Trace element incorporation in stalagmite calcite as a paleo-volcanic record: a chemical perspective

Scott D. Midgley

Submitted for the degree of
Doctor of Philosophy

Department of Chemistry
School of Chemistry, Food and Pharmacy

July 2021

Supervisors:
Dr Ricardo Grau-Crespo
Prof. Dominik Fleitmann

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Scott D. Midgley

"*It would not be much of a universe if it wasn't home to the people you love*."

- Stephen Hawking

To G. S. and F. R. Day.

- The brightest stars in my night sky.

Acknowledgements

Firstly, I must thank my supervisor Dr R. Grau-Crespo. I extremely grateful for his support, mentorship, and friendship over the past four years, and for helping me to make the most of such an incredible opportunity in my life. He is a truly brilliant scientist and it has been a privilege to be a part of his research group. I am grateful to Prof D. Fleitmann for his role as my second supervisor, and for his input in directing my research. I am grateful to the Scenario NERC DTP for my studentship, and to the materials chemistry consortium of the ARCHER UK supercomputing facility for much of the computational resources required for the present work.

Looking back; thanks must go to the science teachers of LPGS (2011-2013), specifically R. Small and J. Fisher, for their recognition and encouragement of my scientific potential. Thanks of this kind are also due to the staff of the University of Reading Chemistry department 2013-2016, for making my undergraduate studies so enjoyable and successful.

I have had the privilege of working around some excellent people at the University of Reading during the past four years. Thanks are due to Ms. W. Neale, Prof K. Shankland, Dr L. Bartók-Pártay, and Prof A. Chippindale, for a variety of reasons. I have had the privilege of being mentored in high-performance computing by Dr J. Saßmannshausen during my PhD, for which I am very grateful. I have also been extremely lucky to have worked with excellent collaborators both locally and internationally, including: Prof H. Colquhoun, Dr K. Butler, Dr D. Di Tommaso, and Dr S. Hamad, to whom I express many thanks.

To Dr M. Spillman, for your patience, mentorship, and for starting me on this journey, I am eternally grateful. To Dr R. Crespo-Otero, thank you for an excellent MSc year and for helping me to develop as a young researcher.

To my friends in the department: Dr J. Ontaneda-Rojas, Dr V. Posligua, Miss S. Grover, Miss E. Gousseva and Mr J. Seymour, thanks are due for various reasons.

To my friends of old: G. Grainger, T. Banerjee, T. Hill, and S. Tabor, thank you for the times we have shared, and the encouragement you have given me over the years.

To my *day-one* friends at UoR: H. Brackley, J. Churchward, A. Fee, J. Devonport, J. Baylis, and B. Harrison, I am forever grateful to have shared with you such a memorable time in our lives.

To Gary, Jax, Paul, Nicky, Grandma, and Grandpa, thank you for your encouragement over the years.

Finally, to those whom I owe the most.

To my mum, dad, and sister, for your love, belief, and providing me with the amazing upbringing and support that has led me to this point, thank you.

To Georgina, during these past years I could have asked for no more from you. Your unwavering patience, support, love, and inspiration has made this all possible. I will be grateful to you forever.

Abstract

This thesis examines the incorporation of oxyanion impurities (sulphate, molybdate and bromate) in calcium carbonate minerals, using atomic-level computer simulations. The motivation is to understand the chemical nature and environment of these impurities in speleothems, as they constitute a record of past volcanic activity that can be used in past climate reconstruction models.

Sulphur in its tetrahedral oxyanion form, sulphate (SO_4^{2-}), shows low thermodynamic tendency to incorporate in bulk crystalline regions of calcite, which is the calcium carbonate phase forming most speleothems. Incorporation thermodynamics in other calcium carbonate phases is highly dependent on the phase density. Incorporation stability in the anhydrous phases follows the order: vaterite > calcite > aragonite. Molybdenum, in its tetrahedral oxyanion form, molybdate (MoO_4^{2-}), shows similar incorporation behaviour to sulphate, though it requires around 40-50% more energy to incorporate across the naturally occurring calcium carbonate host phases. The incorporation in hydrated calcium carbonate phases is also discussed and compared with the behaviour in anhydrous phases.

Given that bulk substitution is thermodynamically unstable, the incorporation of sulphate and molybdate at the calcite/water interface was investigated, using ab initio molecular dynamics. The incorporation at both the calcite (10.4) terraces and at step line defects with the same surface termination was examined. Oxyanion/carbonate exchange energies calculated at the calcite/water interface are far more favourable than those calculated for the bulk, indicating a clear tendency for sulphate and molybdate to accumulate at the surface, and even more strongly at the defect regions. The specific surface area or crystallinity can therefore be expected to play a large role in the uptake of sulphur and molybdenum in speleothem, which should be considered when interpreting experimentally trace element concentrations for paleovolcanic reconstruction models.

Bromine in its trigonal pyramidal oxyanion form, bromate (BrO_3^-), was also studied as a possible anionic trace element in calcium carbonates (focusing on calcite and aragonite in this case). Bromate provides a geochemically stable redox state for bromine whilst inducing lower elastic strain effects on the host crystal compared with sulphate and molybdate, owing to one less apical oxygen atom. The charge imbalance introduced when substituting a singly charged bromate ion for a doubly charged carbonate ion was compensated by the co-incorporation of a singly charged cation. Lithium, sodium, and potassium ions were investigated for their effect on the thermodynamics of substitution in the crystalline bulk of calcite and aragonite. It was

demonstrated that at ambient temperature the binding between the oppositely charged impurity dominated over the configurational entropy tendency to separate the defects. Therefore, the impurities are likely to be found in nearest neighbour configurations in both calcite and aragonite, which could have implications for their detection and measurement. Although bromate is also metastable in calcite in aragonite with respect to phase separation, the thermodynamic driving force for that is much lower than for sulphate and molybdate. Bulk crystalline incorporation of bromate would therefore be expected to be far higher than sulphate and molybdate. This higher stability could be related to the reported superior reliability of bromine in speleothems as a record of past volcanic activity.

This theoretical work offers a first approximation to the chemistry of incorporation of molybdenum and bromine in calcium carbonate minerals, providing clear thermodynamic and structural data that will help geochemists to interpret their speleothem records, and hopefully motivate further research work.

Publications related to this thesis

1. **Molybdenum and sulphur incorporation as oxyanion substitutional impurities in calcium carbonate minerals: A computational investigation**

Authors: Scott D. Midgley, James O. Taylor, Dominik Fleitmann, Ricardo Grau-Crespo

Status: Published in *Chemical Geology*

(Midgley et al., 2020)

2. **Sulphur and Molybdenum Incorporation at the Calcite-Water Interface: Insights from Ab Initio Molecular Dynamics**

Authors: Scott D. Midgley, Devis Di Tommaso, Dominik Fleitmann, Ricardo Grau-Crespo

Status: Published in *ACS Earth and Space Chemistry*.

(Midgley et al., 2021a)

3. **Bromate incorporation in calcite and aragonite**

Authors: Scott D. Midgley, Dominik Fleitmann, Ricardo Grau-Crespo

Status: Submitted to *Geochimica et Cosmochimica Acta*, also available in ChemRxiv.

(Midgley et al., 2021b)

Other publications by the author

During the PhD I have had the opportunity to contribute to other research projects, using the core skills that were honed through my main topic of research on carbonate minerals. These *side* projects have led to publications as follows:

1. **Supramolecular complexation between chain-folding poly(ester-imide)s and polycyclic aromatics: a fractal-based pattern of NMR ring-current shielding**

Authors: Marcus Knappert, Tianqi Jin, Scott D. Midgley, Guanglu Wu, Oren A. Scherman, Ricardo Grau-Crespo and Howard M. Colquhoun

Status: Published in *RSC Polymer Chemistry*

(Knappert et al., 2019)

2. **Single-site binding of pyrene to poly(ester-imide)s incorporating long spacer-units: prediction of NMR resonance-patterns from a fractal model**

Authors: Marcus Knappert, Tianqi Jin, Scott D. Midgley, Guanglu Wu, Oren A. Scherman, Ricardo Grau-Crespo and Howard M. Colquhoun

Status: Published in *RSC Chemical Science*

(Knappert et al., 2020)

3. **Novel WS2-Based Nanofluids for Concentrating Solar Power: Performance Characterization and Molecular-Level Insights**

Authors: Paloma Martínez-Merino, Scott D. Midgley, Elisa I. Martín, Patrice Estellé, Rodrigo Alcántara, Antonio Sánchez-Coronilla, Ricardo Grau-Crespo, and Javier Navas

Status: Published in *ACS Applied Materials & Interfaces*

(Martinez-Merino et al., 2020)

4. **Understanding the Specific Heat Enhancement in Metal-Containing Nanofluids for Thermal Energy Storage: Experimental and Ab Initio Evidence for a Strong Interfacial Layering Effect**

Authors: Iván Carrillo-Berdugo, Scott D. Midgley, Ricardo Grau-Crespo, David Zorrilla, and Javier Navas

Status: Published in *ACS Applied Energy Materials*

(Carrillo-Berdugo et al., 2020)

5. **Bandgap engineering in the configurational space of solid solutions via machine learning:(Mg, Zn) O case study**

Authors: Scott D. Midgley, Said Hamad, Keith T. Butler, Ricardo Grau-Crespo

Status: Published in *The Journal of Physical Chemistry Letters*

(Midgley et al., 2021c)

Table of Contents

Frequently used abbreviations

AIMD	*Ab Initio* Molecular Dynamics
BOA	Born-Oppenheimer Approximation
BOMD	Born-Oppenheimer Molecular Dynamics
CG	Conjugate Gradients (ionic optimisation method)
CP2K	Open source molecular dynamics package
CPU	Central Processing Unit
D3	Grimme D3 dispersion corrections (DFT-D3)
DFT	Density Functional Theory
DIIS	Direct Inversion in the Iterative Subspace
DoF	Degree(s) of Freedom
DZ	Double-Zeta (basis set)
GGA	Generalised Gradient Approximation
GTF	Gaussian Type Function
GTO	Gaussian Type Orbital
HF	Hartree-Fock
KS	Kohn-Sham
LDA	Local Density Approximation
MD	Molecular Dynamics
MO	Molecular Orbital
MOLOPT	CP2K molecule optimised basis sets
OT	Orbital Transformation
PAW	Projector Augmented Wave
PBE	Perdew–Burke-Ernzerhof (a GGA functional)

PES	Potential Energy Surface
QS	Quickstep (fast DFT implementation in CP2K)
RAM	Random Access Memory
RDF	Radial Distribution Function
SCF	Self-Consistent Field
SD	Steepest Descent (ionic optimisation method)
SOD	Site Occupancy Disorder (code)
VASP	Vienna Ab initio Simulation Package
x_m	Stable Molar Fraction or Solubility Limit (terms used interchangeably)
XC	Exchange-Correlation (quantum mechanical forces)

1. Introduction

1.1. Background

Information on Earth's climate spanning millennia is essential for understanding and addressing modern climate change, which is considered one of the greatest challenges faced by our society. Extreme changes in climate are known to have occurred naturally throughout Earth's history, but there is significant evidence indicating that the sharp increase in global temperatures following the industrial revolution is unprecedented. Identifying historical extremes in Earth's climate is therefore crucial to understanding and benchmarking modern climate change. To illustrate this scientific necessity by way of example, the Last Glacial Period (ice age) which is thought to have ended around 11,700 years BP (before present) saw global temperatures at an extreme low at around 24,000 years BP (Rutter et al., 2012). Though the planet has since recovered from this period of extreme cool via natural warming, there is significant evidence that the rate of modern climate change is far higher than ever before. It is therefore vital to gather maximum information on the rate of natural climactic oscillations, for benchmarking of the modern anthropogenic effect.

Modern scientific measurements of Earth climatic conditions span from around the turn of the 20th century to present day, meaning there is no well-documented record of Earth's climate over most of its history. Paleoclimatology (the study of Earth's climate history across geological timescales) involves examination of indicators (proxies) that are found in geological, physical and biological environments (archives). This field aims to provide detailed climate history by developing precisely-dated and highly-resolved records of temperature and precipitation (the most relevant ones among others) that span millennia. Examples of commonly exploited paleoclimate archives include ice cores (Faria et al., 2010), tree-rings (McCarroll and Loader, 2004), lake and ocean sediments (Lowe, 2011; Singer, 1984; Voelker, 2002), corals (Ruiz-Hernandez et al., 2010) and more recently, speleothems (stalagmites and stalactites) (Badertscher et al., 2014; Fairchild and Treble, 2009; Finch et al., 2001; Fleitmann et al., 2009; Göktürk et al., 2011; Ünal-İmer et al., 2015). Paleoclimate research based on these archives has produced Earth temperature time series such as the one shown in Figure 1.1. The figure shows global temperatures from 500-2000 C.E. with the dip around the 1500 C.E mark known informally as the *Little Ice Age*. The green section represents temperatures that have been reconstructed using proxies, while the blue section is derived from instrumental scientific measurements. This section shows a significant upward trend around the time that the industrial revolution began, which is the basis for modern theories of global warming by human-influence. Though this kind of climate history represents great scientific success, the field remains under active research interest with the aim of extending the paleoclimate models

further back in Earth's history, reducing current uncertainties, and increasing resolution of existing information.

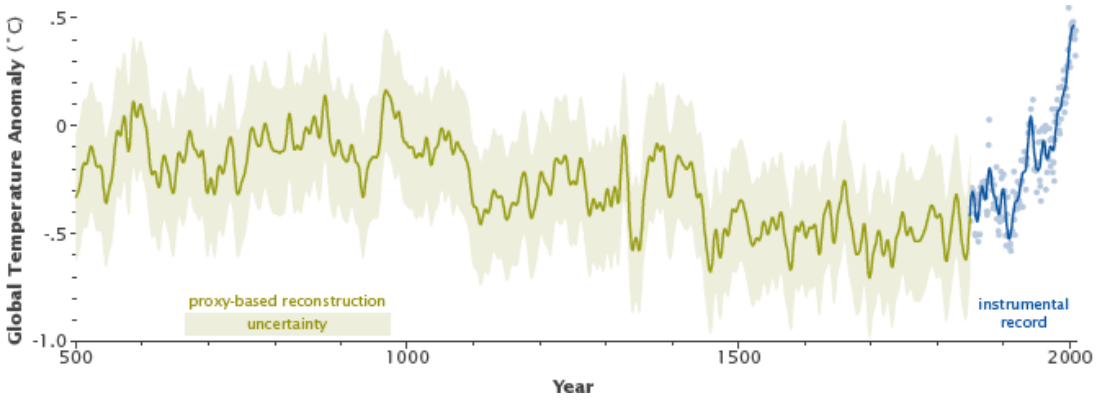

Figure 1.1. Earth temperatures from proxy-based climate reconstruction from years 500-2000 C.E. Representation from (NASA), using data from (Mann et al., 2008).

Ice cores are the *de facto* archive in paleoclimate research due to their natural longevity (Severi et al., 2012; Yalcin et al., 2007; Zielinski et al., 1994). Ice cores are usually taken from polar ice which has often been unchanged for several thousands of years. The ice layer 'grows' in an upward direction with newly formed layers of ice, encapsulating small indicators or tracers of the ancient climatic conditions. Ice cores provide essential information in deducing environmental conditions during major climatic periods such as the Last Glacial Period, by preserving proxies in naturally undisturbed states. The term *proxy* is used for any chemical, biological, or physical indicator of some historical environmental anomaly. Detection of the proxy in a section of ice core that formed millennia in the past is therefore an indicator of environmental condition(s) at the time that the proxy incorporated into the structure. In ice cores, viable proxies range from macroscopically visible particulates to trace ion impurities which are detectable only through chemical analysis. Annually layered ice cores therefore present time-resolved indicators of the local environment spanning geological timescales, which are essential in deducing global environmental conditions, deep into the past.

Earth's climate is governed by multiple external and internal climate factors which range from the core (generating the magnetosphere), Earth's solar orbital parameters, to the upper atmosphere (ozone layer). Here, even very subtle variations can have a profound impact on Earth's interaction with solar radiation, which is a major contributor to temperatures and climates. Volcanic eruptions (external forcing) can have a powerful impact on both the local and global climate. The cooling effect that follows large eruptions can, amplified by internal feedbacks of the Earth's system, in some cases last for decades after the actual eruption

(Robock, 2000; Sigl et al., 2013; Zielinski, 2000). The mechanism for this cooling effect is thought to be based on volcanic particulates (tephra) and sulphur dioxide entering the upper atmosphere following an eruption, which disrupt solar radiation transfer leading to increased reflection. Documented examples of this phenomenon include the year 1816 in Europe became known as the *year without a summer*, because of the so-called *volcanic winter* caused by eruption of Mount Tambora in 1815. Another, more recent example of a volcanic winter involves the eruption of Mt. Pinatubo in 1991 (Stenchikov et al., 1998), when global air temperature dropped by approx. 0.5 °C for 1-2 years.

The history of volcanic events is an integral input parameter for accurate proxy-based climate reconstruction models. Volcanic sulphur-containing and (occasionally) tephra particles accumulate in polar ice sheets and are detectable during analysis of geological structures, including ice cores. However, the limited geographical distribution of ice cores means it very difficult to decipher eruption size and location, prompting a need for alternate archives to assist interpretation (Fairchild and Baker, 2012).

Speleothem-based records of paleovolcanic activity have seen increased research interest in recent years because of their ability to capture chemical impurities. Speleothem-based records such as trace element concentrations (Hartland et al., 2012), oxygen and carbon isotopes of calcite and hydrogen and oxygen isotopes of fluid inclusion water (Affolter et al., 2019; Fleitmann et al., 2009; Marshall et al., 2009) provide valuable information on decadal-to-millennial-scale changes in local precipitation and temperature. Proxies can be identified in speleothems using well-established isotopic dating techniques, to give an indication of paleovolcanism in some past timeframe (Dorale et al., 2004; St Pierre et al., 2009). Though speleothems records are not as well-established as the use of ice cores, they have the advantage of forming readily across a range of climate conditions, making it easier to deduce the specific location of paleovolcanic events, and record regional volcanic activity.

There are several examples in the literature of speleothem-based evidence for paleovolcanism. One notable report by Badertscher and co-workers (Badertscher et al., 2014) matched elevated levels of sulphate, bromate and molybdate found in speleothems formed in the Sofular cave (northern Turkey) to the known Minoan eruption of Santorini at around 1627 BCE (Friedrich et al., 2006). Other studies by Fleitmann (Fleitmann et al., 2009) and Bar-Matthews (Bar-Matthews et al., 1999) also report speleothem-based evidence for this eruption. Some recent work by Theodorakopoulou (Theodorakopoulou et al., 2020) reported the first speleothem-based evidence for a different eruption of Santorini. The eruption known as Hiera, which is

thought to have occurred around 197 BCE according to speleothem records, is well-matched to eruption records from ancient human civilisations in the region.

Chemical anomalies in speleothems therefore provide a previously untapped database of paleovolcanic history. The variability of paleovolcanic chemical indicators is not yet known and is under active investigation across a range of geoscientific disciplines. Research in the field commonly focusses on experimental detection and empirical rationalisation of observed correlations based on geophysical theories of speleothem formation and behaviour. To fully understand observed correlations however, knowledge of atomic-level structural and thermodynamic aspects of chemical anomalies is likely to be important, though mechanistic chemical analysis is rarely reported in this context.

1.2. Calcium carbonates

Calcium carbonate ($CaCO_3$) is a ubiquitous natural mineral and is commonly found as limestone rocks or as exoskeleton constituents in marine organisms. Calcium carbonate forms several structurally distinct crystallographic phases known as *polymorphs*. The polymorphic crystalline phase formed is dictated by the thermodynamic pathway of crystallisation, which depends on a variety of physical factors such as temperature, pressure, and the presence of other chemical species in the local environment.

Calcite, aragonite and vaterite are the three naturally occurring phases of calcium carbonate with the chemical formula, $CaCO_3$. The solid-state chemistry of these three phases is relatively well documented; however geochemical factors affecting the thermodynamics of formation in natural environments is not yet clearly defined in the literature, due to its extreme depth and complexity. Phases of calcium carbonate containing structural water are also possible. Monohydrocalcite ($CaCO_3 \cdot H_2O$) (MHC) and ikaite ($CaCO_3 \cdot 6H_2O$) are known chemically, though are much rarer in geological contexts. Figure 1.2 illustrates the crystal structures of the five calcium carbonate host phases discussed in this section. These are further introduced, in terms of their physical properties, in Chapter 3. Each structure shown is shown as a *supercell* (section 2.1), as used in the work reported in Chapter 3.

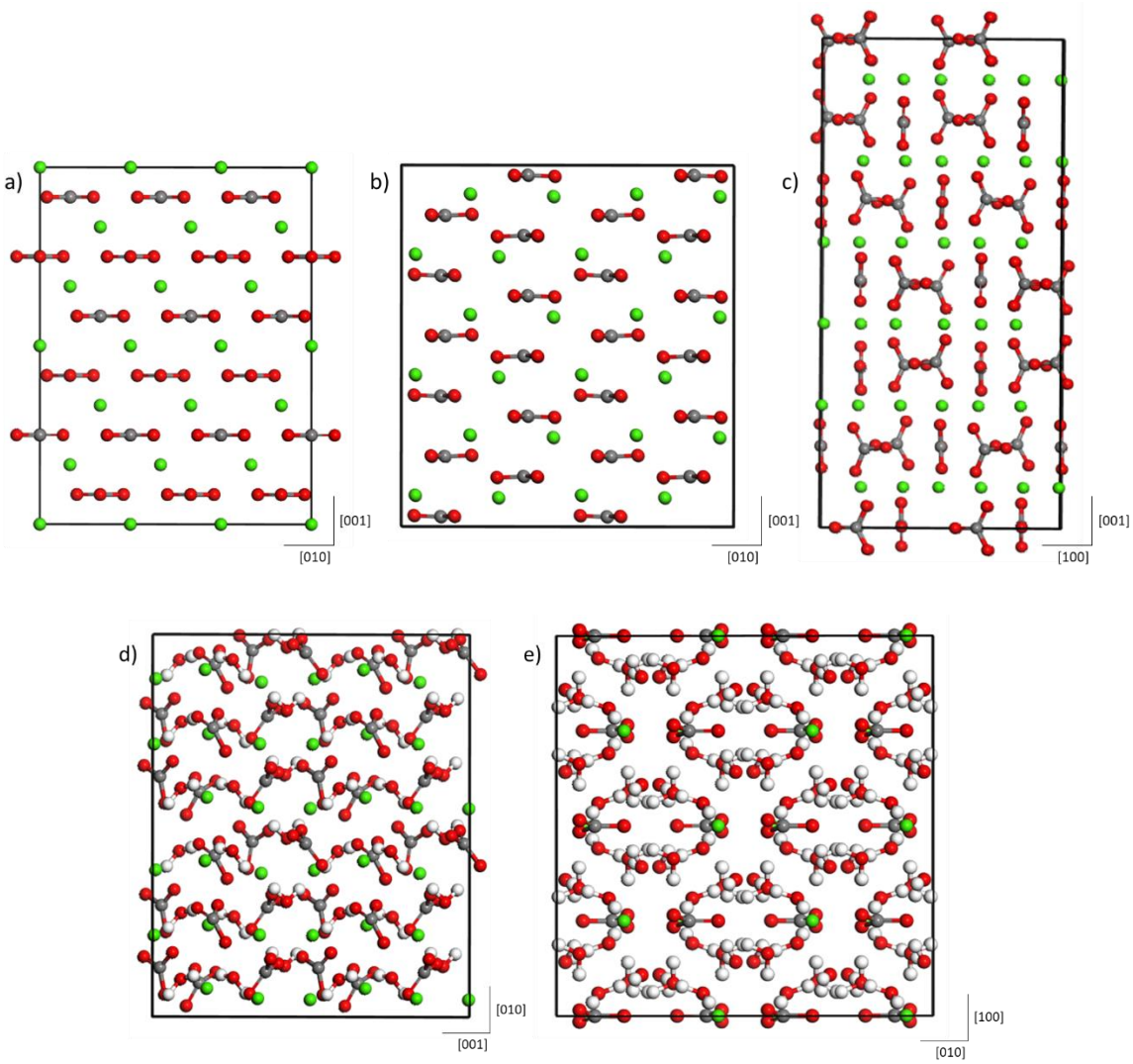

Figure 1.2. a) Calcite 3×3×1 (Markgraf and Reeder, 1985); b) Aragonite 3×2×3 (De Villiers, 1971); c) Vaterite 2×2×1 (Wang and Becker, 2009); d) MHC 2×2×2 (Swainson, 2008); e) Ikaite 2×2×2 (Hesse et al., 1983).

Calcite is the most stable naturally occurring phase of calcium carbonate and comprises the majority of speleothem carbonate material reported in the scientific literature. It is highly abundant in nature and provides the most viable carbonate-based volcanic archive material (Fairchild and Baker, 2012). Calcite is the predominant host phase featured in the present work and is discussed at length following a brief introduction to the four other naturally occurring calcium carbonate phases.

Following calcite, aragonite is a dense, highly metastable phase of calcium carbonate. Aragonite forms under conditions such as an Mg/Ca ratio that is too high for calcite formation but too low for formation of a separate phase of $MgCO_3$. Aragonite may be a suitable paleoclimate archive due to its long-term metastability in the absence of meteoric

precipitation; however its low natural abundance on land means that it is less commonly used as a paleoclimate archive when compared to calcite. Aragonite reported in on-land geological context may be found in rare cases (De Choudens-Sanchez and Gonzalez, 2009; Lachniet, 2015); however other natural occurrences such as in marine animals (e.g., corals) are far more common (Addadi et al., 2003; Railsback et al., 1994).

Vaterite is the least stable phase of anhydrous calcium carbonate and rapidly recrystallises to calcite in the presence of water; however it is known to be stabilised by the presence of some ions such as sulphate. Large quantities of metastable vaterite are often found in geothermal springs, where volcanic sulphur yields a high concentration of sulphate in the local aqueous environment (Grasby et al., 2003). Vaterite is the least dense of the anhydrous $CaCO_3$ crystal phases, which allows for a relatively high degree of structural disorder. This means that though the crystal system (hexagonal) and lattice vectors are known, there is ambiguity in reported ionic structure, as reviewed by (Christy, 2017). Vaterite does not form single crystals that are large enough for clear X-ray diffraction, therefore powder X-ray diffraction and nuclear magnetic resonance (NMR) spectroscopy techniques are needed for structural characterisation. Vaterite is unlikely to be found in speleothems due to its instability across geological timescales, meaning that it has likely recrystallised to calcite or aragonite long before practical use as a paleovolcanic archive (Boulos et al., 2014; Fernández-Díaz et al., 2010; Lacelle et al., 2009).

There exists a complex interplay of chemical factors governing the phase transitions between calcium carbonate phases, as described in various geochemical literature sources (Ogino et al., 1987; Plummer and Busenberg, 1982; Silk and Lewin, 1970). A notable study by Tang (Tang et al., 2012) combined experimental and theoretical techniques to conclude that interaction between the calcite (10.4) surface and the sulphate anion had an inhibitive effect on the calcite growth mechanism, favouring formation of aragonite over calcite. Another study by Boulos (Boulos et al., 2014) describes in detail how temperature factors may strongly affect the recrystallization of vaterite to calcite or aragonite. A computer simulation study by Fernández-Díaz (Fernández-Díaz et al., 2010) details how vaterite is known to recrystallize to a more stable phase in ca. 24 hours, further confirming that vaterite detection in speleothem contexts is unlikely.

The scientific literature does not describe complete geochemical mechanisms for the behaviour of calcium carbonate speleothems; however the presence of trace elements such as

sulphur in the local aqueous environment can clearly influence speleothem carbonate crystal precipitation (Meyer, 1984).

Though calcite, aragonite and vaterite phases dominate calcium carbonate geochemistry, hydrous phases are known despite being extremely rare. Ikaite, the hexahydrate form of calcium carbonate is well known in geochemical research; however it was not reported in the context of speleothems until 2017. Ikaite forms in cold environments below 10° C (although the exact temperature is disputed in the literature) and rapidly recrystallises to other phases when removed from cold water surroundings (Purgstaller et al., 2017). Ikaite is often stabilised in nature by sea ice, where small mineral phases are known to form. To date, ikaite has never been identified in speleothems; however ikaite *pseudomorphs* were reported in speleothems by researchers in 2017 (Field et al., 2018). *Pseudomorphism* is a crystallographic phenomenon that is frequently reported in the context speleothem geochemistry. Pseudomorphs form when ions in a mineral are replaced by substitution processes, but the original crystal system and macroscopic shape is retained (c.f. recrystallization where the mineral is completely reformed) (Jansen et al., 1987). The significance of this was highlighted in 2017, where pseudomorphism was reported in UK-based speleothems (Field et al., 2018). Though the authors reported that the presence of calcite via X-ray diffraction analysis, there were clear pseudomorph indicators that ikaite had once been present. This was the first evidence of ikaite in speleothem carbonates during the early stages of formation.

Monohydrocalcite (MHC) is less common than ikaite in the natural world; however it has been reported in the context of speleothems in some rare cases (Fischbeck and Müller, 1971; Stoffers and Fischbeck, 1974). Its rarity may again be credited to the fact that precipitation of this phase is governed by extremely precise Mg/Ca ratios in the aqueous phase (although other factors may have an important but less substantial effect) (Fischbeck and Müller, 1971).

Calcite is undisputed as the primary calcium carbonate phase in speleothems, but to understand these archives at an atomic scale, it is prudent to consider the chemistry throughout the lifetime at the speleothem, rather than the phase present at the time of analysis (calcite). This includes consideration of kinetically stabilised phases that may have recrystallised to the stable thermodynamic phase, calcite, during the early stages of the speleothem formation. Other chemical factors may also be important to consideration, such as complementary trace-species which are known to encourage trace-element binding by positive feedback.

Calcite forms a rhombohedral particle that grows in predominantly in the crystallographic direction perpendicular to the (10.4) plane. The crystal may be represented using either a hexagonal or rhombohedral structure, though the hexagonal form is used throughout the present work. Chemical and physical factors influencing calcite growth are broad, and are reported throughout the literature, with some of the key points discussed briefly in the following paragraphs.

Calcite growth is strongly affected by magnesium in the aqueous environment. At very high magnesium concentrations, $CaCO_3$ and $MgCO_3$ form a solid solution, where the relative stability of the phase-separated end members is determined by the solubility products, i.e. the relative stability of each phase with respect to dissolution. Lower concentrations of magnesium have been shown to promote aragonite formation (Murray, 1954), inhibiting calcite growth by increasing the solubility of the magnesium-containing calcite (Davis et al., 2000; Mather et al., 2004; Wasylenki et al., 2005a). Many impurities (both organic and inorganic) are known to stunt calcite growth by incorporating at the (10.4) growth plane (Meyer, 1984; Teng et al., 1999; Wasylenki et al., 2005b).

The ability of calcite (and other $CaCO_3$ phases) to incorporate trace impurities (less than 1% concentration) is strongly affected by atomic-scale structural and thermodynamic factors, which are often underexplored in the context of speleothems. Impurities can incorporate in either the crystalline bulk structure of the carbonate, or at surface regions. Bulk incorporation commonly involves divalent cations (Mg^{2+}, Co^{2+}, Sr^{2+}) because they can easily substitute calcium ions in terms of both shape and charge. Anionic bulk incorporation is possible, though it has received less research attention that cation uptake, presumably due to fewer known commercial applications, such as radioactive element deactivation by carbonate incorporation. Examples of computational investigation (similar methods to the present thesis) into both anion and cation substitution in calcium carbonates are reported in the literature (Fernández-Díaz et al., 2010; González-López et al., 2014).

Incorporation of impurities at calcite surfaces has also been investigated extensively using computational methods, with a well-cited example by de Leeuw (de Leeuw, 2002). Trace element affinity for calcite is highly dependent on particle morphology (Didymus et al., 1993). The calcite particle is rhombohedral at equilibrium with the major facets comprised of (10.4) termination (Figure 1.3a). Changes in calcite morphology can affect the trace-element concentrations absorbed by the particle, by changing the ratio of trace-element-accommodating facets. Different trace-element affinity for different facets causes increased

substitution in specific regions of the particle – a phenomenon known as *sector zoning* (Reeder and Grams, 1987).

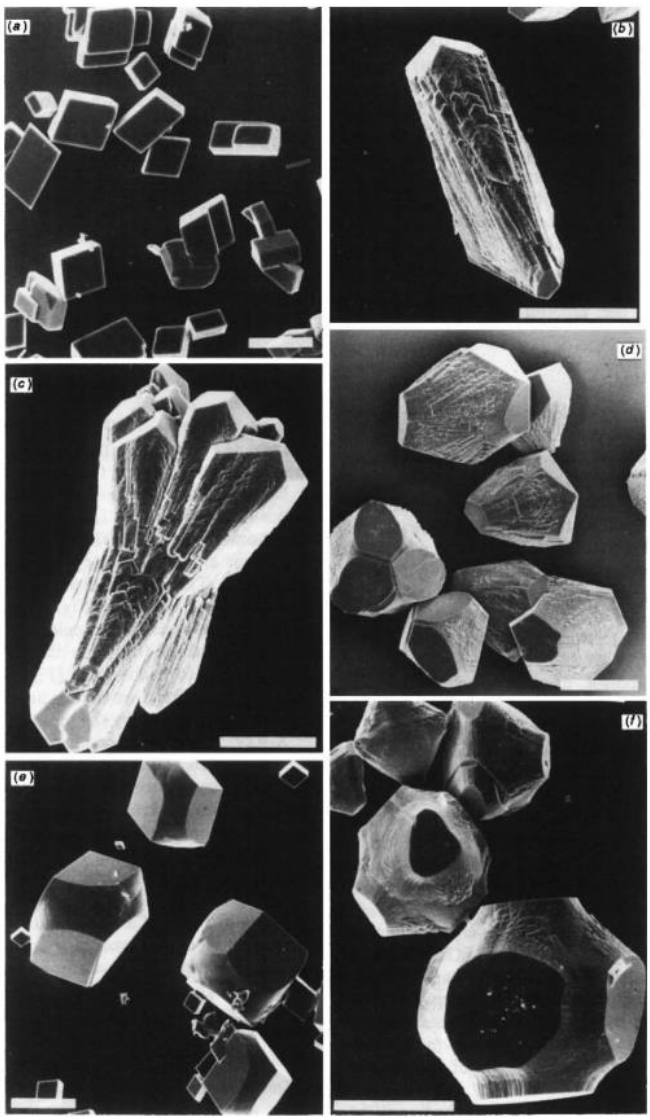

Figure 1.3. Scanning electron micrograph images of calcite particle: **a.** Equilibrium rhombohedral particle, **b-f.** Calcite morphology under a variety of chemical conditions. Figure from (Didymus et al., 1993).

Imperfections on the (10.4) terrace are commonly referred to as *steps* and *kinks* (step shown in Figure 1.4.). These sites are often at the edge of *pits* formed on the calcite terrace as it grows (Liang and Baer, 1997; Liang et al., 1996a). There is some evidence that surface imperfections provide favourable conditions to accommodate impurities into the calcium carbonate structure (Paquette and Reeder, 1995).

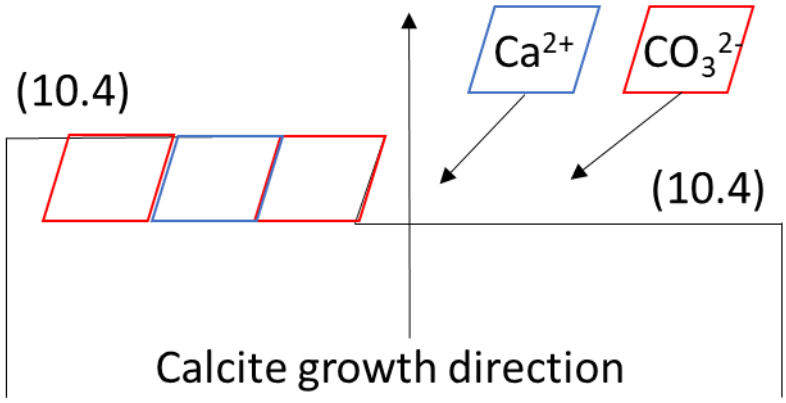

Figure 1.4. Simplified diagram showing calcite growth direction. Aqueous ions bind to an *acute* step in the (10.4) surface, propagating growth in the upward direction.

Sulphate sector zoning in calcite was reported by Staudt (Staudt et al., 1994), where it was concluded that sulphate incorporates in calcite preferentially where there are surface steps. Conclusions given in this study and subsequent citing articles appear to be consistent with the work by Wang and Xu (Wang and Xu, 2001), where it was suggested that trace element partitioning in calcium carbonates is controlled by three main physical factors:

i. Lattice energy (formation) caused by the substitution of impurity.
ii. Lattice strain caused by geometric disparities in the substituting ions.
iii. Change in *chemical potential* due to the substitution.

Favourable trace element incorporation at surface imperfections may logically be attenuated to one or more of these. Though at the time the present research project begun, no atomic-level study had been conducted to provide evidence for the link between these physical factors and the observed trace element zoning in calcite.

Sulphate incorporation in calcite can also be dependent on synergistic chemical factors such as pH, where lower pH is shown to enhance sulphate incorporation (Wynn et al., 2018). The effect of pH on general calcite growth has been reported in the literature, notably in the work by Ruiz-Agudo (Ruiz-Agudo et al., 2011). Here, calcite growth rate decreases with increasing pH, which is in fitting with calcite growth models proposed by Nilsson and Sternbeck (Nilsson and Sternbeck, 1999). The effect of pH specifically in the context of trace element incorporation in speleothems is not yet clear, though the calcite growth speed is likely to play an important role in calcite/trace-element interactions. There is some evidence that incorporation of impurities is correlated, where trace ion impurities facilitate incorporation of

other impurities. Though these effects are highly complex and not yet fully understood, examples are reported in the literature (Davis et al., 2000; Lorens, 1981).

Use of speleothem archives for paleovolcanism is an emerging field and much of the literature reported in geological/archaeological contexts is missing small-scale chemical and thermodynamic insights. However, the solid-state chemistry of calcium carbonates is extremely important for applications in paleoclimatology, and warrants greater research attention in this context. Though not directly related to the speleothem applications of the present thesis, work by Ruiz-Hernandez (Ruiz-Hernandez et al., 2010) documents how computer simulation studies of calcium carbonate systems may be used to derive thermodynamic conclusions, to support paleoclimate reconstruction research. Here the authors use computer simulation to derive thermodynamic behaviour, and therefore paleoenvironmental significance of strontium incorporation in coral aragonites, which draws several parallels to work reported in the present thesis.

1.3. Speleothems

Speleothems such as stalagmites and stalactites are secondary calcium carbonate mineral deposits that form in *karst* caves. The karst primarily develops in limestone, that has been partially eroded in channels by the flow of water. Limestone is a common sedimentary rock that forms from deposition of chemical and biological calcium carbonate particles. *Aquifers* are underground bodies of water and *karst aquifers* are underwater bodies of water in karst rocks. Aquifers are filled mainly by rainwater that has passed through the various layers of soil before reaching the bedrock, thus aquifers can be rich in dissolved impurities from the soil layers. Karst aquifers form caves in carbonate rocks because small amounts of natural carbonic acid (formed from rainwater and carbon dioxide gas) dissolve the limestone to create large fissures. These aquifers are differentiated by *porous* aquifer rocks because water flows much more rapidly through open channels, rather than filtering slowly between particles in substances such as sandstone.

Speleothem formation relies heavily on materials delivered via the karst layers above caves. Drip water from the karst aquifer enters the cave containing dissolved calcium and bicarbonate ions from the limestone rock and calcitic soils. Lower relative partial pressure of carbon dioxide gas in the cave drives a degassing of bicarbonate ions to form a solid precipitate of calcium carbonate with the production of water and carbon dioxide gas. Water entering from the cave ceiling first forms a *stalactite*, which grows by drip water running down the outside of the solid straw precipitating a very thin layer of calcium carbonate. As excess water

drips onto the floor from the stalactite, a *stalagmite* begins to form in the upward direction, via the same process. Impurities that have filtered through the soil (*leeching*), into the karst aquifer and into the cave drip water accumulate are incorporated into the calcium carbonate structures at the time of formation. Such impurities may be volcanic in origin and are the basis for investigation of speleothems in the present context. Trace element concentrations of the order of ppm are generally required for effective paleoclimate mapping, therefore an effective delivery of trace compounds from the soil layer to the cave via fast-flowing karst aquifer systems is essential (Fairchild and Baker, 2012).

Though speleothems are often in seemingly isolated cave environments, they may be influenced by external factors such as seasonal changes in rainfall and chemical composition of cave drip water. Speleothem growth rates are usually observed to be influenced by seasonal changes in rainfall and therefore drip water flow, and it follows that speleothems show seasonal trends in trace impurities due to changes in delivery (rainwater) (Fairchild and Treble, 2009). Aside from being used as a temporal indicator of local environmental impurities, an annual trace element 'pulse' may also be used to assist uranium techniques in the dating of speleothem subsections (Smith et al., 2009).

Layered speleothem growth is one of the main reasons why speleothems are well suited to serve as paleoclimate archives. The speleothem grows 'outward' via cave drip water that runs down the outside of the structure, therefore the newly formed thin layers of calcium carbonate are roughly concentric. These are known as speleothem *laminae* (Figure 1.5). Speleothem laminae are often annual and crucial features for paleoclimate archives because they are naturally discrete sections of the overall structure which capture and preserve impurities from around the time of their formation. Thus, identifying the age of a particular layer identifies the time at which the impurities detected were present in the local environment.

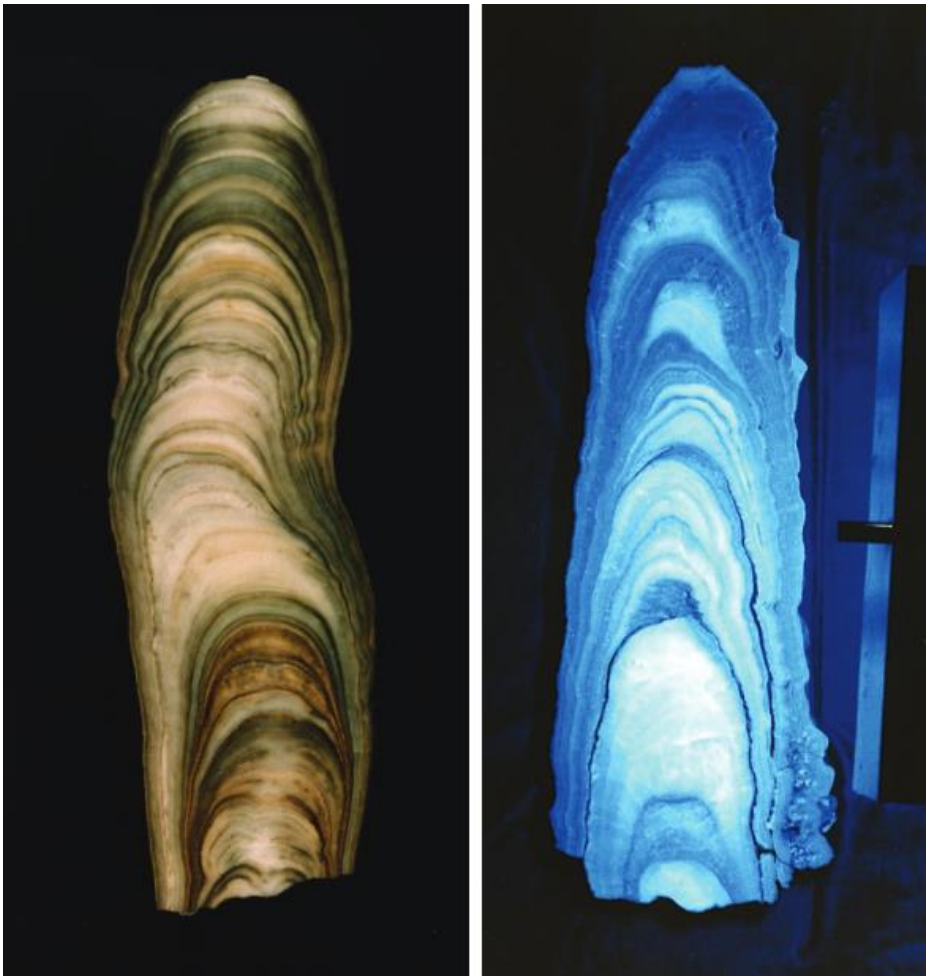

Figure 1.5. Speleothem axial cross section image from (King et al., 2004).

In the previous section, it was described how calcite is the most common phase detected in speleothems, though there are likely to be recrystallisation and phase transition processes happening between formation and analysis. Unstable phases such as vaterite could form where cave drip waters possess the right Mg/Ca ion ratio, which recrystallizes to calcite later in the life of the speleothem. Though there is currently no evidence for this kind of phenomenon in caves, the interaction of the speleothem carbonate with impurities throughout its lifetime, rather than just at the time of analysis is a potentially important consideration for interpreting trace-element proxy concentrations.

Section 1.2 describes in detail how impurities may incorporate chemically into the crystal structure of calcium carbonate. There are also mechanisms of physical incorporation, in which impurities do not necessarily become chemically bonded. One such physical phenomenon known as *fluid inclusion* is the trapping of micrometre-scale pockets of water between or within calcite crystals. These small pockets of water become completely isolated from the atmosphere as the speleothem grows, trapping any chemical impurities from the karst aquifer. Analysis of

fluid inclusions for paleoclimate research contexts is reported in a variety of texts, including (Affolter et al., 2019; Fleitmann et al., 2003; Ghadiri et al., 2020).

Chemical analysis of speleothem carbonates is broadly divided into categories: detection and dating. Detection and identification of impurities may be achieved via different analytical approaches due to the broad range of possible impurities to be detected. Geochemical methods for impurity detection in the solid state is briefly introduced here, due to the relevance to practical applications of the present work. For analysis of the solid crystalline structure, X-ray diffraction based on Bragg's law is frequently used. Powder X-ray diffraction (and appropriate post process refinement modelling) is frequently used in this field because speleothems rarely contain single crystals of sufficient size for single crystal diffraction experiments. Topographical studies of surface and grain boundary regions of the speleothem carbonates are frequently carried out using electron microscopy techniques. Both techniques can provide atomic level structural information, allowing researchers to determine the phases present in the system, as well as the degree of speleothem crystallinity, and the particle shapes/morphologies within the macroscopic mineral structure. Detection of elements in trace concentrations provides a significant analytical challenge. X-ray fluorescence spectroscopy is typically required, with a high flux radiation source such as a synchrotron. X-ray fluorescence can identify trace-elements with great sensitivity by measuring photo-electronic transitions that are unique to each element. This technique was applied in the paper by Badertscher (Badertscher et al., 2014), to identify trace-element pulses consistent with paleovolcanic activity (Frisia et al., 2005; Frisia et al., 2008). Figure 1.6 below shows XRF derived trace element profiles for sulphur, molybdenum, and bromine in speleothems as reported Badertscher and co-workers (Badertscher et al., 2014). Pulses can clearly be seen in the yellow region, which is around the time of the Minoan eruption, which is known to have occurred around the time.

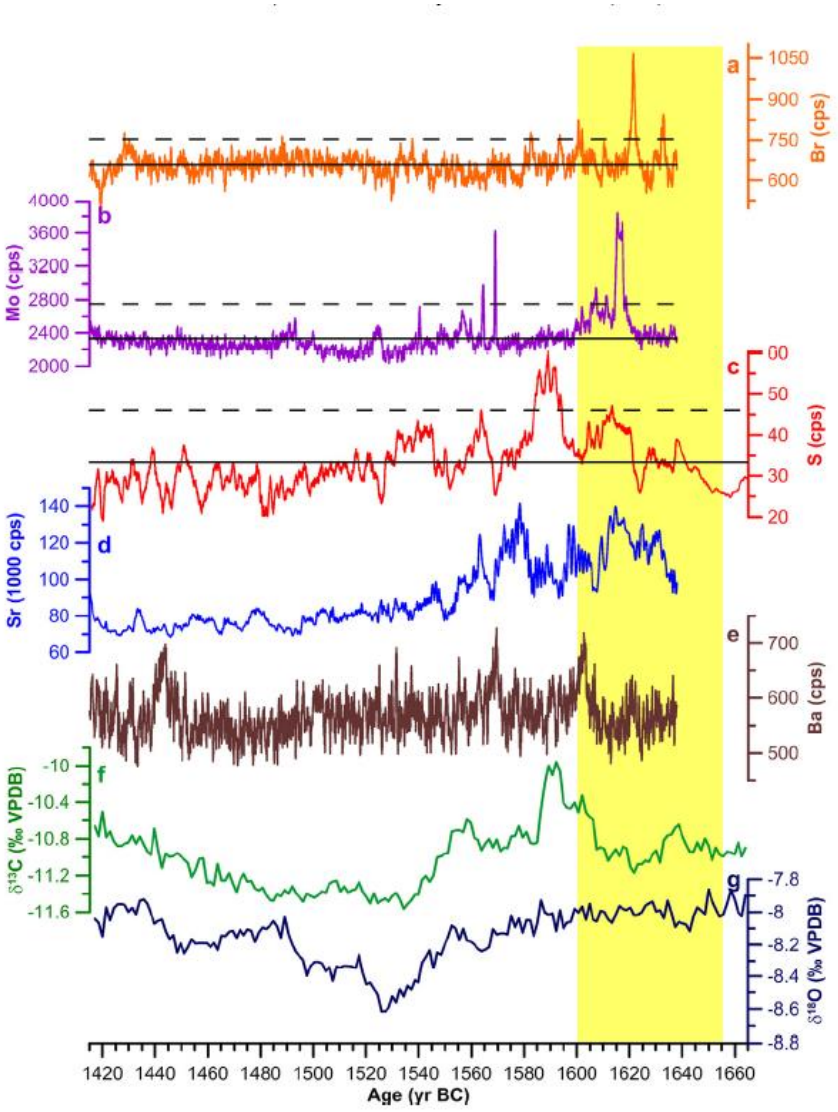

Figure 1.6. XRF trace element profiles for speleothems matching the Minoan eruption around 1600 BCE. Figure from (Badertscher et al., 2014).

Dating speleothem laminae is also essential in matching detected trace-elements to periods in Earth's history. This is typically done using isotopic analysis, which falls into one of three main categories. The first is the use of ^{14}C which is produced in polar atmospheres by the interaction of ^{14}N with neutrons from galactic cosmic rays and solar energetic particles. When carbon-based compounds are in an open system with their environments, the ratio of $^{13}C/^{14}C$ is at equilibrium. When the carbon-based compound is in a closed system and unable to exchange with the surrounding environment (i.e. when a new speleothem layer forms over the last) ^{14}C begins to decay with a half-life of around 5.5 kyr. Because of this relatively low half-life compared to the age of speleothems, there is often a very low percentage of heavy carbon in speleothems, which can hinder the use of carbon-based isotopic techniques for dating.

Radiocarbon dating of speleothems also suffers from the so-called 'hardwater' effect. A key assumption in radiocarbon dating is that a system exchanges heavy carbon with the atmosphere during its lifetime. This is accurate for systems such as photosynthetic organisms or for organisms that feed on photosynthetic plants; however systems such as speleothems do not exchange carbon with the atmosphere, they exchange carbon with a local aqueous environment. This local aqueous environment is likely to be a 'reservoir' to dissolved calcium carbonates, meaning that the carbon exchange equilibrium is quite different to carbon exchange with basal atmospheric levels. This effect is well documented and often leads to inaccurate estimation of the age of marine carbonates.

Uranium series dating (U-Th, U-Pb) is the *de facto* dating technique used in speleothem research. Uranium may be transferred to speleothems by dissolution of bedrock into the karst aquifer, where it becomes incorporated into the mineral structure. The exact chemical nature of uranium species in the speleothem is unclear from the literature, though surface cationic binding of U(VI), or cation binding of UO_2^{2+}, seems plausible based on studies of pure calcite (Kelly et al., 2003). ^{238}U - ^{230}Th (two-step alpha decay) works based on differing solubilities between uranium and thorium. Uranium compounds are highly soluble in water and can be readily transported via cave drip water. Conversely, thorium is highly insoluble therefore it is extremely unlikely to be transported in solution to the speleothem. It is therefore likely that any thorium in the speleothem is a decay product of uranium. U-Pb dating is a family of techniques which rely on multiple alpha and beta decay steps via unstable intermediates between uranium and lead. Some techniques use the different half-lives of ^{238}U and ^{235}U isotopes to determine age via lead isotope analysis. U-Pb dating also makes use of differences in crystal lattice affinity for each radioactive decay member. Uranium is thought to incorporate into calcium carbonates readily while lead does not, meaning that lead detected in crystalline speleothem carbonates is likely to be a decay product of uranium.

1.4. Aims and objectives

From the discussion above, it is clear that speleothem-bound trace-element impurities have a very promising potential to provide long timescale records of paleo-volcanic activity, which is a key input to models for reconstructing past climates. Experimental detection of such impurities via synchrotron radiation-based techniques is at present, unable to give detailed information on the specific chemical nature of the elemental impurity.

The aim of this thesis is therefore to provide atomic-level information about the incorporation of commonly-used chemical tracers of volcanic origin into speleothem-forming minerals. This work focuses on three main chemical tracer elements (Mo, S, Br) reported in the works by Badertscher (Badertscher et al., 2014), in geochemically stable oxidation states, facilitating anionic exchange process with CO_3^{2-}. Calcium carbonate in five naturally occurring phases are considered, though calcite features predominantly, as the most common natural phase. The investigation will use modern theoretical simulation techniques based on density functional theory (DFT), with molecular dynamics for solid/liquid interfaces.

Specific objectives are:

I. To elucidate the relative thermodynamic stabilities of molybdenum in the form of molybdate oxyanions in different calcium carbonate minerals, with respect to the more widely studied sulphate oxyanion.
II. To assess the relative stability of bulk vs. surface incorporation of these elements, and the possible role of surface step defects in accumulated impurities.
III. To understand the charge compensation mechanisms when singly charged anions (in particular bromate) substitute doubly charged carbonate anions, and the effect of these compensation mechanisms on the impurity distribution and thermodynamic stability.
IV. To evaluate the efficiency and limitations of *ab initio* simulations including static optimisation and molecular dynamics in modelling the incorporation of impurities in carbonate minerals.

1.5. Thesis outline

In Chapter 2 we present a discussion of the theoretical methods and simulation codes relevant to work reported in the present thesis. Chapter 3 reports an investigation into use of molybdate as a tracer in multiple calcium carbonate phases, with comparison with the more widely studied sulphate tracer. This chapter is an adaptation of the publication (Midgley et al., 2020). Chapter 4, which is also an adaptation of a publication (Midgley et al., 2021a), builds on ideas presented in Chapter 3 by also studying the incorporation of sulphate and molybdate. Here we focus on calcite as the predominant calcium carbonate phase, while examining the incorporation of sulphate and molybdate at typical surface terminations of this phase, using *ab initio* molecular dynamics to represent the carbonate interface with liquid water. This is a more sophisticated structural model compared with what is used in Chapter 3. Chapter 5 is an adaptation of the pre-printed publication (Midgley et al., 2021b), which has also been submitted to a peer review journal. Here we present the incorporation of bromate in calcium carbonates. Though we revert to a less sophisticated structural model by examining only crystalline bulk bromate substitutions; this allows us to examine the effect of synergistic incorporation of a singly charged cation, to compensate the singly charged bromate anion. There are many possible configurations when incorporating multiple ions, introducing interesting thermodynamic correlations. Chapter 6 contains general discussion and implications of this work, including limitations with ideas and suitability of simulation techniques for the tasks complete. In Chapter 7, we discuss projects that were started but not completed, due to limitations in the amount of supercomputing time available. These partially completed research projects are discussed in the context of suggestions for future work. All references are given at the end of the document.

2. Computational Methods

2.1. Introduction to basic concepts in crystal structures

2.1.1. Periodic materials

Calcium carbonate phases reported in the present thesis form naturally occurring 3D arrays of atoms and molecular ions, which are arranged in *periodically* repeating *unit cells*. Calcium carbonate *phase* (used interchangeably with *polymorph* throughout this thesis) refers to materials with different structure and symmetry, despite having the same stoichiometry. The phase formed by a solid material is dependent on both kinetic and thermodynamic pathways. With geological timescales, thermodynamic driving forces often prevail over temporarily kinetically stabilised phenomena (e.g. vaterite may form as a temporary kinetically-driven phase of calcium carbonate, while after time, the more thermodynamically stable calcite phase is likely to be formed).

Unit cells are the smallest subset of atoms arranged according to the symmetry of the full system. The asymmetric unit is also sometimes used in crystallography, where a sub-unit-cell representation of the system is expressed with appropriate symmetry operators. Unit cells are of important in computational materials science as they are often the smallest repeat unit that can be used to model a crystalline material. When simulating a unit cell, the system is represented as if it is part of an infinitely repeated crystalline array, this is a *periodic* simulation. *Periodic* simulations form one of two major branches of computational materials science, where the other is simulation of finite systems such as discrete molecules. The present thesis studies mainly *periodic* crystalline models, though there are principles from molecular system simulation employed as discussed in the present chapter.

The unit cell has important implications when computing electronic properties for crystalline solids. They may be combined to form supercells, which are used throughout work described in the present thesis. By convention, a supercell consisting of two unit cells in each of the three crystallographic directions (*a, b, c*) is expressed as a 2×2×2 supercell. Supercells are advantageous when representing impurities substituted in crystalline materials because they allow for the avoidance of an unphysical interaction between an impurity and its periodically repeated image during a simulation.

2.1.2. Miller indices

Miller indices are a numerical description of crystallographic planes, used frequently throughout this work to describe the surface termination of calcium carbonate. There are various formalisms when using Miller indices, though in the present thesis, crystallographic

planes are noted using the form (*h k . l*), which is based on the Miller-Bravais formalism for describing hexagonal crystal systems.

2.1.3. Surfaces, interfaces and slab models

Computational materials science frequently deals with the unique chemistry present at crystallographic plane terminations. To models these structures, the crystal is typically cleaved, with a vacuum spaced introduced in the crystallographic direction perpendicular to the surface (*c*). This is because models, as employed in the present thesis, remain periodic in three dimensions despite the formation of a surface termination. In the *c* crystallographic direction, the solid-vacuum (slab) model is periodically repeated, therefore vacuum space must be enough to ensure that the surface does not interact with its periodically repeated image in the *c* direction. The model is also periodically repeated in the *a* and *b* crystallographic planes, making an infinite solid/vacuum slab model in two dimensions.

When simulating interactions between solids/liquids, full quantum mechanical description of the nanoparticle with liquid surrounding would incur extreme computational costs. It is therefore often necessary to form slab models of key solid/liquid interfaces, to capture the fundamental interactions on a minimal number-of-atoms.

When simulating the interface between a solid and a liquid, the model is prepared by cleaving the bulk solid and introducing a vacuum space. The vacuum space is then 'filled' with the appropriate solvent, to create a periodically repeated solid-liquid model, with alternating solid-liquid-solid layers in the *c* direction. The liquid layer must be sufficiently large to be able to accommodate representation of liquid in the bulk and liquid that is adsorbed at the surface of the solid.

2.2. Quantum chemistry

2.2.1. Introduction

Computational chemistry calculates energies of atoms when in different configurations. Interactions between groups of atoms may be modelled as either classical (obeying laws of Newtonian mechanics) or *ab initio* (obeying laws of quantum mechanics). *Ab initio* calculations require, in principle, no prior information to compute the energy for a set of atoms; however modern *ab initio* simulation techniques rely on approximations to allow for computational investigation of large sets of atoms within a viable timeframe. The present thesis describes simulations based mainly on quantum mechanics; therefore, a comprehensive review of contemporary *ab initio* methods is presented forthwith. Semi-empirical and classical methods

are discussed briefly towards the end of the chapter, because these methods were used less frequently throughout the present work.

2.2.2. The Schrödinger equation

In quantum mechanics, the time-independent Schrödinger equation affords energy calculation for a massive particle moving in space, with an external potential that does not change with time.

$$\hat{H}\psi = E\psi \tag{2.1}$$

The wavefunction, ψ, is a mathematical object containing all possible observable attributes relating to a system (such as an electron). It allows for calculation of the probability of physical attributes such as position and momentum. The Hamiltonian operator contains the kinetic and potential energy for the system in question, (Equation 2.2 – Hamiltonian for a particle moving in one dimension). Where the wavefunctions are eigenvectors of the Hamiltonian operator, the energy eigenvalues, E, are the energies of the stationary states (experimentally observable time-independent quantum states) (Atkins et al., 2018; Atkins and Friedman, 2011). The Hamiltonian is expressed as:

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x) \tag{2.2}$$

Where $\hbar$ is h/2π, where h is Planck's constant, m is the mass of the particle and V represents the potential energy.

The form of the wavefunction itself is related to the de Broglie wavelength of the particle, which is a function of its linear momentum. The Heisenberg uncertainty principle tells us that there is a numerical trade-off between knowledge of the position and momentum of a particle. Therefore, a sharper wavefunction contains more positional information, a smoother wavefunction will contain more information on particle momentum. In the Born interpretation, the square modulus of the wavefunction is used to describe the probability density of finding the particle in a region of space (Atkins et al., 2018). The integral of $|\psi|^2 \mathrm{d}x$ is therefore proportional to the probability of finding the particle in that region of space. The wavefunction can be either negative or complex, with the following implications. Where the wavefunction is negative, this allows for destructive interference, which means there can be regions of wavefunction overlap where the total wavefunction is zero. Where wavefunction

amplitude is zero, there is zero probability density of finding a particle in that region of space, this is called a *node*. Wavefunctions are often expressed as complex numbers; while this carries no physical significance, it is a mathematically concise way of expressing wavefunction phase and amplitude, without laborious expression of the attenuated sine and cosine functions.

The Schrödinger equation for a many-body system of electrons and nuclei is reported in Equation 2.3.

$$\left[\frac{\hbar^2}{2m}\sum_{i=1}^{N}\nabla_i^2 + \sum_{i=1}^{N}V(r_i) + \sum_{i=1}^{N}\sum_{j>i}^{N}U(r_i, r_j)\right]\psi = E\psi \tag{2.3}$$

Where the terms in brackets are as follows: sum over electronic kinetic energies in three dimensions, sum over electronic potential energy with all nuclei, sum over all electron-electron potential energies (Atkins and Friedman, 2011).

Increasing the number of electrons and nuclei increases the number of variables required to solve Equation 2.3. In a system where there are more than two particles, the *N*-body problem (known in all of physics) applies, meaning that an analytical solution to the Schrödinger equation is not possible. A numerical solution is therefore obtained using approximations and iterative approaches described in the following sections of this chapter. Near-numerical solutions to many particle systems incur extreme computational costs, often beyond the scope of modern computing. It is therefore necessary for practical applications to impose further approximations to the quantum mechanical problem, so that modern high-performance computers can generate numerical solutions to an acceptable standard.

2.2.3. The Born-Oppenheimer approximation

A commonly exploited and highly effective approach to reduce the complexity of the quantum chemistry computation is known as the Born-Oppenheimer approximation (BOA). The BOA assumes that the high mass of the atomic nucleus compared with the mass of an electron means that electrons move without consequence to the position of the nucleus, therefore nuclear and electronic motion may be separated (Jensen, 2006). In practice, only electrons need to be treated quantum mechanically, giving the electronic Schrödinger equation (where nuclei are treated classically). This approximation is fundamental to computational materials science and has been used throughout all *ab inito* calculations reported in the present thesis.

2.2.4. The potential energy surface

Solution of the electronic Schrödinger equation gives a potential energy surface (PES), also known as the electronic or adiabatic potential energy surface. The PES describes how electronic energy varies as a function of each individual geometric degree of freedom (DoF). Where there are two or more DoF in a system such as a small molecule, the PES surface describing the electronic energy is a function of the same number variables. The overall electronic energy in a multi-DoF system is therefore described by a multi-dimensional potential energy hypersurface (PEHS) that is a function of multiple variables (one for each DoF). The PEHS is often referred to as the PES and for brevity, represented on one dimension, where the electronic energy changes as a function of some conjugate geometric variable (Lewars, 2011).

Nuclei move according to the electronic PES by classical forces calculated using the Hellman-Feynman theorem. The forces on each nucleus are calculated from the negative of the gradient of the PES thus, nuclei move to minimise the electronic potential energy of the system (Atkins and Friedman, 2011; Jensen, 2006).

2.2.5. Hartree-Fock theory

The Hartree-Fock (HF) method is an *ab initio* approach to computing an approximate solution to the electronic Schrödinger equation. In HF theory, the total electronic wavefunction is expressed as a Slater determinant of one-electron wavefunctions called spin-orbitals. The Slater determinant formalism allows the HF wavefunction to obey the antisymmetry principle, which says that the sign of the wavefunction must change when the generalised coordinates (including both position and spin) of two fermionic particles is inverted. This is a general formulation of Pauli's exclusion principle and means that two fermions with the same spin will experience a repulsive force, known as the exchange force, if they are in the same region of space (Jensen, 2006; Lewars, 2011).

$$\Psi(x_1, x_2) = -\Psi(x_2, x_1) \tag{2.4}$$

For the case of two electrons, this is expressed mathematically in Equation 2.4 – rearrangement shows that for two electrons with the same spatial and spin coordinates, the wavefunction $\Psi(x, x)$ must be zero (i.e. the state is impossible). Though HF calculations have not been employed in the present thesis, many principles from HF are used in Kohn-Sham density functional theory, therefore they are briefly reviewed here.

2.2.6. Exchange and correlation

Exchange is one of two quantum mechanical forces between electrons that have no counterpart in classical physics. In classical Coulombic theory, two electrons may only ever repel one another; while the exchange force allows for attraction between two negatively charged electrons of antiparallel spin. Hartree-Fock theory is used most often in modern computational materials science not as a standalone method, but in hybrid methods because it can be used to provide an exact exchange force calculation via the Fock matrix. This exchange term is often used to assist more advanced electronic structure methods (discussed later in this chapter).

The other of the two non-classical forces mentioned above is known as correlation, which is defined by its complete absence from HF theory. Correlation energy is the difference between the HF limit (the energy of the best HF solution with a complete basis set - Section 2.4.2) and the exact total energy (Jensen, 2006; Lewars, 2011). This may be summarised as saying that the overall motion of a group of electrons is not completely described by taking a product (or a Slater determinant) of one-electron wavefunctions. Post-HF methods for calculation of correlation include Moller-Plesset perturbation theory and Coupled Cluster. While these approaches are extremely accurate for calculating electronic structure, they are computationally expensive and are only viable for use on systems of a few atoms.

Though exchange and correlation interactions account for a small overall percentage of the energy, chemical interactions such as covalent bonding and van der Waals can only be explained by considering these interactions. Therefore, great efforts are made in computational materials science to represent these forces accurately.

2.2.7. Variational principle and self-consistent field theory

The variational principle is an important concept computational materials science and theoretical chemistry. It simply states that an energy derived from an approximate wavefunction, will be above or equal to the actual ground state energy, but never below.

The HF wavefunction may not be solved analytically because the Fock operator contains one-electron orbital coefficients, which are in-turn calculated by diagonalization of the Fock matrix. To begin computation, an approximate guess at the one-electron wavefunctions must be made. Initial guesses at one-electron orbitals are used to calculate an approximate solution to the electronic Schrödinger equation, which in-turn generates an updated set of energy eigenvalues and one-electron wavefunctions. This iterative process to minimise the total

energy of the system is known as the self-consistent field (SCF) and continues until the energy is converged to a user-defined sensitivity. By the variational principle, the calculated energy can never be below the true ground state energy, therefore better and better approximations used should yield a lower energy which is closer to that of the true state (Dronskowski, 2005; Lewars, 2011). The SCF approach is extremely successful and has been adopted by other extremely popular *ab initio* methods such as density functional theory (Section 2.4).

2.3. Density functional theory

2.3.1. Introduction

Density functional theory (DFT) is the *de facto* electronic structure technique in modern computational materials science. Like HF theory, it is a method for gaining an approximate solution to the electronic Schrödinger equation. However, it works on the electron density of the system, rather than the multi-electron wavefunction. Conceptually, the electron density is a simpler descriptor of the state of the system, because it only has 3 variables, or 4 if spin is included; whereas the multi-electron wavefunction depends on $4N$ variables, i.e 3 electron coordinates and one spin coordinate per atom.

2.3.2. Hohenberg–Kohn theorems

DFT was born with the development of two key theorems. The first theorem (Hohenberg and Kohn, 1964):

> *'The ground state energy is a unique functional of the electron density'*

This may be interpreted as meaning that the ground state electron density overall determines the energy and the wavefunction. Therefore, the Schrödinger equation may be solved with respect to the electron density rather than the multi-electron wavefunction. Though there is a unique functional that may be used to solve the Schrödinger equation without the need for wavefunctions at all, the mathematical form of such a functional is unknown, therefore further subsequent developments have been necessary to enable practical use of DFT.

The second theorem (Hohenberg and Kohn, 1964):

> *'The electron density that minimizes the energy of the overall functional is the true electron density corresponding to the full solution of the Schrödinger equation'*

The second theorem may be interpreted as meaning that if the energy is minimised, then the ground state electron density is known. However, because the mathematical form of the

functional is not known, the variational principle (Section 2.2.7) may be used in practise to minimise the energy via manipulation of the electron density.

The first, '*pure*' form of DFT (no wavefunctions required), sometimes known as orbital-free DFT was successful in the description of small or no band gap solids, where the electron density is relatively homogeneous (Jensen, 2006). It was not successful however in describing areas of localised electron density such as molecules, therefore uses were reasonably limited.

2.3.3. Kohn-Sham theory

Kohn-Sham (KS) density functional theory was developed so that DFT may be applied to molecular systems of non-homogenous electron density (Kohn and Sham, 1965). The total KS energy may be written as a sum of potential energies:

$$E[\rho(r)] = T_e[\rho(r)] + V_{ne}[\rho(r)] + V_{ee}[\rho(r)] + E_{xc}[\rho(r)] \tag{2.5}$$

In order of terms: electronic kinetic energy, electron-nuclear potential energy, classical electron-electron potential energy and exchange-correlation (XC) electron-electron potential energy. Where $\rho(r)$ is the electron density at point *r*.

KS theory replaces the electron density with a fictitious system of non-interacting electrons, with the same density. This non-interacting electron density is used to calculate the term T_e in Equation 2.5, which comprises around 99% of the total energy. The fictitious density of non-interacting electrons is built from one-electron KS orbitals:

$$\rho(r) = \sum_{i=1}^{N} |\psi_i|^2 \tag{2.6}$$

Each of which have the mathematical form:

$$\left[\frac{h^2}{2m}\nabla^2 + V(r) + V_H(r) + V_{XC}(r)\right]\psi_i = \varepsilon_i\psi_i \tag{2.7}$$

The first term is the electronic kinetic energy (∇^2 is the three-dimensional differential operator). The second $V(\boldsymbol{r})$ term is a potential energy for the interaction of the electron with all nuclei in the system. The third term, known as the Hartree potential, has the form:

$$V_H(r) = \int \frac{\rho(r')}{|r - r'|} dr' \tag{2.8}$$

This is the classical repulsive coulomb potential for the electron with the electron cloud. As the electron itself forms part of the overall electron cloud, there is an unphysical interaction of the electron with itself as part of the cloud, known as the *self-interaction* problem. The self-interaction problem is corrected by a parameter included in the XC (V_{XC}) term, though it can cause problems when simulating areas of very well-defined electron density such as partially filled d orbitals. The exchange-correlation term depends on the XC functional, this contains (approximated) quantum mechanical electron-electron effects.

KS DFT relies on one-electron wavefunctions to build the so-called KS orbitals (Sholl and Steckel, 2009). The orbitals are expanded in terms of basis functions to give a the so-called KS matrix. The KS operator (h^{ks}) is build using an initial guess at the KS orbital coefficients. It is these orbital coefficients which must be optimised to find the ground state KS energy, therefore an iterative problem based on the SCF procedure is reintroduced, analogously to the HF approach. The goal of a KS electronic energy minimisation is therefore to find a numerical solution through iterative minimisation of the one-electron KS wavefunction energy eigenvalues. A schematic diagram of a simplified SCF Kohn-Sham algorithm is shown in Figure 2.1.

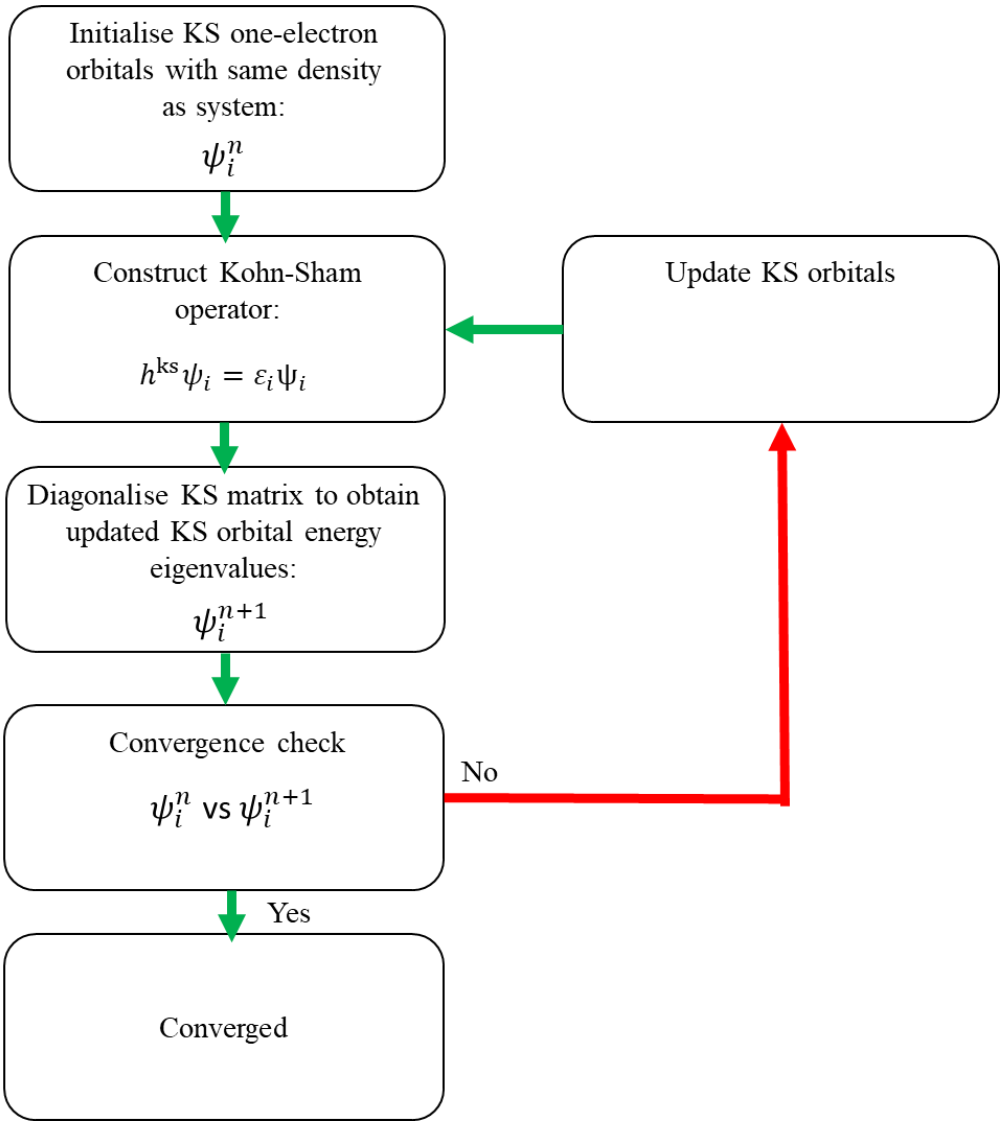

Figure 2.1. Iterative Kohn-Sham algorithm.

2.3.4. Exchange-correlation functionals

The final term in the one-electron KS orbital equation (Equation 3.7), V_{XC}, accounts for <1% of the total KS energy, yet the XC interactions are imperative for describing chemical behaviour in covalent and dispersion interactions. The universal form of the XC functional, $E_{XC}[\{\psi_i\}]$, is not known, therefore approximations must be made (Sholl and Steckel, 2009). The simplest way to model quantum mechanical electron-electron interactions is to approximate the density as a homogeneous electron gas, known as a jellium. The form of the XC interactions can be derived from the jellium because the electron density is uniform at all points in space. The XC functional derived from the jellium is known as the local density approximation (LDA) (Sholl and Steckel, 2009):

$$V_{XC}^{LDA}(r) = V_{XC}[\rho(r)] \tag{2.9}$$

The resulting solution to the Schrödinger equation is therefore subject to significant approximations to the XC interactions as in reality, the Jellium is not a good representation of the electron density in most systems. The LDA is known to cause various systematic errors such as: overestimation of binding energies in molecules, poor description of van de Waals interactions and a poor presentation of highly localised electronic densities such as d orbitals (Dronskowski, 2005). LDA functionals are not truly *ab initio* as they contain empirically derived parameters from more accurate computational methods and are fitted to analytical functions. For this reason, LDA is known to perform anomalously well when representing very specific systems, particularly where there is high homogeneity in the electron density. However, this is indicative of overfitting empirical terms rather than effectiveness of the functional. LDA is therefore unsuitable for many applications in computational materials science, prompting use of more sophisticated approximations to the XC interactions.

The second rung on the so-called *Jacob's ladder* of exchange-correlation functionals is the generalized-gradient approximation (GGA) class of functionals. GGA functionals are similar to LDA however they contain the local gradient of variation to the electron density at any point in the electron gas (Sholl and Steckel, 2009):

$$V_{XC}^{GGA} = V_{XC}[\rho(r), \nabla\rho(r)] \tag{2.10}$$

Where $\nabla\varrho$ is the gradient of the electron density, ϱ, around point *r*. Thus, they can capture electronic interactions in a material of variable electron density, which is representative of systems containing electronic orbitals. GGA functionals are widely used in computational materials science as they present an effective balance between accuracy and computational efficiency. The GGA functional by Perdew, Burke and Ernzerhof (PBE) is a *de facto* when simulating periodic materials with DFT, as it features no empirical fitting parameters (Perdew et al., 1996a). The effectiveness is well documented with several examples and reviews testament to its success, particularly in larger systems comprised of many atoms, where its computational efficiency is essential. (Ernzerhof and Scuseria, 1999; Lee and Martin, 1997) report the efficacy of the PBE functional. There are however some known issues with GGA which come from concessions made to simplify the functional. Lattice parameters in metallic systems and molecular excitation energies are among examples of where a lack of sensitivity in GGA functionals is known to be problematic. The main flaw with LDA and GGA type

functionals is the lack of *non-local correlation*, which is responsible for dispersion forces in chemical systems (Sholl and Steckel, 2009). The derivative of the electron density around a point allows for semi-local electron correlation interactions to be captured; however these are short ranged and do not critically depend on the electron density across the whole system.

Most DFT calculations reported in the present thesis were performed using the PBE functional. When simulating liquid water, the revised PBE (revPBE) functional was employed (Hammer et al., 1999). The revPBE functional is an extension of the standard PBE functional, with some extra fitting parameters for describing liquid water. Several studies have shown that revPBE combined with Grimme's DFT-D3 corrections (Section 2.3.6) is extremely effective in representing water (Lin et al., 2012; Ruiz Pestana et al., 2017; Wang et al., 2011a).

Though GGA represents the highest level of theory that was computationally affordable for systems of interest (< 100 atoms), it is important to comment on the efficacy of the GGA-PBE functional to capture energy differences of the magnitude reported throughout this thesis. There are several examples in the literature of PBE functionals being used to derive *ab initio* thermodynamics for mineral systems, including calcium carbonates. Some illustrative examples by Chaka (Chaka, 2018; Chaka, 2019) and Balan (Balan et al., 2014) show the efficacy of PBE functionals in representing calcium carbonates. Energy variations with respect to other functionals are typically in the order of 0.1 eV, which is lower than most thermodynamic energy differences reported in the present thesis. Though not reporting thermodynamic energies, there are several other reports of a PBE functional being used to probe atomic-level effects in calcium carbonates (+ impurities), with emphasis on accurately representing experimental properties such as unit cell vectors (Feng and Redfern, 2018; González-López et al., 2014). We are therefore confident that a PBE functional is an appropriate level of DFT theory for derivation of structural and thermodynamic insights in calcium carbonates with an accuracy in the order of 0.1 eV. Indeed, for static DFT calculations reported in the present thesis, this level of accuracy is sufficient to validate all observed correlations. Though estimates of the energetic error bars due to the DFT method are difficult to compute for static simulations, we show in later Chapters that our methods effectively reproduce experimental unit cell vectors. Energies computed via *ab initio* molecular dynamics methods may also contain a small systematic uncertainty through use of a PBE functional, however this is likely to be vanishingly small compared with the magnitude of other energetic fluctuations caused by the MD simulation.

More sophisticated XC correlation functionals are known and widely exploited in computational materials science, though the present thesis uses none of these methods that form the *third* and *fourth* rungs of the *Jacob's ladder* of XC functionals. More sophisticated functionals such as those described briefly in the following paragraphs are of extreme computational expense for just a few atoms. They are therefore unsuitable in simulations of many hundreds of atoms, where a very precise description of the electronic structure is not required.

The third rung of the *Jacob's ladder* of exchange correlation functionals is the meta-GGA functionals. Meta-GGA functionals are very similar to standard GGA functionals, though they contain the second derivative of the electron density $\nabla^2\rho(r)$, which is a slight improvement to the GGA functionals.

The fourth rung of the *Jacob's ladder* of exchange correlation functionals is the hybrid functionals. Hybrid functionals are commonly used in molecular DFT applications where the system size is typically small enough to allow these computationally expensive methods, although they are becoming more popular in solid state calculations, especially when accurate bandgaps are required. Hybrid functionals incorporate a fraction of exact exchange energy calculated using a Hartree-Fock-type matrix (Section 2.2.6) formed of KS orbitals. Tailoring the ratio of GGA/HF exchange gives rise to hybrid functionals such as the widely exploited B3LYP functional (Lee et al., 1988), that has been tailored to reproduce band gaps with excellent precision. Other examples of these functionals include the ωb97xd functional (Chai and Head-Gordon, 2008), that features long range corrections to the HF exchange to capture molecular excited states and charge-transfer interactions. The hybrid HSE (Heyd and Scuseria, 2004) functional is often used for studying electronic bands in periodic solids, though there are extreme computational costs associated with this functional. This is in part because hybrid functionals such as B3LYP are non-local, meaning that the electron density and its derivatives at a point cannot be computed without knowledge of the electron density and its derivatives at all other points. Non-local correlation corrections allow for a description of van der Waals interactions, which are imperative for the study of the condensed phases.

The fifth rung of the *Jacob's ladder* of exchange correlation functionals are the so-called double-hybrid functionals, these are not relevant to discussion in the present thesis but the reader is referred to the text by Sholl and Steckel (Sholl and Steckel, 2009), for further information.

2.3.5. Corrections to the self-interaction error

During this chapter, the DFT self-interaction problem has been described as the unphysical interaction between an electron and itself in the full electron cloud during calculation of the Hartree potential. This *extra* repulsion causes improbable delocalisation of highly localised electronic environments, which is particularly prevalent in d orbitals. Hybrid functionals offer a remedy to the self-interaction problem, as the HF exchange component of the XC term cancels self-interaction errors as it does in full HF calculations (Sholl and Steckel, 2009). Hybrid functionals however incur extreme computational cost, and therefore are not a viable option in systems where there are many atoms.

A computationally cheaper option that is often used when simulating periodic materials containing d electrons is known as DFT+U (Anisimov et al., 1991). Here, empirical parameters are used to penalise the unlikely expansion of localised atomic orbitals.

2.3.6. Van der Waals corrections

A class of empirical corrections used throughout work reported in the present thesis are the empirical van der Waals corrections by Grimme (Grimme, 2006; Grimme et al., 2010; Grimme et al., 2011). Dispersion forces arise from electron correlation between two points that are spatially separated by a distance further than what can be captured by the semi-local correlation of GGA functionals. GGA functionals therefore completely neglect long-range dispersion forces. Such forces have minimal effect in crystalline ionic solids, though they are significant in the behaviour of liquids. Non-local correlation functionals such as the vdW-DF functional facilitate an excellent description of dispersion forces. However, their computational demands mean that they are not suitable for use with many-atom systems.

Empirical corrections to the electron correlation are highly successful and widely used in computational materials science. In practical applications, a dispersion energy derived from a Lennard-Jones type ($1/R^6$) interatomic potential is added to the total DFT energy, often giving an excellent representation of van der Waals interactions. One of the most commonly encountered implementations of this method is the dispersion correction by Grimme (Grimme et al., 2010). DFT methods employing the empirical Grimme corrections are often referred to as DFT-D2 or DFT-D3 methods, depending on whether Grimme-D2 or Grimme-D3 corrections are used. Grimme corrections are based on the general form:

$$E_{\text{disp}} = -s_6 \sum_{i=1} \sum_{j=i+1} \frac{C_6^{ij}}{R_{ij}^6} f_{\text{damp}} \tag{2.11}$$

Where s_6, is a scaling factor, C is the dispersion coefficient for atoms i and j, R is the internuclear distance and f is a damping function that switches off the correction at short distances, where the semi-local functional gives correct results. E_{disp} is simply added to E_{KSDFT}. Interatomic parameterisations if the Grimme D2 method are between pairs of atomic elements, irrespective of the chemical environment such as geometry, charge and orbital hybridisation. The Grimme D3 method seeks to improve this by making the dispersion coefficient geometry-dependent, which can crudely but effectively differentiate between elements in different chemical states (Sholl and Steckel, 2009).

2.4. Density functional theory: finite systems

2.4.1. Introduction

DFT communities are broadly divided into those concerning approximate wavefunctions for finite systems, often composed of molecular clusters, and those concerning periodic Bloch wavefunctions for extended solids. Though there is significant overlap between these two DFT 'flavours', (most notably in the CP2K code – Section 2.12.3) their representation of electronic orbitals is discussed separately in the context of this chapter.

2.4.2. Localised basis functions

When solving the one-electron problems posed by both HF and KS theory on a computer, a *basis* is required to expand the molecular orbitals (MO), in such a way that the coefficients of the expansion determine the MO and can be found numerically. MO theory is frequently used to describe the electronic structure of molecules, through a linear combination of atomic orbitals (LCAO). When representing these orbitals mathematically (or computationally), *basis functions* are used to represent electron density around atomic nuclei. There are two main types of basis function typically used: the first is the Slater type function (STF) and the second is the Gaussian type function (GTF). STF's accurately represent natural exponential electron density decay when moving away from atomic nuclei. However, it is computationally expensive to calculate overlap integrals of STF's to form molecular basis functions, meaning that they are rarely used. GTFs do not capture natural changes in electron density as accurately as an STF; however their mathematical convenience when calculating molecular basis functions makes them favourable for most practical applications (Lewars, 2011). Typically, contracted Gaussian functions, which are a weighted sum of Gaussian primitives, are defined to mimic a Slater function, therefore achieving the best of both worlds.

Calculations performed in the CP2K simulation package throughout this thesis were performed with GTFs and attenuated Gaussian type (molecular) orbitals (GTO) for representation of electron density around atomic nuclei. CP2K employs a system of contracted Gaussian basis functions to represent the radial electron density component. The spherical component of the GTO is represented using functions from *spherical harmonics*.

There are varying levels of description contained within basis functions, which weigh the benefits of numerical accuracy with computational cost. The minimum requirement for a basis set is to contain one contracted Gaussian basis function per valence orbital (e.g. a H 1s orbital would be represented by a single s function). Such a basis set is often referred to as being *single zeta*; however it is minimal and often insufficient to accurately represent valence shell interactions. The next step up from this is to use the so-called *double-zeta* (DZ) type basis sets. Here, each valence orbital is described by two Gaussian contractions, one at short range and one at longer range (Jensen, 2006; Lewars, 2011). This has the benefit of being able to represent bonding and charge transfer type electronic interactions at a range of distances from the nucleus. DZ basis sets are used throughout CP2K simulations reported in the present thesis. There are several higher-quality types of basis set; however these would have incurred an inhibitive computational cost for the may hundreds of atoms studied.

Core electrons may be described using a single contracted GTF (as in Pople type basis sets), giving the so-called *all-electron* type representations. All-electron calculations are useful when simulating core level spectroscopy type phenomena, or where valence electrons benefit from highly accurate representation of core electrons. As example of this is when studying molecular excited states, where the tiny perturbations in the valence orbitals due to poor description of the core density would be catastrophic. In situations where extreme accuracy of valence orbitals is not necessary, as in the present thesis, it is sufficient to represent all core electrons and nuclear charge using an effective *pseudopotential* (Section 2.5.4).

Basis sets may be augmented with diffuse functions which allow for electron density representation further from the nucleus. This is achieved by elongating the 'tail' on the gaussian function, which can be important for representing long range valence electronic interactions. Diffusivity is noted using a '+' in Pople type basis sets, or in CP2K, the letters 'SR' (short-ranged), are used to signify low diffusivity in the basis set (Lewars, 2011). Diffusivity functions can add computational cost, therefore in basis sets built for speed such as those implemented in CP2K, low diffusivity (computational speed) is prioritised over high physical accuracy over long ranges.

Polarization functions are often added to the orbital functions to allow for representation of natural anisotropic polarizability in molecular systems. Polarization functions are often built to allow atomic orbitals to adopt the shape of an orbital with one *l* quantum number higher than its own, in the presence of a local potential (e.g. a *p* orbital may adopt a shape reminiscent of a *d* orbital) (Lewars, 2011). These functions can improve description of electronic structure, while incurring very little computational cost.

2.5. Density functional theory: periodic boundary conditions

2.5.1. Introduction

Solid state DFT differs from DFT in finite systems most significantly in the numerical basis used to represent the one-electron KS orbitals. In solid state DFT, *planewaves* are used to represent the electron density in *reciprocal space* (term used interchangeably with *k-space*). In reciprocal space, the position of the electron is defined by its momentum, which is represented using periodically repeating waves (planewaves), where wave frequency and momentum are related by the de Broglie relationship. A periodically repeating wavefunction must satisfy Bloch's theorem (Section 2.5.2), therefore it is necessary to work in reciprocal space when calculating wavefunctions for periodic materials. The relationship between real-space (*r*) and reciprocal space (*k*) is a Fourier Transformation (FT). The *r* and *k*-space vectors are related such that $a_r \cdot a_k = 2\pi$, meaning that $a_r \perp b_k$ and $b_r \perp a_k$. An important concept in *k* space calculations is that of the Brillouin zone (BZ). This represents the FT of the unit cell space, the centre of which is known as the Γ point, where $k = 0$ (Dronskowski, 2005; Sholl and Steckel, 2009).

2.5.2. Bloch's theorem

Where DFT is used to gain an approximate solution to the Schrödinger equation for periodic systems, the solution must satisfy Bloch's theorem. Here, the wavefunction of a system with a periodic potential is expressed with the form (Sholl and Steckel, 2009):

$$\psi_k(r) = e^{ik \cdot r} \mu_k(r) \tag{2.12}$$

where exp(ik·r) is a planewave representing the phase of the periodic wavefunction. The wavefunction is indexed by *k* vector, implying that the Schrödinger equation can be solved for each independent value of *k*. $\mu_k(r)$ is a periodic function related to the lattice vectors in *k* space, which can be expanded as a Fourier series using the general form:

$$\mu_k(r) = \sum_G C_G e^{iG\cdot r} \tag{2.13}$$

Where *G* represents all possible vectors in *k* space and *C* represents *G* space planewave expansion coefficients. Therefore, the wavefunction over all of *G* space may be written as:

$$\psi_k(r) = \sum_G C_{G+k} e^{i(G+k)\cdot r} \tag{2.14}$$

Higher valued *G* space vectors represent higher energy planewaves.

Since *G* can be infinite, a planewave kinetic energy cut off is required in practical DFT applications. Truncation of the planewave kinetic energy discussed further in Section 2.5.4, with respect to pseudopotentials.

2.5.3. *k* points

Calculations in modern computational materials science often employ a method of evaluating integrals in *k* space over a finite number of points, known as *k* points, where the *k* point samples a small region of the Brillouin zone. The Monkhorst-Pack (MP) (Monkhorst and Pack, 1976), implementation of *k* points is the *de facto* standard in modern solid state DFT. In an MP grid, M×M×M *k* points are defined in each of the three *k* space directions. A calculation employing the simplest 1×1×1 *k* point grid is known as a Γ point calculation (Sholl and Steckel, 2009). This type of calculation evaluates the electronic structure (Bloch wavefunction) at the centre of the Brillouin zone only and is therefore in theory the fastest type of periodic DFT calculation. This however, is often unsuitable where a precise description of the electronic band structure is required. An important practical consequence of the relationship between real space and *k* space is the size of the cell in these scenarios. A small cell in real space is large in *k* space, and *vice versa*. If one chooses a MP grid sampling only the Γ point, then the real space cell must be sufficiently large, that the Brillouin zone is small, thus allowing the Γ point integral to sufficiently represent the whole of the Brillouin zone. Work reported in the present thesis frequently makes use of supercells, multiple unit cells to ensure that the Brillouin zone is small enough to justify Γ point calculation. Conversely, where a small unit cell is used, several *k* points may be used to sample a large, inhomogeneous Brillouin zone.

2.5.4. Pseudopotentials and energy cutoff

Pseudopotentials are a simplified description of the nucleus and core electrons, which are not chemically active in most use-cases. Ignoring explicit quantum mechanical description of the core electrons affords significant computational time savings over all-electron descriptions. The description of core electrons is computationally expensive, and it is often prudent to represent only valence electrons fully *ab initio*, because they are responsible for most of the chemical behaviour. Another advantage of pseudopotentials is that they can be parameterised to contain a relativistic correction to the core electrons. These electrons move at speeds that can be comparable to the speed of light when close to the atomic nucleus, but relativistic effects cannot be captured by the time-independent Schrödinger equation; however effective parameters can be included in the pseudopotential (Lewars, 2011).

Pseudopotentials are not truly *ab initio* because they are parameterised using isolated atomic elements; however they are in theory transferable, to all other chemical environments. A key consideration of pseudopotentials is the hardness of the pseudopotential, which is related to the planewave kinetic energy cutoff which must be used to describe the electrons in the atom.

Planewave energy cutoffs are a key contributor to computational expense in periodic DFT calculations. Despite the expense, the planewave energy must always be high enough to represent rapidly oscillating wavefunctions for core electrons close to atomic nuclei. Pseudopotentials reduce the need for high planewave energy cutoffs by removing *ab initio* description of such core electrons, thereby reducing the overall computational expense introduced by high energy planewaves. Soft pseudopotentials such as the ultra-soft pseudopotentials (USPP's) are widely implemented and provide excellent computational efficiency in large systems. Pseudopotentials such as those by Goedecker, Teter and Hutter (GTH) are implemented in the CP2K simulation code and are used frequently in simulations reported in the present thesis.

Another planewave approach is the Projector Augmented Wave (PAW) method (Blochl, 1994). PAW is an all-electron method meaning that core electrons are treated *ab initio*. Though these electrons are frozen for computational efficiency, there is some evidence that *ab initio* description of the core yields greater accuracy when describing the valence states of polar and polarizable systems (Kresse, 1999). Such methods are implanted in the Vienna Ab initio Simulation Package (VASP) and are used frequently in work reported in the present thesis.

2.6. Optimisation and convergence

Numerical optimisation from multi-variable calculus is central to modern computational materials science. Many computational materials science methods amount to finding stationary points on a multi-dimensional function. A key practical consideration when solving KS equations iteratively using the SCF (Section 2.2.7) method, is when to declare convergence, and stop the electronic optimisation, before moving the ions. Convergence is achieved when further manipulation of independent variables leads to no practical improvement of dependant variables. During DFT, the electronic structure (SCF) calculation must converge at each SCF step. At each SCF step, once the SCF cycle is converged, ionic positions are adjusted, and the electronic structure is re-optimised at the updated ionic positions.

One of the most important convergence criteria in *ab initio* simulation of materials is convergence with respect to the potential energy of the system. Convergence with respect to the potential energy means that movements to the nuclei along the electronic PES generate an energy change that is below a predefined threshold. Such convergence indicates that the overall geometry of the system is in a satisfactory minimum on the electronic PES (Section 2.2.4), which is often used to stop the simulation. The DFT calculation may not finish until overall convergence is obtained, where movements to the ions yields no further improvement to the overall energy. Specific algorithms for numerical optimisation of electronic structure and ionic positions are discussed forthwith.

2.7. Electronic optimisation

2.7.1. Introduction

When updating the KS orbitals based on new energy eigenvalues, a DFT program must ensure that the change to the electron density is an improvement to the previous density, (i.e. the energy eigenvalues are being minimised at consecutive steps). The KS SCF scheme is shown in Figure 2.1. DFT programs such as VASP and CP2K employ algorithms to identify how the KS orbitals should be updated at each step, often weighing the benefits of fast vs stable electronic structure optimisation.

2.7.2. Diagonalization

Diagonalization of a square matrix involves transformation to a diagonal matrix, where the diagonal elements are the (energy) eigenvalues. In 'standard' DFT, diagonalization involves finding the energy eigenvalues of the KS Hamiltonian using a Schrödinger equation formulation, which is an eigenvalue-eigenvector problem. Diagonalization in computational

materials science is therefore process of computing the new energy eigenvalues of the KS orbitals to update the density for the next SCF iteration; or in the case of the last iteration, to give the density that minimised the total KS energy (Lewars, 2011). The algorithms implementing this are widely varied according to a trade-off between computational efficiency and numerical accuracy. Most work on the basis of taking small incremental steps along the function, evaluating the function at each step before 'deciding' which direction in which to take the next step. The steepest descent (SD) algorithm is among the crudest of electronic optimisers. It ensures that the next step along the function describing the KS orbital energy does so whilst maximising the gradient of the step (steepest). SD will converge at minima very rapidly, though this often compromises on 'stability' and is not desirable. An *unstable* minimiser means that the algorithm frequently converges at stationary points, rather than the global minimum. The consequence of instability during diagonalization is an inaccurate electronic PES, leading to improbable motion of atomic nuclei. More sophisticated minimisers that do not rapidly converge to the nearest local minimum are often used in DFT calculations. Examples of which are the conjugate gradients algorithm, which is also used for ionic optimisation (Section 2.8), or the Blocked-Davidson minimiser algorithm that has been used when performing VASP calculations throughout work reported in the present thesis (Kresse and Furthmüller, 1996b).

2.7.3. Orbital Transformation

The Orbital Transformation (OT) method (VandeVondele and Hutter, 2003) is unique to DFT quickstep (QS) (VandeVondele et al., 2005) that is implemented in the CP2K code (Section 2.12.3). OT is a fast method that circumvents the computational time expenditure of traditional KS matrix diagonalization, by minimising the KS energy directly with respect to the molecular orbital coefficients (built from atomic basis functions). Minimisation of energy with respect to the molecular orbital coefficients is handled by the Direct Inversion in the Iterative Subspace minimiser (DIIS), a.k.a Pulay method (Pulay, 1980; Pulay, 1982). DIIS extrapolates the function to the next iteration based on a minimisation of error coefficients from the previous step. This method is very fast but is unstable (i.e. the algorithm is unlikely to self-correct improbable minimisation trajectories) due to the dependence on information from prior (potentially incorrect) steps. Therefore, small issues with the electronic PES quickly run out of control, while more stable optimisers such as the aforementioned Blocked-Davidson method, may be able to self-correct, but this added stability comes at the expense of computational speed. CP2K helps to minimise instability of OT/DIIS by the introduction of other parameters, described in Section 2.12.3. OT with any minimiser is unsuitable for use

with metallic systems, where extreme charge redistribution between consecutive iterations causes unsystematic changes to molecular orbital coefficients.

2.8. Ionic optimisation

2.8.1. Introduction

The multi-variable function being minimised in the case of ionic optimisation is the electronic PES, where the potential energy varies simultaneously as a function of each physical DoF in the system. The problem of optimising the PES therefore involves manipulation of the ionic positions, where the global minimum is the ground state equilibrium geometry. There are several optimisation schemes for DFT simulation, though the *de facto*, which has been used throughout work reported in the present thesis, is known as the Conjugate Gradients (CG) method. Moving ions in the way described here is commonly referred to as a geometry optimisation (GO). In a GO calculation, the temperature of the system is 0 K, meaning that the particles do not acquire kinetic energy as the potential energy is minimised. This has the important implication that GO calculations cannot transverse potential energy barriers (maxima). CG type algorithms can often 'escape' potential energy minima based on the assumption that in multi-dimensional hyperspace there is likely to be a low energy barrier in at least one direction. A general principle of multi-dimensional space is that more dimensions mean that each region of space is more closely situated to every other region, because more dimensions mean more possible 'routes' from one region to another. Such a route is known as a *hyperplane*, which are used by CG algorithms to escape stationary points.

2.8.2. Conjugate gradients

Before discussion of CG, it is useful to consider the shortcomings of SD-type algorithms (CG and SD algorithms are available for both electronic and ionic optimisation). In SD, the method seeks to maximise the descent gradient of the function when making consecutive steps. In this type of method, it is likely that the system will converge at stationary points rather than global minimum of a complex functional landscape. Stationary points other than the global minimum may correspond to improbable arrangement of atomic nuclei, because the *phase space* contains all possible configurations. Phase space contains low energy regions which typically correspond to chemical bonding and high energy regions of improbable non-bonding atomic configuration. Thus, the ionic optimiser seeks to find the lowest energy region of phase space, with the global minimum that corresponds to the ground state equilibrium atomic configuration. Optimisation algorithms have a systematic PES trajectory, meaning that the starting geometry (often taken from a database or hand-drawn using visualiser software)

determines the region of phase space in which the optimisation is likely to converge. Starting from ionic positions that are distant in phase space from the equilibrium geometry is unlikely to end up with the global minimum, because the algorithm may converge at stationary points closer to the starting geometry.

Avoiding convergence at the first local minimum encountered is addressed by the CG algorithm. CG does not simply maximise the gradient taken by each step, but instead ensures that every consecutive step is in a PES direction orthogonal to the last. This corrects the 'bias' of the SD algorithm which preferentially moves over regions of phase space regions of extreme PES topology. CG has been shown as an extremely effective optimiser by avoiding getting 'stuck' in improbable PES minima close to the starting geometry, and is said to be a 'stable' optimiser owing to reliable convergence at physically meaningful minima (Sholl and Steckel, 2009). Optimiser stability however comes at the cost of computational speed because even in regions very close to the global minimum, CG will always take orthogonal step directions. Conversely, SD type algorithms will rapidly converge at the global minimum where the geometry is very close to that of equilibrium. There are intermediate algorithms such as the residual minimum method (RMM-DIIS) algorithm, which balance stability and speed.

2.9. Molecular dynamics

2.9.1. Introduction

There are often situations in computational materials science where an optimisation of the ionic positions at 0 K is not an adequate description of the physical system, because the natural motion of atoms at temperatures > 0 K are of high importance. An example of such a situation which is relevant to simulations reported in the present thesis is the representation of liquid water. When simulating water, ionic (static) optimisation to the PES global minimum at 0 K would lead to the solid structure of ice, rather than a liquid, which is unsuitable for representing water at room temperature.

Molecular dynamics (MD) techniques are based on providing some kinetic energy to the system, in correspondence with the desired temperature. The conversion between potential energy and kinetic energy as described by classical theories of physics allow for a system in PES minima to overcome energy barriers in a way that is impossible to systems during static optimisation. MD simulations do not optimise the ionic positions in the system, they simply explore a region of the phase space systematically, as a function of time.

Work reported in the present thesis uses MD simulations where the energy at each step is calculated using DFT. These techniques are therefore as *ab initio* MD (AIMD) or Born-Oppenheimer MD (BOMD) where the PES is a solution to the electronic Schrödinger equation. The terms AIMD and BOMD are used interchangeably throughout the present thesis, because all *ab initio* MD techniques used here employ the Born-Oppenheimer approximation. MD algorithms discussed forthwith are first presented based on the classical theories of physics for which they were first developed. There are significant parallels between classical and *ab initio* MD, with the only major differences existing in the calculation of the energy term, which does not change the general MD algorithm. Specific AIMD considerations are discussed following discussion of the general MD techniques.

2.9.2. Equilibration

Equilibration is a term used throughout MD discussion which refers to the plateauing of MD energies with respect to time. Unlike static simulation techniques where the geometry is optimised, MD simulations do not show convergence in their ionic positions because they are not designed to minimise the energy of the overall system. Equilibration is achieved once the MD simulation has reached a low energy region of phase space with relatively flat PES topology, where the potential energy does not show any further overall trends (ignoring small fluctuations from step-to-step).

Equilibration has no strict mathematical definition and MD simulations do not terminate once equilibration has been achieved. Equilibration is assessed post-simulation once the whole potential energy profile with respect to time has been obtained. Plotting the potential energy vs time allows for identification of equilibration as typically the latter portion of such a plot will be qualitatively flat, provided that the MD simulation has run for sufficient time. Analysis of MD simulations in the present thesis used more sophisticated qualitative analysis, which involved plotting the running averages as a function of MD simulation time. Statistical measures such as variance may be employed for more quantitative analysis.

Well-equilibrated sections of MD simulations are vital, as these regions are used to represent the average energy of the system. A typical MD simulation must be long enough to first achieve equilibration, then run for a sufficiently long period of time to reduce the impact of anomalous or long-timescale-periodic fluctuations in the mean potential energy. When employing extremely expensive techniques such as AIMD, it is common practice to pre-equilibrate the system using a less-expensive computational technique before AIMD is started (Marx and Hutter, 2009).

2.9.3. Periodic boundary conditions in molecular dynamics

MD simulations may be employed both for molecular simulation where there is no periodicity in the models, or for models of periodically repeating systems. Period boundary conditions (PBC) are relevant to all calculations of periodic models reported in the present thesis. They are particularly important during periodic MD simulations containing liquid particles, because of the extra freedom of particles to move between cells (compared to solids). PBC allow for particles approaching the edge of a simulation cell to pass through the barrier and enter the adjacent periodically repeated cell. As the particle leaves the simulation cell, it is reinserted to the simulation cell by the wall on the opposite side of the cell as the same exact process occurs in the periodically repeated image of the cell on the other side (Jensen, 2006).

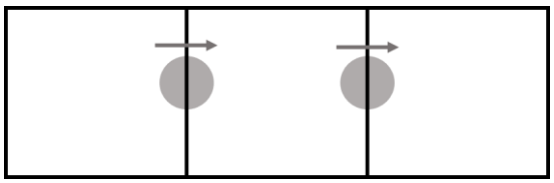

Figure 2.2. Simulation cell showing particle moving through periodic boundary conditions.

2.9.4. Ensembles

2.9.4.1. Microcanonical ensemble

MD simulation in periodic systems involves volume, V, filled with number of atoms N. The evolution of the system with time is described numerically by tracking the atomic positions and velocities. The two quantities used to follow the state of the system are the kinetic and potential energies (Sholl and Steckel, 2009):

$$E_k = \frac{1}{2}\sum_{i=1}^{3N} m_i v_i^2 \tag{2.15}$$

$$F_i = m_i a_i = -\frac{\partial U}{\partial r_i} \tag{2.16}$$

Where $3N$ is the dimensional velocities of each atom, m_i is the mass of the atom at coordinate i and U is the potential energy. Where the total energy ($E_k + U$), the number of particles (N) and the volume of the box (V) are conserved, the system follows the microcanonical ensemble (NVE). At a given temperature (T), the velocities of the particles in the ensemble vary according to a Maxwell-Boltzman distribution, where the average kinetic energy for particles in the system is:

$$\frac{k_{\mathrm{B}}T_{\mathrm{MD}}}{2} = E_{\mathrm{k(ave)}} \tag{2.17}$$

The temperature of the MD simulation is therefore defined by the average kinetic energy of particles in the system. Though the total energy (E_k + U) is conserved, the ratio of kinetic and potential energy varies based on the PES surface position, therefore temperature must fluctuate to accommodate variation in the kinetic energy. When simulating the NVE ensemble, temperature is a variable which equilibrates as the MD simulation evolves.

2.9.4.2. Canonical ensemble

It is often desirable in MD simulation to control the temperature, rather than letting it vary in accordance with the potential energy surface. Many physical properties are highly dependent on the temperature at which they occur (e.g. phase transitions or the kinetics of adsorption processes). The canonical ensemble keeps number of particles (N) and volume (V) constant, as with NVE; however the temperature (T) is now kept constant by allowing kinetic energy (heat) to be exchanged with an external modulator known as a *heat bath*.

Development of the canonical ensemble by Nosé employed *Lagrangian* mechanics, which is a reformulation of standard Newtonian mechanics with improved numerical efficiency. Maximum numerical efficiency is highly important in simulations which already incur a significant computational cost (such as AIMD). The Nosé-Hoover (NH) (Hoover, 1985; Nosé, 1984) thermostat allows efficient modulation of the MD temperature by rescaling particle velocities using a friction term. The rate of particle velocity rescaling is defined as:

$$\frac{d\xi}{dt} = \frac{3Nk_B}{Q}[T_{MD} - T] \tag{2.18}$$

Where the friction term (ξ) is used to rescale the velocities of the MD particles to keep temperature constant, T_{MD} is the instantaneous MD temperature (this will naturally fluctuate before rescaling), T is the target (user defined) temperature for the simulation and Q is a (user defined) parameter that defines the speed at which the temperature is rescaled.

Q is an important parameter in NVT MD simulations because the speed of the feedback between the instantaneous temperature and the target temperature can affect the simulation trajectory. Rescaling the temperature at every timestep is not only computationally expensive but does not allow some flexibility in the kinetic energy which may be required to overcome

energy barriers for access to unexplored regions of the phase space. However, where a simulation can run for too long without thermal regulation, the temperature may increase unphysically, leading to atomic rearrangement. This is particularly important to consider in AIMD simulations, where a high temperature can lead to improbable electronic reorganisation (chemical reactivity).

MD simulation reported in the present thesis was primarily performed within the canonical ensemble and were performed using a NH chain thermostat. The *ergodic* principle in MD is the assumption that the ensemble average is equal to the time average providing that sufficient time has passed (i.e. a sufficient region of the phase space has been sampled for statistical averaging of the properties). The 'average' can refer to any observable time-dependent property of the system, including the potential energy. Standard NH thermostats are known to sometimes introduce non-ergodic sampling of the phase space, which amounts to the simulation becoming 'stuck' in regions of phase space due to the temperature modulation. One popular way this has been rectified is to use a NH *chain* thermostat, where there are typically three nested levels of thermostat. The first heat bath modulates the simulation cell, then the second heat bath modulates the first heat bath, and so on.

2.9.4.3. Isothermal-isobaric ensemble

In the isothermal-isobaric ensemble, number of particles (N), pressure (P) and temperature (T) are constant, while the volume can change. This is particularly desirable in periodic systems where lattice vectors must be equilibrated as well as the atomic positions. By keeping pressure and temperature constant, the volume of the simulation box will approach equilibrium after sufficient MD steps. NPT is often very expensive compared with NVE and NVT ensembles because variation in the cell vectors adds extra degrees of freedom to the problem, as well as expensive calculation of stress tensors. Equilibration of the system with respect to the cell vectors is not only expensive per step, but cell vectors are usually very slow to equilibrate when simulating liquids. This is exacerbated when computing the energy of the system using accurate *ab initio* methods, therefore the canonical ensemble was employed for all AIMD simulation reported in the present thesis. The isothermal-isobaric ensemble was employed using classical MD methods for 300 K equilibration of a calcite supercell, for use in all subsequent AIMD simulation in the canonical ensemble.

In the isobaric-isothermal ensemble, the temperature is controlled identically to the canonical ensemble using a NH thermostat. The pressure is kept constant by a barostat, such as the one developed by Andersen (Andersen, 1980). The barostat works similarly to the thermostat,

employing an extended Lagrangian. The external pressure is rescaled using a fictitious piston which exerts an external force, that partially resists changes to the cell volume. The inertia of the piston is defined by a fictitious mass, which together define the responsiveness of the external pressure in maintaining the internal pressure of the simulation box. This is analogous to the speed of the feedback in velocity rescaling of the thermostat. In practice, the user defines a timescale for the pressure feedback loop, and the program adjusts the mass of the piston accordingly (Marx and Hutter, 2009).

2.9.5. Numerical integration

Application of Newton's second law in its pure form to many particle MD problems cannot give an analytical solution due to the *N*-body problem which has been discussed in preceding sections. To compute forces on atoms at each MD timestep, numerical integration is often the best practical solution. Such a numerical integration involves approximate representation of Equation 2.16 as a truncated Taylor expansion of atomic positions with respect to time, to represent the force (second derivative of particle position with respect to time multiplied by mass) (Sholl and Steckel, 2009):

$$r_i(t \pm \Delta t) = r_i(t) \pm \frac{dr_i(t)}{dt} + \frac{d^2 r_i(t)}{dt^2}\frac{\Delta t^2}{2!} \pm \frac{d^3 r_i(t)}{dt^3}\frac{\Delta t^3}{3!} + O(\Delta t^4) \tag{2.19}$$

Where r_i is position of the i^{th} atom, the first derivative term is the velocity of the i^{th} atom, the second term is the acceleration and the third term is the rate of change of acceleration (aka the *Jerk*), the fourth (constant) term (O) truncates the expansion at the third order derivative. Δt represents the timestep for the MD simulation (Section 2.9.6).

Adding the positive and negative of Equation 2.19 gives the Verlet algorithm for following the positions of particles during MD simulation:

$$r_i(t + \Delta t) \cong 2r_i(t) - r_i(t - \Delta t) + \frac{F_i(t)}{m_i}\Delta t \tag{2.20}$$

Omission of velocity from this algorithm is simplistic, and means low accuracy for the evolution of atomic positions. The velocity Verlet algorithm is based on the standard Verlet method but incorporates both velocity and position at the same value for time (*t*), via the equations (Marx and Hutter, 2009; Sholl and Steckel, 2009):

$$r_i(t + \Delta t) = r_i(t) + v_i(t)\Delta t + F_i(t)\frac{\Delta t^2}{2m_i} + O(\Delta t^4) \quad (2.21)$$

$$v_i(t + \Delta t) = v_i(t) + \frac{\Delta t}{2m_i}[F_i(t + \Delta t) + F_i(t)] \quad (2.22)$$

The velocity Verlet method is employed throughout MD simulations reported in the present thesis.

2.9.6. Timestep

The MD timestep is an important user-defined parameter on which the Verlet integration scheme critically depends. Qualitatively, the timestep is the period of time over which the forces on each atoms may act, before the energy at updated atomic positions is recalculated. A timestep that is too small will typically slow a calculation down, potentially meaning that the simulation will finish before any physically meaningful time-dependent interactions have been represented. A timestep that is too large may mean that atoms moving towards each other that should collide, will instead move 'through' each other before the energy (interaction potential) is recalculated, thereby missing physical atomic interactions. As a rule of thumb, the timestep is usually chosen to be one tenth of the oscillation time period for the fastest vibrational mode in the system. This is usually in the region of 1 fs.

The timestep is an important parameter for balancing computational efficiency with accuracy of the physical representation. This requires careful user consideration, especially when using extremely expensive *ab initio* methods where only a small number of steps may be afforded.

2.9.7. Energy calculation

Generic MD simulation protocols described in previous sections were presented with respect to the classical physics for which they were developed; and the energies defined using classical kinetic and potential energies. Classical physics is convenient for the introduction and discussion of MD algorithms, though classical MD simulations do not feature significantly in work reported in the present thesis.

MD simulations employing DFT (AIMD) are significantly different only in the way that the energy of the configuration is calculated at each step, though Verlet integration and movement of atoms remains classical in AIMD. Here, DFT is used to calculate the energy at the atomic configuration following each timestep, rather than employing classical interatomic potentials. Despite relatively minor differences in the MD algorithm, there is a significant difference in

computing time required when employing classical or AIMD. Modern DFT techniques are effective in accurately describing physical interactions, but computing costs (associated with self-consistent Kohn-Sham method) often prove to be a major practical hinderance. AIMD simulation relies heavily on access to tier 1 and tier 2 high performance computing facilities, meaning that as a technique it is far less accessible than the classical alternatives. Development of the computational efficiency of AIMD is therefore of high research interest.

Classical and AIMD have niche applications and are both used in modern computational materials science. Classical MD has the benefit of being extremely efficient, enabling simulation of large models such as full nanoparticle-solvent systems or large biological proteins. Simulations reported in the present thesis are mainly based on calcium carbonate/water interfaces, which are in principle well parameterised for classical simulation techniques. However, the interaction of calcium carbonate/water with impurities is not well parameterised, particularly in transition metal containing ions. We have therefore employed AIMD which is highly transferrable, and has been exploited far less frequently than classical MD in the computational study of calcium carbonate minerals.

2.9.7.1. Car-Parrinello molecular dynamics

Car-Parrinello MD (CPMD) was the first method of AIMD with sufficient computational efficiency for popular use in computational materials science. CPMD relies on separation of the nuclear and electronic motion, which evolve with respect to time simultaneously. The electronic wavefunction (from KS orbitals) is only converged at the start of the simulation. This is expanded as a Taylor series to give evolution of a fictitious set of nuclear and electronic degrees of freedom with respect to time. Evolution of the fictitious ground state electronic structure yields accurate ionic forces, though the ions to not follow the exact electronic PES because electronic minimisation only happens at the beginning of the simulation. To maintain decoupling of nuclear and electronic degrees of freedom (analogously to the BOA), CPMD requires a particularly small timestep which can slow the simulation significantly.

2.9.7.2. Born-Oppenheimer molecular dynamics

Born-Oppenheimer MD (BOMD) is an alternate method of AIMD to CPMD that has been employed throughout simulations reported in the present thesis. It involves convergent minimisation of the electronic structure at every MD step meaning that ions move accurately, according to the electronic PES (Marx and Hutter, 2009; Sholl and Steckel, 2009). Early implementations involving full diagonalization of the KS matrix were extremely costly and not widely used. The development of DFT quickstep method employing the OT method

(Section 2.7.3) for direct minimisation of the energy with respect to molecular orbital coefficients was intended specifically for fast BOMD in the CP2K code. Modern, fast BOMD is stable when using longer timesteps than CPMD, which combined with direct electronic minimisation means that it is now the *de facto* modern AIMD technique for both speed and reliability.

2.10. Other methods

The majority of simulations reported in the present thesis were performed within the density functional theory. Other, computationally cheaper methods were used in specific scenarios that did not warrant the use of full *ab initio* techniques; particularly for structural refinement or equilibration before employing expensive DFT methods. Though there are no examples of final results generated using these methods, they are discussed here in brief owing to their importance in pre-production runs.

2.10.1. Molecular mechanics

Methods based on interatomic potentials employ classical physics to describe the interaction between atoms which are represented as simple spheres connected by springs. Different chemical bonds are represented by springs with different mechanical properties, which is why these methods are often collectively known as molecular mechanics (MM). Functional chemical groups often contain bonds with mechanical properties that are transferable between chemical environments; for example, the stretching frequency of a carbonyl group is almost always the same, regardless of its chemical setting. It is therefore possible to parameterise the stretching frequency of a carbonyl bond by setting the mechanic properties of this spring. Parameterisations in MM methods are usually based on either experimental or quantum mechanical simulation data, but MM methods cannot work without prior information of the system. MM methods also require explicit bonding information, unlike *ab initio* methods where chemical bonding results from solving the electronic structure. Because there are no electrons in the system, the PES is simply an electronic energy that results from parameterisation of the nuclear coordinates.

In modern practice, the interatomic potential is described by a forcefield (FF) which contains terms describing how the energy varies a function of each mode of distortion (stretch, bend, torsion, intermolecular (vdW and electrostatic) and three-body terms). The simplest form of interatomic potential is a simple harmonic oscillator based on Hooke's law of displacement of a spring from mechanical equilibrium. More sophisticated anharmonic potentials such as the Morse potential have been developed so that at high atomic displacement, the bond-stretch

energy approaches the dissociation energy. Another commonly used approach for describing the intermolecular attraction is a Lennard-Jones type attraction potential, which varies as $1/R^6$ where R is the distance between two atoms that are not covalently bonded (Jensen, 2006). The Lennard-Jones type potential is particularly useful when simulating fluids.

Classical interatomic potential methods can be extremely computationally efficient and can be used to represent systems of many tens-of-thousands of atoms. However, their accuracy relies entirely on the parameterisations used (the chosen FF). Carbonate/water systems are very well parameterised for classical methods, including the Carbonate (Gale et al., 2011) interatomic method as implemented in the General Utility Lattice Program. However, in systems with heavy-metal containing impurities such as molybdate, or charged species with lone pairs such as bromate, there are no effective parameterisations available. Technical difficulties of this kind are why work reported in the present thesis used AIMD instead of classical methods.

2.10.2. Density functional tight binding

Density functional tight binding (DFTB) is an example of a semi-empirical quantum mechanical method which were popular in early quantum chemistry before the widespread use of KS DFT.

Tight binding is a physical approximation originally for calculating band structures of periodic solids by combining localised atomic wavefunctions. DFTB employs a system of parameterisations derived from full DFT calculations to approximate interatomic interactions of these tightly-bound atoms. Atomic wavefunctions of tightly bound atoms are similar to the wavefunction of the atom in vacuum and there is assumed to be minimal overlap between atoms when combined, to give an approximate *crystal Hamiltonian*.

In DFTB the KS DFT energy functional is expanded in a second (or recently, third) order Taylor series containing terms for contributions to the energy, often noted as E^0, E^1 and E^2. A reference density is built on a linear combination of KS orbitals, where the valence electronic contributions from tightly-bound atoms are calculated using a GGA method, represented using minimal atomic basis sets. KS DFT including all interactions up to exchange and correlation contained within the term E^0 are irrespective of charge density fluctuation, i.e. the zero order DFTB KS energy is calculated on a system of tightly bound atoms with no long-range overlap between atoms (Elstner and Seifert, 2014).

The first order term (*repulsive energy*) is fully empirical and is fitted using full DFT at a high level of theory. This term is fully parameterised for distance dependant pairwise interactions for all

atomic kinds in the system, allowing for longer range corrections to the electrostatic potentials calculated in the zero order term. These parameters are contained within the Slater-Koster (SK) libraries. DFTB1 (non self-consistent charge DFTB) contains only E^0 and E^1 (Hourahine et al., 2020).

The second order term is unique to the more modern an effective DFTB2 and DFTB3 methods (a.k.a self-consistent charge SCC-DFTB). Charge density fluctuations (perturbations of the zero order term 'reference density') are calculated using an overlap matrix of atomic Mulliken charges. In the second order term, molecular orbital coefficients are converged self-consistently with respect to the nuclear potentials, for the purpose deriving Mulliken charges at each ionic step. This term introduces important long-range coulombic interactions which are essential when describing molecular systems (Hourahine et al., 2020).

DFTB is often orders of magnitude faster than the equivalent full DFT calculation and can be run effectively on standard Intel Core processors with standard PC memory (c.f. full DFT which often requires high performance computing resources). DFTB is much slower than the equivalent interatomic potentials simulation, though it has several advantages over fully classical methods. Unlike classical methods, DFTB contains a description of the electronic structure, meaning it is possible to simulate photoexcitation (spectra) and band (conductivity) properties. SK files are highly transferrable between chemical environments and are qualitatively simpler than interatomic potentials because they depend only on the two parameterised atoms. Though SK files do not encompass the entire periodic table, libraries such as 3ob or mio contain all organically relevant and many inorganic elements in one single library. This makes DFTB highly user friendly in many cases compared to interatomic potential models where parameters must be carefully chosen in each individual case.

DFTB was particularly useful for pre-equilibration when simulating carbonate systems with sulphate and bromate impurities. There are no interatomic potentials for these full systems, yet all relevant elements are contained within the 3ob SK library.

2.11. Thermodynamics of mixing in solid solutions

In this section, some basic concepts of the thermodynamics governing the behaviour of solid solutions are introduced, to explain in greater detail some of the quantities reported in later Chapters 3 and 5. Solid solutions are a chemical mixture of solids (c.f. a mechanical mixture where there is no chemical bond forming between foreign species). The chemical problem of impurity incorporation in crystalline solids throughout this thesis may be regarded as a solid

solution-type problem, in the limit of low concentrations. We review here some of the elements that are fundamental to work reported in the present thesis.

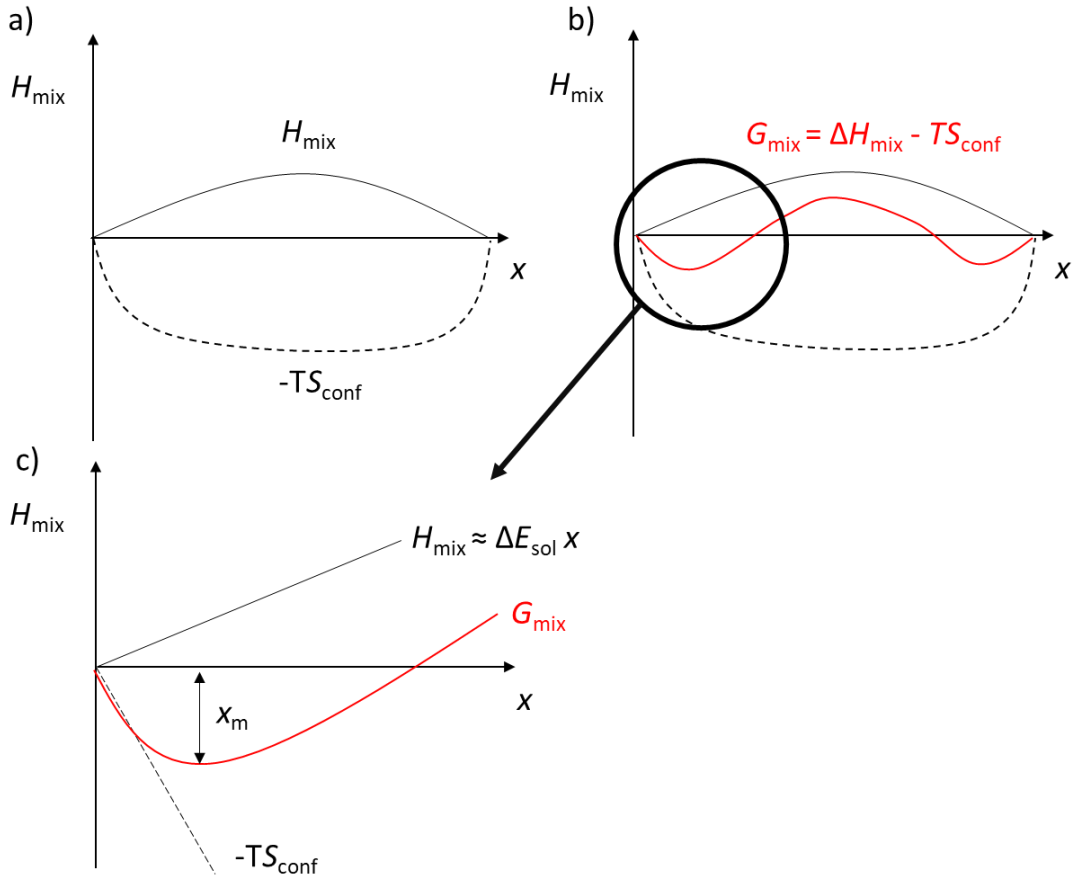

Figure 2.3. Visual representations of the solubility limit x_m as reported in Chapter 3 and 5.

Figure 2.3 represents a simplified diagram of the thermodynamic quantities used for analysis in later chapters. The enthalpy of mixing, H_{mix}, is defined as:

$$H_{\text{mix}} = E[A_{1-x}B_x] - (1-x)E[A] - xE[B] \tag{2.23}$$

Figure 2.3a shows how a typical endothermic mixing enthalpy varies as a function of x, where x is the concentration of foreign species in the solid solution end member. At low concentration, the enthalpy of mixing can be conveniently expressed as a function of the *solution energy*, which is the mixing enthalpy per impurity. Figure 2.3c represents the region of Figure 2.3b that is very close to $x = 0$, which is the only relevant region of the plot for the present application of trace element chemistry. In the region of Figure 2.3c close to $x = 0$, H_{mix} may be approximated as:

$$H_{mix} \approx \Delta E_{sol}\, x \tag{2.24}$$

The entropy term:

$$-TS_{\text{conf}} = k_B T[x \ln x + (1-x)\ln(1-x)] \tag{2.25}$$

is, in reality, only slightly negative; but it has been exaggerated here for clarity. The key feature of the entropy term is that it is very steep in the regions close to $x = 0$ and 1. This means that the free energy of mixing

$$G_{mix} = H_{mix} - TS_{\text{conf}} = \Delta E_{sol}\, x + k_B T[x \ln x + (1-x)\ln(1-x)] \tag{2.26}$$

has always at least a small negative region, close to $x = 0$ (and 1), where the configurational entropy term dominates over the enthalpy term. In this region, we see a negative Gibbs free energy minimum, at the concentration

$$x_m = \exp\left(-\frac{\Delta E_{sol}}{k_B T}\right) \tag{2.27}$$

which *roughly* corresponds to the maximum concentration that is thermodynamically stable against phase separation (roughly, because this is only exact when the enthalpy of mixing is symmetric with respect to x, but otherwise is still a good first approximation).

2.12. Simulation codes

2.12.1. Introduction

Prior sections in Chapter 2 have introduced the methods and theoretical principles that have been used throughout work reported in the present thesis. The following section discusses the simulation codes used, as well as practical considerations and implementations associated with those codes.

2.12.2. Vienna Ab initio Simulation Package

The Vienna Ab initio Simulation Package (VASP) is a commercial DFT code specialising in simulation of periodic systems using plane wave methods (Kresse and Furthmüller, 1996a; Kresse and Furthmüller, 1996b). Work reported in the present thesis made use of version 5.4.4. which was released in April 2017. VASP is highly user friendly with intuitive input file

specification and wide user community support. VASP is often used where numerical accuracy is of primary concern, while computational speed is of close-secondary significance when employing this code. Many important methodological implementations in VASP such as Blocked-Davidson, conjugate gradients electronic and ionic minimisation, and projector augmented wave methods have been discussed throughout the present chapter.

2.12.3. CP2K

CP2K is a free, open-source software package specialising in very fast BOMD simulations, though it has a broad array of simulation features including static/dynamic/meta-dynamic methods, DFT/wavefunction-based *ab initio*/semi-empirical/classical methods and various optimiser algorithms allowing the user to balance numerical accuracy with computational efficiency (Kühne et al., 2020). The primary aim of CP2K is to perform DFT rapidly, which it does by employing unique techniques which differ from most 'standard' DFT implementations such as those found in VASP.

The most significant implementation in CP2K to enhance DFT speed is the orbital transformation direct minimisation method which has already been discussed in the context of electronic optimisation (Section 2.7.3). There are many more features uniquely implemented in the CP2K code, with some key details discussed forthwith.

2.12.3.1. Gaussian and plane waves method

The gaussian and plane wave (GPW) method is a key feature of CP2K enabling fast DFT. The KS one electron wavefunction is built primarily using a basis set of localised gaussian basis functions. Localised basis functions are more computationally efficient than plane waves because they are represented only in small, relevant regions of chemical bonding, rather than being present throughout the simulation cell where integrals are evaluated at finite points (plane waves and k-points). Because the Hartree potential faster to calculate when using a periodic Poisson solver, a plane wave representation of the electron density is also generated (VandeVondele et al., 2005). Gaussian basis functions are mapped to a real space grid where the user defined grid spacing controls the planewave energy cutoff for the auxiliary basis. The Fourier transformation of a gaussian function is also gaussian, therefore gaussians on a uniform grid in reciprocal space produce a periodically repeating planewave, where the energy is defined by the periodicity (grid spacing). CP2K typically uses multiple grids to represent differing planewave energies (four grids by default). These dual-space representations of the KS density matrix are mathematically related, as described in full detail by (Kühne et al., 2020).

2.12.3.2. Basis sets and pseudopotentials

CP2K implements many different types of basis set including both pseudopotential and all-electron (Pople type basis sets). Work reported in the present thesis made use of pseudopotential methods, with the corresponding gaussian basis sets optimised for speed and efficiency. Present simulations employed the DZVP-MOLOPT-SR-GTH pseudopotential, which is explained as follows. Double zeta valence with polarisation functions (DZVP) refers to the general setup of the basis set. The valence shell is split into an 'inner' and 'outer' shell for description of electron density close to and further from atomic nuclei. Polarisation functions (P) allowing s orbitals to polarise with p orbital symmetry and p orbitals to polarise with d orbital symmetry. The so-called 'MOLOPT' (molecule-optimised) basis sets are optimised for systems where there is localised (bonding) electron density. Broadly, MOLOPT is suitable for most problems in computational materials science and chemistry, for elements in the first five rows of the periodic table. The 'SR' (short-ranged) tag refers to low number of diffuse functions, which are computationally expensive and often not required unless refined virtual orbital structure is required. These interactions are typically important when simulating excited state phenomena; however their impact may be quite small in the ground state. SR basis sets therefore ignore long-range ground state electron densities as an approximation to improve simulation speed. GTH refers to the corresponding type of dual-space pseudopotential developed for the CP2K GPW method and intended for use specifically with this for this basis set, by Goedecker, Teter and Hutter (GTH) (Goedecker et al., 1996).

2.12.3.3. Other CP2K details

Preconditioner: is set where the OT-DIIS electronic minimisation is employed, to speed up convergence by expanding the initial KS Hamiltonian in a Taylor series. Preconditioning involves a trade-off between numerical accuracy and computational speed; FULL_SINGLE_INVERSE was used in all AIMD simulations as it presents a reasonable balance and is often used in CP2K AIMD.

Inner and outer SCF loops: The inner loop is the general SCF optimisation algorithm. Here, the KS Hamiltonian eigenvalues are computed at the beginning of the loop by diagonalization, and the preconditioner remains the same until the OT-DIIS iteration has converged. At the outer loop steps, the preconditioner is updated following a new KS Hamiltonian diagonalization.

2.12.4. DFTB+

DFTB+ is a free, open-source code for density functional tight binding simulation (Hourahine et al., 2020). DFTB+ is well optimised for speed and scales well on local multi-threaded machines up to many hundreds of atoms. There are multiple pre-built SK libraries that cover elements across the upper rows of the periodic table making it suitable for quantum chemistry and materials science applications. It is possible to simulate molecular species with implicit solvation environments as well as periodic systems sampled using a Monkhorst-Pack grid.

2.12.5. General Utility Lattice Program

The general utility lattice program (GULP) is a code based on classical interatomic potentials primarily aimed at studying periodic systems (Gale and Rohl, 2003). Though not optimised for speed across multiple processors, GULP has an intuitive input file layout combined with pre-built interatomic potential libraries for organic materials and carbonates.

2.12.6. Site Occupancy Disorder code

The site occupancy disorder (SOD) code assists materials modelling for disordered solids such as alloys and solid solutions (Grau-Crespo et al., 2007). Work reported in the present thesis mainly makes use of just one substitutional impurity, meaning that (in calcite and aragonite) there is only one substitution site. In this case all anion sites are equivalent by symmetry. In systems where there is more than one substitutional impurity, there can be many symmetrically inequivalent sites at which the two impurities could be located relative to each other. In Chapter 5, bromate substitution in the bulk of both calcite and aragonite is investigated. The charge discrepancy introduced by this ion we propose could be balanced by a singly charged cation, which introduces a question of *disorder* (i.e. where these two ions are likely to be located relative to each other).

SOD calculates all substitutional configurations using symmetry operators for the host structure's crystal space group. The number of configurations is often extremely high, which scales with both cell size (number of possible substitution sites) and number of substituting species. The total number of configurations is given by the equation:

$$N_{\text{config}} = \frac{N_s!}{n!\,(N_s - n)!} \tag{2.28}$$

Where N_s is the number of substitution sites in the host structure and n is the number of substituting ions. The probability of any configuration occurring is calculated by a Boltzmann

distribution. The program generates all symmetrically inequivalent configurations, which significantly reduces the computational cost of investigating symmetrical duplicates. Each inequivalent configuration is assigned a weighted average based on the probability that each configuration occurs in the system; this represents the relative abundance of the configuration. SOD pairs directly with GULP and VASP simulation codes to generate structure files for either classical or *ab inito* calculation thereafter.

2.13. Computing facilities

Work reported in the present thesis made use of a high-performance computing tower consisting of a 16 core Intel Xeon Broadwell CPU with 64 Gb RAM, purchased via the SCENARIO NERC DTP (grant ref. NE/L002566/1).

We also made extensive use of high-performance computing facilities as follows:

ARCHER, the UK's tier one national scientific supercomputing facility via membership to the materials chemistry consortium. ARCHER employs Cray XC30 hardware composed of 24 core nodes with dual 12 core Intel Ivybridge processors. InfiniBand parallelisation is implemented with openMPI libraries available for software compilation. ARCHER is the primary research resource used in the present thesis and is essential to running AIMD and DFT on large molecular models where many hundreds of processors are required in parallel.

The Reading Academic Computing Cluster (RACC) has been used to complement ARCHER calculations. Though not InfiniBand ready, the RACC offers a range of hardware with single nodes of up to 24 cores and wall clock times of up to one month.

3. Molybdenum and sulphur incorporation as oxyanion substitutional impurities in calcium carbonate minerals: a computational investigation

3.1. Summary

This chapter reports a computational investigation of sulphate and molybdate incorporated in the crystalline bulk phases of calcium carbonate. Molybdate and sulphate substitution thermodynamics are compared here, with the effect of the carbonate phase density also assessed. This chapter is an adaptation of the publication (Midgley et al., 2020), where S. D. Midgley is the primary author and contributor to the experimental procedures and analyses. Co-authors provided helped with the initial project idea and refinement of the manuscript, but methods and analysis represent the work of the primary author. RGC was involved in calculation of the maximum stable molar fractions. Research data relating to this chapter is available at https://doi.org/10.5281/zenodo.5041725.

3.2. Introduction

Volcanic eruptions are often rich in elements (particularly sulphur) which eventually become incorporated into speleothem calcite. Sulphur is frequently exploited in paleoclimate research as a tracer of geophysical anomalies. In addition to sulphur, the concentrations of molybdenum and bromine in stalagmites have recently been identified as important tracers of volcanic activity (Badertscher et al., 2014), using synchrotron-based X-ray fluorescence techniques. To rationalize the involved correlations, knowledge about atomic-level structural and thermodynamic aspects of the impurity incorporation is needed. Calcite, the most common polymorph of calcium carbonate, forms the majority of speleothems but given certain conditions, aragonite is also stable (De Choudens-Sanchez and Gonzalez, 2009; Given and Wilkinson, 1985; Railsback et al., 1994). Although much rarer, vaterite can also form when the temperature within the cave is low enough, or if specific ion concentrations are prevalent in the aqueous phase forming the solid, but it is known to be unstable with respect to recrystallization to calcite and/or aragonite (Lacelle et al., 2009). Monohydrocalcite $CaCO_3 \cdot H_2O$, (MHC), is a rare hydrous phase of calcium carbonate which has been discovered in cave environments in a few rare cases (Fischbeck and Müller, 1971; Stoffers and Fischbeck, 1974). Ikaite $CaCO_3 \cdot 6H_2O$ is a hydrous phase of calcium carbonate commonly referred to as the "cold weather" phase. It is found in arctic environments, but is rare in caves. Field (Field et al., 2018) reported the discovery of ikaite pseudomorphs in speleothems from the UK, which is the first evidence for ikaite being a constituent of speleothems. This rarity is partially because ikaite is known to be unstable with respect to recrystallisation at temperatures above 10°C. Despite their rarity, the hydrated phases of calcium carbonate are still an important

consideration in our investigation, due to the dynamic nature of calcium carbonate phase transformations over long timescales.

The chemistry of sulphur incorporation in calcium carbonates has been studied in previous work, for example: the isotopic ratios of sulphur in carbonates (Balan et al., 2014), sulphur incorporation as sulphate ions $(SO_4)^{2-}$ in the bulk of different $CaCO_3$ polymorphs (Fernández-Díaz et al., 2010), its effect on the growth of calcite (Vavouraki et al., 2008) and the reactions of SO_2 with the calcite surface (Böke et al., 1999; Malaga-Starzec et al., 2004). It is clear from these studies that carbonate substitution by sulphate is increasingly unfavourable in the order vaterite – calcite – aragonite. In contrast, very little is known about the incorporation of molybdenum in carbonate minerals. By analogy with sulphur, one can assume that incorporation will occur by molybdate substitution at the carbonate site, but since the $(MoO_4)^{2-}$ anion is bigger than the $(SO_4)^{2-}$ anion, the energetics of the substitution can be expected to be quite different.

In this paper, the incorporation of molybdate ions $(MoO_4)^{2-}$ (and $(SO_4)^{2-}$, for comparison) into $CaCO_3$ polymorphs is discussed. Quantum-mechanical simulations are used to elucidate the geometry and thermodynamics of incorporation of these anions in carbonate minerals: calcite, aragonite, vaterite, ikaite and monohydrocalcite (MHC).

3.3. Computational methods

Calcium carbonates and their interactions with other phases have been widely studied using computer simulation methods based on classical interatomic potentials (e.g. (de Leeuw and Parker, 1998; Raiteri et al., 2010; Ruiz-Hernandez et al., 2010; Wang et al., 2011b)). However, because in this and following work it was intended to study a range of impurities for which currently there are no interatomic potentials available, and to gain in accuracy, calculations using quantum- mechanical methods based on the density functional theory (DFT) have been performed, as implemented in the VASP code (Kresse and Furthmüller, 1996a; Kresse and Furthmüller, 1996b). The generalized gradient approximation (GGA) was used for the exchange-correlation term, with the functional by (Perdew et al., 1996a). The projector augmented wave (PAW) method (Blochl, 1994; Kresse and Joubert, 1999) was used to describe the interactions between the cores (including electrons in orbitals up to 1s for C and O, 2p for Ca and S, and 3d for Mo, which were frozen at the atomic reference states) and the valence electrons. The number of plane waves in the basis was determined by a kinetic energy cutoff of 520 eV, which is 30% above the recommended value for the set of PAW potentials used, to minimize Pulay stress errors.

We considered five different calcium carbonate phases. Calcite is trigonal with a space group of *R*-3*c* (167) (Markgraf and Reeder, 1985) and is modelled here using a hexagonal cell. Aragonite is orthorhombic with space group *Pmcn* (62) (De Villiers, 1971). Vaterite is an interesting case as there are several proposed models with different space groups. Its structure has been described using an orthorhombic crystal system with space groups *Ama*2 (40) and *Pbnm* (62) but also using a hexagonal crystal system with space groups $P65_22$ (179), $P3_221$ (154) or $P6_3/mmc$ (194). We chose the hexagonal crystal system with space group $P65_22$, as reported by Wang and Becker (Wang and Becker, 2009), which has the advantage of being fully ordered (no fractional site occupancies). In this structure there are two possible positions for the carbon atom: 6a and 12c, using the Wyckoff notation. Incorporation at both sites has been investigated. The experimental structure for MHC was reported by Swainson (Swainson, 2008), and is described using a trigonal crystal system with a $P3_1$ (144) space group. It contains three inequivalent carbon positions, all possessing the same 3a Wyckoff symbol. We henceforth refer to these three unique positions as 3a1, 3a2 and 3a3. Finally, ikaite is monoclinic with space group *C* 1 2/*c* 1(15) (Hesse et al., 1983), with no non-standard structural features. Table 3.1 below shows that there is good agreement between the calculated and experimental unit cell parameters for each of the five phases studied.

Table 3.1. Relaxed cell parameters and symmetry notation for the unit cell of each polymorph in the bulk.

		Calcite	Aragonite	Vaterite	MHC	Ikaite
Crystal System (space group)		Hexagonal (*R*-3*c*)	Orthorhombic (*Pmcn*)	Hexagonal ($P65_22$)	Trigonal ($P3_1$)	Monoclinic (*C* 1 2/*c* 1)
a/Å	Calc	5.049	5.011	7.290	10.632	8.839
	Exp	4.988	4.961	7.273	10.555	8.792
b/Å	Calc	5.049	8.019	7.290	10.632	8.313
	Exp	4.988	7.967	7.273	10.555	8.310
c/Å	Calc	17.218	5.800	25.450	7.640	10.953
	Exp	17.061	5.687	25.499	7.564	11.021
β/°	Calc	-	-	-	-	108.6
	Exp	-	-	-	-	110.5

When simulating the anion substitutions, supercells were used to minimize the interaction between an impurity and its periodically repeated image. A single $(CO_3)^{2-}$ anion from each

supercell was replaced with a tetrahedral $(XO_4)^{2-}$ anion, similarly to previous work reported in the literature (Arroyo-de Dompablo et al., 2015; Fernández-Díaz et al., 2010). The supercells we used were: 3×3×1 for calcite (54 formula units of $CaCO_3$), 3×2×3 for aragonite (72 formula units), and 2×2×1 for vaterite (72 formula units). For the hydrated calcium carbonate phases (MHC and ikaite), 2×2×2 supercells were used, which led to cell sizes similar to those used for the anhydrous phases. As a preliminary investigation of surface incorporation behavior, we considered anion substitution in the calcite (10.4) plane, which is the dominant and most stable surface termination of calcite (de Leeuw and Parker, 1998; Rohl et al., 2003). The (10.4) surface was modelled using a slab consisting of four layers and a vacuum space of 12 Å; the bottom two layers were fixed to their bulk positions.

The stability of anion substitutions is characterized using two different measures. The first one is the *ion exchange energy*, ΔE_{exch}, which is the energy required to exchange a $(CO_3)^{2-}$ anion with an $(XO_4)^{2-}$ from the gas phase:

$$\mathrm{Ca}_n(\mathrm{CO}_3)_n + (\mathrm{XO}_4^{2-})_{\mathrm{gas}} \rightarrow \mathrm{Ca}_n(\mathrm{CO}_3)_{n-1}\mathrm{XO}_4 + (\mathrm{CO}_3^{2-})_{\mathrm{gas}}$$

$$\Delta E_{\mathrm{exch}} = E[\mathrm{Ca}_n(\mathrm{CO}_3)_{n-1}XO_4] - E[\mathrm{Ca}_n(\mathrm{CO}_3)_n] + \left(E\left[(XO_4^{2-})_{\mathrm{gas}}\right] - E\left[(\mathrm{CO}_3^{2-})_{\mathrm{gas}}\right]\right). \quad (3.1)$$

Here $E[\mathrm{Ca}_n(\mathrm{CO}_3)_{n-1}XO_4]$ is the energy of the calcium carbonate supercell containing one (XO_4^{2-}) anion in a carbonate site; $E[\mathrm{Ca}_n(\mathrm{CO}_3)_n]$ is the energy of the pure calcium carbonate supercell (or hydrated calcium carbonate in the case of MHC and ikaite); $E[(XO_4^{2-})_{\mathrm{gas}}]$ and $E[(\mathrm{CO}_3^{2-})_{\mathrm{gas}}]$ are the energies of the isolated anions, which were obtained in a cell with a positive charge background to restore charge neutrality. Because the density of the background charge depends on the supercell size, the $E[(XO_4^{2-})_{\mathrm{gas}}]$ - $E[(\mathrm{CO}_3^{2-})_{\mathrm{gas}}]$ term of Equation 3.1 was extrapolated to the limit of an infinite-size cell from calculations in supercells of increasing size (Figure 3.1). We use the gas phase (in the limit of isolated anions) as a reference state for computational convenience, and obviously not in connection with any relevant geochemical process. The purpose here is to use a common reference state to compare the stability of the substitution anions across different carbonate phases, including the hydrated ones. Only the relative values of ΔE_{exch} between different phases are relevant; the absolute values do not carry useful information.

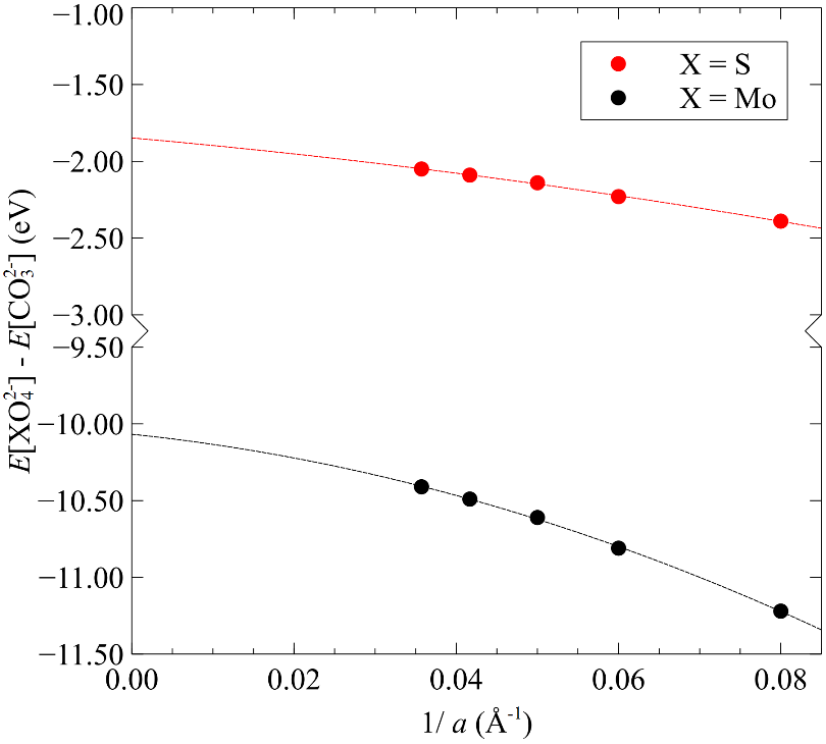

Figure 3.1. Variation of the energy difference $E[XO_4^{2-}]$ - $E[CO_3^{2-}]$ with the inverse of the lattice parameter *a*, and extrapolations to the infinite cell using quadratic polynomials.

The second measure used to evaluate the stability of substitution is the *solution energy*, ΔE_{sol}, which is the energy of substituting the anion with reference to relevant competing phases ($CaXO_4$ solids):

$$\frac{(n-1)}{n}\,\mathrm{Ca}_n(\mathrm{CO_3})_n \;+\; \mathrm{CaXO_4} \;\rightarrow\; \mathrm{Ca}_n(\mathrm{CO_3})_{n-1}\mathrm{XO_4}$$

$$\Delta E_{\mathrm{sol}} \;=\; E[\mathrm{Ca}_n(\mathrm{CO_3})_{n-1}XO_4] - \frac{(n-1)}{n}\,E[\mathrm{Ca}_n(\mathrm{CO_3})_n] - E[\mathrm{CaXO_4}] \qquad (3.2)$$

where $E[Ca_n(CO_3)_n]$ is the energy of the pure calcium carbonate supercell, and $E[CaXO_4]$ is the energy per formula unit of the competing phase: powellite, $CaMoO_4$, for molybdate substitutions, and anhydrous $CaSO_4$ for sulphate substitutions. The ΔE_{sol} energies, which correspond to the solid solution mixing energies (per impurity atom) are not calculated for the hydrated phases, because the end-members of the solid solutions exhibit different degrees of hydration. However, the absolute values of ΔE_{sol} provide some insight on the stability of the impurity incorporation with respect to the formation of competing phases, and therefore on how likely bulk substitution is in each phase.

3.4. Results and discussion

The calculated ion exchange energies are shown in Table 3.2. For sulphate/carbonate substitutions in the anhydrous calcium carbonate polymorphs, the ion exchange energies follow the order vaterite < calcite < aragonite, which is in agreement with previous work by Arroyo-de Dompablo (Arroyo-de Dompablo et al., 2015). This previous work did not consider the hydrated phases. Our calculations of the hydrated phases show that the sulphate/carbonate exchange energies for both MHC and ikaite are of similar values as for calcite. The lowest-energy substitution occurs in the low-density polymorph vaterite, in particular in the 12c sites, where sulphate/carbonate exchange is much more stable than in calcite by ~0.8 eV. On the other side of the ranking, the ion exchange requires much more energy in the high-density polymorph aragonite (2 eV more than in vaterite 12 c sites, and 1.3 eV more than in calcite).

Table 3.2. Calculated ion exchange energies of substitution for sulphate and molybdate anions in different calcium carbonate phases.

	ΔE_{exch} /eV	
polymorph	sulphate	molybdate
calcite	1.15	1.45
aragonite	2.40	3.18
vaterite 6a	0.68	1.10
vaterite 12c	0.37	0.48
MHC 3a1	1.48	1.95
MHC 3a2	1.14	2.13
MHC 3a3	1.10	1.93
ikaite	1.28	2.01

For molybdate/carbonate exchange, the trend is found to be like the one observed for sulphate/carbonate exchange, with the order vaterite < calcite < aragonite among the anhydrous phases, and the 12c site of vaterite being the most favourable. Substitution in calcite and aragonite is less stable than in vaterite (12c) by 1 eV and 2.7 eV, respectively. In this case the exchange energies at the hydrated phases MHC and ikaite are intermediate between those at calcite and at aragonite, but closer to the former. Figure 3.2 provides a graphic illustration of the similarity between sulphate and molybdate exchange energies. There is an

approximately linear correlation between the two sets of values. However, exchange energy differences between phases are larger (by ~44%) for molybdate than for sulphate substitution. This can be rationalised considering the larger size of the molybdate anion, which is therefore more difficult to accommodate in the smaller (and flat) carbonate sites, in hosts with higher density.

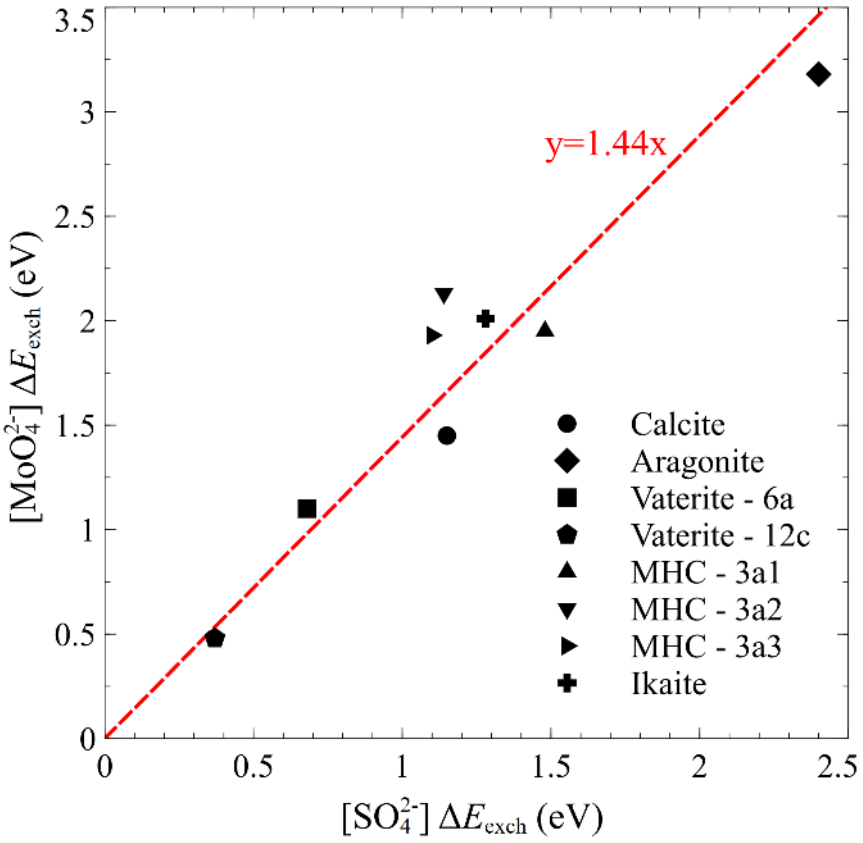

Figure 3.2. ΔE_{exch} energies of molybdate vs. sulphate substitutions in calcium carbonate phases.

The density of the host phase is clearly a key factor determining the relative substitution energies of large tetrahedral anions in the carbonate sites across calcium carbonate phases of the same composition. However, when including the hydrated phases in the comparison, trends cannot be established, because these phases have different chemical compositions and cannot be placed in the same density scale as the others. The densities of MHC and ikaite are much lower than those of the anhydrous phases because they have space occupied by lighter atoms. Therefore, the stability of the substitution must be rationalised in more general terms, for example, as function of the space available around the anion and the interactions with neighbouring atoms.

In Figures 3.3-3.8 we visualize the geometric environments around each type of impurity in each of the carbonate phases, including the most relevant interatomic distances and angles.

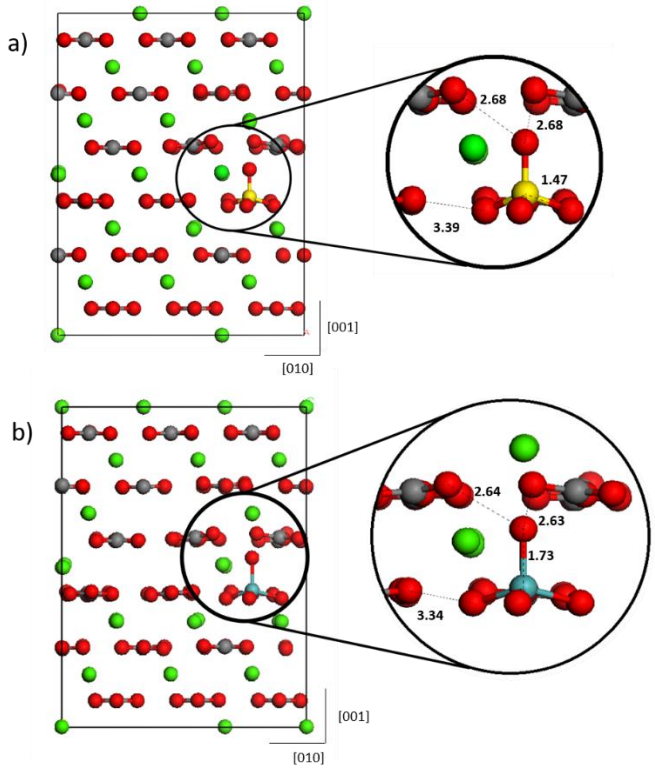

Figure 3.3. Optimized geometries for the incorporation of **a)** sulphate and **b)** molybdate oxyanions into the bulk of the calcite supercell. Here, X-O and O-O bond lengths represent distance in Å.

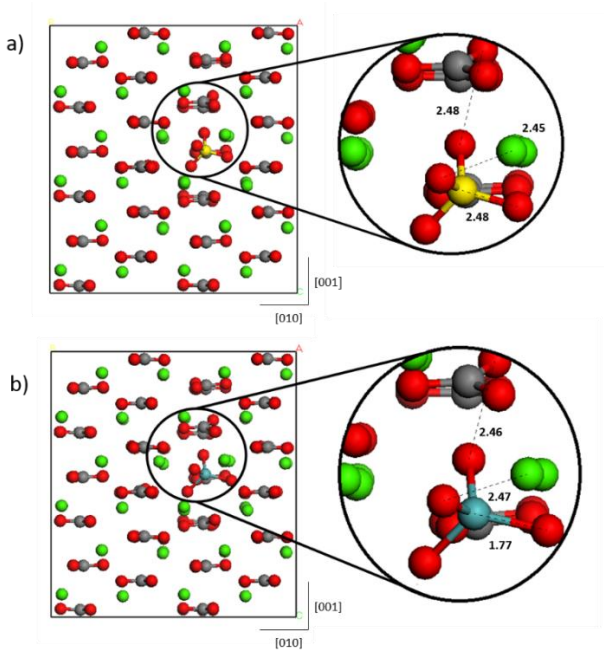

Figure 3.4. Optimized geometry for the incorporation of **a)** sulphate and **b)** molybdate oxyanions into the bulk of aragonite. Interatomic distances are represented in Å.

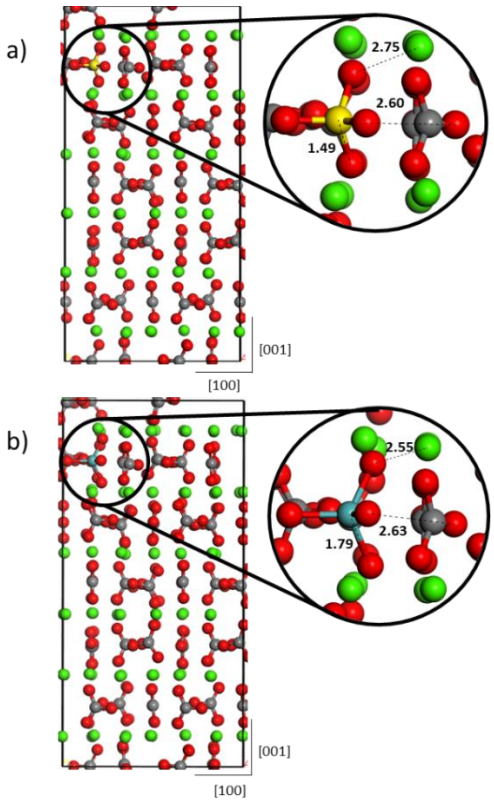

Figure 3.5. Optimized geometry following incorporation of **a)** sulphate and **b)** molybdate into the bulk of vaterite at the 12c carbonate sites. Interatomic distances shown are in Å.

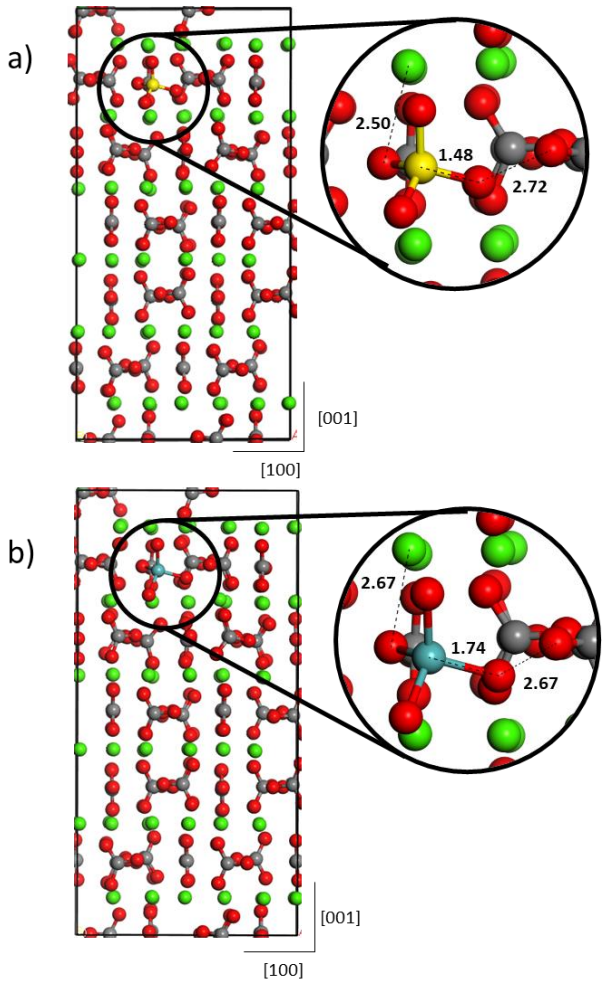

Figure 3.6. Optimized geometry for the incorporation of **a)** sulphate and **b)** molybdate into the bulk of vaterite at the 6a carbonate sites. Interatomic distances shown are in Å.

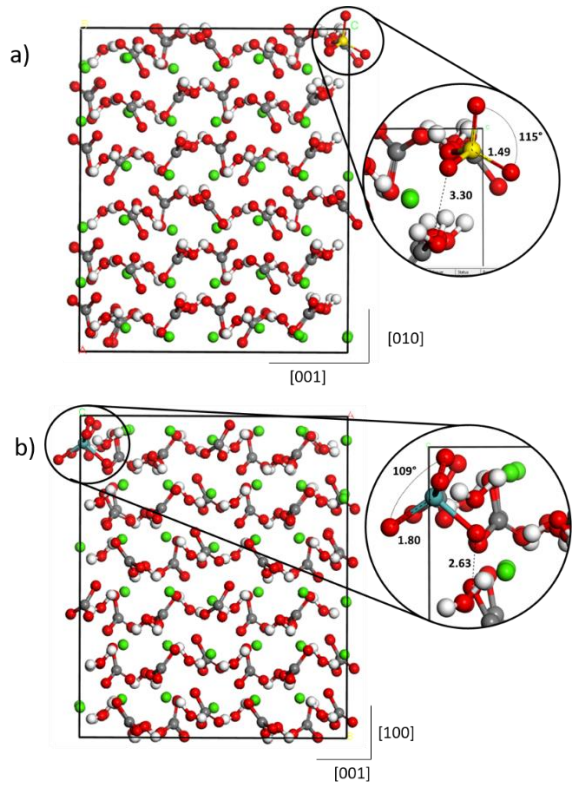

Figure 3.7. Optimized geometry following incorporation of **a)** sulphate and **b)** molybdate incorporation into MHC at carbonate site 3a1. Interatomic distances are shown in Å.

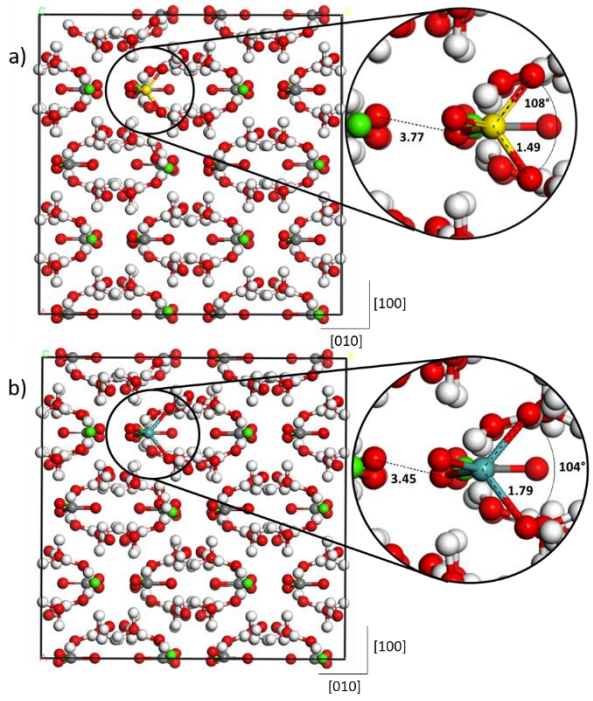

Figure 3.8. Optimized geometry following incorporation of **a)** sulphate and **b)** molybdate incorporation into the bulk of ikaite. Interatomic distances are shown in Å.

Elastic effects may be seen visually in Figures 3.3-3.8 as disturbances of local ionic environment surrounding the anion impurity. The most obvious pattern that emerges is the inverse correlation between the shortest distance from impurity oxygen to carbonate oxygen, and the energy cost of anion exchange (Figure 3.9). These trends make sense on the basis that the energetic cost of the substitution mainly comes from accommodating the fourth oxygen in the tetrahedral ion, which suffers the strongest electrostatic repulsion from the carbonate oxygen atoms. In general, values for molybdate are shifted left with respect to the corresponding values for sulphate, indicating that it is more difficult for the molybdate impurity to escape the repulsion from the host carbonate anions. The plot in Figure 3.9 also rationalises the intermediate values of the exchange energies at the hydrated phases, for which a correlation with density cannot be established: the shortest impurity oxygen-carbonate oxygen distances for the hydrated phases are indeed intermediate between those of calcite and aragonite. The simple correlation in Figure 3.9 is; however unable to account for the different behaviour of sites 6a and 12c of vaterite (the latter are the squares at around 2.6 Å, which are the obvious outliers in the trend). We have examined this case in more detail by calculating the electrostatic-only energy (using quantum-mechanically derived charges on the ions) but including not only the shortest O-O distance but all the carbonate neighbours within a certain radius. Such analysis reveals that the preference for site 12c over 6a is still electrostatic in origin, it just requires the consideration of the full geometric environment around the impurity ions rather than just the shortest distance.

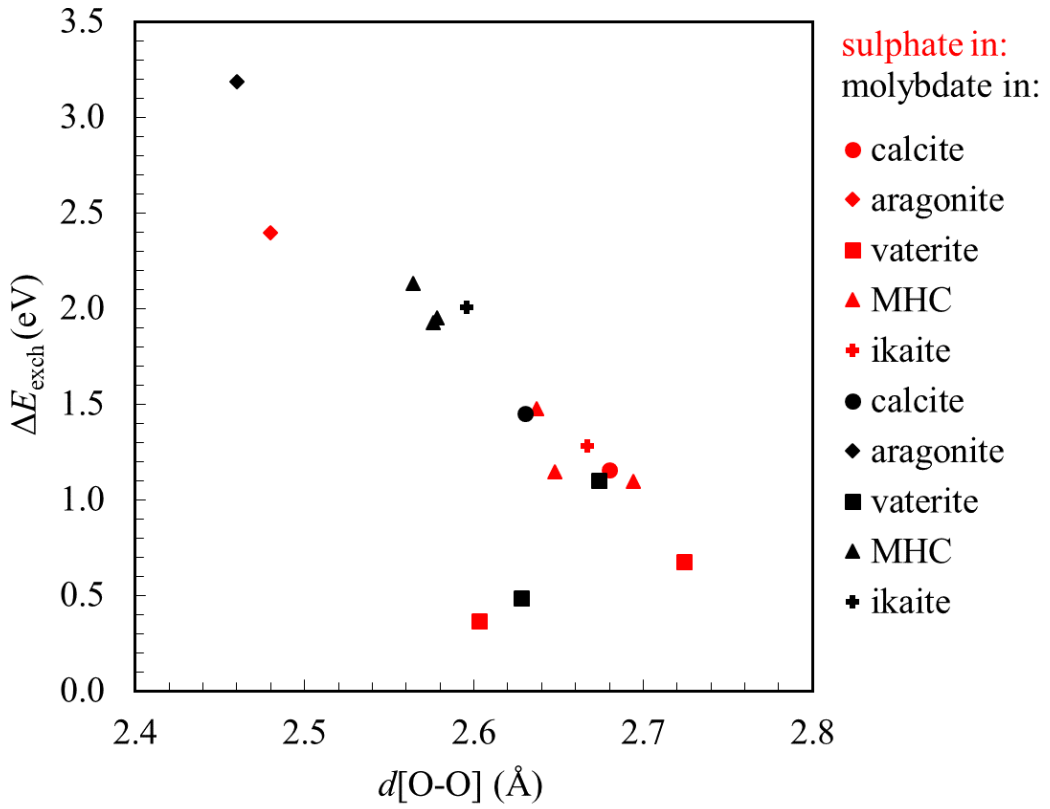

Figure 3.9. ΔE_{exch} energies of molybdate and sulphate substitutions in calcium carbonate phases versus the shortest distance from impurity oxygen to carbonate oxygen in the optimized structure.

The exchange energies reported above are obtained with respect to an arbitrary (but convenient) reference state of the impurity ions. That allowed for comparison across polymorphic phases, but the absolute values do not contain any information on stability with respect to competing phases. The energies of incorporation of sulphate and molybdate into calcium carbonate was also calculated with respect to a competing solid phase (calcium sulphate or calcium molybdate), and the results are shown in Table 3.3.

Table 3.3. Solution energies (ΔE_{sol}) and maximum equilibrium molar fraction (x_m) for S or Mo incorporation as oxyanions in calcium carbonate phases. The reference endmember phase is powellite $CaMoO_4$ for molybdate and anhydrous $CaSO_4$ for sulphate anions. For vaterite, x_m values are calculated assuming that only 12c sites are occupied.

	ΔE_{sol} /eV		$x_m = \exp(-\Delta E_{sol}/k_B T)$	
polymorph	sulphate	molybdate	sulphate	molybdate
calcite	1.04	2.05	3×10^{-18}	4×10^{-35}
aragonite	2.41	3.91	3×10^{-41}	2×10^{-66}
vaterite 12c	0.29	1.13	9×10^{-6}	7×10^{-20}
vaterite 6a	0.60	1.76	-	-

The positive ΔE_{sol} values indicate that for both sulphate and molybdate anions it is energetically preferable to form the competing phases rather than to incorporate as impurities in the bulk of calcium carbonate phases. At any temperature, a small amount of impurities can still be incorporated in thermodynamic equilibrium thanks to the configurational entropy contribution. However, the solution energies are so strongly positive that the maximum stable molar fraction (x_m) is extremely small. To illustrate this point, values for x_m, assuming isolated substitutions in bulk, can be estimated from the minimum of the free energy:

$$\Delta G = x\Delta E_{\mathrm{sol}} + k_B T\{x\ln(x) + (1-x)\ln(1-x)\} \tag{3.3}$$

where k_B is Boltzmann's constant and T is the temperature (we use ambient temperature of 300 K in our analysis). We ignore vibrational and pressure-volume effects here as they are typically much smaller contributions to the free energy. The minimum of the configurational free energy is at:

$$x_{\mathrm{m}} = \frac{\exp(-\Delta E_{\mathrm{sol}}/k_{\mathrm{B}}T)}{1+\exp(-\Delta E_{\mathrm{sol}}/k_{\mathrm{B}}T)} \approx \exp(-\Delta E_{\mathrm{sol}}/k_{\mathrm{B}}T) \tag{3.4}$$

and the values are shown in Table 3.3. Only in the case of sulphur impurities in vaterite, a significant level of substitution (~9 ppm) is thermodynamically stable against separation into the competing phase. Values for molybdate substitutions are many orders of magnitude higher than for sulphate substitutions, reflecting the higher energetic cost of accommodating the larger molybdate anions in the calcium carbonate lattices.

Of course, actual sulphate incorporation in carbonate phases can be much higher than those obtained from the above analysis: they are in the parts per million for non-biogenic calcite and aragonite (Busenberg and Niel Plummer, 1985) and in the parts per thousands in biogenic carbonates (Kampschulte et al., 2001). Our analysis show that sulphate incorporation by substituting carbonate anions in bulk lattice sites must account for a very small fraction of the total concentrations measured. Some level of metastable lattice substitution is possible, as sulphate impurities can enter the lattice from the solution from which the carbonate precipitates and become trapped there by kinetic barriers to diffusion, even if there is a thermodynamic preference to form a competing phase. But based on the calculated stabilities, we conclude that other incorporation pathways, including surface substitution/adsorption or stabilization via defects or other impurities in the bulk, should account for most of the detected sulphate ions in carbonates. In the case of molybdate, this conclusion is even stronger, because lattice $(MoO_4)^{2-}/(CO_3)^{2-}$ substitutions at the bulk are prevented by a huge energy cost.

As an example of how much easier it would be to have sulphate or molybdate impurities at the surfaces than at the bulk of the carbonates, we have also calculated the energetic cost of $(XO_4)^{2-}/(CO_3)^{2-}$ substitutions at the (10.4) surface of calcite. In this case, the fourth oxygen in the tetrahedral anion can protrude from the surface, and therefore be accommodated farther from the repulsion of other oxygen species (Figure 3.10), leading to lower strain. The corresponding solution energies, i.e. the energetic costs of bringing the impurity anion from a competing phase to the surface, are 0.30 eV for sulphate and 0.44 eV for molybdate, which are 20-30% of the bulk values.

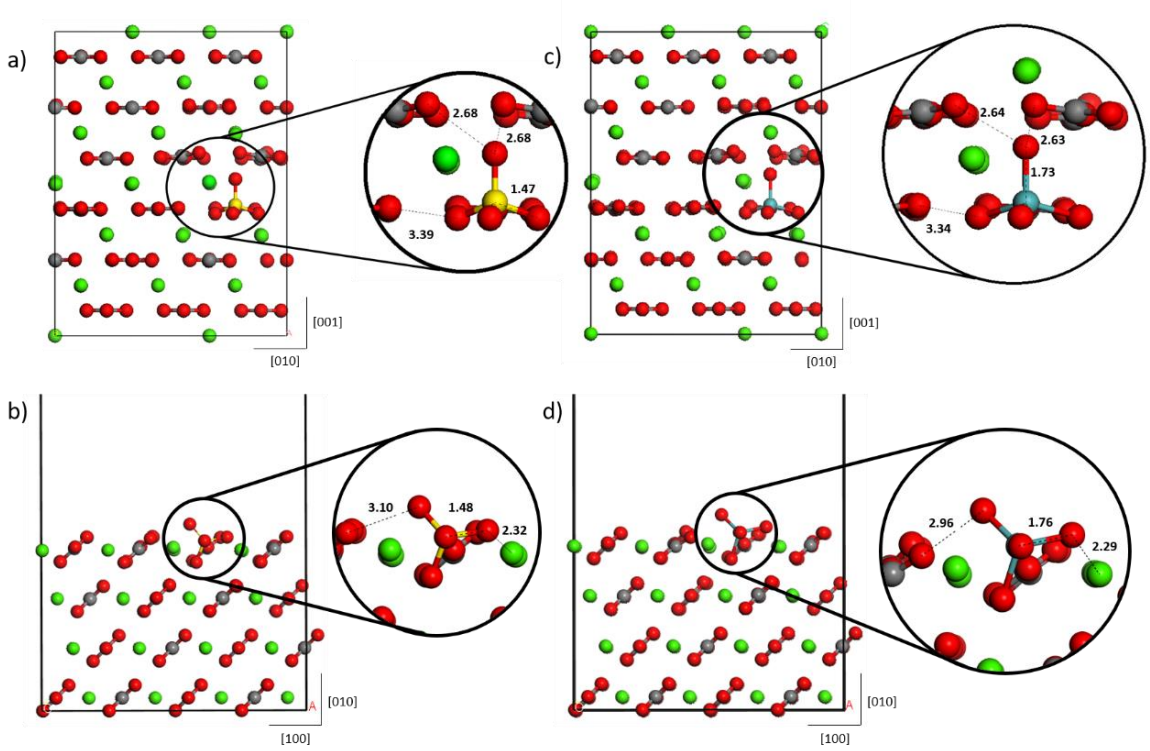

Figure 3.10. Comparison of the geometry of incorporation sulphate (a, b) and molybdate (c,d) anions in the bulk (a, c) and (10.4) surface (b, d) of calcite. Interatomic distances are represented in Å.

3.5. Conclusions

We have presented a computer simulation study, based on density functional theory, of the incorporation of sulphate and molybdate impurities in calcium carbonates. The trend in incorporation stability of molybdate anions in anhydrous calcium carbonate phases is similar to the one obtained here, and in previous work, for sulphate anions (vaterite > calcite > aragonite). The 12c site of the $P65_22$ structure of vaterite is the most stable substitution site for both sulphate and molybdate anions. We have also included for the first time the hydrated phases as possible hosts for these ions, and we show that they exhibit $(XO_4)^{2-}$ / $(CO_3)^{2-}$ exchange energies which are intermediate between those of calcite and aragonite. We found that there is a strong linear correlation between the exchange energies for molybdates and sulphates. While the correlation between the stability of a tetrahedral impurity and the density of the host is confirmed for the anhydrous phases, it cannot be extended to the hydrated phases which have much lower density. We demonstrate here that a measure of the space available in the site for substitution, in this case the shortest distance from impurity oxygen to carbonate oxygen, can be used instead to rationalize the order of stabilities of the impurities across all the phases, included the hydrated ones.

We have also obtained the solution energies as a measure of stability against the formation of a competing phase (calcium sulphate or calcium molybdate, respectively), and found that they are strongly positive, much more so for molybdate than for sulphate. This is consistent with the larger size of the molybdate ion, which makes it more difficult to accommodate a fourth oxygen atom at the host site. We estimated the maximum molar fraction that is thermodynamically stable against phase separation and found them to be negligible (below typical detection limits) in all cases except for sulphate in vaterite. Whilst some level of metastable lattice substitution is possible, based on the low impurity stability we conclude that other incorporation pathways, including surface substitution/adsorption or stabilization via defects or other impurities in the bulk, must account for most of the detected sulphate ions in carbonates. In the case of molybdate, this conclusion is even stronger, because the stability of the impurities with respect to phase separation is much lower. Preliminary calculations of sulphate and molybdate substitutions at a calcite surface (the (10.4) termination) indeed find that solution energies are 20-30% of the bulk values.

Results suggest that the magnitude of incorporation of trace amounts of sulphur and molybdenum in speleothems is critically affected by factors like the crystallinity of the carbonate phases present. It is likely that in samples with lower crystallinity (higher surface area), a higher concentration of trace-element impurity may be detected, regardless of the magnitude of the paleo-volcanic activity responsible for the atypical trace-element abundance at a particular time. Future investigations of the chemistry of sulphur and molybdenum incorporation into calcium carbonate speleothems should probably focus on surface and grain boundary incorporation rather than on lattice bulk substitutions.

3.6. Reflection

3.6.1. Limitations

The present chapter employed state-of-the art DFT simulations which were well received by reviewers for *Chemical Geology*. The methodology is well-formed, and the primary limitation of the chapter is in the approximations used to model the physical system. The main approximation used in this chapter is the substituting ion used to calculate the exchange energies, which was modelled as being isolated in vacuum. While this effectively served as a common reference phase for energy comparison across calcium carbonate phases, the natural process does not involve vacuum space, but instead involves ion exchange from an aqueous phase. It is possible that small, yet meaningful thermodynamic contributions could come from

the liquid water, meaning that a better physical model should involve explicitly represented water molecules at the carbonate surface.

3.6.2. Further work

Improving the physical accuracy of the ion in vacuum reference phase could be addressed to provide improved exchange energy calculation. Explicit representation of liquid water at the calcium carbonate terminus using molecular dynamics methods is an effective way to achieve this, as well as providing a non-zero Kelvin temperature environment, which is more reflective of the standard geological conditions which we are trying to capture.

Investigation of sulphur and molybdenum incorporation in calcium carbonates should focus on surface substitution, rather than the crystalline bulk, following conclusions discussed in this chapter. These ideas are developed in work reported in Chapter 4.

Other tracer elements such as bromine have been experimentally identified, but there exists very little atomic-level work on carbonate incorporation for bromine-containing tracers. Bromate therefore makes an interesting case study of a paleovolcanic trace ion which is geometrically different to sulphate and molybdate. These ideas are developed in work presented in Chapter 5.

4. Sulphur and Molybdenum Incorporation at the Calcite-Water Interface: Insights from Ab Initio Molecular Dynamics

4.1. Summary

This chapter is a logical extension of work reported in Chapter 3. Here we investigate the incorporation of sulphate and molybdate in the bulk of calcite, at the (10.4) terrace and at a step commonly found in the (10.4) terrace. The reference phase used here is the ion in a box of water. We also include water at the surfaces of calcite, mimicking the edge of a calcite particle as found in nature, which represents a more sophisticated structural model than was used in Chapter 3. This chapter is an adaptation of the article (Midgley et al., 2021a), which is published in *ACS Earth and Space Chemistry*. Here, S. D. Midgley is the primary author and contributor to the experimental procedures and analyses. Co-authors provided helped with the initial project idea and refinement of the manuscript, but methods and analysis represent the work of the primary author. Research data for this chapter is available via Zenodo, at https://doi.org/10.5281/zenodo.5041442.

4.2. Introduction

The calcium carbonate in speleothems is mainly in calcite form, with aragonite also present in some cases (De Choudens-Sanchez and Gonzalez, 2009; Given and Wilkinson, 1985; Railsback et al., 1994). Other calcium carbonate phases can also be found in speleothems but are very rare. Recently, we used computer simulations to study the lattice incorporation of both sulphur and molybdenum in the form of tetrahedral molybdate anions, in all the naturally occurring bulk phases of calcium carbonate (Midgley et al., 2020). We found that both sulphate and molybdate anions are unstable in the bulk of calcite, aragonite or vaterite, with respect to the formation of naturally occurring competing phases. In the case of sulphate ions, only the least dense vaterite polymorph provides a somewhat stable host for lattice substitution at low concentrations. A preliminary calculation of substitution at the calcite (10.4) surface (ignoring dynamic aspects and assuming a simple interface with vacuum) showed that substitution at the surface, where there is more space to accommodate the tetrahedron anion, is energetically more favourable than in the bulk (Midgley et al., 2020). If indeed trace elements are mainly incorporated at the surface or grain boundaries, then the crystallinity of the calcium carbonate in the stalagmite will directly impact detected trace-element concentrations. This effect is important to interpret the speleothem record and requires a more careful and detailed investigation, considering some of the complexity of mineral surfaces and interfaces, which is the purpose of the present work.

The (10.4) surface of calcite investigated here is very stable and typically the most prominent in natural calcite particle morphology (Blanchard and Baer, 1992; Didymus et al., 1993; MacInnis and Brantley, 1992), making it a good representative surface for the present study. Calcite growth occurs through the formation and motion of monolayer steps at the surfaces (Gratz et al., 1993; Liang et al., 1996b). This growth is well known to be inhibited by the incorporation of cation impurities (most notably Mg^{2+}) at the steps (Davis et al., 2000; Meyer, 1984; Nielsen et al., 2013; Wasylenki et al., 2005b). Meyer and subsequent citing articles have indicated that the sulphate anion is also a calcite growth inhibitor (Meyer, 1984). It is therefore clear that these steps can provide a more favourable environment for the incorporation of impurities than the bulk or the perfect surfaces (Paquette and Reeder, 1995; Staudt et al., 1994).

We report here density functional theory (DFT) molecular dynamics simulations of the incorporation of sulphate and molybdate oxyanions both at terrace and step sites of the calcite (10.4) surface in contact with water. There is some experimental evidence for the presence of OH^- groups from water dissociation at calcite surfaces (Kuriyavar et al., 2000; Neagle and Rochester, 1990; Stipp and Hochella Jr, 1991), and recent theoretical work based on cluster models has considered the interaction of H^+ cations with calcite models at low pH conditions (Ghatee and Koleini, 2017). However, the nature and location of the dissociated species is still debated, and DFT computer simulations indicate that water adsorbs without dissociation at the calcite (10.4) terrace and steps, requiring charged defects such as anion vacancies to dissociate (Lardge et al., 2010). Therefore, in this work we will only consider the interaction of the (10.4) surface with water in a non-dissociative scenario.

4.3. Computational methods

The bulk phase of calcite, a trigonal crystal with space group of *R*-3*c* (167), (Markgraf and Reeder, 1985) was modelled using a 3×3×1 supercell of the hexagonal unit cell, to minimise interactions between trace elements and their periodic images (minimum distance between images around 15 Å). Slab models were used to represent the calcite/water interface. The slab for the (10.4) perfect surface consisted of four $CaCO_3$ layers in the crystallographic direction perpendicular to the surface, which is the typical thickness used in the DFT simulation of this surface (Andersson et al., 2016; Hafshejani et al., 2021). Using static optimisation test calculations, we check that increasing the thickness to six molecular layers, changed the substitution energy for sulphate by only ~0.1 eV. Parallel to the surface, the model is infinite (periodic in 2D), and a 2×3 supercell of the rectangular surface unit cell was used, which led

to a model with 48 $CaCO_3$ formula units for the pure calcite slab, and a minimum lateral distance between impurity images of around 15 Å in the substituted slabs. The gap between slabs was filled with water molecules at a density of 1 g cm^{-3}. An example of the structural model generated for the perfect (terraced) interface is shown in Figure 4.1.

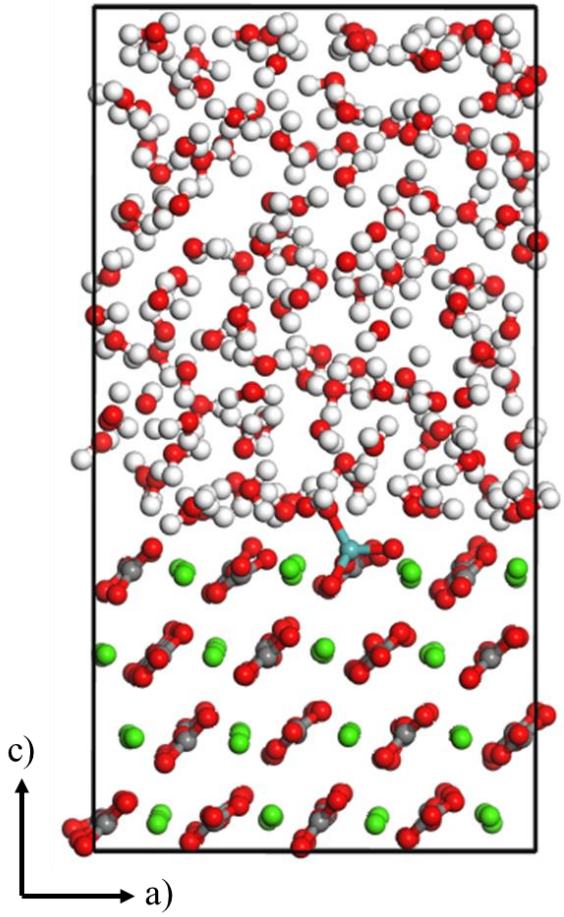

Figure 4.1. Terraced structural model containing molybdate impurity.

In order to illustrate the role of surface defects on the incorporation of impurities, we also considered 'imperfect' or stepped surfaces, to represent formation of pits on the calcite (10.4) growth termination. To represent this in the (10.4) surface models, we created a vicinal (31.8) surface, giving an acute step in the standard (10.4) terrace. The model has similar thickness and lateral dimensions as that of the terrace, with 52 $CaCO_3$ formula units in the case of the pure calcite slab. In addition to the acute step, it would be possible, in principle, to consider other point defects (cation or anion vacancies) or line defects, such as obtuse steps, to understand their role in oxyanion impurity incorporation. However, the high computational cost of the ab initio molecular dynamics (AIMD) technique used in this work limits the number of defect models that can be practically simulated. Obtuse steps, which are not

considered here, might in fact be able to accommodate some impurities more easily than acute steps (de Leeuw, 2002), but also exhibit more complex behaviour due to faster kinetics of dissolution (de Leeuw et al., 1999). We therefore limit our examination of the role of defects to the case of the acute step. A future detailed investigation of the effect of specific surface defects on the incorporation of these impurities will require the implementation of more efficient simulation approaches, either via the extension of existing force fields (Archer et al., 2003; Jackson and Price, 1992; Pavese et al., 1992) to include molybdate and sulphate species, or perhaps by using modern on-the-fly machine-learning force field generation from AIMD (Jinnouchi et al., 2019).

Three compositions were investigated in each slab model: pure calcium carbonate, calcium carbonate with $(SO_4)^{2-}/(CO_3)^{2-}$ substitution, and calcium carbonate with $(MoO_4)^{2-}/(CO_3)^{2-}$ substitution. In each substituted case, a single carbonate ion was replaced with the relevant trace-element containing oxyanion. At the perfect surface, the substitution was performed at the top layer, at the solid/water interface. At the stepped surface, we considered substitution only at the apex of the step as this has most space to accommodate the ion impurity, maximising the thermodynamic benefits of this surface topography. The vicinal surface cleavage left two symmetrically inequivalent sites at the apex of the step, which we called C1 and C2 arbitrarily. Initial investigation revealed negligible energetic differences between these two anion sites, therefore we proceeded to consider just the C1 site. An example of a structural model generated for the acute stepped surface is shown in Figure 4.2.

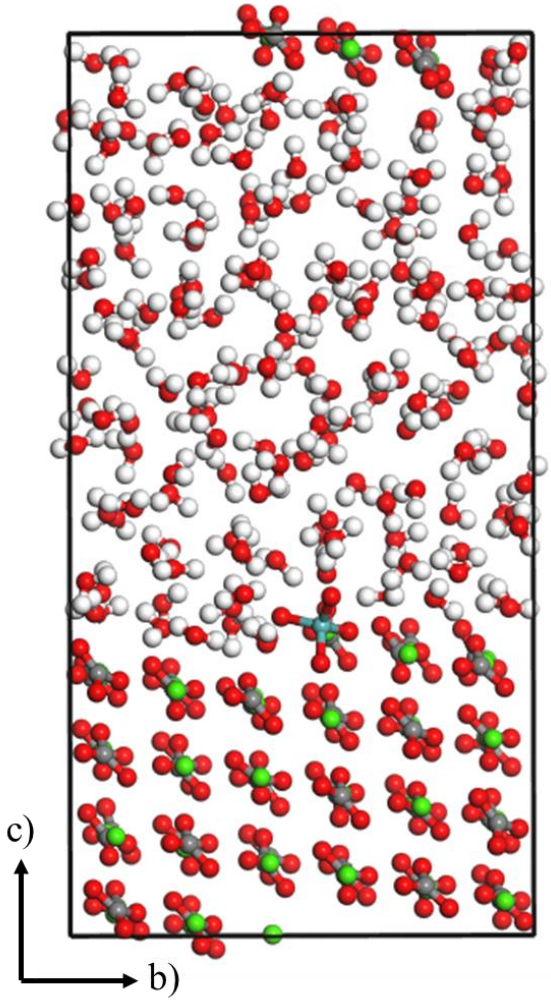

Figure 4.2. Acute stepped surface structural model containing molybdate impurity.

As a reference phase, each anion (carbonate, sulphate and molybdate) was placed into a ca. 15 Å box containing water molecules at a density of 1 g cm^{-3}. Each anion in a box of water introduces a double negative charge to simulation cell; this was corrected using a homogeneous charge background. An example of a structural model generated as a reference phase is shown in Figure 4.3.

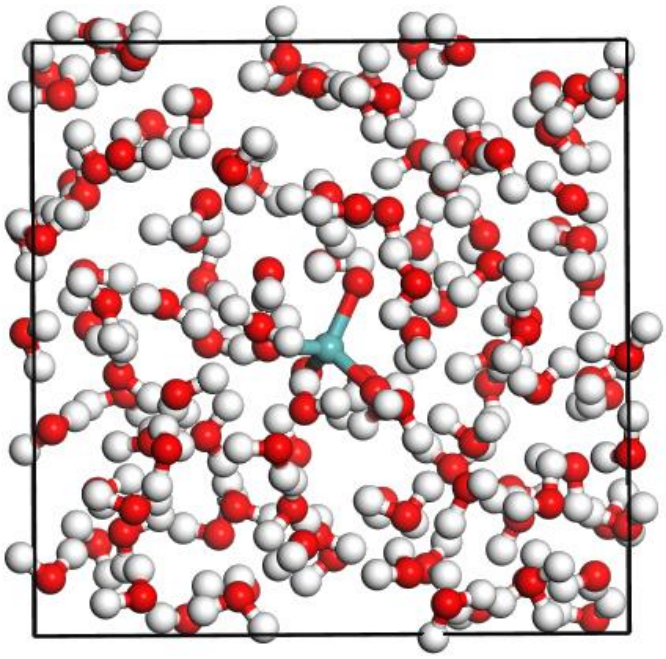

Figure 4.3. Molybdate ion in box of water (reference phase).

The ab initio molecular dynamics (AIMD) simulations were performed at constant volume and temperature using the canonical NVT ensemble implemented in the CP2K software package (https://www.cp2k.org/). To find the equilibrium volume and cell parameters at 300 K, while avoiding a computationally expensive constant-pressure (NPT) simulation at *ab initio* level, we followed a two-step procedure. First, we optimised the bulk structure statically (i.e. at 0 K and ignoring zero-point effects) using CP2K. To calculate the temperature-correction to the 0 K cell parameters, we then performed NPT classical molecular dynamics (MD) simulations at 300 K, using a calcite 3×3×1 bulk supercell. The classical simulations were carried out using the GULP code (Gale and Rohl, 2003) with the carbonates forcefield library by (Fisler et al., 2000). Cell vectors simulated in GULP at 0 K and 300 K were used to calculate the average linear thermal expansion from 0 K to 300 K, which was then applied as a correction to the cell vectors optimised at 0 K using the CP2K code. Table 4.1 shows that the predicted cell parameters are in reasonably good agreement with crystallographic measurements. The negative thermal expansion in the *a*-axis is consistent with previous reports (Romao, 2017).

Table 4.1: Calcite cell vector lengths obtained from CP2K at 0 K, and after finite temperature correction, in comparison with experimental values.

	***T* = 0 K**	***T* = 300K**	**Exp. (297 K)**(Markgraf and Reeder, 1985)
***a* (Å)**	5.22	5.07	4.99
***c* (Å)**	17.59	17.74	17.06

In the AIMD simulations, electronic minimisation was carried out using the Quickstep implementation of DFT (VandeVondele et al., 2005) where the orbital transformation method was employed (VandeVondele and Hutter, 2003). All DFT calculations used the generalised gradient approximation (GGA) in the form of the revised Perdew-Burke-Ernzerhof (revPBE) exchange-correlation functional (Zhang and Yang, 1998). The revPBE functional has been shown to be particularly effective for describing the liquid structure of water, especially compared to the standard PBE functional (Ruiz Pestana et al., 2017). For representation of dispersion corrections, which is particularly important when simulating liquids, Grimme's D3 corrections were applied (Grimme et al., 2011). In all calculations the short ranged, double-zeta, molecule-optimised basis sets were used with polarisation on heavy atoms (DZVP-MOLOPT-SR-GTH) (VandeVondele and Hutter, 2007). The Goedecker-Teter-Hutter (GTH) pseudopotentials were used to represent atomic cores (Goedecker et al., 1996). The Nose-Hoover thermostat was used for the NVT simulations, with a time constant of 13 fs (Nosé, 1984). A time step of 1 fs was used in all simulations, with the initial temperature set to 300 K. Each interface system was subject to ca. 120,000 steps, giving a total simulation time of 120 ps. For the *ab initio* MD simulations of $E[Ca_n(CO_3)_{(n-1)}\ XO_4]_{(slab,aq)}$ each time step required, on average, around 6-8 seconds on 576 cores of the UK National Supercomputing Service ARCHER. The *ab initio* MD simulation of interfaces models reported herein required in the region of 134,000 CPU hours (wall-clock time × number of processors).

The thermodynamics of surface incorporation were characterized using three interdependent parameters (any two of them determine the value of the third). The first one is the *surface/liquid exchange energy*, which is the energy required to exchange one carbonate and one impurity anion between the surface and the aqueous phase:

$$\Delta E_{\text{exch}}^{\text{surf}} = E[\text{Ca}_n(\text{CO}_3)_{n-1}XO_4]_{(\text{slab,aq})} + E[(\text{CO}_3)^{2-}]_{(\text{isol,aq})} - (E[\text{Ca}_n(\text{CO}_3)_n]_{(\text{slab,aq})} + E[(XO_4)^{2-}]_{(\text{isol,aq})}) \quad (4.1)$$

where $E[\text{Ca}_n(\text{CO}_3)_{n-1}XO_4]_{(\text{slab,aq})}$ is the energy of the calcite/water slab model containing one (XO_4^{2-}) anion in a carbonate position, $E[\text{Ca}_n(\text{CO}_3)_n]_{(\text{slab,aq})}$ is the energy of the pure calcite/water slab, and $E[(\text{CO}_3)^{2-}]_{(\text{isol,aq})}$ and $E[(XO_4)^{2-}]_{(\text{isol,aq})})$ are the energies of the isolated anions in water. In our calculations, there are n=48 formula units in the terrace surface slab supercell and n=52 formula units in the stepped surface supercell.

The second calculated parameter is the *bulk/liquid exchange energy* – the energy required to exchange one carbonate/trace ion from the calcite bulk with aqueous phase:

$$\Delta E_{\text{exch}}^{\text{bulk}} = E[\text{Ca}_m(\text{CO}_3)_{m-1}XO_4]_{(\text{bulk})} + E[(\text{CO}_3)^{2-}]_{(\text{isol,aq})} - (E[\text{Ca}_m(\text{CO}_3)_m]_{(\text{bulk})} + E[(XO_4)^{2-}]_{(\text{isol,aq})}) \quad (4.2)$$

where $E[\text{Ca}_m(\text{CO}_3)_{m-1}XO_4]$ is the energy of the calcite bulk 3×3×1 supercell model containing one (XO_4^{2-}) anion in a carbonate position, $E[\text{Ca}_m(\text{CO}_3)_m]$ is the energy of the pure calcite unsubstituted supercell. Finally, the *segregation energy*, which is the energy required to exchange one carbonate/trace ion between the bulk and the (hydrated) surface of calcite, is given by:

$$\Delta E_{\text{seg}} = E[\text{Ca}_n(\text{CO}_3)_{n-1}XO_4]_{(\text{slab,aq})} + E[\text{Ca}_m(\text{CO}_3)_m]_{(\text{bulk})} - (E[\text{Ca}_n(\text{CO}_3)_n]_{(\text{slab,aq})} + E[\text{Ca}_m(\text{CO}_3)_{m-1}XO_4]_{(\text{bulk})}) = \Delta E_{\text{exch}}^{\text{surf}} - \Delta E_{\text{exch}}^{\text{bulk}} \quad (4.3)$$

The absolute energies used in equations 4.1-4.3 to define $\Delta E_{\text{exch}}^{\text{surf}}$, $\Delta E_{\text{exch}}^{\text{bulk}}$, and ΔE_{seg} were obtained from the average energies of the MD simulations after equilibration. Equilibrated sections of the MD simulation were found to be around the second half of the MD simulations, ca. 90-100 ps of MD run time. In most cases, the first 20 ps of the MD simulation involved equilibration, meaning that the production period consists of ca. 90-100 ps of MD run time in all cases.

We also extracted radial distribution functions (RDFs) from the AIMD simulations. The RDF $g(r)$ is defined in such a way that the number of atoms of the given type at a distance between r and r+dr from the central atom is proportional to g(r)dr, and the function is normalised so that g(r) = 1 when $r \rightarrow \infty$. The RDFs were calculated using every frame from the final 20k AIMD steps, and using bin sizes of 0.1 Å, up to a cutoff of 6 Å.

4.4. Results and discussion

The numerous AIMD simulations obtained as part of this investigation are each very similar. For brevity, only simulations containing molybdenum are shown here as an example of each type of simulation (ion in a box of water, ion in the bulk crystal, ion at the terrace, and ion at the acute step). Figures 4.4-4.7 show both AIMD energy plotted as a function of time, and a running average of the AIMD energy in increments of 1000. For example, Figure 4.4b has a point at 8000 on the x scale, this corresponds to an energy averaged over the final 8000 steps of the AIMD simulation. Bulk MD reported in Figure 4.4 was substantially shorter than all other AIMD simulations reported in this chapter, this is because the solid system shows extremely small energy fluctuations when compared to softer (water-containing) systems. From Figure 4.4b, running averages taken over 6,000 to 10,000 AIMD steps show fluctuations of less than one tenth of an eV, which is extremely small. We can therefore assume that the average energy of the equilibrated system is captured by the latter ~8,000 AIMD steps of this simulation.

A similar approach can be applied for the softer, water-containing systems. Here the running averages tend to increase towards lower values of 'last N steps', which is because for the molybdate-containing models, the energy vs time plots each happen to be on an 'upward swing' when the simulation was terminated. In Figure 4.5b, the running average stays relatively stable when from N = 120,000 to N = 40,000. This means that taking the average energy from the latter 100,000 AIMD steps is a good representation of the average energy of the system.

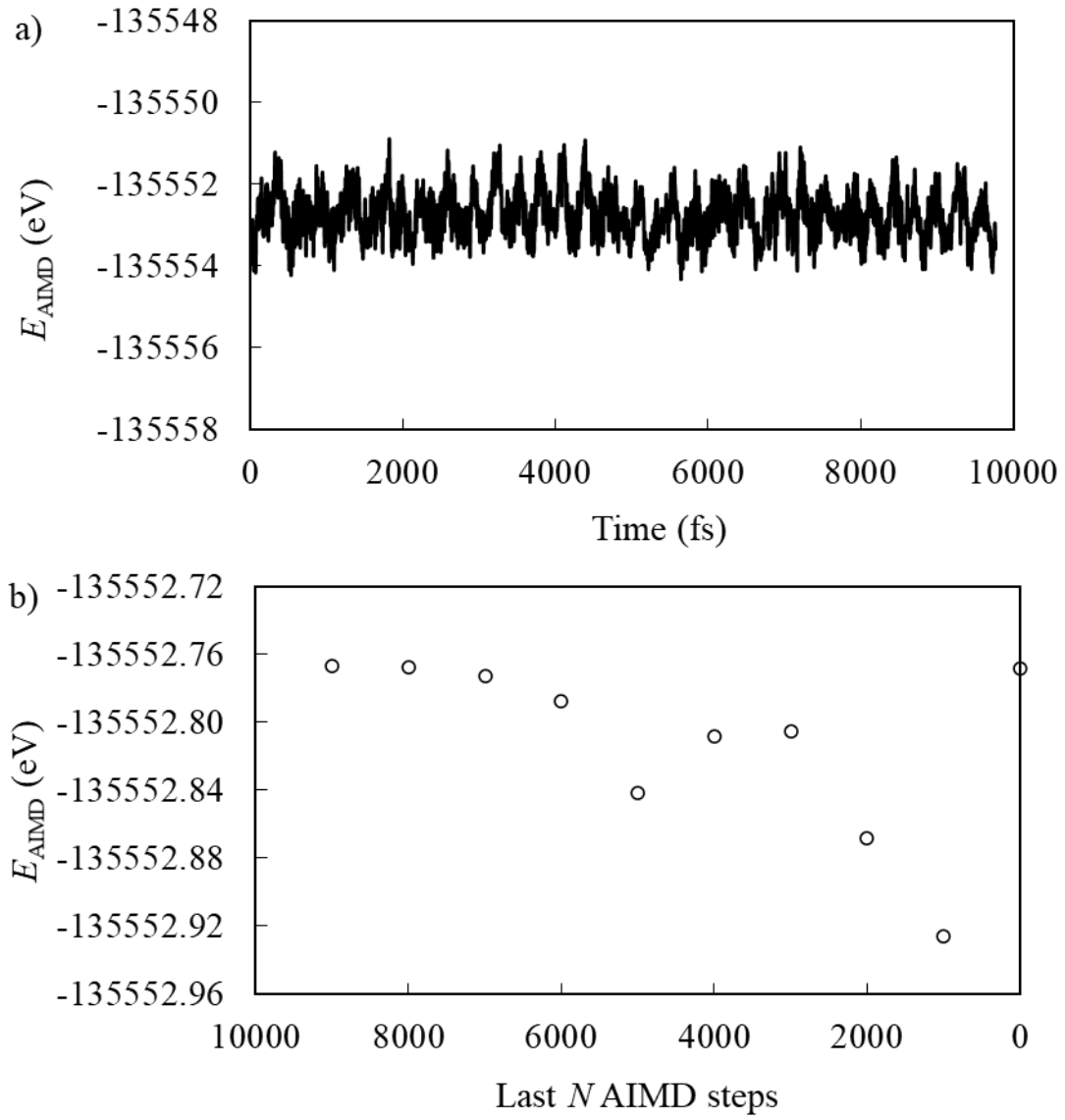

Figure 4.4. Bulk calcite AIMD with molybdate substitution; **a)** AIMD Energy vs Time, **b)** Running average of AIMD energy vs Time.

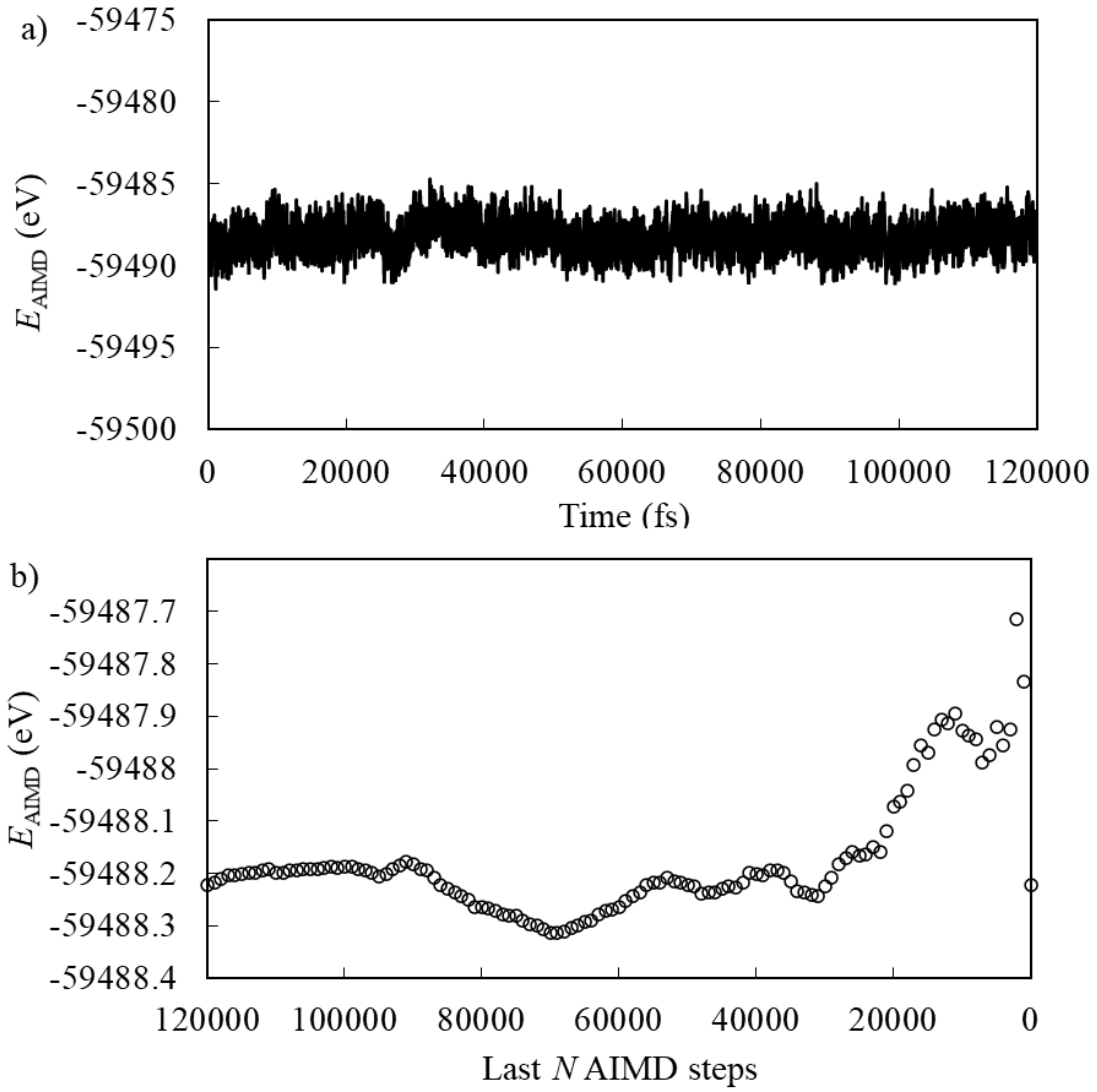

Figure 4.5. Molybdate ion in box of water AIMD; **a)** AIMD Energy vs Time, **b)** Running average of AIMD energy vs Time.

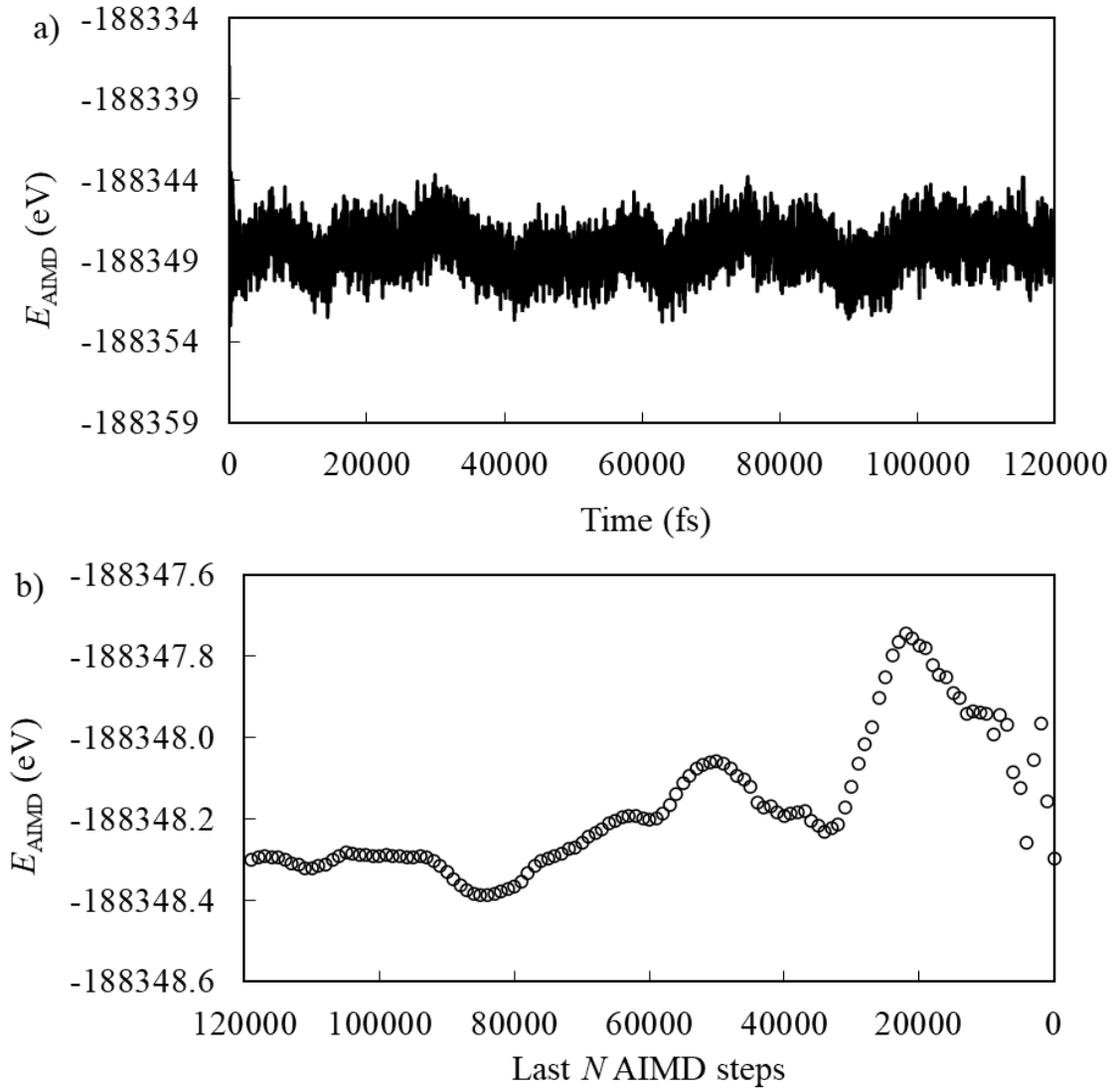

Figure 4.6. Calcite (10.4) terrace interface AIMD with molybdate substitution; **a)** AIMD Energy vs Time, **b)** Running average of AIMD energy vs Time.

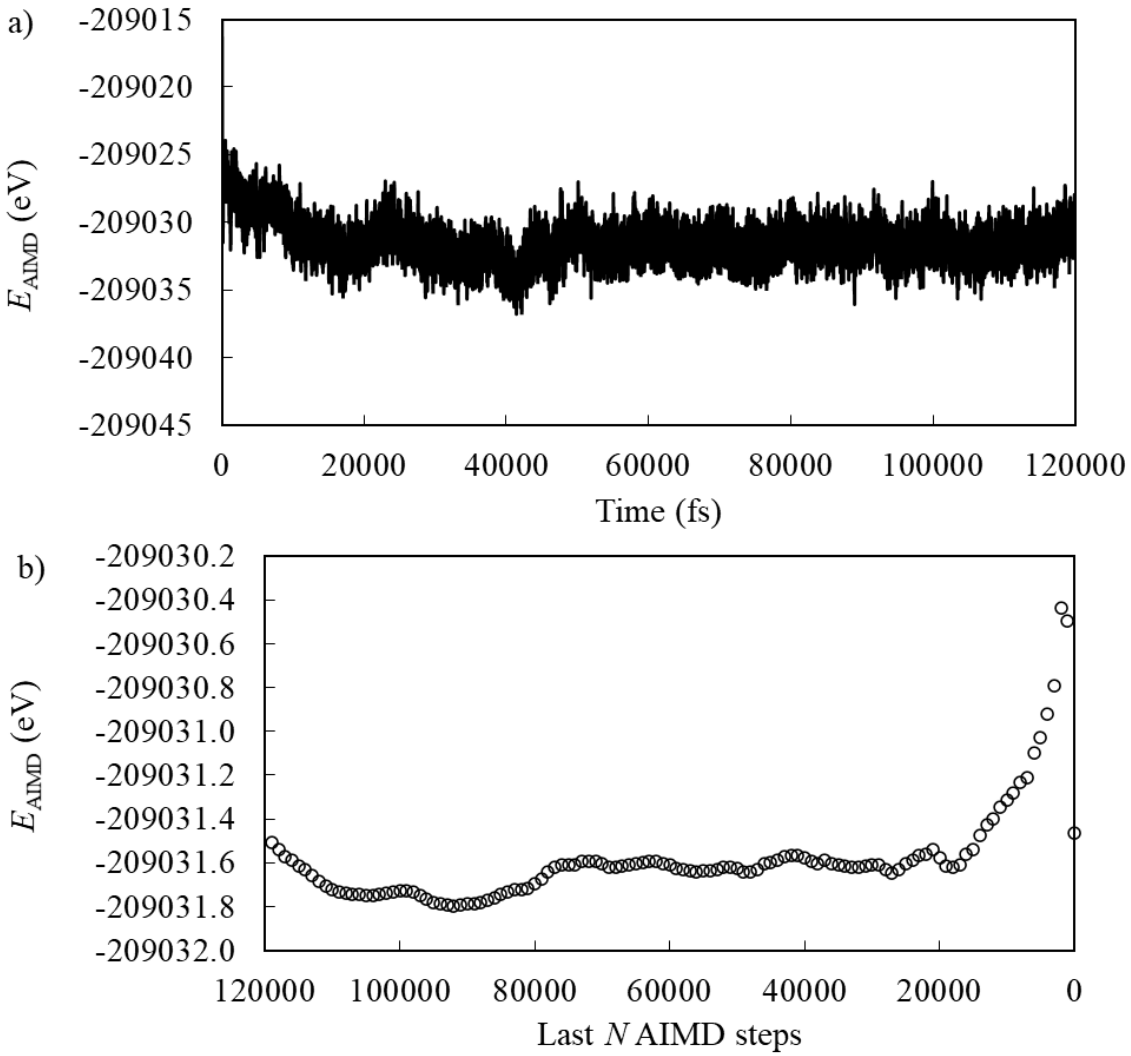

Figure 4.7. Calcite acute step interface AIMD with molybdate substitution; **a)** AIMD Energy vs Time, **b)** Running average of AIMD energy vs Time.

Table 4.2 summarises the exchange and segregation energies obtained for molybdate and sulphate substitutions in calcite. The bulk/liquid exchange energies in calcite (1.08 eV and 1.62 eV for sulphate and molybdate, respectively) show the same trend to the bulk/vacuum exchange energies reported in work reported in Chapter 3 (1.15 eV and 1.45 eV, respectively) (Midgley et al., 2020), i.e. molybdate substitution incurred a greater energy penalty than sulphate substitution, due to the larger size of the former anion. Though the trend in the substitution energies was the same, the difference between sulphate and molybdate substitution was found here to be different (larger) than what was reported in Chapter 3. This is partly because we now use a different thermodynamic reference phase, i.e. oxyanions in water instead of oxyanions in vacuum, and partly because of the dynamic nature of the simulations here, in contrast with the static calculations in Chapter 3. We will see below that,

indeed, the incorporation of oxyanion impurities in calcite contains distinctive dynamic features that cannot be captured by static calculations, even for bulk substitution.

Table 4.2. Exchange and segregation energies of molybdate and sulphate species in calcite.

Species	ΔE_{exch}^{bulk} (eV)	ΔE_{exch}^{surf} (eV)		ΔE_{seg} (eV)	
	bulk	(10.4) terrace	(10.4) step	(10.4) terrace	(10.4) step
sulphate	1.08	0.79	0.27	-0.29	-0.81
molybdate	1.62	0.48	-0.07	-1.14	-1.69

Exchange energies at the (10.4) terrace are reduced when compared to equivalent bulk exchange energies, indicating overall greater thermodynamic stability of the impurities at the surface. Unlike the bulk exchange process, molybdate incorporation at the terrace is more favourable than sulphate, which is surprising when considering that the former is a much larger species. These observations will be rationalized below, based on the discussion of the geometries of the substituted structures. Exchange energies at the step are even more favourable than at the terrace, indicating further thermodynamic favourability than bulk and terrace substitution. Like at the terrace, molybdate substitution is favoured over sulphate substitution at the step, with a comparable energy difference of ca. 0.3 eV.

Overall, the stability of ion exchange in calcite with sulphate and molybdate follows the order: stepped surface > terrace > bulk. The tetrahedral oxyanion impurity does not fit the planar carbonate anion site in calcite. As a result, significant lattice strain is caused in the bulk, which is reduced at surface regions due to the increased interfacial space around the substitution site. The step affords even greater interfacial space than the terrace, which is why the trend in exchange energies at the bulk, terrace and step follow the expected order. The crystallinity of the speleothem and factors affecting the stability of imperfect calcite surface topographies may therefore have a significant impact on detected concentrations of these trace elements in speleothems. Therefore, this should be an important consideration for the interpretation of impurity concentration data collected from speleothems.

To get further insights and rationalise the results from the thermodynamics analysis, we now discuss the MD trajectories and average geometries in equilibrium. Figure 4.8 shows snapshots from the MD simulations, taken after equilibration and at energies close to the equilibrium

average. Figures 4.8 a and b illustrate why the bulk of calcite is not well suited to accommodate the large tetrahedral oxyanion species, as the apical oxygen atom is unable to fit within the planar carbonate anion site. Interestingly, the tetrahedral oxyanions is found to rotate significantly with respect to the plane of the carbonate anions. For example, in the snapshot of the simulation of sulphate in bulk calcite (Figure 4.8a), two of the four oxygen atoms occupy the same oxygen positions as in the vacant carbonate ion, whereas the other two oxygen atoms are above and below the substituting plane. This contrasts with the equilibrium geometry found in static DFT work reported in Chapter 3, where three of the sulphate oxygen atoms occupied the three vacant carbonate oxygen sites, and the fourth (apical) sulphate oxygen pointing out of the plane above the incorporation site. The observation of oxyanion rotation suggests that static calculations are not fully able to capture the behaviour of tetrahedral oxyanion dopants in calcite, even for bulk simulations. This is important because static DFT (or forcefield-based) calculations are still the most widely used method for the simulation of impurities in carbonate minerals.

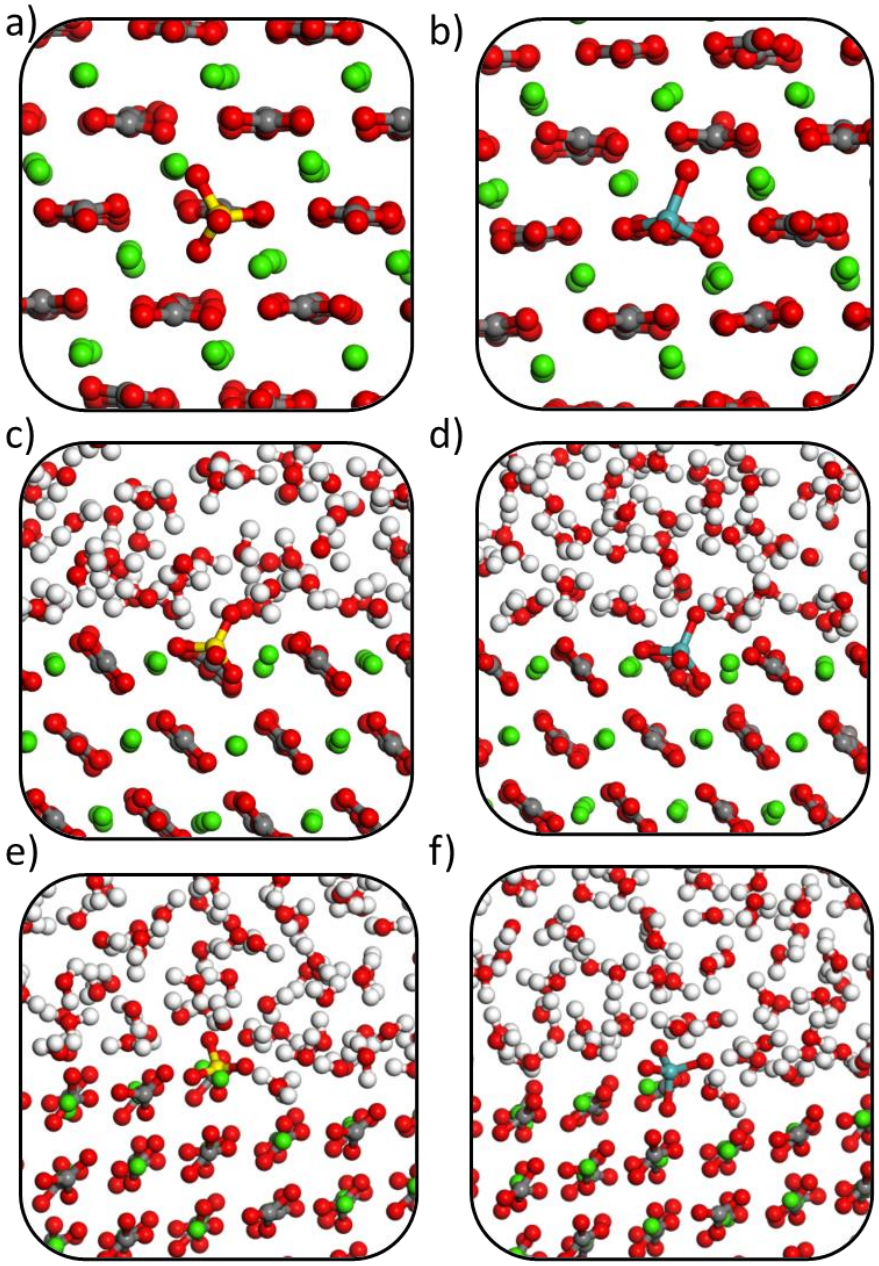

Figure 4.8. MD trajectory snapshots for calcite substituted with a) sulphate at the bulk, b) molybdate at the bulk, c) sulphate at the terrace, d) molybdate at the terrace, e) sulphate at the step, f) molybdate at the step.

In contrast with the bulk substitution, when the oxyanion impurities are substituted at the surface, the apical oxygen atom can stick out, which not only decreases the elastic cost of the substitution, but also leads to extra stabilization by a favourable interaction with water at the interface. To describe this effect in a more quantitative way, we turn to the calculated radial distribution functions (RDFs) from the MD simulations.

We start with the analysis of RDFs for X-O_w (i.e. from C, S or Mo atom in the oxyanion to the oxygen atoms O_w in water molecules), for the simulation of isolated oxyanions in water are reported in Figure 2a. The three anions have the same charge but different ion sizes (carbonate < sulphate < molybdate), as illustrated by the peak positions between 3 and 4 Å in the X-O_w RDFs. This means that the strength of the interaction with water is expected to follow the opposite trend (carbonate > sulphate > molybdate), in agreement with experimental hydration enthalpies, e.g. the hydration of a carbonate anion is ~2.6 eV more exothermic than the hydration of the larger sulphate ion (Smith, 1977). Compared with the case of sulphate, the first hydration sphere around molybdate is shifted right and broadened. There is also no-clearly defined second hydration sphere, proving that polarization on the surrounding water is much weaker for molybdate. The weaker interaction of molybdate with water can be expected to lead to a faster kinetics of water exchange in and out the hydration spheres of the anions. This is consistent with our calculation of the mean residence time (MRT) of water molecules surrounding the sulphate and molybdate ions in aqueous solution. These are obtained as $MRT = t_{sim} \times CN_{ave} / N_{ex}$, where t_{sim} is the ab initio MD simulation time, CN_{ave} is the mean coordination number for sulphate and molybdate, and N_{ex} is the number of exchange events involving the first hydration sphere during the simulation time (only those with a duration longer than 0.5 ps are counted). MRTs for sulphate and molybdate were calculated as 0.8 and 0.5 ps respectively, which indicate that the hydration environments of both oxyanions are highly labile (comparable MRTs to that of the Cl^- ion) (Wang et al., 2020) but more so for molybdate than for sulphate.

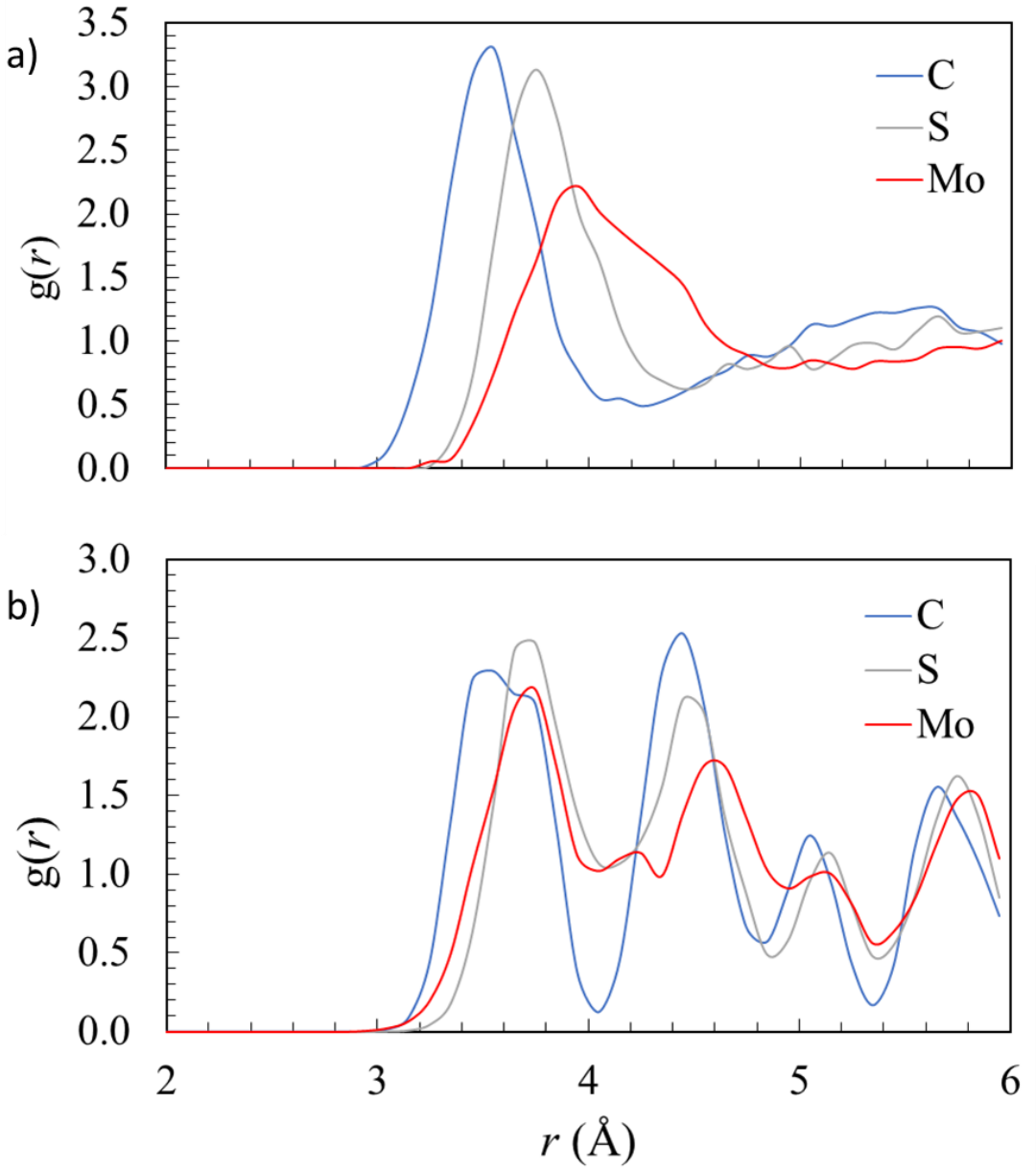

Figure 4.9. a) Radial distribution function X-O_w, i.e. from the X = (C, S, Mo) atom in the $(XO_n)^{2-}$ oxyanion to the oxygen atoms (O_w) in water molecules, for the simulation of isolated oxyanions in water, **b)** Radial distribution function X-O_c, i.e. from the X = (C, S, Mo) atom in the $(XO_n)^{2-}$ oxyanion to the oxygen atom (O_c) of the carbonate anions, for the simulation of bulk calcite.

The X-O_c RDFs, i.e. from the X = (C, S, Mo) atom in the $(XO_n)^{2-}$ oxyanion to the oxygen atom (O_c) of the carbonate anions, for the simulation of bulk calcite is shown in Figure 4.9b. The first peak at around 3.5 Å corresponds to X-O distances in the monolayer above and below the plane of the anion, and this clearly illustrates the structural impact of the impurity on the surrounding ions. Peaks for sulphate and molybdate are shifted right compared with

carbonate, which is consistent with crystallographic strain and repulsion with neighbouring carbonate ions, caused by the tetrahedral impurities. The second peak at around 4.5 Å corresponds to the *X*-O distances in the neighbouring carbonate ions in the plane of the subject anion. Here, the sulphur and molybdenum peaks are shifted right, again indicating crystallographic strain and repulsion of the neighbouring carbonate ions by the presence of the impurity. Where *X* = Mo, oxygen atoms are repelled most significantly, due to the large size of the molybdate anion.

At the surface, the unfavourable steric strain in the lattice caused by tetrahedral oxyanion substitution is reduced, as the apical oxygen atom can protrude from the solid surface rather than causing strain on neighbouring ions in the crystal. This effect is further extended at the step where the substituting anion is afforded maximal interfacial space, meaning very low structural strain the carbonate solid. At the step, three out of four oxygen atoms can protrude from the carbonate solid, while still allowing the anion to be chemically bound. This explains why overall, the surface substitutions are more favourable than the equivalent bulk substitutions.

Comparison of RDFs for molybdate and sulphate at the terrace and step may be made from Figures 4.10 a and b, which explain the lower exchange energies for both ions at the step than the terrace, as reported in Table 4.2. When substituting the impurity at the terrace, there is around 180° of arc which is in contact with the water (the other 180° is in contact with the lattice). When substituting at the step however, the angle of arc in contact with the water is around 270°, which allows for the formation of larger partial hydration spheres, which drive thermodynamic favourability. We discuss molybdenum as an example (sulphate may be assumed to behave in the same way, but with effects less pronounced because of its smaller size). RDFs reported in Figure 4.10 show the partial hydration environment when molybdate is incorporated at the terrace and the step. From the Figure 4.10, the first hydration sphere around molybdate at 3.8 Å is clearly better defined at the step (4b) than at the terrace (4a). This is because the molybdate ion at the step has more contact with liquid water, causing greater polarization of the aqueous layer and more favourable partial hydration spheres. A combination of decreased lattice strain and increased hydration sphere formation at the step when compared to the terrace explain why there is a thermodynamic preference for substitution at the step.

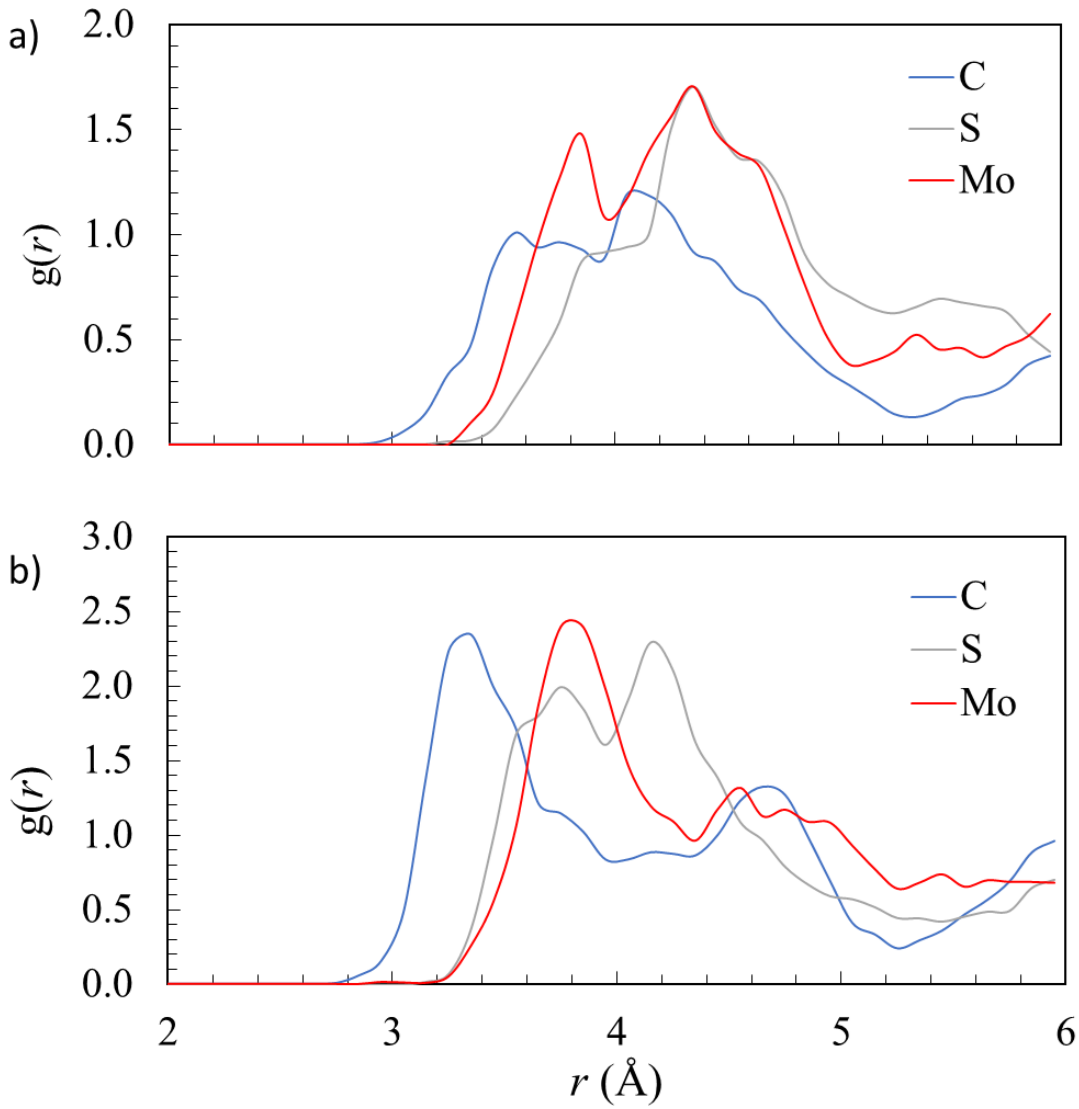

Figure 4.10. Radial distribution function X-O_w, i.e. from the X = (C, S, Mo) atom in the $(XO_n)^{2-}$ oxyanion to the oxygen atoms (O_w) in water molecule, for the simulations of oxyanions substitutions at a) the calcite terrace and b) the calcite step.

Another significant trend in the exchange energies (Table 4.2) is that molybdate appears more favourable than sulphate when substituting both surface sites, which is a reversal of the bulk incorporation energy trends. This means that despite greater steric strain caused by molybdate over sulphate, other factors are significant in driving the thermodynamic exchange energy. A possible explanation comes from the different levels of hydration of the substituting ions at the surface. As the molybdate ion protrudes more readily from the surface, aside from reducing lattice strain this also allows for the formation of stronger partial hydration spheres. Hydration sphere formation is favourable and drives the thermodynamic process of substitution. Figure 4.10a reports RDFs showing the partial hydration spheres formed around sulphate and molybdate ions substituted at the calcite terrace. Molybdate shows clear first peat

at around 3.8 Å, while sulphate shows only a very small shoulder at this region. The explanation for this is that the molybdate ion protrudes from the surface more significantly than the sulphate meaning that the central atom (Mo) is closer to the water than the central atom (S) in sulphate, which is more deeply embedded into the lattice. Similar trends are reported in Figure 4.10b - RDFs showing partial hydration environments for sulphate and molybdate substituted at the calcite step. The molybdate shows a stronger, more well-defined first hydration sphere when compared to sulphate, because it is repelled from the carbonate crystal more strongly and is therefore in closer contact with the surrounding water. This analysis provides more evidence to the notion that the molybdate is more strongly ejected from the terrace while remaining chemically bound. Our proposed mechanism of stronger patrial surface rehydration for molybdate over sulphate also gives another thermodynamic contributor to the relative favourability of molybdate substitution over sulphate at this site.

Finally, while the discussion above focuses on the thermodynamics of impurity incorporation, the presence of sulphate and molybdate impurities is likely to significantly impact the kinetics of surface processes, including the dynamics of water adsorption/desorption from the surface and, at a longer timescales, the dynamics of mineral growth/dissolution. These processes are beyond the scope of this work, and their consideration would require free energy calculations using, for example, metadynamics techniques, which are very computationally expensive for the size and number of models considered here. However, to gain some preliminary insight into the kinetics of water adsorption at the interface, and the effect of molybdate and sulphate impurities on that kinetics, we have performed an analysis of the water exchange rate at the calcite water interface, in the absence and presence of sulphate/molybdate impurities. We counted the number of water molecules going in and out of the first hydration spheres of surface Ca^{2+} cations, with radius of ~3 Å according to the Ca-water RDF. As in the analysis for the individual anions in water presented above, exchange events had to last at least 0.5 ps to be counted, i.e. rapid oscillations at the boundary were not included. The results are shown in Table 4.3. The rate of water exchange events increases from the pure carbonate surface to the sulphate-substituted one, and then increases even more for the molybdate-substituted one. This behaviour can again be rationalised in terms of the weaker polarisation by the molybdate anion of the surrounding water molecules, compared with the case of sulphate, as discussed above for the cations of the isolated ions in water.

Table 4.3. Rates of water exchange events at the pure and substituted (10.4) terraces of calcite (per picosecond and normalised by the number of surface cations).

Surface	Water exchange events per ps
Pure carbonate	3.19
Sulphate substituted	4.27
Molybdate substituted	4.56

4.5. Conclusions

We have presented an ab initio molecular dynamics investigation of the incorporation of sulphate and molybdate as substitutional impurities in the bulk and surface of calcite. Results from this investigation indicate a strong tendency for ion exchange at the calcite/water interface regions of the speleothem when compared to exchange with the crystalline bulk. When comparing the incorporation of a given impurity species across the bulk, terrace and step substitution sites, the ability of the solid to accommodate the tetrahedral oxyanion is still the key thermodynamic factor. In order to understand the differences between sulphate and molybdate substitution thermodynamics, we need to pay attention to other factors, for example the partial hydration of the impurity at the calcite surface.

Bulk ion exchange of molybdate with liquid water is more thermodynamically unfavourable than sulphate, which is due to the increased lattice strain caused by the much larger molybdate ion. Exchange at both the terraced and stepped interfaces is more favourable for molybdate than for sulphate, which is a reversal in the trend observed for the two ions in the bulk. We rationalise this trend reversal by considering that the molybdate ion is ejected from the solid surface more significantly due to higher lattice strain, but is able to remain chemically bound to the solid due to its longer *X*-O bonds than sulphate. This proposed mechanism is supported by the analysis of radial distribution functions indicating stronger partial rehydration of the molybdate impurity at the surfaces, because the ionic centre (Mo) is less deeply embedded in the carbonate solid, meaning closer proximity to surrounding water than the analogous S atom. Although our analysis here is based on thermodynamics, the preferential substitution of the anion impurities at the surface and steps of calcite must also have important consequences for the growth kinetics. Calcite crystals growth is dominated by steps dynamics,

and the presence of impurities, particularly during extreme events such as volcanic eruptions, at the steps is known to stunt calcite growth (de Leeuw, 2002; Wasylenki et al., 2005b). The interplay beyond thermodynamic and kinetic factors in this phenomenon is complex and beyond the scope of the present study.

These calculations demonstrate and quantify the importance of surfaces and surface defects in the incorporation of oxyanion impurities in calcite. Since Mo and S impurities in speleothem calcites constitute a reliable record of past volcanic activity, the understanding gained of the factors controlling the concentrations of these impurities in calcite will be useful for a quantitative interpretation of speleothem records in the future.

4.6. Reflection

4.6.1. Limitations

We used slab models to represent solid/water interfaces, because full representation of a nanoparticle/water systems would have incurred in a prohibitively large computational cost when using *ab initio* techniques. Though these slab models perform well, they show small unphysical elastic interactions between impurities and their periodically repeated images, which may have a small systematic effect in energies obtained. A more accurate physical model would be to represent the calcite particle in a very large box of water, rather than using the periodic slab model approach. Though at present, the computational demands of AIMD mean that this kind of 'whole-particle' model would not be feasible, advances in exascale computing however may facilitate this kind of model in the future.

AIMD simulations were subject to pre-equilibration where possible, using density functional tight binding techniques, though this was not possible for molybdate containing systems due to a lack of known parameters. Sulphate and molybdate containing models are structurally analogous, meaning that the sulphate models can serve as a good starting point for molybdate containing models. However, molybdate containing structures still require a large degree of reorganisation in the water layers surrounding the impurity, which means that molybdate containing models must still be subject to extended AIMD runs.

A noteworthy methodological contribution of this chapter is to comment on the cost-effectiveness of the *ab initio* molecular dynamics to represent the carbonate-water systems. Though AIMD techniques can, in principle, accurately represent the carbonate-water-impurity systems investigated, their extreme computational costs have two significant practical implications for research endeavours of this kind:

I. Preliminary investigation for this chapter revealed that even when using the national tier 1 supercomputing facility, ARCHER, it is only reasonable to simulate up to 700-800 total atoms using AIMD, before the computational cost becomes inhibitively high. Relatively small models such as this are subject to very slow equilibration due to the high magnitude of oscillations with respect to the size of the cell, meaning that extended AIMD runs are required for effective equilibration, incurring an extreme computational cost.

II. Due to the extreme computational cost of each individual MD step, running many steps is extremely expensive in terms of supercomputing time. We had originally estimated that 40 ps of AIMD would be enough to achieve equilibration; however statistical analysis revealed that a minimum of 100 ps was in fact required in most cases.

In the present investigation we were able to mitigate the above consequences of the extreme cost of AIMD by running extended simulations on tier 1 supercomputing facilities. This however was at the expense of using computing time that had originally been intended for other research projects relating to the present thesis. The extreme cost of AIMD simulations therefore significantly reduced the volume of AIMD-based work that could be completed throughout the present PhD; and this may be an important consideration for research groups thinking of employing AIMD to soft chemical systems such as solid/liquid interfaces.

4.6.2. Future work

The above limitations mean that AIMD study of *soft* carbonate-water interfaces is only recommended where researchers have access to extreme high performance computing facilities. Though the advent of exascale supercomputing may mean that AIMD can be better used in this kind of investigation, there are currently practical limitations to extended AIMD simulations. It may therefore be prudent for researchers wishing to study these systems to develop effective interatomic potentials, which would allow for very long MD simulations to be obtained at a fraction of the cost of AIMD. This however is not a trivial, and there would be significant technological challenges in developing effective potentials, particularly for the heavy metallic atoms such as molybdenum, which contain highly diffuse orbitals. Energies derived from parameterisations and classical MD would also not allow direct comparison between impurity energies, due to the poor transferability of interatomic potentials between systems. Transferability is one of the major benefits in using the DFT-based simulations in the present thesis.

Though the calcite (10.4) and acute stepped surface are common under geochemical conditions, there are many other calcite particle facets which could provide interesting alternate thermodynamic pathways to trace element incorporation. Calcite often possesses other (less stable) steps, including the obtuse step (de Leeuw et al., 1999). Another prominent calcite surface is the (00.1) basal termination. This surface was preliminarily investigated as part of the present PhD project; however the investigation was discontinued due to insufficient computing resources.

Aside from alternative calcite incorporation sites, aragonite also provides a reasonable means of trace element incorporation in speleothems, though it is less commonly encountered than calcite. Aragonite bulk and surface substitution was preliminarily investigated as described in Chapter 7; however this investigation was discontinued due to attenuated computing costs.

5. Bromate incorporation in calcite and aragonite

5.1. Summary

Here the incorporation of BrO_3^- in calcite and aragonite was investigated, with analysis of the disorder introduced by the singly charged cation incorporated to balance overall charge. The multi-configurational problem introduced by two substituting ions meant that AIMD study of the carbonate/water interface was not computationally viable, therefore we decided to revert to a simpler structural model, only representing bulk substitutions, similar to those reported in Chapter 4. This chapter is an adaptation of the article (Midgley et al., 2021b), which has also been submitted to *Geochimica et Cosmochimica Acta*. Here, S. D. Midgley is the primary author and contributor to the experimental procedures and analyses. Co-authors provided helped with the initial project idea and refinement of the manuscript, but methods and analysis represent the work of the primary author. Research data for this chapter is available via Zenodo, at https://doi.org/10.5281/zenodo.5035293. The term x_m as discussed in this section has the same mathematical definition as x_m in Chapter 3. In Chapter 3 we referred to this value as the 'maximum stable molar fraction', while in the present chapter we refer to this value as the 'solubility limit'. Though these values are the same, we decided to use an alternative name to reflect the different context in which x_m is used. These names can accurately be used interchangeably for any instance of x_m.

5.2. Introduction

Major volcanic eruptions affect the chemical composition of the atmosphere and pedosphere (soil layers). Elements such as sulphur, molybdenum and bromine are known to be in high relative abundance following an eruption, and these may eventually become incorporated as impurities into the structures of speleothems. However, little is known about the chemical nature of these species in speleothem carbonate minerals. Recent computer modelling research by our group has considered the thermodynamics of sulphate and molybdate incorporation in speleothems, finding that bulk lattice incorporation of these anions is thermodynamically unfavourable, owing to significant elastic strain on the crystal caused by the large, tetrahedral anion geometries (Midgley et al., 2020), and that surfaces (or more accurately, the mineral-water interface) provide a more favourable environment for hosting the oxyanion impurities, especially in the presence of surface defects (Midgley et al., 2021a).

We extend here our computational studies to the investigation of bromine incorporation in calcite and aragonite. Badertscher and co-workers (Badertscher et al., 2014) have argued that bromine is probably the most effective speleothem-based paleovolcanic proxy among those investigated so far. Its fast leaching through soil layers made it detectable in speleothems found

in the Sofular cave, within +/- 1-2 years of the Minoan eruption that occurred in Northern Turkey at around 1621 BCE (Friedrich et al., 2006). Very little is known about how bromine incorporates in calcium carbonate minerals. Elevated quantities of bromine in the form of hypobromite (BrO) are dumped into the local environment following a major volcanic eruption (Bobrowski et al., 2003; Hörmann et al., 2013; Theys et al., 2009). Hypobromite has been shown to play a critical role in tropospheric and stratospheric chemistry, with ozone depletion known to be a significant environmental consequence of this species if airborne. Its stability with respect to oxidation to bromate is low under most conditions (Engel and Perlmutter-Hayman, 1953; Lee and Lister, 1971). It is therefore reasonable to assume that incorporation of bromine in speleothems occurs in the form of the monovalent anion bromate, $(BrO_3)^-$, in analogy with the incorporation of iodine as iodate, $(IO_3)^-$, in calcium carbonate minerals (Feng and Redfern, 2018; Kerisit et al., 2018; Podder et al., 2017).

Bromate has a trigonal pyramidal geometry due to the presence of a lone pair, and therefore it is not an ideal fit to substitute the trigonal planar carbonate (CO_3^{2-}) anion. However, its flatter shape in comparison with that of tetrahedral oxyanions sulphate $(SO_4)^{2-}$ and molybdate $(MoO_4)^{2-}$ may allow easier lattice substitution. The chemistry of the aliovalent bromate/carbonate substitution is different from that of the isovalent molybdate/carbonate or sulphate/carbonate substitutions, and the charge compensating mechanisms can be expected to play a significant role in the thermodynamics of bromate incorporation. (Kerisit et al., 2018) indeed found that, in the case of iodate substitution in calcite, there is an energetic preference for the compensating cation (H^+ or Na^+ in their study) to be placed near the iodate impurity. The ease of incorporation also depends on the host carbonate phase, and it was found by Feng and Redfern (Feng and Redfern, 2018) that iodate can incorporate into crystalline bulk calcium carbonate with increasing difficulty in the order vaterite, calcite and aragonite, following the relative density of these phases.

In the present chapter we use first principles simulation techniques to investigate for the first time the incorporation of bromate in calcite and aragonite (as the phases most interesting for our understanding of speleothems), and the relative distribution of bromate anions and compensating cations in both phases. We will report the thermodynamic parameters describing the ease of incorporation in each case, both with respect to isolated ions and to competing phases. We will discuss the impurity distribution problem in terms of two thermodynamic driving forces: the binding energy between the two oppositely charged impurity defects that tends to keep the impurities paired, and the configurational entropy effects that favour unpaired impurities at low concentrations.

5.3. Computational methods

Computer simulations within the density functional theory (DFT) were performed using the VASP code (Kresse and Furthmüller, 1996a; Kresse and Furthmüller, 1996b), under the generalized gradient approximation (GGA) in the form of the PBE exchange-correlation functional (Perdew et al., 1996b). The projector augmented wave (PAW) method (Blochl, 1994; Kresse and Joubert, 1999) was used to describe the interaction between valence electrons and the core. The number of plane waves used in the basis was determined by a kinetic energy cutoff of 520 eV, which is 30% above the recommended value for the set of PAW potentials used, to minimize Pulay stress errors.

Calcite has a trigonal crystal system and a space group of *R*-3*c* (No. 167) (Markgraf and Reeder, 1985), and was modelled here using a hexagonal 3×3×1 supercell (54 formula units of $CaCO_3$) as shown in Figure 5.1a. Aragonite is orthorhombic with the space group *Pmcn* (No. 62) (De Villiers, 1971), and was modelled using a 3×2×3 supercell (72 formula units; Figure 5.1b). The use of relatively large supercells is needed to maximise the distance between each impurity and its periodic images (at least ~15 Å apart), as well as to explore the configurational space of relative positions of cation and anion impurities. The symmetrically inequivalent substitution configurations were obtained using the methodology implemented in the Site Occupancy Disorder program (SOD) (Grau-Crespo et al., 2007), which has been useful before in the investigation of impurity distribution in carbonate minerals (González-López et al., 2014; Ruiz-Hernandez et al., 2010; Wang et al., 2011b). Two configurations are equivalent if they are related by a symmetry operator of the parent (unsubstituted) crystal structure. Each configuration was relaxed to full geometry optimization in VASP, and the final energies were used for statistical analysis of the impurity distribution.

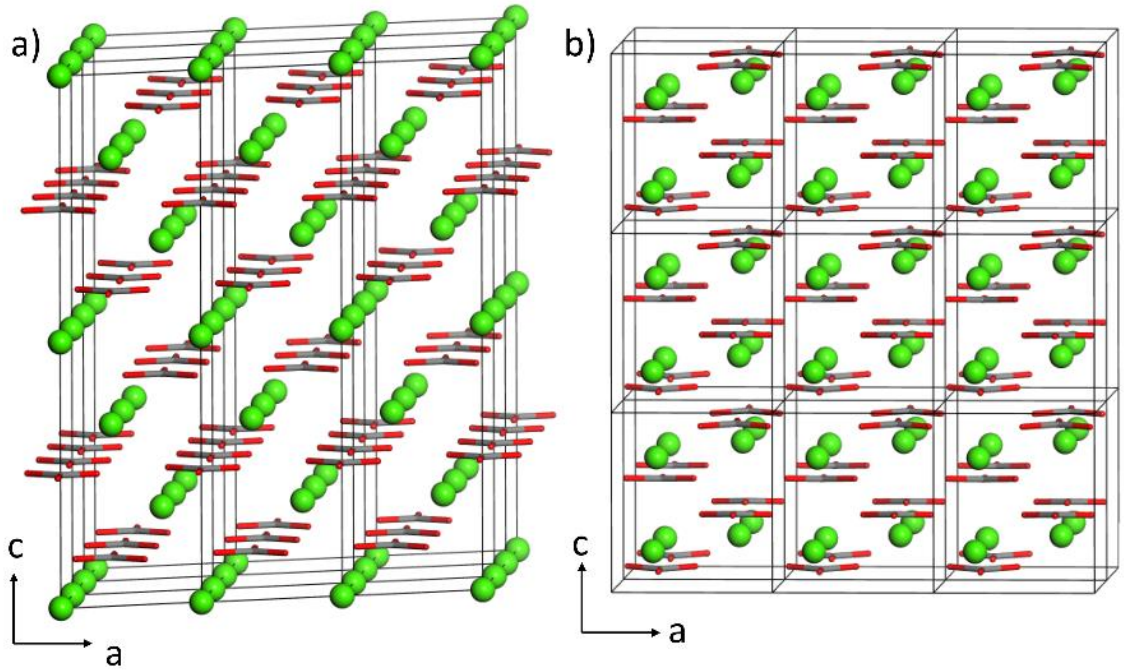

Figure 5.1. Supercells of **a)** calcite (3×3×1) and **b)** aragonite (3×2×3) employed in this work. Colour code: Ca= green; C=grey; O=red.

The SOD/VASP analysis elucidated the minimum energy configuration of $M\mathrm{BrO_3}$ in calcite and aragonite, which was always the proximal ion-pair configuration. To calculate the energy of $\mathrm{CaCO_3}/M\mathrm{BrO_3}$ substitution in calcite, we computed two sets of energy values. The first we call the ion exchange energy E_{exch}, and is the energy of exchanging $\mathrm{Ca^{2+}}$ and $\mathrm{CO_3^{2-}}$ ions from the bulk calcium carbonate structure with M^+ and $\mathrm{BrO_3^-}$ ions from vacuum:

$$\Delta E_{\mathrm{exch}} = E\left[\mathrm{Ca}_{n-1}M(\mathrm{CO_3})_{n-1}\mathrm{BrO_3}\right] + \left(E\left[(\mathrm{Ca^{2+}})_{\mathrm{gas}}\right] + E\left[(\mathrm{CO_3^{2-}})_{\mathrm{gas}}\right]\right) - E\left[\mathrm{Ca}_n(\mathrm{CO_3})_n\right] + \left(E\left[(M^+)_{\mathrm{gas}}\right] + E\left[(\mathrm{BrO_3^-})_{\mathrm{gas}}\right]\right) \quad (5.1)$$

where $E[\mathrm{Ca}_{n-1}M(\mathrm{CO_3})_{n-1}\mathrm{BrO_3}]$ is the energy of the calcium carbonate supercell containing one ($M^+ + \mathrm{BrO_3^-}$) substitution, $E[\mathrm{Ca}_n(\mathrm{CO_3})_n]$ is the energy of the pure calcium carbonate supercell, $E[(M^+)_{\mathrm{gas}}]$ and $E[(\mathrm{BrO_3^-})_{\mathrm{gas}}]$/ $E[(\mathrm{CO_3^{2-}})_{\mathrm{gas}}]$ are the energies of the isolated ions. This analysis allows for comparison between cation substitutions within the same carbonate phase, and comparison of the same cation substitution between carbonate phases. As these calculations are benchmarked against the arbitrary reference of ions in vacuum, the absolute values of these energies carry no physical meaning. The purpose of E_{exch} is to provide a common reference point for comparison of energies between each of the systems examined.

To calculate Ca^{2+} + CO_3^{2-} and M^+ + BrO_3^- as isolated ions in vacuum, within the context our periodic DFT calculations, we calculated the energy of the ion pair in the limit of an infinitely large cell. Models were created where the cation was placed in the centre of the cell and the anion was placed at the corner of the cell. The cell was expanded from a = 4 Å to a = 28 Å in 4 Å increments. We plotted $1/a$ vs energy as shown in Figure 5.2 below. In each case, the y-intercept of the linear relationship shown corresponds to the energy of the non-interacting ion pair in an infinitely large cell.

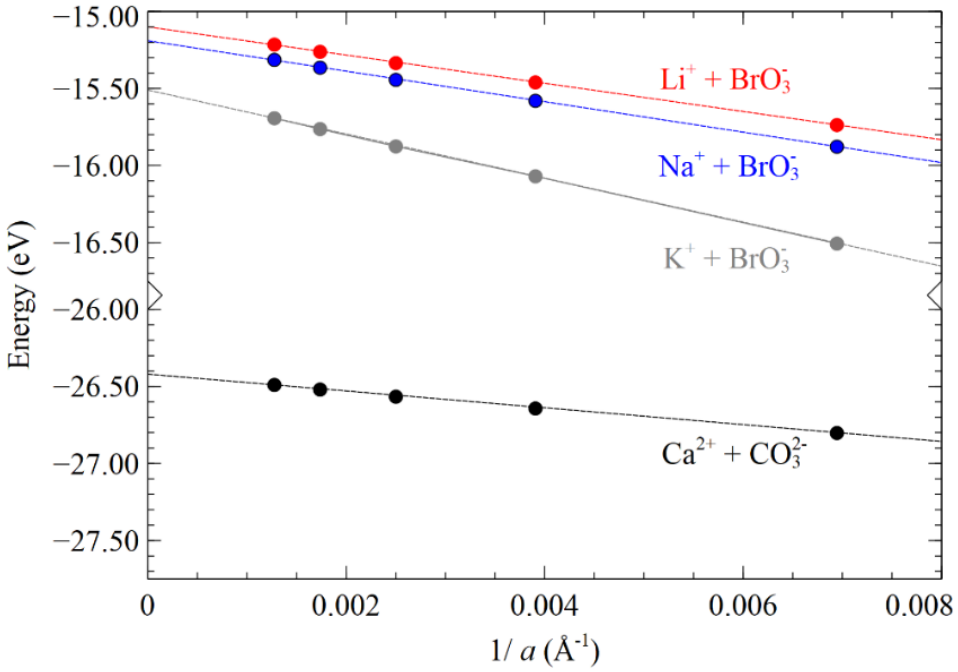

Figure 5.2. Determination of the energy of separated anion-cation pairs: the energy per cell of a periodic system containing one cation at the origin and one anion at the centre of the cell is plotted vs the inverse of the lattice parameter a, keeping interatomic distances constant within the oxyanion. The linear extrapolation when $1/a$ tends to zero gives the energy of the pair at infinite separation.

To assess the stability of bromate species in the calcium carbonate phases, we also calculated the solution energies, E_{sol}, which are relative to the thermodynamic competing phase (sodium bromate or potassium bromate):

$$\Delta E_{\text{sol}} = E[\text{Ca}_{n-1}M(\text{CO}_3)_{n-1}\text{BrO}_3] - (n-1)E[\text{CaCO}_3] - E[M\text{BrO}_3] \quad (5.2)$$

where E[$CaCO_3$] is the energy per formula unit of the pure calcium carbonate, and E[$M$$BrO_3$] is the energy per formula unit of the competing phase: sodium bromate and potassium

bromate. As pure $LiBrO_3$ is not known, we have excluded lithium from consideration here. The absolute magnitude of the solution energy calculated in this way is more physically meaningful. However, since the ΔE_{sol} energies are not calculated with respect to a common reference, they should not be used for comparison across different polymorphs or compensating cations.

5.4. Exchange and solution energies

For calcite we found 11 inequivalent configurations of one cation/anion pair substitution per supercell, whereas in the aragonite supercell the number of inequivalent configurations is 48. The energies of each configuration in calcite, after full geometry optimisations, are reported in Table 5.1. They are given relative to the overall lowest total energy configuration, which is configuration 1 in each case and corresponds to the nearest-neighbour anion-cation configuration.

Table 5.1. Energies of $M^+ + (BrO_3)^-$ incorporation in calcite (M = Li, Na and K), relative to the configuration with nearest-neighbour pairing.

Configuration	d_{min} (Br-M^+) (Å)	E_{Li} (eV)	E_{Na} (eV)	E_K (eV)
1	3.25	0.00	0.00	0.00
2	4.31	0.60	0.40	0.24
3	6.00	0.51	0.29	0.08
4	6.64	0.57	0.39	0.24
5	6.64	0.43	0.28	0.08
6	7.74	0.63	0.43	0.27
7	7.85	0.56	0.36	0.21
8	9.24	0.59	0.40	0.25
9	9.75	0.59	0.40	0.25
10	9.75	0.59	0.41	0.28
11	10.56	0.63	0.36	0.28

In Table 5.1, d_{min} represents the minimum distance between Br and M *before* relaxation (i.e. between the C and the Ca atoms being substituted in the pure calcite structure). Where the anion-cation distance is the same in more than one case (configurations 4&5 and 9&10), the configurations are still symmetrically inequivalent by virtue of the oxygen atoms, and in these cases, anion-cation distances diverge following structural relaxation.

The binding energies (ε) between the impurities, calculated as the difference between the lowest and highest energies in the configurational spectrum, are shown in Table 2 together with the exchange energies for each host phase and compensating cation, calculated using Eq. (1) under the assumption that the bromate anions and the compensating cations are in relative NN positions.

Table 5.2. Binding energies (ε) between oppositely charged impurity defects, and ion exchange energies (ΔE_{exch}) for calcite and aragonite assuming impurity pairing in nearest-neighbour positions.

Compensating cation	calcite		aragonite	
	ΔE_{exch} (eV)	ε (eV)	ΔE_{exch}	ε (eV)
Li^+	5.59	0.63	5.38	1.08
Na^+	6.42	0.43	6.31	0.62
K^+	7.33	0.28	6.95	0.54

Table 3 list solution energies, calculated using Equation 5.2. Solution energies provide thermodynamic insight into formation energy with respect to the competing phase, for example sodium bromate ($NaBrO_3$) or potassium bromate ($KBrO_3$). Here the correct procedure is to evaluate the energy of the energy of the mixed system in the absence of NN pairing, because as we will show below, solubility limits (shown as x_m in Table 5.3) are well below the pairing threshold.

Table 5.3. Solution energies (ΔE_{sol}) and solubility limits (x_m) for alkali metal-compensated bromate substitution in calcite and aragonite.

	calcite		aragonite	
Competing phase	ΔE_{sol} (eV)	x_m	ΔE_{sol} (eV)	x_m
$NaBrO_3$	1.41	2×10^{-12}	1.71	5×10^{-15}
$KBrO_3$	1.79	1×10^{-15}	1.88	2×10^{-16}

5.5. Results and discussion

There is a clear energy difference between the configuration with nearest-neighbouring impurities and the rest (Table 5.1), whereas the energy differences between the rest of the configurations are smaller. This is illustrated in Figure 5.3, for the case of bromate/lithium

substitution in calcite. The energy distribution across the configurations can be roughly approximated as a sharp drop in energy (the binding energy) at nearest neighbour position, and a constant energy for all the other configurations. This behaviour is observed for both calcite and aragonite (the latter not shown in a table due to the much larger number of configurations). The binding energies (ε) between the impurities follow the trends Li>Na>K, and aragonite> calcite, which we will attempt to rationalise below.

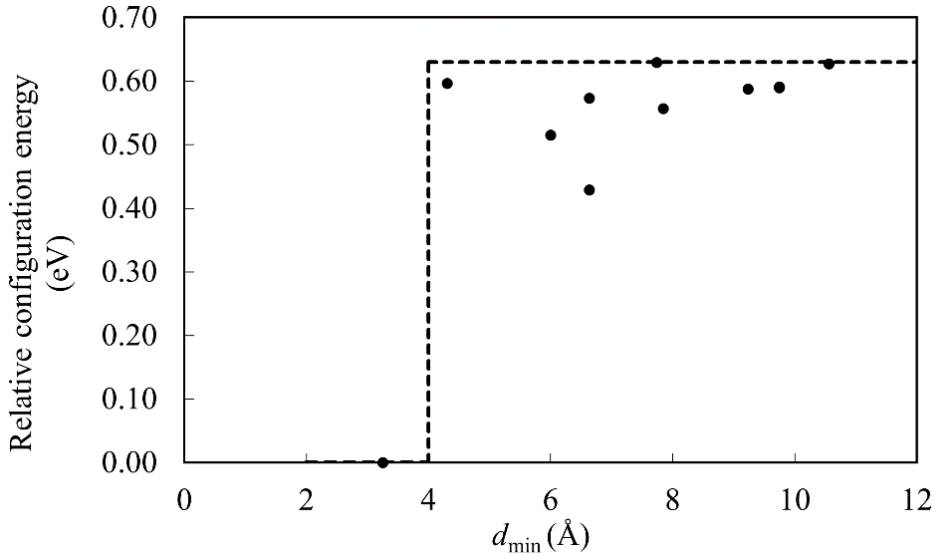

Figure 5.3. Plot of E vs d_{min} in the configurational space of bromate and lithium substituted calcite. The dotted line represents the simplified interaction model used for the statistical mechanical analysis: the horizontal dotted line is used to illustrate the pairing energy from this plot (0.63 eV). The vertical dotted line separates the region of nearest-neighbouring pairing from the longer distances at which interactions are considered negligible in the simplified interaction model (the precise position of this vertical line is arbitrary).

This simple "binding energy" approximation is useful for a discussion of the probability of nearest-neighbour pairing of bromates and compensating cations, based on statistical mechanics. Let us consider a very large supercell of n $CaCO_3$ formula units, where we introduce one cation and one anion substitution, in such a way that the molar fraction of substitution is $x=1/n$. This supercell can be in one of two states: the binding or pairing state, where the cation and anion substitution form a nearest-neighbour (NN) pair, which has energy $-\varepsilon$ and degeneracy χn (where ε is the binding energy between the oppositely charged substitutions, and χ is the coordination number between NN pairs); and the non-binding state, where the cation and anion substitutions are not in NN positions, which has energy zero and

degeneracy n^2-zn. Using a Boltzmann's canonical distribution, the probability of NN pairing is then:

$$P_{\text{pair}} = \frac{zn \exp\left(\frac{\varepsilon}{k_B T}\right)}{zn \exp\left(\frac{\varepsilon}{k_B T}\right) + n(n-z)} = \frac{1}{1 + (\frac{1}{zx} - 1) \exp\left(-\frac{\varepsilon}{k_B T}\right)} \quad (5.3)$$

where k_B is Boltzmann's constant and T is the absolute temperature of equilibration. The above expression is consistent with the equation for energy-dependent site occupancy derived by (Catlow, 1978). In principle it is possible to use a more sophisticated numerical approach, going beyond the minimal supercell to take into account mutual siting of many pairs (Smith et al., 2010), but the more approximate treatment is sufficient here to elucidate the pairing trends. In the limit of strong binding ($\varepsilon >> k_B T$) and not very low concentration x, the probability of pairing tends to 1, i.e. each substituted cation has a substituted anion in NN position. On the other hand, when the concentration of substitutions is very low, the probability is simply linear on x:

$$P_{\text{pair}} \approx zx \ \exp\left(\frac{\varepsilon}{k_B T}\right) \quad (5.4)$$

The probabilities of pairing at ambient temperature T=300 K as function of impurity concentration for each phase and compensating cation are shown in Figure 5.4. At very low concentrations the probability of pairing increases linearly with x following Eq (5.4). Above certain threshold concentrations, the pairing probability saturates at 1, which means that all impurity pairs will be found in nearest neighbour configurations. For calcite, these threshold concentrations are ~10^{-11} for Li, ~10^{-7} for Na, and ~10^{-5} for K. For aragonite, the corresponding threshold concentrations for pairing are even lower, following the same trend (Li>Na>K) with the nature of the compensating cation.

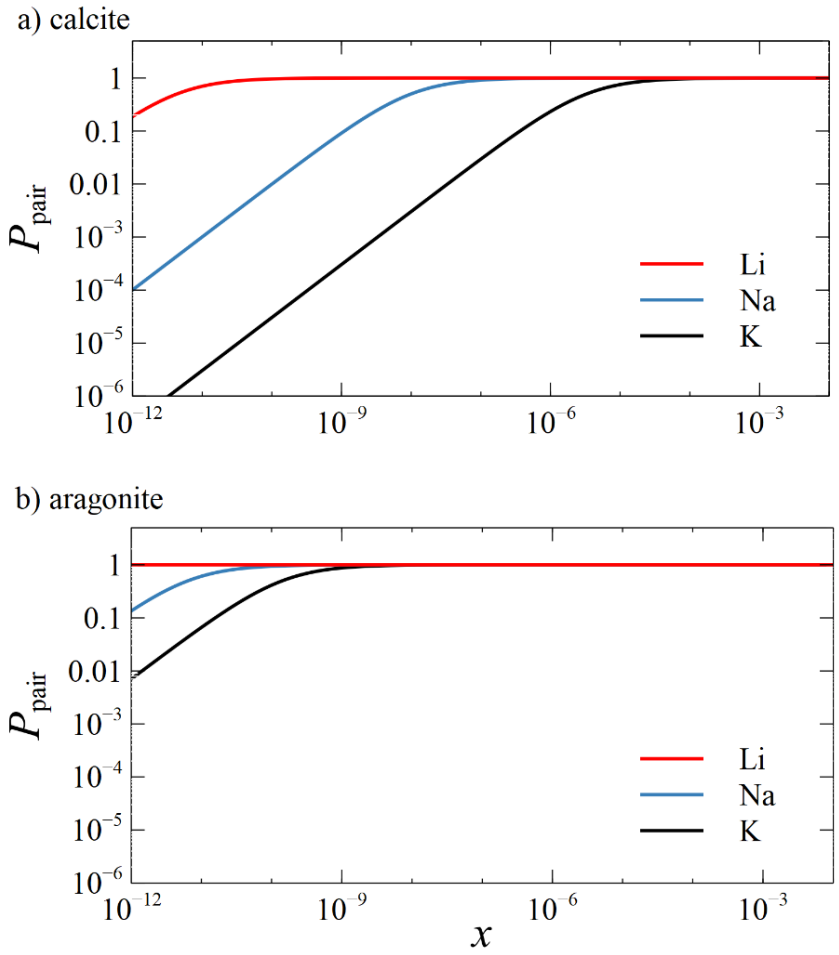

Figure 5.4. Pairing probability P_{pair} vs impurity (M^+ + BrO_3^-, where M=Li, Na, or K) concentration x at ambient temperature for a) calcite b) aragonite.

Precise bromate concentrations in naturally occurring speleothems are not reported in the literature. The most commonly studied trace element, sulphur, is known to appear in a range of concentrations from roughly 10 to 40 ppm (Frisia et al., 2008) following volcanic activity, though this figure could vary dramatically depending on magnitude of volcanic eruption. Marine carbonates have been shown to take up sulphate in the order of 10^3 ppm, two orders of magnitude higher than in speleothem archives (Busenberg and Niel Plummer, 1985). Our analysis above suggests that for concentrations of impurities in the order of several ppm, bromate impurities can be expected to be located in the immediate vicinity of a compensating cation.

The absolute values of the exchange energies reported in Table 5.2 are irrelevant, with the very high values mainly reflecting the loss of electrostatic stability associated with the aliovalent substitutions, when the isolated ions are taken as reference. But the trends across different cations and host phases are interesting. Compensated bromate incorporation in aragonite is slightly favoured over incorporation in calcite, which is somewhat unexpected

given that aragonite is a denser phase than calcite. This is in contrast with the behaviour observed for isovalent substitution of carbonates by tetrahedral oxyanions (Midgley et al., 2020), where elastic effects, and therefore density, was the primary driver of increased exchange energy.

The stability of the double substitution (anion + cation) is controlled by two factors: 1) a significant favourable electrostatic stabilisation is gained when the distance between the substituting ions is minimised; 2) an unfavourable elastic (lattice) strain energy that comes from distortion of the host crystal, this is particularly unfavourable when the distortion is accumulated in a small region of the crystal, i.e. when the ions are in a proximal configuration. Comparing the impurity binding energies of aragonite and calcite with the corresponding exchange energy differences, indicates that the main reason for a lower exchange energy in aragonite is the stronger binding energy between the two impurities. The binding energy is greater in aragonite because the higher phase density affords a shorter anion-cation distance, which is around 2.9 Å in aragonite and 3.2 Å in calcite, prior to structural relaxation. On the other hand, the higher density of aragonite also increases the elastic strain effects, which explains why the exchange energies for aragonite are still similar to those for calcite. The strain effects on the local structures around the impurities are illustrated in Figure 5.5 for calcite and in Figure 5.6 for aragonite. After relaxation, there is significant ionic reorganisation in aragonite, while in calcite ions occupy the lattice sites of calcium and carbonate without much distortion.

Another trend seen in Table 5.2 is the increasing substitution energy with increasing cation radius. This may be explained by considering the increased elastic strain imparted on the host crystal when incorporating a larger cation. In the case of K^+ compensation, in addition to the relatively high elastic cost of inserting a large cation, there is a reduction in the binding energy stabilization, due to the combined effect of the distortion fields of both substitutions. It is therefore likely that in geochemical environments, smaller cations are more favourable for charge compensation mechanisms.

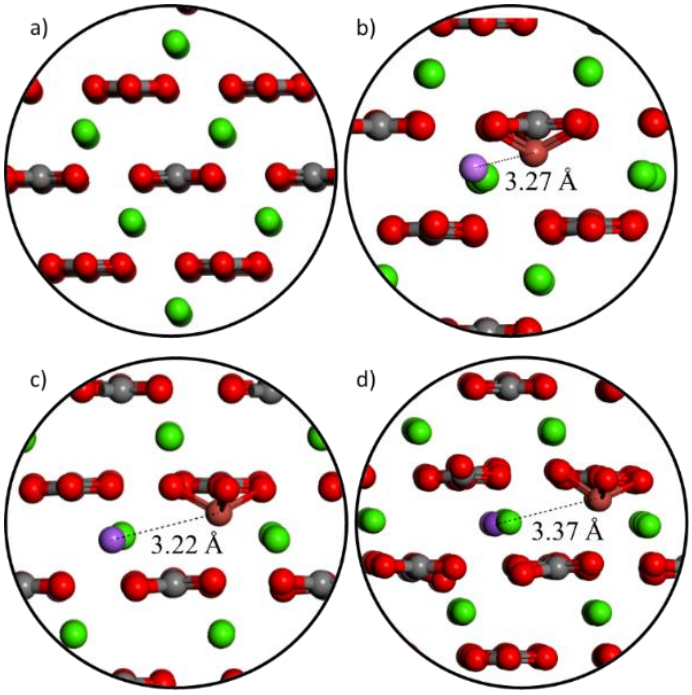

Figure 5.5. Optimized local geometries around the double substitution of M^+ and $(BrO_3)^-$ into calcite where **a)** is pure calcite and the others are bromate-substituted compensated by **b)** M = Li **c)** M = Na **d)** M = K.

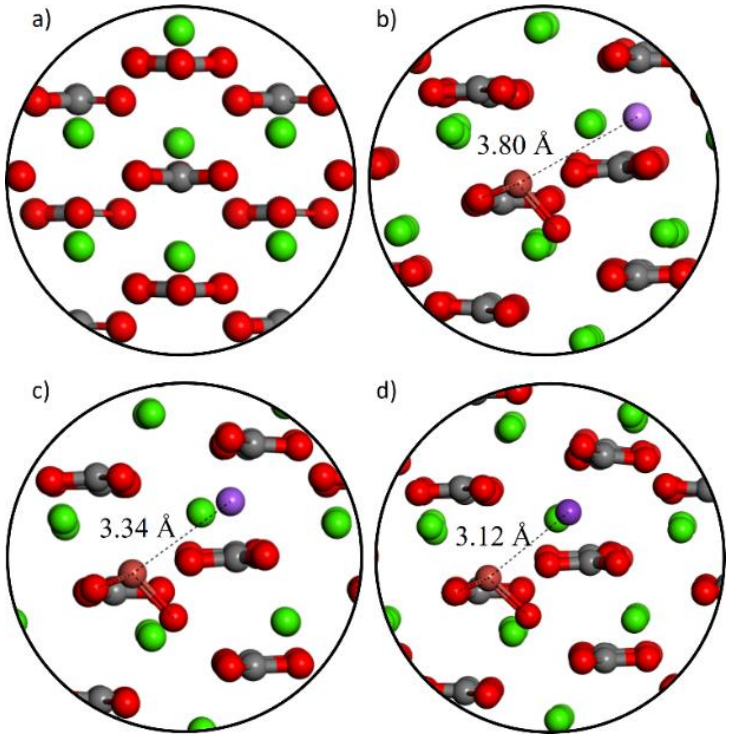

Figure 5.6. Optimized local geometries around the double substitution of M^+ and $(BrO_3)^-$ into aragonite where **a)** is pure aragonite and the others are bromate-substituted compensated by **b)** M = Li **c)** M = Na **d)** M = K.

Finally, we discuss the implications of the calculated solution energies, presented in Table 5.3. Whether a certain level of trace impurity incorporation is stable or not with respect to separation into a competing phase depends on the balance between the enthalpy of mixing:

$$\Delta H_{\mathrm{mix}} \approx x\Delta E_{\mathrm{sol}} \tag{5.5}$$

and the configurational entropy effect:

$$-TS_{\mathrm{conf}} = 2k_{\mathrm{B}}T[x\ln x + (1-x)\ln(1-x)] \tag{5.6}$$

that stabilises the mixing. The factor of 2 in eq. (5.6) appears because there are two independently disordered species with concentration x (when the concentration is well below the pairing threshold). As in previous work (Grau-Crespo et al., 2011; Midgley et al., 2020) the solubility limit or maximum concentration of impurity thermodynamically stable against phase separation can therefore be estimated from the solution energy:

$$x_{\mathrm{m}} \approx \exp(-\frac{\Delta E_{\mathrm{sol}}}{2k_{\mathrm{B}}T}) \tag{5.7}$$

For Na-compensated bromate substitution in calcite, this value is ~2×10^{-12}, and for the same substitution in aragonite is ~5×10^{-15}. The values for K-compensated substitutions are even lower (see Table 5.3). The low values suggest that bromate substitution at ppm level and even at ppb level in bulk lattice sites is metastable with respect to phase separation, at least when compensated by Na or K. This conclusion should be accompanied by some caveats. First, the solubility limits via other charge compensating schemes (e.g. with Li^+, or even H^+ which has not been considered here) might be much higher, leading to thermodynamically stable bromate substitution. Second, metastable bulk incorporation can still occur due to the complicated kinetics of mineral growth. For example, it is known that sulphate can partially incorporate in the bulk of calcite, even if the calculated solubility limits obtained for sulphate (Midgley et al., 2020) are very low (in fact much lower than the solubility limits obtained here for bromate). The higher solution limit of bromate compared to sulphate and molybdate is mainly due to the lower strain around the oxyanion impurity when the fourth, apical oxygen is not present, even if the trigonal pyramidal bromate ion is not a perfect fit to substitute the trigonal planar carbonate anion.

5.6. Conclusions

We have presented a computer simulation study of the incorporate of bromate ions as aliovalent substitutional impurities in the anion sites of calcium carbonate minerals calcite and aragonite. The chemistry associated to the incorporation of these species is more sophisticated than that of isovalent substitutions of carbonate by sulphate or molybdate, that we investigated previously (Midgley et al., 2021a; Midgley et al., 2020), due to the additional complexity associated to the possible charge compensating schemes, and the interaction of bromate impurities with the charge compensating anions. Our calculations provide a prediction of the pairing thresholds, i.e. the concentrations above which the binding energy between oppositely charged impurities at nearest-neighbour positions dominate over the entropic tendency to disassociate the impurities (at ambient temperature). We have demonstrated that the pairing is slightly stronger in aragonite than in calcite, and that if decreases down the group of alkali metals as compensating cations (Li>Na>K). Assuming thermodynamic equilibrium, it can be expected that bromate substitution at ppm-level occurs with a very high degree of pairing of the bromate ion with the compensating cations in nearest-neighbour positions in the lattice.

Based on our simulations, we expect cations of smaller radii to preferentially incorporate into the carbonate structure as they cause lower degrees of lattice strain on the host. Where there is low concentration of sodium or lithium in the local aqueous environment, bromate is unlikely to be stabilised in the carbonate phase, therefore counterion abundance may be an important consideration when interpreting bromine-containing trace-element records. An important scenario for future investigation is charge compensation by protons, which leads to interesting chemistry as the thermodynamics becomes pH dependent. The solution energy calculations indicate that bromate substitution compensated by sodium or potassium cations is metastable with respect to phase separation of the impurities as solid $NaBrO_3$ or $KBrO_3$ phases, respectively. However, the solubility limits of bromate in calcite and aragonite are still higher than those calculated for sulphate or molybdate, so the thermodynamic driving force for bromate to separate into a competing phase is relatively weaker. This analysis suggests a larger degree of bulk substitution of bromate in calcite and aragonite compared to sulphate and molybdate.

Our findings provide structural and atomic-level insight that may the interpretation of bromine records in speleothems. Our work highlights that bromine abundance may be carbonate phase specific, i.e. relative abundances of calcite and aragonite are important

considerations. Also, monovalent ions should be detected proximal to bromine impurities, where there is a strong thermodynamic preference for smaller ions. Abundance of lithium in the bromine-containing speleothem is likely to be representative of the amount of lithium in the local aqueous environment around the time of speleothem formation, which may significantly skew the detected levels of bromine compared with the deduced volcanic eruption magnitude. These are all key considerations when interpreting speleothem archive data.

5.7. Reflection

5.7.1. Limitations

The main limitation in the physical model used here is that only bulk substitutions were considered. Although we have reported already that substitution at carbonate-water interfaces is a better representation of the natural process for tetrahedral oxyanions, running molecular dynamics for the full configurational space of surface substitutions at the carbonate water interface would have introduced an inhibitively high computational cost. We discovered here; however that the bromate ion produces less elastic strain when incorporating in the bulk of calcite and aragonite, compared with sulphate and molybdate, which provides mitigation for our choice of simplified structural model (bulk substitution only).

5.7.2. Future work

Work presented in this chapter scratches the surface on a branch of potential computational geochemical investigation, which is highly relevant to the present thesis – disorder of ions in a doped crystal. Strong correlations between the distributions of oppositely charged impurities have been shown, because of the interplay between elastic and Coulombic effects. Though there is evidence in the literature for correlated substitution of multiple ions in calcium carbonates, this is something not touched upon in the present thesis. It would however, be possible to apply similar approaches to those reported in this chapter, to investigate the problem of structural disorder in calcite where there are multiple anion substitutions. Building on principles described in the present chapter, SOD could be used to investigate a double (or triple etc.) substitution event in the calcite bulk. As an example of an interesting investigation of this kind; two bromate ions could be incorporated in the calcite bulk, which would be charge-corrected by a calcium ion vacancy (or double monovalent anion substitution). Electrostatically, the two ions would be attracted to the calcium ion vacancy; however accumulation of the impurities would likely introduce unfavourable elastic effects, similarly to those reported in the present chapter.

Understanding this kind of chemical synergy is vital in the interpretation of speleothem records, as it is likely to be a key factor in determining the location of chemical tracers in the mineral due to positive (or negative) feedback influences from impurities. As an aside, developing machine learning methods could provide a computationally inexpensive route to energy calculation of the entire configurational space where multiple correlated substitutions produce a high number of asymmetric configurations. Though not specific to the present geochemical applications, this concept has been proven by our group in alternative solid solution-based problems using the SOD code (Midgley et al., 2021c).

6. Conclusions

6.1. Overall conclusions from previous chapters

This thesis work has offered insights into the thermodynamics of anionic trace element incorporation in calcium carbonates, in particular: bromate, sulphate, and molybdate impurities. Though this is only a very small subset of possible tracer species, the principles described in the preceding chapters may be applicable to a much wider range of substituting ions in calcium carbonates and possibly other minerals. Useful insights into how oxyanion tracers might incorporate into calcium carbonates have been provided, though we have only scratched the surface in uncovering the necessary links to equate experimentally detected speleothem-based trace elements with details of the volcanic eruption.

It has been shown that molybdenum may incorporate into calcium carbonate in its tetrahedral oxyanion form, molybdate, with similar structural and thermodynamic behaviour as the more widely studied sulphate impurity. It has also been demonstrated that bulk crystalline incorporation of both sulphate and molybdate is unlikely, but that thermodynamics of incorporation are strongly influenced by the host phase density. Vaterite is able to incorporate some sulphate at concentrations up to several ppm while remaining thermodynamically stable with respect to phase separation with formation of calcium sulphate. However, more generally the incorporate of both sulphate and molybdate in calcium carbonate minerals is only metastable with respect to phase separation. Some bulk crystalline incorporation may still be expected due to a high kinetic barrier to anion diffusion, despite an overall thermodynamic driving force favouring the formation of a separate calcium sulphate/molybdate phase.

Molybdate and sulphate incorporate at the calcite (10.4) terrace preferentially to the crystalline bulk. Common stepped imperfections in calcite (10.4) encourage trace element incorporation to an even greater extent than the terrace. Though tetrahedral oxyanions are unsuited to bulk incorporation, surface regions and grain boundaries within speleothems can provide an effective means of oxyanion incorporation in the speleothem. We have also demonstrated that liquid water can have a potentially important effect on the thermodynamics of exchange, suggesting that dynamics methods in conjunction with explicit solvents are important for simulations of this kind of interface.

Whereas sulphur and molybdenum levels in speleothems are useful records of past volcanic activity, experimental work has concluded that bromine seems to provide the most reliable record. Our simulations have helped to understand the different chemistry of incorporation of bromine with respect to sulphate and molybdate. It has been demonstrated here that bromine, in its trigonal pyramidal form, bromate, may be better suited to crystalline bulk

incorporation when compared with sulphate and bromate, because absence of an apical oxygen atom causes less structural distortion in the calcium carbonate host phase. Bromate incorporation introduces a charge imbalance to the crystal, which is likely corrected in nature by incorporation of a singly charged cation. Whether this cation is lithium, sodium, or potassium, speleothem samples are likely to be found with proximal M^+/BrO_3^- configurations in the crystal structure. Distal configurations stabilised by the configurational entropy may only found at much lower concentrations (<<1 ppm in calcite and << ppb in aragonite). The solubility limit allows for direct comparison of thermodynamic stability with values obtained for sulphate and molybdate in Chapter 3. Solubility limits were found to be much higher in bromate (+ sodium) compared to bulk incorporation of sulphate (and molybdate), indicating that bromate substitution in calcite is more stable with respect to its own competing phase, compared to sulphate. It is known that at least some sulphate incorporates in the crystalline bulk despite a thermodynamic driving force to phase separation. It is therefore possible (and likely) that bromate (+ sodium) should bulk incorporate to a greater extent than sulphate. This is in line with conclusions drawn in the work by Badertscher (Badertscher et al., 2014), where it was suggested that bromine is more sensitive as a paleovolcanic tracer when compared to sulphur and molybdenum.

The field of paleoclimatology concerned with using speleothems as archives is relatively new and at present is largely unexplored from a chemical perspective. Though exciting and pioneering works by various research groups around the world are underway, extraction of high-resolution paleo-atmospheric information from speleothems is still not possible. We have demonstrated in prior chapters that detected trace-element concentrations are extremely unlikely to show linear correlations with eruption parameters such as magnitude or location. There are multiple complex chemical factors that affect the journey of the trace element species from volcano to signal on an XRF measurement, many of which are not fully understood yet. Earth scientists have begun to investigate various different aspects of this complex link such as atmospheric dynamics of the volcanic ejecta and how this materials leaches through soil and subterranean layers. Recent research, including this thesis, has also demonstrated that impurity incorporation in speleothems is critically dependent on the precise chemistry of the trace element and speleothem host. Examples of such chemical considerations may be: impurity oxidation state, impurity coordination geometry, and the microstructure of the calcium carbonate.

Consideration of the speleothem microstructure is essential for rescaling detected trace-element concentrations, to accurately deduce the concentration of impurity in the local

environment following the volcanic eruption. For example: low concentrations of sulphate in highly crystalline calcite may be misleading. As we have shown in preceding chapters, high crystallinity is unlikely to accommodate trace tetrahedral oxyanions, therefore the true concentration of paleo-atmospheric sulphate may be much higher than indicated by a highly crystalline speleothem sample. The attenuated eruption may therefore have been larger, or closer to the speleothem site than it may seem. Another example of this kind of consideration is that minerals that have undergone a phase transition from vaterite to calcite early in their lifetimes may show higher trace element concentrations compared with minerals which formed calcite directly. This is because the temporary presence of vaterite may have held tracers such as sulphate in the local chemical environment for a longer time, due to higher trace element affinity for this phase.

Geochemically, we know that some specific ions can stabilise the formation of calcite steps, and it has been shown in the present work that such steps can encourage substitution of certain anions. There may therefore exist a complex interplay where small amounts of sulphate in the local aqueous environment are able to stabilise surface defects, which in turn will encourage higher-than-usual concentrations of trace element to become incorporated in the carbonate. Such complex chemical synergy is likely to have an effect on detected trace element concentrations, though this is not yet well understood.

Another implication of the present work is that surface defects may affect the partitioning of trace impurities differently to each other. For example, the stepped calcite surface strongly stabilises molybdate impurities, while the same impurity is highly unstable in the crystalline bulk. By contrast, the stepped surface does not affect the partitioning of sulphate as significantly, meaning that a higher proportion of sulphate could be found in crystalline bulk regions of the carbonate mineral, compared to molybdate. The two impurity ions are therefore affected differently by the same host structure, meaning that ratios of trace ions could be skewed from their true paleo-atmospheric abundance.

Chemical form and impurity location in the host crystal are clearly important factors which should be considered when analysing speleothem archives. At present, speleothem chemical analysis techniques employed in this field are only used to detect presence of elements, while chemical factors such as impurity oxidation state and physical location in the host crystal are mostly overlooked. Work presented here has shown the importance of such atomic-scale and thermodynamic considerations in extracting maximum paleoclimate information from speleothem archives. By incorporating the kind of research detailed in the present thesis with

experimental measurements, paleoclimatologists may be able to uncover more of the sophisticated link between detected impurities and eruption parameters, maximising information obtainable from speleothem-based archives. However, until the precise geochemistry of the speleothem-impurity system is understood, there will remain a high degree of uncertainty in paleovolcanic insights obtainable from such archives.

The task of relating speleothem-bound impurities to specific paleoenvironmental conditions therefore poses a significant challenge to researchers. How then, can speleothems be effectively used as archives of paleovolcanism? The answer is perhaps that far greater volumes of research are needed in the highly underexplored field of geochemical paleoclimatology.

6.2. Assessment of methods

One of the objectives outlined at the beginning of the thesis was to assess the advantages and limitations of quantum-mechanical methods to investigate the investigation of these impurities in carbonate minerals. Summarised here are the main conclusions on this point.

This work employed simulations based on density functional theory (DFT) in the Kohn-Sham (KS) formulation. KS-DFT is a state-of-the art technique for approximate quantum-mechanical quality solutions to the electronic Schrödinger equation. The methods described in all chapters were well received by reviewers during the publication process.

One novel feature of this project was the use of *ab initio* molecular dynamics (AIMD) for the specific application of geochemistry. AIMD has historically been the domain of very small physical models, specifically requiring accurate dynamical information in a very short time period. AIMD has not seen widespread use until very recently because of the extreme computing costs associated even a small number of steps. It follows that the computational study of calcium carbonates often involves interatomic potential methods, which are very well parameterised for carbonate/water systems as well as being far lower in computational cost compared to AIMD, often by several orders of magnitude.

The present thesis employed AIMD methods to take advantage of recent developments in algorithms and the availability of extremely fast supercomputers (all the work was performed in Tier 1 and Tier 2 national supercomputer facilities), and because the chemical impurity species that were investigated are not well parameterised in the context of calcium carbonates. Transition metals (such as molybdate) and species containing an axially oriented lone pair (such as bromate) are typically very difficult to parameterise using interatomic potentials, because their complex structures cannot be easily approximated using ball-and-spring type

approaches. Though it may be possible to generate interatomic potentials following a small amount of *ab initio* calculation, it was decided, early on, to commit to a method of running full-length AIMD simulations, with the anticipation of highly accurate energies obtained. The benefit of running AIMD for all systems (assuming a consistent level of theory used) is that all results are directly comparable. By contrast, classical methods tend to be far less transferable, as they are fitted for very specific chemical systems.

Static calculations like those performed for the bromate-substituted systems are appropriate for that kind of configurational analysis. This investigation looked at bromate incorporation in crystalline solid calcium carbonates, which was found to be more likely than the previously investigated sulphate and molybdate ions, due to the different shape of the anion. Configurational analysis in the SOD code also meant that many hundreds of DFT calculations had to be run. Attempting this approach using AIMD methods would have incurred an inhibitively high computational cost.

The main limitations in the methods and outcomes have already been discussed in each chapter, though it is worth reiterating here as one of the methodological conclusions from this thesis, before making practical recommendations on future research relating to the context of the present thesis (Chapter 7). In particular, the use of AIMD for the study of calcium carbonate – water interfaces, which is not commonly employed when studying these systems. Though AIMD provides excellent transferability for systems involving exotic species such as molybdate and bromate, the study of these systems using AIMD has some practical issues which were not known to our group prior to this work. These relate to the extreme computational cost of running long-enough AIMD simulations to achieve convergence of thermodynamic quantities. Although we managed to perform very long simulations (up to 120 ps) for the results presented here, this limitation meant that we were not able to study other chemical scenarios (e.g. bromate at surfaces, other oxyanions, other combinations of impurities and host phases) that we would have liked to investigate.

To summarise the issues with computational cost, the computational cost per timestep of AIMD required that no more than ~800 atoms be used for a solid/liquid interface system, even when running on the UK's national Tier 1 supercomputer, ARCHER (details of this machine are given in Chapter 2). The use of small models to accommodate longer computing times actually makes the system very slow to equilibrate with respect to MD simulation time, because the magnitude of energy fluctuations are larger with respect to the total energy of the

cell. It is therefore necessary to run upwards of 100 ps of AIMD, to achieve concurrent MD energies.

In this regard, my conclusion from the work in this thesis is that AIMD, although an excellent technique for many areas of computational materials science, is still not suitable as a stand-alone technique for the large-scale investigation of the complex chemical environments of impurities at the mineral-water interface. It is instead recommended that despite the significant time investment and technical challenge, researchers develop effective interatomic potentials describing the systems of interest, to extend the simulations to more complex environments.

Development of interatomic potentials is a difficult undertaking, and may not be possible for systems with severely aspherical geometry, such as bromate. However, the functional benefits of developing interatomic potentials are significant for the generation of statistically converged MD simulations at a low computational cost, as well as minimising the environmental impact of high-performance computing. Research efforts into the development of interatomic potentials are highly recommended, despite advances in computing power which make AIMD an attractive technique. A strategy that could work well here is to use machine-learned interatomic potentials, which can be integrated efficiently with AIMD. For example, Jinnouchi and co-workers (Jinnouchi et al., 2019) have described a strategy for on-the-fly machine learning force field generation, which is being implemented in the VASP code, but this is still not available in the distributed version.

The above recommendations are based on current supercomputing facilities available to UK-based researchers. However in the next ~10 years, exascale computing may mean that running AIMD up to millions of steps may become trivial, allowing for more complete AIMD representation of very large carbonate-water interface models (>> 800 atoms). This could address shortcomings of AIMD for the present use-case, which have been highlighted throughout this chapter.

6.3. Final comments

Collaborative projects between archaeologists (paleoclimatologists) and materials scientists are relatively rare, though there is the potential for large amounts of high-impact research from this type of collaboration. Speleothems present a myriad of geochemical phenomena including crystal substitution as studied in the present thesis, but also with phenomena such as fluid inclusions where paleoclimate tracers are found in small aqueous environments. Future work by computational materials scientists in the field of paleoclimatology has the potential to be expansive, covering many aspects of speleothem chemistry that are not currently understood. By increasing interdisciplinary research of this kind, it is plausible that geological archives such as speleothems could one-day provide high-resolution climate data spanning millennia.

7. Recommendations for Future Work

7.1. Introduction

Recommendations for future work are given here. Some of these recommendations are based on work attempted during the course of the present PhD, but which could not be completed due to limitations in computational cost. The magnitude of the computational costs involved in running AIMD for the study of geochemical interfaces was not known to our group *a priori*. As a result of unforeseen computational expenses, we had to repurpose much of the supercomputing time that was originally intended for other research projects so as to be able to obtain statistically converged AIMD runs for Chapter 4. However, with alternative computational strategies, these partially completed research endeavours could provide interesting insights on the behaviour of trace element incorporation in speleothems, fitting with the context of the present thesis. A specific list of recommendations for future work is given here:

I) To develop interatomic potentials that can accurately describe sulphate, molybdate, and bromate at a calcium carbonate/water interface.
II) To consider sulphate and molybdate incorporation at the obtuse calcite (10.4) step.
III) To consider alternative tracer species which are isostructural to bromate, sulphate, and molybdate.
IV) To consider alternative calcite terminations as possible substitution sites for sulphate and molybdate.
V) To consider possible surface incorporation mechanisms in aragonite.

7.2. Development of interatomic potentials

It has already been mentioned in Chapter 6 that the development of interatomic potentials could be beneficial for the present area of research, to alleviate extreme computational costs associated with AIMD. An example of why effective parameterisations in the contexts of oxyanion-carbonate-water systems could be extremely beneficial for future work following this thesis, is given here.

Parameterisation of bromate in aqueous carbonate environments would be extremely useful for future work in the context of this thesis, because we have shown that it has excellent potential as a paleo-volcanic chemical tracer. We have not investigated bromate bulk/surface partitioning behaviour in the present thesis, but this would be an interesting investigation. Based on our analysis in Chapter 5, the attenuated charge correction could logically involve a singly charged cation located in a proximal lattice configuration to the bromate anion. The

atypical charge on these species would likely cause large amounts of structural reorganisation at the carbonate/water interface, therefore extremely long MD simulations would be required. This is an excellent example of why bromate parameterization for aqueous carbonate environments would be extremely beneficial, and how such parameterizations could be effectively employed.

Sulphate, molybdate, and bromate all have excellent potential and volcanic tracers, and are being detected already by synchrotron radiation techniques for elemental detection. It is vital that chemical effects continue to be probed using atomic-level simulation so that researchers can better understand the behaviour of these tracers. Effective parameterisation could therefore allow in depth investigation of dynamic binding effects at both calcite and aragonite hosts, whilst incurring a fraction of the computational cost associated with AIMD.

7.3. Sulphate and molybdate incorporation at the obtuse step

Chapter 4 reports sulphate and molybdate incorporation at the calcite (10.4) terrace at the calcite acute step. It was originally intended to also consider the calcite obtuse step, which forms typical 'pits' in the calcite (10.4) terrace along with the acute step. A structural model of this kind is shown in Figure 7.1.

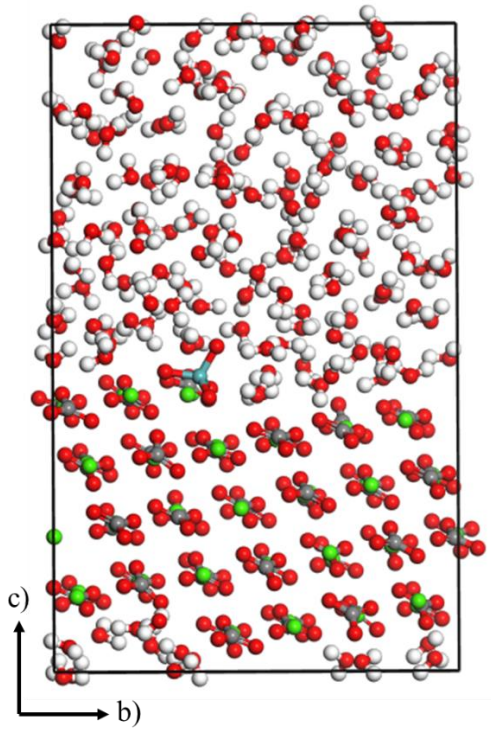

Figure 7.1. Obtuse stepped surface structural model containing molybdate impurity.

These structural models were subject to AIMD in the CP2K code using identical procedures as those reported in Chapter 4; however only 30-40 ps of AIMD was obtainable using the available computing resources. Preliminary exchange energies are therefore not reported here, as these are derived from MD simulations that are not fully equilibrated and are not reliable.

The decision to proceed only with the acute step was driven primarily by restricted computing resources and the requirement for extended AIMD simulations to achieve well-equilibrated systems. We chose to proceed with the acute step over the obtuse step because of the greater stability of the acute step with respect to dissolution, which in theory makes it a better candidate structure for the host phase (de Leeuw et al., 1999).

7.4. Oxyanion trace elements at calcite (10.4) surface

Following the early indication that calcite surfaces offered enhanced thermodynamic potential for trace element incorporation compared to the crystalline bulk, other trace element impurities were investigated at the calcite (10.4) terrace for their potential as geochemical tracers. These elements, though not yet exploited as paleovolcanic tracers, could have the desired thermodynamic behaviour to be potential volcanic tracers. The idea of this project was to help identify materials ejected during volcanic activity that could act as potential new tracers to be exploited by experimental researchers. To do this, their sensitivity compared to know tracer impurities was assessed, based on calculated thermodynamics of incorporation.

Three main categories of oxyanionic trace element were studied: those with single negative charge (nitrate, chlorate, bromate, iodate), double negative charge (selenate, chromate, tungstate) and triple negative charge (phosphate and arsenate). The single negative charged species were structurally analogous to either carbonate (trigonal planar) or bromate (trigonal pyramidal), which we showed in Chapter 5 to be a suitable substituting ion for carbonate. All doubly and triply charged ions were structurally analogous to sulphate and molybdate (tetrahedral).

All ions were subject to AIMD in the CP2K code, analogously to simulations reported in Chapter 4. Charge imbalances introduced by substitution of ions were corrected using the homogeneous charge background provided by CP2K. All results discussed here are 'preliminary', i.e. they have not been derived from fully equilibrated AIMD simulations. The present figures were calculated from around 30-40 ps of AIMD in each case, and should be interpreted with caution. They do; however provide some basic indication for how incorporation thermodynamics is affected by charge imbalances.

Table 7.1 reports energy values for the ion exchange process involving the bulk crystal, calcite (10.4) terrace and the aqueous phase. Ion exchange energies were calculated using identical methodology to that reported in Chapter 4. The ions are presented as ascending from monovalent ions through divalent ions to trivalent ions. Within each valency subset, elements are ordered by their *X*-O equilibrium bond length. The ion exchange energy vs the charge on the trace-element oxyanion is plotted in Figure 7.1. There is a clear trend between the exchange energy and the charge on the substituting anion and the exchange energy.

Table 7.1. Preliminary Bulk/Liquid, Surface/Liquid and Bulk/Surface exchange energies.

Species	**d[*X*-O] (Å)**	**ΔE_{bulk} (eV)**	**ΔE_{surf} (eV)**	**ΔE_{seg} (eV)**
Nitrate	1.26	-4.29	-0.74	3.55
Chlorate	1.51	-4.37	-1.75	2.63
Bromate	1.68	-4.53	-1.20	3.32
Iodate	1.85	-3.64	1.38	5.03
Sulphate	1.51	0.45	0.36	-0.09
Chromate	1.66	0.43	-0.19	-0.62
Selenate	1.68	1.55	0.83	-0.72
Molybdate	1.80	1.55	0.30	-1.25
Tungstate	1.84	1.33	-0.25	-1.57
Phosphate	1.59	5.01	0.77	-4.24
Arsenate	1.76	5.78	1.52	-4.26

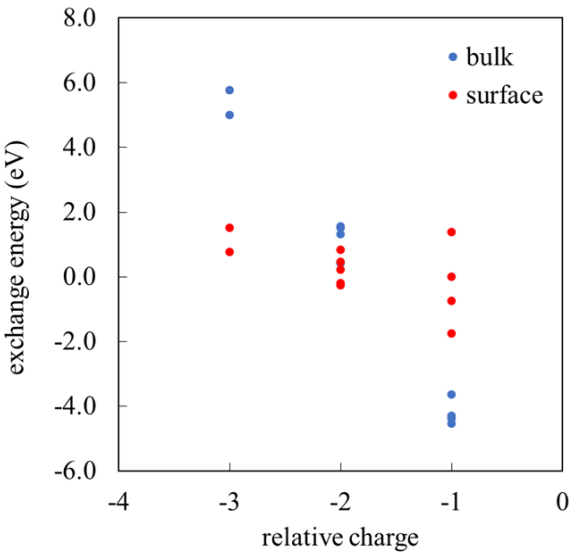

Figure 7.2. Preliminary ion exchange energies vs ion charge.

Exchange energies involving the monovalent ions appear to be extremely favourable. We assessed the possibility that this was a product of the homogeneous charge background provided by CP2K, though we only found evidence of an effect in the order of < 0.1 eV here. It is possible then that the extremely low exchange energies may be attenuated to the charge discrepancies between the substituting ions and the relative hydration thermodynamics. Nitrate offers the simplest explanation. As it is trigonal planar in geometry (like carbonate), it incorporates in the host crystal causing very little elastic strain. This can be seen in the MD snapshot geometries reported in Figure 7.3. The main energetic driver to its incorporation therefore comes from the fact that the carbonate ion is more stable in the aqueous environment, because it is a double negative ion compared with the single negative nitrate ion, which should command a comparatively weaker hydration environment.

The reverse effect is also a possible explanation for trends seen in the trivalent anions. Here, the impurity is structurally analogous to all of the divalent (tetrahedral) oxyanions, yet they show much higher exchange energies, particularly in the bulk. There are possible inaccuracies from incomplete AIMD simulations; however there does seem to be a trend at this preliminary stage. The explanation for this is again likely to be because of the difference in hydration environments in the aqueous phase. In this case however, the trivalent ion is stabilised in the aqueous phase compared to the divalent carbonate, which creates an energy barrier to substitution. These effects are lesser when substitution occurs at the surface, because hydration effects for the reference ion are mitigated by partial hydration effects surrounding the impurity at the surface. Such partial hydration effects have been introduced in Chapter 4.

Hydration effects discussed here could be confirmed using RDF analysis; however we have not included this analysis in the present section because AIMD simulations are not fully equilibrated, therefore subtle differences between RDFs would be meaningless. The incomplete AIMD simulations may also be impacting the calculated preliminary exchange energies, and may partially explain the very low energies obtained for the monovalent substitutions.

Small differences between ions of the same charge number have not been discussed here, as these differences are subtle, and are likely to be influenced heavily by the incomplete AIMD simulations. Overall, strong conclusions should not be drawn from this work; however there may be a vague indication that monovalent ions are best-suited as tracer species because of their favourable preliminary thermodynamics of incorporation.

Figures 7.3-7.7 below contain snapshots from AIMD simulations reported in this section. Though these are not fully equilibrated yet, the equilibration is likely to be in the soft partial hydration environments around each of the impurities. The geometry of the solid components is quite reliable, as solids typically equilibrate their structures much faster than liquids. It can be seen throughout that there are extreme structural similarities between many of the oxyanions, meaning that their incorporation thermodynamics may be expected to be quite similar.

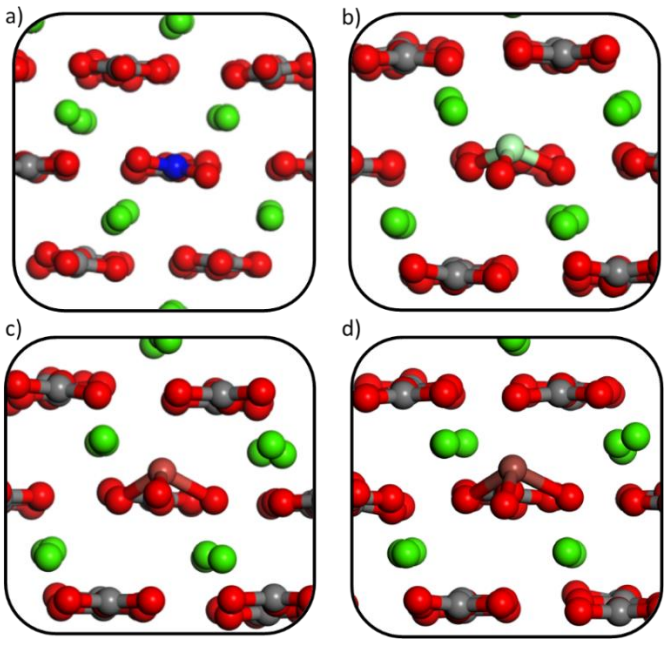

Figure 7.3. MD snapshot for XO_3^- in bulk calcite, where X = **a:** N, **b:** Cl, **c:** Br, **d:** I.

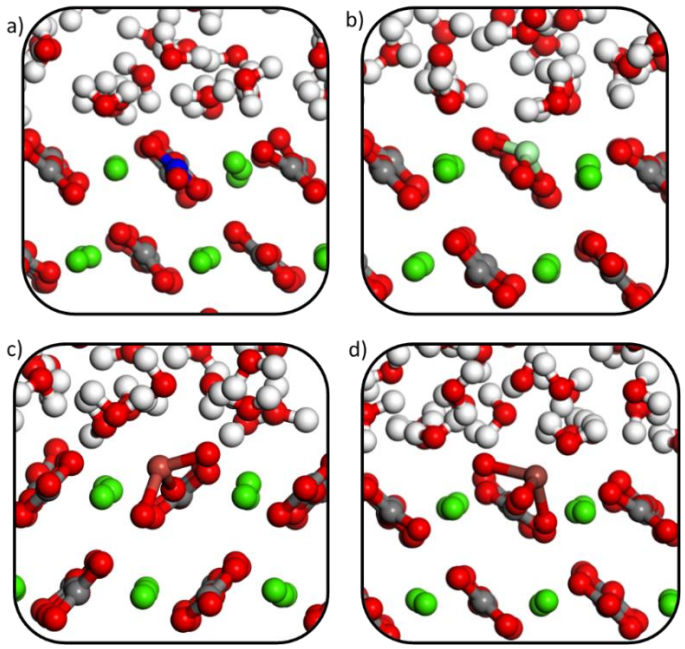

Figure 7.4. MD snapshot for XO_3^- at calcite (10.4) terrace, where X = **a:** N, **b:** Cl, **c:** Br, **d:** I.

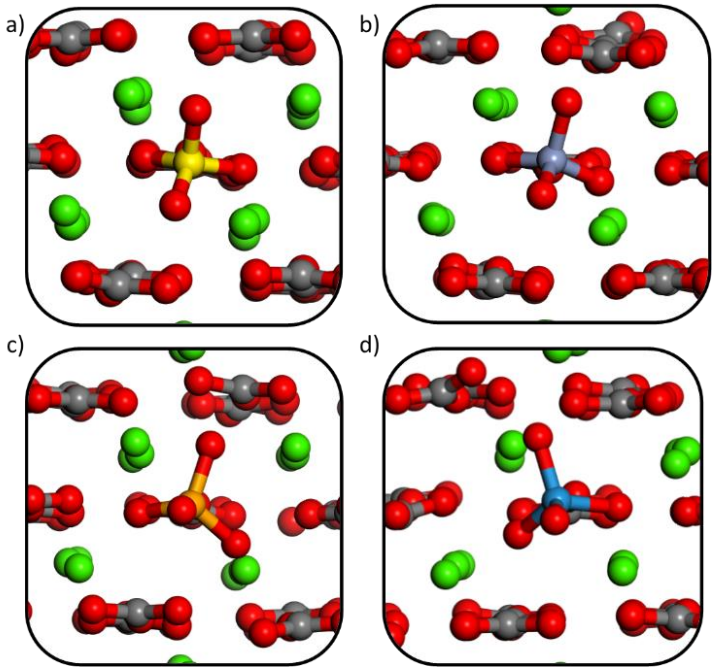

Figure 7.5. MD snapshots for XO_4^{2-} in bulk calcite, where X = **a:** S, **b:** Cr, **c:** Se, **d:** W. N.B geometry for molybdate excluded for brevity.

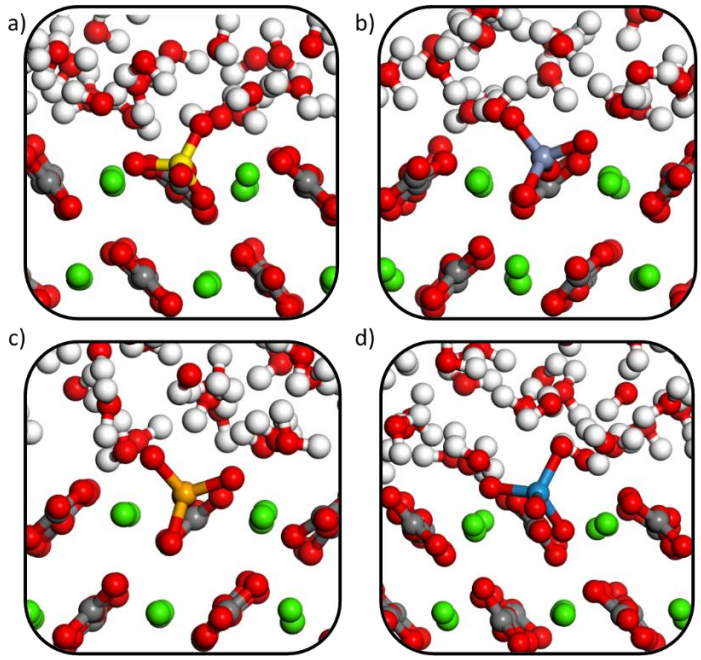

Figure 7.6. MD snapshots for XO_4^{2-} at calcite (10.4) terrace, where X = **a:** S, **b:** Cr, **c:** Se, **d:** W. N.B geometry for molybdate excluded for brevity.

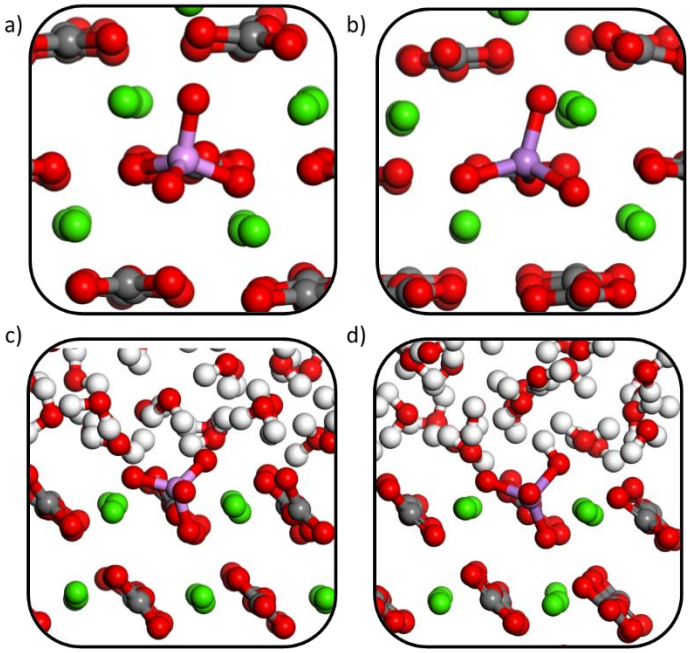

Figure 7.7. MD snapshots for XO_4^{3-} in calcite bulk and (10.4) terrace, where X = **a:** P_{bulk}, **b:** As_{bulk}, **c:** P_{surf}, **d:** As_{surf}.

The incomplete study of oxyanions at the (10.4) terrace is an interesting case which presents an opportunity for identifying new and interesting potential tracers, as well as defining the effect of impurity-ion charge on incorporation behaviour. For this project, there is no obvious classical MD analogue, because there is a very high number of heavy elements for which it is difficult to generate effective parameterisations. Empirically derived potentials would also mean that direct comparison could not be made between different trace-element species. Another way to address the investigation while reducing computational expense would be to run AIMD simulations of surface models without the presence of liquid water. Though the physical model would be less accurate, the removal of the liquid would drastically reduce the computational expense required for each simulation, as solids are known to equilibrate very quickly. This kind of 'dry AIMD' approach as introduced in section 7.4 would still require vast computational resources and run time, though it may reduce overall computing costs significantly while allowing for interesting thermodynamic behaviour to be discovered.

7.5. Calcite (00.1)

The calcite (00.1) termination was considered as an alternative site to the (10.4) termination used throughout the present thesis. As the (00.1) termination can either be calcium-terminating or carbonate-terminating, the periodic slab model generated upon cleavage is unphysically dipolar. To address this in the model, a method of surface reconstruction is employed, which is commonplace in computational materials science. Half of the atoms from the top surface are manually moved to the bottom surface, alleviating dipolarity. This however has the effect of leaving some surface vacancies in the model, which cause massive ionic reorientation during MD simulation. Such structural reorganisation of the solid surface is itself slow, but also the surrounding water cannot form stable hydration layers until the solid has structurally equilibrated. This effect can be seen in Figure 7.8, which shows a snapshot of the AIMD simulation after 20 ps for the calcium-terminating (00.1) interface model. A significant reorientation of the subsurface carbonate ions can be seen, which is a primary driver to the extreme computational cost of these preliminary simulations. In models of the (10.4) surface, the solid component of the interface is typically well equilibrated after a short period of simulation time; however this is not achieved in the (00.1) models. It is likely therefore that these simulations would require substantially more than the 100-120 ps required to equilibrate the (10.4) surface models. We therefore discontinued this project owing to extreme computing requirements.

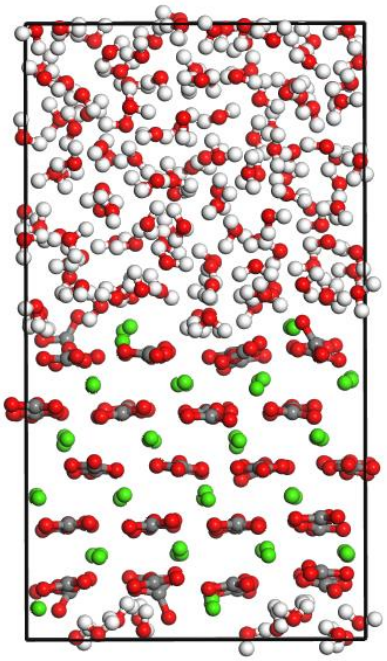

Figure 7.8. MD snapshot of the calcite (00.1) interface following 20 ps of simulation time.

7.6. Aragonite surfaces

Calcium carbonate surfaces have exclusively been represented by the calcite phase throughout the present thesis. Aragonite surfaces were investigated as a potential host environment by a summer project student, Aidan Robinson. This summer studentship was funded by the Scenario NERC DTP at UoR and was primarily supervised by Scott Midgley.

This work used analogous AIMD techniques to those reported in Chapter 4 to represent the carbonate/water interface for the aragonite (11.0) and (01.1) surfaces. Aragonite provides an interesting case study for the bulk/surface segregation mechanisms of sulphate impurities reported in the present thesis, because the higher phase density of aragonite should provide even stronger partitioning of the impurity between bulk and surface regions.

The original idea of this project was to employ DFBT semi-empirical methods to study the incorporation of sulphate in the bulk and at the surface of aragonite, with reference to an ion in water. However, representation of liquid water by DFTB is not accurate enough for production-level simulations, hindering this method for the present use. Full AIMD in CP2K was attempted to solve this issue; however computing costs associated with this kind of simulation meant that this project was unable to continue.

References

ADDADI, L., RAZ, S. & WEINER, S. 2003. Taking Advantage of Disorder: Amorphous Calcium Carbonate and Its Roles in Biomineralization. *Advanced Materials,* 15**,** 959-970.

AFFOLTER, S., HÄUSELMANN, A., FLEITMANN, D., EDWARDS, R. L., CHENG, H. & LEUENBERGER, M. 2019. Central Europe temperature constrained by speleothem fluid inclusion water isotopes over the past 14,000 years. *Science Advances,* 5**,** eaav3809.

ANDERSEN, H. C. 1980. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics,* 72**,** 2384-2393.

ANDERSSON, M. P., DIDERIKSEN, K., SAKUMA, H. & STIPP, S. L. S. 2016. Modelling how incorporation of divalent cations affects calcite wettability–implications for biomineralisation and oil recovery. *Scientific Reports,* 6**,** 28854.

ANISIMOV, V. I., ZAANEN, J. & ANDERSEN, O. K. 1991. Band theory and Mott insulators: Hubbard U instead of Stoner I. *Physical Review B,* 44**,** 943-954.

ARCHER, T. D., BIRSE, S. E. A., DOVE, M. T., REDFERN, S. A. T., GALE, J. D. & CYGAN, R. T. 2003. An interatomic potential model for carbonates allowing for polarization effects. *Physics and Chemistry of Minerals,* 30**,** 416-424.

ARROYO-DE DOMPABLO, M. E., FERNÁNDEZ-GONZÁLEZ, M. A. & FERNÁNDEZ-DÍAZ, L. 2015. Computational investigation of the influence of tetrahedral oxoanions (sulphate, selenate and chromate) on the stability of calcium carbonate polymorphs. *RSC Adv.,* 5**,** 59845-59852.

ATKINS, P. W., DE PAULA, J. & KEELER, J. 2018. *Atkins' Physical Chemistry,* Oxford, UK, Oxford University Press.

ATKINS, P. W. & FRIEDMAN, R. 2011. *Molecular Quantum Mechanics,* Oxford, Oxford University Press.

BADERTSCHER, S., BORSATO, A., FRISIA, S., CHENG, H., EDWARDS, R. L., TÜYSÜZ, O. & FLEITMANN, D. 2014. Speleothems as sensitive recorders of volcanic eruptions – the Bronze Age Minoan eruption recorded in a stalagmite from Turkey. *Earth and Planetary Science Letters,* **392,** 58-66.

BALAN, E., BLANCHARD, M., PINILLA, C. & LAZZERI, M. 2014. First-principles modeling of sulfate incorporation and 34S/32S isotopic fractionation in different calcium carbonates. *Chemical Geology,* 374-375**,** 84-91.

BAR-MATTHEWS, M., AYALON, A., KAUFMAN, A. & WASSERBURG, G. J. 1999. The Eastern Mediterranean paleoclimate as a reflection of regional events: Soreq cave, Israel. *Earth and Planetary Science Letters,* 166**,** 85-95.

BLANCHARD, D. L. & BAER, D. R. 1992. The interactions of Co, Mn and water with calcite surfaces. *Surface Science,* 276**,** 27-39.

BLOCHL, P. E. 1994. Projector Augmented-Wave Method. *Physical Review B,* 50**,** 17953-17979.

BOBROWSKI, N., HÖNNINGER, G., GALLE, B. & PLATT, U. 2003. Detection of bromine monoxide in a volcanic plume. *Nature,* 423**,** 273-276.

BÖKE, H., GÖKTÜRK, E. H., CANER-SALTıK, E. N. & DEMIRCI, Ş. 1999. Effect of airborne particle on SO_2–calcite reaction. *Applied Surface Science,* 140**,** 70-82.

BOULOS, R. A., ZHANG, F., TJANDRA, E. S., MARTIN, A. D., SPAGNOLI, D. & RASTON, C. L. 2014. Spinning up the polymorphs of calcium carbonate. *Scientific Reports,* 4**,** 3616.

BUSENBERG, E. & NIEL PLUMMER, L. 1985. Kinetic and thermodynamic factors controlling the distribution of SO_3^{2-} and Na^+ in calcites and selected aragonites. *Geochimica et Cosmochimica Acta,* 49**,** 713-725.

CARRILLO-BERDUGO, I., MIDGLEY, S. D., GRAU-CRESPO, R., ZORRILLA, D. & NAVAS, J. 2020. Understanding the Specific Heat Enhancement in Metal-Containing Nanofluids for Thermal Energy Storage: Experimental and Ab Initio Evidence for a Strong Interfacial Layering Effect. *ACS Applied Energy Materials,* 3**,** 9246-9256.

CATLOW, C. 1978. Ion distribution functions for complex solids and their application to the conductivity of glasses. *physica status solidi (a),* 46**,** 191-198.

CHAI, J.-D. & HEAD-GORDON, M. 2008. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Physical Chemistry Chemical Physics,* 10**,** 6615-6620.

CHAKA, A. M. 2018. Ab Initio Thermodynamics of Hydrated Calcium Carbonates and Calcium Analogues of Magnesium Carbonates: Implications for Carbonate Crystallization Pathways. *ACS Earth and Space Chemistry,* 2**,** 210-224.

CHAKA, A. M. 2019. Quantifying the Impact of Magnesium on the Stability and Water Binding Energy of Hydrated Calcium Carbonates by Ab Initio Thermodynamics. *The Journal of Physical Chemistry A,* 123**,** 2908-2923.

CHRISTY, A. G. 2017. A Review of the Structures of Vaterite: The Impossible, the Possible, and the Likely. *Crystal Growth & Design,* 17**,** 3567-3578.

DAVIS, K. J., DOVE, P. M. & YOREO, J. J. D. 2000. Resolving the Control of Magnesium on Calcite Growth: Thermodynamic and Kinetic Consequences of Impurity Incorporation for Biomineral Formation. *MRS Proceedings,* 620**,** M9.5.1.

DE CHOUDENS-SANCHEZ, V. & GONZALEZ, L. A. 2009. Calcite and Aragonite Precipitation Under Controlled Instantaneous Supersaturation: Elucidating the Role of $CaCO_3$ Saturation State and Mg/Ca Ratio on Calcium Carbonate Polymorphism. *Journal of Sedimentary Research,* 79**,** 363-376.

DE LEEUW, N. H. 2002. Molecular Dynamics Simulations of the Growth Inhibiting Effect of Fe^{2+}, Mg^{2+}, Cd^{2+}, and Sr^{2+} on Calcite Crystal Growth. *The Journal of Physical Chemistry B,* 106**,** 5241-5249.

DE LEEUW, N. H. & PARKER, S. C. 1998. Surface Structure and Morphology of Calcium Carbonate Polymorphs Calcite, Aragonite, and Vaterite: An Atomistic Approach. *The Journal of Physical Chemistry B,* 102**,** 2914-2922.

DE LEEUW, N. H., PARKER, S. C. & HARDING, J. H. 1999. Molecular dynamics simulation of crystal dissolution from calcite steps. *Physical Review B,* 60**,** 13792-13799.

DE VILLIERS, J. P. R. 1971. Crystal Structures of Aragonite, Strontianite, and Witherite. *American Mineralogist,* 56**,** 758-767.

DIDYMUS, J. M., OLIVER, P., MANN, S., DEVRIES, A. L., HAUSCHKA, P. V. & WESTBROEK, P. 1993. Influence of low-molecular-weight and macromolecular organic additives on the morphology of calcium carbonate. *Journal of the Chemical Society, Faraday Transactions,* 89**,** 2891-2900.

DORALE, J. A., EDWARDS, R. L., CALVIN-ALEXANDER JR, E., SHEN, C. C., RICHARDS, D. A. & CHENG, H. 2004. *Studies of Cave Sediments,* Boston, MA, Springer US.

DRONSKOWSKI, R. 2005. *Computational Chemistry of Solid State Materials,* Germany, Wiley.

ELSTNER, M. & SEIFERT, G. 2014. Density functional tight binding. *Philos Trans A Math Phys Eng Sci,* 372**,** 20120483.

ENGEL, A. O. & PERLMUTTER-HAYMAN, B. 1953. The Decomposition of Hypobromite and Bromite Solutions. *Journal of the American Chemical Society,* 76**,** 2010-2015.

ERNZERHOF, M. & SCUSERIA, G. E. 1999. Assessment of the Perdew–Burke–Ernzerhof exchange-correlation functional. *The Journal of Chemical Physics,* 110**,** 5029-5036.

FAIRCHILD, I. J. & BAKER, A. 2012. *Speleothem Science : From Process to Past Environments,* Hoboken, UNITED KINGDOM, John Wiley & Sons, Incorporated.

FAIRCHILD, I. J. & TREBLE, P. C. 2009. Trace elements in speleothems as recorders of environmental change. *Quaternary Science Reviews,* 28**,** 449-468.

FARIA, S. H., FREITAG, J. & KIPFSTUHL, S. 2010. Polar ice structure and the integrity of ice-core paleoclimate records. *Quaternary Science Reviews,* 29**,** 338-351.

FENG, X. & REDFERN, S. A. T. 2018. Iodate in calcite, aragonite and vaterite $CaCO_3$: Insights from first-principles calculations and implications for the I/Ca geochemical proxy. *Geochimica et Cosmochimica Acta,* 236**,** 351-360.

FERNÁNDEZ-DÍAZ, L., FERNÁNDEZ-GONZÁLEZ, Á. & PRIETO, M. 2010. The role of sulfate groups in controlling $CaCO_3$ polymorphism. *Geochimica et Cosmochimica Acta,* 74**,** 6064-6076.

FIELD, L. P., MILODOWSKI, A. E., SHAW, R. P., STEVENS, L. A., HALL, M. R., KILPATRICK, A., GUNN, J., KEMP, S. J. & ELLIS, M. A. 2018. Unusual morphologies

and the occurrence of pseudomorphs after ikaite ($CaCO_3 \cdot 6H_2O$) in fast growing, hyperalkaline speleothems. *Mineralogical Magazine,* 81**,** 565-589.

FINCH, A. A., SHAW, P. A., WEEDON, G. P. & HOLMGREN, K. 2001. Trace element variation in speleothem aragonite: potential for palaeoenvironmental reconstruction. *Earth and Planetary Science Letters,* 186**,** 255-267.

FISCHBECK, R. & MÜLLER, G. 1971. Monohydrocalcite, hydromagnesite, nesquehonite, dolomite, aragonite, and calcite in speleothems of the Fränkische Schweiz, Western Germany. *Contributions to Mineralogy and Petrology,* 33**,** 87-92.

FISLER, D. K., GALE, J. D. & CYGAN, R. T. 2000. A shell model for the simulation of rhombohedral carbonate minerals and their point defects. *American Mineralogist,* 85**,** 217-224.

FLEITMANN, D., BURNS, S. J., NEFF, U., MANGINI, A. & MATTER, A. 2003. Changing moisture sources over the last 330,000 years in Northern Oman from fluid-inclusion evidence in speleothems. *Quaternary Research,* 60**,** 223-232.

FLEITMANN, D., CHENG, H., BADERTSCHER, S., EDWARDS, R. L., MUDELSEE, M., GÖKTÜRK, O. M., FANKHAUSER, A., PICKERING, R., RAIBLE, C. C., MATTER, A., KRAMERS, J. & TÜYSÜZ, O. 2009. Timing and climatic impact of Greenland interstadials recorded in stalagmites from northern Turkey. *Geophysical Research Letters,* 36.

FRIEDRICH, W. L., KROMER, B., FRIEDRICH, M., HEINEMEIER, J., PFEIFFER, T. & TALAMO, S. 2006. Santorini eruption radiocarbon dated to 1627-1600 BC. *Science,* 312**,** 548-548.

FRISIA, S., BORSATO, A., FAIRCHILD, I. J. & SUSINI, J. 2005. Variations in atmospheric sulphate recorded in stalagmites by synchrotron micro-XRF and XANES analyses. *Earth and Planetary Science Letters,* 235**,** 729-740.

FRISIA, S., BORSATO, A. & SUSINI, J. 2008. Synchrotron radiation applications to past volcanism archived in speleothems: An overview. *Journal of Volcanology and Geothermal Research,* 177**,** 96-100.

GALE, J. D., RAITERI, P. & VAN DUIN, A. C. T. 2011. A reactive force field for aqueous-calcium carbonate systems. *Physical chemistry chemical physics : PCCP,* 13**,** 16666-16679.

GALE, J. D. & ROHL, A. L. 2003. The General Utility Lattice Program (GULP). *Molecular Simulation,* 29**,** 291-341.

GHADIRI, E., AFFOLTER, S., BRENNWALD, M. S., FLEITMANN, D., HÄUSELMANN, A. D., CHENG, H., MADEN, C., LEUENBERGER, M. & KIPFER, R. 2020. Estimation of temperature – altitude gradients during the Pleistocene–Holocene transition from Swiss stalagmites. *Earth and Planetary Science Letters,* 544**,** 116387.

GHATEE, M. H. & KOLEINI, M. M. 2017. Bonding, structural and thermodynamic analysis of dissociative adsorption of H3O+ ion onto calcite $$ \left(10\overline{1}4\right) $$surface: CPMD and DFT calculations. *Journal of Molecular Modeling,* 23**,** 331.

GIVEN, R. K. & WILKINSON, B. H. 1985. Kinetic Control of Morphology, Composition, and Mineralogy of Abiotic Sedimentary Carbonates. *SEPM Journal of Sedimentary Research,* Vol. 55**,** 109-119.

GOEDECKER, S., TETER, M. & HUTTER, J. 1996. Separable dual-space Gaussian pseudopotentials. *Physical Review B,* 54**,** 1703-1710.

GÖKTÜRK, O. M., FLEITMANN, D., BADERTSCHER, S., CHENG, H., EDWARDS, R. L., LEUENBERGER, M., FANKHAUSER, A., TÜYSÜZ, O. & KRAMERS, J. 2011. Climate on the southern Black Sea coast during the Holocene: implications from the Sofular Cave record. *Quaternary Science Reviews,* 30**,** 2433-2445.

GONZÁLEZ-LÓPEZ, J., RUIZ-HERNÁNDEZ, S. E., FERNÁNDEZ-GONZÁLEZ, Á., JIMÉNEZ, A., DE LEEUW, N. H. & GRAU-CRESPO, R. 2014. Cobalt incorporation in calcite: Thermochemistry of $(Ca,Co)CO_3$ solid solutions from density functional theory simulations. *Geochimica et Cosmochimica Acta,* 142**,** 205-216.

GRASBY, S. E., ALLEN, C. C., LONGAZO, T. G., LISLE, J. T., GRIFFIN, D. W. & BEAUCHAMP, B. 2003. Biogeochemical sulphur cycle in an extreme environment—

lifebeneath a high arctic glacier, Nunavut, Canada. *Journal of Geochemical Exploration,* 78-79**,** 71-74.
GRATZ, A. J., HILLNER, P. E. & HANSMA, P. K. 1993. Step dynamics and spiral growth on calcite. *Geochimica et Cosmochimica Acta,* 57**,** 491-495.
GRAU-CRESPO, R., DE LEEUW, N. H., HAMAD, S. & WAGHMARE, U. V. 2011. Phase separation and surface segregation in ceria–zirconia solid solutions. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences,* 467**,** 1925-1938.
GRAU-CRESPO, R., HAMAD, S., CATLOW, C. R. A. & LEEUW, N. H. D. 2007. Symmetry-adapted configurational modelling of fractional site occupancy in solids. *Journal of Physics: Condensed Matter,* 19**,** 256201.
GRIMME, S. 2006. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of Computational Chemistry,* 27**,** 1787-1799.
GRIMME, S., ANTONY, J., EHRLICH, S. & KRIEG, H. 2010. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of Chemical Physics,* 132**,** 154104.
GRIMME, S., EHRLICH, S. & GOERIGK, L. 2011. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *Journal of Computational Chemistry,* 32**,** 1456-1465.
HAFSHEJANI, T. M., WANG, W., HEGGEMANN, J., NEFEDOV, A., HEISSLER, S., WANG, Y., RAHE, P., THISSEN, P. & WÖLL, C. 2021. CO adsorption on the calcite(10.4) surface: a combined experimental and theoretical study. *Physical Chemistry Chemical Physics,* 23**,** 7696-7702.
HAMMER, B., HANSEN, L. B. & NØRSKOV, J. K. 1999. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Physical Review B,* 59**,** 7413-7421.
HARTLAND, A., FAIRCHILD, I. J., LEAD, J. R., BORSATO, A., BAKER, A., FRISIA, S. & BAALOUSHA, M. 2012. From soil to cave: Transport of trace metals by natural organic matter in karst dripwaters. *Chemical Geology,* 304-305**,** 68-82.
HESSE, K. F., KÜPPERS, H. & SUESS, E. 1983. Refinement of the structure of Ikaite, $CaCO_3$ - $6H_2O$. *Zeitschrift für Kristallographie - Crystalline Materials.*
HEYD, J. & SCUSERIA, G. E. 2004. Efficient hybrid density functional calculations in solids: Assessment of the Heyd–Scuseria–Ernzerhof screened Coulomb hybrid functional. *The Journal of Chemical Physics,* 121**,** 1187-1192.
HOHENBERG, P. & KOHN, W. 1964. Inhomogeneous Electron Gas. *Physical Review,* 136**,** B864-B871.
HOOVER, W. G. 1985. Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A,* 31**,** 1695-1697.
HÖRMANN, C., SIHLER, H., BOBROWSKI, N., BEIRLE, S., PENNING DE VRIES, M., PLATT, U. & WAGNER, T. 2013. Systematic investigation of bromine monoxide in volcanic plumes from space by using the GOME-2 instrument. *Atmos. Chem. Phys.,* 13**,** 4749-4781.
HOURAHINE, B., ARADI, B., BLUM, V., BONAFÉ, F., BUCCHERI, A., CAMACHO, C., CEVALLOS, C., DESHAYE, M. Y., DUMITRICĂ, T., DOMINGUEZ, A., EHLERT, S., ELSTNER, M., VAN DER HEIDE, T., HERMANN, J., IRLE, S., KRANZ, J. J., KÖHLER, C., KOWALCZYK, T., KUBAŘ, T., LEE, I. S., LUTSKER, V., MAURER, R. J., MIN, S. K., MITCHELL, I., NEGRE, C., NIEHAUS, T. A., NIKLASSON, A. M. N., PAGE, A. J., PECCHIA, A., PENAZZI, G., PERSSON, M. P., ŘEZÁČ, J., SÁNCHEZ, C. G., STERNBERG, M., STÖHR, M., STUCKENBERG, F., TKATCHENKO, A., YU, V. W. Z. & FRAUENHEIM, T. 2020. DFTB+, a software package for efficient approximate density functional theory based atomistic simulations. *The Journal of Chemical Physics,* 152**,** 124101.
HTTPS://WWW.CP2K.ORG/, T. C. K. D. G. A. A. *CP2K Developers Group* [Online]. [Accessed 2021].
JACKSON, R. & PRICE, G. 1992. A transferable interatomic potential for calcium carbonate. *Molecular simulation,* 9**,** 175-177.

JANSEN, J. H. F., WOENSDREGT, C. F., KOOISTRA, M. J. & VAN DER GAAST, S. J. 1987. Ikaite pseudomorphs in the Zaire deep-sea fan: An intermediate between calcite and porous calcite. *Geology,* 15**,** 245-248.

JENSEN, F. 2006. *Introduction to Computational Chemistry*, John Wiley & Sons, Inc.

JINNOUCHI, R., KARSAI, F. & KRESSE, G. 2019. On-the-fly machine learning force field generation: Application to melting points. *Physical Review B,* 100**,** 014105.

KAMPSCHULTE, A., BRUCKSCHEN, P. & STRAUSS, H. 2001. The sulphur isotopic composition of trace sulphates in Carboniferous brachiopods: implications for coeval seawater, correlation with other geochemical cycles and isotope stratigraphy. *Chemical Geology,* 175**,** 149-173.

KELLY, S. D., NEWVILLE, M. G., CHENG, L., KEMNER, K. M., SUTTON, S. R., FENTER, P., STURCHIO, N. C. & SPÖTL, C. 2003. Uranyl Incorporation in Natural Calcite. *Environmental Science & Technology,* 37**,** 1284-1287.

KERISIT, S. N., SMITH, F. N., SASLOW, S. A., HOOVER, M. E., LAWTER, A. R. & QAFOKU, N. P. 2018. Incorporation modes of iodate in calcite. *Environmental science & technology,* 52**,** 5902-5910.

KING, D., WILLIAMS, P. & SALINGER, J. 2004. Reconstructing past environmental changes using speleothems. *Water and Atmosphere,* 12**,** 14-15.

KNAPPERT, M., JIN, T., MIDGLEY, S. D., WU, G., SCHERMAN, O. A., GRAU-CRESPO, R. & COLQUHOUN, H. M. 2019. Supramolecular complexation between chain-folding poly (ester-imide) s and polycyclic aromatics: a fractal-based pattern of NMR ring-current shielding. *Polymer Chemistry,* 10**,** 6641-6650.

KNAPPERT, M., JIN, T., MIDGLEY, S. D., WU, G., SCHERMAN, O. A., GRAU-CRESPO, R. & COLQUHOUN, H. M. 2020. Single-site binding of pyrene to poly (ester-imide) s incorporating long spacer-units: prediction of NMR resonance-patterns from a fractal model. *Chemical science,* 11**,** 12165-12177.

KOHN, W. & SHAM, L. J. 1965. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review,* 140**,** A1133-A1138.

KRESSE, G. 1999. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B,* 59**,** 1758-1775.

KRESSE, G. & FURTHMÜLLER, J. 1996a. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science,* 6**,** 15-50.

KRESSE, G. & FURTHMÜLLER, J. 1996b. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B,* 54**,** 11169-11186.

KRESSE, G. & JOUBERT, D. 1999. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B,* 59**,** 1758-1775.

KÜHNE, T. D., IANNUZZI, M., DEL BEN, M., RYBKIN, V. V., SEEWALD, P., STEIN, F., LAINO, T., KHALIULLIN, R. Z., SCHÜTT, O., SCHIFFMANN, F., GOLZE, D., WILHELM, J., CHULKOV, S., BANI-HASHEMIAN, M. H., WEBER, V., BORŠTNIK, U., TAILLEFUMIER, M., JAKOBOVITS, A. S., LAZZARO, A., PABST, H., MÜLLER, T., SCHADE, R., GUIDON, M., ANDERMATT, S., HOLMBERG, N., SCHENTER, G. K., HEHN, A., BUSSY, A., BELLEFLAMME, F., TABACCHI, G., GLÖß, A., LASS, M., BETHUNE, I., MUNDY, C. J., PLESSL, C., WATKINS, M., VANDEVONDELE, J., KRACK, M. & HUTTER, J. 2020. CP2K: An electronic structure and molecular dynamics software package - Quickstep: Efficient and accurate electronic structure calculations. *The Journal of Chemical Physics,* 152**,** 194103.

KURIYAVAR, S. I., VETRIVEL, R., HEGDE, S. G., RAMASWAMY, A. V., CHAKRABARTY, D. & MAHAPATRA, S. 2000. Insights into the formation of hydroxyl ions in calcium carbonate: temperature dependent FTIR and molecular modelling studies. *Journal of Materials Chemistry,* 10**,** 1835-1840.

LACELLE, D., LAURIOL, B. & CLARK, I. D. 2009. Formation of seasonal ice bodies and associated cryogenic carbonates in caverne de l'ours, Québec, Canada: Kinetic isotope effects and pseudo-biogenic crystal structures. *Journal of Cave and Karst Studies,* 71**,** 48-62.

LACHNIET, M. S. 2015. Are aragonite stalagmites reliable paleoclimate proxies? Tests for oxygen isotope time-series replication and equilibrium. *GSA Bulletin,* 127**,** 1521-1533.
LARDGE, J. S., DUFFY, D. M., GILLAN, M. J. & WATKINS, M. 2010. Ab Initio Simulations of the Interaction between Water and Defects on the Calcite (10$\bar{1}$4) Surface. *The Journal of Physical Chemistry C,* 114**,** 2664-2668.
LEE, C., YANG, W. & PARR, R. G. 1988. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical Review B,* 37**,** 785-789.
LEE, C. L. & LISTER, M. W. 1971. The Decomposition of Aqueous Sodium Bromite. *Canadian Journal of Chemistry,* 49**,** 2822-2826.
LEE, I.-H. & MARTIN, R. M. 1997. Applications of the generalized-gradient approximation to atoms, clusters, and solids. *Physical Review B,* 56**,** 7197-7205.
LEWARS, E. G. 2011. *Computational Chemistry,* Springer.
LIANG, Y. & BAER, D. R. 1997. Anisotropic dissolution at the CaCO3(101̄4)—water interface. *Surface Science,* 373**,** 275-287.
LIANG, Y., BAER, D. R., MCCOY, J. M. & LAFEMINA, J. P. 1996a. Interplay between step velocity and morphology during the dissolution of $CaCO_3$ surface. *Journal of Vacuum Science & Technology A,* 14**,** 1368-1375.
LIANG, Y., LEA, A. S., BAER, D. R. & ENGELHARD, M. H. 1996b. Structure of the cleaved $CaCO_3$ (1014) surface in an aqueous environment. *Surface Science,* 351**,** 172-182.
LIN, I. C., SEITSONEN, A. P., TAVERNELLI, I. & ROTHLISBERGER, U. 2012. Structure and Dynamics of Liquid Water from ab Initio Molecular Dynamics—Comparison of BLYP, PBE, and revPBE Density Functionals with and without van der Waals Corrections. *Journal of Chemical Theory and Computation,* 8**,** 3902-3910.
LORENS, R. B. 1981. Sr, Cd, Mn and Co distribution coefficients in calcite as a function of calcite precipitation rate. *Geochimica et Cosmochimica Acta,* 45**,** 553-561.
LOWE, D. J. 2011. Tephrochronology and its application: A review. *Quaternary Geochronology,* 6**,** 107-153.
MACINNIS, I. N. & BRANTLEY, S. L. 1992. The role of dislocations and surface morphology in calcite dissolution. *Geochimica et Cosmochimica Acta,* 56**,** 1113-1126.
MALAGA-STARZEC, K., PANAS, I. & LINDQVIST, O. 2004. Model study of initial adsorption of SO_2 on calcite and dolomite. *Applied Surface Science,* 222**,** 82-88.
MANN, M. E., ZHANG, Z., HUGHES, M. K., BRADLEY, R. S., MILLER, S. K., RUTHERFORD, S. & NI, F. 2008. Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proceedings of the National Academy of Sciences,* 105**,** 13252.
MARKGRAF, S. & REEDER, R. 1985. High-temperature structure refinements of calcite and magnesite. *Am. Mineral.,* 70**,** 590-600.
MARSHALL, D., GHALEB, B., COUNTESS, R. & GABITIES, J. 2009. Preliminary paleoclimate reconstruction based on a 12,500 year old speleothem from Vancouver Island, Canada: stable isotopes and U-Th disequilibrium dating. *Quaternary Science Reviews,* 28**,** 2507-2513.
MARTINEZ-MERINO, P., MIDGLEY, S. D., MARTIN, E. I., ESTELLE, P., ALCANTARA, R., SANCHEZ-CORONILLA, A., GRAU-CRESPO, R. & NAVAS, J. 2020. Novel WS2-based nanofluids for concentrating solar power: performance characterization and molecular-level insights. *ACS applied materials & interfaces,* 12**,** 5793-5804.
MARX, D. & HUTTER, J. 2009. *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods,* Cambridge, Cambridge University Press.
MATHER, T. A., ALLEN, A. G., DAVISON, B. M., PYLE, D. M., OPPENHEIMER, C. & MCGONIGLE, A. J. S. 2004. Nitric acid from volcanoes. *Earth and Planetary Science Letters,* 218**,** 17-30.
MCCARROLL, D. & LOADER, N. J. 2004. Stable isotopes in tree rings. *Quaternary Science Reviews,* 23**,** 771-801.
MEYER, H. J. 1984. The influence of impurities on the growth rate of calcite. *Journal of Crystal Growth,* 66**,** 639-646.

MIDGLEY, S. D., DI TOMMASO, D., FLEITMANN, D. & GRAU-CRESPO, R. 2021a. Sulfate and Molybdate Incorporation at the Calcite–Water Interface: Insights from Ab Initio Molecular Dynamics. *ACS Earth and Space Chemistry*.

MIDGLEY, S. D., FLEITMANN, D. & GRAU-CRESPO, R. 2021b. Bromate incorporation in calcite and aragonite. *ChemRxiv*.

MIDGLEY, S. D., HAMAD, S., BUTLER, K. T. & GRAU-CRESPO, R. 2021c. Bandgap Engineering in the Configurational Space of Solid Solutions via Machine Learning: (Mg,Zn)O Case Study. *The Journal of Physical Chemistry Letters,* 12**,** 5163-5168.

MIDGLEY, S. D., TAYLOR, J. O., FLEITMANN, D. & GRAU-CRESPO, R. 2020. Molybdenum and sulphur incorporation as oxyanion substitutional impurities in calcium carbonate minerals: a computational investigation. *Chemical Geology*.

MONKHORST, H. J. & PACK, J. D. 1976. Special points for Brillouin-zone integrations. *Physical Review B,* 13**,** 5188-5192.

MURRAY, J. W. 1954. The Deposition of Calcite and Aragonite in Caves. *The Journal of Geology,* 62**,** 481-492.

NASA. *https://earthobservatory.nasa.gov/features/GlobalWarming/page3.php* [Online]. [Accessed 2020].

NEAGLE, W. & ROCHESTER, C. H. 1990. Infrared study of the adsorption of water and ammonia on calcium carbonate. *Journal of the Chemical Society, Faraday Transactions,* 86**,** 181-183.

NIELSEN, L. C., DE YOREO, J. J. & DEPAOLO, D. J. 2013. General model for calcite growth kinetics in the presence of impurity ions. *Geochimica et Cosmochimica Acta,* 115**,** 100-114.

NILSSON, Ö. & STERNBECK, J. 1999. A mechanistic model for calcite crystal growth using surface speciation. *Geochimica et Cosmochimica Acta,* 63**,** 217-225.

NOSÉ, S. 1984. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics,* 81**,** 511-519.

OGINO, T., SUZUKI, T. & SAWADA, K. 1987. The formation and transformation mechanism of calcium carbonate in water. *Geochimica et Cosmochimica Acta,* 51**,** 2757-2767.

PAQUETTE, J. & REEDER, R. J. 1995. Relationship between surface structure, growth mechanism, and trace element incorporation in calcite. *Geochimica et Cosmochimica Acta,* 59**,** 735-749.

PAVESE, A., CATTI, M., PRICE, G. D. & JACKSON, R. A. 1992. Interatomic potentials for $CaCO_3$ polymorphs (calcite and aragonite), fitted to elastic and vibrational data. *Physics and Chemistry of Minerals,* 19**,** 80-87.

PERDEW, J. P., BURKE, K. & ERNZERHOF, M. 1996a. Generalized gradient approximation made simple. *Physical Review Letters,* 77**,** 3865-3868.

PERDEW, J. P., BURKE, K. & ERNZERHOF, M. 1996b. Generalized gradient approximation made simple. *Physical review letters,* 77**,** 3865.

PLUMMER, L. N. & BUSENBERG, E. 1982. The solubilities of calcite, aragonite and vaterite in CO_2-H_2O solutions between 0 and 90°C, and an evaluation of the aqueous model for the system $CaCO_3$-CO_2-H_2O. *Geochimica et Cosmochimica Acta,* 46**,** 1011-1040.

PODDER, J., LIN, J., SUN, W., BOTIS, S., TSE, J., CHEN, N., HU, Y., LI, D., SEAMAN, J. & PAN, Y. 2017. Iodate in calcite and vaterite: Insights from synchrotron X-ray absorption spectroscopy and first-principles calculations. *Geochimica et Cosmochimica Acta,* 198**,** 218-228.

PULAY, P. 1980. Convergence acceleration of iterative sequences. the case of scf iteration. *Chemical Physics Letters,* 73**,** 393-398.

PULAY, P. 1982. Improved SCF convergence acceleration. *Journal of Computational Chemistry,* 3**,** 556-560.

PURGSTALLER, B., DIETZEL, M., BALDERMANN, A. & MAVROMATIS, V. 2017. Control of temperature and aqueous Mg2+/Ca2+ ratio on the (trans-)formation of ikaite. *Geochimica et Cosmochimica Acta,* 217**,** 128-143.

RAILSBACK, L. B., BROOK, G. A., CHEN, J., KALIN, R. & FLEISHER, C. J. 1994. Environmental Controls on the Petrology of a Late Holocene Speleothem from Botswana

with Annual Layers of Aragonite and Calcite. *SEPM Journal of Sedimentary Research,* Vol. 64A**,** 147-155.

RAITERI, P., GALE, J. D., QUIGLEY, D. & RODGER, P. M. 2010. Derivation of an Accurate Force-Field for Simulating the Growth of Calcium Carbonate from Aqueous Solution: A New Model for the Calcite−Water Interface. *The Journal of Physical Chemistry C,* 114**,** 5997-6010.

REEDER, R. J. & GRAMS, J. C. 1987. Sector zoning in calcite cement crystals: Implications for trace element distributions in carbonates. *Geochimica et Cosmochimica Acta,* 51**,** 187-194.

ROBOCK, A. 2000. Volcanic eruptions and climate. *Reviews of Geophysics,* 38**,** 191-219.

ROHL, A. L., WRIGHT, K. & GALE, J. D. 2003. Evidence from surface phonons for the (2 × 1) reconstruction of the $(10\bar{1}4)$ surface of calcite from computer simulation. *American Mineralogist,* 88**,** 921-925.

ROMAO, C. P. 2017. Anisotropic thermal expansion in flexible materials. *Physical Review B,* 96**,** 134113.

RUIZ-AGUDO, E., PUTNIS, C. V., RODRIGUEZ-NAVARRO, C. & PUTNIS, A. 2011. Effect of pH on calcite growth at constant a Ca^{2+}/aCO_3^{2-} ratio and supersaturation. *Geochimica et Cosmochimica Acta,* 75**,** 284-296.

RUIZ-HERNANDEZ, S. E., GRAU-CRESPO, R., RUIZ-SALVADOR, A. R. & DE LEEUW, N. H. 2010. Thermochemistry of strontium incorporation in aragonite from atomistic simulations. *Geochimica et Cosmochimica Acta,* 74**,** 1320-1328.

RUIZ PESTANA, L., MARDIROSSIAN, N., HEAD-GORDON, M. & HEAD-GORDON, T. 2017. Ab initio molecular dynamics simulations of liquid water using high quality meta-GGA functionals. *Chemical Science,* 8**,** 3554-3565.

RUTTER, N., CORONATO, A., HELMENS, K., RABASSA, J. & ZÁRATE, M. 2012. *Glaciations in North and South America from the Miocene to the Last Glacial Maximum: Comparisons, Linkages and Uncertainties* Springer.

SEVERI, M., UDISTI, R., BECAGLI, S., STENNI, B. & TRAVERSI, R. 2012. Volcanic synchronisation of the EPICA-DC and TALDICE ice cores for the last 42 kyr BP. *Climate of the Past.,* 8**,** 509-517.

SHOLL, D. S. & STECKEL, J. A. 2009. *Density functional theory : a practical introduction,* Hoboken, N.J., Wiley.

SIGL, M., MCCONNELL, J. R., LAYMAN, L., MASELLI, O., MCGWIRE, K., PASTERIS, D., DAHL-JENSEN, D., STEFFENSEN, J. P., VINTHER, B., EDWARDS, R., MULVANEY, R. & KIPFSTUHL, S. 2013. A new bipolar ice core record of volcanism from WAIS Divide and NEEM and implications for climate forcing of the last 2000 years. *Journal of Geophysical Research: Atmospheres,* 118**,** 1151-1169.

SILK, S. T. & LEWIN, S. Z. 1970. X-Ray Diffraction Analysis of the Calcium Carbonate Polymorphs. *Advances in X-ray Analysis,* 14**,** 29-37.

SINGER, A. 1984. The paleoclimatic interpretation of clay minerals in sediments — a review. *Earth-Science Reviews,* 21**,** 251-293.

SMITH, C. L., FAIRCHILD, I. J., SPÖTL, C., FRISIA, S., BORSATO, A., MORETON, S. G. & WYNN, P. M. 2009. Chronology building using objective identification of annual signals in trace element profiles of stalagmites. *Quaternary Geochronology,* 4**,** 11-21.

SMITH, D. W. 1977. Ionic hydration enthalpies. *Journal of Chemical Education,* 54**,** 540.

SMITH, K. C., FISHER, T. S., WAGHMARE, U. V. & GRAU-CRESPO, R. 2010. Dopant-vacancy binding effects in Li-doped magnesium hydride. *Physical Review B,* 82**,** 134109.

ST PIERRE, E., ZHAO, J.-X. & REED, E. 2009. Expanding the utility of Uranium-series dating of speleothems for archaeological and palaeontological applications. *Journal of Archaeological Science,* 36**,** 1416-1423.

STAUDT, W. J., REEDER, R. J. & SCHOONEN, M. A. A. 1994. Surface structural controls on compositional zoning of SO_4^{2-} and SeO_4^{2-} in synthetic calcite single crystals. *Geochimica et Cosmochimica Acta,* 58**,** 2087-2098.

STENCHIKOV, G. L., KIRCHNER, I., ROBOCK, A., GRAF, H.-F., ANTUÑA, J. C., GRAINGER, R. G., LAMBERT, A. & THOMASON, L. 1998. Radiative forcing from the

1991 Mount Pinatubo volcanic eruption. *Journal of Geophysical Research: Atmospheres,* 103**,** 13837-13857.

STIPP, S. L. & HOCHELLA JR, M. F. 1991. Structure and bonding environments at the calcite surface as observed with X-ray photoelectron spectroscopy (XPS) and low energy electron diffraction (LEED). *Geochimica et Cosmochimica Acta,* 55**,** 1723-1736.

STOFFERS, P. & FISCHBECK, R. 1974. Monohydrocalcite in the sediments of Lake Kivu (East Africa). *Sedimentology,* 21**,** 163-170.

SWAINSON, I. P. 2008. The structure of monohydrocalcite and the phase composition of the beachrock deposits of Lake Butler and Lake Fellmongery, South Australia. *American Mineralogist,* 93**,** 1014-1018.

TANG, J., NIEDERMAYR, A., KÖHLER, S. J., BÖHM, F., KıSAKÜREK, B., EISENHAUER, A. & DIETZEL, M. 2012. Sr2+/Ca2+ and 44Ca/40Ca fractionation during inorganic calcite formation: III. Impact of salinity/ionic strength. *Geochimica et Cosmochimica Acta,* 77**,** 432-443.

TENG, H. H., DOVE, P. M. & DEYOREO, J. J. 1999. Reversed calcite morphologies induced by microscopic growth kinetics: insight into biomineralization. *Geochimica et Cosmochimica Acta,* 63**,** 2507-2512.

THEODORAKOPOULOU, K., KYRIAKOPOULOS, K., ATHANASSAS, C. D., GALANOPOULOS, E., ECONOMOU, G., MANIATIS, Y., GODELITSAS, A., DOTSIKA, E., MAVRIDIS, F. & DARLAS, A. 2020. First Speleothem Evidence of the Hiera Eruption (197 BC), Santorini, Greece. *Environmental Archaeology***,** 1-13.

THEYS, N., VAN ROOZENDAEL, M., DILS, B., HENDRICK, F., HAO, N. & DE MAZIÈRE, M. 2009. First satellite detection of volcanic bromine monoxide emission after the Kasatochi eruption. *Geophysical Research Letters,* 36.

ÜNAL-İMER, E., SHULMEISTER, J., ZHAO, J.-X., TONGUÇ UYSAL, I., FENG, Y.-X., DUC NGUYEN, A. & YÜCE, G. 2015. An 80 kyr-long continuous speleothem record from Dim Cave, SW Turkey with paleoclimatic implications for the Eastern Mediterranean. *Scientific Reports,* 5**,** 13560-13560.

VANDEVONDELE, J. & HUTTER, J. 2003. An efficient orbital transformation method for electronic structure calculations. *Journal of Chemical Physics,* 118**,** 4365-4369.

VANDEVONDELE, J. & HUTTER, J. 2007. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *Journal of Chemical Physics,* 127.

VANDEVONDELE, J., KRACK, M., MOHAMED, F., PARRINELLO, M., CHASSAING, T. & HUTTER, J. 2005. QUICKSTEP: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Computer Physics Communications,* 167**,** 103-128.

VAVOURAKI, A. I., PUTNIS, C. V., PUTNIS, A. & KOUTSOUKOS, P. G. 2008. An Atomic Force Microscopy study of the growth of calcite in the presence of sodium sulfate. *Chemical Geology,* 253**,** 243-251.

VOELKER, A. H. L. 2002. Global distribution of centennial-scale records for Marine Isotope Stage (MIS) 3: a database. *Quaternary Science Reviews,* 21**,** 1185-1212.

WANG, J. & BECKER, U. 2009. Structure and carbonate orientation of vaterite (CaCO3). *American Mineralogist,* 94**,** 380-386.

WANG, J., ROMÁN-PÉREZ, G., SOLER, J. M., ARTACHO, E. & FERNÁNDEZ-SERRA, M. V. 2011a. Density, structure, and dynamics of water: The effect of van der Waals interactions. *The Journal of Chemical Physics,* 134**,** 024516.

WANG, Q., GRAU-CRESPO, R. & DE LEEUW, N. H. 2011b. Mixing Thermodynamics of the Calcite-Structured $(Mn,Ca)CO_3$ Solid Solution: A Computer Simulation Study. *The Journal of Physical Chemistry B,* 115**,** 13854-13861.

WANG, X., TOROZ, D., KIM, S., CLEGG, S. L., PARK, G.-S. & DI TOMMASO, D. 2020. Density functional theory based molecular dynamics study of solution composition effects on the solvation shell of metal ions. *Physical Chemistry Chemical Physics,* 22**,** 16301-16313.

WANG, Y. & XU, H. 2001. Prediction of trace metal partitioning between minerals and aqueous solutions: a linear free energy correlation approach. *Geochimica et Cosmochimica Acta,* 65**,** 1529-1543.

WASYLENKI, L. E., DOVE, P. M. & DE YOREO, J. J. 2005a. Effects of temperature and transport conditions on calcite growth in the presence of Mg^{2+}: Implications for paleothermometry. *Geochimica et Cosmochimica Acta,* 69**,** 4227-4236.

WASYLENKI, L. E., DOVE, P. M., WILSON, D. S. & DE YOREO, J. J. 2005b. Nanoscale effects of strontium on calcite growth: An in situ AFM study in the absence of vital effects. *Geochimica et Cosmochimica Acta,* 69**,** 3017-3027.

WYNN, P. M., FAIRCHILD, I. J., BORSATO, A., SPÖTL, C., HARTLAND, A., BAKER, A., FRISIA, S. & BALDINI, J. U. L. 2018. Sulphate partitioning into calcite: Experimental verification of pH control and application to seasonality in speleothems. *Geochimica et Cosmochimica Acta,* 226**,** 69-83.

YALCIN, K., WAKE, C. P., KREUTZ, K. J., GERMANI, M. S. & WHITLOW, S. I. 2007. Ice core paleovolcanic records from the St. Elias Mountains, Yukon, Canada. *Journal of Geophysical Research,* 112**,** D08102-D08102.

ZHANG, Y. K. & YANG, W. T. 1998. Comment on "Generalized gradient approximation made simple". *Physical Review Letters,* 80**,** 890-890.

ZIELINSKI, G. A. 2000. Use of paleo-records in determining variability within the volcanism–climate system. *Quaternary Science Reviews,* 19**,** 417-438.

ZIELINSKI, G. A., MAYEWSKI, P. A., MEEKER, L. D., WHITLOW, S., TWICKLER, M. S., MORRISON, M., MEESE, D. A., GOW, A. J. & ALLEY, R. B. 1994. Record of Volcanism Since 7000 B.C. from the GISP2 Greenland Ice Core and Implications for the Volcano-Climate System. *Science (New York, N.Y.),* 264**,** 948-52.