improbable configuration. Sampling can be improved in efficiency by biasing the choice of destination to areas immediately bordering the nanoparticle.

To do this, a border list is maintained of void sites which have occupied neighbours. This is fairly easy to update when a move is made. Around the old site the neighbour count of all the neighbouring sites is decremented and if a void site no longer has any neighbours it is removed from the list. At the new site, the neighbour count of the neighbouring sites is incremented, and if any of these sites is both void and has exactly one neighbour it is added to the list. If it has more than one neighbour it would already be on the border list.

The destination site is chosen from the border list with probability p_S, typically around 95% — otherwise a void site is chosen from those sites without occupied neighbours. No attempt is made to bias selection within the surface list itself — a site is chosen from the N_s elements at random. If the full grid has N_G sites in total, and $x_0 \in (0, 1)$ again a (new) uniform random number:

$$p_{CD} = \begin{cases} \dfrac{p_S}{N_S} & \text{if } x_0 < p_S \\ \dfrac{1 - p_S}{N_G - N - N_S} & \text{if } x_0 > p_S \end{cases} \tag{6.2}$$

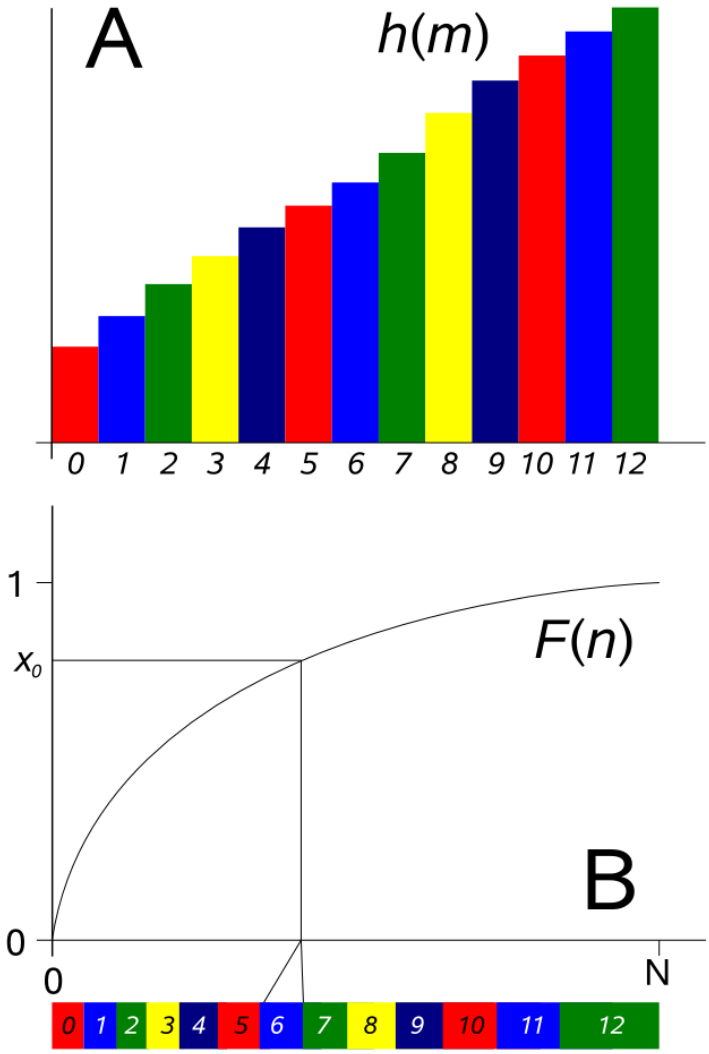

Figure 6.3: Choice of ion.**A.** Cartoon histogram $h(m)$ for an ISMC nanoparticle. **B.** Picking a site. Random number x_0 generated, then correct column of $h(m)$ (6 in this case) found. Thus an ion with 6 neighbours would be attempted to be moved in this example.

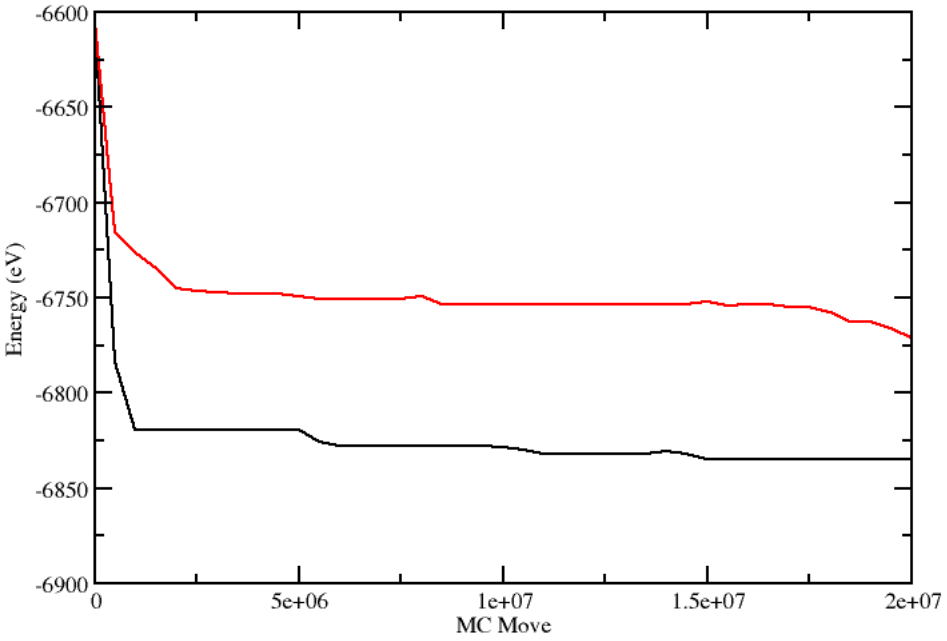

Figure 6.4: Energy of nanoparticle over course of ISMC simulations at 5000 K, for unbiased (RED) and biased (BLACK) move selection criteria.

Efficiency of CBMC scheme

To illustrate the efficiency of the CBMC scheme, I ran two short simulations of 147 formula units of calcite, each beginning from the same roughly spherical initial configuration. Of 20,000,000 move attempts at 5000 K, the version of the code with bias accepted 303. This is a very small fraction, but nevertheless represents a significant improvement over the unbiased variant where only 125 moves were accepted. Figure 6.4 illustrates the difference this made to the search for a low energy configuration — the potential energy of the nanoparticle generated using biased move selection criteria was substantially lower throughout the simulation.

6.4.2 Parallel tempering

As discussed in the section 2.5.3, parallel tempering is a technique which facilitates accelerated exploration of an ensemble. Multiple simulations are ran in parallel at different temperatures, and depending on the energy differences between them, configurations are swapped in a process called replica exchange. This was implemented into the ISMC programs.

6.4.3 Partial Disorder Version

Vaterite, in common with many materials, has considerable disorder within the unit cell; each carbonate ion can point in any one of three directions. This makes the surfaces of vaterite extremely complicated — some local orientational ordering may or may not occur at each surface for example, and might permeate some layers through the structure. The ISMC method was modified slightly, to cope with such disorder. This was done by introducing the additional move type of altering the orientation of a carbonate group. No attempt at biasing the move was made: a random carbonate ion was reorientated, the energy for the new configuration calculated, compared to the old energy, and accepted according to the standard Metropolis criteria. The frequency of rotation as opposed to translation moves is an adjustable parameter, with 10% of moves the value typically used. The extra degrees of freedom does enlarge the size of the potential table, but the full table is still manageably small.

The unit cell used in the vaterite code is that of Kamhi [1963], though some details of the vaterite structure remain controversial — see section 1.1.3 for details. This cannot be relaxed using the potential beforehand because of the local disorder.

The other difference between the vaterite version and the other two polymorphs is the temperature progression scheme. As discussed in section 6.4.2 extremely high temperatures are necessary to give the high energy ion displacement moves a chance of success. The reorientation moves have a much smaller effect on the energy however, and at 3000 K, the carbonate orientations there is too much thermal energy for any real ordering. Instead a linear temperature progression was used from 1 K to 12000 K across the different processors, which allows low temperature simulations to anneal carbonate orientations in nanoparticle morphologies generated by high temperature simulations.

6.4.4 Applicability of the technique to the study of voids

As well as generating low energy nanoparticle conformations, the ISMC technique also predicts the shape of vacuum voids in bulk aragonite and calcite, subject to the usual assumption that relaxation is irrelevant. Consider a large block of either aragonite or calcite, as a subsection of an effectively infinite bulk lattice. This block has lattice energy designated E. Removing one calcium ion, denoted j, would lead to a correction in the energy:

$$E_{\mathrm{Ca}} = \sum_{\text{all ions } i,\neq j} U_{ij} \tag{6.3}$$

This quantity is independent of which calcium ion j, is removed for these polymorphs when all sites are identical. Similarly, removing a single carbonate ion would lead to a correction $E_{\mathrm{CO_3}}$ independent of site choice. If however one removes both a single calcium and a single carbonate ion then the energy of the block would be:

$$E_{\mathrm{tot}} = E - E_{\mathrm{Ca}} - E_{\mathrm{CO_3}} + U_{ij} \tag{6.4}$$

The additional term U_{ij} is necessary as this term has been counted twice in the E_X terms. Note that if one seeks to minimise E_{tot} through choosing different ions to remove, this reduces to minimising U_{ij}. This is exactly identical to the ISMC formulation for a single formula unit. Extending inductively, it follows that for n calcium and carbonate vacancies:

$$E_{\mathrm{tot}} = E - nE_{\mathrm{Ca}} - nE_{\mathrm{CO_3}} + \sum_{\text{all vacancy pairs } i,j} U_{ij} \tag{6.5}$$

Again, on the right hand side all terms are constant except the summation over all vacancy pairs. This term is again the same quantity that the ISMC algorithm seeks to minimise for n formula units. Subject to the assumption that relaxation around the void is irrelevant, the void shape minimising the total energy is simply the 'negative image' of the nanoparticle obtained through the regular procedure.

This argument is not valid for the partially disordered vaterite cells, where the quantities E_{Ca} and $E_{\mathrm{CO_3}}$ are dependent on the orientation of the surrounding carbonates and hence are not constant. Even in this case, the negative image seems likely to present a good first approximation if one does not require information on the surrounding carbonate group orientations.

Chapter 7

Nanoparticle generation results

In this chapter I present the results of the calcium carbonate nanoparticle generation methods, and investigate the stability of the resulting nanoparticles in water.

7.1 Surface energies and Wulff Constructions

7.1.1 Vacuum surface energies

In order to proceed with the Wulff construction the surface energies in vacuum were calculated as described in section 10.4.1. Tables 7.1 and 7.2 give the potential energy values both before and after the geometry of the surfaces has been optimised to minimise their energy.

Table 7.3 shows the 10 lowest energy surfaces when relaxed, together with the energy of the calcite $(10\bar{1}4)$ surface. Many of the aragonite surfaces have similar energies, and none are as low in energy as the calcite surface. With the bulk energies of calcite and aragonite very similar, these higher energy surfaces imply nanoparticles of aragonite will be much higher in potential energy and hence less stable than calcite nanoparticles.

7.1.2 Vacuum Wulff Constructions

Figure 7.1 shows the idealised large-size shape of the Wulff construction for aragonite based on the vacuum surface energies calculated above. The exposed surfaces are (101), (001), (011), (210) and (211).

Using all these surfaces, I attempted to produce charge neutral nanoparticles

			Surface Energy (Jm^{-2})	
Surface	**Shift**	**Dhkl (Å)**	**Unrelaxed**	**Relaxed**
$(10\bar{1})$	0.00	4.56	1.574	1.204
$(10\bar{1})$	0.25		2.242	1.339
$(10\bar{1})$	0.50		0.995	0.841
$(10\bar{1})$	0.75		2.242	1.246
(011)	0.00	4.22	4.400	1.224
(011)	0.50		1.357	1.035
(001)	0.00	3.99	1.358	1.147
(001)	0.25		3.904	1.496
(111)	0.00	3.36	2.947	1.321
(111)	0.50		1.820	1.221
$(10\bar{2})$	0.00	3.24	2.134	1.302
$(10\bar{2})$	0.42		3.263	1.295
$(10\bar{2})$	0.50		3.019	1.862
$(10\bar{2})$	0.58		3.263	1.295
(100)	0.00	2.77	1.198	1.016
(100)	0.25		2.648	1.614
$(11\bar{2})$	0.00	2.72	2.609	1.289
$(11\bar{2})$	0.26		2.356	1.521
$(11\bar{2})$	0.50		2.165	1.323
$(11\bar{2})$	0.74		2.356	1.215
(201)	0.00	2.62	2.625	1.085
(201)	0.50		1.857	1.114
(010)	0.00	2.49	2.486	1.742
(210)	0.00	2.42	1.133	0.878
(210)	0.50		1.281	0.936
(103)	0.00	2.40	3.557	1.279
(103)	0.28		4.029	1.469
(103)	0.50		3.309	1.638
(103)	0.72		4.029	1.437
(013)	0.00	2.35	2.551	1.641
(013)	0.50		2.884	1.365
$(2\bar{1}\bar{1})$	0.00	2.32	2.142	1.173
$(2\bar{1}\bar{1})$	0.17		2.395	1.316
$(2\bar{1}\bar{1})$	0.50		1.211	0.896
$(2\bar{1}\bar{1})$	0.83		2.239	1.204
$(1\bar{2}1)$	0.00	2.18	2.881	1.415
$(1\bar{2}1)$	0.25		3.337	1.166
$(1\bar{2}1)$	0.50		2.549	1.567
$(1\bar{2}1)$	0.75		3.337	1.385

Table 7.1: Surface energies for the surfaces for the 14 highest $D\{hkl\}$ values (the distance between two adjacent lattice planes)

			Surface Energy (Jm^{-2})	
Surface	**Shift**	**Dhkl (Å)**	**Unrelaxed**	**Relaxed**
$(11\bar{3})$	0.00	2.16	2.380	1.506
$(11\bar{3})$	0.25		2.805	1.530
$(11\bar{3})$	0.50		2.241	1.397
$(11\bar{3})$	0.75		2.805	1.380
$(2\bar{1}\bar{2})$	0.00	2.07	1.792	1.196
$(2\bar{1}\bar{2})$	0.40		2.007	1.170
$(2\bar{1}\bar{2})$	0.50		2.395	1.107
$(2\bar{1}\bar{2})$	0.60		2.007	1.230
$(1\bar{2}\bar{2})$	0.00	1.97	4.509	1.262
$(1\bar{2}\bar{2})$	0.42		3.011	1.317
$(1\bar{2}\bar{2})$	0.50		1.626	1.104
$(1\bar{2}\bar{2})$	0.58		3.011	1.319
$(20\bar{3})$	0.00	1.92	2.844	1.295
$(20\bar{3})$	0.25		2.569	1.266
$(20\bar{3})$	0.50		2.844	1.295
$(20\bar{3})$	0.75		2.569	1.266
$(10\bar{4})$	0.00	1.88	4.521	1.477
$(10\bar{4})$	0.50		2.014	1.406
$(2\bar{2}0)$	0.00	1.88	1.703	1.200
$(2\bar{2}0)$	0.25		1.883	1.329
$(2\bar{2}1)$	0.00	1.80	2.003	1.300
$(2\bar{2}1)$	0.50		2.094	1.371
$(30\bar{1})$	0.00	1.80	1.736	1.111
$(30\bar{1})$	0.22		3.382	1.308
$(30\bar{1})$	0.50		2.948	1.293
$(30\bar{1})$	0.78		3.382	1.184
$(21\bar{3})$	0.00	1.79	2.048	1.275
$(21\bar{3})$	0.50		2.673	1.315
$(11\bar{4})$	0.00	1.76	2.695	1.501
$(11\bar{4})$	0.23		2.657	1.524
$(11\bar{4})$	0.50		2.292	1.404
$(11\bar{4})$	0.77		2.657	1.508
$(1\bar{2}3)$	0.00	1.73	4.635	1.217
$(1\bar{2}3)$	0.22		3.687	1.422
$(1\bar{2}3)$	0.50		2.635	1.338
$(1\bar{2}3)$	0.78		3.687	1.412

Table 7.2: Surface energies for the surfaces for the next 11 highest values $D\{hkl\}$ (the distance between two adjacent lattice planes)

Aragonite Surface	**Shift**	**Surface Energy (Jm^{-2})**
$(10\bar{1})$	0.50	0.8406
(210)	0.00	0.8780
$(2\bar{1}\bar{1})$	0.50	0.8955
(210)	0.50	0.9358
(100)	0.00	1.0161
(011)	0.50	1.0350
(201)	0.00	1.0848
$(1\bar{2}2)$	0.50	1.1041
$(2\bar{1}\bar{2})$	0.50	1.1068
$(30\bar{1})$	0.00	1.1115
(201)	0.50	1.1136
(001)	0.00	1.1467
$(1\bar{2}1)$	0.25	1.1662
Calcite Surface	**Shift**	**Surface Energy (Jm^{-2})**
$(10\bar{1}4)$	0.00	0.7113

Table 7.3: Lowest energy surfaces for aragonite with calcite $(10\bar{1}4)$ surface for comparison

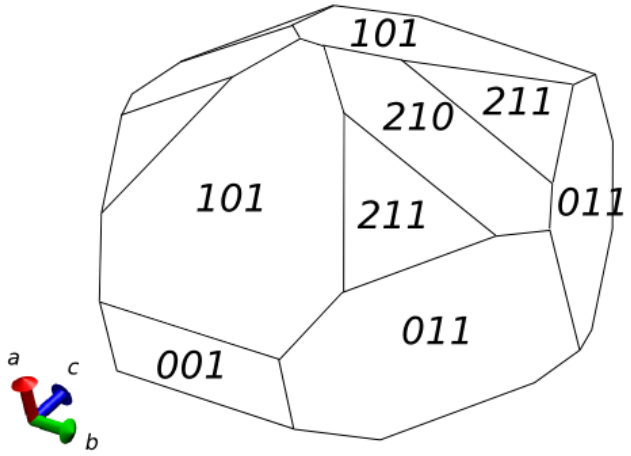

Figure 7.1: Vacuum Wulff construction morphology for aragonite, using the potential of Raiteri and Gale [2010]

using the program described in section 10.4.1, however, none could be found. This is a result of the finite size effects discussed in section 6.2: as the nanoparticle grows the significance of any small charge on the nanoparticle would diminish. It was necessary to exclude the (001) surfaces from the program, which after an extensive search (varying origin within the unit cell and base cleave radius) lead to some charge neutral nanoparticles, of sizes 75, 82, 84, 85, 114, 127 and 162 formula units. Figure 7.2 shows example configurations. The energy of these nanoparticles is discussed in section 7.4, and atomic configurations can be found in appendix B.

7.1.3 Hydrated surface energies

These simulations were initially conducted under the assumption that the entropic contribution to the surface free energy would be negligible, and hence the surface enthalpy (as defined in equation 10.1) would be a reasonable estimate of this quantity. Intriguingly however, many of the surfaces have negative surface enthalpy (table 7.4, plotted in figure 7.3). If the surface free energies were enthalpy dominated, this would imply that aragonite would be more stable as a clay with a series of films of water and aragonite, or, because so many surfaces have negative energy, perhaps as either free ions or ion pairs in solution. It appears that the water—mineral bonds are stronger than the water—water bonds. This is consistent with the negative enthalpy of dissolution of calcite (-12.54 kJ/mol).

There is however significant ordering of the water around these surfaces, with figure 7.4 showing a particularly extreme example. In this case water has penetrated through the first layer of calcium carbonate ions, and remains there over the course of the 5 ns simulation. There are several layers of ordered water, and the configurations of two of these are shown in 7.5 with a disordered layer for comparison. The disordered layer has small density fluctuations around the bulk value of 1000 kg/m^3, whereas for the ordered layers most x, y locations have zero density, with occasional spikes. It appears that the entropic contribution of such a layer is far from negligible. The entropy of a surface is difficult to evaluate directly, but we may estimate its order of magnitude by considering the enthalpy of fusion of ice, which has a value of 334.0 kJ/kg [Lide, 2004]. For a layer of ice 5 Å thick this equates to an addition to the surface free energy of 0.168 Jm^{-2} at 300 K from the entropic term. Methods exist of quantifying the true surface free energy for solid-liquid interfaces [Handel et al., 2008], but unfortunately it was not possible to apply them in the time frame available. Until surface free energy estimates are obtained

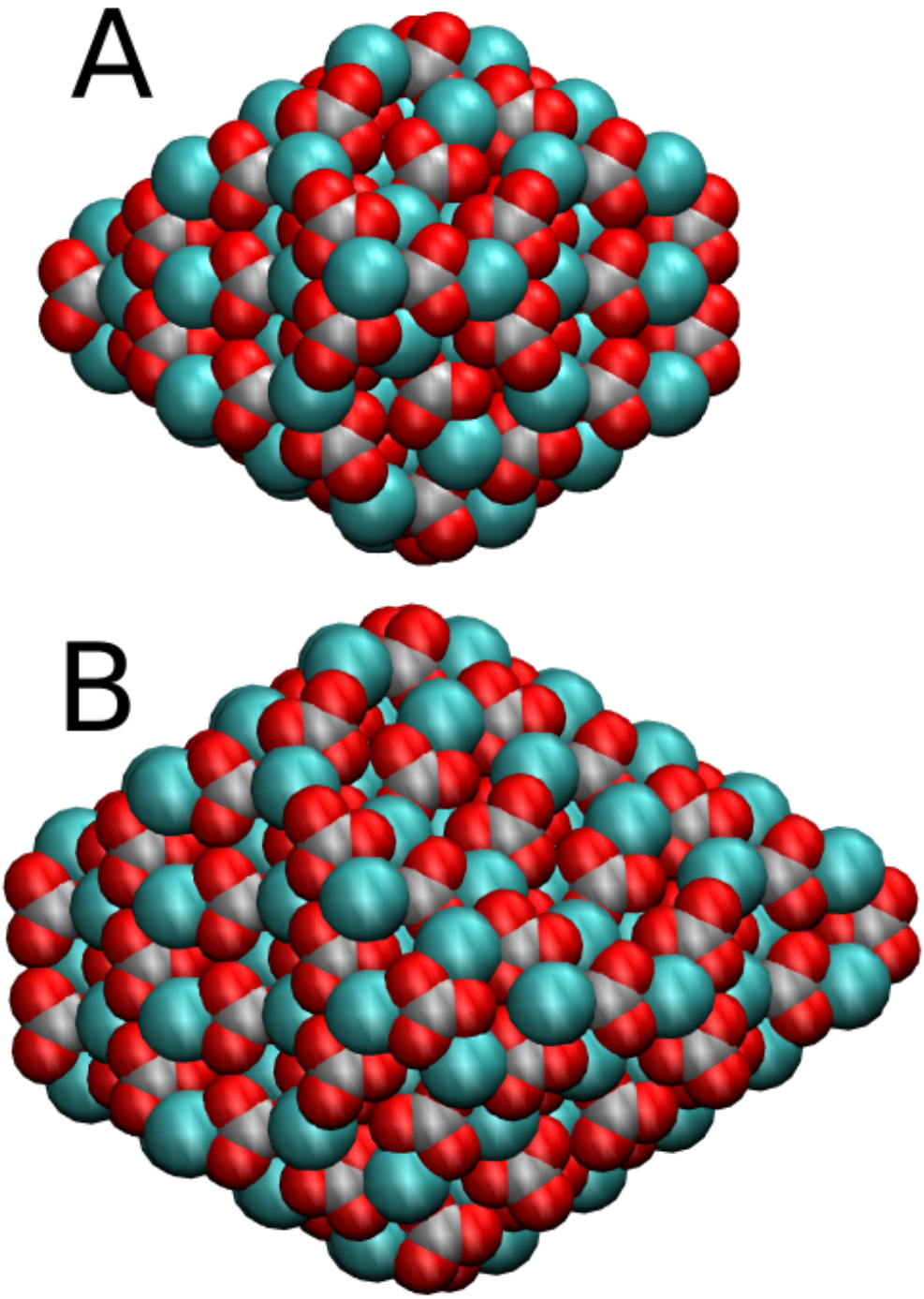

Figure 7.2: Aragonite Wulff constructions for (**A**) 85 and (**B**) 162 formula units.

Surface	Shift	Dhkl (Å)	Surface Enthalpy (Jm^{-2})	Error
(101)	0.00	4.56	-0.032	0.012
(101)	0.25		-0.058	0.012
(101)	0.50		-0.007	0.011
(101)	0.75		0.002	0.018
(011)	0.00	4.22	-0.172	0.018
(011)	0.50		-0.007	0.014
(001)	0.00	3.99	-0.052	0.014
(001)	0.25		-0.090	0.015
(111)	0	3.36	-0.092	0.0171
(111)	0.5		0.015	0.0238
(102)	0.00	3.24	-0.105	0.014
(102)	0.08		-0.030	0.013
(102)	0.50		0.002	0.013
(102)	0.91		-0.130	0.012
(100)	0.00	2.77	0.025	0.012
(100)	0.25		0.039	0.044
(112)	0	2.72	-0.006	0.0192
(112)	0.26		0.011	0.0172
(112)	0.5		-0.038	0.0200
(112)	0.72		-0.001	0.0184
(201)	0.00	2.62	-0.063	0.014
(201)	0.50		-0.031	0.016
(010)	0.00	2.49	0.002	0.013
(210)	0.00	2.42	0.013	0.014
(210)	0.50		-0.068	0.016
(103)	0.00	2.40	-0.092	0.013
(103)	0.28		-0.081	0.012
(103)	0.50		-0.090	0.016
(103)	0.71		-0.123	0.012
(013)	0.00	2.35	-0.014	0.013
(013)	0.50		-0.185	0.012
(211)	0	2.32	-0.007	0.0159
(211)	0.18		-0.031	0.0158
(211)	0.5		-0.039	0.0169
(211)	0.82		-0.088	0.0184
(121)	0	2.18	-0.022	0.0148
(121)	0.25		-0.038	0.0163
(121)	0.5		-0.025	0.0159
(121)	0.75		-0.038	0.0157

Table 7.4: Hydrated surface enthalpies for the highest D(hkl) aragonite surfaces

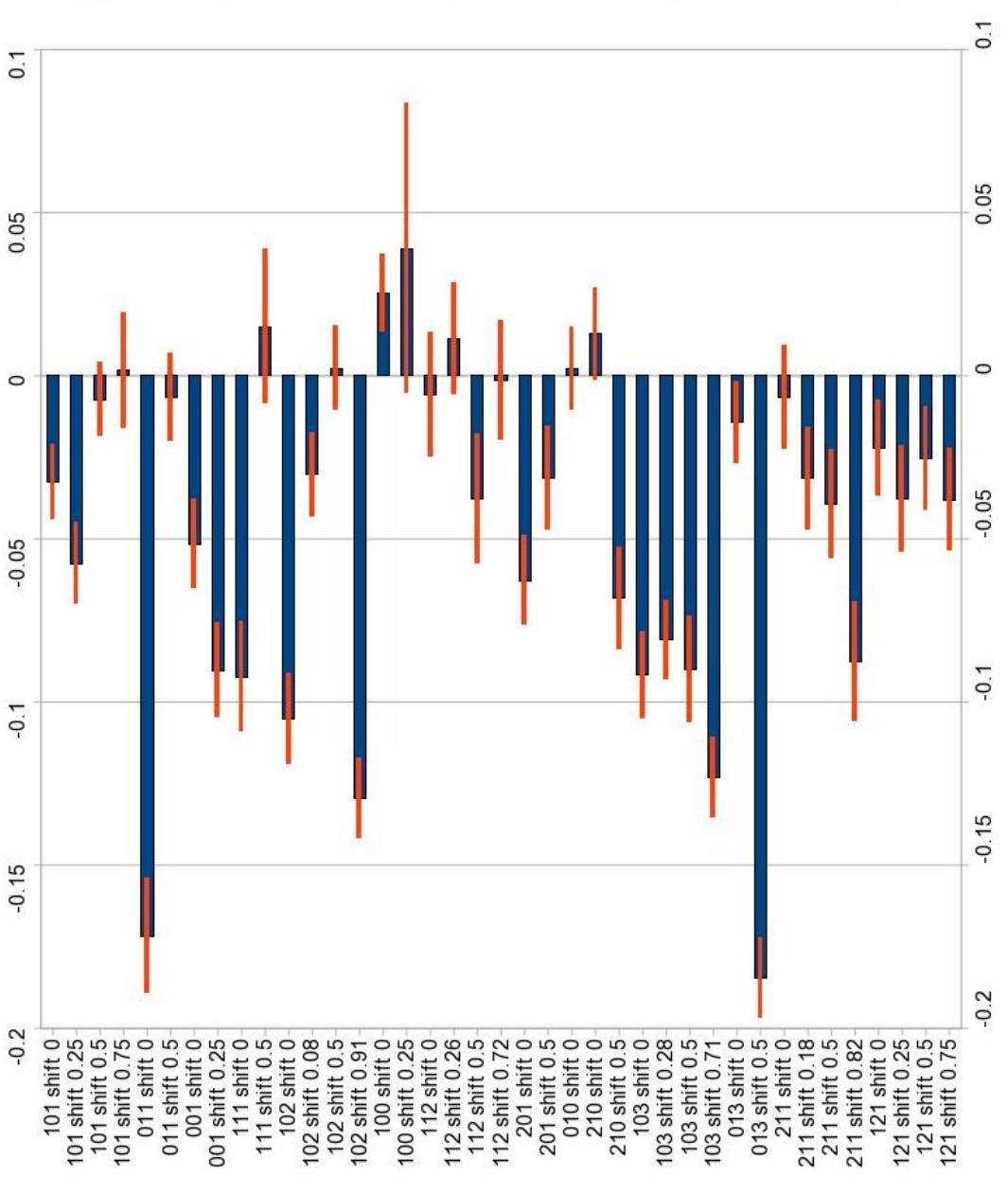

Figure 7.3: hydrated surface enthalpies for the highest D(hkl) aragonite surfaces.

however, it will be impossible to apply the Wulff construction methodology to find nanoparticle morphologies specific to hydrated surfaces. de Leeuw and Parker [1998] were only able to calculate such morphologies because their potential model used different terms and hence produced positive surface energies.

Separately, the enthalpy of the calcite $(10\bar{1}4)$ surface was calculated using the same procedure. The value obtained was -0.010 Jm^{-2} with an error of 0.015 Jm^{-2}, and again it appears that the entropic penalty of ordered water layers must significantly contribute to the surface free energy. Whereas for the vacuum surface energies the calcite surface was the lowest in surface energy, for the hydrated systems most of the aragonite surfaces explored have lower surface enthalpy than the calcite $(10\bar{1}4)$ surface. If this is still the case after correcting for entropic effects, this has significant implications for crystallisation: calcite nanoparticles would only become stable with respect to aragonite when large enough for the bulk free energy difference to overcome the penalty of the unfavourable calcite surface.

Calculation of the hydration energies (figure 7.5) for aragonite confirms the strength of the water—mineral interactions. All the hydration enthalpies are strongly negative, demonstrating the energetic favourability of a hydrated surface over a bare vacuum surface. These values are again enthalpies rather than free energies, and hence it would be necessary to correct for entropic effects to get the most relevant quantities.

As well as entropic effects, there is another possibility concerning the calculated negative surface energies which should be considered. It may be the case that these values best represent not the importance of entropy to surface energy calculations, but a failure of the potential used. Any empirical force-field, no matter how carefully it has been fitted, can only give an approximation to the full quantum mechanical forces experienced by atoms and any attempt to apply a forcefield to a situation to which it has not been fitted must be viewed with caution, as in this case. The water—mineral interactions of the force-field of Raiteri and Gale [2010] were not fitted using mineral surfaces but by using single carbonate or calcium ions in solution. A mineral ion coordinated entirely by water is in a different environment from one coordinated partly by other mineral ions and partly by water. The results are not necessary implausible, but they are surprising, and a degree of caution is warranted.

Surface	Shift	Dhkl (Å)	Hydration Enthalpy (Jm^{-2})	Error
(101)	0.00	4.56	-1.24	0.016
(101)	0.25		-1.40	0.018
(101)	0.50		-0.85	0.016
(101)	0.75		-1.24	0.025
(011)	0.00	4.22	-1.40	0.025
(011)	0.50		-1.04	0.019
(001)	0.00	3.99	-1.20	0.020
(001)	0.25		-1.59	0.021
(111)	0	3.36	-1.41	0.024
(111)	0.5		-1.21	0.034
(102)	0.00	3.24	-1.41	0.020
(102)	0.08		-1.33	0.018
(102)	0.50		-1.86	0.018
(102)	0.91		-1.42	0.018
(100)	0.00	2.77	-0.99	0.017
(100)	0.25		-1.57	0.063
(112)	0	2.72	-1.29	0.027
(112)	0.26		-1.51	0.024
(112)	0.5		-1.36	0.028
(112)	0.72		-1.22	0.026
(201)	0.00	2.62	-1.15	0.019
(201)	0.50		-1.15	0.022
(010)	0.00	2.49	-1.74	0.018
(210)	0.00	2.42	-0.87	0.020
(210)	0.50		-1.00	0.022
(103)	0.00	2.40	-1.37	0.019
(103)	0.28		-1.55	0.017
(103)	0.50		-1.73	0.023
(103)	0.71		-1.56	0.018
(013)	0.00	2.35	-1.66	0.018
(013)	0.50		-1.55	0.018
(211)	0	2.32	-1.18	0.023
(211)	0.18		-1.35	0.022
(211)	0.5		-0.93	0.024
(211)	0.82		-1.29	0.026
(121)	0	2.18	-1.44	0.021
(121)	0.25		-1.20	0.023
(121)	0.5		-1.59	0.022
(121)	0.75		-1.42	0.022

Table 7.5: Hydration enthalpy for the highest D(hkl) aragonite surfaces

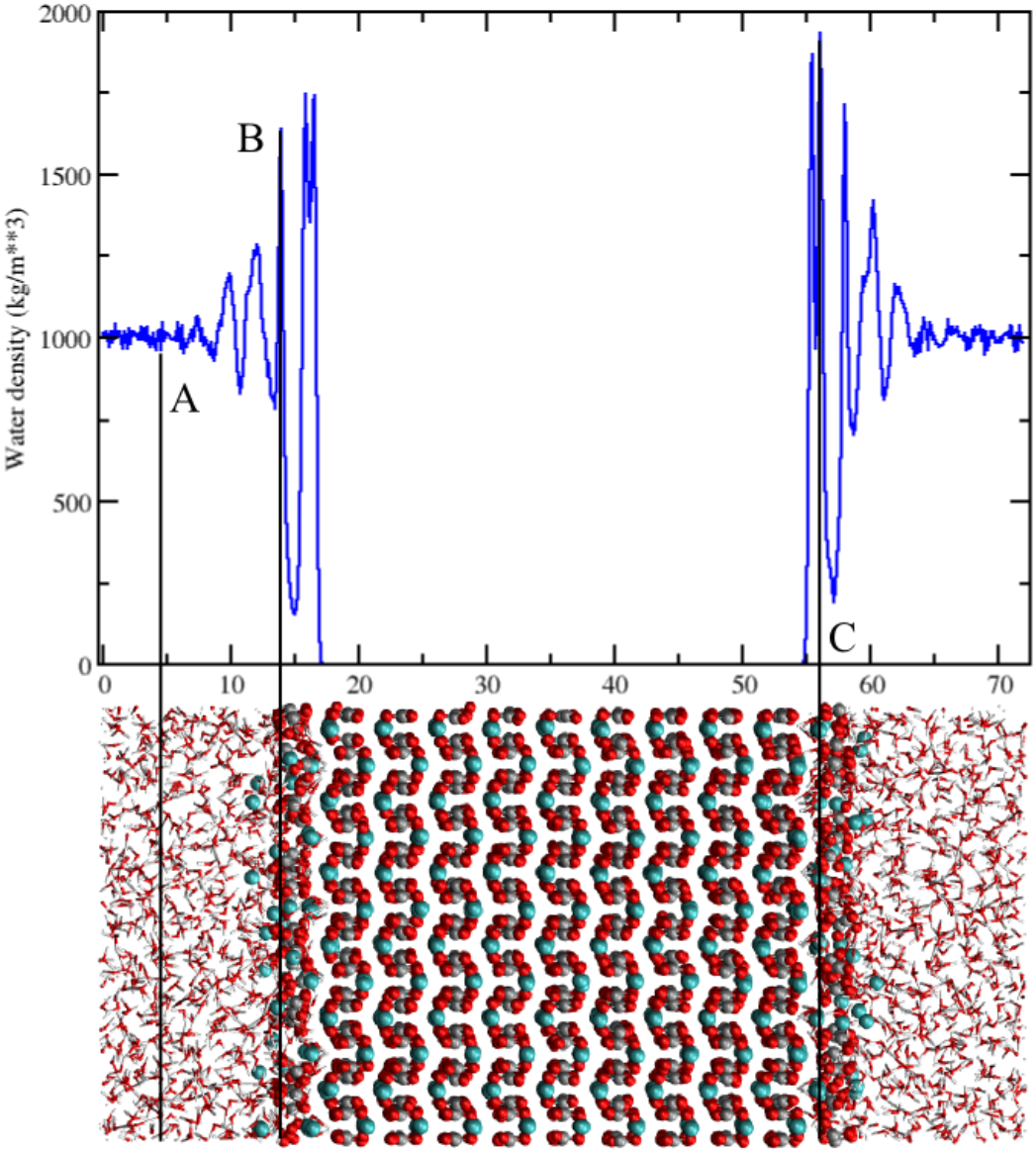

Figure 7.4: Water Z-density for the (011) aragonite surface (shift 0.0), averaged over 2 ns trajectory, with aligned sample configuration. The penetration of the water into the first layer of calcium carbonate structure is clearly visible. Layer A, B and C refer to figures 7.5 and 7.6 where the x, y water density is considered.

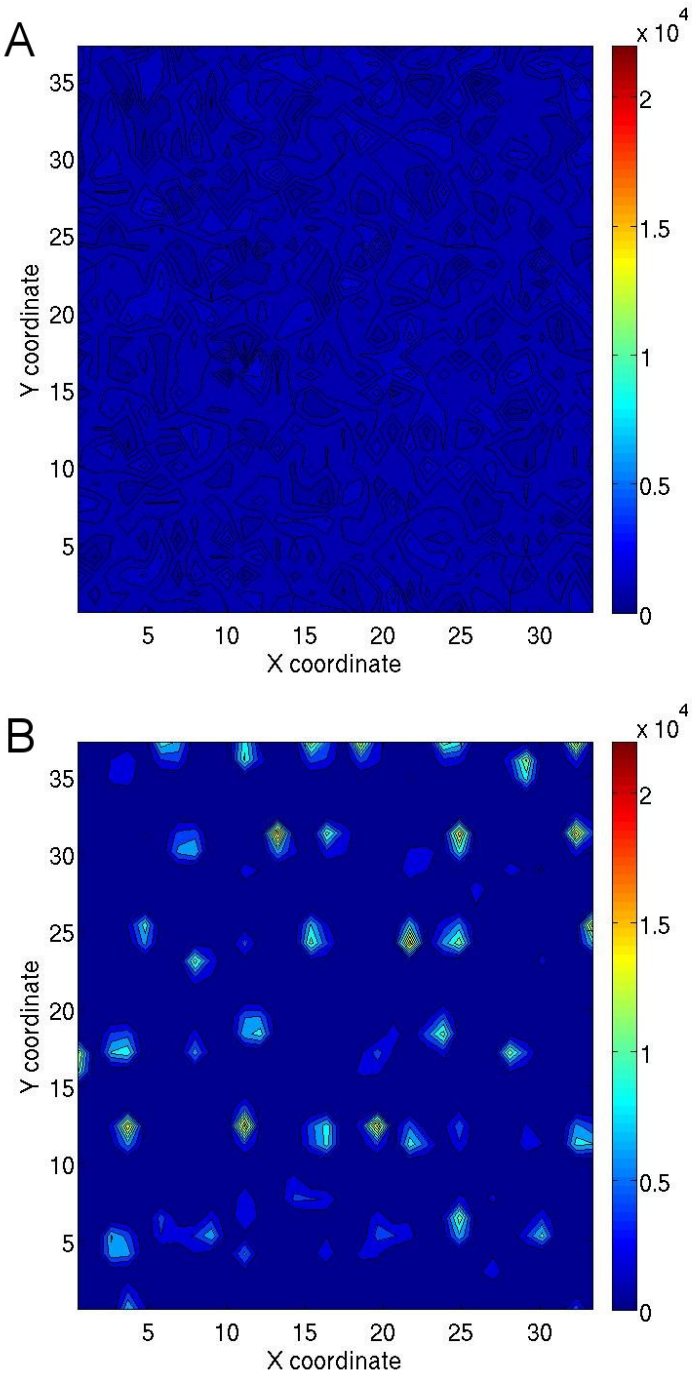

Figure 7.5: Water density in x, y planes parallel to the (011) aragonite surface (shift 0.0) as shown in figure 7.4. Both disordered (**A**) and ordered (**B**) layers are shown on the same scale. Density is in kg/m^3.

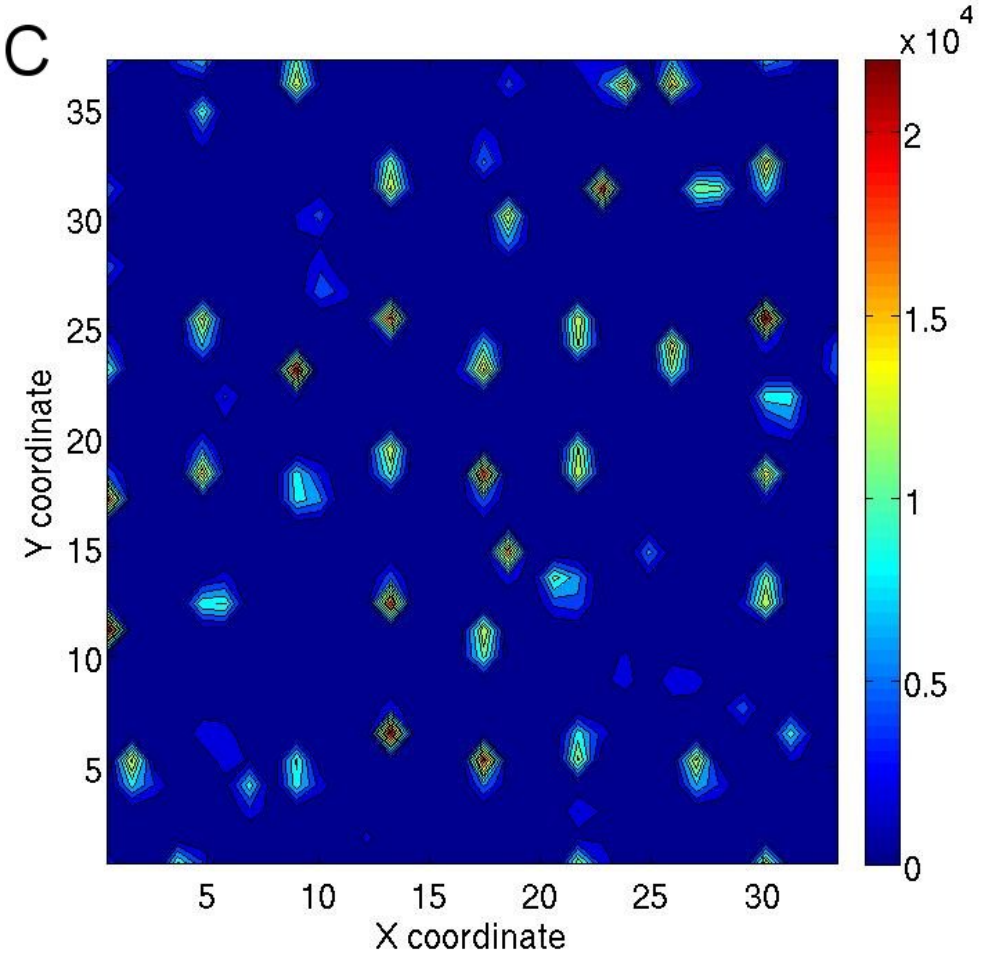

Figure 7.6: Water density in x, y planes parallel to the (011) aragonite surface (shift 0.0) as shown in figure 7.4. Ordered layer C shown on the same scale as figure 7.5. Density is in kg/m^3.

7.2 Random structure searching

Around 20,000 candidate structures of 75 formula units were generated and the nearest local energy minimum located for each, as described in section 6.3. When subjected to local structure analysis however none of these showed any evidence of crystalline order — a small fraction showed one or two ions as belonging to a crystalline polymorph, but there was never widespread crystallinity. The size of phase space grows exponentially with the number of ions, and the crystalline polymorphs are a very small fraction of it. It appears that this technique is currently impractical for discovering crystalline conformations for nanoparticles of this size. The technique is more commonly applied to periodic systems, with a smaller number of atoms [Pickard and Needs, 2011].

7.3 ISMC nanoparticles

Nanoparticles of the three crystalline polymorphs were produced using the ISMC method described in section 6.4 across the size range of 80 to 180 formula units. The complete set of atomic configurations can be found in appendix B.

7.3.1 Aragonite

Figure 7.7 shows examples of the structures found using the ISMC technique in vacuum, before any relaxation. there is considerable variation in the shape of the nanoparticles, though in all cases the same (101) and (011) surfaces dominate the structures. Figure 7.8 shows the surfaces exposed in one of these examples, and they are very similar to the Wulff constructions (figure 7.1), though the (210) cut is not made. The key difference between the structures produced by the two methodologies is that the ISMC structures tend to be smoother, with fewer sharp edges. The greater flexibility of the ISMC methodology allows an exploration of configurations, which although allowing greater exposure of surfaces which are higher in energy than might be desirable, allows the structure to minimise flaws and sharp edges, for a net energetic gain.

7.3.2 Calcite

For calcite the $(10\bar{1}4)$ surface dominates the Wulff construction and experimentally observed structures utterly, and it is supportive of the ISMC methodology that, bar-

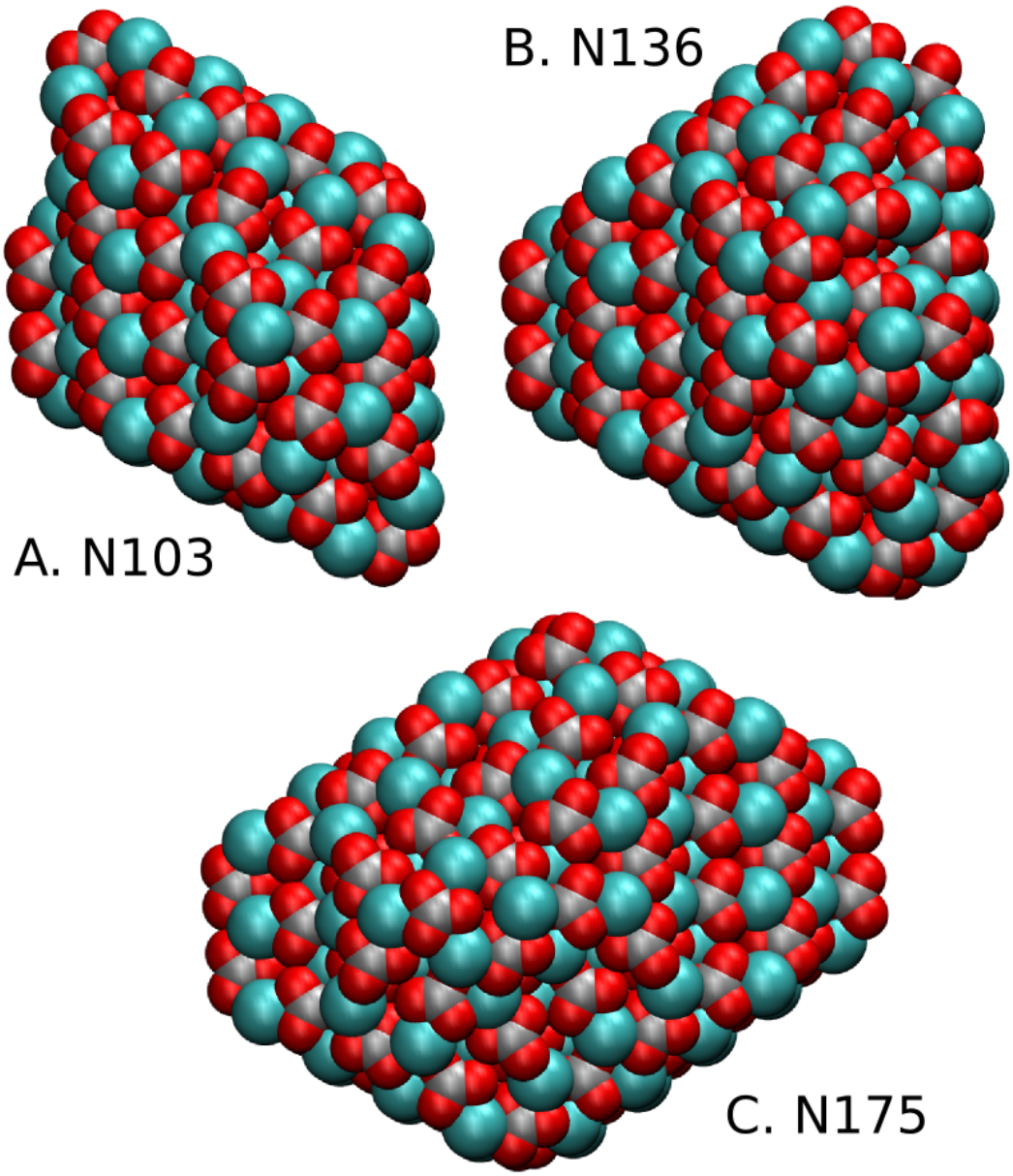

Figure 7.7: Example unrelaxed vacuum aragonite ISMC nanoparticles for (**A**) 103, (**B**) 136, (**C**) 175 formula units.

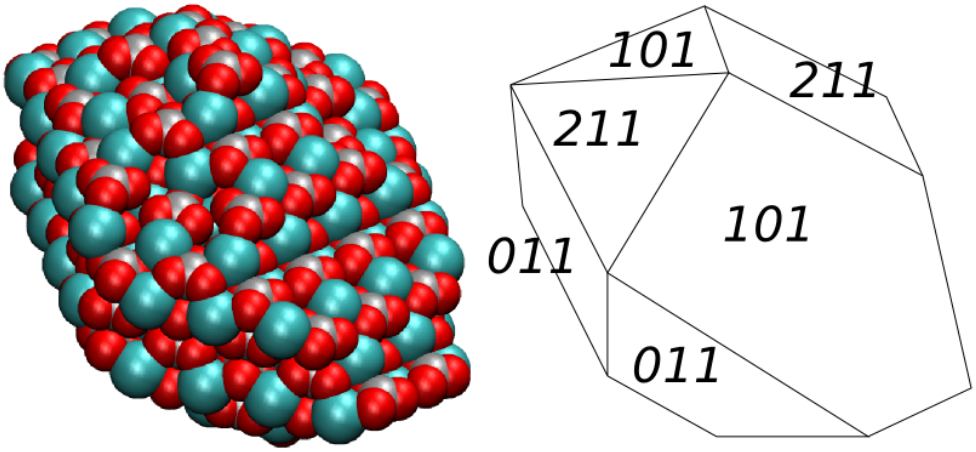

Figure 7.8: 175 formula unit aragonite nanoparticle with exposed surfaces. the (211) surface is actually more of a curved surface around those faces.

ring steps and edges, this is the only surface recovered in the ISMC nanoparticles. Figure 7.9 shows some example configurations, all in the familiar calcite rhombohedron shape. Ideal calcite Wulff constructions all have formula unit numbers of the form $\frac{(2i)^3}{2}$, $\frac{i(i+1)^2}{2}$ or $\frac{i^2(i+1)}{2}$ for integer i, to give a perfect flawless rhombohedron with roughly equal side lengths. For such formula unit numbers the ISMC technique recovers the Wulff construction exactly. Between such numbers there is a trade off between having as few flaws as possible and exposing the smallest possible surface area. A 150 formula unit nanoparticle could be constructed as either a regular rhombohedron with sides of 5x6x10 ions, or it might be lower in energy to have a nanoparticle with less exposed surface such as a 6x7x7 rhombohedron with an additional six ions added to one side. In this case the flawless structure is favoured.

7.3.3 Vaterite

The vaterite structures, such as those in figure 7.10 all appear to have distinct points at opposite ends, and are otherwise relatively smooth with few large distinct surfaces. The carbonate groups are arranged so as to avoid lone oxygen atoms pointing out of the structure.

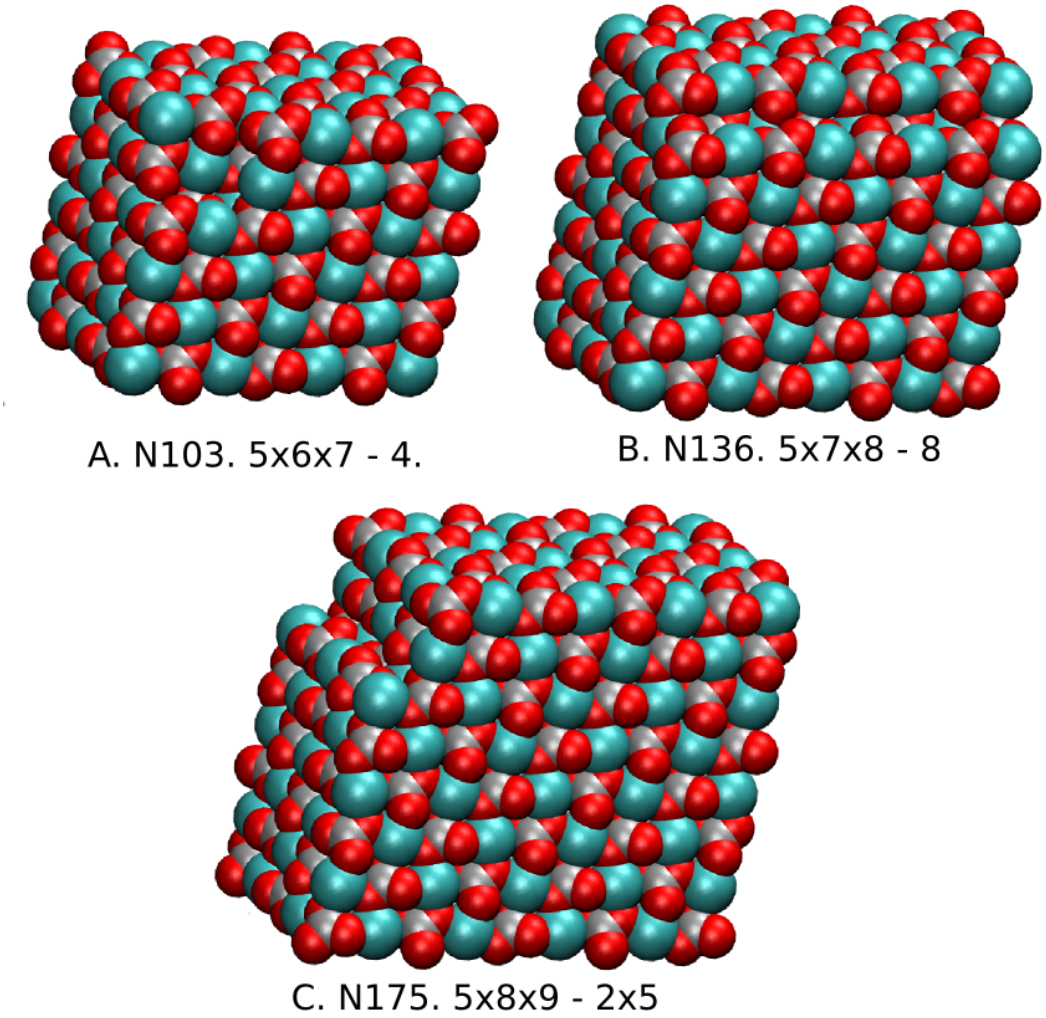

Figure 7.9: Example calcite ISMC nanoparticles for (**A**) 103, (**B**) 136, (**C**) 175 formula units, with side lengths of structures.

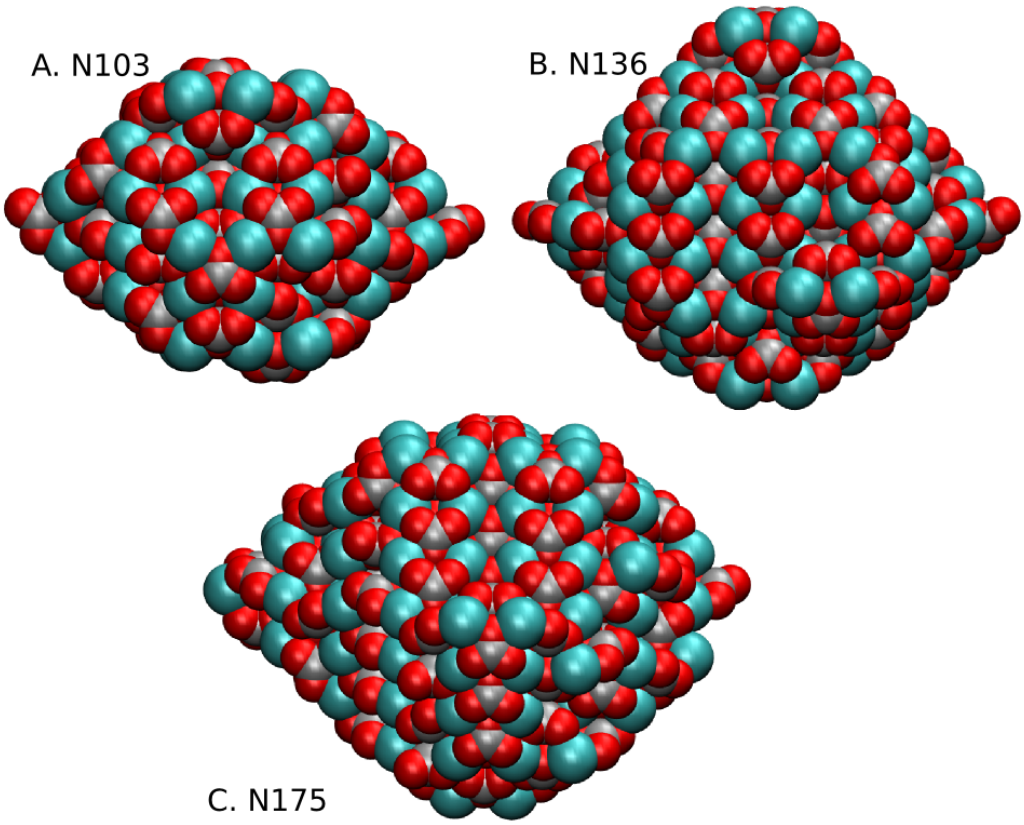

Figure 7.10: Example vaterite ISMC nanoparticles for (**A**) 103, (**B**) 136, (**C**) 175 formula units.

7.3.4 Zero Kelvin ordering of polymorphs

Figure 7.11 show the potential energy per atom for nanoparticles of the three calcium carbonate polymorphs generated using the ISMC technique. For every nanoparticle size, calcite is always the lowest energy structure, followed by aragonite and then vaterite. The sharp dips in the calcite graph are at the location of perfect Wulff construction rhombohedra, at 90, 126, and 147 formula units for example, where the IMSC nanoparticle is identical to the Wulff construction. It is notable that the energy difference between nanoparticles is far larger than the energy difference between the bulk polymorphs, indicating the importance of surface energy in nanoparticles. The zero Kelvin potential energy difference between bulk calcite and aragonite is just 0.0028 eV/Atom using this potential, compared to around 0.05 eV/Atom for these nanoparticles. This is presumably due to the higher surface energy of the aragonite surfaces, though line and edge energies where surfaces meet might also contribute significantly at these sizes. Also included on this graph are the results for the aragonite Wulff constructions - they are considerably higher in energy than the ISMC structures of the same size.

Figure 7.12 show the potential energy per atom for nanoparticles of the three calcium carbonate polymorphs generated using the ISMC technique, after relaxation. The energetic ordering of the three polymorphs is the same as before at most sizes, although the energy difference between vaterite and aragonite is much lower, and for 80 formula units vaterite is actually lower in energy than aragonite. Vaterite may have relaxed so much further due to the structures used in the ISMC code: whereas calcite and aragonite were based on crystal structures relaxed using the same potential, due to the partial disorder in the vaterite structure that version of the ISMC code used the unrelaxed crystal structure of Kamhi [1963]. The large gap in energy between calcite and the other polymorphs in vacuum implies that where aragonite or vaterite is favoured at small sizes, this must be due to the presence of solvent.

The aragonite Wulff constructions are again higher in energy than the aragonite ISMC structures. This suggests the ISMC method is more appropriate than the Wulff construction methodology for nanoparticles in this size range. A disadvantage of the ISMC technique is that is does not facilitate nanoparticle relaxation, yet for aragonite at least, even following relaxation the ISMC nanoparticles are still lower in energy.

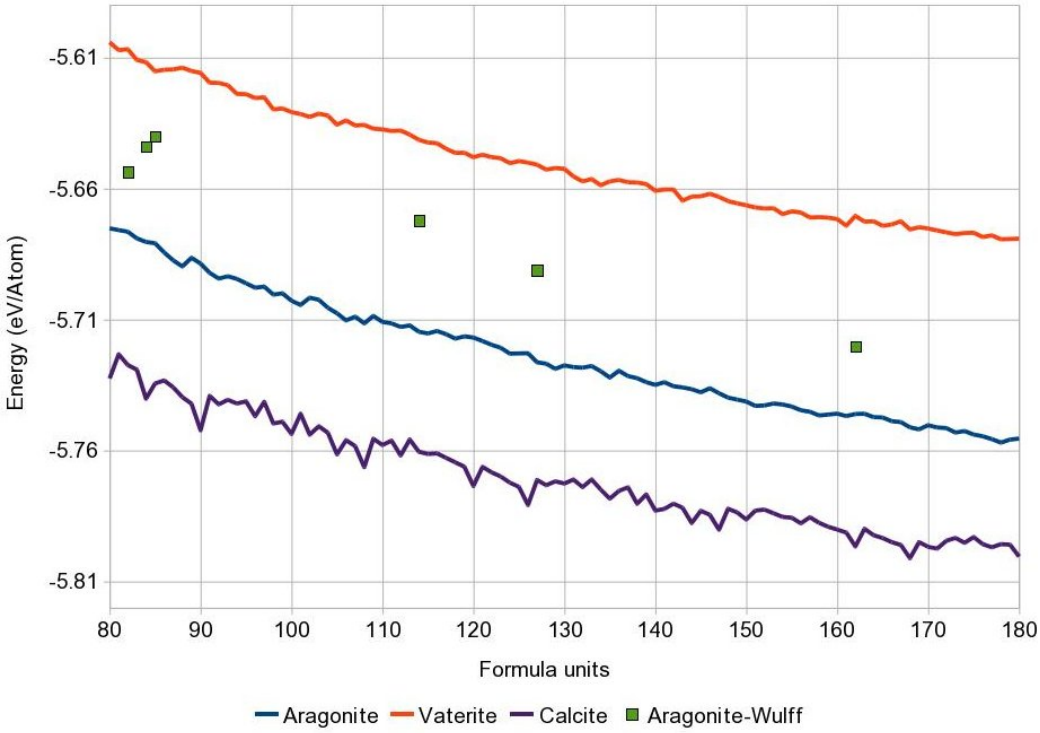

Figure 7.11: Energy per atom for ISMC calcite, aragonite and vaterite before relaxation, with aragonite Wulff constructions for comparison. Calcite Wulff constructions exactly coincide with calcite ISMC nanoparticles.

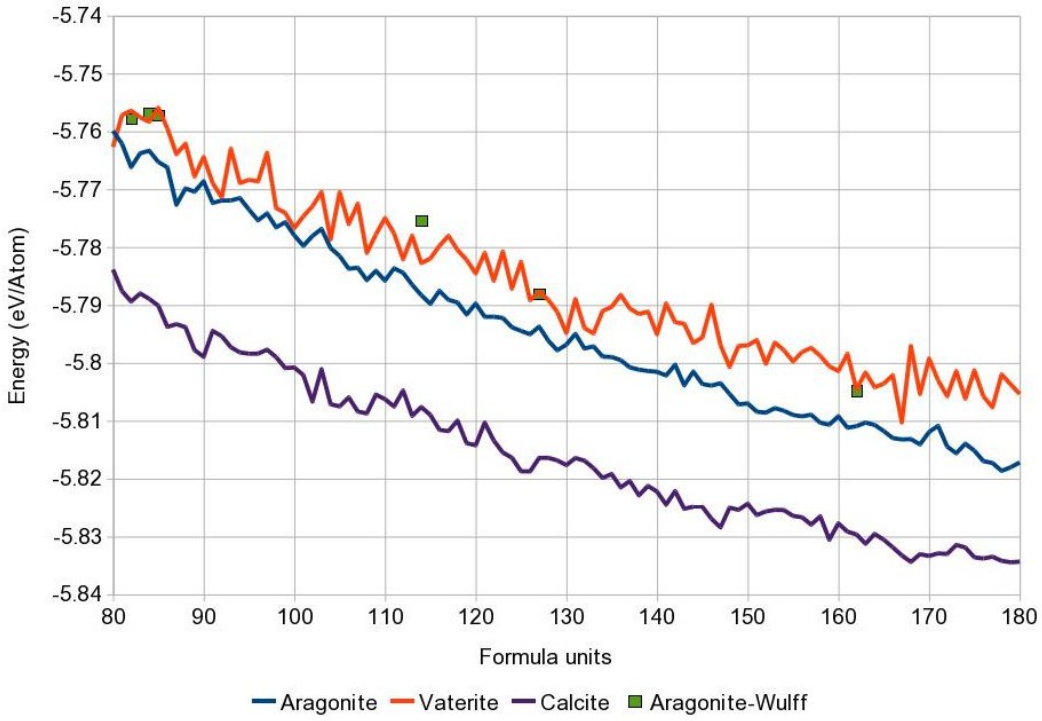

Figure 7.12: Energy per atom for ISMC calcite, aragonite and vaterite after relaxation, with aragonite Wulff constructions for comparison.

7.4 Nanoparticle stability

Having generated nanoparticles and considered their potential energy at 0 K, the stability of these nanoparticles at room temperature when immersed in water was investigated. As detailed in the introduction (section 1.5) Cooke and Elliott [2007] found that water stabilised calcite rhombohedra as small as 18 formula units over the course of a 1 ns simulation. However, that work used a different potential, and I have been able to perform longer simulations, as detailed in section 10.7. I am not aware of any prior simulation of aragonite or vaterite nanoparticles.

The results are characterised by large variations in the simulation enthalpies. Figures 7.13 and 7.14 shows enthalpies over the course of some representative simulations. Large fluctuations in the instantaneous enthalpies are to be expected in an (N, P, T) simulation with fluctuations in the instantaneous temperature estimator, yet these graphs are of the rolling average taken over 1 ns and 1000 data points. For these nanoparticle simulations there appear to be processes proceeding over nanosecond time scales causing reasonably large enthalpy changes. From visual inspection of the trajectories of these simulations this appears to be disruption of the nanoparticles edges (figure 7.15) with chains of ions moving around, with subsequent movement of the water molecules around them. In some simulations of calcite and aragonite, ions or ion pairs were shown to break off the nanoparticles altogether (table 7.6). The calcite example of 192 formula units (figure figure 7.16) started as a perfect rhombohedron with side lengths 6x8x8, yet at the end of the simulation a carbonate ion had detached and many of the edges were disordered. Table 7.6 is a full list of the simulations attempted and the average total enthalpy taken over the final 2.5 ns of the simulation.

When plotted, the rolling average enthalpy shows a general downward trend in all simulations, and in no case can it be convincingly argued that the simulation has reached equilibrium. This is the case even for the largest calcite nanoparticles simulated, such as the 256 formula unit 8x8x8 rhombohedron of figure 7.14. Even if this were not the case, a much longer simulation would be necessary to average out the effects of the chain movement as this appears to be a very slow process. Nevertheless, it is intriguing that in the results obtained to date the energetic ordering of the polymorphs is heavily dependent on particle size, at least in terms of enthalpy. For 90 formula units aragonite appears to be lower in enthalpy than the other two polymorphs, whereas for 100 formula units vaterite is lower in enthalpy

(figure 7.13). While graphs such as these where the variations in enthalpy are as large or larger than the enthalpy difference between polymorphs, in some cases aragonite was observed to be substantially lower in enthalpy than calcite, such as the 147 formula unit nanoparticles (figure 7.14, upper). The calcite 147 formula unit nanoparticle again begins as a perfect rhombohedron with sides 7x7x6. It is notable that although the aragonite nanoparticles were optimised for vacuum conditions, they are actually more competitive with calcite when hydrated, at least in terms of enthalpy.

Fewer vaterite simulations were attempted, but it was notable that in no case did ions detach themselves from the nanoparticle, though the same kind of loose chain formation was observed. Figure 7.17 shows a representative configuration of 148 formula units.

One unknown factor is the extent to which the general downwards drift in enthalpy represents an amorphisation of whole nanoparticles or just the corners, edges, and surfaces. Applying local structure analysis to final configurations (table 7.7) does indicate that the smaller nanoparticles do become sufficiently amorphous that they can no longer be confirmed as belonging to the initial polymorph, though residual order clearly remains. For example, the 84 formula unit calcite rhombohedron of figure 7.15 has no ions that can be conclusively identified as calcite-like, although in the initial configuration 40 ions were calcite-like. Figure 7.18 is a scatter plot of residual crystalline order for all 35 simulations, and shows how for calcite in particular the larger nanoparticles appear to be more stable. However, it is quite possible that this is simply because the larger nanoparticles take longer to become amorphous. It may also be the case that regardless of nanoparticle size a central crystalline core will be surrounded by a thick layer of less ordered ions when equilibrium is reached. From this graph and the values in table 7.7, it would be tempting to draw the conclusion that aragonite and vaterite nanoparticles are more stable than calcite nanoparticles of the same size, however it must be considered that the sensitivity and specificity of the local structure analysis varies considerably between the three polymorphs as detailed in chapter 5.

One aspect of these results still to be explored is the entropic contribution to the energetic ordering of the polymorphs. For the hydrated surfaces the restriction and ordering of water molecules appears to be crucial for the ranking of the surface free energies. Furthermore, it is notable that all 35 of the attempted hydrated nanoparticle simulations had lower enthalpy than the equivalent quantities of bulk

Polymorph	Size	Fin. Enthalpy (eV)	Notes
CALC	75	-4833.71 ± 2	
ARAG	80	-4981.99 ± 1.91	
CALC	80	-4978.166 ± 1.65	
ARAG	84	-5102.06 ± 1.99	
CALC	84	-5098.08 ± 1.8	
ARAG	89	-5237.37 ± 1.8	
CALC	89	-5237.59 ± 1.7	
VAT	89	-5236.28 ± 2.04	
ARAG	90	-5273 ± 1.95	
CALC	90	-5264.15 ± 1.6	
VAT	90	-5264.51 ± 2.43	
ARAG	91	-5299.74 ± 1.63	Free carbonate ion
CALC	91	-5292.655 ± 1.7	
VAT	91	-5295.36 ± 1.83	
ARAG	96	-5445.45 ± 1.7	Three free ions
CALC	96	-5446.13 ± 2.05	
ARAG	100	-5562.22 ± 1.7	
CALC	100	-5558.21 ± 2.3	
VAT	100	-5566.45 ± 1.9	
ARAG	126	-6312.16 ± 2	
CALC	126	-6313.08 ± 1.64	
ARAG	146	-6898.53 ± 1.58	Three free ion pairs
CALC	146	-6886.72 ± 1.73	
VAT	146	-6888.62 ± 2.3	
ARAG	147	-6940.18 ± 1.16	
CALC	147	-6918.38 ± 2.06	
ARAG	148	-6968.72 ± 1.3	
CALC	148	-6946.31 ± 1.9	Three free ions
VAT	148	-6945.72 ± 1.9	
CALC	171	-7614.2 ± 3.22	Free carbonate ion
ARAG	175	-7735.83 ± 1.6	
CALC	175	-7731.71 ± 1.9	
CALC	192	-8222.93 ± 2	Free carbonate ion
CALC	196	-8354.19 ± 1.26	
CALC	256	-10086.95 ± 1.31	

Table 7.6: Complete list of hydrated nanoparticle simulations. Final enthalpy is average over final 2.5 ns of 12.5 ns simulation

		Crystalline ions		
Polymorph	**Size**	**End**	**Start**	**ΔH bulk**
CALC	75	2	28	-18.53 ± 2.71
ARAG	80	7	17	-22.32 ± 2.65
CALC	80	3	36	-18.5 ± 2.47
ARAG	84	2	19	-26.8 ± 2.71
CALC	84	0	40	-22.82 ± 2.57
ARAG	89	13	22	-17.62 ± 2.58
CALC	89	1	40	-17.84 ± 2.51
VAT	89	7	18	-16.53 ± 2.75
ARAG	90	14	25	-24.35 ± 2.68
CALC	90	6	48	-15.5 ± 2.44
VAT	90	9	11	-15.86 ± 3.05
ARAG	91	10	24	-22.19 ± 2.46
CALC	91	4	47	-15.11 ± 2.51
VAT	91	13	21	-17.81 ± 2.6
ARAG	96	10	26	-23.41 ± 2.51
CALC	96	3	44	-24.09 ± 2.76
ARAG	100	17	29	-24.59 ± 2.51
CALC	100	3	46	-20.58 ± 2.95
VAT	100	6	21	-28.82 ± 2.65
ARAG	126	34	42	-23.18 ± 2.73
CALC	126	20	80	-24.1 ± 2.48
ARAG	146	34	59	-31.59 ± 2.45
CALC	146	29	85	-19.78 ± 2.55
VAT	146	30	57	-21.68 ± 2.97
ARAG	147	39	61	-44.34 ± 2.21
CALC	147	35	69	-22.54 ± 2.79
ARAG	148	39	60	-43.98 ± 2.29
CALC	148	50	88	-21.57 ± 2.67
VAT	148	39	57	-20.98 ± 2.67
CALC	171	56	120	-24.8 ± 3.74
ARAG	175	57	81	-30.84 ± 2.48
CALC	175	29	120	-26.72 ± 2.68
CALC	192	45	140	-26.67 ± 2.76
CALC	196	71	146	-42.34 ± 2.29
CALC	256	113	208	-41.21 ± 2.35

Table 7.7: Complete list of hydrated nanoparticle simulations. Local structure and enthalpy comparison to bulk

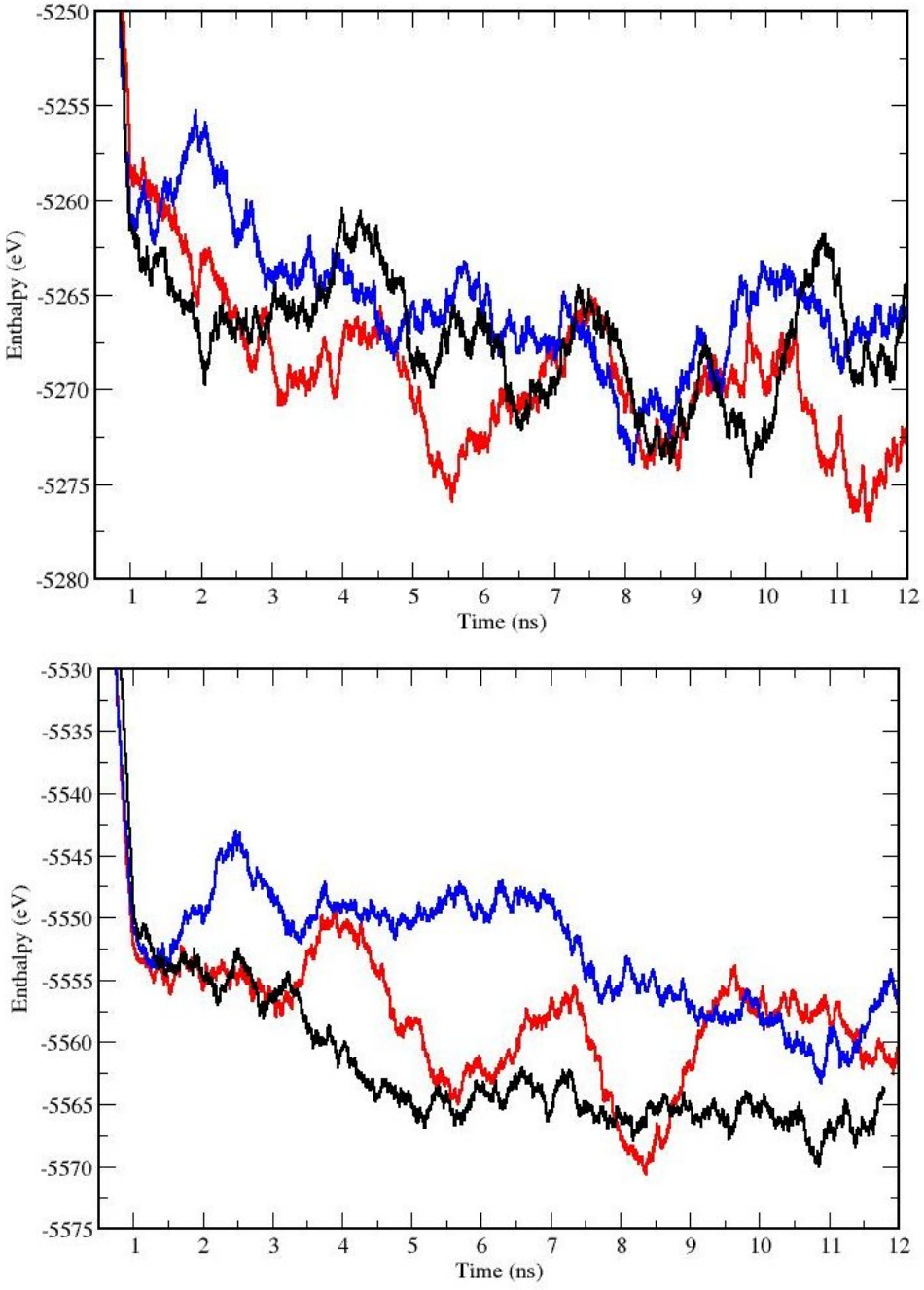

Figure 7.13: simulation enthalpy for nanoparticles of aragonite (RED), calcite (BLUE), and vaterite (BLACK). **Upper** 90 formula units. **Lower** 100 formula units. Values are rolling averages over 1 ns, 1000 data points.

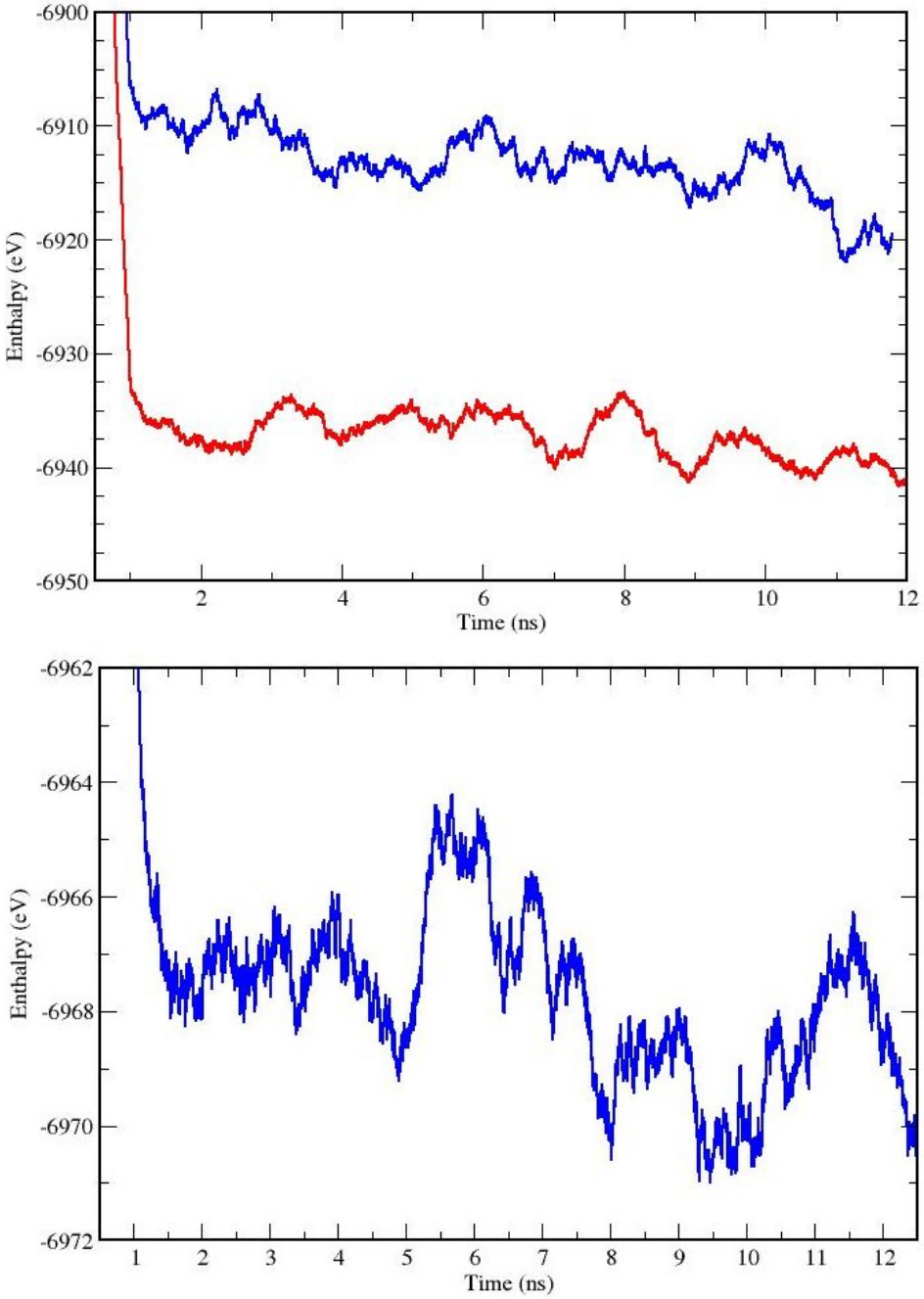

Figure 7.14: **Upper** simulation enthalpy for 147 formula unit nanoparticles of aragonite (RED), and calcite (BLUE). **Lower** simulation enthalpy for 256 formula unit nanoparticle of calcite. Values are rolling averages over 1 ns, 1000 data points.

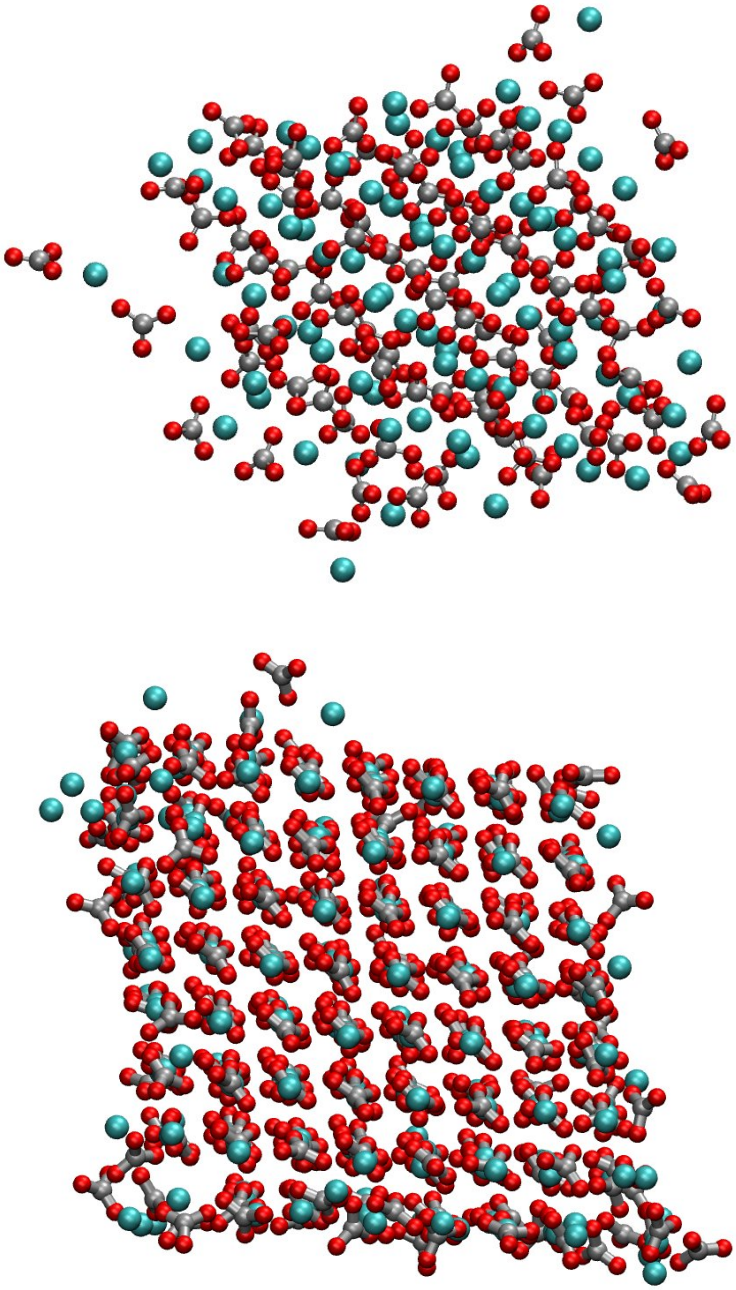

Figure 7.15: Example configuration of 84 (**upper**) and 256 (**lower**) formula unit nanoparticle of calcite after 12.5 ns simulation, showing ions chains trailing from nanoparticles. Similar chains were found in aragonite and vaterite simulations.

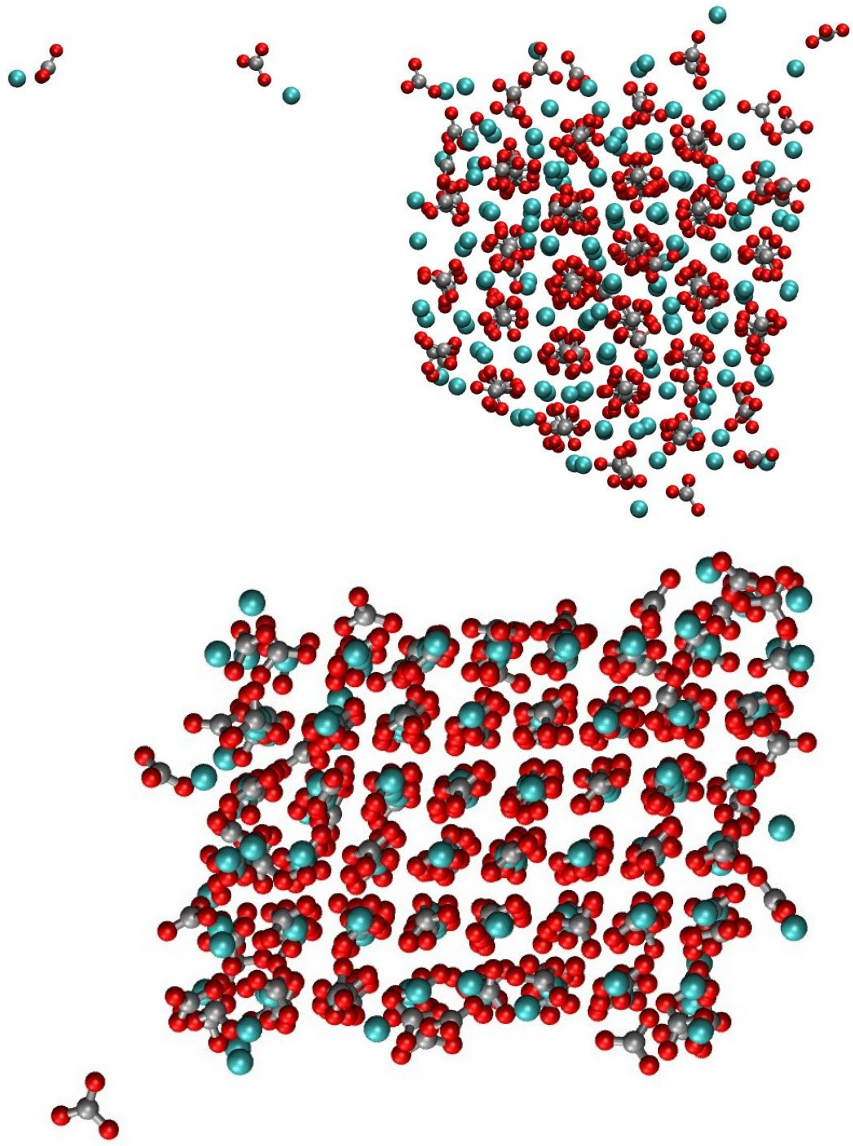

Figure 7.16: Example configuration of 146 formula unit aragonite nanoparticle (**upper**) and 192 formula unit calcite nanoparticle (**lower**) after 12.5 ns simulation, showing free ions which have broken off the nanoparticles.

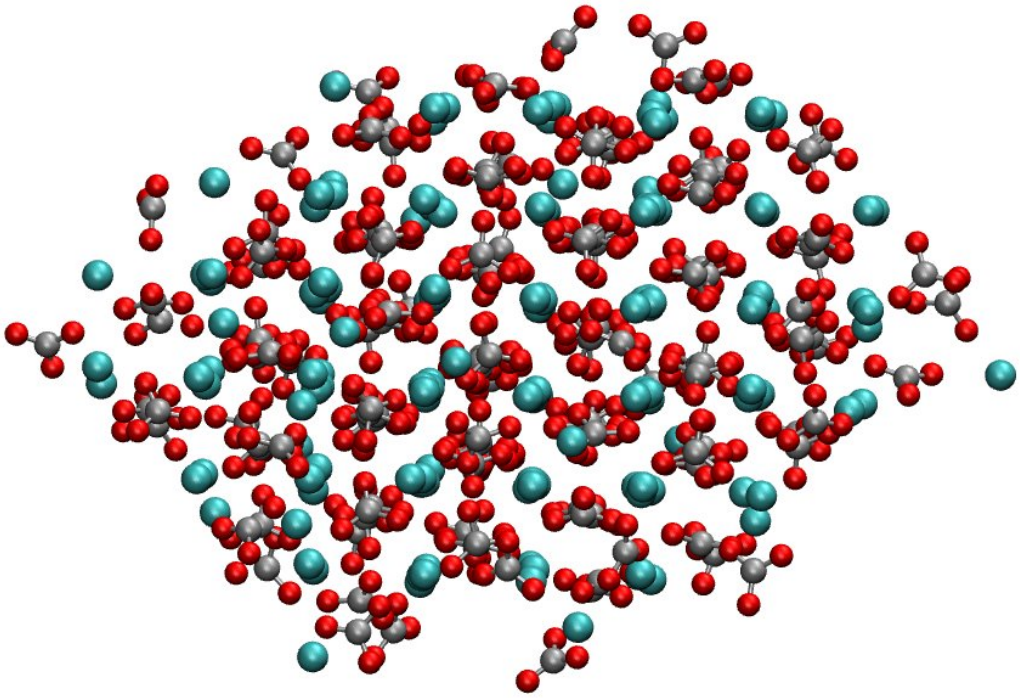

Figure 7.17: Example configuration of 148 formula unit nanoparticle of vaterite after 12.5 ns simulation.

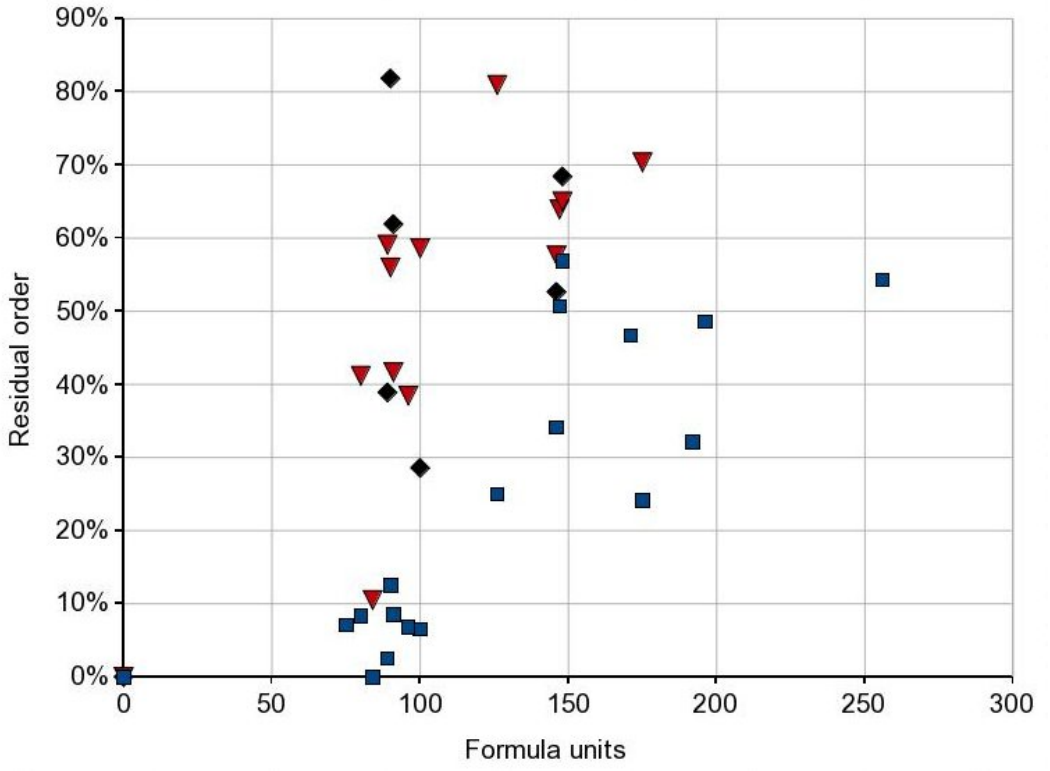

Figure 7.18: Scatter plot of residual order for all 35 hydrated nanoparticle simulations. Residual order is defined as number of crystalline ions at end of simulation divided by number of crystalline ions at the end of the simulations for the relevant polymorphs, expressed as a percentage. RED TRIANGLES: aragonite. BLUE SQUARES: calcite. BLACK DIAMONDS: vaterite.

polymorph and bulk water, as detailed in table 7.6. There is a general trend that increasing size, and hence surface area, leads to a greater energy difference between nanoparticle and bulk.

It would seem to be unwarranted to ignore entropic effects when considering the ordering of the nanoparticles. Metadynamics provides a tool to give a good approximation to the correct free energy differences, and having produced plausible nanoparticles for all three polymorphs we are now in a position to evaluate order parameters for this technique much more comprehensively than before. In the next chapter trajectories are generated between polymorphs to test the ability of order parameters to distinguish between nanoparticle configurations.

Chapter 8

Nanoparticle polymorph transitions

8.1 Introduction

Having produced nanoparticles for all three polymorphs, and investigated their stability in water, a natural next step is to understand how transitions between these states can occur. Metadynamics can provide an estimate of the free energy surface, but requires a set of order parameters with a very specific characteristics as discussed in section 2.7, covering all the important slow processes in the system. A method of evaluating order parameters would be extremely helpful.

The full transition path ensemble — the set of all trajectories between two endpoints — can be explored and sampled using transition path sampling (TPS), first introduced by Dellago et al. [1998]. This technique has been applied to nanoparticle systems [Grünwald, 2007], studying the solid-solid phase transition between two crystalline polymorphs. From a representative set of trajectories, the technique of maximum likelihood estimation has been applied to evaluate order parameters [Peters et al., 2007].

Transition path sampling begins with a seed trajectory between the two endpoints, and obtaining this trajectory can present a significant challenge. The trajectory does not need to be particularly representative of the ensemble as a whole. There simply needs to be a set of time slices along the pathway, which when simulated with a positive timestep travel to one endpoint, and with a negative timestep travel to the other. This trajectory is then used to initialise a Monte-Carlo

sampling of trajectories through phase space.

I have applied the nudged elastic band (NEB) method to try and find such a seed trajectory for the calcite–aragonite transition. The reasons this attempt was ultimately unsuccessful reveals details of the phase space of the calcium carbonate nanoparticle system which are of interest, and the resulting trajectories provide a powerful tool for evaluating order parameters. In this chapter I give an outline of the method, describe the local implementation, and present the results.

8.2 The nudged elastic band method

Reviewed extensively in chapter 16 of Berne et al. [1998] the nudged elastic method is well established method of identifying the lowest energy path between two stable configurations on a reasonably smooth energy landscape. In this case the minima are two different polymorphs of calcium carbonate nanoparticles and the energy landscape is the full configuration space between them.

The technique as implemented involves a series of n atomic configurations, each of N formula units. I denote one endpoint $\mathbf{r}_1$, and the other $\mathbf{r}_n$. The intermediate configurations are initially set as linear interpolations between these two configurations ie:

$$\mathbf{r}_i = \frac{i-1}{n-1}\mathbf{r}_n + \frac{n-i}{n-1}\mathbf{r}_i \tag{8.1}$$

The different configurations can be considered as beads along an elastic band, stretched over the energy surface between the two fixed endpoints, with the initial configuration a straight line between them. The configurations, or beads, are subjected to forces from the regular force-field as well as harmonic spring forces between each bead and its neighbours. For each non-endpoint configuration $\mathbf{r}_i$, a tangent vector is defined:

$$\mathbf{r}_i^\tau = \frac{\dfrac{(\mathbf{r}_{i+1}-\mathbf{r}_i)}{\|\mathbf{r}_{i+1}-\mathbf{r}_i\|} + \dfrac{(\mathbf{r}_i-\mathbf{r}_{i-1})}{\|\mathbf{r}_i-\mathbf{r}_{i-1}\|}}{\left\| \dfrac{(\mathbf{r}_{i+1}-\mathbf{r}_i)}{\|\mathbf{r}_{i+1}-\mathbf{r}_i\|} + \dfrac{(\mathbf{r}_i-\mathbf{r}_{i-1})}{\|\mathbf{r}_i-\mathbf{r}_{i-1}\|} \right\|} \tag{8.2}$$

The final force on each non-endpoint bead is:

$$f(\mathbf{r}_i) = -\left(\nabla V(\mathbf{r}_i) - \langle \nabla V(\mathbf{r}_i).\mathbf{r}_i^\tau \rangle \, \mathbf{r}_i^\tau\right) + k \, \langle (\mathbf{r}_{i+1}-\mathbf{r}_i).\mathbf{r}_i^\tau \rangle \, \mathbf{r}_i^\tau + k \, \langle (\mathbf{r}_{i-1}-\mathbf{r}_i).\mathbf{r}_i^\tau \rangle \, \mathbf{r}_i^\tau \tag{8.3}$$

The three terms of this equation are the force due to the inter-atomic potential, adjusted to act normal to the tangent vector, and the spring forces to the adjoining beads. The spring constant k is a parameter, though the configuration is not sensitive to the value used. This is because the configuration which balances equation 8.3 would have the last two terms exactly equal and opposite in value, and both are scaled by k.

The set of beads is then relaxed subject to these forces, to produce a series of configurations distributed along a low energy pathway between the two endpoints.

8.3 Mapping between configurations

8.3.1 Method

One limitation of the technique when applied to atomistic systems is that the nature of the landscape is heavily dependent on the mapping of atoms between the endpoints. If for example the mapping required all the atoms to travel from one end of the nanoparticle to the other, then the resulting trajectory would be very different, and likely higher in energy along the pathway, than one where each atom was required to move just a couple of Angstroms. For my purposes, I make the assumption that the best possible map is the one which minimises the root-mean squared distance (RMSD) over all ions between the start and end configurations. This is not entirely satisfactory as it seems more likely that the phase transition proceeds through sliding and distortion of the layers described in section 1.1. However, this would appear to be the most straightforward approximation when using a technique such as this, or the related finite temperature string method [W. et al., 2005], where a one-to-one mapping between the endpoints is required.

The global search of all maps is completely impractical, as there are $n!$ possibilities. Instead I have derived a simulated annealing scheme to find a 'good enough' map.

Initially both the beginning and end configurations are translated so their centre of mass is at the origin, then the inertial tensor is calculated for both configurations. The endpoint configuration is then rotated so the principal moments of inertia are aligned and the distributions of mass are as equal as possible. I then perform a separate simulated annealing procedure for both the Ca^{2+} and CO_3^{2-} ions, in each case reordering the ions to minimise the RMSD over the full map. In detail,

for each ion type:

- I define a one-to-one map, or permutation, $g(i)$ from the set of integers $\{1, ..., N\}$ to itself. Furthermore, I define an 'energy' function:

$$E = \sum_{i=1}^{N} \|\mathbf{r}_1^i - \mathbf{r}_n^{g(i)}\|^2 \tag{8.4}$$

  Here $\mathbf{r}_1^i$ refers to the coordinates of i-th ion of the starting configuration, and $\mathbf{r}_n^i$ the i-th ion of the end configuration. The location of the carbonate ions is taken as the location of the carbon atom. The initial ordering of the ions is arbitrary. The goal of the algorithm is to search the space of possible functions g to minimise E, and hence minimise the RMSD. I initialise with simply $g(i) = i$.

- Each MC move proceeds by attempting to swap two destinations of the mapping g, and calculating the change in the value E that would result. the move is accepted with the standard Metropolis Hastings probability, i.e.:

$$p = \min\left(1, \exp\left(-\frac{\Delta E}{T}\right)\right) \tag{8.5}$$

  The quantity ΔE can be calculated without evaluating the full energy function, vastly increasing efficiency. Since this is an annealing procedure, the 'temperature' T drops linearly over the course of the minimisation, from $T_{\max}$, to $T_{\min}$. The best map, with the lowest E value, is stored.

- From the best map g the ions in the endpoint configuration are reordered. There are six possible mappings between the oxygen atoms of each carbonate, and these are tested exhaustively, to minimise the amount of rotation of each carbonate group.

8.3.2 Results

After some experimentation, values of $T_{\max}$ and $T_{\min}$ of 4 and 0.01 Å^2 were found to be appropriate, with the annealing proceeding over 20,000,000 MC moves. Figure 8.1 shows how the quantity E changes over the course of a minimisation, where the beginning and end points are minimised calcite and vaterite ISMC structures

respectively. With 146 formula units, a value of E of 1250 Å^2 equates to an RMSD of 2.93 Å per ion.

When ran repeatedly the algorithm always produced RMSD values of this order, and does not do significantly better if ran for a great deal longer. There are likely to be a whole range of solutions which give RMSD values close to the minimum, and it seems probable that the mappings produced by this algorithm are amongst these.

8.4 Local implementation

The local implementation of the NEB method uses a quasi-Newtonian minimiser, with a velocity-Verlet integrator of equation 8.3 modified such that if, for velocity $\mathbf{v}$, $\langle \mathbf{v}.\mathbf{f} \rangle < 0$, then the velocities are set to zero. This gives a straightforward and effective, if slightly crude, method of finding the local minimum in the $3n(N-2)$ -dimensional energy landscape. The method has been implemented using David Quigley's interface to the force evaluation routines in DL_POLY, enabling straightforward calculation of these terms. These are then modified to the forces of equation 8.3. A copy of the code is included in appendix B.

8.5 Trajectories

Trajectories were produced between relaxed ISMC nanoparticles (see chapter 6) of all three crystalline polymorphs, each of 146 formula units. Trajectories of 49 frames were produced, with a spring constant of 10 amu.ps^{-2}, and a timestep length of 0.0025 fs, and the minimisation proceeding over 50000 timesteps. This was found to give good convergence. Coordinate files along each trajectory are included in appendix B.

In all the graphs presented in this section, the pathway is described as moving along a reaction coordinate. For each frame $\mathbf{r}_i$, this is defined as:

$$\frac{(\mathbf{r}_i - \mathbf{r}_1).(\mathbf{r}_n - \mathbf{r}_1)}{||\mathbf{r}_n - \mathbf{r}_1||^2} \tag{8.6}$$

if the frame is projected onto the direct linear path between the endpoints, this quantity is the normalised distance along that path for that frame. There is a possibility that two configurations along the pathway could be linked by a vector

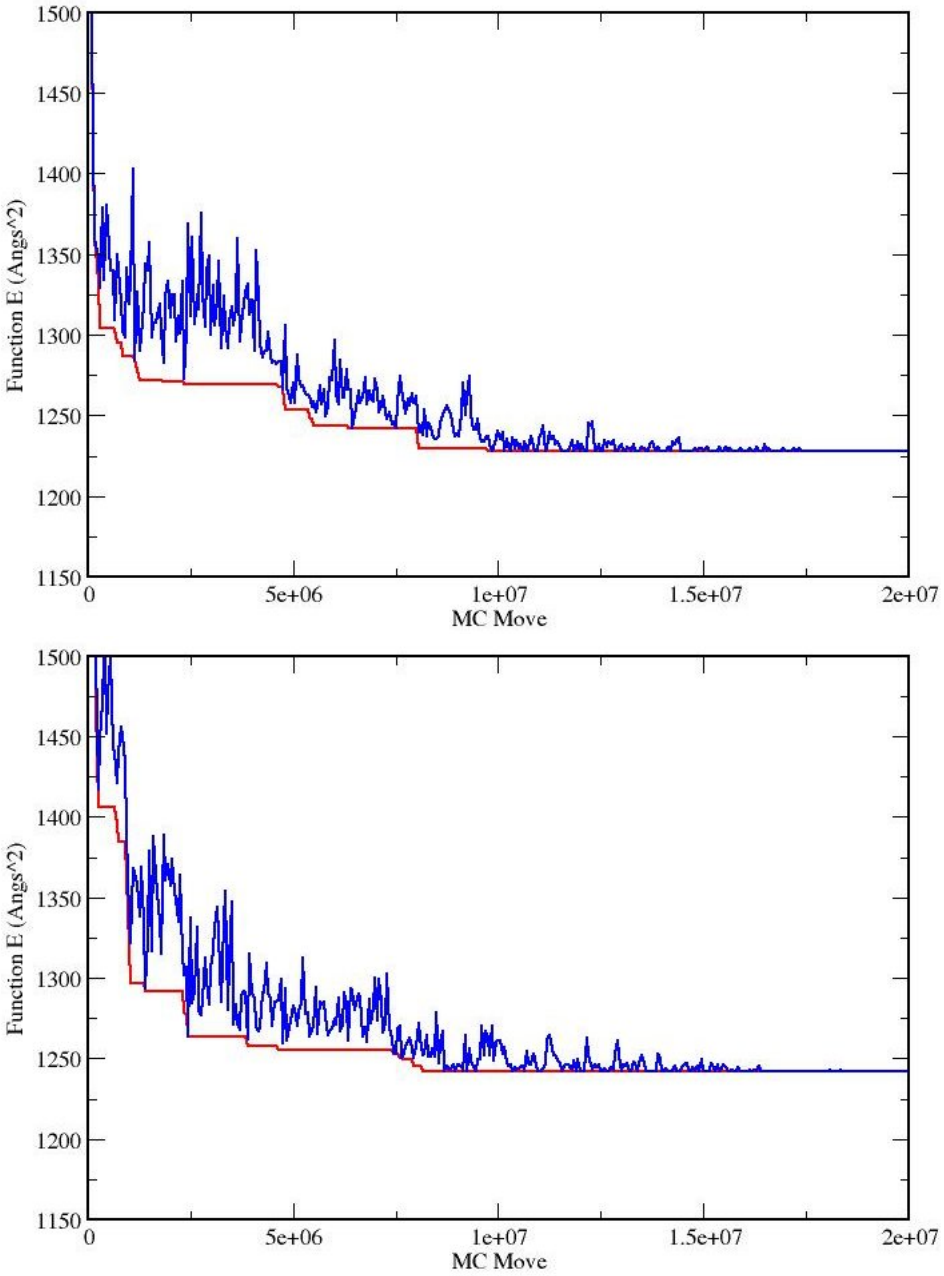

Figure 8.1: Quantity E (equation 8.4) over the course of minimisation. Current value (BLUE) with lowest value (RED), for Ca^{2+} (**Upper**) and CO_3^{2-} (**Lower**) ion lists.

perpendicular to the direct linear path, but this has not occurred for the results presented here.

Figure 8.2 shows how the potential energy varies along the three pathways. All three pathways move through a high energy configuration between the two endpoints, with each midpoint of approximately equal energy. Figure 8.3 shows how the local structure changes along the pathways, and reveals that the midpoints are amorphous, or at least not representative of any of the three crystalline polymorphs, a result confirmed by visual inspection, and consideration of the global order parameters as below.

Plotting how the three sets of order parameters tested in this thesis change over the course of the pathway (figures 8.4 to 8.6) reveal that for the most part they smoothly change throughout each trajectory. More instructive are projections of the trajectories onto 2D surfaces of order-parameter pairs. While for the Steinhardt Q_4 parameters (fig. 8.9) the trajectories move smoothly across the surface, with the three polymorphs all well separated, for the inertia tensor and ring statistics order parameters (figs. 8.7 and 8.8) the trajectories move erratically across the surface repeating crossing themselves, showing how the order parameters are unable to distinguish between some transition states along the pathway. The 2D Steinhardt Q_4 graph also show how all three trajectories pass through the same region with order parameters characteristic of the amorphous phase. For the inertia tensor parameters however, the amorphous intermediate between aragonite and calcite can be distinguished from the other two amorphous intermediate states. This is because the vaterite nanoparticle is not as similar in shape to calcite and aragonite as those structures are to each other, and hence the nanoparticle shape does not have to change as dramatically.

Having shown that the Steinhardt Q_4 parameters are most appropriate for the study of nanoparticle phase transitions, we can use the trajectories to study the relationship between them. The metadynamics simulations of Quigley and Rodger have biased 5 Q4 order parameters, which, since each additional parameter causes an exponential growth in the time for the FES approximation to converge, is an undesirably large number. Table 8.1 shows the correlation coefficient between all pairs of order parameters. If two order parameters are highly correlated this indicates that biasing both is probably unnecessary. The Q4 parameters separate neatly into two groups — the Ca–Ca and C–C OPs are highly correlated, as are the Ca–C, Ca-O and C–O OPs. Thus is would seem that it is only necessary to bias one of

each of these groups in a metadynamics simulation. Such simulations are considered in the next section.

Attempts to recover molecular dynamics trajectories between crystalline polymorphs were unsuccessful: the program was modified to output frames with velocities with a direction towards the next frame. When simulations were run from these trajectories they never progressed much beyond the start point — amorphous nanoparticles remained amorphous, crystalline nanoparticles remained crystalline and intermediate nanoparticles stayed intermediate. The energy surface along the trajectory appears to contain numerous deep minima, and this rugged landscape hinders phase transitions even when initialised with velocities chosen to cross the next barrier.

	Q4 Ca-C	**Q4 Ca-O**	**Q4 C-C**	**Q4 C-O**	**I3**
Q4 Ca-Ca	0.48	0.08	0.81	0.41	-0.24
Q4 Ca-C		0.90	0.89	0.95	0.67
Q4 Ca-O			0.64	0.92	0.88
Q4 C-C				0.85	0.31
Q4 C-O					0.73
	I2	**I1**	**n4**	**n6**	**nc**
Q4 Ca-Ca	0.53	0.21	-0.01	0.97	0.16
Q4 Ca-C	-0.45	0.93	0.84	0.47	0.92
Q4 Ca-O	-0.76	0.97	0.98	0.10	0.98
Q4 C-C	-0.04	0.72	0.56	0.81	0.70
Q4 C-O	-0.49	0.96	0.87	0.45	0.93
I3	-0.88	0.83	0.90	-0.24	0.83
I2		-0.66	-0.83	0.49	-0.72
I1			0.94	0.24	0.96
n4				0.03	0.98
n6					0.20

Table 8.1: Correlation coefficients for order parameters along C–V trajectory

8.6 Metadynamics simulation of calcium carbonates

Repeating the work of Quigley and Rodger with the new force-field it was found that for a hydrated system, any attempt to bias nanoparticle configurations caused ions to rapidly disassociate (figure 8.10). As shown in chapter 7, even unbiased simulations would occasionally have ions separate from the main nanoparticle. It appears

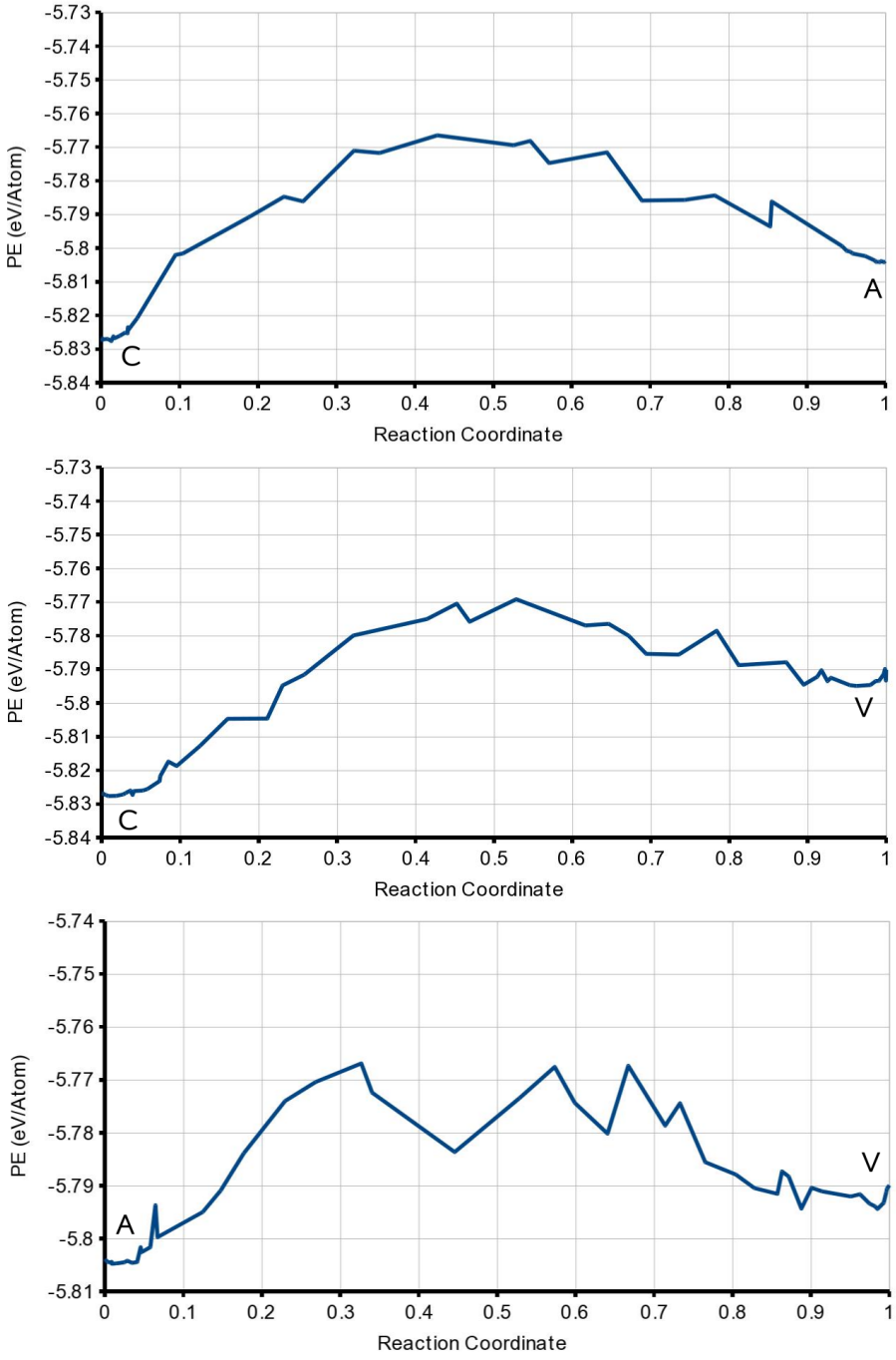

Figure 8.2: Potential energy for each frame, for the calcite–aragonite, calcite–vaterite, and aragonite–vaterite trajectories

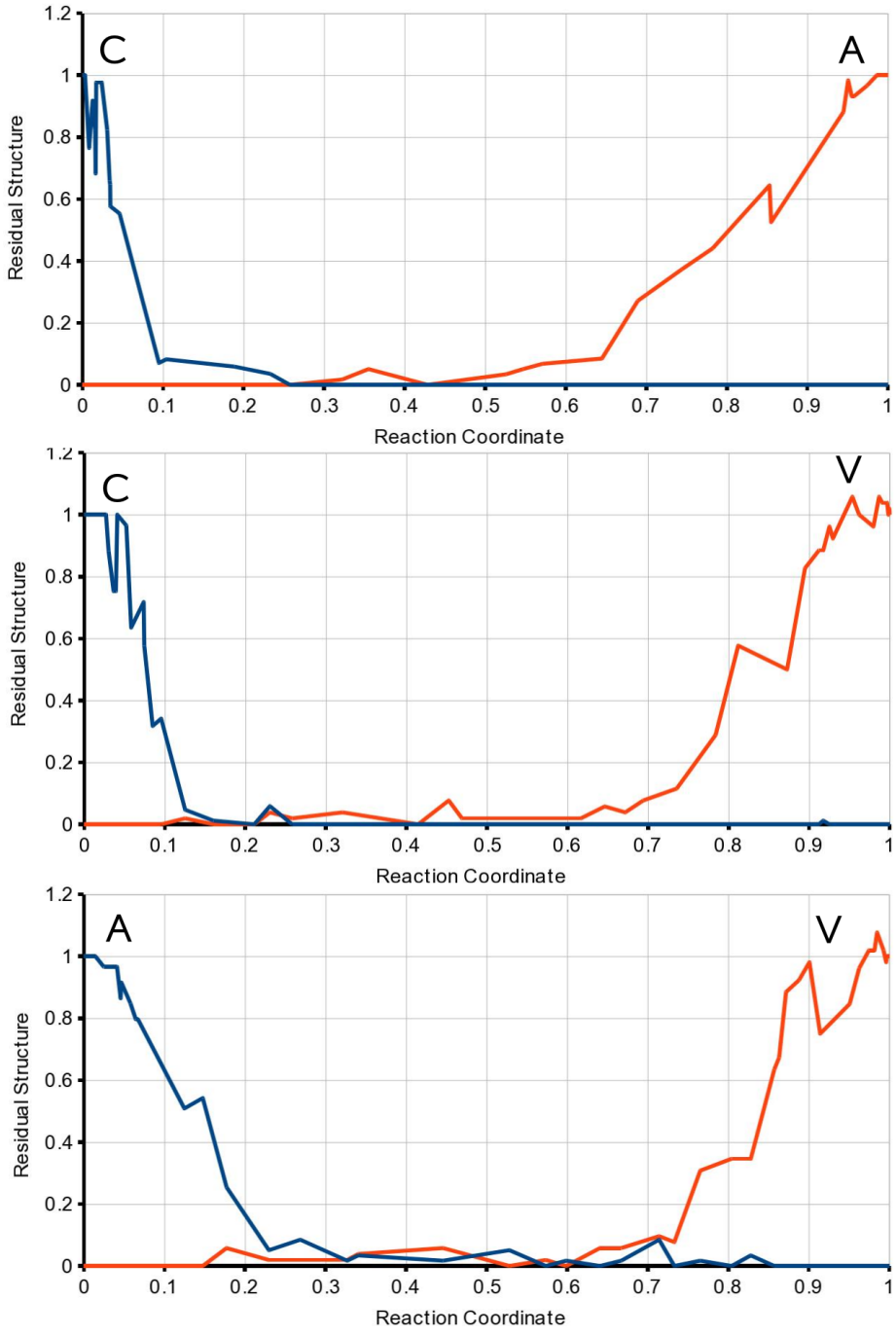

Figure 8.3: Ions identified as crystalline, normalised to numbers for the start and endpoints.

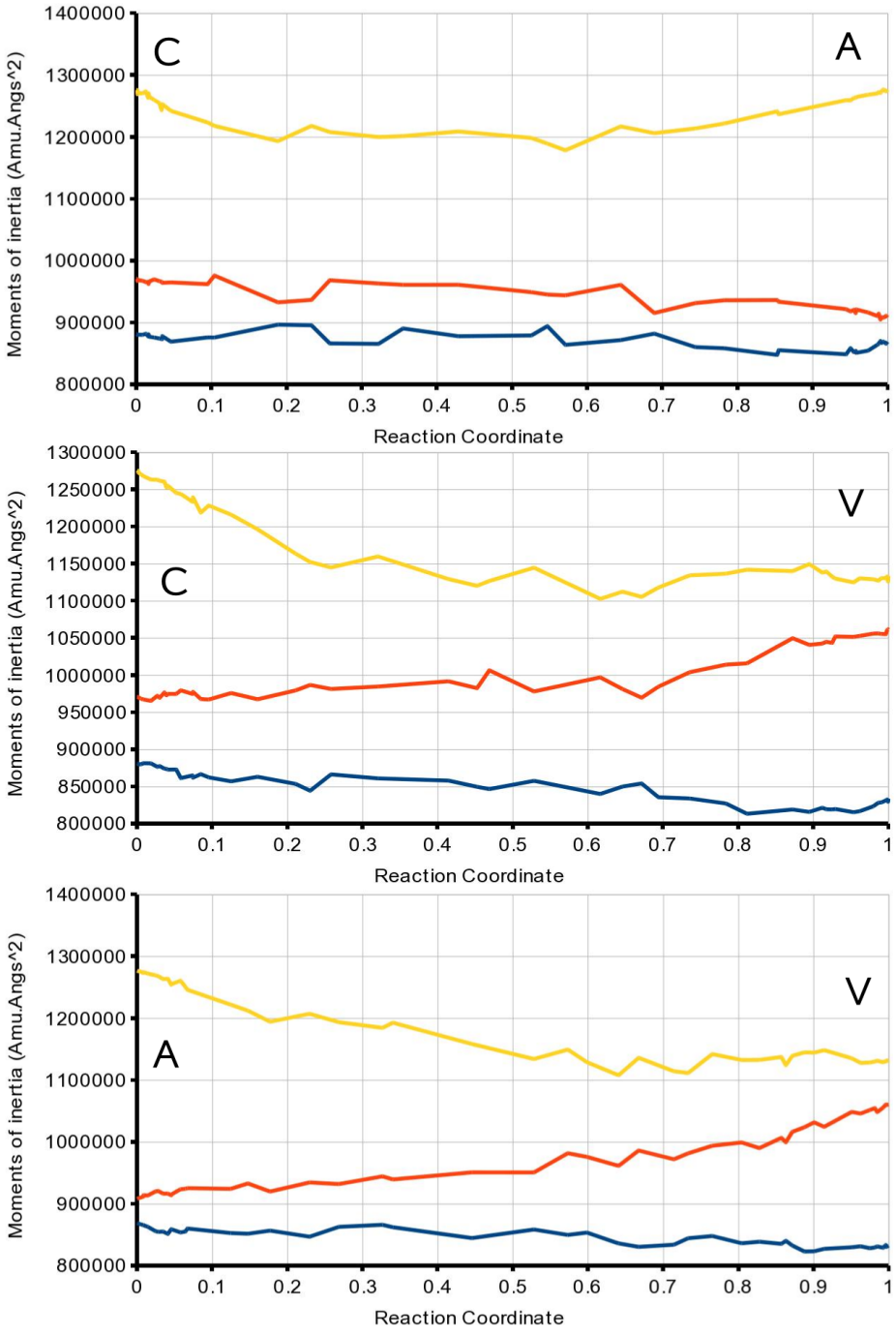

Figure 8.4: Inertia tensor order parameters, for the calcite–aragonite, calcite–vaterite, and aragonite–vaterite trajectories. These are the first (yellow), second (orange) and third (blue) principle moments of inertia.

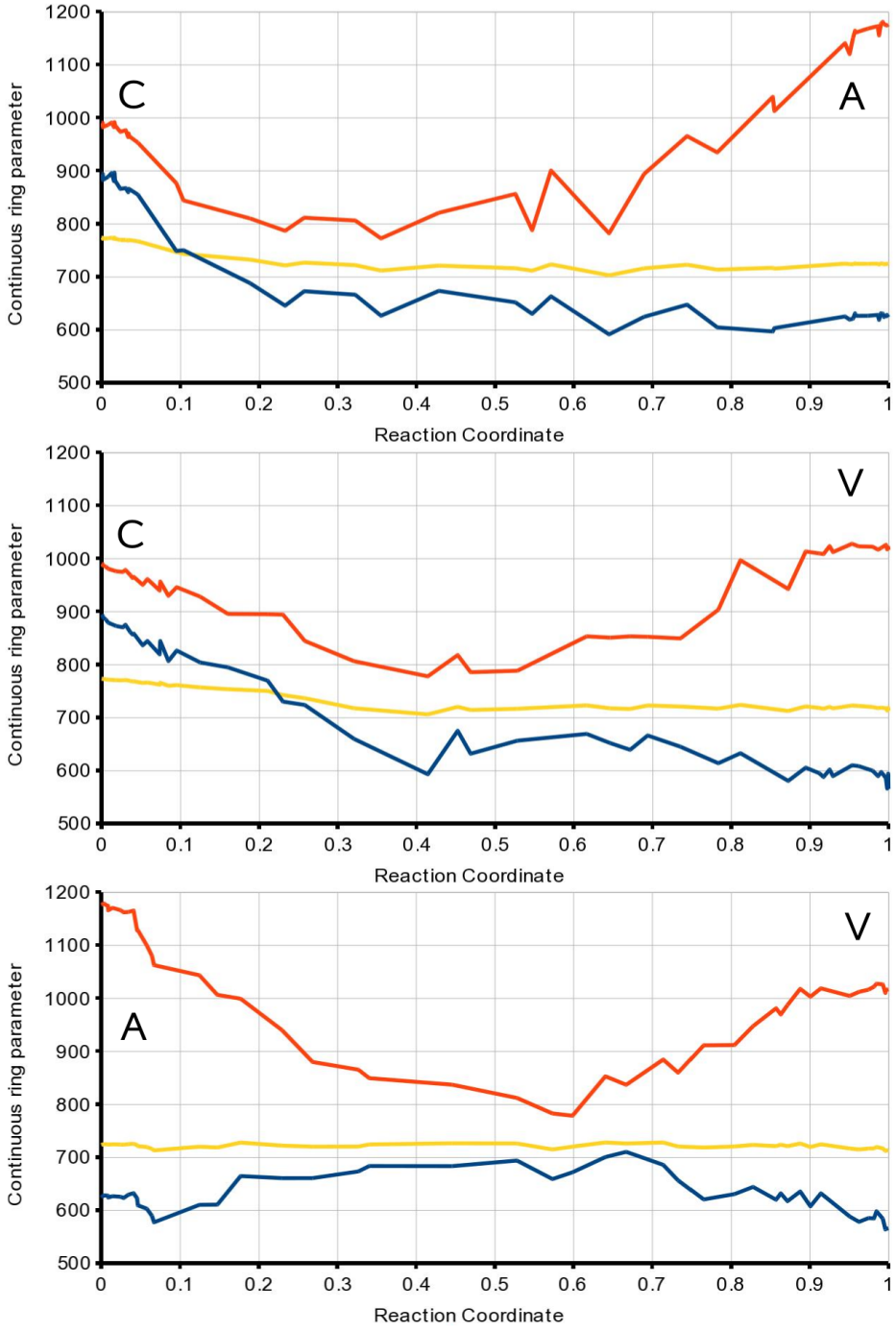

Figure 8.5: Rings statistics, for the calcite–aragonite, calcite–vaterite, and aragonite–vaterite trajectories. These are the n_4 (blue), n_6 (orange) and n_C (yellow) order parameters.

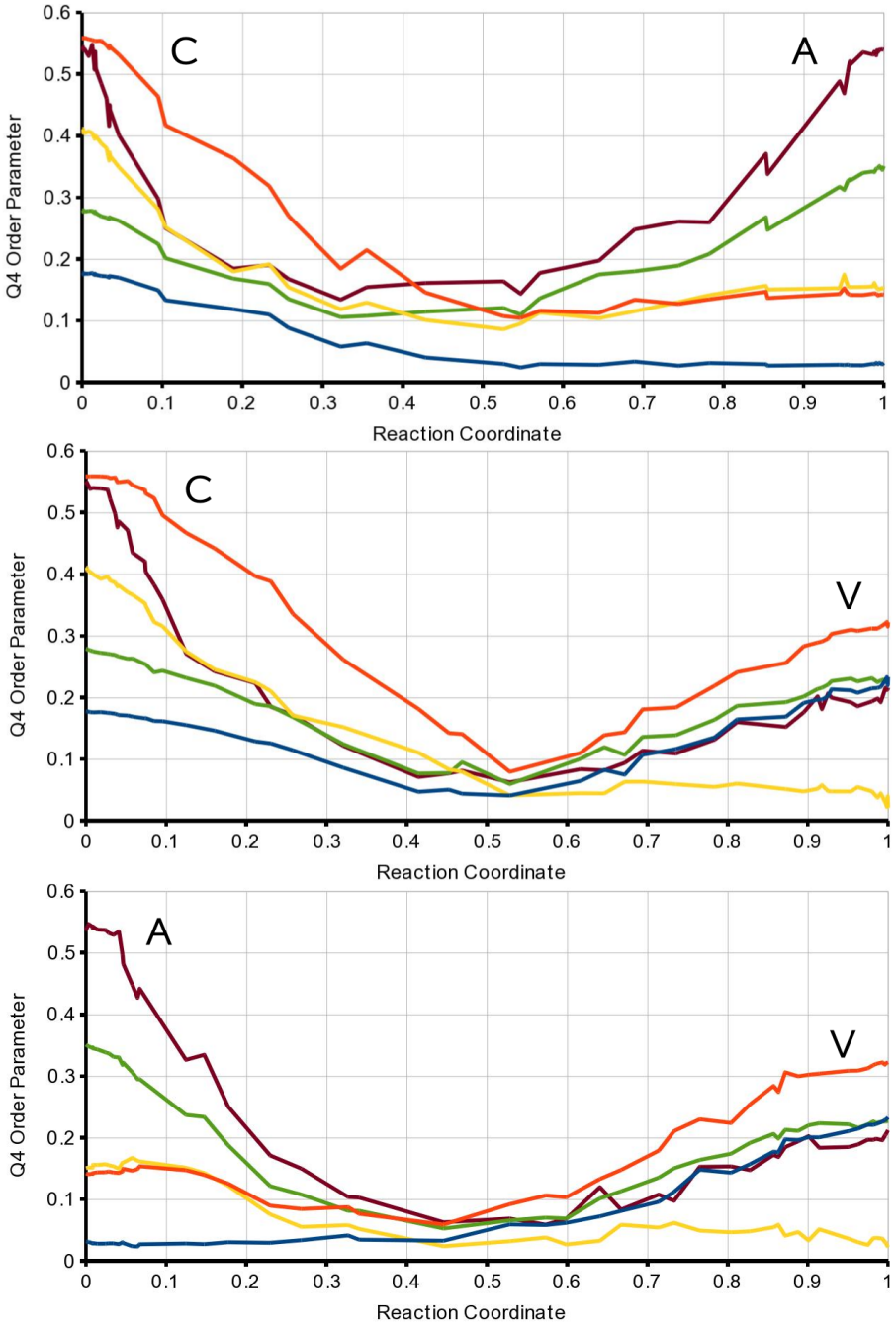

Figure 8.6: Q4 Steinhardt parameters, for the calcite–aragonite, calcite–vaterite, and aragonite–vaterite trajectories. These are the Ca–Ca (blue), Ca–C (orange), Ca-O (yellow), C–C (green) and C–O (brown) order parameters, with cut-offs as for Quigley et al. [2009].

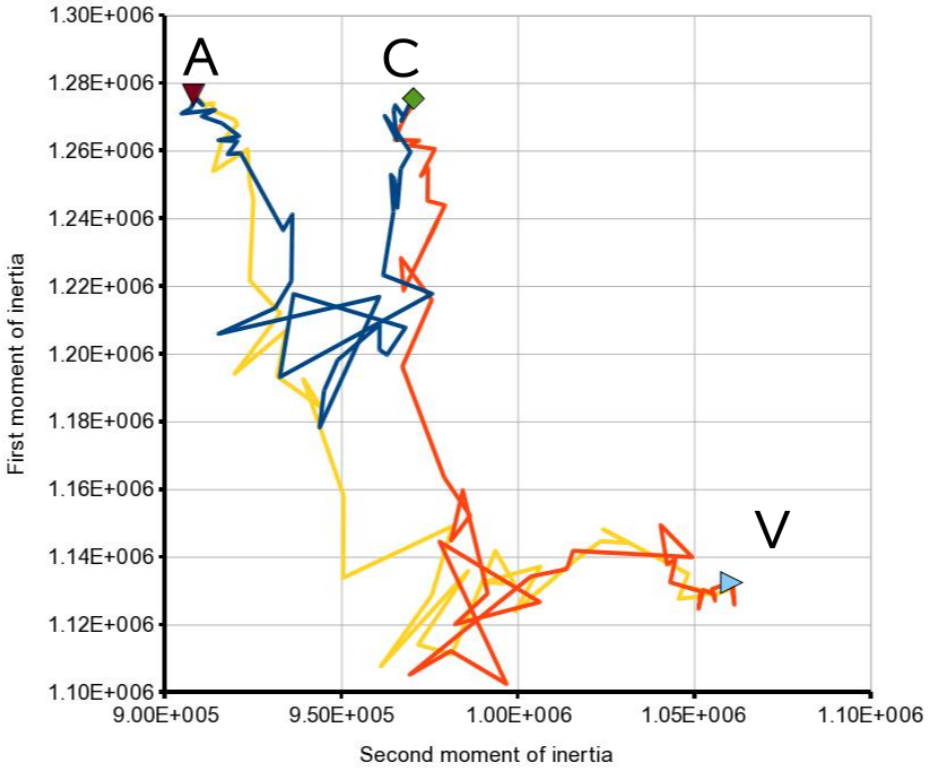

Figure 8.7: Trajectories between the three structures projected onto 2D surface of the first and second principal moments of inertia

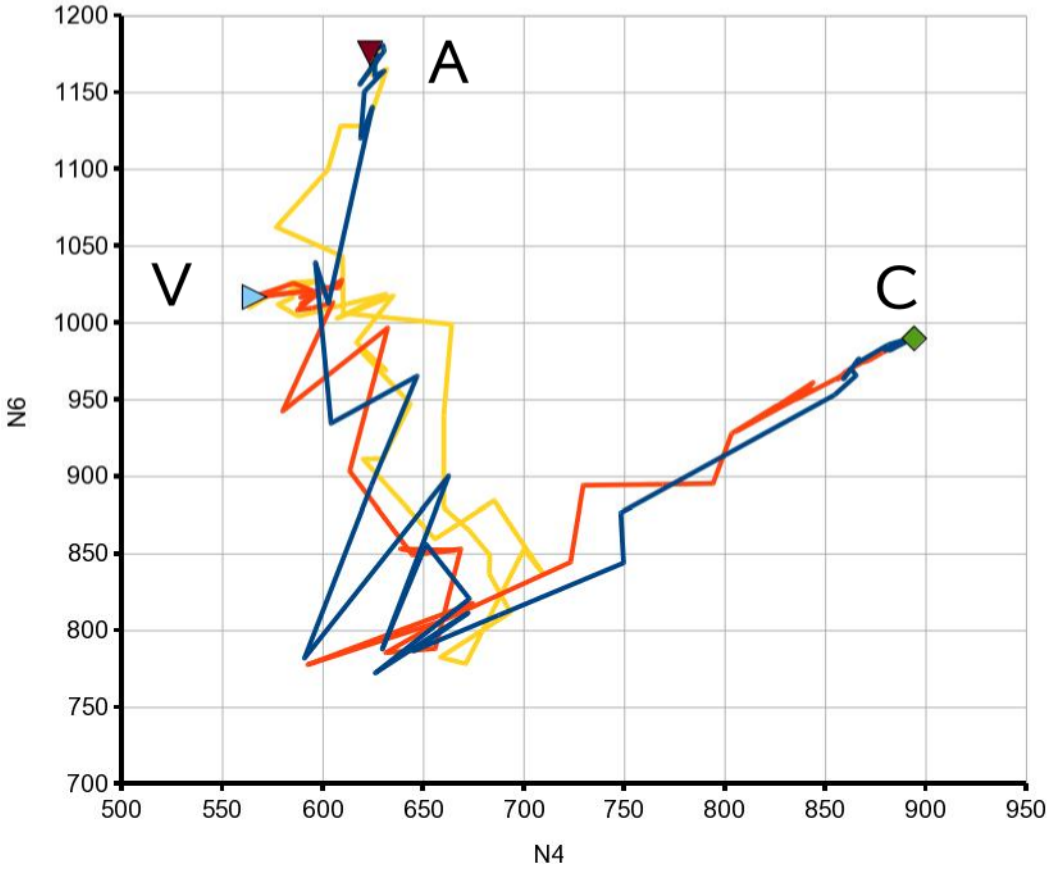

Figure 8.8: Trajectories between the three structures projected onto 2D surface of n_4 and n_6 order parameters.

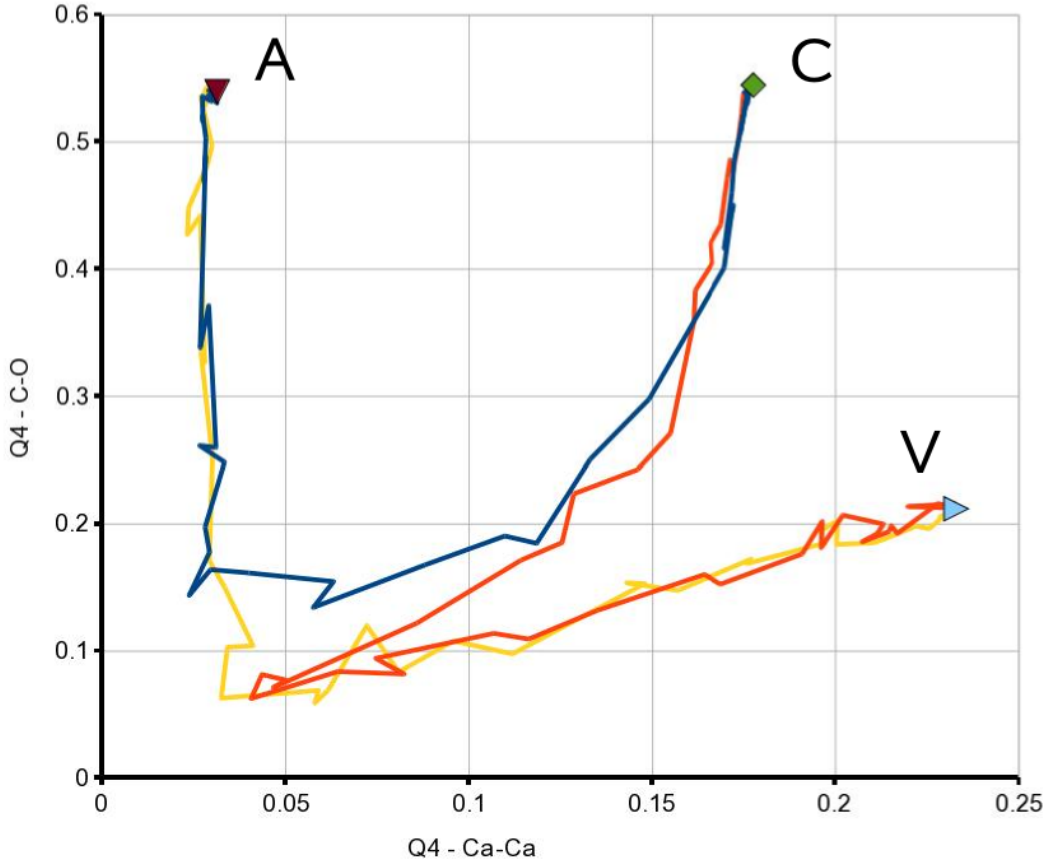

Figure 8.9: Trajectories between the three structures projected onto 2D surface of Ca–Ca and C–O order parameters.

that where a large biasing force is applied, encouraging rapid severe conformational change in a nanoparticle, the separation of ions is accelerated. The movement of free ions around the solution is however a slow process on metadynamics timescales, and has not been biased. It might be possible to introduce some kind of constraint on the nanoparticle ensuring it remains contiguous, but this would render the simulation less representative of physical reality. More interesting would be biasing the water–mineral phase interaction, exploring the behaviour of ion adsorption and dissolution. Ideally such a scheme could enable the recovery of a free energy surface constraining not just anhydrous ACC and the crystalline phases, but anhydrous and hydrous ACC. A candidate order parameter might be some measure of the sum total number of water–mineral interactions, though a more sophisticated scheme might well be necessary. For future work, the recent reconnaissance metadynamics technique [Tribello et al., 2010] might be employed to test candidate order parameters. It might however be the case that the complicated phase space encompassing so many characteristics cannot be sensibly reduced to the small number of dimensions available to a metadynamics simulation.

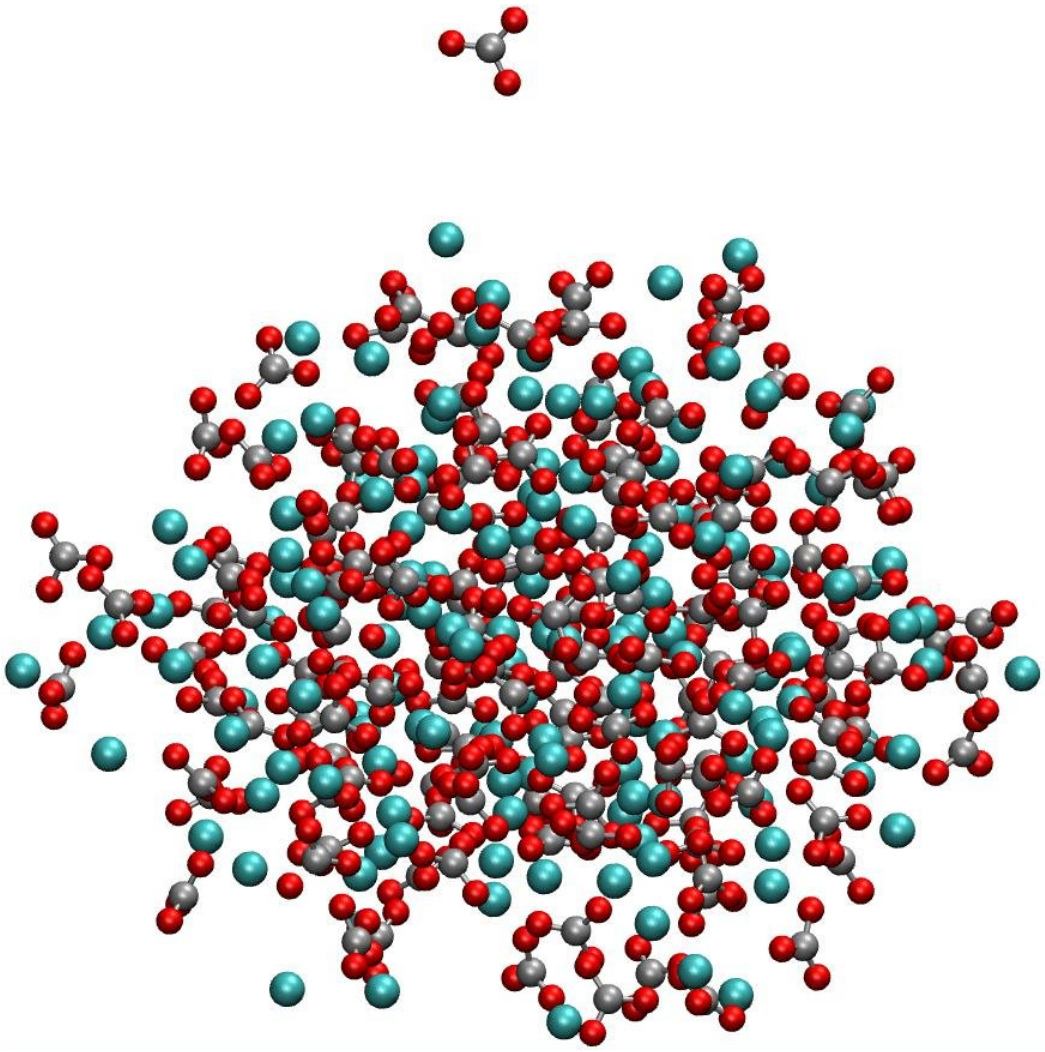

Figure 8.10: Configuration of 147 formula unit nanoparticle after brief metadynamics simulation.

Chapter 9

Conclusion

9.1 Nanoparticles of tricalcium phosphate

The available potentials for calcium phosphates have been reviewed and tested for their applicability to the polymorphs of TCP. Where the shell model is not too expensive the potential of Pedone et al. [2007] was found to be the most appropriate, while where a shell model is undesirable, as in this work, a simplification of the force-field by Mkhonto and de Leeuw [2002] was also adequate. Both of these force-fields gave good agreement with experimental structures and bulk structural properties. The development of a new force-field would however be desirable. Raiteri et al. [2010] have persuasively argued that where polymorph selection is being studied it is important to fit the potential not just to structural properties, but to thermodynamic differences between the polymorphs. A new force-field using the Ca^{2+} terms of the Raiteri 2010 force-field could be designed to be interoperable with that force-field. Since the protonation state is so important for phosphate systems it might also be useful to develop a ReaxFF [Chenoweth et al., 2008] forcefield allowing the formation and breakage of covalent bonds.

Having selected a force-field I performed a variety of metadynamics simulations on TCP nanoparticles. The inertia tensor order parameters were found to encourage the nanoparticle to break into pieces, and in any case it seems improbable that they would adequately separate crystalline and amorphous systems. Metadynamics simulations of TCP using Steinhardt parameters identified two stable states for small nanoparticles in vacuum, separated by a small energy barrier. The intrinsic complication in the TCP unit cells does however mean that identifying these

states with bulk polymorphs would be extremely difficult. Given the complicated nature of these bulk polymorphs the simulation study of TCP nucleation and growth is likely to continue to lag behind calcium carbonate systems. A ReaxFF system would allow the simulation of a much greater variety of compounds, including many more relevant to the formation of mammalian bones and teeth.

9.2 Application of topology principals to biomineral simulation

In this work I present the first application of topological principles to analyse local order in biomineral systems, and throughout the work I have demonstrated the ability of this technique to identify all three crystalline polymorphs successfully. An ideal local structure analysis method would be able to identify smaller areas of crystalline order in aragonite and vaterite systems, and hence assign ions closer to surfaces, yet even as implemented the parameters are perfectly adequate for the purposes for which they have been utilised.

The extension of such techniques to continuous, smooth, global order parameters was less successful. When applied to biomineral systems they were able to drive conformational change, however they gave poor separation between crystalline states, and crucially, the transition states and amorphous states. They were however able to drive solid–solid phase transitions in nanoparticles of cadmium selenide, with trajectories moving between the rock-salt and zincblende structures. For further work, it would be interesting to compare the performance of these parameters with the related ones of Pietrucci and Andreoni [2011] and Donadio et al. [2005] on the same systems, to gain a better understanding of the strengths and weaknesses of the different methodologies.

9.3 Nanoparticle generation and evaluation

For the first time with the new forcefield by Raiteri and Gale [2010], surface energies have been calculated for aragonite and calcite. For anhydrous systems the surface energies have been used to generate aragonite nanoparticles using the Wulff construction technique. These crystalline nanoparticle surfaces are (101), (001), (011), (210) and (211).

The negative surface enthalpies for the hydrated systems are consistent with the negative enthalpy of calcium carbonate dissolution. The entropic penalty of ordered water around the mineral phase ions is crucial to the solubility of the calcium carbonate system. One cannot however use the calculated negative surface enthalpies in the Wulff construction methodology. To extend this work, it might be possible to use experimental morphologies of aragonite nanoparticles to calculate the surface free energies for some surfaces and hence estimate the entropic correction for some surfaces. However, any additives within the solvent might dramatically alter the surface free energies, and hence nanoparticle morphologies.

The ISMC technique has been used to predict low energy nanoparticle conformations for all three crystalline polymorphs across a continuous range of nanoparticle sizes. For aragonite, the recovered nanoparticles are lower in energy than the Wulff constructions. Following relaxation in vacuum, calcite is the lowest energy polymorph, followed by aragonite, then vaterite, though the gap between aragonite and vaterite is small.

The ISMC technique may have wider applicability: there may be many systems where the cluster size of interest is larger than that accessible to global search methods, but small enough so that edge and corner effects cause Wulff constructions to fail in predicting the lowest energy conformations. The ability to account for disorder within the unit cell may also be advantageous for some systems, where calculation of surface energies becomes intractably complicated. Extensions could include producing a suite of candidate conformations which then undergo basin hopping [Wales and Doye, 1997], finding lower energy nanoparticles perhaps at the cost of producing nanoparticles with more amorphous surface layers.

When hydrated these nanoparticles were found to undergo a lengthy relaxation and equilibration process, which extended beyond time scales I was able to simulate. The eventual endpoint of this process remains a mystery. At particle edges and corners chains and links of ions were observed to form, and in some cases broke away from the nanoparticles. This may represent a continuous, dynamic process, in line with classical nucleation theory, of ragged edges gaining and losing ions, with occasional deposition or dissolution of whole surface layers. Alternatively, the process may simply be one of amorphisation: Raiteri and Gale [2010] have shown that hydrous ACC nanoparticles are lower in energy than calcite up to 480 formula units, and it could be that the crystalline phases aren't even metastable at the sizes considered in this work.

Nevertheless, from the values obtained through hydrated nanoparticle simulations it does appear that the neglect to date of aragonite and vaterite nanoparticle simulations is unwarranted. While varying heavily across particle size, vaterite or aragonite nanoparticles frequently had lower enthalpy than calcite nanoparticles, and appeared to undergo no more amorphisation than the calcite nanoparticles, the focus of study to date. One key unanswered question is whether the energetic ordering of the nanoparticles remains when entropic contributions have been considered.

9.4 Nanoparticle phase-transitions

Trajectories between the three calcium carbonate crystalline polymorphs have been produced using the nudged elastic band method. All trajectories pass through an intermediate amorphous phase, higher in energy than the crystalline polymorphs. Trajectories started from these midpoints did not return to either endpoint, instead remaining amorphous, and the energy landscape between any two crystalline polymorphs is extremely rugged. This renders the calcium carbonate system unsuitable for the transition path sampling technique, where the vast majority of trajectories generated would become trapped in amorphous local minima without reaching either crystalline end-point.

Nevertheless the NEB trajectories themselves provide a highly useful tool for studying order parameters. Calculating the correlation coefficient between them illustrates when an additional order parameter gives no new information, while plotting the trajectories on a 2D CV surface enables qualitative judgements about whether those CVs give good separation states along a pathway. The Steinhardt parameters were shown to outperform the others on this measure, and as I have shown correlations between them mean metadynamics simulations using them need only two representative OPs, not the five of work such as Quigley et al. [2011].

When tested with the modern potential however, hydrated nanoparticle systems were shown to lose ions to solution, when biased using these parameters. This nanoparticle dissolution, and indeed the expulsion of water from ACC nanoparticles, is undoubtedly a slow process in metadynamics terms, yet is not encapsulated by the OPs used. Further metadynamics simulations of the hydrated calcium carbonate system would have to address this problem, probably through the use of additional order parameters. A crude measure of the number of water–mineral bonds might be a reasonable first attempt.

9.5 Implications for crystal growth and nucleation

The recent potential of Raiteri and Gale [2010] has dramatically different characteristics from those derived from Pavese et al. [1992], and in this work we have revealed some of the implications of this. The work of Cooke and Elliott [2007] has crystallinity being stabilised by water, yet in this work water facilitates at least some nanoparticle amorphisation, or perhaps even the beginnings of dissolution. The ions breaking off is actually more consistent with classical nucleation theory than non-classical theories such as nanoparticle aggregation. In the series of metadynamics simulations beginning with Quigley et al. [2009] aragonite is not found, yet I have shown that it is more stable in terms of enthalpy than calcite for many nanoparticle sizes. I have also shown that when considering the phase space of nanoparticle confirmations using metadynamics, it is necessary to bias water–mineral phase interactions.

Chapter 10

Method

This chapter gives more detail about the exact nature of the simulations ran for this work.

10.1 Tricalcium phosphate

Since the potential chosen for use in the TCP nanoparticle simulations was not originally used for molecular dynamics, and has been modified (see chapter 3) it was necessary to choose an appropriate timestep length and thermostat relaxation parameter. This section also includes details of the metadynamics parameters used in TCP simulations, and a description of the method used to generate amorphous nanoparticles.

10.1.1 Choice of time step

An appropriate choice of time step is important to ensure simulation accuracy while maximising computational efficiency. With too long a time step the total energy will start to drift in an (N, V, E) simulation, as errors in the integrator build up. Too short a timestep and expensive computer time is needlessly wasted.

To investigate the appropriate choice for this potential I simulated a large block of β-TCP (9 unit cells — 2457 atoms) for at least 2 ns of (N, V, E) simulation at 300 K using timesteps from 1.0 fs to 6 fs. I then monitored the total energy to see if it drifted over time. For timesteps less than 5 fs the total energy is constant at -27250 eV over the entire length of the simulation to within 1 eV. Only above this value does the energy start to drift upwards as integration errors mount, rising by

about 7 eV over the course of a 2.5 ns simulation. A time step size of 4 fs is quite large for biomineral simulations, and this simulation was of a bulk stable mineral phase rather than a nanoparticle with its complicated surfaces. I therefore selected a time step of 2 fs for further simulations. It is likely that incorporating a water model, with the rapid fluctuations of the hydrogen atoms, would require further lowering of the time step.

10.1.2 Thermostat relaxation parameter

The thermostat relaxation time for the Nosé-Hoover thermostat must be carefully selected to ensure adequate sampling of the canonical, or (N, V, T), ensemble. This parameter controls how tightly atomic velocities are coupled to the thermostat. A low value increases the efficiency of the sampling, yet too low a value causes the sampling to break down. Too high a value and the motion of the thermostat and particle can become decoupled. The key diagnostic is whether the instantaneous temperatures are sampled from a normal distribution with standard deviation $T\sqrt{\frac{2}{3N}}$. I ran a variety of test values with the same large block of β-TCP as before for 2 ns. Figure 10.1 shows the histogram for the value 10 fs, with fitted Gaussian. The width of 4.92 K gives good agreement with the desired value, for 2497 atoms at 300 K, of 4.90 K, suggesting this value is adequate. All (N, V, T) TCP simulations in this work use at least this value or greater.

10.1.3 Production of nanoparticles

It was considered desirable to produce an amorphous nanoparticle as a starting point for metadynamics simulations. Following Quigley and Rodger [2008a] I first simply took a β-TCP unit cell in vacuum and ran simulations at a hot enough temperature to remove all order from within that cell. Even at 300 K however, exposing the (001) surface at the edge of the unit cell was so undesirable that it caused grossly unphysical distortion in the phosphate groups. After much experimentation it was found that the following protocol was sufficient to produce an amorphous TCP nanoparticle of 273 atoms:

- Take one unit cell of β-TCP, place in vacuum and find the 0 K local minimum. I performed this step using GULP, and note that the particle distorted to bury the (001) surface.

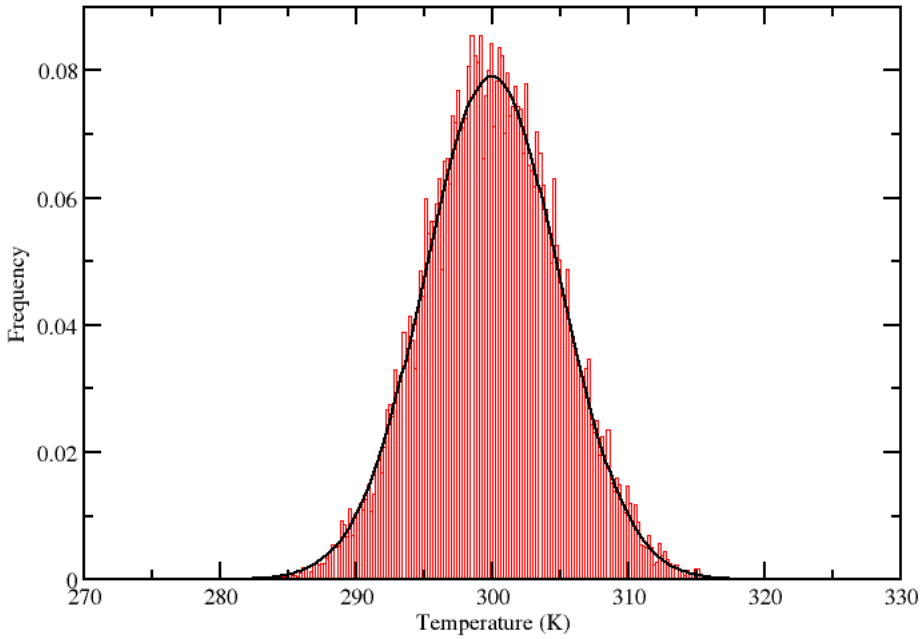

Figure 10.1: Normalised histogram of instantaneous temperature samples from a β-TCP simulation. The thermostat relaxation time was 0.01 ps. The fitted Gaussian has mean 300 K and standard deviation 4.92 K.

- Convert the phosphate groups to rigid units, and simulate the nanoparticle at 1500 K for at least 2 ns. Using the flexible potential, I was never able to run at sufficient temperature to ensure rapid amorphisation without causing the phosphate groups to disassociate. By using a model with perfect, rigid, tetrahedra, but leaving the intermolecular terms unchanged this issue was avoided.

- Simulate at 300 K with the rigid potential for a short time to anneal the structure, and then switch back to the flexible potential to verify the stability of the resulting nanoparticle.

From consideration of the radial distribution functions (figure 10.2) the resulting nanoparticle appears to be amorphous. Both the Ca–Ca and Ca–P RDFs exhibit only one coordination shell peak before losing any order. The P–P RDF does show two peaks put they are poorly separated, and after 6 Å, the RDF again shows no order. Comparison with the RDFs for the crystalline TCP phases (figure 10.3) shows that the bulk phases have a much greater amount of long range order.

10.1.4 Inertia tensor metadynamics parameters

The code to perform these simulations was written by David Quigley using an interface to DL_POLY 2.20. The simulations were performed in the (N, V, T) ensemble using a Nosé-Hoover thermostat with a time step of 2 ns and a thermostat relaxation time of 0.5 ps — as justified above this is easily long enough to ensure adequate sampling from the canonical ensemble.

Gaussians of height 5 $k_B T$ were deposited every 1000 timesteps. Width was chosen to be of the same order as the fluctuations in the CVs over the course of short regular MD simulation. For this 273 atom simulation this width was about 6000 amu.Å^2 (figure 10.4). Simulations were conducted for at least 5,000,000 timesteps (10 ns — though biased).

10.1.5 Steinhardt metadynamics parameters

In their previous work on calcium carbonate, Quigley and Rodger [2008a] used a system of six CVs, with five Steinhardt parameters, plus the potential energy. Each additional order parameter causes an exponential increase in the time taken for adequate sampling, and so this number is more than is desirable. As the phosphate

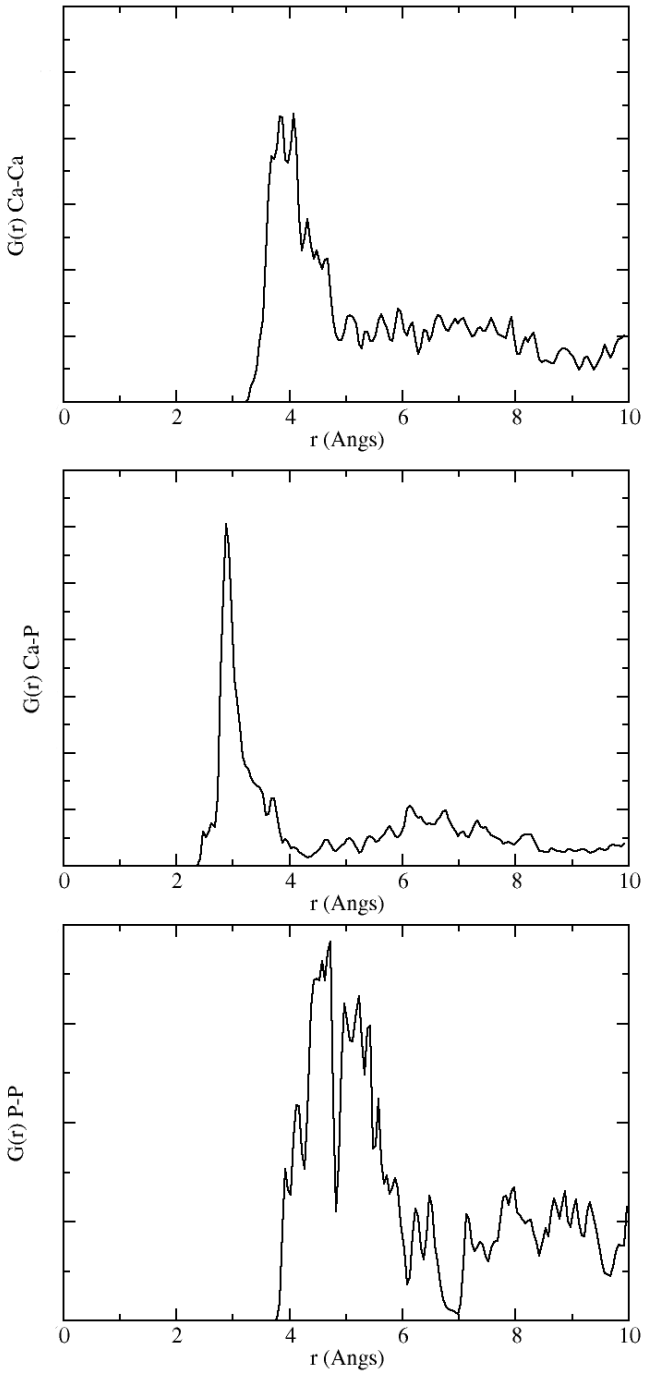

Figure 10.2: Radial distribution functions for TCP nano particle. Since the volume term in the RDF formulation is ill-defined for a nanoparticle in an infinite vacuum, the y-axis values have been removed.

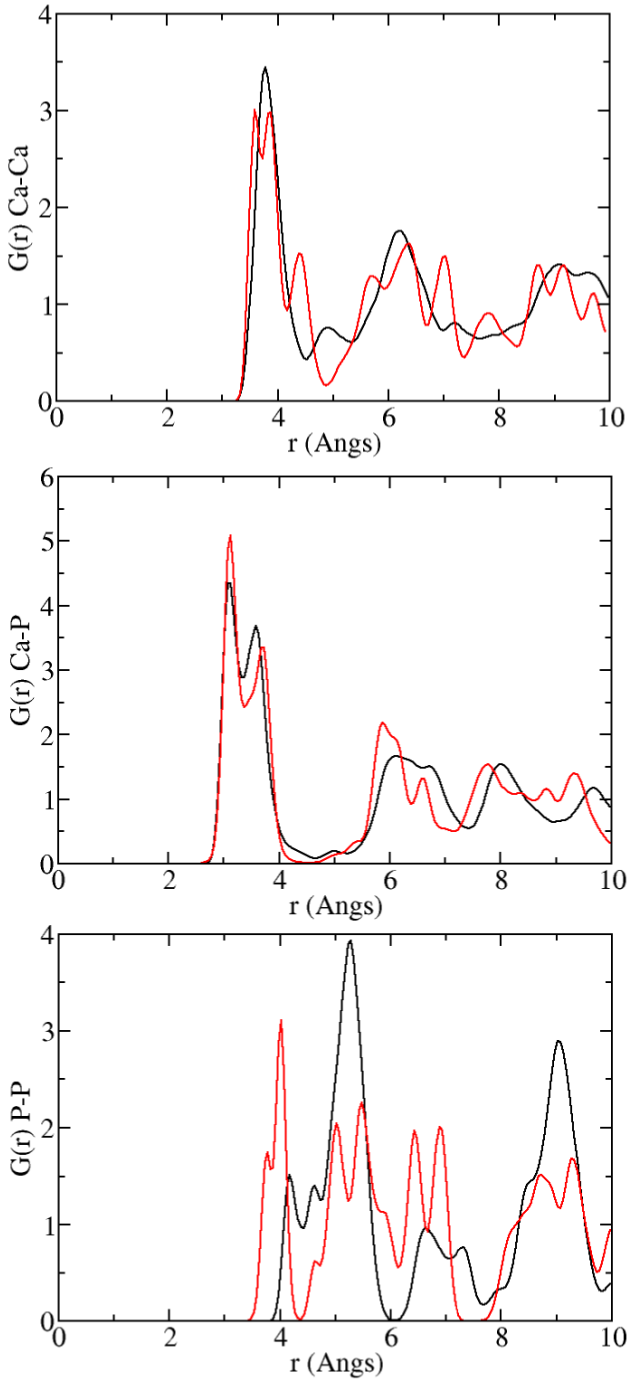

Figure 10.3: Radial distribution functions for bulk α-TCP (black) and β-TCP (red).

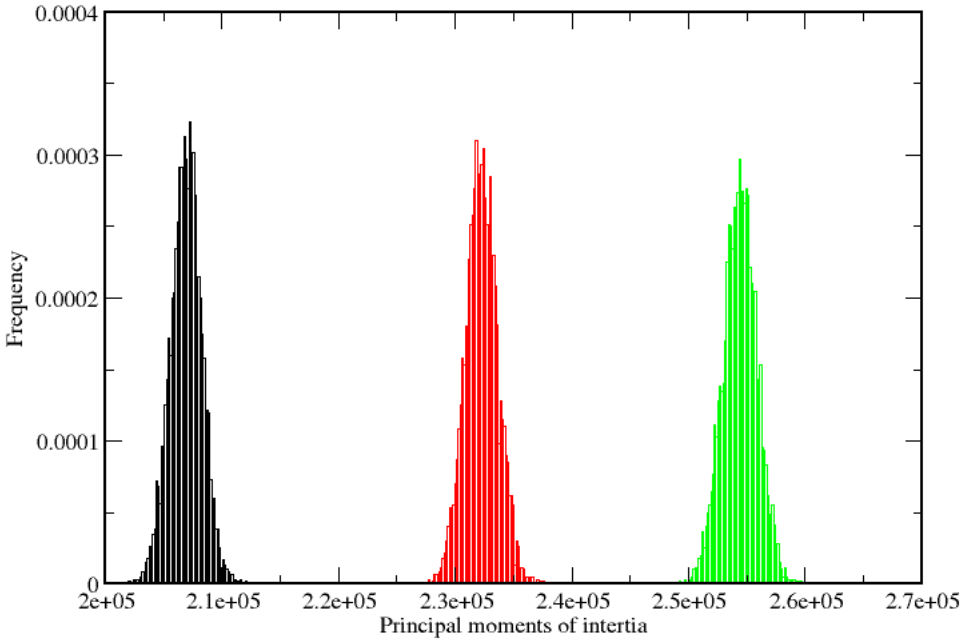

Figure 10.4: Normalised histograms of the principal moments of inertia in an unbiased MD simulation of a 273 atom nanoparticle of TCP. The units on the x-axis are amu.Å$_2$

groups are tetrahedral, and more freely able to rotate than the planar carbonate groups, they might be able to find their preferred orientation without additional explicit bias. Thus it was reasonable to exclude the oxygen atoms, and use only the Ca–Ca, Ca–P, and P–P atom pairs for Steinhardt order parameters, with the addition of the potential energy of the nanoparticle to make four CVs in total. Simulations were ran using a modified version of DL_POLY.

The Steinhardt parameters required the choice of inner and outer cut-off for the function which determines how the contribution of each pair varies with pair distance. Following Quigley and Rodger [2008a], these cut-offs would be between the first and second coordination shells of the bulk polymorphs. From the radial distribution functions of the α and β-TCP simulations (figure 10.3), this was repeated (table 10.1.5), though the more complicated nature of the unit cells mean this region is not as clearly defined.

Pair	r_c	r_o
Ca–Ca	4.6	5.2
Ca–P	4.3	4.9
P–P	5.7	6.2

Table 10.1: Inner and outer cut-offs for the three Steinhardt parameters

The next parameters to select were the width and height of the Gaussians deposited, and their deposition rate. As with the inertia tensor CVs the width was chosen to be of the same order as that of the thermal fluctuations in an unbiased simulation. For the Steinhardt parameters this width was approximately 0.167 (figure 10.5), while for the potential energy this width was of the order of 20000 eV (figure 10.6). For the Gaussian height early simulations used a height of 5 k_BT and thoroughly explored the energy landscape, giving an energy difference of 900 K_BT between the lowest minimum and the boundary of the explored area. These simulations indicated two minima separated by about 50 k_BT but otherwise no significant features on the energy landscape. To explore this region with greater accuracy well-tempered metadynamics was used (see section 2.7). In this technique the energy landscape is only 'filled up' to a pre-chosen depth, with the height of deposited Gaussians reduced as the basin around the start point fills. A maximum height of 200 k_BT was chosen with initial height 5 k_BT. Simulations were run in the (N, V, T) ensemble at a temperature of 300 K, with Gaussians deposited every

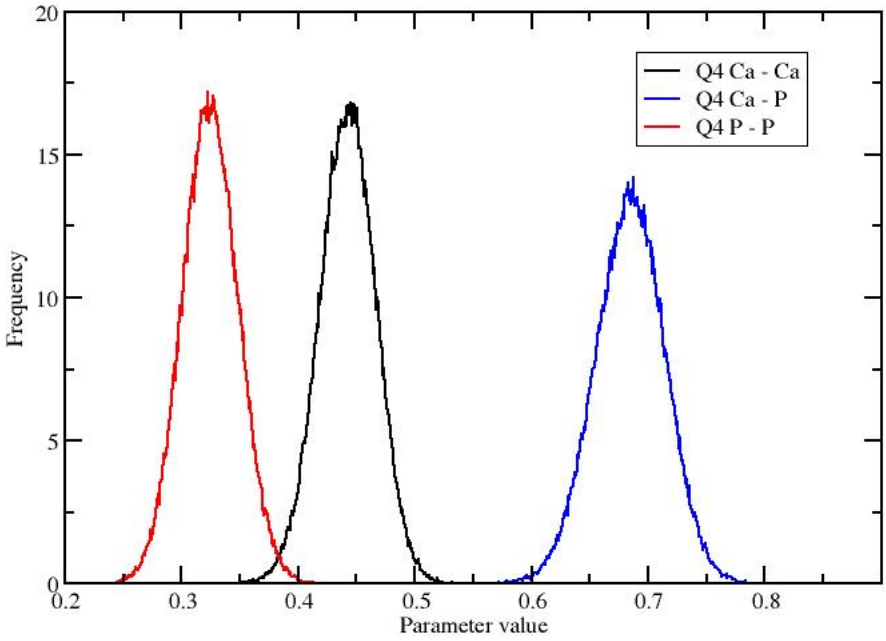

Figure 10.5: Normalised histograms for the three Steinhardt Q4 parameters.

1000 timesteps.

10.2 Calcium carbonates

All calcium carbonate simulations used the potential of Raiteri and Gale [2010], and followed that work in using a timestep of 1 fs, a thermostat relaxation time of 0.1 ps, and a barostat relaxation time of 1 ps.

10.3 Metadynamics using continuous topological order parameters

Metadynamics code was produced by adapting the principal moments of inertia metadynamics code written by David Quigley discussed above, which is itself based

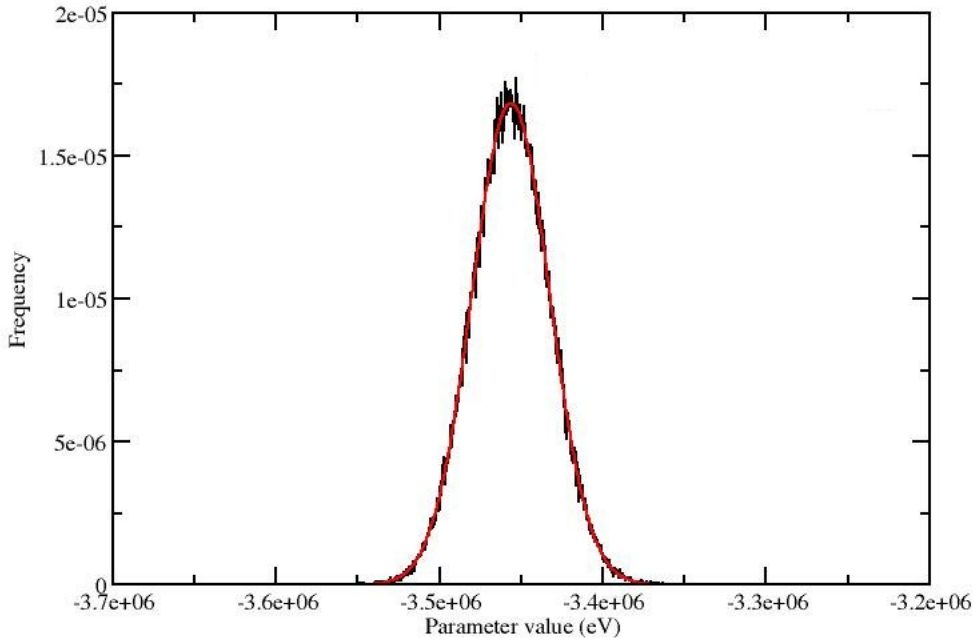

Figure 10.6: Normalised histogram for the potential energy (BLACK), with fitted normal distribution (RED - mean $-3.45e6$ eV, standard deviation $2.37e5$ eV).

on an interface written by Dr Quigley to the force calculation routines of DL_POLY, This code simulates the (N, V, T) ensemble using a Nosé Hoover thermostat. Several variations on the same algorithm to calculate these order parameters were tried as discussed in appendix A, varying in how and whether parallel programming techniques were applied. I note that appendix contains the first description of efficiently calculating ring statistics within any kind of MD simulation of which I am aware: Donadio et al. [2005] likely used similar techniques, however the details were not included in that work. Parallel and serial versions of the code are included in appendix B. All of these simulations were ran in vacuum.

Calcium carbonates

There are a great number of parameters to choose using topological order parameters. As well as the standard metadynamics parameters of Gaussian width, height, and deposition interval, there are the choice of which atoms to include on the graph, which links to include, and which cut-offs to choose. Numerous variations were tried: here I describe the most successful. As with the local structure analysis methodology, only calcium and carbon atoms are nodes on the graph, and only Ca—C pairs contribute links to the graph. Variants with C—C and Ca—Ca links were tried, but did not perform better and were computationally far less efficient as far more rings were present in the graph structure. The inner and outer cutoffs for use in equation 5.2 are 3 and 5 Å respectively, allowing the metadynamics forces to push atoms a considerable distance. The simulation was performed at 300 K with a 90 formula-unit calcium carbonate nanoparticle in vacuum. Based on fluctuations in the quantities in the unbiased simulation, a width of 20 rings was selected for n_4, the number of 4-membered rings, 18 rings for n_6, and 4 for n_c the coordination number. Gaussian height was 5 k_BT, with Gaussians deposited every 500 timesteps, until more than 10000 Gaussians had been deposited for a total simulation length of 5 (biased) nanoseconds. The initial configuration was a calcite rhombohedron of the kind predicted by Wulff constructions (section 6.2).

Cadmium Selenide

The simulation used the potential of Rabani [2002], a combination of electrostatics with Lennard-Jones terms for the Van der Waals interactions. The simulation was conducted at 300 K with a thermostat relaxation time of 100 fs. Both cadmium and

selenium atoms were considered as nodes on the graph, with only the cross bonds between them forming edges, so as with the carbonate system, only four and six membered rings would need to be found. The inner and outer cutoffs for the bond length were 2.5 Å and 4.0 Å respectively. The simulation was continued for more than 2,500,000 timesteps of length 1 fs, and the nanoparticle contained 134 CdSe formula units. The initial configuration was generated to be in the wurtzite structure. Gaussian height was 5 k_BT. Based on fluctuations in an unbiased simulation, the Gaussian width was 100 4-rings and 6-rings, and 1 n_C unit for the coordination number, ad the deposition interval was 500 timesteps.

10.4 Wulff Constructions

10.4.1 Calculation of Surface Energies

Vacuum surface energies were calculated using GULP [Gale and Rohl, 2003]. $D\{hkl\}$ is defined as the distance between two adjacent lattice planes for the (h, k, l) surface. Using the program GDIS [Fleming and Rohl, 2005] the 25 surfaces with highest $D\{hkl\}$ values were generated, with two distinct regions. The upper layer was allowed to relax, while the lower layer was fixed in position. The results were converged with respect to these two regions, with the upper region at least 10 Å thick, and the lower region at least 8 Å. The energy was calculated was the potential energy of the surface at 0 K, calculated using the relaxed local minimum of the starting structure. This is calculated by comparing the energy of this structure to the energy of the bulk polymorph.

The calculation of the hydrated surface energies is more involved. Here a slab of each surface approximately 40 Å thick, with sides of a similar length, was generated using a Perl program written for the task. A copy of the program is in appendix B. A water film of 15 Å thickness was then placed at the top and bottom of the slab. Using periodic boundaries this produced simulation boxes of approximately 40 x 40 x 70 Å^3, with a water thickness of approximately 30 Å. Equilibration was conducted by freezing the mineral phase and allowing the water to position itself around it for 0.25 ns. During this phase of (N, P, T) simulation the box could only contract in the z direction. During data collection the box could expand and contract independently in each direction, with a barostat maintaining atmospheric pressure. Data collection proceeded for at least 1 ns during which the

enthalpy and surface area were recorded every picosecond. These calculations were performed in the program LAMMPS [Plimpton, 1995]. The energy calculated was the surface enthalpy at 300 K.

Separate simulations of bulk aragonite and bulk water were conducted and the average enthalpies per formula unit (H_A and H_W) calculated. If the surface simulation has overall average enthalpy H_F, with two surfaces each of area A_S, n formula units of $CaCO_3$ and m units of h_2O, then surface enthalpy is:

$$H_S = \frac{H_F - (nH_A + mH_W)}{2A_S} \tag{10.1}$$

Hydration enthalpies were calculated in a similar way: from the energy of a bulk vacuum slab H_V:

$$H_{\text{Hyd}} = \frac{H_F - (H_V + mH_W)}{2A_S} \tag{10.2}$$

Errors were calculated using block averaging of enthalpy and surface area values. Due to the computational expense of these simulations, the surface enthalpies were only calculated for the surfaces with the 14 highest $D\{hkl\}$ values.

Many of the surfaces have several different end terminations, and where this was the case, the surface energies for all non-polar end terminations were calculated. While in many cases in the literature it has been sufficient to differentiate between different end terminations by the uppermost atom [de Leeuw and Parker, 1998], this would be ambiguous for many of the surfaces considered here. The $(10\bar{2})$ has four different end terminations for example, two of which are terminated by carbonate groups, two by calcium ions, with all four terminations giving different surface energies. To remove ambiguity it is convenient to define these end terminations mathematically. As mentioned above $D\{hkl\}$ gives the distance between two identical images of the lattice plane specified by the Miller indices h, k and l. If $\hat{\mathbf{n}}$ is the unit vector normal to those lattice planes then $D\{hkl\}\hat{\mathbf{n}}$ is a translation vector between two such identical planes. $s * D\{hkl\}\hat{\mathbf{n}}$ for some s between 0 and 1 can then uniquely define any plane with any possible end termination. This s is by convention referred to as the 'shift' with a shift of 0 corresponding to the plane intersecting the origin.

10.4.2 Generating Wulff Constructions

For aragonite Wulff constructions, a program in Perl was written (a copy is in appendix B). Given a scaling factor it picks a site within a unit cell at the centre of a very large lattice. The surface energies are hard-coded into the program, and for each surface in each direction the lattice is cleaved at the closest corresponding plane. The overall charge of the resulting nanoparticle is then measured, and if zero, the structure is saved. Each run of the program tries 100 random start sites within the unit cell, to help mitigate against the rounding issues discussed above. For vacuum Wulff constructions the program was ran for scaling values (very roughly corresponding to particle radius) between 5 and 15 $\text{Å}\text{m}^2\text{J}^{-1}$, at intervals of 0.1 $\text{Å}\text{m}^2\text{J}^{-1}$. It was not possible to produce Wulff constructions from the hydrated surface enthalpies for reasons discussed in chapter 7.

Wulff constructions for calcite were produced using Fortran 90 code supplied by D. Quigley. The method used by this code is similar — a polyhedron of the Wulff construction shape is produced, and filled with atoms. Different centre points for this polyhedron are tried in order to find charge neutral particles. In this case the $(10\bar{1}4)$ surface is so completely dominant that other surfaces need not be considered, a result supported by both theoretical and experimental results [de Leeuw and Parker, 1998].

10.5 Random Structure Searching

I have implemented RSS on nanoparticles of calcium carbonate. The particles were positioned within a sphere in vacuum (the radius of this sphere is a parameter). The carbonates were positioned first, with their orientation chosen at random. Each atom was coplanar with the other atoms of the carbonate with a C–O distance of 1.19 Å, and a O–C–O bond angle of 120 °. The minimum distance between two carbon atoms of different carbonate groups was set at twice the C–O distance plus 1.5 Å, ensuring the closest possible distance for two oxygen atoms is 1.5 Å. The calcium ions are positioned at random within the sphere, again at a minimum distance from all other atoms of 1.5 Å.

Energy minimisation was performed using DL_POLY software using a zero temperature molecular dynamics run, or by a conjugate gradient based method using an interface to DL_POLY written by D. Quigley, as described in section 2.6.

10.6 ISMC

The ISMC procedure was run for all particle sizes in the range of 80 to 180 formula units for calcite, aragonite, and vaterite. Each simulation was of at least 12 billion trial moves, with most finding the lowest energy configuration within 3 billion moves.

Rather than calculate all potential energy terms explicitly, the code exploits the restricted nature of the ensemble to tabulate all possible values of the potential function between any two pairs of ions. This speeds up calculation by a factor of 20 at 200 formula units. As the carbonate groups cannot relax all intramolecular terms can be neglected. Written in Fortran 90, copies of all three versions of this code can be found in appendix B.

Replica exchange was attempted on average every 50000 MC moves, with the interval drawn from the exponential distribution, to ensure replica exchange is a Poisson process with events occurring independently. For calcite and aragonite the temperature of the lowest simulation was 3000 K, rising in a geometric progression to 12000 K in the highest temperature simulation. These high values were due to the highly restricted nature of the ensemble which only allows very specific movements. This meant local minima were separated by very high energy barriers, and few moves were accepted at temperatures below 6000 K. Above around 11000 K the system undergoes a phase transition to a lattice gas and ceases to be a contiguous nanoparticle. For vaterite a linear progression in temperature was used, from 1 K to 12000 K, for reasons discussed in section 6.4.3. Each simulation was ran at either 8 or 12 temperatures simultaneously.

10.7 Investigating nanoparticle stability

A selection of the generated ISMC nanoparticles for all three polymorphs were relaxed in vacuum, and then immersed in water, in a periodic box 62 Å to a side. For each simulation the mineral phase was fixed for a short simulation while the water was allowed to equilibrate around it, and the particle was then simulated in the (N, P, T) ensemble for at least 12.5 ns using the thermostat and barostat parameters of Raiteri and Gale [2010], at 300 K and standard pressure. Some larger Wulff constructions of calcite were also simulated using this procedure. A full list of nanoparticles simulated can be found in table 7.6. These simulations were conducted using LAMMPS. Enthalpy values were calculated after every picosecond, and the

values presented are rolling averages over a nanosecond length scale, except where stated otherwise.

Appendix A

Ring statistics algorithms

Calculation of ring statistics can be an involved, computationally expensive process. Here I describe the local implementation of the algorithm of Hobbs et al. [1998] for local structure analysis, and extensions I have made to enable the efficient calculation of the continuous order parameters within a metadynamics simulation.

A.1 Local Structure Algorithm

I have implemented the algorithm of Hobbs et al. [1998] locally using Fortran 90. The algorithm proceeds as follows:

Initialise: build graph. Having read in the atomic coordinate data, the nodes are extracted — in the calcium carbonate case the carbon and calcium atoms are the only nodes on the graph. By looping over all pairs of nodes and calculating the distance between them in angstroms, the appropriate nodes are linked by edges, to populate an edge-list data structure. This is a list of all the neighbours of each of the nodes. Once this is done the following stages are performed for each node on the graph for which the local cluster is to be calculated.

1. **Find nearby nodes.** The maximum ring length $maxl$ is a parameter — typically a value of six was found to be sufficient. Nodes further than $maxl/2$ links from the home node cannot be part of rings passing through the home node, and hence need not be considered. A local list of atoms of this distance or less is compiled.

2. **Build distance matrix.** From this list of D nodes, a DxD matrix is produced

of the smallest number of edges between any two nodes. This is accomplished with a variant of Dijkstra algorithm [Dijkstra, 1959].

3. **Breadth first search of paths.** A path is a series of linked nodes. At this stage essentially every possible path starting at the home node is followed to find all rings of length less than or equal to $maxl$. This proceeds through the use of a first-in, first-out queue system. The first path in the queue is simply the home node. The following procedure is then followed until the queue is empty:

   - Unqueue a path. Check the length of the path: if it is equal to $maxl$, then discard this path and move onto next. Else, loop over each of the neighbours of the end node of the path, creating a new path for each by adding the neighbour to the unqueued path.
   - For each new path, check if the new end-node has already been visited by the path. If it has not then enqueue the new path.
   - If it has, then check if the end-node is the home node. If it is the home node, then providing the path length is more than two a ring has been found and can be added to a list. If it is not the home node then the path is intercepting itself, and can be discarded and not enqueued.

   One additional check is made before adding a ring to the list: this method double counts each ring as there is an implied directionality in the path process. For example there is a ring from nodes $1-2-3-4-1$, the algorithm would record both $1-2-3-4$ and $1-4-3-2$ as rings. To prevent this, a ring is only recorded if the second node on the ring has a lower index than the final node, so in this case $1-2-3-4$ would be recorded.

4. **Check primitivity of rings**. For each ring, loop over all pairs of nodes and check the distance around the ring is equal to the value in the DxD matrix of distances. If it is not, then the ring is not primitive as there must exist a short-cut across it. Else a primitive ring has been found and can be added to a list.

The details of the local cluster (the number of nodes, and the number of primitive rings of each length) are then easily extracted from the primitive ring list.

A.2 Metadynamics Extension

While suitable for the calculation of a single local cluster the algorithm above is inefficient for the purposes of finding the full set of primitive rings in the whole graph — the same rings are explored multiple times with useful information discarded after finding each cluster. For the purposes of local structure analysis efficiency might be desirable, but it is not essential — for nanoparticles typically considered in this chapter, the code described above will complete the local structure analysis in a few seconds on modern hardware. This would however be inadequate as a subroutine in a metadynamics code that has to run every timestep. An additional complication is the need to calculate the full order parameters from the compiled ring statistics, as well as the partial derivatives of these quantities for every atom that is a node on the graph.

Yuan and Cormack [2002] claim that their algorithm is more efficient, particularly in the case where all primitive rings are to be found rather than one local cluster, however this algorithm finds only primitive rings, and so is unsuitable for this work where 'primitivity' is a continuous quantity. Goetzke and Klein [1991] also have an algorithm suitable for finding the whole network, though again their algorithm only finds primitive rings, and, as argued by Hobbs et al. [1998], fails to find all of them.

A.2.1 Serial Version

Here we divide the algorithm into three separate sections:

- **Initialisation at start of simulation.** As in the local structure code, the first step is to calculate the edge-list. As this involves calculating the distance in terms of the atomic coordinates, this is also a convenient stage to calculate the f-scores of each edge, as well as the partial derivatives in each of the x, y and z directions. The list of all rings is then populated. Working from a copy of the edge-list, this proceeds by looping over all nodes. For each node, all the rings as short or shorter than $maxl$ containing that node are found by a similar path exploration technique to that in the local structure code. After this has been done all the edges attached to that node are removed from the copy of the edge-list. In this way each ring is found only once and the copy of the edge-list becomes more and more sparse as the rings are located. This gradual removal of edges is similar to that conducted by Goetzke and Klein [1991]. A

major difference between this algorithm and the local structure code is that the full matrix of distances between every pair of nodes is not calculated, as where every ring, primitive or not, is of interest, it is unnecessary and would be computationally expensive to calculate for the full graph.

- **Update rings every MD timestep.** The updating process proceeds by recalculating the full edge-list and all f-scores and their partial derivatives. Once this has been accomplished the new edge list is compared to the version calculated the previous time step. Two lists are compiled: old edges, which are no longer on the graph, and new edges. All the rings on the ring list featuring old edges can be deleted. More complicated is the finding of new rings featuring the new edges. This involves the same path exploration technique as before, only using each new edge in turn as the seed path and not just a node. By restricting the expensive searching for rings to only those cases where it is strictly necessary the algorithm is vastly more efficient than a new search every timestep would be.

- **Calculation of order parameters** This can occur after either initialisation or updating of the ring list. The calculation of q scores for each ring is straightforward and proceeds by simply looping over every edge in the ring. The p scores are more complicated, and for ease of implementation are dealt with differently for rings of different lengths. Three membered rings are inherently always primitive. For four membered rings it is only necessary to check if there exists an edge between the pairs of far corners. For 5-rings there are five possible direct shortcuts and all are checked. For 6-rings the situation is more complicated again: there are 9 possible shortcuts in all, of which three can be of two edges in length, and all must be checked and noted for contribution to the p score for that ring. Support for rings larger than this has not been implemented. The shortcutting nodes are also recorded for their contribution to the partial derivatives of each order parameter.

Implementation is complex, and to ensure accuracy a test suite was produced. The tests were as follows:

- **Exact answers for a small system.** For a rough cube of atoms linked at the edges the order parameters were calculated independently, and compared to those given by the program.

- **Update procedure test.** The update procedure was tested by perturbing a configuration at random, then testing to see if the updating subroutines gave the same answers as a full re-initialisation of all rings.

- **Partial derivative test.** The accuracy of the partial derivative calculation was tested by calculating an estimate using the finite difference method for each of the atomic coordinates in each of the three Cartesian directions.

A.2.2 Parallel Versions

For some of the systems studied the serial version of the code performed extremely slowly. There is however considerable potential to speed up the algorithm by parallelising it, and indeed this has been performed. Two different variants were written, both using the MPI protocol, with both subjected to the same test suite as the serial version of the code. These are the first attempt at any parallel ring finding algorithm of which I am aware.

The simplest scheme involves just parallelising the calculation of order parameter values. This subroutine involved a simple loop over all rings, and as each processor has the same copy of the ring list each can work on a small subset of it. The data is then combined once this has been performed.

More complicated was the parallelisation of the ring finding section of the code. A version of the algorithm was written where each processor maintained only a subset of the full ring list. During initialisation the processors would find rings starting from different nodes, with a system of carefully sequenced edge deletion ensuring that each ring would be found by one and only one processor. During updates before every MD step, each processor would take it in turns to process new edges, and add the resulting new rings to its ring list. Each processor would then calculate the order parameters just for its own ring list, after which these values would be combined. This variant of the code lowers the memory required by each processor, and in all likelihood the amount of memory swapped to the processor cache. Counteracting this benefit, the number of rings can vary considerably between processors, and the speed of the program will proceed at the rate of the slowest processor. Ring finding is also actually a small cost compared to the OP calculation section, as during most MD steps there aren't actually any new or old edges, while the order parameters need to be calculated at every step. As a result there was no significant performance difference between the two variants.

Appendix B

Included programs

The files detailed in this appendix are available on CD only, and are described in more detail in the relevant sections of this thesis.

- **Chapter 3 — Potentials of tricalcium phosphate.** DL_POLY `FIELD` files for both simplified and full versions of the Mkhonto and de Leeuw [2002] potential are included.
- **Chapter 5 — Application of graph theory to molecular simulation.** Local order parameter calculation tool '`locstruc_caco3.f90`', based on discrete topological rings statistics. Continuous ring statistics order parameter code is also included, with two variants of the parallel code and the serial code. As I have not included David Quigley's interface to the force calculation routines the makefiles compile against the test suite. The code in the parallel 1 directory parallelises only the OP value calculation sections, while the parallel 2 variant parallelises the ring finding section.
- **Chapter 7 — Nanoparticle generation results.** This directory includes ISMC code for calcite (`ISMC_calc.f90`) , aragonite (`ISMC_arag.f90`), and vaterite (`ISMC_vat.f90`) as well as Wullf construction code for aragonite (`make_arag_wulff.pl`). IMSC nanoparticles, both in DL_POLY `CONFIG` and XYZ file formats, across the size range from 80 to 180 formula units, are also included for all three polymorphs. The aragonite Wulff constructions are also included in XYZ format.
- **Chapter 8 — nanoparticle polymorph transitions.** Nudged elastic band code (`nano_neb.f90`) is supplied, though this also requires D.Quigley's inter-

face to DL_POLY to run. The pathways between each of the three polymorphs analysed in this thesis are included in PSF and DCD format.

Bibliography

L. Addadi, S. Raz, and S. Weiner. Taking advantage of disorder: Amorphous calcium carbonate and its roles in biomineralization. *ChemInform*, 34(33):no–no, 2003. ISSN 1522-2667. URL http://dx.doi.org/10.1002/chin.200333237.

A. Barducci, G. Bussi, and M. Parrinello. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Physical Review Letters*, 100(2):020603–, January 2008. URL http://link.aps.org/doi/10.1103/PhysRevLett.100.020603.

C. Bealing, G. Fugallo, R. Martonak, and C. Molteni. Constant pressure molecular dynamics simulations for ellipsoidal, cylindrical and cuboidal nano-objects based on inertia tensor information. *Physical Chemistry Chemical Physics*, 12(30):8542–8550, 2010. ISSN 1463-9076. URL http://dx.doi.org/10.1039/C004053C.

B. J. Berne, G. Ciccotti, and D. F. Coker. *Classical and quantum dynamics in condensed phase simulations: proceedings of th International School of Physics "Computer Simulation of are Events and the Dynamics of Classical and Quantum Condensed-Phase Systems", Lerici, Villa Marigola, 7 July- 18 July 1997*. World Scientific, 1998. URL http://books.google.co.uk/books?id=v2eAcgAACAAJ.

J. Bhimasenachar. Elastic constants of apatite. *Proceedings Mathematical Sciences*, 22(4):209–214, October 1945. URL http://dx.doi.org/10.1007/BF03170956.

L. Boyer. PhD thesis, Institut National Polytechnique de Toulouse, France, 1998.

L. Boyer, J. Carpena, and J. L. Lacout. Synthesis of phosphate-silicate apatites at atmospheric pressure. *Solid State Ionics*, 95(1-2):121–129, February 1997. ISSN 0167-2738. URL http://dx.doi.org/10.1016/S0167-2738(96)00571-1.

F. Bruneval, D. Donadio, and M. Parrinello. Molecular dynamics study of the solvation of calcium carbonate in water. *The Journal of Physical Chemistry B*, 111(42):12219–12227, October 2007. ISSN 1520-6106. URL http://dx.doi.org/10.1021/jp0728306.

C. R. A. Catlow. Point defect and electronic properties of uranium dioxide. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 353(1675):533–561, 1977. doi: 10.1098/rspa.1977.0049. URL http://rspa.royalsocietypublishing.org/content/353/1675/533.abstract.

C. R. A. Catlow, M. J. Norgett, and T. A. Ross. Ion transport and interatomic potentials in the alkaline-earth-fluoride crystals. *Journal of Physics C: Solid State Physics*, (10):1627, 1977. URL http://stacks.iop.org/0022-3719/10/1627.

R. Cerf and J. Picard. *The Wulff crystal in Ising and percolation models: Ecole d'Et''e de Probabilités de Saint-Flour XXXIV, 2004*. Springer, 2006. URL http://books.google.co.uk/books?id=rm8wwGB-HK8C.

K. Chenoweth, A. C. T. van Duin, and W. A. Goddard. ReaxFF reactive force field for molecular dynamics simulations of hydrocarbon oxidation. *The Journal of Physical Chemistry A*, 112(5):1040–1053, January 2008. ISSN 1089-5639. URL http://dx.doi.org/10.1021/jp709896w.

Y.-C. Chien, M. T. Hincke, H. Vali, and M. D. McKee. Ultrastructural matrix-mineral relationships in avian eggshell, and effects of osteopontin on calcite growth in vitro. *The Journal of Structural Biology*, 163(1):84–99, July 2008. ISSN 1047-8477. URL http://dx.doi.org/10.1016/j.jsb.2008.04.008.

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. URL http://dx.doi.org/10.1063/1.2756840.

L. Dai, E. P. Douglas, and L. B. Gower. Compositional analysis of a polymer-induced liquid-precursor (PILP) amorphous $CaCO_3$ phase. *The Journal of Non-Crystalline Solids*, 354(17):1845–1854, April 2008. ISSN 0022-3093. URL http://dx.doi.org/10.1016/j.jnoncrysol.2007.10.022.

S. D'Arco and M. S. Islam. Defect and dopant properties of the oxyfluoride super-

conductor % Sr_2 CuO_2 $f_{2+\delta}$s. *Physical Review B*, 55(5):3141–, February 1997. URL http://link.aps.org/abstract/PRB/v55/p3141.

N. H. de Leeuw. A computer modelling study of the uptake and segregation of fluoride ions at the hydrated hydroxyapatite (0001) surface: introducing a % $Ca_{10}(PO_4)_6(OH)_2$ potential model. *Physical Chemistry Chemical Physics*, 6(8): 1860–1866, 2004. URL http://dx.doi.org/10.1039/b313242k.

N. H. de Leeuw and S. C. Parker. Surface structure and morphology of calcium carbonate polymorphs calcite, aragonite, and vaterite: An atomistic approach. *The Journal of Physical Chemistry B*, 102(16):2914–2922, April 1998. ISSN 1520-6106. URL http://dx.doi.org/10.1021/jp973210f.

J. P. R. de Villiers. Crystal structures of aragonite, strontianite and witherite. *American Mineralogist*, 56:758–767, 1971.

C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler. Transition path sampling and the calculation of rate constants. *The Journal of Chemical Physics*, 108(5): 1964–1977, 1998. URL http://dx.doi.org/doi/10.1063/1.475562.

B. G. Dick and A. W. Overhauser. Theory of the dielectric constants of alkali halide crystals. *Physical Review*, 112(1):90–103, October 1958. URL http://link.aps.org/doi/10.1103/PhysRev.112.90.

E. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1):269–271, 1959. ISSN 0029-599X. URL http://dx.doi.org/10.1007/BF01386390.

D. Donadio, P. Raiteri, and M. Parrinello. Topological defects and bulk melting of hexagonal ice. *The Journal of Physical Chemistry B*, 109(12):5421–5424, March 2005. ISSN 1520-6106. URL http://dx.doi.org/10.1021/jp050690z.

S. Dorozhkin. Calcium orthophosphates. *Journal of Materials Science*, 42(4):1061–1095, February 2007. URL http://dx.doi.org/10.1007/s10853-006-1467-8.

D. J. Earl and M. W. Deem. Parallel tempering: Theory, applications, and new perspectives. *Physical Chemistry Chemical Physics*, 7(23):3910–3916, 2005. ISSN 1463-9076. URL http://dx.doi.org/10.1039/B509983H.

P. P. Ewald. Die berechnung optischer und elektrostatischer gitterpotentiale. *Annals of Physics*, 369(3):253–287, 1921. ISSN 1521-3889. URL http://dx.doi.org/10.1002/andp.19213690304.

V. C. Farmer. *Infrared Spectra of Minerals*. London Mineralogical Society, 1974.

S. Fleming and A. Rohl. GDIS: a visualization program for molecular and periodic systems. *Zeitschrift für Kristallographie*, 220(5-6-2005):580–584, May 2005. ISSN 0044-2968. URL http://dx.doi.org/10.1524/zkri.220.5.580.65071.

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, August 2007. ISSN 1932-7447. URL http://dx.doi.org/10.1021/jp071887p.

C. L. Freeman, J. H. Harding, and D. M. Duffy. Simulations of calcite crystallization on self-assembled monolayers. *Langmuir*, 24(17):9607–9615, September 2008. ISSN 0743-7463. URL http://dx.doi.org/10.1021/la800389g.

C. L. Freeman, J. H. Harding, D. Quigley, and P. M. Rodger. Structural control of crystal nuclei by an eggshell protein. *Angewandte Chemie International Edition*, 49(30):5135–5137, 2010. ISSN 1521-3773. URL http://dx.doi.org/10.1002/anie.201000679.

D. Frenkel and B. Smit. *Understanding molecular simulation: from algorithms to applications*. Academic Press, 2002. URL http://books.google.co.uk/books?id=5qTzldS9ROIC.

J. D. Gale. GULP: A computer program for the symmetry-adapted simulation of solids. *Faraday Transactions*, 93(4):629–637, 1997. doi: 10.1039/a606455h. URL http://dx.doi.org/10.1039/a606455h.

J. D. Gale and A. L. Rohl. The General Utility Lattice Program (GULP). *Molecular Simulation*, 29(5):291–341, May 2003. ISSN 0892-7022. URL http://www.tandfonline.com/doi/abs/10.1080/0892702031000104887.

J. D. Gale, P. Raiteri, and A. C. T. van Duin. A reactive force field for aqueous-calcium carbonate systems. *Physical Chemistry Chemical Physics*, 13(37):16666–16679, 2011. ISSN 1463-9076. URL http://dx.doi.org/10.1039/C1CP21034C.

D. Gebauer, A. Völkel, and H. Cölfen. Stable prenucleation calcium carbonate clusters. *Science*, 322(5909):1819–1822, 2008. doi: 10.1126/science.1164271. URL http://www.sciencemag.org/content/322/5909/1819.abstract.

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. Edén, and N. Hedin. Proto-calcite and proto-vaterite in amorphous calcium carbonates. *Angewandte Chemie International Edition*, 49(47):8889–8891, 2010. ISSN 1521-3773. URL http://dx.doi.org/10.1002/anie.201003220.

S. Giralt, R. Julia, and J. Klerkx. Microbial biscuits of vaterite in lake Issyk-Kul (Republic of Kyrgyzstan). *The Journal of Sedimentary Research*, 71(3): 430–435, 2001. doi: 10.1306/2DC40951-0E47-11D7-8643000102C1865D. URL http://jsedres.sepmonline.org/cgi/content/abstract/71/3/430.

K. Goetzke and H.-J. Klein. Properties and efficient algorithmic determination of different classes of rings in finite and infinite polyhedral networks. *Journal of Non-Crystalline Solids*, 127(2):215–220, February 1991. ISSN 0022-3093. URL http://www.sciencedirect.com/science/article/pii/002230939190145V.

M. Grünwald. An efficient transition path sampling algorithm for nanoparticles under pressure. *The Journal of Chemical Physics*, 127(15):154718–, 2007. URL http://dx.doi.org/doi/10.1063/1.2790431.

M. Grünwald and C. Dellago. Transition state analysis of solid-solid transformations in nanocrystals. *The Journal of Chemical Physics*, 131(16):164116, 2009. URL http://dx.doi.org/10.1063/1.3253700.

R. Handel, R. L. Davidchack, J. Anwar, and A. Brukhno. Direct calculation of solid-liquid interfacial free energy for molecular systems: TIP4P ice-water interface. *Physical Review Letters*, 100(3):036104–, January 2008. URL http://link.aps.org/doi/10.1103/PhysRevLett.100.036104.

L. Hark and W. Cai. Ewald summation for coulomb interactions in a periodic supercell. Department of Mechanical Engineering, Stanford University, CA 94305-4040, January 2009.

S. Hauptmann, H. Dufner, J. Brickmann, S. M. Kast, and R. S. Berry. Potential

energy function for apatites. *Physical Chemistry Chemical Physics*, 5(3):635–639, 2003. URL http://dx.doi.org/10.1039/b208209h.

R. F. S. Hearmon. The elastic constants of anisotropic materials II. *Advances in Physics*, 5(19):323–382, 1956. URL http://www.informaworld.com/10.1080/00018735600101205.

S. B. Hendricks, M.E. Jefferson, and V.M. Mosley. The crystal structures of some natural and synthetic apaite-like substances. *Zeitschrift für Kristallographie*, 81: 352–369, 1932.

K. Henriksen, J. R. Young, P. R. Bown, and S. L. S. Stipp. Coccolith biomineralisation studied with atomic force microscopy. *Palaeontology*, 47(3):725–743, 2004. ISSN 1475-4983. URL http://dx.doi.org/10.1111/j.0031-0239.2004.00385.x.

L. W. Hobbs, C. E. Jesurum, V. Pulim, and B. Berger. Local topology of silica networks. *Philosophical Magazine A*, 78(3):679–711, 1998. URL http://www.informaworld.com/10.1080/01418619808241930.

W. G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A*, 31(3):1695–1697, March 1985. URL http://link.aps.org/doi/10.1103/PhysRevA.31.1695.

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. URL http://dx.doi.org/10.1063/1.1683075.

C. J. Horowitz, M. A. Pérez-García, D. K. Berry, and J. Piekarewicz. Dynamical response of the nuclear Pasta in neutron star crusts. *Physical Review C*, 72(3):035801–, September 2005. URL http://link.aps.org/doi/10.1103/PhysRevC.72.035801.

J. M. Hughes, M. Cameron, and K. D. Crowley. Structural variations in natural F, OH, and Cl apatites. *American Mineralogist*, 74(7-8):870–876, 1989. URL http://ammin.geoscienceworld.org/cgi/content/abstract/74/7-8/870.

K. W. Jeon, M. Friedlander, and J. Jarvik. *International Review of Cytology*. Elsevier Science, 1993. URL http://books.google.co.uk/books?id=wfSM70CuYqYC.

S. Kamhi. On the structure of vaterite $CaCO_3$. *Acta Crystallographica*, 16(8): 770–772, 1963. URL http://dx.doi.org/10.1107/S0365110X63002000.

S. Kerisit, D. J. Cooke, D. Spagnoli, and S. C. Parker. Molecular dynamics simulations of the interactions between water and inorganic solids. *Journal of Materials Chemistry*, 15(14):1454–1462, 2005. ISSN 0959-9428. URL http://dx.doi.org/10.1039/B415633C.

A. Laio and F. L Gervasio. Metadynamics: a method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Reports on Progress in Physics*, (12):126601, 2008. URL http://stacks.iop.org/0034-4885/71/126601.

A. Laio and M. Parrinello. Escaping free-energy minima. *Proceedings of the National Academy of Sciences*, 99(20):12562–12566, October 2002. URL http://www.pnas.org/content/99/20/12562.abstract.

W. T. Lee, M. T. Dove, and E. K. H. Salje. Surface relaxations in hydroxyapatite. *Journal of Physics: Condensed Matter*, (48):9829, 2000. URL http://stacks.iop.org/0953-8984/12/9829.

L. Liang, P. Rulis, and W.Y. Ching. Mechanical properties, electronic structure and bonding of α- and β-tricalcium phosphates with surface characterization. *Acta Biomaterialia*, 6(9):3763–3771, September 2010. ISSN 1742-7061. URL http://www.sciencedirect.com/science/article/pii/S1742706110001595.

D. R. Lide. *CRC handbook of chemistry and physics: a ready-reference book of chemical and physical data*. CRC Press, 2004. URL http://books.google.co.uk/books?id=WDll8hA006AC.

P. J. D. Lindan and M. J. Gillan. Shell-model molecular dynamics simulation of superionic conduction in CaF_2. *Journal of Physics: Condensed Matter*, 5(8):1019–, 1993. ISSN 0953-8984. URL http://stacks.iop.org/0953-8984/5/i=8/a=005.

B. Marie, G. Luquet, J.-P. Pais De Barros, N. Guichard, S. Morel, G. Alcaraz, L. Bollache, and F. Marin. The shell matrix of the freshwater mussel *Unio pictorum*

(Paleoheterodonta, Unionoida). *FEBS Journal*, 274(11):2933–2945, 2007. ISSN 1742-4658. URL http://dx.doi.org/10.1111/j.1742-4658.2007.05825.x.

E. N. Maslen, V. A. Streltsov, and N. R. Streltsova. X-ray study of the electron density in calcite, $CaCO_3$. *Acta Crystallographica Section B*, 49(4):636–641, 1993. doi: doi:10.1107/S0108768193002575. URL http://dx.doi.org/10.1107/S0108768193002575.

M. Mathew, L. W. Schroeder, B. Dickens, and W. E. Brown. The crystal structure of α-$Ca_3(PO_4)_2$. *Acta Crystallographica Section B*, 33(5):1325–1333, 1977. doi: doi:10.1107/S0567740877006037. URL http://dx.doi.org/10.1107/S0567740877006037.

S. K. Medeiros, E. L. Albuquerque, F. F. Maia Jr., E. W. S. Caetano, and V. N. Freire. First-principles calculations of structural, electronic, and optical absorption properties of CaCO3 vaterite. *Chemical Physics Letters*, 435(1-3):59–64, February 2007. ISSN 0009-2614. URL http://www.sciencedirect.com/science/article/pii/S0009261406018598.

C. Meis, J. D. Gale, L. Boyer, J. Carpena, and D. Gosset. Theoretical study of Pu and Cs incorporation in a mono-silicate neodymium fluoroapatite % $Ca_9Nd(Sio_4)(PO_4)_5F_2$. *The Journal of Physical Chemistry A*, 104(22):5380–5387, June 2000. ISSN 1089-5639. URL http://dx.doi.org/10.1021/jp000096j.

N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equations of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21:1087–1091, 1953.

R. A. Metzler, M. Abrecht, R. M. Olabisi, D.l Ariosa, C. J. Johnson, B. H. Frazer, S. N. Coppersmith, and P. U. P. A. Gilbert. Architecture of columnar nacre, and implications for its formation mechanism. *Physical Review Letters*, 98(26):268102–, June 2007. URL http://link.aps.org/doi/10.1103/PhysRevLett.98.268102.

H. J. Meyer. *Fortschritte der Mineralogie*, 38:187–188, 1960.

H. J. Meyer. Struktur und fehlordnung des vaterits. *Zeitschrift für Kristallographie*, 128(3-6):183–212, March 1969. ISSN 0044-2968. URL http://dx.doi.org/10.1524/zkri.1969.128.3-6.183.

D. Mkhonto and N. H. de Leeuw. A computer modelling study of the effect of water on the surface structure and morphology of fluorapatite: introducing a % $Ca_{10}(PO_4)_6F_2$ potential model. *The Journal of Materials Chemistry*, 12(9): 2633–2642, 2002. URL http://dx.doi.org/10.1039/b204111a.

M. Niederberger and H. Colfen. Oriented attachment and mesocrystals: Non-classical crystallization mechanisms based on nanoparticle assembly. *Physical Chemistry Chemical Physics*, 8(28):3271–3287, 2006. ISSN 1463-9076. URL http://dx.doi.org/10.1039/B604589H.

S. Nosé. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*, 81(1):511–, 1984. URL http://dx.doi.org/10.1063/1.447334.

J. J. Novoa, D. Braga, and L. Addadi. *Engineering of crystalline materials properties: state of the art in modeling, design and applications*. Springer, 2008. URL http://books.google.co.uk/books?id=xh4YYoKet00C.

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. ISSN 0342-1791. URL http://dx.doi.org/10.1007/BF00198605.

A. Pavese, M. Catti, S. C. Parker, and A. Wall. Modelling of the thermal dependence of structural and elastic properties of calcite, $CaCO_3$. *Physics and Chemistry of Minerals*, 23(2):89–93, 1996. ISSN 0342-1791. URL http://dx.doi.org/10.1007/BF00202303.

A. Pedone, M. Corno, B. Civalleri, G. Malavasi, M. C. Menziani, U. Segre, and P. Ugliengo. An ab initio parameterized interatomic force field for hydroxyapatite. *The Journal of Materials Chemistry*, 17(20):2061–2068, 2007. URL http://dx.doi.org/10.1039/b617858h.

B. Peters. Competing nucleation pathways in a mixture of oppositely charged colloids: Out-of-equilibrium nucleation revisited. *The Journal of Chemical Physics*, 131(24):244103, 2009. URL http://dx.doi.org/10.1063/1.3271024.

B. Peters, G. T. Beckham, and B. L. Trout. Extensions to the likelihood maximiza-

tion approach for finding reaction coordinates. *The Journal of Chemical Physics*, 127(3):034109, 2007. URL http://dx.doi.org/10.1063/1.2748396.

C. J. Pickard and R. J. Needs. Ab initio random structure searching. *Journal of Physics: Condensed Matter*, 23(5):053201–, 2011. ISSN 0953-8984. URL http://stacks.iop.org/0953-8984/23/i=5/a=053201.

F. Pietrucci and W. Andreoni. Graph theory meets ab initio molecular dynamics: Atomic structures and transformations at the nanoscale. *Physical Review Letters*, 107(8):085504–, August 2011. URL http://link.aps.org/doi/10.1103/PhysRevLett.107.085504.

Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1–19, 1995. ISSN 0021-9991. doi: 10.1006/jcph.1995.1039. URL http://www.sciencedirect.com/science/article/pii/S002199918571039X.

E. M. Pouget, P. H. H. Bomans, J. A. C. M. Goos, P. M. Frederik, G. de 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, March 2009. URL http://www.sciencemag.org/content/323/5920/1455.abstract.

W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, September 2007.

L. Qiao, Q.L. Feng, and Y. Liu. A novel bio-vaterite in freshwater pearls with high thermal stability and low dissolubility. *Materials Letters*, 62(12-13):1793–1796, April 2008. ISSN 0167-577X. URL http://www.sciencedirect.com/science/article/pii/S0167577X07010099.

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. URL http://dx.doi.org/10.1063/1.2940322.

D. Quigley and P. M. Rodger. Metadynamics simulations of ice nucleation and growth. *The Journal of Chemical Physics*, 128(15):154518, 2008b. URL http://dx.doi.org/doi/10.1063/1.2888999.

D. Quigley, P. M. Rodger, C. L. Freeman, J. H. Harding, and D. M. Duffy. Metadynamics simulations of calcite crystallization on self-assembled monolayers. *The Journal of Chemical Physics*, 131(9):094703, 2009. URL http://dx.doi.org/10.1063/1.3212092.

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. URL http://dx.doi.org/doi/10.1063/1.3530288.

E. Rabani. An interatomic pair potential for cadmium selenide. *The Journal of Chemical Physics*, 116(1):258–262, 2002. URL http://dx.doi.org/doi/10.1063/1.1424321.

J. A. L. Rabone and N. H. de Leeuw. Interatomic potential models for natural apatite crystals: Incorporating strontium and the lanthanides. *The Journal of Computational Chemistry*, 27(2):253–266, 2006. URL http://dx.doi.org/10.1002/jcc.20323.

P. Raiteri and J. D. Gale. Water is the key to nonclassical nucleation of amorphous calcium carbonate. *The Journal of the American Chemical Society*, 132(49):17623–17634, December 2010. ISSN 0002-7863. URL http://dx.doi.org/10.1021/ja108508k.

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, April 2010. ISSN 1932-7447. URL http://dx.doi.org/10.1021/jp910977a.

E. E. Santiso and B. L. Trout. A general set of order parameters for molecular crystals. *The Journal of Chemical Physics*, 134(6):064109, 2011. URL http://dx.doi.org/doi/10.1063/1.3548889.

M. C. Sha, Z. Li, and R. C. Bradt. Single-crystal elastic constants of fluorapatite, $Ca_5(PO_4)_3F$. *The Journal of Applied Physics*, 75(12):7784–7787, 1994. URL http://dx.doi.org/10.1063/1.357030.

R. Shetty, F. A. Escobedo, D. Choudhary, and P. Clancy. A novel algorithm for characterization of order in materials. *The Journal of Chemical Physics*, 117(8): 4000–4009, 2002. URL http://dx.doi.org/doi/10.1063/1.1494986.

W. Smith, Forester T. R., and Todorov I. T. *The DL_POLY Classic User Maual*. Daresbury Laboratory, United Kingdom.

P. J. Steinhardt, D. R. Nelson, and M. Ronchetti. Bond-orientational order in liquids and glasses. *Physical Review B*, 28(2):784–805, July 1983. URL http://link.aps.org/doi/10.1103/PhysRevB.28.784.

Y. Sugita and Y. Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, 314(1-2):141–151, November 1999. ISSN 0009-2614. URL http://www.sciencedirect.com/science/article/pii/S0009261499011239.

M. G. Taylor, S. F. Parker, K. Simkiss, and P. C. H. Mitchell. Bone mineral: evidence for hydroxy groups by inelastic neutron scattering. *Physical Chemistry Chemical Physics*, 3(8):1514–1517, 2001. URL http://dx.doi.org/10.1039/b005666i.

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, August 2009. ISSN 1520-6106. URL http://dx.doi.org/10.1021/jp902606x.

G. A. Tribello, M. Ceriotti, and M. Parrinello. A self-learning algorithm for biased molecular dynamics. *Proceedings of the National Academy of Sciences*, 107(41):17509–17514, 2010. doi: 10.1073/pnas.1011511107. URL http://www.pnas.org/content/107/41/17509.abstract.

E W., W. Ren, and E. Vanden-Eijnden. Finite temperature string method for the study of rare events. *The Journal of Physical Chemistry B*, 109(14):6688–6693, February 2005. ISSN 1520-6106. URL http://dx.doi.org/10.1021/jp0455430.

D. J. Wales and J. P. K. Doye. Global optimization by basin-hopping and the lowest energy structures of lennard-jones clusters containing up to 110 atoms. *The Journal of Physical Chemistry A*, 101(28):5111–5116, July 1997. ISSN 1089-5639. URL http://dx.doi.org/10.1021/jp970984n.

J. Wang and U. Becker. Structure and carbonate orientation of vaterite ($CaCO_3$). *American Mineralogist*, 94(2-3):380–386, 2009. doi: 10.2138/am.2009.2939. URL http://ammin.geoscienceworld.org/cgi/content/abstract/94/2-3/380.

S. E. Wolf, J. Leiterer, M. Kappl, F. Emmerling, and W. Tremel. Early homogenous amorphous precursor stages of calcium carbonate and subsequent crystal growth in levitated droplets. *The Journal of the American Chemical Society*, 130(37):12342–12347, September 2008. ISSN 0002-7863. URL http://dx.doi.org/10.1021/ja800984y.

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. URL http://dx.doi.org/10.1063/1.2136877.

G. Wulff. Zur frage der geschwindigkeit des wachstums und der auflösung der kristallflächen. *Zeitschrift für Kristallographie und Mineralogie*, 34:449, 1901.

M. Yashima, A. Sakai, T. Kamiyama, and A. Hoshikawa. Crystal structure analysis of [beta]-tricalcium phosphate $Ca_3(PO_4)_2$ by neutron powder diffraction. *Journal of Solid State Chemistry*, 175(2):272–277, November 2003. ISSN 0022-4596. URL http://dx.doi.org/10.1016/S0022-4596(03)00279-2.

X. Yin, M. J. Stott, and A. Rubio. Alpha and beta tricalcium phosphate: A density functional study. *Physical Review B*, 68(20):205205–, November 2003. URL http://link.aps.org/doi/10.1103/PhysRevB.68.205205.

X. Yuan and A. N. Cormack. Efficient algorithm for primitive ring statistics in topological networks. *Computational Materials Science*, 24(3):343–360, June 2002. ISSN 0927-0256. URL http://dx.doi.org/10.1016/S0927-0256(01)00256-7.