# Is the reactivity of calcium carbonate fully understood?

## A step forward using an atomic scale perspective

**Janou A. Koskamp**

**Utrecht Studies in Earth Sciences**

**№ 240**

Members of the Dissertation Committee:

**Prof. dr. Steve. Parker**

University of Bath

United Kingdom

**Dr. Helen King**

Utrecht University

The Netherlands

**Prof. dr. René van Roij**

Utrecht University

The Netherlands

**Prof. dr. Martyn Drury**

Utrecht University

The Netherlands

**Prof. dr. Caroline Slomp**

Utrecht University

The Netherlands

ISBN: 978-90-6266-607-2

Utrecht Studies in Earth Sciences (USES) No. 240

Printed by: Proefschriftmaken

Layout by: Janou Koskamp

Cover design: Margot Stoete

Cover photo: © iStock

# Is the reactivity of calcium carbonate fully understood?

A step forward using an atomic scale perspective

**Is de reactiviteit van calcium carbonaat volledig begrepen?**

Een stap in de goede richting op atoom schaal

(met een samenvatting in het Nederlands)

**Proefschrift**

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof.dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 29 oktober 2021 des middags te 12.15 uur

door

**Janou Aranka Koskamp**

geboren op 4 juni 1992
te Boxmeer

## Promotoren:

Prof. dr. N.H. de Leeuw
Prof. dr. J.B.M. Middelburg

## Copromotoren:

Dr. M. Wolthers
Dr. A. Raoof

This study was conducted at the Geochemistry Group, Department of Earth Sciences, Faculty of Geosciences, Utrecht University, Utrecht, The Netherlands.

This study was funded by the Industrial Partnership Programme i32 Computational Sciences for Energy Research that is carried out under an agreement between Shell and the Netherlands Organization for Scientific Research (NWO).

# Contents

# Chapter 1 Introduction

1

## 1.1 The importance of calcium carbonate

In the Earth's surface environment, calcium carbonates are abundant (Sekkal and Zaoui, 2013), in particular in carbonate-rich sedimentary rocks, coral reefs, stalactites, and stalagmites. $CaCO_3$ is also the largest long-term stable sink for $CO_2$ in the global carbon cycle. Moreover, dissolved calcium and carbonate can be found, for instance, dissolved in natural waters (rivers and oceans), and are the main building blocks of many biominerals. Many organisms including primary producers such as coccolithophorids and cyanobacteria, as well as secondary producers foraminifera and diverse animals (bivalves, pteropods), process $CaCO_3$ to create the required polymorph for their biominerals in a (more or less) controlled manner (Endo et al., 2018; Meldrum and Cölfen, 2008; Morse et al., 2007). This formation is called biomineralization and can lead to surprising structures and polymorphs due to the modification of the environment such that thermodynamic conditions for $CaCO_3$ precipitation are optimized (Blue et al., 2017; Von Euw et al., 2017) The most common way this mineral is preserved and found in nature is in carbonate-rich sandstones, marble, or chalk and limestone that are composed of the remains of shells and skeletons of sea organisms and/or calcium carbonate matrix. Such rocks host roughly half of the earth's oil and gas reserves (Liteanu et al., 2013; Roehl and Choquette, 1990). Next to its natural role on Earth, calcium carbonates have been reported at the surface of Mars, using remote sensing techniques and studying weathering profiles (Bultel et al., 2019; Wray et al., 2016). These observations suggest the existence of carbonic acids dissolved in liquid water, more than 3.7 billion years ago, that reacted and changed Mars' surface, hinting towards an environment suitable for some form of life (Bultel et al., 2019).

Furthermore, research on calcium carbonate formation is important in a broad range of (applied) research and engineering fields and industry. From money-driven (e.g. oil industry (Olajire, 2015), paper industry (Vashistha et al., 2021), geothermal heat extraction (Pandey et

al., 2018), drinking water systems (Richards et al., 2018), and desalinization of (drinking) water (Burn et al., 2015; Zhao et al., 2018)), and the preservation of culture and nature, to the search for novel sustainable applications. Next, an overview of some of the most relevant research areas interested in the calcium carbonate system is given. Recovery of gas from calcium carbonate rocks with typically relatively high-porosity-low-permeability has a major hold back in the search for an optimal stimulation technique for enhanced oil and gas recovery (EOR/EGR) (Hawez and Ahmed, 2014; Liteanu et al., 2013). For example, accurate predictions of the outcome of the stimulation techniques are ambitious and there is still no consensus on the mechanisms behind it (Liteanu et al., 2013). Therefore, one of the fields of research applications of calcium carbonate is to improve the understanding of the dissolution and growth mechanisms to implement and correct pore network (Raoof et al., 2013, 2012) and continuum models (Golfier et al., 2002) for improved outcome-prediction and optimization of stimulation techniques for EOR/EGR. Moreover, the same approach/research is used to investigate the storage of $CO_2$ in such rock in depleted calcium carbonate reservoirs (Liteanu et al., 2012) where it can be combined with the enhanced recovery of oil and gas (Gozalpour et al., 2005) besides e.g. the sequestration in concrete (Parvan et al., 2021), to compensate for the impact of EOR/EGR and the consumption of oil and gas on the emission of $CO_2$ gas into the atmosphere, since $CO_2$ is a significant greenhouse gas that causes global warming and ocean acidification. Trapping and long-time storage of this gas is, therefore, a constant topic in both politics as science (Koytsoumpa et al., 2018).

Another important research area of calcium carbonate is marine calcifying organisms, such as corals and other sea(shell)-organisms in a changing environment. Many organisms depend on the surrounding conditions when biomineralizing their skeleton or shell, while their environment changes due to the ongoing natural and human-related changes in their environment (Caldeira and Wickett, 2003). Due to the rise in the emission of carbon dioxide,

the pH of the ocean decreased by transforming the extra $CO_2$ with water into bicarbonate and an extra proton. This process is called acidification of the ocean and triggers, among other things, a depletion of resources like the concentration of carbonate, and enhances the dissolution of biominerals due to the decreasing saturation index (Ω) with respect to calcium carbonate minerals (Bednaršek et al., 2012; Roleda et al., 2012). Yet, there is a significant knowledge gap when it comes to, for example, the impact on the final products of biomineralization (e.g. (Lemasson et al., 2017; Meng et al., 2019)). Understanding the effects of ocean acidification by understanding the molecular mechanisms has, at first place a great importance for the environment, but has also an economical factor related to the production of food and the shell-fish industry (Meng et al., 2019).

Beyond the description of natural processes, understanding biomineralization in nature gives us the opportunity to learn from it and use the knowledge in the development of new applications. Through the synthesis of novel bio(inspired)materials it is possible to guide or control their structure, morphology and characteristics, and they can be implemented in the industrial, environmental and medical fields (Di Tommaso et al., 2017; Endo et al., 2018; Li et al., 2017; Pasteris et al., 2008; Skinner, 2005).

The control over where, when, and how nucleation and growth take place is crucial in producing artificial minerals (bioinspired or not), and has applications in the production of advanced materials such as composite materials and photonic structures (Kaplan et al., 2017). A way of controlling them is by selection of different $CaCO_3$ polymorphs as a substrate that have different effects on the growth of other carbonate salts due to templating effects. This together with the control in supersaturation, makes it possible to compose an artificial mineral by directing the crystal formation via local supersaturations (Li et al., 2018). Another way to control the selection of the polymorph and the morphology is presence of impurities (e.g. $Mg^{2+}$ or phosphate (Noel et al., 2013)) with or without the combination of amino acids (Xie et al.,

2007). Where high concentrations of $Mg^{2+}$ in solution favours the growth of aragonite against the thermodynamic equilibrium conditions of the local growth media (Blue et al., 2017)

An example of bio-inspired insights are in the prevention of the formation of minerals that has led to the design of new anti-scaling agents with the use of small bioorganic molecules (Montanari et al., 2016). This is essential, for instance, when clogging due to scaling can lead to repair costs and interruption of industrial processes (Li et al., 2014). But also in tanks, heat exchangers and boreholes scaling endangers the safety and causes efficiency and economical losses (Li et al., 2018; Senthilmurugan et al., 2010).

Another interesting biomineralization process that may lead to applications in other fields is the formation of bird eggshell. The creation of an eggshell is more complicated than one initially might think, as it must meet strict criteria for an organism to survive. The main goal of the hard shell is the protection it provides for external physical stress, but it also allows the exchange of water and gas, it buffers fluctuations in external temperature and it provides calcium for the skeleton of the chick. Having such a complex set of functionalities makes the biominerals in an eggshell an outstanding source of inorganic and bio-organic materials. Given this, the required knowledge about the underlying mechanisms of the formation and functions of the eggshell may be implemented in several industries. Think about, the food industry, as in food supplements and due to the affordability and biodegradability, it has important applications in the chemical industry (Guru and Dash, 2014). Plus, it has led to various applications in the cleaning industry in the treatment of environmental pollutants like textile dyes and ionized heavy metals. Because of fundamental studies (Innocenti Malini et al., 2017) about the detailed processes that take place, during biomineralization, this biomaterial could expand its usage in other areas (Guru and Dash, 2014).

Finally, calcium carbonate systems are used as a steppingstone to next-level applications in for example the semiconductor industry. The complex shaped calcium carbonate crystals, for example, think of the magnificent corals, can be converted into lead halide perovskite semiconductors with adjustable bandgaps using ion-exchange reactions. Having the new material with a three-dimensional structure of $CaCO_3$ contributes to the development of next-generation optoelectronics and new forms of catalysts (Holtus et al., 2018). All these findings broadened our understanding, made it possible to mimic, and ultimately expand upon nature's mineralization strategies and contribute to the development of functional (microscale) materials.

The following sections start with a description of the different forms of calcium carbonate that we encounter in nature. Subsequently, the theory behind crystal formation is briefly discussed, to be able to embed the different studies in the crystallization landscape. Then, the background is given to explain why molecular dynamics (MD) can provide complementing insights.

## 1.2 Polymorphs

Calcium carbonate has six different polymorphs: three forms are crystalline, two are hydrated phases, and one is non-crystalline.

Amongst the crystalline phases, firstly, calcite is thermodynamically the most stable polymorph at ambient temperature and pressure and therefore most commonly found in nature, in the Earth's surface environment. Calcite is hexagonal, a trigonal system, and its space group is $R\bar{3}c$ (Rachlin et al., 1992). $Ca^{2+}$ is coordinated with six carbonate oxygens and the structure is built up of alternating calcium and carbonate groups, the carbonate's plane perpendicular and parallel to the c-axis. Its most stable cleavage plane is $\{10\bar{1}4\}$, which results in a perfect rhombohedral shape (Rachlin et al., 1992). Secondly, aragonite is the second most commonly found polymorph of calcium carbonate and is less stable than calcite at ambient conditions.

Aragonite is often detected in biomineralization processes, for instance in freshwater cultured pearls (Jacob et al., 2011), in the shell of molluscs (Weiss et al., 2002), stony corals (Von Euw et al., 2017), and in embryos of a pulmonate snail (Marxen et al., 2003). The crystal system of aragonite and space group are, respectively, orthorhombic and *Pmcn* and $Ca^{2+}$ has a nine-fold coordination with $CO_3^{2-}$. Thirdly, the least stable crystalline polymorph is vaterite and occurs mainly as an intermediate phase before transforming into aragonite or calcite (Rodriguez-Blanco et al., 2011). The crystalline structure of vaterite has been difficult to capture and while it has been studied extensively, there is no complete consensus. A decade ago, two consistent vaterite structures have been identified: a hexagonal and monoclinic crystal system with space groups $C_2$ and $P3_121$. Both structures are supported by computational *ab initio* calculations and experimental methods, like Nuclear Magnetic Resonance (NMR), X-ray diffraction (XRD) and Raman (Costa et al., 2016).

There are two hydrated calcium carbonate phases. Firstly, one of the two hydrated phases, monohydrocalcite ($CaCO_3$–$H_2O$) (MHC), is a metastable phase that transforms into calcite or aragonite by losing its water molecules in a dehydration process, commonly at elevated temperatures. The MHC crystal is classified as a hexagonal crystal system and its space group is $P3_1$ (Effenberger, 1981). The second hydrated phase of $CaCO_3$ is Ikaite ($CaCO_3$–$6H_2O$) and is, like MHC, less often found in nature compared to the anhydrous polymorphs and serves as a precursor for the more stable crystalline forms. This monoclinic system has a C2/c space group (Dickens and Brown, 1970) and converts into vaterite when reaching the maximum density of water around 277 K. Nevertheless, this polymorph occurs in marine sediments and in sea ice as ikaite and is stabilized with a combination of low temperatures, below 278 K, a high pH and the presence of nucleation inhibitors of anhydrous crystals such as phosphate (Bischoff et al., 1993; Hu et al., 2015). $Mg^{2+}$ is another inhibitor proven to stabilize ikaite by retaining the transformation to calcite (Tollefsen et al., 2018).

Lastly, besides the crystalline forms and hydrated phases, calcium carbonate also exists in amorphous states all labelled as amorphous calcium carbonate (ACC). The fact that there are different polyamorphs of ACC became clear after the observation of a change in properties, like solubility, from one structure to the other (Gebauer et al., 2010, 2008). Only under lower supersaturation (concentrations lower than ~10 mM $CaCO_3$) (Michel et al., 2008) pro-vaterite and pro-calcite are observed in ACC (Gebauer et al., 2010), driven by the conditions during nucleation and further crystallization of this ACC resulted in one of the crystal polymorphs (Cartwright et al., 2012). ACC is known to play an important role in biomineralization, as it occurs as a precursor towards the more stable crystalline forms. In short, usually, biomineralization is thought to proceed roughly in three stages (Addadi et al., 2003): first, hydrated ACC is formed, upon dehydration the second stage is characterized by the rather unstable anhydrous ACC that rapidly leads to the third stage and (semi) final biomineralized crystal (Addadi et al., 2003). Whereas the most stable polymorph, calcite, is formed directly or via vaterite at room temperatures and colder, aragonite is observed under higher temperatures again via vaterite (Rodriguez-Blanco et al., 2011). During $CaCO_3$ nucleation, all three polymorphs have been observed simultaneously in liquid cell transmission electron microscopy (TEM) (Nielsen et al., 2014).

## 1.3 Crystallization theory in a nutshell

The crystallization of calcium carbonate has been studied for decades (De Leeuw and Parker, 1997; Reddy and Nancollas, 1971; Sand et al., 2016; Schmidt et al., 2019; Smeets et al., 2015; Ueckert et al., 2020; Wolthers et al., 2008), the mineral profiled itself as an easy-to-work-with model system that can be studied and unravelled in many different settings. Thus, the acquired knowledge can, where possible, be extrapolated to other materials to coordinate crystallographic characteristics and open the door for new applications, including the contribution to “green chemistry” (Gower, 2008).

The theory behind crystallization can be divided into two approaches, the classical and non-classical theory. The classical nucleation theory describes nucleation and further crystal growth as two separate topics. The nucleation of new crystals is depending on the level of supersaturation with respect to $CaCO_3$ in the solution. Nucleation can occur homogeneously or heterogeneously, where the first one requires higher activation energy. In homogenous nucleation, the nucleus appears spontaneously, however this only happens when the supersaturation is high enough.

When the solution is supersaturated, nucleation is only successful for further growth if the size of the nucleus is beyond its critical size, in other words, the nucleus is stable and doesn't dissolve again. This is because of the energy balance between a molecule in the bulk of the newly nucleated phase (lower than a dissolved molecule) and a molecule at the interface (higher free energy than a dissolved molecule). Below this critical size, the free energy penalty coming from the interfacial energy, is too high due to the high the nucleus' interface/bulk ratio to be a stable nucleus. Modifying the solution composition can alter the interfacial energy and therewith alter the critical nucleus size. Also increasing the supersaturation increases the probability and shortens the induction time for successful nucleation and further growth. Whenever the nucleation starts on an existing surface, less energy is required and because of the use of an external element it is a heterogeneous process. The formation of the crystal, as described in the classical nucleation theory, is controlled by ion diffusion to and on the crystal surface, followed by their incorporation in the crystal lattice.

The non-classical crystallization theory is overall quite similar to the classical nucleation theory, but deviates from the classical theory in explaining the crystallization via an amorphous precursor (Addadi et al., 2003; De Yoreo and Vekilov, 2003) or the appearance of mesocrystals (Meldrum and Cölfen, 2008). The latter are obtained by the formation of liquid like calcium carbonate phases such as a dynamically-ordered liquid-like oxyanion polymer (DOLLOP)

(Demichelis et al., 2011) or a Polymer-Induced Liquid-Precursor (PILP) (Gower, 2008). Since the formation of the three main polymorphs of calcium carbonate, calcite, aragonite and vaterite, occurs after the formation of ACC, understanding this amorphous precursor is an important step in elucidating the crystallization process and polymorph selection. ACC is unstable in natural systems and can convert within 2 minutes into one of the crystalline calcium carbonate polymorphs (Rodriguez-Blanco et al., 2011). There is still an open debate on which theory most accurately explains $CaCO_3$ formation (Gebauer et al., 2018; Smeets et al., 2017).

## 1.4 Molecular Dynamics

In the last decades, Earth scientists embraced computational modelling as a tool to study Earth materials and elucidate reaction mechanism. Exploring and analysing materials at the molecular level provides in-depth knowledge of the mechanisms that drive processes at the micro- and macroscopic levels. With computational techniques it is feasible to go beyond the experimental limitations. For instance, molecular dynamic (MD) simulations have been carried out to help in the characterization of minerals, and their growth and dissolution mechanism (Cygan and Kubicki, 2001). It is also used to study the behaviour of isotopes, their stabilities, compositional preferences and the interaction with organic (macro)-molecules (Eiler et al., 2014; W. Zhang et al., 2020).

Despite the impact of MD simulations in Earth sciences, its applicability is limited by three main factors: the accuracy of the model, in particular the force fields selected, model domain size and the achievable simulation times that are limited by the available computer power. At the expense of time scale and system size, greater accuracy can be obtained combining MD with density functional theory (DFT) based methods, also called *ab initio* MD (AIMD). Materials can be characterized by calculating the properties of various compositions. However, the limitations on the number of atoms and the short timescales in the order of picoseconds,

prohibit or complicate the comparison with more complex systems studied in experiments. Nonetheless, AIMD and DFT are used to study for instance relative stabilities (Li et al., 2009), adsorption behavior (Zivković et al., 2021), and crystal defects (Živković and De Leeuw, 2020). In the need to study larger-scale problems where AIMD methods fall short, atomistic force fields are generally a suitable alternative. Such force fields describe the interactions of the atoms in terms of potentials in which the electrons are implicitly accounted for (this method is called classical MD). With these methods, one can simulate systems containing millions of atoms for on the order of (hundreds of) nanoseconds. It must be noted that this method is less accurate compared to AIMD and depends heavily on the quality of the parameters and the potential forms used to describe the system. The parameters are usually derived from either *ab initio* methods or are obtained empirically by fitting to experimental data. In both cases, the force field is constructed by tweaking the potential parameters in such a way that it mimics the properties observed experimentally or with *ab initio* as accurately as possible. Common properties to fit are the structure, vibrational frequencies, phonon dispersion curves, elastics, and dielectric constants.

The main technique used in this thesis is classical molecular dynamics to get a glimpse into the molecular scale evolution of calcium carbonate systems. There are a few important aspects to consider when processing MD data. On the one hand, when simulating non-equilibrium systems, the result of each trajectory may depend on the initial conditions, making the initial setup very important for the quality of the simulation (output) and a potential cause for simulation bias. On the other hand, a great advantage of equilibrium simulations is that the outcome is independent of the initial conditions, making the results reproducible and robust. Last but not least, it is important to realise that a force field might not be capable of showing certain events, independent of the simulation time. For example, non-dissociative force fields are not able to capture a (de)protonation mechanism. Despite the drawbacks of classical

molecular dynamics, it is still believed that MD can complement experiments and that their the combined efforts have led to resolving fundamental reaction mechanisms (Orme et al., 2001; Reischl et al., 2019).

## 1.5 Research objective

This thesis contributes to determine the behaviour of calcium and what its role is in nucleation, growth, and dissolution of calcium carbonate from an atomistic perspective. Chapter 2 briefly describes the theory behind the computational technique used for our molecular dynamic simulations. The purpose of this research is to strengthen our knowledge and complement experimental observations with mechanistic insights to fill knowledge gaps (see below) by gathering information at a molecular level. This molecular insight can be used in larger-scale models (Bracco et al., 2013; Kurganskaya and Rohlfs, 2020; Wolthers et al., 2012a) in which smaller details are usually ignored, while such details are at times important to explain disagreement between real-life practice and theory.

To achieve this, we started from the very first step of any calcium carbonate mineralization process, a calcium ion dissolved in pure water, to study the water dynamics around the ion and to find an answer on how important the dehydration of $Ca^{2+}$ upon mineralization actually is (Chapter 3). Looking at the dehydration of cations is not new, however, large unexplained discrepancies remain within classical MD, depending on the method, and compared to *ab initio* or experimental methods. We addressed this dispute by using *ab initio* dynamics and classical molecular dynamics and compared different data processing methods and force fields. Our model was represented by calcium in a small box surrounded by liquid water and by using periodic boundary conditions (PBC) we could describe the $Ca^{2+}$ as being fully dissolved in an infinitely large ocean of pure water. Among others, the results indicate that calcium

dehydration kinetics is faster than previously considered, in agreement with neutron scattering experiments.

Additionally, we studied the next phase of mineralization where calcium reacts with carbonate to form amorphous calcium carbonate clusters in the presence or absence of one or more biomolecules. Using this approach, we could see the formation of clusters and could characterize them by size, charge, shape, and degree of hydration. By adding biomolecules, we could see how these initial steps in the nucleation process are influenced by the presence of organic molecules. Initially, we simulated this process using accepted combinations of force fields (Chapter 4). In this process, we observed that, contrasting to experimental evidence, calcium does not bind preferentially to the carboxyl groups of the biomolecules. In chapter 5, we refined the force field and repeated some of the simulations to investigate the impact of the refined force field on the outcome of chapter 4.

In chapter 6, we investigated the material formed upon nucleation, transformation, and growth to calcite crystals. We focused on the dynamic equilibrium between solid $CaCO_3$ and a solution with dissolved constituents, which is known to result in (equal rates of) dissolution and growth of calcite crystals. We examined this dynamic equilibrium in terms of the thermodynamic and kinetic stability of calcium ions positioned in different sites in the most stable $\{10\bar{1}4\}$ surface. In addition, by placing a small island of 16 $CaCO_3$ units we can observe the stability of calcium adsorbed onto the crystal. Besides classical MD, we also used Umbrella Sampling (US) to explore the energy landscape. In the first chapters, the solids under investigation consisted purely of $Ca^{2+}$ and $CO_3^{2-}$ (whether or not) together with traces of organic molecules. In contrast, in nature, impurities in the crystal structure are the norm rather than the exception. To investigate the impact of impurities on the stability of calcite, we disseminated magnesium in the calcite slab; magnesium is a very common impurity. However, studying a $Mg^{2+}$-rich system is not straightforward, due to the extremely slow water kinetics around the magnesium ions,

which would need unachievably long simulation times in classical MD. Therefore, instead of focussing on $Mg^{2+}$, we concentrated on the effect of this impurity on the stability of surface $Ca^{2+}$. Via this approach, we tried to explain the non-linear impact of Mg on the solubility of the calcite surface.

Throughout this thesis our focus has been on analysing the effect of different settings on $Ca^{2+}$ (dissolved versus clustering, with and without biomolecules, versus crystalline surface, with and without impurities) using computational methods. We did this to provide mechanistic insight that complements experimental observations and can help to update transport-reaction models (at larger scales) with our new outcomes.

# 1

2

# Chapter 2 Computational background and methodology

## 2.1 Molecular Dynamics



Since the mid-seventies, simulations at atomic level play an active role in science as how we know it today, with the description of the interaction between the ions in alkali halides as one of the first instructive examples (Catlow et al., 1977). Classical mechanics was implemented to describe the interactions between the atoms in the simulations, via a set of interatomic potential equations and their associated constants. Such a set of potentials and parameters is called a forcefield. In the beginning of the nineties, force fields for oxyanions, like carbonate and sulphate, were developed, which opened the door to more complex simulations of a larger variety of crystalline materials. Besides time-resolved simulations (Classical Molecular Dynamics (MD)), force fields can also be used to obtain energy minimisations, where the local minimum is explored by an algorithm considering all described interactions. Accounting for all energy contributions, we use the Born model as starting point. In this model the long and short-range interactions are included and calculated via different methods including the Ewald summation, and several potentials to describe short range interactions.

### 2.1.1 The Born model

The Born model (Born and Huang, 1954) describes the total energy of the interaction of two atoms as a combination of long-range electrostatic forces and short-range forces by the following equation:

$$U(r_{ij}) = \sum_{ij} \frac{q_i q_j}{4\pi\varepsilon_0 (r_{ij} + 1)} + \sum_{ij} \Phi_{ij}(r_{ij}) \qquad (1)$$

where $q_i$ and $q_j$ are the charges on atoms $i$ and $j$ respectively, $r_{ij}$ is the distance between atoms $i$ and $j$, and $\varepsilon_0$ is the permittivity of free space. $\Phi_{ij}$ is the potential describing the short-range interaction. In the first term, the Coulombic electrostatic energy of the interaction is considered

a long-range force. This electrostatic energy is an important term in ionic and inorganic systems but has a significant bottle neck due to its slow convergence. To circumvent this problem, the Ewald summation and closely related particle-particle particle-mesh solvers are used to approach the corresponding long-range energies as accurate as possible.



### 2.1.2 Ewald sum and particle-particle particle-mesh solver

Published in 1921, the Ewald summation provided an adequate way to sum the coulombic interactions between atoms, taking in consideration all infinite periodic images (Ewald, 1921). The method approaches the potential energy by replacing the slow convergent term (first part in equation 1) by two series that converge fast and an extra constant:

$$U_{Ewald} = U^r + U^m + U^0 \tag{2}$$

where $U^r$ is the sum in the real space, the reciprocal sum is given by $U^m$, and the constant, known as the self-term is ($U^o$) that removes the interaction of each of the artificial counter-charges with itself. The three separate terms can be written in more detail as:

$$U^r = \frac{1}{2}\sum_{i,j}^{N}\sum_{n} q_i q_j \frac{erfc(\alpha r_{ij,n})}{r_{ij,n}} \tag{3}$$

$$U^m = \frac{1}{2\pi V}\sum_{i,j}^{N} q_i q_j \sum_{m\neq 0} \frac{e^{-\left(\frac{\pi m}{\alpha}\right)^2 + 2\pi i m(r_i - r_j)}}{m^2} \tag{4}$$

$$U^0 = \frac{-\alpha}{\sqrt{\pi}}\sum_{i=1}^{N} q_i^2 \tag{5}$$

where α is the convergence parameter, which controls the distribution of the summation between real and reciprocal space. As α increases, the real-space sum converges more rapidly and the reciprocal sum more slowly. $V$ is the volume of the system; $m$ is the reciprocal space vector, and $N$ the number of particles. The cell-coordinate vector is $n=(n_1,n_2,n_3)=n_1Lx+n_2Ly+n_3Lz$, where $x$, $y$, and $z$ are the Cartesian coordinate unit

vectors. The origin cell is located at *n=(0, 0, 0)* with image cells located at *Ln* intervals in all three dimensions as *n* goes to infinity.

The complimentary error function, that can be evaluated and truncated in real space decreases monotonically as *x* increases and is defined by:

$$erfc(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} du \qquad (6)$$

Another method to deal with the slow convergence of the electrostatic forces has been developed in 1989 and is called the particle-particle particle-mesh solver (PPPM) (Hockney and Eastwood, 1988; Luty et al., 1994; Rajagopal and Needs, 1994) Which maps the charges of the particles in a 3D map and solves the equation of Poisson on the mesh with Fast Fourier Transformations. Via interpolation of the electric field, the electrostatic force is calculated on the location of the particles. This technique is closely related to the particle-mesh Ewald summation, however the PPPM solver scales better due to the FFTs and is therefore usually faster (Pollock and Glosli, 1996).

### 2.1.3 Short range forces

The second term in equation 1 is the energetic contribution of the short-range interaction. At short range both, repulsive and attractive forces are included: the former are due to the overlap of ion charge clouds, the latter represent van der Waals attraction between adjoining electron charge clouds. The covalent bond potentials that represent the repulsion and attraction are captured in two, three and four-body potentials, respectively, the bond, the angle, and the dihedral. The two-body potential ($\Phi_{ij}(r_{ij})$) is the main component in the short-range energy and is only dependent on the distance between two atoms. The non-covalent bond potentials describe the interatomic interactions and have a repulsion and attraction term; commonly used potentials are the Buckingham and Lennard-Jones potential (see below for an explanation). The

three-body potential ($\Phi_{ijk}(r_{ijk})$) is a function of three atoms and is the main contributor in the vibrational characteristics. The four-body potential ($\Phi_{ijkl}(r_{ijkl})$) is responsible for the three-dimensional shape of the molecule. A more detailed description and definition of the potentials used within this study is given in the following subsections.



### 2.1.3.1 Two body potential functions

#### 2.1.3.1.1 Body harmonic function

The simplest two-body potential is the body harmonic function: a straightforward harmonic, proportional to the square of the difference of the distance between two atoms and an equilibrium value:

$$U(r_{ij}) = \frac{k_{ij}}{2}(r_{ij} - r_0)^2 \qquad (7)$$

where $k_{ij}$ is the bond force constant between atoms $i$ and $j$, with $r_{ij}$ the distance between atoms $i$ and $j$, and $r_0$ the equilibrium distance between the atoms.

#### 2.1.3.1.2 Buckingham potential

The Buckingham potential is the most extensively used form of the two body potential functions for ionic solids, with the exponential repulsive term before the attractive van-der-Waals term:

$$U(r_{ij}) = A_{ij}e^{-\frac{r_{ij}}{\rho_{ij}}} - \frac{C_{ij}}{r_{ij}^6} \qquad (8)$$

$A_{ij}$ is related to the size of the atoms involved and $\rho_{ij}$ to the electron density. The exponential term in the Buckingham potential is computationally rather expensive to calculate, however this potential is providing a more accurate estimation of the repulsive forces and is therefore regularly implemented for the more critical interactions.

2.1.3.1.3 Lennard-Jones potential

The widely used Lennard-Jones potential is, like the Buckingham potential, mostly for non-bonded interactions and is made up of a repulsive $r^{-12}$ (dominant at shorter distances, close to zero) and an attractive $r^{-6}$ term (dominant at larger distances). The equilibrium distance depends on the values of the two variable parameters $A$ and $B$, which are usually derived empirically. A and B can be rewritten in terms of $\varepsilon_{ij}$ representing the depth of the energy well, and the minimum separation of two atoms, $\sigma$ (Figure 2.1):

$$A = 4\epsilon\varepsilon_{ij}\sigma^{12} \tag{9}$$

$$B = 4\epsilon\varepsilon_{ij}\sigma^{6} \tag{10}$$

where $\sigma$ is:

$$\sigma = \frac{r_{ij}}{2^{\frac{1}{6}}} \tag{11}$$

Lennard-Jones potentials are commonly chosen in large systems over Buckingham potential since the computational costs are lower.

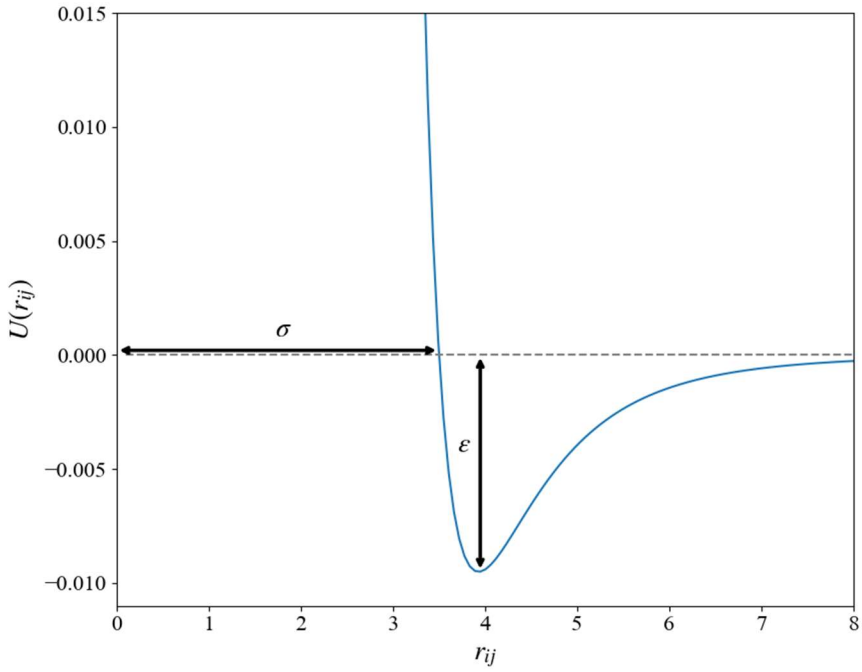



Figure 2.1 an example of a Lennard-Jones potential

An adaptation of pair-wise van-der-Waal interactions is regularly implemented to speed up the simulations. The use of a taper on the tail of the potential from 0.6 to 0.9 nm sets the value of the resulting energy of the potential to zero to circumvent calculations that lead to insignificant contributions to the energy. This can be done for both the Buckingham and the Lennard-Jones potential (Raiteri et al., 2010).

### 2.1.3.2 Three body potential

We use three-body forces to describe the angle between three atoms within a molecule. Having a potential instead of a fixed angle allows the molecule to bend. The extent of bending and the equilibrium angle is given by a simple harmonic spring with the $k_{ijk}$ as the bend constant for the three-body bond and $\theta_0$ as the equilibrium angle:

$$U(\theta_{ijk}) = \frac{k_{ijk}}{2}(\theta_{ijk} - \theta_0)^2 \qquad (12)$$

where $\theta_{ijk}$ is the angle between the two bonds i–j and j–k.

### 2.1.3.3 Four-body potential

Guaranteeing the three-dimensional structure, four-body potentials are often applied to constrain the dihedral angle made by four atoms. This torsional angle is the angle between the two planes built between i, j and k and j, k, and l, respectively. A typical four-body potential is written as:

$$U(\phi_{ijkl}) = k_{ijkl}\left(1 - scos(n\phi_{ijkl})\right) \quad (13)$$

where $k_{ijkl}$ is the torsional force constant, $s = \pm 1$ depending on which conformation is the lowest in energy, and $n$ is the periodicity of the torsion.

Another way to enforce planarity is by using a potential on the improper angle. The potential used in this study is based on the following equations:

$$U(\phi_{ijkl}) = U_i + U_{aa} \quad (14)$$

$$U_i = K\left[\frac{\chi_{ijkl} + \chi_{ijkl} + \chi_{ljik}}{3} - \chi_0\right]^2 \quad (15)$$

$$U_{aa} = M_1(\theta_{ijk} - \theta_1)(\theta_{kjl} - \theta_3) + M_2(\theta_{ijk} - \theta_1)(\theta_{ijl} - \theta_2) + M_3(\theta_{ijl} - \theta_2)(\theta_{kjl} - \theta_3) \quad (16)$$

where $U(\phi_{ijkl})$ is the improper term and $U_{aa}$ is an angle-angle term. The four atoms in an improper potential are ordered i, j, k and l. The three $\chi$ terms in equation 15 are an average over the three out-of-plane angles. For example, $\chi_{ijkl}$ represents the angle between the plane defined by i, j and k and the plane defined by j, k and l. The various bond angles, θ, are given for different combinations of the four atoms together with their angle in equilibrium denoted with numbers 1, 2 and 3. As j appears in all bond and plane angles at the same position, this atom is labelled as the symmetry centre.

### 2.1.4 Transferability of potential and potential parameters

In general, potential parameters are derived from empirically or from *ab initio* calculations for that atom and system specifically. However, sometimes this data is not available. The transfer of potentials from one system to another can be very beneficial in combining systems and creating different ones. It has been satisfactorily done in several cases, for example when the potentials of binary oxides were correctly transferred to ternary oxides (Bush et al., 1994; Lewis and A Catlow, 1985; Purton and Catlow, 1990). Also, extrapolating potential parameters for one element to another is not uncommon, for example the transfer from alkali halides to the divalent strontium halides provided reasonable reproduction of the experimental data (Telfer et al., 1997). This set a precedent for confidently and carefully combining force fields and extrapolating parameters where necessary.

### 2.1.5 Shell model

Usually, the force field of the system describes every atom by a point charge, without a flexible dipole moment. However, in some models, the electronic polarizability of an atom is accounted for via a shell model (Dick B G and Overhauser, 1958) to improve the description of the system. In such a shell model, every polarisable atom, in our case the oxygen atoms of a water molecule, is represented by a core and a shell. While the core represents the nucleus and the core electrons, the massless shell describes the valence electron of which the position has no physical meaning (Harding, 1990). The core and the shell are connected via a harmonic spring.

$$\Phi(r_i) = \frac{1}{2} k_i r_i^2 \qquad (17)$$

$k_i$ is the force constant for the harmonic spring, connecting the core and shell. $r_i$ is the distance between the core and the shell. The total charge of the ion is given by the charge of the core plus the charge of the shell. Subsequently, the polarizability of the ion can be determined using the spring constants and both charges:

$$\alpha_i = \frac{Y^2}{4\pi\varepsilon_0 k_i} \tag{18}$$



Here, $Y$ is the charge on the shell, $k_i$ the force constant for the spring connecting the core and the spring, and $\varepsilon_0$ is the permittivity of free space. The values of $Y$ and $k$ are obtained by empirical fitting to experimental properties such as dielectric and elastic constants. The polarizability is described in terms of the displacement of the shell relative to the core. Short-range forces are assumed to act between the shells, with coulombic interactions between the shells and cores of different atoms.

The electronic polarizability is implemented slightly differently in the codes in this thesis. The general principle is the same and in this thesis the variant in the DL_POLY (Todorov et al., 2006) code is used where the shell is given a mass, which is negligibly small compared to the mass of the core, to ensure that there is no exchange of energy between the core and shell during the simulation.

## 2.2 Software dedicated to MD

Common software employed to carry on MD simulations are DL_POLY (Todorov et al., 2006) and LAMMPS (Plimpton, 1995). In the current work, we used both. In chapter 3, we used DL_POLY because of the implementation of the shell model. In chapter 4-6, we moved to LAMMPS looking for performance and faster calculations. In DLPOLY we used the Ewald summations and in LAMMPS we used the PPPM.

# 2

3

# Chapter 3 Reconsidering Calcium Dehydration as the Rate-determining Step in Calcium Mineral Growth

Accepted as: **Koskamp J. A., Ruiz-Hernandez S. E., Di Tommaso D., Elena A. M., De Leeuw N. H., Wolthers, M.** Reconsidering Calcium Dehydration as the Rate-Determining Step in Calcium Mineral Growth. *The Journal of Physical Chemistry C*

**Abstract:** The dehydration of cations is generally accepted as being the rate-limiting step in many processes. Molecular dynamics (MD) can be used to investigate the dynamics of water molecules around cations, and two different methods exist to obtain trajectory-based water dehydration frequencies. Here, these two different post-processing methods (direct method *versus* survival function) have been implemented to obtain calcium dehydration frequencies from a series of trajectories obtained using a range of accepted Force Fields. None of the method combinations reproduced the commonly accepted experimental water exchange frequency of $10^{-8.2}$ $s^{-1}$. Instead, our results suggest much faster water dynamics, comparable with more accurate *ab initio* MD simulations and with experimental values obtained using neutron scattering techniques. We obtained the best agreement using the survival function method to characterize the water dynamics, and we show that different method combinations significantly affect the outcome. Our work strongly suggests that the fast water exchange kinetics around the calcium ions is not rate limiting for reactions involving dissolved/solvated calcium. Our results further suggest that, for alkali and most of the earth alkali metals, mechanistic rate laws for growth, dissolution, and adsorption, which are based on the principle of rate-limiting cation dehydration, need careful reconsideration.

## 3.1 Introduction

The dynamics of water molecules surrounding reacting cations is generally considered as rate-determining for the reactions involving these ions in aqueous solutions (Andersson et al., 2016; Di Tommaso et al., 2014; Joswiak et al., 2018; Kowacz et al., 2007; Nielsen and Toft, 1984; Stumm and Morgan, 1996). Almost sixty years ago, ultrasonic sound adsorption experiments performed by Eigen, indicated a large range of water exchange frequencies around cations (Eigen and Wilkens, 1965; Petrucci, 1971). Years later, Nielsen (1984) found a correlation between the frequency of ions entering surface kink sites and the dehydration frequencies of the cations in this surface, where the attachment frequency of the ions was approximately $10^{-3}$ times its dehydration frequency. Nielsen proposed this difference was due to the ion diffusion from solution to a kink site, a process included in his overall attachment frequency. Nevertheless, as anions dehydrate faster, cations are considered to control the reaction rate (Nielsen, 1984; Piana et al., 2006; Pokrovsky and Schott, 2002). These general findings have been incorporated into mechanistic models explaining crystal dissolution (e.g. (Lasaga and Lüttge, 2004, 2003)) and crystal growth (Bracco et al., 2012; Christoffersen and Christoffersen, 1990; Hong and Teng, 2014; Larsen et al., 2010; Nielsen et al., 2012; Sand et al., 2016; Stack and Grantham, 2010; Wolthers et al., 2012b; Zhang and Nancollas, 1998). Essentially, in such models it is assumed that the exchange frequency of water molecules between the first hydration shell around cation surface sites and bulk water controls the reaction rate.

The dynamics of water molecules can be investigated using Molecular Dynamics (MD) simulations, for example of water molecules around dissolved ions (Di Tommaso et al., 2014; Larentzos and Criscenti, 2008; Rode et al., 2005; Rode and Hofer, 2006), at mineral-water interfaces (De La Pierre et al., 2016; Raiteri et al., 2015; Sena et al., 2015; Wolthers et al., 2012a) and in amorphous materials (Ruiz Hernandez et al., 2015). Several studies have employed MD to study ion complexation and dynamics in water (e.g. (Di Tommaso and de

Leeuw, 2010; Joung and Cheatham, 2009; Lee and Rasaiah, 1996; Pham et al., 2015)), whereas the technique has also been used to analyse adsorption of small organics on mineral-water interfaces (Ruiz Hernandez et al., 2015; Ruiz Hernandez and De Leeuw, 2015) , the structural changes in (water near) mineral interfaces (Spagnoli et al., 2006a; Wolthers et al., 2013) and the influence of ion impurities on the kinetic stability of amorphous materials (Koishi et al., 2018).



In the case of calcite, combined experimental and computational work has shown that electrolyte ions impact water exchange frequencies and crystal growth rate (Di Tommaso et al., 2014). MD simulations also confirmed that the formation of the kink sites on the dissolving edge of the obtuse step of calcite is the rate limiting step and the dissolution of this edge is thermodynamically favoured (de Leeuw et al., 1999) and the addition of the first units of calcium carbonate is enthalpically unfavourable (de Leeuw and Parker, 2001). Also, there is good agreement in the description of the structure of water at the calcite-water interface, as observed computationally (De Leeuw and Parker, 1998; Fenter et al., 2013), using force-modulated Atomic Force Microscopy (Marutschke et al., 2014; Söngen et al., 2018), or small-angle X-ray reflectivity (Fenter et al., 2013) and neutron diffraction (Heberling et al., 2011).

Despite the promising agreement on the structure of the calcite-water interface obtained computationally and experimentally, large discrepancies remain in observations of water dynamics. Recent studies using *ab initio* MD (AIMD) calculations (Di Tommaso and De Leeuw, 2010; Hofer et al., 2004) show that exchange frequencies around dissolved calcium are almost three orders of magnitude higher than Eigen's value. Comparably high frequencies have also been observed around dissolved calcium using incoherent quasi-elastic neutron scattering (IQUENS (Hewish et al., 1983)), from which the authors concluded that the water exchange frequency around calcium ions is faster than $10^{10}$ $s^{-1}$. Other MD simulations have indicated frequencies much closer to those measured by Eigen (De La Pierre et al., 2016), while classical

MD simulations have resulted in water exchange frequencies around dissolved calcium and calcium in calcite surfaces (Wolthers et al., 2013) that are similar to those observed in *ab initio* MD and IQUENS. Besides, it is still debated if dehydration of the attaching species is rate limiting in crystal growth (Andersson et al., 2016) or that step-migration of carbonate ions is rate limiting for kink nucleation (De La Pierre et al., 2017).



While an evaluation of the differences between the various experimentally observed frequencies is beyond the scope of the paper, the large range in computationally determined water exchange frequencies may be explained by differences between the force fields used in MD methods, and/or the approach followed to obtain exchange frequencies from the MD output. Here we present our evaluation of both potential areas for discrepancy. We have simulated identical systems of dissolved calcium using four sets of force fields, some of which have previously provided contradictory results. We next evaluated two generally accepted methods to determine water exchange frequencies. The results from our simulations and calculations provide new insights into the impact of the water dynamics around calcium on the reactivity of materials containing this cation. Our results dispute that calcium dehydration is the rate-limiting step in mineral formation and dissolution and suggest that other processes are rate-determining instead. We discuss implications for mechanistic growth models.

## 3.2 Computational methods

### 3.2.1 Classical molecular dynamics

The MD simulations were conducted on a cubic simulation cell containing 809 water molecules and a single calcium ion. They were performed with DL_POLY 4.09 (Todorov et al., 2006) which was modified to include a tail required for the SPC/fw-tail Force Field (Raiteri et al., 2015). The Verlet leap-frog scheme and the Nosé-Hoover algorithm (Hoover, 1985; Nosé, 1984) were used to integrate the equations of motion with a time step of 0.2 fs and to maintain

the temperature at 300 K with 0.5 ps relaxation time. The simulations were carried out in the isothermal-isobaric (constant NPT) ensemble using a Nosé-Hoover thermostat to maintain the average temperature and pressure at T = 300K and P = 1 atm. The total simulation time for each of the systems was 144 ns including 200 ps of equilibration. An 8.0 Å cut-off was used for the van der Waals forces, except for the SPC/fw-tail Force Field in which the cut-off was defined at 9.0 Å with a tail from 6.0 Å (Raiteri et al., 2015).



We tested four different force fields to analyse the water dynamics around the calcium cation: first, the shell-model (SM-lj) developed by de Leeuw and Parker and revised by Kerisit and Parker; (de Leeuw and Parker, 1998; Kerisit and Parker, 2004) and second, the simple point charge flexible water model (SPC/fw-tail (Wu et al., 2006),). Comparison between these first two is the focus of the main paper. However, in order to guarantee the robustness and comparability of these two water force fields, we also included an in-depth assessment of the dynamics and structures produced by SM-lj and SPC/Fw in comparison with two further force fields, i.e. SWM4-NDP (Bruneval et al., 2007; Lamoureux et al., 2006) and TIP3P (Jorgensen et al., 1983; Jorgensen and Jenson, 1998), which are also commonly used for simulations of calcium-containing systems or in electrolyte solutions. The full assessment of all four force fields can be found in the Supplementary Information (SI), which includes a study of the different structural parameters, including the dipole moment, and the three- and four-body order parameters. The dynamics of the system, such as diffusion coefficients and vibrational spectra can also be found in the SI, as well as further details and a table with parameters for the different force fields used in this study.

### 3.2.2 *Ab initio* molecular dynamics

AIMD simulations of one $Ca^{2+}$ in 63 water molecules were conducted with the electronic structure code CP2K/Quickstep, version 2.7 (VandeVondele et al., 2005). CP2K implements

density functional theory (DFT) based on a hybrid Gaussian plane wave. The BLYP, PBE and revPBE generalized gradient approximations for the exchange and correlation terms were used together with the general dispersion-correction DFT-D3 (Becke, 1988; Ernzerhof and Scuseria, 1999; Grimme et al., 2010; Lee et al., 1988; Zhang and Yang, 1998). Goedecker-Teter-Hutter pseudopotentials (Goedecker et al., 1996) were used to describe the core–valence interactions. All atomic species were represented using a double-zeta valence polarized basis set. The plane wave kinetic energy cut off was set to 1000 Ry. k-sampling was restricted to the Γ point of the Brillouin zone. Simulations were carried out with a wave function optimization tolerance of $10^{-6}$ Hartree that allows for 1.0 fs time steps with reasonable energy conservation. Periodic boundary conditions were applied throughout. Starting from the last configuration of 200 ps of classical MD simulation of 64 water molecules, one $H_2O$ molecule was replaced by one $Ca^{2+}$ and we conducted 1 ns of classical MD simulations in the NPT ensemble (T = 300 K and P = 1 atm). The last configuration was taken as the starting point for AIMD simulations, which were carried out in the NVT ensemble using a Nosé-Hoover chain thermostat to maintain the average temperature at T = 300K. The total simulation times for the different functionals were 500 ps for BLYP, PBE and 40 ps revPBE. For comparison, AIMD simulations of pure bulk water were performed following the same approach.



### 3.2.3 Calculations of the water exchange frequencies

To investigate the water dynamics around dissolved calcium ions, the mean residence time (MRT) of water molecules in the first hydration shell of the cation was determined, using both the “direct method” (Hofer et al., 2004) and the Impey method (Impey et al., 1983). An in-house python code was developed to calculate the MRT using the “direct method” from the classical and *ab initio* MD trajectories (see SI). The software tracks all the water molecules that leave or enter the first shell and, subsequently, stay outside or inside the first shell for a time longer than $\tau_{DM}$ = 0.5 ps (Hofer et al., 2004). In the “direct method” (DM), $\tau_{DM}$ = 0.5 ps is a

generally accepted value, based on the average life time of a hydrogen bond between the solvent molecules (Di Tommaso et al., 2014; Hofer et al., 2004; Wolthers et al., 2012a). The first hydration shell is taken to be within the distance from Ca ion to the position of the first minimum in the $Ca^{2+}$-$O_W$ radial distribution function (RDF) (Hofer et al., 2004) and is 3.5 Å for all force fields used here (see also (Di Tommaso and De Leeuw, 2010), for the specific case of calcium). The inverse of the average exchange frequency is taken to calculate the MRT:



$$MRT = \frac{t_{sim} CN_{av}}{N_{ex}^{\tau_{DM}}} \quad (19)$$

where $t_{sim}$ is the total simulation time, $CN_{av}$ is the average coordination number for calcium (i.e. the number of water molecules around calcium obtained from integrating the RDF of Calcium-Oxygen); $N_{ex}^{\tau_{DM}}$ is the number of exchanges counted that lasted longer than 0.5 ps.

The Survival Function (SF) method (or Impey method) determines the survival function for water molecules in the first shell. The first shell is defined in the same way as in the direct method, within 3.5 Å from the calcium ion. The SF method measures the residence time (RT) for each water molecule that enters the first shell. RT is recorded only when the leaving time, $\tau_{SF}$, of a water molecule from the first shell meets the minimum time limit. The $\tau_{SF}$ is advised to be set at 2 ps to guarantee that the water molecule truly moved to the bulk (Impey et al., 1983; Kerisit and Parker, 2004), although a $\tau_{SF}$ of 1 ps has sometimes been used as well (De La Pierre et al., 2016). In this study we adopt the initial proposal of the method and used $\tau_{SF}$ = 2 ps. After extracting all the residence times for all water molecules around calcium over the entire simulation period, the frequency (occurrences) of each RT is normalized by the total number of RT counted. When plotting the normalized frequencies against time, the survival function is constructed. This function gives the probability that a water molecule stays in the first shell as a function of time. A best fit exponential decay can then be plotted and the MRT can be calculated by taking the integral of the function:

$$MRT = \int_0^\infty e^{\frac{-rt}{MRT}} dt \qquad (20)$$

## 3.3 Results and Discussion

Four different force fields were assessed with respect to structural parameters, diffusion coefficients and vibrational properties (see SI for all details). Here, we focus on the results regarding water dynamics, as reflected by calculated exchange frequencies and mean residence times (MRT) for water molecules in the first shell around the calcium cation, and the implications for crystal growth rate.



### 3.3.1 Water dynamics around calcium cations

The MRT values obtained using the "direct method" (Hofer et al., 2004) and the "survival function" (Impey et al., 1983) are compared with other theoretical and experimental MRT evaluations in Table 3.1. There are large differences in MRT values obtained with different water force fields while using the same methodology to determine the residence time. For example, $MRT_{SF}$ from SM-lj (56 ps) is less than half the one obtained from MD simulations using the same $Ca^{2+}$-$O_W$ interaction potential but the SPC/fw-buck water model (125 ps). The longest $MRT_{SF}$ is obtained based on the SPC/fw-tail (333 ps). The $MRT_{SF}$ value obtained from SM-lj output is the only value to agree with the experimental IQUENS results. The $MRT_{DM}$ follow a similar trend, where the value obtained from SPC/fw-buck output reproduces the *ab initio* results.

It was mentioned in earlier work that the higher number of water molecules in the first shell of calcium in SM-lj causes more exchanges and therefore shorter MRT, suggesting that the higher water density simulated with this force field is responsible for the higher coordination number (CN) of $Ca^{2+}$ (De La Pierre et al., 2016). We can add that the structure of the water around the cation in SM-lj is also more distorted with a CN = 8.4, whereas according to *ab initio* MD, the

coordination geometry is octahedral (CN = 6 (Di Tommaso and De Leeuw, 2010)); the addition of extra molecules probably destabilizes this ideal geometry around $Ca^{2+}$, contributing to the shorter $MRT_{SF}$ for SM-lj.

When comparing the two methodologies used to extract MRT from MD simulations, we note that the values obtained using the DM are consistently smaller than those produced by the SF, irrespective of the force field used in the MD simulations (Table 3.1). Since the definition of the size of the first shell is the same for both DM and SF (3.5 Å, see section 3.2.3), the observed discrepancy can be explained by differences between the two methodologies to calculate (mean) residence times, as well as the different "time constraint" considered ($\tau_{DM}$ versus $\tau_{SF}$).

Table 3.1 Mean Residence Times calculated using either the Direct Method ($MRT_{DM}$) or the Survival Function ($MRT_{SF}$); $N_{H2O}/Ca^{2+}$ is the number of water molecules per calcium in the simulation cell; the coordination number (CN) was used to normalize the exchange frequency for calculation of $MRT_{DM}$ and the correlation coefficient $R^2$ is for $MRT_{SF}$. [*] Obtained using the exchange frequency of a molecule[4,5] using equation 1

| Method | $N_{H2O}/Ca^{2+}$ | $MRT_{DM}$ (ps) | CN | $MRT_{SF}$ (ps) | $R^2$ | Ref. |
|---|---|---|---|---|---|---|
| SM-lj | 809 | 14.9 | 8.4 | 56 | 0.997 | This work |
| SPC/fw-tail | 809 | 116 | 7.0 | 333 | 0.999 | This work |
| SPC/fw-buck | 809 | 46 | 6.9 | 125 | 0.999 | This work |
| BLYP-D3 / CP2K | 63 | 73.8 | 6.8 | - | - | This work |
| revPBE-D3 / CP2K | 63 | 60.2 | 6.1 | - | - | This work |
| CPMD | 53 | 23.2 | 6 | - | - | (Di Tommaso and De Leeuw, 2010) |
| QM/MM-MD | 199 | 42.5 | 7.8 | - | - | (Di Tommaso et al., 2014) |
| SPC/fw-tail | - | - | - | 200 | - | (De La Pierre et al., 2016) |
| SPC/E/TIP4P | - | - | - | 158 | - | (Hofmann et al., 2012) |
| SM-buck | 2027 | 7.0 | - | - | - | (Wolthers et al., 2013) |
| Adsorption exp. | - | 6310 | - | - | - | [*],(Eigen and Wilkens, 1965; Petrucci, 1971) |
| IQUENS exp. | - | <100 | - | <100 | - | (Hewish et al., 1983) |

In the direct method, an exchange event is counted when a water molecule enters or leaves the first shell around the cation, and this water molecule remains in/outside the first shell for at least 0.5 ps, a value for $\tau_{DM}$ that is commonly used (Hofer et al., 2004) (see also section 3.2.3). The $MRT_{DM}$ is the mean of all residence times (RT) observed, assuming the RT distribution is a simple Gaussian distribution. However, the SF method shows that this distribution follows an exponential decay function (Figure 3.1a), suggesting the Gaussian distribution is not a valid assumption. As a result, the $MRT_{DM}$ preferentially reflects fast exchanging molecules, resulting in shorter values of $MRT_{DM}$.

In contrast, in the SF method, the $MRT_{SF}$ takes into account the actual distribution of normalized RT frequencies. Such a distribution is shown in Figure 3.1a, based on MD simulations using SPC/fw-tail (SF based on results from SM-lj are shown in Figure S3.4). The $MRT_{SF}$ was calculated by taking the integral of the survival function (i.e. the integral of the fit to these graphs). In our survival functions for all RT probability distributions, a combination of two exponential functions gives the best fit to explain the data (e.g. Figure S3.6). This fit indicates that two underlying distributions are responsible for the observed $MRT_{SF}$, one distribution corresponding to short $MRT_{SF}$ (<20 ps) and a second group of long $MRT_{SF}$ (>20 ps). A similar bi-exponential distribution of RT's for the calcium solvation shell was previously reported by De la Pierre et al. (De La Pierre et al., 2016), who related the short RT's to water librations and excluded those in their $MRT_{SF}$ calculations. Additionally, this bi-modal $MRT_{SF}$ distribution may be due to differences in the average $Ca^{2+}$-$O_W$ distance. When the average distance is plotted against the RT (Figure 3.1b) one can see that, while generally all long RT are found for water molecules at a distance of ≤2.40 Å, short RT are found for water molecules in the full range of distances, from 2.50Å to shorter than 2.40 Å. This suggests that another property affects RT besides the $Ca^{2+}$-$O_W$ distance, for example the H-bonding network. To investigate if this is the case, we determined the percentage of H-bonds between first-shell water molecules and those outside the first shell (referred to as bulk water molecules) for short versus long RT water molecules (Figure 3.2). In that following the definition of a hydrogen bond proposed by Chandra (Chandra, 2000). We used an in-house python script to count the number of hydrogen bonds in the trajectory file. The molecules having a short RT show a larger number of H-bonds with bulk water molecules than those with long RT, suggesting that the H-bond network may affect residence time as well. Furthermore, SPC/fw-tail shows a larger number of H-bonds with bulk water than SM-lj (Figure 3.2 and SI; Figure S3.7a). This more

pronounced H-bond network around the cation leads to a less disturbed water structure and could therefore explain the longer $MRT_{SF}$ obtained from SPC/fw simulations.

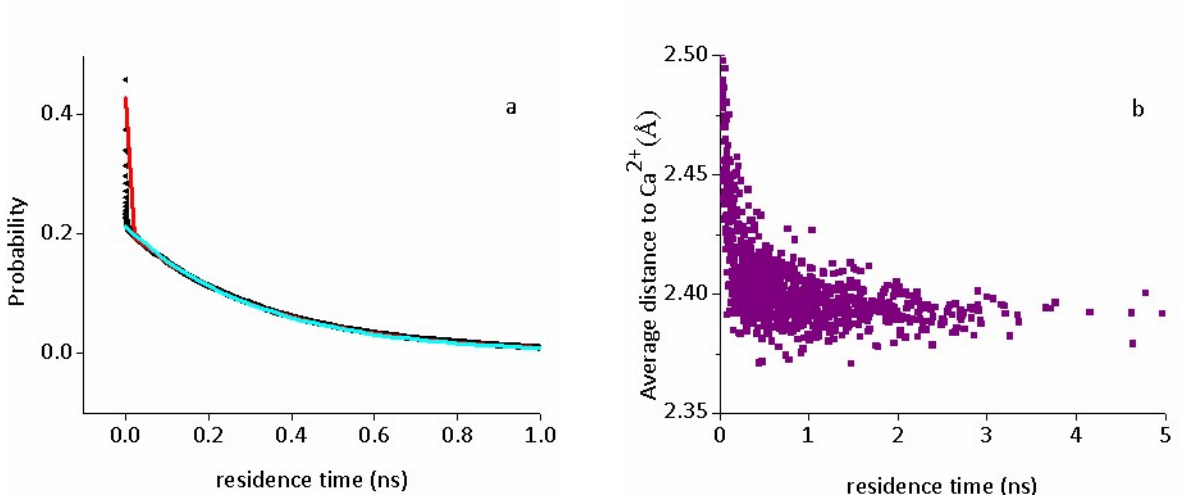



Figure 3.1 a) The probability distribution of different residence times (RT, black triangles) for first-shell water molecules obtained using SPC/fw-tail, where the Survival Function is represented by the red and blue lines, which are two exponential fits. b) Average distance between a coordinated water molecule and the calcium cation as a function of the residence time of that water molecule.

If we group the 'distance-to-cation' data based on their RT, the average distance of the short RTs is 2.97 Å, while the average distance of the long RTs is only 2.48Å (averaged over the first 20 ps after entering the first shell, or in case of short RT, water molecules that left the first shell within these 20ps, as long as they remained).

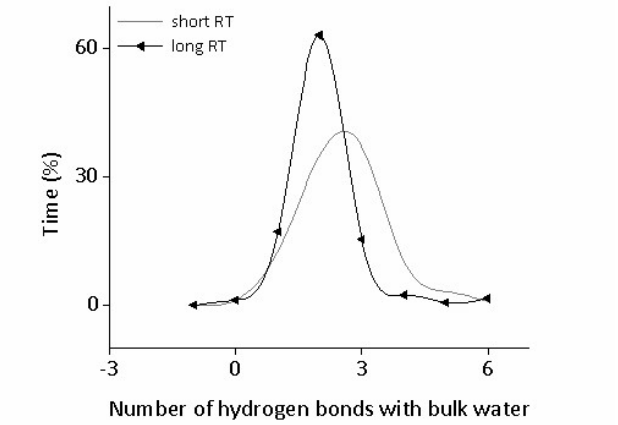



Figure 3.2 The relative time a water molecule has H-Bonds with bulk water molecules, during the 20 ps of each residence time (RT); the two gaussians represent the two different populations (short and long RT $H_2O$).

Besides H-bond network and $Ca^{2+}$-$O_W$ distance, it is worth considering the structure of the first coordination shell. The average coordination number of SPC/fw water is found to be 7.3 (Table 3.1). For 96% of the configurations, the short RT is only observed when the CN is +1 above this average (Table 3.2). As mentioned above, CN is larger for SM-lj (Table 3.1) and the coordination distances are on average also larger (Figure 3.3), leading to greater distortion of the ideal bipyramid (octahedral) coordination for calcium. Moreover, the width of the first coordination shell is also controlled by the van der Waals interaction. When comparing the different interactions using the same water model (SPC/fw-buck and SPC/fw-tail, Figure 3.3). Using Buckingham potential and the related parameters, lead to a wider first shell and the first maximum in the RDF is at a slightly larger distance from the cation compared to the tailed Lennard-Jones interaction proposed by Raiteri et al. (Raiteri et al., 2015). Based on the results for both force fields, it may be suggested that longer $Ca^{2+}$-$O_W$ coordination distances reflect first-shell water molecules in excess of the ideal for $Ca^{2+}$ (CN=6) that will be more liable to exchange than in the rigid model. Therefore, force fields yielding higher CN for calcium expectedly give rise to shorter average MRT.

Table 3.2 The coordination number of calcium in relation to residence time of the water molecules in the first shell using SPC/fw-tail.

| | **First shell configuration** | | |
|---|---|---|---|
| **CN** | **only long RT-water** | **1>= short RT-water** | **all** |
| 6 | 1% | - | 1% |
| 7 | 75% | 1% | 70% |
| 8 | 25% | 96% | 29% |
| 9 | 0% | 3% | 0.16% |

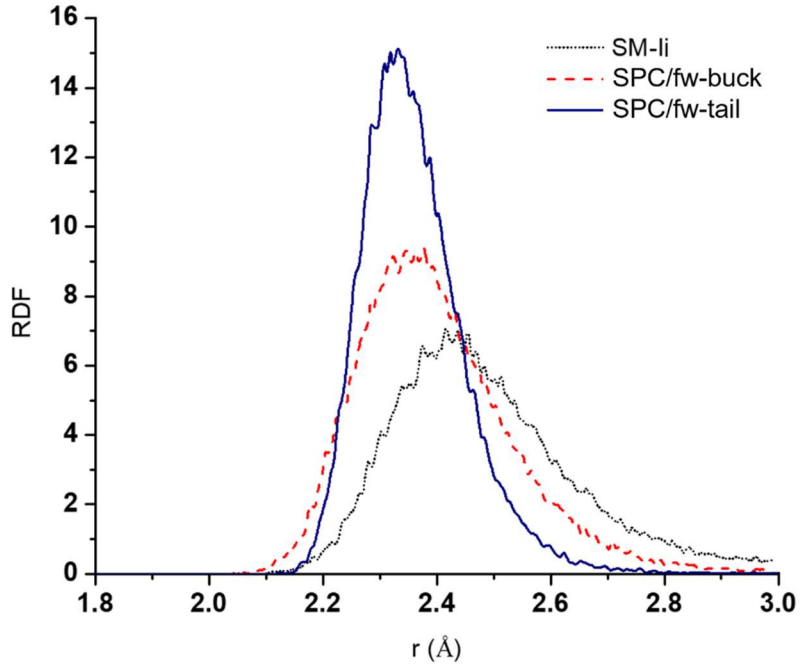





Figure 3.3 Radial Distribution Function reflecting $Ca^{2+}$-$O_W$ coordination distances for the different models.

Both the Direct and Survival Function Methods have advantages and disadvantages. The Direct Method results in shorter mean residence times, because the fast exchanges are over-represented. Nevertheless, the relative differences in $MRT_{DM}$ we observe between the force fields tested (Table 3.1) are very similar to those obtained with $MRT_{SF}$. Moreover, the DM needs shorter simulation times than SF to obtain statistically meaningful output and it can therefore also be used to investigate shorter trajectories generated by *ab initio* and QM/MM MD methods. However, due to the definition of a water exchange event in the DM method, it is likely to be less accurate. Nonetheless, this method can provide insight in the relative variations in local dynamics, such as a comparative assessment of the reactivity of (i.e. water dynamics around) different ions in aqueous solution, or of change of the reactivity of an ion in aqueous solutions with changing composition (Di Tommaso et al., 2014).

In contrasting, the Survival Function needs simulation times long enough to capture enough long residence times for the probability distribution function. Depending on the system

investigated, this may be hundreds of nanoseconds, even microseconds. However, the SF results facilitate detailed differentiation of the MRT distribution in relation to CN, geometry, and H-bond network. But most importantly, the SF approach closely resembles the data post-processing, in terms of fitting exponential decay functions to determine MRT, used in experiments of Eigen and reviewed by Petrucci (Eigen and Tamm, 1962; Eigen and Wilkens, 1965; Petrucci, 1971) It is also similar to the theory behind the interpretation of IQUENS signals (Hewish et al., 1983) to obtain MRT from neutron scattering data. As such, the SF results are in principle more comparable to available experimental data for 2 to 3 molal calcium dissolved in water.



When we compare the MRT's extracted from our MD results with those reported from experiments (Table 3.1), neither methodology (irrespective of the force fields adopted) is able to reproduce the estimation of Eigen (Eigen and Wilkens, 1965; Petrucci, 1971) The MRT values obtained here are closer to IQUENS measurements (Hewish et al., 1983), *ab initio* (Di Tommaso and De Leeuw, 2010) and QM/MM MD simulations (Hofer et al., 2004). The different computer simulation methods still generate a range of water exchange frequencies, and comparability is hampered because: (i) the SF cannot be used on output generated by computationally demanding AIMD and QM/MM MD methods, which are generally limited to few hundreds of ps, and (ii) partly because of the intrinsic limitations of the classical MD force fields, such as the different ways of expressing the structure of the water with the different interatomic potentials.

The simulations conducted in this study considered a calcium ion in pure liquid water. However, Di Tommaso et al. (Di Tommaso et al., 2014) have shown that the frequency of water exchange in the first hydration shell of $Ca^{2+}$ is reduced in the presence of halide ions; this effect is observed not only when the halide anions are directly coordinated to calcium, but also when the alkali and halide ions are placed at or outside the second coordination shell of

$Ca^{2+}$. As natural aqueous solutions are far from pure water, but rich in ions, in such environments the water-exchange dynamics around $Ca^{2+}$ are likely to be influenced by the formation of contact and solvent shared ion pairs. An extended model, also considering the effect of solution speciation on the residence times of water around $Ca^{2+}$, will therefore be part of future studies.

To summarize, residence times are affected by Ca-$O_W$ distance and H-bonding network and therefore depend on which force field is used. The Direct Method can be used on MD as well as AIMD and QM/MM-MD output to obtain relative differences in dynamics. Based on the current comparison, we recommend the SPC/fw-buck force field to determine $MRT_{DM}$, since it most closely reproduces the *ab initio* results in $MRT_{DM}$. However, the $MRT_{SF}$ based on this force field is too long compared to experimental IQUENS data, where the SM-lj provides the best agreement. Furthermore, if at all feasible, we suggest that the use of the SF method to extract the $MRT_{SF}$ from MD simulation output obtains the most accurate (absolute) value of the mean residence time.



### 3.3.2 Comparison with water exchange kinetics around carbonate ions

Our atomistic simulations agree with recent experiments, supporting exchange frequency values for calcium that are two orders of magnitude larger than previously thought (Eigen). Consequently, the water exchange kinetics around $Ca^{2+}$ approach expected values around anions, e.g. the carbonate ion ($>10^{10}$ $s^{-1}$) (Wolthers et al., 2013). To test this hypothesis, we ran a 13 ns simulation using SPC/FW and carbonate (Details in SI), to determine the MRT around a carbonate ion. We obtained an $MRT_{SF}$ of 8.4 ps ($R^2$ 0.999) and $MRT_{DM}$ of 9.2, which are both a factor of 5-40 shorter than for water residence times around calcium obtained with the same force field. As such, the difference in water exchange kinetics around the carbonate ion is not orders of magnitude faster than around the calcium ion, although it remains noticeably

faster. The slightly longer MRT around calcium is expected, because the anion interactions are through hydrogen-bonds with a relatively shorter lifetime (~2.5 ps) (Kumar et al., 2009), than the metal-oxygen interactions as estimated by *ab initio* MD (Di Tommaso and De Leeuw, 2010). Furthermore, the free energy of hydration for calcium is a larger than for carbonate (-1447±2 and -1301±2 $kJmol^{-1}$ respectively) (Raiteri et al., 2010), indicating a stronger interaction between calcium and water than between carbonate and water.



### 3.3.3 Implications for crystal growth rate

In order to discuss the impact of a smaller difference in water dynamics around calcium and carbonate ions than has been suggested thus far, we relate to crystal growth via ion-by-ion attachment. While there is strong evidence of ion-pair attachment during growth, for example by the attachment of $CaCO_3$(aq) pairs during calcite growth (Andersson et al., 2016; De La Pierre et al., 2017; Genovese et al., 2016; Rodriguez-Navarro et al., 2016), such a growth mechanism cannot explain the dependency of growth rate on the solution stoichiometry at a constant degree of supersaturation observed in bulk (Nehrke, 2007; Tai et al., 2005) and microscopically. (Bracco et al., 2012; Hong and Teng, 2014; Sand et al., 2016) According to thermodynamic equilibrium calculations, the concentration of the $CaCO_3$(aq) ion pair is constant at a constant degree of supersaturation, irrespective of the solution stoichiometry, and the growth step velocity and bulk growth rate dependency on solution stoichiometry is therefore assumed to be caused by an ion-by-ion growth mechanism (Larsen et al., 2010; Nielsen et al., 2012; Stack and Grantham, 2010; Wolthers et al., 2012b).

In the ion-by-ion growth mechanism, the ion attachment frequencies are controlled by several, potentially simultaneous processes: cation and anion diffusion to lattice sites (surface anions and cations, respectively), dehydration of solvated ions and surface lattice sites. Various experiments investigating calcite {10-14} surface growth strongly suggest that these processes

are different for the structurally distinct growth steps at the calcite surface (Bracco et al., 2012; Hong and Teng, 2014; Larsen et al., 2010; Sand et al., 2016; Stack and Grantham, 2010). While obtuse step edges propagate through direct attachment of ions delivered via solution diffusion, the acute steps grow by attachment of ions delivered through surface diffusion. Moreover, Hong and Teng (Hong and Teng, 2014) have proposed, based on their Atomic Force Microscopy (AFM) observations, that obtuse step growth is limited by calcium dehydration, and acute step propagation is limited by surface adsorption, reorientation and rearrangement of carbonate ions. We would like also to indicate that interpretation of the kinetics of crystal growth, deconvolved from experimental data, using molecular simulations computer models is very challenging, because the mechanisms controlling these processes (nucleation and growth) can be significantly more complex than those considered in simplified computer models.



One of the first direct relationships between crystal growth kinetics and ion dehydration was presented by Nielsen [8]. He obtained the values for the cation dehydration frequency, $\nu_{dh}$, related to various minerals from Petrucci et al. (Petrucci, 1971) and obtained the correlation with ion integration frequency, $\nu_{in}$, as illustrated with blue dots and line in Figure 3.4. Expressing this correlation as an Eyring equation, it was found that the integration rate is approximately a thousand times smaller than the dehydration frequency:

$$\nu_{in} = \nu_{dh} e^{\frac{-\Delta G_D^{\neq}}{kT}} = \frac{\nu_{dh}\nu_D}{\nu_0} \qquad (21)$$

where $\nu_D$ is the diffusion frequency, $\Delta G_D^{\neq}$ the Gibbs free energy of diffusion and the ratio between diffusion rate and $\nu_0 = \frac{kT}{h}$, with k and h being the Boltzmann and Planck's constant, respectively, $\frac{\nu_D}{\nu_0}$ =$10^{-3}$ (Nielsen, 1984).

This correlation (Figure 3.4) supported the notion that the cation dehydration is the limiting factor for growth (and dissolution). The cations retain water molecules longer than anions, and are therefore rate limiting (De La Pierre et al., 2016; Di Tommaso et al., 2014; Wolthers et al., 2013).

In general, the MRT's obtained with DM and SF support this notion, although the difference between cation and anion dehydration is not as extreme (orders of magnitude) as considered previously. Both cation and anion dehydration frequencies are in the order of, or slightly smaller than, the reported diffusion rate ($\nu_D$) ~$10^{10}$ $s^{-1}$ (Nielsen, 1984). However, adding our water exchange frequencies as well as the other values from Table 3.1 to Figure 3.4, reveals that the correlation previously reported by Nielsen (Nielsen, 1984) is not very strong, in particular in the case of cations with a dehydration frequency higher than $10^8$ $s^{-1}$, (Figure 3.4).

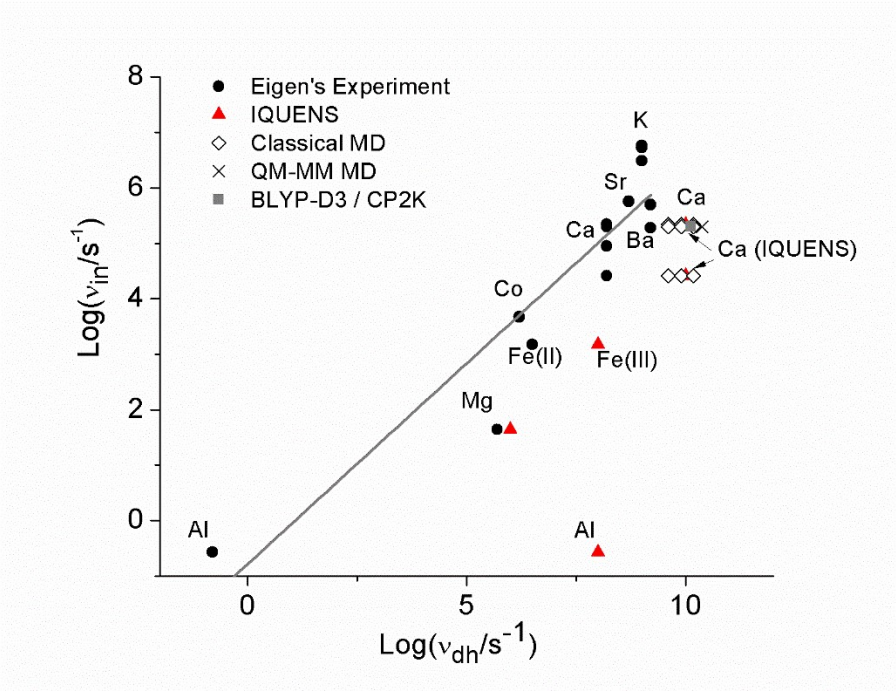



Figure 3.4 The logarithm of the cation integration frequency versus the logarithm of the cation dehydration frequency. Circles: as determined from parabolic function fitted to the dehydration frequency (Nielsen, 1984); triangles based on IQUENS measurements (Helm and Merbach, 1999; Hewish et al., 1983); cross based on QM/MM-MD simulations and DM (Di Tommaso et al., 2014); gray square based on *ab initio* calculations (BLYP-D3 functional) and DM (this study); diamond from classical MD using SF (this study).

Furthermore, if we use the updated values of dehydration frequencies, measured with the IQUENS method, and the values of the ion integration frequency, $\nu_{in}$, then the term $\frac{\nu_D}{\nu_0}$ (equation 3) is not $10^{-3}$ for calcium anymore, but a factor $10^{-5}$ is needed to correlate $\nu_{dh}$ with $\nu_{in}$. This indicates that another process is affecting $\nu_{in}$. Furthermore, the dehydration frequency derived from *ab initio* MD (Table 3.1) is smaller than the diffusion frequency ($\nu_D$), which suggests that the diffusion of the cation into the lattice site from a neighbouring position is the rate limiting step, in agreement with the work of various experimental works (Bracco et al., 2012; Hong and Teng, 2014; Stack et al., 2012). Additionally, Petrucci (Petrucci, 1971) proposed, based among others on data from Eigen et al. (Eigen and Wilkens, 1965), that for some ions, including $Ca^{2+}$, dehydration was not the limiting factor. Petrucci concluded that substitution of water by any ligand is the rate limiting step only if the ligand's exchange rate is less than $10^7$ $s^{-1}$. The latter is undoubtedly the case for most cations with small radii (<100 pm) such as $Mg^{2+}$, $Ni^{2+}$, $Co^{3+}$, $Zn^{2+}$ and $Fe^{2+}$. Petrucci (Petrucci, 1971) compared the substitution rates of different ligands; for $Mg^{2+}$, which has a very low water exchange frequency compared to $Ca^{2+}$, there is indeed no difference in the reaction rate with a range of ligands, indicating that the dehydration of magnesium is the rate controlling step. In contrast, he observed different reaction rates between $Ca^{2+}$ and the same selection of ligands, suggesting that calcium dehydration is not the rate determining step. Alternatively, it may also suggest that calcium dehydration can be affected differently by different ligands, as has been shown for halide ions (Di Tommaso et al., 2014).

To summarize, the water exchange kinetics of $Ca^{2+}$ and $CO_3^{2-}$ are more comparable than previously thought, which is substantiated by maximum bulk growth rates and step propagation rates in near-stoichiometric solutions (Nehrke, 2007; Wolthers et al., 2012b). Furthermore, the MRT's reported here, combined with the experimental IQUENS values for the dehydration frequency of calcium (Hewish et al., 1983), and the ligand experiments (Petrucci, 1971) shed

a different light on the most likely rate limiting step for calcium carbonate crystal growth. Rather than cation dehydration, (surface) diffusion of the cation into the lattice site from a neighbouring position is the rate limiting step, which finding is in agreement with the work from Hong and Teng (Hong and Teng, 2014), Bracco et al. (Bracco et al., 2012), Stack et al. (Stack et al., 2012). Alternatively, our results may support the proposition of De La Pierre et al. (De La Pierre et al., 2017) that the attachment of the carbonate ion to the kink site, is the rate determining step of the calcium carbonate crystal growth.



## 3.4 Conclusions

The mean residence time of water molecules around dissolved calcium ions, obtained from our classical MD simulations, does not agree with the experimental mean residence time values around calcium estimated by Eigen (Eigen and Wilkens, 1965), as used in Nielsen's analysis of crystal growth (Nielsen, 1984) and adopted by many recent calcite growth rate studies. In contrast, our MRT values are in the same order of magnitude as more accurate *ab initio* MD, whereas QM/MM-MD agree with the limit obtained from IQUENS experiments. This finding indicates that calcium falls in the group of cations with fast water exchanges and its dehydration is therefore not rate limiting during calcite crystal growth. This study also revealed that of the two methods we tested to calculate the MRT from MD simulation output, only the survival function method provides accurate absolute values and can distinguish two different groups of water molecules around calcium, revealing that ion-water distance and H-bonding networks affect MRT values. Additionally, our work suggests that the direct method can be used to obtain relative differences in MRT, rather than accurate values, and is most suitable to post-process output from MD simulations of larger systems and *ab initio* calculations, as it can be implemented on shorter production runs.

## 3.5 Supporting Information.

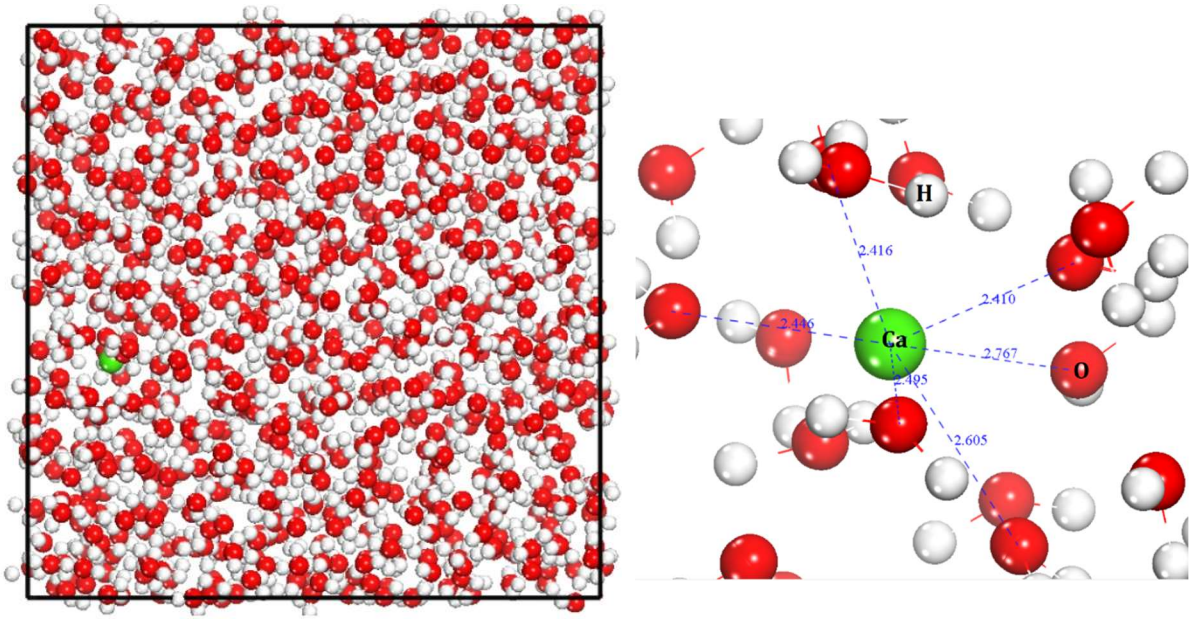
This supplementary information contains tables with the parameters of the Force Fields used in this work, a study of the different structural parameters including the dipole moment, and the three- and four-body order parameters, the dynamics of the system, such as diffusion coefficients and vibrational spectra for the four water models, a python code to calculate the MRT using the "direct method" from the trajectory files and the probability distribution figures for the Ca-SM-lj system.



### 3.5.1 Computational methods

#### 3.5.1.1 Structural models.

To study the liquid water, we have created a simulation cubic box consisting of 809 water molecules. We inserted a calcium cation in the water box. The cation quickly binds water (Figure S3.1) reaching the coordination number and distances established by each force field employed.




(a) (b)

Figure S3.1 (a) One frame of the simulation box including the calcium ion. (b) Zoom around $Ca^{2+}$ showing coordination distances in de SM-lj water model. Color scheme, $Ca^{2+}$ green, Oxygen red, carbon gray.

### 3.5.1.2 Classical molecular dynamics

The MD simulations were performed with of DL_POLY 4.09 (Todorov et al., 2006) modified to include a tail required for a forcefield. The verlet leap-frog scheme and the Nosé-Hoover algorithm (Hoover, 1985; Nosé, 1984) were used to integrate the equations of motion with a time step of 0.1 fs and to maintain the temperature at 300 K with 0.5 ps relaxation time. The simulations were carried out in the NPT ensemble. The total simulation time for all liquid water systems was 160 ns including 200 ps of equilibration. An 8.0 Å cut-off was used for the van der Waals forces.



### 3.5.1.3 Water models

We used the following force fields to describe the water molecule interactions:

(1) A shell-model (SM-lj), developed by de Leeuw and revised by Kerisit and Parker replacing the Buckingham potential with a Lennard-Jones potential (de Leeuw and Parker, 1998; Kerisit and Parker, 2004). The motion of oxygen shells was treated using a small mass (0.2 au) to simulate the fast adaptation of the electronic polarization to the change in ionic positions. SM-lj is the most complete description of a single water molecule, it is a vibrating water molecule and it has a shell on the oxygen to count the polarizability of the molecule, represented by the harmonic term $O_W$ core—$O_W$ shell, according to the shell model of Dick and Overhauser in the description of the force field. The higher density of the SM-lj model is partially corrected by Kerisit and Parker (Kerisit and Parker, 2004) in which they modified the hydrogen bond description, replacing a more attractive Buckingham potential by a 9-6 Lennard-Jones. This is the only force field considered in the present study where explicitly describe van der Waals interactions between an oxygen and a hydrogen of a different molecule.

(2) A simple water model with negative Drude polarization (SWM4-NDP). The water model has a negatively charged massless dummy particle at a fixed distance from the oxygen site and

a negative Drude particle attached to the oxygen with a spring constant representing the polarizability of the water molecule (Lamoureux et al., 2006).

(3) The simple point charge flexible water model (SPC/fw) and

(4) the transferable intermolecular potential with 3 points (TIP3P).

The latter two do not use core/shell configurations, and therefore cannot describe the polarizability of water molecules. As a result, though, these simulations scale better in parallelisation and are less compute demanding (Wu et al., 2006).



#### 3.5.1.4 Calcium water interactions

Two sets of $Ca^{2+}$—$O_W$ potential parameters were employed in this study. One of them was developed by de Leeuw and Parker (De Leeuw and Parker, 1997) and extensively used in the description of both calcite and aragonite surface systems (De Leeuw and Parker, 1998; Ruiz-Hernandez et al., 2012; Spagnoli et al., 2006a); Table S3.1. The other interaction was developed by Raiteri et al. (Raiteri et al., 2015), and it is being extensively used to investigate crystalline and amorphous (hydrated) $CaCO_3$ (De La Pierre et al., 2016; Kirch et al., 2018; Smeets et al., 2017).

Table S3.1 Water potential parameters

| **TIP3P** | | |
|---|---|---|
| **Species** | **Charge (e)** | |
| | **core** | **shell** |
| $O_W$ | -0.8340 | — |
| $H_W$ | +0.417 | |
| **Harmonic Potential** | ***k* (ev Å$^{-2}$)** | ***r*$_0$ (Å)** |
| $O_W$—$H_W$ | 47.960699 | 0.957 |
| **Three-Body Potential** | ***k* (ev rad$^{-2}$)** | **Θ$_0$** |
| $H_W$— $O_W$— $H_W$ | 8.672821 | 104.52 |
| **Lennard-Jones Potential** | **ε (eV)** | **σ (Å)** |
| $O_W$— $O_W$ | 0.00659134 | 3.1507524 |

| **SPC/fw** | | |
|---|---|---|
| **Species** | **Charge (e)** | |
| | **core** | **shell** |
| $O_W$ | -0.82 | — |
| $H_W$ | +0.41 | |
| **Harmonic Potential** | ***k* (ev Å$^{-2}$)** | ***r*$_0$ (Å)** |
| $O_W$—$H_W$ | 45.9296231 | 1.012 |
| **Three-Body Potential** | ***k* (ev rad$^{-2}$)** | **Θ$_0$** |
| $H_W$— $O_W$— $H_W$ | 3.29134 | 113.24 |
| **Lennard-Jones Potential** | **ε (eV)** | **σ (Å)** |
| $O_W$— $O_W$ | 0.006739769454 | 3.165492 |

Table S3.1 continued. Water potential parameters

| SM-lj | | | | |
|---|---|---|---|---|
| **Species** | **Charge (e)** | | | |
| | **core** | **shell** | **Core-shell interaction (eV.Å$^{-2}$)** | |
| Ow | +1.250 | -2.050 | 209.449602 | |
| Hw | +0.40 | | | |
| **Lennard-Jones Potential** | **$A_{ij}$ (eV. Ån)** | **$B_{ij}$ (eV. Åm)** | **n** | **m** |
| Ow— Ow | 39344.98 | 42.15 | 12 | 6 |
| Ow— Hw | 24.0 | 6.0 | 9 | 6 |
| **Morse Potential** | **$D_{ij}$ (eV)** | **$\alpha_{ij}$ (Å$^{-1}$)** | **$r_0$ (Å)** | **Coul.** |
| Ow— Hw | 6.203713 | 2.22003 | 0.92367 | 0.50 |
| Hw — Hw | 0.000000 | 2.840499 | 1.5 | 0.50 |
| **Three-Body Potential** | ***k* (ev rad$^{-2}$)** | **$\Theta_0$** | | |
| Hw— Ows — Hw | 4.19978 | 108.693195 | | |

| SWM4-NDP | | | |
|---|---|---|---|
| **Species** | **Charge (e)** | | |
| | **core** | **shell** | **Core-shell interaction (eV.Å$^{-2}$)** |
| OW | +1.71636 | -1.71636 | 43.3275 |
| Hw | +0.55733 | | |
| D | -1.11466 | | |
| **Harmonic Potential** | ***k* (ev Å$^{-2}$)** | ***$r_0$* (Å)** | |
| Ow—Hw | ∞ | 0.957 | |
| **Three-Body Potential** | ***k* (ev rad$^{-2}$)** | **$\Theta_0$** | |
| Hw— Ow — Hw | ∞ | 104.52 | |
| **Lennard-Jones Potential** | **ε (eV)** | **σ (Å)** | |
| Ow— Ow | 0.00914665 | 3.18395 | |

| Calcium—water interactions | | | |
|---|---|---|---|
| **De Leeuw-Parker (-buck)** | | | |
| **Buckingham Potential** | **$A_{ij}$ (eV)** | **$\rho_{ij}$ (Å)** | **$C_{ij}$ (eV. Å$^6$)** |
| Ca—Ow | 1186.6 | 0.297 | 0 |
| **Raiteri et. al. (-tail) *** | | | |
| **Lennard-Jones Potential** | **ε (eV)** | **σ (Å)** | |
| Ca—Ow | 0.000950 | 3.35 | |

*The interaction was set to zero over the range 6 – 9 Å

### 3.5.2 Analysis of simulation

#### 3.5.2.1 Dimer-structure

A dimer is the most energetic favourable configuration of two water molecules forming a hydrogen bond. In literature, evaluating the dimer-structure is a common method to validate a new force field (Burnham and Xantheas, 2002; de Leeuw and Parker, 1998; Mukhopadhyay et al., 2015), as this basic building block influences the reproducibility of the dipole moment of the liquid, which is very important to accurately describe its interaction with charged species like ions or interfaces. The structure of the water dimer is shown in Figure S3.2.

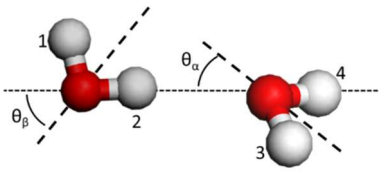



Figure S3.2 Dimer of water. To find the right water-model for our study on the calcite surface one of the approaches is the examination of the dimers published in literature and compare the different models with our own data.

The five requirements that need to be full filled are (Burnham and Xantheas, 2002; de Leeuw and Parker, 1998; Lane, 2013; Mukhopadhyay et al., 2015):

(1) the hydrogen must be in the first shell of neighbouring O

(2) the angle between the planes molecules must be 90°

(3) the angle $O_a$-$H_a$---$O_b$ angle should be quasi linear

(4) Dihedral, the angle between proton 1 and 3 or 4, should be around 120°

(5) the angle α must be below 70°

### 3.5.2.2 Three-body and Four body order parameters

Besides the structure of the dimer there are other more general parameters to evaluate the different water models, such as the angular order parameter or three-body order parameter (F3) (BÁEZ and CLANCY, 1994) and the four-body structural order parameter F4 order parameter (Wu et al., 2016). These two parameters are both widely used in clathrate hydrate but as far as we are aware no data are published for the water models in the scope of this study. The F3 gives an indication about the angular ordering of the water molecules and is defined as follows (BÁEZ and CLANCY, 1994):



$$F3 = \frac{1}{N_{H_2O}} \left\langle \sum_{1}^{N_{H_2O}} \sum_{1}^{CN} [(|\cos\theta| \cos\theta + \cos^2(109.47°))]^2 \right\rangle \qquad (22)$$

where θ is the angle between a water oxygen atom and two water oxygens in its first solvation shell (Figure 3.3a). By summing up all the possible angles around the centre oxygen within the first shell (at a cut-off of 3.5 Å) and averaging that number by all the water molecules present in the simulation cell, one gets the F3 order parameter. F3 is designed to measure deviations from the tetrahedral angle and has a value of zero for an ideal tetrahedral ice-like water structure network and 0.8 for liquid water. Therefore, the F3 parameter gives an indication of how liquid-like the water generated with the different force fields is.

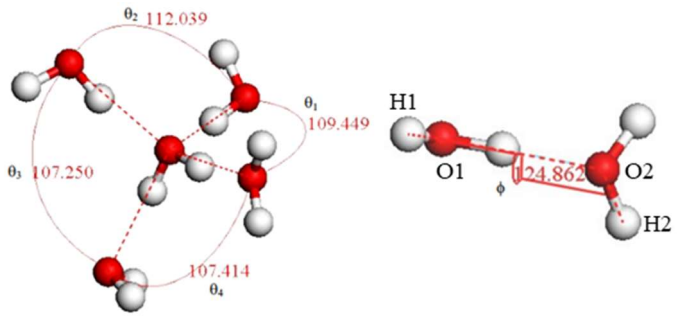



Figure S3.3 a) an illustration of how the angles around the central oxygen are determined for the F3 order parameter, b) the F4 parameter is determined with the dihedral angle as illustrated with the phi

The F4 parameter distinguishes between different tetrahedral networks the water may adopt. The F4 order-parameter is considering the angle of the dihedral, or torsion angle, formed by two adjacent water molecules and measures the planarity of the water pairs according to

$$F4 = \cos 3\phi \tag{23}$$

where ϕ is the dihedral angle H1–O1…..O2–H2 from two neighbouring water molecules (within the cut-off of 3.5 Å), Figure 3.3b with H1 and H2 are the farthest hydrogen atoms on each of the water molecules pairs. The F4 for liquid water is 0 and a negative value indicates the development of ice-like structures (Wu et al., 2016).



### 3.5.2.3 Diffusion coefficient of water

DL_POLY provides a tool to calculate the mean square displacement (MSD) of the different species. The MSD can be defined as:

$$MSD = \langle R^2(t) \rangle = \frac{1}{N} \sum_{i=1}^{N} |R_i(t) - R_i(0)|^2 = 6Dt \tag{24}$$

where $R_i$ is the position vector of atom i, and N the total number of atoms *i*, D is the diffusion coefficient (Wang and Hou, 2011).

### 3.5.2.4 Vibrational spectrum

The vibrational spectrum of the different water models is built from the sum of all the vibrational density of states (VDOS) of each atom involved (Scott et al., 1991). The VDOS can be calculated using the Fourier transformation of the velocity-autocorrelation function (VACF) of the individual atoms in the water molecule. We used the following equation to calculate the VACF:

$$VACF(t) = \frac{1}{N_{tsteps} N_{atm}} \sum_{j=1}^{N_{tsteps}} \sum_{i=1}^{N_{atm}} v_j(t)_j \times v_i(t_j + t) \tag{25}$$

where $N_{tsteps}$ and $N_{atm}$ are the number of timesteps and the number of atoms, oxygen and hydrogen in this case, respectively; $v_i$ is the velocity vector of O or H atoms in the $i^{th}$ water molecule.

### 3.5.3 Comparison of water models

We compared four forcefields to find the best compromise between accuracy of the results—that is, capable of reproducing experimentally/independently determined water-calcite interface characteristics—and computational cost. As the models are extensively used, their full characteristics can be found elsewhere (de Leeuw and Parker, 1998; Jorgensen et al., 1983; Kerisit and Parker, 2004; Lamoureux et al., 2006, 2003; Paesani et al., 2006; Wu et al., 2006). The most striking differences (Table S2) between the two groups of models (flexible core-shell models versus rigid water-molecule models) were found in the density and the diffusivity. The TIP3P model results in water self-diffusion coefficients that are twice the experimental value, caused by the intra-molecular parameters: TIP3P has a higher spring constant in both the two- and three-body potential terms, which leads to a higher internal energy and thus a more mobile molecule. The SM-lj exhibited the higher density previously reported (Kerisit and Parker, 2004). Furthermore, the diffusion coefficient of SM-lj decreased with time after 40 ns, to a value of 1.87 x 10-9 m2s-1, while the volume decreased, leading to the higher density of water.

Both diffusion coefficient and density are essential to describe the dissolution and reactivity of calcite. Particularly, the density of water has been presumed as critical to determine the MRT of a water molecule in a calcium site of a calcite surface (De La Pierre et al., 2016), although this point is discussed in this manuscript. All models, except for SM-lj, exhibit comparable dipole moments and are all lower than experimental values. SM-lj and SWM4-NDP reproduce accurately the experimental dipole in gas phase. Only the SWM4-NDP results in both the gas phase dipole moment as the assumed dipole moment for liquid water. Furthermore, this force

field shows an induced dipole under influence of calcium of 21% which may give a good indication of the polarizability of the electron cloud around water. It is worth noting that the rigid models were fitted to reproduce the estimated experimental value for the dipole of water molecules in liquid water, the SM-lj was fitted to reproduce the experimental for the dipole of water molecules value in gas the phase.



Table S3.2 Properties of water molecules in the various force fields, and water characteristics obtained over 160 ns of production time using the different water models. 1ai ab initio (Badyal et al., 2000) 2ai ab initio MD (Silvestrelli and Parrinello, 1999)

| Model property | SM-lj | SWM4-NDP | SPC/FW | TIP3P | Experimental |
|---|---|---|---|---|---|
| Polarizable | • | • | | | • |
| Dummy particle | | • | | | |
| Flexible bonds | • | | • | • | • |
| Water characteristics | | | | | |
| Density (300K) | 1.27 (Kerisit and Parker, 2004) | 1.00 (Stukan et al., 2013) | 1.01 (Wu et al., 2006) | 1.00 (Jorgensen and Jenson, 1998) | 1.00 (Tanaka et al., 2001) |
| Diffusion coefficient ($10^{-9}$ $m^2s^{-1}$ 300K) | 2.3 (Kerisit and Parker, 2004) | 2.33 (Stukan et al., 2013) | 2.32 (Wu et al., 2006) | 5.30 (Wu et al., 2006) | 2.3 (Krynicki et al., 1978) |
| Dipole moment liquid | 2.00 | 2.43 | 2.40 | 2.44 | 2.95[1ai] |
| Dipole moment vapor | 1.85 | 1.85 | 2.20 | 2.23 | 1.85 (Clough et al., 1973) (1.85[2ai]) |
| Dipole moment around Ca | 2.09 | 2.93 | 2.46 | 2.47 | — |
| Induced dipole | 0.10 | 0.50 | 0.06 | 0.17 | — |
| Relative induced dipole | 5% | 21% | 3% | 0% | — |

The dipole moment is important to describe interactions with the solid surfaces (Carrasco et al., 2012; Meng et al., 2004; Michaelides et al., 2003), and water dimers influence the overall dipole. Therefore, we also investigated the structures of water pairs. The values for the different

dimer parameters are in a good agreement among the different models (Table S3.3). All the classic models failed to reproduce the ab initio and experimental values for α (Figure S3.2), while they reproduce the value for β (Figure S3.2). Only the average number of dimers per timeframe differs between the runs with the various water models. The SM-lj water shows less dimers compared to the water simulated with the other force fields. This can be related to the higher density of the SM-lj water, which makes it more difficult to fulfil the geometric characteristics of the dimer structure.



Table S3.3 Dimer properties. Distance O-O, values of α and β for the dimers modelled by the different water models. The number of dimers was averaged over our simulation time.

| | **Distance** | **α** | **β** | **Number of dimers** |
|---|---|---|---|---|
| **SM-lj** | 2.90 | 37.11 | 49.87 | 5.12 |
| **SWM4-NDP** | 2.8 | 38.72 | 50.33 | 10.72 |
| **SPC/fw** | 2.8 | 39.0 | 54.8 | 11.97 |
| **TIP3P** | 2.78 | 37.2 | 49.9 | 8.73 |
| **Ab initio** (Kerisit and Parker, 2004) | 2.82 | 53.9 | 56.6 | |
| **Experimental** (Odutola and Dyke, 1980) | 2.98 | 57 | 51 | |

Nevertheless, it cannot be concluded that the network distortions of the SM-lj due to the higher density will be the sole responsible of the disruption the structure of the interface as the F3 and F4 parameters (Table S3.4) indicate that all models express a clear liquid behaviour of the water.

Table S3.4 F3 and F4 parameters for the different water models

| | F4 | F3 ($OW^{-1}$) | No tetrahedral (OW%) * |
|---|---|---|---|
| **SM-lj** | 0.0 | 0.64 | 7.6 |
| **SWM4NDP** | 0.0 | 0.51 | 3.3 |
| **TIP3P** | 0.0 | 0.46 | 1.4 |
| **SPC/fw** | 0.0 | 0.50 | 1.7 |
| **SM-lj (Ca)†** | 0.0 | 0.64 | 0.2 |
| **SWM4NDP (Ca)†** | 0.0 | 0.51 | 4.5 |
| **TIP3P (Ca)†** | 0.0 | 0.59 | 4.6 |
| **SPC/fw-buck (Ca)†** | 0.0 | 0.51 | 3.8 |
| **SPC/fw-tail (Ca)‡** | 0.0 | 0.52 | 3.9 |

†A single $Ca^{2+}$ ion was included in the water box. The $Ca^{2+}$— $O_W$ interaction is from de Leeuw and Parker
‡ A single $Ca^{2+}$ ion was included in the water box. The $Ca^{2+}$— $O_W$ interaction is from Raiteri et al.
* This value is averaged over the simulation time.

We have performed an ANOVA test, to compare the data for F3 and F4 of the different models at once. Then a two-tailed t-test with Bonferroni-correction was ran to avoid false rejection of the hypothesis, stating all models have the same percentage of a non-tetrahedral structure (Miller and Miller, 2005). The ANOVA test confirms that the data of the four models is statistically different from each other, with a t-test p-value much lower 0.0125. Also, the percentages of tetrahedrally coordinated water molecules (Table S3.4) are significantly different between the various force fields. The F3 obtained with the SM-lj is closest to the value reported for liquid water (0.8 (BÁEZ and CLANCY, 1994)). Together with the F4 being equal to zero, this result indicates this water model is strictly speaking not “overstructured” as it was reported previously (Van Maaren and Van Der Spoel, 2001).

When one calcium ion is present in the box of water, the ANOVA and t-test show the same result as in absence of calcium. The impact of the $Ca^{2+}$ to the tetrahedral structure of the water is only noticeable in the rigid models by a slightly more disrupted tetrahedral structure. This is in line with the increased percentage of water molecules not involved in any tetrahedron (Table S3.4). It is also worth noting that the results were consistent when the interaction Ca- $O_W$ was changed. This suggests the influence of calcium is independent of the choice of the van de Waals interaction parameters, and the alteration of the tetrahedral structure of liquid water by the cation is a consequence of the induced dipoles, due to cation presence (Table S3.2). Moreover, for SM-lj, the percentage of water molecules involved in forming tetrahedra increases strongly, and the dipole of the molecules in the first hydration shell of calcium approach the value of liquid water.



### 3.5.4 VACF and VDOS

After focusing on the water structure, insights into the dynamics of water were obtained from the analysis of the velocity-autocorrelation function (VACF) of the water oxygen atoms. The VACF reveals the combination of movement and vibrations for every group of molecules. A water molecule vibrates faster than it travels in space, so for every real travel step the molecule also moved back and forth around its position. The bending and librational modes, where the $O_W$—$H_W$ interaction "pulls and pushes" the molecule around the liquid network, are also faster than the translation.

While three of the tested force fields display similar VACF profiles, the SWM4-NDP model displays longer wavelengths in the profile and a distinct a minimum (Figure S3.4), representative of a so-called "cage effect" (slow displacements of the molecules reflected in longer wavelengths). This results in a peak in the Vibrational Density of States (VDOS, Figure S3.4b) in a low wavenumber range in the spectrum. It is less noticeable in VACF profiles

generated using the other models, as they show shorter wavelengths generated by bond vibrations, that oscillate around the long wavelength caused by the same "cage effect". This long wavelength can be affected by the water structure in two ways. 1) the amplitude can vary. When it is higher, more water molecules feel the same cage effect and move according to the same frequency. This is an indication of a more rigid H-bond network. 2) the wavelength varies, causing a shift in the vibration spectrum. According to previous studies, confinement causes a little shift to a higher wavenumber (Di Tommaso et al., 2017). The SM-lj shows a smaller amplitude and a higher frequency, indicating that in this case the caging effect is less, in agreement with the higher ordering parameter F3, while the higher density of the water results in smaller cage volumes. Therefore, less water molecules bounce around with a higher frequency. This effect is similar to the confinement effect seen by Di Tommaso (Di Tommaso et al., 2017).

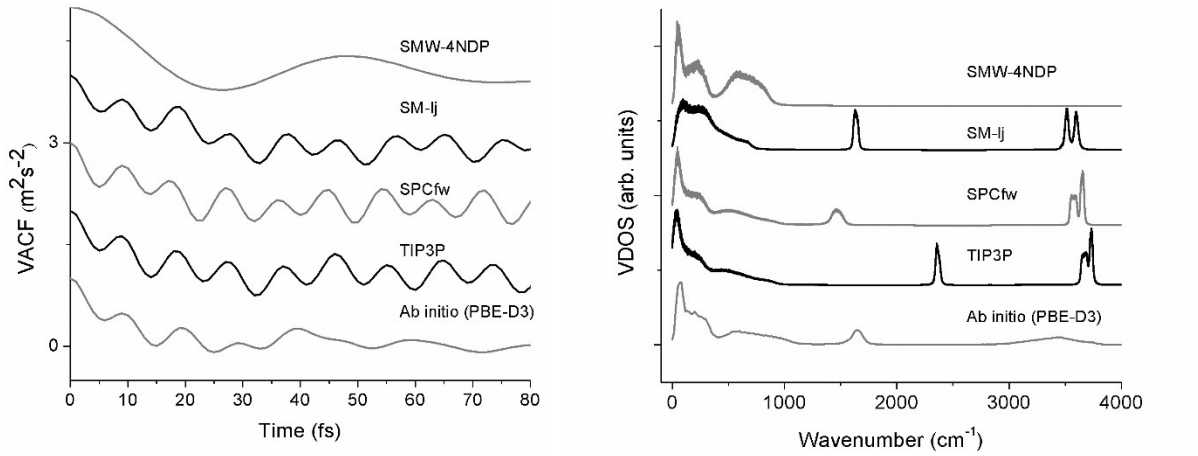



Figure S3.4 left: Velocity autocorrelation function (VACF) of the oxygen atoms of water. Right: Vibrational spectrum of the different water models

The vibrational spectrum of the TIP3P model fails to reproduce the frequencies obtained using ab initio MD, while SM-lj and the SPC/fw water are in good agreement with the ab initio obtained spectrum (main text Fig 5). Compared to the experimental Raman spectrum, the bending and stretching frequency is slightly better in the SM-lj model, the peak positions of

the SPC/fw are shifted to the left and right side of the spectrum, respectively, indicating that the SPC/fw water vibrates faster but bends slower (Bertie and Lan, 1996; Pudlas et al., 2010; Willow et al., 2015). The reason for this shift is the parameterization of the force field. In case of the bending vibration, the angular spring constant is smaller for SPC/fw water compared to SM-lj, resulting in a lower bending frequency. While in the case of the stretching frequency, although the two models use different potentials to describe the bonds, the spring constant for SPC/fw is higher relative to the spring constant of the SM-lj model, which results in a higher frequency of the stretching vibration of the water. To summarize, at structural level all the tested water models behave similar, but when considering the vibrational spectra of the simulated water, the SM-lj force field provides the best agreement with experimental vibrational data. In what follows, comparisons between SPC/fw and SM-lj are made, and other force fields were not tested.

### 3.5.5 Comparison of the RDF ($O_w$–$O_w$) obtained using AIMD with experiments

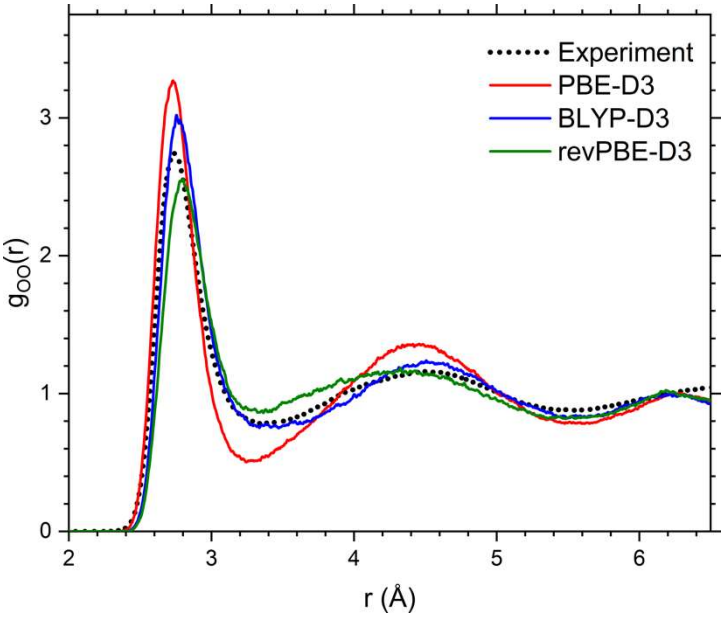



Figure S3.5 Oxygen–oxygen radial distribution function, gOO(r), for pure liquid water obtained from the AIMD trajectories using the PBE-D3, BLYP-D3, and revPBE-D3 methods, and neutron diffraction experiments.

Information regarding the structural properties of water can be quantified from the oxygen–oxygen (O–O) radial distribution function (RDF), which represent the probability of finding two oxygen atoms at a distance *r*. Previous AIMD simulations using pure GGA functionals yield over structured water compared to experiment and slow dynamics at T = 300 K, suggesting that the freezing point of the simulated water is around 400 K (Schwegler et al., 2004; Sit and Marzari, 2005), which reflects the weakness of pure GGA approximations for the simulation of liquid systems. In this study, the PBE, BLYP, and revPBE functionals were, therefore, augmented by the Grimme dispersion correction term DFT-D3 (Grimme et al., 2010). Figure S3.5 compares the O–O RDFs of liquid water obtained from the AIMD trajectories using the PBE-D3, BLYP-D3, and revPBE-D3 methods with the experimental profile obtained from neutron diffraction data (Soper, 2000). The inclusion of van der Walls corrections to GGA functionals generally leads to a softer liquid structure (Lin et al., 2012; J. Wang et al., 2011) but Figure S3.5 shows that the water simulated with the PBE-D3 functional

is still over-structured compared to experiment, whilst better is obtained using the BLYP-D3 and revPBE-D3 density functional methods with included van der Waals interactions.

### 3.5.6 Water dynamics around calcium

Figure S3.6 shows the survival function of the SM-lj model, in the inset average distance is plotted against the Residence Time (RT). The hydrogen bond distribution of the groups, long and short RT, is different compared to the SPC/fw water model. For the SM-lj there is a minimal difference between the two groups (Figure S3.7b). This shows a lack of a hydrogen bond network and may clarify the lower MRT compared to the other water models (Table S3.1). The boxplots, however, reveal similar trend as for the SPC/fw water model (Figure S3.7c, shorter distance of for long RT compared to longer distance for short RT. Though, the distribution is wider as expected from the distance plot (Figure S3.6b).



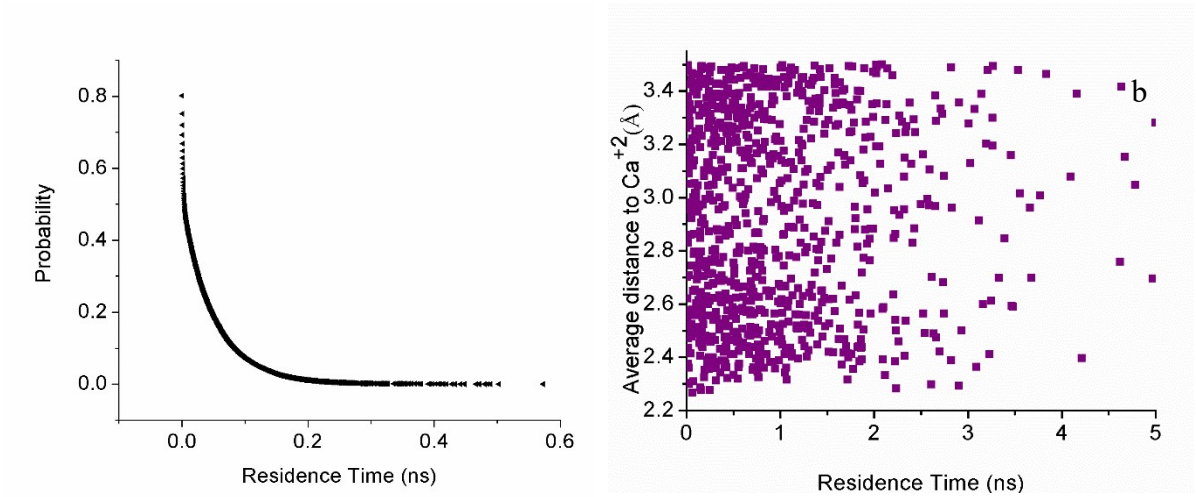
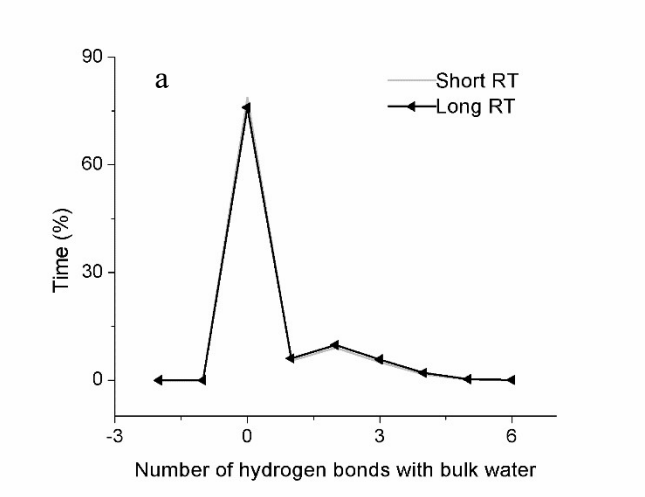



Figure S3.6 a) The probability distribution of different residence times (RT) for first-shell water molecules obtained using SM-lj, which are two exponential fits. b) Average distance between a coordinated water molecule and the calcium cation as a function of the residence time of that water molecule.



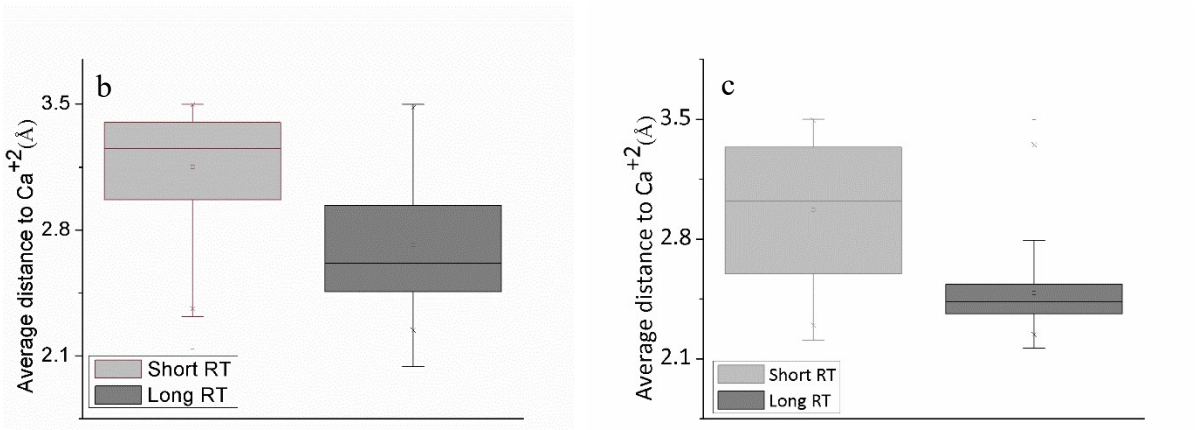




Figure S3.7 a) the relative time a water molecule is in a specific configuration with the number of hydrogen bonds formed with bulk of second shell water minus the number of hydrogen bonds formed coordinated waters. The two gaussians represent the two different distributions (based on output obtained using SM-lj). b) two boxplots indicating the difference in distance between the calcium and the coordinated water for short and long staying water molecules in the first 20 ps of the simulation (also obtained from data generated with the SM-lj). c) Same as b) only with data generated with SPC/fw water model.

The abundance of coordination numbers (CN) for $Ca^{2+}$ during the simulation, and the distinction if that CN contained at least one water molecule with a short rt, can be found in Table S3.5. The average coordination number of SM-lj water is found to be 8.3. For 73% of the configurations, the short RT only happens when the CN is +1 above the average (Table S3.5). This agrees with the results for the SPC/fw water model (data not shown), but the difference is less pronounced.

Table S3.5 The coordination number of SM-lj calcium in relation to residence time of first-shell water molecules (3.5Å cut-off).

| CN | all configurations | configurations with only long RT-water | configurations with a short RT-water |
|---|---|---|---|
| 6 | 0% | 0% | |
| 7 | 3% | 3% | 1% |
| 8 | 62% | 63% | 24% |
| 9 | 35% | 34% | 73% |
| 10 | 0% | 0% | 2% |

### 3.5.7 Python codes

```
##/usr/bin/env python

import math
import numpy
import csv
from numpy import linalg



##-----------------------------------------#
## I/O files to open                      #
##-----------------------------------------#

history = open('HISTORY', 'r')
step=open('step','r') #### File with number of steps in HISTORY file (grep 'timestep'
HISTORY>step)

output_DM=open('output_DM.txt','w')  ####  Output Direct method exchange frequency------
-----
gout=open('survivalfunction.txt','w')#### Output file with exponential decay MRT is in the
inverse of the exponent of the best fit of the data
dist=csv.writer(open('Distances.csv','w')) #### Distance between cation and water molecule
for every timestep of residence



##-----------------------------------------#
## Input parameters                       #
##-----------------------------------------#

selected_ca = numpy.arange(1,2) #### List of calciums to calculate the MRT for
atm_1 = 'Ca'
atom1 = 1000 #### Maximum amount of Calcium molecules
atm_2 = 'OW'
atom2 = 809 #### number of water molecules in CONFIG
for line in step:
    nstep = int(line)#### to do: read from grep text file
rcut  = 3.5 #### Maximum distance of the first shell
tau   = 2.0 #### Tau for the IMPEY Survival Function method
tau1  = 0.5 #### Tau for Direct method


##-----------------------------------------#
## Definitions of the used Functions      #
##-----------------------------------------#

def M33INV(A):
##-----------------------------------------#
```

```
## To calculate the inverse of a matrix        #
##-------------------------------------------#
    COFACTOR = numpy.zeros((3,3))
    OK_FLAG = False
    EPS = 1.0E-10

    DET = A[0][0]*A[1][1]*A[2][2] - A[0][0]*A[1][2]*A[2][1] - A[0][1]*A[1][0]*A[2][2] +
A[0][1]*A[1][2]*A[2][0] + A[0][2]*A[1][0]*A[2][1] - A[0][2]*A[1][1]*A[2][0]
    if abs(DET) <= EPS:
        AINV = 0.0E0
        OK_FLAG = False
    else:
        COFACTOR[0][0] = +(A[1][1]*A[2][2]-A[1][2]*A[2][1])
        COFACTOR[0][1] = -(A[1][0]*A[2][2]-A[1][2]*A[2][0])
        COFACTOR[0][2] = +(A[1][0]*A[2][1]-A[1][1]*A[2][0])

        COFACTOR[1][0] = -(A[0][1]*A[2][2]-A[0][2]*A[2][1])
        COFACTOR[1][1] = +(A[0][0]*A[2][2]-A[0][2]*A[2][0])
        COFACTOR[1][2] = -(A[0][0]*A[2][1]-A[0][1]*A[2][0])

        COFACTOR[2][0] = +(A[0][1]*A[1][2]-A[0][2]*A[1][1])
        COFACTOR[2][1] = -(A[0][0]*A[1][2]-A[0][2]*A[1][0])
        COFACTOR[2][2] = +(A[0][0]*A[1][1]-A[0][1]*A[1][0])

        AINV = numpy.transpose(COFACTOR) / DET
        OK_FLAG = True

    return (AINV)

def minimum_distance(r_i, r_j, h):
    ##-------------------------------------------#
    ## Compute the inverse of the matrix          #
    ##-------------------------------------------#

    h_inv = M33INV (h)

    ##-------------------------------------------#
    ## Compute fractional coordinates of the atoms #
    ##-------------------------------------------#

    s_i = numpy.matmul(h_inv,r_i)
    s_j = numpy.matmul(h_inv,r_j)

    ##-------------------------------------------#
    ## Compute the distance between the atoms      #
    ##-------------------------------------------#
    s_ij = s_i - s_j
    s_ij = s_ij - numpy.around(s_ij)

    r_ij = numpy.matmul(h, s_ij)
```

```
    d_ij = linalg.norm(r_ij)

    return (d_ij)

def
crossover_SF(rml,i,j,s,r,ntau,rcut,nex_out,nex_in,count_in,count_out,i_in,i_out,j_in,j_out):
    ##-----------------------------------------#
    ## Impey survival method                  #
    ##   Exchange in the MIDDLE of HISTORY    #
    ##-----------------------------------------#
    if rml[i][j][s] < rcut and rml[i][j][s+1] > rcut:

        for p in range(s+1, int(s+ntau+1)):

            if rml[i][j][p] > rcut:

                r = r + 1

                if r == ntau:
                    ex_out=numpy.array([i,j,s])
                    nex_out.append(ex_out)

                    count_out.append(s)
                    i_out.append(i)
                    j_out.append(j)

                    break


    elif rml[i][j][s] > rcut and rml[i][j][s+1] < rcut:
                ex_in=numpy.array([i,j,s])
                nex_in.append(ex_in)

                count_in.append(s)
                i_in.append(i)
                j_in.append(j)

    return(nex_in,nex_out,count_in,count_out)

def
crossover_SF_end(rml,i,j,s,r,nstep,rcut,exchange_transitionlist_out,exchange_transitionlist_i
n):
    ##-----------------------------------------#
    ## Impey survival method                  #
    ##   Exchange at the END of HISTORY       #
    ##-----------------------------------------#

    if rml[i][j][s] < rcut and rml[i][j][s+1] > rcut:
```

```
        for p in range(s+1, int(nstep)):

            if rml[i][j][p] > rcut:

                r = r + 1

                if r == (nstep-(s+1)):
                    ex_out=numpy.array([i,j,s,r])
                    exchange_transitionlist_out.append(ex_out)

                    break


    elif rml[i][j][s] > rcut and rml[i][j][s+1] < rcut:

                ex_in=numpy.array([i,j,s])
                exchange_transitionlist_in.append(ex_in)

def
crossover_SF_start(rml,i,j,s,r,ntau,rcut,exchange_transitionlist_start_out,exchange_transitionl
ist_start_in):
    ##-------------------------------------------#
    ## Impey survival method                    #
    ##   Exchange at the START of HISTORY       #
    ##-------------------------------------------#
    if rml[i][j][s] < rcut and rml[i][j][s+1] > rcut:



        for p in range(s+1, int(s+ntau+1)):

            if rml[i][j][p] > rcut:

                r = r + 1

                if r == ntau:
                    ex_out=numpy.array([i,j,s,r])
                    exchange_transitionlist_start_out.append(ex_out)

                    break


    elif rml[i][j][s] > rcut and rml[i][j][s+1] < rcut:

                ex_in=numpy.array([i,j,s,r])
                exchange_transitionlist_start_in.append(ex_in)

def
crossover_DM(rml,i,j,s,r,ntau1,rcut,nex_out1,nex_in1,count_in1,count_out1,i_in1,i_out1,j_in
1,j_out1):
```

```
    ##-------------------------------------------#
    ## Direct method                            #
    ##   Exchange in the MIDDLE of HISTORY      #
    ##-------------------------------------------#
    if rml[i][j][s] < rcut and rml[i][j][s+1] > rcut:



       for p in range(s+1, int(s+ntau1+1)):

          if rml[i][j][p] > rcut:
             r = r + 1


             if r == ntau1:
                ex_out1=numpy.array([i,j,s])
                nex_out1.append(ex_out1)

                count_out1.append(s)
                i_out1.append(i)
                j_out1.append(j)

                break
          elif rml[i][j][p] <= rcut:
              break


    elif rml[i][j][s] > rcut and rml[i][j][s+1] < rcut:

      for p in range(s+1, int(s+ntau1+1)):

         if rml[i][j][p] < rcut:

            r = r + 1

            if r == ntau1:
               ex_in1=numpy.array([i,j,s])
               nex_in1.append(ex_in1)

               count_in1.append(s)
               i_in1.append(i)
               j_in1.append(j)

         elif rml[i][j][p] >= rcut:
              break
    return(nex_in1,nex_out1,count_in1,count_out1,r)

def
crossover_DM_end(rml,i,j,s,r,ntau1,nstep,rcut,exchange_transitionlist_out1,exchange_transit
ionlist_in1):
```

```
    ##-------------------------------------------#
    ## Direct method                            #
    ##   Exchange in the END of HISTORY         #
    ##-------------------------------------------#
    if rml[i][j][s] < rcut and rml[i][j][s+1] > rcut:

        for p in range(s+1, int(nstep)):

            if rml[i][j][p] > rcut:

                r = r + 1

                if r == (nstep-(s+1)):
                    ex_out1=numpy.array([i,j,s,r])
                    exchange_transitionlist_out1.append(ex_out1)
                elif r == ntau1:
                    ex_out1=numpy.array([i,j,s,r])
                    exchange_transitionlist_out1.append(ex_out1)

                    break
            elif rml[i][j][p] <= rcut:
                break


    elif rml[i][j][s] > rcut and rml[i][j][s+1] < rcut:
        for p in range(s+1, int(nstep)):

            if rml[i][j][p] < rcut:

                r = r + 1

                if r == (nstep-(s+1)):
                    ex_in1=numpy.array([i,j,s,r])
                    exchange_transitionlist_in1.append(ex_in1)
                elif r == ntau1:
                    ex_in1=numpy.array([i,j,s,r])
                    exchange_transitionlist_in1.append(ex_in1)
                    break

            elif rml[i][j][p] >= rcut:
                break
    return(r)

def
crossover_DM_start(rml,i,j,s,r,ntau1,rcut,exchange_transitionlist_start_out1,exchange_transit
ionlist_start_in1,event_start_out1,event_start_in1):
    ##-------------------------------------------#
    ## Direct method                            #
    ##   Exchange in the START of HISTORY       #
    ##-------------------------------------------#
```

```
    if rml[i][j][s] < rcut and rml[i][j][s+1] > rcut:
        event_out=numpy.array([i,j,s])
        event_start_out1.append(event_out)

        for p in range(s+1, int(s+ntau1+1)):

            if rml[i][j][p] > rcut:

                r = r + 1

                if r == ntau1:
                    ex_out1=numpy.array([i,j,s,r])
                    exchange_transitionlist_start_out1.append(ex_out1)


                    break
            elif rml[i][j][p] <= rcut:
                break


    elif rml[i][j][s] > rcut and rml[i][j][s+1] < rcut:
        event_in=numpy.array([i,j,s])
        event_start_in1.append(event_in)
        for p in range(s+1, int(s+ntau1+1)):

            if rml[i][j][p] < rcut:

                r = r + 1

                if r == ntau1:
                    ex_in1=numpy.array([i,j,s,r])
                    exchange_transitionlist_start_in1.append(ex_in1)

            elif rml[i][j][p] >= rcut:
                break
    return(r)

def who_in_end(rml, i,j, s,rcut,who_is_in_end):
    ##-------------------------------------------#
    ## Impey survival method                    #
    ##   Determines the atoms in the first shell  #
    ##     in the END of the HISTORY              #
    ##   To connect two CONSECUTIVE HISTORY-files  #
    ##-------------------------------------------#

    if rml[i][j][s] < rcut:
        win=numpy.array([i,j,s])
        who_is_in_end.append(win)

def who_in_start(rml, i,j, s,rcut,who_is_in_start):
```

```
    ##-------------------------------------------#
    ## Impey survival method                    #
    ##   Determines the atoms in the first shell  #
    ##     in the START of the HISTORY           #
    ##   To connect two CONSECUTIVE HISTORY-files  #
    ##-------------------------------------------#

    if rml[i][j][s] < rcut:
      wins=numpy.array([i,j,s])
      who_is_in_start.append(wins)

def print_distances(rml,rt_ts,k_out,k_in, keyi,keyj):
    ##-----------------------------------------------------#
    ##   Write distance between cation and water molecule   #
    ##   In every timestep during residence                 #
    ##-----------------------------------------------------#

    k=0
    for i in range(k_in, k_out+1):
       k+=1
       dist.writerow([keyj, rt_ts, k, rml[keyi][keyj][i]])

def
print_restart_file_SF(frame,R_t,nstep,ntau,exchange_transitionlist_out,exchange_transitionlis
t_in,exchange_transitionlist_out1,exchange_transitionlist_in1,exchange_transitionlist_start_o
ut,exchange_transitionlist_start_in,exchange_transitionlist_start_out1,exchange_transitionlist
_start_in1,exchange_transitionlist_startswith_out,
exchange_transitionlist_endswith_in,tsim,who_is_in_end,who_is_in_start):
    ##-------------------------------------------#
    ## Impey survival method                    #
    ##   Write file to connect two              #
    ##   CONSECUTIVE HISTORY files              #
    ##-------------------------------------------#

    restart=csv.writer(open(frame+'_SF.csv','w'))
    restart.writerow(['Exchanges at the Boundary to determine an event over HISTROY-files'])
    restart.writerow(['Start of file: nstep:',nstep,'tau:', ntau,'simulation time',tsim])
    restart.writerow(['Survival function'])
    restart.writerow(['Survival function: start in'])
    for x in range(0,len(exchange_transitionlist_start_in)):
       restart.writerows([exchange_transitionlist_start_in[x]])
    restart.writerow(['Survival function: start out'])
    for x in range(0,len(exchange_transitionlist_start_out)):
       restart.writerows([exchange_transitionlist_start_out[x]])
    restart.writerow(['Survival function: end in'])
    for x in range(0,len(exchange_transitionlist_in)):
       restart.writerows([exchange_transitionlist_in[x]])
    restart.writerow(['Survival function: end out'])
    for x in range(0,len(exchange_transitionlist_out)):
       restart.writerows([exchange_transitionlist_out[x]])
```

```
    restart.writerow(['Survival function: HISTORY ends with an IN event'])
    for x in range(0,len(exchange_transitionlist_endswith_in)):
        restart.writerows([exchange_transitionlist_endswith_in[x]])
    restart.writerow(['Survival function: HISTORYs first OUT event'])
    for x in range(0,len(exchange_transitionlist_startswith_out)):
        restart.writerows([exchange_transitionlist_startswith_out[x]])
    restart.writerow(['Survival function: who is in at the end'])
    for x in range(0,len(who_is_in_end)):
        restart.writerows([who_is_in_end[x]])
    restart.writerow(['Survival function: who is in at the start'])
    for x in range(0,len(who_is_in_start)):
        restart.writerows([who_is_in_start[x]])
    restart.writerow(['Survival function: data R_t'])
    for k in range(0,len(R_t)):
        if numpy.sum(R_t[k])!=0:
          restart.writerows([R_t[k]])

def
print_restart_file_DM(dt,frame,nstep,ntau1,exchange_transitionlist_out,exchange_transitionli
st_in,exchange_transitionlist_out1,exchange_transitionlist_in1,exchange_transitionlist_start_
out,exchange_transitionlist_start_in,exchange_transitionlist_start_out1,exchange_transitionlis
t_start_in1,exchange_transitionlist_startswith_out,
exchange_transitionlist_endswith_in,event_start_out1,event_start_in1,tsim,count_in_out1,wh
o_is_in_end,who_is_in_start):
    ##-----------------------------------------#
    ## Direct method                          #
    ##   Write file to connect two            #
    ##   CONSECUTIVE HISTORY files            #
    ##-----------------------------------------#
    restart1=csv.writer(open(frame+'_DM.csv','w'))
    restart1.writerow(['Exchanges at the Boundary to determine an event over HISTROY-
files'])
    restart1.writerow(['Start of file: nstep: ',nstep,'tau:',ntau1,'total exchanges before
correction',count_in_out1,'dt',dt,'simulation time',tsim])
    restart1.writerow(['Direct method'])
    restart1.writerow(['Direct method: start in'])
    for x in range(0,len(exchange_transitionlist_start_in1)):
        restart1.writerows([exchange_transitionlist_start_in1[x]])
    restart1.writerow(['Direct method: start out'])
    for x in range(0,len(exchange_transitionlist_start_out1)):
        restart1.writerows([exchange_transitionlist_start_out1[x]])
    restart1.writerow(['Direct method: end in'])
    for x in range(0,len(exchange_transitionlist_in1)):
        restart1.writerows([exchange_transitionlist_in1[x]])
    restart1.writerow(['Direct method: end out'])
    for x in range(0,len(exchange_transitionlist_out1)):
        restart1.writerows([exchange_transitionlist_out1[x]])
    restart1.writerow(['Direct method: start event_out'])
    for x in range(0,len(event_start_out1)):
        restart1.writerows([event_start_out1[x]])
```

```
    restart1.writerow(['Direct method: start event_in'])
    for x in range(0,len(event_start_in1)):
        restart1.writerows([event_start_in1[x]])
    restart1.writerow(['Direct method: who is in at the end'])
    for x in range(0,len(who_is_in_end)):
        restart1.writerows([who_is_in_end[x]])
    restart1.writerow(['Direct method: who is in at the start'])
    for x in range(0,len(who_is_in_start)):
        restart1.writerows([who_is_in_start[x]])

def
RT(nex_in,nex_out,i_in,j_in,count_in_out,rml,exchange_transitionlist_endswith_in,exchange
_transitionlist_startswith_out):
    ##-----------------------------------------#
    ## Process survival function              #
    ##-----------------------------------------#
    #....Initialize.
    dist.writerow(['OW', 'RT','Time', 'Distance'])
    rt=False
    N_0=0 #### total number of collected residence times
    y=0
    p=0
    i_in_dic={}
    j_in_dic={}
    key_prev_i=-1
    key_prev_j=-1
    R_t=numpy.zeros(((int(count_in_out)),5),dtype=numpy.int)
    j_in=sorted(set(j_in))
    i_in=sorted(set(i_in))



    for line in nex_out:
        ##-----------------------------------------#
        ## Impey survival method                  #
        ##   First out-event in HISTORY for moleucle i #
        ##   To connect two CONSECUTIVE HISTORY files  #
        ##-----------------------------------------#
        key_out_i=line[0]
        key_out_j=line[1]

        if key_out_i!=key_prev_i or key_out_j!=key_prev_j:

            key_prev_i=key_out_i
            key_prev_j=key_out_j
            k_out_start=numpy.array([key_out_i,key_out_j,line[2]])
            exchange_transitionlist_startswith_out.append(k_out_start)
```

```
    for keyi in i_in:
        #### For loop to get cation number of in-event
        rt=False
        a=0
        b=0
        for keyj in j_in:
            #### For loop to get water number of in-event
            rt=False
            a=0
            b=0


            for x in range(y,(len(nex_in))):


                if y==(len(nex_in)) or p==(len(nex_out)):

                  break

                #### Search timestep of in-event with between ith cation and jth watermolecule
                if nex_in[x][0]==keyi:

                  if nex_in[x][1]==keyj:

                    if nex_out[p-1][0]==keyi and nex_out[p-1][1]==keyj and
nex_in[x][2]<nex_out[p-1][2]:
                      #### Not the right in-event, alter For loop
                      continue

                    #### Right in-event, get timestep
                    k_in=nex_in[x][2]


                    #### If in-event is found
                    rt = True

                    #### To prevent next timestep in-event to be smaller than current out-event:
                    y=x

                    if rt == True:
                      for q in range(p,(len(nex_out))):

                        if nex_out[q][0]==keyi:
                          if nex_out[q][1]==keyj:
                            if k_in<nex_out[q][2]:
                              #### If molecule number, cation number matches and timestep in-
event is smaller than out event
                              k_out = nex_out[q][2]
```

```
                p=q+1
                rt_ts = k_out-k_in
                R_t[N_0][0] = keyi
                R_t[N_0][1] = keyj
                R_t[N_0][2] = k_in
                R_t[N_0][3] = k_out
                R_t[N_0][4] = rt_ts

                ##### for the distance plot
                print_distances(rml,rt_ts,k_out,k_in, keyi,keyj)

                #### Reset flag
                rt=False

                N_0  += 1
                break


            if q == (len(nex_out)-1) and a==0:
                ##------------------------------------------#
                ## Impey survival method                    #
                ##   In-event in HISTORY for moleucle i      #
                ##    without an out-event                   #
                ##   To connect two CONSECUTIVE HISTORY files  #
                ##------------------------------------------#

                a=a+1
                k_in_end=numpy.array([keyi,keyj,k_in])
                exchange_transitionlist_endswith_in.append(k_in_end)



##-------------------------------------------------#
## Calculate frequencies of every residence time    #
##  And the decrease in probability in time         #
##-------------------------------------------------#

freq  = numpy.bincount(R_t[:,4])
bin = numpy.nonzero(freq)[0]
freq[1]=N_0-freq[1]
for f in range (2, len(freq)):
    freq[f] = freq[f-1]-freq[f]
binsurv=numpy.vstack((bin,freq[bin])).T


##-------------------------------------------------#
## Write total number of residence times            #
## And write probability of residence time in time  #
##-------------------------------------------------#
```

```
    gout.write('%3s%9i\n'% ('N_0',int(N_0)))

    for i in range(0,binsurv.shape[0]):
       gout.write('%9i%9i\n'%(binsurv[i][0],binsurv[i][1]))


    ##-------------------------------------------#
    ##   To connect two CONSECUTIVE HISTORY files  #
    ##-------------------------------------------#
    return (R_t)

def mrt (rml, n1, n2, nstep, istart, rcut, tau,tau1, dt,rx_ca,ry_ca,rz_ca,frame):
   ##-------------------------------------------#
   ## Impey survival method and Direct method     #
   ##   Determines the frequencies of the        #
   ##    residence times (Impey)                 #
   ##   And the number of exchanges (Direct)      #
   ##-------------------------------------------#

    ####.... Initialize.
    kend = nstep - istart
    nex_in,nex_in1=[],[]
    nex_out,nex_out1=[],[]
    i_in,i_in1=[],[]
    i_out,i_out1 =[],[]
    j_in,j_in1=[],[]
    j_out,j_out1 =[],[]
    count_in,count_in1=[],[]
    count_out,count_out1 =[],[]
    ntau = round(tau/dt)
    ntau1 = round(tau1/dt)
    tsim = (kend - ntau1)*dt

    #### To connect two CONSECUTIVE HISTORY files

    exchange_transitionlist_out,exchange_transitionlist_in=[],[]
    exchange_transitionlist_out1,exchange_transitionlist_in1=[],[]
    exchange_transitionlist_start_out,exchange_transitionlist_start_in=[],[]
    exchange_transitionlist_start_out1,exchange_transitionlist_start_in1=[],[]
    event_start_out1,event_start_in1=[],[]
    exchange_transitionlist_startswith_out, exchange_transitionlist_endswith_in =[],[]
    who_is_in_end,who_is_in_start=[],[]


    if ntau > 0:

      ####.... Compute the number of exchange events
      for i in range(0,len(selected_ca)):
      #### For loop over all Cations
```

```
for j in range(0, n2):
#### For loop over all Water molecules
   #Initialize counters
   s = -1
   s1 = -1

   for k in range (0, kend+1):
   #### For loop over all frames (mdsteps)
      #Initialize counters
      r = 0
      r2 = 0
      r4 = 0

      s = s + r + 1
      s1 = s1 + r1 + 1

      r1 = 0
      r3 = 0
      r5 = 0


      ####c---- Check if there is a cross-point by comparing two consecutives
      ####c---- elements of RML, i.e. rml(i,j,k) and rml(i,j,k+1).
      if (s+ntau) <= kend:
         ##-------------------------------------------#
         ## Impey survival method                    #
         ##   Exchange in the MIDDLE of HISTORY      #
         ##-------------------------------------------#
         nex_in,nex_out,count_in,count_out=crossover_SF(rml, (selected_ca[i]-1), j, s,
r, ntau, rcut, nex_out, nex_in,count_in,count_out,i_in,i_out,j_in,j_out)

      if (s1+ntau) <= kend:
         ##-------------------------------------------#
         ## Direct method                            #
         ##   Exchange in the MIDDLE of HISTORY      #
         ##-------------------------------------------#

         nex_in1,nex_out1,count_in1,count_out1,r1=crossover_DM(rml,
(selected_ca[i]-1), j, s1, r1, ntau1, rcut, nex_out1,
nex_in1,count_in1,count_out1,i_in1,i_out1,j_in1,j_out1)



      #### To connect two CONSECUTIVE HISTORY files

      if s > (kend-ntau) and s<kend:
         ##-------------------------------------------#
         ## Impey survival method                    #
         ##   Exchange in the END of HISTORY         #
```

```
                    ##-------------------------------------------#
                    crossover_SF_end(rml, (selected_ca[i]-1), j, s, r2, nstep,
rcut,exchange_transitionlist_out,exchange_transitionlist_in)


                if s1 > (kend-ntau)and s1<kend:
                    ##-------------------------------------------#
                    ## Direct method                            #
                    ##   Exchange in the END of HISTORY         #
                    ##-------------------------------------------#
                    r3=crossover_DM_end(rml, (selected_ca[i]-1), j, s1, r3, ntau1,nstep,
rcut,exchange_transitionlist_out1,exchange_transitionlist_in1)


                if s <=ntau+1:
                    ##-------------------------------------------#
                    ## Impey survival method                    #
                    ##   Exchange in the START of HISTORY       #
                    ##-------------------------------------------#
                    crossover_SF_start(rml, (selected_ca[i]-1), j, s, r4, ntau,
rcut,exchange_transitionlist_start_out,exchange_transitionlist_start_in)

                if s1<=ntau:
                    ##-------------------------------------------#
                    ## Direct method                            #
                    ##   Exchange in the START of HISTORY       #
                    ##-------------------------------------------#
                    r5=crossover_DM_start(rml, (selected_ca[i]-1), j, s1, r5, ntau1,
rcut,exchange_transitionlist_start_out1,exchange_transitionlist_start_in1,event_start_out1,ev
ent_start_in1)


                if s1==kend:
                #### Impey survival method; check who is in in last frame
                    who_in_end(rml, (selected_ca[i]-1),j, s1,rcut,who_is_in_end)


                if s1==0:
                #### Impey survival method; check who is in in first frame
                    who_in_start(rml,(selected_ca[i]-1),j, s,rcut,who_is_in_start)



    ##-----------------------------------------#
    ## Write output Direct method              #
    ##-----------------------------------------#
    count_in_out1=len(count_in1)+len(count_out1)
```

```
        output_DM.write('%-20s%15.9f%9s%9i%3s%15.9f\n'%('simulation
time',tsim,'nex_tot',count_in_out1,'tau',(ntau1*dt)))


        ##-----------------------------------------#
        ## Process survival function              #
        ##-----------------------------------------#

        count_in_out=len(count_in)+len(count_out)
        R_t=RT(nex_in,
nex_out,i_in,j_in,count_in_out,rml,exchange_transitionlist_endswith_in,exchange_transitionl
ist_startswith_out)


        #### To connect two CONSECUTIVE HISTORY files


print_restart_file_SF(frame,R_t,nstep,ntau,exchange_transitionlist_out,exchange_transitionlis
t_in,exchange_transitionlist_out1,exchange_transitionlist_in1,exchange_transitionlist_start_o
ut,exchange_transitionlist_start_in,exchange_transitionlist_start_out1,exchange_transitionlist
_start_in1,exchange_transitionlist_startswith_out,
exchange_transitionlist_endswith_in,tsim,who_is_in_end,who_is_in_start)

print_restart_file_DM(dt,frame,nstep,ntau1,exchange_transitionlist_out,exchange_transitionli
st_in,exchange_transitionlist_out1,exchange_transitionlist_in1,exchange_transitionlist_start_
out,exchange_transitionlist_start_in,exchange_transitionlist_start_out1,exchange_transitionlis
t_start_in1,exchange_transitionlist_startswith_out,
exchange_transitionlist_endswith_in,event_start_out1,event_start_in1,tsim,count_in_out1,wh
o_is_in_end,who_is_in_start)

def history_parser (nstep, atm_1, atom1, atm_2, atom2,ipcal,ipcal_interval, fichero):
    ##-------------------------------------------------------------------------------#
    ## Reading of the DL_POLY HISTORY file                                          #
    ##   And collect distances of the molecules of interest in all frames in 3D matrix  #
    ##-------------------------------------------------------------------------------#

     ####.... Initialize.
        contador = 0
        av=[]
        bv=[]
        cv=[]
        ac=[]
        maxi=0
        frame=[]
        caja = False
        md_step = 0
        natms_1 = 0
        natms_2 = 0
        switch = 0
        contH=0
```

```
        rx = numpy.zeros((5,max(atom1,atom2)))
        ry = numpy.zeros((5,max(atom1,atom2)))
        rz = numpy.zeros((5,max(atom1,atom2)))
        r_i,r_j = numpy.zeros((3)),numpy.zeros((3))


        #### Read HISTORY line by line
        for linea in fichero:
            contador += 1
            if md_step==nstep+1:
              break
            if linea.find("timestep")!=-1:
                frame=(linea.split()[1])
                tstep =(linea.split()[5])
                caja= True

                md_step +=1


                if md_step == 1:
                  step_1 = int(frame)

                elif md_step == 2:

                  step_2 = int(frame)
                  dt = (step_2 - step_1)*float(tstep)
                  rx_ca = rx[0]
                  ry_ca = ry[0]
                  rz_ca = rz[0]



            if linea.find("timestep")==-1 and linea[0]!=" ":
            #### Reads the line with the atom name
                ccord=False
                linea=linea[:-1]

                atom=(map(str,linea.split()))[0]
            if contador > 6 and contador%2==0 and linea[0]==" " and caja==False and
ccord==False:
            #### Reads the coordinates of the atom
                ccord=True
                linea=linea[:-1]
                ac=(map(float,linea.split()))

                if atom == atm_1:

                  if (natms_1+1) in selected_ca:
```

```
            rx[0][natms_1] = ac[0]
            ry[0][natms_1] = ac[1]
            rz[0][natms_1] = ac[2]

        natms_1 = natms_1 + 1

    elif atom == atm_2:

        flag=True

        rx[1][natms_2] = ac[0]
        ry[1][natms_2] = ac[1]
        rz[1][natms_2] = ac[2]
        natms_2 = natms_2 + 1

        if natms_2==atom2:

          if md_step==1:
            #### Initialize distance matrix
            rml = numpy.zeros((natms_1, natms_2, nstep))
          for i in range(0,len(selected_ca)):

            r_i[0] = rx[0][(selected_ca[i]-1)]
            r_i[1] = ry[0][(selected_ca[i]-1)]
            r_i[2] = rz[0][(selected_ca[i]-1)]

            for j in range(0,(natms_2)):
              r_j[0] = rx[1][j]
              r_j[1] = ry[1][j]
              r_j[2] = rz[1][j]

              d_ij = minimum_distance (r_i, r_j, h)

              rml[i][j][md_step-1] = d_ij



          rx = numpy.zeros((5,max(atom1,atom2)))
          ry = numpy.zeros((5,max(atom1,atom2)))
          rz = numpy.zeros((5,max(atom1,atom2)))
          if md_step<nstep:
            #### Reset counters
            natms_1 = 0
            natms_2 = 0

if caja == True and contador>2 and linea[0]==" ":
    maxi +=1
    h = numpy.zeros((3,3))
```

```
            if maxi ==1:
            #### Reads vector A
              av=(map(float,linea.split()))
            if maxi ==2:
            #### Reads vector B
              bv=(map(float,linea.split()))
            if maxi ==3:
            #### Reads vector C
              cv=(map(float,linea.split()))
              maxi=0
              caja=False
              h[0][0] = av[0]
              h[1][0] = av[1]
              h[2][0] = av[2]

              h[0][1] = bv[0]
              h[1][1] = bv[1]
              h[2][1] = bv[2]

              h[0][2] = cv[0]
              h[1][2] = cv[1]
              h[2][2] = cv[2]


        mrt (rml, natms_1, natms_2, nstep, 1, rcut, tau,tau1, dt,rx_ca,ry_ca,rz_ca,frame)


##-------------------------------------------#
##          MAIN BLOCK OF CODE               #
##-------------------------------------------#

print "-------------------------------------------------------------"
print "------------------python version-------------------------"
print "--------------DL_POLY Mean Residence Time-----------------"
print "---------------Impey's Survival function & ---------------"
print "---------------Direct method Exchange frequency-----------"
print "----------------jakoskamp@gmail.com---------------------"
print "-------------------------------------------------------------"
print "-------------------------------------------------------------"
print "The process may take several hours,..."
print "Depending on de size of the HISTORY file"
print "-------------------------------------------------------------"
print "-------------------------------------------------------------"

history_parser(nstep, atm_1, atom1, atm_2, atom2,1,1, history)

history.close()
```

4

# Chapter 4 Influence of biomolecules on amorphous Calcium Carbonate formation: a molecular dynamics investigation

Almost ready to be submitted as: **Koskamp J. A., Ruiz-Hernandez S. E., Baken A., De Leeuw N. H., Wolthers, M.** Influence of biomolecules on amorphous Calcium Carbonate formation: a molecular dynamics investigation. *Crystal Growth and Design*

**Abstract:** Biomolecules are considered to play an important role during different stages of biomineralization. We performed molecular dynamics simulations of different aspartic acid-based biomolecules and the aspartate-containing protein ovocleidin-17, together with a high concentration of calcium ($Ca^{2+}$) and carbonate ($CO_3^{2-}$) ions dissolved in water. In these systems, we investigated the mechanism of biomolecule-mediated formation of amorphous calcium carbonate (ACC), a precursor phase for several crystal polymorphs. The results reveal that, among other things, (i) $CaCO_3$ ion-pairs are formed instantly without a significant role of the biomolecule. (ii) During cluster growth, the biomolecules inhibit dehydration of the ACC clusters and (iii) alter the stoichiometry in the cluster, creating overall negatively charged ($CO_3$ excess) particles as confirmed by ζ-potential experiments. (iv) The different biomolecules direct the shape of the final cluster in different ways. And (v) the positively charged arginine residues in the two main $\alpha$-helices of the protein bind $CO_3^{2-}$ and promote aggregation of ions around them. Altogether, this study shows the importance of the biomolecules in the initial steps of the formation of ACC by altering the mechanism of ion aggregation, guiding shape and water content, prior to any crystallization in the biomineralization process. It also suggests that amino acids can be used to guide water content and/or shape of ACC to tailor $CaCO_3$ crystal growth.

## 4.1 Introduction

Calcium carbonate ($CaCO_3$) is a commonly found biomineral and an abundantly present mineral with a wide range of morphologies, composition and therewith purpose, functionality, and usability in nature. The new formation of $CaCO_3$ particles is topic of an ongoing debate (Demichelis et al., 2011; Gebauer et al., 2018; Smeets et al., 2017). Crystallization of $CaCO_3$ from calcifying fluids in living cells goes through different stages, from forming the first clusters to the formation of amorphous calcium carbonate (ACC) nanoparticles to eventually one of the crystalline forms (aragonite, calcite, ikaite, monohydrocalcite or vaterite) (De Nooijer et al., 2014; Gower, 2008; Radha et al., 2010; Weiner and Addadi, 2011) Most biomineralization processes are suspected to include an amorphous phase as precursor prior to crystallization into the final desired crystal structure (Addadi et al., 2003; Cartwright et al., 2012; R et al., 2005; Radha et al., 2010). At all stages, different biomolecules can enhance or inhibit the process of precursor formation and ripening by for example stabilizing the current situation or favouring the next step by lowering activation energies (Belcher et al., 1996; Berman et al., 1993; Mann et al., 1993). This gives the living organism control over the mineralization and production on demand (Sano et al., 2005; Stephenson et al., 2008) while it is difficult to reproduce the processes in the calcifying fluids in experimental setting (Stephenson et al., 2008).



Biomolecules, such as proteins, interact with ions in solutions via their charged sites. For example, experiments showed that lysozyme, an eggshell protein, increases the nucleation rate of $CaCO_3$ crystals (Mass et al., n.d.). Recent computer simulations contributed to study the aggregation mechanism of charged lysozyme on dissolved $Ca^{2+}$ and $CO_3^{2-}$ ions (Rani and Saharay, 2019). Charged lysozyme was observed to attract carbonate ions, trigger aggregation of ions and formation of ACC around the positively charged side chains. The negatively charged chains, however, had a minimal impact on the growth of ACC. This is in contrast with

what has been suggested in experiments with coral skeleton proteins in which several studies showed that the negative charge on the protein stabilizes ACC: aggregation facilitated by this stabilization is thought to govern a non-classical crystallization (Falini et al., 2015; Gal et al., 2015; Mass et al., 2014; Reggi et al., 2014), a process more than 100 times faster than classical ion-by-ion growth from solution, even outside the organism. The role of the Ovocleidin-17 (OC-17) protein has also been studied using computer simulations. This work revealed that the affinity of the protein with the mineral decreases after ACC is transformed to crystalline $CaCO_3$, suggesting a catalytic function of the protein in the crystallization process (Freeman et al., 2015, 2010). Although negatively charged biopolymers are suspected to trap $Ca^{2+}$ by forming $Ca^{2+}$ globules and further enhance aggregation and the formation of ACC (Smeets et al., 2015), the positive side chains in OC-17 has been shown to overrule this effect and will bind the negative $CO_3^{2-}$ instead (Freeman et al., 2011). The different amino acids this protein is constructed of also showed an impact on the crystallization process. Single amino acids modified the nucleation and growth of hydroxyapatite (Tavafoghi Jahromi et al., 2013). For example, aspartic acid (a negative biopolymer) is found to reduce the symmetry in the crystal and hardens the crystal by inhibiting the motion of dislocations in the crystal as found in stress tests (Kim et al., 2016; Orme et al., 2001). Single aspartic acids also trigger a significant effect on the growth and morphology of $CaCO_3$ crystals by roughening the edges and rounding the corners (Montanari et al., 2016). Therefore, beside focussing on the larger biomolecules like proteins, studying the effect of single amino acids will help in filling in the gaps in knowledge needed to unravel the impact of these biomolecules on biomineral formation (Picker et al., 2012).

While experiments at the smallest scales are not yet able to follow the molecular mechanism (Innocenti Malini et al., 2017), molecular dynamic (MD) simulations (Rani and Saharay, 2019) revealed the mechanisms behind the first steps of cluster formation, contradicting conclusions

drawn on the role of the positive and negative binding sites (Innocenti Malini et al., 2017; Rani and Saharay, 2019). To study the mechanism and impact of aspartic acid-based biomolecules on calcium carbonate cluster formation at a molecular level, we simulated the effect of both isomers D and L aspartic acid (D-Asp and L-Asp, respectively), poly aspartic acid (poly-D-Asp) and a peptide containing aspartic acid. Besides that, we investigated the same process in the presence of an eggshell protein OC-17, to compare the effects of a small biomolecule with a macromolecule and contribute with knowledge about the initial steps of biomineralization. The protonation state of the biomolecules mimics a solution pH above 10.5. At this pH, the amino acids are negatively charged. The obtained results confirm the stabilizing effect of the biomolecules on $CaCO_3$ cluster formation and a dominant interaction with the amine groups. This can be suitable for tools to guide biocalcification and potentially calcium carbonate synthesis for industrial applications.



## 4.2 Methods

To study the effect of biomolecules on $CaCO_3$ clustering and formation of ACC we simulated nine different dissolved calcium carbonate systems with aspartic acid-based biomolecules and two systems without biomolecules. Five smaller cells were simulated with and without the smaller biomolecules. Six supercells were also simulated, one with the large OC-17 protein and the others with and without the smaller biomolecules.

The large-scale atomic/molecular massively parallel simulator (LAMMPS) was used to perform the classical molecular dynamics simulations (MD) (Plimpton, 1995). The simulation temperature and pressure were kept at 300 K and 1 atm in NPT ensemble with the Nosé-Hoover (Hoover, 1985; Nosé, 1984) with 0.1 and 1 ps relaxation time of the thermostat and the barostat, respectively. The equations of motion were integrated using the velocity-Verlet algorithm

(Hoover, 1985; Nosé, 1984) with a time step of 1 fs. Every system was simulated for 100 ns with 1 ns of equilibration period.

**System details.** We simulated five small boxes with 15831 water molecules and 255 $Ca^{2+}$ and 255 $CO_3^{2-}$ ions randomly distributed. Four out of the five boxes have one aspartic acid-rich biomolecule, namely D- or L-aspartic acid, a poly-D-aspartic acid with a length of five units, or a peptide (IUPAC code: DYDSD). The concentration of $CaCO_3$ was 0.89 M and the biomolecule concentration was 3.5 mM. A snapshot of one of the simulation boxes is given as an example in supplementary information Figure S4.1. We simulated the same boxes in a supercell of eight times (times two along every orthogonal axis) the small size resulting in a simulation cell with eight biomolecules, 2040 $Ca^{2+}$, 2040 $CO_3^{2-}$ ions and 126648 water molecules. Due to the large size of the OC-17 protein, this biomolecule was only simulated in the supercell with 126648 water molecules and 4080 ions, resulting in a biomolecule concentration of 0.44 mM. Additionally, OC-17 had a modified serine amino acid in which the side chain was phosphorylated and 14% rich in arginine, present in various domains in the protein. How the side arginine chains were positioned into the water depended on the tertiary structure of the protein. The labels of the atoms in the functional groups of interest (aspartate, glutamate, arginine and Serine (phosphorylated)) are shown in Figure S4.2. A table listing all systems can be found in the SI in Table S4.1.

**Force Field.** The description of the biomolecules was taken from the AMBER forcefield using the tleap program to generate all the structures and extract the topology and interaction parameters (D.A. Case, K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III et al., 2020). The amino acids were C-terminated and completed with a proton on the nitrogen to reproduce the final structure (see SI Figure S4.3). The $Ca^{2+}$ and $CO_3^{2-}$ ions were simulated using the parameters from (Demichelis et al., 2011) and the water parameters corresponding to the current SPC/fw interactions (Wu et al., 2006) were selected to not

compromise on the $CaCO_3$ interaction (Raiteri et al., 2010) and because TIP3P (incorporated in AMBER), has a diffusion coefficient twice the experimental value (Koskamp et al., 2019). Other relevant interaction parameters between water, the ions and biomolecules are listed in the SI Table S4.2 and were either constructed from combination rules or assigned based on similarities with atom types available in the force field, as it was tested before in similar organic inorganic interactions (Ruiz Hernandez and De Leeuw, 2015).

**Cluster size and solvent accessible surface area.** An in-house python script was used to calculate the number of clusters and their size at every timestep during the simulation to follow the change in cluster size distribution in time. The size is given in terms of number of ions and the gyration radius, which variance showed how liquid-like the particle behaved. The gyration radius ($R_G$) is defined as:



$$R_{gyration} = \sqrt{\frac{\sum_{i=1}^{N} m_i (r_i - r_{COM})^2}{\sum_{i=1}^{N} m_i}} \quad (1)$$

where $r_i$ and $r_{com}$ are the vectors of the $i$th atom in the cluster and of the centre of mass of the cluster, respectively. $m_i$ is the mass of the $i$th atom and $N$ is the number of atoms in the cluster. The flexibility of the clusters was measured by the weighted average relative standard deviation (RSD) of the $R_G$ for its size in terms of ions of the largest cluster in every time frame measured. The weight was defined as the relative frequency of a certain cluster size in terms of number of ions:

$$Weighted\ average\ RSD = \sum_{j=1}^{F} w_j \times RSD_j \quad (2)$$

where $w_j$ is the ratio between the frequency of clusters $j$ and the total number of largest clusters. And $RSD_j$ is the relative standard deviation of $R_G$ of cluster $j$. The cluster size increase (growth) rate at different time periods was calculated by a linear regression of the size of the largest cluster in terms of number of ions in time.

To study the assumed amount of water in and around the clusters, the solvent accessible surface area (SASA) was computed over the full trajectory to follow the change in time, using the MDTraj (McGibbon et al., 2015) module in the Anaconda Python distribution. The method consisted of exploring the area around all Ca and $CO_3$ ions with a probe with the size of a water molecule (radius=0.14 nm). This was done by running the Shrake and Rupley (Shrake and Rupley, 1973) algorithm.

**Cluster composition.** The stoichiometry of the formed clusters was calculated by dividing the number of $Ca^{2+}$ ions by the total number of ions ($Ca^{2+}$ + $CO_3^{2-}$) in the cluster. A relative value of 0.5 therefore represents a neutral, stoichiometric cluster. The presented data were an average of the last 10 ns (10000 snapshots) of the simulation. The local density of ACC was calculated from the integral of the radial distribution function from 0 to 0.9 nm divided by the volume corresponding to that radius.



**Cluster structure.** The Tanimoto index and Tanimoto distance were both used to evaluate the similarity of the internal structure of the final cluster compared with and without biomolecule, while a third Tanimoto analysis provided insight in the long-distance similarity. The formula used for the Tanimoto index is:

$$T_{index}(A,B) = \frac{|A \cap B|}{|A \cup B|} \quad (3)$$

Where $A$ and $B$ were lists containing all distances (nm) rounded at 2 decimals between all $CaCO_3$ ions in two different systems. In this case, A was a set of distances in the ACC cluster formed in the system without biomolecules, and B the distances in the cluster formed with the presence of biomolecules.

Tanimoto distance, used before to evaluate if nanoparticles where significantly different (Pérez-Badell and Montero, 2010), used the same lists as the Tanimoto index, only now, due

to the pairwise comparison the list is ordered according to the distance to the centre of mass (COM) of the cluster. The list contains first the distance between the atom closest to the COM of the cluster and the rest of the atoms in the cluster followed by the distances between the atom closest to the first atom and the rest of the atoms and so on.

$$T_{distance}(A,B) = \frac{A \times B}{\|A\|^2 + \|B\|^2 - A \times B} \quad (4)$$

To analyse the longer-range similarity, the Tanimoto analysis was performed using a histogram of all possible distances between the ions, comparing how frequent certain ion-ion distances occur in A and B.



**Cluster shape.** In order to describe the effect of the solvation shell of the biomolecule on the shape of the final cluster, we calculated the overall induced dipole moment by taking the sum of the dipole vectors of each water molecule coordinated to the biomolecule.

**Cluster crystallinity.** The overall crystallinity of the cluster was studied by a comparison of the RDFs of the pure common crystalline structures of $CaCO_3$. The initial crystal structure of calcite (Graf, 1961), aragonite (Negro and Ungaretti, 1971), vaterite (Meyer, 1969; Wang and Becker, 2009) and ikaite (Hesse and Suess, 1983) were taken from the American Mineralogist Crystal Structure Database. The structures were simulated for 1 ns using the same force field (Demichelis et al., 2013; Feng and Redfern, 2018; Nada et al., 2016), before the RDFs were constructed.

**Zeta potential measurements.** Zeta potential measurements were performed using a Zetasizer Nano ZS Ultra (Malvern). Two growth solutions were prepared: one containing $Ca^{2+}$ ($CaCl_2$, Merck) and one containing $CO_3^{2-}$ ($Na_2CO_3$, Sigma-Aldrich) with and without L-aspartic acid (Sigma-Aldrich). The pH of the latter solution was adjusted through the addition of a 50 mM NaOH (Merck) solution, so that the mixed solution would be of pH~8.7. This pH was selected

to be near the point of zero charge for $CaCO_3$ (calcite). The solutions were prepared with ultrapure water (UPW) (ISO 3696 standard grade, 1−18 mΩ) and the concentration of $CaCO_3$ in the growth solutions was 10 mM. In the solution with the biomolecule, the concentration L-aspartic acid was 0.67 mM. Before starting the measurement, 3.5 mL of each of the growth solutions was mixed in a beaker and transferred into a new DTS1070 cell using a 20 mL syringe. This cell was cleaned with, respectively, ethanol and MilliQ water prior to experiment. The ζ-potential measurements were performed at 20°C and included 25 runs, with a pause of 300 s after every run. The experiments were duplicated, and the average of both experiments is reported in the result section.



## 4.3 Results

The initial step to form a cluster that can lead to the formation of a final crystal is the formation of ion pairs. In our simulation the ion pairs were formed rapidly, after one ns of simulation (equilibration time) ~90% of the ions were paired (Table 4.1). After the first ion pairs were formed, all systems showed further agglomeration of the ions and ion pairs, and interactions with the biomolecules. In the coming sections we will present the results we obtained from this agglomeration process.

Table 4.1 Percentage of paired $Ca^{2+}$ and $CO_3^{2-}$ ions after 1 ns of equilibration

| System | Small cell | Super cell |
|---|---|---|
| **D-Asp** | 89.6 | 93.5 |
| **L-Asp** | 92.5 | 92.7 |
| **Poly-Asp** | 90.8 | 92.6 |
| **Peptide** | 90.6 | 92.6 |
| **Protein OC-17** | n.a. | 93.5 |
| **Pure** | 92.4 | 94.2 |

### 4.3.1 Cluster size

#### 4.3.1.1 Cluster distribution and evolution in time

In all simulations, the ions formed clusters. And in all systems with a biomolecule(s), the biomolecule(s) was found incorporated in the largest cluster. These clusters grew by ion or ion-pair attachment or aggregation of multiple clusters. Figure 4.1 and Figure 4.2 show the cluster distribution, in terms of number of ions, in the small simulation cell with or without one biomolecule. The probability plots in Figure 4.1 displays the cluster size distribution by calculating the chance for an ion to be part of a cluster with a certain size over 100 ns of simulation.



In the small cells, all systems with a biomolecule started with small clusters containing less than 100 ions. Then, after 20 ns larger clusters appeared and rapidly converged into the final cluster distribution (from 50 ns onwards). At the end of the 100 ns simulation, a large cluster of ~420 ions had formed in all systems and in the peptide system a slightly larger cluster of ~435 ions. This means that 82-85% of all ions in the small cells was associated with the largest cluster. Despite these similarities, differences were observed in the clustering with simulation time as shown in Figure 4.2. The system containing D-Asp revealed clusters of ~65 ions and 300-325 ions between 25 and 55 ns with only after 60 ns the appearance of the final larger cluster.

L-aspartic acid showed a slightly larger cluster sooner and the final cluster formed within 50 ns. For poly-D-Asp, an intermediate cluster like in D-Asp was absent, and the growth pattern is quite similar to L-Asp, as clearly visible in the cumulative distribution at 40 ns in Figure 4.2. The peptide system showed the fastest formation of a large cluster of ~405 ions within ~40 ns and after 75 ns the cluster slowly grew to a final cluster with ~450 ions. Contrastingly, in the pure $CaCO_3$ system, we found sharply stepped lines in the cumulative probability profiles

(Figure 4.2), indicating narrower size distributions with time. Besides this, the largest cluster formed (460 ions) was larger than in any of the biomolecule-containing systems and consisted of up to ~93% of the ions in the system. Overall, in the small systems, pure and D-Asp converged the slowest to the final cluster size distribution, then the poly-D-Asp system. The peptide and L-Asp converged fastest to the final cluster size distribution. The growth rate of the largest cluster in all small cells can be divided into two domains, a high (from 1 to 30 ns) and a low rate (from 30 to 100 ns). The highest rate in the first domain corresponds to the pure system (14.2 ions/ns). The D-Asp with 10.3 ions/ns is the slowest system observed. The growth rate in the second domain decreases to less than one ion per ns for all systems.



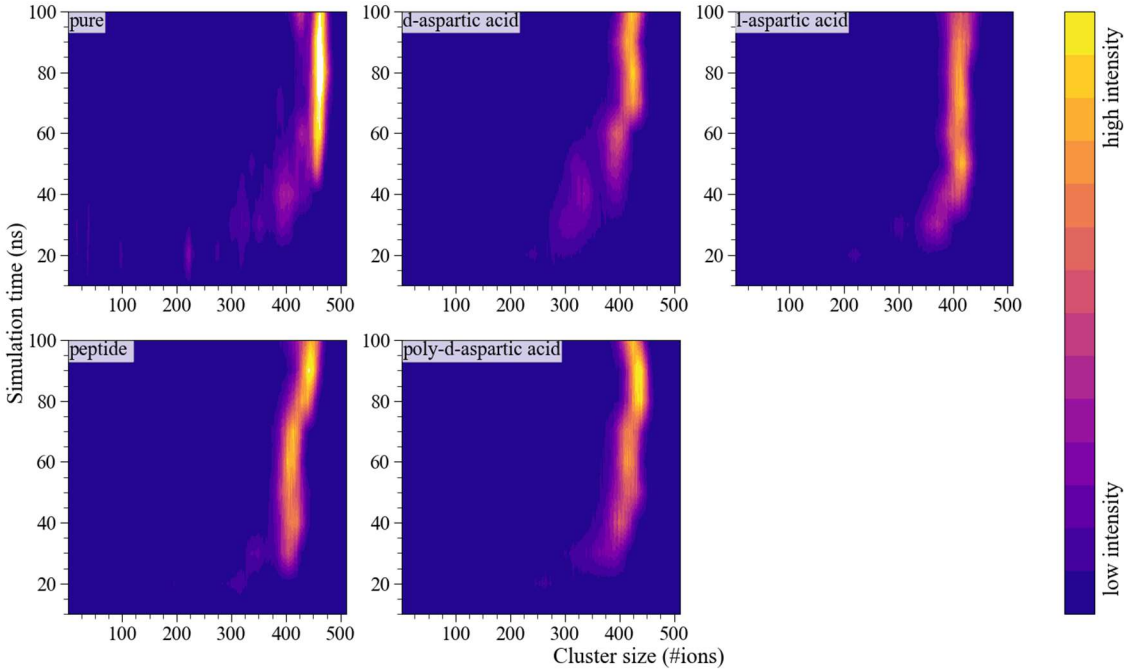



Figure 4.1 Probability intensity plots of different biomolecule-$CaCO_3$-systems showing the probability of an ion to be in a cluster with a certain size in the small cell.

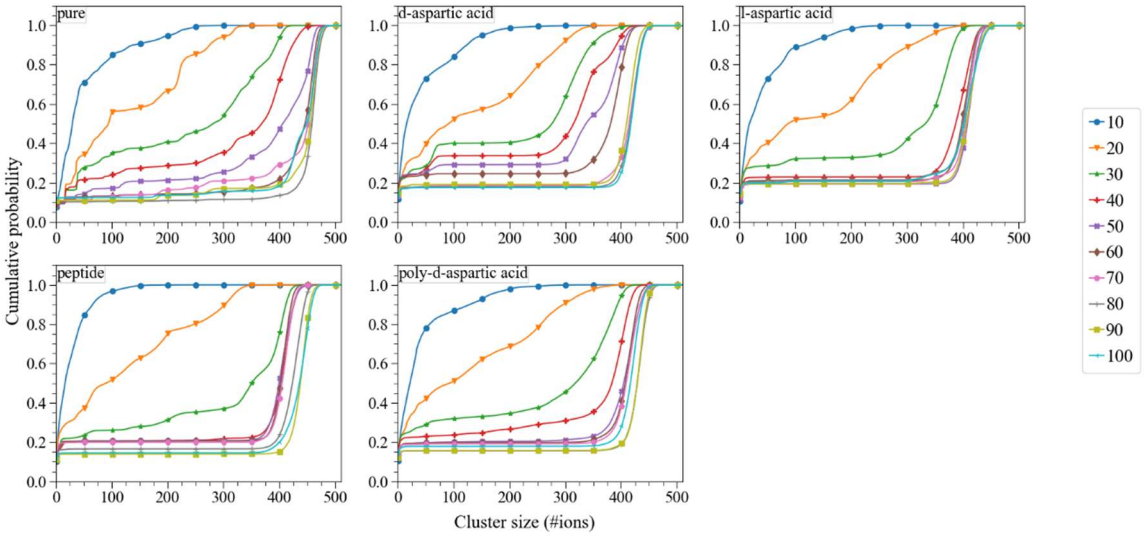



Figure 4.2 Cumulative probability profiles of the cluster sizes in terms of ions of different systems for every 10 ns of simulation of the small cell.



We created a super cell to have a system with the same concentration but the ability to build larger clusters with the extra free ions that can agglomerate to the same cluster. Additionally, we could analyse the potential impact of multiple biomolecules in a cluster. The supercell contained eight times more ions and eight biomolecules, except for the OC-17 protein system, which had one biomolecule. Consequently, the clusters became larger (Figure 4.3 and Figure 4.4) and needed more time to form than the largest clusters in the small cells. As can be seen in Figure 4.3 and Figure 4.4, all types of biomolecules showed a similar effect on the cluster size distribution compared to the pure system. In the systems with biomolecules, the final configuration showed two dominant cluster sizes, an ion was either part of an ion pair or part of the largest cluster; no intermediate sized clusters were observed. Contrastingly, the pure system showed again a less smooth profile with a larger cluster and clusters of intermediate sizes toward the end of the simulation. Unlike the small cell, all super cell systems still showed evolving cluster size distributions after 100 ns of simulation. The final largest cluster size varied from 1900 to 2800 ions in the biomolecule-systems and ~3500 ions in the pure system, meaning that only 47-69% of all ions were part of the largest cluster when a biomolecule was

present compared to 86% for the pure system. Like the small cell, the D-Asp showed the slowest convergence. The super cell with the single OC-17 protein initially showed slower clustering in the first 60 ns, and subsequently converged more rapidly in the last 40 ns of the simulation. The growth rate of the largest cluster in the super cells of pure, L-Asp, poly-D-Asp, peptide, and OC-17 can be divided into three domains over time, an intermediate, high, and low rate. In the initial segment of the simulation (first 20 ns), the pure system experienced the highest ion addition rate (39.3 ions/ns) and the systems with biomolecules averaged at 22.9 ion/ns. Then in the second period, all systems experienced a sudden increase in rate whereas, the extreme cases are represented by OC-17 (149.3 ions/ns) and pure (82.3 ions/ns); this rate was maintained for shorter (5 ns) or longer (36 ns), respectively. After this high growth rate and in the last period of simulation, the clusters grew with a lower rate. Pure showed the lowest rate of 5.4 ions/ns, while L-Asp, poly-D-Asp, and protein OC-17 had a similar rate of 13.8 ions/ns. The D-Asp system only exhibited an intermediate rate of ~26.5 ions/ns during the whole simulation.

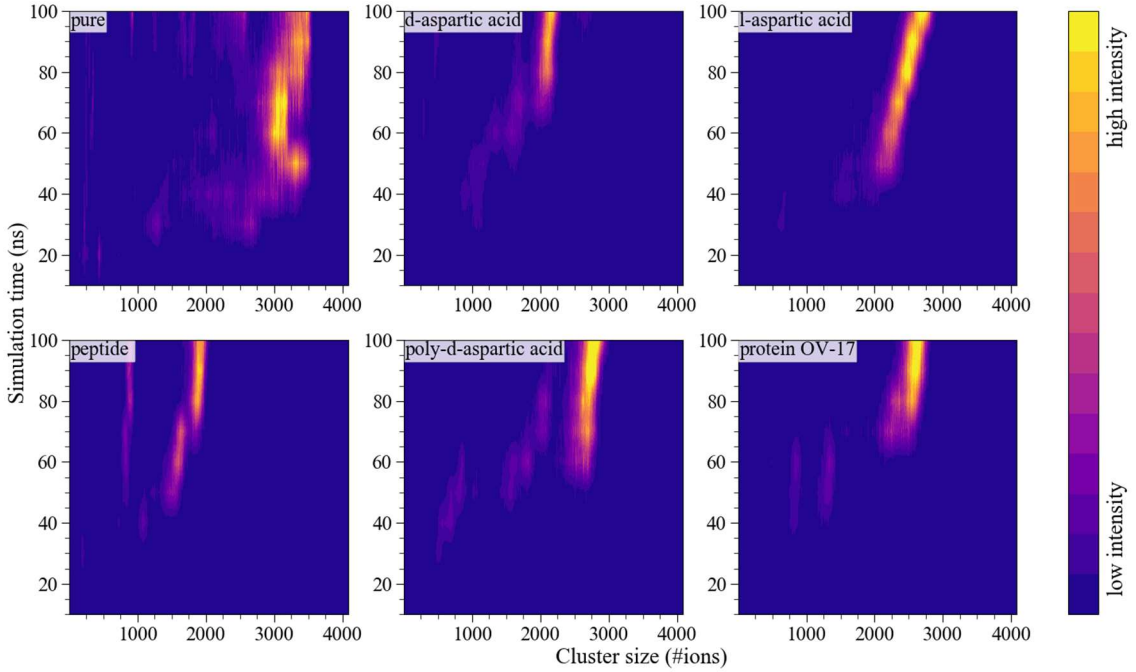



Figure 4.3 Probability intensity plots of different biomolecule-$CaCO_3$-systems and the pure system, showing the probability of an ion to be in a cluster with a certain size in the super cell.



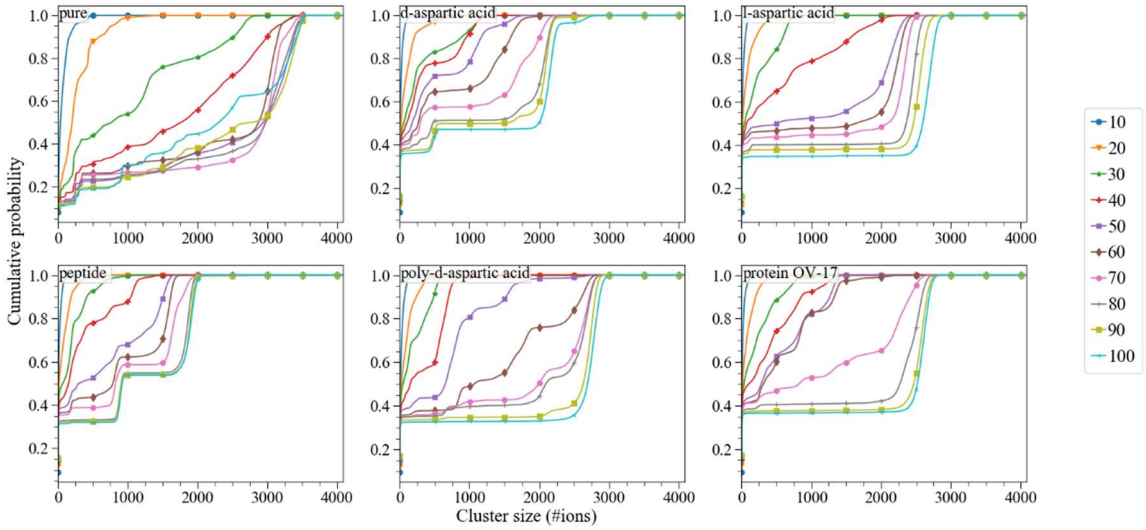



Figure 4.4 Cumulative probability profiles of the cluster sizes in terms of ions of different systems for every 10 ns of simulation of the super cell.

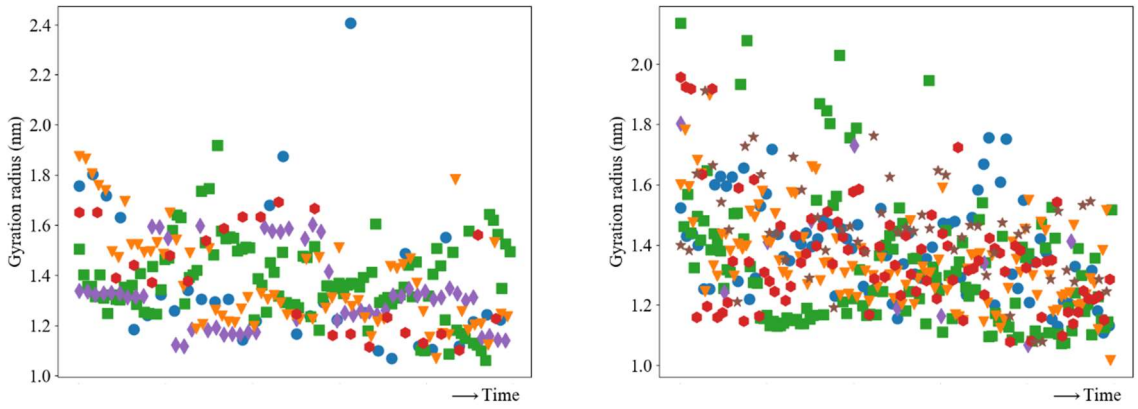
### 4.3.1.2 Gyration radius

We collected the $R_g$ of the largest cluster in all eleven systems and followed the change in time.

Doing so we could study the persistence of dynamically ordered liquid-like oxyanion polymers

(DOLLOPs) (cf. (Raiteri et al., 2012)) and the effect of the biomolecule(s). We choose clusters with 72 ions of $Ca^{2+}$ and $CO_3^{2-}$ in each of the simulated systems, according to the criteria of (Demichelis et al., 2011). When following clusters with this number of ions, we observed that the $R_g$ slightly decreased, albeit not linearly, over the course of the simulation while the number of ions remained constant (Figure 4.5). To give an indication of the variability in Rg for all clusters, irrespective of the number of ions we calculated the weighted average relative standard deviation (RSD) of the $R_g$ (Table 4.2). We observed a lower RSD for the small systems with a biomolecule compared to the small system without a biomolecule. This impact of biomolecules on the variability of Rg is not observed in the super cell.






(a) (b)

Figure 4.5 $R_G$ distribution over time of a nanoparticles containing 72 units of $Ca^{2+}$ and $CO_3^{2-}$. Left: the small cell and right: the supercell. Pure (♦), D-Asp (●), L-Asp (▼), poly-D-Asp (⬢), peptide (■), protein OC-17 (★).

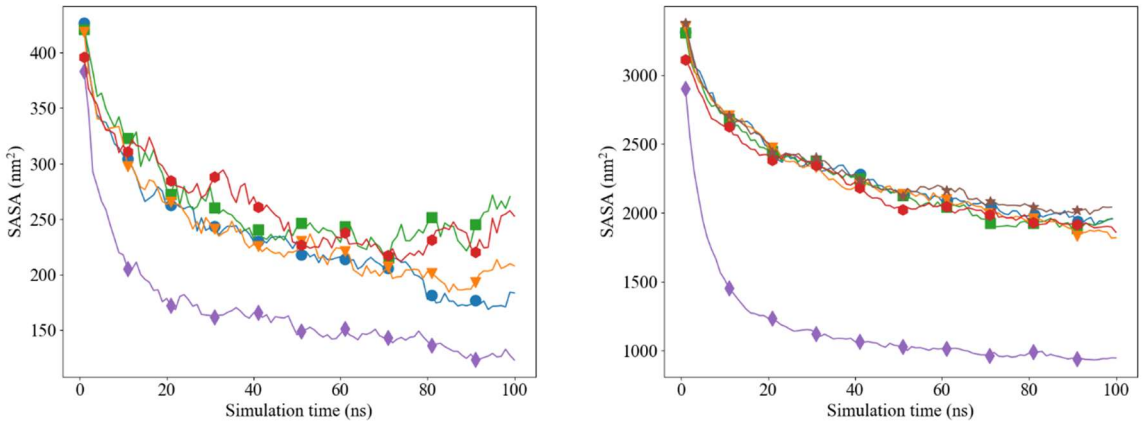
Table 4.2 Weighted average RSD of $R_G$ of all cluster sizes in terms of number of ions.

| System | Small cell | Super cell |
|---|---|---|
| d-aspartic acid | 7.99 | 8.59 |
| l-aspartic acid | 6.00 | 7.13 |
| poly-d-aspartic acid | 6.69 | 8.00 |
| peptide | 4.77 | 7.79 |
| protein OC-17 | n.a. | 8.97 |
| pure | 9.18 | 7.12 |

#### 4.3.1.3 Solvent accessible surface area

The solvent accessible surface area (SASA) was followed over time to study the effect of cluster formation on the surface area exposed to water. Figure 4.6 shows an exponential decrease in SASA over time in both the small simulation cell and the super cell for all simulated systems. In the small cell, the total decrease in SASA deviates between 36% to 57%, this is a wider range compared to the super cell in which the deviation is 40% to 46% (Table 4.3). In both cases, the larger biomolecule developed clusters with a higher SASA. The order was similar for small- and super-cell, with the notable exception of D-Asp, which showed a stronger decrease in SASA in the supercell than L-Asp. Furthermore, in the small cell, the smaller dimensions of the single aspartic acid led to a smaller SASA compared to the larger poly-D-Asp and the peptide. The SASA for both pure systems shrank significantly more over time compared to the biomolecule systems, leaving a much dryer cluster.






(a) (b)

Figure 4.6 SASA analysis using the Shrake and Rupley algorithm. Left the SASA profiles of the small simulation boxes and right the SASA profiles of the super cell simulation boxes. Pure (purple; ♦), D-Asp (blue; ●), L-Asp (orange; ▼), poly-D-Asp (red; ⬢), peptide (green; ■), protein OC-17 (brown; ★).

Table 4.3 SASA reduction in 100 ns simulation time.

| System | Small cell | Super cell |
|---|---|---|
| d-aspartic acid | 57% | 41% |
| l-aspartic acid | 50% | 46% |
| poly-d-aspartic acid | 36% | 40% |
| peptide | 36% | 41% |
| protein OC-17 | n.a. | 40% |
| pure | 68% | 67% |

### 4.3.2 Cluster composition

#### 4.3.2.1 Stoichiometry



Figure 4.7 shows the fraction of Ca over the total ions in a cluster versus the cluster size distribution in the last 10 ns of the 100-ns simulation. The clusters formed in the systems with the biomolecule all showed a tendency of forming negatively charged clusters with on average ~4% less calcium ions than carbonate ions. Contrastingly, the clusters formed in the pure $CaCO_3$ system had a stoichiometry of $Ca:CO_3 = 1:1$ or a slight excess Ca, resulting in neutral or slightly positively charged clusters.

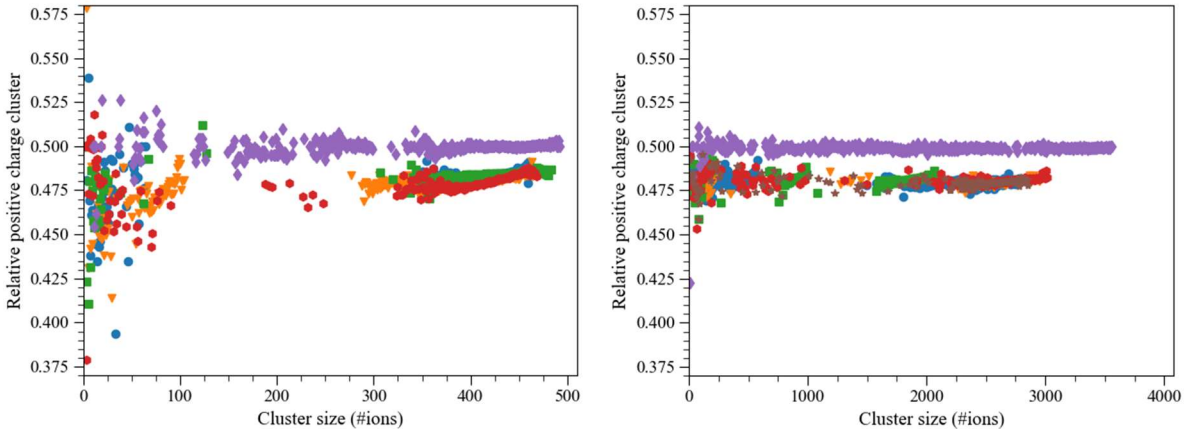



Figure 4.7 Fraction of Ca over the total ions in a cluster ($Ca^{2+}\cdot(Ca^{2+}+CO_3^{2-})^{-1}$) over the last 10 ns of the total simulation (100 ns). Left: the small cell and right: the supercell. Pure (♦), D-Asp (●), L-Asp (▼), poly-D-Asp (⬢), peptide (■), protein OC-17 (★).

#### 4.3.2.2 Density

The average $CaCO_3$ mass density of the cluster formed near the functional groups of the biomolecules is shown in Table 4.4 and Table 4.5. In both small and super cell, the densities in the pure $CaCO_3$ system were consistently ~0.4 g/cm$^3$ higher than in the biomolecule systems. In all biomolecule systems (both the small and super cell) the density of the ions around the biomolecule increased with simulation time, however L-Asp and, to a lesser extent, D-Asp, showed more fluctuations throughout the simulation. In most of the cases, the densities in the small cell were lower compared to in the super cells. Again, L-Asp was an exception to the rule: after 20 ns, the density was higher in the small cell.



The density around the side chains of aspartate, glutamate, arginine, and phosphorylated serine in the protein are listed in Table 4.5. Although a similarly increasing density was observed during the simulation, the average densities were ~40% lower compared to the densities around the smaller biomolecules (Table 4.4). Only the phosphorylated serine side chain (that pointed into the solution) highest densities, comparable to those calculated for the smaller biomolecules. The equal density around Asp, Glu and Arg in the beginning of the simulation diverted during the simulation with Arg having a slightly higher $CaCO_3$ density. The average density around all $Ca^{2+}$ was similar for all biomolecules, except for the protein system where densities were slightly higher. The lower densities for the supercell compared to the small cell, agreed with the trend found in the cluster size distributions (Figure 4.2 and Figure 4.4). That is, the small systems converged faster to their final cluster sizes than the supercells, where all small systems were converged before 70 ns the supercells converged only after 90 ns or were not able to fully converge yet within 100 ns of simulation.

Table 4.4 Density of $Ca^{2+}$ and $CO_3^{2-}$ ions around the functional groups of the biomolecules of the small (S) and super cell (XL) at different times in the simulation around different functional groups of the biomolecule and for the average density around all Ca ions.

| Density (g/cm³) | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $O_{carboxyl\ group}$ | | | | $N_{amine}$ | | | | $Ca^{2+}$ | | | |
| **Simulation time (ns)** | **20** | | **100** | | **20** | | **100** | | **20** | | **100** | |
| **System size / Biomolecule** | **S** | **XL** | **S** | **XL** | **S** | **XL** | **S** | **XL** | **S** | **XL** | **S** | **XL** |
| **L-Asp** | 0.78 | 0.73 | 0.79 | 1.01 | 0.83 | 0.73 | 0.91 | 0.99 | 0.46 | 0.37 | 0.65 | 0.62 |
| **D-Asp** | 0.58 | 0.77 | 1.00 | 1.04 | 0.58 | 0.79 | 0.85 | 1.00 | 0.47 | 0.39 | 0.73 | 0.61 |
| **Poly-D-Asp** | 0.54 | 0.69 | 0.79 | 0.88 | 0.64 | 0.79 | 0.91 | 0.95 | 0.44 | 0.39 | 0.67 | 0.62 |
| **Peptide** | 0.65 | 0.91 | 0.82 | 0.91 | 0.71 | 0.76 | 0.89 | 0.94 | 0.46 | 0.38 | 0.70 | 0.63 |
| **Pure** | n.a. | | n.a. | | n.a. | | n.a. | | 0.63 | 0.73 | 0.89 | 1.00 |

4

Table 4.5 Density of $Ca^{2+}$ and $CO_3^{2-}$ ions around the target side chains of the amino acids in the protein at different times in the simulation of the super cell (XL).

| Density (g/cm³) | | | |
|---|---|---|---|
| **Simulation Time (ns)** | | **20** | **100** |
| **Atom label** | | **XL** | **XL** |
| **Aspartate** | **OD1** | 0.35 | 0.53 |
| | **OD2** | 0.37 | 0.57 |
| | **CG** | 0.37 | 0.55 |
| **Glutamate** | **OE1** | 0.35 | 0.53 |
| | **OE2** | 0.34 | 0.53 |
| | **CD** | 0.34 | 0.53 |
| **Arginine** | **NE** | 0.36 | 0.57 |
| | **NH1** | 0.37 | 0.61 |
| | **NH2** | 0.37 | 0.60 |
| | **CZ** | 0.37 | 0.59 |
| **Serine (phosphorylated)** | **P** | 0.45 | 0.98 |
| **$Ca^{2+}$** | | 0.46 | 0.69 |

### 4.3.3 Cluster structure

#### 4.3.3.1 Radial distribution function

The radial distribution functions (RDFs) between different species are reported in Figure 4.8. Table 4.6 shows the coordination distance (CD) and number (CN) between the $Ca^{2+}$, $CO_3^{2-}$ and oxygen of water ($O_w$) of the last 10 ns of the simulation of the super cell (the small cell RDFs are shown in the SI Figure S4.4). The RDFs of the ions and water of the small and supercell are similar, yet, due to the increased statistics in the supercell results in little variation in intensity. The water structure around $Ca^{2+}$ and $CO_3^{2-}$ were identical for all systems, although the coordination number of water molecules around the ions is different from the pure $CaCO_3$ system. In particular, this is reflected in the lower peak intensities for the systems in which biomolecules were present. Also, the RDF's between the ions were close to identical for all biomolecules. Again, the pure system showed a discrepancy not only in higher CN but also in the structuring of the ions around each other (Figure 4.8, Table 4.6). In particular, the second peak in the $Ca^{2+}$ — $C_c$ (carbon of carbonate) and $Ca^{2+}$ — $O_c$ (oxygen of carbonate) showed more structuring when a biomolecule was present. The shortest distance found between $O_c$—$Na^+$ was 0.45 nm. The evolution of the RDFs over time of $Ca^{2+}$ and $CO_3^{2-}$ with water showed no significant change in peak position, however the intensity slightly decreased over time, while the peak intensity in the $Ca^{2+}$ — $O_c$ increased simultaneously. The $C_c$—$C_c$ RDF (Figure S4.5) showed increasing peak intensities with time. Additionally, these peaks became more defined over time.

The RDF and CN of the ions in solution with the different functional groups of the biomolecules are shown in Figure 4.9 (the small cell RDFs are shown in SI Figure S4.6). Again, small variations in peak position and intensity in the RDFs can be explained by an increase of statistics and the possibility of biomolecules to directly interact. In all systems and to both

amine, and carboxyl functional groups, the average coordination distance for $Ca^{2+}$ is similar: ~0.46 nm with a small peak around ~0.30 nm for OD — $Ca^{2+}$. For $CO_3^{2-}$, the coordination distance to both functional groups was much shorter, with a minimum distance around ~0.33 nm for $C_c$ and an average coordination distance of N — $O_c$ of 0.30 nm. Furthermore, the RDF of the single aspartic acids is different compared to the larger biomolecules. The protein showed more distinguishable peaks towards the ions in general. Besides that, the development over time in peak shape, position and intensity was more complex than for the other biomolecules (SI Figure S4.7, Figure S4.8 and Figure S4.9). Other side chains in the protein all showed a shorter distance to $CO_3^{2-}$ (~0.35 nm) compared to $Ca^{2+}$ (>0.45 nm), independent of the residue type (e.g. acidic or basic; SI Figure S4.10 until Figure S4.12). Both, $Ca^{2+}$ and $CO_3^{2-}$, ions had the same coordination distance to the phosphor according to Figure S4.13. In the RDF of the different arginine functional groups of OC-17, two groups were distinguished: eleven arginine residues pointed into the solution and showed a higher density of water around it; the other eight were interacting more with other protein residues and showed a lower density of water. A smaller, yet distinguishable, difference was found in the water density around the arginine groups where cluster formation continued versus where cluster formation stopped within the simulation time: around the arginine group where the cluster formation was not continuous, the water density was 19% higher (0.77 g/m$^3$, SI Figure S4.14 and Table S4.3).

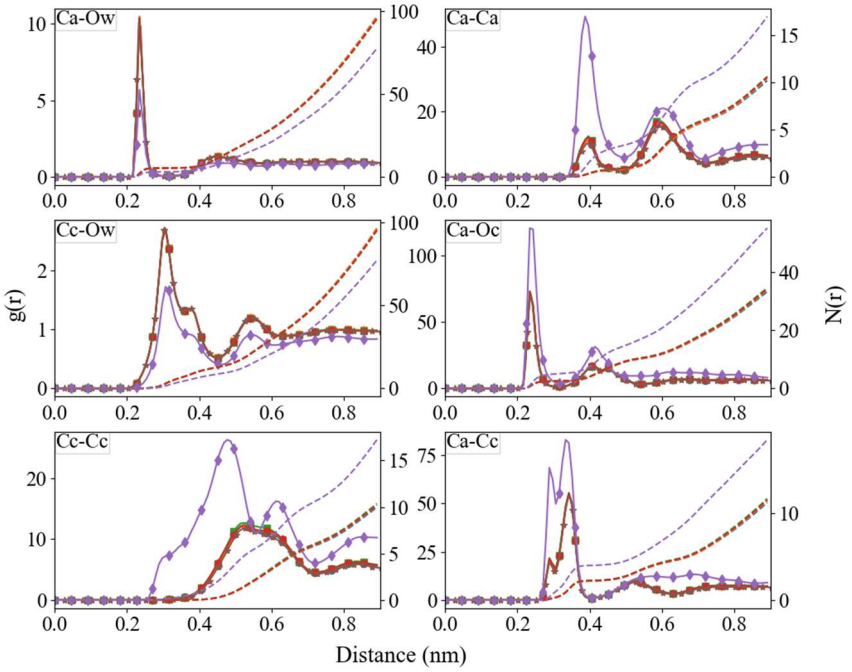
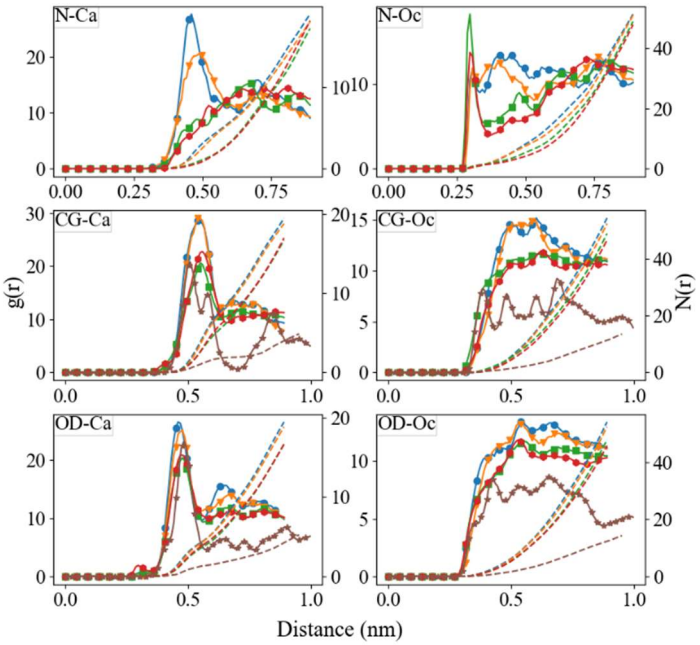




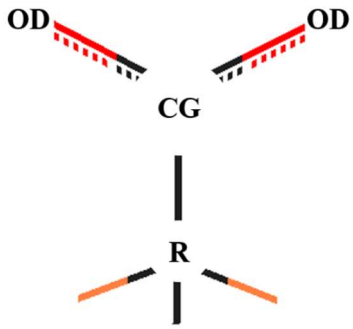

Figure 4.8 Radial distribution function (left axis) and corresponding integral (N(r)) (right axis) between $Ca^{2+}$, $C_c$, $O_c$, and $O_w$, after 100 ns of simulation of the super cell. Pure (purple; ♦), D-Asp (blue; ●), L-Asp (orange; ▼), poly-D-Asp (red; ⬢), peptide (green; ■), protein OC-17 (brown; ★).

Table 4.6 The first coordination distance (CD) and coordination numbers (CN) in the ion–ion and ion–water radial distribution functions.

| | D-Asp | | L-Asp | | Poly-D-Asp | | Peptide | | Protein OC-17 | | Pure | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Atom Pair** | **CD** | **CN** | **CD** | **CN** | **CD** | **CN** | **CD** | **CN** | **CD** | **CN** | **CD** | **CN** |
| **$Ca^{2+}$—$Ca^{2+}$** | 0.39 | 0.9 | 0.39 | 0.9 | 0.39 | 0.9 | 0.39 | 0.9 | 0.39 | 0.9 | 0.39 | 3.4 |
| **$Ca^{2+}$—$C_C$** | 0.29, 0.34 | 2.3 | 0.29, 0.34 | 2.3 | 0.29, 0.34 | 2.3 | 0.29, 0.34 | 2.3 | 0.29, 0.34 | 2.3 | 0.29, 0.34 | 4.0 |
| **$Ca^{2+}$—$O_C$** | 0.24 | 2.3 | 0.24 | 2.3 | 0.24 | 2.3 | 0.24 | 2.3 | 0.24 | 2.3 | 0.24 | 5.2 |
| **$Ca^{2+}$—$O_W$** | 0.24 | 5.3 | 0.24 | 5.3 | 0.24 | 5.3 | 0.24 | 5.3 | 0.24 | 5.3 | 0.24 | 3.0 |
| **$C_C$—$C_C$** | 0.54 | 6.0 | 0.54 | 6.0 | 0.54 | 6.0 | 0.54 | 6.0 | 0.54 | 6.0 | 0.48, 0.62 | 10.5 |
| **$C_C$—$O_W$** | 0.31 | 12.7 | 0.31 | 12.7 | 0.31 | 12.7 | 0.31 | 12.7 | 0.31 | 12.7 | 0.31 | 8.5 |




(a)




(b)

Figure 4.9 (a) Radial distribution function between $Ca^{2+}$, $C_c$, $O_c$, with the amine (N) and carboxylate (CG and OD) groups in the aspartic acid (part of the biomolecule) after 100 ns of simulation of the super cell. D-aspartic acid (blue; ●), L-Asp (orange; ▼), poly-D-Asp (red; ⬢), peptide (green; ■), protein OC-17 (brown; ★). (b) The labels for the atoms within the carboxylate group used in the RDFs presented in (a).

### 4.3.3.2 Tanimoto index, distance, and analysis

A similarity analysis between clusters was conducted to quantify the difference in internal cluster structure between pure and biomolecule-associated clusters using Tanimoto analysis. Using the Tanimoto index we tried to capture unique features of the clusters depending on the presence or absence of a biomolecule. Contrastingly, the Tanimoto distance gives an indication of the similarity of the overall cluster in terms of ion-pair distances, while the Tanimoto analysis calculated over the distance histogram (Figure 4.10) quantifies the similarity in the

distribution of ion-ion distances in the entire cluster. The Tanimoto indices for the small system (Table 4.7) showed a high similarity between the pure and biomolecule rich clusters. However, the clusters in the supercell for D-Asp and peptide had a significantly lower Tanimoto index which was consistent with the trend in the Tanimoto analysis (Table 4.7 and Figure 4.10), indicating a lower similarity in ion-pair distances in the clusters associated with D-Asp or peptide and the largest pure cluster. The Tanimoto distance showed a similarity of ~ 0.75 for all biomolecule systems compared to the pure system. It is worth noting that in the small systems, the longest $Ca^{2+}$-$CO_3^{2-}$ distances are similar for all clusters, irrespective of biomolecule presence. While a cluster of more ions was formed in the pure small cell than in the biomolecule systems, this cluster was of comparable size with a higher density and lower water content (SASA), resulting in similar longest distances (Figure 4.10a). Comparably, in the supercell, the higher density and lower water content resulted in a higher intensity of the shorter distances (Figure 4.10b). The higher intensity at longer distances for the pure system matched with the shape of the final cluster (see next section and Figure S4.15). The Tanimoto indices and distances between other (ion) pairs were also calculated and can be found in SI Table S4.4.

Table 4.7 Outcome of the different Tanimoto analysis of the largest cluster after 100 ns of simulation between the pure system (system A) and a biomolecule system (system B) for the small and super cell.

| System A | System B | Tanimoto index | | Tanimoto distance | | Tanimoto analysis | |
|---|---|---|---|---|---|---|---|
| $Ca^{2+}$ and $CO_3^{2-}$ | | S | XL | S | XL | S | XL |
| pure | D-Asp | 0.94 | 0.76 | 0.74 | 0.70 | 0.93 | 0.82 |
| | L-Asp | 0.91 | 0.98 | 0.74 | 0.79 | 0.91 | 0.86 |
| | peptide | 0.95 | 0.78 | 0.73 | 0.73 | 0.79 | 0.69 |
| | poly-D-Asp | 0.95 | 0.97 | 0.76 | 0.74 | 0.91 | 0.91 |
| | protein OC-17 | | 0.98 | | 0.75 | | 0.86 |



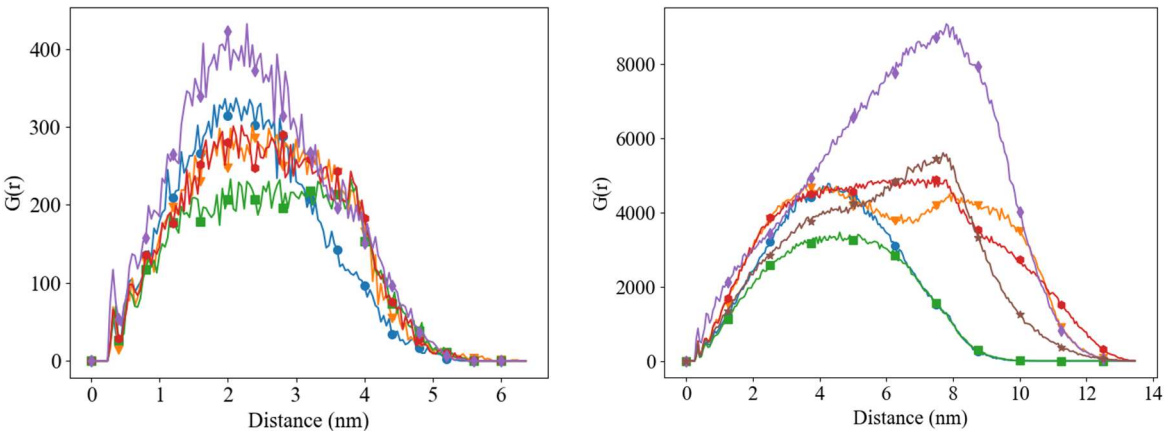



Figure 4.10 Histogram of distances between all ions in the largest cluster at 100 ns. Left: the small cell and right: the supercell. Pure (♦), D-Asp (●), L-Asp (▼), poly-D-Asp (⬢), peptide (■), protein OC-17 (★).

### 4.3.4 Cluster shape

#### 4.3.4.1 Small biomolecules

There were some distinct differences in cluster shape due to the impact of the different biomolecules. In the small cells, the large cluster formed generally had a tabular shape, meaning an anisotropic growth (Figure S4.16). Furthermore, the cluster in L-Asp and the peptide had an elongated shape crossing the periodic boundary condition (PBC), thereby

creating a tabular shaped cluster with an infinite height. Contrastingly, D- and poly-D-Asp both showed a more disc like shape that did not cross PBCs. The cluster in the pure system was more elongated, yet short enough to not cross the PBC.

In the super cells, a clear effect of the biomolecule was also observed (Figure S4.15). The D-Asp showed a smoother, more disc-shaped cluster compared to the more elongated cluster with L-Asp. However, they both did not create an infinite cluster by crossing the opposite boundaries of the PBC like in the small cell. The same held for the peptide system. Poly-D-Asp triggered the formation of a more elongated cluster than in the small cell. All systems with biomolecules resulted in a rather smooth cluster with many free ions, ion pairs and small clusters still in solution, contrary to the elongated rugose shaped cluster in the pure system. The latter cluster contained the largest number of ions, leaving only a few free ions (or ion pairs) in solution.



In general, the cluster shapes were oriented relative to the overall induced dipole moment of water in the solvation shell of the small biomolecule. In all small biomolecule systems, the resulting dipole moment pointed away from the cluster, in other words, in the direction where the lowest number of ions was present (see SI Table S4.5). When relating the resulting dipole with the shape of the clusters, the dipole was found either parallel (D-Asp) or normal (L-Asp) to the face of the tabular shaped clusters (SI Table S4.5). Additionally, the significant difference in the organization of the water caused the distinct dipoles (Figure S4.17) of the solvation water around the L- and D- aspartic acid. Although the dipole-cluster orientations over the course of the simulation were less obvious for the larger biomolecules, the resulting dipole of the final configuration did show similarities with the smaller biomolecules. Poly-D-Asp showed an induced dipole moment that was pointing parallel to the face of the tabular cluster, comparable to single D-Asp. Moreover, the single peptide system had its resulting dipole normal to the face of the tabular shaped largest cluster, like L-aspartic acid.

#### 4.3.4.2 Protein

The cluster formation and shape evolution in the protein system is shown in Figure 4.11. Initially, ion pairs formed before an interaction with the protein was observed. During the simulation, ions were found to distribute fairly evenly over the protein surface, covering the protein with a 1 nm layer of ions (SI Figure S4.18). This ionic layer around the protein contained ~4% more $CO_3^{2-}$ compared to $Ca^{2+}$. The final cluster was tabularly shaped and had two main branches that grew into the solution away from the protein and converging at ~3.5 nm from the protein, creating a donut like circle (SI Figure S4.15a and Figure S4.18e,f, and g). Those branches were rod shaped, with a similar thickness as the protein in the shortest direction (Figure 4.11d) and a length of 8.5 nm. Furthermore, we found that in one of the two α-helices, with more arginine exposed to the ionic solution, the agglomeration of ions was initially promoted compared to the agglomeration above the other α-helix (SI Figure S4.18b, c and d). After 75 ns, both branches were formed and were located above the two main α-helices present. (SI Figure S4.18b, c, and d).

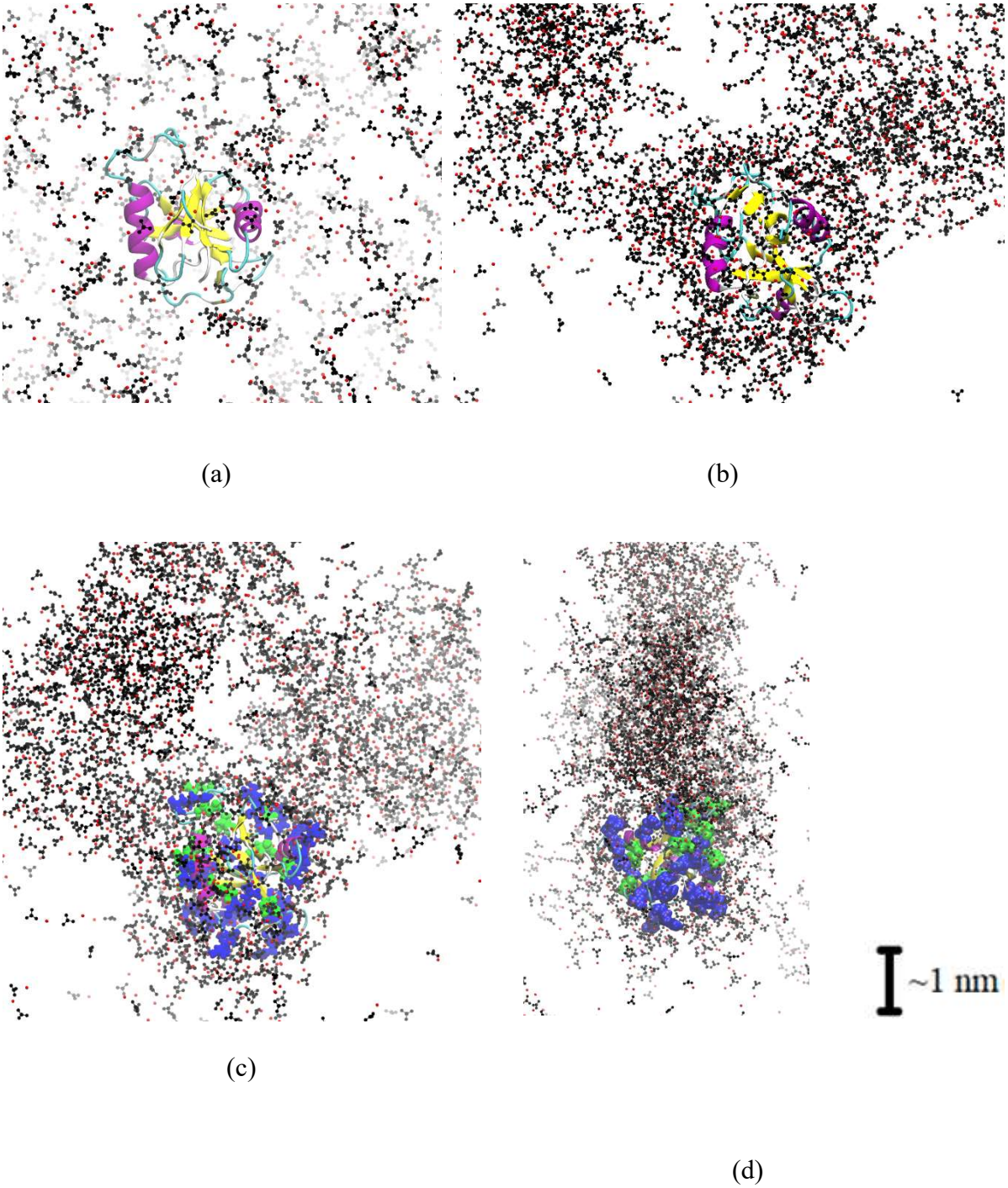



Figure 4.11 Snapshot of the $CaCO_3$ system with protein OC-17, $Ca^{2+}$ in red, $CO_3^{2-}$ in black and the protein coloured by its secondary structure using VMD, after (a) 1 ns of equilibration, (b) 100 ns of simulation, (c) 100 ns of simulation with residues visible in blue (arginine) and green (aspartate and glutamate) and (d) 90° rotation of (c). For visibility, no water molecules are shown. The scale bar for all four snapshots can be found in the lower right corner

### 4.3.5 Cluster crystallinity

In the presence of biomolecules, the clusters formed ($ACC_{bio}$) all showed very similar internal organization of $CaCO_3$, independent of which biomolecule was present (Figure 4.12). The pure cluster ($ACC_{pure}$) showed distinct differences from all $ACC_{bio}$ (RDF, Figure 4.12). So, here we compare internal structures of $ACC_{bio}$ with $ACC_{pure}$ and with pure crystalline polymorphs of $CaCO_3$, calcite, aragonite, vaterite and ikaite.

The RDFs for the crystalline polymorphs and the ACC clusters formed revealed similarities in the direct contact distances (Figure 4.12, SI Figure S4.19), in particular at the shortest distances. The Ca – $O_c$ coordination distance was identical for all structures, while the split peak at the shortest distance between Ca – $C_c$ in our ACC was most comparable to ikaite, the polymorph with crystalline water in its structure. Also, the long-range Ca– Ca, most of Ca – $C_c$ and Ca– $O_c$ RDFs match with ikaite. With increasing distance, less structuring is observed in our ACC compared to the crystalline materials. For example, the peak positions in Ca – Ca are slightly shifted and not all peaks are present. While the $C_c$ – $C_c$ RDF is characterised by an absence of any clearly defined peak in our ACC clusters, the presence of a small increase in intensity around 0.3, 0.47 and 0.6 nm of $ACC_{pure}$ followed the peak pattern of aragonite. The $C_c$– $C_c$ $ACC_{bio}$ showed again the highest resemblance with the RDF of ikaite.

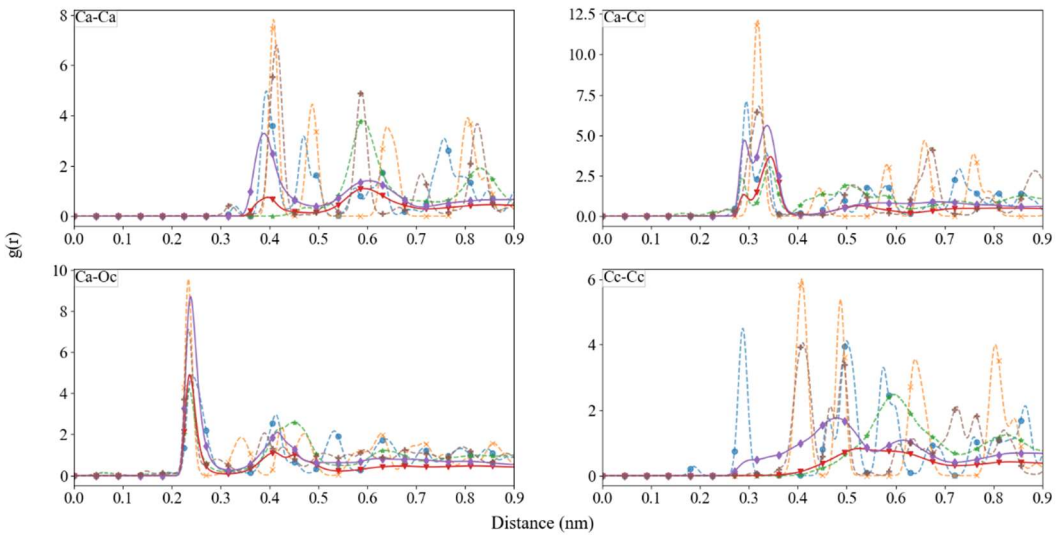



Figure 4.12 Radial distribution function between $Ca^{2+}$, $C_c$ and $O_c$. $ACC_{pure}$ (purple; ♦), $ACC_{bio}$ (red; ▼), calcite (orange-dashed; ×), aragonite (blue-dashed; ●), vaterite (brown-dashed; □), ikaite (green-dashed; ★)



### 4.3.6 Zeta potential measurements

The $CaCO_3$ growth solutions showed reproducible results; initially a negative and positive populations were measured, which alternately represent the largest population in the system. These merged into a population of particles carrying slightly negative ζ-potential (Figure 4.13a). Most likely, these reflect the $CO_3^{2-}$ and $Ca^{2+}$ species that aggregated into $CaCO_3$ particles. The growth solution with L-Asp was less reproducible during the first six timesteps, however after this initial stabilization period the ζ-potential converged, the average of timestep eight until fifteen had a negative potential of -13.06 ± 2.08 mV (Figure 4.13b). This was more negative compared to the solution without L-Asp (-5.58 ± 2.76 mV).

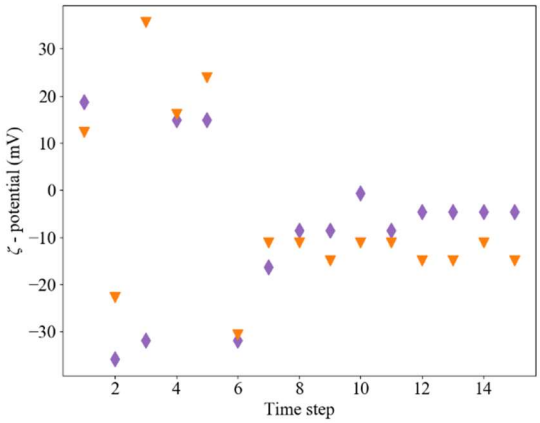



Figure 4.13 ζ-potential over time for the systems without (purple; ♦) and with L-Asp (orange; ▼).



## 4.4 Discussion

In order to contribute to the debate on the significance of $Ca^{2+}$ binding effects of biomolecules (Allouche et al., 1999; Bulo et al., 2007; Jiang et al., 2017; Morse et al., 2007; Ning et al., 2019; Pipich et al., 2008; Teng and Dove, 1997) versus (Freeman et al., 2015; Innocenti Malini et al., 2017; Rani and Saharay, 2019). We will discuss our findings regarding the role of different biomolecules in the first steps of crystallization of $CaCO_3$, starting from dissolved ions.

In the early stages of the simulations, the randomly distributed dissolved ions paired almost instantly (Table 4.1), which is expected considering the high concentration (0.89 M) chosen to see cluster growth within MD-accessible timescales. This observation is in agreement with experiments using a slightly lower concentration of 0.50 M $CaCO_3$ (Bots et al., 2012) and even lower concentrations (0.05 M) (Rieger et al., 2007; Wolf et al., 2008) and previous MD simulations (0.53 M) (Tribello et al., 2009). Because of this very rapid process, the pairing of the ions happened before they interacted with the biomolecule.

In the next step, the clustering continues beyond ion-pairing by aggregation of ion pairs into small clusters and, for the simulations with biomolecules present, the full incorporation of the biomolecule into the largest cluster. Further growth of the clusters is substantially slower. On

the one hand, this growth phase is slower because the number of ions, ion pairs and clusters free in solution decreases and, consequently, the agglomeration frequency goes down. On the other hand, the lower diffusivity of the larger clusters slows down further growth and makes the agglomeration process and formation of ACC diffusion limited (Martin et al., 2006).

In all systems under investigation, all biomolecules were surrounded by ions and ion pairs early in the simulations, and these ions ultimately formed the largest cluster in the system. Strikingly, the density of the cluster surrounding the functional groups was found to be higher in the supercells. Most likely, this is due to the presence of more than one biomolecule in the vicinity of the clusters, and the fact that they act as agglomeration seeds.



### 4.4.1 Biomolecule-ion interactions

The aspartic-acid-containing biomolecules have two distinctive functional groups to be involved with the biomineralization process, where the hydrogen from the amine group attracts the carbonate, and the oxygen of the carboxyl group was suggested to be a $Ca^{2+}$ trap (Smeets et al., 2015). As aspartic acid is thought to associate $Ca^{2+}$ in order to regulate the formation of a crystalline face, our results did not confirm any tendency of the aspartic acid or aspartic acid-based biomolecules to bind $Ca^{2+}$. On the contrary, $CO_3^{2+}$ ions approached first and closer to all functional groups (amine and carboxyl) of the biomolecule than $Ca^{2+}$. Our results are consistent with MD simulations of aspartic acid interacting with ACC via the amine group in the amino acid (Innocenti Malini et al., 2017) (although they used a different Ca-$O_{carboxylate}$ interaction) and the role of aspartic acid observed in Arg-rich proteins (Freeman et al., 2015; Rani and Saharay, 2019). However, other research obtained different results regarding the approach of $Ca^{2+}$ to the biomolecule, when using a more pronounced interaction between Ca and $O_{carboxylate}$ (from CHARMM (Raiteri et al., 2012)). The difference could be ascribed to either the choice of force field and/or the difference in configuration (e.g. the concentration of dissolved ions).

Density functional theory (DFT) with implicit water showed a distance between the Ca and $O_{carboxylate}$ of 0.245 nm (Tang and Skibsted, 2016), although in this case the important contribution of the hydrogen bond network was not taken into account, due to the use of implicit water (Iyemperumal and Deskins, 2017). X-ray experiments showed distances of ~0.241 - 0.254 nm (Marchand and Roux, 1998). However, it must be noted that this distance was only applicable when $Ca^{2+}$ is fully incorporated in the folded protein structure (Marchand and Roux, 1998). Other experimental work showed the complexation of Ca with different proteins and single aspartic acids, for example by using flow dialysis (using calmodulin protein) (Haiech et al., 1981; Maune et al., 1992), ion selective electrode (ISE) (Vavrusova and Skibsted, 2013) and titration (Picker et al., 2012). Yet, only a weak interaction between the aspartic acid and Ca was observed with Nuclear Magnetic Resonance (NMR), which was confirmed by a lack of impact of calcium in the Raman and infrared (IR) spectra for aspartic acid (Maeda et al., 1990). The association energy of Ca and $O_{carboxylate}$ in aspartic acid, -11.6 kJ/mol (based on the Ca — Asp metal ligand formation constant of $10^{1.96}$ (Covington and Danish, 2009)) or slightly smaller (Blaquiere and Berthon, 1987; Maeda et al., 1990) is within the range of the energy of a hydrogen bond and is therefore expectedly very unstable (Maeda et al., 1990). Aside from the similarities between the biomolecule with respect to their ability to serve as a $Ca^{2+}$ trap and their differences with the pure system, differences between different aspartic acid-based biomolecule are more subtle. For example, the distances and quantities between $Ca^{2+}$ and/or $CO_3^{2-}$ and water are extremely similar (Figure 4.8).

### 4.4.2 Modification of cluster hydration and composition by the biomolecules

During the agglomeration of the clusters, biomolecules are playing a significant role that leads to different agglomeration rates and different final configurations compared to the pure ACC systems. Over the simulation period, the cluster size distribution changed from ion pairs and small clusters into larger nanoparticles. In the pure system, a more stable cluster size over time

is observed (narrower and more constant size distribution, Figure 4.3), implying less change in cluster size via addition or removal of ions, ion pairs or small clusters. Moreover, the higher abundancy of intermediate cluster sizes in the pure system, and the stepped increase in the cumulated probability plots (Figure 4.2 and Figure 4.4) suggests the agglomeration of clusters may be the dominant growth mechanism.

The systems with biomolecules showed far less defined cluster sizes with time (more diffuse lines in Figure 4.1 and Figure 4.3), with more variability in numbers of ions associated with the cluster(s). There is also a variability in density of these clusters with time, most likely due to variable water content (Figure 4.6), possibly causing some atoms to move too far from the cluster to be considered bound to a cluster (according to the criteria in our script). Furthermore, the lack of intermediate cluster sizes in the biomolecule systems might imply that addition of ion pairs is the dominant growth mechanism, rather than growth via agglomeration. This is most clearly illustrated by the cumulative probability plots (Figure 4.2 and Figure 4.4), where the $ACC_{bio}$ clusters grow more smoothly. A reason for this is that the biomolecules attract the ions in such a way it inhibits the formation of intermediate sized clusters, therefore, the growth mechanism by ion addition is an indirect effect of this. In short, the biomolecules appear to guide cluster formation and growth via ion (pair) additions rather than agglomeration of intermediate-sized clusters.



All large clusters formed in our simulations showed a slight decrease in $R_g$ with time and with temporary increases in $R_g$ for some particles (Figure 4.5). This overall decrease in $R_g$ with time, while the number of ions in these clusters is constant, implies either cluster compaction or dehydration. Both of these processes are supported by an increase in particle density with time (Table 4.4 and Table 4.5) and the shift and change of intensities of the peaks in the RDF. The decrease of SASA for all clusters with time (Figure 4.6), supports particle dehydration, in particular for the $ACC_{pure}$. This continuous decreasing $R_g$ over time is contrary to dynamically

ordered liquid-like oxyanion polymers, or DOLLOPs, published in other studies where they reported back and forth changes in the $R_g$ over time (Demichelis et al., 2011; Raiteri et al., 2012). DOLLOPs are clusters that are remarkable flexible and can change their shape easily and reversibly, due to the flat energy landscape they maintain; they therefore behave like liquid droplets. We compared a cluster of 72 ions with one of the same size observed before (Demichelis et al., 2011). Typically, the $R_g$ of DOLLOPs with 72 ions are in the range of 0.9 to 1.2 nm. In our simulations, the $R_g$ of such clusters is initially larger than the $R_g$ of a DOLLOP (Figure 4.5), partially explained by the variable $Ca:CO_3$ stoichiometry in the clusters with 72 ions observed in our system. The $R_g$ decreases during the simulation due to restructuring and dehydration. The $R_g$ values observed here overlap with the range of other DOLLOPs observed before (Demichelis et al., 2011). However, our cluster never reached the $R_g$ of wet ACC in our simulation time, following the trend observed by (Demichelis et al., 2011). The formed DOLLOP needs to overcome a larger energy barrier to convert into wet ACC ($R_g$ 0.75 – 0.88 nm (Demichelis et al., 2011)) than the energy barrier to add another ion to the cluster. Given the high ionic concentration in the current study, the clusters continue to grow in the timescales of our simulations, rather than convert completely to wet ACC with lower $R_{g\ values}$.

In addition, the lower weighted RSD of the $R_g$ for our $ACC_{bio}$ compared to $ACC_{pure}$ in the small cell, combined with the much stronger decrease in SASA for $ACC_{pure}$, indicates that the biomolecules inhibit compaction and dehydration of the clusters. This is consistent with, in particular, the stronger association of L-asp with more solvated clusters (Raiteri et al., 2012). Most likely, the presence of the biomolecule(s) affects cluster properties by inhibiting dehydration, potentially by creating an extra energy barrier that could be overcome at a later stage in the biomineralization process, for example by (catalytic) reactions (Addadi et al., 2003; Krajewska, 2018; Murai et al., 2013) provided by the complex biomatrix. Alternatively, the biomolecules may lower the (surface) energy of the hydrated clusters, thereby deepening the

local energy minimum for cluster metastability and increasing the energy barrier (relatively) that needs to be overcome for cluster dehydration.

Our results are in agreement with previous work that showed pure $CaCO_3$ systems show less stabilization of water incorporated ACC compared to biomolecule-rich systems (Addadi et al., 2003). However, the observation of the inclusion of water in pure ACC clusters is contradicting both experiments (Gebauer et al., 2008; Pouget et al., 2009; Wolf et al., 2008) and thermodynamics (Tribello et al., 2009), and is probably mainly due to the high supersaturation in our simulations, the fast cluster formation and the lower water kinetics around the calcium (although, see also (Koskamp et al., 2019)). This brings our observations in the regime of spinodal decomposition, where the cluster grows continuously as the system evolves via a monotonically decreasing free energy pathway (Wallace et al., 2013). This implies that our systems did not achieve an equilibrium state during our simulation time of 100 ns.



The biomolecules also affected the stoichiometry of the clusters formed. Clusters in the pure system showed stoichiometric or slightly Ca-excess compositions, similar to both experiments (Gebauer et al., 2008) and MD simulations (Tribello et al., 2009). In contrast, clusters formed in the presence of biomolecules were Ca-limited, resulting in negatively charged $CaCO_3$ clusters. This last result is counter intuitive as the aspartic acid-based biomolecules are overall negatively charged and, aspartic acid has been suggested to be a $Ca^{2+}$ trap, due to its acidity (Borukhin et al., 2012; Haiech et al., 1981; Marchand and Roux, 1998; Tang and Skibsted, 2016). Moreover, biomineralization, in corals for example, often happens in elevated pH (Brownlee et al., 2021; De Nooijer et al., 2014; Holcomb et al., n.d.), the pH of the biomolecule in our simulations is slightly higher (>10.5 versus ~9.3) compared to the pH of calcifying fluids within the coral compartments when in the light (Al-Horani et al., 2003). At this pH, the carboxyl group of aspartic acid is negatively charged, and the Ca coordinated directly to $O_c$ without the interfering protons. However, ζ-potential experiments of $CaCO_3$ clustering with a

0.67 mM L-Asp confirmed the MD results by indicating the development of (more) negatively charged clusters in the presence of L-Asp than in the pure systems. It is worth mentioning that the ζ-potential measurements were conducted at pH 8.7. At this pH, calcite is known to have a ζ-potential of around -8 mV (Cicerone et al., 1992; Wolthers et al., 2008). In our pure experiments, we record a ζ-potential of around -5.6 ±2.8 mV, while the experiment with L-Asp generates particles with an average ζ-potential of around -13.1±2.1 mV. This is in agreement with the more negative ζ-potential values observed when large L-aspartic acid polymers are present (Zou et al., 2017). Note also that the $CaCO_3$ concentration is much lower in the experiments compared to the simulations to slow down the precipitation in the ζ-potential cell. At pH 10.5, the pH our MD simulations reflect, stoichiometric calcite will have a more strongly negative ζ-potential (e.g. (Cicerone et al., 1992; Wolthers et al., 2008)) and our simulations suggest that, in the presence of Asp-rich biomolecules, $CaCO_3$ particles will be even more negatively charged. Also, the formation of $Ca^{2+}$ limiting clusters in the presence of Arg-rich OC-17, is in agreement with other protein ζ-potential experiments (Hassani et al., 2013). The measurements in the presence of lysozyme, also Arg-rich, showed the formation of negative $CO_3^{2-}$ -rich double layers favoured by the negative ζ-potential around lysozyme (Hassani et al., 2013). Both our simulations and experimental data indicate the formation of Ca-limited clusters.

### 4.4.3 Impact of aspartic acid on cluster formation rate and shape

In the absence of biomolecules, the cluster shape is altered by the more rapid dehydration of the pure ACC clusters, this faster dehydration leads to a more rugose surface of the final cluster compared to the systems with biomolecule. Other differences in shape and trends in size evolution were observed in the presence of biomolecules. Based on size, the small biomolecules can be divided in two groups, the single aspartic acid and the poly amino acid (i.e. poly-D-Asp and peptide). The significance of this will be explained below. The estimated

growth rates for all small cells were nearly the same, although, single amino acids also affect the clustering process in a different way, delaying the formation of the final cluster in higher extent than all other biomolecules (Figure 4.7 and Figure 4.9). The functional groups in the single aspartic acid can interact more straightforward with ions in their environment, resulting in a clear difference in cluster shape between the flat tabular L-Asp and the rounded tabular D-Asp in the small cells. Contrastingly, the larger biomolecules have a more complex interaction with the surrounding, leading to an ambiguous impact on the shape of the cluster formed.

A factor influencing the shape of the cluster is the dipole moment. For all investigated biomolecules the preserved dipole moment of the biomolecule itself induced a dipole in the solvation shell around the biomolecules, pointing away from the cluster that is forming. At the positive side of the dipole, aggregation of the carbonate on the biomolecule was mainly observed. This suggests that the induced dipole in the solvation shell of the biomolecule affects cluster formation. For the Asp (L and D), the clear orientation of the tabular clusters, either perpendicular or normal to this dipole, suggests that the dipole also affects the shape of the ACC clusters formed. The growth in the direction of the dipole appears inhibited, a dipole parallel to the face of the tabular shape (i.e. in D-Asp) promotes growth normal to that same face. This creates a more uniform growth and therefore it results in a more rounded tabular shape of the final cluster compared to L-Asp. The complexity of the structure of poly-D-aspartic acid and peptide created local alterations of the dipole over the biomolecule. Therefore, no correlation was found between the induced dipole of the solvation water and the shape of the cluster.

Our initial and final growth rate of 39.3 ions/ns and 5.4 ions/ns, respectively, observed in our pure super cell represents 0.017 $mol \cdot L^{-1} \cdot ns^{-1}$ and 0.002 $mol \cdot L^{-1} \cdot ns^{-1}$, respectively. In experiments with concentration of 0.5 $mol \cdot L^{-1}$ dissolved $CaCO_3$ an initial growth of minimal $7.083 \cdot 10^{-12}$ $mol \cdot L^{-1} \cdot ns^{-1}$ was observed (Bots et al., 2012). Note that this value is the minimum

rate since the first measured point is after one minute and this approximation assumes that the free ions mainly contribute to the growth of the largest cluster. Between one and two minute the rate slows down to $1.167 \cdot 10^{-13}$ mol·L$^{-1}$·ns$^{-1}$ whereafter it stabilizes. Our rates are more than 10 orders of magnitude higher and although the almost two times higher initial concentration contributes to this, the main discrepancy between our simulations and experiments is minimum timescale that can be measured.

The shape of the clusters also played a role in the observed cluster growth rate in the super cell. In the initial stages of the simulation many ion pairs and small clusters were formed. Then, when the clusters obtained a larger size, the distance between the free ions and the largest cluster decreased. This resulted in an acceleration in the growth rate that is followed by a decrease in rate most likely due to the decrease in concentration of free ions and therewith the saturation index. This impact of cluster shape on rate depends on biomolecule presence. Firstly, the cluster shape in the pure system is rather elongate and covers almost the whole simulation cell. Because of this stretched shape, free ions, ion pairs and small clusters are all relatively close to the growing cluster, facilitating a consistent high growth rate of this largest cluster. However, due to the narrow diameter of this elongated shape, the attractive force at longer distances remains lower than for the tabular shaped clusters as observed in the systems with biomolecules. This results in a rate that was lower than the highest rates in the other systems, but which could be maintained for a longer time period before all the small ions and clusters were consumed. The L-Asp, poly-D-Asp and peptide showed a similar increase in growth rate with simulation time, although higher rates were observed than in the pure system, the period of fast growth was shorter. Most likely, the bulkier clusters have a larger volume affecting smaller (dissolved) species with their attraction forces. Contrastingly, D-Asp had a constant rate during the simulation, similar to the initial, intermediate rate in the other systems. As described above, the dipole moment influences the shape of the cluster, resulting in a shape of

the clusters that is more rounded in the presence of D-Asp compared to L-Asp. This contracted cluster structure probably led to an insignificant acceleration growth rate of the largest cluster.

### 4.4.4 Self-organization of $CaCO_3$ in the presence of protein OC-17

The protein OC-17 is rich in glutamate and arginine. The side chain of glutamate is similarly acidic as aspartic acid and is therefore expected to have a similar binding affinity for $Ca^{2+}$ and $CO_3^{2-}$, whereas arginine has a basic nitrogen-rich side chain. Due to the location of the arginine groups in the protein with respect to the solution, the arginine is playing a significant role in the clustering and the final cluster shape and, as published before for other arginine-rich proteins (Rani and Saharay, 2019): arginine overrules any smaller impact of the aspartate in the protein. Clustering of one or two layers of $CaCO_3$ occurred predominantly around the arginine groups and in the vicinity of arginine, the final density of ions was highest in the cluster formed (Table S4.3). Additionally, further growth of the cluster is only seen around the α-helices that are rich in arginine even though this is not always the shortest route of a dissolved ion to the $ACC_{bio}$. This clustering around the *α*-helices contrasts with the mechanism thought to be the most favourable: "clamping" by the protein of an existing ACC nanoparticle by the arginine groups present in the outer loops (these are depicted at the bottom of the protein in Figure 4.11) (Freeman et al., 2010). As with the small biomolecules, $CO_3^{2+}$ ions approach first and closer to the side chains of Asp, Arg and Glu (Figure S4.10 until Figure S4.12) in disagreement with experimental findings of $Ca^{2+}$ being trapped by the biomolecule (Borukhin et al., 2012). The clustering rate of the $CaCO_3$ is lower in all systems with biomolecules compared to the pure system, however, a direct comparison between the systems with 8 biomolecules or 1 protein is not possible. A clear inhibition of cluster growth due to the interference of the protein and the inclusion of water lead to a smoother final cluster compared to the pure system. Due to the fact that there are only two α-helices the overall cluster is tabularly shaped.

### 4.4.5 Crystallinity of the clusters

Studies, using cryo-TEM, SEM, and in situ WAXS (Nielsen et al., 2014; Rieger et al., 2007; Wolf et al., 2008), observed a hydrated ACC phase that rapidly converted into crystalline polymorphs under high supersaturation conditions. However, the very short lifetime and overall amorphous structure of transient ACC make a characterization experimentally challenging. Hence, experimental attempts to detect this transient ACC relied on quenching techniques (Gebauer et al., 2014) and studies that concentrated on the description of the stable form of ACC (Addadi et al., 2003; Michel et al., 2008), where the later found a structure that did not resemble any of the known polymorphs (Michel et al., 2008). In the RDFs for our ACC nanoparticles (in agreement with ACC's RDFs observed in previous computational work (Raiteri and Gale, 2010)) (Figure 4.8), $ACC_{pure}$ and $ACC_{bio}$ show resemblance with the ikaite structure. Furthermore, the RDFs for all $ACC_{bio}$ were indistinguishable from each other. Besides, in the less hydrated $ACC_{pure}$, some domains of the cluster resembled aragonite more than calcite or vaterite. Note that, aragonite features have also been found in microdomains in other MD studies (Tribello et al., 2009). While beyond the scope of this study, the difference in crystalline-like microdomains in presence of the different biomolecules is worth looking into, to find a possible link with between the ACC formed and the eventual preferred polymorph of $CaCO_3$.

### 4.4.6 Implications

Aspartic acid-based biomolecules stabilize hydrous ACC, as confirmed in this study. Such an effect of biomolecules could be favourable for biomineralizing organisms, because it prolongs the time the organism has to cast or move and aggregate hydrated clusters to the final site and shape of the biomineral it wants to form (Gong et al., 2012). This effect could be used in several applications, for example to guide the formation of the polymer-induced liquid-precursor (PILP) in order to form $CaCO_3$ films and coatings for e.g. metal wires (Gower and Odom,

2000) and potentially to optimize antiscalant products (Li et al., 2014) as it stabilizes hydrated nanoparticle (temporarily) and inhibits the further crystal growth (Montanari et al., 2016). Besides, it could be the mechanism behind the formation of $CaCO_3$ crystal polymorphs in the presence or with the assistance of bacteria is found in nature (Wei et al., 2015) and has also industrial applications in bio-cemented soils (Yu et al., 2020). The subjected bacteria commonly contain/produce L-aspartic acid and as experiments already revealed changes in crystal shape, the role of this amino acid helps in elucidating the mechanism in microbial induced calcite precipitation (Braissant et al., 2003). Once the mechanism is known, one can mimic and design organic molecules with the required set of functional groups that facilitate controlling any crystallization pathway of interest (Deng et al., n.d.). Via this way it might be possible to answer to the increasing need for environmental and energy friendly development of new materials and applications (Gower, 2008). Biogenic applications could then for example be found in the search for sustainable options for soil treatment, reducing the current carbon footprint. For example, in protecting sand dunes against erosion (Ashraf et al., 2017) and cement in constructions (Chang et al., 2016). This knowledge is not limited to $CaCO_3$ since $CaCO_3$ can be seen as an analogue for other systems by extrapolating the gained insights, for example, the crystallization of magnesium carbonate (Harrison et al., 2019; Toroz et al., 2021) or calcium phosphate (Q. Wang et al., 2011).



## 4.5 Conclusion

The impact of aspartic acid-based biomolecules and the aspartic acid containing protein Ovocleidin-17 on the formation of ACC clusters was studied by performing MD simulations of solutions with dissolved calcium and carbonate ions in the absence and presence of biomolecules. Our work provides an insight of the role these biomolecules play upon the formation of a dense amorphous phase via the strong interaction between its functional groups and the dissolved ions and ions in the forming clusters. The main observations were:

- $CaCO_3$ pairs formed rapidly, during the first nanosecond of equilibration time, without any visible impact of the biomolecule
- Initially, all ACC clusters formed (predominantly) by accumulation of ion pairs
- The largest $ACC_{pure}$ clusters grew larger by aggregation of intermediate sized clusters
- The $ACC_{bio}$ clusters grew by ion (pair) addition rather than by cluster aggregation
- $CO_3^{2-}$ was the ion interacting with the biomolecule first and strongest
- The biomolecules were embedded in/remained associated with the $ACC_{bio}$ clusters that became the largest in the system
- $ACC_{bio}$ clusters remained more hydrated than $ACC_{pure}$ over the course of our simulations
- $ACC_{pure}$ cluster shapes were more rugose
- $ACC_{bio}$ cluster shapes were tabular and oriented to the induced dipole in the solvent shell of the functional groups in the biomolecules
- The ACC clusters did not have a distinct crystalline structure but the alikeness with ikaite and longer-range structuring tendency to aragonite suggests the existence of crystalline domains
- $ACC_{pure}$ clusters were either stoichiometric or slightly enriched (0.6%) in calcium
- $ACC_{bio}$ clusters were nonstoichiometric, composed of ~ 4% more carbonate than calcium, resulting in a net negative charge of the clusters
- This net negative charge was confirmed for ACC formed in the presence of biomolecules, using ζ-potential measurements.

- Protein OC-17 clusters the ions predominantly around the arginine, the guanidine group traps the $CO_3^{2-}$ that is paired with $Ca^{2+}$.
- Cluster growth happens around the α-helices that are rich in arginine pointing into the solution, where the water density is lower compared to other superficial arginine.

All in all, aspartic acid-based biomolecules could be useful tools in biomineralization processes to keep ACC hydrated longer.

## 4.6 Supplementary information

### 4.6.1 Structural models

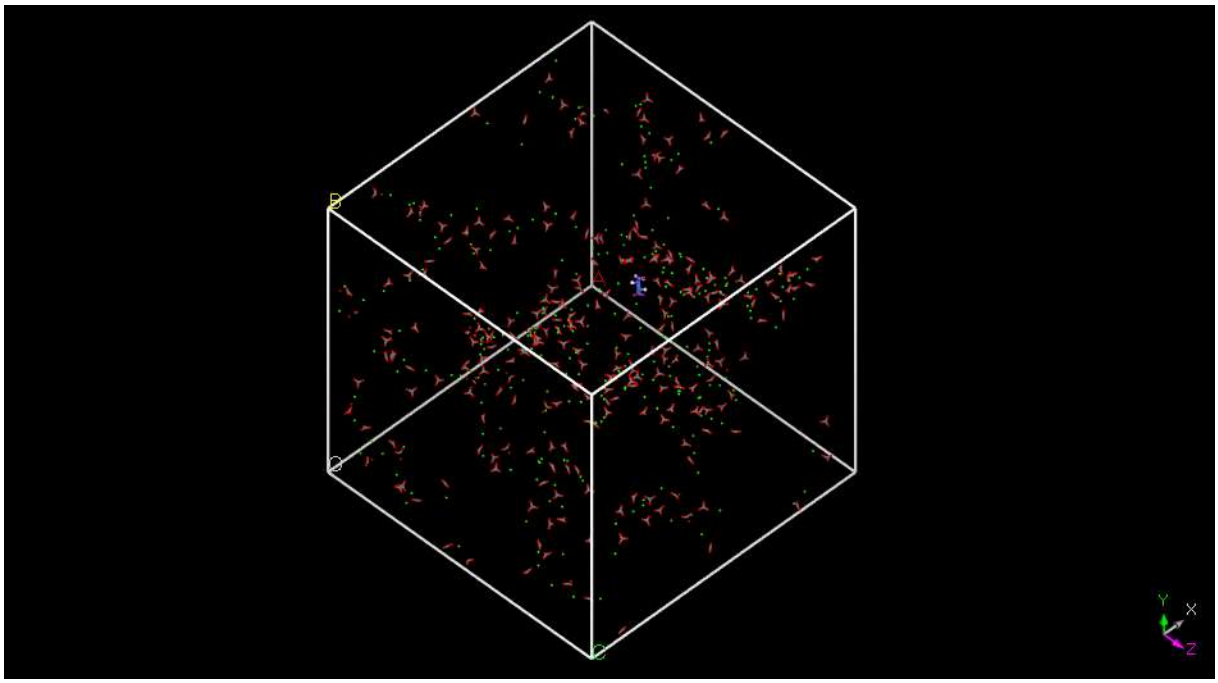
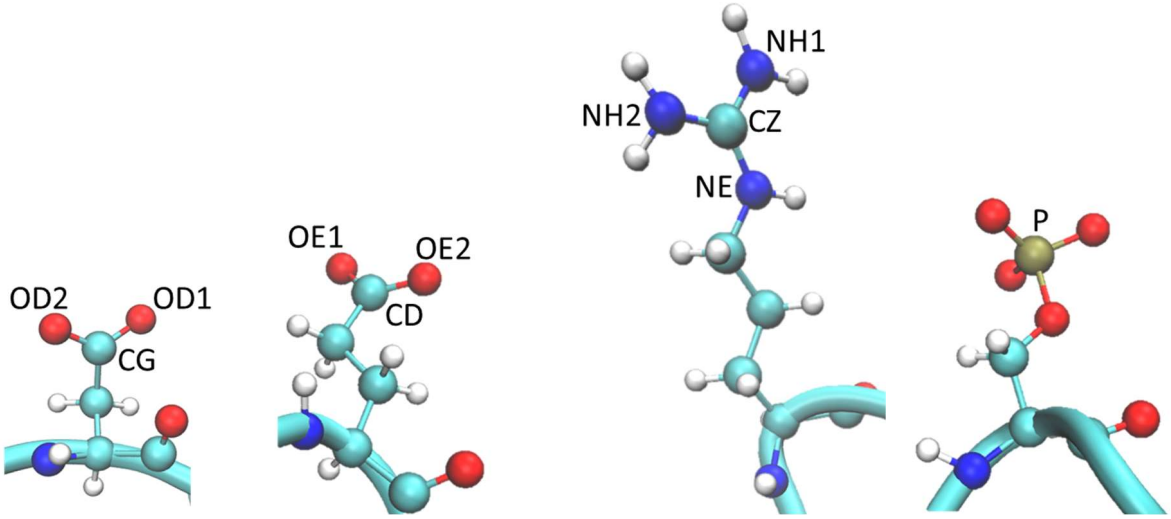


Figure S4.1 Snapshot of the small simulation cell with 255 $Ca^{2+}$ and $CO_3^{2-}$ and an aspartic acid biomolecule, water molecules were left out for visibility. The size of the cell is 7.8 $nm^3$.




Aspartate Glutamate Arginine Serine (phosphorylated)

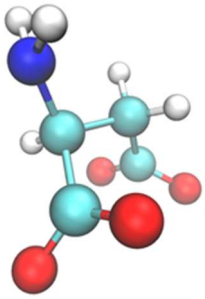
Figure S4.2 Molecular structure of the side chains of aspartate, glutamate, arginine and phosphorylated serine including the atom labels used in the RDF and in the main text for the protein OC-17. The atoms are coloured: Nitrogen (blue), Carbon (cyan), Oxygen (red),

Phosphorus (ochre) and Hydrogen (white). The tube (cyan) represents the backbone of the protein. Images were created using VMD software package.

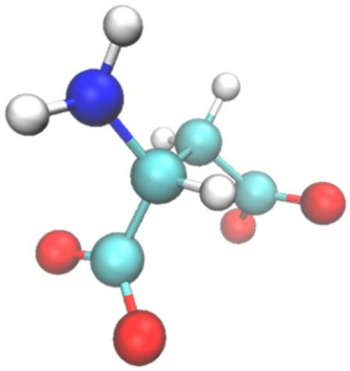
Table S4.1 Overview of all simulation cells

| | SYSTEM | $Ca^{2+}$ | $CO_3^{2-}$ | BIOMOLECULE | $H_2O$ |
|---|---|---|---|---|---|
| **SMALL CELL** | Pure | 255 | 255 | n / a | 15831 |
| | D-Asp | 255 | 255 | 1 D-aspartic acid | 15831 |
| | L-Asp | 255 | 255 | 1 L-aspartic acid | 15831 |
| | Poly-D-Asp | 255 | 255 | 1 Poly-D-aspartic acid | 15831 |
| | Peptide | 255 | 255 | 1 Peptide (DYDSD) | 15831 |
| **SUPER CELL** | Pure | 2040 | 2040 | n / a | 126648 |
| | D-Asp | 2040 | 2040 | 8 D-aspartic acid | 126648 |
| | L-Asp | 2040 | 2040 | 8 L-aspartic acid | 126648 |
| | Poly-D-Asp | 2040 | 2040 | 8 Poly-D-aspartic acid | 126648 |
| | Peptide | 2040 | 2040 | 8 Peptide (DYDSD) | 126648 |
| | Protein-OC-17 | 2040 | 2040 | 1 OC-17 | 126648 |

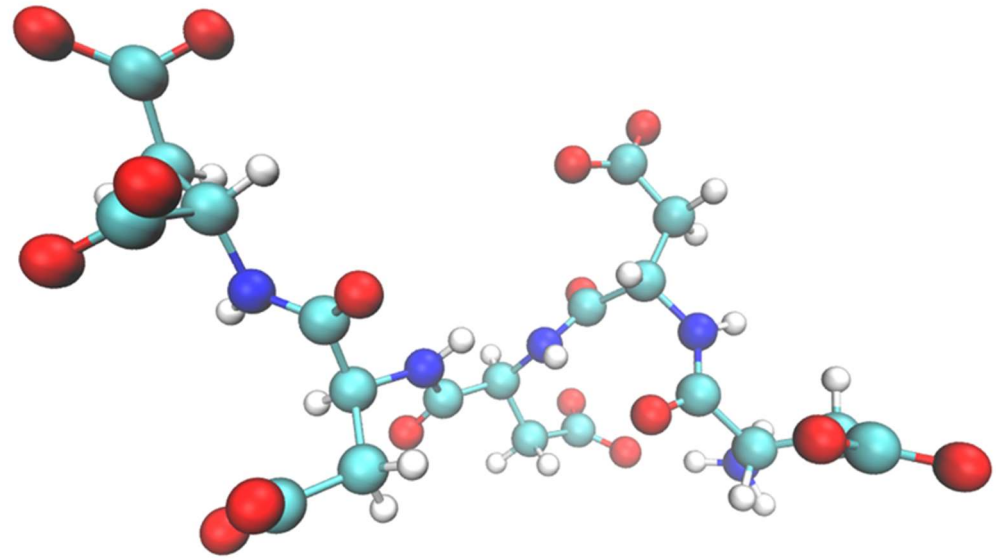




*D-aspartic acid*

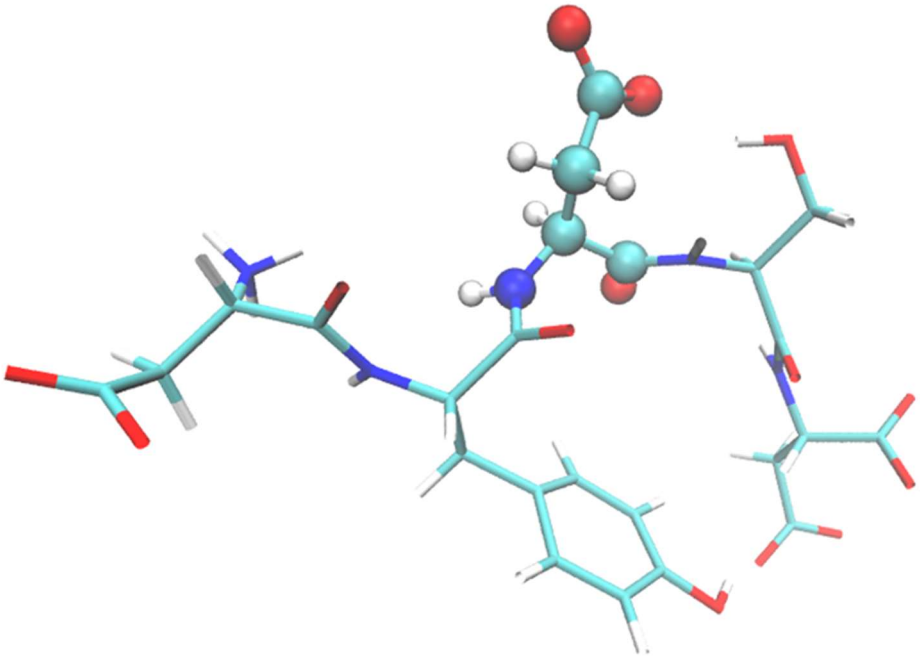


*L-aspartic acid*

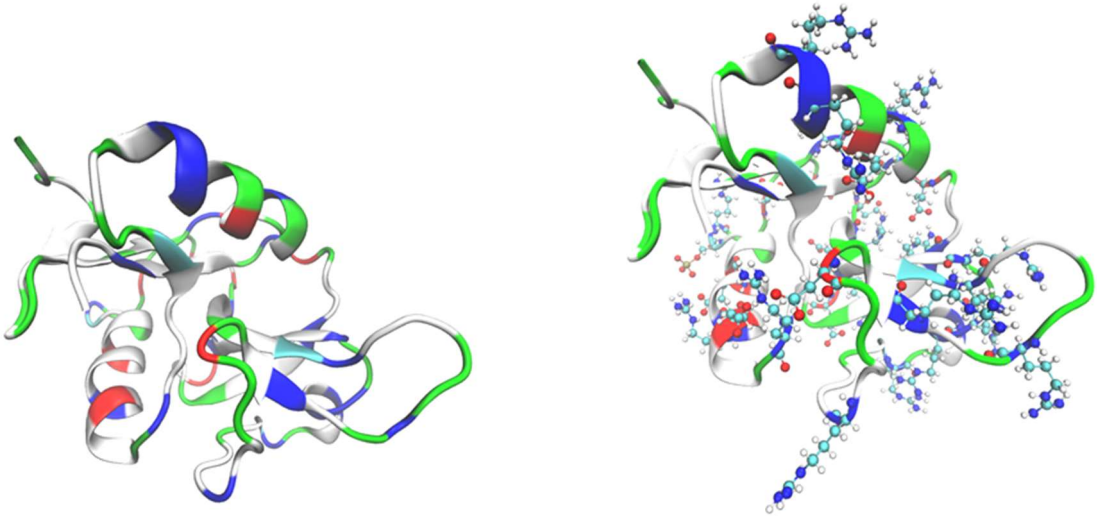


*Poly-d-aspartic acid*





*Peptide*



*Protein OC-17*

Figure S4.3 Molecular structure of the studied biomolecules. Nitrogen (blue), Carbon (cyan), Oxygen (red) and Hydrogen (white). In the structure of peptide, the aspartate is highlighted by visualizing the rest of the peptide in sticks. The protein is shown, both with and without aspartate, glutamate, arginine, and phosphorylated serine side chains. Different regions in the protein are visualized; hydrophobic region (white), acidic region (red), basic region (blue), polar region (green). Images were created using VMD software package.

### 4.6.2 Force field

Table S4.2 Inter and intraatomic potential parameters.

| Water (SPC/fw) | | | | | |
|---|---|---|---|---|---|
| Species | | Molecule/Atom | | | Charge (e) |
| Ca | | Calcium | | | +2.000 |
| $C_c$ | | Carbon from carbonate | | | +1.135 |
| Oc | | Oxygen from carbonate | | | -1.045 |
| $O_W$ | | Oxygen from water | | | -0.820 |
| $H_W$ | | Hydrogen from water | | | +0.410 |
| Bond styles | | $k$ (ev Å$^{-2}$) | | | $r_0$ (Å) |
| $C_c$—$O_c$ | Harmonic | 40.8493 | | | 1.012 |
| $O_W$—$H_W$ | Harmonic | 45.9296231 | | | 1.012 |
| Angle styles | | $k$ (ev rad$^{-2}$) | | | $\Theta_0$ |
| $H_w$— $O_W$ — $H_W$ | Harmonic | 3.29134 | | | 113.24 |
| | | $\Theta_0$ | $K_2$ (ev rad$^{-2}$) | $K_3$ (ev rad$^{-2}$) | $K_4$ (ev rad$^{-2}$) |
| $O_c$— $C_c$ — $O_c$ | class2 | 120.0 | 6.617 | 0.0 | 0.0 |
| | | M (ev Å$^{-2}$) | $r_1$ (Å) | $r_2$ (Å) | |
| $O_c$— $C_c$ — $O_c$ | class2 bb | 12.818 | 1.3042 | 1.3042 | |
| | | $N_1$ (ev Å$^{-2}$) | $N_2$ (ev Å$^{-2}$) | $r_1$ (Å) | $r_2$ (Å) |
| $O_c$— $C_c$ — $O_c$ | class2 ba | 1.53319 | 1.53319 | 1.3042 | 1.3042 |
| Improper styles | | $K_2$ (ev rad$^{-2}$) | | | $K_4$ (ev rad$^{-2}$) |
| $O_c$— $C_c$ — $O_c$-$O_c$ | distance | 13.647 | | | 360.0 |
| Interatomic interactions | | | | | |
| Lennard-Jones Potential | | ε (eV) | | | σ (Å) |
| $O_W$ — $O_W$ | | 0.006739769454 | | | 3.165492 |
| Ca—$O_W$ * | | 0.000950 | | | 3.35 |
| Buckingham Potential | | A (eV) | ρ (Å) | | C (eV Å$^6$) |
| $O_c$—$O_W$ * | | 12534.455133 | 0.202 | | 12.09 |
| $O_c$—$H_W$ * | | 396.0 | 0.217 | | 0.0 |
| Ca—$O_c$* | | 3161.6335 | 0.271511 | | 0.0 |
| Ca—$C_c$ * | | 120000000 | 0.12 | | 0.0 |
| $O_c$— $O_c$* | | 63840.199 | 0.198913 | | 27.89901 |

*The interaction was set to zero over the range 6 – 9 Å

### 4.6.3 Radial distribution functions

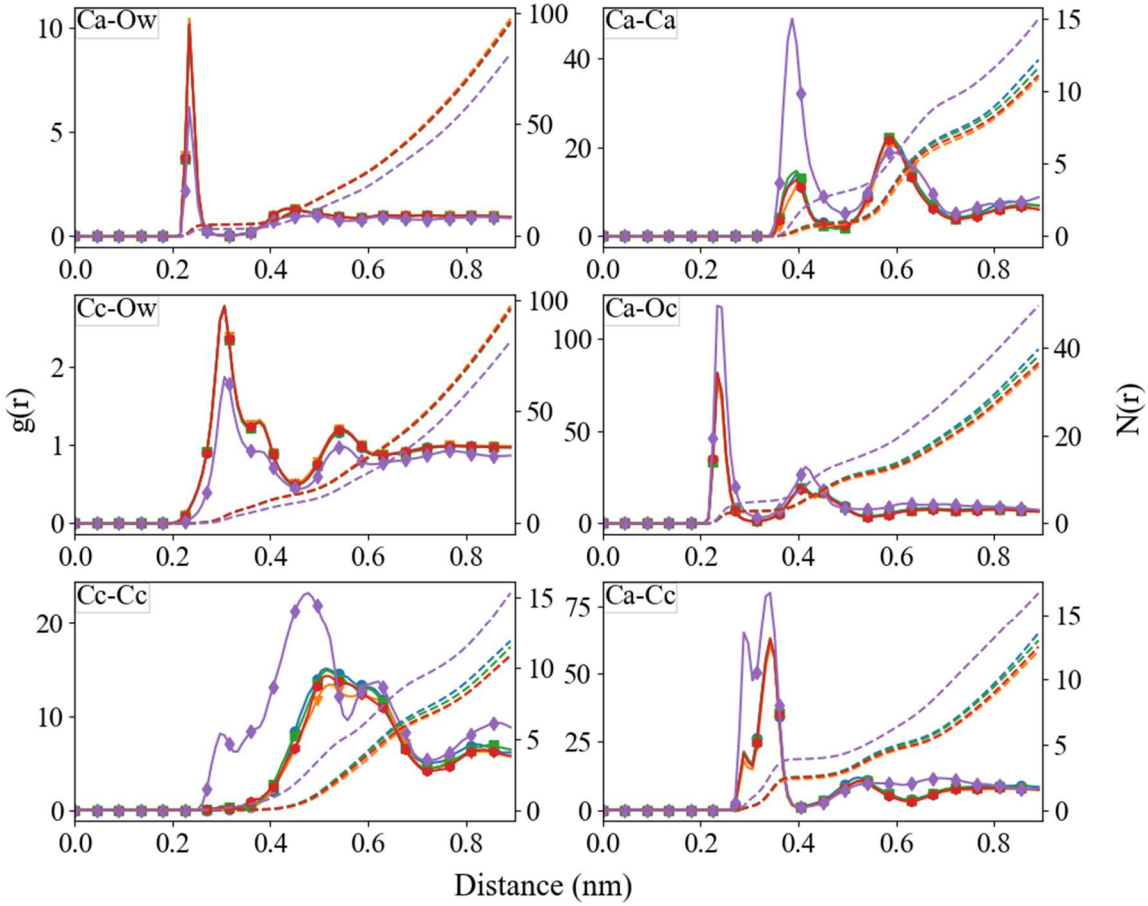



Figure S4.4 RDF (left axis) and corresponding integral (N(r)) (right axis) between $Ca^{2+}$, $C_c$, $O_c$, and $O_w$, after 100 ns of the simulation of the small cell. Pure (purple; ♦), D-Asp (blue; ●), L-Asp (orange; ▼), poly-D-Asp (red; ⬢), peptide (green; ■), protein OC-17 (brown; ★)

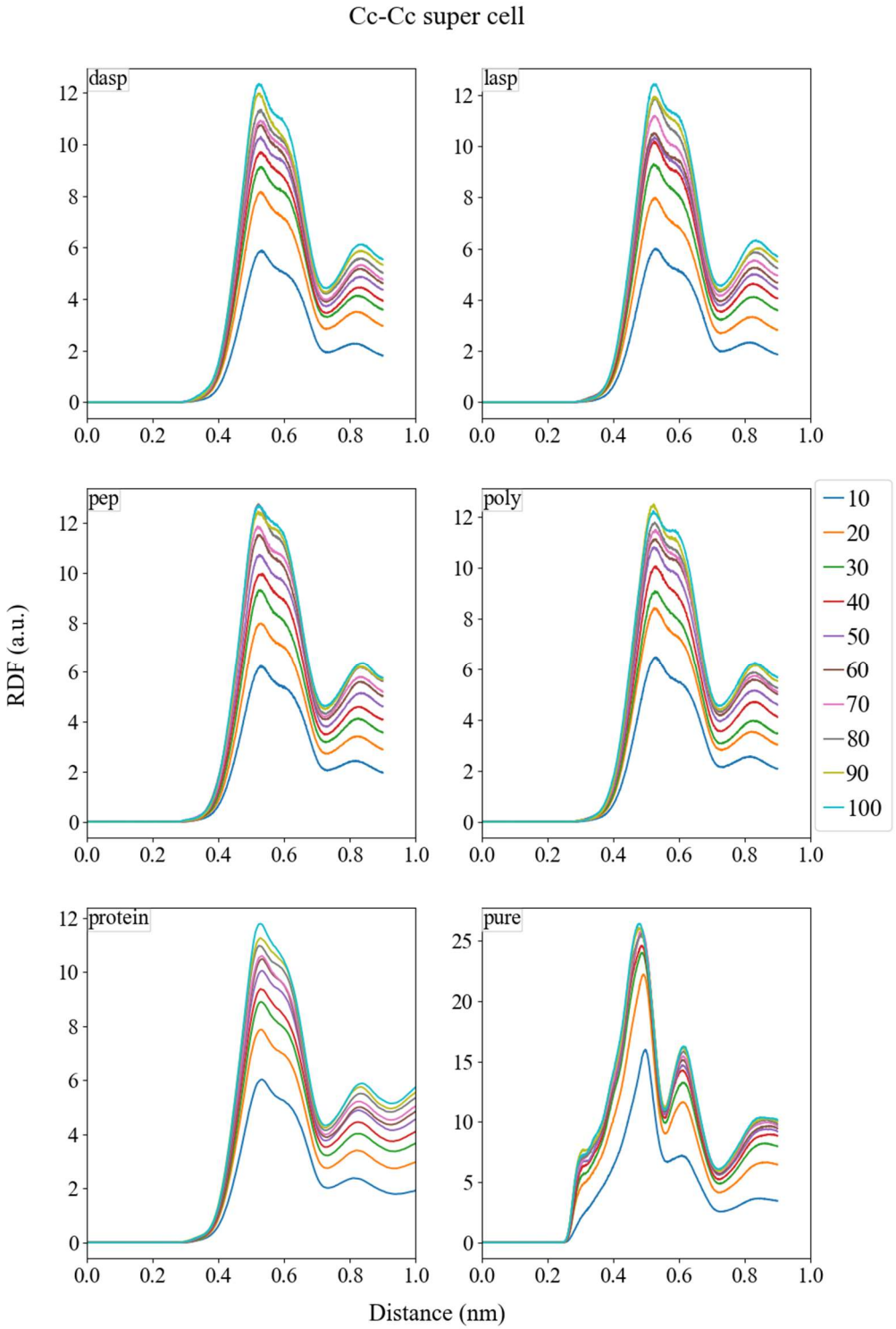



Figure S4.5 RDF between $C_c$ and $C_c$ of every 10 ns in the super cell simulations.

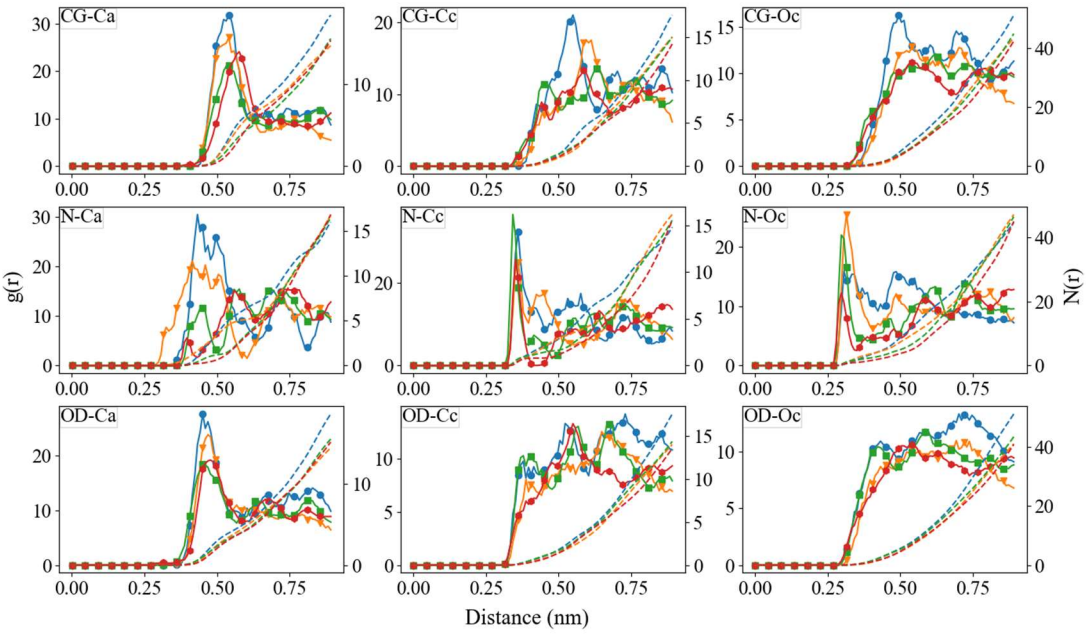



Figure S4.6 RDF between the different functional groups in the aspartic acid part of the biomolecule (see Figure S4.2 for the atom labels) and $Ca^{2+}$, $C_c$, $O_c$, after 100 ns of the simulation of the small cell. D-aspartic acid (blue; ●), L-Asp (orange; ▼), poly-D-Asp (red; ⬢), peptide (green; ■), protein OC-17 (brown; ★).

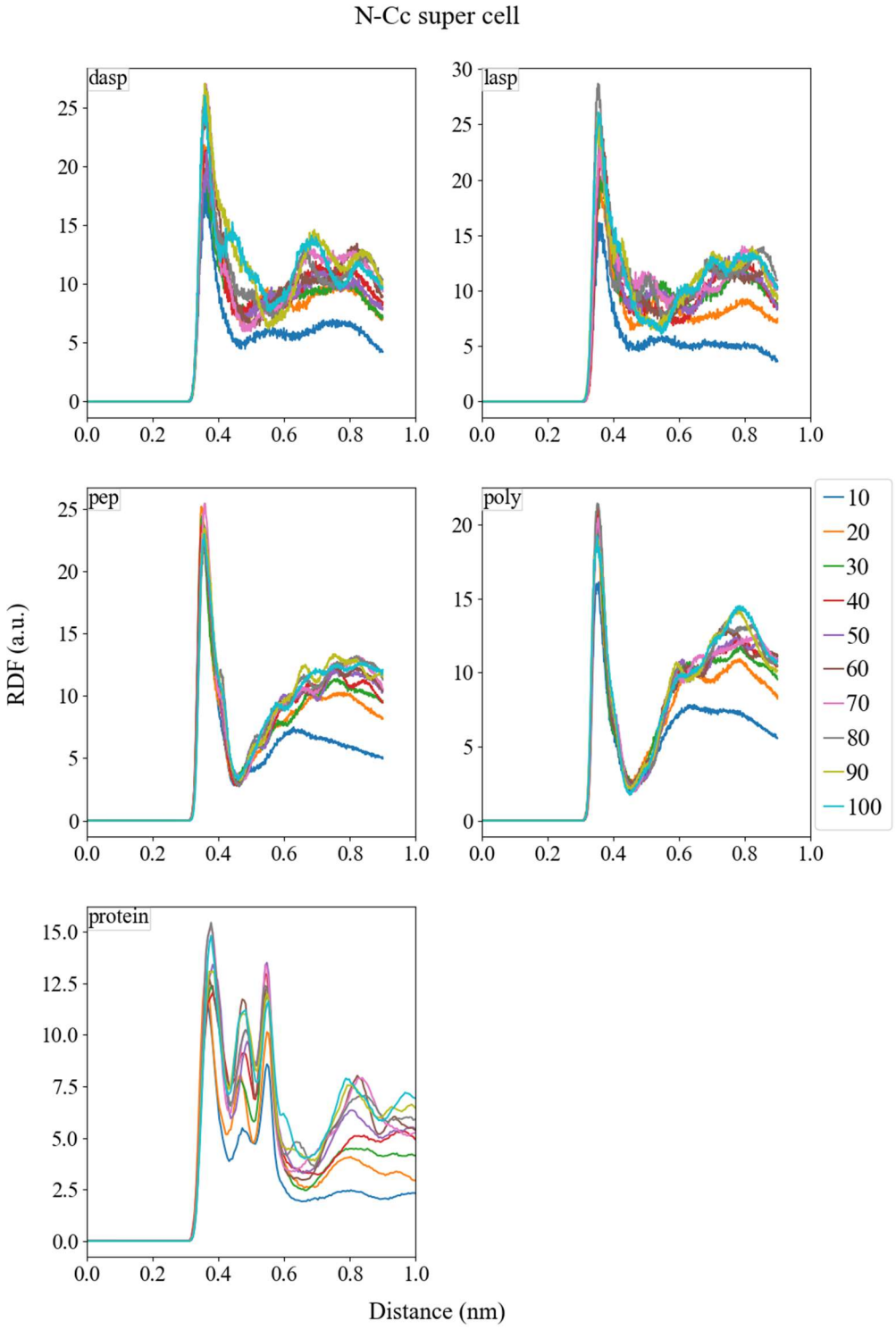



Figure S4.7 RDF between $N_{bio}$ and $C_c$ of every 10 ns in the super cell simulations.

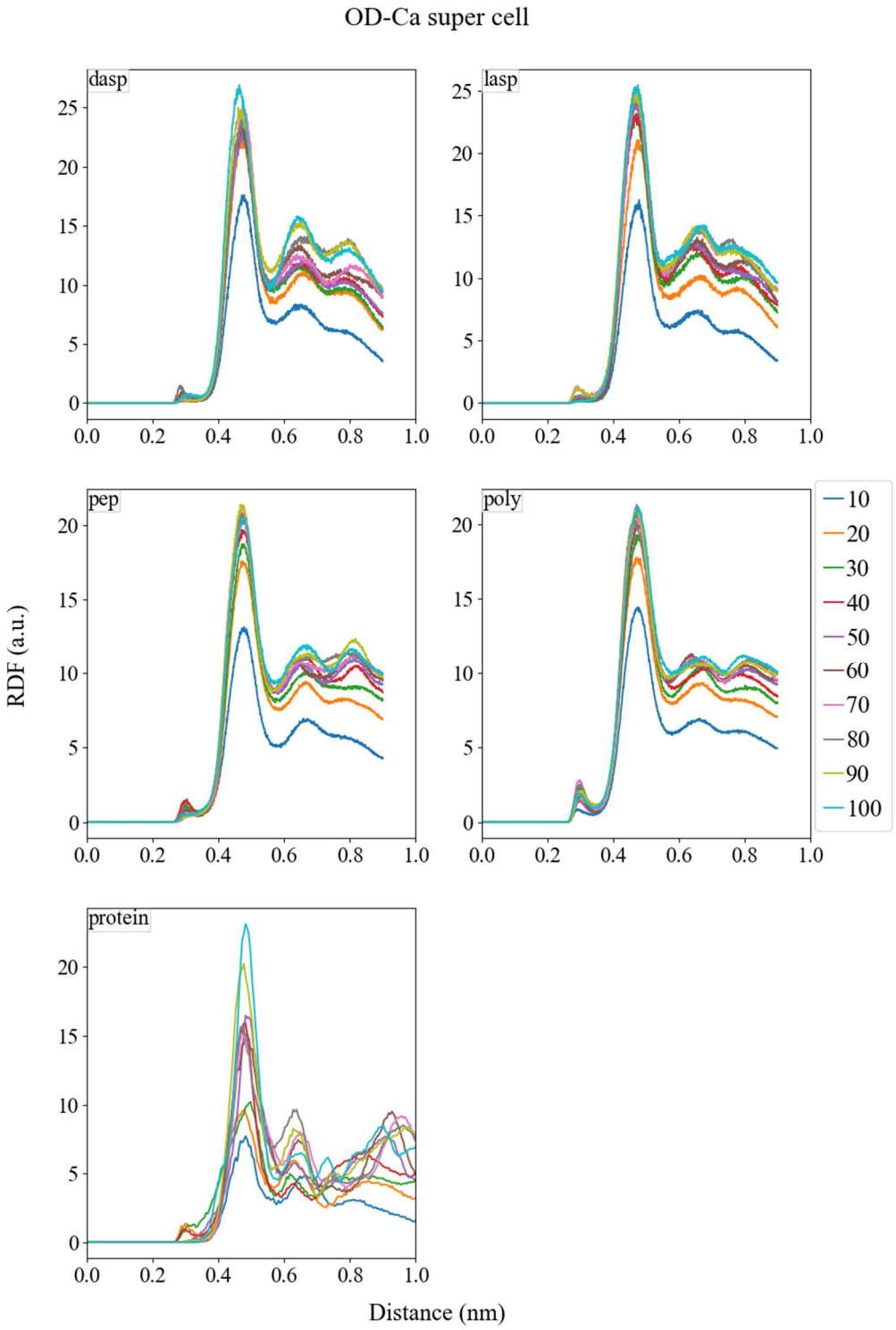



Figure S4.8 RDF between $O_{bio}$ (Oxygen of the carboxyl group(s)) and $Ca^{2+}$ of every 10 ns in the super cell simulations.

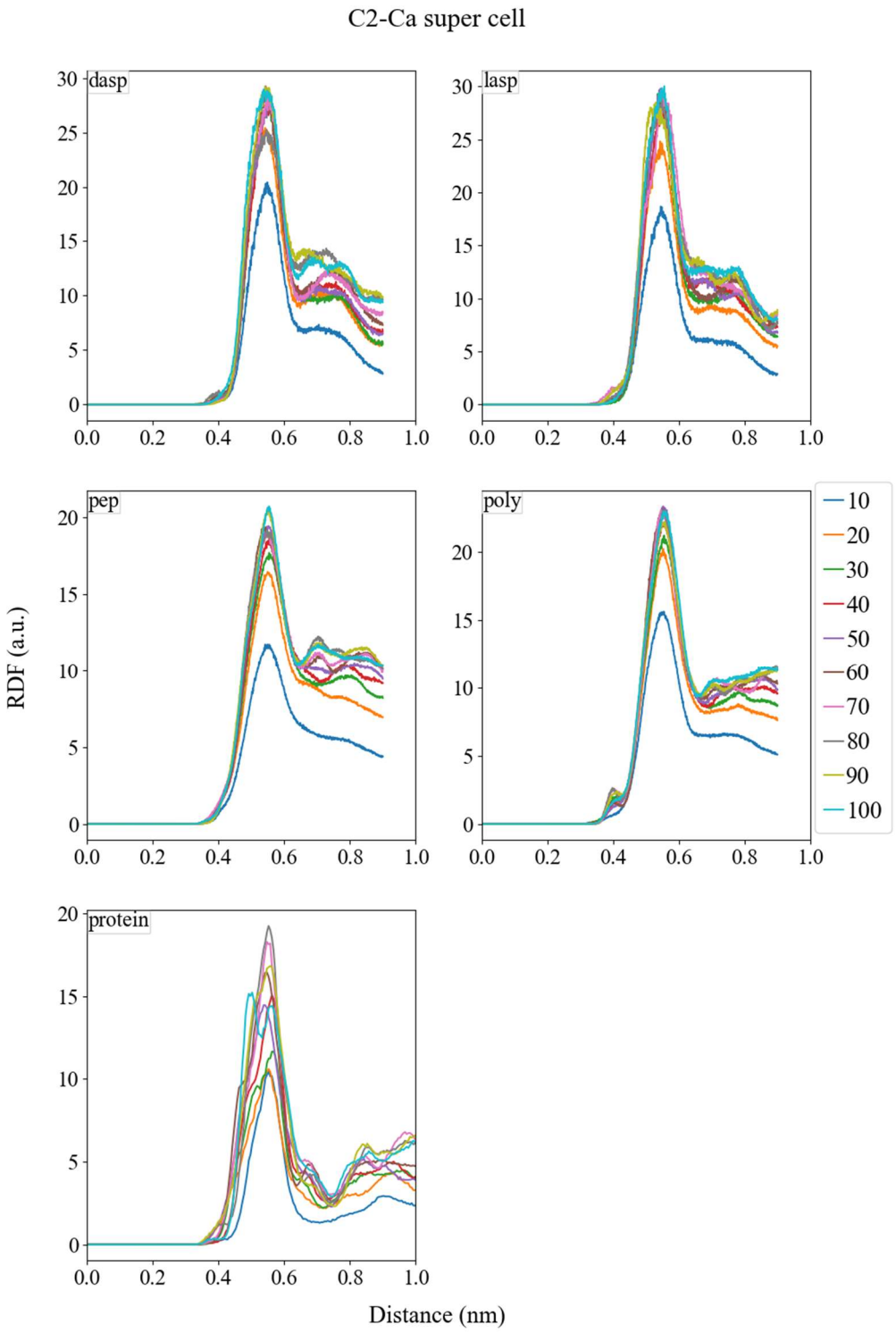



Figure S4.9 RDF between C2 (Carbon of the carboxyl group(s)) and $Ca^{2+}$ of every 10 ns in the super cell simulations.

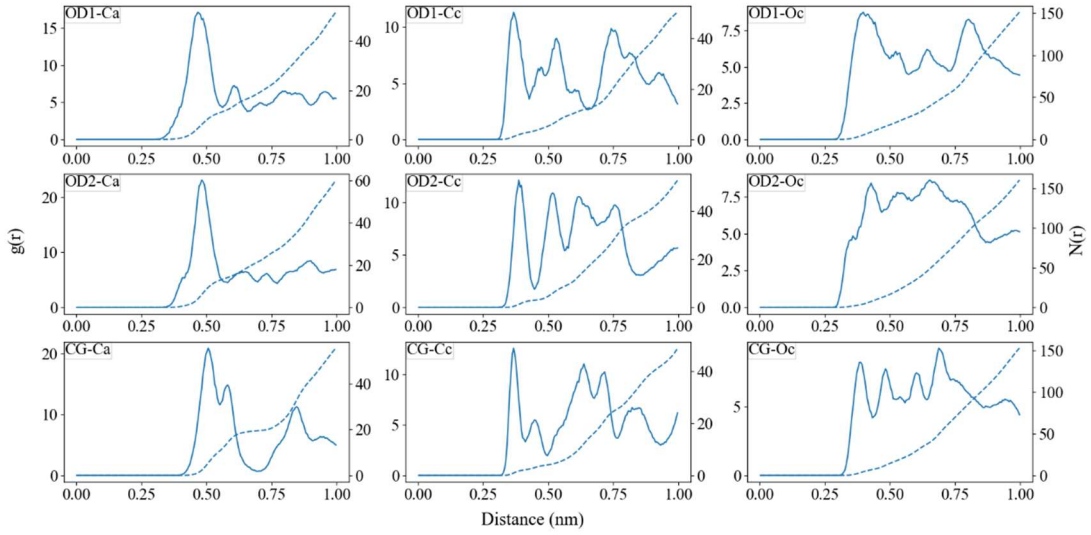



Figure S4.10 RDF of the main atoms (see Figure S4.2 for the atom labels) in the **aspartate** side chain with $Ca^{2+}$, $C_c$ and $O_c$



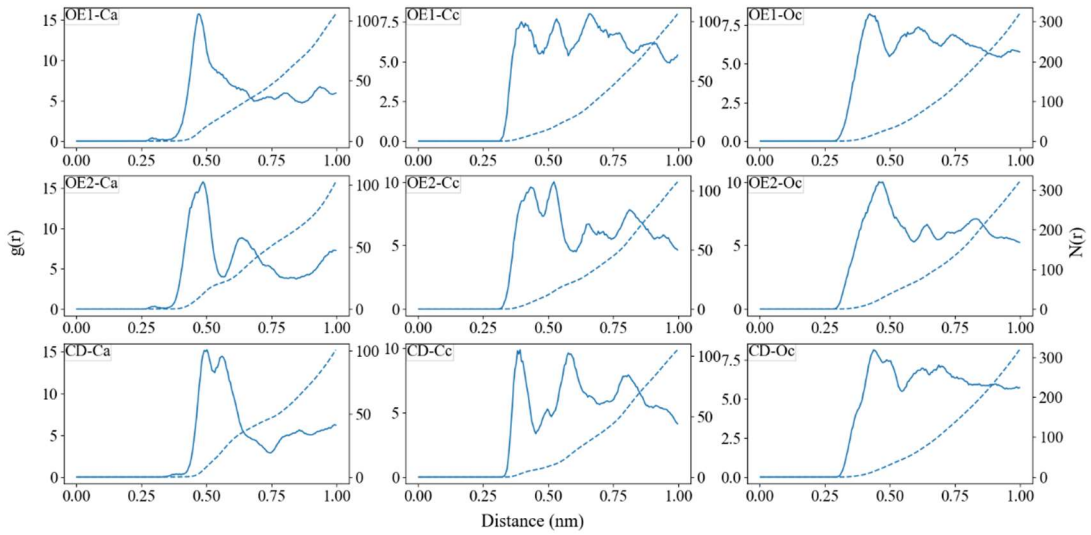



Figure S4.11 RDF of the main atoms (see Figure S4.2 for the atom labels) in the **glutamate** side chain with $Ca^{2+}$, $C_c$ and $O_c$

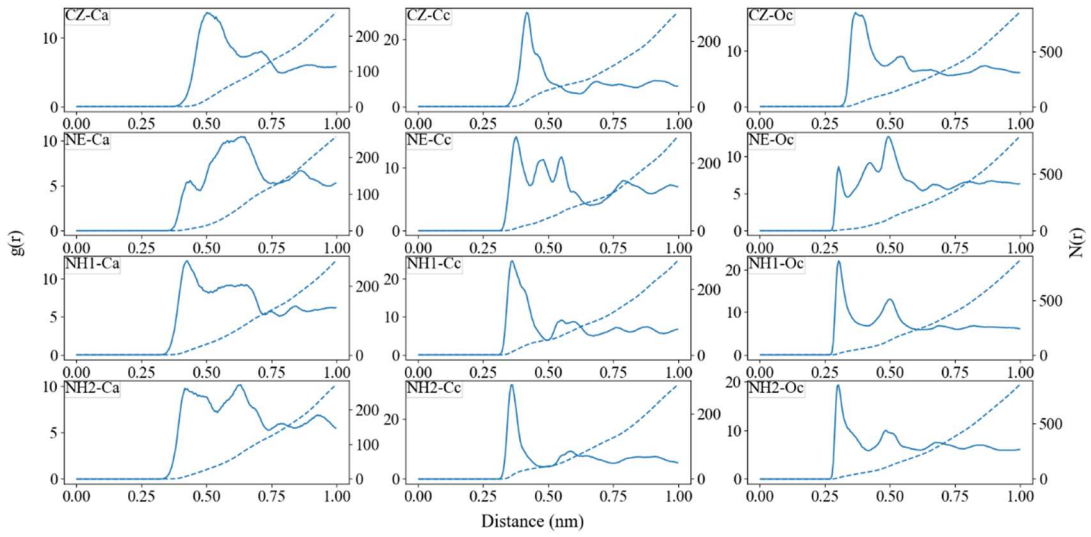



Figure S4.12 RDF of the main atoms (see Figure S4.2 for the atom labels) in the **arginine** side chain with $Ca^{2+}$, $C_c$ and $O_c$

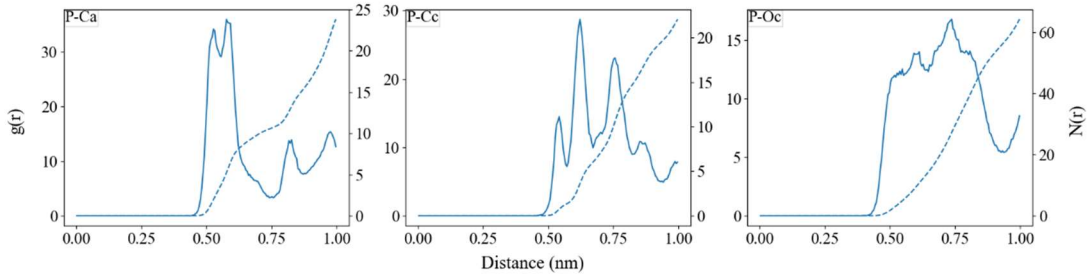
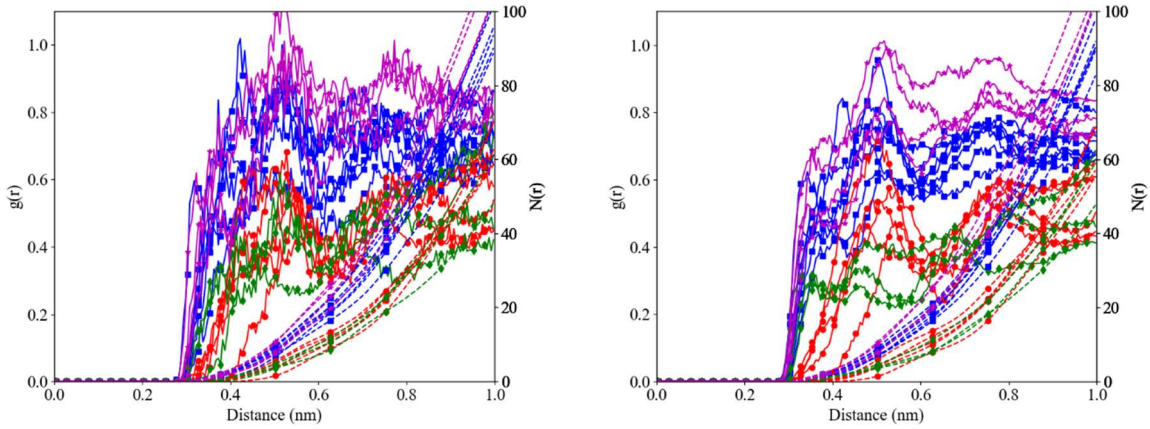



Figure S4.13 RDF of the main atoms (see Figure S4.2 for the atom labels) in the **phosphorylated serine** side chain with $Ca^{2+}$, $C_c$ and $O_c$




(a) (b)

Figure S4.14 RDF between NE (see Figure S4.2 for the atom label) and $O_w$ of different arginine groups in the protein grouped by their position in the protein after before the extended clusters were formed. Arginine groups encapsulated in the protein (red; ●, green; ♦), arginine with their residue pointing in solution (blue; ■) and residues pointing in solution in part of the protein on which the growth of the cluster stopped (purple; ★) (a) after 1 ns of simulation and (b) after 10 ns of simulation.



Table S4.3 Water density around different arginine residues with different residue identities (resid) in the protein at different times in the simulation

| Density (g/cm$^3$) | | | |
|---|---|---|---|
| Resid \ Simulation time (ns) | Label | 1 | 10 |
| 43, 71, 78 | green; ♦ | 0.40 | 0.37 |
| 31, 86, 105, 114, 125 | red; ● | 0.44 | 0.43 |
| 19, 25, 32, 49, 59, 109, 130 | blue; ■ | 0.66 | 0.65 |
| 83, 94, 100, 106 | purple; ★ | 0.79 | 0.77 |

### 4.6.4 Tanimoto index and distance

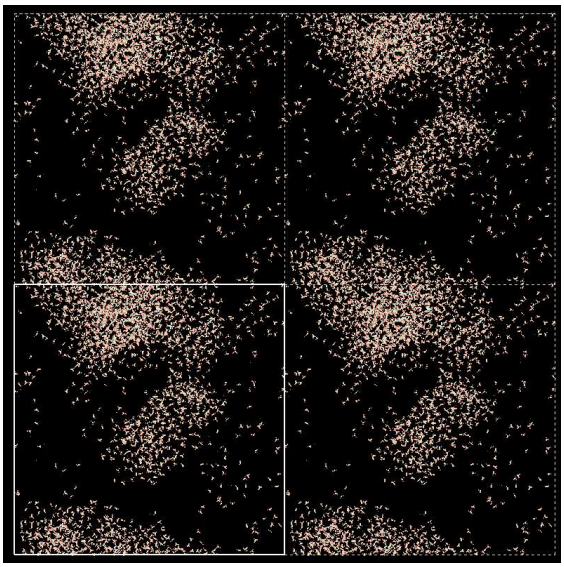
Table S4.4 Tanimoto index and distance of the final cluster between system A and system B.

| System A | System B | Tanimoto index | | Tanimoto distance | |
|---|---|---|---|---|---|
| **$Ca^{2+}$ and $CO_3^{2-}$** | | **Small Cell** | **Super cell** | **Small Cell** | **Super cell** |
| pure | D-Asp | 0.94 | 0.76 | 0.74 | 0.70 |
| | L-Asp | 0.91 | 0.98 | 0.74 | 0.79 |
| | peptide | 0.95 | 0.78 | 0.73 | 0.73 |
| | poly-D-Asp | 0.95 | 0.97 | 0.76 | 0.74 |
| | protein OC-17 | | 0.98 | | 0.75 |
| D-Asp | L-Asp | 0.90 | 0.75 | 0.76 | 0.71 |
| | peptide | 0.93 | 0.96 | 0.76 | 0.76 |
| | poly-D-Asp | 0.93 | 0.74 | 0.75 | 0.68 |
| L-Asp | peptide | 0.91 | 0.77 | 0.75 | 0.74 |
| | poly-D-Asp | 0.91 | 0.98 | 0.78 | 0.74 |
| poly-D-Asp | peptide | 0.94 | 0.75 | 0.75 | 0.74 |
| protein OC-17 | D-Asp | | 0.75 | | 0.73 |
| | L-Asp | | 0.98 | | 0.76 |
| | peptide | | 0.77 | | 0.73 |
| | poly-D-Asp | | 0.98 | | 0.76 |
| **Only $Ca^{2+}$** | | | | | |
| pure | D-Asp | 0.85 | 0.75 | 0.77 | 0.71 |
| | L-Asp | 0.86 | 0.97 | 0.79 | 0.77 |
| | peptide | 0.87 | 0.76 | 0.77 | 0.73 |
| | poly-D-Asp | 0.87 | 0.96 | 0.77 | 0.79 |
| | protein OC-17 | | 0.97 | | 0.76 |
| D-Asp | L-Asp | 0.83 | 0.74 | 0.78 | 0.69 |
| | peptide | 0.83 | 0.95 | 0.77 | 0.77 |
| | poly-D-Asp | 0.84 | 0.72 | 0.77 | 0.71 |
| L-Asp | peptide | 0.85 | 0.75 | 0.76 | 0.76 |
| | poly-D-Asp | 0.85 | 0.97 | 0.78 | 0.79 |
| poly-D-Asp | peptide | 0.86 | 0.73 | 0.78 | 0.76 |
| protein OC-17 | D-Asp | | 0.74 | | 0.73 |
| | L-Asp | | 0.97 | | 0.77 |
| | peptide | | 0.75 | | 0.79 |
| | poly-D-Asp | | 0.96 | | 0.77 |

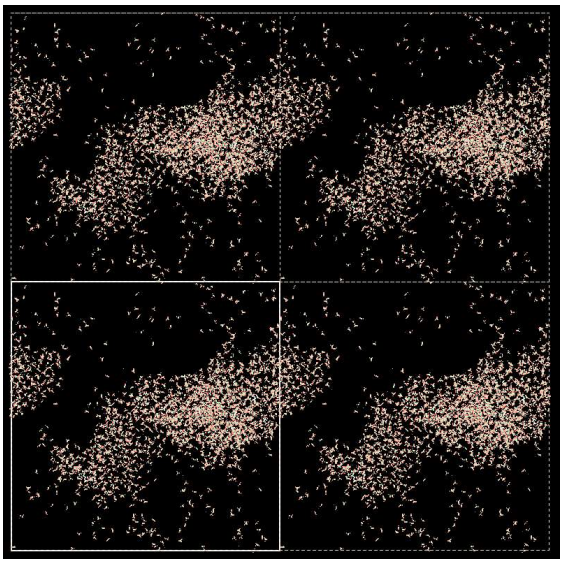
Table S4.4 continued. Tanimoto index and distance of the final cluster between system A and system B.

| | | | | | |
|---|---|---|---|---|---|
| **Only $CO_3^{2-}$** | | | | | |
| pure | D-Asp | 0.87 | 0.75 | 0.76 | 0.72 |
| | L-Asp | 0.86 | 0.96 | 0.76 | 0.83 |
| | peptide | 0.88 | 0.75 | 0.76 | 0.76 |
| | poly-D-Asp | 0.88 | 0.95 | 0.76 | 0.78 |
| | protein OC-17 | n.a. | 0.96 | n.a. | 0.76 |
| D-Asp | L-Asp | 0.86 | 0.74 | 0.79 | 0.70 |
| | peptide | 0.87 | 0.95 | 0.75 | 0.77 |
| | poly-D-Asp | 0.87 | 0.72 | 0.80 | 0.74 |
| L-Asp | peptide | 0.86 | 0.75 | 0.75 | 0.76 |
| | poly-D-Asp | 0.86 | 0.96 | 0.78 | 0.77 |
| poly-D-Asp | peptide | 0.87 | 0.73 | 0.83 | 0.75 |
| protein OC-17 | D-Asp | n.a. | 0.74 | n.a. | 0.74 |
| | L-Asp | n.a. | 0.97 | n.a. | 0.77 |
| | peptide | n.a. | 0.75 | n.a. | 0.74 |
| | poly-D-Asp | n.a. | 0.96 | n.a. | 0.76 |

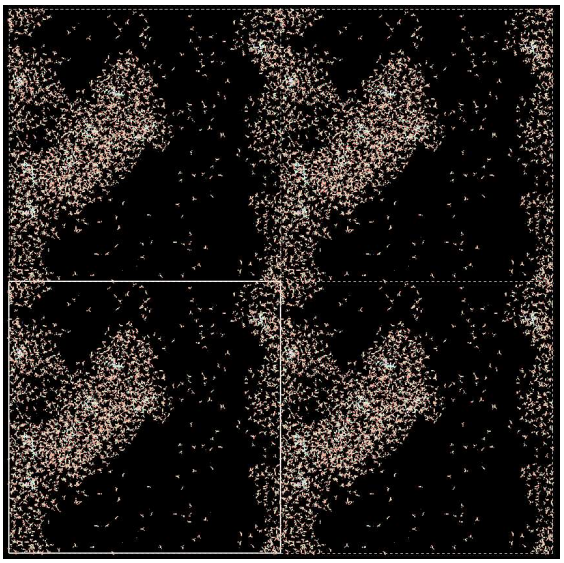
### 4.6.5 Cluster shapes



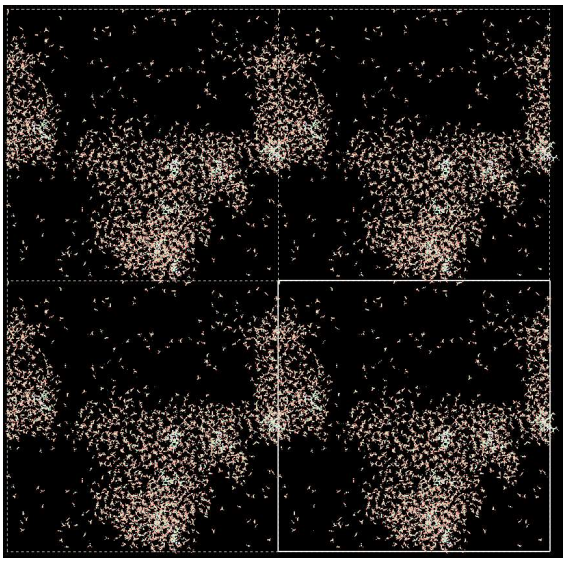
(a)



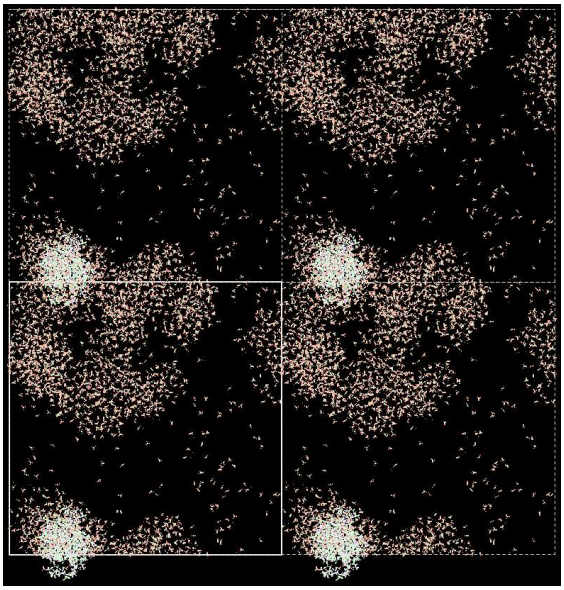
(b)



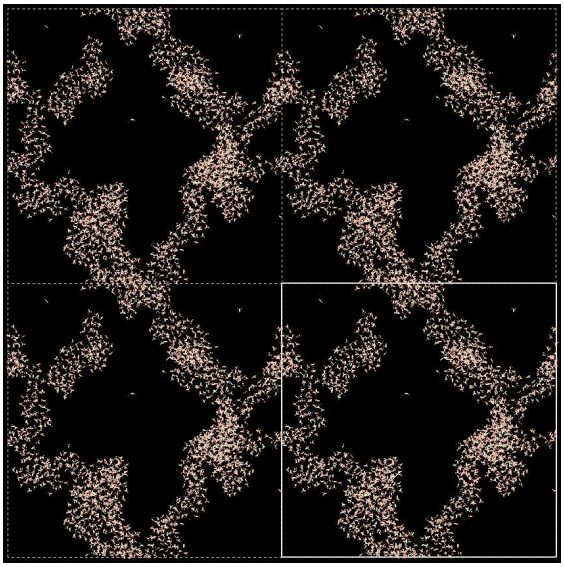
(c)



(d)



**(e)**



**(f)**

Figure S4.15 final cluster shape in the super cell with eight biomolecules; (**a**) disc shaped cluster with D-Asp, (**b**) elongated cluster with L-Asp, (**c**) elongated cluster with poly-D-Asp, (**d**) elongated cluster with peptide, (**e**) disc shaped cluster with OC-17 and (**f**) elongated rugose tubed cluster without biomolecule. The larger biomolecules stand out as highlighted in white (c-e).

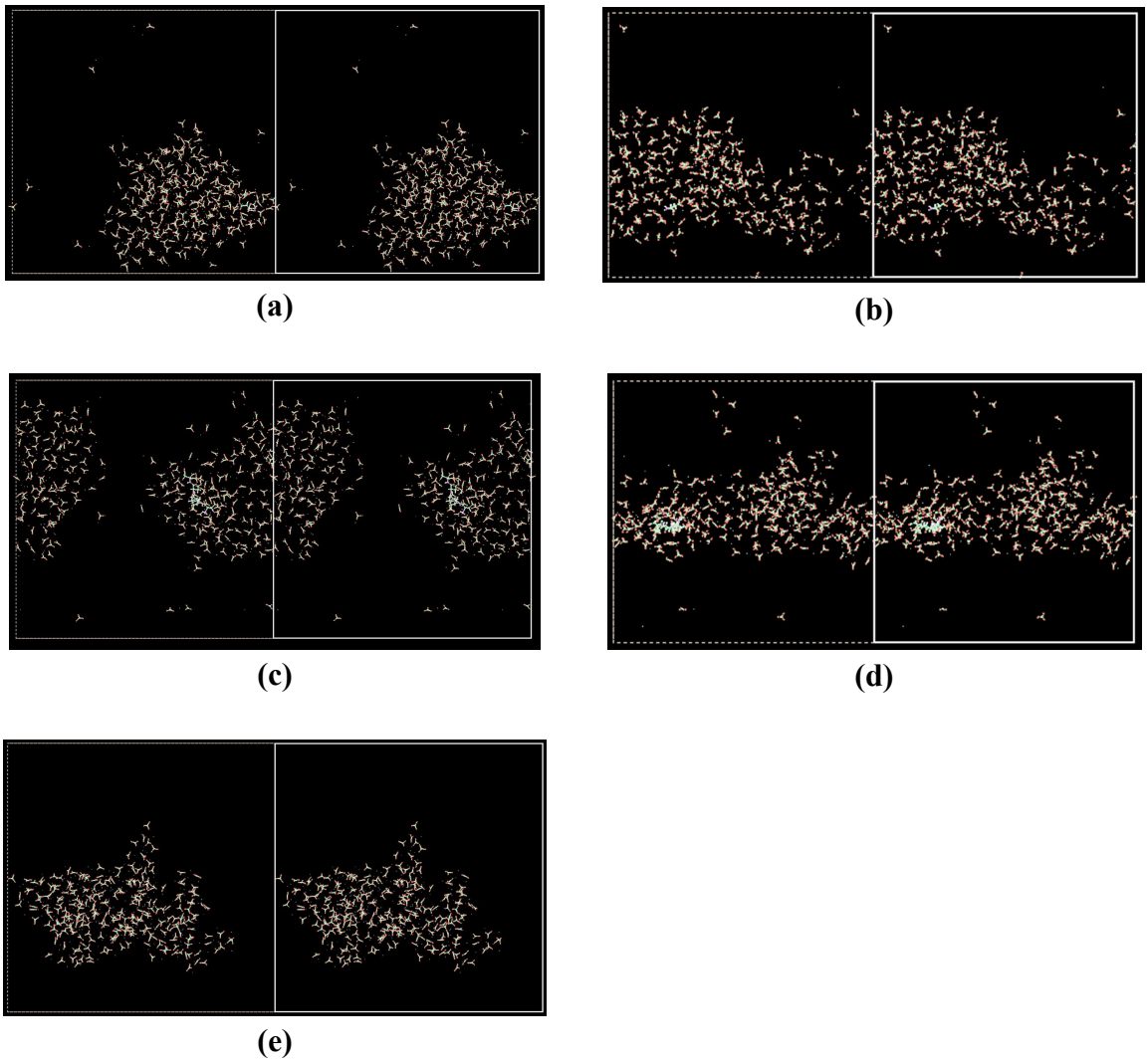
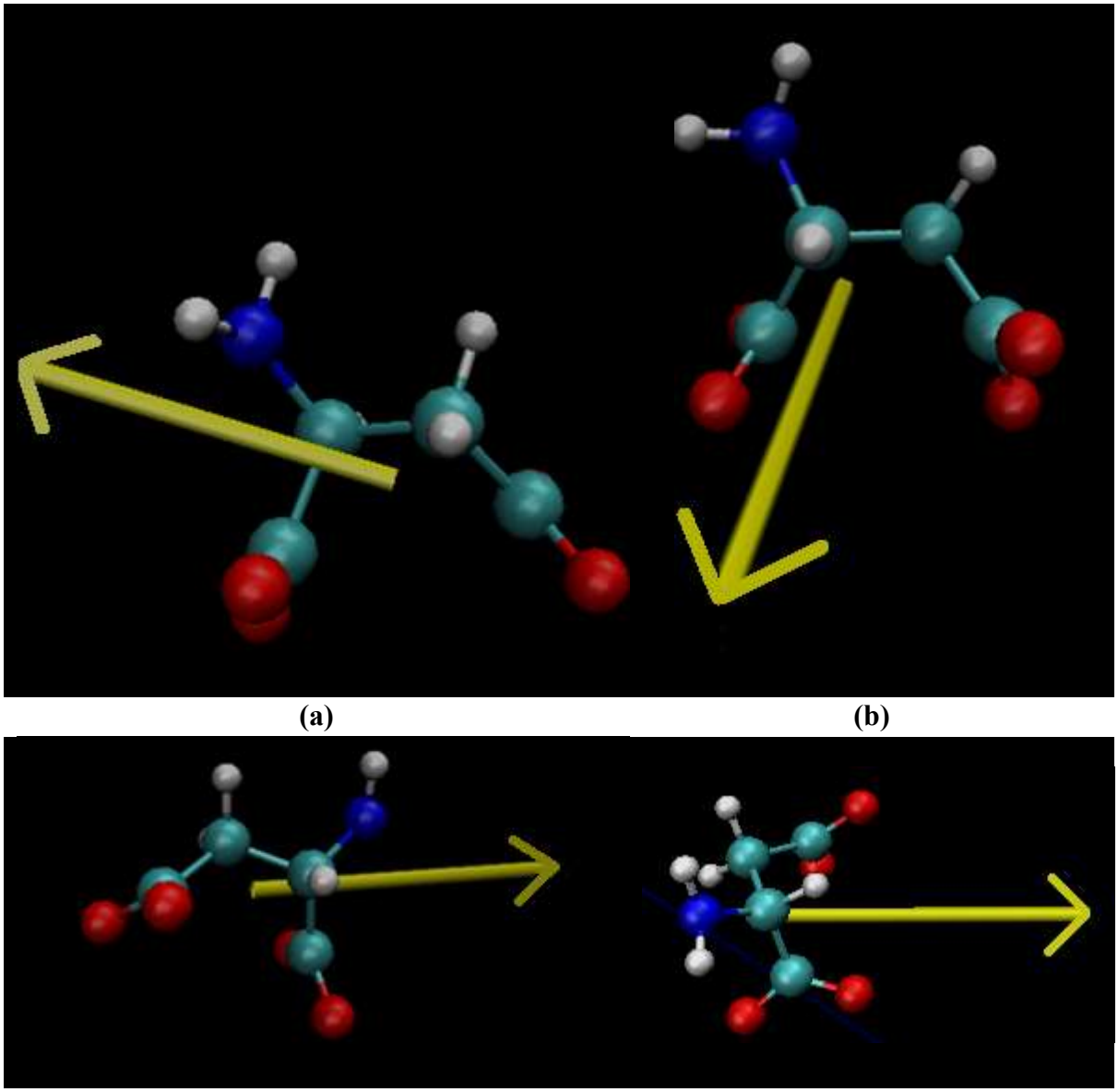



Figure S4.16 final cluster shape small cell; (**a**) rounded tabular shaped cluster with D-Asp, (**b**) elongated tabular cluster with L-Asp, (**c**) tabular shaped cluster with poly-D-Asp, (**d**) elongated cluster with peptide and (**f**) elongated rugose cluster without biomolecule.



**(a)** **(b)**

**(c)** **(d)**

Figure S4.17 Small biomolecule with the resulting induced dipole (yellow arrows) of its solvation shell in the small cells. The atoms are: Nitrogen (blue), Carbon (cyan), Oxygen (red) and Hydrogen (white).). Two projections are presented for (**a**) and (**c**) D-Asp and (**b**) and (**d**) L-Asp. In the projection of (**a**) and (**b**), focus was on representing the biomolecule in the same way, while the projection in (**c**) and (**d**) was oriented so the dipole moment lies exactly in the plane of view (2D representation).

Table S4.5 Induced dipole moment of solvation water around the biomolecule at different times during the simulation. * is simulation time; [a] pointing away from the cluster; ∥ is parallel; ⊥ is normal. * the dipole is pointing in the direction the cluster growth was the second least instead of the least.

| System | Simulation time (ns) | Resulting Dipole (D) | average dipole (D) | #water molecules | normalized dipole | Dipole direction | Dipole direction at face |
|---|---|---|---|---|---|---|---|
| D-Asp | 25 | 9.05 | 2.50 | 23 | 0.39 | away[a] | ∥ |
| D-Asp | 50 | 14.00 | 2.41 | 28 | 0.50 | away[a] | ⊥ |
| D-Asp | 75 | 9.15 | 2.41 | 21 | 0.44 | away[a] | ∥ |
| D-Asp | 100 | 7.73 | 2.46 | 24 | 0.32 | away[a] | ∥ |
| L-Asp | 25 | 8.15 | 2.44 | 24 | 0.34 | away[a] | ⊥ |
| L-Asp | 50 | 4.60 | 2.42 | 26 | 0.18 | away[a] | ⊥ |
| L-Asp | 75 | 3.88 | 2.44 | 22 | 0.18 | away[a] | ⊥ |
| L-Asp | 100 | 6.93 | 2.49 | 23 | 0.30 | away[a] | ⊥ |
| peptide | 25 | 13.85 | 2.41 | 61 | 0.23 | away[a] | ∥ |
| peptide | 50 | 13.34 | 2.42 | 61 | 0.22 | away[a] | ∥ |
| peptide | 75 | 14.65 | 2.45 | 57 | 0.26 | away[a] | ⊥ |
| peptide | 100 | 13.50 | 2.45 | 67 | 0.20 | away[a] | ⊥ |
| Poly-D-Asp | 25 | 17.25 | 2.46 | 64 | 0.27 | n.d. | ∥ |
| poly-D-Asp | 50 | 7.26 | 2.40 | 62 | 0.12 | away[a] | ∥ |
| poly-D-Asp | 75 | 10.77 | 2.44 | 62 | 0.17 | away[a*] | ⊥ |
| poly-D-Asp | 100 | 8.28 | 2.43 | 66 | 0.13 | away[a] | ∥ |

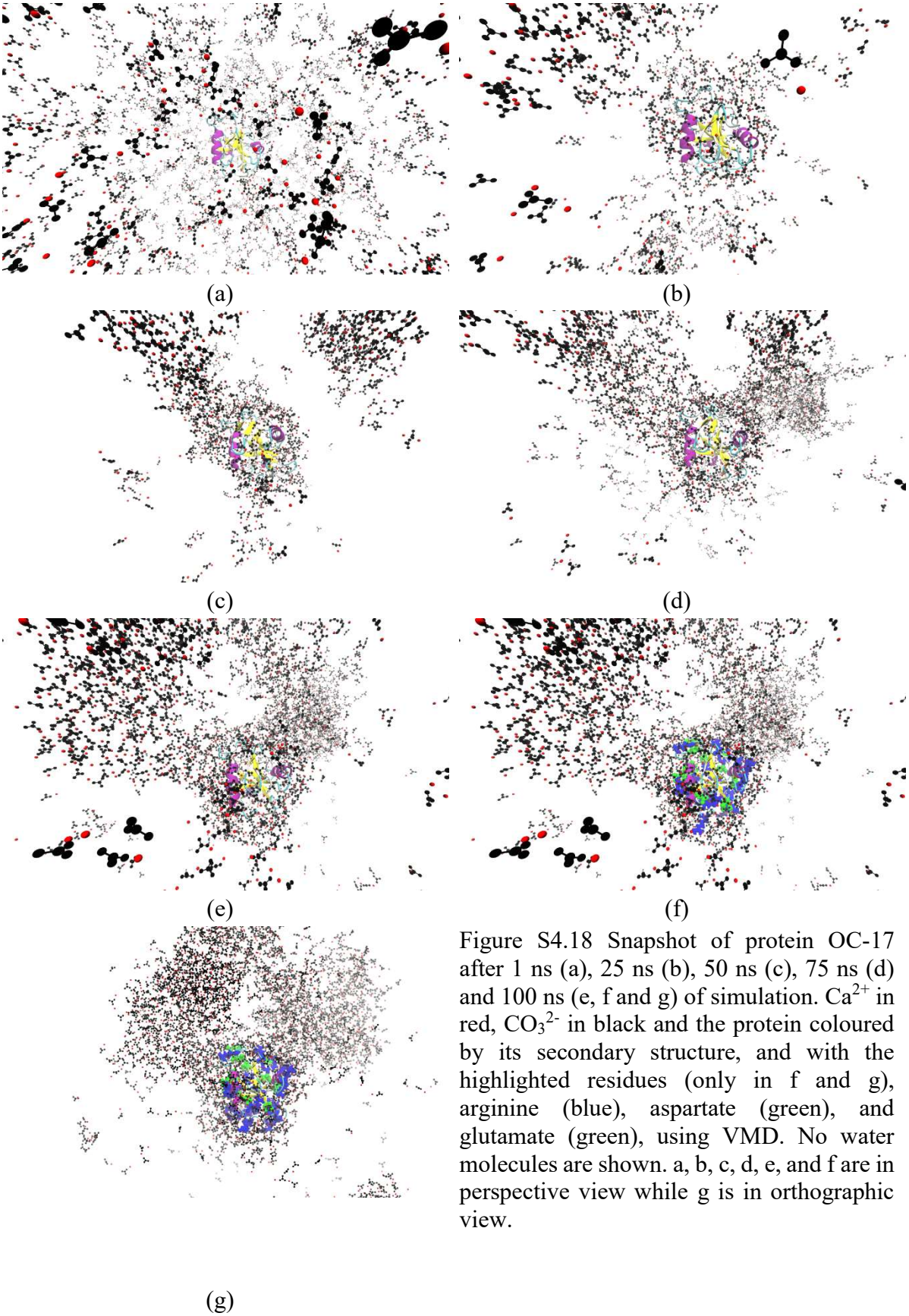



Figure S4.18 Snapshot of protein OC-17 after 1 ns (a), 25 ns (b), 50 ns (c), 75 ns (d) and 100 ns (e, f and g) of simulation. $Ca^{2+}$ in red, $CO_3^{2-}$ in black and the protein coloured by its secondary structure, and with the highlighted residues (only in f and g), arginine (blue), aspartate (green), and glutamate (green), using VMD. No water molecules are shown. a, b, c, d, e, and f are in perspective view while g is in orthographic view.

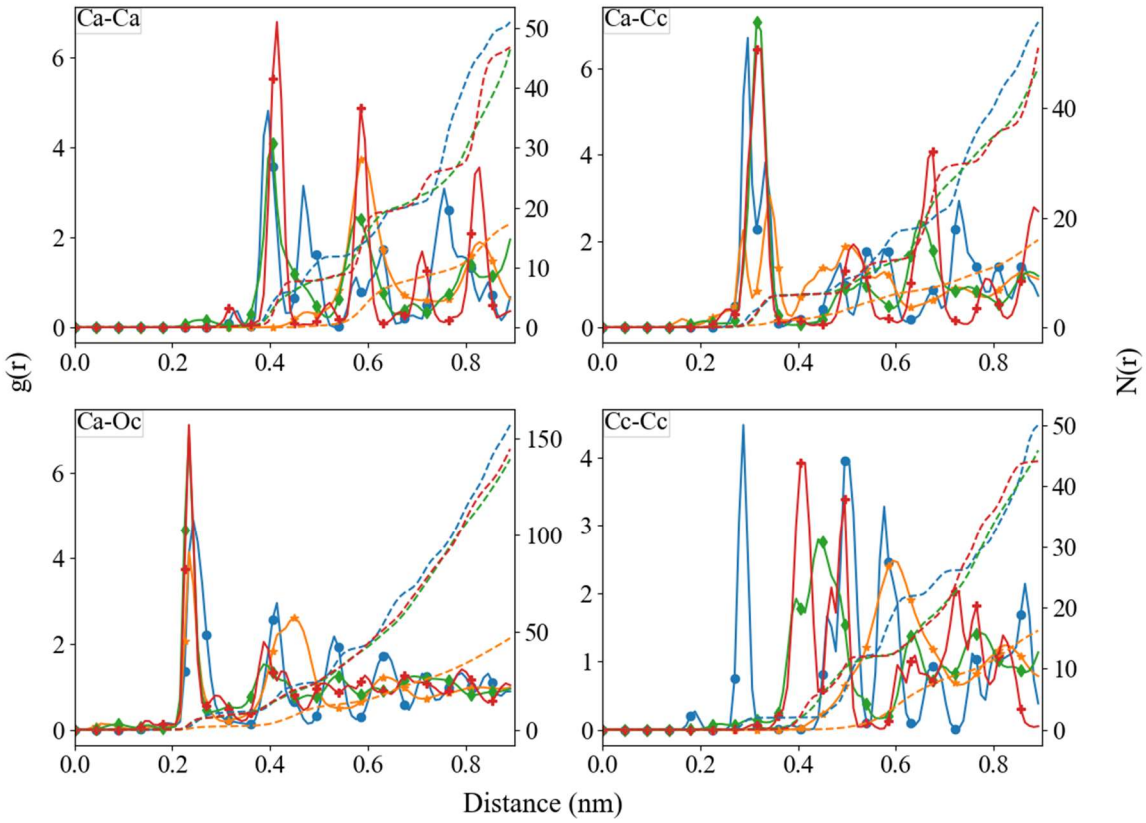



Figure S4.19 RDF (left axis) and corresponding integral (N(r)) (right axis) between $Ca^{2+}$, $C_c$, $O_c$, after 1 ns of the simulation of different $CaCO_3$ polymorphs; Calcite (purple; ♦), Aragonite (blue; ●), Vaterite (Experimental (Meyer, 1969)) (orange; ▼), Vaterite (DFT (Wang and Becker, 2009)) (red; ⬢), ikaite (green; ■) Using the forcefield published before in (Feng and Redfern, 2018; Nada et al., 2016).

# Chapter 5 Recalibrating the calcium trap in amino acid carboxyl groups in classical molecular dynamics simulations

5

Almost ready to be submitted as: **Koskamp J. A., Ruiz-Hernandez S. E., De Leeuw N. H., Wolthers, M.** Recalibrating the calcium trap in amino acid carboxyl groups in classical molecular dynamics simulations. *RSC Advances*

**Abstract:** To use classical molecular dynamics to complement experiments, it is important to use robust descriptions of the system. Studying the interaction between biomolecules, like aspartic and glutamic acid, and dissolved ions is often executed using biomolecular standardized force-fields, while the interactions between biomolecules and cations are often not explicitly parameterized. In this study, we investigated different interactions of Ca with aspartic and glutamic acid and constructed the free energy profiles of $Ca^{2+}$ - carboxylate association with metadynamics simulations. Starting from a generally accepted AMBER-based force field, the association was substantially over and under-estimated, depending on the choice of water model (TIP3P and SPC/fw, respectively). To rectify this description, we replaced the default calcium parameters to the one used in $CaCO_3$ forcefields. Additionally, we modified the $\sigma_{ij}$ value in the heteroatomic Lennard-Jones interaction by 0.5% to further improve the interaction between Ca and carboxylate, based on comparison with experimentally determined association constant for Ca with the carboxylate group of L-aspartic acid. The corrected description retrieved the structural properties of the ion pair in agreement with the original biomolecule – $Ca^{2+}$ interaction in AMBER, while having an association constant comparable to experimental observations.

## 5.1 Introduction

Calcium is an important ion in many biological processes, as it interacts with biomolecules, for example, to regulate enzyme activity(Lee et al., 2007). It is also an important building block in biominerals such as in the bone mineral apatite and as $CaCO_3$ minerals to create for example the shell of molluscs, corals, and eggs. The conditions under which such biominerals are formed is complex and some of the biomineralization mechanisms remain unresolved. Atomistic computational approaches can be, and have been, used to provide insights on structure and dynamics between calcium ions and complex biomolecules that are considered to play a role in biomineralization.

Biomolecules are formed by many atoms. The treatment of such many electrons using pure *ab initio* or density functional theory approximations (DFT) is very expensive and in some cases unaffordable computationally. Therefore, classical methods are still widely used to simulate proteins and large peptides and their interactions with other systems. Multiple force fields and software packages were developed to accurately describe biomolecules: AMBER (Hornak et al., 2006), CHARMM (MacKerell et al., 1998) and GROMOS (Hermans et al., 1984) are software packages with their own set of biomolecular force fields and are all widely used these days to perform MD. Over the years the force fields have been improved and updated along with the modernization of experimental techniques and the insights from *ab initio* calculations (Henriques et al., 2015; Klauda et al., 2010; Li and Brüschweiler, 2010; Nerenberg and Head-Gordon, 2018; Reif et al., 2012; Robertson et al., 2015).

Generally, biomineralization is thought to be guided by proteins that are rich in glutamic and aspartic acid and trap calcium (Addadi et al., 2003; Mann et al., 1993; Schulze and Vogel, 1998). However, in Classical Molecular Dynamics (MD) simulations this phenomenon is not observed around the same amino acids in OC-17, neither in small aspartic acid-based

biomolecules (Innocenti Malini et al., 2017; Rani and Saharay, 2019) [Koskamp et al., 2021b]. This is contradicting the conclusions from experiments with biomolecules binding Ca (Colowick and Womack, 1969; Maune et al., 1992; Su et al., 2013; Tang and Skibsted, 2016) and *ab initio* (DFT) calculations with implicit water, which reports a shorter distance between oxygen of the carboxyl group in the amino acid ($O_{carboxylate}$) and Ca than for carbonate (Tang and Skibsted, 2016). To obtain meaningful insights into the calcium trapping behaviour of these specific amino acids it is important to address why computational outcomes differ from experimental work.

Primarily, for molecular simulations it is crucial to have an accurate description of the free energy surface (FES) of the system (König et al., 2014). Although *ab initio* calculations give the most reliable results, the high computational costs make it unfeasible to calculate the FES for these complex biosystems (König et al., 2014). By using a force field, as in MD approaches, one mimics the FES and therewith physical insights are acquired. The development of such force fields (see examples above) can be done in different manners resulting in multiple ways to describe the same system. In the commonly used force fields the non-bonding interactions are obtained using one of the Lorentz-Berthelot mixing rules and in many cases the parameters for the non-biological interactions in the systems (i.e. between ions, water, etc.) are transferred from other published developments, in individual cases this may lead to an unphysical or incorrect description of the FES (Church et al., 2015).

In this work, we investigated the interaction of Ca with $O_{carboxylate}$ in glutamic and aspartic acid dissolved in water, by using well-tempered metadynamics. We compared the binding energies of our AMBER based force field with experimental data. Using different Ca descriptions, our results show that the standard AMBER parameters combined with a Ca description reported previously (Raiteri et al., 2010), respectively under or over-estimate the Ca-$O_{carboxylate}$ binding energy. We therefore propose a refinement of the Ca-$O_{carboxylate}$ Lennard-Jones (LJ) parameters

for when SPC/fw water and Ca description based on previously reported parameters (Raiteri et al., 2010) are used in combination with AMBER-based biomolecules with carboxyl groups. The impact of this refinement on the interactions between aspartic acid and forming amorphous calcium carbonate clusters is, among other things, improving the ion pair configurations, increasing the cluster growth rate in some cases, and altering the stoichiometry of the clusters.

## 5.2 Methods

For the calculations of the free energy, and the refinement of the interatomic potential parameters for dissolved calcium and the carboxyl-oxygen in specific biomolecules, we created two different simulation cells, both containing 300 water molecules and one $Ca^{2+}$. The first cell contains an L-aspartic acid (L-Asp) and the second cell contains a glutamic acid. To study the impact of the refined description on the interaction between Ca and the biomolecule, we simulated two boxes with 15831 water molecules, 255 $Ca^{2+}$ and 255 $CO_3^{2-}$ ions randomly distributed. The box contained either one L-Asp or D-Asp. The large-scale atomic/molecular massively parallel simulator (LAMMPS) was used to perform the MD simulations (Plimpton, 1995). The simulation temperature and pressure were kept at 300 K and 1 atm in NPT ensemble with the Nosé-Hoover (Hoover, 1985; Nosé, 1984) with 0.1 and 1 ps relaxation time of the thermostat and the barostat, respectively. The equations of motion were integrated using the velocity-Verlet algorithm (Hoover, 1985; Nosé, 1984) with a time step of 1 fs. Every system was simulated for 30 ns with 1 ns of equilibration period.

The description of the biomolecules was taken from the AMBER forcefield using tleap program to generate all the structures and extract the topology and interaction parameters (D.A. Case, K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III et al., 2020). The amino acids were C-terminated and completed with a proton on the nitrogen to reproduce the final structure (Figure S5.1) [Koskamp et al., 2021b]. The $Ca^{2+}$ and $CO_3^{2-}$ ions were

simulated using the parameters from (Demichelis et al., 2011) and the parameters for the water are the SPC/fw interactions (Wu et al., 2006), as they are preferred given the validated $CaCO_3$ interactions (Raiteri et al., 2015, 2010) and because TIP3P (incorporated in AMBER), has a diffusion coefficient twice the experimental value (Koskamp et al., 2019). The missing interaction parameters between water, the ions and biomolecules are listed in Table S5.1and were either constructed from combination rules or assigned based on similarities with atom types available in the force field, as it was tested before in similar organic inorganic interactions (Ruiz Hernandez and De Leeuw, 2015)[Koskamp et al., 2021b].

Several parameter-sets for the description of Ca using Lennard-Jones potential can be found in literature. We tested two different Ca parameters, Set_1 and Set_2. For completeness, the Ca parameters of Set_1 were tested using SPC/fw and AMBER's original TIP3P water models. The single Ca parameters based on a previously reported force field (Set_2) (Raiteri et al., 2010) were extracted from the $O_W$– $O_W$ and $O_W$– Ca interaction parameters. The modification in the force field, to refine the Set_2 interaction Ca - amino acid, were done by changing the $\sigma$ parameter (Church et al., 2015; Kahlen et al., 2014; Project et al., 2008). In the LJ-interaction, $\sigma$ alters the distance of the well and thus the equilibrium distance between the ion pair, this may also change the depth of the well. We increased the $\sigma$ LJ-parameter for $Ca^{2+}$ - $O_{carboxylate}$ stepwise up to +2.0% (see Table 5.3) to decrease the well depth to match experimental observations similar to the method described in (Church et al., 2015).

The free energy profiles were constructed using well-tempered metadynamics simulations for every modified set of interactions. The PLUMED 2.5.3 (Tribello et al., 2014) plug-in for LAMMPS was used to perform all free energy calculations. The collective variable (CV) was defined as the distance between Ca and the central carbon of the carboxylate group. Gaussian hills were deposited every 10 ps with a hill height of 1.0 kJ $mol^{-1}$ and a width of 0.02 nm. The bias factor was set on 10. Since Ca is considered fully dissolved at approximately 1 nm, we

explored the CV up to 1.6 nm distance, and we placed an upper wall at 1.6 nm with a force constant ($\kappa$) of 2000. The total simulation time was 100 ns per metadynamics simulation to achieve convergence. The obtained free energy profiles were normalized by setting the asymptotic region to zero. To be concrete, the average energy of the DV distance at range from 1 to 1.4 nm was set to zero and the correction value was subtracted from the energy profile.

The comparison of the free energy profiles from the metadynamics simulations with the experimentally obtained association constant ($K_a$) was done by predicting the $K_a$ from the profiles. $K_a$ is a stability constant, in literature also referred to as the formation or binding constant and in terms of activities, $K_a$ is defined as:

$$K_a = \frac{a_{AC}}{a_A a_C} \quad (1)$$

where $a_{AC}$ is the activity of the anion (ligand)-cation complex, while $a_A$ and $a_C$ are the activities of the free anions and cations in solution, respectively. The association constant of calcium to the ligand as seen in experiments covers a multistep process (Eigen and Tamm, 1962). Upon association, the calcium comes closer to the ligand, expressing the multiple metastable association structures: the solvent-separated ion pair (SSIP) describing a loose ion pair, solvent-shared ion pairs (SIP) and the contact ion pair (CIP). It is therefore important to consider all states of association when processing MD results. Another way of expressing $K_a$ is (Chialvo et al., 1995):

$$K_a = 4\pi \int_0^{R_U} g^{id}(r) r^2 \mathrm{d}r \quad (2)$$

with $g^{id}(r)$ as the radial distribution function (RDF) of the ligand and the cation, $r$ the radial distance and $R_U$ the upper limit corresponding to the distance between the cation and the ligand in the dissociated state. The dissociated state is defined as the first local minimum in the $g^{id}(r)$ beyond the SIP maximum (around 0.8 nm). To get the $K_a$ from our own MD simulations, we

converted the energy profile into a radial distribution function $g^{id}(r)$, as we assumed an infinite dilution of the ion pair:

$$g^{id}(r) = e^{-\frac{dE(r)}{k_b T}} \quad (3)$$

where $dE(r)$ is the free energy profile in kJ/mol along the CV distance, the Boltzmann constant is given by $k_b$ (kJ/K) and T the temperature in K. The obtained $K_a$ could then be converted into Gibbs free energy (ΔG):

$$\Delta G_{\text{association}} = -RT \ln(K_a) \quad (4)$$

where R is the gas constant. The advantage of this method over computing the $\Delta G_{association}$ ($\Delta G_a$) as the difference between the free energy minimum and the asymptotic limit for long-range separation, is that now the width and depth of the local wells are accounted for and thus the whole profile is of more importance compared to taking a single minimum and an asymptotic maximum.



**Impact on clustering.** To investigate the impact of our modified Ca ligand interaction of the AMBER based forcefield, we analysed the formation of an amorphous calcium carbonate (ACC) cluster in presence of L- or D-aspartic acid. An in-house python script was used to calculate the number of clusters and their size in terms of ions over the course of the simulation. We used the same method as [Koskamp et al., 2021b] to compare the evolution in cluster size distribution with time. The size is given in terms of number of ions and the gyration radius, which variance showed how liquid-like the particles behaved. The gyration radius ($R_G$) is defined as:

$$R_{gyration} = \sqrt{\frac{\sum_{i=1}^{N} m_i (r_i - r_{COM})^2}{\sum_{i=1}^{N} m_i}} \quad (5)$$

where $r_i$ and $r_{com}$ are the vectors of the $i$th atom in the cluster and of the centre of mass of the cluster, respectively. $m_i$ is the mass of the $i$th atom and $N$ is the number of atoms in the cluster. To study the possibility of water in and around the clusters, the solvent accessible surface area (SASA) was computed over the full trajectory to follow the change in time, using the MDTraj (McGibbon et al., 2015) module in the Anaconda Python distribution. The method consisted of exploring the area around all $Ca^{2+}$ and $CO_3^{2-}$ ions with a probe with the size of a water molecule (radius=0.14 nm).This was done by running the Shrake and Rupley (Shrake and Rupley, 1973) algorithm. The stoichiometry of the formed clusters was calculated by dividing the number of $Ca^{2+}$ ions by the total number of ions in the cluster. A relative charge of 0.5 was therefore equal to a neutral cluster. The presented data were an average of the last 10 ns (10000 snapshots) of the simulation. The local density of ACC ($Ca^{2+}$ and $CO_3^{2-}$) was calculated from the integral of the radial distribution function until 0.9 nm divided by the volume corresponding to that radius.



## 5.3 Results

### 5.3.1 Calcium description

The description of $Ca^{2+}$ in an interatomic potential for MD simulations varies over different studies involving a calcium ion. Some of the calcium descriptions that can be found in literature are listed in Table 5.1. As can be seen from the table, ε varies more than one order of magnitude while the maximum difference in σ is approximately 0.1 nm. The intrinsic Pauli repulsion is described by the A parameter ($A = 4\pi\varepsilon\sigma^{12}$), while the attraction between two Ca ions is described by the B parameter ($B = 4\pi\varepsilon\sigma^{6}$). For the highest and the lowest value of ε, the A parameter of Set_1 (Hornak et al., 2006) is 2.5 orders of magnitude larger than the Ca repulsion of Set_2. The attraction that compensates this repulsion, was also 1.4 order of magnitude larger, which resulted in more similar balance between repulsion and attraction as seen in the A/B

values. In the continuing sections we compared the two most extreme descriptions for Ca, Set_1 and Set_2.

Table 5.1 Comparison of different Ca parameters

| $Ca^{2+}$ | acronym | ε (eV) | σ (Å) | A | B | A/B |
|---|---|---|---|---|---|---|
| Hornak, 2006 (Hornak et al., 2006) | Set_1 | 0.019937830 | 3.426200 | 208685 | 129 | 1618 |
| Mamatkulov, 2013 (Mamatkulov et al., 2013) AMBER | | 0.019899403 | 3.050000 | 51582 | 64 | 805 |
| Aqvist, 1990 (Åqvist, 1990) | | 0.019692117 | 2.410000 | 3024 | 15 | 196 |
| Wang, 2008 (Wang et al., 2008) | | 0.011457217 | 2.789015 | 10152 | 22 | 471 |
| Mamatkulov, 2013 (Mamatkulov et al., 2013) | | 0.009742416 | 2.410000 | 1496 | 8 | 196 |
| Martinek, 2017 (Kohagen et al., 2014; Martinek et al., 2018) | | 0.005256800 | 2.819600 | 5309 | 11 | 502 |
| Beglov, 1994 (Beglov and Roux, 1994) CHARMM | | 0.005182136 | 2.430000 | 879 | 4 | 206 |
| Dang, 1995 (Dang and Smith, 1995) | | 0.004336300 | 2.895000 | 6011 | 10 | 589 |
| Shen (Shen et al., 2013) | | 0.003316575 | 2.450000 | 621 | 3 | 216 |
| Babu and Lim, 2006 (Babu and Lim, 2006) | | 0.001243713 | 3.250000 | 6908 | 6 | 1178 |
| This work (based on (Raiteri et al., 2015)) | Set_2 | 0.001104240 | 3.247210 | 6071 | 5 | 1172 |

### 5.3.2 Free energy of association

The potential parameters of Ca were used to describe the interaction with the aspartic acid in our MD simulations. To evaluate and improve this interaction, the MD results were reviewed considering the experimental $K_a$ and the coordination distance from *ab initio* MD. As a starting point we constructed the free energy profiles using the unmodified Set_1 and Set_2 interactions, Figure 5.1a. We combined the parameters from Set_1 with TIP3P and SPC/fw water. The combination with TIP3P resulted in the deepest well in the energy profile compared

to all other parameters set published here. The well ends at 0.78 nm and had four local minima at 0.290, 0.346, 0.513 and 0.677 nm. The minima correspond with different Ca-ligand-water structures, the first and second minima corresponded to a CIP that was split in a bidentate and a monodentate configuration. In the bidentate (biCIP) configuration, Ca was coordinated to both oxygens in the carboxylate group, while in the monodentate (monoCIP) Ca was only coordinated to one of the oxygens. The third and the fourth minima resembled SIP and SSIP, respectively.

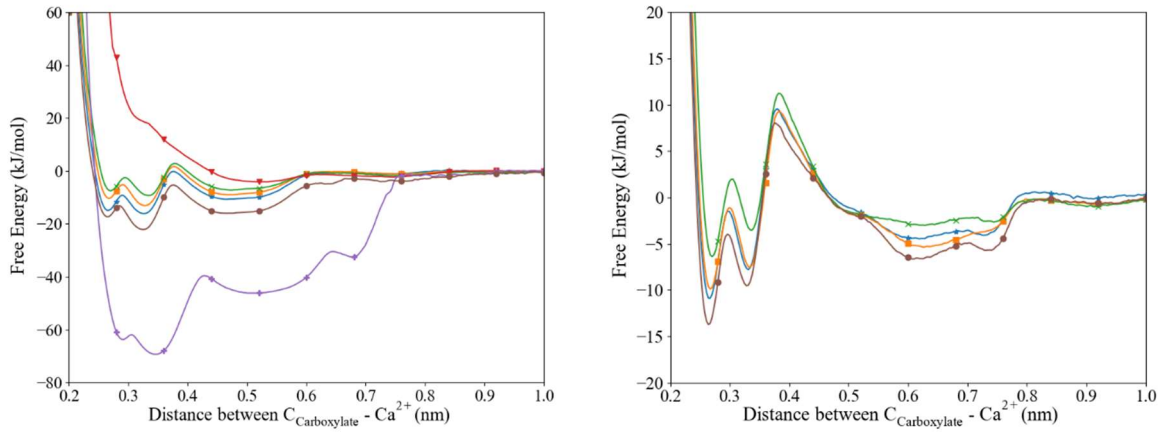



Figure 5.1 Free energy profiles for the interaction between $Ca^{2+}$ and the $O_{carboxylate}$. (a) L-Asp and (b) glutamic acid. Set_1$_{TIP3P}$ (purple; □), Set_1$_{SPC/fw}$ (red; ▼), Set_2$_{unmodified}$ (brown; ●), Set_2$_{+0.5\%\sigma}$ (blue; ★), Set_2$_{+1.0\%\sigma}$(orange; ■), Set_2$_{+2.0\%\sigma}$ (green; ×).

From the free energy profile, we extracted the highest $\Delta G_a$ (62.3 kJ/mol) and $K_a$ ($7.07 \cdot 10^{10}$) compared to the other parameter sets and literature values (Table 5.2). The coordination distance between Ca and $O_{carboxylate}$ was slightly longer than observed in literature. In our previous work [Koskamp et al., 2021b], TIP3P water was replaced by SPC/fw water and set_1 was used, since this water model is found more realistic (Koskamp et al., 2019). The corresponding free energy profile of this system did not show clear minima Figure 5.1a. The global minimum was at 0.546 nm which meant a 0.466 nm Ca-$O_{carboxylate}$ distance. The shallow minimum in the energy profile led to the lowest observed $K_a$ (Table 5.2). Both water models in combination with the parameters from Set_1 were not able to reproduce the experimentally

observed $K_a$ and distance between the Ca and the ligand, although the distance for Ca-$O_{carboxylate}$ with set_1 and SPCF/W compared favourably with previously published distanced obtained from MD.

The association free energy obtained from the unmodified set_2 simulations was in better agreement with experimental values than those from set_1 (Table 5.2). For the free energy profile of $Set\_2_{unmodified}$, we observed three distinct minima: around 0.28, 0.33 and 0.50 nm (Figure 5.1a), corresponding to biCIP, monoCIP, and SIP, respectively. A less distinguishable minimum was observed around 0.8 nm that correlates with SSIP. The deepest minimum of the energy profile was found at the second well at 0.33 nm. The resulting association free energy was calculated to be 15.7 kJ/mol, approximately 4.4 kJ/mol higher than the latest experimental value. The coordination distance was closer to the distances reported in literature (based on DFT or experimental studies) compared to the $Set\_1_{TIP3P}$ and $Set\_1_{SPC/fw}$ interactions, Table 5.2.

Table 5.2 Association free energy ($\Delta G_a$), association constant ($K_a$) and coordination distance (r) extracted from free energy profiles in this work and literature experimental and computational values. [1] no experimental reference

| Ca | Water model | acronym | Asp | | | Glu | | |
|---|---|---|---|---|---|---|---|---|
| | | | $r_{Ca-OD}$ (nm) | $-\Delta G_a$ | $K_a$ | $r_{Ca-OD}$ (nm) | $-\Delta G_a$ | $K_a$ |
| (Hornak et al., 2006) | TIP3P | $Set\_1_{TIP3P}$ | 0.274 | 62.3 | 7.07E10 | n.a. | n.a. | n.a. |
| | SPC/fw | $Set\_1_{SPC/fw}$ | 0.466 | 4.9 | 7.07 | n.a. | n.a. | n.a. |
| **based on** (Raiteri et al., 2015) | SPC/fw | $Set\_2_{unmodified}$ | 0.238 | 15.7 | 548.22 | 0.242 | 7.7 | 22.06 |
| | SPC/fw | $Set\_2_{+0.5\%\sigma}$ | 0.242 | 10.4 | 65.46 | 0.242 | 6.0 | 10.98 |
| | SPC/fw | $Set\_2_{+1.0\%\sigma}$ | 0.242 | 8.1 | 25.87 | 0.245 | 6.3 | 12.43 |
| | SPC/fw | $Set\_2_{+2.0\%\sigma}$ | 0.245 | 6.5 | 13.58 | 0.248 | 5.0 | 7.40 |
| **Experiments** | | | | | | | | |
| 2009 (Covington and Danish, 2009) | | | | 11.3 | 91.20 | | 7.4 | 19.05 |
| 1990 (Maeda et al., 1990) | | | | 11.4 | 97.50 | | | |
| 1987 (Blaquiere and Berthon, 1987) | | | | 6.5 | 13.64 | | | |
| 2012 (Raiteri et al., 2012)[1] | | | 0.243 – 0.261 | | | | | |
| 1998 (Marchand and Roux, 1998) | | | 0.241 – 0.254 | | | | | |
| **Simulations** | | | | | | | | |
| DFT(Tang and Skibsted, 2016) | | | 0.245 | | | 0.243 | | |
| 2012 MD-CHARMM(Raiteri et al., 2012) | | | 0.244 – 0.260 | 12.2 | 133.10 | | | |
| 2017 MD-AMBER-Buck (Innocenti Malini et al., 2017) | | | 0.459 | 4.5 | 6.07 | | | |
| Modified CHARMM (Church et al., 2015) | | | | | | | 6.1 | 11.54 |

Subsequently, we optimized this interaction as we modified the $\sigma_{Ca\text{-}OD}$ values in the LJ interaction by increasing it with 0.5, 1 or 2% (see Table 5.3 for the absolute values). The energy profile and corresponding $\Delta G_a$ and $K_a$ (Figure 5.1a and Table 5.2) revealed that a modified σ of +0.5% yielded in the best match with the latest experimental data (Covington and Danish, 2009). Moreover, the equilibrium coordination distance was almost the same as published using *ab initio* simulations. Higher increase of σ resulted in a lower $\Delta G_a$, however they were within the range of experimentally observed values. Increasing σ with 2.0% resulted in a distance exactly matching previously reported values (Tang and Skibsted, 2016), however the association free energy is slightly lower than the lowest experimental value (Blaquiere and Berthon, 1987). For all modifications, three distinct minima were observed corresponding to the different ion pair configurations biCIP, monoCIP and SIP. Increasing σ changed the difference between biCIP and monoCIP. σ + 0.5% reduced this difference the most from 4.8 kJ/mol difference in the unmodified profile to 1.2 kJ/mol. As for $Set\_2_{unmodified}$, the modified interactions generated the lowest minimum for monoCIP.



Table 5.3 Absolute heteroatomic Lennard-Jones σ values for the $Ca^{2+}$–$O_{carboxylate}$ pair for the combined Set_1 and Set_2 force-field for the Asp/Glu – $Ca^{2+}$ systems

| Ca-OD | σ (Å) |
|---|---|
| $Set\_2_{unmodified}$ | 3.10356595082234 |
| $Set\_2_{+0.5\%\sigma}$ | 3.11908378057645 |
| $Set\_2_{+1.0\%\sigma}$ | 3.13460161033056 |
| $Set\_2_{+2.0\%\sigma}$ | 3.16563726983879 |

Regarding the glutamic acid, a similar approach yielded in energy profiles shown in Figure 5.1b. We observed four minima around 0.26, 0.33, 0.61 and 0.74 nm. The first minimum is found to be the deepest and corresponded to the biCIP. When following the same modification method as for aspartic acid, the CIP well became shallower and the separation of SIP and SSIP became more nebulous upon increasing σ. For glutamic acid, the association energies were

lower compared to aspartic acid. When comparing to previous studies the unmodified interaction slightly overestimated the association energy (Covington and Danish, 2009), while the association energy calculated for σ + 0.5% was lower than experimental and computational published values (Church et al., 2015; Covington and Danish, 2009). The value of σ + 1% appeared to be between experimental and computational association energies. Since σ + 0.5% is closer to the unmodified value and it is most comparable to the more recent MD simulations, we proceed with evaluating the effect on the water structure only for the Set_$2_{unmodified}$ and + 0.5% modification.

## 5.4 Water structure

We investigated the details of the solvent structuring around the ion pair in the CIP coordination mode. We calculated the radial distribution function (RDF) between the carboxylate-oxygen atom and the water oxygen atom and compared our findings for both the Set_$2_{unmodified}$ and our +0.5% modification, as shown in Figure 5.2. Two peaks in the RDF were apparent, the first and most predominant at a $O_{carboxylate}$–$O_W$ separation of ~0.29 nm, that corresponded to the first solvation shell, and a second peak at ~0.46 nm that indicated the second solvation shell. These data showed that our modification did not induce a significant difference on these RDFs, for both Asp and Glu. The angular distribution function (ADF) in Figure 5.3 confirms that also the hydrogen-bond angle between the coordinated water and $O_{carboxylate}$ stayed unaltered upon our modification when Ca-$O_{carboxylate}$ were in CIP state. To complete our analysis, we studied the RDF between Ca and water when Ca was paired with the biomolecule. Figure 5.4 shows that also the RDF and coordination of water around calcium remained the same and in both cases a coordination number of 5 was found. However, note the slight increase of number of water molecules around Ca in CIP with glutamic acid, as the CN increases with 0.33 for Set_$2_{+0.5\%\sigma}$ compared to the unmodified forcefield.

Overall, the Set_2$_{unmodified}$ interaction in combination the AMBER description of the amino acid appears to be more realistic than the Set_1 Ca description. However, it still over-estimates the binding of Ca to the carboxylate group when compared with experiments (Table 5.2). Modifying this single interaction individually to Set_2$_{+0.5\%\sigma}$ helped with improving the description of the binding energy of the ligand-Ca pair in water even further (Table 5.2), without affecting the water structure around it (Figure 5.2 until Figure 5.4).

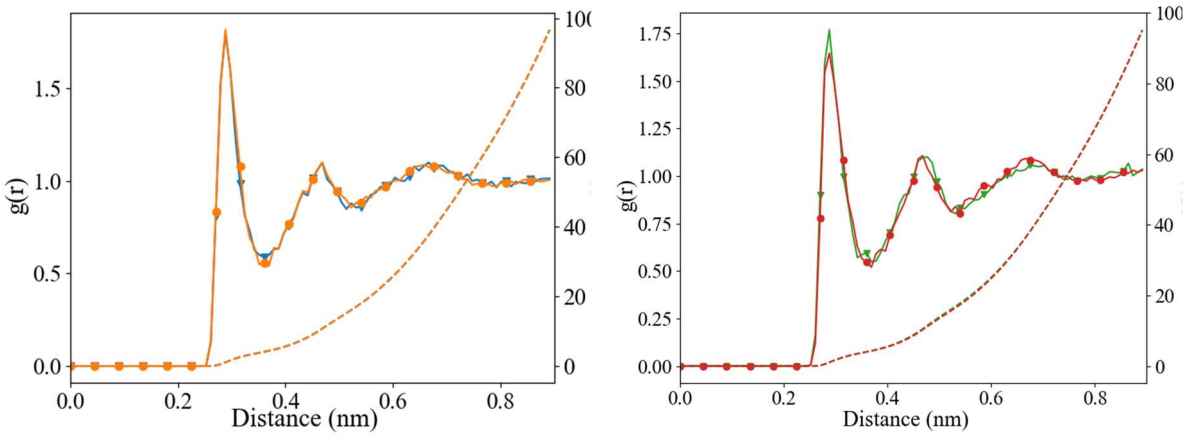



Figure 5.2 Radial distribution function (left axis) and corresponding integral (N(r)) (right axis) between $O_{carboxylate}$ and $O_w$, after 10 ns of classical MD simulation (a) L-Asp, Set_2$_{unmodified}$ (orange; ●), Set_2$_{+0.5\%\sigma}$ (blue; ▼), and (b) glutamic acid, Set_2$_{unmodified}$ (red; ●), Set_2$_{+0.5\%\sigma}$ (green; ▼) when Ca – $O_{carboxylate}$ were in CIP state.

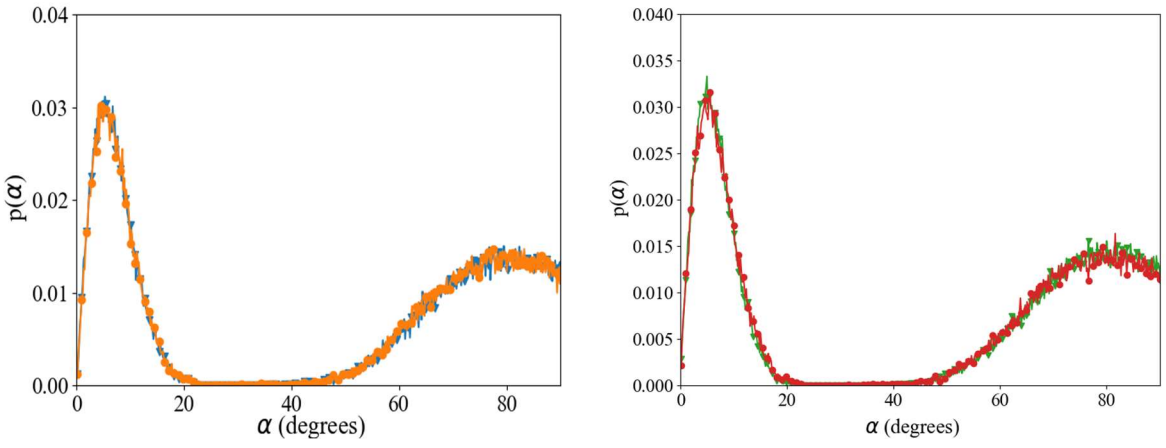
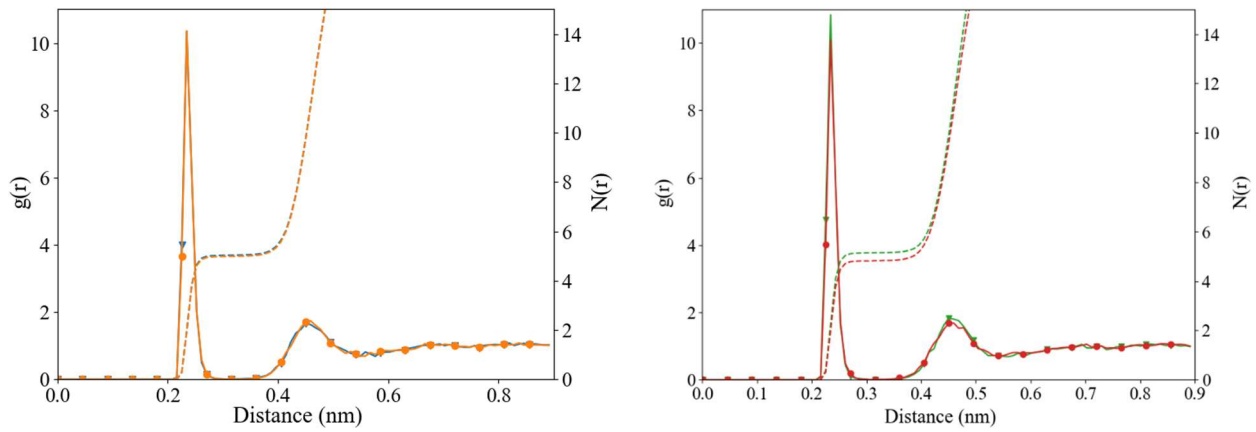



Figure 5.3 Angular distribution function between $O_{carboxylate}$, $O_w$, $H_w$ after 10 ns of classical MD simulation (a) L-Asp, Set_2$_{unmodified}$ (orange; ●), Set_2$_{+0.5\%\sigma}$ (blue; ▼), and (b) glutamic acid, Set_2$_{unmodified}$ (red; ●), Set_2$_{+0.5\%\sigma}$ (green; ▼) when Ca – $O_{carboxylate}$ were in CIP state.




a b

Figure 5.4 Radial distribution function between $Ca^{2+}$ and $O_w$, after 10 ns of classical MD simulation (a) L-Asp, Set_2$_{unmodified}$ (orange; ●), Set_2$_{+0.5\%\sigma}$ (blue; ▼), and (b) glutamic acid, Set_2$_{unmodified}$ (red; ●), Set_2$_{+0.5\%\sigma}$ (green; ▼) when Ca – $O_{carboxylate}$ were in CIP state.

## 5.5 Association Mechanism

Insights in the association and dissociation mechanism of $Ca^{2+}$ with the carboxyl group of Asp and Glu were provided by the free energy calculation when plotted as a function of 2 CV's, in this case the distance $Ca^{2+}$ and the $O_{carboxylate}$ and the coordination number of $Ca^{2+}$ and Ow. Figure 5.5 shows that in solution at >0.7 nm, the highest water coordination number for Ca was between 7 and 9. When the distance between Ca and the biomolecule was reduced, the range of possible coordination numbers increased, ranging from 5 to 8 at around 0.5 nm. Though, it is most probable to find a coordination number of 5.5 around this distance. In associated state with the carboxyl group we could distinguish both bi- and monoCIP states and the number of water molecules around Ca could be between 4 until 7, with most abundantly a coordination of 5.4.

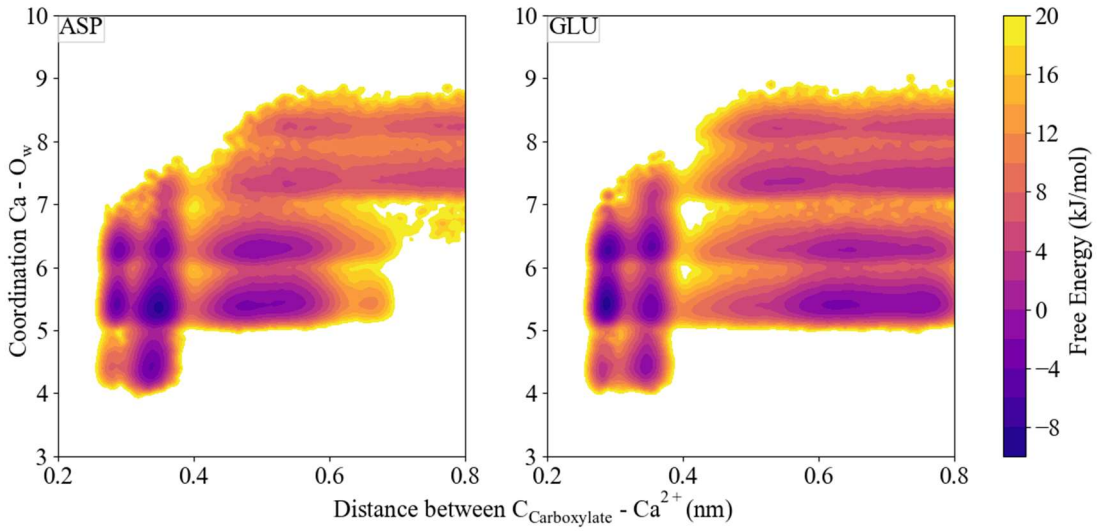



Figure 5.5 Free energy as a function of distance $Ca^{2+}$ and $O_{carboxylate}$ and coordination number of $Ca^{2+}$ and $O_W$.

## 5.6 ACC formation in presence of L/D-aspartic acid

The further implications of our proposed modification of the interatomic potential between $Ca^{2+}$ and $O_{carboxylate}$ (Set_$2_{+0.5\%\sigma}$) were investigated in a more complex system. We studied the interaction of freely dissolved $Ca^{2+}$ and $CO_3^{2-}$ and D-aspartic acid or L-aspartic acid in a system highly supersaturated with respect to ACC, comparable to Koskamp et al. (2021b) and compared the development of cluster size, the cluster composition, and the cluster structure with previously published results where Set_$1_{SPC/fw}$ was used.

### 5.6.1 ACC cluster size distribution and their solvent accessible surface area

The freely dissolved ions and one L/D-aspartic acid aggregated over the course of the simulation. The agglomeration was followed by calculating the cluster size distribution every 10 ns of simulation. The cumulative distribution functions in Figure 5.6 showed no drastic changes in the formation of the clusters between the proposed and former forcefield, though the distribution in presence of D-Asp shifted to higher cluster sizes after 20 (orange) and 30 (green) ns while keeping a similar shape of the curve. Additionally, we observed intermediate

cluster sizes of ±23, ±90 and ±280 ions, these cluster sizes were not visible for L-Asp. The complete cumulative probability profiles for all tested σ can be found in the SI Figure S5.2 and Figure S5.3.

The solvent accessible surface area (SASA) of the $Ca^{2+}$ and $CO_3^{2-}$ -ions in solution was followed over time to study the effect of clustering on the surface area that was exposed to water with the new proposed interaction. The volume within the cluster was assumed accessible and/or occupied by water molecules when it was equal or bigger than the volume of one water molecule (approximately 0.14 nm). As observed in Figure 5.7 (the data for the other σ can be found in SI Figure S5.4 and Figure S5.5), the new interaction increased the SASA for both biomolecules compared to the former interaction. Moreover, there was a difference between (L and D) Asp: clusters formed with L-Asp showed a higher SASA than D-Asp with the refined force field, while the SASA was indistinguishable for both enantiomers when Set_$1_{SPC/fw}$ was used Koskamp et al. (2021b).



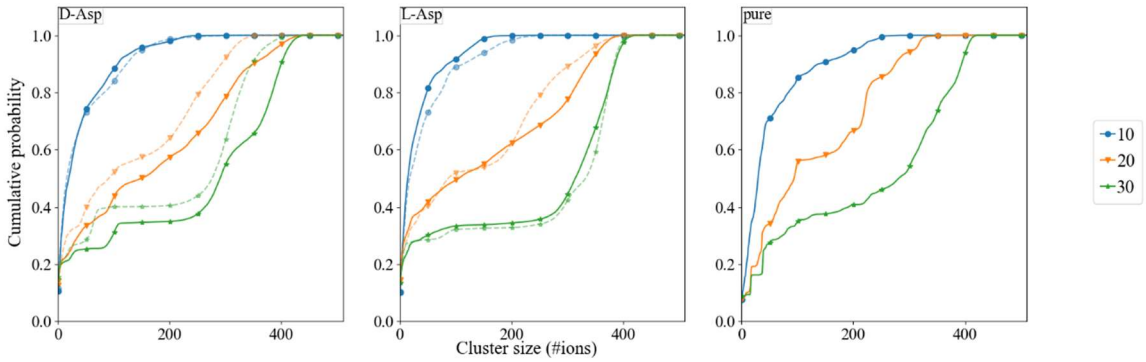



Figure 5.6 Cumulative probability profiles of different $CaCO_3$ systems for every 10 ns of simulation using different forcefields: Set_$1_{SPC/fw}$ (dashed lines) and Set_$2_{+0.5\%\sigma}$ (continuous lines) with D-Asp or L-Asp and without biomolecule (pure) for comparison.

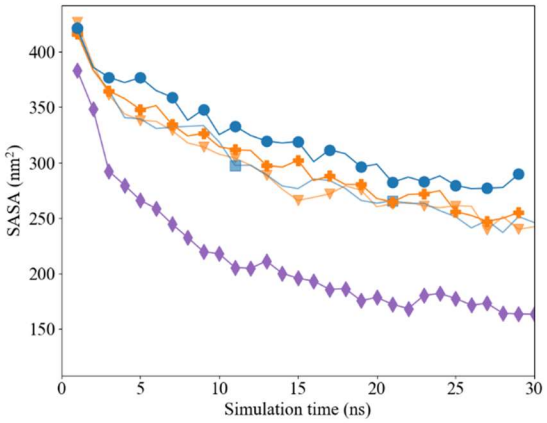



Figure 5.7 SASA analysis using the Shrake and Rupley algorithm of Pure (♦), L-Asp with Set_1$_{SPC/fw}$ (transparent blue; ■), L-Asp Set_2$_{+0.5\%\sigma}$ (blue; ●), D-Asp Set_1$_{SPC/fw}$ (transparent orange; ▼), D-Asp Set_2$_{+0.5\%\sigma}$ (orange; ✚)

### 5.6.2 ACC cluster composition and density

The composition of the clusters showed more $CO_3^{2-}$ in the clusters, after 30 ns of simulation, making it overall negatively charged, Figure 5.8 (the data for the other σ can be found in SI Figure S5.6). As with the SASA, the new interaction had more impact on the stoichiometry of the clusters for L-Asp compared to D-Asp, with most of the smaller clusters systematically more negative than when the Set_1$_{SPC/fw}$ were used. The final large clusters show a convergence of stoichiometry between L- and D-Asp systems, with comparable excess in $CO_3^{2-}$ as for the Set_1 simulations (Koskamp et al., 2021b).

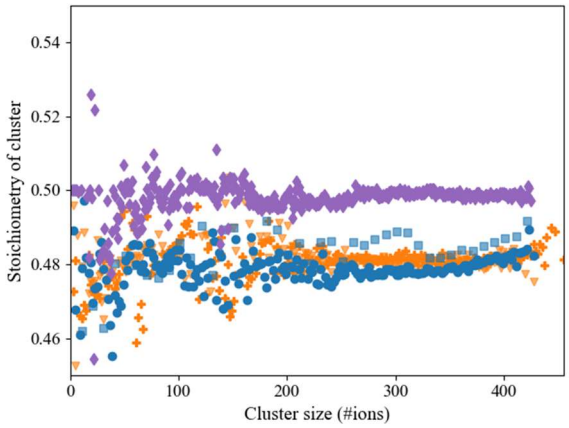



Figure 5.8 Average cluster stoichiometry ($Ca^{2+}\cdot(Ca^{2+}+CO_3^{2-})^{-1}$) over 30 ns of simulation. Pure (♦), L-Asp with Set_1$_{SPC/fw}$ (transparent blue; ■), L-Asp Set_2$_{+0.5\%\sigma}$ (blue; ●), D-Asp Set_1$_{SPC/fw}$ (transparent orange; ▼), D-Asp Set_2$_{+0.5\%\sigma}$ (orange; ✚)

We calculated the density of $CaCO_3$ around different groups in the biomolecule as well as the average density of the ACC clusters by calculating the density around every $Ca^{2+}$ in the system (Table 5.4). An increase in density around the functional groups of the biomolecule, due to the new interaction, was most noticeable for D-Asp with 38%, while the average ACC density was very comparable for both sets of simulations. Contrastingly, L-Asp and the formed ACC for both systems the measured densities were less affected by this new interaction.

Table 5.4 Density of $Ca^{2+}$ and $CO_3^{2-}$ ions around the functional groups of the biomolecules and around $Ca^{2+}$. *hydrated ACC

| | Density (g/cm³) | | |
|---|---|---|---|
| **Biomolecule** | **$O_{carboxyl\ group}$** | **$N_{amine}$** | **$Ca^{2+}$** |
| **L-Asp** | 0.73 | 0.78 | 0.47 |
| **D-Asp** | 0.80 | 0.80 | 0.50 |
| **Literature** | | | |
| **L-Asp** [Koskamp et al., 2021b] | 0.78 | 0.83 | 0.46 |
| **D-Asp** [Koskamp et al., 2021b] | 0.58 | 0.58 | 0.47 |
| **Pure** [Koskamp et al., 2021b] | n.a. | n.a. | 0.63 |
| **L-Asp** (Innocenti Malini et al., 2017) | | | ~1.1*-1.4 |
| **Aspartate in lysozyme** (Rani and Saharay, 2019) | ~0.71 | | |



### 5.6.3 ACC cluster structure

As observed in the ACC RDFs, the distances remained unaltered upon potential refinement, yet the peak intensities were somewhat lower, especially for $C_c$-$C_c$ (carbon of carbonate) RDFs (Figure 5.9), indicating slightly lower densities of carbonate groups at those distances. The RDFs between the different parts of Asp as shown in Figure 5.10 revealed a clear difference between the force fields for the distance between the carboxyl oxygen OD and the coordinated the Ca, which decreased from 0.46 nm to 0.24 nm. In addition, for L-Asp a clear separation between biCIP and monoCIP was observed at 0.288 and 0.349 nm, respectively. This was more indistinct for D-Asp. The refined OD-Ca potential parameters also resulted more structure between the amine group and Ca, as reflected by the more pronounced minima in the N-Ca and especially the N-$C_c$ RDFs. Data from the other tested σ values can be found in SI Figure S5.7, Figure S5.8, Figure S5.9 and Figure S5.10)

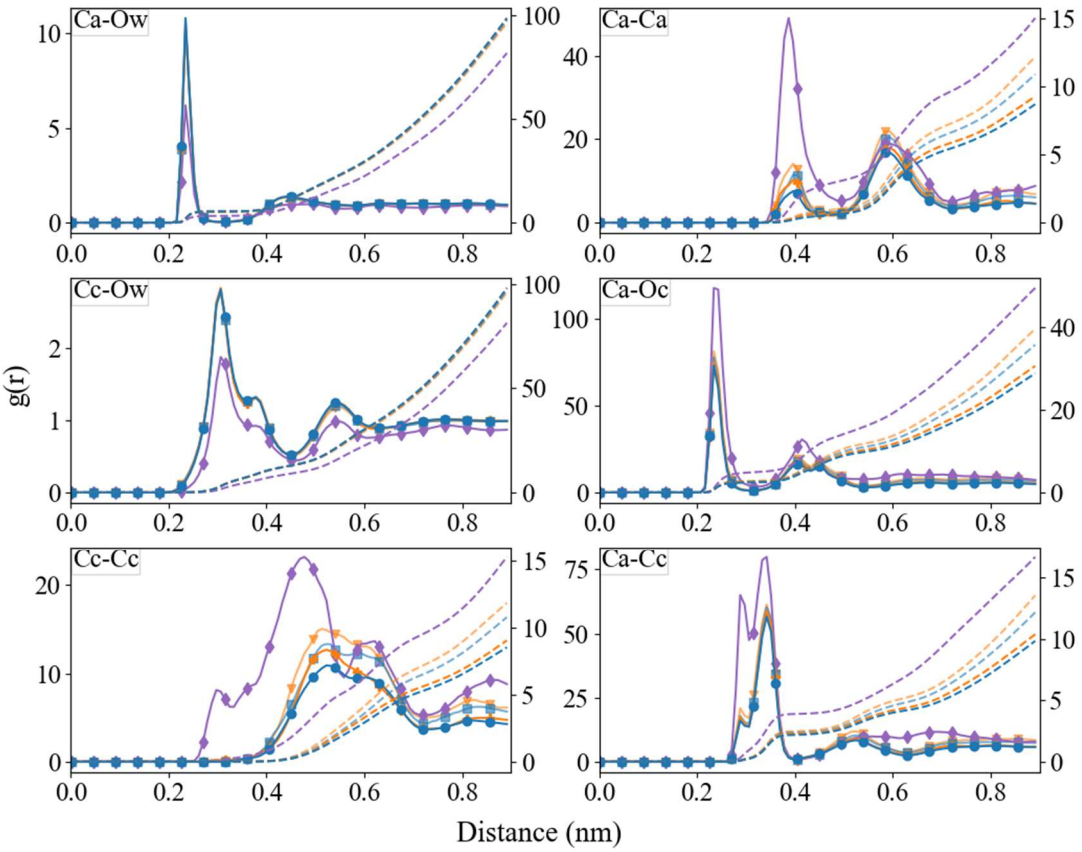



Figure 5.9 Radial distribution function (left axis) and corresponding integral (N(r)) (right axis) between $Ca^{2+}$, $C_c$, $O_c$, and $O_w$, after 30 ns of simulation. Pure (purple; ♦), L-Asp with Set_1$_{SPC/fw}$ (transparent blue; ■), L-Asp Set_2$_{+0.5\%\sigma}$ (blue; ●), D-Asp Set_1$_{SPC/fw}$ (transparent orange; ▼), D-Asp Set_2$_{+0.5\%\sigma}$ (orange; ✚)

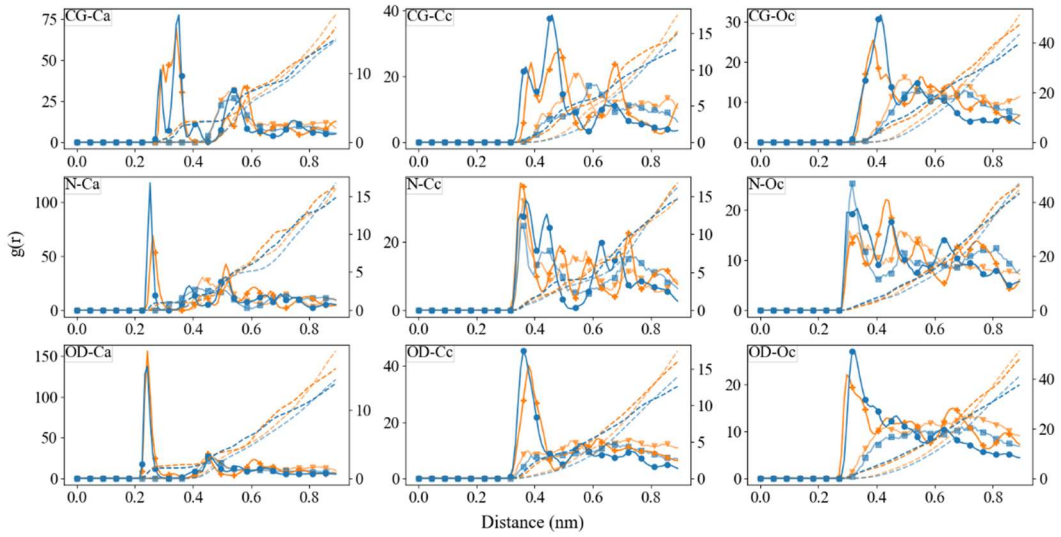



Figure 5.10 Radial distribution function between $Ca^{2+}$, $C_c$, $O_c$, with the different functional groups in aspartic acid after 30 ns of simulation. L-Asp with Set_1$_{SPC/fw}$ (transparent blue; ■), L-Asp Set_2$_{+0.5\%\sigma}$ (blue; ●), D-Asp Set_1$_{SPC/fw}$ (transparent orange; ▼), D-Asp Set_2$_{+0.5\%\sigma}$ (orange; ✚)

## 5.7 Discussion

### 5.7.1 Interaction of calcium with carboxyl groups in aspartic and glutamic acid

In the past years, several studies have pointed out the importance of a re-evaluation of the parametrization of the interactions of charged amino-acid side chains with water and ions to achieve a more accurate description of the system (Kahlen et al., 2014; Project et al., 2008; Reif et al., 2012). In previous simulations that used AMBER based force field parameters [Koskamp et al., 2021b] (Innocenti Malini et al., 2017), the conclusion was that it is mainly the amine group that interacts with the $CO_3^{2-}$ or an ACC surface instead of the carboxyl group, which was expected to be the main interacting functional group based on experimental evidence. The current study into the carboxyl-oxygen-calcium interaction revealed that Set_1

has one of the highest repulsion terms in the Lennard-Jones potential, and although the attraction term is also higher, the overall result is a stronger repulsion in the other potentials parameters (Åqvist, 1990; Babu and Lim, 2006; Beglov and Roux, 1994; Dang and Smith, 1995; Kohagen et al., 2014; Mamatkulov et al., 2013; Martinek et al., 2018; Shen et al., 2013; Wang et al., 2008). This has implications for simulating single amino acids and larger biomolecules that are rich in aspartic and/or glutamic acid (Freeman et al., 2015; Picker et al., 2012),[ Koskamp et al., 2021b]. Another description of Ca, derived from a refined $CaCO_3$-water$_{SPC/fw}$ force field (Raiteri et al., 2015) is on the other side of the spectrum with the lowest repulsion and attraction values, yet the ratio is in the same order of magnitude as the Ca in Set_1, explaining the comparable results for similar simulations (Raiteri et al., 2015), [Koskamp et al., 2021b].

A very common water model in simulations with biomolecules is the TIP3P water models (Mínguez-Toral et al., 2020; Solanke et al., 2019; Wahedi et al., 2020; Wang et al., 2020; Wolf et al., 2020). The energy profile obtained for the Ca description from Set_1 with TIP3P water overestimates the association constant of Ca with both amino acids by nine orders of magnitude, as seen before in AMBER and CHARMM based force fields (Church et al., 2015; Yoo et al., 2016). This is mainly due to lack of well-defined, explicit parameterization of the interaction between the amino acid and calcium (Church et al., 2015). However, this interaction is able to reproduce the different configurations upon association (Chialvo et al., 1995). Unfortunately,TIP3P has been shown to have a self-diffusion coefficient twice the experimental value (Koskamp et al., 2019). Moreover, other water models have been considered more suitable for $CaCO_3$ systems in terms of water dynamics including mean residence time and water structuring (De La Pierre et al., 2016; Fenter et al., 2013; Koskamp et al., 2019). For this reason, in our previous work [Koskamp et al., 2021b] we combined the parameters from Set_1 with SPC/fw water. With this force field, we obtained an energy profile

very comparable to previously published results (Innocenti Malini et al., 2017). The only difference was the choice of a Buckingham potential, in their study, over a less expensive LJ potential to describe the interaction between Ca and biomolecule. A Buckingham potential is proven to be wider leading to a less structured solvation shell (Migliorati et al., 2017). This therefore contributes to the difference of -0.4 kJ/mol between their (Innocenti Malini et al., 2017) and our results [Koskamp et al., 2021b] (Table 5.2). Unlike the energy profile of Set_1$_{TIP3P}$, the Set_1$_{SPC/fw}$ force field is unable to capture the structural configurations of the ion pair and shows with global minimum at ~0.466 nm indicating that the SIP configuration would be preferred over CIP. Nevertheless, Set_1$_{SPC/fw}$ shows overall improved -$\Delta G_a$ (Table 5.2), which is why, in the current work, we decided to refine this force field to restore the carboxyl-Ca configuration, as known from experiments.



In the refined Set_2, further improved agreement between estimated -$\Delta G_a$ and experimental values (Table 5.2) was found (Covington and Danish, 2009; Raiteri et al., 2012). Note that, with this refined force field, the Ca – $O_{carboxylate}$ distance is slightly shorter (Marchand and Roux, 1998; Raiteri et al., 2012; Tang and Skibsted, 2016). When increasing the σ values of the LJ interaction between Ca – $O_{carboxylate}$ further, a systematic decrease in well-depth is observed, comparable to other studies using the same method for glutamic acid, γ-carboxyglutamic acid (CHARMM) (Church et al., 2015). In the case of aspartic acid an increase of σ of 0.5% is sufficient to be within the reported experimental values for the association constant and closer to the latest value for Ca-OD bond distance (Covington and Danish, 2009; Raiteri et al., 2012).

With the refined force field for aspartic acid, the distance between carboxyl-oxygen and calcium corresponds to the monoCIP structure, although the difference between monoCIP and biCIP minima is, with 1.22 kJ/mol compared to 4.79 kJ (Set_2$_{unmodified}$), the smallest for Set_2$_{+0.5\%\sigma}$. In terms of stability, this means that a monoCIP configuration is favoured over a

biCIP. This agrees with experimental observations using $^{13}C$ nuclear magnetic resonance (NMR) (Kondoh and Oi, 1998). Moreover, the water structure, in terms of RDF and hydrogen bond angle around the $O_{carboxylate}$) compares well with those previously reported for $O_{carboxylate}$ in acetate (Kahlen et al., 2014). In short, the refined force field captures known interaction energies and structures for calcium with the carboxyl-group of aspartic acid.

For glutamic acid, the association energy is lower compared to aspartic acid: due to the extra $CH_2$ group, the charge density in Glu is lower, therefore the attraction of $Ca^{2+}$ is weaker. Although the unmodified interaction gives a $-\Delta G_a$ that better compares to values obtained with an unmodified CHARMM based forcefield(Church et al., 2015), it still slightly overestimates experimental observations (Covington and Danish, 2009) and refined computational observations (Church et al., 2015). Increasing σ with 1 % agrees best with experiments and computations in terms of $-\Delta G_a$, yet it loses structural information between SIP and SSIP. Additionally, the well is slightly shallower around CIP distance, indicating a lower close-contact association energy. The structural information is preserved for Set_$2_{+0.5\%\sigma}$ which gives the best match with other simulations. In all cases the biCIP is more stable compared to monoCIP, this is opposite to aspartic acid but in agreement with other MD simulations for Ca-Glu interactions (Church et al., 2015).

In the first step towards association of a dissolved calcium ion with another (dissolved) biomolecule such as Asp or Glu, Ca must lose one of its ~8.3 water molecules. This first transition state can be characterized with coordination number 7. The most probable mechanism for this first step is that a water molecule leaves the solvation shell before Ca can approach closer. This in contrast with the second step, where the transition state upon full association with the biomolecule is rather broad: it ranges from coordinated to 5 to 7 water molecules (see e.g. Figure 5.5). Only when in contact with a dissolved biomolecule, Ca can drop another water molecule, although it is not the most stable state (Figure 5.5). All three

associative, dissociative and interchange exchange mechanisms, as previously described (Helm and Merbach, 1999), are possible Ca association mechanism and observed here for Asp and Glu (Figure 5.5). There is a debatable preference for an associative mechanism, considering the deepest wells at a horizontal line around coordination 5.5 for both aspartate and glutamate.

### 5.7.2 Implications for ACC cluster formation in presence and absence of aspartic acids

The simulation of ACC cluster formation in the presence of D and L-Asp shows some clear differences between the enantiomers. When the refined force field is used for the simulations, the stronger attraction of Ca with the carboxyl groups resulted in clear changes in RDFs between Ca and the functional group of both enantiomers. The bi and monoCIP separation is more abundant in L-Asp than the D- form, just as in the simpler calcium-aspartic acid simulations. In D-Asp the formation and growth of clusters is faster than the unmodified interaction yet appears to follow the same distribution pattern [Koskamp et al., 2021b], and has a higher density around the biomolecule. The final cluster size in presence of L-Asp is ~125 ions smaller which correspondents to 31% of the cluster size compared to the pure system. This seems contradicting the findings in experiments where they reported only minimal effect of aspartic acid polymers on the particle size (Zou et al., 2018). However, this difference of 31% is most likely due to the small system size and time-period of the simulation. Both for L- and D-Asp systems, the particles formed are somewhat enriched in carbonate, and therefore more negatively charged, with ~1% more carbonate than previously observed for L-Asp [Koskamp et al., 2021b]. This potentially leads to a slightly lower density, since the charged particles are likely stabilized with water, as reflected by the higher SASA (Figure 5.7).

Overall, the aspartic acid enantiomers have the same impact on ACC cluster formation as reported previously [Koskamp et al., 2021b] and cluster stabilization observed in computational (Finney and Rodger, 2012) and experimental work (Picker et al., 2012).

Compared to the pure system, ACC clusters in the presence of Asp formed more slowly, predominantly by accumulation of ions and/or ion pairs, remained more hydrated and were nonstoichiometric, composed of ~4% (D-Asp) and 5% (L-Asp) more carbonate than calcium, resulting in a net negative charge of the clusters.

Other implications of the insights presented in this work for biomineralization can be for example related to the role of biomolecules in combination with impurities like Mg. Previous studies have shown that the presence of biomolecules accelerate the growth rate of a calcite surface (Stephenson et al., 2008). Our simulations showed that L/D-aspartic acid has a stabilizing effect on the formation of hydrated ACC and could therewith provide a rather fluid-like/malleable reservoir of calcium and carbonate that can facilitate in a fast attachment and growth of the crystal. Furthermore, in the light of impurity uptake in the crystal, it is found that Mg is more strongly affected by the biomolecule compared to Ca (Stephenson et al., 2008). A plausible mechanism based on our findings could be as follows: the presence of aspartic acid leading to $Ca^{2+}$: $CO_3^{2-}$ ratio smaller than one, confirming the more negative zeta-potential values observed when aspartic acid polymers are present (Zou et al., 2017). The excess of carbonate or exclusion of Ca creates a negative ion-cloud that could enhance the uptake of smaller and higher charge dense Mg. It could therefore change the Ca/Mg ratio close to the growth side and therewith the Mg signature in the crystal, as observed for peptides experimentally (Stephenson et al., 2008).



## 5.8 Conclusion

In this study we investigated the interaction of Ca with aspartic and glutamic acid with classical MD using various refined force fields. We can conclude that it is of great importance to choose the parameters carefully so that one is able to describe the association of the ions and the different structural stages upon ion pairing. Using the accepted parameters from Set_1, the

force field is not able to describe the association energy, when done with TIP3P water, or the structural details, when using SPC/fw, correctly. Replacing the description of Ca to the one used in $CaCO_3$ improved the force field considerably, yet small modifications on the specific Ca – $O_{carboxylate}$ parameter were necessary to overcome the overestimation of the free energy of association. We therefore corrected the heteroatomic LJ parameters between $Ca^{2+}$ and $O_{carboxylate}$ by increasing $\sigma_{ij}$ by 0.5%. This slight increase adequately enhanced the binding distance and the $\Delta G_a$ without disrupting the solvent structure around the biomolecule and the ion. This new description shows a strong impact in more complex systems in where Ca-binding is now favoured over $CO_3^{2-}$, contradicting previously reported results while supporting experimental findings. Therefore, we recommend using Set_$2_{+0.5\%\sigma}$ interaction for Ca – $O_{carboxylate}$ when conducting MD simulations of systems including $Ca^{2+}$ and AMBER-based biomolecules with carboxyl functional groups.

## 5.9 Supplementary information

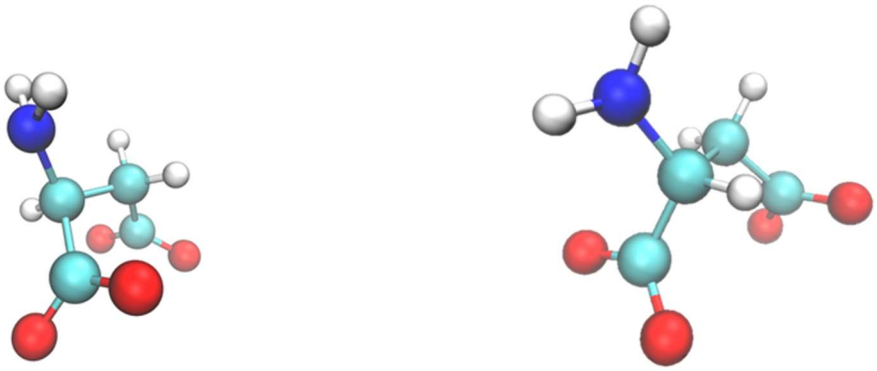
### 5.9.1 Force field

Table S5.1 Inter and intraatomic potential parameters.

| **Water (SPC/fw)** | | | | | |
|---|---|---|---|---|---|
| **Species** | | **Molecule/Atom** | | **Charge (e)** | |
| Ca | | Calcium | | +2.000 | |
| $C_c$ | | Carbon from carbonate | | +1.135 | |
| $O_c$ | | Oxygen from carbonate | | -1.045 | |
| $O_W$ | | Oxygen from water | | -0.820 | |
| $H_W$ | | Hydrogen from water | | +0.410 | |
| **Bond styles** | | ***k* (ev Å$^{-2}$)** | | ***r*$_0$ (Å)** | |
| $C_c$—$O_c$ | Harmonic | 40.8493 | | 1.012 | |
| $O_W$—$H_W$ | Harmonic | 45.9296231 | | 1.012 | |
| **Angle styles** | | ***k* (ev rad$^{-2}$)** | | **Θ$_0$** | |
| $H_w$— $O_W$ — $H_W$ | Harmonic | 3.29134 | | 113.24 | |
| | | **Θ$_0$** | ***K*$_2$ (ev rad$^{-2}$)** | ***K*$_3$ (ev rad$^{-2}$)** | ***K*$_4$ (ev rad$^{-2}$)** |
| $O_c$— $C_c$ — $O_c$ | class2 | 120.0 | 6.617 | 0.0 | 0.0 |
| | | **M (ev Å$^{-2}$)** | ***r*$_1$ (Å)** | ***r*$_2$ (Å)** | |
| $O_c$— $C_c$ — $O_c$ | class2 bb | 12.818 | 1.3042 | 1.3042 | |
| | | **N$_1$ (ev Å$^{-2}$)** | **N$_2$ (ev Å$^{-2}$)** | ***r*$_1$ (Å)** | ***r*$_2$ (Å)** |
| $O_c$— $C_c$ — $O_c$ | class2 ba | 1.53319 | 1.53319 | 1.3042 | 1.3042 |
| **Improper styles** | | ***K*$_2$ (ev rad$^{-2}$)** | | ***K*$_4$ (ev rad$^{-2}$)** | |
| $O_c$— $C_c$ — $O_c$-$O_c$ | distance | 13.647 | | 360.0 | |
| **Interatomic interactions** | | | | | |
| **Lennard-Jones Potential** | | **ε (eV)** | | **σ (Å)** | |
| $O_W$ — $O_W$ | | 0.006739769454 | | 3.165492 | |
| Ca—$O_W$ * | | 0.000950 | | 3.35 | |
| **Buckingham Potential** | | **A (eV)** | **ρ (Å)** | **C (eV Å$^6$)** | |
| $O_c$—$O_W$ * | | 12534.455133 | 0.202 | 12.09 | |
| $O_c$—$H_W$ * | | 396.0 | 0.217 | 0.0 | |
| Ca—$O_c$* | | 3161.6335 | 0.271511 | 0.0 | |
| Ca—$C_c$ * | | 120000000 | 0.12 | 0.0 | |
| $O_c$— $O_c$* | | 63840.199 | 0.198913 | 27.89901 | |

*The interaction was set to zero over the range 6 – 9 Å

### 5.9.2 Supplementary figures



***D-aspartic acid*** ***L-aspartic acid***

Figure S5.1 Molecular structure of the studied biomolecules. Nitrogen (blue), Carbon (cyan), Oxygen (red) and Hydrogen (white).



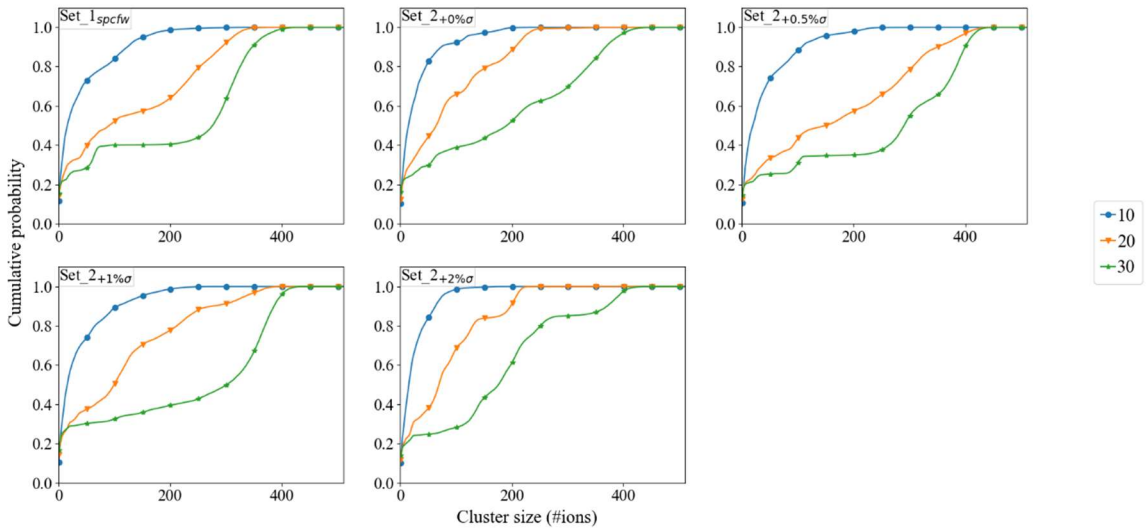



Figure S5.2 Cumulative probability profiles of different $CaCO_3$ systems for every 10 ns of simulation using different forcefields: $Set\_1_{SPC/fw}$, $Set\_2_{+0\%\sigma}$, $Set\_2_{+0.5\%\sigma}$, $Set\_2_{+1\%\sigma}$, $Set\_2_{+2\%\sigma}$ with **D-Asp**.

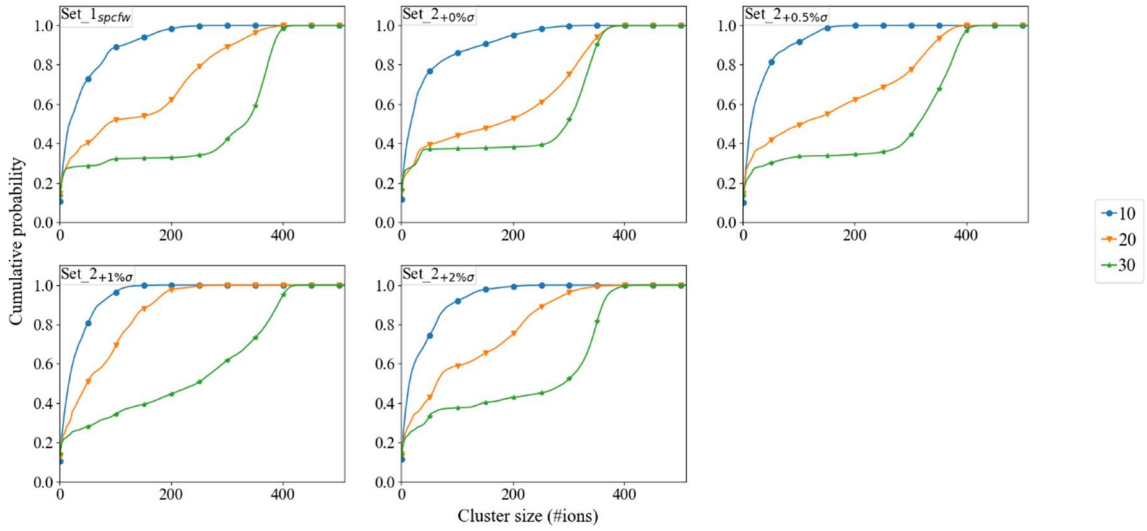



Figure S5.3 Cumulative probability profiles of different $CaCO_3$ systems for every 10 ns of simulation using different forcefields: $Set\_1_{SPC/fw}$, $Set\_2_{+0\%\sigma}$, $Set\_2_{+0.5\%\sigma}$, $Set\_2_{+1\%\sigma}$, $Set\_2_{+2\%\sigma}$ with **L-Asp**.

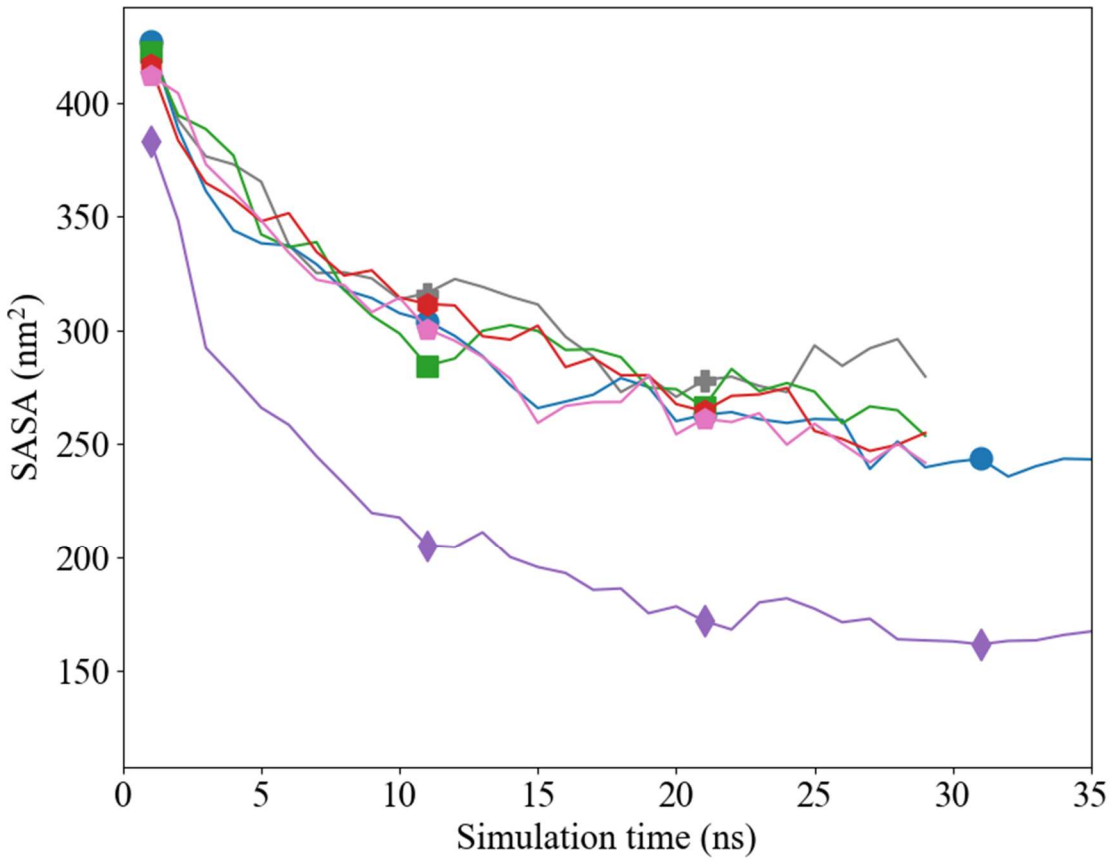



Figure S5.4 SASA analysis using the Shrake and Rupley algorithm of pure (♦) and **D-Asp** systems using different forcefields**;** Set_1$_{SPC/fw}$ (blue; ●), Set_2$_{+0‰σ}$ (pink; ⬟), Set_2$_{+0.5‰σ}$ (red; ⬢), Set_2$_{+1‰σ}$ (green; ■), and Set_2$_{+2‰σ}$ (gray; ✚).

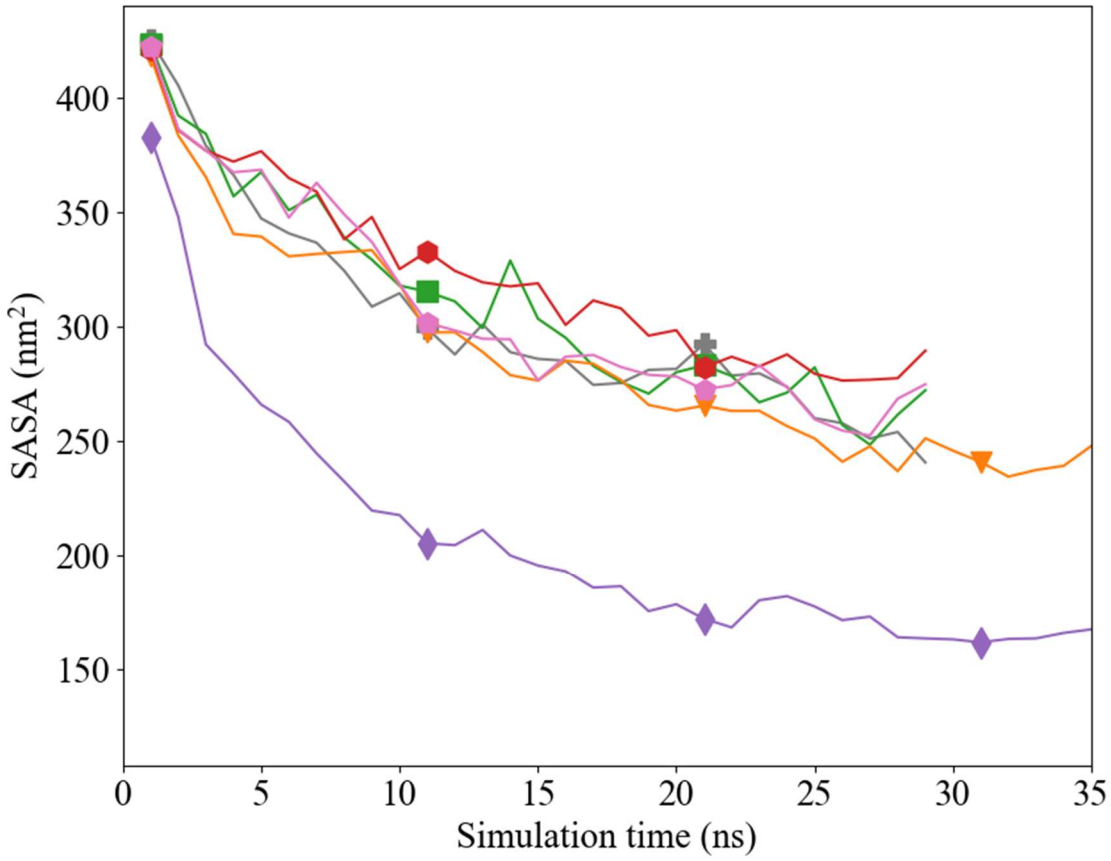



Figure S5.5 SASA analysis using the Shrake and Rupley algorithm of pure (♦) and **L-Asp** system using different forcefields**;** Set_1$_{SPC/fw}$ (blue; ●), Set_2$_{+0\%\sigma}$ (pink; ⬟), Set_2$_{+0.5\%\sigma}$ (red; ⬢), Set_2$_{+1\%\sigma}$ (green; ■), and Set_2$_{+2\%\sigma}$ (gray; ✚).

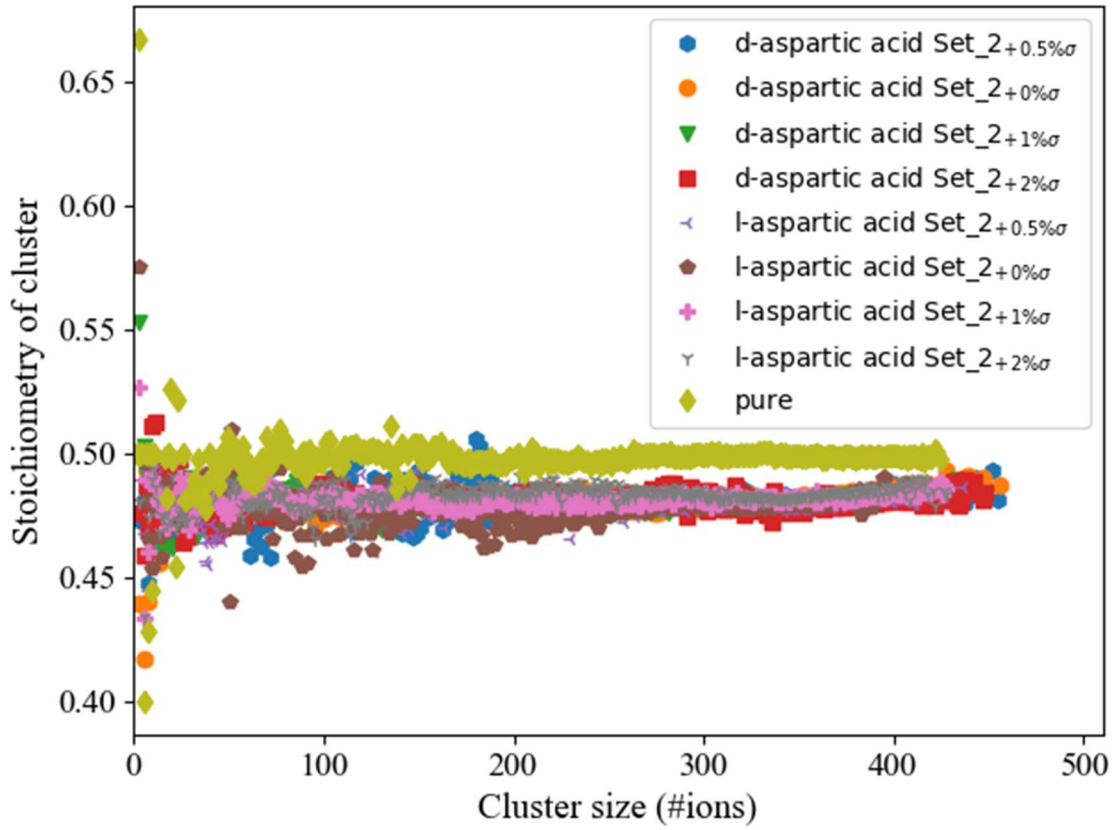



Figure S5.6 Average cluster stoichiometry ($Ca^{2+}\cdot(Ca^{2+}+CO_3^{2-})^{-1}$) over 30 ns of simulation.

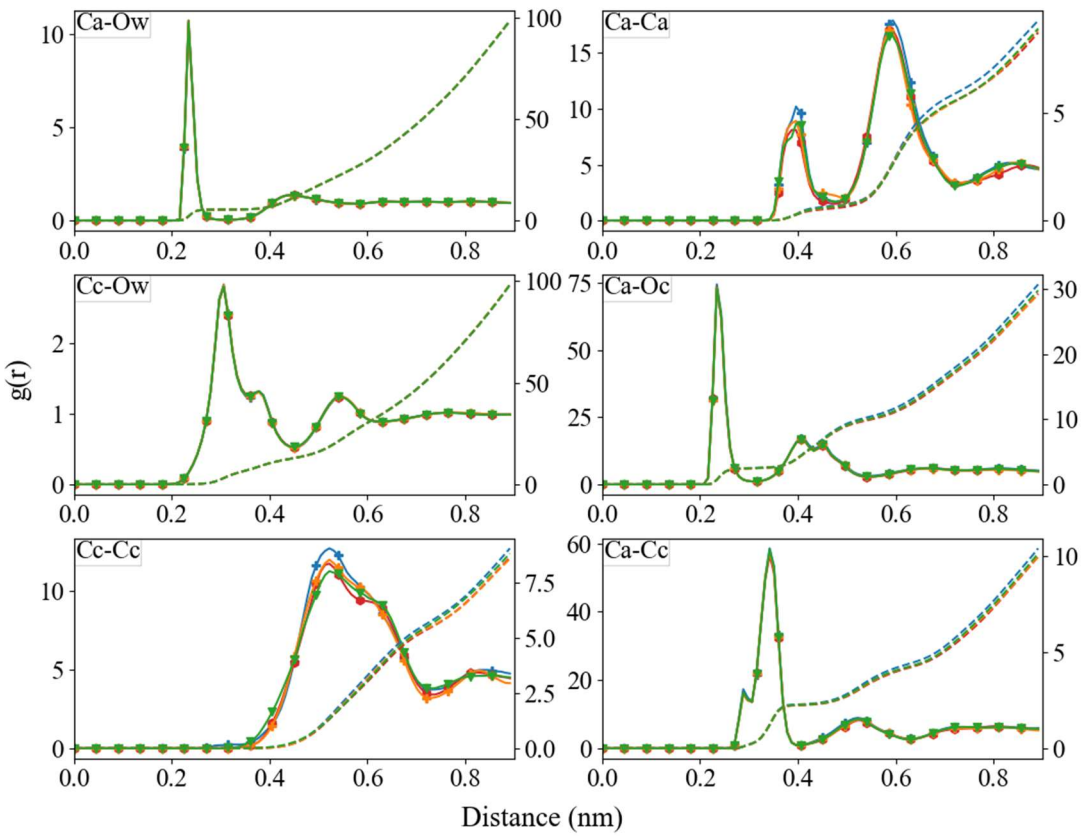



Figure S5.7 Radial distribution function (left axis) and corresponding integral (N(r)) (right axis) between $Ca^{2+}$, $C_c$, $O_c$, and $O_w$, after 30 ns of simulation. **D-Asp** using different forcefields**;** Set_2$_{+0\%\sigma}$ (red; ⬢), Set_2$_{+0.5\%\sigma}$ (blue; ●), Set_2$_{+1\%\sigma}$ (orange; ✚), and Set_2$_{+2\%\sigma}$ (green; ▼)

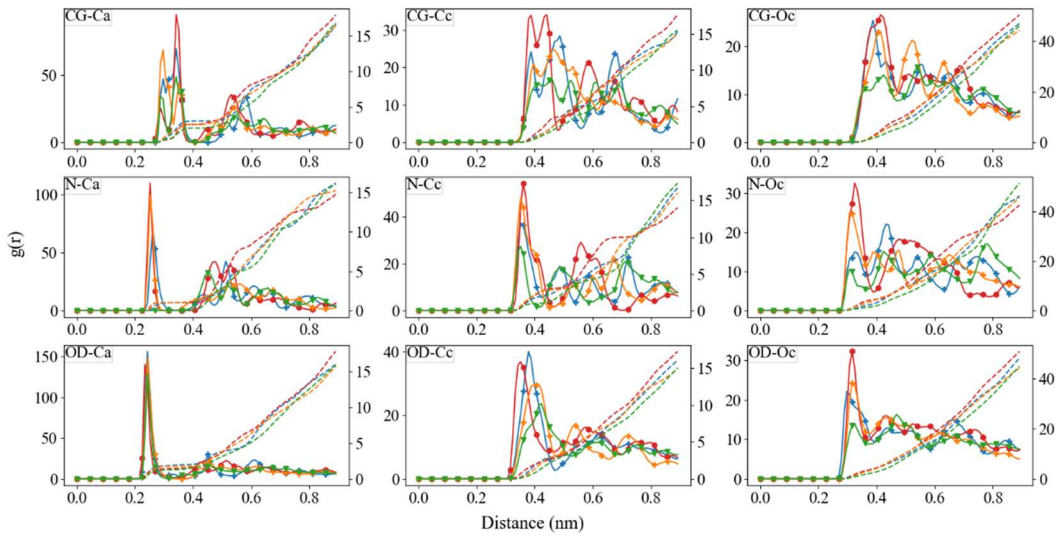



Figure S5.8 Radial distribution function between $Ca^{2+}$, $C_c$, $O_c$, with the different functional groups in aspartic acid after 30 ns of simulation. **D-Asp** using different forcefields**;** Set_2$_{+0\%\sigma}$ (red; ⬢), Set_2$_{+0.5\%\sigma}$ (blue; ●), Set_2$_{+1\%\sigma}$ (orange; ✚), and Set_2$_{+2\%\sigma}$ (green; ▼)

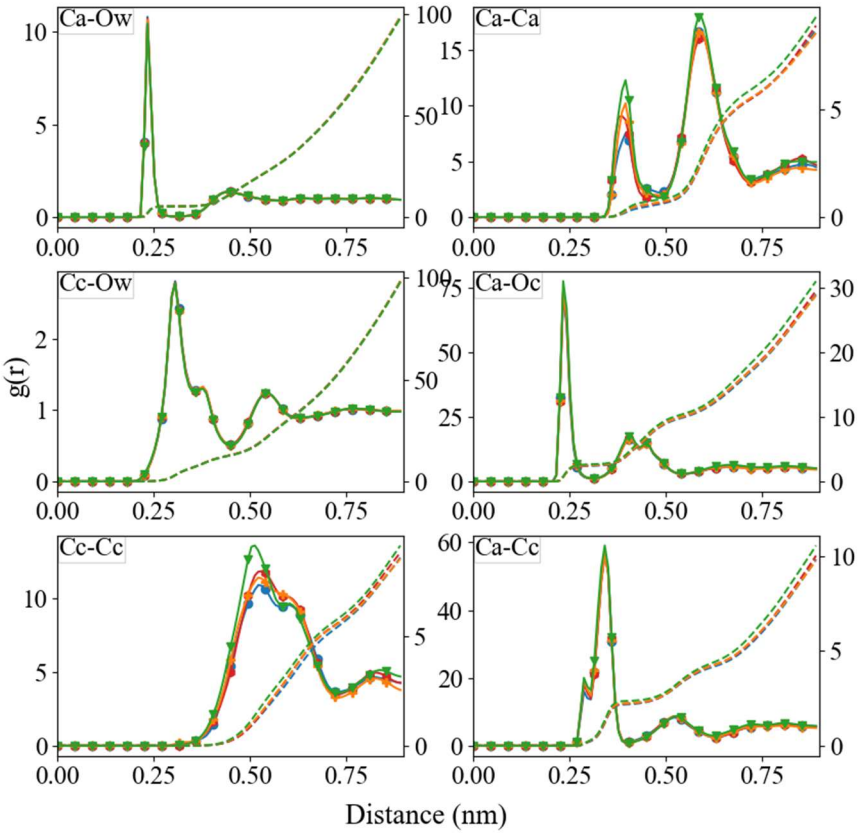



Figure S5.9 Radial distribution function (left axis) and corresponding integral (N(r)) (right axis) between $Ca^{2+}$, $C_c$, $O_c$, and $O_w$, after 30 ns of simulation. **L-Asp** using different forcefields**;** Set_2$_{+0\%\sigma}$ (red; ⬢), Set_2$_{+0.5\%\sigma}$ (blue; ●), Set_2$_{+1\%\sigma}$ (orange; ✚), and Set_2$_{+2\%\sigma}$ (green; ▼)

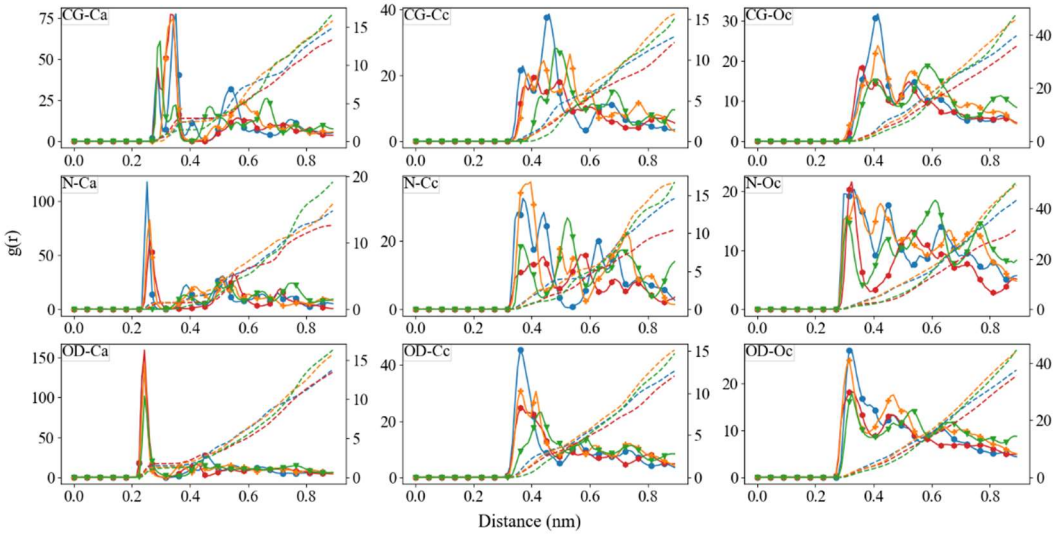



Figure S5.10 Radial distribution function between $Ca^{2+}$, $C_c$, $O_c$, with the different functional groups in aspartic acid after 30 ns of simulation. **L-Asp** using different forcefields**;** Set_2$_{+0\%\sigma}$ (red; ⬢), Set_2$_{+0.5\%\sigma}$ (blue; ●), Set_2$_{+1\%\sigma}$ (orange; ✚), and Set_2$_{+2\%\sigma}$ (green; ▼)

# Chapter 6 First Steps towards Understanding the Non-Linear Impact of Mg on Calcite Solubility: A Molecular Dynamics Study



Accepted as: **Koskamp J. A., Ruiz-Hernandez S. E., De Leeuw N. H., Wolthers, M.** First steps towards understanding the non-linear impact of mg on calcite solubility: A molecular dynamics study. *Minerals*

**Abstract:** Magnesium ($Mg^{2+}$) is one of the most common impurities in calcite and is known to have a non-linear impact on the solubility of magnesian calcites. Using molecular dynamics (MD), we observed that $Mg^{2+}$ impacts overall surface energies, local free energy profiles, interfacial water density, structure, and dynamics and, at higher concentrations, it also causes crystal surface deformation. Low Mg concentrations did not alter the overall crystal structure, but stabilised $Ca^{2+}$ locally and tended to increase the etch pit nucleation energy. As a result, Ca-extraction energies over a wide range of 39 kJ/mol were observed. Calcite surfaces with an island were less stable compared to flat surfaces, and the incorporation of $Mg^{2+}$ destabilised the island surface further, increasing the surface energy and the calcium extraction energies. In general, $Ca^{2+}$ is less stable in islands of high $Mg^{2+}$ concentrations. The local variation in free energies depends on the amount and distance to nearest Mg in addition to local disruption of interfacial water and the flexibility of surface carbonate ions to rotate. The result is a complex interplay of these characteristics that cause variability in local dissolution energies. Taken together, these results illustrate molecular scale processes behind the non-linear impact of $Mg^{2+}$ concentration on the solubility of magnesium-bearing calcites.

## 6.1 Introduction

Calcium carbonate minerals, and in particular calcite, are extensively found in nature, for example as the main component in marble, limestone, and chalk, and it is therefore one of the most abundant non-silicate minerals at the Earth's surface. Furthermore, calcite has proven its functionality and usability in different contexts in both natural and industrial processes (Ahr, 2008). To enhance the understanding of the mechanisms behind observations on the formation or dissolution of calcite, theoretical models describe calcite growth and dissolution processes at many different scales. Pore-scale and continuum scale models are becoming more and more sophisticated, yet there are still important discrepancies between model outcomes and experimental observations, e.g., (Agrawal et al., 2020; Cao et al., 2018; Molins et al., 2020, 2014). Similarly, statistical (e.g., (Hellevang et al., 2014)), kinetic (e.g., (Kurganskaya and Luttge, 2016)) and ion-by-ion (e.g., (Bracco et al., 2012; Larsen et al., 2010; Sand et al., 2016; Wolthers et al., 2012b)) models have also revealed discrepancies between model and experimental observations. Possible reasons for those discrepancies are further elucidated using atomistic-scale approaches and ab initio models, for example in (Fenter et al., 2013; Kirch et al., 2018; Lardge et al., 2010).



Impurities are more the rule than the exception in natural calcite and are commonly found not only in different concentrations but also in a range of different types of impurities (Harstad and Stipp, 2007). One of the most common impurities found in calcite crystals is the magnesium ion, $Mg^{2+}$. Magnesium carbonate (magnesite) is isomorphic with calcite and a wide range of solid solutions are therefore formed between these endmembers. Generally, two categories are employed to differentiate between low and high $Mg^{2+}$ concentrations, with arbitrary limits (Lenders et al., 2012; Long et al., 2014; Müller and Tietz, 1966; Wang et al., 2012), i.e., low-$Mg^{2+}$ calcite in which a maximum of 2% of the $Ca^{2+}$ is replaced by $Mg^{2+}$, and high-$Mg^{2+}$ calcite

with 8% to 40% of $Mg^{2+}$ substituted for $Ca^{2+}$ (Stanienda, 2016). Beyond 40% we cannot speak of calcite anymore but should consider the mineral as dolomite or huntite instead.

According to computer simulations, the incorporation of magnesium impurities into the step edges of a calcite surface is initially thermodynamically very favourable and is independent of the growth site's topography (obtuse or acute edges) (N. H. De Leeuw, 2002; de Leeuw and Parker, 2001). It has been shown experimentally that during calcite growth $Mg^{2+}$ incorporates more into acute edges than obtuse steps (Paquette and Reeder, 1995), although in more recent computer simulations it was found that $Mg^{2+}$ is randomly distributed over the calcite crystal (Wang, 2011). Surface studies have also shown that high concentrations of impurities in the growth edge will poison the edge and thereby inhibit further crystal growth (N. H. De Leeuw, 2002; Nora H. De Leeuw, 2002; de Leeuw and Parker, 2001). Furthermore, it was shown that it is thermodynamically unfavourable to sequentially incorporate $CaCO_3$ units next to a $MgCO_3$ edge corner in a growing calcite crystal (N. H. De Leeuw, 2002; de Leeuw and Parker, 2001). In contrast, a mechanism that was shown to be thermodynamically feasible is to incorporate magnesium via diffusion of $Mg^{2+}$ towards the surface, followed by $Ca^{2+}$–$Mg^{2+}$ exchange with an existing surface calcium site (Nora H. De Leeuw, 2002). Experimental work has shown an increase in initial growth rate when a metallic ($Mn^{2+}$, $Zn^{2+}$, $Co^{2+}$ and $Cd^{2+}$) ion was adsorbed or exchanged at the calcite surface (Brady et al., 1996). Other studies concluded that calcium units were dissolving faster than $Mg^{2+}$ units (Busenberg and Plummer, 1982; Pokrovsky and Schott, 2001), resulting in non-stoichiometric (incongruent) dissolution. Other, more recent, experimental work has concluded that the dissolution rate of calcite is strongly but non-linearly dependent on the $Mg^{2+}$ concentration in the crystal (Subhas et al., 2018).

Arguably one of the most difficult variables to capture with dissolution rate laws is the impact of impurity ions such as $Mg^{2+}$ in the solution and in the dissolving calcite crystal. Experimental work using atomic force microscopy (AFM) (Astilleros et al., 2010; Ruiz-Agudo and Putnis,

2012; Xu and Higgins, 2011; X. Zhang et al., 2020) has shown alterations in crystal geometries as a result of variations in dissolution velocities of calcite due to the incorporated magnesium and its distinct preference for (interactions with) different surface sites. Additionally, the effect of salt ions on water structure and dynamics (Di Tommaso et al., 2014) and the consequence for calcite dissolution, with an increase in water viscosity and a resulting decrease in diffusion coefficients (Kirch et al., 2018; Wasylenki et al., 2005), has been widely suggested as an explanation for observed variations in dissolution behaviour.

Molecular dynamics (MD) simulations have been used extensively to study reactivity of calcite crystals under different conditions. Among others, these studies have investigated the importance of variability in reactivity of edges and kink sites on a $\{10\bar{1}4\}$ surface and defined rate equations in up-scaled models (De La Pierre et al., 2017, 2016; de Leeuw et al., 1999; Spagnoli et al., 2006b; Wolthers et al., 2013) or the impact of confinement and solution composition (Kirch et al., 2018; Koleini et al., 2018; Mutisya et al., 2017; Spagnoli et al., 2006a). Moreover, MD was used to study the free energy landscape of adsorption and dissolution of metal and carbonate ions on calcite surfaces (De La Pierre et al., 2017; Kerisit and Parker, 2004; Spagnoli et al., 2006b). As in experiments (Bracco et al., 2012; Larsen et al., 2010; Liang et al., 1996) and unbiased MD simulations (De La Pierre et al., 2016; N. H. De Leeuw, 2002; Kerisit et al., 2003; Wolthers et al., 2013), free energy studies do not show consensus in their observations on the acute and obtuse edges (Andersson et al., 2016; De La Pierre et al., 2017; de Leeuw et al., 1999).



The aim of this work is to provide atomistic insight into the non-linear impact of $Mg^{2+}$ on the solubility of calcite, and to contribute to the understanding of the mechanism(s) in place and responsible for the variable reactivity with Mg concentration in calcite. In the literature, the focus has been on explaining the reactivity of $Mg^{2+}$ and its poisoning effect on the reactivity (Davis et al., 2000; N. H. De Leeuw, 2002; de Leeuw and Parker, 2001; Nielsen et al., 2013;

Xu et al., 2013). Using molecular dynamics, simulating $Mg^{2+}$ rich systems is complicated due to the slow exchange rates of water around this cation, necessitating timescales beyond what is currently feasible to gather statistically meaningful data. Consequently, we describe the reactivity of the crystal surface in terms of stability of the surface calcium ions instead. We have studied the residence time of water coordinated to surface $Ca^{2+}$, interfacial water structure, the water diffusion, the deformation of the crystal, the surface energies and eventually the free energy profiles of the extraction of surface calcium ions. Our results reveal that $Mg^{2+}$ incorporation at a low concentration does lead to a local stabilisation of surface $Ca^{2+}$, whereas a higher Mg concentration disturbs the crystal structure, leading to local variabilities in the crystal surface, which are expressed in local destabilisation of $Ca^{2+}$, thereby facilitating the dissolution of an adjacent $CO_3^{2-}$ and enhancing dissolution [39] compared to pure calcite. At the macroscopic scale, these observations imply a lower solubility than pure calcite when the concentration of Mg in calcite is low (<2%) and a higher solubility with increasing Mg, in agreement with the non-linear solubility observed experimentally.

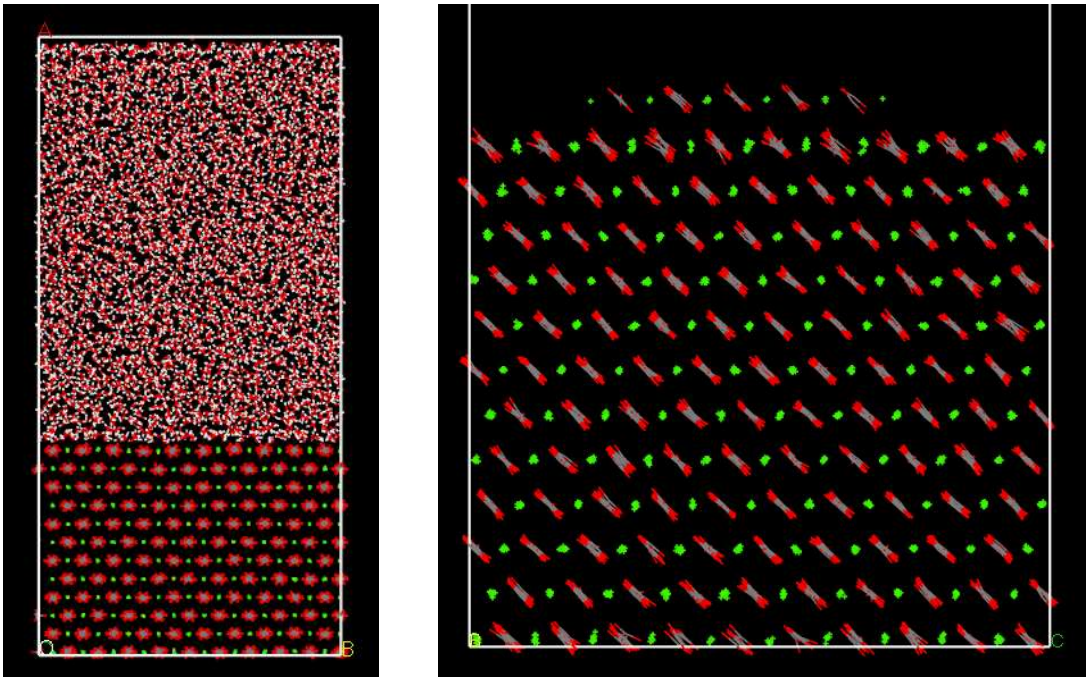


## 6.2 Methods

We simulated two different calcite $\{10\bar{1}4\}$ crystal surfaces in contact with pure water. One of the crystals had an island of 16 crystal cation-anion units adsorbed onto the flat $\{10\bar{1}4\}$ surface and an etch pit of the same size on the opposite side of the crystal slab, whereas the other surface was atomically flat. Since the edges of the small pit are prone to interact (Wolthers et al., 2013), we have focused on the edges of an island that can also represent the edges of a bigger pit. Both crystals were simulated with and without 30% magnesium randomly distributed over the entire crystal slab, representing the high $Mg^{2+}$ calcite system. In addition, we simulated a flat $\{10\bar{1}4\}$ surface with one magnesium impurity in the surface, representing the low $Mg^{2+}$ calcite. In total we thus had five different molecular systems. The classical molecular dynamics (MD) simulations were performed in the large-scale atomic/molecular

massively parallel simulator (LAMMPS) (Plimpton, 1995). We used the velocity-Verlet integrator (Hoover, 1985; Nosé, 1984) and the Nosé-Hoover thermostat (Hoover, 1985; Nosé, 1984) to integrate the equations of motion with a time step of 1 fs and to maintain the temperature at 300 K. The simulations were carried out in constant number of particles, volume and temperature (NVT) ensemble with a relaxation time for the thermostat of 0.1 ps. The unbiased simulation time was 10 ns with 1 ns of equilibration period.

**System details.** The calcite slab was 3.3 nm thick, which corresponds to 12 atomic layers of material, and was previously shown to be thick enough to have no interactions between the surfaces through the crystal slab (Fenter et al., 2013). A surface of 4.8 nm by 4.0 nm was exposed to a water column of 6.5 nm ensuring bulk-like water behaviour in the centre of the column. After an initial equilibration period, eight out of the twelve layers in the centre of the crystal were immobilised and kept in the same position, for two reasons: (i) to speed up the calculations; and (ii) to simulate the bulk crystal and prevent the box from deviating from its reference point, so comparison of the different Mg-doped systems was facilitated. Initially, we exchanged a randomly selected single $Ca^{2+}$ by a $Mg^{2+}$ in the surface of the material, to be used as a reference of low $Mg^{2+}$ concentrations (LMg). Based on previous computer simulations, there is no preference position for $Mg^{2+}$ in a calcite crystal (Wang, 2011). For the high-Mg calcite, we therefore randomly distributed $Mg^{2+}$ over the entire crystal, replacing 30% of the $Ca^{2+}$ ions with $Mg^{2+}$ (from here on referred as high $Mg^{2+}$ or HMg). In the slab with an island on top, the cations in the acute and obtuse corners were kept as $Ca^{2+}$, and $Mg^{2+}$ was disseminated on the island around the corners. The flat surface under the island maintained a 30% random distribution of $Mg^{2+}$. An image of our defective surface model can be seen in Figure 6.1.



(**a**) (**b**)

Figure 6.1 Snapshot of simulation cells after equilibration. (**a**) full simulation box with the calcite {$10\bar{1}4$} slab in the bottom, carbonate groups in grey (C) and red (O) and $Ca^{2+}$ in green. The 6.5 nm water column with red (O) and white (H). (**b**) the calcite slab with a 16 cation-anion units island; water molecules were left out for visibility.



**Force field.** As described previously (Di Tommaso et al., 2014; Koskamp et al., 2019), the simple point-charge flexible water (SPC/fw) in combination with a Buckingham potential (de Leeuw and Parker, 1998; Kerisit and Parker, 2004) provided the best water dynamics around calcium in solution when compared to ab initio calculations. For comparison with more recent work on calcite surfaces (De La Pierre et al., 2017, 2016; Kirch et al., 2018; Mutisya et al., 2017; Reischl et al., 2019, 2013), the force field to describe the inter- and intra-molecular interactions was taken from Raiteri et al. (Raiteri et al., 2015), including the SPC/fw water description (Wu et al., 2006). An 0.9 nm cut-off was used for the van der Waals forces, except for the SPC/fw-tail force field, in which the cut-off was defined at 0.9 nm, but with a tail from 0.6 nm (Raiteri et al., 2015). The exact values used in our simulations can be found in Table S6.1.

**Coordination number.** The number of molecules directly coordinated to a central ion was reported as the coordination number. The coordination number is determined from the integral of the radial distribution function (RDF). To include the whole first coordination shell, the cut-off distance was set at 0.35 nm as this distance included the full first peak in the RDF (whether or not split in the case of oxygen in the carbonate ($O_c$) in the high $Mg^{2+}$ calcite), unless stated otherwise.

**Surface energy.** An estimation of the surface energy was obtained by simulating bulk pure calcite, LMg and HMg at zero K (a "dry run" using T = 10 K using the energy of the minimised structure). The energy of the system was then subtracted from the energy of a zero K simulation of the surfaces, where each surface system was first relaxed with water on top at 300 K for 1 ns before the dry zero K run was conducted to minimise the energy. The energy in $J/m^2$ is the estimated surface energy of the water-relaxed-dry surfaces:

$$\sigma = \frac{U_s - U_b}{2 \times A} \tag{26}$$

$U_s$ is the energy of the crystal with expressed $\{10\bar{1}4\}$ surface after relaxation at 0 K, $U_b$ is the energy of the bulk crystal, $A$ is the surface area of one side of the crystal. The surface energy indicates the stability of the surface, with a low positive value indicating a stable surface, whereas the more positive the value, the less stable is the surface. An estimation of the "wet" surface energy was calculated after hydrating the surface by placing a water molecule approximately every 1 $nm^2$. Only the first four layers could relax, as the rest of the slab was fixed according to its energetic minimum of the dry surface. The energy of one water molecule multiplied by the number of water molecules on the surface was subtracted from $U_s$ according to Equation (27):

$$\sigma = \frac{U_s - U_b - nU_{H_2O}}{2 \times A} \tag{27}$$

where $U_{H_2O}$ is the energy of one liquid water molecule in bulk water, $U_s$ is the energy of the configuration excluding the water and $U_b$ is the energy of the bulk crystal.

We have also used Equation (26) to calculate the surface energy, employing the final configuration resulting from the dynamic simulation including water, but excluding the water in the estimation of $U_s$. The standard deviation was calculated to demonstrate the variance of the surface energy and thus, indirectly, the mobility of the ions in each system studied.

**Free energy profiles.** The free energy calculations were performed to extract information about the reactivity of Mg-calcite surfaces and to compare them to the pure calcite surface. We simulated various single calcium ion extractions from different positions in a crystal surface (with or without an island on top and with different amounts of Mg in the vicinity) into the pure liquid water layer 1.2 nm away from the surface. The PLUMED (Tribello et al., 2014) plug-in for LAMMPS was used to perform all free energy calculations. In a single steered MD simulation, we extracted a specific calcium out of the surface and dissolved it. To have a converged simulation for construction of the energy profile, the frames of the steered MD were used to perform biased umbrella sampling at different heights above the initial surface position. In our exploration of the free energy, we selected the distance perpendicular to the surface as the collective variable (CV), based on the assumption that this distance would affect the free energy the most. Due to convergence issues in the directions parallel to the plane, the energy profile was not stable after 120 ns of simulation and we therefore decided to limit the energy exploration only in the direction orthogonal to the surface, by restricting the movement in the plane parallel to the surface with a strong harmonic spring (spring constant of 5.0 eV/Å$^2$). With this restriction in place, the free energy calculations were carried out for 10 ns. The calculation consisted of more than 61 harmonic umbrella potentials with spring constants of 0.5 eV/Å$^2$, or 5.0 eV/Å$^2$ in positions where necessary to obtain a satisfactory overlap in the sampling along

the CV of neighbouring windows (see section 6.6.2 in the supplementary materials and Figure S6.1 for more details). The unbiased free energy profiles were obtained through self- consistent histogram re-weighting, using the weighted histogram analysis method (WHAM) code (Grossfield, n.d.). For the interpretation of the energy profiles, the $\Delta G_{extraction}$ was used, which is defined as the Helmholtz free energy of extraction of a calcium ion from bulk water (~1.3 nm) to its position in the crystal structure. The positive energy indicates that the cation is favoured in the surface position.

**Diffusion.** The mean square displacement (MSD) was calculated with a modified version of the Fortran code written by W. Smith from the UK Daresbury Laboratory (Smith, 1996). Thereafter, the self-diffusion coefficient was derived as the slope of the MSD as a function of time divided by the number of dimensions times two (for both positive and negative direction). The self-diffusion coefficient of interfacial water parallel to the calcite surface was calculated by adjusting this code to only account for water molecules present in the interface (e.g., below 0.35 nm from the calcite surface) at the scanned time. Due to the small size of the interfacial water layer, water molecules leave the specified layers within short times. To improve the statistics, we modified the code to consider a linear extrapolation of the MSD when the water molecule entered the interfacial water layer between the time-origin and the remaining simulation time, and in the input file we set the time interval between MSD origins to one frame. In this way, water molecules that stayed longer in the specified interface area had more weight in the MSD, and short visiting molecules could be included without wrongly interfering with the MSD (Koskamp et al., 2019).



**Vibrational spectrum.** The vibrational spectrum at different distances from the surface was built from the sum of all the vibrational density of states (VDOS) of each atom involved (Allen and Tildesley, 1987). The VDOS can be calculated using the Fourier transformation of the

velocity-autocorrelation function (VACF) of the individual atoms in the water molecule. We used the following equation to calculate the VACF:

$$VACF(t) = \frac{1}{N_{tsteps} N_{atm}} \sum_{j=1}^{N_{tsteps}} \sum_{i=1}^{N_{atm}} v_j(t)_j \times v_i(t_j + t), \qquad (28)$$

where $N_{tsteps}$ and $N_{atm}$ are the number of timesteps and the number of atoms, oxygen and hydrogen in this case; $v_i$ is the velocity vector of O or H atoms in the $i^{th}$ water molecule.

To produce the vibrational spectrum of interfacial water (<0.35 nm from the surface), the water molecules in the interface were flagged in the starting configuration, which was compared with all configurations (the final 40 ps of the simulation) to make sure that the flagging was still valid (that is, that the water molecules still remained in the interface).

**Water density.** The water density profile perpendicular to the surface was determined to visualise the layering of water at the interface. The density is calculated based on the distance from the surface to the centre of mass of the water.

$$g(x) = \frac{dn_x}{Vx} \qquad (29)$$

where $dn_x$ is the number of particles within a water layer with volume $Vx$. To visualise the water $z$-density profile above a certain surface position, we only considered the water molecules in the column on top of the position within a rectangular box with width of 0.2 nm by 0.2 nm.

**Number of hydrogen bonds between water molecules.** Another feature to characterise the interfacial water structure is the number of hydrogen-bonds that a water molecule, coordinated to the surface, forms with other water molecules. This interaction stabilises the configuration and is fundamental to create an intermolecular network at the solid-water interface, as well as in the bulk solution. The coordination distances were defined based on the first minimum in the RDF and was set at 0.28 and 0.40 nm for the cation–oxygen and anion–oxygen, respectively (Chandra, 2000). A hydrogen-bond was defined with a proton–oxygen distance of 0.245 nm

and an angle of less than 30 degrees between the oxygen acceptor, oxygen donor and proton donor (Chandra, 2000).

**Water exchange frequency.** To investigate the water dynamics around dissolved calcium ions, the water exchange frequency ($N_{ex}$) of water molecules in the first hydration shell of the cation was determined, using the "direct method" as described in previous work (Koskamp et al., 2019). The average exchange frequency for the whole surface was calculated by dividing the total number of exchanges by the number of Ca ions in the surface.

## 6.3 Results

### 6.3.1 Water Density, Structure and Dynamics in the Interface

The water *z*-density profile on top of the pure, low $Mg^{2+}$ flat surface (LMgF) and high $Mg^{2+}$ flat surface (HMgF) calcite $\{10\bar{1}4\}$ surfaces revealed the layering of the interfacial water (Figure 6.2). The water profile of LMgF was identical to the profile of water on top of pure calcite, but there was a clear difference observed in the water layering above the HMgF surface compared to the pure and LMgF calcite surfaces. For HMgF, the first peak had a lower intensity and showed a shoulder at shorter distance, indicating higher densities closer to the surface.

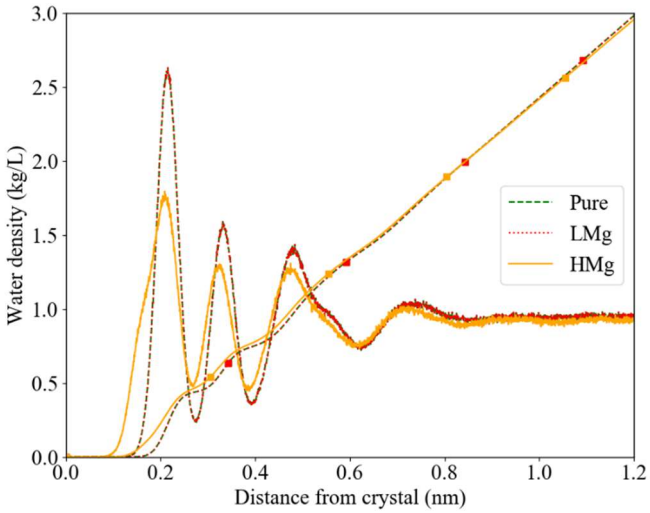



Figure 6.2 Overall water z-density profile of pure (dashed green, overlaps with dotted red line), low $Mg^{2+}$ flat surface (LMgF; dotted red) and high $Mg^{2+}$ flat surface (HMgF; yellow) with respect to the flat crystal surface (= 0.0 nm). The lines with the marker (■) is the integral of the water density.

The hydrogen-bond network varied with the level of Mg incorporation, while the total number of hydrogen-bonds with water coordinated to Ca was comparable between the different interfaces (Table 6.1). Our analysis of the interactions between oxygens in water and carbonate molecules showed that water molecules coordinated to $Mg^{2+}$ form on average 1.37 hydrogen-bonds with $O_c$, compared to 1.02 for water molecules coordinated to $Ca^{2+}$. The latter value is irrespective of whether the surface $Ca^{2+}$ is located in a pure or HMgF calcite. Almost half of the water molecules coordinated to surface Mg (40.8%) showed more than one H-bond to at least one of the neighbouring surface carbonate oxygen ($O_c$) (Table 6.1), in contrast to water molecules coordinated to $Ca^{2+}$, where this was the case for only 15–16%. The lifetime of the formed H-bond on top of $Ca^{2+}$ on the HMgF calcite was 6.5% longer than on the pure calcite. On top of $Mg^{2+}$, this was 19.4% shorter compared to a water molecule on top of pure calcite. Furthermore, the water dipole angle with the surface is slightly smaller (0.24°) for water in the interface with HMgF (Section 6.6.3 of the supplementary materials for further details). The vibrational spectrum (VDOS) showed a difference in the behaviour of the water molecules present in the HMgF interface compared to the water molecules present in the pure calcite interface. The lower intensity of the peaks for intermolecular stretching and intermolecular O–

O–O bonding motion (Figure S6.2 and Figure S6.3) indicates that there are fewer hydrogen-bonds that vibrate with the same frequency as seen in the pure calcite interface.

Table 6.1 Interfacial water dipole angle and H-bond properties.

| Property | $Ca^{2+}$ in Pure Calcite | $Ca^{2+}$ in HMgF Calcite | $Mg^{2+}$ in HMgF Calcite | $CO_3^{2-}$ in Pure Calcite | $CO_3^{2-}$ in HMgF Calcite |
|---|---|---|---|---|---|
| Water dipole angle with surface (°) | 76.13 | 75.89 | | | |
| Average number of hydrogen bonds between water molecules | 0.633 | 0.628 | 0.485 | 0.942 | 0.910 |
| 0 H-bond (%) | 49.9 | 49.9 | 57.7 | 40.8 | 42.3 |
| 1 H-bond (%) | 37.8 | 38.0 | 36.3 | 34.0 | 33.9 |
| 2 H-bond (%) | 11.5 | 11.3 | 5.8 | 16.9 | 15.9 |
| 3 H-bond (%) | 0.8 | 0.7 | 0.2 | 7.1 | 6.7 |
| H-bond Lifetime | 3.1 | 3.3 | 2.5 | 3.5 | 3.9 |
| Average number of H-bond between cation—$O_w$–$H_w$–$O_c$ | 1.01 | 1.02 | 1.37 | | |
| Water forming >1 H-bond with $O_c$(%) | 15.8 | 15.3 | 40.8 | | |
| Distance $H_W$–$O_c$ | 1.853 | 1.855 | 1.860 | | |

The water dynamics were also altered when a high amount of Mg was incorporated. The exchange frequencies for water molecules coordinated to $Ca^{2+}$ showed a broader range in HMgF than in LMgF and pure calcite (Figures S4 and S5), and the average frequency was slightly higher (31.4 versus 25.7 $ns^{-1}$ for HMgF versus pure calcite, Table S6.2), whereas LMgF showed similar dynamics to the pure surface. The translational diffusion coefficient parallel to the interface ($D_{x/y}$) was also comparable for LMgF and pure calcite, in contrast to HMgF, which showed an increase of 38.9%. Further details can be found in section 6.6.5 in the Supplementary Materials.



### 6.3.2 Surface Energies and Structural Relaxation

The energies of the surfaces are shown in Table 6.2 (see alternative way of measuring this in Table S6.3 and literature values in Table S6.4). LMgF calcite crystals had a slightly higher

surface energy compared to the $\{10\bar{1}4\}$ surface of pure calcite for both dry and wet surfaces. In contrast HMgF calcite had a slightly lower surface energy for the dry surface and a decrease of 0.1 J/m$^2$ on the wet surface was observed compared to the pure wet calcite surface. Moreover, the surface energy of a surface with an island (PureI and HMgI) and etch pit in the opposite side of the slab, was higher for the dry island. This trend was in contrast with the wet HMgI surface which showed the lowest surface energy (Table 6.2). To summarise, the overall dry $\{10\bar{1}4\}$ surface energy varied according to HMgF < Pure < LMgF < PureI < HMgI and the overall wet surface energy was ordered as HMgF < HMgI < Pure < LMgF < PureI.

Table 6.2 Overview of the surface energies of the simulated systems

| **$\{10\bar{1}4\}$ System** | **Acronym** | **Surface Energies (J/m$^2$)** | |
|---|---|---|---|
| | | Dry | Wet |
| Pure calcite | Pure | 0.63 | 0.30 |
| Low $Mg^{2+}$ flat surface | LMgF | 0.64 | 0.31 |
| High $Mg^{2+}$ flat surface | HMgF | 0.62 | 0.20 |
| Island pure calcite | PureI | 0.66 | 0.35 |
| Island high $Mg^{2+}$ | HMgI | 0.68 | 0.29 |

Evidence of the degree of relaxation in the surfaces can be seen in the distances between the constituent ions, depicted in Figure 6.3 (and Figure S6.6). The average cation-carbon distances ($C_c$) in the surfaces, observed during the simulations, showed variations depending on the presence of Mg (Figure 6.3a), in contrast to the cation–$O_c$ distances (Figure S6.6). In Figure 6.3a, the second peak (~0.41–0.43 nm) represents the distance between the cation and carbon of the first carbonate along the *c*-axis of calcite, as illustrated in Figure 6.3b. This distance was shortened upon Mg incorporation. The RDF of LMg had similar Ca-$C_c$ distances compared to pure calcite and the Mg–$C_c$ distances were the same as in HMg.

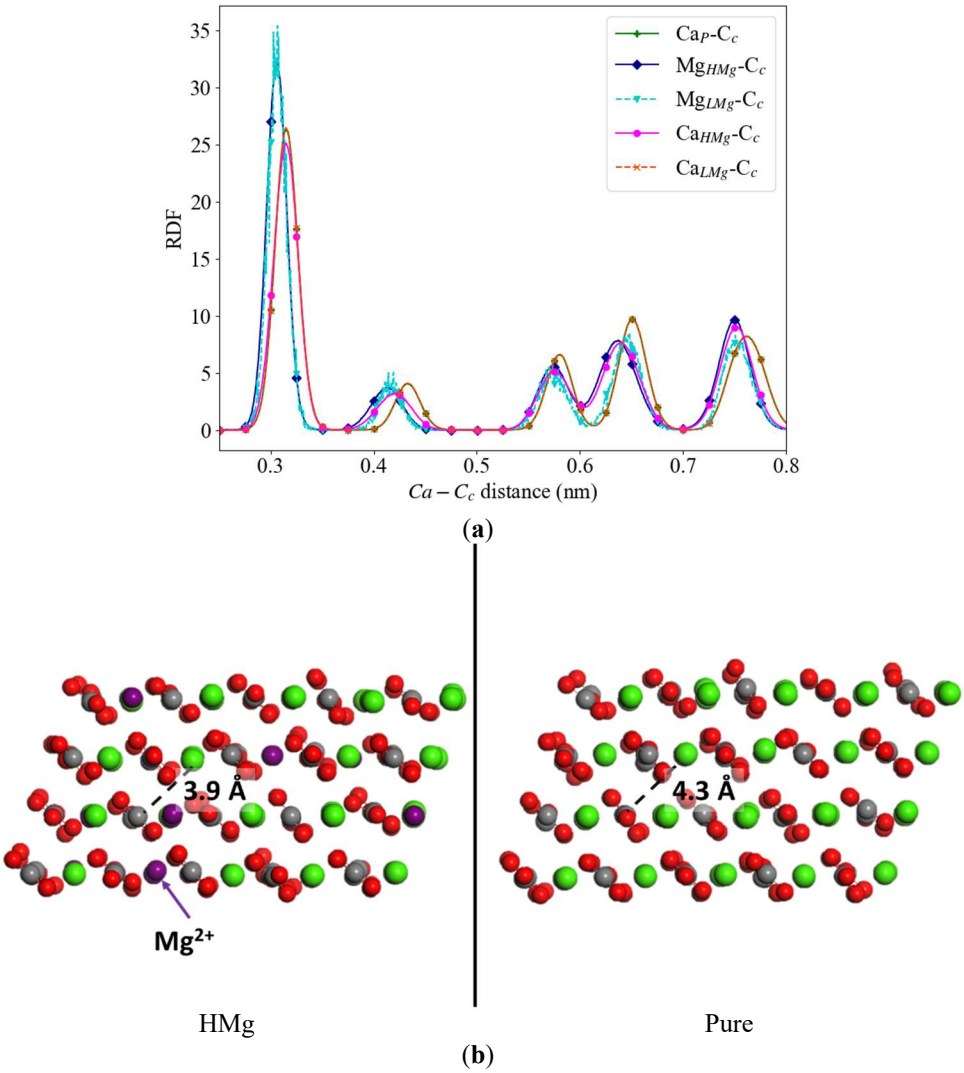



Figure 6.3 (**a**) Radial distribution function for the cation with the carbon of carbonate in pure, low $Mg^{2+}$ concentration (LMg) and high $Mg^{2+}$ concentration (HMg) calcite crystals. (**b**) Side view snapshots of the calcite crystal showing the deformation along the *c*-axis in the HMg (left) compared to pure calcite (right—for average distances see Table S6.5 and Figure 6.3a); calcium ions are in green, magnesium ions in purple, carbonate oxygen and carbon in red and grey, respectively.

### 6.3.3 Free Energy Profiles

The free energy difference between a fully solvated $Ca^{2+}$ and a $Ca^{2+}$ in the crystal position, $\Delta G_{extraction}$, was calculated for eight different environments on a flat $\{10\bar{1}4\}$ surface, in pure, LMgF and HMgF calcite. The values were in the range ~191–230 kJ/mol, where the $\Delta G_{extraction}$ for different crystal positions with respect to $Mg^{2+}$ ranged from ~197-223 kJ/mol and the $\Delta G_{extraction}$ of a $Ca^{2+}$ from the $\{10\bar{1}4\}$ flat calcite crystal surface was 204 kJ/mol (Table 6.3). The profiles of the different flat surfaces showed a steep increase in free energy until 0.5 nm

and converged at around 0.9 nm, corresponding to the second interfacial water layer and the start of bulk water behaviour, respectively. In Figure 6.4, we present free energy profiles obtained from the extraction of $Ca^{2+}$ from a pure and LMgF calcite $\{10\bar{1}4\}$ surface. All calcium ions in the nearest surroundings of $Mg^{2+}$, i.e., without another cation in-between, had a higher $\Delta G_{extraction}$ compared to the $\Delta G_{extraction}$ of $Ca^{2+}$ in pure calcite. For HMgF, the $\Delta G_{extraction}$ for surface calcium varies with the local environment of the crystal between 191 to 230 kJ/mol. Higher energy barriers were found when a $Mg^{2+}$ ion was in the direct vicinity of the target $Ca^{2+}$. Two out of three $Ca^{2+}$ gave a $\Delta G_{extraction}$ higher than the energy needed to extract one $Ca^{2+}$ from a pure system (Figure 6.5). To describe in more detail the conditions leading to a difference in $\Delta G_{extraction}$, we used the labels in Table 6.3 to distinguish between the different $Ca^{2+}$ environments, both with respect to the number of Mg neighbours and the surface topography.

Table 6.3 Selected calcium sites for forced molecular dynamics (MD) extraction, including the allocated colour code used in Figure 6.4 until Figure 6.6, label and free energy of full extraction. The second free energies reported for the HMgI system are for simulations in which the island $CO_3^{2-}$ molecules were frozen.

| Calcium Site | Colour Code | Label | $\Delta G_{extraction}$ (kJ/mol) |
|---|---|---|---|
| $Ca^{2+}$ in pure calcite | Green (✚) | $Ca_P$ | 204 |
| **Low $Mg^{2+}$ flat surface (LMgF; Figure 6.4)** | | | |
| $Ca^{2+}$with 1 $Mg^{2+}$ in first cation shell* | Turquoise (▼) | $Ca_{LMgF1}$ | 218 |
| | Pink (●) | $Ca_{LMgF2}$ | 223 |
| | Orange (▬) | $Ca_{LMgF3}$ | 222 |
| $Ca^{2+}$with 1 $Mg^{2+}$ in second cation shell | Purple (■) | $Ca_{LMgF4}$ | 197 |
| **High $Mg^{2+}$ flat surface (HMgF; Figure 6.5)** | | | |
| $Ca^{2+}$with 3 $Mg^{2+}$ in first cation shell and 2 $Mg^{2+}$ in second | Purple (■) | $Ca_{HMgF3\text{-}5}$ | 209 |
| $Ca^{2+}$with 1 $Mg^{2+}$ in first cation shell and 1 $Mg^{2+}$ in second | Pink (●) | $Ca_{HMgF1\text{-}2}$ | 230 |
| $Ca^{2+}$with 0 $Mg^{2+}$ in first cation shell and 0 $Mg^{2+}$ in second | Turquoise (▼) | $Ca_{HMgF0\text{-}0}$ | 191 |
| **Island-surface, pure calcite (PureI, Figure S6.7)** | | | |
| $Ca^{2+}$ in an Acute Corner | Purple (■) | $Ca_{PAC}$ | 43 |
| $Ca^{2+}$ in an Obtuse Corner | Pink (●) | $Ca_{POC}$ | 39 |
| $Ca^{2+}$ in an Acute Edge | Turquoise (▼) | $Ca_{PAE}$ | 56 |
| $Ca^{2+}$ in an Obtuse Edge | Orange (▬) | $Ca_{POE}$ | 76 |
| **Island-surface, high $Mg^{2+}$ calcite (HMgI; Figure 6.6)** | | | |
| $Ca^{2+}$ in an Acute Corner | Purple (■) | $Ca_{HMgIAC}$ | 37; 97 |
| $Ca^{2+}$ in an Obtuse Corner | Pink (●) | $Ca_{HMgIOC}$ | 24; 90 |
| $Ca^{2+}$ in an Acute Edge | Turquoise (▼) | $Ca_{IHMgIAE}$ | 72; 206 |
| $Ca^{2+}$ in an Obtuse Edge | Orange (▬) | $Ca_{HMgIOE}$ | 60; 175 |

The bold text indicates the different systems investigated * first cation shell is defined here as the first "cation" shell and therefore ignoring carbonate ion groups that are located more closely to the central metal ion.

### 6.3.3.1 Low $Mg^{2+}$ Calcite

The $\Delta G_{extraction}$ profiles for the LMgF system showed narrower and shallower local minima (Figure 6.4b) than in the pure calcite (green line, Figure 6.4b). Furthermore, while the average distances in LMgF were the same as in pure calcite, locally there were small but visible variations (Table S6.6). Figure 6.4b illustrates the small local variations in the free energy profiles, water densities and distances for four $Ca^{2+}$ ions located at a unique distance from the

surface $Mg^{2+}$. Note that the water densities (right axis in Figure 6.4b) represent the local water *z*-density of the unbiased simulation, i.e., the water density before calcium extraction. $Ca_{LMgF1}$ had the shortest distance to the $Mg^{2+}$ and was the only $Ca^{2+}$ for which the two neighbouring $Ca^{2+}$ (at ~0.4–0.5 nm) were not at the same distance (the closest at 0.4525 nm and second nearest at 0.4925 nm). The $\Delta G_{extraction}$ profile for this surface calcium shows a small dent at the distance where the highest water density is observed (Figure 6.4b), although there is no strong correlation between the water density and free energy profiles. For $Ca_{LMgF2}$ and $Ca_{LMgF3}$, the free energy profiles, water density profiles and distances to neighbours were comparable (Figure 6.4b). The three $Ca^{2+}$ nearest to $Mg^{2+}$ had $\Delta G_{extraction}$ which was higher than pure calcite. For $Ca_{LMgF4}$, the second $Ca^{2+}$ distance was slightly longer than the other distances and it was the only site with lower $\Delta G_{extraction}$ than pure calcite.

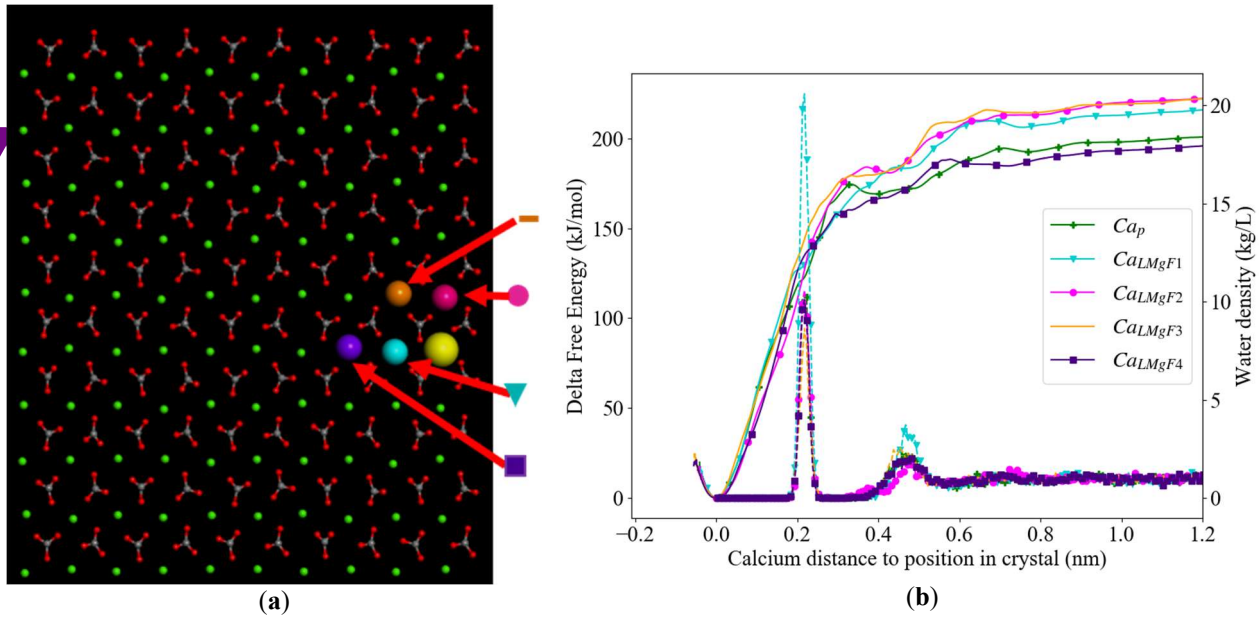
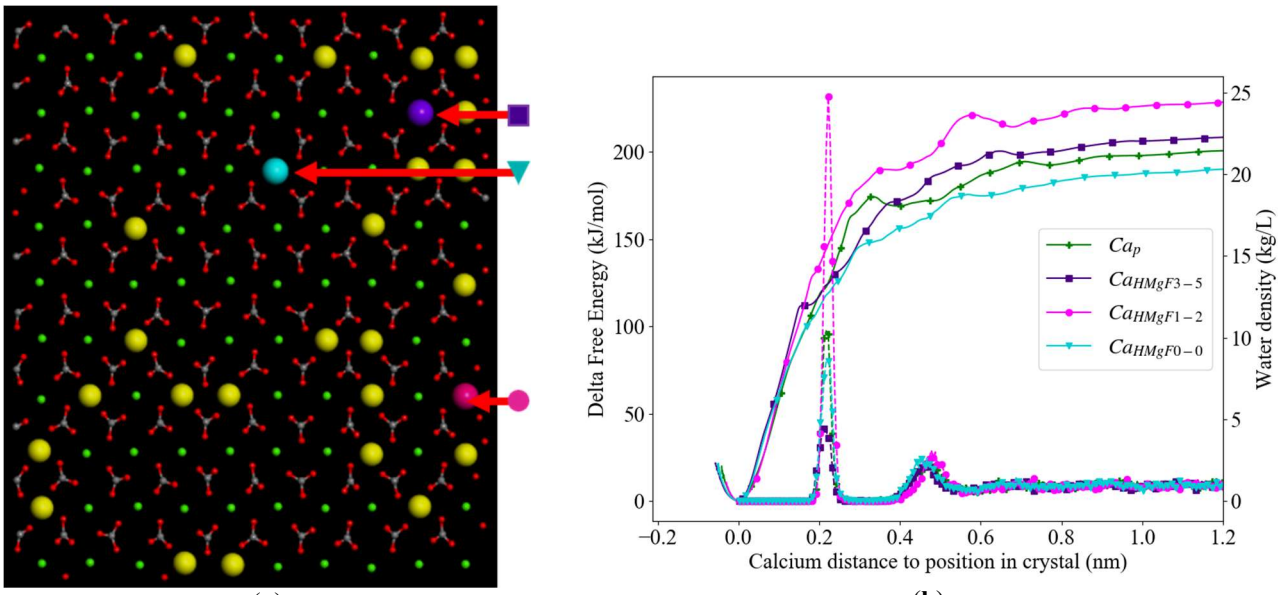



Figure 6.4 (**a**) Top view of the LMgF {10$\bar{1}$4} surface with $Mg^{2+}$ in yellow, carbonate groups in grey (C) and red (O) and $Ca^{2+}$ in small green or colour coded (cf. Table 6.3) larger spheres to link them to the free energy profiles in turquoise (▼), pink (●), orange (no marker) and purple (■). (**b**) Overlay of free energy profile (continuous line, left *y*-axis) of $Ca^{2+}$ extracted from its equilibrium position in the crystal (0.0 nm) with the water *z*-density (dashed line, right *y*-axis) along the same axis. The colours are referring to the corresponding $Ca^{2+}$equilibrium position in the left image and $Ca^{2+}$ in pure calcite in green (+).

#### 6.3.3.2 High $Mg^{2+}$ Calcite

For the free energy profiles determined in the HMg system, three calcium sites were selected that could be differentiated by the number of $Mg^{2+}$ in the vicinity. The first site ($Ca_{HMgF0-0}$; Figure 6.5a) showed average coordination distances to $O_c$ (0.232 nm) close to the average coordination distance in pure calcite. The energy profile showed a shallower energy minimum at its relaxed position in the surface, resulting in a total $\Delta G_{extraction}$ that was 13 kJ/mol lower compared to the extraction of $Ca^{2+}$ from a pure system. The first and second peak in the water density profile correlated to moderation of the slope of the energy profile, but no other local minima were observed upon extraction. The second $Ca^{2+}$ ($Ca_{HMgF1-2}$) had one $Mg^{2+}$ next to it and another at a slightly larger distance (Figure 6.5a). Furthermore, two $Mg^{2+}$ ions were present nearby in the second crystal layer. The average coordination distance to $O_c$ (0.282 nm) was somewhat shorter than for pure calcite (Table S6.5). In the free energy profile of the second $Ca^{2+}$ ($Ca_{HMgF1-2}$; Figure 6.5a), shallow local minima on both sides of the second peak of higher water density (at 0.45 nm) were observed. With a total $\Delta G_{extraction}$ that was 26 kJ/mol higher than the pure calcite, this Ca showed the highest energy needed for extraction. Also, the water column on top showed the highest water density at the first peak in the density profile of all calcium sites. This high water density was comparable to $Ca_{LMgF1}$, which also had a $Mg^{2+}$ at 0.39 nm (cf. Figure 6.4a and Figure 6.5a). Moreover, in both cases the second peak in the water density profile was slightly shifted away from the flat surface.




(**a**) (**b**)

Figure 6.5 (**a**) Top view of HMgF {10$\bar{1}$4} surface with $Mg^{2+}$ in yellow, carbonate groups in grey (C) and red (O) and $Ca^{2+}$ in small green or colour coded (cf. Table 6.3) larger spheres to link them to the free energy profiles in purple (■), turquoise (▼) and pink (●). (**b**) Overlay of free energy profiles (continuous line, left *y*-axis) of calcium ions extracted from their equilibrium position in the crystal (0.0 nm) with the water *z*-density (dashed line, right *y*-axis) along the same axis. The colours are referring to the corresponding $Ca^{2+}$ equilibrium position in the left image and $Ca^{2+}$ in pure calcite in green (+).

The last investigated site on HMgF ($Ca_{HMgF3-5}$; Figure 6.5) had the most $Mg^{2+}$ ions in its vicinity and a total of five $Mg^{2+}$ in the surface layer. It also had one $Mg^{2+}$ in its vicinity in the second layer. It is worth noting that the distance of the closest $Mg^{2+}$ and $Ca_{HMgF3-5}$ of 0.385 nm was similar to the distance between $Ca^{2+}$ and $Mg^{2+}$ in dolomite (0.387 nm). The average distance of the other two $Mg^{2+}$ and one $Mg^{2+}$ in the second layer was 0.486 nm. The average coordination distance to $O_c$ (0.234 nm) is comparable with pure calcite. The Ca site was the only site that showed a clear shallow local minimum at <0.2 nm. The total $\Delta G_{extraction}$ had an intermediate value of 5 kJ/mol, which was the closest to pure calcite. This Ca-site showed the lowest water density in the first water layer.

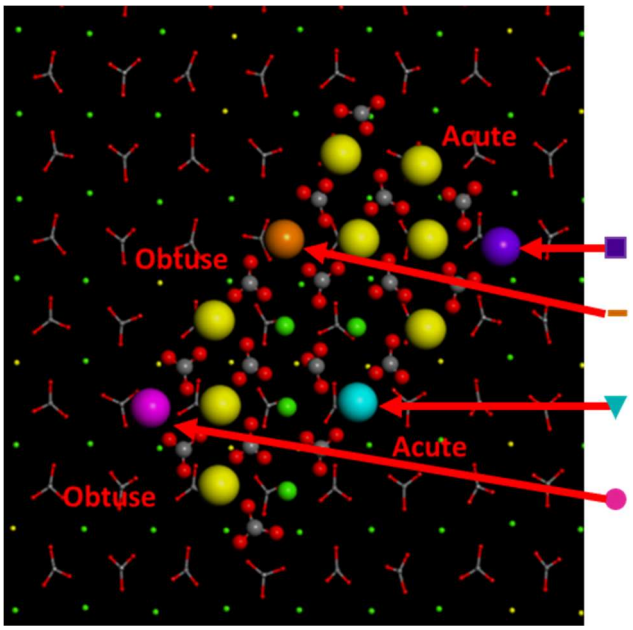
### 6.3.3.3 Calcite with an Island

In the Mg-doped island topography on the {10$\bar{1}$4} surface (HMgI, Figure 6.6a), we focused on four different sites of the 4×8 atoms island. Note that initially the $Mg^{2+}$ impurities were

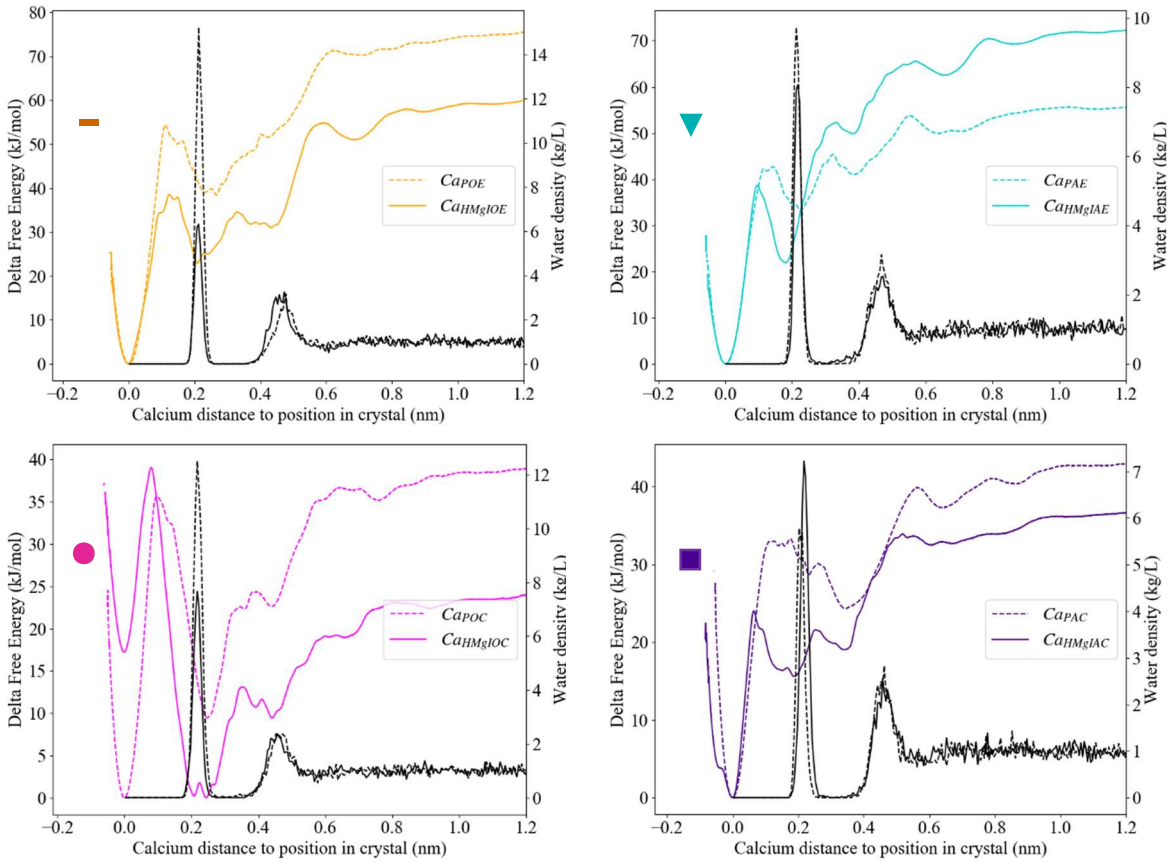
randomly distributed, causing a 50% impurity level in the island. Subsequently, we moved a few $Mg^{2+}$ closer to the corners to make them have equal amounts of impurities before starting the simulations. The Mg-rich island showed relaxation of the surface structure (Table S6.7), where the local symmetry of the island changed and the distances and positions were altered relative to the pure island. Consequently, the edges in HMgI did not have the same symmetry in the crystal as observed in the PureI edges. The RDFs between island-surface $Ca^{2+}$ and $O_c$ of high HMg (Figure 6.6) showed more distinguishable peaks compared to island-surface $Ca^{2+}$ in pure calcite (in Figure 6.7).

In general, the energy profiles of PureI and HMgI (Figure 6.6) showed similar shapes with local minima and maxima at the same positions, yet the minima in the HMgI profiles were deeper, for example creating a more substantial minimum around the second peak of water density (~0.5 nm) on the obtuse edge. The total $\Delta G_{extraction}$ (Table 6.3) for the obtuse edge in PureI is 16 kJ/mol higher than the same edge in HMgI. Furthermore, the extraction of $Ca^{2+}$ is less favourable from the obtuse edge compared to the acute edge. For the acute edge, the PureI had a total $\Delta G_{extraction}$ that was 16 kJ/mol lower than HMgI and was therefore the most easily extracted (dissolved). For the corner calcium sites in HMgI, the free energy profile (Figure 6.6b) showed several shallow minima and clear local minima on both sides of the water density peaks. Exceptionally, for the calcium site at the obtuse corner of HMgI, these minima were lower in free energy than the minimum for $Ca^{2+}$ in the crystal structure. Although the difference in extraction energy (Table 6.3) between the position in the crystal and fully solvated was only 7 kJ/mol, the total $\Delta G_{extraction}$ for this calcium was 24 kJ/mol, due to the low second minimum, yielding a free energy of extraction 15 kJ/mol lower than its pure calcite equivalent. The difference between total $\Delta G_{extraction}$ of HMgI and PureI at the acute corner site is 6 kJ/mol and thereby the smallest difference observed on the island.

The HMgI edges showed a similar density for the obtuse and the acute edge, both lower than the PureI edges. The main differences in the water density profiles were observed in the first peak, i.e., the first layer of interfacial water. The water density on top of the obtuse corner in PureI had the highest water density of the island corners. The lowest density was observed on the acute corner of PureI, which also showed a small shift of the peak towards the surface. As in the trend for the corners, the obtuse calcium edge sites $Ca_{POE}$ showed the highest water density in the first peak.




(**a**)




(**b**)

Figure 6.6 (**a**) Top view of the 4×8 atoms island on top of HMg $\{10\bar{1}4\}$ surface with $Mg^{2+}$ in yellow, carbonate groups in grey (C) and red (O) and $Ca^{2+}$ in small green or colour and symbol coded larger spheres to link them to the free energy profiles in orange (obtuse edge, ▬), turquoise (acute edge, ▼), pink (obtuse corner, ●) and purple (acute corner, ■). (**b**) Overlay of free energy profile with the water density along the same axis. The dashed line represents the results for the equivalent calcium site in the pure system, i.e., the $Ca^{2+}$ are in a similar position as in (a), but no $Mg^{2+}$ is present in the surface.

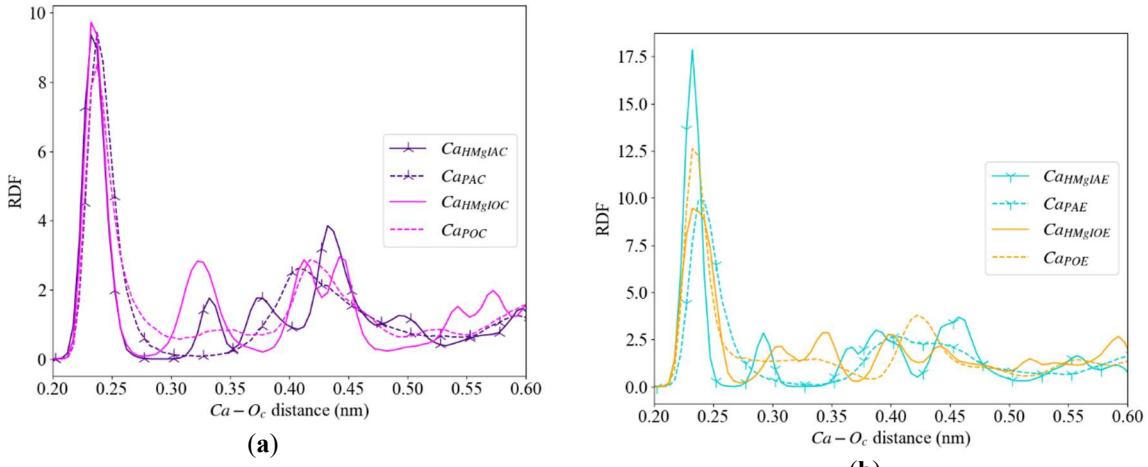



Figure 6.7 Radial distribution function (RDF) of the different calcium ions with $O_c$. (**a**) The acute and obtuse calcium corners and (**b**) the acute (with tripod marker) and obtuse calcium edges of the pure (dashed line) and HMg (continuous line) with the corresponding colours as presented in Figure 6.6.

### 6.3.4 Impact of Surface Carbonate on Energy Profiles

When determining the free energy profiles for calcium extraction from the island surfaces, we observed relaxation of carbonate ions in the vicinity of the extracted calcium sites. To study the impact on $\Delta G_{extraction}$, we froze all carbonate ions in the island and compared the free energy profiles (i.e., the carbonate groups were fixed in position and could not rotate or vibrate, Figure 6.8). The second energy value for the HMgI surface in Table 6.3 corresponds to the $\Delta G_{extraction}$ for frozen $CO_3^{2-}$.

Due to the frozen carbonate groups, the energy profiles changed significantly, the first minimum was deeper and the profiles are more similar in shape to the profiles obtained from the flat surface. In addition, the shapes and locations of the local minima showed little or no correlation to the unfrozen profiles. When the $CO_3^{2-}$ groups were frozen in their position, the $\Delta G_{extraction}$ for the extraction from the corners were calculated at 90 and 97 kJ/mol for the obtuse and the acute corners respectively. These energies are 66 and 60 kJ/mol higher, respectively, than in the system where the carbonate ions were not frozen. To remove calcium from the edges, we calculated the total $\Delta G_{extraction}$ to be 175 kJ/mol for the obtuse and 206 kJ/mol for the

acute edges, respectively. As such, the free energy required to remove calcium from the acute edge was within the range of free energies obtained for the flat HMg surface.

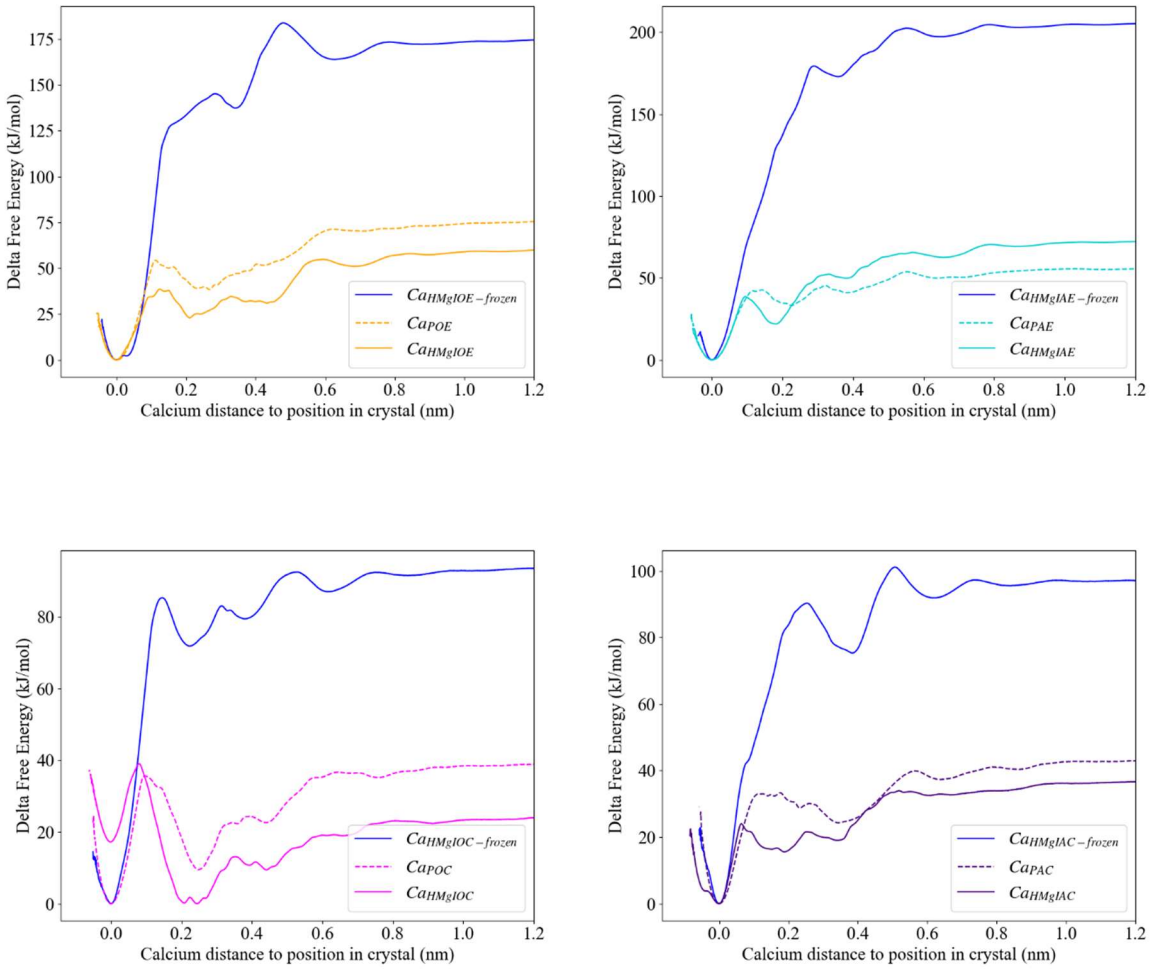



Figure 6.8 Free energy profiles of $Ca^{2+}$ extraction from PureI (dashed lines), HMgI (continuous line) and HMgI in which the island-carbonate molecules are frozen (continuous blue line).

## 6.4 Discussion

Before we focus on the impact of Mg, we compared our observations for the pure system with those previously published. Note that an extensive study of the force field used here, and the resulting $\{10\bar{1}4\}$ surface interaction with water, was compared to X-ray data and our results compare well with these previously published values (Fenter et al., 2013; Koleini et al., 2018). The surface energy we obtained for the atomically flat dry $\{10\bar{1}4\}$ surface of a pure calcite

compares well with previously published values (0.23–0.86 J/m$^2$, Table S6.4) (Bruno et al., 2013), including the value reported for the same forcefield (0.71 J/m$^2$) (Sekkal and Zaoui, 2013). The difference with this last value is possibly due to the extra relaxation by the MD simulation with water prior to the dry optimization of the slab in our methodology. Experimentally dry surfaces are never completely free of water, and hence lower surface energies were measured. The surface energies for wet surfaces are more comparable to the experimental values for the dry surfaces (0.23–0.34 J/m$^2$); this discrepancy can be explained by the fact that experimental surfaces have defects (covered with water) that are not reproduced in the atomistic models (Røyne et al., 2011).

### 6.4.1 Influence of Magnesium on Calcite Solubility and Dissolution

It is important to note that we do not see a systematic trend in a single parameter with changing Mg concentration in the solid, which suggests that the impact of increasing Mg content is a complex interplay between the different system characteristics. Consequently, upscaling to, for example, the meso-scale leads to numerous issues (see also (Bracco et al., 2013; Kurganskaya and Rohlfs, 2020)). This complexity is inherent in the microscopic scale, since there are many parameters that can play a role (simultaneously). The non-linear correlation of solubility with Mg concentration may suggest that the controlling parameters might change with increasing Mg.

The overall crystal and surface structures, as well as the interfacial water, for LMgF was very similar to pure calcite, suggesting that the local environment controls the stability of specific Ca-sites in the surface. For HMgF, the explanation becomes more complex since there are more parameters in the interfacial water and the crystal structure that are distinct from pure calcite. In the sections below we discuss these parameters consecutively, showing trends and possible explanations for why some calcium sites are more prone to dissolve than others.

#### 6.4.1.1 Interfacial Water Structure

The interfacial water is key in the growth and dissolution of any material (e.g., (Dewan et al., 2013; Hancer et al., 2001)). The structure and dynamics of the water in direct contact with the material can be described by the hydrogen-bond structure (Di Tommaso et al., 2017). In this section, we focus on the flat surfaces to discuss the impact of Mg on the interfacial water, to avoid the added interplay of roughness. A clear layering of the water until 0.8 nm from the surface was observed for our pure calcite system, which has been seen before on calcite and other minerals in both theoretical (Du and De Leeuw, 2006; Maureen I. McCarthy et al., 1996) and experimental studies (Cheng et al., 2001). While the interfacial water structure in LMgF was indistinguishable from that in pure system, the water structure around the $Mg^{2+}$ in the HMgF surface showed clear differences. As a result of the higher magnesium content, HMgF attracts water closer to the calcite surface, thereby disturbing the water layering (Figure 6.2). This behaviour agrees with that shown in previous computer simulations, where $Mg^{2+}$ was adsorbed onto the surface (de Leeuw and Parker, 1998; Forbes et al., 1992; Kerisit and Parker, 2004). With this disruption in the water layer, the hydrophilic and hydrophobic zones in the layering are both shifted 0.01 nm towards the surface.



The hydrogen-bond network for interfacial water is highly affected by the interaction with the crystal surface, with all vibrational modes shifting to higher energies (Figure S6.2). The higher intensities in our VDOS spectra indicate a stronger H-bond network directly over the surfaces compared to bulk liquid water for pure, LMg and HMg systems. For HMgF, the somewhat lower vibrational intensities around 110 and 400 $cm^{-1}$ suggest that $Mg^{2+}$ disrupts the H-bond network. This agrees with the lower number in H-bonds on top of $Mg^{2+}$ (Table 6.1) and the altered dipole angle, indicating that the plane of water molecules is positioned more parallel to the surface (Section 6.6.3 in the supplementary materials and Table 6.1). Combined with the shorter $Mg^{2+}$ to water oxygen distances (Figure 6.2, see c), this results in a lower tendency to

form H-bonds with other water molecules from, for example, the second water layer. Instead, water molecules coordinated to surface $Mg^{2+}$ are more likely to have proton-oxygen interactions with neighbouring carbonates. As a result, water molecules are librating towards neighbouring water molecules and $O_c$ that are at the H-bond distance. Furthermore, the increase in librations (Figure S6.3, ~600 $cm^{-1}$) for HMgF interfacial water suggests that there are more possibilities to form H-bonds, in agreement with the higher average water density (Figure 6.2) and shorter H-bond lifetimes. Due to the disruptions in the water layer, surface $Ca^{2+}$ sites experience a weaker interaction with water molecules, resulting in higher water exchange frequencies and higher water diffusion coefficients to and from the surface calcium sites (Table S6.2).

The differences in intensities in the water *z*-density profiles do not correlate with the energy profiles or the $\Delta G_{extraction}$. The effect of water on the stability of the $Ca^{2+}$ in the crystal is therefore minor. This observation is in agreement with experiments (Xu et al., 2013), although the disruption of the interfacial water structure in HMgF, resulting in higher water exchange frequencies and diffusion rates near surface calcium sites, are likely to contribute towards the release of $Ca^{2+}$ and thereby enhance dissolution of the crystal.



#### 6.4.1.2 Surface Energies and Structural Relaxation

Since higher Mg-bearing calcites are generally observed to be more soluble than pure calcite, the surface energy for HMgF was expected to be higher than for pure calcite (Gao et al., 2012). Counter-intuitively, the lower surface energy for HMgF obtained indicates a more stable flat $\{10\bar{1}4\}$ surface compared to pure calcite. This is most likely due to the strong relaxation observed in the crystal surface, in combination with the re-positioning of water molecules in the interface. Alternatively, in the atomically flat low-Mg calcite, no alteration of the global crystal structure and interfacial water was observed, and a small local deformation at the

surface resulted in a 0.01 J/m$^2$ increase of the surface energy relative to the pure system. Again, this was counter-intuitive, given that low-Mg calcites generally show lower solubility than pure calcite. Despite the differing methodologies, the higher surface energy for LMgF compared well to the trend seen in the literature, where a calcite slab with one $Mg^{2+}$ had a slightly higher energy compared to a pure calcite $\{10\bar{1}4\}$ surface (Bruno et al., 2018).

In the case of a $\{10\bar{1}4\}$ surface with a small island, which may be considered more comparable to realistic surface topographies, we obtained dry surface energies indicating that HMgI calcite is less stable (Table 6.2). Due to the random Mg distribution, the island contains higher concentrations of $Mg^{2+}$ than the average HMgI slab, causing local alterations in the crystal structure (Table S6.7) that cannot be compensated fully, causing higher stress at the interface.

As can be seen from the surface energies of the wet surfaces (Table 6.2), water is stabilizing the interface as it interacts with the surface and in particular with surface $Mg^{2+}$. This agrees with previous work where dolomite and magnesite have a lower wet surface energy than pure calcite and the water showed a greater impact on the surface energies for the Mg-bearing minerals (Nora H. De Leeuw, 2002). Due to the stabilizing effect of water on the surface energy, there is no indication that the HMgF surface is more prone to dissolution than the pure and LMgF. We therefore explored the variability of the surface energy during the simulation relative to the optimised bulk configuration (cf. (Lardge et al., 2009; Parker et al., 2003)). Note, that the average dynamic surface energies and standard deviations obtained (Table S6.3) are significantly higher than the surface energies in Table 6.2, due to the absence of energy minimization. Yet, the order of the values is the same as for the dry surface energies, implying again that the HMgI is the least stable system. The larger variance in surface energy for the surfaces with an island (Table S6.3), indicates a more mobile surface, with the largest variance for the HMgI indicating a very dynamic surface. As expected, the structure further into the slab

and the interfacial water structure and dynamics were affected, which may imply that dissolving a surface island on a HMgI calcite is easier compared to a pure calcite and may be related to the deformed crystal template and the difficulty to find an island configuration that matches the crystal layer below.

The deformation of the crystal due to $Mg^{2+}$ incorporation has been mentioned before in experiments, and is generally thought to be the reason for the increase in solubility of high-Mg calcite (Ferreira, 2013). Furthermore, anhydrous experiments found that discrepancies in reactivity can be explained by stress in the crystal caused by the presence of $Mg^{2+}$, independent of the water-material interaction (Xu et al., 2013). Ab initio simulations supported both experiments and found a carbonate ion tilt by 6% due to the introduction of $Mg^{2+}$ into the crystal (Elstnerová et al., 2010). We also observed some minor tilting of the carbonate ions in the surface, although we did not quantify this behaviour (Figure 6.3b), while the compression observed for HMgF calcite along the *c*-axis (Figure 6.3a) is related to local alterations of $Ca^{2+}$–C distances (Figure 6.3b).



To summarise, the atomically flat surfaces show trends in overall surface energies that are opposite to those expected from known solubility trends, while the variability in surface energy with time does follow the solubility trend, suggesting that a more dynamic surface reflects a more soluble surface. Increasing structural deformation with increasing Mg content confirms previously reported structural destabilization of the crystal structure.

#### 6.4.1.3 Local Variability in Surface Free Energies

The trend in the surface energies of the flat surfaces, showing HMgF being more stable due to the structural rearrangement, appears to contradict experimental observations. However, when also considering more realistic surface structures such as edges and kinks/corners, the surface energies, crystal deformation and interfacial water distortion suggest that once the surface

contains imperfections, the system may become thermodynamically less stable, increasing the solubility product for HMgF as measured in experiments (Bischoff et al., 1987; Busenberg and Niel Plummer, 1989; Langmuir, 1997; Morse and Mackenzie, 1990; Robie and Hemingway, 1995). In order to investigate the impact of magnesium impurities, it is worth investigating the influence on the actual mechanism of dissolution, i.e., extraction of calcium ions from surface edge and kink sites as well as from atomically flat surfaces.

In general, calcite crystals will dissolve faster from terrace corners, edges and kink sites (e.g., (Bibi et al., 2018; Lasaga and Lüttge, 2003; Ruiz-Agudo and Putnis, 2012)) than from a flat surface. The extraction of a cation from an edge or corner generates new kinks (strongly under-coordinated surface sites). The free energy profile for such calcium edge or kink site extraction was investigated in HMgI and pure calcite with an island (PureI). With the notable exception of the acute edge calcium site, the overall $\Delta G_{extraction}$ was consistently lower for HMg island surface sites than for the same sites in the pure material (Table 6.3 and Figure 6.6b). Generally, it may therefore be concluded that rough HMg surfaces will dissolve more easily via calcium detachment than pure rough surfaces. The first minimum in the energy profiles corresponds to the equilibrium position in the crystal surface, the second minimum is at the same height as the first water layer near the surface (Figure 6.6b), which shows that the interfacial water molecules slightly stabilise extracted calcium ions. Note that there are subtle differences in interfacial water structure, density and local energy minima (Figure 6.6; and potentially carbonate molecule flexibility, see below) and that this interplay may contribute to defining the overall $\Delta G_{extraction}$ for surface calcium ions. For extraction of PureI and HMgI surface corner and edge calcium sites, the local energy minima within <0.5 nm from the surface suggest that there is stabilisation of the extracted calcium, most likely as an inner sphere complex. For the obtuse corner on HMgI, this inner sphere complex is energetically more stable than its position in the crystal.

The dissolution of flat surfaces first needs new etch pit nucleation events, which is an important (rate limiting) dissolution mechanism at near equilibrium conditions (Teng et al., 2000; Wolthers, 2015). The extraction of a single ion from a pure flat calcite face represents an unassisted etch pit nucleation event (i.e., without related sub-surface structural defect), which needs 204 kJ/mol according to our simulations (Table 6.3). The extracted $Ca^{2+}$ from the flat surfaces shows a similar energy profile as published before, although the shallow energy minimum observed around 0.3 nm is more pronounced in our simulation and the total $\Delta G_{extraction}$ at ~250 kJ/mol is higher (Spagnoli et al., 2006b). A possible explanation can be the use of a different water force field, whereas our water force field also gave a slightly lower density in the coordination shell of $Ca^{2+}$ (Kerisit and Parker, 2004; Koskamp et al., 2019). In both our results and the literature, the free energies converge at >0.5 nm away from the surface. The extracted ion can still interact with the vacant site, even though there are layers of water in between. The lack of a local minimum closer than 0.5 nm to the surface suggests that there is no stabilisation of an inner sphere complex at the flat surface. The formation of outer sphere complexes (i.e., the energy minima at >0.5 nm from the surface) initiates the convergence of the free energy profiles. This observation is also seen in the impure flat surfaces, where the shallow minima/maxima are positioned between the first and second water layer. Structural imperfections, such as (local) deformation due to Mg impurities, can assist or inhibit nucleation of new etch pits. Accordingly, in LMgF it is expected that calcium extraction generally needs more energy than from pure calcite. Depending on the local environment of the extracted calcium, we observed higher $\Delta G_{extraction}$ for LMgF than the pure mineral, albeit not for all calcium extractions (Table 6.3).

The more favourable $\Delta G_{extraction}$ for one of the LMgF sites may be linked to the flexibility of the coordinated carbonate molecules, since we observed that the substitution of Ca by smaller Mg ions generates more space in the crystal structure and therefore more flexibility in the

carbonate ions coordinated to that substitutional ion. This flexibility may enhance calcium extraction, and we observed a major increase in the $\Delta G_{extraction}$ by 300% when interfacial $CO_3^{2-}$ ions were frozen (Table 6.3). Such flexibility of surface carbonate groups has been found previously in X-ray reflective and diffractive experiments, where the flexibility is mainly expressed in the standard deviation of the tilting angle of carbonate, which was reportedly 13 times larger for the surface layer compared to the bulk (Fenter and Sturchio, 2012; Heberling et al., 2011). Although this variation could be due to the two main orientation directions of $CO_3^{2-}$ observed in AFM (Liang et al., 1996) and MD (e.g., Figure 6.1) or the surface roughness expected in every non-ideal, cleaved surface, there was no evidence found in the X-ray data for the existence of two distinguishable groups (Fenter and Sturchio, 2012). Furthermore, based on the 'roughness factor' introduced by Fenter et al., the impact of the surface roughness is negligible for X-ray reflectivity data interpretation (Fenter and Sturchio, 2012). It is very likely that the observed $CO_3^{2-}$ flexibility in X-Ray Diffraction (XRD) is due to imperfections and thereby indicates the presence of rotating carbonate sites on the calcite surface at geometrically more open surface sites (i.e., edge and kink sites). This is supported by the small flexibility observed in MD calculations of a perfectly flat calcite surface here and in previous investigations (Fenter et al., 2013), indicating that the increased coordination of corner and edge sites with water molecules might not be the only reason for more favourable dissolution energetics and higher reactivities of such sites, but that more opportunity for carbonate molecule rotation and flexibility plays a (potential key) role as well. When $Mg^{2+}$ is in the direct vicinity of $Ca^{2+}$, the shared coordinated carbonate molecules are held in place more strongly by the stronger bond with $Mg^{2+}$. Consequently, carbonate is less flexible and less likely to assist in the extraction of calcium. When distortion occurs in the local crystal environment, for example due to the longer $Ca^{2+}$-distance (Table S6.6), the carbonate molecules are more flexible, resulting in a $\Delta G_{extraction}$ that is slightly below the energy value of $Ca^{2+}$ in pure calcite.

The lowest value for assisted etch pit nucleation (191 kJ/mol) was observed for the flat HMg surface, although not all calcium extractions led to lower $\Delta G_{extraction}$ than in pure calcite (Table 6.3). These results indicate that it may be energetically more favourable to nucleate new etch pits on HMg calcite than pure calcite surfaces, although this strongly depends on the local magnesium distribution. Two out of three $\Delta G_{extraction}$ derived from the HMgF agree with the calculated surface energy, indicating a more stable surface. In one of these sites, smaller distances between Ca and $O_c$ are observed, which causes the $Ca^{2+}$ to be better charge-compensated and increases $\Delta G_{extraction}$ by 26 kJ/mol. In the other site, the local environment is structurally more similar to dolomite than high-Mg calcite and shows an increase in $\Delta G_{extraction}$ by 5 kJ/mol, despite the larger Ca–$O_c$ distances and higher water-exchange frequencies (Table S6.2). The calcium site with the most favourable $\Delta G_{extraction}$ showed unaltered bond distances compared to pure calcite, so it appears that here there is no extra compensation from the crystal, unless the nearby presence of Mg ions alters the local charge distribution. Note that classical MD is not capable of revealing such differences at the electronic level and previous ab initio calculations have not reported alterations in charge distributions upon Mg incorporation, but is found to stiffen the calcite structure (Elstnerová et al., 2010; Zhu et al., 2013) Alternatively, it may be that this calcium site is less stable than $Ca^{2+}$ in pure calcite, due to the characteristics of interfacial water: with a higher diffusion (Table S6.2) and less structured water (Figure 6.2 and Figure S6.3), $Ca^{2+}$ is more easily extracted and solvated.

To conclude, calcium extraction is generally more favourable from edge, kink, and flat surface sites in a high-magnesium surface than from pure calcite, depending on the subtle interplay between crystal surface and interfacial water structure and density that affects local energy minima. Moreover, the flexibility and opportunity for rotation of surface carbonate molecules may contribute to this complex interplay. This behaviour is particularly clear on atomically flat

surfaces, where magnesium impurities limit this flexibility, in particular in low-Mg calcite, thereby potentially playing a key role in the non-linear impact of Mg on calcite solubility.

### 6.4.2 Implications for the Influence of Magnesium on Calcite Growth

The forced MD may also be considered to represent the molecular scale energetics of ion approach and attachment/adsorption, when regarding the reversed process of ion extraction. For example, in contrast to the enhanced dissolution of a high-magnesium island surface, it will be more difficult to grow such a feature on a HMgF surface compared to growth on pure calcite. In particular, the deformed crystal template makes it more complicated to match the structure and propagate a step edge. AFM experiments have revealed similar behaviour; the first monolayer of $Mg^{2+}$ calcite grew normally on a pure calcite template crystal, but subsequent monolayers grew significantly more slowly (Astilleros et al., 2010). In addition, AFM experiments have shown that $Mg^{2+}$ is found to be a kink poisoner, by occupying a growth site, which drastically slows down growth (Arvidson et al., 2004, 2006; Nielsen et al., 2013; Ruiz-Agudo et al., 2009). In case of growth, our results strongly suggest $Ca^{2+}$ is less stable on the high-magnesium surface (island) and less prone to approach and attach compared to $Ca^{2+}$ on a pure calcite crystal. Moreover, we observed that the formation of adsorbed inner sphere complexes of calcium likely precedes incorporation of that calcium in acute corners and edge sites of PureI and HMgI, as well as the obtuse edges of the pure calcite island. In contrast, on the HMgI obtuse edge, the formation of such an inner sphere complex may be energetically more favourable than incorporation into the edge, potentially inhibiting the growth of the obtuse edge on high-Mg calcites.

## 6.5 Conclusion

We have performed molecular dynamics simulations of $\{10\bar{1}4\}$ calcite surfaces with and without different $Mg^{2+}$ concentrations and with and without an island on the crystal surface. With these simulations, the impact of $Mg^{2+}$ was determined on the overall and local solid-water interface structure, energetics, as well as the free energy profiles for the onset of calcite dissolution via $Ca^{2+}$ removal.

The overall $\{10\bar{1}4\}$ surface energy varies according to HMgF < Pure < LMgF < PureI < HMgI. Strong relaxation of the crystal structure was observed in HMgF and HMgI. LMg calcite relaxed its structure to a lesser extent than HMg calcite, showing a crystal structure very similar to pure calcite.

The average free energies for calcium ion extraction from the pure and Mg-doped surfaces followed roughly the opposite trend as the overall surface energy differences. The average energy needed to remove calcium ions increased from HMgI ≤ PureI < Pure < LMgF = HMgF, although very large local differences (from −13 kJ/mol to + 26 kJ/mol) were observed. The large variations and the resulting appearance of lower $\Delta G_{extraction}$, compared to pure calcite, lowers the threshold for unassisted nucleation of new etch pits locally, in particular for HMgF.

The local variation in free energy (minima) depends on the amount and distance to the nearest Mg, in addition to local disruption of interfacial water and the flexibility for carbonate ions to rotate. Local configurations were observed to be less stable when $Mg^{2+}$ was nearby in the surface, supporting experimental data in which calcite with higher percentages of $Mg^{2+}$ has a higher solubility compared to pure calcite up until the dolomite ratio (i.e, Ca:Mg = 1). Some of the free energy profiles showed a local energy minimum where the highest interfacial water density was observed (reflecting the hydrophilic nature of calcium).

Based on the interfacial water structure and dynamics, the surface energies and the $\Delta G_{extraction}$, the low-Mg calcite surface is comparable to pure calcite, although locally $Mg^{2+}$ induces stabilization of neighbouring $Ca^{2+}$, which results in slightly unfavourable new etch pit nucleation energetics. We can conclude that low concentrations of $Mg^{2+}$ tend to stabilise $Ca^{2+}$ and increase etch pit nucleation energies, and thereby decrease the calcite solubility, whereas higher concentrations of $Mg^{2+}$ lead to deformation of the surface crystal and interfacial water structure, which leads to local variabilities. The result is a thermodynamically less stable crystal and a higher solubility of high-Mg calcite compared to pure calcite, including more facile unassisted etch pit nucleation. Taken together, these results illustrate the molecular scale processes and demonstrate the first steps towards understanding of the non-linear impact of $Mg^{2+}$ on the solubility of magnesium-bearing calcites.

## 6.6 Supplementary Information

### 6.6.1 Force field

All potential parameters used in this study are listed in Table S6.1 (Raiteri et al., 2015, 2010; Wu et al., 2006).

Table S6.1 Potential parameters

| **Water (SPC/fw)** | | | | | |
|---|---|---|---|---|---|
| **Species** | | **Molecule/Atom** | | **Charge (e)** | |
| Mg | | Magnesium | | +2.000 | |
| Ca | | Calcium | | +2.000 | |
| $C_c$ | | Carbon from carbonate | | +1.135 | |
| $O_c$ | | Oxygen from carbonate | | -1.045 | |
| $O_W$ | | Oxygen from water | | -0.820 | |
| $H_W$ | | Hydrogen from water | | +0.410 | |
| **Bond styles** | | ***k* (ev Å$^{-2}$)** | | $r_0$ (Å) | |
| $C_c$—$O_c$ | Harmonic | 40.8493 | | 1.012 | |
| $O_W$—$H_W$ | Harmonic | 45.9296231 | | 1.012 | |
| **Angle styles** | | ***k* (ev rad$^{-2}$)** | | **$\Theta_0$** | |
| $H_w$— $O_W$ — $H_W$ | Harmonic | 3.29134 | | 113.24 | |
| | | **$\Theta_0$** | ***K*$_2$ (ev rad$^{-2}$)** | ***K*$_3$ (ev rad$^{-2}$)** | ***K*$_4$ (ev rad$^{-2}$)** |
| $O_c$— $C_c$ — $O_c$ | class2 | 120.0 | 6.617 | 0.0 | 0.0 |
| | | **M (ev Å$^{-2}$)** | ***r*$_1$ (Å)** | ***r*$_2$ (Å)** | |
| $O_c$— $C_c$ — $O_c$ | class2 bb | 12.818 | 1.3042 | 1.3042 | |
| | | **N$_1$ (ev Å$^{-2}$)** | **N$_2$ (ev Å$^{-2}$)** | ***r*$_1$ (Å)** | ***r*$_2$ (Å)** |
| $O_c$— $C_c$ — $O_c$ | class2 ba | 1.53319 | 1.53319 | 1.3042 | 1.3042 |
| **Improper styles** | | ***K*$_2$ (ev rad$^{-2}$)** | | ***K*$_4$ (ev rad$^{-2}$)** | |
| $O_c$— $C_c$ — $O_c$-$O_c$ | distance | 13.647 | | 360.0 | |
| Interatomic interactions | | | | | |
| **Lennard-Jones Potential** | | **ε (eV)** | | **σ (Å)** | |
| $O_W$ — $O_W$ | | 0.006739769454 | | 3.165492 | |
| Ca—$O_W$ * | | 0.000950 | | 3.35 | |
| Mg—$O_W$ * | | 0.001137 | | 2.82 | |
| **Buckingham Potential** | | **A (eV)** | **ρ (Å)** | **C (eV Å$^6$)** | |
| $O_c$—$O_W$ * | | 12534.455133 | 0.202 | 12.09 | |
| $O_c$—$H_W$ * | | 396.0 | 0.217 | 0.0 | |
| Ca—$O_c$* | | 3161.6335 | 0.271511 | 0.0 | |
| Ca—$C_c$ * | | 120000000 | 0.12 | 0.0 | |
| $O_c$— $O_c$* | | 63840.199 | 0.198913 | 27.89901 | |
| Mg—$O_c$* | | 3944.8613 | 0.23816 | 0.0 | |

*The interaction was set to zero over the range 6 – 9 Å

### 6.6.2 Umbrella sampling

An example of and input file used to perform the steered MD calculations prior to the umbrella sampling:

```
UNITS LENGTH=A

p: POSITION ATOM=1191

dx: COMBINE ARG=p.x PERIODIC=-47.70,47.70

MOVINGRESTRAINT

 ARG=dx

 STEP0=0        AT0=40.7  KAPPA0=0.0

 STEP1=1000     AT1=41.2  KAPPA1=20000.0

 STEP2=2000     AT2=41.7  KAPPA2=20000.0

 STEP3=3000     AT3=42.2  KAPPA3=20000.0

 STEP4=4000     AT4=42.7  KAPPA4=20000.0

 STEP5=5000     AT5=43.2  KAPPA5=20000.0

 STEP6=6000     AT6=43.7  KAPPA6=20000.0

 STEP7=7000     AT7=44.2  KAPPA7=20000.0

 STEP8=8000     AT8=44.7  KAPPA8=20000.0

 STEP9=9000     AT9=45.2  KAPPA9=20000.0

 STEP10=10000 AT10=45.7  KAPPA10=20000.0
```

STEP11=11000 AT11=46.2 KAPPA11=20000.0

STEP12=12000 AT12=46.7 KAPPA12=20000.0

STEP13=13000 AT13=47.2 KAPPA13=20000.0

STEP14=14000 AT14=47.7 KAPPA14=20000.0

STEP15=15000 AT15=48.2 KAPPA15=20000.0

STEP16=16000 AT16=48.7 KAPPA16=20000.0

STEP17=17000 AT17=49.2 KAPPA17=20000.0

STEP18=18000 AT18=49.7 KAPPA18=20000.0

STEP19=19000 AT19=50.2 KAPPA19=20000.0

STEP20=20000 AT20=50.7 KAPPA20=20000.0

STEP21=21000 AT21=51.2 KAPPA21=20000.0

STEP22=22000 AT22=51.7 KAPPA22=20000.0

STEP23=23000 AT23=52.2 KAPPA23=20000.0

STEP24=24000 AT24=52.7 KAPPA24=20000.0

STEP25=25000 AT25=53.2 KAPPA25=20000.0

STEP26=26000 AT26=53.7 KAPPA26=20000.0

STEP27=27000 AT27=54.2 KAPPA27=20000.0

STEP28=28000 AT28=54.7 KAPPA28=20000.0

STEP29=29000 AT29=55.2 KAPPA29=20000.0

STEP30=30000 AT30=55.7 KAPPA30=20000.0

STEP31=31000 AT31=56.2 KAPPA31=20000.0

STEP32=32000 AT32=56.7 KAPPA32=20000.0

STEP33=33000 AT33=57.2 KAPPA33=20000.0

... MOVINGRESTRAINT

PRINT ARG=* STRIDE=100 FILE=COLVAR

After the steered MD, the different frames with the calcium at different positions above the surface were taken as the input configuration for the different umbrella calculations. Below an example of an umbrella calculation

UNITS LENGTH=A

p: POSITION ATOM=1191

dx: COMBINE ARG=p.x PERIODIC=-47.70,47.70

dy: COMBINE ARG=p.y PERIODIC=-23.984,23.984

dz: COMBINE ARG=p.z PERIODIC=-20.150,20.150

RESTRAINT ARG=dx AT=40.7 KAPPA=48.2

RESTRAINT ARG=dy AT=23.89 KAPPA=482.0

RESTRAINT ARG=dz AT=8.91 KAPPA=482.0

PRINT ARG=* STRIDE=10 FILE=COLVAR

To evaluate if there was a satisfactory overlap between the sampling of the umbrellas a histogram was constructed. Figure S6.1 shows an example of the histograms obtained over all the umbrellas used to construct the energy profile. The higher intensities indicate the stronger spring constant of that umbrella. All systems showed similar overlap.

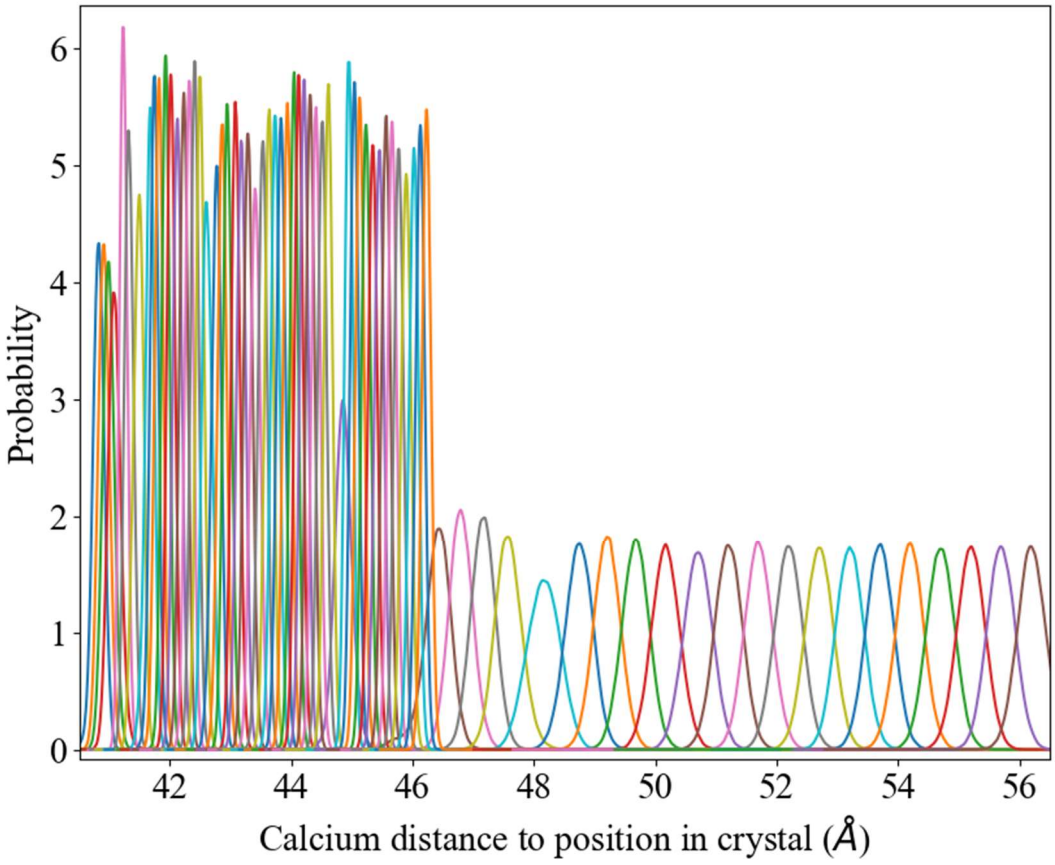



Figure S6.1 An example of the histograms obtained from all umbrellas placed along the CV

### 6.6.3 Water dipole angle relative to the surface

The angle between the dipole of the water coordinated to the cations in the surface with the surface plane revealed the orientation of the water molecules relative to the surface. An angle of 90 degrees means that the water molecule is standing normal to the surface, with protons pointing towards the solution. An angle of zero means that a water molecule is lying parallel to the surface. As Kerisit and Parker (Kerisit and Parker, 2004) showed, this angle changes

while molecules move to or from the surface, indicating that a water molecule is rotating (Kerisit and Parker, 2004). The dipole angles of water on top of the different surfaces are shown in Table 6.1 in the main text. The average angle on the HMgF calcite was 0.24° smaller compared to the average angle on the pure calcite, indicating that the water molecules were positioned (almost) perpendicular to the surface in the HMgF interface.

### 6.6.4 Vibrational spectrum

Insights in the dynamics of interfacial water were obtained from the analysis of the Vibrational Density of States (VDOS). The VDOS was taken as the sum of the Fourier transformed Velocity AutoCorrelation Function (VACF) of both oxygen and hydrogen atoms of the water. At high frequencies (Figure S6.2 on the left) the bending and stretching vibrations of water (~1500 $cm^{-1}$ and ~3700 $cm^{-1}$, respectively) were represented. For both, pure calcite and HMgF calcite, interfacial water (<3.5 angstrom from the surface) showed similar intramolecular vibrations. Also, the intensity and position were comparable to the VDOS of bulk water for the bending, but interfacial water expressed a little shift to a higher energy for the stretching vibration (~3700 $cm^{-1}$). The wavenumber range related with vibrations and librations of the hydrogen bond network were expressed in the region below 1000 $cm^{-1}$, with, from left to right, first the intermolecular motion of the O-O-O bonding (~ 50 $cm^{-1}$ in bulk water), second the O–O intermolecular stretching (~ 250 $cm^{-1}$ in bulk water) and third the intermolecular rotational modes (~300 – 1000 $cm^{-1}$ in bulk water), Figure S6.3 is a zoom in this region. The vibrational modes representing the H-bond network of interfacial water on top of pure and HMgF calcite system were shifted towards higher energies compared to bulk water. Also, the intermolecular stretching of the hydrogen bond is more pronounced (~400 $cm^{-1}$) compared to a pure water system. The hydrogen bond network above HMgF calcite showed less intermolecular stretching and intermolecular O-O-O bonding motion. The librational (rotational) mode on the

other hand was slightly more expressed in the interfacial water on top of a HMgF calcite system.

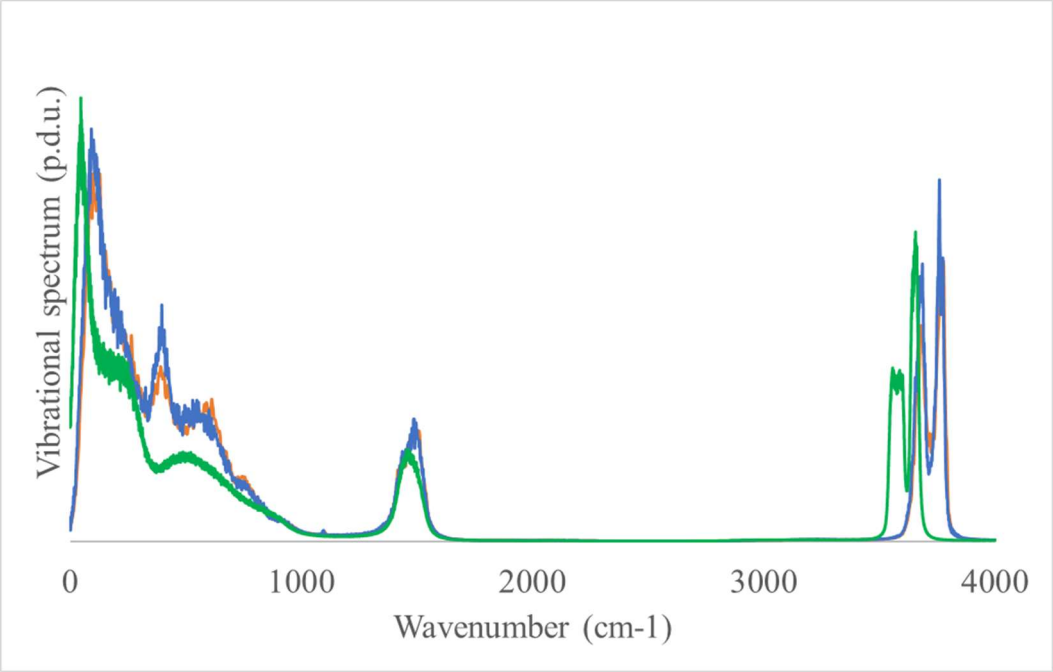



Figure S6.2 VDOS of interfacial water on top of a pure calcite (blue) and HMgF calcite (orange). For comparison and forcefield specific peak positions, in green the VDOS spectrum of pure bulk water



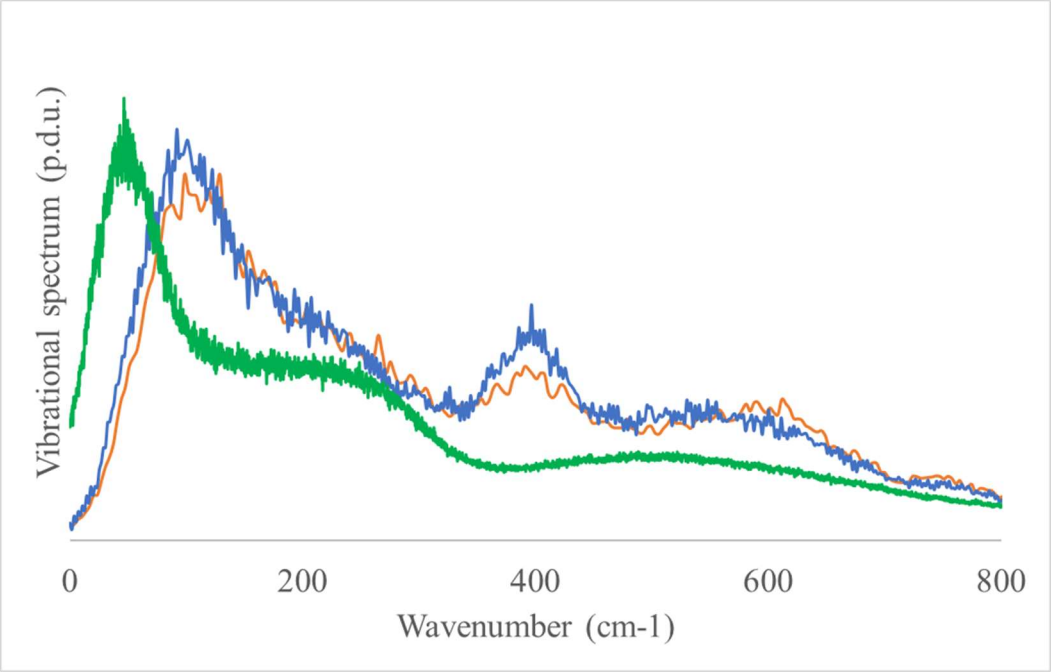



Figure S6.3 Zoom of VDOS spectrum range related to characteristics of the hydrogen bond network. Interfacial water on top of a pure calcite (blue) and HMgF calcite (orange). Pure bulk water (green).

### 6.6.5 Exchange frequency and diffusion coefficient

The average exchange frequency of water in the coordination shell of $Ca^{2+}$ in the flat surface is shown in Table S6.2. The number of exchanges per ns near surface calcium ions was higher when $Mg^{2+}$ was present in the surface. Because $Mg^{2+}$ was randomly distributed over the surface, local environment could vary strongly, from calcium ions without any $Mg^{2+}$ in the direct vicinity to being fully surrounded by $Mg^{2+}$. To visualize, the variation in exchange frequency on top of pure and HMgF is shown in Figure S6.4 and Figure S6.5. The range of exchange frequencies of water on top of pure calcite and low $Mg^{2+}$ calcite varied from 6 to 54 $ns^{-1}$ while the range for HMgF calcite is wider and goes from 6 to 98 $ns^{-1}$. In addition, the diffusion coefficient of the interfacial water (set at a thickness of 0.350 nm, based on the first minimum in the $Ca^{2+}$ — $O_w$ RDF for calcium in solution) was calculated and shown in Table 6.1 in the main text. The three components (x, y, and z) were split. This way we can exclude the diffusion perpendicular (z-axis) to the surface, because, due to the small finite size in the z direction, the accuracy of the diffusion coefficient in this direction is very low. As can be seen in the table, the diffusion coefficient parallel to the surface was higher when water was in contact with HMgF calcite. Apart from that, the diffusion along the x-axis was, in both pure and HMgF calcite, approximately two times the diffusion coefficient along the y-axis. Along the x-axis, the distance between two cations was around 0.490 nm while the distance between two cations along the y-axis was 0.398 nm.

Table S6.2 Average exchange frequencies and diffusion coefficients of water coordinated to $Ca^{2+}$ in a flat calcite surface

| System | $N_{ex}$·ns$^{-1}$ | $D_x$ ($10^{-9}$ m$^2$s$^{-1}$) | $D_y$ ($10^{-9}$ m$^2$s$^{-1}$) | $D_z$ ($10^{-9}$ m$^2$s$^{-1}$) |
|---|---|---|---|---|
| Pure calcite | 25.7 | 0.031 | 0.017 | 0.003 |
| Low $Mg^{2+}$ flat surface | 24.9 | 0.031 | 0.018 | 0.003 |
| High $Mg^{2+}$ flat surface | 31.4 | 0.046 | 0.022 | 0.004 |
| Values from literature | | | | |
| Pure_DM[a] | 78 | | | |
| Pure_SF[b] | 0.52 | | | |
| Pure_ST[c] | 3.3 | 0.6 | 0.6 | >0.2 |
| Pure_ME[d] | - | 0.1 | 0.1 | - |

[a] Obtained using Direct method (Wolthers et al., 2013); [b] Obtained using Survival Function (De La Pierre et al., 2016); [c] Diffusion obtained using short (5 ps) trajectories (Kerisit and Parker, 2004); [d] Obtained using the modified Einstein equation (Mutisya et al., 2017).

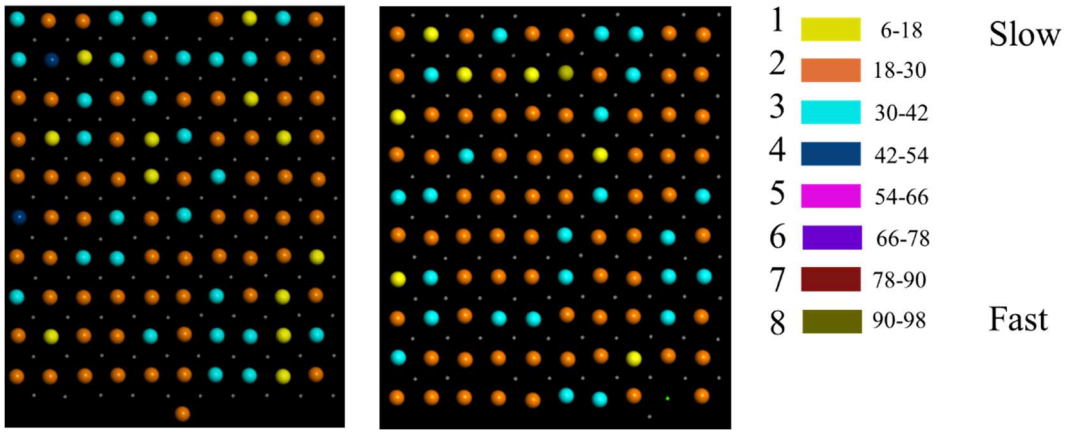



Figure S6.4 Variation in exchange frequency of water on top of pure calcite. The colours follow the exchange frequencies indicated in the legend. The highest $N_{ex}$ found in pure calcite was 54 s$^{-1}$

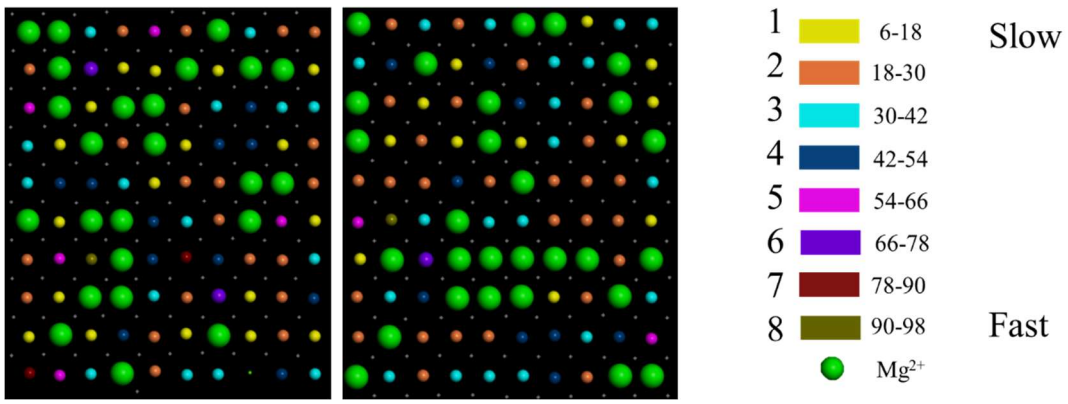



Figure S6.5 Variation in exchange frequency of water on top of HMgF calcite. The colours follow the exchange frequencies indicated in the legend. The green spheres represent $Mg^{2+}$ as indicated with the red arrow, where no exchanges occurred over the course of the simulations.

### 6.6.6 Dynamic surface energies and bulk distances

The dynamic surface energies (Table S6.3) showed the same trend as the surface energy of the dry systems (Table 6.2 in the main text). However, the difference between the systems is less subtle with the dynamic surface energy of HMgI being almost twice as high as the flat surfaces. In addition, the standard deviation was higher for the slabs with an island on top, where PureI had a standard deviation closer to the flat surfaces than to HMgI.

Table S6.3 Surface energies of after relaxation at 300K with water on top. The $U_s$ in equation 1 was calculated as the energy of the configuration without water

| $\{10\bar{1}4\}$ System | Surface Energies ($J/m^2$) | Standard deviation |
|---|---|---|
| Pure calcite | 1.75 | 0.013 |
| Low $Mg^{2+}$ flat surface | 1.78 | 0.013 |
| High $Mg^{2+}$ flat surface | 1.73 | 0.014 |
| Island pure calcite | 1.83 | 0.021 |
| Island high $Mg^{2+}$ | 3.25 | 0.036 |

Table S6.4 A summary of computed relaxed surface energies of calcite

| Surface Energies ($J/m^2$) | Reference |
|---|---|
| 0.510 | (Akiyama et al., 2011) |
| 0.503 | (Bruno et al., 2010) |
| 0.860 | (Kvamme et al., 2009) |
| 0.380 | (Kerisit et al., 2003) |
| 0.590 | (De Leeuw and Cooper, 2004) |
| 0.430 | (Kerisit et al., 2003) |
| 0.590 | (Kerisit et al., 2003) |
| 0.590 | (Parker et al., 2003) |
| 0.520 | (Parker et al., 2003) |
| 0.420 | (Parker et al., 2003) |
| 0.534 | (Rohl et al., 2003) |
| 0.706 | (Braybrook et al., 2002) |
| 0.590 | (Cooper and De Leeuw, 2002) |
| 0.590 | (Nora H. De Leeuw, 2002) |
| 0.863 | (Hwang et al., 2001) |
| 0.322 | (Wright et al., 2001) |
| 0.590 | (De Leeuw and Parker, 1998) |
| 0.590 | (Titiloye et al., 1998) |
| 0.570 | (Nygren et al., 1998) |
| 0.600 | (De Leeuw and Parker, 1997) |
| 0.530 | (Parker et al., 1997) |
| 0.230 | (Parker et al., 1993) |

To evaluate how stressed the local structure was around the extracted $Ca^{2+}$, we compared its cation distances with the ones in the structure of pure calcite and dolomite. Average $Ca^{2+}$ — $I^{+/-}$ distances (where $I^{+/-}$ is $O_c$, $Ca^{2+}$) in LMg were equal to the pure calcite. This is illustrated by the RDF's in Figure 2a of the main text and in Table S6.2, where the shortest average distances in the crystal between the ions are listed.

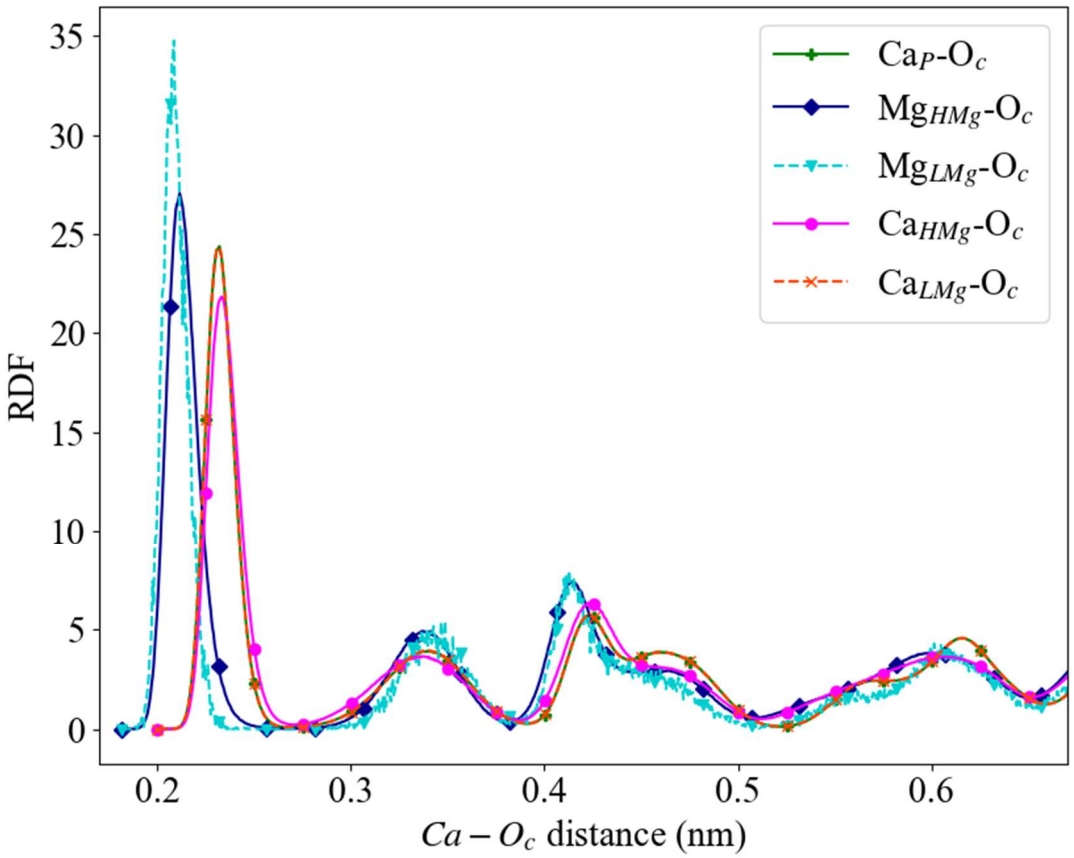



Figure S6.6 Radial distribution function for the cation with the oxygen of carbonate ($O_c$) in pure, LMg and HMg calcite crystals

Table S6.5 Coordination distances in the bulk crystal structure

| Element pair | Distance (nm) |
|---|---|
| **Calcite** | |
| **$Ca^{2+}$ — $Ca^{2+}$ (c-axis) *** | 0.405; 0.411±0.020 |
| **$Ca^{2+}$ — $Ca^{2+}$ (b-axis) *** | 0.499; 0.505±0.023 |
| **$Ca^{2+}$— $O_c$*** | 0.233; 0.233±0.006 |
| **Dolomite** | |
| **$Mg^{2+}$— $Ca^{2+}$** | 0.387 |
| **$Ca^{2+}$— $Ca^{2+}$** | 0.482 |
| **$Mg^{2+}$— $O_c$** | 0.212 |
| **$Ca^{2+}$— $O_c$** | 0.241 |
| **Low $Mg^{2+}$ flat surface** | |
| **$Mg^{2+}$— $Ca^{2+}$ (c-axis)** | 0.388±0.014 |
| **$Mg^{2+}$— $Ca^{2+}$ (b-axis)** | 0.491±0.025 |
| **$Ca^{2+}$— $Ca^{2+}$ (c-axis)** | 0.411±0.020 |
| **$Ca^{2+}$— $Ca^{2+}$ (b-axis)** | 0.505±0.023 |
| **$Mg^{2+}$— $O_c$** | 0.209±0.006 |
| **$Ca^{2+}$— $O_c$** | 0.233±0.007 |
| **High $Mg^{2+}$ flat surface** | |
| **$Mg^{2+}$— $Ca^{2+}$ (c-axis)** | 0.392±0.012 |
| **$Mg^{2+}$— $Ca^{2+}$ (b-axis)** | 0.485 ±0.012 |
| **$Ca^{2+}$— $Ca^{2+}$ (c-axis)** | 0.400±0.013 |
| **$Ca^{2+}$— $Ca^{2+}$ (b-axis)** | 0.479±0.013 |
| **$Mg^{2+}$— $O_c$** | 0.213±0.007 |
| **$Ca^{2+}$— $O_c$** | 0.235±0.008 |

*the second reported distance is the value retrieved form our own simulations



### 6.6.7 LMg local distances

Table S6.6 Relevant distances between surface calcium and the neighbouring atoms and calcium coordination numbers at the different $Ca^{2+}$ positions near $Mg^{2+}$ in the low $Mg^{2+}$ $\{10\bar{1}4\}$ calcite surface

| Element pair | Distance to $Mg^{2+}$ (nm) | Distance to $Ca^{2+}$ (nm) | Distance to $O_c$(nm) |
|---|---|---|---|
| $Ca_{LMgF1}$ | 0.3925 | 0.3975; 0.4525;0.4925 0.6275 | 0.2325 |
| $Ca_{LMgF2}$ | 0.4525 | 0.3975; 0.4875; 0.6325 | 0.2325 |
| $Ca_{LMgF3}$ | 0.6225 | 0.3975; 0.4975; 0.6325 | 0.2325 |
| $Ca_{LMgF4}$ | 0.7975 | 0.3975; 0.5175; 0.6325 | 0.2325 |

### 6.6.8 Surface with an island

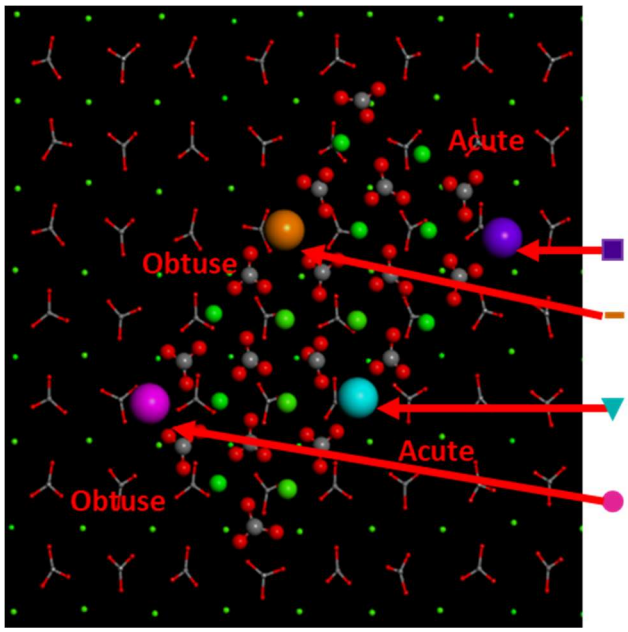



Figure S6.7 Top view of the 4x8 atoms island on top of Pure $\{10\bar{1}4\}$ surface carbonate groups in grey (C) and red (O) and $Ca^{2+}$ in small green or colour and symbol coded larger spheres to link them to the free energy profiles (Figure 5b) in orange (obtuse edge, —), turquoise (acute edge, ▼), pink (obtuse corner, ●) and purple (acute corner, ■).

Table S6.7 Comparison of relevant distances and coordination in the surface of PureI versus HMgI. Between the selected calcium (Figure S6.7) and the neighbouring atoms. *For calcium in the obtuse edge ($Ca_{POE}$ and $Ca_{HMgIOE}$) a cut-off of 0.30 nm was used for the coordination number. Based on the first minimum in the corresponding RDF's.

| Element pair | Distance (nm) | Coordination Number (<0.350 nm*) |
|---|---|---|
| **Calcium acute corner** | | |
| $Ca_{PAC}$ — $O_c$ | 0.2375 | 4.02 |
| $Ca_{HMgIAC}$ — $O_c$ | 0.2325 | 4.03 |
| $Ca_{PAC}$ — $O_W$ | 0.2375 | 4.38 |
| $Ca_{HMgIAC}$ — $O_W$ | 0.2375 | 4.96 |
| $Ca_{HMgIAC}$ — Nearest $Mg^{2+}$ on island | 0.4125 | |
| $Ca_{PAC}$ — Nearest $Ca^{2+}$ on island | 0.4325 | |
| $Ca_{HMgIAC}$ — Second $Mg^{2+}$ on island | 0.6375 | |
| $Ca_{PAC}$ — Second $Ca^{2+}$ on island | 0.6225 | |
| **Calcium obtuse corner** | | |
| $Ca_{POC}$ — $O_c$ | 0.2375 | 5.04 |
| $Ca_{HMgIOC}$ — $O_c$ | 0.2325 | 5.66 |
| $Ca_{POC}$ — $O_W$ | 0.2375 | 3.94 |
| $Ca_{HMgIOC}$ — $O_W$ | 0.2325 | 4.02 |
| $Ca_{HMgIOC}$ — Nearest $Mg^{2+}$ on island | 0.3875 | |
| $Ca_{POC}$ — Nearest $Ca^{2+}$ on island | 0.4075 | |
| $Ca_{HMgIOC}$ — Second $Mg^{2+}$ on island | 0.6075 | |
| $Ca_{POC}$ — Second $Ca^{2+}$ on island | 0.6325 | |
| **Calcium acute edge** | | |
| $Ca_{PAE}$ — $O_c$ | 0.2375 | 5.09 |
| $Ca_{HMgIAE}$ — $O_c$ | 0.2325 | 5.12 |
| $Ca_{PAE}$ — $O_W$ | 0.2375 | 3.50 |
| $Ca_{HMgIAE}$ — $O_W$ | 0.2325 | 4.03 |
| $Ca_{HMgIAE}$ — Nearest $Ca^{2+}$ on island | 0.4075 | |
| $Ca_{HMgIAE}$ — Nearest $Ca^{2+}$ on island | 0.4425 | |
| $Ca_{PAE}$ — Nearest $Ca^{2+}$ on island | 0.4325 | |
| $Ca_{HMgIAE}$ — Second $Mg^{2+}$ on island | 0.5775 | |
| $Ca_{HMgIAE}$ — Second $Ca^{2+}$ on island | 0.6075 | |
| $Ca_{HMgIAE}$ — Second $Ca^{2+}$ on island | 0.6575 | |
| $Ca_{PAE}$ — Second $Ca^{2+}$ on island | 0.6225 | |

Table S6.7 continued. Comparison of relevant distances and coordination in the surface of PureI versus HMgI. Between the selected calcium (Figure S6.7) and the neighbouring atoms. *For calcium in the obtuse edge ($Ca_{POE}$ and $Ca_{HMgIOE}$) a cut-off of 0.30 nm was used for the coordination number. Based on the first minimum in the corresponding RDF's.

| Element pair | Distance (nm) | Coordination Number (<0.350 nm*) |
|---|---|---|
| **Calcium obtuse edge** | | |
| $Ca_{POE}$ — $O_c$ | 0.2325 | 5.45* |
| $Ca_{HMgIOE}$ — $O_c$ | 0.2325 | 4.70* |
| $Ca_{POE}$ — $O_W$ | 0.2325 | 2.20* |
| $Ca_{HMgIOE}$ — $O_W$ | 0.2325 | 2.37* |
| $Ca_{HMgIOE}$ — Nearest $Mg^{2+}$ on island | 0.4025 | |
| $Ca_{HMgIOE}$ — Nearest $Ca^{2+}$ on island | 0.5025 | |
| $Ca_{POE}$ — Nearest $Ca^{2+}$ on island | 0.4383 | |
| $Ca_{HMgIOE}$ — Second $Mg^{2+}$ on island | 0.5925 | |
| $Ca_{HMgIOE}$ — Second $Mg^{2+}$ on island | 0.6325 | |
| $Ca_{HMgIOE}$ — Second $Ca^{2+}$ on island | 0.6275 | |
| $Ca_{POE}$ — Second $Ca^{2+}$ on island | 0.6375 | |

# Chapter 7 Synthesis

In this thesis, the reactivity of calcium ions and the formation of calcium carbonate clusters and the (Mg-doped) calcite surface using classical Molecular Dynamics (MD) and molecular mechanics (MM) simulations was investigated. The output of the MD simulations was post-processed to obtain information on the energetics, dynamics, structure, and other defining characteristics of the various systems under investigation.

The performed simulations were on atomic scale, and the description of the systems was based on a combination of interatomic potentials described in a forcefield. For the validation of our results, we compared our findings with experimental and *ab initio* studies. This research elucidated the kinetics and thermodynamics of calcium ions dissolved in pure water, dissolved in supersaturated $CaCO_3$ water, during cluster formation with and without biomolecules present, and in different positions in and adsorbed on a calcite surface with and without Mg impurities.

The first two chapters are a general introduction and a brief discussion of the theoretical background of the main computational method used in the thesis. The introduction addresses the various research areas and applications of $CaCO_3$, significant previous work and the scope and research questions of the thesis. The remainder of the introduction summarizes the major findings of each study discussed in chapters 3 to 6.

Chapter 3, also published as (Koskamp et al., 2019) reports the water dynamics around dissolved $Ca^{2+}$ with classical molecular dynamics (MD). The dynamics of water and the related dehydration of cations is assumed to be the rate-limiting step in many reactions involving those cations. We used molecular dynamics to generate trajectories of which we could extract water exchange frequencies. We evaluated several conventional force fields and compared two different post-processing methods, namely the direct method and the method based on the survival function. Our results show that the choice in force field and post-processing method

have a significant impact on the outcome. However, all our results also suggest that the water dynamics around calcium is significantly higher than the well-known experimental observations. The two orders of magnitudes difference place our results in the same range as the exchange frequencies observed with experimental neutron scattering and *ab initio* MD simulations in which the description of the interactions is more accurate. The survival function method gave the most comparable results and it allows characterization of the dynamics of water molecules, in particular, by distinguishing between fast and slow exchanges. This study heavily suggests that the rapid water exchange kinetics in the coordination shell of dissolved calcium is not rate limiting for any further reaction, which includes a dehydration step. This implies that the mechanistic rate laws for adsorption, growth, and dissolution using this calcium dehydration as the rate-limiting step need to be revaluate. This most likely also applies for other alkali and most of the earth alkali metals.

In chapter 4 we studied the first stage in the $CaCO_3$ precipitation process: the formation of $CaCO_3$ clusters and the interaction with aspartic acid-based biomolecules. We followed the formation of amorphous calcium carbonate (ACC) clusters in a supersaturated system in liquid water, by for instance, monitoring the cluster size distribution, $Ca:CO_3$ stoichiometry of the formed clusters, and the solvent accessible surface area. The same analysis was performed when aspartic acid-based biomolecules were added to the system to study the impact of these biomolecules upon the initial steps of mineralization. Among the biomolecules was the protein ovocleidin-17, found in eggshells, as this protein contains aspartic amino acids. The first finding is that ion pairs of calcium and carbonate were formed instantly and independently of the presence of biomolecules. Other key findings of this research mainly point out the difference between pure $CaCO_3$ and the biomolecule-containing system: upon cluster formation, the main impact of the biomolecule was observed by inhibition of ACC dehydration, compared to the system where the biomolecule is absent. Moreover, the biomolecules stimulate

the formation of particles containing more $CO_3^{2-}$ than $Ca^{2+}$ ions, therefore creating clusters that have an overall negative charge. The last key finding is that the biomolecules lead to distinctive cluster shapes. In conclusion, our results proof that biomolecules change the way the ions aggregate in the early stages prior to any (bio)crystallization.

Chapter 5 is devoted to reparametrizing and optimizing the force field interactions between the biomolecule and the cation. As discussed in the introduction of this thesis, a solid description of the system is crucial to get reliable results. In the work presented in chapter 4, we used a generally accepted description in our force field that is standardized for biomolecular systems. However, these interactions are often not explicitly parameterized, and fall short in accuracy. In this research, we studied different descriptions of calcium in combination with aspartic and glutamic acid. We used meta-dynamic molecular dynamic calculations to extract the energy profiles and, additionally, we calculated the corresponding constants of $Ca^{2+}$ - carboxylate association. Besides replacing the calcium Lennard-Jones parameters with the parameters used in $CaCO_3$ forcefield, we rectified the $Ca^{2+}$ - carboxylate by modifying the $\sigma_{ij}$ value (i.e. the minimum distance between the two atoms) in the heteroatomic Lennard-Jones interaction by 0.5% in order to have a consensus with experimental measurements of the association constant. This new description of the system simultaneously improved the structural properties of the ion pair. Mainly, this work concluded that the reparameterization was successful and the force field as used in chapter 4 should be updated accordingly. The impact on the results of chapter 4 are briefly mentioned.

Finally, my last work in this thesis is discussed in chapter 6, where we investigated the reactivity of $CaCO_3$ (see also (Koskamp et al., 2021)) and the impact of Mg impurities. We studied a calcite surface and the interaction with water by looking at the stability of calcium positioned in the crystal surface and adsorbed in the form of a small $CaCO_3$ island that can represent either a growth island or an edge of a larger etch pit. We investigated the impact of

disseminated magnesium, a very common impurity in calcium carbonate salts, in the crystal slab. This is again an MD study where we combined it with umbrella sampling to retrieve thermodynamical data via free energy profiles. Our results showed that this impurity deformed the crystal and changed the overall surface energy, local free energy profiles, interfacial water density, structure, and dynamics when Mg concentration is high (~30% of Ca replaced by Mg). At low concentrations of Mg, the overall crystal properties remained unchanged, however, locally, $Ca^{2+}$ stabilised and we observed a tendency in increased etch pit nucleation energy. The $Ca^{2+}$ ions adsorbed as part of a growth island were, as expected, thermodynamically less stable compared to in a flat surface. Moreover, with the presence of Mg, the island surface destabilised even further. This resulted in an increased surface energy and lower extraction energies. Locally, a complex interplay of several different factors, like the distance from the extracted Ca to the nearest Mg, the possibility for the carbonate ions to rotate, and the local interfacial water structure, resulted in a divergence in local extraction energies. All in all, the results of our MD study delineate the non-linear impact of Mg on a calcium ion and explains the non-linear relation between Mg impurity concentration and the solubility of the crystal.

As an overall conclusion, the four chapters demonstrate the impact of different components during different stages of the (bio)mineralization process. Subsequent research can be done on chapter 4 and 5 where we could look at the existence of prenucleation clusters and if there is a tendency of the formation of proto-calcite, proto-vaterite and or proto-aragonite. Furthermore, studying the impact of a combination of biomolecules with dissolved Mg in the initial steps of amorphous calcium carbonate (ACC) formation might give us different insights in how we can tailor crystal formation and how calcifying organisms control crystallization (e.g. (Kellock et al., 2020; Meldrum and Cölfen, 2008)). Although these molecular mechanisms are far below our sense of reality in our daily life, they are the motor of what we see. Consequently, the insights from this thesis are important to comprehend natural and laboratory observations and

are essential in targeted searches for new applications such as tailoring crystal formation in a sustainable way.

# Chapter 8 Nederlandse samenvatting

8

Het aardoppervlak is rijk aan calciumcarbonaten (Sekkal and Zaoui, 2013), bijvoorbeeld in carbonaatrijke sedimentaire gesteenten, koraalriffen, stalactieten en stalagmieten. Onderzoek naar de vorming van calciumcarbonaat is belangrijk in een breed scala van (toegepaste) onderzoeks- en engineeringgebieden als ook in de industrie. Van geldgedreven (onder andere olie-industrie (Olajire, 2015), papierindustrie (Vashistha et al., 2021), aardwarmtewinning (Pandey et al., 2018), drinkwatersystemen (Richards et al., 2018), en ontzilting van (drink)water (Burn et al., 2015; Zhao et al., 2018), en het conserveren van cultuur en natuur, tot het zoeken naar nieuwe duurzame toepassingen.

In dit proefschrift is gekeken naar de processen op molecuul schaal en energie verschillen die bepalen hoe en hoe snel calcium carbonaat (nano)deeltjes kunnen vormen, en wat de invloed is van magnesium en amino zuren op deze processen. Met dit doel is de reactiviteit van calciumionen, de invloed van aminozuren op de vorming van calciumcarbonaatclusters en die van magnesium op de stabiliteit van het calciet oppervlak onderzocht. Hiervoor zijn simulaties uitgevoerd met behulp van klassieke Molecular Dynamics (MD) en Moleculaire Mechanica (MM). Uit de MD simulaties werd informatie geëxtraheerd over bijvoorbeeld de energieën, dynamiek, structuur en andere bepalende kenmerken van de verschillende onderzochte systemen.

De simulaties werden uitgevoerd op atomaire schaal en de beschrijving van de systemen was gebaseerd op een combinatie van interatomaire potentialen beschreven in een zogenoemd *forcefield*. Voor de validatie van onze resultaten werden de bevindingen vergeleken met experimentele en *ab initio* studies. Deze studie richtte zich op het verhelderen van de kinetiek en thermodynamica van calciumionen opgelost in zuiver water, opgelost in oververzadigd $CaCO_3$ water, tijdens clustervorming met en zonder de aanwezigheid van biomoleculen, en op verschillende posities in en geadsorbeerd aan een calcietoppervlak met en zonder Mg onzuiverheden.

De eerste twee hoofdstukken zijn een algemene inleiding en een korte bespreking van de theoretische achtergrond van de belangrijkste simulatiemethoden die in het proefschrift werden gebruikt. De inleiding behandelt de verschillende onderzoeksgebieden en toepassingen van $CaCO_3$, significante onderzoeksresultaten van eerdere studies, het doel en de onderzoeksvragen van het proefschrift. De rest van de inleiding geeft een opsomming van de belangrijkste bevindingen van elk onderzoek dat in de hoofdstukken 3 tot en met 6 is besproken.

Hoofdstuk 3, ook gepubliceerd als (Koskamp et al., 2019) behandeld de water dynamica rond opgelost $Ca^{2+}$ met klassieke Moleculaire Dynamica (MD). De dynamiek van water en de daarmee samenhangende dehydratatie van kationen wordt verondersteld de snelheidsbepalende stap te zijn in veel reacties waarbij die kationen betrokken zijn. De gegenereerde MD trajecten werden gebruikt om de water uitwisselingsfrequenties vast te stellen. Verschillende conventionele *forcefields* werden getoetst en twee verschillende data verwerkingsmethode werden vergeleken, namelijk de directe methode en de methode gebaseerd op de overlevingsfunctie. De resultaten laten zien dat de keuze in *forcefield* en analysemethode een significante impact hebben op de uitkomst. Al onze resultaten suggereren echter ook dat de water dynamiek rond calcium significant hoger is dan de bekende en meest gebruikte experimentele observaties. De huidige resultaten vallen wel in dezelfde orde van grootte als de uitwisselingsfrequenties waargenomen met experimentele neutronenverstrooiing en *ab initio* MD-simulaties waarin de beschrijving van de interacties nauwkeuriger is. De overlevingsfunctiemethode gaf de meest vergelijkbare resultaten en maakt het mogelijk om de dynamiek van watermoleculen te karakteriseren, met name door onderscheid te maken tussen snelle en langzame uitwisselingen. Dit resultaat suggereert sterk dat de snelle water uitwisselingssnelheden in de coördinatieschil van opgelost calcium niet snelheidsbeperkend is voor enige verdere reactie, waar de dehydratatie stap deel van uit maakt. Dit houdt in dat de mechanistische snelheidswetten voor adsorptie, groei en ontbinding met deze calcium

dehydratatie als snelheidsbeperkende stap opnieuw moeten worden geëvalueerd. Dit geldt hoogstwaarschijnlijk ook voor andere alkalische en de meeste aardalkalimetalen.

In hoofdstuk 4 wordt de eerste fase in de $CaCO_3$ neerslagreactie bestudeerd: de vorming van $CaCO_3$ clusters en nanodeeltjes en de interactie met asparaginezuur-houdende biomoleculen. De vorming van clusters van amorf calciumcarbonaat (ACC) in een oververzadigd systeem in water werd geanalyseerd, door bijvoorbeeld het monitoren van de clustergrootte verdeling, $Ca:CO_3$-stoichiometrie van de gevormde clusters en de grootte van het oppervlak dat toegankelijk is voor water. Dezelfde analyse werd uitgevoerd voor de systemen waar de biomoleculen werden toegevoegd om de impact van deze biomoleculen op de eerste stappen van mineralisatie te bestuderen. Onder de biomoleculen bevond zich het eiwit ovocleidin-17, dat in eierschalen voorkomt, dit eiwit bevat asparaginezuren . De eerste bevinding is dat ionenparen van calcium en carbonaat onmiddellijk en onafhankelijk van de aanwezigheid van biomoleculen werden gevormd. Andere belangrijke resultaten van dit onderzoek wijzen voornamelijk op het verschil tussen puur $CaCO_3$ en de systemen met biomoleculen: bij clustervorming werd de belangrijkste impact van het biomolecuul waargenomen door de vertraging van ACC-dehydratie, vergeleken met het systeem waar het biomolecuul afwezig is. Bovendien stimuleren de biomoleculen de vorming van clusters die meer $CO_3^{2-}$ dan $Ca^{2+}$-ionen bevatten, waardoor clusters ontstaan die over het geheel genomen een negatieve lading hebben. De laatste belangrijke bevinding is dat de biomoleculen een effect hebben op de vorm van de clusters. De algemene conclusie van dit hoofdstuk is dat biomoleculen de manier veranderen waarop de ionen aggregeren in de vroege stadia voorafgaand aan enige (bio)kristallisatie.



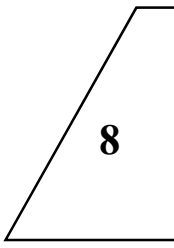
Hoofdstuk 5 is gewijd aan het herparametriseren en optimaliseren van de *forcefield*-interacties tussen het biomolecuul en het kation. Zoals besproken in de inleiding van dit proefschrift, is een gedegen beschrijving van het systeem cruciaal om betrouwbare resultaten te krijgen. In het werk gepresenteerd in hoofdstuk 4 is een algemeen aanvaarde beschrijving in ons *forcefield*

gebruikt die gestandaardiseerd is voor biomoleculaire systemen. Deze interacties zijn echter vaak niet expliciet geparametriseerd en schieten tekort in nauwkeurigheid. In dit onderzoek werden verschillende beschrijvingen van calcium in combinatie met asparaginezuur en glutaminezuur nauwkeuriger bestudeerd. Metadynamische MD simulaties werden gebruikt om de energieprofielen te reconstrueren en daarnaast werden de corresponderende constanten van de associatie van $Ca^{2+}$ met carboxylgroepen berekend. Naast het vervangen van de Lennard-Jones-calciumparameters door de parameters die worden gebruikt in het $CaCO_3$- *forcefield*, werd het $Ca^{2+}$ - carboxylaat gecorrigeerd door de $\sigma_{ij}$-waarde (dat de de minimale afstand tussen de twee atomen beschrijft) in de heteroatomaire Lennard-Jones-interactie met 0,5% te wijzigen om te zorgen dat de associatieconstante beter overeenkomt met experimentele waarden. Deze nieuwe beschrijving van het systeem verbeterde tegelijkertijd de structurele eigenschappen van het ionenpaar. De voornaamste conclusie van deze studie is dat de herparameterisering succesvol was en dat het *forcefield* zoals gebruikt in hoofdstuk 4 dienovereenkomstig moet worden bijgewerkt. De impact op de resultaten van hoofdstuk 4 wordt kort genoemd.

Ten slotte wordt het laatste werk in dit proefschrift besproken in hoofdstuk 6, waar de reactiviteit van $CaCO_3$ (zie ook (Koskamp et al., 2021)) en de impact van Mg onzuiverheden werd onderzocht. Een calcietoppervlak en de interactie met water werd bestudeerd door te kijken naar de stabiliteit van calcium gepositioneerd in het kristaloppervlak en geadsorbeerd in de vorm van een klein $CaCO_3$-eiland dat ofwel een groei-eiland kan vertegenwoordigen, of een rand van een grotere ets-put, waar door oplossing (etsen) een deel van het oppervlak is verwijderd. De impact van ingebouwd magnesium, een veel voorkomende onzuiverheid in calciumcarbonaat zouten, in de kristalstructuur werd onderzocht. Dit is opnieuw een MD onderzoek, nu gecombineerd met *Umbrella Sampling* om thermodynamische gegevens te extraheren via de constructie van de verschillende energieprofielen. Onze resultaten toonden aan dat magnesium inbouw het kristal vervormde en de totale oppervlakte-energie, lokale

energieprofielen, water dichtheid in het grensvlak, structuur en dynamiek veranderde wanneer de Mg concentratie hoog is (~30% van Ca vervangen door Mg). Bij lage concentraties Mg bleven de algehele kristaleigenschappen onveranderd, maar plaatselijk stabiliseerde $Ca^{2+}$ en verhoogde de vormingsenergie voor nieuwe ets putjes. De $Ca^{2+}$-ionen in een groei-eiland waren, zoals verwacht, thermodynamisch minder stabiel in vergelijking met een plat oppervlak. Bovendien destabiliseerde het eilandoppervlak met de aanwezigheid van Mg nog verder. Dit resulteerde in een verhoogde oppervlakte-energie en lagere extractie energieën. Lokaal resulteerde een complex samenspel van deze verschillende factoren in een verschil in lokale extractie energieën. Al met al impliceerde de resultaten van de MD-studie de niet-lineaire impact van Mg op een calciumion en geven daarmee een verklaring voor de niet-lineaire relatie tussen de Mg concentratie in calciet en de oplosbaarheid van het kristal.

Als algemene conclusie tonen de vier hoofdstukken de impact van verschillende componenten tijdens verschillende stadia van het (bio)mineralisatieproces. Vervolgonderzoek kan gedaan worden op hoofdstuk 4 en 5 waar er gekeken zou kunnen worden naar het bestaan van prenucleatieclusters en of er een neiging is tot vorming van protocalciet, protovateriet en/of proto-aragoniet. Bovendien zou het bestuderen van de impact van een combinatie van biomoleculen met opgelost Mg in de eerste stappen van de vorming van amorf calciumcarbonaat (ACC) verschillende inzichten kunnen geven in hoe kristalvorming gestuurd kan worden en hoe verkalkende organismen de kristallisatie beheersen (bijv. (Kellock et al., 2020; Meldrum and Cölfen, 2008)). Hoewel deze moleculaire mechanismen in ons dagelijks leven ver beneden onze realiteitszin liggen, zijn ze de motor van wat we zien. Bijgevolg zijn de inzichten uit dit proefschrift belangrijk om natuurlijke en laboratoriumobservaties te begrijpen en essentieel bij het gericht zoeken naar nieuwe toepassingen, zoals het op een duurzame manier controleren van kristalvorming.


8

# References

Addadi, L., Raz, S., Weiner, S., 2003. Taking advantage of disorder: Amorphous calcium carbonate and its roles in biomineralization. Adv. Mater. https://doi.org/10.1002/adma.200300381

Agrawal, P., Raoof, A., Iliev, O., Wolthers, M., 2020. Evolution of pore-shape and its impact on pore conductivity during CO2 injection in calcite: Single pore simulations and microfluidic experiments. Adv. Water Resour. 136, 103480. https://doi.org/10.1016/j.advwatres.2019.103480

Ahr, W.M., 2008. Geology of Carbonate Reservoirs, Geology of Carbonate Reservoirs: The Identification, Description, and Characterization of Hydrocarbon Reservoirs in Carbonate Rocks. Wiley, New Jersey, NJ. https://doi.org/10.1002/9780470370650

Akiyama, T., Nakamura, K., Ito, T., 2011. Atomic and electronic structures of surfaces. Phys. Rev. B - Condens. Matter Mater. Phys. 84, 085428. https://doi.org/10.1103/PhysRevB.84.085428

Al-Horani, F.A., Al-Moghrabi, S.M., de Beer, D., 2003. The mechanism of calcification and its relation to photosynthesis and respiration in the scleractinian coral Galaxea fascicularis. Mar. Biol. 142, 419–426. https://doi.org/10.1007/s00227-002-0981-8

Allen, M.P., Tildesley, D.J., 1987. Computer simulation of liquids. Oxford University Press, New York, NY, USA.

Allouche, D., Parello, J., Sanejouand, Y.-H., 1999. Ca 2+ /Mg 2+ exchange in parvalbumin and other EF-hand proteins. A theoretical study. J. Mol. Biol. 285, 857–873.

https://doi.org/10.1006/jmbi.1998.2329

Andersson, M.P., Dobberschütz, S., Sand, K.K., Tobler, D.J., De Yoreo, J.J., Stipp, S.L.S., 2016. A Microkinetic Model of Calcite Step Growth. Angew. Chemie Int. Ed. 55, 11086–11090. https://doi.org/10.1002/anie.201604357

Åqvist, J., 1990. Ion-water interaction potentials derived from free energy perturbation simulations. J. Phys. Chem. 94, 8021–8024. https://doi.org/10.1021/j100384a009

Arvidson, R., Davis, K., … M.C.-P. of the 11th, 2004, U., 2004. Etch pit morphology and magnesium inhibition of calcite dissolution. (Wanty, R.B., Seal, R.R., Eds.), Saratoga Springs, New York, Balkema.

Arvidson, R.S., Collier, M., Davis, K.J., Vinson, M.D., Amonette, J.E., Luttge, A., 2006. Magnesium inhibition of calcite dissolution kinetics. Geochim. Cosmochim. Acta 70, 583–594. https://doi.org/10.1016/j.gca.2005.10.005

Ashraf, M.S., Azahar, S.B., Yusof, N.Z., 2017. Soil Improvement Using MICP and Biopolymers: A Review, in: IOP Conference Series: Materials Science and Engineering. Institute of Physics Publishing, p. 012058. https://doi.org/10.1088/1757-899X/226/1/012058

Astilleros, J.M., Fernández-Díaz, L., Putnis, A., 2010. The role of magnesium in the growth of calcite: An AFM study. Chem. Geol. 271, 52–58. https://doi.org/10.1016/j.chemgeo.2009.12.011

Babu, C.S., Lim, C., 2006. Empirical force fields for biologically active divalent metal cations in water. J. Phys. Chem. A 110, 691–699. https://doi.org/10.1021/jp054177x

Badyal, Y.S., Saboungi, M.-L., Price, D.L., Shastri, S.D., Haeffner, D.R., Soper, A.K., 2000.

Electron distribution in water. J. Chem. Phys. 112, 9206. https://doi.org/10.1063/1.481541

BÁEZ, L.A., CLANCY, P., 1994. Computer Simulation of the Crystal Growth and Dissolution of Natural Gas Hydrates. Ann. N. Y. Acad. Sci. 715, 177–186. https://doi.org/10.1111/j.1749-6632.1994.tb38833.x

Becke, A.D., 1988. Density-functional exchange-energy approximation with correct asymptotic behavior. Phys. Rev. A 38, 3098–3100. https://doi.org/10.1103/PhysRevA.38.3098

Bednaršek, N., Tarling, G.A., Bakker, D.C.E., Fielding, S., Jones, E.M., Venables, H.J., Ward, P., Kuzirian, A., Lézé, B., Feely, R.A., Murphy, E.J., 2012. Extensive dissolution of live pteropods in the Southern Ocean. Nat. Geosci. 5, 881–885. https://doi.org/10.1038/ngeo1635

Beglov, D., Roux, B., 1994. Finite representation of an infinite bulk system: Solvent boundary potential for computer simulations. J. Chem. Phys. 100, 9050–9063. https://doi.org/10.1063/1.466711

Belcher, A.M., Wu, X.H., Christensen, R.J., Hansma, P.K., Stucky, G.D., Morse, D.E., 1996. Control of crystal phase switching and orientation by soluble mollusc-shell proteins. Nature 381, 56–58. https://doi.org/10.1038/381056a0

Berman, A., Hanson, J., Leiserowitz, L., Koetzle, T.F., Weiner, S., Addadi, L., 1993. Biological control of crystal texture: A widespread strategy for adapting crystal properties to function. Science (80-. ). 259, 776–779. https://doi.org/10.1126/science.259.5096.776

Bertie, J.E., Lan, Z., 1996. Infrared Intensities of Liquids XX: The Intensity of the OH Stretching Band of Liquid Water Revisited, and the Best Current Values of the Optical Constants of H 2 O(l) at 25°C between 15,000 and 1 cm −1. Appl. Spectrosc. 50, 1047–

1057. https://doi.org/10.1366/0003702963905385

Bibi, I., Arvidson, R., Fischer, C., Lüttge, A., 2018. Temporal Evolution of Calcite Surface Dissolution Kinetics. Minerals 8, 256. https://doi.org/10.3390/min8060256

Bischoff, J.L., Fitzpatrick, J.A., Rosenbauer, R.J., 1993. The solubility and stabilization of ikaite (CaCO3.6H2O) from 0° to 25°C: environmental and paleoclimatic implications for thinolite tufa. J. Geol. 101, 21–33. https://doi.org/10.1086/648194

Bischoff, W.D., Mackenzie, F.T., Bishop, F.C., 1987. Stabilities of synthetic magnesian calcites in aqueous solution: Comparison with biogenic materials. Geochim. Cosmochim. Acta 51, 1413–1423. https://doi.org/10.1016/0016-7037(87)90325-5

Blaquiere, C., Berthon, G., 1987. Speciation studies in relation to magnesium bioavailability. Formation of Mg(II) complexes with glutamate, aspartate, glycinate, lactate, pyroglutamate, pyridoxine and citrate, and appraisal of their potential significance towards magnesium gastrointestin. Inorganica Chim. Acta 135, 179–189. https://doi.org/10.1016/S0020-1693(00)81296-7

Blue, C.R., Giuffre, A., Mergelsberg, S., Han, N., De Yoreo, J.J., Dove, P.M., 2017. Chemical and physical controls on the transformation of amorphous calcium carbonate into crystalline CaCO3 polymorphs. Geochim. Cosmochim. Acta 196, 179–196. https://doi.org/10.1016/j.gca.2016.09.004

Born, M., Huang, K., 1954. Dynamical theory of crystal lattices, Oxford Uni. ed. Oxford.

Borukhin, S., Bloch, L., Radlauer, T., Hill, A.H., Fitch, A.N., Pokroy, B., 2012. Screening the Incorporation of Amino Acids into an Inorganic Crystalline Host: the Case of Calcite. Adv. Funct. Mater. 22, 4216–4224. https://doi.org/10.1002/adfm.201201079

Bots, P., Benning, L.G., Rodriguez-Blanco, J.-D.D., Roncal-Herrero, T., Shaw, S., 2012. Mechanistic insights into the crystallization of amorphous calcium carbonate (ACC). Cryst. Growth Des. 12, 3806–3814. https://doi.org/10.1021/cg300676b

Bracco, J.N., Grantham, M.C., Stack, A.G., 2012. Calcite Growth Rates As a Function of Aqueous Calcium-to-Carbonate Ratio, Saturation Index, and Inhibitor Concentration: Insight into the Mechanism of Reaction and Poisoning by Strontium. Cryst. Growth Des. 12, 3540–3548. https://doi.org/10.1021/cg300350k

Bracco, J.N., Stack, A.G., Steefel, C.I., 2013. Upscaling calcite growth rates from the mesoscale to the macroscale. Environ. Sci. Technol. 47, 7555–7562. https://doi.org/10.1021/es400687r

Brady, P. V., Krumhansl, J.L., Papenguth, H.W., 1996. Surface complexation clues to dolomite growth. Geochim. Cosmochim. Acta 60, 727–731. https://doi.org/10.1016/0016-7037(95)00436-X

Braissant, O., Cailleau, G., Dupraz, C., Verrecchia, E.P., 2003. Bacterially induced mineralization of calcium carbonate in terrestrial environments: The role of exopolysaccharides and amino acids. J. Sediment. Res. 73, 485–490. https://doi.org/10.1306/111302730485

Braybrook, A.L., Heywood, B.R., Jackson, R.A., Pitt, K., 2002. Parallel computational and experimental studies of the morphological modification of calcium carbonate by cobalt. J. Cryst. Growth 243, 336–344. https://doi.org/10.1016/S0022-0248(02)01439-2

Brownlee, C., Langer, G., Wheeler, G.L., 2021. Coccolithophore calcification: Changing paradigms in changing oceans R. Acta Biomater. 120, 4–11. https://doi.org/10.1016/j.actbio.2020.07.050

Bruneval, F., Donadio, D., Parrinello, M., 2007. Molecular dynamics study of the solvation of calcium carbonate in water. J. Phys. Chem. B 111, 12219–12227. https://doi.org/10.1021/jp0728306

Bruno, M., Bittarello, E., Massaro, F.R., Aquilano, D., 2018. The effect of impurities on the structure and energy of a crystal surface: Mg impurities in calcite as a case study. CrystEngComm 20, 4556–4564. https://doi.org/10.1039/c8ce00750k

Bruno, M., Massaro, F.R., Pastero, L., Costa, E., Rubbo, M., Prencipe, M., Aquilano, D., 2013. New estimates of the free energy of calcite/water interfaces for evaluating the equilibrium shape and nucleation mechanisms. Cryst. Growth Des. 13, 1170–1179. https://doi.org/10.1021/cg3015817

Bruno, M., Massaro, F.R., Prencipe, M., Aquilano, D., 2010. Surface reconstructions and relaxation effects in a centre-symmetrical crystal: The {00.1} form of calcite (CaCO3). CrystEngComm 12, 3626–3633. https://doi.org/10.1039/c002969f

Bulo, R.E., Donadio, D., Laio, A., Molnar, F., Rieger, J., Parrinello, M., 2007. “Site Binding” of Ca 2+ Ions to Polyacrylates in Water: A Molecular Dynamics Study of Coiling and Aggregation. https://doi.org/10.1021/ma062467l

Bultel, B., Viennet, J., Poulet, F., Carter, J., Werner, S.C., 2019. Detection of Carbonates in Martian Weathering Profiles. J. Geophys. Res. Planets 124, 989–1007. https://doi.org/10.1029/2018JE005845

Burn, S., Hoang, M., Zarzo, D., Olewniak, F., Campos, E., Bolto, B., Barron, O., 2015. Desalination techniques-A review of the opportunities for desalination in agriculture. https://doi.org/10.1016/j.desal.2015.01.041

Burnham, C.J., Xantheas, S.S., 2002. Development of transferable interaction models for

water. I. Prominent features of the water dimer potential energy surface. J. Chem. Phys. 116, 1479–1492. https://doi.org/10.1063/1.1423940

Busenberg, E., Niel Plummer, L., 1989. Thermodynamics of magnesian calcite solid-solutions at 25°C and 1 atm total pressure. Geochim. Cosmochim. Acta 53, 1189–1208. https://doi.org/10.1016/0016-7037(89)90056-2

Busenberg, E., Plummer, N., 1982. The kinetics of dissolution of dolomite in CO2-H2O systems at 1.5 to 65oC and 0 to 1 atm PCO2. Am. J. Sci. 282, 45–78.

Bush, T.S., Gale, J.D., Catlow, C.R.A., Battle, P.D., 1994. Self-consistent interatomic potentials for the simulation of binary and ternary oxides. J. Mater. Chem. 4, 831–837. https://doi.org/10.1039/jm9940400831

Caldeira, K., Wickett, M.E., 2003. Anthropogenic carbon and ocean pH. Nature 425, 365. https://doi.org/10.1038/425365a

Cao, B., Stack, A.G., Steefel, C.I., DePaolo, D.J., Lammers, L.N., Hu, Y., 2018. Investigating calcite growth rates using a quartz crystal microbalance with dissipation (QCM-D). Geochim. Cosmochim. Acta 222, 269–283. https://doi.org/10.1016/j.gca.2017.10.020

Carrasco, J., Hodgson, A., Michaelides, A., 2012. A molecular perspective of water at metal interfaces. Nat. Mater. 11, 667–674. https://doi.org/10.1038/nmat3354

Cartwright, J.H.E., Checa, A.G., Gale, J.D., Gebauer, D., Sainz-Díaz, C.I., 2012. Calcium carbonate polyamorphism and its role in biomineralization: How many amorphous calcium carbonates are there? Angew. Chemie - Int. Ed. https://doi.org/10.1002/anie.201203125

Catlow, C.R.A., Diller, K.M., Norgett, M.J., 1977. Interionic potentials for alkali halides. J.

Phys. C Solid State Phys. 10, 1395–1412. https://doi.org/10.1088/0022-3719/10/9/013

Chandra, A., 2000. Effects of Ion Atmosphere on Hydrogen-Bond Dynamics in Aqueous Electrolyte Solutions. Phys. Rev. Lett. 85, 768–771. https://doi.org/10.1103/PhysRevLett.85.768

Chang, I., Im, J., Cho, G.-C., 2016. Introduction of Microbial Biopolymers in Soil Treatment for Future Environmentally-Friendly and Sustainable Geotechnical Engineering. Sustainability 8, 251. https://doi.org/10.3390/su8030251

Cheng, L., Fenter, P., Nagy, K.L., Schlegel, M.L., Sturchio, N.C., 2001. Molecular-Scale Density Oscillations in Water Adjacent to a Mica Surface. Phys. Rev. Lett. 87, 156103. https://doi.org/10.1103/PhysRevLett.87.156103

Chialvo, A.A., Cummings, P.T., Cochran, H.D., Simonson, J.M., Mesmer, R.E., 1995. Na+-Cl- ion pair association in supercritical water. J. Chem. Phys. 103, 9379–9387. https://doi.org/10.1063/1.470707

Christoffersen, J., Christoffersen, M.R., 1990. Kinetics of spiral growth of calcite crystals and determination of the absolute rate constant. J. Cryst. Growth 100, 203–211. https://doi.org/10.1016/0022-0248(90)90623-S

Church, A.T., Hughes, Z.E., Walsh, T.R., 2015. Improving the description of interactions between Ca2+ and protein carboxylate groups, including γ-carboxyglutamic acid: revised CHARMM22∗ parameters. RSC Adv. 5, 67820–67828. https://doi.org/10.1039/c5ra11268k

Cicerone, D.S., Regazzoni, A.E., Blesa, M.A., 1992. Electrokinetic properties of the calcite/water interface in the presence of magnesium and organic matter. J. Colloid Interface Sci. 154, 423–433. https://doi.org/10.1016/0021-9797(92)90158-I

Clough, S.A., Beers, Y., Klein, G.P., Rothman, L.S., 1973. Dipole moment of water from Stark measurements of H 2 O, HDO, and D 2 O. J. Chem. Phys. 59, 2254–2259. https://doi.org/10.1063/1.1680328

Colowick, S.P., Womack, F.C., 1969. Binding of diffusible molecules by macromolecules: rapid measurement by rate of dialysis. J. Biol. Chem. 244, 774–777. https://doi.org/10.1016/s0021-9258(18)94419-4

Cooper, T.G., De Leeuw, N.H., 2002. Adsorption of methanoic acid onto the low-index surfaces of calcite and aragonite, in: Molecular Simulation. Taylor & Francis Group , pp. 539–556. https://doi.org/10.1080/08927020290030125

Costa, S.N., Freire, V.N., Caetano, E.W.S., Maia, F.F., Barboza, C.A., Fulco, U.L., Albuquerque, E.L., 2016. DFT Calculations with van der Waals Interactions of Hydrated Calcium Carbonate Crystals CaCO3·(H2O, 6H2O): Structural, Electronic, Optical, and Vibrational Properties. J. Phys. Chem. A 120, 5752–5765. https://doi.org/10.1021/acs.jpca.6b05436

Covington, A.K., Danish, E.Y., 2009. Measurement of Magnesium Stability Constants of Biologically Relevant Ligands by Simultaneous Use of pH and Ion-Selective Electrodes. J. Solution Chem. 38, 1449–1462. https://doi.org/10.1007/s10953-009-9459-3

Cygan, R.T., Kubicki, J.D. (Eds.), 2001. Molecular Modeling Theory, Molecular Modeling Theory. De Gruyter. https://doi.org/10.1515/9781501508721

D.A. Case, K. Belfon, I.Y. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D.C., T.A. Darden, R.E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz,R Harris, S. Izadi, S.A.I., K. Kasavajhala, A. Kovalenko, R. Krasny, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J.L., T. Luchko, R. Luo, V. Man, K.M. Merz, Y.

Miao, O. Mikhailovskii, G. Monard, H. Nguyen, A. Onufriev, F., Pan, S. Pantano, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, N.R., Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, L. Wilson, R.M. Wolf, X. Wu, Y. Xiong, Y.X., D.M. York and P.A. Kollman (2020), AMBER 2020, University of California, S.F., 2020. AMBER 2020. J. Chem. Inf. Model. 53, 1689–1699.

Dang, L.X., Smith, D.E., 1995. Comment on '"Mean force potential for the calcium–chloride ion pair in water"' [J. Chem. Phys. 99, 4229 (1993)]. J. Chem. Phys. 102, 3483–3484. https://doi.org/10.1063/1.468572

Davis, K.J., Dove, P.M., De Yoreo, J.J., 2000. The role of Mg2+ as an impurity in calcite growth. Science 290, 1134–7.

De La Pierre, M., Raiteri, P., Gale, J.D., 2016. Structure and Dynamics of Water at Step Edges on the Calcite {1014} Surface. Cryst. Growth Des. 16, 5907–5914. https://doi.org/10.1021/acs.cgd.6b00957

De La Pierre, M., Raiteri, P., Stack, A.G., Gale, J.D., 2017. Uncovering the Atomistic Mechanism for Calcite Step Growth. Angew. Chemie Int. Ed. 56, 8464–8467. https://doi.org/10.1002/anie.201701701

De Leeuw, N. H., 2002. Molecular dynamics simulations of the growth inhibiting effect of Fe2+, Mg2+, Cd2+, and Sr2+ on calcite crystal growth. J. Phys. Chem. B 106, 5241–5249. https://doi.org/10.1021/jp014488h

De Leeuw, Nora H., 2002. Surface structures, stabilities, and growth of magnesian calcites: A computational investigation from the perspective of dolomite formation. Am. Mineral. 87, 679–689. https://doi.org/10.2138/am-2002-5-610

De Leeuw, N.H., Cooper, T.G., 2004. A Computer Modeling Study of the Inhibiting Effect of

Organic Adsorbates on Calcite Crystal Growth. Cryst. Growth Des. 4, 123–133. https://doi.org/10.1021/cg0341003

de Leeuw, N.H., Parker, S.C., 2001. Surface–water interactions in the dolomite problem. Phys. Chem. Chem. Phys. 3, 3217–3221. https://doi.org/10.1039/b102928m

de Leeuw, N.H., Parker, S.C., 1998. Molecular-dynamics simulation of MgO surfaces in liquid water using a shell-model potential for water. Phys. Rev. B - Condens. Matter Mater. Phys. 58, 13901–13908. https://doi.org/10.1103/PhysRevB.58.13901

De Leeuw, N.H., Parker, S.C., 1998. Surface Structure and Morphology of Calcium Carbonate Polymorphs Calcite, Aragonite, and Vaterite: An Atomistic Approach. J. Phys. Chem. B 102, 2914–2922. https://doi.org/10.1021/jp973210f

De Leeuw, N.H., Parker, S.C., 1997. Atomistic simulation of the effect of molecular adsorption of water on the surface structure and energies of calcite surfaces. J. Chem. Soc. - Faraday Trans. 93, 467–475. https://doi.org/10.1039/a606573b

de Leeuw, N.H., Parker, S.C., Harding, J.H., 1999. Molecular dynamics simulation of crystal dissolution from calcite steps. Phys. Rev. B 60, 13792–13799. https://doi.org/10.1103/PhysRevB.60.13792

De Nooijer, L.J., Spero, H.J., Erez, J., Bijma, J., Reichart, G.J., 2014. Biomineralization in perforate foraminifera. Earth-Science Rev. https://doi.org/10.1016/j.earscirev.2014.03.013

De Yoreo, J.J., Vekilov, P.G., 2003. Principles of Crystal Nucleation and Growth. Rev. Mineral. Geochemistry 54, 57–93. https://doi.org/10.2113/0540057

Demichelis, R., Raiteri, P., Gale, J.D., Dovesi, R., 2013. The multiple structures of vaterite.

Cryst. Growth Des. 13, 2247–2251. https://doi.org/10.1021/cg4002972

Demichelis, R., Raiteri, P., Gale, J.D., Quigley, D., Gebauer, D., 2011. Stable prenucleation mineral clusters are liquid-like ionic polymers. Nat. Commun. 2, 590. https://doi.org/10.1038/ncomms1604

Deng, H., Shen, X.-C., Wang, X.-M., Du, C., n.d. Calcium carbonate crystallization controlled by functional groups: A mini-review. https://doi.org/10.1007/s11706-013-0191-y

Dewan, S., Yeganeh, M.S., Borguet, E., 2013. Experimental correlation between interfacial water structure and mineral reactivity. J. Phys. Chem. Lett. 4, 1977–1982. https://doi.org/10.1021/jz4007417

Di Tommaso, D., de Leeuw, N., 2010. Modelling the nucleation of metal carbonates: The importance of the hydration shell in the monomer formation. Geochim. Cosmochim. Acta 74, 236–236.

Di Tommaso, D., De Leeuw, N.H., 2010. First principles simulations of the structural and dynamical properties of hydrated metal ions Me2+ and solvated metal carbonates (Me = Ca, Mg, and Sr). Cryst. Growth Des. 10, 4292–4302. https://doi.org/10.1021/cg100055p

Di Tommaso, D., Prakash, M., Lemaire, T., Lewerenz, M., de Leeuw, N.H., Naili, S., 2017. Molecular Dynamics Simulations of Hydroxyapatite Nanopores in Contact with Electrolyte Solutions: The Effect of Nanoconfinement and Solvated Ions on the Surface Reactivity and the Structural, Dynamical, and Vibrational Properties of Water. Crystals 7, 57. https://doi.org/10.3390/cryst7020057

Di Tommaso, D., Ruiz-Agudo, E., De Leeuw, N.H., Putnis, A., Putnis, C. V., 2014. Modelling the effects of salt solutions on the hydration of calcium ions. Phys. Chem. Chem. Phys. 16, 7772–7785. https://doi.org/10.1039/C3CP54923B

Dick B G, J., Overhauser, A.W., 1958. Theory of the dielectric constants of alkali halide crystals. Phys. Rev. 112, 90.

Dickens, B., Brown, W.E., 1970. The Crystal Structure of Calcium Carbonate Hexahydrate at ~ –120°. Inorg. Chem. 9, 480–486. https://doi.org/10.1021/ic50085a010

Du, Z., De Leeuw, N.H., 2006. Molecular dynamics simulations of hydration, dissolution and nucleation processes at the α-quartz (0001) surface in liquid water. J. Chem. Soc. Dalt. Trans. 2623–2634. https://doi.org/10.1039/b516258k

Effenberger, H., 1981. Kristallstruktur und Infrarot-Absorptionsspektrum von synthetischem Monohydrocalcit, CaCO3·H2O. Monatshefte für Chemie 112, 899–909. https://doi.org/10.1007/BF00905061

Eigen, M., Tamm, K., 1962. Schallabsorption in Elektrolytösungen als Folge chemischer Relaxation. I. Relaxationstheorie der mehrstufigen Dissoziation. Zeitschrift für Elektrochemie 66, 93–121. https://doi.org/10.1002/bbpc.19620660205

Eigen, M., Wilkens, R.G., 1965. The Kinetics and Mechanism of Formation of Metal Complexes. pp. 55–80. https://doi.org/10.1021/ba-1965-0049.ch003

Eiler, J.M., Bergquist, B., Bourg, I., Cartigny, P., Farquhar, J., Gagnon, A., Guo, W., Halevy, I., Hofmann, A., Larson, T.E., Levin, N., Schauble, E.A., Stolper, D., 2014. Frontiers of stable isotope geoscience. Chem. Geol. 372, 119–143. https://doi.org/10.1016/j.chemgeo.2014.02.006

Elstnerová, P., Friák, M., Fabritius, H.O., Lymperakis, L., Hickel, T., Petrov, M., Nikolov, S., Raabe, D., Ziegler, A., Hild, S., Neugebauer, J., 2010. Ab initio study of thermodynamic, structural, and elastic properties of Mg-substituted crystalline calcite. Acta Biomater. 6, 4506–4512.

Endo, K., Kogure, T., Nagasawa, H. (Eds.), 2018. Biomineralization, Biomineralization. Springer Singapore, Singapore. https://doi.org/10.1007/978-981-13-1002-7

Ernzerhof, M., Scuseria, G.E., 1999. Assessment of the Perdew–Burke–Ernzerhof exchange-correlation functional. J. Chem. Phys. 110, 5029–5036. https://doi.org/10.1063/1.478401

Ewald, P.P., 1921. Die Berechnung optischer und elektrostatischer Gitterpotentiale. Ann. Phys. 369, 253–287. https://doi.org/10.1002/andp.19213690304

Falini, G., Fermani, S., Goffredo, S., 2015. Coral biomineralization: A focus on intra-skeletal organic matrix and calcification. Academic Press. https://doi.org/10.1016/j.semcdb.2015.09.005

Feng, X., Redfern, S.A.T., 2018. Iodate in calcite, aragonite and vaterite CaCO3: Insights from first-principles calculations and implications for the I/Ca geochemical proxy. Geochim. Cosmochim. Acta 236, 351–360. https://doi.org/10.1016/j.gca.2018.02.017

Fenter, P., Kerisit, S., Raiteri, P., Gale, J.D., 2013. Is the Calcite–Water Interface Understood? Direct Comparisons of Molecular Dynamics Simulations with Specular X-ray Reflectivity Data. J. Phys. Chem. C 117, 5028–5042. https://doi.org/10.1021/jp310943s

Fenter, P., Sturchio, N.C., 2012. Calcite (104)-water interface structure, revisited. Geochim. Cosmochim. Acta 97, 58–69. https://doi.org/10.1016/j.gca.2012.08.021

Ferreira, S.O., 2013. Advanced Topics on Crystal Growth, Advanced Topics on Crystal Growth. InTech, Rijeka, Croatia. https://doi.org/10.5772/46151

Finney, A.R., Rodger, P.M., 2012. Probing the structure and stability of calcium carbonate pre-nucleation clusters. Faraday Discuss. 159, 47–60. https://doi.org/10.1039/c2fd20054f

Forbes, R.T., York, P., Fawcett, V., Shields, L., 1992. Physicochemical Properties of Salts of

p-Aminosalicylic Acid. I. Correlation of Crystal Structure and Hydrate Stability. Pharm. Res. An Off. J. Am. Assoc. Pharm. Sci. 9, 1428–1435. https://doi.org/10.1023/A:1015806812349

Freeman, C.L., Harding, J.H., Quigley, D., Rodger, P.M., 2015. How does an amorphous surface influence molecular binding? – ovocleidin-17 and amorphous calcium carbonate. Phys. Chem. Chem. Phys. 17, 17494–17500. https://doi.org/10.1039/C5CP00434A

Freeman, C.L., Harding, J.H., Quigley, D., Rodger, P.M., 2011. Simulations of ovocleidin-17 binding to calcite surfaces and its implications for eggshell formation. J. Phys. Chem. C 115, 8175–8183. https://doi.org/10.1021/jp200145m

Freeman, C.L., Harding, J.H., Quigley, D., Rodger, P.M., 2010. Structural control of crystal nuclei by an eggshell protein. Angew. Chemie - Int. Ed. 49, 5135–5137. https://doi.org/10.1002/anie.201000679

Gal, A., Weiner, S., Addadi, L., 2015. A perspective on underlying crystal growth mechanisms in biomineralization: Solution mediated growth versus nanosphere particle accretion. CrystEngComm. https://doi.org/10.1039/c4ce01474j

Gao, Z.Y., Sun, W., Hu, Y.H., Liu, X.W., 2012. Anisotropic surface broken bond properties and wettability of calcite and fluorite crystals. Trans. Nonferrous Met. Soc. China (English Ed. 22, 1203–1208. https://doi.org/10.1016/S1003-6326(11)61306-X

Gebauer, D., Gunawidjaja, P.N., Ko, J.Y.P., Bacsik, Z., Aziz, B., Liu, L., Hu, Y., Bergström, L., Tai, C.-W., Sham, T.-K., Edén, M., Hedin, N., 2010. Proto-Calcite and Proto-Vaterite in Amorphous Calcium Carbonates. Angew. Chemie Int. Ed. 49, 8889–8891. https://doi.org/10.1002/anie.201003220

Gebauer, D., Kellermeier, M., Gale, J.D., Bergströ, L., Cö Lfen, H., 2014. Pre-nucleation

clusters as solute precursors in crystallisation. 2348 | Chem. Soc. Rev 43, 2348. https://doi.org/10.1039/c3cs60451a

Gebauer, D., Raiteri, P., Gale, J.D., Cölfen, H., 2018. On classical and non-classical views on nucleation. Am. J. Sci. 318, 969–988. https://doi.org/10.2475/09.2018.05

Gebauer, D., Völkel, A., Cölfen, H., 2008. Stable prenucleation calcium carbonate clusters. Science (80-. ). 322, 1819–1822. https://doi.org/10.1126/science.1164271

Genovese, D., Montalti, M., Otálora, F., Gómez-Morales, J., Sancho-Tomás, M., Falini, G., García-Ruiz, J.M., 2016. Role of CaCO3° neutral pair in calcium carbonate crystallization. Cryst. Growth Des. 16, 4173–4177. https://doi.org/10.1021/acs.cgd.6b00276

Goedecker, S., Teter, M., Hutter, J., 1996. Separable dual-space Gaussian pseudopotentials. Phys. Rev. B 54, 1703–1710. https://doi.org/10.1103/PhysRevB.54.1703

Golfier, F., Zarcone, C., Bazin, B., Lenormand, R., Lasseux, D., Quintard, M., 2002. On the ability of a Darcy-scale model to capture wormhole formation during the dissolution of a porous medium. J. Fluid Mech. 457, 213–254. https://doi.org/10.1017/S0022112002007735

Gong, Y.U.T., Killian, C.E., Olson, I.C., Appathurai, N.P., Amasino, A.L., Martin, M.C., Holt, L.J., Wilt, F.H., Gilbert, P.U.P.A., Performed, P.U.P.A.G., 2012. Phase transitions in biogenic amorphous calcium carbonate 109. https://doi.org/10.1073/pnas.1118085109/-/DCSupplemental

Gower, L.B., 2008. Biomimetic model systems for investigating the amorphous precursor pathway and its role in biomineralization. Chem. Rev. 108, 4551–4627. https://doi.org/10.1021/cr800443h

Gower, L.B., Odom, D.J., 2000. Deposition of calcium carbonate films by a polymer-induced liquid-precursor (PILP) process. J. Cryst. Growth 210, 719–734. https://doi.org/10.1016/S0022-0248(99)00749-6

Gozalpour, F., Ren, S.R., Tohidi, B., 2005. CO2 Eor and Storage in Oil Reservoir. Oil Gas Sci. Technol. 60, 537–546. https://doi.org/10.2516/ogst:2005036

Graf, D.L., 1961. Crystallographic tables for the rhombohedral carbonates. Am. Mineral. 46, 1283–1316.

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. J. Chem. Phys. 132, 154104. https://doi.org/10.1063/1.3382344

Grossfield, A., n.d. "WHAM: the weighted histogram analysis method" version 2.0.9 [WWW Document]. version 2.0.9. URL http://membrane.urmc.rochester.edu/wordpress/?page_id=126 (accessed 6.17.20).

Guru, P.S., Dash, S., 2014. Sorption on eggshell waste—A review on ultrastructure, biomineralization and other applications. Adv. Colloid Interface Sci. 209, 49–67. https://doi.org/10.1016/j.cis.2013.12.013

Haiech, J., Klee, C.B., Demaille, J.G., 1981. Effects of Cations on Affinity of Calmodulin for Calcium: Ordered Binding of Calcium Ions Allows the Specific Activation of Calmodulin-Stimulated Enzymes. Biochemistry 20, 3890–3897. https://doi.org/10.1021/bi00516a035

Hancer, M., Celik, M.S., Miller, J.D., 2001. The significance of interfacial water structure in soluble salt flotation systems. J. Colloid Interface Sci. 235, 150–161. https://doi.org/10.1006/jcis.2000.7350

Harding, J.H., 1990. Computer simulation of defects in ionic solids. Reports Prog. Phys. https://doi.org/10.1088/0034-4885/53/11/002

Harrison, A.L., Mavromatis, V., Oelkers, E.H., Bénézeth, P., 2019. Solubility of the hydrated Mg-carbonates nesquehonite and dypingite from 5 to 35 °C: Implications for CO2 storage and the relative stability of Mg-carbonates. Chem. Geol. 504, 123–135. https://doi.org/10.1016/j.chemgeo.2018.11.003

Harstad, A.O., Stipp, S.L.S., 2007. Calcite dissolution: Effects of trace cations naturally present in Iceland spar calcites. Geochim. Cosmochim. Acta 71, 56–70. https://doi.org/10.1016/j.gca.2006.07.037

Hassani, L.N., Hindré, F., Beuvier, T., Calvignac, B., Lautram, N., Gibaud, A., Boury, F., 2013. Lysozyme encapsulation into nanostructured CaCO3 microparticles using a supercritical CO2 process and comparison with the normal route. J. Mater. Chem. B 1, 4011–4019. https://doi.org/10.1039/c3tb20467g

Hawez, H., Ahmed, Z., 2014. Enhanced oil recovery by CO2 injection in carbonate reservoirs. WIT Trans. Ecol. Environ. 186, 547–558. https://doi.org/10.2495/ESUS140481

Heberling, F., Trainor, T.P., Lützenkirchen, J., Eng, P., Denecke, M.A., Bosbach, D., 2011. Structure and reactivity of the calcite–water interface. J. Colloid Interface Sci. 354, 843–857. https://doi.org/10.1016/j.jcis.2010.10.047

Hellevang, H., Miri, R., Haile, B.G., 2014. New insights into the mechanisms controlling the rate of crystal growth. Cryst. Growth Des. 14, 6451–6458. https://doi.org/10.1021/cg501294w

Helm, L., Merbach, A.E., 1999. Water exchange on metal ions: Experiments and simulations. Coord. Chem. Rev. 187, 151–181. https://doi.org/10.1016/S0010-8545(99)90232-1

Henriques, J., Cragnell, C., Skepö, M., 2015. Molecular Dynamics Simulations of Intrinsically Disordered Proteins: Force Field Evaluation and Comparison with Experiment. J. Chem. Theory Comput. 11, 3420–3431. https://doi.org/10.1021/ct501178z

Hermans, J., Berendsen, H.J.C., Van Gunsteren, W.F., Postma, J.P.M., 1984. A consistent empirical potential for water-protein interactions. Biopolymers 23, 1513–1518. https://doi.org/10.1002/bip.360230807

Hesse, K.F., Suess, E., 1983. Refinement of the structure of Ikaite, CaC03 · 6 H20. Zeitschrift fur Krist. - New Cryst. Struct. 163, 227–231. https://doi.org/10.1524/zkri.1983.163.3-4.227

Hewish, N.A., Enderby, J.E., Howells, W.S., 1983. The dynamics of water molecules in ionic solution. J. Phys. C Solid State Phys. 16, 1777–1791. https://doi.org/10.1088/0022-3719/16/10/009

Hockney, R.W., Eastwood, J.W., 1988. Computer Simulation Using Particles, Computer Simulation Using Particles. IOP Publishing Ltd. https://doi.org/10.1887/0852743920

Hofer, T.S., Tran, H.T., Schwenk, C.F., Rode, B.M., 2004. Characterization of Dynamics and Reactivities of Solvated Ions by Ab Initio Simulations. J. Comput. Chem. 25, 211–217. https://doi.org/10.1002/jcc.10374

Hofmann, A.E., Bourg, I.C., DePaolo, D.J., 2012. Ion desolvation as a mechanism for kinetic isotope fractionation in aqueous systems. Proc. Natl. Acad. Sci. U. S. A. 109, 18689–18694. https://doi.org/10.1073/pnas.1208184109

Holcomb, M., Venn, A.A., Tambutté, E., Tambutté, S., Allemand, D., Trotter, J., Mcculloch, & M., n.d. Coral calcifying fluid pH dictates response to ocean acidification. https://doi.org/10.1038/srep05207

Holtus, T., Helmbrecht, L., Hendrikse, H.C., Baglai, I., Meuret, S., Adhyaksa, G.W.P., Garnett, E.C., Noorduin, W.L., 2018. Shape-preserving transformation of carbonate minerals into lead halide perovskite semiconductors based on ion exchange/insertion reactions. Nat. Chem. 10, 740–745. https://doi.org/10.1038/s41557-018-0064-1

Hong, M., Teng, H.H., 2014. Implications of solution chemistry effects: Direction-specific restraints on the step kinetics of calcite growth. Geochim. Cosmochim. Acta 141, 228–239. https://doi.org/10.1016/J.GCA.2014.06.023

Hoover, W.G., 1985. Canonical dynamics: Equilibrium phase-space distributions. Phys. Rev. A 31, 1695–1697. https://doi.org/10.1103/PhysRevA.31.1695

Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., Simmerling, C., 2006. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins Struct. Funct. Bioinforma. 65, 712–725. https://doi.org/10.1002/prot.21123

Hu, Y.-B., Marië Tte Wolthers, †, Wolf-Gladrow, D.A., Nehrke, G., 2015. Effect of pH and Phosphate on Calcium Carbonate Polymorphs Precipitated at near-Freezing Temperature. https://doi.org/10.1021/cg500829p

Hwang, S., Blanco, M., Goddard, W.A., 2001. Atomistic simulations of corrosion inhibitors adsorbed on calcite surfaces I. Force field parameters for calcite. J. Phys. Chem. B 105, 10746–10752. https://doi.org/10.1021/jp010567h

Impey, R.W., Madden, P.A., McDonald, I.R., 1983. Hydration and Mobility of Ions in Solution. J. Phys. Chem. 87, 5071–5083. https://doi.org/10.1021/j150643a008

Innocenti Malini, R., Finney, A.R., Hall, S.A., Freeman, C.L., Harding, J.H., 2017. The Water–Amorphous Calcium Carbonate Interface and Its Interactions with Amino Acids. Cryst.

Growth Des. 17, 5811–5822. https://doi.org/10.1021/acs.cgd.7b00874

Iyemperumal, S.K., Deskins, N.A., 2017. Evaluating Solvent Effects at the Aqueous/Pt(111) Interface. ChemPhysChem 18, 2171–2190. https://doi.org/10.1002/cphc.201700162

Jacob, D.E., Wirth, R., Soldati, A.L., Wehrmeister, U., Schreiber, A., 2011. Amorphous calcium carbonate in the shells of adult Unionoida. J. Struct. Biol. 173, 241–249. https://doi.org/10.1016/j.jsb.2010.09.011

Jiang, W., Pacella, M.S., Athanasiadou, Di., Nelea, V., Vali, H., Hazen, R.M., Gray, J.J., McKee, M.D., 2017. Chiral acidic amino acids induce chiral hierarchical structure in calcium carbonate. Nat. Commun. 8, 15066. https://doi.org/10.1038/ncomms15066

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L., 1983. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79, 926–935. https://doi.org/10.1063/1.445869

Jorgensen, W.L., Jenson, C., 1998. Temperature dependence of TIP3P, SPC, and TIP4P water from NPT Monte Carlo simulations: Seeking temperatures of maximum density. J. Comput. Chem. 19, 1179–1186. https://doi.org/10.1002/(SICI)1096-987X(19980730)19:10<1179::AID-JCC6>3.0.CO;2-J

Joswiak, M.N., Doherty, M.F., Peters, B., 2018. Ion dissolution mechanism and kinetics at kink sites on NaCl surfaces. Proc. Natl. Acad. Sci. 115, 656–661. https://doi.org/10.1073/pnas.1713452115

Joung, I.S., Cheatham, T.E., 2009. Molecular Dynamics Simulations of the Dynamic and Energetic Properties of Alkali and Halide Ions Using Water-Model-Specific Ion Parameters. J. Phys. Chem. B 113, 13279–13290. https://doi.org/10.1021/jp902584c

Kahlen, J., Salimi, L., Sulpizi, M., Peter, C., Donadio, D., 2014. Interaction of Charged Amino-Acid Side Chains with Ions: An Optimization Strategy for Classical Force Fields. J. Phys. Chem. B 118, 3960–3972. https://doi.org/10.1021/jp412490c

Kaplan, C.N., Noorduin, W.L., Li, L., Sadza, R., Folkertsma, L., Aizenberg, J., Mahadevan, L., 2017. Controlled growth and form of precipitating microsculptures. Science (80-. ). 355, 1395–1399. https://doi.org/10.1126/science.aah6350

Kellock, C., Cole, C., Penkman, K., Evans, D., Kröger, R., Hintz, C., Hintz, K., Finch, A., Allison, N., 2020. The role of aspartic acid in reducing coral calcification under ocean acidification conditions. Sci. Rep. 10, 12797. https://doi.org/10.1038/s41598-020-69556-0

Kerisit, S., Parker, S.C., 2004. Free energy of adsorption of water and metal ions on the {10$\bar{1}$4} calcite surface. J. Am. Chem. Soc. 126, 10152–10161. https://doi.org/10.1021/ja0487776

Kerisit, S., Parker, S.C., Harding, J.H., 2003. Atomistic simulation of the dissociative adsorption of water on calcite surfaces. J. Phys. Chem. B 107, 7676–7682. https://doi.org/10.1021/jp034201b

Kim, Y.Y., Carloni, J.D., Demarchi, B., Sparks, D., Reid, D.G., Kunitake, M.E., Tang, C.C., Duer, M.J., Freeman, C.L., Pokroy, B., Penkman, K., Harding, J.H., Estroff, L.A., Baker, S.P., Meldrum, F.C., 2016. Tuning hardness in calcite by incorporation of amino acids. Nat. Mater. 15, 903–910. https://doi.org/10.1038/nmat4631

Kirch, A., Mutisya, S.M., Sánchez, V.M., De Almeida, J.M., Miranda, C.R., 2018. Fresh Molecular Look at Calcite-Brine Nanoconfined Interfaces. J. Phys. Chem. C 122, 6117–6127. https://doi.org/10.1021/acs.jpcc.7b12582

Klauda, J.B., Venable, R.M., Freites, J.A., O'Connor, J.W., Tobias, D.J., Mondragon-Ramirez,

C., Vorobyov, I., MacKerell, A.D., Pastor, R.W., 2010. Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. J. Phys. Chem. B 114, 7830–7843. https://doi.org/10.1021/jp101759q

Kohagen, M., Mason, P.E., Jungwirth, P., 2014. Accurate description of calcium solvation in concentrated aqueous solutions. J. Phys. Chem. B 118, 7902–7909. https://doi.org/10.1021/jp5005693

Koishi, A., Fernandez-Martinez, A., Ruta, B., Jimenez-Ruiz, M., Poloni, R., di Tommaso, D., Zontone, F., Waychunas, G.A., Montes-Hernandez, G., 2018. Role of Impurities in the Kinetic Persistence of Amorphous Calcium Carbonate: A Nanoscopic Dynamics View. J. Phys. Chem. C 122, 16983–16991. https://doi.org/10.1021/acs.jpcc.8b05189

Koleini, M.M., Mehraban, M.F., Ayatollahi, S., 2018. Effects of low salinity water on calcite/brine interface: A molecular dynamics simulation study. Colloids Surfaces A Physicochem. Eng. Asp. 537, 61–68. https://doi.org/10.1016/j.colsurfa.2017.10.024

Kondoh, A., Oi, T., 1998. Interaction of Alkaline Earth Metal Ions with Carboxylic Acids in Aqueous Solutions studied by 13c NMR Spectroscopy. Zeitschrift fur Naturforsch. - Sect. A J. Phys. Sci. 53, 77–91. https://doi.org/10.1515/zna-1998-1-212

König, G., Hudson, P.S., Boresch, S., Woodcock, H.L., 2014. Multiscale free energy simulations: An efficient method for connecting classical MD simulations to QM or QM/MM free energies using non-Boltzmann Bennett reweighting schemes. J. Chem. Theory Comput. 10, 1406–1419. https://doi.org/10.1021/ct401118k

Koskamp, J.A., Ruiz-Hernandez, S.E., Di Tommaso, D., Elena, A.M., De Leeuw, N.H., Wolthers, M., 2019. Reconsidering Calcium Dehydration as the Rate-Determining Step in Calcium Mineral Growth. J. Phys. Chem. C 123, 26895–26903.

https://doi.org/10.1021/acs.jpcc.9b06403

Koskamp, J.A., Ruiz Hernandez, S.E., De Leeuw, N.H., Wolthers, M., 2021. First steps towards understanding the non-linear impact of mg on calcite solubility: A molecular dynamics study. Minerals 11, 407. https://doi.org/10.3390/min11040407

Kowacz, M., Putnis, C.V., Putnis, A., 2007. The effect of cation:anion ratio in solution on the mechanism of barite growth at constant supersaturation: Role of the desolvation process on the growth kinetics. Geochim. Cosmochim. Acta 71, 5168–5179. https://doi.org/10.1016/J.GCA.2007.09.008

Koytsoumpa, E.I., Bergins, C., Kakaras, E., 2018. The CO 2 economy: Review of CO 2 capture and reuse technologies. J. Supercrit. Fluids 132, 3–16. https://doi.org/10.1016/j.supflu.2017.07.029

Krajewska, B., 2018. Urease-aided calcium carbonate mineralization for engineering applications: A review. J. Adv. Res. https://doi.org/10.1016/j.jare.2017.10.009

Krynicki, K., Green, C.D., Sawyer, D.W., 1978. Pressure and temperature dependence of self-diffusion in water. Faraday Discuss. Chem. Soc. 66, 199. https://doi.org/10.1039/dc9786600199

Kumar, P.P., Kalinichev, A.G., Kirkpatrick, R.J., 2009. Hydrogen-Bonding Structure and Dynamics of Aqueous Carbonate Species from Car−Parrinello Molecular Dynamics Simulations. J. Phys. Chem. B 113, 794–802. https://doi.org/10.1021/jp809069g

Kurganskaya, I., Luttge, A., 2016. Kinetic Monte Carlo Approach To Study Carbonate Dissolution. J. Phys. Chem. C 120, 6482–6492. https://doi.org/10.1021/acs.jpcc.5b10995

Kurganskaya, I., Rohlfs, R.D., 2020. Atomistic to meso-scale modeling of mineral dissolution:

Methods, challenges and prospects. Am. J. Sci. 320, 1–26. https://doi.org/10.2475/01.2020.02

Kvamme, B., Kuznetsova, T., Uppstad, D., 2009. Modelling excess surface energy in dry and wetted calcite systems. J. Math. Chem. 46, 756–762. https://doi.org/10.1007/s10910-009-9548-y

Lamoureux, G., Harder, E., Vorobyov, I. V., Roux, B., MacKerell, A.D., 2006. A polarizable model of water for molecular dynamics simulations of biomolecules. Chem. Phys. Lett. 418, 245–249. https://doi.org/10.1016/j.cplett.2005.10.135

Lamoureux, G., MacKerell, A.D., Roux, B., 2003. A simple polarizable model of water based on classical Drude oscillators. J. Chem. Phys. 119, 5185–5197. https://doi.org/10.1063/1.1598191

Lane, J.R., 2013. CCSDTQ optimized geometry of water dimer. J. Chem. Theory Comput. 9, 316–323. https://doi.org/10.1021/ct300832f

Langmuir, D., 1997. Aqueous environmental geochemistry. Prentice Hall, Upper Saddle River, N.J. :

Lardge, J.S., Duffy, D.M., Gillan, M.J., 2009. Investigation of the interaction of water with the calcite (10.4) surface using ab initio simulation. J. Phys. Chem. C 113, 7207–7212. https://doi.org/10.1021/jp806109y

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. J. Phys. Chem. C 114, 2664–2668. https://doi.org/10.1021/jp909593p

Larentzos, J.P., Criscenti, L.J., 2008. A molecular dynamics study of alkaline earth metal-

chloride complexation in aqueous solution. J. Phys. Chem. B 112, 14243–14250. https://doi.org/10.1021/jp802771w

Larsen, K., Bechgaard, K., Stipp, S.L.S., 2010. Modelling spiral growth at dislocations and determination of critical step lengths from pyramid geometries on calcite {1014} surfaces. Geochim. Cosmochim. Acta 74, 558–567. https://doi.org/10.1016/j.gca.2009.10.002

Lasaga, A.C., Lüttge, A., 2004. Mineralogical approaches to fundamental crystal dissolution kinetics. Am. Mineral. 89, 527–540. https://doi.org/10.2138/am-2004-0407

Lasaga, A.C., Lüttge, A., 2003. A model for crystal dissolution. Eur. J. Mineral. 15, 603–615. https://doi.org/10.1127/0935-1221/2003/0015-0603

Lee, C., Yang, W., Parr, R.G., 1988. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. Phys. Rev. B 37, 785–789. https://doi.org/10.1103/PhysRevB.37.785

Lee, S., Park, H.I., Sang, Q.X.A., 2007. Calcium regulates tertiary structure and enzymatic activity of human endometase/matrilysin-2 and its role in promoting human breast cancer cell invasion. Biochem. J. 403, 31–42. https://doi.org/10.1042/BJ20061390

Lee, S.H., Rasaiah, J.C., 1996. Molecular Dynamics Simulation of Ion Mobility. 2. Alkali Metal and Halide Ions Using the SPC/E Model for Water at 25 °C †. J. Phys. Chem. 100, 1420–1425. https://doi.org/10.1021/jp953050c

Lemasson, A.J., Fletcher, S., Hall-Spencer, J.M., Knights, A.M., 2017. Linking the biological impacts of ocean acidification on oysters to changes in ecosystem services: A review. https://doi.org/10.1016/j.jembe.2017.01.019

Lenders, J.J.M., Dey, A., Bomans, P.H.H., Spielmann, J., Hendrix, M.M.R.M., De With, G.,

Meldrum, F.C., Harder, S., Sommerdijk, N.A.J.M., 2012. High-magnesian calcite mesocrystals: A coordination chemistry approach. J. Am. Chem. Soc. 134, 1367–1373. https://doi.org/10.1021/ja210791p

Lewis, G. V., A Catlow, C.R., 1985. Potential models for ionic oxides. J. Phys. C Solid State Phys. 18, 1149–1161. https://doi.org/10.1088/0022-3719/18/6/010

Li, D.-W., Brüschweiler, R., 2010. NMR-Based Protein Potentials. Angew. Chemie Int. Ed. 49, 6778–6780. https://doi.org/10.1002/anie.201001898

Li, G., Guo, S., Zhang, J., Liu, Y., 2014. Inhibition of scale buildup during produced-water reuse: Optimization of inhibitors and application in the field. Desalination 351, 213–219. https://doi.org/10.1016/j.desal.2014.08.003

Li, L., Brodholt, J., Alfè, D., 2009. Structure and elasticity of hydrous ringwoodite: A first principle investigation. Phys. Earth Planet. Inter. 177, 103–115. https://doi.org/10.1016/j.pepi.2009.07.007

Li, L., Fijneman, A.J., Kaandorp, J.A., Aizenberg, J., Noorduin, W.L., 2018. Directed nucleation and growth by balancing local supersaturation and substrate/nucleus lattice mismatch. Proc. Natl. Acad. Sci. U. S. A. 115, 3575–3580. https://doi.org/10.1073/pnas.1712911115

Li, M., Zhu, X., Mukherjee, A., Huang, M., Achal, V., 2017. Biomineralization in metakaolin modified cement mortar to improve its strength with lowered cement content. J. Hazard. Mater. 329, 178–184. https://doi.org/10.1016/j.jhazmat.2017.01.035

Liang, Y., Lea, A.S., Baer, D.R., Engelhard, M.H., 1996. Structure of the cleaved CaCo3 (10$\bar{1}$4) surface in an aqueous environment. Surf. Sci. 351, 172–182. https://doi.org/10.1016/0039-6028(95)01298-2

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. J. Chem. Theory Comput. 8, 3902–3910. https://doi.org/10.1021/ct3001848

Liteanu, E., Niemeijer, A., Spiers, C.J., Peach, C.J., De Bresser, J.H.P., 2012. The effect of CO2 on creep of wet calcite aggregates. J. Geophys. Res. Solid Earth 117. https://doi.org/10.1029/2011JB008789

Liteanu, E., Spiers, C.J., De Bresser, J.H.P., 2013. The influence of water and supercritical CO2 on the failure behavior of chalk. Tectonophysics 599, 157–169. https://doi.org/10.1016/j.tecto.2013.04.013

Long, X., Ma, Y., Qi, L., 2014. Biogenic and synthetic high magnesium calcite - A review. J. Struct. Biol. https://doi.org/10.1016/j.jsb.2013.11.004

Luty, B.A., Davis, M.E., Tironi, I.G., Van Gunsteren, W.F., 1994. A comparison of particle-particle, particle-mesh and ewald methods for calculating electrostatic interactions in periodic molecular systems. Mol. Simul. 14, 11–20. https://doi.org/10.1080/08927029408022004

MacKerell, A.D., Bashford, D., Bellott, M., Dunbrack, R.L., Evanseck, J.D., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F.T.K., Mattos, C., Michnick, S., Ngo, T., Nguyen, D.T., Prodhom, B., Reiher, W.E., Roux, B., Schlenkrich, M., Smith, J.C., Stote, R., Straub, J., Watanabe, M., Wiórkiewicz-Kuczera, J., Yin, D., Karplus, M., 1998. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B 102, 3586–3616. https://doi.org/10.1021/jp973084f

Maeda, M., Okada, K., Tsukamoto, Y., Wakabayashi, K., Ito, K., 1990. Complex formation of calcium(II) with amino acids under physiological conditions. J. Chem. Soc. Dalt. Trans. 2337. https://doi.org/10.1039/dt9900002337

Mamatkulov, S., Fyta, M., Netz, R.R., 2013. Force fields for divalent cations based on single-ion and ion-pair properties. J. Chem. Phys. 138, 024505. https://doi.org/10.1063/1.4772808

Mann, S., Archibald, D.D., Didymus, J.M., Douglas, T., Heywood, B.R., Meldrum, F.C., Reeves, N.J., 1993. Crystallization at Inorganic-organic Interfaces: Biominerals and Biomimetic Synthesis. Science (80-. ). 261, 1286–1292. https://doi.org/10.1126/science.261.5126.1286

Marchand, S., Roux, B., 1998. Molecular dynamics study of calbindin D(9k) in the Apo and singly and doubly calcium-loaded states. Proteins Struct. Funct. Genet. 33, 265–284. https://doi.org/10.1002/(SICI)1097-0134(19981101)33:2<265::AID-PROT10>3.0.CO;2-I

Martin, P., Spagnoli, D., Marmier, A., Parker, S.C., Sayle, D.C., Watson, G., 2006. Application of molecular dynamics DL_POLY codes to interfaces of inorganic materials. Mol. Simul. 32, 1079–1093. https://doi.org/10.1080/08927020601013817

Martinek, T., Duboué-Dijon, E., Timr, Š., Mason, P.E., Baxová, K., Fischer, H.E., Schmidt, B., Pluhařová, E., Jungwirth, P., 2018. Calcium ions in aqueous solutions: Accurate force field description aided by ab initio molecular dynamics and neutron scattering. J. Chem. Phys. 148, 222813. https://doi.org/10.1063/1.5006779

Marutschke, C., Walters, D., Cleveland, J., Hermes, I., Bechstein, R., Kühnle, A., 2014. Three-dimensional hydration layer mapping on the (10.4) surface of calcite using amplitude

modulation atomic force microscopy. Nanotechnology 25, 335703. https://doi.org/10.1088/0957-4484/25/33/335703

Marxen, J.C., Becker, W., Finke, D., Hasse, B., Epple, M., 2003. Early mineralization in Biomphalaria glabrata: Microscopic and structural results. J. Molluscan Stud. 69, 113–121. https://doi.org/10.1093/mollus/69.2.113

Mass, T., Drake, J.L., Peters, E.C., Jiang, W., Falkowski, P.G., 2014. Immunolocalization of skeletal matrix proteins in tissue and mineral of the coral Stylophora pistillata. Proc. Natl. Acad. Sci. U. S. A. 111, 12728–12733. https://doi.org/10.1073/pnas.1408621111

Mass, T., Giuffre, A.J., Sun, C.-Y., Stifler, C.A., Frazier, M.J., Neder, M., Tamura, N., Stan, C. V, Marcus, M.A., Gilbert, P.U.P.A., n.d. Amorphous calcium carbonate particles form coral skeletons. https://doi.org/10.1073/pnas.1707890114

Maune, J.F., Klee, C.B., Beckingham, K., 1992. Ca2+ binding and conformational change in two series of point mutations to the individual Ca2+-binding sites of calmodulin. J. Biol. Chem. 267, 5286–5295. https://doi.org/10.1016/s0021-9258(18)42764-0

Maureen I. McCarthy, *, Gregory K. Schenter, Carol A. Scamehorn, and, Nicholas, J.B., 1996. Structure and Dynamics of the Water/MgO Interface. https://doi.org/10.1021/JP961373I

McGibbon, R.T., Beauchamp, K.A., Harrigan, M.P., Klein, C., Swails, J.M., Hernández, C.X., Schwantes, C.R., Wang, L.P., Lane, T.J., Pande, V.S., 2015. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. Biophys. J. 109, 1528–1532. https://doi.org/10.1016/j.bpj.2015.08.015

Meldrum, F.C., Cölfen, H., 2008. Controlling mineral morphologies and structures in biological and synthetic systems. Chem. Rev. https://doi.org/10.1021/cr8002856

Meng, S., Wang, E.G., Gao, S., 2004. Water adsorption on metal surfaces: A general picture from density functional theory studies. Phys. Rev. B - Condens. Matter Mater. Phys. 69, 1–13. https://doi.org/10.1103/PhysRevB.69.195404

Meng, Y., Guo, Z., Yao, H., Yeung, K.W.K., Thiyagarajan, V., 2019. Calcium carbonate unit realignment under acidification: A potential compensatory mechanism in an edible estuarine oyster. Mar. Pollut. Bull. 139, 141–149. https://doi.org/10.1016/j.marpolbul.2018.12.030

Meyer, H.J., 1969. Struktur und fehlordnung des vaterits. Zeitschrift fur Krist. - New Cryst. Struct. 128, 183–212. https://doi.org/10.1524/zkri.1969.128.3-6.183

Michaelides, A., Ranea, V.A., de Andres, P.L., King, D.A., 2003. General Model for Water Monomer Adsorption on Close-Packed Transition and Noble Metal Surfaces. Phys. Rev. Lett. 90, 4. https://doi.org/10.1103/PhysRevLett.90.216102

Michel, F.M., MacDonald, J., Feng, J., Phillips, B.L., Ehm, L., Tarabrella, C., Parise, J.B., Reeder, R.J., 2008. Structural characteristics of synthetic amorphous calcium carbonate. Chem. Mater. 20, 4720–4728. https://doi.org/10.1021/cm800324v

Migliorati, V., Serva, A., Terenzio, F.M., D'Angelo, P., 2017. Development of Lennard-Jones and Buckingham Potentials for Lanthanoid Ions in Water. Inorg. Chem. 56, 6214–6224. https://doi.org/10.1021/acs.inorgchem.7b00207

Miller, J.N., Miller, J.C., 2005. Statistics and chemometrics for analytical chemistry. Pearson Education.

Mínguez-Toral, M., Cuevas-Zuviría, B., Garrido-Arandia, M., Pacios, L.F., 2020. A computational structural study on the DNA-protecting role of the tardigrade-unique Dsup protein. Sci. Rep. 10, 1–18. https://doi.org/10.1038/s41598-020-70431-1

Molins, S., Soulaine, C., Prasianakis, N.I., Abbasi, A., Poncet, P., Ladd, A.J.C., Starchenko, V., Roman, S., Trebotich, D., Tchelepi, H.A., Steefel, C.I., 2020. Simulation of mineral dissolution at the pore scale with evolving fluid-solid interfaces: review of approaches and benchmark problem set. Comput. Geosci. 1–34. https://doi.org/10.1007/s10596-019-09903-x

Molins, S., Trebotich, D., Yang, L., Ajo-Franklin, J.B., Ligocki, T.J., Shen, C., Steefel, C.I., 2014. Pore-scale controls on calcite dissolution rates from flow-through laboratory and numerical experiments. Environ. Sci. Technol. 48, 7453–7460. https://doi.org/10.1021/es5013438

Montanari, G., Lakshtanov, L.Z., Tobler, D.J., Dideriksen, K., Dalby, K.N., Bovet, N., Stipp, S.L.S., 2016. Effect of Aspartic Acid and Glycine on Calcite Growth. Cryst. Growth Des. 16, 4813–4821. https://doi.org/10.1021/acs.cgd.5b01635

Morse, J.W., Arvidson, R.S., Lüttge, A., 2007. Calcium Carbonate Formation and Dissolution. Chem. Rev. 107, 342–381. https://doi.org/10.1021/cr050358j

Morse, J.W., Mackenzie, F.T., 1990. Geochemistry of Sedimentary Carbonates. Elsevier Science, New York, NY, USA.

Mukhopadhyay, A., Cole, W.T.S., Saykally, R.J., 2015. The water dimer I: Experimental characterization. Chem. Phys. Lett. 633, 13–26. https://doi.org/10.1016/j.cplett.2015.04.016

Müller, G., Tietz, R., 1966. Recent dolomitization of quaternary biocalcarenites from fuerteventura (Canary Islands). Contrib. to Mineral. Petrol. 13, 89–96. https://doi.org/10.1007/BF00518020

Murai, K., Higuchi, M., Kinoshita, T., Nagata, K., Kato, K., 2013. Calcium carbonate

biomineralization utilizing a multifunctional β-sheet peptide template. Chem. Commun. 49, 9947–9949. https://doi.org/10.1039/c3cc44947e

Mutisya, S.M., Kirch, A., De Almeida, J.M., Sánchez, V.M., Miranda, C.R., 2017. Molecular Dynamics Simulations of Water Confined in Calcite Slit Pores: An NMR Spin Relaxation and Hydrogen Bond Analysis. J. Phys. Chem. C 121, 6674–6684. https://doi.org/10.1021/acs.jpcc.6b12412

Nada, H., Nishimura, T., Sakamoto, T., Kato, T., 2016. Heterogeneous growth of calcite at aragonite {001}- and vaterite {001}-melt interfaces: A molecular dynamics simulation study. J. Cryst. Growth 450, 148–159. https://doi.org/10.1016/j.jcrysgro.2016.06.042

Negro, A.D., Ungaretti, L., 1971. Refinement of the crystal structure of aragonite. Am. Mineral. J. Earth Planet. Mater. 56, 768–772. https://doi.org/10.1180/minmag.1985.049.350.12

Nehrke, G., 2007. Calcite Precipitation from Aqueous solution: Transformation from Vaterite and Role of Solution Stoichiometry. Geol. Ultraiectina; Meded. van Fac. Geowetenschappen, Univ. Utr. 273.

Nerenberg, P.S., Head-Gordon, T., 2018. New developments in force fields for biomolecular simulations. Curr. Opin. Struct. Biol. https://doi.org/10.1016/j.sbi.2018.02.002

Nielsen, A.E., 1984. Electrolyte crystal growth mechanisms. J. Cryst. Growth 67, 289–310. https://doi.org/10.1016/0022-0248(84)90189-1

Nielsen, A.E., Toft, J.M., 1984. Electrolyte Crystal Growth Kinetics. J. Cryst. Growth 67, 278–288. https://doi.org/10.1016/0022-0248(84)90188-X

Nielsen, L.C., De Yoreo, J.J., DePaolo, D.J., 2013. General model for calcite growth kinetics in the presence of impurity ions. Geochim. Cosmochim. Acta 115, 100–114.

https://doi.org/10.1016/j.gca.2013.04.001

Nielsen, L.C., DePaolo, D.J., De Yoreo, J.J., 2012. Self-consistent ion-by-ion growth model for kinetic isotopic fractionation during calcite precipitation. Geochim. Cosmochim. Acta 86, 166–181. https://doi.org/10.1016/j.gca.2012.02.009

Nielsen, M.H., Aloni, S., De Yoreo, J.J., 2014. In situ TEM imaging of CaCO3 nucleation reveals coexistence of direct and indirect pathways. Science (80-. ). 345, 1158–1162. https://doi.org/10.1126/science.1254051

Ning, Y., Han, L., Douverne, M., Penfold, N.J.W., Derry, M.J., Meldrum, F.C., Armes, S.P., 2019. What Dictates the Spatial Distribution of Nanoparticles within Calcite? J. Am. Chem. Soc. 141, 2481–2489. https://doi.org/10.1021/jacs.8b12291

Noel, E.H., Kim, Y.Y., Charnock, J.M., Meldrum, F.C., 2013. Solid state crystallization of amorphous calcium carbonate nanoparticles leads to polymorph selectivity. CrystEngComm 15, 697–705. https://doi.org/10.1039/c2ce26529j

Nosé, S., 1984. A unified formulation of the constant temperature molecular dynamics methods. J. Chem. Phys. 81, 511–519. https://doi.org/10.1063/1.447334

Nygren, M.A., Gay, D.H., Catlow, C.R.A., Wilson, M.P., Rohl, A.L., 1998. Incorporation of growth-inhibiting diphosphonates into steps on the calcite cleavage plane surface. J. Chem. Soc. - Faraday Trans. 94, 3685–3693. https://doi.org/10.1039/a806585c

Odutola, J.A., Dyke, T.R., 1980. Partially deuterated water dimers: Microwave spectra and structure. J. Chem. Phys. 72, 5062–5070. https://doi.org/10.1063/1.439795

Olajire, A.A., 2015. A review of oilfield scale management technology for oil and gas production. https://doi.org/10.1016/j.petrol.2015.09.011

Orme, C.A., Noy, A., Wierzbicki, A., McBride, M.T., Grantham, M., Teng, H.H., Dove, P.M., DeYoreo, J.J., 2001. Formation of chiral morphologies through selective binding of amino acids to calcite surface steps. Nature 411, 775–779. https://doi.org/10.1038/35081034

Paesani, F., Zhang, W., Case, D.A., Cheatham, T.E., Voth, G.A., 2006. An accurate and simple quantum model for liquid water. J. Chem. Phys. 125, 184507. https://doi.org/10.1063/1.2386157

Pandey, S.N., Vishal, V., Chaudhuri, A., 2018. Geothermal reservoir modeling in a coupled thermo-hydro-mechanical-chemical approach: A review. https://doi.org/10.1016/j.earscirev.2018.09.004

Paquette, J., Reeder, R.J., 1995. Relationship between surface structure, growth mechanism, and trace element incorporation in calcite. Geochim. Cosmochim. Acta 59, 735–749. https://doi.org/10.1016/0016-7037(95)00004-J

Parker, S.C., Kelsey, E.T., Oliver, P.M., Titiloye, J.O., 1993. Computer modelling of inorganic solids and surfaces. Faraday Discuss. 95, 75–84. https://doi.org/10.1039/FD9939500075

Parker, S.C., Kerisit, S., Marmier, A., Grigoleit, S., Watson, G.W., 2003. Modelling Inorganic Solids and Their Interfaces: A Combined Approach of Atomistic and Electronic Structure Simulation Techniques. ChemInform 35, no-no. https://doi.org/10.1002/chin.200406248

Parker, S.C., Oliver, P.M., De Leeuw, N.H., Titiloye, J.O., Watson, G.W., 1997. Atomistic simulation of mineral surfaces: Studies of surface stability and growth. Phase Transitions 61, 83–107. https://doi.org/10.1080/01411599708223731

Parvan, M.-G., Voicu, G., Badanoiu, A.-I., Nicoara, A.-I., Vasile, E., Lamastra, R., 2021. CO 2 Sequestration in the Production of Portland Cement Mortars with Calcium Carbonate Additions. https://doi.org/10.3390/nano

Pasteris, J.D., Wopenka, B., Valsami-Jones, E., 2008. Bone and Tooth Mineralization: Why Apatite? Elements 4, 97–104. https://doi.org/10.2113/GSELEMENTS.4.2.97

Pérez-Badell, Y., Montero, L.A., 2010. Multiple minima hypersurfaces studies of aluminosilicate hydration. Int. J. Quantum Chem. 110, 586–594. https://doi.org/10.1002/qua.22292

Petrucci, S., 1971. Ionic Interactions; From Dilute Solutions to Fused Salts, The Journal of Chemical Thermodynamics. Elsevier Science. https://doi.org/10.1016/0021-9614(72)90085-7

Pham, T.T., Lemaire, T., Capiez-Lernout, E., Lewerenz, M., To, Q.D., Christie, J.K., Di Tommaso, D., de Leeuw, N.H., Naili, S., 2015. Properties of water confined in hydroxyapatite nanopores as derived from molecular dynamics simulations. Theor. Chem. Acc. 134, 1–14. https://doi.org/10.1007/s00214-015-1653-3

Piana, S., Jones, F., Gale, J.D., 2006. Assisted desolvation as a key kinetic step for crystal growth. J. Am. Chem. Soc. 128, 13568–13574. https://doi.org/10.1021/ja064706q

Picker, A., Kellermeier, M., Seto, J., Gebauer, D., Cölfen, H., 2012. The multiple effects of amino acids on the early stages of calcium carbonate crystallization, in: Zeitschrift Fur Kristallographie. De Gruyter, pp. 744–757. https://doi.org/10.1524/zkri.2012.1569

Pipich, V., Balz, M., Wolf, S.E., Tremel, W., Schwann, D., 2008. Nucleation and growth of CaCO3 mediated by the egg-white protein ovalbumin: A time-resolved in situ study using small-angle neutron scattering. J. Am. Chem. Soc. 130, 6879–6892. https://doi.org/10.1021/ja801798h

Plimpton, S., 1995. Fast parallel algorithms for short-range molecular dynamics. J. Comput. Phys. 117, 1–19. https://doi.org/10.1006/jcph.1995.1039

Pokrovsky, O.S., Schott, J., 2002. Surface chemistry and dissolution kinetics of divalent metal carbonates. Environ. Sci. Technol. 36, 426–432. https://doi.org/10.1021/es010925u

Pokrovsky, O.S., Schott, J., 2001. Kinetics and mechanism of dolomite dissolution in neutral to alkaline solutions revisited. Am. J. Sci. 301, 597–626. https://doi.org/10.2475/ajs.301.7.597

Pollock, E.L., Glosli, J., 1996. Comments on P3M, FMM, and the Ewald method for large periodic Coulombic systems. Comput. Phys. Commun. https://doi.org/10.1016/0010-4655(96)00043-4

Pouget, E.M., Bomans, P.H.H., Goos, J.A.C.M., Frederik, P.M., de With, G., Sommerdijk, N.A.J.M., 2009. The Initial Stages of Template-Controlled CaCO3 Formation Revealed by Cryo-TEM. Science (80-. ). 323, 1455–1458. https://doi.org/10.1126/science.1169434

Project, E., Nachliel, E., Gutman, M., 2008. Parameterization of Ca+2–protein interactions for molecular dynamics simulations. J. Comput. Chem. 29, 1163–1169. https://doi.org/10.1002/jcc.20876

Pudlas, M., Koch, S., Bolwien, C., Walles, H., 2010. Raman spectroscopy as a tool for quality and sterility analysis for tissue engineering applications like cartilage transplants. Int. J. Artif. Organs 33, 228–237. https://doi.org/10.1021/j100606a013

Purton, J., Catlow, C.R.A., 1990. Computer simulation of feldspar structures, American Mineralogist.

R, L., JS, J., RM, K., S, V., 2005. Structure-function relationship of avian eggshell matrix proteins: a comparative study of two major eggshell matrix proteins, ansocalcin and OC-17. Biomacromolecules 6. https://doi.org/10.1021/BM049423+

Rachlin, A.L., Henderson, G.S., Goh, M.C., 1992. An atomic force microscope (AFM) study of the calcite cleavage plane: image averaging in Fourier space. Am. Mineral. 77, 904–910.

Radha, A. V., Forbes, T.Z., Killian, C.E., Gilbert, P.U.P.A., Navrotsky, A., 2010. Transformation and crystallization energetics of synthetic and biogenic amorphous calcium carbonate. Proc. Natl. Acad. Sci. U. S. A. 107, 16438–16443. https://doi.org/10.1073/pnas.1009959107

Raiteri, P., Demichelis, R., Gale, J.D., 2015. Thermodynamically Consistent Force Field for Molecular Dynamics Simulations of Alkaline-Earth Carbonates and Their Aqueous Speciation. J. Phys. Chem. C 119, 24447–24458. https://doi.org/10.1021/acs.jpcc.5b07532

Raiteri, P., Demichelis, R., Gale, J.D., Kellermeier, M., Gebauer, D., Quigley, D., Wright, L.B., Walsh, T.R., 2012. Exploring the influence of organic species on pre- and post-nucleation calcium carbonate. Faraday Discuss. 159, 61. https://doi.org/10.1039/c2fd20052j

Raiteri, P., Gale, J.D., 2010. Water is the key to nonclassical nucleation of amorphous calcium carbonate. J. Am. Chem. Soc. 132, 17623–17634. https://doi.org/10.1021/ja108508k

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. J. Phys. Chem. C 114, 5997–6010. https://doi.org/10.1021/jp910977a

Rajagopal, G., Needs, R.J., 1994. An optimized ewald method for long-ranged potentials. J. Comput. Phys. 115, 399–405. https://doi.org/10.1006/jcph.1994.1205

Rani, R.S., Saharay, M., 2019. Molecular dynamics simulation of protein-mediated

biomineralization of amorphous calcium carbonate. RSC Adv. 9, 1653–1663. https://doi.org/10.1039/C8RA08459A

Raoof, A., Nick, H.M., Hassanizadeh, S.M., Spiers, C.J., 2013. PoreFlow: A complex pore-network model for simulation of reactive transport in variably saturated porous media. Comput. Geosci. 61, 160–174. https://doi.org/10.1016/j.cageo.2013.08.005

Raoof, A., Nick, H.M., Wolterbeek, T.K.T., Spiers, C.J., 2012. Pore-scale modeling of reactive transport in wellbore cement under CO2 storage conditions. Int. J. Greenh. Gas Control 11, S67–S77. https://doi.org/10.1016/j.ijggc.2012.09.012

Reddy, M.M., Nancollas, G.H., 1971. The crystallization of calcium carbonate. I. Isotopic exchange and kinetics. J. Colloid Interface Sci. 36, 166–172. https://doi.org/10.1016/0021-9797(71)90161-5

Reggi, M., Fermani, S., Landi, V., Sparla, F., Caroselli, E., Gizzi, F., Dubinsky, Z., Levy, O., Cuif, J.P., Dauphin, Y., Goffredo, S., Falini, G., 2014. Biomineralization in mediterranean corals: The role of the intraskeletal organic matrix. Cryst. Growth Des. 14, 4310–4320. https://doi.org/10.1021/cg5003572

Reif, M.M., Hünenberger, P.H., Oostenbrink, C., 2012. New interaction parameters for charged amino acid side chains in the GROMOS force field. J. Chem. Theory Comput. 8, 3705–3723. https://doi.org/10.1021/ct300156h

Reischl, B., Raiteri, P., Gale, J.D., Rohl, A.L., 2019. Atomistic Simulation of Atomic Force Microscopy Imaging of Hydration Layers on Calcite, Dolomite, and Magnesite Surfaces. J. Phys. Chem. C 123, 14985–14992. https://doi.org/10.1021/acs.jpcc.9b00939

Reischl, B., Watkins, M., Foster, A.S., 2013. Free energy approaches for modeling atomic force microscopy in liquids. J. Chem. Theory Comput. 9, 600–608.

https://doi.org/10.1021/ct3008342

Richards, C.S., Wang, F., Becker, W.C., Edwards, M.A., 2018. A 21st-Century Perspective on Calcium Carbonate Formation in Potable Water Systems. Environ. Eng. Sci. https://doi.org/10.1089/ees.2017.0115

Rieger, J., Frechen, T., Cox, G., Heckmann, W., Schmidt, C., Thieme, J., 2007. Precursor structures in the crystallization/precipitation processes of CaCO3 and control of particle formation by polyelectrolytes. Faraday Discuss. 136, 265–277. https://doi.org/10.1039/b701450c

Robertson, M.J., Tirado-Rives, J., Jorgensen, W.L., 2015. Improved Peptide and Protein Torsional Energetics with the OPLS-AA Force Field. J. Chem. Theory Comput. 11, 3499–3509. https://doi.org/10.1021/acs.jctc.5b00356

Robie, R., Hemingway, B., 1995. Thermodynamic Properties of Minerals and Related Substances at 298.15 K and 1 bar (105 Pascals) Pressure and at Higher Temperatures. Government Printing Office, Washington, DC, USA. https://doi.org/10.1016/0016-7037(79)90232-1

Rode, B.M., Hofer, T.S., 2006. How to access structure and dynamics of solutions: The capabilities of computational methods. Pure Appl. Chem. https://doi.org/10.1351/pac200678030525

Rode, B.M., Schwenk, C.F., Hofer, T.S., Randolf, B.R., 2005. Coordination and ligand exchange dynamics of solvated metal ions. Coord. Chem. Rev. 249, 2993–3006. https://doi.org/10.1016/j.ccr.2005.03.032

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

265–271. https://doi.org/10.1039/c0nr00589d

Rodriguez-Navarro, C., Burgos Cara, A., Elert, K., Putnis, C. V., Ruiz-Agudo, E., 2016. Direct Nanoscale Imaging Reveals the Growth of Calcite Crystals via Amorphous Nanoparticles. Cryst. Growth Des. 16, 1850–1860. https://doi.org/10.1021/acs.cgd.5b01180

Roehl, P.O., Choquette, P.W., 1990. Carbonate petroleum reservoirs. Sediment. Geol. 66, 305. https://doi.org/10.1016/0037-0738(90)90081-4

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. Am. Mineral. 88, 921–925. https://doi.org/10.2138/am-2003-5-622

Roleda, M.Y., Boyd, P.W., Hurd, C.L., 2012. BEFORE OCEAN ACIDIFICATION: CALCIFIER CHEMISTRY LESSONS1. J. Phycol. 48, 840–843. https://doi.org/10.1111/j.1529-8817.2012.01195.x

Røyne, A., Bisschop, J., Dysthe, D.K., 2011. Experimental investigation of surface energy and subcritical crack growth in calcite. J. Geophys. Res. 116, B04204. https://doi.org/10.1029/2010JB008033

Ruiz-Agudo, E., Putnis, C. V., 2012. Direct observations of mineral fluid reactions using atomic force microscopy: the specific example of calcite. Mineral. Mag. 76, 227–253. https://doi.org/10.1180/minmag.2012.076.1.227

Ruiz-Agudo, E., Putnis, C. V., Jiménez-López, C., Rodriguez-Navarro, C., 2009. An atomic force microscopy study of calcite dissolution in saline solutions: The role of magnesium ions. Geochim. Cosmochim. Acta 73, 3201–3217. https://doi.org/10.1016/j.gca.2009.03.016

Ruiz-Hernandez, S.E., Grau-Crespo, R., Almora-Barrios, N., Wolthers, M., Ruiz-Salvador, A.R., Fernandez, N., De Leeuw, N.H., 2012. Mg/Ca partitioning between aqueous solution and aragonite mineral: A molecular dynamics study. Chem. - A Eur. J. 18, 9828–9833. https://doi.org/10.1002/chem.201200966

Ruiz Hernandez, S.E., De Leeuw, N.H., 2015. Effect of Chondroitin 4-Sulfate on the Growth and Morphology of Calcium Oxalate Monohydrate: A Molecular Dynamics Study. Cryst. Growth Des. 15, 4438–4447. https://doi.org/10.1021/acs.cgd.5b00747

Ruiz Hernandez, S.E., Streeter, I., De Leeuw, N.H., 2015. The effect of water on the binding of glycosaminoglycan saccharides to hydroxyapatite surfaces: A molecular dynamics study. Phys. Chem. Chem. Phys. 17, 22377–22388. https://doi.org/10.1039/c5cp02630j

Sand, K.K., Tobler, D.J., Dobberschütz, S., Larsen, K.K., Makovicky, E., Andersson, M.P., Wolthers, M., Stipp, S.L.S., 2016. Calcite Growth Kinetics: Dependence on Saturation Index, Ca2:CO32- Activity Ratio, and Surface Atomic Structure. Cryst. Growth Des. 16, 3602–3612. https://doi.org/10.1021/acs.cgd.5b01792

Sano, Y., Shirai, K., Takahata, N., Hirata, T., Sturchio, N.C., 2005. Nano-SIMS Analysis of Mg, Sr, Ba and U in Natural Calcium Carbonate, ANALYTICAL SCIENCES.

Schmidt, I., Zolotoyabko, E., Lee, K., Gjardy, A., Berner, A., Lakin, E., Fratzl, P., Wagermaier, W., 2019. Effect of Strontium Ions on Crystallization of Amorphous Calcium Carbonate. Cryst. Res. Technol. 54, 1900002. https://doi.org/10.1002/crat.201900002

Schulze, S., Vogel, H., 1998. Aspects of the Safe Storage of Acrylic Monomers: Kinetics of the Oxygen Consumption. Chem. Eng. Technol. 21, 829–837. https://doi.org/10.1002/(SICI)1521-4125(199810)21:10<829::AID-CEAT829>3.0.CO;2-T

Schwegler, E., Grossman, J.C., Gygi, F., Galli, G., 2004. Towards an assessment of the accuracy of density functional theory for first principles simulations of water. II. J. Chem. Phys. 121, 5400–5409. https://doi.org/10.1063/1.1782074

Scott, R., Allen, M.P., Tildesley, D.J., 1991. Computer Simulation of Liquids. Math. Comput. 57, 442. https://doi.org/10.2307/2938686

Sekkal, W., Zaoui, A., 2013. Nanoscale analysis of the morphology and surface stability of calcium carbonate polymorphs. Sci. Rep. 3, 1–10. https://doi.org/10.1038/srep01587

Sena, M.M., Morrow, C.P., Kirkpatrick, R.J., Krishnan, M., 2015. Supercritical Carbon Dioxide at Smectite Mineral-Water Interfaces: Molecular Dynamics and Adaptive Biasing Force Investigation of $CO_2/H_2O$ Mixtures Nanoconfined in Na-Montmorillonite. Chem. Mater. 27, 6946–6959. https://doi.org/10.1021/acs.chemmater.5b01855

Senthilmurugan, B., Ghosh, B., Kundu, S.S., Haroun, M., Kameshwari, B., 2010. Maleic acid based scale inhibitors for calcium sulfate scale inhibition in high temperature application. J. Pet. Sci. Eng. 75, 189–195. https://doi.org/10.1016/j.petrol.2010.11.002

Shen, J.W., Li, C., Vegt, N.F.A. Van Der, Peter, C., Van Der Vegt, N.F.A., Peter, C., Vegt, N.F.A. Van Der, Peter, C., 2013. Supplementary Information for " Understanding the Control of Mineralization by Polyelectrolyte Additives : Simulation of Preferential Binding to Calcite Surfaces ." J. Phys. Chem. C 117, 1–7. https://doi.org/10.1021/jp402341w

Shrake, A., Rupley, J.A., 1973. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. J. Mol. Biol. 79, 351–371. https://doi.org/10.1016/0022-2836(73)90011-9

Silvestrelli, P.L., Parrinello, M., 1999. Water molecule dipole in the gas and in the liquid phase.

Phys. Rev. Lett. 82, 3308–3311. https://doi.org/10.1103/PhysRevLett.82.3308

Sit, P.H.-L., Marzari, N., 2005. Static and dynamical properties of heavy water at ambient conditions from first-principles molecular dynamics. J. Chem. Phys. 122, 204510. https://doi.org/10.1063/1.1908913

Skinner, H.C.W., 2005. Biominerals. Mineral. Mag. 69, 621–641. https://doi.org/10.1180/0026461056950275

Smeets, P.J.M., Cho, K.R., Kempen, R.G.E., Sommerdijk, N.A.J.M., De Yoreo, J.J., 2015. Calcium carbonate nucleation driven by ion binding in a biomimetic matrix revealed by in situ electron microscopy. Nat. Mater. 14, 394–399. https://doi.org/10.1038/nmat4193

Smeets, P.J.M.M., Finney, A.R., Habraken, W.J.E.M.E.M., Nudelman, F., Friedrich, H., Laven, J., De Yoreo, J.J., Rodger, P.M., Sommerdijk, N.A.J.M.J.M., 2017. A classical view on nonclassical nucleation. Proc. Natl. Acad. Sci. U. S. A. 114, E7882–E7890. https://doi.org/10.1073/pnas.1700342114

Smith, W., 1996. Mean Square Displacement [WWW Document]. URL https://www.scd.stfc.ac.uk/Pages/DL_POLY.aspx (accessed 4.10.21).

Solanke, C.O., Trapl, D., Šućur, Z., Mareška, V., Tvaroška, I., Spiwok, V., 2019. Atomistic simulation of carbohydrate-protein complex formation: Hevein-32 domain. Sci. Rep. 9, 18918. https://doi.org/10.1038/s41598-019-53815-w

Söngen, H., Reischl, B., Miyata, K., Bechstein, R., Raiteri, P., Rohl, A.L., Gale, J.D., Fukuma, T., Kühnle, A., 2018. Resolving Point Defects in the Hydration Structure of Calcite (10.4) with Three-Dimensional Atomic Force Microscopy. Phys. Rev. Lett. 120, 116101. https://doi.org/10.1103/PhysRevLett.120.116101

Soper, A.K., 2000. The radial distribution functions of water and ice from 220 to 673 K and at pressures up to 400 MPa. Chem. Phys. 258, 121–137. https://doi.org/10.1016/S0301-0104(00)00179-8

Spagnoli, D., Cooke, D.J., Kerisit, S., Parker, S.C., 2006a. Molecular dynamics simulations of the interaction between the surfaces of polar solids and aqueous solutions. J. Mater. Chem. 16, 1997–2006. https://doi.org/10.1039/b600808a

Spagnoli, D., Kerisit, S., Parker, S.C., 2006b. Atomistic simulation of the free energies of dissolution of ions from flat and stepped calcite surfaces. J. Cryst. Growth 294, 103–110. https://doi.org/10.1016/j.jcrysgro.2006.05.030

Stack, A.G., Grantham, M.C., 2010. Growth Rate of Calcite Steps As a Function of Aqueous Calcium-to-Carbonate Ratio: Independent Attachment and Detachment of Calcium and Carbonate Ions. Cryst. Growth Des. 10, 1409–1413. https://doi.org/10.1021/cg901395z

Stack, A.G., Raiteri, P., Gale, J.D., 2012. Accurate Rates of the Complex Mechanisms for Growth and Dissolution of Minerals Using a Combination of Rare-Event Theories. J. Am. Chem. Soc. 134, 11–14. https://doi.org/10.1021/ja204714k

Stanienda, K.J., 2016. Carbonate phases rich in magnesium in the Triassic limestones of the eastern part of the Germanic Basin. Carbonates and Evaporites 31, 387–405. https://doi.org/10.1007/s13146-016-0297-2

Stephenson, A.E., Deyoreo, J.J., Wu, L., Wu, K.J., Hoyer, J., Dove, P.M., 2008. Peptides enhance magnesium signature in calcite: Insights into origins of vital effects. Science (80-. ). 322, 724–727. https://doi.org/10.1126/science.1159417

Stukan, M.R., Asmadi, A., Abdallah, W., 2013. Bulk properties of SWM4-NDP water model at elevated temperature and pressure. J. Mol. Liq. 180, 65–69.

https://doi.org/10.1016/j.molliq.2012.12.023

Stumm, W., Morgan, J.J., 1996. Aquatic Chemistry : Chemical Equilibria and Rates in Natural Waters. Wiley.

Su, J., Liang, X., Zhou, Q., Zhang, G., Wang, H., Xie, L., Zhang, R., 2013. Structural characterization of amorphous calcium carbonate-binding protein: An insight into the mechanism of amorphous calcium carbonate formation. Biochem. J. 453, 179–186. https://doi.org/10.1042/BJ20130285

Subhas, A. V., Rollins, N.E., Berelson, W.M., Erez, J., Ziveri, P., Langer, G., Adkins, J.F., 2018. The dissolution behavior of biogenic calcites in seawater and a possible role for magnesium and organic carbon. Mar. Chem. 205, 100–112. https://doi.org/10.1016/j.marchem.2018.08.001

Tai, C.Y., Lu, J.-H., Wu, J.-K., 2005. Crystal Growth Rate of Calcite in a Constant-Composition Environment. J. Chinese Inst. Chem. Eng. 36, 443–450. https://doi.org/10.6967/JCICE.200509.0443

Tanaka, M., Girard, G., Davis, R., Peuto, A., Bignell, N., 2001. Recommended table for the density of water between 0 °c and 40 °c based on recent experimental reports. Metrologia 38, 301–309. https://doi.org/10.1088/0026-1394/38/4/3

Tang, N., Skibsted, L.H., 2016. Calcium binding to amino acids and small glycine peptides in aqueous solution: Toward peptide design for better calcium bioavailability. J. Agric. Food Chem. 64, 4376–4389. https://doi.org/10.1021/acs.jafc.6b01534

Tavafoghi Jahromi, M., Yao, G., Cerruti, M., 2013. The importance of amino acid interactions in the crystallization of hydroxyapatite. J. R. Soc. Interface 10. https://doi.org/10.1098/rsif.2012.0906

Telfer, G.B., Gale, J.D., Roberts, K.J., Jackson, R.A., Wilde, P.J., Meenan, P., 1997. A Transferable Interatomic Potential for Alkali Chlorates and Bromates. Acta Crystallogr. Sect. A Found. Crystallogr. 53, 415–420. https://doi.org/10.1107/S0108767397003097

Teng, H.H., Dove, P.M., 1997. Surface site-specific interactions of aspartate with calcite during dissolution: Implications for biomineralization. Am. Mineral. 82, 878–887. https://doi.org/10.2138/am-1997-9-1005

Teng, H.H., Dove, P.M., De Yoreo, J.J., 2000. Kinetics of calcite growth: Surface processes and relationships to macroscopic rate laws. Geochim. Cosmochim. Acta 64, 2255–2266. https://doi.org/10.1016/S0016-7037(00)00341-0

Titiloye, J.O., De Leeuw, N.H., Parker, S.C., 1998. Atomistic simulation of the differences between calcite and dolomite surfaces. Geochim. Cosmochim. Acta 62, 2637–2641. https://doi.org/10.1016/S0016-7037(98)00177-X

Todorov, I.T., Smith, W., Trachenko, K., Dove, M.T., 2006. DL_POLY_3: new dimensions in molecular dynamics simulations via massive parallelism. J. Mater. Chem. 16, 1911. https://doi.org/10.1039/b517931a

Tollefsen, E., Stockmann, G., Skelton, A., Mörth, C.-M., Dupraz, C., Sturkell, E., 2018. Chemical controls on ikaite formation. Mineral. Mag. 82, 1119–1129. https://doi.org/10.1180/mgm.2018.110

Toroz, D., Song, F., Chass, G.A., Tommaso, D. Di, 2021. New insights into the role of solution additive anions in Mg 2+ dehydration: implications for mineral carbonation †. https://doi.org/10.1039/d1ce00052g

Tribello, G.A., Bonomi, M., Branduardi, D., Camilloni, C., Bussi, G., 2014. PLUMED 2: New feathers for an old bird. Comput. Phys. Commun. 185, 604–613.

https://doi.org/10.1016/j.cpc.2013.09.018

Tribello, G.A., Bruneval, F., Liew, C.C., Parrinello, M., 2009. A molecular dynamics study of the early stages of calcium carbonate growth. J. Phys. Chem. B 113, 11680–11687. https://doi.org/10.1021/jp902606x

Ueckert, M., Wismeth, C., Baumann, T., 2020. Crystallization of calcium carbonate in a large-scale push–pull heat storage test in the Upper Jurassic carbonate aquifer. Geotherm. Energy 8, 7. https://doi.org/10.1186/s40517-020-0160-5

Van Maaren, P.J., Van Der Spoel, D., 2001. Molecular Dynamics Simulations of Water with Novel Shell-Model Potentials. J. Phys. Chem. B 105, 2618–2626. https://doi.org/10.1021/jp0038431

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. Comput. Phys. Commun. 167, 103–128. https://doi.org/10.1016/j.cpc.2004.12.014

Vashistha, P., Singh, S.K., Kumar, V., 2021. Sustainable Utilization of Paper-Industry Lime Sludge through Development of Metakaolinite-Based Cementitious Binder. J. Mater. Civ. Eng. 33, 04020492. https://doi.org/10.1061/(asce)mt.1943-5533.0003589

Vavrusova, M., Skibsted, L.H., 2013. Calcium binding to dipeptides of aspartate and glutamate in comparison with orthophosphoserine. J. Agric. Food Chem. 61, 5380–5384. https://doi.org/10.1021/jf400741e

Von Euw, S., Zhang, Q., Manichev, V., Murali, N., Gross, J., Feldman, L.C., Gustafsson, T., Flach, C., Mendelsohn, R., Falkowski, P.G., 2017. Biological control of aragonite formation in stony corals. Science (80-. ). 356, 933–938.

https://doi.org/10.1126/science.aam6371

Wahedi, H.M., Ahmad, S., Abbasi, S.W., 2020. Stilbene-based natural compounds as promising drug candidates against COVID-19. J. Biomol. Struct. Dyn. 1–10. https://doi.org/10.1080/07391102.2020.1762743

Wallace, A.F., Hedges, L.O., Fernandez-Martinez, A., Raiteri, P., Gale, J.D., Waychunas, G.A., Whitelam, S., Banfield, J.F., De Yoreo, J.J., 2013. Microscopic evidence for liquid-liquid separation in supersaturated CaCO3 solutions. Science (80-. ). 341, 885–889. https://doi.org/10.1126/science.1230915

Wang, D., Hamm, L.M., Giuffre, A.J., Echigo, T., Rimstidt, J.D., De Yoreo, J.J., Grotzinger, J., Dove, P.M., 2012. Revisiting geochemical controls on patterns of carbonate deposition through the lens of multiple pathways to mineralization. Faraday Discuss. 159, 371–386. https://doi.org/10.1039/c2fd20077e

Wang, J., Becker, U., 2009. Structure and carbonate orientation of vaterite (CaCO3). Am. Mineral. 94, 380–386. https://doi.org/10.2138/am.2009.2939

Wang, J., Hou, T., 2011. Application of molecular dynamics simulations in molecular property prediction II: Diffusion coefficient. J. Comput. Chem. 32, 3505–3519. https://doi.org/10.1002/jcc.21939

Wang, J., Román-Pérez, G., Soler, J.M., Artacho, E., Fernández-Serra, M.-V., 2011. Density, structure, and dynamics of water: The effect of van der Waals interactions. J. Chem. Phys. 134, 024516. https://doi.org/10.1063/1.3521268

Wang, J., Xu, X., Zhou, X., Chen, P., Liang, H., Li, X., Zhong, W., Hao, P., 2020. Molecular simulation of SARS-CoV- 2 spike protein binding to pangolin ACE2 or human ACE2 natural variants reveals altered susceptibility to infection. J. Gen. Virol. 101, 921–924.

https://doi.org/10.1099/jgv.0.001452

Wang, Q., 2011. A computational study of calcium carbonate.

Wang, Q., Wang, X.M., Tian, L.L., Cheng, Z.J., Cui, F.Z., 2011. In situ remineralizaiton of partially demineralized human dentine mediated by a biomimetic non-collagen peptide. Soft Matter 7, 9673–9680. https://doi.org/10.1039/c1sm05018d

Wang, X., Han, Y., Lin, L., Fuji, M., Endo, T., Watanabe, H., Takahashi, M., 2008. A molecular dynamics study on aqueous solutions for preparation of hollow CaCO3 particles. Model. Simul. Mater. Sci. Eng. 16. https://doi.org/10.1088/0965-0393/16/3/035006

Wasylenki, L.E., Dove, P.M., Wilson, D.S., De Yoreo, J.J., 2005. Nanoscale effects of strontium on calcite growth: An in situ AFM study in the absence of vital effects. Geochim. Cosmochim. Acta 69, 3017–3027. https://doi.org/10.1016/j.gca.2004.12.019

Wei, S., Cui, H., Jiang, Z., Liu, H., He, H., Fang, N., 2015. Biomineralization processes of calcite induced by bacteria isolated from marine sediments. https://doi.org/10.1590/S1517-838246220140533

Weiner, S., Addadi, L., 2011. Crystallization Pathways in Biomineralization. Annu. Rev. Mater. Res. 41, 21–40. https://doi.org/10.1146/annurev-matsci-062910-095803

Weiss, I.M., Tuross, N., Addadi, L., Weiner, S., 2002. Mollusc larval shell formation: Amorphous calcium carbonate is a precursor phase for aragonite. J. Exp. Zool. 293, 478–491. https://doi.org/10.1002/jez.90004

Willow, S.Y., Salim, M.A., Kim, K.S., Hirata, S., 2015. Ab initio molecular dynamics of liquid water using embedded-fragment second-order many-body perturbation theory towards its

accurate property prediction. Sci. Rep. 5, 14358. https://doi.org/10.1038/srep14358

Wolf, S., Lickert, B., Bray, S., Stock, G., 2020. Multisecond ligand dissociation dynamics from atomistic simulations. Nat. Commun. 11, 1–8. https://doi.org/10.1038/s41467-020-16655-1

Wolf, S.E., Leiterer, J., Kappl, M., Emmerling, F., Tremel, W., 2008. Early homogenous amorphous precursor stages of calcium carbonate and subsequent crystal growth in levitated droplets. J. Am. Chem. Soc. 130, 12342–12347. https://doi.org/10.1021/ja800984y

Wolthers, M., 2015. How minerals dissolve. Science (80-. ). https://doi.org/10.1126/science.aad0852

Wolthers, M., Charlet, L., Van Cappellen, P., 2008. The surface chemistry of divalent metal carbonate minerals: A critical assessment of surface charge and potential data using the charge distribution multi-site ion complexation model. Am. J. Sci. 308, 905–941. https://doi.org/10.2475/08.2008.02

Wolthers, M., Di Tommaso, D., Du, Z., de Leeuw, N.H., 2012a. Calcite surface structure and reactivity: molecular dynamics simulations and macroscopic surface modelling of the calcite–water interface. Phys. Chem. Chem. Phys. 14, 15145. https://doi.org/10.1039/c2cp42290e

Wolthers, M., Di Tommaso, D., Du, Z., De Leeuw, N.H., 2013. Variations in calcite growth kinetics with surface topography: molecular dynamics simulations and process-based growth kinetics modelling. CrystEngComm 15, 5506. https://doi.org/10.1039/c3ce40249e

Wolthers, M., Nehrke, G., Gustafsson, J.P., Van Cappellen, P., 2012b. Calcite growth kinetics: Modeling the effect of solution stoichiometry. Geochim. Cosmochim. Acta 77, 121–134.

https://doi.org/10.1016/j.gca.2011.11.003

Wray, J.J., Murchie, S.L., Bishop, J.L., Ehlmann, B.L., Milliken, R.E., Wilhelm, M.B., Seelos, K.D., Chojnacki, M., 2016. Orbital evidence for more widespread carbonate-bearing rocks on Mars. J. Geophys. Res. E Planets 121, 652–677. https://doi.org/10.1002/2015JE004972

Wright, K., Cygan, R.T., Slater, B., 2001. Structure of the (1014) surfaces of calcite, dolomite and magnesite under wet and dry conditions. Phys. Chem. Chem. Phys. 3, 839–844. https://doi.org/10.1039/b006130l

Wu, J.-Y., Chen, L.-J., Chen, Y.-P., Lin, S.-T., 2016. Molecular dynamics study on the nucleation of methane + tetrahydrofuran mixed guest hydrate. Phys. Chem. Chem. Phys. 18, 9935–9947. https://doi.org/10.1039/C5CP06419H

Wu, Y., Tepper, H.L., Voth, G.A., 2006. Flexible simple point-charge water model with improved liquid-state properties. J. Chem. Phys. 124, 024503. https://doi.org/10.1063/1.2136877

Xie, A. jian, Shen, Y. hua, Li, X. yan, Yuan, Z. wei, Qiu, L. guang, Zhang, C. yan, Yang, Y. feng, 2007. The role of Mg2+ and Mg2+/amino acid in controlling polymorph and morphology of calcium carbonate crystal. Mater. Chem. Phys. 101, 87–92. https://doi.org/10.1016/j.matchemphys.2006.02.019

Xu, J., Yan, C., Zhang, F., Konishi, H., Xu, H., Teng, H.H., 2013. Testing the cation-hydration effect on the crystallization of Ca-Mg-CO3 systems. Proc. Natl. Acad. Sci. 110, 17750–17755. https://doi.org/10.1073/pnas.1307612110

Xu, M., Higgins, S.R., 2011. Effects of magnesium ions on near-equilibrium calcite dissolution: Step kinetics and morphology. Geochim. Cosmochim. Acta 75, 719–733.

https://doi.org/10.1016/J.GCA.2010.10.018

Yoo, J., Wilson, J., Aksimentiev, A., 2016. Improved model of hydrated calcium ion for molecular dynamics simulations using classical biomolecular force fields. Biopolymers 105, 752–763. https://doi.org/10.1002/bip.22868

Yu, T., Souli, H., Péchaud, Y., Fleureau, J.-M., 2020. Optimizing protocols for microbial induced calcite precipitation (MICP) for soil improvement–a review. Eur. J. Environ. Civ. Eng. 1–16. https://doi.org/10.1080/19648189.2020.1755370

Zhang, J., Nancollas, G.H., 1998. Kink Density and Rate of Step Movement during Growth and Dissolution of an AB Crystal in a Nonstoichiometric Solution. J. Colloid Interface Sci. 200, 131–145. https://doi.org/10.1006/JCIS.1997.5357

Zhang, W., Shen, B., Chen, Y., Wang, T., Chen, W., 2020. Molecular dynamics simulations about isotope fractionation of methane in shale nanopores. Fuel 278, 118378. https://doi.org/10.1016/j.fuel.2020.118378

Zhang, X., Guo, J., Wu, S., Chen, F., Yang, Y., 2020. Divalent heavy metals and uranyl cations incorporated in calcite change its dissolution process. Sci. Rep. 10, 16864. https://doi.org/10.1038/s41598-020-73555-6

Zhang, Y., Yang, W., 1998. Comment on "Generalized Gradient Approximation Made Simple." Phys. Rev. Lett. 80, 890–890. https://doi.org/10.1103/PhysRevLett.80.890

Zhao, J., Wang, M., Lababidi, M.S., Al-Adwani, H., Gleason, K.K., 2018. A review of heterogeneous nucleation of calcium carbonate and control strategies for scale formation in multi-stage flash (MSF) desalination plants. https://doi.org/10.1016/j.desal.2018.05.008

Zhu, L.F., Friák, M., Lymperakis, L., Titrian, H., Aydin, U., Janus, A.M., Fabritius, H.O., Ziegler, A., Nikolov, S., Hemzalová, P., Raabe, D., Neugebauer, J., 2013. Ab initio study of single-crystalline and polycrystalline elastic properties of Mg-substituted calcite crystals. J. Mech. Behav. Biomed. Mater. 20, 296–304. https://doi.org/10.1016/j.jmbbm.2013.01.030

Živković, A., De Leeuw, N.H., 2020. Exploring the formation of intrinsic p -type and n -type defects in CuO. Phys. Rev. Mater. 4, 074606. https://doi.org/10.1103/PhysRevMaterials.4.074606

Zivković, A., Somers, M., Camprubi, E., King, H.E., Wolthers, M., de Leeuw, N.H., 2021. Changes in co2 adsorption affinity related to ni doping in fes surfaces: A dft-d3 study. Catalysts 11. https://doi.org/10.3390/catal11040486

Zou, Z., Bertinetti, L., Politi, Y., Fratzl, P., Habraken, W.J.E.M., 2017. Control of Polymorph Selection in Amorphous Calcium Carbonate Crystallization by Poly(Aspartic Acid): Two Different Mechanisms. Small 13, 1603100. https://doi.org/10.1002/smll.201603100

Zou, Z., Polishchuk, I., Bertinetti, L., Pokroy, B., Politi, Y., Fratzl, P., Habraken, W.J.E.M.E.M., 2018. Additives influence the phase behavior of calcium carbonate solution by a cooperative ion-association process. J. Mater. Chem. B 6, 449–457. https://doi.org/10.1039/c7tb03170j

[Koskamp et al., 2021b]:

Janou A. Koskamp, Sergio E. Ruiz Hernandez, Annet Baken and Mariette Wolthers, Chapter 4 of thesis: Influence of biomolecules on amorphous Calcium Carbonate formation: a molecular dynamics investigation (to be submitted)

# Acknowledgements

After blood, sweat and tears we are – finally - at the end of the quest. This is something I could not have achieved without the help, advice, and guidance of the people around me.

First and foremost, I would like to thank my supervisor Mariette Wolthers for giving me the opportunity to do research and for introducing me to the world of computer science and molecular dynamics in particular. Her enthusiasm is contagious, and her extensive knowledge and abundant experience have encouraged me from the first day until the last dot was placed. I greatly appreciate the freedom and confidence she gave me to find my own path and the guidance she offered at the right moments. Under her supervision, I developed not only professionally and beyond my own research (as she encouraged me to join the central works council) but also personally. It was a delight being a member of the diverse team she constructed.

I am thankful to my promoters, Nora de Leeuw and Jack Middleburg for letting me under their scientific wings while giving me the freedom to perform my research. Moreover, I would like to show my appreciation to Nora de Leeuw for her constructive feedback on the work presented in chapters 3, 4, 5 and 6.

I would also like to take a moment to thank my co-supervisor, Amir Raoof. Although his main expertise is on a higher scale than I was focusing on, I appreciated the brainstorming sessions and his energetic personality.

I was also very lucky to have Sergio as my work buddy, who dealt with all the problems that I felt like discussing, and always had a minute to give me constructive feedback and supporting me until the end, which I highly appreciate. Sergio also contributed to my development as a person with lots of insights into history, politics and many other subjects and points of view. I

am grateful you decided to join Mariette's group, leaving Cardiff behind and settle for a while in our beloved little frogland. I hope you have a good time again in Great Britain.

Furthermore, I would like to thank Devis di Tommaso for his scientific contribution, I enjoy working with you and I hope we can do that again during my postdoc.

Priyanka, my PhD partner, we started together as the first two PhD students of Mariette, and we had an instant click. From talking about Bollywood movies to the differences in our cultures to doing yoga in our office in the good old pp9. I had so much fun with you, thank you very much. I learned a lot from you, and I am very happy that I met you. As Priyanka said, a good support system is important to surviving and staying sane during the long journey of (Covid influenced) PhD.

I would like to thank all together, Priyanka, Sergio, Sergej and Alemeh, for being my team members. I am still missing the long lunches we had with the intense discussions on work and non-work-related random topics.

I also want to thank all the past and current members of the geochemistry group for always being so helpful and friendly. Lukas and Theo, I thank you for the nice chats and helping me get everything working on the supercomputer Eejit.

I would also like to show gratitude towards Ian B. Butler, Christine V. Putnis and Helen King, for letting me do experiments in their groups. Although that data did not make it to the thesis, I am still grateful that you allowed me to not lose touch with the experimental side of me.

And a thank you to Nora and Richard's research group in Cardiff for welcoming me and introducing me to the memorable British pub nights.

During part of my PhD, I joined the NWO, former FOM, COR (central work council). I would like to thank all members of the COR and especially Ineke van der Vegt for her warm and cheerful personality always being there to help me out. Also, Sjoerd Wouda for his dedication and for showing me all the in's and out's of being a COR member and guiding the COR in the turbulent years after the merge.

In addition, I would like to take a moment to show my gratitude to the people that were there when I was just a little Janoutje. To my primary school teacher, Kees Groos, who motivated me to learn and enjoy school. And to Els Vissers, who always sees the opportunities rather than the limitations of people and who created the best environment in a primary school you can imagine. Then much later in my career, Hennie Boonen, Robin Willemse and Pieter Iedema, for inspiring me to go from my internship in the industry back to academia and apply for a PhD position.

And of course, I would like to thank my family and friends for all their love and encouragement. For my parents who raised me and supported me in all my pursuits. Thanks to you I became the person I am today, happy with my life and proud of the things I have achieved. My mom, who I can call at any time of the day to ask for advice or to just chat for hours and who gives me the feeling that I am top priority. My dad, who is always there for me, for example, to help me casually moving two times in one and a half years. For my little brothers, Maico, Tjacco, Giano, Dahli and David, although the time is long gone that I could call the first four "little". We do not spend a lot of time together, but when we do, we have all the fun of the world whereafter I feel completely recharged and I feel lucky to have you guys. I also want to thank Irma, sister from another mister (and mother). With her I did not only follow the same career in where we challenged each other continuously, but we also share friends and family time together. I am glad that you want to be my paranymph. I do not want to forget to thank all the friends in my life which I cherish incredible moments in different stages of my PhD and life:

Vera, Rosalie, Vienna, Judith, Suzan, Dan, Tyler, Hannah, and Mel. I would always be thankful for making my life easier, nicer and way more joyful.

Last but not least, I would like to thank you, Josefina, my partner and best friend, for helping me when I just need to keep focus. After disrupting everything in my life when you made your entrance, you brought structure in my life. Even when I thought I was in a never-ending loop, especially during the finalizing stage, you brought structure in my texts. I would never forget your support in the roughest moments and your company in the nicest. Ik hou van jou.

# Curriculum vitae

| | |
|---|---|
| 4 June 1992 | Born in Beugen, The Netherlands |
| 2013- 2015 | Master Analytical Chemistry, Vrije Universiteit Amsterdam |
| 2015-2016 | Research Scientist, Océ Technologies B.V. Venlo |
| 2016-2021 | PhD Geochemistry group, Utrecht University, Utrecht, The Netherlands |
| 2021-present | Postdoctoral Researcher, Geochemistry group, Utrecht University, Utrecht, The Netherlands |