UNIVERSITY COLLEGE LONDON

DEPARTMENT OF CHEMISTRY

Computational Study of the Nucleation of Calcium Phosphate

THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY (PHD), AUTHORED BY:

Giulia MANCARDI

Supervisor:

Prof. Nora H. DE LEEUW

June, 2018

Declaration

I, Giulia Mancardi, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Giulia Mancardi

Abstract

Hydroxyapatite (HA), the main mineral phase of mammalian tooth enamel and bone, originates in body fluids from amorphous calcium phosphate (ACP). The early stage of ACP formation from solution is the object of the investigations presented in this thesis.

I have first performed *ab initio* molecular dynamics simulations to shed light on the structure and stability of the calcium phosphate (CaP) prenucleation clusters, $[Ca(HPO_4)_3]^{4-}$, in aqueous solution. The calcium is seven coordinated by two water molecules, two bidentate phosphates and one monodentate phosphate. Free energy profiles obtained using umbrella sampling simulations show that the complex with a Ca-to-P ratio of 1:3 is the most energetically favoured, and thermodynamically more stable than the free ions.

In order to be able to study a larger system, I have employed shell-model molecular dynamics simulations to investigated the aggregation and clustering of calcium and phosphate ions in water. ACP presents short-range order in the form of small domains with size of 0.9 nm and chemical formula $Ca_9(PO_4)_6$, known as Posner's clusters. Calcium phosphate aggregates form in solution with compositions and Ca coordination that are similar to those found in Posner's cluster, but the stoichiometry of these species is dependent on the ionic composition of the solution: calcium-deficient clusters in solutions with low Ca/P ratio; cluster containing protonated phosphate groups in neutral solutions; sodium ions partially substituting calcium in solutions containing a mixture of sodium and calcium ions. These Posner-like clusters can be connected by phosphate groups, which act as a bridge between their central calcium ions. The simulations of the aggregation in solution of calcium phosphate clusters are an unbiased and unequivocal validation of Posner's model and reveal for the first time the structure and composition of the species that form during the early stages of ACP nucleation at a scale still inaccessible to experiment.

Lastly, I have investigated the heterogeneous nucleation of CaP on a titanium implant. Titanium is commonly employed in orthopaedic and dental surgery due to its good mechanical properties. In order to promote the integration of the metallic implant with the biological tissues, titanium is passivated by a thin layer of oxide and covered by a bioactive material, normally HA; HA can originate on the oxide during a process called biomimetic deposition which consists in soaking the implant in simulated body fluids, that are supersaturated with respect to HA. This method allows to efficiently cover implants with complex shapes and to create a porous coating similar to the bone tissue. Here, I have used molecular dynamics and interatomic potentials to study the deposition of calcium and phosphate species on the titanium dioxide anatase. Different force fields developed for calcium phosphate, titanium dioxide and water were combined and the new parameters were benchmarked against DFT data. Calcium phosphate interaction with the (101) and (100) surfaces of anatase was successfully investigated and the force field here proposed can be used to study the nucleation of calcium phosphate on other titanium dioxide polymorphs and on common surface defects.

Acknowledgement

First of all, I would like to thank my supervisor, Prof. Nora de Leeuw for giving me the opportunity to join her fantastic research group for three years and for her advice and guidance during this time.

A big thank to Dr. Umberto Terranova which taught me how to start doing computational chemistry coming from a completely experimental chemistry background. I am also very grateful to another Italian researcher, Dr. Devis Di Tommaso, with whom I collaborated at Queen Mary University of London, and who gave me scientific ideas as well as confidence in myself as a doctoral student.

And, most importantly, I want to express my gratitude to my family who supported me from Italy giving me the strength to overcome the most discouraging moments.

Contents

List of Publications

- G. Mancardi, U. Terranova, N.H. de Leeuw, "Calcium phosphate prenucleation complexes in water by means of ab initio molecular dynamics simulations", *Crystal Growth & Design*, **2016**, 16 (6), pp.3353-3358. (10.1021/acs.cgd.6b00327)

- G. Mancardi, C.E. Hernandez Tamargo, D. Di Tommaso, N.H. de Leeuw, "Detection of Posner's clusters during calcium phosphate nucleation: a molecular dynamics study", *Journal of Materials Chemistry B*, **2017**, 5, pp.7274-7284. (10.1039/C7TB01199G)
  Invited for preparing the cover page of the same issue

- G. Mancardi, C.E. Hernandez Tamargo, U. Terranova, N.H. de Leeuw, "Calcium phosphate deposition on planar and stepped (101) surfaces of anatase TiO_2", submitted

List of Conferences

Oral Contributions

- Crystal Growth of Organic Materials and British Association for Crystal Growth annual conference (CGOM12/BACG2016), Leeds, 26th-30th June 2016

  Talk title: Calcium Phosphate Prenucleation Complexes in Water by means of ab initio Molecular Dynamics

- Methods to Simulate Nucleation and Growth from Solution (MSNGS), Sheffield, 4th-5th August 2016

  Talk title: Calcium Phosphate Prenucleation Complexes in Water by Means of ab initio Molecular Dynamics Simulations

- Goldschmidt 2017, Paris, 13th-18th August 2017

  Talk title: Detection of Posner's clusters during calcium phosphate nucleation: a molecular dynamics study

- South West Computational Chemists annual meeting 2017 (SWCC2017), Cardiff University, 14th September 2017

  Talk title: Detection of Posner's clusters during calcium phosphate nucleation: a molecular dynamics study

Poster Presentations

- CDT annual industry day, University College London, 9th July 2015

  Poster title: Calcium Phosphate Prenucleation Clusters in Water by Means of ab initio Molecular Dynamics

- Congress of Theoretical Chemists of Latin Expression (CHITEL), Torino (Italy), 26th-31st July 2015

  Poster title: Cluster di prenucleazione del Fosfato di Calcio in Acqua mediante Dinamica Molecolare ab initio (Italian language)

- Fifth European Conference on Crystal Growth (ECCG5), Bologna (Italy), 9th-11th September 2015

  Poster title: Study of the early stages of Calcium Phosphate Nucleation in water by means of ab initio Molecular Dynamics

- Materials Chemistry Consortium (MCC), Cardiff, 6th-8th April 2016

  Poster title: Calcium Phosphate Prenucleation Complexes in Water by means of ab initio Molecular Dynamics

- University College London Departmental Postgraduate poster session, 21st June 2016

  Poster title: Calcium Phosphate Prenucleation Complexes in Water by means of ab initio Molecular Dynamics

- CDT annual industry day, University College London, 6th July 2016

  Poster title: Nucleation of Calcium Phosphate in water at body fluid pH: a Molecular Dynamics study

- CDT annual industry day, University College London, 6th July 2017

  Poster title: Nucleation of Calcium Phosphate in water at body fluid pH: a Molecular Dynamics study

List of Figures

List of Tables

List of Abbreviations

ACP Amorphous Calcium Phosphate

AFM Atomic Force Microscopy

AIMD Ab Initio Molecular Dynamics

AO Atomic Orbital

BCP Biphase mixture of Calcium Phosphate

BSSE Basis Set Superposition Error

CaP Calcium Phosphate

CDHA Calcium Deficient Hydroxyapatite or Dahllite

CIF Crystallographic Information File

CN Coordination Number

CNT Classical Nucleation Theory

CPMD Car-Parrinello Molecular Dynamics

DCACP Dispersion Corrected Atom Centred Potential

DCPA DiCalcium Phosphate Anhydrate or Monetite

DCPD DiCalcium Phosphate Dihydrate or Brushite

DFT Density Functional Theory

DLCA Diffusion Limited Cluster Aggregation

DZ Double-Zeta

ECP Effective Core Potential

FFT Fast Fourier Transforms

FTIR Fourier Transform Infrared Spectroscopy

GAFF Generalized Amber Force Field

GAPW Gaussian and Augmented Plane Wave

GGA Generalised Gradient Approximation

GPW Gaussian-Plane Waves

GTH Goedecker-Teter-Hutter

GTO Gaussian Type Orbital

GULP General Utility Lattice Program

HA Hydroxyapatite

KS Kohn-Sham

LDA Local Density Approximation

LDSAED Low-Dose Selected Area Electron Diffraction

LJ Lennard-Jones

MCPA MonoCalcium Phosphate Anhydrate

MCPM MonoCalcium Phosphate Monohydrate

MD Molecular Dynamics

METADISE Minimum Energy Technique Applied to Dislocations, Interfaces and Surfaces Energies

MO Molecular Orbital

μVT constant chemical potential, Volume and Temperature

NMR Nuclear Magnetic Resonance

NPT constant Number of particles, Pressure and Temperature

NVE constant Number of particles, Volume and Energy

NVT constant Number of particles, Volume and Temperature

OCP OctaCalcium Phosphate

PAW Projector Augmented Wave

PBC Periodic Boundary Conditions

PBE Perdew-Burke-Ernzerhof

PBG Phosphate Based Glasses

PC Posner's cluster

PDB Protein Data Bank

PES Potential Energy Surface

PNC PreNucleation Cluster

PS Pseudopotential

RDF Radial Distribution Function

RLCA Reaction Limited Cluster Aggregation

SAXS Small-Angle Synchrotron X-ray Scattering

SIBLING Small Integrin-Binding-Ligand-N-linked Glycoprotein

SBF Simulated Body Fluid

SM Shell Model

SPC/E Extended Single Point Charge

STO Slater Type Orbital

SV Split Valence

TCP TriCalcium Phosphate

TEM Transmission Electron Microscopy

TPS Trajectory Path Sampling

TTCP TetraCalcium Phosphate

TZ Triple-Zeta

US Umbrella Sampling

VASP Vienna Ab-initio Simulation Package

VMD Visual Molecular Dynamics

WAXS Wide-Angle Synchrotron X-ray Scattering

WHAM Weighted Histogram Analysis Method

XANES X-ray Absorption Near Edge structure Spectroscopy

XRD X-Ray Diffraction

Chapter 1

Introduction

1.1 Calcium Phosphates

Calcium and phosphorus are widely distributed elements on our planet and a combination of oxides of these two elements with or without incorporation of water gives different calcium phosphates. Geologically, calcium phosphates (CaP) are present as fluoroapatite deposits ($Ca_{10}(PO_4)_6F_2$) and phosphorites. In the apatite structures, calcium ions may be partially replaced by Sr, Ba, Mg, K, Na or Fe whereas phosphate ions may be replaced by ${AsO_4}^{3-}$, ${CO_3}^{2-}$ and ${VO_4}^{3-}$; F^- ions can be substituted by OH^-, Cl^-, Br^- and ${CO_3}^{2-}$. The main industrial application of calcium phosphate minerals is in the production of agricultural fertilisers, as calcium and phosphorus are among the macronutrients needed by plants.[19]

Calcium phosphates are the main components of vertebrates bone and tooth tissues, along with collagen and water, and they originate in the *Biomineralisation* process, which means biological formation of minerals by living organisms.[20] Fish enameloid and some species of shells also contain calcium phosphates.[21]

Amorphous calcium phosphate (ACP) is the precursor phase of apatite crystals in bone and tooth tissues.[22] The nucleation of ACP in body fluids is thought to proceed by aggregation of prenucleation clusters (PNCs), identified as calcium triphosphate complexes by some authors[23] or, alternatively, by combination of Posner's clusters.[12,24]

In the bone and tooth tissue, calcium phosphate is mainly present in the form of carbonated hydroxyapatite and confers hardness and mechanical properties to these organs. In

the human body, calcium phosphate is also responsible for pathological crystallisation leading to several common diseases like atherosclerosis, dental caries, osteoporosis and kidney stones. Atherosclerotic plaques are a composite of calcium phosphate and cholesterol; caries originate when part of the apatite in the teeth tissue is replaced by calcium hydrogen phosphate, which is more soluble[21] and some kidney stones are also calcium phosphate aggregates, although the majority is made of calcium oxalate.[25]

Generally, calcium phosphates are more soluble and acidic for lower Ca/P ratio (0.5-1.5), whereas less soluble compositions such as hydroxyapatite have a Ca/P ratio of 1.5-2.[26]

Below there is a list of pure calcium phosphate compounds ordered in increasing Ca/P ratio:

- Monocalcium phosphate monohydrate (MCPM) $Ca(H_2PO_4)_2 \cdot H_2O$, Ca/P=0.5
  Employed as a fertiliser,[27] component of some phosphate cements in medicine, also used as a supplement for baking powders, foods, and beverages.

- Monocalcium phosphate anhydrate (MCPA) $Ca(H_2PO_4)_2$, Ca/P=0.5
  The anhydrous form of MCPM, crystallises above 100°C, it is used as a fertiliser.[27]

- Brushite (DCPD) $Ca(HPO_4) \cdot 2H_2O$, Ca/P=1.0
  Found in some pathological calcifications,[28] used in anti-caries toothpaste and calcium supplements,[29] and as a fertiliser.[27]

- Monetite (DCPA) $Ca(HPO_4)$, Ca/P=1.0
  Anhydrate form of DCPD, crystallises above 80°C. Used in calcium phosphate cements,[30] food supplements, toothpaste.

- Octacalcium phosphate (OCP) $Ca_8(HPO_4)_2(PO_4)_4 \cdot 5H_2O$, Ca/P=1.33
  It is a component of human dental and urinary calculi.[31]

- α-tricalcium phosphate (α-TCP) α-$Ca_3(PO_4)_2$, Ca/P=1.5
  Derives from calcination of β-TCP at temperatures higher than 1125°C, it is used as a fertiliser.[27]

- β-tricalcium phosphate (β-TCP) β-$Ca_3(PO_4)_2$, Ca/P=1.5
  It does not precipitate from solution and it is therefore prepared by high temperature calcination (>800°C), it is used in bone cements.[32]

- Amorphous calcium phosphate (ACP) $Ca_x(PO_4)_y \cdot nH_2O$, 1.2<Ca/P<2.2

  It is the first phase precipitating from aqueous solution and its Ca/P ratio depends on the solution pH: the higher the pH the higher is the Ca/P ratio. ACP shows a short range order, but appears amorphous in X-ray diffraction experiments.[24] It is found in pathological calcifications and it is used in some cements.[33]

- Calcium deficient hydroxyapatite (CDHA) $Ca_{10-x}(HPO_4)_x(PO_4)_{6-x}(OH)_{2-x}$ (0<x<1), Ca/P=1.5-1.67

  It is also called *dahllite*, and it can be seen as a highly substituted hydroxyapatite, with calcium and hydroxy vacancies. Typical substitutions are Na^+, K^+, Sr^{2+} and Mg^{2+} for calcium, carbonate for phosphate and F^-, Cl^- and carbonate for hydroxide. It is the main component of physiological and pathological calcifications.[33]

- Hydroxyapatite (HA) $Ca_{10}(PO_4)_6(OH)_2$, Ca/P=1.67

  It is the most stable and least soluble among the calcium phosphates, pure HA crystallises in the monoclinic space group $P2_1/b$ at a temperature below 250°C and in the hexagonal group $P6_3/m$ above 250°C. Impurities stabilise the hexagonal form which is the most common in nature.[34] It is used as a coating for metallic prostheses and dental implants and in calcium phosphate cements. It can not be used alone for bone and tooth replacements due to its poor mechanical properties.[35]

- Tetracalcium phosphate (TTCP) $Ca_4(PO_4)_2O$, Ca/P=2.0

  It is not stable in water, it is prepared by a solid-state reaction above 1300°C. TTCP is employed in calcium phosphate cements.[30]

1.2 Calcium phosphates in the human body

Biomineralisation began to take place around 570 million years ago when the first animals with skeleton started to appear. Nowadays, the fossils of those organisms are studied to get information on paleotemperature (through $^{18}O/^{16}O$ isotopic ratio) and sea salinity (through the measurement of trace elements).[36]

In the human body, the two major mineralised tissues are bone and tooth. Bone tissue is a composite material of calcium phosphate ($\approx 65\%$ weight), collagen and other proteins. Biological apatite and molecules of type I collagen form the mineralised bone matrix, where the needle-like crystals insert themselves in a parallel way into the collagen fibrils and are elongated along the crystallographic *c*-axis.

Collagen is a protein made of about 1000 amino acids, which assumes a characteristic triple helix secondary structure (see Figure 1.1). Collagen contains a high number of glycine (Gly) residues, which are regularly spaced in the backbone in the sequence:

$$Gly - X - Y \tag{1.1}$$

where *X* and *Y* are typically proline (Pro) and 4-hydroxyproline (Hyp) but can vary along the protein. The wrapping of the three helices originates the *tropocollagen*, which is a compact super helix having 3.3 residues per turn; the interaction among the three helices is stabilised by hydrogen bonds, covalent bonds, bridging water molecules and steric locking. Collagen *fibrils* originate from the replication of the superhelix in space, whose not perfectly periodic arrangement create hole zones and grooves where calcium phosphate nucleate and grow. Collagen fibrils organise in *sheets*, which are arranged together with bone cells into cylindrical structures called *osteons*. Osteons are stacked to form *haversian* bone which is the elementary unit of both spongy bone and compact bone[36] (see Figure 1.2 taken from Ref.[2] for the hierarchic structure of bone and Figure 1.3 from Ref.[3] for the macroscopic structure of a long bone) The mineralisation of the collagen matrix is assisted by Small Integrin-Binding-Ligand-N-linked Glycoprotein (SIBLING);[37] among them, bone sialoprotein and phosphophoryn are calcium phosphate nucleating agents whereas osteopontin is an inhibitor.[38] Bone HA is situated in the groove areas of the collagen, where it forms small crystals which assume a plate or needle-like shape, with their *c*-axis parallel to the protein helix axis; furthermore, HA crystals have the same

periodicity of the collagen fibres.[39] The carbonate content in bone HA increases with the age of the subject from ≈4% to ≈8% at the expense of hydrogen phosphate.[40]

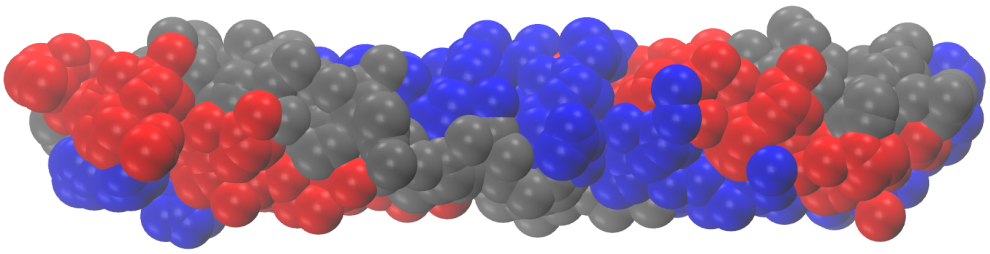


Figure 1.1 Representation of a collagen type I molecule, the coordinates are taken from RSCB protein database (PDB ID 1CGD) and the structure is rendered using VMD.[1] The three helices are represented in red, blue and grey.

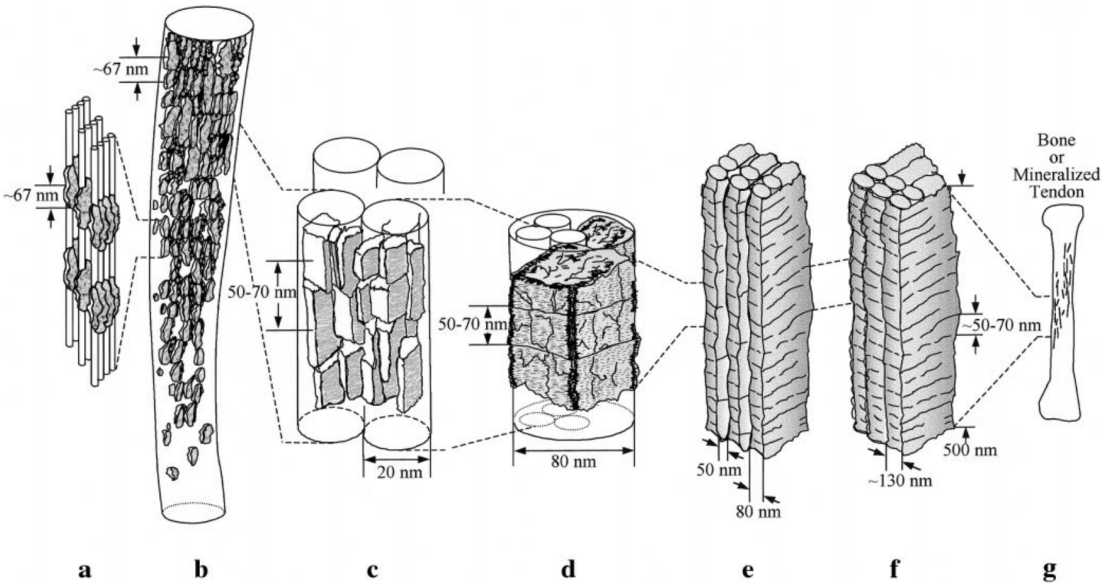


Figure 1.2 Hierarchical structure of mineralised bone: (a) collagen fibrils with HA crystals in the grooves, (b) HA crystals growing along their c-axis, (c-f) partial fusion of small crystals into plates and growth, (g) the final lamellar structure of bone (or mineralised tendon). Taken from Ref.[2]

The human skeleton is formed of 206 bones, which can be classified according to their shape in flat (e.g.: sternum), long (e.g.: femur), sesamoid (e.g.: knee patella), irregular (e.g.: vertebrae) and short (e.g.: cuneiform). The central part of long bones (see Figure 1.3) is called diaphysis and contains a medullary cavity, made of compact bone, which hosts the yellow bone marrow, that is made of fat cells and works mainly as energy storage. The extremities of long bones are called epiphyses and are made of spongy or can-

cellous bone which pores are filled with red bone marrow; the red marrow consists mainly of hematopoietic tissue and participates in the production of several blood components. Between the diaphysis and the epiphyses there is a region called metaphysis which is the bone growth region active during childhood and puberty. Flat bones contain a core of spongy bone surrounded by compact bone. Compact bone is dense and provides support, whereas spongy bone is lighter and its trabeculae are aligned along the stress lines to give strength to the tissue.[41]

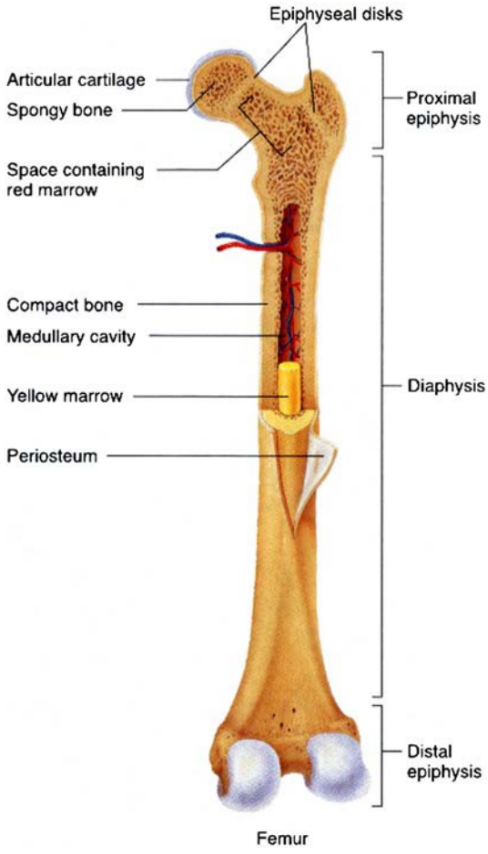


Figure 1.3 Representation of a femur bone, from Ref.[3]

Bone tissue hosts four kinds of cells: osteoblasts, osteoclasts, osteocytes and osteogenic cells. Osteogenic cells originate osteoblasts, which function is to develop new bone tissue, before becoming osteocytes when trapped in the calcified matrix; osteoclasts resorb old bone. Bones remodel themselves continuously,[42] in a process controlled by many factors such as age, nutrition, physical exercise and hormones. The physiological regulation of the bone remodelling process can be disrupted in diseases like osteoporosis which occurs when the rate of bone resorption is higher than the rate of new bone formation, resulting

in a decrease in bone density, which raises the probability of fractures. Osteoporosis is common in women after menopause when oestrogen, one of the hormones involved in bone remodelling, decreases.[43] The formation of HA takes place inside matrix vesicles which accumulate calcium and phosphate ions promoting the precipitation of the mineral phase. The vesicles containing HA are then released into the extracellular fluid by the osteocytes, liberating the HA crystals, which start to bind the collagen fibrils.[44]

Teeth (see Figure 1.4) are enclosed in the gums and kept in place by the periodontal ligament. The portion of tooth emerging from the gums is called crown, whereas the other part is named root; the core of the tooth is called pulp cavity and consists of connective tissue protecting nerves and blood vessels.[41] The crown consists of two different biominerals: enamel (outside) and dentin (inside), both containing calcium phosphate. The enamel is considered as the most resistant biological material, it is much harder than bone, as it contains $\approx 95\%$ of hydroxyapatite which crystals are bigger than those of bone. It does not contain living cells and therefore it is not capable of regenerating itself.[40] Dentin is softer and its composition resembles that of bones. Fluoride ions are incorporated in the HA lattice and reduce the solubility of the mineral.[21] In the root, the dentin layer is covered by cementum, which is another bone-like material.[41] Bone and dentine form on a collagenous matrix whereas enamel originates by crystallisation of HA on non-collagenous proteins. Odontoblasts and ameloblasts are responsible for teeth mineralisation.[21] Tooth decay is a mechanical and chemical erosion which leads to the thinning of the enamel layer; the common cavities are caused by the acid products of bacteria's metabolism of sugar, which dissolve the calcium phosphate crystals in the enamel.

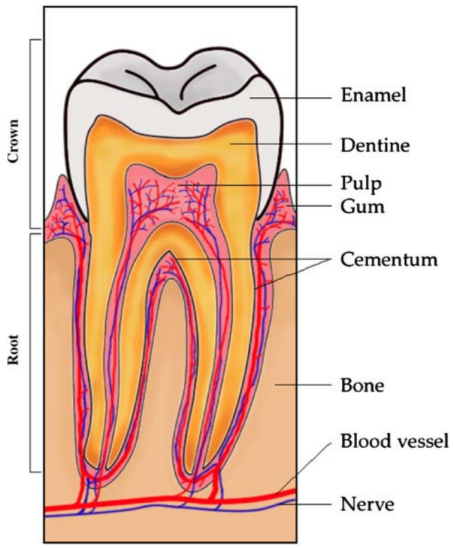


Figure 1.4 Representation of a tooth, from Ref.[3]

Body fluids are supersaturated with respect to HA, but ectopic mineralisation is inhibited *in vivo* by several biomolecules. Pathological mineralisation in the form of vascular calcifications is a process regulated by vascular smooth muscle cells instead of bone cells and the mineralised material is very similar in composition to bone.[45] The process starts with the accumulation of proteoglycans which bind low-density lipoproteins at the lesioned site. Inflammatory signals induce death of the smooth muscle cells of the blood vessels which start to release vesicles containing HA crystals; the latter act as crystallisation seeds. Kidney stones are a common disease related to pathological mineralisation, where calcium oxalate, calcium phosphates (apatite or brushite) and other minerals form abnormal precipitates in the kidney, due to metabolic and environmental-nutritional circumstances.[25]

Because of the ubiquity of calcium phosphates in the human body, this material is often used as a biomaterial, particularly for hard-tissue regeneration. Bone HA is very difficult to prepare in the laboratory due to its non-stoichiometry, therefore other types of calcium phosphates are commonly employed in reconstructive surgery. Calcium phosphates are bioactive ceramics, which means that they are biocompatible and easily integrated into bone, as they release calcium and phosphate ions to form the new bone tissue when placed in contact with the body fluids. The bioactivity of pure crystalline HA has been enhanced by preparing substituted apatites, which are more similar to dahllite: carbonate and stron-

tium make the material more soluble and easily resorbable, whereas silicon substituting part of the phosphorus atoms enhance bioactivity and mechanical strength.[40] Another option to fill bone cavities is to employ calcium phosphate cements, which are injected into the site in paste form and become solids thereafter, thus having the advantages, over the bioceramics, of not having to be prepared to match the sometimes complex shape of the defect and of being injectable, so that invasive surgical procedures are not needed.[46] The first commercialised cement is the Constantz's cement which is prepared from a dry mixture of α-TCP, MCPM and calcium carbonate, to which liquid sodium monoacid phosphate is added to make the paste.[47] This cement becomes solid in just 10 minutes from the injection and is gradually replaced by new bone tissue. Calcium phosphate cements are also used to deliver different drugs (antibiotics, anti-inflammatory, analgesic, anticancer and growth factors) to the healing site to promote the formation of new bone tissue and to prevent infections.[48] This is facilitated by the ability of calcium phosphates to adsorb various chemicals on their surface.[46] Biphase mixtures of calcium phosphates (BCP) have been also considered for bone replacement, in particular, the mixture of HA and β-TCP evolves to carbonate apatite in physiological conditions.[49,50] Some studies have focused on organic-calcium phosphate composites, where the organic matrix can be gelatin, collagen[51] or a biodegradable polymer such as poly(L-lactide).[52] Other widely studied materials for bone replacement are phosphate-based glasses (PBGs), which are amorphous and have composition $CaO{-}Na_2O{-}P_2O_5$. They are employed as templates to repair soft and hard tissues and for drug delivery; the molar ratio of the components determine their bioactivity and dissolution rate.[53] PBGs do not contain silica and thus differ in composition from the original Bioglass developed by Hench in 1971, which was containing 45% SiO_2, 24.5% Na_2O, 24.5% CaO and 6% P_2O_5.[54]

To make calcium phosphate materials osteoinductive (*i.e.* able to induce new bone formation), it is important to synthesise them with appropriate geometry and porosity to facilitate their dissolution in body fluids or combining them with growth factors or bioactive proteins such as collagen.[55]

Unfortunately, the poor mechanical properties of calcium phosphate materials limit their applications to non-load bearing materials, such as coatings on metallic prostheses, cements for healing small bone defects, nano-powders for spinal fusion and matrices for

drug delivery.[56]

1.3 Titanium dioxide as a biomaterial

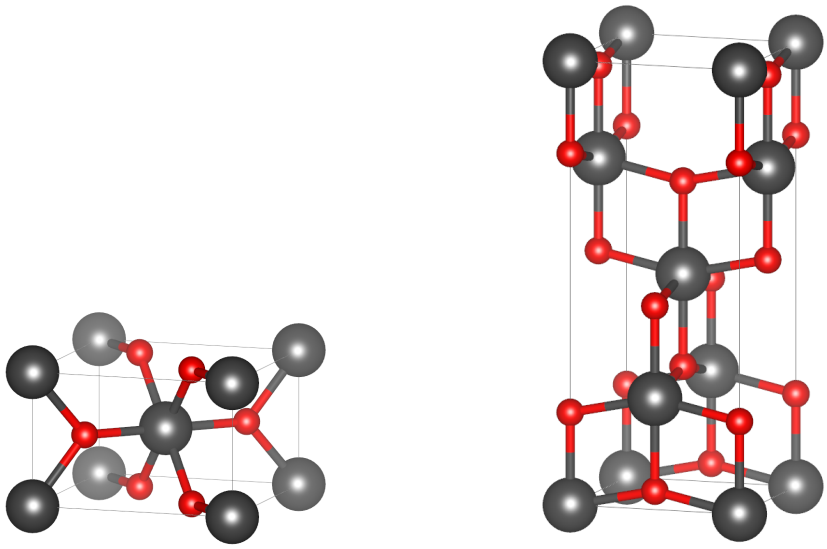
Calcium phosphates are often used as coatings on titanium and titanium alloys implants, which are otherwise bioinert but have the required mechanical strength to be employed in load-bearing applications such as hip replacements.[57] HA and BCP are also used as abrasives to roughen the surface of metal implants. Titanium alloys are the most corrosion-resistant material currently used for implants.[58] Other advantages are the ability to bond to the HA coating,[59] the low density and the similarity between their coefficient of thermal expansion (9-10 $\times 10^{-6}$ °C), and that of HA (12 $\times 10^{-6}$ °C).[60] If not coated by a bioactive material, the titanium implant is encapsulated by a fibrous tissue, which causes looseness of the prosthesis with consequent needing of a new surgical operation. The HA coating allows optimal integration with the bone at the site of the implant, thus minimising adverse reactions due to metallic ions release into the body.[35] Porous HA is commonly deposited as a 40-200 μm thick layer to the metallic surface using the plasma spraying technique. Thicker coatings have poor mechanical properties whereas too thin layers are resorbed too fast; moreover, the porosity of HA favours the integration of the implant with bones. Plasma spraying consists in ionising a stream of mixed gases through a direct current electric arc generated by two electrodes; HA powders are suspended in the generated plasma and sent towards the surface of the metallic implant.[61] Calcium phosphate coatings can also be prepared through electrophoretic deposition,[62] sol–gel deposition,[63] electrochemical deposition,[64] sputtering[65] and biomimetic deposition.[66] Other calcium phosphate materials such as bioglasses are as well employed as coatings for titanium implants.[60]

Commercially pure Titanium is used in dentistry,[67] whereas the alloy Ti-6Al-4V, which contains 5.5-6.75 wt.% of Al and 3.5-4.5 wt.% of V, is chosen for bone implants due to its better mechanical properties and corrosion resistance over pure Ti.[68]

To improve the corrosion resistance of titanium, and therefore the durability of the implant, the surface is commonly passivated by a thin layer of TiO_2; passivation is achieved employing a series of techniques (sol-gel, anodisation and thermal oxidation) and the

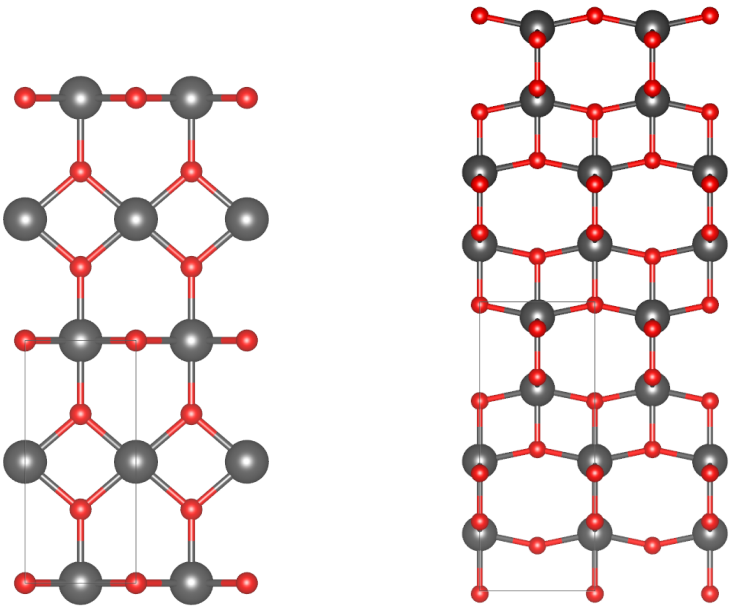
thicker the layer the better the protection against corrosion. Moreover, the oxide layer bonds stronger to HA than the bare metal and protects the implant surface also after the HA coating has been dissolved.[63]

The three titanium dioxide polymorphs are rutile (tetragonal, space group D_{4h}^{14}-$P4_2/mnm$, a=b=4.584Å, c=2.953Å),[69] anatase (tetragonal, space group D_{4h}^{19}-$I4_1/amd$, a=b=3.782Å, c=9.502Å) and brookite (rhombohedral, space group D_{2h}^{15}-Pbca, a=5.436Å, b=9.166Å, c=5.135Å); other crystalline phases can as well be synthesized in extreme temperature/pressure conditions. Only rutile and anatase are commonly employed as biomaterials, in both Ti coordinates six oxygen atoms in a distorted octahedral arrangement, the degree of distortion is higher in anatase.[70] The unit cells of rutile and anatase are shown in Figure 1.5.



(a) **(b)**

Figure 1.5 (a) TiO_2 rutile unit cell, (b) TiO_2 anatase unit cell. Colour key: Ti:grey, O:red.



(a) **(b)**

Figure 1.6 (a) TiO_2 rutile (110) 2×2×1 surface, view from the top; (b) TiO_2 anatase (101) 2×2×1 surface, view from the top. Colour key: Ti:grey, O:red.

Anatase is 1.2-2.8 kcal/mol less stable than rutile and the phase transition occurs between 700 °C and 1000 °C; the most thermodynamically stable surfaces are the (110) for rutile and the (101) for anatase, the latter has the lowest surface energy,[71] so that the most thermodynamically stable phase for small particles (< 14 nm) is anatase.[72] Top views of rutile (110) and anatase (101) surfaces are depicted in Figure 1.6.

Depending on the technique used to passivate the titanium surface, a different polymorph can be obtained: rutile is formed through thermal oxidation, whereas anatase originates from conventional anodic oxidation in solution;[73] the last method allows also to incorporate impurities ions such as calcium and phosphates, which enhance cell growth.[74]

1.3.1 Step-edges on titanium dioxide anatase

Metal oxides surface reactivity is influenced by the presence of defects.[70] Monoatomic step-edges on anatase (101) surface are common defects which have a characteristic trapezoidal shape.[75] The parallel sides of the trapezoidal islands are oriented along the [010] direction, whereas the lateral sides are oriented along the $[\bar{1}11]$ or the $[11\bar{1}]$ directions. Figure 1.7 shows an STM image of step-edges on anatase (101) surface and their preferred

orientation.

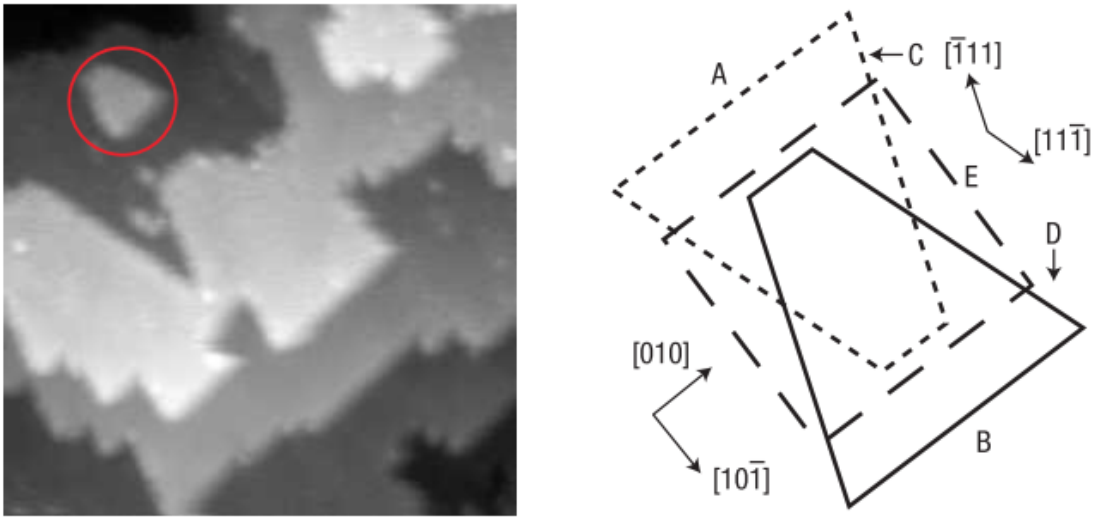


Figure 1.7 STM image showing the preferential orientation of monoatomic step-edges on an anatase (101) surface and a representation of the possible step-edges orientations. The trapezoidal island depicted with a bold line represents the most energetically favoured step-edge. The five possible step orientations are labelled A-E. The picture is taken from the article of Gong *et al.*[4]

A step can be seen as a narrow slice of the corresponding facet; the lowest step formation energy scales with the surface energy of the corresponding extended surface, and B and D present the lowest formation energy, meaning that the trapezoid island depicted in bold in Figure 1.7 is the one observed in the experiment. Contrariwise, steps E and therefore rectangular islands are never observed.[4,76] Due to the importance of step-edges in the reactivity of anatase, we investigated the adsorption of calcium and phosphate on the (100) surface of anatase, which is the most thermodynamically stable surface after the (101)[77] and is the corresponding facet of the step-edge type B in Figure 1.7. A complete explanation of the steps simulation is reported in chapter 5.

1.4 Classical Nucleation Theory

Classical Nucleation Theory (CNT) is the simplest and most widely used theory that describes the nucleation process. This theory was originally derived for condensation of a vapour into a liquid, but it is used also to explain precipitation of crystals from supersatu-

rated solutions and melts. The thermodynamic description of this process was developed at the end of 19th century by Gibbs, who defined the free energy change required for cluster formation (ΔG) as sum of the free energy change for the phase transformation (ΔG_v) and the free energy change for the formation of a surface (ΔG_s). A schematic representation is shown in Figure 1.8:

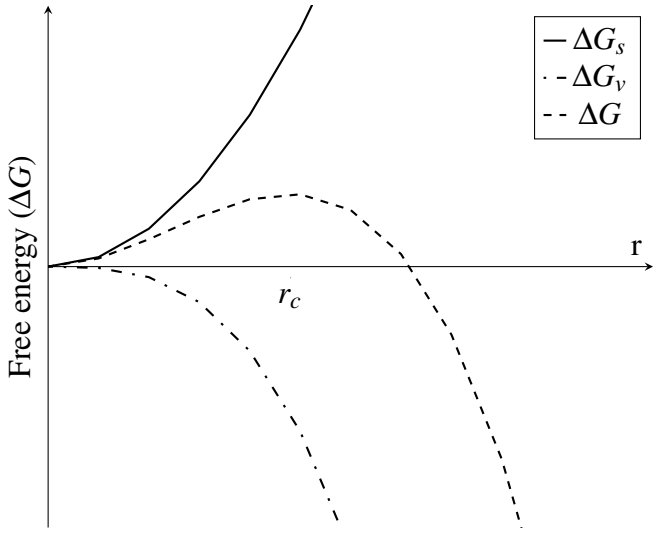


Figure 1.8 Free Energy diagram for nucleation according to CNT

Since the solid state is more stable than the liquid, ΔG_v becomes negative and thus decreases the Gibbs free energy of the system. On the other hand, introduction of a solid/liquid interface increases the free energy by an amount proportional to the surface area of the cluster. As a result, the growth of clusters depends on the competition between a decrease in ΔG_v, which favours growth, and an increase in ΔG_s which favours dissolution. The positive surface free energy ΔG_s term dominates at small radii. As cluster size increases, the total free energy reaches a maximum at a critical size (r_c), above which the total free energy decreases and growth becomes energetically favourable, resulting in the formation of crystal nuclei.

According to CNT,[78] nuclei (defined as the minimum amount of a new phase capable of independent existence) form in supersaturated solutions as a consequence of microscopic density fluctuations. These fluctuations occur by random collisions of the dissolved constituents and may be seen as monomer association in pseudo-equilibrium. CNT is based on five assumptions:

1. Droplet model: the clusters are modelled as spherical droplets with density equal to

the macroscopic density of the bulk condensed phase.

2. Capillarity approximation: the surface tension of a liquid droplet is equal to the respective value of this quantity for a stable coexistence of both phases at an infinite planar interface.

3. The growth of clusters takes place by addition of one monomer at a time.

4. Steady-state kinetics: the nucleation rate is time-independent.

5. The clusters are incompressible and the vapour surrounding them is an ideal gas with constant pressure.

Unfortunately, in many cases, CNT fails to predict nucleation rates, due to the simplifications outlined above. Furthermore, CNT is only valid for homogeneous nucleation and it assumes that clusters evolve in size attaching a single molecule at a time, thus neglecting the collision between two clusters. Theoretical and experimental studies show that the nucleation may involve the assembly of pre-formed clusters, and that clusters can exist even in undersaturated solutions (whilst CNT is valid only for supersaturated solutions).[78,79]

1.5 Nucleation of Calcium Phosphates

A $Ca_9(PO_4)_6$ cluster, now known as Posner's cluster, was proposed by Posner and Betts in the 1970s as a possible model for the structural unit of ACP.[12,24] They obtained the X-ray radial distribution for ACP and the results suggested the existence of ordered domains with diameter $\approx$ 9.5 Å and Ca/P ratio of 1.5. These clusters are also identifiable in the structure of crystalline HA. On the *ab* plane of HA, a Posner's cluster overlaps with the vicinal clusters, sharing two phosphate groups; a hexagonal packing of Posner's clusters leads to the HA framework, two OH groups and one Ca are incorporated in the voids among the clusters, forming positive and negative columns parallel to the *c* axis[80] (see Figure 1.9). The HA structure contains two distinct Posner's clusters with different chirality, the aggregation of three Posner's clusters with the same chirality around the positive column leads to the formation of another Posner's cluster of opposite chirality.[81] Calcium ions in HA are named Ca1 or Ca2 according to their coordination environment. There

are 4 Ca1 in the HA unit cell, they coordinate nine oxygen from phosphates whereas the remaining 6 Ca2 coordinate six oxygen from phosphates and one from a hydroxy group. Ca1 are stacked along the *c*-direction, Ca2 form hexagonal channels along the same direction.[82] Each Ca1 site is a centre of a Posner's cluster; along the *c* direction, a cluster centred on Ca1 at level $c = 0$ shares half of its volume with the cluster on the top (centred on a Ca1 at level $c = +1/2$) and half volume with the cluster on the bottom (centred on a Ca1 at level $c = -1/2$).[81]

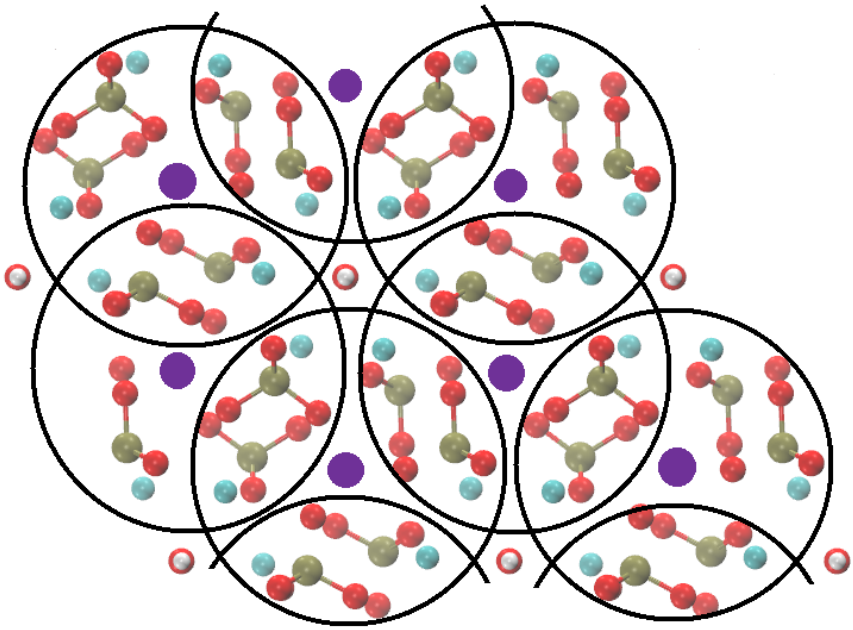


Figure 1.9 Schematic representation of the hydroxyapatite crystal projected on the *ab* plane. Posner's clusters sharing part of their volume are circled in black. Colour key: Ca1:purple, Ca2:light blue, P:green, O:red, H:white

In the past, it was believed that the ACP formation in solution involved the formation of Posner's clusters which then aggregated randomly to produce larger spherical particles, with the intercluster space filled with water. Hydroxyapatite originates from ACP in a process of solution and renucleation.[24]

More recently, a cluster-growth model for the formation of ACP has been proposed: the first step is the aggregation of *prenucleation clusters* (PNCs), which are already present in body fluids before nucleation.[22] To form the unit cell of hydroxyapatite, two OH species and a Ca atom must be incorporated in the voids between clusters.[80] Although *Classical Nucleation Theory* has been revised and expanded in the past years to contemplate the

involvement of prenucleation clusters in the nucleation path,[83] it still appears as a too simplified model to describe the real process.

Experimental[22,83–85] and computational[23,86] studies evidence the role of prenucleation clusters in the nucleation of several materials. In terms of energy, these clusters can be metastable or stable with respect to the free ions from which originate.

The nucleation rate J can be expressed as:

$$J = A \cdot exp(-E_A/k_BT) \cdot exp(-\Delta G_{ex}/k_BT) \tag{1.2}$$

where the first exponential term refers to the kinetic barrier and the second one to the thermodynamic barrier. The kinetic effects are difficult to quantify and therefore are typically neglected, whereas the ΔG_{ex} in the thermodynamic term is calculated using the capillarity approximation in the CNT (*i.e.* the small clusters have the same bulk structure and interfacial energy of the corresponding macroscopic phase). In the CNT the expression for ΔG_{ex} is:

$$\Delta G_{ex}(S) \sim \alpha\gamma^3\phi^{-2} \tag{1.3}$$

where γ is the size-independent interfacial free energy, α is a shape factor and ϕ is an affinity term depending on ion activities and the solubility. In order to take into account the two possibilities (stable and metastable clusters) Hu *et al.*[83] added a term C in the ΔG_{ex} expression:

$$\Delta G_{ex}(S) \sim \gamma^3(\phi \pm C)^{-2} \tag{1.4}$$

where C depends on the shape, size and free energy of the clusters. The sign is + in the case of metastable clusters and - in the case of stable clusters.

Regarding calcium phosphate, small clusters have been observed in simulated body fluids, as well as in solutions undersaturated with respect to ACP and OCP, but supersaturated with respect to HA.[80]

XANES and XRD experiments of the early stage of calcium phosphate crystallisation have suggested that an idealised cluster with formula $Ca_9(PO_4)_6(H_2O)_{30}$ may act as the structural unit of ACP,[87] whereas the most abundant clusters detected in solution at the early stage of nucleation have been reported to be in the form $Ca(\eta^2-PO_4{}^{3-})_2L_2$ (L=H_2O or $\eta^1-PO_4{}^{3-}$, with η^1 and η^2 indicating monodentate and bidentate binding of the phos-

phate groups).[88] The idealised cluster appears to have the same stoichiometry of a solvated Posner's cluster and a Ca/P ratio of 1.5, as in ACP.

Habraken *et al.*[23] studied the biomimetic precipitation of calcium phosphate in a buffered solution under a constant ionic strength using several *in situ* analysis techniques. The experiments reveal that PNCs are in fact ion-association complexes, with formula $[Ca(HPO_4)_3]^{4-}$, which aggregate in solution to form polymeric structures. Above the solubility limit of ACP, they take up calcium ions from solution to form insoluble *postnucleation clusters* with formula $[Ca_2(HPO_4)_3]^{2-}$ that precipitate as ACP. The aggregation of the complexes into polymeric assemblies is driven by the gain in entropy associated with the release of hydration water. Continued calcium uptake converts ACP into octacalcium phosphate and subsequently into apatite, which both contain the calcium triphosphate complex as their basic structural unit. Interestingly, the Posner's cluster can be seen as two deprotonated prenucleation complexes in which all negative charges are compensated by complexing calcium ions. According to the authors, the calcium triphosphate complexes are metastable with respect to the free ions, they base this assumption on the fractality of the assemblies, which suggests an RLCA (reaction limited) and not DLCA (diffusion limited) aggregation process. The above consideration implies that the aggregating entities possess a surface energy, in agreement with the CNT assumptions (i.e. capillarity approximation). The existence of prenucleation clusters, which have an excess free energy, alters the nucleation pathway by making amorphous phases accessible at concentrations for which classical nucleation theory would predict the exclusive formation of the more stable crystalline phases. By the contrary, the formation of stable clusters may seem odd in the nucleation process, because they make the nucleation thermodynamic barrier higher; however stable prenucleation clusters have been already observed in materials like calcium carbonate. The authors[84,86] state that the prenucleation clusters have "solute character", which means that hydration energy and solvent effects are important, but the clusters do not have the surface tension contemplated by the classical theory. Moreover, the dehydration of ions to form clusters results in a state that is more similar to the nucleated phase (in our case ACP), lowering the kinetic activation barrier for nucleation with respect to the dissociated ions.[79]

The occurrence of stable solute species in homogeneous solutions has been evidenced for

calcium phosphates, calcium carbonate,[84,86] iron oxides and hydroxides[89] and silica.[90] The prenucleation clusters have the following characteristics:[79]

1. PNCs have the same chemical composition of the forming solid, but can also contain additional chemical species.
2. PNCs are small and thermodynamically stable and there is no phase boundary between them and the solution.
3. PNCs are molecular precursors to the phase nucleation from solution and hence participate in the process of phase separation.
4. PNCs are dynamic entities, so they change configuration on timescales typical for molecular rearrangements in solution.
5. PNCs can have motifs resembling one of the corresponding crystalline polymorphs.

The experimental technique used in literature to study the calcium phosphate nucleation process are mainly:[23,88]

- Cryogenic transmission electron microscopy analysis (cryo-TEM) allows to obtain images with nanometric resolution and therefore to follow the morphological changes of the material in the early stages of formation;
- Atomic Force Microscopy (AFM);
- Low-dose selected area electron diffraction (LDSAED) gives information about the short and long-range structural order;
- Wide-angle synchrotron X-ray scattering (WAXS) and small-angle synchrotron X-ray scattering (SAXS) give structural information;
- Fourier transform infrared spectroscopy (FTIR) provides structural information and allows to see the chemical transformations at each stage;
- X-ray absorption near edge structure spectroscopy (XANES) gives detailed information on the coordination environment of an atom (e.g. calcium);
- Analytical techniques like titration, dilution experiments and pH monitoring are employed to evaluate the chemical ratio between components (e.g. Ca/P ratio).

1.6 Research questions and objectives of the work

This thesis is aimed at unravelling the mechanism of calcium phosphate nucleation during the Biomineralisation process which takes place in the human body, using computational chemistry tools. Theoretical simulations are a successful way to complement the experimental techniques which capacity at the atomic scale is limited. Many approximations have been done in this work in order to get significant results from computer simulations of the system:

- Body fluids are complex solutions containing ions and proteins; only water molecules were used in the simulations (computationally less expensive).
- In order to consider the electronic effects, *ab initio* methods must be employed (see chapter 3), and this leads to a decrease of the system size that can be treated.
- Increasing the size of the system under investigation requires some approximations: MD and interatomic potentials were used to significantly reduce the computational time in chapters 4 and 5.
- Strongly supersaturated solutions were simulated to observe calcium phosphate aggregation in a reasonable amount of computing time, moreover, a physiological concentration would require a prohibitively large system size.

The main *research questions* are

- Which is the structure of the calcium phosphate prenucleation clusters observed in the experiment;
- What is the pattern of homogeneous nucleation of calcium phosphate in aqueous solution;
- How the heterogeneous nucleation of calcium phosphate takes place on titanium dioxide anatase, which is passivating titanium implants used in orthopaedic procedures.

Chapter 2

Computational Methods

Because the nucleation event takes place at atomic scale, computer simulation is a very powerful tool to assist the interpretation of experimental data. Molecular simulations are capable to provide a mechanistic understanding at the atomic level and also allow to measure kinetic and thermodynamic quantities.
The Computational Chemistry Codes used in this work are:

- CP2K versions 2.4, 2.6 and 3.0,[91–101] for DFT and *ab initio* Molecular Dynamics
- DL_POLY version 4.07,[102] for classical Molecular Dynamics
- Vienna Ab-initio Simulation Package (VASP) version 5.4.1,[103–106] for DFT
- General Utility Lattice Program (GULP) version 4.4,[107–109] for optimisation of bulk and surfaces using interatomic potentials
- Minimum Energy Technique Applied to Dislocations, Interfaces and Surfaces Energies (METADISE) version 5.64,[110] for cutting surfaces according to the Miller indexes

In the following pages, the theory behind the computational methods used is described briefly, with a reference to the chapters where the specific tool is employed.

2.1 The Schrödinger equation

The fundamental postulate of quantum mechanics states that to every chemical system is associated a wave function Ψ, and appropriate operators acting on Ψ return the observable properties of the system. The product of the wave function with its complex conjugate (i.e. $|\Psi^*\Psi|$) is a density of probability; when considering *real* wave functions, the integral of $|\Psi|^2$ gives the probability of the system to be found in a specific region of space. The Hamiltonian operator H is the operator that, acting on the wave function, returns the system energy E:

$$H\Psi = E\Psi \tag{2.1}$$

which is the time-independent Schrödinger equation.[111] The Hamiltonian is made of five contributions to the total energy of the system: (i) the kinetic energy of the electrons, (ii) the kinetic energy of the nuclei, (iii) the electrons-nuclei attraction, (iv) the electron-electron repulsion and (v) the repulsion between the nuclei:

$$H = -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k}\nabla_k^2 - \sum_i \sum_k \frac{e^2 Z_k}{r_{ik}} + \sum_{i<j} \frac{e^2}{r_{ij}} + \sum_{k<l} \frac{e^2 Z_k Z_l}{r_{kl}} \tag{2.2}$$

where $\hbar$ is Planck's constant divided by 2π, i and j refer to electrons, k and l to nuclei, m_e is the mass of the electron and e its charge, m_k is the mass of the nucleus and Z its atomic number, ∇^2 is the Laplacian operator and r is a distance between particles. Ψ is a function of 3N coordinates, where N is the total number of particles in the system.

2.2 The Born-Oppenheimer approximation

Nuclei are much heavier than electrons ($m_k >> m_e$), so it is reasonably possible to consider the nuclei fixed when the electrons move. This means that electrons move in a static potential generated by the nuclei, whereas nuclei are affected by an average potential generated by electrons. It is, therefore, possible to decouple electronic and nuclear motions and compute the electronic energies for fixed nuclear positions; the term (ii) of equation 2.2 is now constant for a given configuration. The Schrödinger equation becomes:

$$(H_{elec} + V_N)\Psi_{elec}(q_i;q_k) = E_{elec}\Psi_{elec}(q_i;q_k) \tag{2.3}$$

where H_{elec} is the electronic Hamiltonian, which includes terms (i), (iii) and (iv) of equation 2.2, V_N is the nucleus-nucleus repulsion energy (term (v) in equation 2.2) and is a constant for a given set of fixed nuclear coordinates, q_i are the electronic coordinates and q_k are the nuclear coordinates, which are included as parameters. The eigenvalue of the electronic Schrödinger equation, E_{elec}, is the electronic energy and depends on the coordinates of the nuclei.[112] The approximation of separating electronic and nuclear motions is called the Born-Oppenheimer approximation;[113] according to it, nuclei move in a potential energy surface (PES) which is the solution of the electronic Schrödinger equation (equation 2.3).

2.3 Density Functional Theory

Unfortunately, the methods based on the wave function, such as the Hartree-Fock, become too expensive in terms of computer time when dealing with solids and other large systems: the electronic wave function of an n-electrons molecule depends on $3n$ spatial and n spin coordinates. Density Functional Theory (DFT) is nowadays one of the most used theoretical approximations in quantum chemistry, thanks to its versatility in terms of systems that can be treated (molecules and solids), and its high accuracy. DFT allows calculating the electronic properties of the ground state knowing the electron density.
In the present work, DFT has been used in chapter 3 to compute the relative energies of different calcium triphosphate structures and as a method to compute the electronic energy and forces at each *ab initio* Molecular Dynamics step. The DFT optimisation of calcium and phosphate ions on a TiO_2 surface was performed in chapter 5 to benchmark the ion-surface interactions computed with interatomic potentials.

2.3.1 Hohenberg and Kohn's Theorems

In 1964, Hohenberg and Kohn proved that for molecules with a nondegenerate ground state, the ground state molecular energy, wave function and all other electronic properties are uniquely determined by the electronic density of the ground state $\rho(x,y,z)$.[114,115]

This theorem also proves that the ground state electron density ρ_0 determines the external potential V_{ext} and the number of electrons. Therefore ρ_0 defines H_{elec}, the energy and the other properties. The ground state energy E_0 is a functional of $\rho(x,y,z)$:

$$E_0 = E_0[\rho_0(x,y,z)] \tag{2.4}$$

where the square brackets stand for a functional. A functional $F[f]$ is a rule that associates a number with each function $f(x)$.[115] The electronic Hamiltonian $\hat{H}_{elec}$ is given by:

$$\hat{H}_{elec} = -\frac{1}{2}\sum_{i}^{N}\nabla_i^2 + \sum_{i=1}^{N}V_{ext} + \sum_{i}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} \tag{2.5}$$

with

$$V_{ext} = \sum_{A}^{M}\frac{Z_A}{r_{iA}} \tag{2.6}$$

V_{ext} depends only on the coordinates of the electron i and on the coordinates of the nuclei (which are constant in the Born-Oppenheimer approximation). V_{ext} is called external potential because it is generated by charges external to the electronic system. Each of the terms in equation 2.5 is determined by the ρ_0. Known ρ_0, V_{ext} (term (ii)) is known, but the kinetic energy term (i) and the electron-electron repulsion term (iii) remain unknown. The energy functional of the electronic density is given by:

$$E[\rho(r)] = \int V_{ext}\rho(r)dr + F[\rho(r)] \tag{2.7}$$

where $F[\rho(r)]$ is independent of V_{ext} and includes the kinetic energy and electron-electron repulsion energy.

A way to determine ρ_0 is given by the variational theorem of Hohenberg and Kohn which states that the real ground state electronic density corresponds to the minimum of the energy functional $E[\rho(r)]$. The way to calculate ρ_0 was found by Kohn and Sham one year after.

The electronic density has some important properties:

- $\rho(r)$ is positive
- $\rho(r)$ depends on the spatial variables: $\rho(x,y,z)$
- $\rho(r)$ is zero at an infinite distance
- integrated gives the total number of electrons in the system $\int \rho(r)dr = N$

- its value can be obtained through X-ray diffraction experiments
- $\rho(r)$ has a singularity in correspondence of the nuclei

2.3.2 Kohn-Sham Method

In order to calculate ρ_0, Kohn and Sham considered a fictitious system of non-interacting electrons, affected by the same V_{ext} of the original system, so that the density of the fictitious system is equal to that of the real system ($\rho_s(r) = \rho_0(r)$, $\rho_s(r)$ is the electron density of the fictitious system).[116] Hohenberg and Kohn demonstrated that ρ_0 determines V_{ext} so, knowing ρ_s, $V_{ext}[\rho_s(r)]$ is determined. Because the electrons do not interact in the fictitious system, the Hamiltonian $\hat{H}_s$ is:

$$\hat{H}_s = \sum_{i}^{N}[-\frac{1}{2}\nabla_i^2 + V_{ext}] \equiv \sum_{i=1}^{N}\hat{h}_i^{KS} \tag{2.8}$$

where $\hat{h}_i^{KS}$ is the one-electron Kohn-Sham Hamiltonian. The Hamiltonian of the fictitious system can be related to the one of the real system according to the following relation:

$$\hat{H}_\lambda = T + \sum_i V_\lambda(r_i) + \lambda V_{ee}; 0 \leq \lambda \leq 1 \tag{2.9}$$

where T is the kinetic energy and V_{ee} is the electron-electron repulsion terms; λ is equal to zero for the non-interacting electrons of the fictitious system and it is equal to one for the real system. V_λ is the external potential which equates the electronic density of the system with Hamiltonian $\hat{H}_\lambda$ to that of the ground state of the real system.

The ground state wavefunction of the fictitious system, $\Psi_{s,0}$, is the Slater determinant of the lowest energy Kohn Sham spin-orbitals u_i^{KS} of the fictitious system, where the spatial part $\theta_i^{KS}(r_i)$ of each spin-orbital is an eigenfunction of the operator $\hat{h}_i^{KS}$:

$$\hat{h}_i^{KS}\theta_i^{KS} = \varepsilon_i^{KS}\theta_i^{KS} \tag{2.10}$$

where ε_i^{KS} are the Kohn-Sham orbital energies.

Now we define the difference in ground state electronic kinetic energies between real and fictitious systems as:

$$\Delta T[\rho] \equiv T[\rho] - T_s[\rho] \tag{2.11}$$

and the difference in electronic repulsion in the two systems as:

$$\Delta V_{ee}[\rho] \equiv V_{ee}[\rho] - \frac{1}{2}\iint \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 \tag{2.12}$$

where r_{12} is the distance between two points with electronic densities ρ_1 and ρ_2. At this point it is possible to rewrite the original Hohenberg and Kohn formula (2.4) as follows:

$$\begin{aligned} E_0 = E[\rho_0] &= \int \rho_0(r)V(r)dr + T[\rho_0] + V_{ee}[\rho_0] \\ &= \int \rho(r)V(r)dr + T_s[\rho_0] + \frac{1}{2}\iint \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 + \Delta T[\rho] + \Delta V_{ee}[\rho] \end{aligned} \tag{2.13}$$

the functionals $\Delta T[\rho]$ and $\Delta V_{ee}[\rho]$ are unknown and they constitute the exchange-correlation energy E_{xc}:

$$E_{xc} \equiv \Delta T[\rho] + \Delta V_{ee}[\rho] \tag{2.14}$$

substituting equation 2.14 in 2.13 we obtain:

$$E_0 = E[\rho_0] = \int \rho(r)V(r)dr + T_s[\rho_0] + \frac{1}{2}\iint \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 + E_{xc}[\rho] \tag{2.15}$$

The first three terms of equation 2.15 are obtainable knowing the electron density, which is given by the formula:

$$\rho = \rho_s = \sum_{i=1}^{occ} |\theta_i^{KS}|^2 \tag{2.16}$$

so the equation 2.15 becomes:

$$\begin{aligned} E_0 = -\sum_A Z_A \int \frac{\rho(r_1)}{r_{1A}} dr_1 - \frac{1}{2}\sum_i^{occ} < \theta_i^{KS}|\nabla_1^2|\theta_i^{KS} > \\ + \frac{1}{2}\iint \frac{\rho(r_1)\rho(r_2)}{r_{12}} dr_1 dr_2 + E_{xc}[\rho] \end{aligned} \tag{2.17}$$

The Kohn-Sham orbitals are obtained according to the variational theorem of Hohenberg and Kohn: given an approximate electronic density ρ' which integrated in the configurational space gives the number of electrons, $\int \rho'(r)dr = N$, the energy obtained applying the exact functional to the approximate density is larger than or equal to the exact density,

$$E_0[\rho'] \geq E_0[\rho_0] \tag{2.18}$$

the ground state energy is obtained varying the Kohn-Sham orbitals θ_i^{KS}, but maintaining their orthonormality, to minimise the functional $E[\rho]$. The electronic density which

corresponds to the lowest energy is obtained solving equations formally similar to the equations of Fock:

$$[-\frac{1}{2}\nabla_1^2 - \sum_A \frac{Z_A}{r_{1A}} + \int \frac{\rho(r_2)}{r_{12}} dr_2 + V_{xc}]\theta_i^{KS} = \varepsilon_i^{KS}\theta_i^{KS} \tag{2.19}$$

$$\hat{h}_i^{KS}\theta_i^{KS} = \varepsilon_i^{KS}\theta_i^{KS} \tag{2.20}$$

the exchange-correlation functional V_{xc} is obtained as:

$$V_{xc}(r) \equiv \frac{\delta E_{xc}[\rho(r)]}{\delta\rho(r)} \tag{2.21}$$

however, the functional $E_{xc}[\rho]$ is unknown, and various approximations are made in DFT calculations.[115] In particular, there are three kinds of Exchange-correlation functionals:

- LDA (Local Density Approximation), or local functionals
- GGA (Generalised Gradient Approximation), or semi-local functionals
- Hybrid functionals

2.3.3 The Generalised Gradient Approximation (GGA)

GGA exchange-correlation functionals allow a more accurate description of the system, because they include the gradient in the description of the exchange-correlation energy E_{xc}:

$$E_{xc}^{GGA}[\rho^\alpha\rho^\beta] = \int f(\rho^\alpha(r), \rho^\beta(r), \nabla\rho^\alpha(r), \nabla\rho^\beta(r))dr \tag{2.22}$$

E_{xc}^{GGA} is split into exchange and correlation parts:

$$E_{xc}^{GGA} = E_x^{GGA} + E_c^{GGA} \tag{2.23}$$

In this work, the nonempirical GGA functional of Perdew-Burke-Ernzerhof[100] (PBE) has been used.

2.3.4 Basis Sets

The Basis set is a set of mathematical functions which reproduce the wave function of the system. The choice of the basis set is, therefore, one of the first steps for running a

calculation. The basis set functions describe the Molecular Orbitals (MOs), the MOs are fully described by a complete basis set. However, a complete basis set is computationally demanding, and finite basis sets are used. In a finite basis set only the components of the MO along the coordinate axes corresponding to the selected basis functions can be represented. The accuracy of the representation depends on the size of the basis set and on the type of basis functions used. Important examples of types of basis functions are:

- Slater-Type Orbitals (STO)[117]
- Gaussian-Type Orbitals (GTO)[118]
- Plane waves (see section 2.3.5.2 for details)

Slater-type orbitals have the following functional form:

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{n-1}e^{-\zeta r} \tag{2.24}$$

where N is a normalisation constant and $Y_{l,m}$ are spherical harmonic functions. These orbitals are physically correct, as they reproduce the exact orbitals for the hydrogen atom, so they are used for calculating the atomic functions; however, three- and four- centre two-electron integrals are not possible to calculate analytically.

Gaussian-type orbitals have the form:

$$\chi_{\zeta,n,l,m}(r,\theta,\varphi) = NY_{l,m}(\theta,\varphi)r^{2n-2-l}e^{-\zeta r^2} \tag{2.25}$$

and in cartesian coordinates:

$$\chi_{\zeta,n,l,m}(x,y,z) = Nx^{l_x}y^{l_y}z^{l_z}e^{-\zeta r^2} \tag{2.26}$$

where $l_x + l_y + l_z$ determines the type of orbital (0:s-orbital, 1:p-orbital, etc.) GTO are not physically adequate for two reasons: (1) they present a maximum (and not a cusp) at the nucleus position and (2) the term e^{-r^2} is making the function to fall off too rapidly far from the nucleus.

Despite the aforementioned problems, most of the calculation software utilise a linear combination of GTO, because the product of two gaussian functions is still a gaussian function and the integrals of gaussian functions are analytical and fast.

There is a classification of basis sets based on the number of functions used:

- Minimum basis set: contains the smallest number of functions, consists of one function for each inner shell and valence-shell atomic orbital (AO) of each atom
- Double-zeta (DZ): each AO is represented by two functions with different ζ
- Triple zeta (TZ): each AO is represented by three functions
- Split valence (SV): corresponds to a minimum basis set for the core orbitals and a DZ for the valence orbitals
- Polarized (P): functions with a quantum number higher than the one corresponding to the highest occupied orbital are added to the basis set (p functions on hydrogen, d on elements of the second line of the periodic table, f on transition metals). When the polarisation concerns the second line elements is indicated with a *, when it concerns hydrogen with **.

In chapter 3 we employed a DZVP basis set (Double Zeta Valence Polarized) for the atoms Ca, P, O and H, both for DFT optimisation and for *ab initio* Molecular Dynamics, its characteristics (two GTO for each core orbital, DZ for valence orbitals plus a polarisation function) mean a good compromise between accuracy and calculation speed.

2.3.5 Periodic Solids

The periodic systems are usually described as repetition of the fundamental unit cell in one, two or three dimensions. The unit cell in the direct space is described by three vectors named $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$. The same unit cell in the reciprocal space is defined by the vectors $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ which derive from the direct cell vectors and obey the orthonormality condition $\mathbf{a}_i\mathbf{b}_j = 2\pi\delta_{ij}$. The reciprocal space vectors are computed as follows:

$$\mathbf{b}_1 = 2\pi\frac{\mathbf{a}_2 \times \mathbf{a}_3}{V^3}; \mathbf{b}_2 = 2\pi\frac{\mathbf{a}_3 \times \mathbf{a}_1}{V^3}; \mathbf{b}_3 = 2\pi\frac{\mathbf{a}_1 \times \mathbf{a}_2}{V^3} \tag{2.27}$$

where V is the volume of the cell in the direct space. The first Brillouin zone is the equivalent of a unit cell in the reciprocal space. Points in the direct space are described by r vectors, and points in the reciprocal by k vectors, which are called *wave vectors*.[119]

2.3.5.1 Bloch's theorem

An infinite expansion of the unit cell means that there is an infinite number of electrons in the system. Bloch's theorem allows describing the whole system considering a finite number of electrons, simplifying considerably the problem. The Bloch's theorem states that the values of the wave function ϕ at equivalent positions in different cells are related by a phase factor:[120]

$$\phi(r+T) = e^{ik\cdot T}\phi(r) \tag{2.28}$$

where T is the lattice vector in the direct space and i is the imaginary unit. The same theorem defines a *crystalline orbital* $\phi_{n,k}(r)$ for the nth band in the unit cell as formed by a plane wave-like part and a function φ called *Bloch orbital* which has the same periodicity of the crystal:

$$\phi_{n,k}(r) = e^{ik\cdot r}\varphi_{n,k}(r) \tag{2.29}$$

Applying the monoelectronic Hamiltonian $\hat{h}$ to the crystalline orbital it is possible to obtain energy values:

$$\hat{h}\phi_k(r) = \varepsilon_k\phi_k(r) \tag{2.30}$$

the eigenvalue energy ε_k varies continuously with respect to k, this means that it is necessary to calculate the energy only in the k points with special symmetry, included in the first Brillouin zone.

2.3.5.2 Plane Waves

The periodic function within the Bloch orbital $\varphi_{n,k}(r)$ can be expressed as a Fourier series:

$$\varphi_{n,k}(r) = \sum_G c_{k,G}e^{iG\cdot r} \tag{2.31}$$

where c corresponds to the plane-wave expansion coefficient and G is the reciprocal lattice vector, which is linked to the direct space vector T by the relation $e^{iG\cdot T} = 1$. The Bloch orbital function $\phi_{n,k}(r)$ can be rewritten as:

$$\phi_{n,k}(r) = \sum_G c_{k,G}e^{i(k+G)r} \tag{2.32}$$

which means that the Bloch orbitals can be expanded as a sum of plane waves weighted by the coefficients $c_{k,G}$. In the practice, the infinite series is truncated choosing a kinetic energy cutoff for the plane waves.

The advantage of plane-waves over the gaussian basis set are:

- plane-waves are independent of the position of the nuclei;
- there is no Basis Set Superposition Error (BSSE);
- only one parameter controls the basis set size (kinetic energy cutoff);
- plane-waves are orthogonal;
- plane-waves are numerical efficient using the Fast Fourier Transforms (FFT);
- plane-waves are naturally periodic

however, with plane-waves, it is necessary to use pseudopotentials (see section 2.3.6) and there is a loss of chemical insight.

2.3.5.3 Gaussian-Plane Waves method (GPW)

The advantage and disadvantage of using plane-waves and pure gaussian basis sets are modulated in the Gaussian-Plane Waves (GPW)[94] approach employed in the QUICKSTEP DFT method of the software CP2K.[91–101] The GPW method can be used in gas and condensed phases and for *ab initio* Molecular Dynamics, where the computational cost is linearly scaled with the system size, allowing the use of large basis set also for big systems.

In the GPW method, the wave functions are described by atom-centred Gaussian-type basis, but the density is described by an auxiliary plane-wave basis. The plane-waves representing the density allows using the FFT to efficiently solve the Poisson equation and obtain the Hartree energy. With the GPW it is still necessary to use pseudopotentials, whereas the Gaussian and augmented-plane-wave (GAPW)[121] method allows for all-electron calculations. In chapter 3 we employed CP2K QUICKSTEP method, which makes use of the GPW approach, for performing both DFT optimisations in gas phase and *ab initio* Molecular Dynamics.

2.3.6 The Pseudopotential Approximation

A plane-wave basis set of a reasonable size is not able to describe the behaviour of the core electrons and the singularity in the nuclear potential (see 2.3.4), therefore it is necessary to use pseudopotentials (PS) (also called Effective Core Potentials (ECP)). The PS are analytical functions which describe the core electrons up to a distance from the nucleus given by the *core radius,* r_c. The PS function and its first and second derivatives have to match those of the all-electron wave function at r_c. Depending on the value of r_c, pseudopotentials can be classified as

- *hard* when r_c is small;
- *soft* when r_c is large

hard PS require a larger number of plane wave basis functions, whereas *soft* PS require a smaller number of plane-waves, as more electrons are described by the PS; this means that *soft* PS are less computationally expensive but also less transferable.[119]
According to the properties that PS function have to match with the real function, PS can be classified as

- *norm-conserving*: require that the wave function norm is conserved in the range 0-r_c;[122]
- *ultrasoft*: are not norm-conserving, use larger r_c thanks to the addition of augmentation functions which compensate for the valence charge density in the core region.[123]

PS fail to describe phenomena in which core electrons play an important role, such as X-ray photoelectron spectroscopy and NMR spin-spin coupling constants. Another method that is usually listed in the PS, although it considers all the core electrons, is the *Projector Augmented Wave* (PAW).[124,125] The PAW wave function contains two terms: a valence term expanded in the plane-wave basis set and a term describing the region close to the nucleus, which is evaluated on a grid and accounts for all-electron properties of the wave function, such as its nodal features and orthogonality to the core states.

The pseudopotential parameters usually have to be optimised according to the DFT functional used. In chapter 3 we have used the Goedecker-Teter-Hutter (GTH) pseudopotentials[101] for all the atoms considered; these were taken from the CP2K potential library, which provides GTH potentials already optimised for the use with the PBE functional and a plane-wave basis set.

2.4 Molecular Dynamics

Molecular Dynamics (MD) allows to simulate the behaviour of a system in time; a specific temperature and pressure can also be set. An MD simulation consists in propagating the initial set of coordinates and velocities of the system in a series of time steps. The output of the run is a trajectory which contains the positions and the velocities of the atoms at each time step. The atoms are free to move according to Newton's second law:

$$\frac{d^2x_i}{dt^2} = \frac{F_{x_i}}{m_i} \tag{2.33}$$

which describes the motion of a particle with mass m_i along the coordinate x_i under the effect of the force F_{x_i}. When the forces acting on each atom are calculated using DFT, the MD is called *ab initio* Molecular Dynamics (AIMD).

In the present work, the acronym MD in chapter 3 refers to Born Oppenheimer Molecular Dynamics, where nuclei are classical particles and their behaviour obeys to Newton's second law. In chapter 4, MD refers to classical Molecular Dynamics, where the forces acting on the atoms are calculated through interatomic potentials (see section 2.5.3)

An alternative to AIMD is to employ the Car-Parrinello Molecular Dynamics (CPMD) method[126] where the wavefunctions are not required to fully converge at each time step as they do in AIMD, and the starting wave functions coefficients are given in the CPMD input as fictitious masses.

A different type of simulation is the Monte Carlo method, which consists in (I) perturbating the starting coordinates of the system, (II) calculating the energy of the new system, (III) if the energy is lower than the starting point, repeat I and II, if the energy is higher, the new geometry is accepted only if the Boltzmann factor is lower than a random number between 0 and 1.[119]

2.4.1 The Verlet Algorithm

In a system containing many atoms, the motions of all the particles are coupled together, and the equations of motion are integrated using a finite difference method, which allows to break down the integration in stages separated in time by a δt. The dynamic properties of the system can be expanded as Taylor series:[127]

$$x(t+\delta t) = x(t) + \delta t v(t) + \frac{1}{2}\delta t^2 a(t) + \frac{1}{6}\delta t^3 b(t) + \frac{1}{24}\delta t^4 c(t) + ... \tag{2.34}$$

$$v(t+\delta t) = v(t) + \delta t a(t) + \frac{1}{2}\delta t^2 b(t) + \frac{1}{6}\delta t^3 c(t) + ... \tag{2.35}$$

$$a(t+\delta t) = a(t) + \delta t b(t) + \frac{1}{2}\delta t^2 c(t) + ... \tag{2.36}$$

$$b(t+\delta t) = b(t) + \delta t c(t) + ... \tag{2.37}$$

where x is the position, v is the velocity (first derivative of the positions with respect to the time), a is the acceleration (second derivative of the positions with respect to the time), b is the third derivative and c is the fourth derivative. The Verlet algorithm[128] is one of the most used to integrate equations of motions in MD simulations. The new positions of the particles $x(t+\delta t)$ are obtained using the positions at time t and the positions from the previous step $x(t-\delta t)$:

$$x(t+\delta t) = x(t) + \delta t v(t) + \frac{1}{2}\delta t^2 a(t) + ... \tag{2.38}$$

$$x(t-\delta t) = x(t) - \delta t v(t) + \frac{1}{2}\delta t^2 a(t) - ... \tag{2.39}$$

combining the equations 2.38 and 2.39:

$$x(t+\delta t) = 2x(t) - x(t-\delta t) + \delta t^2 a(t) \tag{2.40}$$

The acceleration is then given by

$$a(t) = \frac{F}{m} = -\frac{1}{m}\frac{dV}{dx} \tag{2.41}$$

where V is the scalar potential function; the velocity is calculated

$$v(t) = [x(t+\delta t) - x(t-\delta t)]/2\delta t \tag{2.42}$$

The limitations of the Verlet algorithm are that (i) the velocity does not appear explicitly in formula 2.40, so velocities are not very simple to calculate; (ii) Verlet is not a self-starting algorithm.

To overcome these problems, in this work we employed an integration algorithm called velocity Verlet,[129] which is implemented in the software we used for MD simulations: CP2K for AIMD (chapter 3) and DL_POLY for classical MD (chapters 4 and 5). Positions, velocities and accelerations are obtained directly from equations 2.43 and 2.46 for every MD step.

$$x(t+\delta t) = x(t) + \delta t v(t) + \frac{1}{2}\delta t^2 a(t) \tag{2.43}$$

$$v(t+\frac{1}{2}\delta t) = v(t) + \frac{1}{2}\delta t a(t) \tag{2.44}$$

The new forces $F(t+\delta t)$ are computed as:

$$ma(t+\delta t) = F(t+\delta t) \tag{2.45}$$

$$v(t+\delta t) = v(t) + \frac{1}{2}\delta t[a(t) + a(t+\delta t)] \tag{2.46}$$

the velocity is updated in two stages (equations 2.44 and 2.46).

2.4.2 Ensembles

Depending on which properties are of interest, MD simulations can be performed in different ensembles:

- Microcanonical ensemble (NVE)

  number of particles, volume and total energy are constant.

- Canonical ensemble (NVT)

  number of particles, volume and temperature are constant. The total energy is not constant because the system is allowed to exchange energy with a heat bath.[130]

- Isothermal-isobaric ensemble (NPT)

  number of particles, pressure and temperature are constant. The NPT ensemble better reproduces the experimental conditions, as chemical reactions are usually performed at a constant pressure and the total energy is unlikely to be constant.

- Grand canonical (μVT)

  temperature, volume and chemical potential are constant. It represents a system that can exchange particles and energy with a reservoir.[130]

There is only a limited number of possible ensemble because not all the thermodynamic quantities are independent, for example it is not possible to fix simultaneously volume and pressure. Since the natural ensemble for performing MD simulations is the Micro-canonical, in order to run simulations in a different ensemble, it is necessary to employ a thermostat (to regulate the temperature) and/or a barostat (to regulate the pressure).

2.4.2.1 Nosé-Hoover thermostat

A thermostat consists essentially in a coupling of the system with a heat bath which exchanges energy with the system at specific intervals of time specified by the time constant. In this work, the *Nosé-Hoover thermostat*[131,132] has been used every time a temperature control was necessary (*i.e.* in simulations in NVT and NPT ensembles). In this thermostat, a fictitious variable η is introduced to change the velocity of the particles in order to maintain the target temperature. The Newton'sequation of motion 2.33 becomes:

$$\frac{d^2x_i}{dt^2} = \frac{F_{x_i}}{m_i} - \eta v_i \tag{2.47}$$

$$\frac{d\eta(t)}{dt} = \frac{2E_{kin}(t) - 2\sigma}{Q} \tag{2.48}$$

$$Q = 2\sigma\tau_T^2 \tag{2.49}$$

where η is the fictitious variable introduced by Nosé, which has a physical meaning of a friction, Q is associated to η and represents the "mass" of the thermostat, σ is the target thermostat energy, T is the target temperature and τ is the time constant. Our main reason for using Nosé-Hoover thermostat is that it is employable both in NVT and NPT ensembles, allowing to keep the simulations consistent when switching to a different ensemble.

2.4.2.2 Nosé-Hoover barostat

We employed the Nosé-Hoover barostat implemented in the DL_POLY software, where are implemented the modifications introduced by Melchionna *et al.*[133] and Martyna *et*

al.[134] for coupling the thermostat and the barostat.

$$\frac{d^2x_i}{dt^2} = \frac{F_{x_i}}{m_i} - \left[\chi(t) + \left(1 + \frac{3}{f}\right)\eta(t)\right] v_i \tag{2.50}$$

$$\frac{d\chi(t)}{dt} = \frac{2E_{kin}(t) + b\eta(t)^2 - 2\sigma - k_B T_{ext}}{Q} \tag{2.51}$$

$$\frac{d\eta(t)}{dt} = 3V(t)\frac{P(t) - P_{ext}}{b} + 3\frac{2E_{kin(t)}}{f}\frac{1}{b} - \chi(t)\eta(t) \tag{2.52}$$

$$b = (f+3)k_B T_{ext}\tau_b^2 \tag{2.53}$$

$$\frac{dH(t)}{dt} = \eta(t)H(t) \tag{2.54}$$

$$\frac{dV(t)}{dt} = [3\eta(t)]V(t) \tag{2.55}$$

where η is the barostat friction coefficient, $\chi(t)$ is the centre of mass of the system at time t, Q is the thermostat mass given in equation 2.49, b is the barostat mass, τ_b is the time constant for pressure fluctuations, P is the instantaneous pressure, V is the volume of the system and H is the cell matrix.

2.4.3 Radial Distribution Functions

The term Radial Distribution Function (RDF) usually refers to the pair distribution function, which measures the probability of finding a particle x as a function of distance from a particle y, with respect to the ideal gas distribution:

$$g(r,\Delta r) = \frac{V}{N}\frac{\langle N(r,\Delta r)\rangle}{4\pi r^2 \Delta r} \tag{2.56}$$

$N(r,\Delta r)$ is the number of particle x within a spherical shell of thickness Δr at a distance r from the particle y. The RDF is a useful tool in MD simulations, as it provides distances and bond lengths values averaged in time, thus allowing to easily compare values from different simulations, from static optimisation and from experiments (X-ray diffraction). RDFs are particularly useful to obtain information from simulations of liquids, as they provide a picture of average distances between particles in a constantly moving system. Although RDFs are usually computed up to a cutoff distance from particle x, from equation 2.56 it is possible to see that at an infinite distance from particle x, $g(r_\infty) = 1$, as the ratio $\frac{\langle N(r,\Delta r)\rangle}{4\pi r^2 \Delta r}$ becomes equal to $\frac{N}{V}$.

Assuming the pairwise additivity of the forces, some thermodynamic properties can be calculated integrating the RDFs, but in most cases, they are obtained directly from the simulation.

2.5 Born model of solids

In this work, we make use of the Born model of solids, which describe the interactions among ions in a periodic system as composed by short- and long-range interactions. In this simplification, the potential energy $V(x)$ becomes:

$$V(x_1,x_2,...,x_n) = \frac{1}{4\pi\varepsilon_0}\sum\frac{q_1q_2}{x_{12}} + V_{short}(x_1,x_2,...,x_n) \tag{2.57}$$

The first term of equation 2.57 describes the long-range interactions as a sum of the coulombic interactions between all the ion pairs, whereas the second term describes the short-range interactions.

2.5.1 Periodic Boundary Conditions

Periodic Boundary Conditions (PBC) allow representing a large system with a relatively small box, which is replicated in space, so that the atoms close to the borders of the box are in contact with a replica and do not dissolve into the vacuum. For representing a crystalline solid, it is sufficient a small box, by the contrary, the correct simulation of amorphous solids and liquids requires a large box, in order to avoid the introduction of a long-range order which does not agree with reality. Different boxes are used depending on the shape of the system in analysis: cubic, orthorhombic, hexagonal prism, truncated octahedron, rhombic dodecahedron and the elongated dodecahedron.[135]

In this work, we employed a cubic box for AIMD simulations of a single calcium phosphate prenucleation cluster (chapter 3), the choice was dictated by the necessity of having the smallest system possible for reducing the computational cost, but at the same time the PNC of the central box had to keep at “safe" distance from the PNCs in the neighbour cells. We used again a cubic box for classical MD simulations of calcium and phosphate aggregation (chapter 4), with each atom as far as possible from its periodic image, in order to limit the errors due to the introduction of an unphysical long-range order, intrinsic to

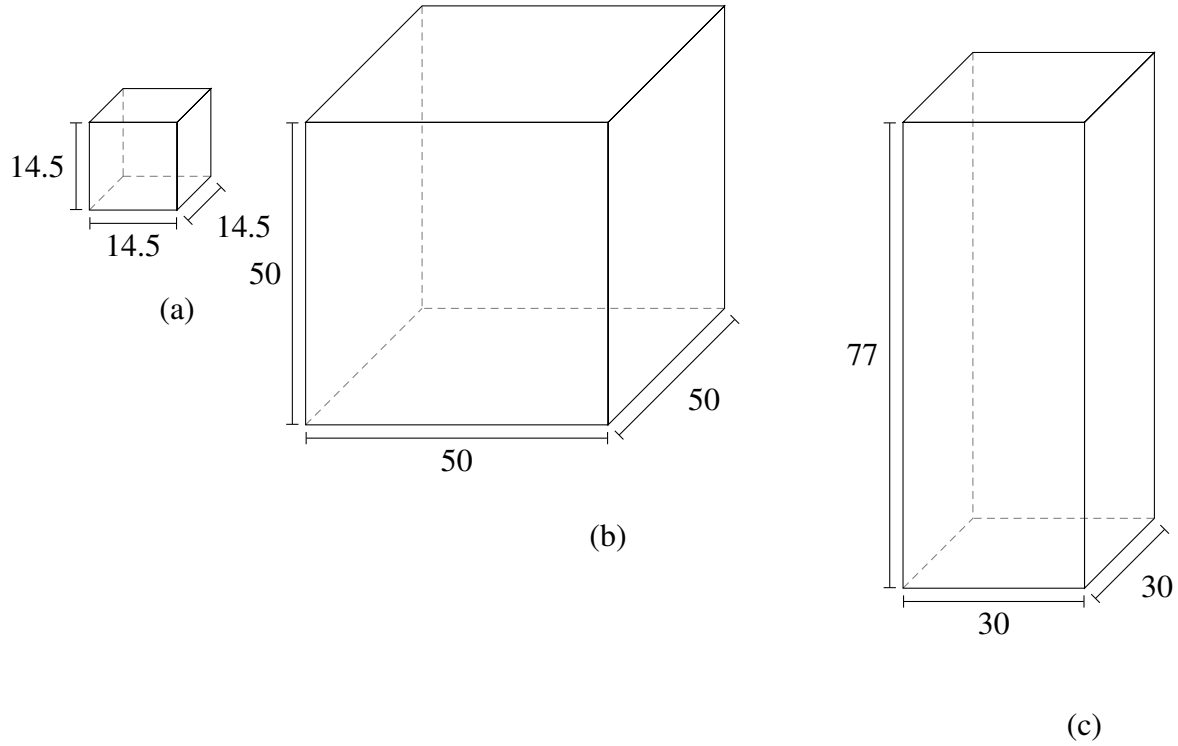


Figure 2.1 Boxes used in this work, all the measures are in Å.

the PBC conditions. In chapter 5, we employed an orthorhombic box in order to maximise the space between two TiO_2 slabs and minimise their unphysical interaction, this means we had to elongate the *c*-axis of the box. Figure 2.1 illustrates the dimensions of the (a) cubic box used in chapter 3, the (b) cubic box used in chapter 4 and the (c) orthorhombic box used in chapter 5.

2.5.2 The Ewald summation

The Ewald method[136] consists in splitting the long-range interaction in equation 2.57, which is a conditionally convergent series, in two series, one calculated in the real space and the other in the reciprocal space, so that both converge quickly. This is achieved by surrounding each ion by a spherical Gaussian having opposite charge and then superimposing a second Gaussian having the same charge of the central ion. The short-range coulombic interaction between the point ions and their first Gaussian functions are computed in the real space, whereas the superimposed Gaussians potential is computed in the reciprocal space.

2.5.3 Interatomic Potentials

Classical MD makes use of interatomic potentials, that are mathematical functions which describe the interactions among the particles in the system. This approximation allows simulating large systems at a much lower computational cost when compared to AIMD; the Schrödinger equation does not have to be solved and the dynamics of the atoms obey to Newton's second law. The main steps of developing a new force field are:

1. Select properties to use for the parametrisation (density, dipole moment, bond length...)
2. Find data on chosen properties from experiment/ab initio calculations
3. Chose the functional form for the force field
4. Chose a procedure for obtaining the parameter (trial and error/use least-square fitting[137])

Developing new parameters requires a lot of time and resources, and the parameters fitted on a specific system may not be used for another system. For organic molecules, several highly transferable force fields such as the Generalized Amber Force Field (GAFF)[138] have been developed; for inorganic systems, the transferability is more difficult to achieve because metal ions show a wide range of coordination and geometries.
An atom which has been assigned a set of parameters is called *atom type* and its parameters can be transferred to another system where the same atom is in a similar environment. In this work, we have employed a force field that had been parametrised to reproduce the structural and mechanical properties of the orthorhombic phase of phosphorus pentoxide P_2O_5,[5] to perform molecular simulations of phosphate-based glasses. The fitting was made on both experimental and *ab initio* data. This force field describes the ions polarisation through the shell model (SM) (see section 2.5.4); the use of formal charges makes the parameters easily transferable to other system containing the PO_4 unit, such as amorphous calcium phosphate. This force field is based on the Born model of solids (see section 2.5), the interactions between ions are expressed by Buckingham potentials and electrostatic energy; for negative oxygen ions, both coulombic and short-range forces act on the shells, whereas only coulombic forces act on the cores.

2.5.3.1 Short Ranged (van der Waals) Potentials

The van der Waals energy describes the interaction between atoms that are not connected. Although many equations are available to describe short-range forces, in this work we employed only a limited number:

- Buckingham potential[139]

$$U(r_{ij}) = Aexp\left(-\frac{r_{ij}}{\rho}\right) - \frac{C}{r_{ij}^6} \tag{2.58}$$

- Lennard-Jones potential

$$U(r_{ij}) = 4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \tag{2.59}$$

- 12-6 potential

$$U(r_{ij}) = \left(\frac{A}{r_{ij}^{12}}\right) - \left(\frac{B}{r_{ij}^6}\right) \tag{2.60}$$

- n-m potential[140,141]

$$U(r_{ij}) = \frac{E_0}{(n-m)}\left[m\left(\frac{r_0}{r_{ij}}\right)^{n} - n\left(\frac{r_0}{r_{ij}}\right)^{m}\right] \tag{2.61}$$

where r_{ij} is the distance between the atoms i and j, the Buckingham potential is made of a repulsive term $Aexp\left(-\frac{r_{ij}}{\rho}\right)$ and an attractive term $-\frac{C}{r_{ij}^6}$. In the Lennard-Jones potential, σ is the distance at which the inter-particle potential is zero and ε is the depth of the potential well. The 12-6 potential is a simplified formulation of the Lennard Jones potential, where $A = 4\varepsilon\sigma^{12}$, $B = 4\varepsilon\sigma^6$ and $\varepsilon = B^2/4A$. The n-m potential is another variation of the Lennard-Jones, where n and m can have different values.

2.5.3.2 Three-Body Potentials

Some force fields contain, in addition to the short-range potentials, three-body potentials, which are useful to restraint the value of angles to physically reasonable values; the force field of Ainsworth *et al.* makes use of three body potential equations to regulate the P-O-P and the O-P-O angles. The mathematical form chosen is the Harmonic:

$$U(\theta_{ijk}) = \frac{K}{2}(\theta_{ijk} - \theta_0)^2 \tag{2.62}$$

where θ_{ijk} is the angle formed by atoms i, j and k and K is the three body spring constant.

2.5.4 Electronic Polarisability via the Core-Shell approximation

When modelling charged species, it is important to consider the electrostatic effect of the neighbour ions on each particle. In the force field of Ainsworth *et al.* the polarisation of the electron cloud around the oxygen ions is taken into account through the adiabatic shell-model,[142–144] where the oxygen ion is represented by a positive core surrounded by a negative shell which has a very small mass; core and shell are connected by a spring.[145] The same formalism was also employed for describing water molecules.[7,8] Figure 2.2 shows a representation of the core-shell formalism: the positive core is connected to a negative shell by a spring.

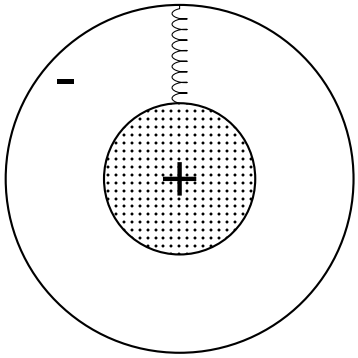


Figure 2.2 Representation of the core-shell formalism

The use of the shell-model in this work allowed to accurately simulate the behaviour of the polarisable ions under the influence of a polarising medium. The disadvantages are the very short timestep due to the high frequency of the core-shell vibration and the increase of computational cost arising from the larger number of particles that have to be considered.

2.6 Water models

Water is the solvent of a large number of chemical and biological reactions, so many models have been developed to simulate it. The very first model dates back to 1933,[146] when a water molecule was represented by a negative charge located on the bisector of the HOH angle and two positive charges on the H atoms. The models used nowadays make use of three-point charge sites on the nuclei (e.g.: SPC, TIP3P[147]), four sites (e.g.:

TIP4P[147]) or five sites (TIP5P[148]), with an increase in computational cost. The electronic repulsion between different water molecules is normally represented by a Lennard-Jones potential centred on the oxygen. Water models are validated against specific properties of liquid water, both at ambient conditions and at different temperatures; the properties more often measured are the density, the heat of vaporisation (enthalpy difference between the vapour and the liquid in coexistence with each other), the self-diffusion coefficient, the structure (from Radial Distribution Functions, see section 2.4.3), the temperature of maximum density, the temperature of freezing and boiling and the dielectric constant.[149] Most of the existing models are empirical, *i.e.* the parameters have been fitted on experimental properties of water, this means that they are reproducing very well the properties they were fitted on, but they may not reproduce other properties of water, or may not give good results at temperatures different from ambient.

In this work, all the MD simulations have been done in water. In chapter 3 we have used the PBE functional for the DFT calculations, which is known to give an overstructured water system with a lower diffusivity compared to those of experiments;[150,151] hybrid functionals can improve the structural and dynamic properties of water,[152] but their application in AIMD simulations would have been too computationally demanding. In chapters 4 and 5 we have employed a polarisable model for water[7,8] where the polarisability of the oxygens is described through the shell-model[142] (see section 2.5.4). The choice of a polarisable model was made on its better compatibility with the polarisable force field used to simulate phosphate ions; moreover, it well reproduces experimental structure (first peak in the radial distribution functions (RDFs) for the pairs O−O, O−H and H−H), dipole moment, energy of vaporisation and vibrational properties of the water monomer, and the binding energy of the dimer.[7,8]

2.7 Enhanced Sampling

The first theoretical models employed to study nucleation were the *hard sphere model*, in which particles are considered as impenetrable spheres that cannot overlap in space, and the *soft particle model*; the latter has a continuous short-ranged interaction potential (e.g.: Lennard-Jones), the exponents of the interaction potential determine the degree

of softness of the particle. These models have been used to investigate the nucleation from the melt. Regarding the nucleation from solution, two kinds of particles must be considered: solute and solvent, moreover the temperature and the pressure need to be appropriate, i.e. the solute must crystallise whilst the solvent remains in the liquid state. The occurrence of nucleation can be enhanced by applying a high chemical potential (by way of supercooling, application of pressure or supersaturation) to the system, but care has to be taken interpreting the results.

Nucleation is a rare event, so it is quite challenging to study: the conventional molecular dynamics has limited time and length scales, so several advanced techniques have been employed to tackle nucleation. The techniques can be biased or unbiased.

Among the directed simulation methods, in which an order parameter is gradually altered to drive the system towards nucleation, the most used are *umbrella sampling* (US, applies a bias potential to one of the degrees of freedom of the system),[153] *constrained dynamics* and *metadynamics* (continuously disfavours regions in phase space such that otherwise stable states are abandoned).[154] The main problem affecting these biased methods is the choice of the order parameter which must not alter the nucleation pathway and it must not lead to a specific crystalline form, because the system may not nucleate directly to the final observed structure. For this reason, there are no applications to molecular crystals other than ice nucleation.[155]

An alternative unbiased approach is *trajectory path sampling* (TPS)[156] and its variations such as *transition interface* or *forward flux sampling*. These alternative methods focus on a single trajectory that captures the nucleation process. The initial configuration of the system is perturbed and propagated in both directions in time by MD simulations. If the resulting trajectory covers the nucleation event, it is chosen to create further trajectories. The collection of successful trajectories represents the possible pathway by which nucleation could proceed. Unfortunately, these unbiased methods are not the best choice for nucleation from solution (which is likely to be limited by diffusion), as in the case of nucleation of ACP in body fluids.

The nucleation of very low solubility compounds (e.g.: CaF_2) from solution has been successfully investigated in an iterative way in which the initial aggregate consists of a single ion and each growth step represents the association of a further ion. The new

adsorption site is identified by means of a Monte Carlo-type approach, and then the cluster and the solvent are relaxed during an MD run.[157,158]

Since in the present work the Umbrella Sampling has been widely employed to obtain free energy profiles, in the following paragraphs the theory beyond this technique is presented in more detail.

2.7.1 Umbrella Sampling

The Umbrella Sampling (US) method allows sampling both low and high energy regions in the configuration space, in order to obtain a free energy profile along a chosen reaction coordinate. Normally the high-energy regions are not sampled accurately during unbiased molecular dynamics, so a bias potential is introduced to ensure efficient sampling along the whole reaction coordinate. Each MD simulation is centred on a reaction coordinate value ("target"), and the system oscillates around that position producing a gaussian-like profile called US window. The effect of the bias potential ("restraint") to connect energetically separated regions in phase space gave rise to the name of this method. The potential ω_i of window i used in this work is harmonic with strength K:

$$\omega_i(\xi) = \frac{1}{2}K(\xi - {\xi_i}^{ref})^2 \tag{2.63}$$

where ${\xi_i}^{ref}$ is the target of each window and ξ is the value assumed by the reaction coordinate at each MD step. The reaction coordinate can be any parameter providing a difference between two thermodynamic states, such as a distance, an angle, or a torsion;[159] in this thesis, the reaction coordinate was always a distance between a Ca^{2+} ion and a phosphorous atom, as this was able to describe the process of interest (see chapter 3). K must be chosen before each simulation and it has to be large enough to drive the system over the barrier, but at the same time not too large, otherwise the overlap between adjacent windows (required by the successive Weighted Histogram Analysis (WHAM)[160–163]) is not guaranteed.[159]

The bias potential ω_i is an additional energy term, which depends only on the reaction coordinate:

$$E^b(r) = E(r) + \omega_i(\xi) \tag{2.64}$$

The superscript b denotes biased quantities, whereas quantities without superscript are

unbiased.

The canonical partition function Q of a system can be calculated via an integral over the whole phase space:

$$Q = \int exp[-\beta E(r)]d^N r \tag{2.65}$$

with $\beta = 1/(k_B T)$, k_B is the Boltzmann's constant, T is the absolute temperature, and N is the number of degrees of freedom of the system. Defining a reaction coordinate ξ, the probability distribution of the system along ξ can be calculated by integrating out all degrees of freedom but ξ:

$$Q(\xi) = \frac{\int \delta[\xi(r) - {\xi_i}^{ref}]exp[-(\beta E)]d^N r}{\int exp[-(\beta E)]d^N r} \tag{2.66}$$

It is now possible to calculate the free energy along the reaction coordinate $A(\xi)$ also called potential of mean force (PMF):

$$A(\xi) = -(1/\beta)lnQ(\xi) \tag{2.67}$$

If every point in phase space is visited during the simulation, i.e. the system is ergodic, $Q(\xi)$ is equal to $P(\xi)$ which is the distribution of the system along the reaction coordinate:

$$P(\xi) = \lim_{t\to\infty} \frac{1}{t} \int_0^t \rho[\xi(t)]dt \tag{2.68}$$

where t denotes the time and ρ counts the occurrence of ξ in a given interval. In order to obtain the unbiased free energy $A_i(\xi)$, the unbiased distribution is needed, according to Eq.(2.66):

$$P_i(\xi) = \frac{\int exp[-\beta E(r)]\delta[\xi(r) - {\xi_i}^{ref}]d^N r}{\int exp[-\beta E(r)]d^N r} \tag{2.69}$$

But the MD simulations provide the biased distribution along the reaction coordinate, so substituting the Eq.(2.64) inside the Eq.(2.69):

$$P_i^b(\xi) = \frac{\int exp\{-\beta[E(r) + \omega_i(\xi(r))]\}\delta[\xi(r) - {\xi_i}^{ref}]d^N r}{\int exp\{-\beta[E(r) + \omega_i(\xi(r))]\}d^N r} \tag{2.70}$$

Because the bias depends only on ξ and the integration in the enumerator is performed over all degrees of freedom but ξ,

$$P_i^b(\xi) = exp[-\beta\omega_i(\xi)]\frac{\int exp[-\beta E(r)]\delta[\xi(r) - {\xi_i}^{ref}]d^N r}{\int exp\{-\beta[E(r) + \omega_i(\xi(r))]\}d^N r} \tag{2.71}$$

using Eq.(2.69) results in:

$$\begin{aligned} P_i(\xi) &= P_i^b(\xi) exp[\beta\omega_i(\xi)] \times \frac{\int exp\{-\beta[E(r)+\omega_i(\xi(r))]\}d^N r}{\int exp[-\beta E(r)]d^N r} \\ &= P_i^b(\xi) exp[\beta\omega_i(\xi)] \times \frac{\int exp[-\beta E(r)] exp\{-\beta\omega_i[\xi(\overrightarrow{r})]\}d^N r}{\int exp[-\beta E(r)]d^N r} \\ &= P_i^b(\xi) exp[\beta\omega_i(\xi)] \langle exp[-\beta\omega_i(\xi)]\rangle \end{aligned} \tag{2.72}$$

where $P_i^b(\xi)$ is obtained from MD simulation of biased system, and $\omega_i(\xi)$ is given analytically (see Eq.(3.1)). At this point $A_i(\xi)$ can be evaluated:

$$A_i(\xi) = -(1/\beta) ln P_i^b(\xi) - \omega_i(\xi) + F_i \tag{2.73}$$

where $F_i = -(1/\beta) ln\langle exp[-\beta\omega_i(\xi)]\rangle$ is an additive constant that is different for each US window and it is calculated in the succeeding WHAM analysis. The worthiness of US lies in the fact that there are no approximation apart from the assumption that the sampling in each window is sufficient.

F_i can be calculated using WHAM analysis[159–163] or Umbrella Integration.[164] In this work WHAM analysis has been used to obtain the final unbiased free energy profiles.

This method minimises the statistical error of $P(\xi)$. The global distribution is calculated by a weighted average of the distributions of the individual windows:

$$P(\xi) = \sum_i^{windows} p_i(\xi) P_i(\xi) \tag{2.74}$$

the weights p_i are chosen in order to minimise the statistical error of P:

$$\frac{\partial \sigma^2(P)}{\partial p_i} = 0 \tag{2.75}$$

under the condition $\sum p_i = 1$. This leads to:

$$p_i = \frac{a_i}{\sum_j a_j} \tag{2.76}$$

$$a_i(\xi) = N_i exp[-\beta\omega_i(\xi) + \beta F_i] \tag{2.77}$$

where N_i is the total number of steps sampled for window i, and

$$\begin{aligned} exp(-\beta F_i) &= \langle exp[-\beta\omega_i(\xi)]\rangle \\ &= \int P(\xi) exp[-\beta\omega_i(\xi)] d\xi \\ &= \int exp\{-\beta[A(\xi)+\omega_i(\xi)]\} d\xi \end{aligned} \tag{2.78}$$

At this stage it is possible to calculate F_i by Eq.(2.78):

$$exp(-\beta F_i) = \int P(\xi) exp[-\beta \omega_i(\xi)] d\xi \tag{2.79}$$

Equations (2.74), (2.76) and (2.79) have to be iterated until convergence.

The US technique was extensively employed in chapter 3 to obtain free energy profiles.

Chapter 3

Calcium phosphate prenucleation complexes in water by means of *ab initio* Molecular Dynamics simulations

3.1 Introduction

Hydroxyapatite[165] originates from an amorphous calcium phosphate (ACP) precursor, in a process of dissolution and reprecipitation.[24] In the 1970s, Posner and Betts proposed that structural units of the form $Ca_9(PO_4)_6$ would aggregate randomly, with the inter-cluster space filled with water, into larger spherical particles of ACP.[12] More recently, a cluster-growth model for the formation of ACP was proposed,[80] whereas aggregating clusters were found to be present in body fluids even before nucleation.[22] XANES and XRD experiments of the early stages of calcium phosphate crystallisation have suggested that an idealised cluster with formula $Ca_9(PO_4)_6(H_2O)_{30}$ may act as the structural unit of ACP,[87] whereas the most abundant clusters detected in solution at the early stage of nucleation have been reported to be of the form $Ca(\eta^2-PO_4{}^{3-})_2L_2$ (L=H_2O or $\eta^1-PO_4{}^{3-}$, where η^1 and η^2 stand for monodentate and bidentate binding of the phosphate groups).[88] Habraken *et al.*[23] showed that the prenucleation clusters (PNCs), aggregating in solution to form polymeric structures, have formula $[Ca(\eta^2-HPO_4)_3]^{4-}$. Interestingly, a Posner's cluster can be seen as two deprotonated PNCs in which all negative charges are compensated by complexing calcium ions. In order to explain the formation of ACP at the

supersaturations used in their experiment, the PNCs were proposed to have an excess free energy over that of free ions, which reduces dramatically the nucleation barrier.[83]

Despite the calcium phosphate PNCs represent a key factor to understand ACP formation, due to the difficulties in accessing experimental data at such small scale, there is no general consensus about their structures. Here, we therefore present a theoretical investigation based on *ab initio* molecular dynamics MD simulations, which suggests that the resulting PNC geometry differs from the existing models in the literature. In addition, we provide an estimate for the free energy cost required to decompose a PNC into isolated free ions, from which we deduce that calcium triphosphate PNCs are more stable entities than other Ca-to-P ratios.

3.2 Computational details

We performed preliminary calculations in gas phase considering explicitly four water molecules and fixing the coordination number of calcium at six, seven or eight in order to examine the structural variations in the prenucleation complexes $[Ca(HPO_4)_3]^{4-}$; the ratio between phosphate and water ligands was also changed. The geometry optimisations were carried out at the DFT level using a hybrid Gaussian plane wave (GPW) scheme that combines a Gaussian basis for the wave functions with an auxiliary plane wave basis set for the density as implemented in the CP2K/QUICKSTEP program.[91–101] We have used Goedecker-Teter-Hutter (GTH) type pseudopotentials,[93,98,101] together with the PBE functional.[100]

The complex $[Ca(HPO_4)_3]^{4-}$ was then optimised in gas phase with DFT without any water molecules and the resulting structure was solvated using the Visual Molecular Dynamics (VMD) solvation tool.[1] Two calcium ions were added at the extremity of the box in order to neutralise the 4- charge of the complex. The obtained square “box" containing 93 water molecules and with a volume of $\approx 15 \times 15 \times 15$ Å was equilibrated in an NPT ensemble for 57.5 ps at a constant temperature of 300 K and pressure 1 bar in order to regulate the density of the water at the chosen conditions. The timestep was set at 0.5 fs. The system was then further equilibrated for 60 ps in an NVT ensemble.

PBE belongs to the GGA class of functionals, which are known to perform scarcely dur-

ing AIMD simulations of liquid water. In particularly they lead to an overstructured liquid and a very low diffusivity compared to the experiments. This disagreement is due to the inability of the GGA to account for the proton quantum effect and the incapacity to describe hydrogen bonds and dispersion interactions.[150,151] In order to obtain RDFs and diffusivity in agreement with experiments, one should run the MD at a temperature $\approx$ 20% higher.[150] However, free energy barriers depends on temperature, so this solution was not suitable for our work. Hybrid functionals give a better description of hydrogen bonding in water,[152] but they are far more expensive than PBE in terms of computing time. Dispersion-corrected atom centred potentials (DCACPs) and empirical van der Waals corrections used to include the dispersion interactions within DFT give RDFs in a better agreement with experiments at a given temperature, but the fluid still appears slightly overstructured.[166] In any case, it is not clear how the inclusion of dispersion interactions affects the free energy barriers. Nevertheless, the structural properties of water and water dimer described by PBE are in good agreement with experiments.[167] Therefore our choice was dictated by the balance between the computing time and the free energy profile estimation at a given temperature.

A step-by-step analysis of the coordination of the calcium ion has been done in the NVT trajectory (coordination number, denticity of the phosphate groups and number of water molecules), then the calculation of the RDFs was used to evaluate the average distances between atom pairs.

The stability of the calcium triphosphate complex in water has been evaluated through the analysis of free energy profiles obtained with the Umbrella Sampling (US) technique.[153,159] The system was equilibrated at each target value for at least 7.5 ps and only the last 5 ps were considered for the WHAM analysis. The potential used in this work was harmonic with strength k:

$$\omega_i(\xi) = k/2(\xi - {\xi_i}^{ref})^2, \tag{3.1}$$

where ${\xi_i}^{ref}$ is the target of each window i. A value of 100 kcal/mol/Å^2 was chosen for k in most of the US windows. Around the minima, lower values of k were sufficient to keep the reaction coordinate oscillating around the target, whereas in the steepest regions of free energy, k was increased up to 150 kcal/mol/Å^2.

3.3 Structure and stability in gas-phase

The structure of the complex $[Ca(HPO_4)_3]^{4-}$ was studied in gas phase considering explicitly four water molecules. The coordination of the central calcium was fixed at six, seven and eight, and the number of water molecules coordinated to calcium was varied from zero to four, thus changing the denticity of the phosphate groups. Figure 3.1 shows the relative energies of the optimised structures:

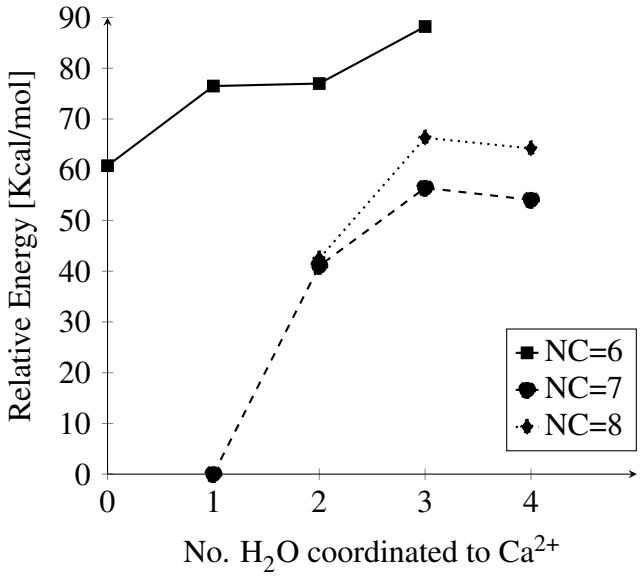


Figure 3.1 Relative energies of the optimised structures with different calcium coordination numbers.

In Figure 3.1 it is possible to see that the coordination number for calcium leading to lowest energies is seven. The most stable structure we found is very similar to the one calculated by Habraken *et al.*[23] with three bidentate phosphate groups and one water molecule, indicated as $Ca(\eta^2-HPO_4{}^{2-})_3(H_2O)$. The water molecules not directly coordinated to calcium were engaged in hydrogen bonding with the phosphate groups. The Ca coordination number six was clearly disfavoured, as the structure with the lowest energy was $\approx$ 60 Kcal/mol above the most stable structure for Ca coordination number seven.

3.4 Structure and stability in water

The initial geometry obtained after gas phase DFT optimisation without water molecules and the final structure after equilibration in water are shown in Figure 3.2. The central

calcium coordination number was six in the structure optimised in the gas phase, with the three phosphate groups coordinated in a bidentate way, (Figure 3.2(I)). A similar structure with three bidentate phosphate groups, and ligating one water molecule, has been found by both *ab initio* calculations with a number of explicit water molecules[23] and classical MD simulations in water.[168] During the dynamics, however, two water molecules entered the first coordination sphere, changing significantly the structure of the complex (Figure 3.2(II)). The change took place already during the equilibration in an NPT ensemble: two water molecules entered in the first calcium coordination sphere, thus increasing the space between two phosphate groups. To describe the phenomenon in the time we visualised a plane defined by three atoms: the calcium and the two phosphorus belonging to the phosphate groups which opened during the dynamic (we can refer to the latter as "lateral groups"); we considered the angle θ formed by the normal to this plane and the line passing through the central calcium and the third phosphorus (Figure 3.3). In the starting configuration, the angle θ was $\approx 90^\circ$, but it decreased rapidly in time, this means that the structure opened quickly to make space for a water molecule which entered after only 0.24 ps. θ oscillated in the interval 10° - 70° during the first 10 ps, then it stabilised around 10°. The oscillation did not mean that the structure closed again, but that the central phosphate group was moving rapidly establishing hydrogen bonds with different water molecules. At the end of the NPT trajectory, the central ligand was interacting permanently with one of the lateral phosphates via a direct hydrogen bond and a bridging water and with the other phosphate via a single bridging water (see Figure 3.2(II)).

For almost 90% of the simulation in an NVT ensemble, the complex remained in a $Ca(\eta^2-HPO_4^{2-})_2(\eta^1-HPO_4^{2-})(H_2O)_2$ distorted pentagonal bipyramidal geometry (Figure 3.2(II)). A similar arrangement of the calcium ligands has been found during classical MD for several types of calcium-phosphate ion pairs.[168] Here, two $\eta^2-HPO_4^{2-}$ groups and one water molecule occupy the equatorial positions, whereas the $\eta^1-HPO_4^{2-}$ oxygen and a second water molecule occupy the axial sites. As the two equatorial groups could undergo a $(\eta^2-HPO_4^{2-})$ to $(\eta^1-HPO_4^{2-})$ transition, we observed a calcium coordination of six for approximately 10% of the simulation time (Figure 3.2(III)). The relative appearances of the two complexes are described by the time evolution of the six $Ca-O_{ph}$ distances during the simulation (Figure 3.4).

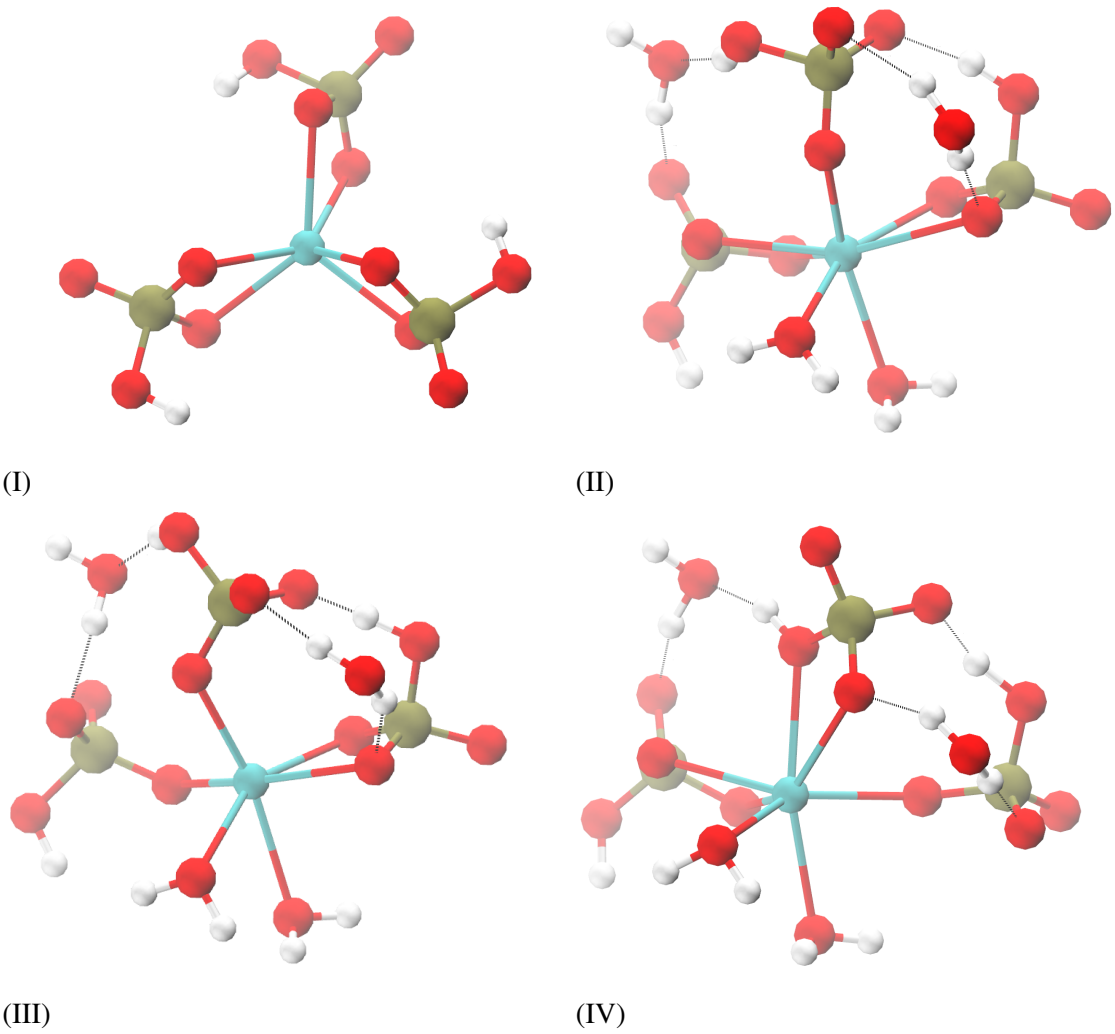


Figure 3.2 Structures obtained in the gas phase (I) and in water (II, III, IV). The dotted lines represent hydrogen bonds. Colour key: Ca:cyan, P:green, O:red, H:white.

The histograms in Figure 3.5 summarise the percentage of appearance of a specific calcium coordination number, number of water molecules and number of monodentate and bidentate groups during the NVT trajectory. It is evident the preference for the calcium for a coordination 7. The data were obtained calculating the calcium coordination number in each trajectory frame and are reported as percentage of the total number of frames. Because of the large computational cost, we ran only one MD simulation and therefore it is not possible to report error bars on the histograms. The data presented in Figure 3.5 refer to a system previously equilibrated in an NPT ensemble, thus the uncertainty should be low; however, a longer trajectory or running more than one simulation from the same starting configuration is required to give an accurate estimate of the errors on each parameter.

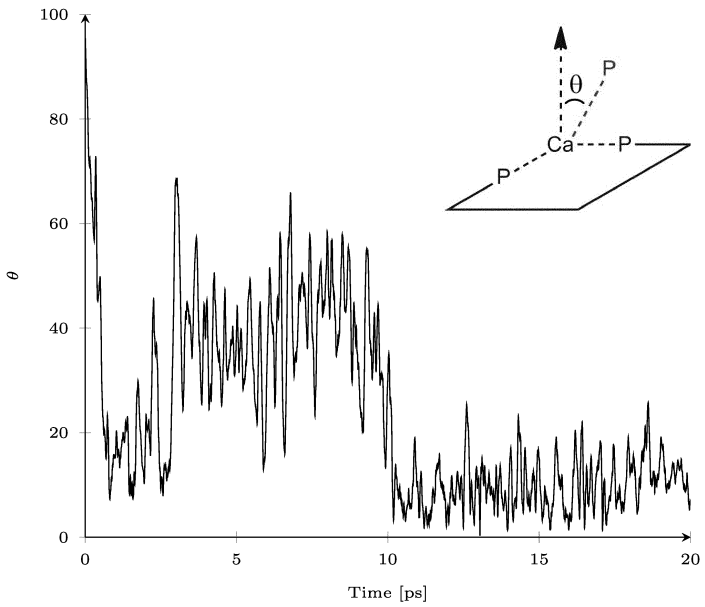


Figure 3.3 Variations in the angle θ during the NPT equilibration, with a schematic representation of the angle θ.

The structural modifications that the complex underwent in water stress the importance of the properties of the solvent. Liquid water forms different hydrogen bond configurations: double-donor, single-donor and non-donor.[169] It is evident that hydrogen bonding is established between water molecules and complex, and this leads to important structural effects. All the water-complex hydrogen bond types found in the NVT trajectory are presented in Figure 3.6. In the cases (a-b) a single hydrogen bond is formed, in the cases (c-h) two bonds are formed. The water molecule bridges two oxygen atoms belonging to the same phosphate group in (c) and (e); in (d) the phosphate acts as donor and acceptor of protons, in (f) a phosphate oxygen accepts two protons of the same water molecule, in (g) and (h) a water molecule bridges two different phosphates. We observed the presence of two bridging waters of type (g) in Figure 3.6 and one of type (h) in Figure 3.6 for $\approx$ 98.5% of the trajectory. The complex is surrounded by 21 water molecules for $\approx$ 56.5% of the time, this number includes the water molecules directly coordinated to calcium.
We show in Figure 3.7 the RDFs $g(r)$, obtained during the entire 60 ps of NVT trajectory, for the pairings Ca$-$O, Ca$-$P, P$-$O, O$-$O and P$-$P. The first maxima (corresponding to the first coordination sphere of Ca^{2+}) of the Ca$-$$O_{ph}$ and Ca$-$$O_{wt}$ pairings are located at 2.45 and 2.55 Å respectively. This result is in line with the findings of Almora-Barrios and de Leeuw for the nucleation of hydroxyapatite at a collagen template by means of

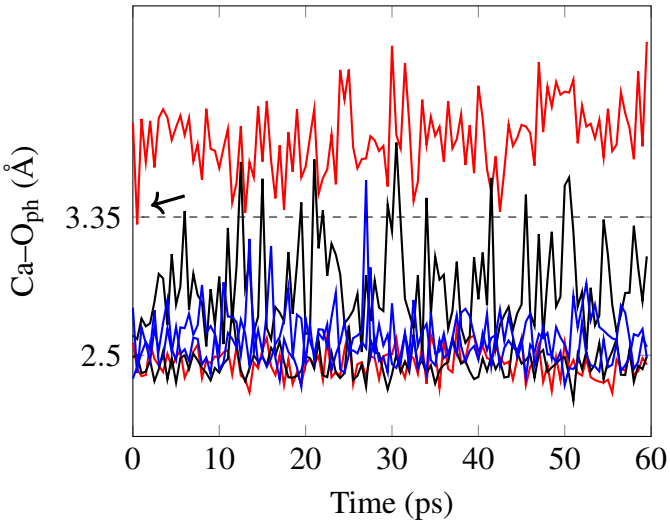


Figure 3.4 Time evolution of the Ca−O$_{ph}$ distances during the NVT run. The red lines refer to the oxygen atoms of the phosphate group at the axial site, the black and the blue lines to those of the two groups at the equatorial sites. The dashed horizontal line highlights the bond distance cutoff of 3.35 Å corresponding to the minimum in the Ca−O$_{ph}$ RDF. One attempt to form a eight-coordinated structure is evidenced by the black arrow.

classical MD.[168] They reported Ca−O$_{ph}$ and Ca−O$_{wt}$ distances for various calcium phosphate complexes in water of 2.2–2.6 Å, and 2.4 Å, respectively. Slightly shorter Ca−O distances (Ca−O$_{wt}$ 2.35 Å, Ca−O$_{bicarb}$ 2.36 Å, and Ca−O$_{carb}$ 2.38 Å) have been predicted by Car-Parrinello simulations of the calcium ion, calcium bicarbonate and calcium carbonate complexes in water,[170] which could be attributed to the lower coordination of calcium (six) or to the less pronounced steric effect of the (bi)carbonate ligand. The Ca−P $g(r)$ shows two maxima, signalling bidentate and monodentate phosphate ligands, whose Ca−P distances are respectively 3.15 and 3.65 Å. The contributions to the first coordination sphere of the phosphorus atom are given by the four oxygens directly ligated, whereas the water oxygens are relegated to the second coordination sphere, as shown by the P−O $g(r)$. The Ca−O, P−O and P−P distances are in good agreement with the *ab initio* calculations of the structure proposed by Habraken *et al.*, whereas the Ca−P and O−O ones are approximately 0.1 Å larger than their upper limits.[23] In general, the computational data tend to overestimate the distances for Ca−O, P−O, and P−P with respect to the experiments on ACP,[12] but the latter were obtained in the solid phase, without solvent effects. Our results are summarised in Table 3.1, where the coordination numbers, obtained by in-

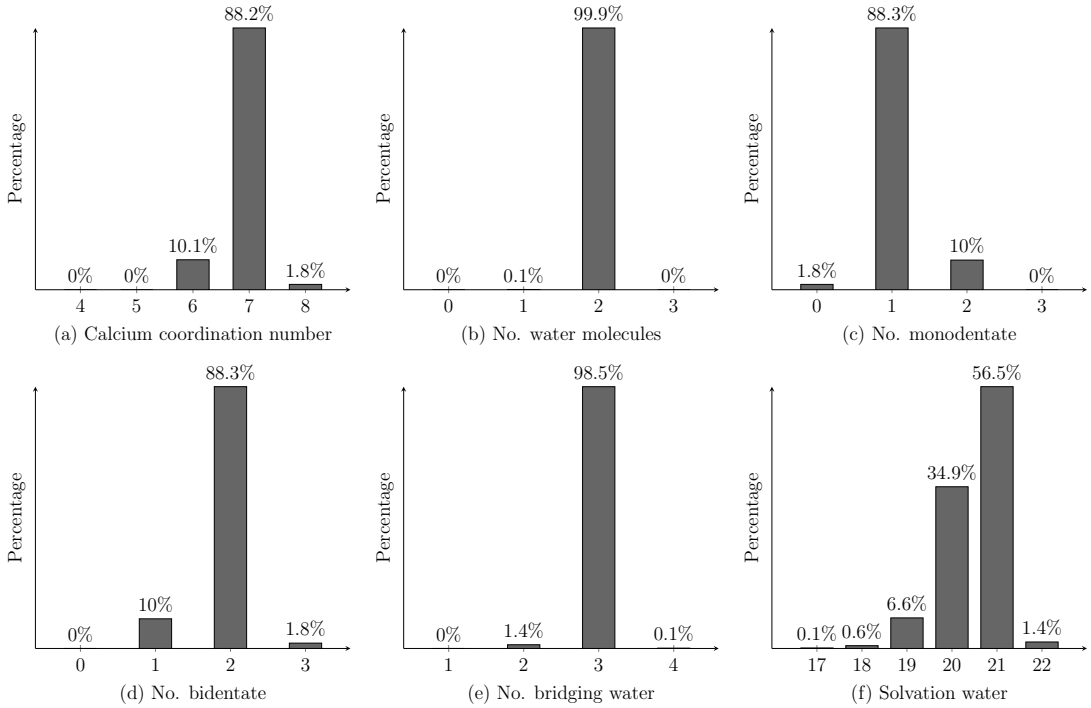


Figure 3.5 (a) Coordination number of the calcium ion during the NVT run, (b) Number of water molecules, (c) Number of monodentate phosphate groups, (d) Number of bidentate phosphates, (e) Number of bridging waters and (f) First sphere solvation water. The cutoff for Ca−O interaction was set at 3.35 Å.

tegrating $g(r)$ between the relevant minima, are consistent with the distorted pentagonal bipyramidal geometry described above. To estimate the accuracy of the data presented in the table, it is necessary to run more simulations from equivalent starting configurations. Because of the non-negligible computational time required by our unbiased AIMD simulations, data were collected on a single 60 ps trajectory.

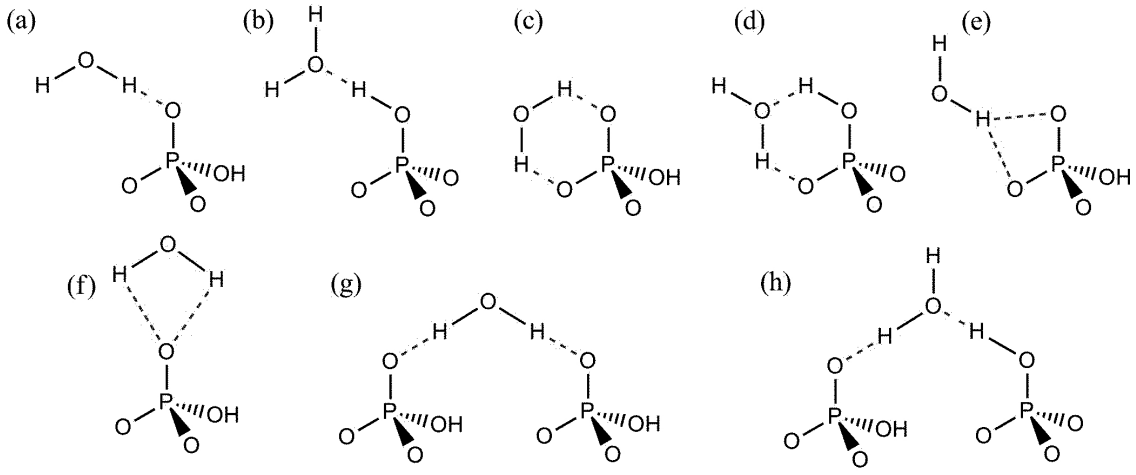


Figure 3.6 Representation of the hydrogen bonds between phosphate groups and water molecules.

Table 3.1 Coordination numbers and distances for the first coordination sphere of Ca−O, Ca−P, P−O, O−O, P−P pairings. O_{ph} refers to oxygen atoms bonded to phosphorus, whereas O_{wt} are the oxygen atoms of water.

Pairing	**CN**$_1$	**d**$_1$
Ca−O_{ph}	5.0	2.45 Å
Ca−O_{wt}	2.0	2.55 Å
Ca−P	3.0	3.15-3.65 Å
P−O_{ph}	4.0	1.55 Å
P−O_{wt}	6.9	3.75 Å
O_{ph}−O_{ph}	3.2	2.55 Å
O_{ph}−O_{wt}	1.2	2.65 Å
P−P	1.3	5.25 Å

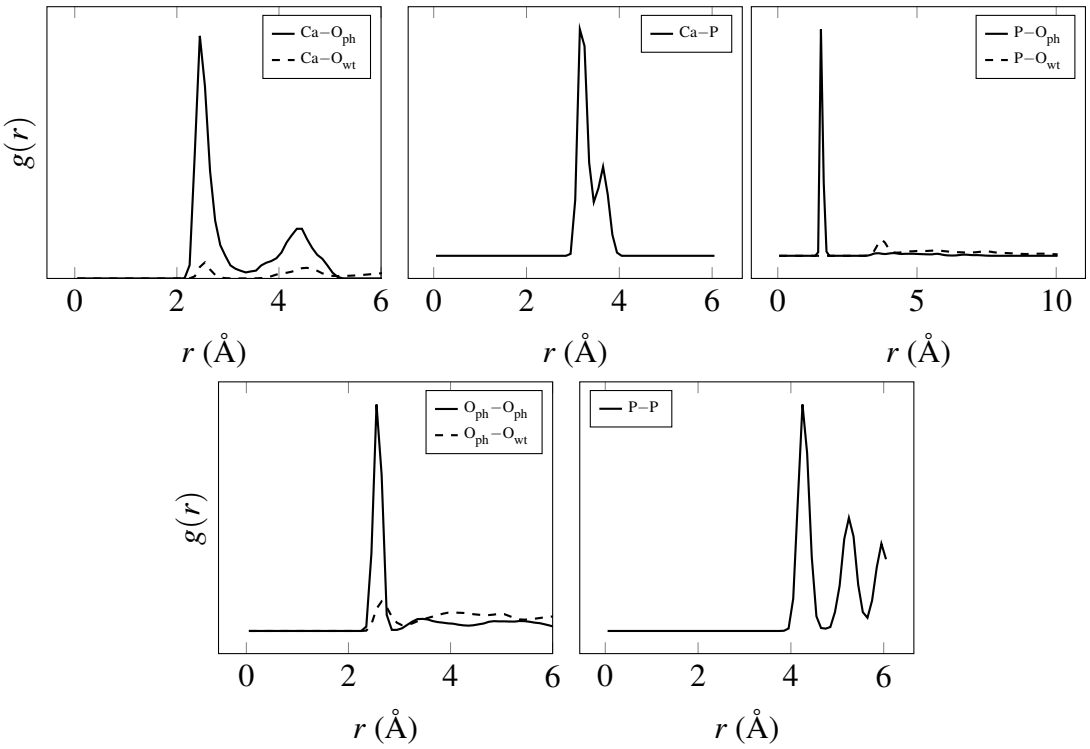


Figure 3.7 Radial distribution functions for the Ca−O, Ca−P, P−O, O−O, P−P pairings.

3.4.1 Stability of secondary configurations

In order to better evaluate the stability of the six-coordinated PNC (Figure 3.2(III)) present during about 10% of the MD trajectory, we have calculated the free energy profile for the $Ca(\eta^2-HPO_4^{2-})_2(\eta^1-HPO_4^{2-})$ to $Ca(\eta^2-HPO_4^{2-})(\eta^1-HPO_4^{2-})_2$ transition. We show it in the upper panel of Figure 3.8, where the reaction coordinate is given by the distance between the calcium and a phosphorus from an equatorial group. The minimum of the six-coordinated structure (Figure 3.2(III)) is only 1.9 kcal/mol higher in energy than that of the seven-coordinated structure (Figure 3.2(II)). An additional 5 ps simulation with the reaction coordinate restrained at the target, corresponding to the six-coordinated minimum, revealed no significant coordination changes, confirming that a $Ca(\eta^2-HPO_4^{2-})(\eta^1-HPO_4^{2-})_2$ PNC can exist in aqueous solution.

During the entire NVT run, two water molecules were tightly ligated to the PNC, while we observed a number of attempts by the three phosphate groups to bind simultaneously to calcium in a bidentate mode (Figure 3.4). In order to shed light on the possibility of formation of a $Ca(\eta^2-HPO_4^{2-})_3$ PNC, we have studied the free energy profile corresponding to the monodentate-bidentate transition of the central phosphate group. We present it in the lower panel of Figure 3.8, where the reaction coordinate is the distance between the central calcium and the phosphorus of the monodentate phosphate. We observe that the structure found after equilibration (Figure 3.2(II)) is slightly more stable than the structure at shorter Ca−P distances (Figure 3.2(IV)), in which the central phosphate group is bidentate. However, during the US, we observed the $\eta^2-HPO_4^{2-}$ to $\eta^1-HPO_4^{2-}$ transition of an equatorial phosphate, which prevented the formation of a $Ca(\eta^2-HPO_4^{2-})_3$ PNC, and led to the formation of a $Ca(\eta^2-HPO_4^{2-})_2(\eta^1-HPO_4^{2-})$ cluster. This behaviour can be explained by considering the bridging between the phosphate groups: in both minima, the central phosphate is linked to an equatorial phosphate via a direct hydrogen bond and a bridging water and to the other equatorial group via a second bridging water (types (g) and (h) respectively in Figure 3.6). In order to retain the bridging structures, the lateral bidentate phosphate in the structure in Figure 3.2(II) had to become monodentate in the structure in Figure 3.2(IV).

Other than water bridging between phosphate groups, the constant presence of one monodentate ligand can be explained in terms of charge transfer: when the 2+ cation is sur-

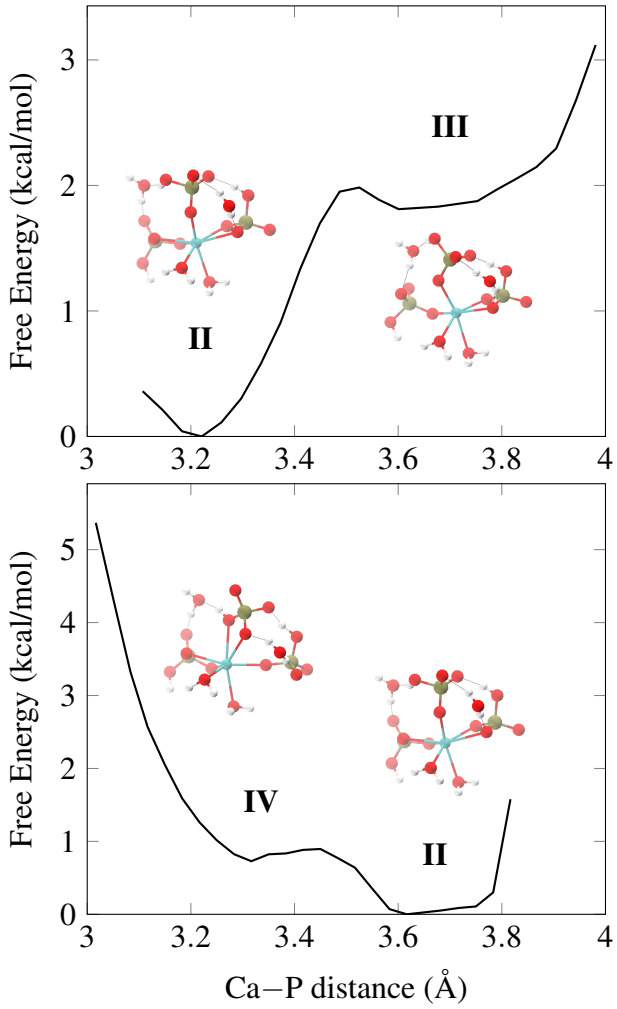


Figure 3.8 Free energy profiles obtained when moving away one of the bidentate groups from calcium (upper panel) and pulling the monodentate group (lower panel) toward calcium. The minima are labelled consistently with the structures in Figure 3.2.

rounded by three 2− anions, its partial positive charge is decreased, and therefore there is a loss of charge-charge interaction with the ligands.[171] The weakening of metal-phosphate interactions increases the importance of other factors, such as the effect of hydrogen-bonds with water, which favours the monodentate binding mode, as it maximises the interaction with the first solvation sphere.

The $Ca(\eta^2-HPO_4^{2-})_2(\eta^1-HPO_4^{2-})(H_2O)_2$ geometry emerging from our MD simulations differs from that suggested by Habraken *et al.*[23] in the substitution of a $(\eta^2-HPO_4^{2-})$ with $(\eta^1-HPO_4^{2-})H_2O$. Zhang *et al.* have been able to detect the presence of CaP PNCs at the early stage of ACP formation in solution.[88] On the basis of calcium K-edge XANES spectrum features, they have speculated the presence of

$Ca(\eta^2-HPO_4{}^{2-})_2L_2$ PNCs, where L is either a $\eta^1-HPO_4{}^{2-}$ or a water molecule, with calcium in a coordination of six. However, we note that a seven-coordinated calcium remains consistent with these experimental findings, as coordinations of six and seven give similar XANES edge positions.[172]

3.4.2 Estimation of the stability of the complex

Next, we have evaluated the stability of the PNC by deriving the free energy profile for the dissociation of the $\eta^1-HPO_4{}^{2-}$ from the central calcium ion:

$$[Ca(HPO_4)_3]^{4-}_{aq} \rightarrow [Ca(HPO_4)_2]^{2-}_{aq} + (HPO_4)^{2-}_{aq} \tag{3.2}$$

We set the reaction coordinate to be the distance between the central calcium ion and the phosphorus atom of the $\eta^1-HPO_4{}^{2-}$ detaching group. Starting from the equilibrated structure (Figure 3.2(II)), a series of US simulations were performed with increasing target distances, up to the complete removal of the phosphate from the Ca^{2+} first coordination sphere.

The free energy profile depicted in the upper panel of Figure 3.9 confirms that the PNC is very stable in water, as the energetic cost for moving the $\eta^1-HPO_4{}^{2-}$ to the second coordination sphere of calcium is approximately 22 kcal/mol. The calcium biphosphate complex in Figure 3.10(V) appears to be in a shallow minimum located at a distance of 6.4 Å. Due to a third water molecule replacing the $\eta^1-HPO_4{}^{2-}$, the calcium coordination is still seven, and the complex is in a $Ca(\eta^2-HPO_4{}^{2-})_2(H_2O)_3$ arrangement. The detached $HPO_4{}^{2-}$ group interacts with the calcium biphosphate via two bridging water molecules (type (g) in Figure 3.6), and a direct hydrogen-bond, as shown in Figure 3.10.

The significant increase in free energy for the dissociation process described above suggests that the PNC is stable with respect to the isolated free ions. In order to verify this hypothesis, we continued the fragmentation process, removing a second phosphate group from the calcium, according to the following reaction:

$$[Ca(HPO_4)_2]^{2-}_{aq} \rightarrow [Ca(HPO_4)]_{aq} + (HPO_4)^{2-}_{aq} \tag{3.3}$$

We have plotted the corresponding free energy profile in the middle panel of Fig-

ure 3.9. Here, it is possible to observe three minima: the first, at a Ca−P distance of 3.2 Å, corresponds to the starting Ca(η^2−HPO_4^{2-})$_2$, the second at ≈ 3.6 Å is Ca(η^2−HPO_4^{2-})(η^1−HPO_4^{2-}), and the third corresponds to Ca(η^2−HPO_4^{2-}), with (η^1−HPO_4^{2-}) not in the first coordination sphere of calcium. The free energy barrier is ≈ 9 kcal/mol, and the free energy difference between the Ca(η^2−HPO_4^{2-})(η^1−HPO_4^{2-}) and Ca(η^2−HPO_4^{2-})$_2$ is 2.6 kcal/mol, a value which is slightly higher than that of the transition in the central phosphate (see Figure 3.8, lower panel). Furthermore, the position of the minimum for the dissociated structure (5.1 Å) is lower than that in the upper panel of Figure 3.9 (6.4 Å), where two bulky ligands form hydrogen-bonds and water bridges with the leaving group. These interactions oppose the separation of the leaving ion and are reflected in the shape of the free energy curve. Only after breaking the bridges with one of the equatorial phosphates does the curve smoothe down to a shallow minimum. The structure with the calcium ion coordinating one bidentate phosphate and five water molecules (Figure 3.10(VI)) has already been found during classical MD simulations of calcium phosphate complexes.[168] The authors reported a similar arrangement of the ligands and distances Ca−O_{ph} of 2.5 Å. We found that, for the Ca(η^2−HPO_4^{2-})$(H_2O)_5$, the Ca−O_{ph} distances oscillate between 2.45 and 2.55 Å, with the distance between the calcium and the oxygen atom carrying the proton shifted towards the upper value of the range.

Finally, starting from the minimum in Figure 3.10(VI), we have detached the last phosphate group, obtaining the isolated ions in solution:

$$[Ca(HPO_4)]_{aq} \rightarrow Ca^{2+}_{aq} + (HPO_4)^{2-}_{aq} \tag{3.4}$$

The corresponding free energy profile is illustrated in the lower panel of Figure 3.9. Similar to the profile in the middle panel of Figure 3.9, its main features are a minimum at ≈ 3.2 Å (bidentate binding mode), a flattening at ≈ 3.6 Å (monodentate binding mode) and a third depression corresponding to the egress of the phosphate group from the first coordination sphere of the calcium. The free energy barrier is ≈ 10 kcal/mol, whereas the free energy difference between the monodentate and bidentate binding modes is 2.7 kcal/mol. We note that all the positions and energy differences of the minima involved in reactions 3.3 and 3.4 are very similar. The isolated calcium ion shown in Figure 3.10(VIIa) is

six-coordinated and the structure is a distorted octahedron. Di Tommaso and de Leeuw reported an equivalent structure during Car-Parrinello molecular dynamics simulations of calcium carbonates in water.[170] Despite the similarity, we found a $Ca-O_{wt}$ average distance of 2.45 Å, 0.1 Å longer than theirs, which presumably is due to the different choice of pseudopotentials.

Taking into account the three free energy profiles in Figure 3.9, we can deduce that each process of a phosphate leaving the calcium coordination shell has an important energetic cost ($\geq$ 9 kcal/mol) and that the complex is more stable than the free ions. It is worth noting that the energetic expense of the detachment of the first phosphate group was approximately two times more costly than the subsequent detachments, suggesting that calcium triphosphate complexes are more stable entities than other Ca-to-P ratios, in agreement with experiment.[23]

The free energy differences between the complex and the isolated ions can be explained considering the entropic effects of solvation water: the aggregation of single ions brings along the partial release of the structured solvation water and thus an increase of the less organised bulk water which makes higher the entropic contribution to the Helmholtz energy (F):

$$F = E - TS \tag{3.5}$$

This differs from the findings of Habraken *et al.*,[23] because they concluded that PNCs are metastable with respect to free ions.

However, our results suggest that the calcium triphosphate prenucleation complexes fit in the PNC definition of Gebauer *et al.*[79] already listed in section 1.4, since they satisfy definitions (i), (ii), (iv) and (v):

(i) PNCs are composed of constituent ions of the forming solid, in fact, calcium triphosphate is the basic structural unit of amorphous calcium phosphate (ACP).

(ii) PNCs are small and thermodynamically stable solutes, with no phase boundary with the solution. We demonstrated here the stability of the $[Ca(HPO_4)_3]^{4-}$ complex computing the free energy profile for the removal of the three phosphate groups, which showed an important energetic cost required for the dissociation.

(iii) PNCs are molecular precursors of the phase nucleating from solution and hence participate in the process of phase separation. Unfortunately, we can not ponder on this point since no studies further than prenucleation stages have been done in this work.

(iv) PNCs are highly dynamic entities and change configuration on timescales typical for molecular rearrangements in solution. We observed that other less stable configuration do exist in solution and lie only 0.5 - 1.9 Kcal/mol over the minimum.

(v) PNCs have encoded structural motifs resembling, or relating to, one of the corresponding crystalline polymorphs. Calcium triphosphate is the structural unit of octacalcium phosphate and apatites.[23]

Since stable clusters can never lie on the thermodynamic path of nucleation, we can assume that, as in the case of calcium carbonate, the nucleation takes place for kinetic reasons.[79,173]

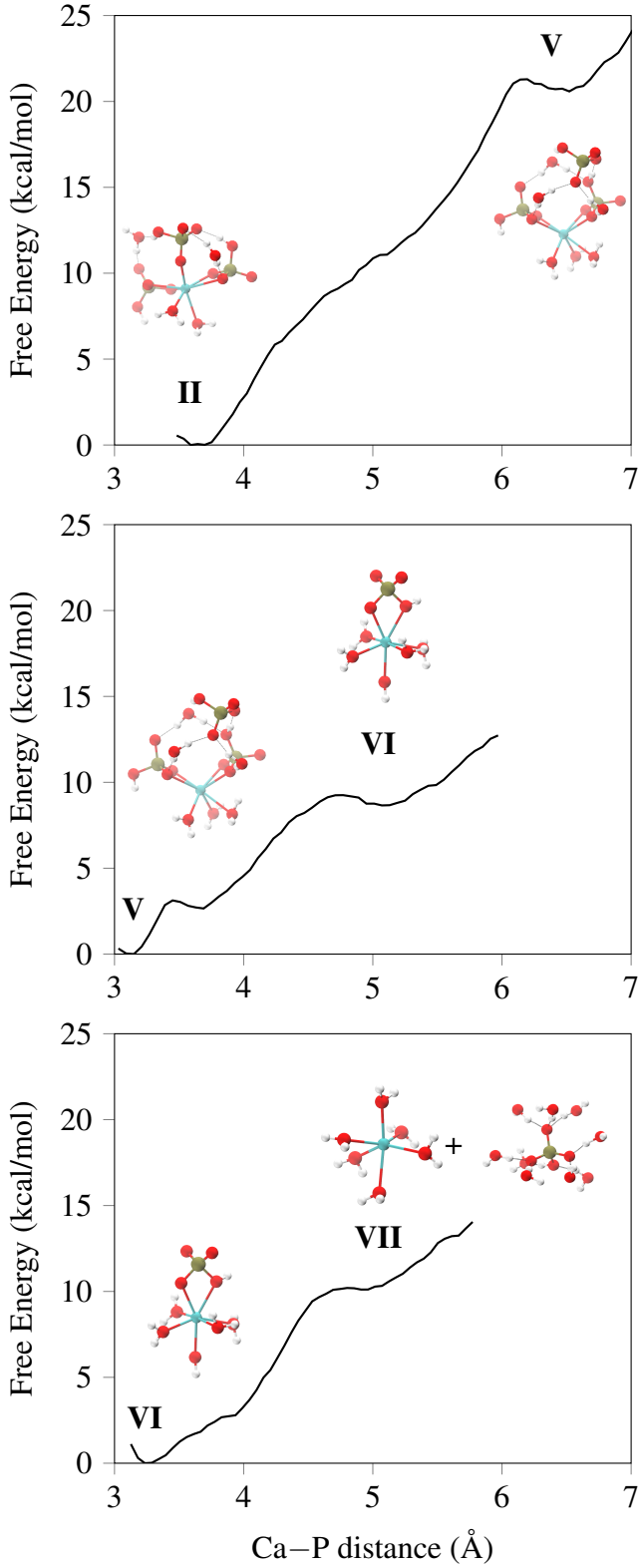
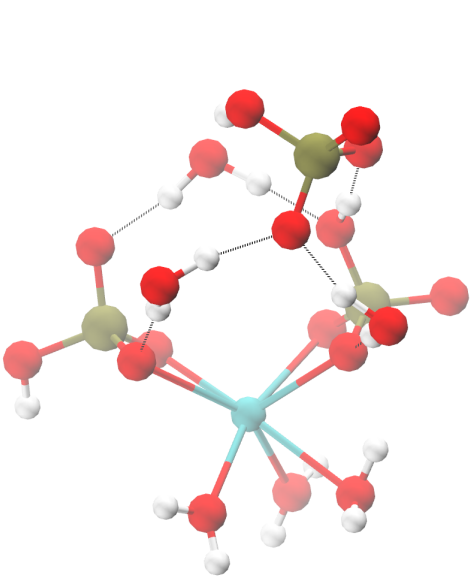
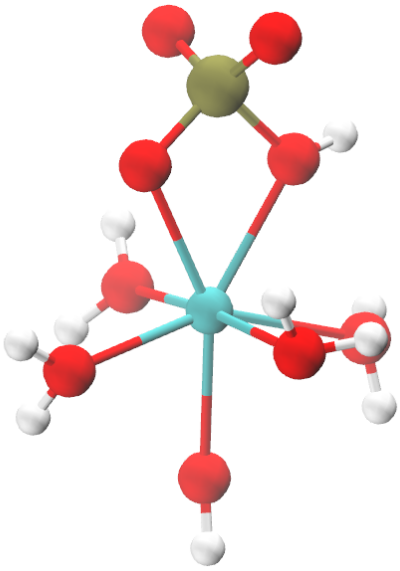


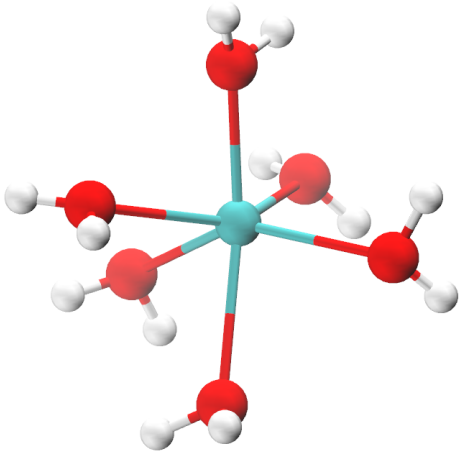
Figure 3.9 Free energy profiles for the break down of the PNC into free ions. Each graph corresponds to the detachment of a phosphate group. The minima are labelled consistently with the structures in Figures 3.2 and 3.10. Colour key: Ca:cyan, P:green, O:red, H:white.



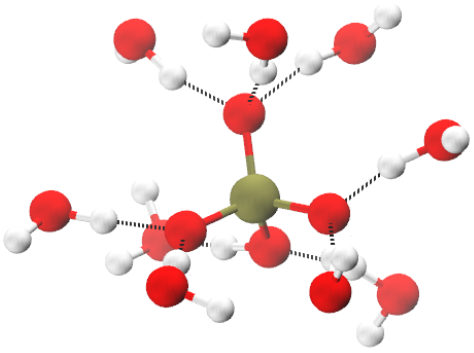
(V)



(VI)



(VIIa)



(VIIb)

Figure 3.10 Metastable structures obtained during the dissociation of the PNC into free ions. Colour key: Ca:cyan, P:green, O:red, H:white. The dotted lines represent hydrogen bonds.

3.5 Conclusions

We have studied the structures and stabilities of $[Ca(HPO_4)_3]^{4-}$ prenucleation complexes in water by means of *ab initio* molecular dynamics simulations and umbrella sampling techniques. We have found that, in the most stable configuration, one phosphate group ligated to the calcium is monodentate, while two are bidentate and placed at the opposite sides of the calcium. In addition, two water molecules also bind the calcium ion, which prefers to stay in a $Ca(\eta^2-HPO_4{}^{2-})_2(\eta^1-HPO_4{}^{2-})(H_2O)_2$ arrangement. The calcium triphosphate PNC is more stable than the isolated ions, as evidenced by the free energy profiles simulating the dissociation process. Moreover, our data suggest that a Ca-to-P ratio of 1:3 is thermodynamically favoured over other ratios, supporting the experimental findings in the literature.[23]

Chapter 4

Detection of Posner's clusters during calcium phosphate nucleation: a Molecular Dynamics study

4.1 Introduction

Calcium phosphates are the main inorganic components of bone tissue, where they are mainly present as crystalline HA. Under conditions of normal temperature and pressure, HA is the thermodynamically most stable phase of calcium phosphate,[174] but despite its supersaturation in body fluids, the direct nucleation of this important biomineral is inhibited *in vivo*, where other calcium phosphate phases precipitate before HA.[175] In particular, ACP has been proposed as the first phase to nucleate from aqueous solution.[176] Owing to the importance of calcium phosphate as a biomaterial, several experimental[23,87,177] and theoretical[168] studies have focused on its nucleation and growth. In particular, using X-ray experiments, Betts and Posner discovered in the 1970s that, despite its lack of long-range order, ACP is characterised by spherical domains of approximately 1 nm in size, with a Ca/P molar ratio of 1.5 and chemical formula $Ca_9(PO_4)_6$. These clusters, which were later named Posner's clusters, aggregate randomly in solution and form spherical particles, with water molecules filling the intercluster space.[12]

The formation of ACP from the aggregation of Posner's clusters was investigated using several experimental techniques. Onuma and Ito used *in situ* atomic force microscopy

(AFM) to probe the crystal growth of HA in supersaturated body fluid solutions, and found the presence of calcium phosphate clusters with sizes in the range of 0.7 to 1.0 nm.[80] They proposed that suitably oriented Posner's clusters are the building units of HA.[80,178] Pure HA crystallises in the monoclinic space group $P2_1/b$ at temperature below 250°C and in the hexagonal group $P6_3/m$ above 250°C.[21] However, naturally occurring impurities stabilise bone HA in the hexagonal crystalline form, where the OH^- groups are positioned in a statistically disordered fashion along the z-axis.[179] The presence of Posner's clusters can be identified clearly in the crystal structure of HA, where they have trigonal symmetry (C3 in the Schönflies notation) and are stacked along the z-direction, sharing half of their volume.[81] In 2010, using high-resolution cryo-TEM, Dey *et al.* proved the existence of nanometric clusters in simulated body fluids, and found that the formation of apatite on an arachidic acid monolayer involves the aggregation of calcium phosphate clusters of size 0.87 ± 0.2 nm that densify in the proximity of the monolayer.[22] In 2012, using *in situ* AFM to follow the nucleation of ACP on a calcium carbonate surface, Wang *et al.* detected nanometric calcium phosphate aggregates that were linked to the presence of Posner's clusters.[180] More recently, calcium K-Edge XANES experiments further supported Posner's hypothesis by proposing a hydrated form $Ca_9(PO_4)_6(H_2O)_{30}$ as the structural building unit of ACP. Based on their interpretation of the XANES spectra during the early stages of calcium phosphate nucleation, the authors also considered deviations from the idealised cluster and proposed the existence of protonated and Ca-deficient Posner-like clusters, as well as clusters containing monovalent counterions (e.g.: Na^+, Cl^-) to compensate for their negative or positive charges. The authors also observed that phosphate groups can bridge two central calcium ions, thus connecting different clusters.[87,177]

So far few computational studies have investigated the structure of the Posner's cluster. Yin and Stott performed *ab initio* DFT calculations to study the structure and the stability of a Posner's cluster in vacuum and in the presence of some ions normally contained in body fluids (H^+, OH^-, Na^+ and Cl^-). Relaxing the structure of a cluster taken from crystalline HA (*i.e.* having C3 symmetry) the authors found that the cluster loses its symmetry. Moreover, the cluster contracts due to the rearrangement of the ions and the number of Ca−O bonds increases. When the studied cluster binds six protons, it regains the C3 symmetry with an increase in stability; the further addition of six OH^- brings

about the formation of three water molecules, conserving the C3 symmetry; one or two sodium ions can also bind to the cluster, but without providing the same effect on the stability observed in the case of proton addition. The addition of six chlorine ions to the six protonated cluster leads to an energy decrease and retention of the C3 symmetry, but the authors report that the Cl^- are weakly bound to the cluster.[14] Treboux *et al.*[181] used *ab initio* methods (Hartree-Fock level of theory) to compute the stability of $[Ca_3(PO_4)_2]$, its dimer and trimer form and of the Posner's cluster. They found that in vacuum the most stable cluster is the Posner's, with a S6 symmetry, although there are several isomers within an energy range of a few kcal/mol. However, these calculations were conducted in vacuum, using pre-assumed structures for the calcium phosphate clusters. Simulations of the behaviour of Posner's clusters in explicit water are thus required in order to obtain a better understanding of the early stages of ACP formation.

This study reports a molecular dynamics investigation of the aggregation and clustering of calcium and phosphate ions from aqueous solution using a force field model, developed in our group,[5] to describe interatomic interactions in calcium phosphate materials, where polarisation effects are included through a shell-model. The aim is to obtain *unbiased* structural information concerning Posner's cluster that could support the interpretation of experimental data on the processes surrounding the early stages of ACP nucleation and growth.

4.2 Methods

4.2.1 Simulation details

MD simulations were performed using version 4.07 of the DL_POLY computational package.[102] Our MD protocol consisted in an initial 25 ps equilibration in an NVE ensemble (constant number of particles, constant volume and constant energy), followed by a second 25 ps equilibration in an NVT ensemble (constant number of particles, constant volume and constant temperature). The Ca, Na and P atoms were kept frozen to prevent aggregation in a system not yet in equilibrium. Finally, all the ions were let free to diffuse for 5 ns in an NPT ensemble (constant number of particles, constant pressure and con-

stant temperature). A similar simulation protocol has been used previously.[7,168,182] The simulation temperature was set at 310 K (body temperature) and the pressure at 1 bar. The timestep was set to 0.05 fs, which is compatible with the frequency of vibration of the core-shell units. The system cutoff was 8 Å, the Nosé-Hoover algorithm,[131,132] with a relaxation time of 0.1 ps, was employed in both NVT and NPT ensembles to ensure consistency between the equilibration and the production runs. Analysis of the temporal behaviour of the total energy, cell volume and temperature shows that all systems considered in the present study have reached equilibrium within the first 200 ps of the production run.

4.2.2 Interatomic potential model

The interatomic potential model developed by Ainsworth *et al.*[5] was used to describe the intramolecular and interatomic interactions of calcium-sodium-phosphate species. In this force field, phosphate is described as the sum of a Morse and a Coulombic potential, phosphate bond angles by a harmonic potential, and non-bonded interactions by Buckingham potentials. This force field makes use of a shell model in order to model the atom's electronic polarisability, in which each oxygen anion in the phosphate and hydroxyl groups consists of both a core and a massless shell connected by a spring.

The water molecules were described using two models: TIP3P,[147] and the core-shell potential originally developed by De Leeuw and Parker[7] with the modification introduced by Kerisit and Parker.[8] The density at 1 bar and the temperature at 310 K was calculated from a simulation of a 50 × 50 × 50 Å test box in an NPT ensemble for 1 ns for each water model (see Table 4.2). The RDFs of both types of water are reported in Figures 4.1, 4.2, 4.3 and 4.4, a comparison with the experiment is reported in Tables 4.3 and 4.4. The density of water described by core-shell interatomic potentials is 1.263 kg/dm^3, much higher than the experimental value.[183] This value was obtained implementing the hydrogen bond Lennard-Jones potential introduced by Kerisit and Parker[8] to modify the original shell-model of De Leeuw and Parker,[7] which was describing the Ow$_s$-Hw interaction using a Buckingham potential. This update allows to well describe the water-water hydrogen bond and the properties of the water dimer, avoiding the unphysical freezing of water at 300 K and matching the diffusion coefficient with the experimental value. The

structure of the hydration shell and the water residence time in the first hydration shell of calcium and other cations were tested by the authors and they are in very good agreement with the *ab initio* MD and X-ray diffraction results, thus validating the high transferability of the parameters.[8]

Table 4.1 Force field parameters used in this work, from ref.[5,7,8]

Potential parameters				
Atom		**Mass (u)**		**Charge (e)**
Ca		40.078		+2.0
Na		22.9898		+1.0
P		30.9738		+5.0
$O_{ph}c$		15.80		+0.84819
$O_{ph}s$		0.20		-2.84819
Oh_c		15.80		+0.90
Oh_s		0.20		-2.30
H		1.01		+0.40
Ow_c		15.80		+1.25
Ow_s		0.20		-2.05
Hw		1.00		+0.4
Core-Shell Spring $\frac{1}{2}k_{c-s}r^2$				
				k_{c-s} ($\frac{eV}{A^2}$)
$O_{ph}c$	$O_{ph}s$			74.92038
Oh_c	Oh_s			74.92038
Ow_c	Ow_s			209.45
Buckingham				$Ae^{-r/\rho} - Cr^{-6}$
		A (eV)	ρ(Å)	C (eV/A^6)
$O_{ph}s$	$O_{ph}s$	22764.3	0.149	27.88
$O_{ph}s$	Oh_s	22764.0	0.149	13.94
P	$O_{ph}s$	1020.0	0.34322	0.03
P	Oh_s	814.2	0.34322	0.03
P	Ow_s	465.25	0.34322	0.03

Ca	O_{phs}	2152.3566	0.309227	0.09944
Ca	Oh_s	1250.0	0.3437	0.0
Ca	Ow_s	1186.6	0.297	0.0
Na	O_{phs}	56465.3453	0.193931	0.0
Na	Oh_s	858.79	0.3065	0.0
Na	Ow_s	2334.72	0.2387	0.0
O_{phs}	Ow_s	23987.77	0.213	12.09
O_{phs}	Hw	758.468	0.23	0.0
12-6				$(\frac{A}{r_{ij}})^{12} - (\frac{B}{r_{ij}})^6$
		A (eV)	B (eV)	
Ow_s	Ow_s	39344.98	42.15	
nm				$\frac{E_0}{(n-m)}[m(\frac{r_0}{r_{ij}})^n - n(\frac{r_0}{r_{ij}})^m]$
		E_0 (eV)	n m	r_0(Å)
Ow_s	Hw	0.0555555	9 6	1.817121
schrm				$\frac{k}{2}(\theta_{jik} - \theta_0)^2 exp[-\frac{r_{ij}}{\rho_1} + \frac{r_{ik}}{\rho_2}]$
	k	θ_0	ρ_1	ρ_2
O_{phs} P O_{phs}	3.3588	109.47	10000000.0	10000000.0

Table 4.2 Density of the water obtained after equilibration at 1 bar and 310 K.

Model	**Final density [kg/dm^3] at 1 bar 310 K**
Shell-model	1.263
TIP3P	1.026
Experiment[183]	0.993

No significant differences in terms of either aggregation or shape of the calcium phosphate species were observed when comparing the results obtained using the TIP3P and core-shell water models. The results reported in the manuscript have been obtained using the water shell model,[8] as the polarisability of the oxygen ions is explicitly included,[142] and it allows better compatibility with the force field for the phosphate ions, which are also described using a core-shell potential. The parameters of this force field are listed in Table 4.1. We also simulated a single Posner's cluster by means of *ab initio* MD in a water box

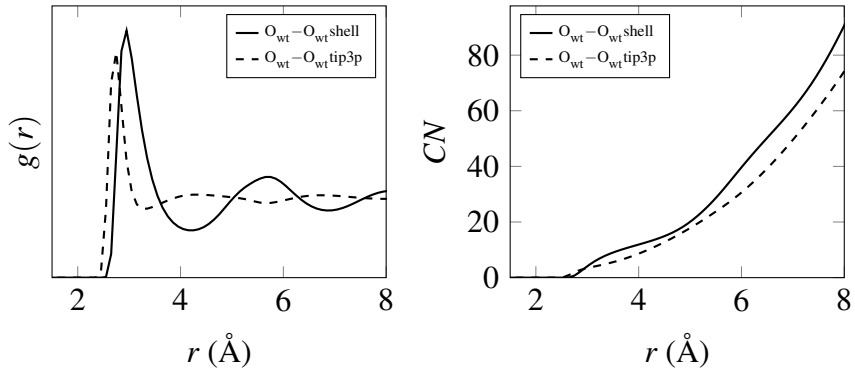


Figure 4.1 Radial distribution function and integration number for the pairing $O_{wt}-O_{wt}$

Table 4.3 $O_{wt}-O_{wt}$ distances for first and second coordination sphere: comparing the two water model tested with the experiment.[9]

Water model	**I peak (Å)**	**I minimum (Å)**	**II peak (Å)**	**II minimum (Å)**
Shell-model	2.95	4.25	5.75	6.85
TIP3P	2.75	3.35	4.25	5.75
Experiment[9]	2.73	3.41	4.44	5.51

containing 101 water molecules using the same level of theory employed in chapter 3 and compared the RDFs with those produced by the classical MD. Considering the atom pairs Ca_c-Ca, Ca_c-O, Ca_c-P, $P-P$, $P-O$ and $O-O$ (where Ca_c is the central calcium of the Posner's cluster), we saw that the interatomic potentials (IP) reproduce very well the positions of the first maxima and fairly well the second coordination spheres. The comparisons are reported in Figure 4.5 and Tables 4.5, 4.6 and 4.7.

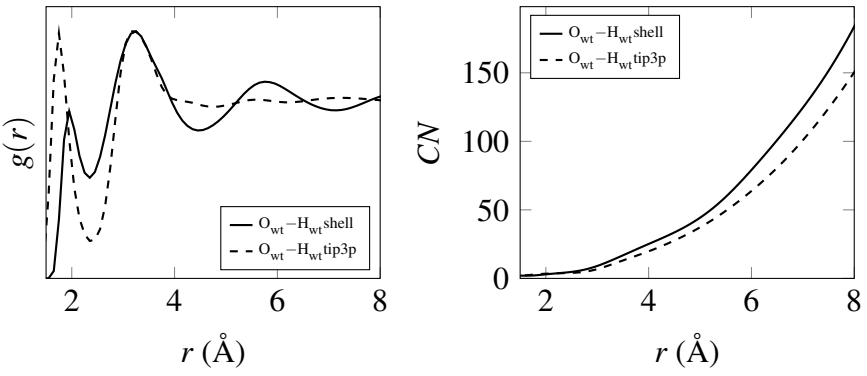


Figure 4.2 Radial distribution function and integration number for the pairing $O_{wt}-H_{wt}$

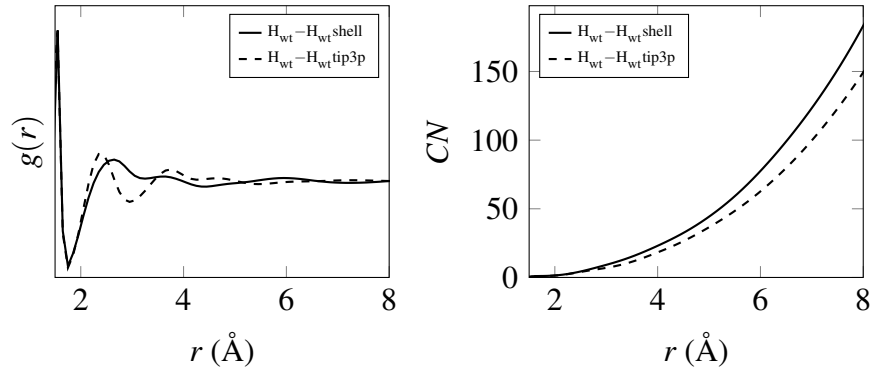


Figure 4.3 Radial distribution function and integration number for the pairing $H_{wt}-H_{wt}$

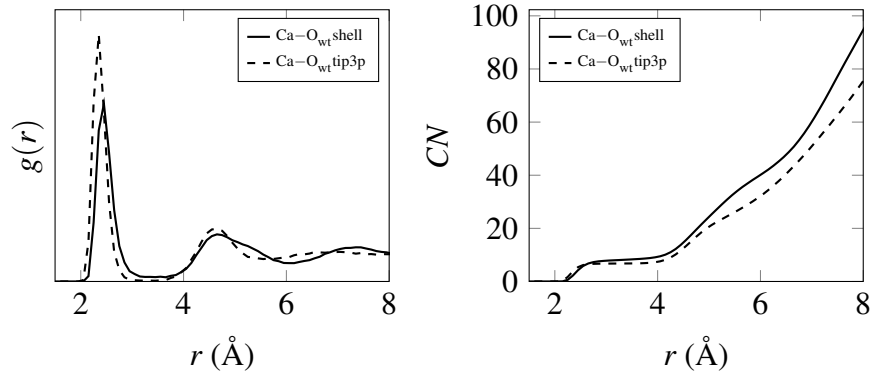


Figure 4.4 Radial distribution function and integration number for the pairing $Ca-O_{wt}$

Table 4.4 $Ca-O_{wt}$ distances for first and second coordination sphere: comparing the two water model tested with the experiment.[10,11]

Water model	**I peak (Å)**	**I min (Å)**	**II peak (Å)**	**II min (Å)**	**I coord. no.**
Shell-model	2.45	3.45	4.65	6.05	7.3
TIP3P	2.35	3.25	4.65	5.65	6.7
Experiment	2.4	3.1			7.1

Table 4.5 Comparison of the atom pair average distances obtained with *ab initio* MD as described in chapter 3, Classical MD and the experimental values of Betts and Posner.[12] All distances are reported in Å.

	Ca_c-Ca	**Ca_c-O_{ph}**	**Ca_c-P**	**$P-P$**	**$P-O_{ph}$**	**$O_{ph}-O_{ph}$**
Ab initio	3.65-4.15	2.45	3.75	4.75-5.65	1.55	2.55
IP	3.75	2.45	3.55	4.85-5.25	1.55	2.55
Exp. (ACP)	–	2.3	3-6	3-6	1.5	2.3

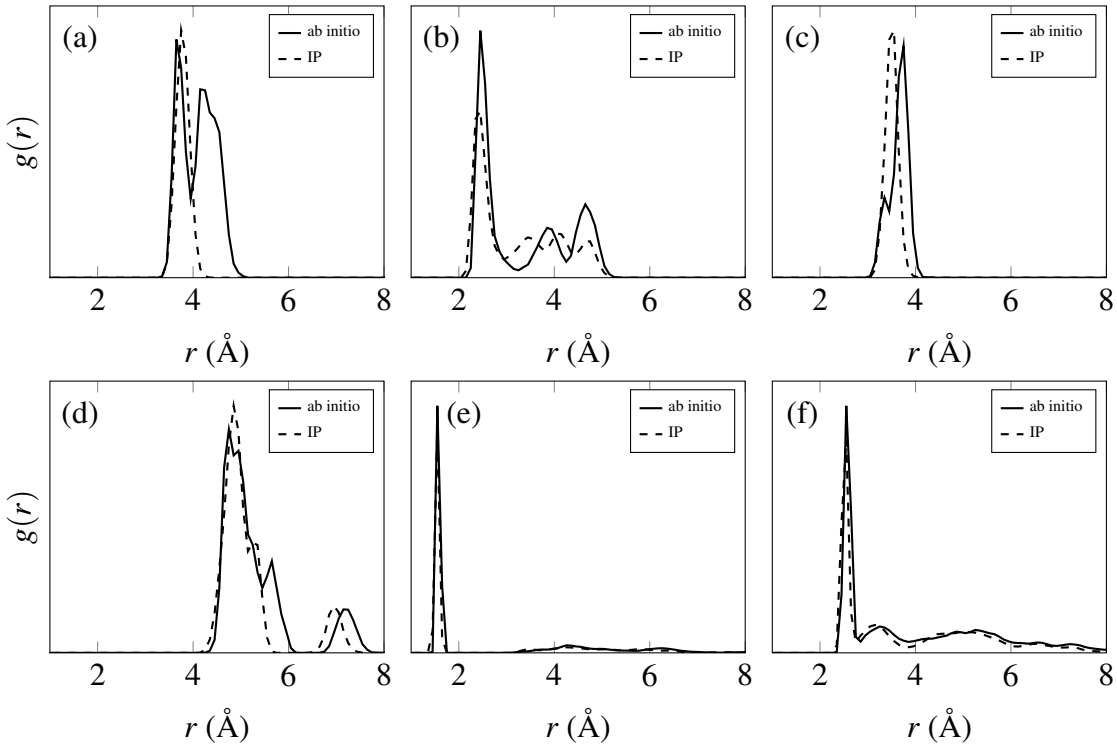


Figure 4.5 Radial distribution functions for the atom pairs (a) Ca_c-Ca, (b) Ca_c-O_{ph}, (c) Ca_c-P, (d) P−P, (e) $P-O_{ph}$ and (f) $O_{ph}-O_{ph}$ in a single Posner's cluster in water.

Table 4.6 Comparison of the coordination numbers obtained with *ab initio* MD as described in chapter 3, Classical MD and the values proper of crystalline HA.[13] All distances are reported in Å.

	$\mathbf{Ca_c-Ca}$	$\mathbf{Ca_c-O_{ph}}$	$\mathbf{Ca_c-P}$	**P−P**	$\mathbf{P-O_{ph}}$	$\mathbf{O_{ph}-O_{ph}}$
Ab initio	3.95 - 8.00	4.75	5.55	3.80	4.00	1.95
IP	8.00	3.75	5.55	3.75	4.00	2.70
HA	8	6	3-6	4	4	0

Table 4.7 Comparison of the Posner's cluster radius in *ab initio* MD in water as described in chapter 3, Classical MD in water and DFT in gas phase performed by Yin *et al.*[14]

Method	**Radius (Å)**
Ab initio MD in water	5.25[a]
Classical MD in water	5.25[a]
DFT in gas phase	4.35

[a]The value refers to the last minimum in the $Ca-O_{ph}$ RDF of a single solvated Posner's cluster.

4.2.3 Computational protocol

Supersaturation is the thermodynamic driving force for the crystal nucleation of a solid phase from a solution[184] and body fluids are supersaturated with respect to HA. However, a simulation box truly representative of the concentrations of calcium and phosphate ions found in physiological solutions (2.5 mmol/l) is computationally unfeasible as the number of particles would be too large and the timescale required to observe spontaneous aggregation of calcium phosphate particles too long. As such, artificially high supersaturation conditions were applied to induce the nucleation process.[185] Table 4.8 summarises the characteristics of the simulated solutions considered in the present study. Solution I corresponds to a calcium concentration, $[Ca^{2+}]$, equal to 0.65 mol/l generated by adding 44 Ca^{2+} ions in a cubic box dimension 50 $\times$ 50 $\times$ 50 Å, together with a mixture of $H_2PO_4^-$ and HPO_4^{2-} with ratio 2:3, both normally present at pH 7.4 (i.e. body fluid condition).[186] Lower Ca^{2+} concentrations of 0.1 mol/l (8 Ca^{2+}) and 0.3 mol/l (20 Ca^{2+}) were tested running an MD for 5 ns, but, although ion aggregation and clustering were observed, a statistically relevant population of clusters was not formed. Solution II only contains Ca^{2+} and PO_4^{3-} ions in order to be consistent with the composition of ACP. Solution III contained higher $[Ca^{2+}]$ and $[PO_4^{3-}]$ in order to induce the formation of larger calcium phosphate clusters. In body fluids and plasma, Na^+ ions are far more concentrated than Ca^{2+}.[187] Consequently, we considered the effect of sodium ions on the aggregation of phosphate species in solution by replacing half Ca^{2+} in solutions I, II and III with Na^+. In Table 1, these solutions are named Ia, IIa and IIIa, respectively.

Table 4.8 Summary of the six prepared solutions. The starting cell length was 48.846 Å in all cases. The reported $[Ca^{2+}]$ and $[Na^{+}]$ concentrations are calculated using the volume of the cell after 1 ns in an NPT ensemble

	S (I)	S (II)	S (III)	S (Ia)	S (IIa)	S (IIIa)
Ca^{2+}	44	45	150	22	22	75
$[Ca^{2+}]$ (mol/l)	0.65	0.66	2.32	0.33	0.32	1.17
Na^{+}	0	0	0	44	46	150
$[Na^{+}]$ (mol/l)	0.00	0.00	0.00	0.66	0.67	2.34
$H_2PO_4^{-}$	22	0	0	22	0	0
HPO_4^{2-}	33	0	0	33	0	0
PO_4^{3-}	0	30	100	0	30	100
H_2O	4497	4679	4136	4477	4652	4044
pH	7.4	14	14	7.4	14	14

4.2.4 Structural criteria for the detection of Posner's clusters

The MD trajectories of the solutions considered in the present study were analysed in order to identify the portions of the calcium-phosphate aggregates forming during the simulations which resemble the Posner's cluster in Figure 4.6(a). In our simulations in water, however, many factors can alter the structure of the highly symmetric cluster identifiable in the crystal:

- Isolated, the cluster loses its symmetry;[81] our Classical MD using the interatomic potentials described in this chapter and *ab initio* MD of a single Posner's cluster in water with the level of theory described in chapter 3 confirmed this finding;
- At pH 7.4, in our simulations the Ca/P ratio, dictated by the balance of positive and negative charges, was equal to 0.8, so only Ca-deficient clusters were obtainable;
- Other ions present in solution (in our case Na^{+}) may substitute calcium.

It is evident that in our MD simulation and, generally, in physiological solution, we can not expect to detect perfectly symmetric clusters having the same stoichiometry of those found in crystalline HA, therefore, we introduced the concept of *Posner-like clus-*

ter, which is a distorted, non-stoichiometric cluster which resemble the original Posner's cluster. In particular, a group of ions were defined as a Posner-like cluster when 6 to 8 peripheral (Ca + Na) ions and 6 phosphate groups surrounded a central calcium ion (Ca_c). Cation-deficient clusters (*i.e.* 6 or 7 out of 8 cations) were also considered, because these calculations focus on the early stages of aggregation and further Ca binding occurs at a second stage.[23] A calcium (or a sodium) ion or a phosphate group was considered to be part of a Posner-like cluster when the Ca_c-Ca, Ca_c-Na and Ca_c-P distances were shorter than 5.0 Å. This value is larger than the position of the first minimum of the RDFs of the Ca_c-Ca (4.35 Å) and Ca_c-P (4.15 Å) pairs obtained from MD simulations of a single Posner's cluster in water. As such, slightly distorted Posner-like clusters were also considered during the analysis.

Phosphate groups can bridge different Ca_c, thus contributing to the formation of more than one Posner-like cluster. The $P-Ca_c$ coordination number (*CN*) during the simulations was computed using the formula:

$$CN(P-Ca_c) = \frac{N(\text{clusters}) \times 6}{N(P)} \tag{4.1}$$

where N(clusters) is the number of detected Posner-like clusters at each MD step and N(P) is the total number of phosphorus atoms (equivalent to the number of phosphate groups) that are part of Posner-like clusters. The value of $CN(P-Ca_c)$ equal to 1 corresponds to completely isolated clusters, *i.e.* not linked to another calcium-phosphate species by a phosphate group.

Crystalline HA is made up of superimposed Posner's clusters with opposite chirality. Each phosphate group bridges two Ca_c in the *xy* plane, as shown in Figure 1.9. Along the *z*-direction, Posner's clusters are stacked, *i.e.* when considering a cluster centred at the Ca_c site at level z=0, its neighbours are centred respectively at level z=+1/2 and level z=-1/2 and share three phosphate groups with the one centred at level z=0; meaning that in the *z*-direction each phosphate bridges two additional Ca_c.[81] Given that in a single unit cell of HA there are six phosphorus atoms and four Ca_c sites, which are at the centre of different Posner's clusters, then in the crystal structure the $CN(P-Ca_c)$ is equal to:

$$CN(P-Ca_c)_{HA} = \frac{4 \times 6}{6} = 4 \tag{4.2}$$

which means that each phosphate group in HA coordinates four different Ca_c.

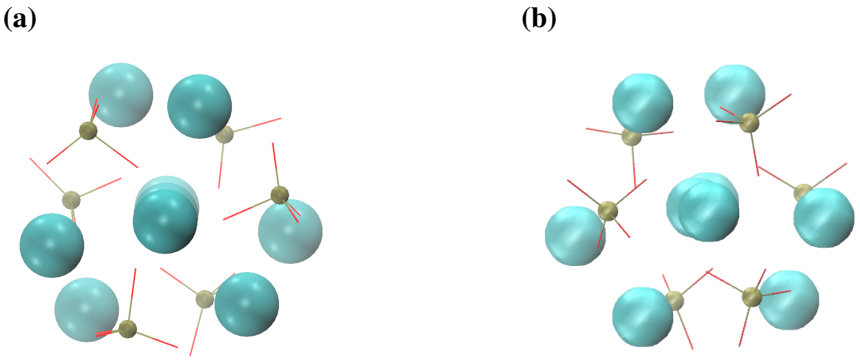


Figure 4.6 Projection on the xy plane of (a) Posner's cluster in hydroxyapatite, C3 symmetry; (b) Posner's cluster in water, simulated with the force field of Ainsworth *et al.*[5] Colour key Ca:light blue, P:green, O:red.

4.3 Results

4.3.1 Symmetry of the Posner's cluster

The unit cell of HA, $Ca_{10}(PO_4)_6(OH)_2$, contains 10 Ca atoms, four of which can be named, according to their coordination environment, Ca1 and the remaining six Ca2 (see Figure 4.7).[82] Ca1 form columns that are perpendicular to the xy plane, whereas Ca2 are arranged in triangles around the hydroxyl groups, which form hexagonal channels along the z-direction. Each Ca1 is surrounded by 9 oxygen atoms belonging to six phosphate groups, whereas the Ca2 coordinate to six oxygen atoms from phosphates and one hydroxyl group.

Figure 4.6(a) shows the structure of a Posner's cluster within the crystalline structure of HA. The cluster consists of three calcium ions of type Ca1 stacked in the centre, with the middle one (Ca_c) coordinating six phosphate groups. The negative charge is then compensated by further coordination of six calcium ions of type Ca2 in HA; the symmetry of the cluster is C3. Because of the superimposition of the Posner clusters, in crystalline HA each Ca1 is also a Ca_c. The SymmetryTool plugin[1] of the Visual Molecular Dynamics (VMD) software was used to analyse the symmetry of the Posner's cluster in Figure 4.6(a), with coordinates taken from the Crystallographic Information File of HA.[13] The

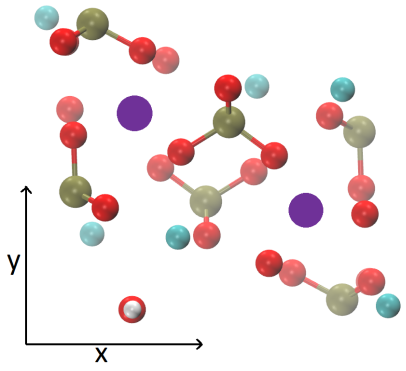


Figure 4.7 Representation of a single unit cell of HA ($Ca_{10}(PO_4)_6(OH)_2$) projected on the xy plane. Each purple circle represents two aligned Ca1. Colour key Ca1:purple, Ca2:light blue, P:green, O:red, H:white.

Ca atoms alone, or the combination Ca+P, gives C2 symmetry, whereas a C3 symmetry is obtained when considering only P atoms, only O atoms, Ca+O atoms, P+O atoms or the entire cluster.

MD simulations of a single Posner's cluster in a 50 Å water box were then conducted using the force field of Ainsworth *et al.*[5] when a loss of symmetry occurs as a result of the bonding of water to the outer layer atoms of the cluster (see Figure 4.6(b)). According to Yin *et al.*,[14] in vacuum the relaxation of the Posner's cluster causes a loss of the C3 symmetry, a contraction of the cluster radius to 4.35 Å, and an increase of the Ca−O coordination number. Conversely, the cluster radius increases to $\approx$ 5.25 Å during the simulation in water, which is due to the decrease of the Ca−O intracluster coordination as the outer layer of Ca ions interacts with water molecules (see Tables 4.5, 4.6 and 4.7). Integration of the $Ca_c - O_{wt}$ RDF (Ca_c is the central calcium of the Posner's cluster and O_{wt} is the oxygen of water) gives an average of 25 water molecules in the first solvation sphere of the Posner's cluster. A recent Ca K-edge XANES study on wet-ACP redefined the idealised cluster model for the structural unit of ACP originally postulated by Du *et al.*[87] as a cluster composed by a central Ca^{2+} and eight peripheral Ca^{2+} ions, each of which is coordinated to three to four water molecules.[177] This leads to a total number of water molecules in the range 24-32, in agreement with the results obtained from classical MD simulation.

4.3.2 Aggregation of Ca^{2+}, $H_2PO_4^-$ and HPO_4^{2-} ions in solution

In this section, the aggregation of calcium and phosphate ions in solution at neutral pH, which is close to physiological conditions, is reported (solution I in Table 4.8).

In this solution the Ca/P ratio is equal to 0.8. Experimental studies have shown that the aggregation of calcium and phosphate ions results in the formation of Ca-deficient clusters, which then incorporate other calcium ions from the solution, thereby increasing the Ca/P ratio, releasing protons and consequently lowering the pH of the solution.[23,188] Xie and co-workers have also suggested that Ca-deficient clusters form, because of the lower dehydration energy of the hydrogen phosphate (-299 kcal mol^{-1}) and dihydrogen phosphate (-68 kcal mol^{-1}) ions compared with Ca^{2+} (-381 kcal mol^{-1}).[188] In the original Posner's cluster, the Ca/P ratio is 1.5 but such a large amount of calcium is not available in the solutions I and Ia and Posner-like clusters made of six phosphates and nine cations were therefore not observed during the simulations. Thus, the minimum amount of calcium ions in a calcium-phosphate aggregate to be considered a Posner-like cluster was set to seven (Ca_c + 6 Ca). This criterion is based on the symmetry considerations mentioned above, namely six phosphate groups are enough to provide Posner's C3 symmetry and on the fact that Ca-deficient clusters have been detected experimentally during the early stages of CaP nucleation.[188] For consistency, the same thresholds for Ca-Ca and Ca-P coordination numbers were used to analyse the other solutions considered in this study.

Figure 4.8 reports the variation of the Ca−Ca and Ca−P coordination numbers during the simulation of solution I. The colour scale corresponds to the fraction of the total Ca having the coordination number specified on the *y* axis. An increase in coordination number is due to the loss of water molecules coordinated to Ca^{2+} ions and the aggregation process. Note that this process takes place during the first 0.5 ns. Throughout the simulation, the Ca-Ca coordination remains low, as more than 50% of Ca ions have four or fewer other Ca ions in close proximity (Figure 4.8(a)), 35% or more Ca ions are coordinated to four phosphate groups and about 25% were surrounded by six phosphates (Figure 4.8(b)), as in the Posner's cluster (Figure 4.6(a)). Analysis of the simulation of solution I detect an average of fewer than two clusters, probably because of the low Ca/P ratio (0.8) in the solution.

These clusters have symmetry C1 and are characterised by only 6-7 calcium ions, rather than 8 as in the Posner's model,[12] and they are mainly formed of HPO_4^{2-} rather than $H_2PO_4^-$ ions. This is probably due to the larger population at neutral pH of the hydrogen phosphate anions, but it could also be explained in terms of the preference of Ca^{2+} to coordinate to the double negatively charged monohydrogenphosphate ion.

Figure 4.9(a) displays the structure of the calcium phosphate aggregate at 5 ns, with the atoms forming Posner-like clusters in colour and the atoms not participating in the clusters in grey. In Figure 4.9(b) it is possible to discern a single Posner-like cluster with formula $[Ca_7(H_2PO_4)(HPO_4)_5]^{3+}$, where the central Ca coordinates six phosphate groups. The phosphates are arranged in a pseudo-octahedral geometry, whereas the outer layer is composed of six calcium ions which do not show any particular order. The three Ca ions in the centre of the cluster are not properly aligned as in the HA crystal (see Figure 4.6), but the effect of the solvating water needs to be taken into account, which displaces the ions causing the cluster to lose the C3 symmetry, as observed during the MD simulation of a single Posner's cluster in water (Figure 4.6(b)).

(a) **(b)**

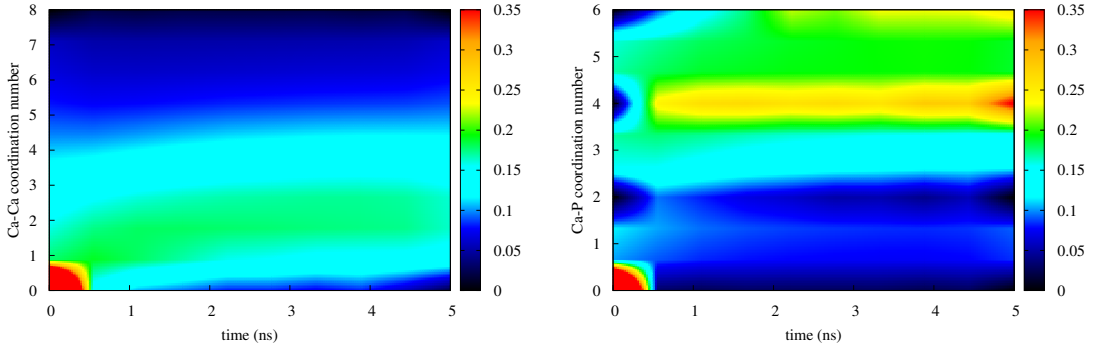


Figure 4.8 Solution (I): (a) Ca-Ca and (b) Ca-P coordination during 5 ns of simulation in an NPT ensemble. The *y* axis measures the number of Ca and P coordinated to a given Ca. The colour scale represents the fraction of the total Ca ions with a given coordination number at x time.

Bone HA is a highly substituted material. The amounts of CO_3^{2-}, F^-, Na^+ and Mg^+ depend on the bone site and the age of the subject.[189] Half of the sodium present in a human body is stored in the bones, where it preferentially substitutes Ca2 in the HA lattice and is mainly located at the bone surface.[190]

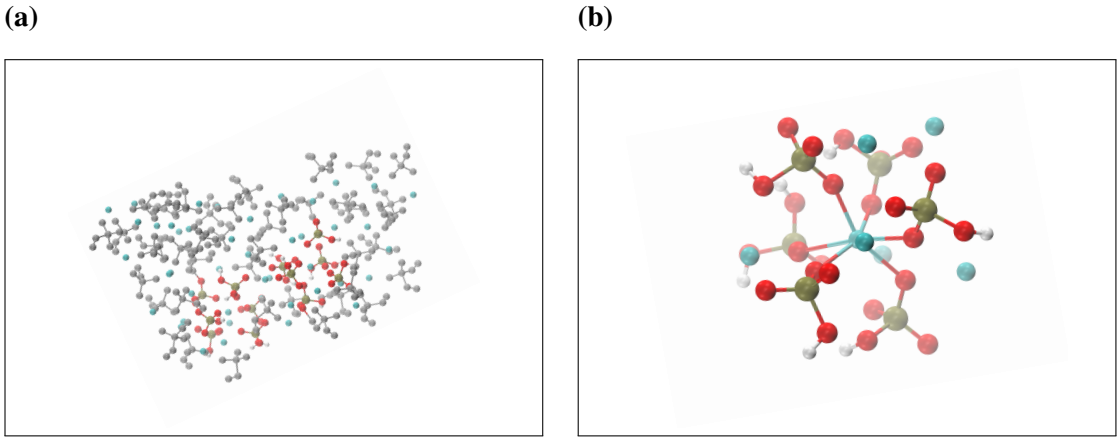


Figure 4.9 Solution (I): (a) Aggregates after 5 ns; (b) Detail of one of the Posner-like clusters formed during the simulation with formula $[Ca_7(H_2PO_4)(HPO_4)_5]^{3+}$. Colour key Ca:light blue, P:green, O:red, atoms not forming Posner's clusters:grey.

Figure 4.10 reports clusters obtained from the simulation of Solution Ia (see Table 4.8), where half of Ca^{2+} are substituted by Na^+. The presence of Na^+ ions in solution inhibits the formation of large aggregates and only two medium size aggregates (≈15 Å and 24 Å in their longest direction) grow during the simulation (Figure 4.10(a)). Posner-like clusters contain up to three Ca ions around the central one and Na ions occupy some of the positions vacated by Ca. In Figure 4.10(b) a cluster with formula $[Ca_3Na_4(H_2PO_4)_3(HPO_4)_3]^+$ is shown, where only one of the Ca^{2+} forming the outer layer of the cluster is aligned with the Ca_c and the other cations are packed in a disorderly fashion around the phosphate groups. Another example of a cluster obtained during the simulation is shown in Figure 4.10(c), $[CaNa_6(HPO_4)_6]^{-4}$, where only one central calcium ion is surrounded by six phosphate groups with six sodium ions partially counterbalancing the negative charge. Lastly, in Figure 4.10(d) two clusters that have three phosphate groups bridging the two Ca_c are presented. In this structure, the average $CN(P{-}Ca_c)$ is equal to 1.33 according to Equation 4.1.

The formation of positively and negatively charged clusters should not come as a surprise, as they are part of a larger aggregate that is not classified as a Posner-like structure, according to the criteria discussed in subsection 4.2.4. Moreover, charged entities in solution have been detected experimentally during the early stages of calcium phosphate formation.[23]

Figure 4.11 shows the lifespan of the Posner-like clusters during 6.5 ns of trajectory. Each stripe refers to a Posner-like cluster centred on a different calcium ion. A stripe is shadowed when the Posner-like cluster is satisfying the criteria for Ca-P and Ca-Ca distances and coordination numbers, discussed in section 2.4, whereas it is white for frames where the Posner-like cluster is not satisfying any of the criteria (i.e. it is not yet formed or it suffers structural distortions). In a solution of Ca^{2+}, $H_2PO_4^-$ and HPO_4^{2-}, a total of six different clusters appears (six horizontal stripes in the plot), whereas after introducing Na^+ ten clusters are formed. As such, the plots in Figure 4.11 suggest that sodium not only favours the formation of Posner-like clusters in a solution containing protonated phosphate species, but it also increases the structural stability of the clusters, evidenced by stripes with more regions coloured in grey or black. We reason that sodium ions are able to effectively counter-balance the low negative charges of the protonated phosphates and are a valid substitute to calcium.

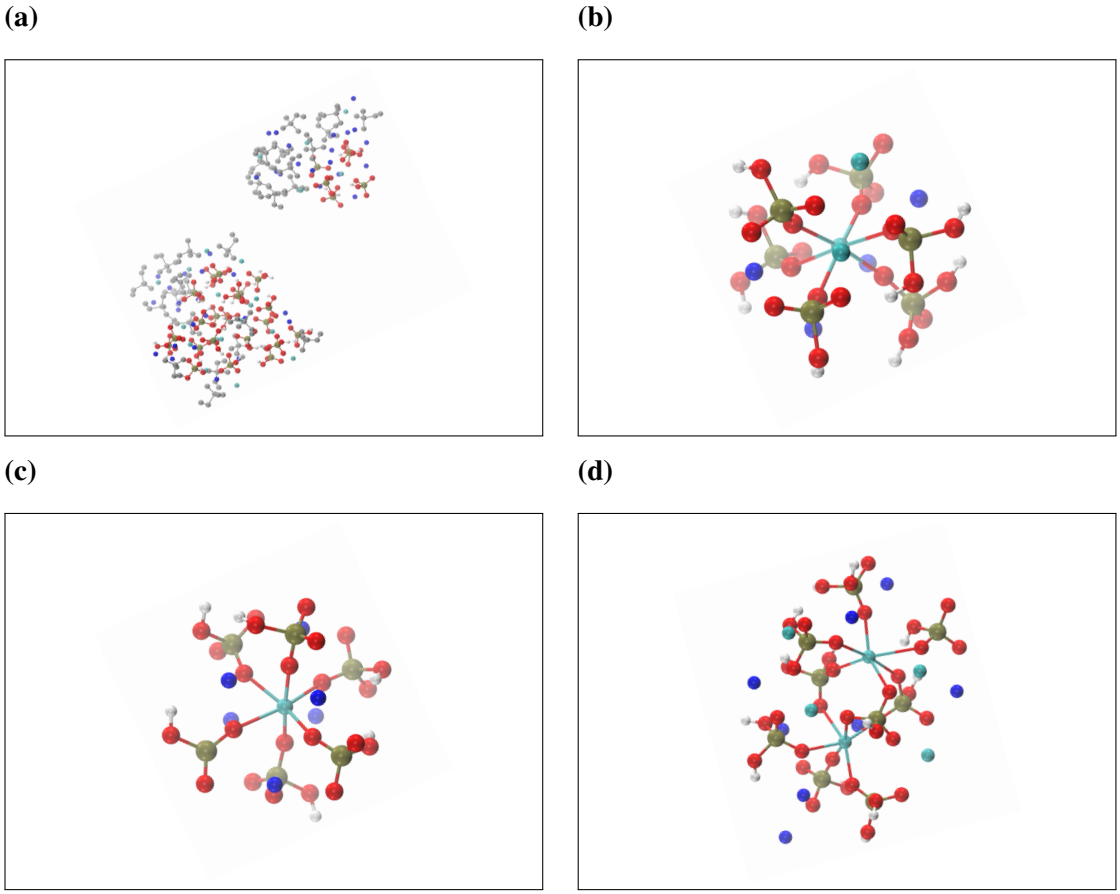


Figure 4.10 Solution (Ia): (a) Aggregates after 5 ns; Posner-like clusters with chemical formulas (b) $[Ca_3Na_4(H_2PO_4)_3(HPO_4)_3]^+$ and (c) $[CaNa_6(HPO_4)_6]^{-4}$. (d) Two Posner-like clusters sharing three phosphate groups, formula $[Ca_6Na_8(H_2PO_4)_4(HPO_4)_5]^{+6}$. Colour key Ca:light blue, Na:dark blue, P:green, O:red, Atoms not forming Posner's clusters:grey.

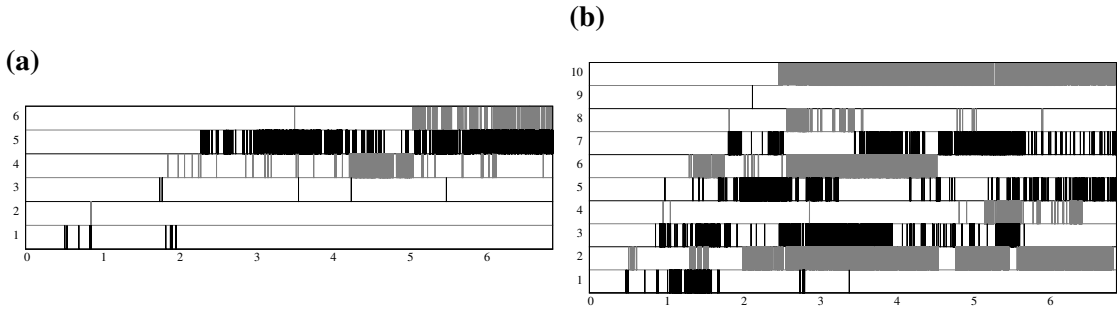


Figure 4.11 Number of different clusters and their lifespan during 6.5 ns of trajectory, (a) Solution (I); (b) Solution (Ia). Each stripe is shaded in grey or black for the time intervals when the cluster is formed. For a better view, clusters are depicted alternatively in black or grey colour. The number of stripes corresponds to the total number of observed clusters in the trajectory.

4.3.3 Aggregation of Ca^{2+} and PO_4^{3-}

Our simulations of the aggregation of Ca^{2+} and PO_4^{3-} mimic the subsequent stages of the aggregation process, after the release of the protons bound to the phosphate species. No OH^- ions were added to the solution, as the first complexes originating from solution are calcium phosphate ion associations and the OH^- incorporates in a further stage of HA crystallisation.[168]

Solution II only contains Ca^{2+} and PO_4^{3-} ions and the Ca/P ratio corresponds to that in the original Posner's cluster and in ACP (Ca/P=1.5). After 5 ns of simulation, five different Posner-like clusters form in solution: neutral ones with the same stoichiometry as Posner's cluster (see Figure 4.12(b)) and Ca-deficient complexes, $[Ca_8(PO_4)_6]^{2-}$, occurring when adjacent clusters share phosphate groups.

(a) **(b)**

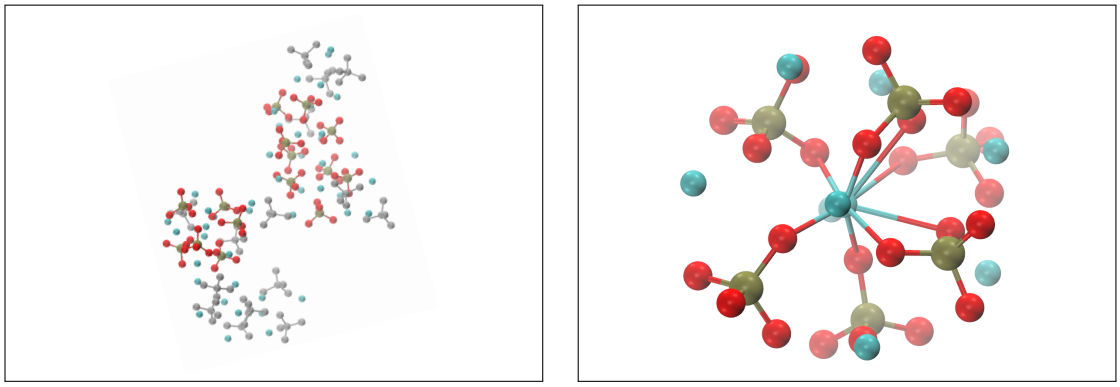


Figure 4.12 Solution (II): (a) Aggregates after 5 ns; (b) Detail of one of the Posner-like clusters formed with formula $Ca_9(PO_4)_6$. Colour key Ca:light blue, P:green, O:red, Atoms not forming Posner's clusters:grey.

When substituting half of the calcium ions by sodium (Solution IIa), a Posner-like cluster with formula $[Ca_3Na_6(PO_4)_6]^{-6}$ forms during the first 4.5 ns of the trajectory (Figure 4.13(b)). This cluster is formed by six phosphate groups and only three Ca ions aligned in the centre of the cluster, whereas six sodium ions complete the clusters occupying the outer cationic positions. A second cluster, sharing part of the phosphate ions with the first one, appears after 4.5 ns, giving a final chemical formula $Ca_7Na_{16}(PO_4)_{10}$ and a $CN(P-Ca_c)$ of 1.20.

Calcium and sodium have similar ionic radii, and thus Na could potentially occupy any

cationic position in the Posner-like clusters. Nevertheless, analysis of the MD simulations shows a preference by Na to position itself in the outer cationic positions, *i.e.* the HA Ca2 sites, confirming the findings of El Feki *et al.* who experimentally studied substitutions of sodium and carbonate in HA.[191] This also agrees with the preference of other impurity cations such as Mg^{2+} and Zn^{2+} for the Ca2 sites.[82,192]

Contrary to the case discussed in section 4.3.2, sodium disfavours clustering upon complete deprotonation of the phosphate groups. This opposite tendency can be explained in terms of the higher negative charge of PO_4^{3-}, compared to HPO_4^{2-} and $H_2PO_4^{-}$, which require doubly charged cations to counterbalance the charge. Consequently, substitution of Ca^{2+} by Na^{+} provokes a reduction in the clustering process.

(a) **(b)**

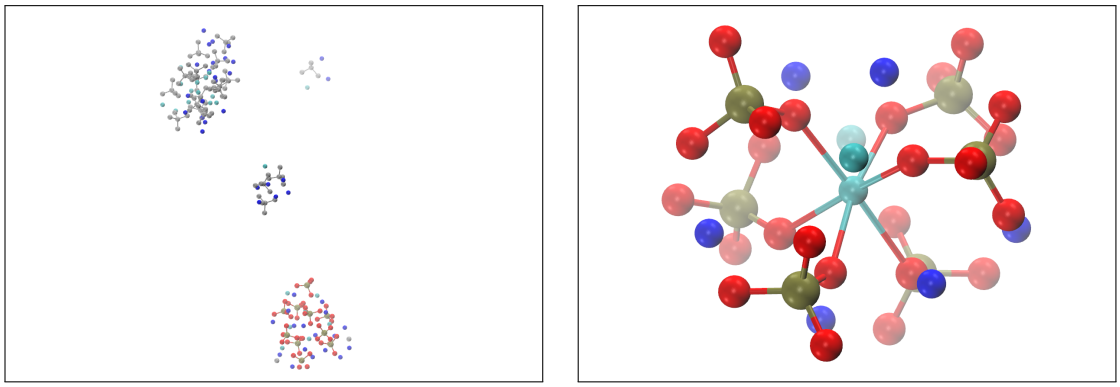


Figure 4.13 Solution (IIa): (a) Aggregates after 5 ns; (b) Detail of one of the Posner-like clusters formed with formula $[Ca_3Na_6(PO_4)_6]^{-6}$. Colour key Ca:light blue, Na:dark blue, P:green, O:red, Atoms not forming Posner's clusters:grey.

In order to promote the formation of larger clusters, the concentration of calcium ions in solution was increased to approximately $[Ca^{2+}]$=2.3 mol/l. The formation of several Posner-like clusters takes place already during the first ns of simulation. These clusters are stable throughout the simulation, although their structure distorts during the dynamics and does not, therefore, satisfy the criteria previously imposed for them to be classified as Posner's cluster. Figure 4.14(a) shows the structure of the aggregates after 5 ns, whose shape is needle-like instead of spherical as in the solution containing Ca^{2+}, HPO_4^{-} and $H_2PO_4^{2-}$ (Figure 4.9(a)) and in the solution containing Ca^{2+}, Na^{+}, HPO_4^{-} and $H_2PO_4^{2-}$ (Figure 4.10(a)). Figure 4.14(b) shows an example of the cluster with formula $Ca_9(PO_4)_6$

obtained during the simulation, where three Ca1 are properly aligned and the central calcium ion is coordinated to nine oxygens as in HA. Comparison of the distances with those reported by Laurencin *et al.* for HA[82] indicates that the Ca1-O distances, shown in Figure 4.14(c), are slightly longer. In particular, we have only three oxygens at less than 2.55 Å and two at distances larger than 3.0 Å from the calcium, whereas these authors have reported six out of nine oxygens below 2.55 Å and three oxygens at 2.77 Å. The symmetry analysis with the VMD Symmetry tool gives an Oh point group when considering the phosphorus atoms or the calcium ions, with a tolerance of 0.23; taking into account both Ca and P at the same time, we obtained the Oh point group with a tolerance of 0.28.

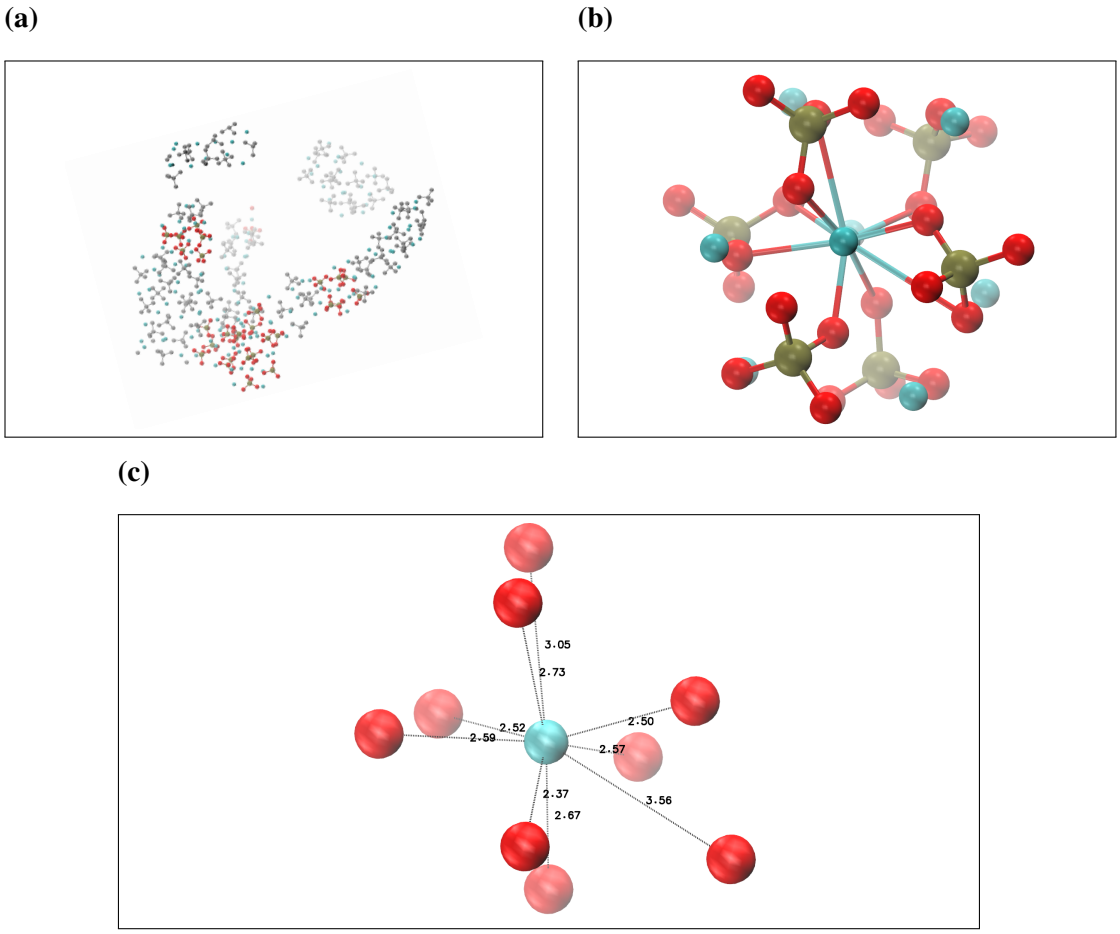


Figure 4.14 Solution (III): (a) Aggregates after 5 ns; (b) Detail of one of the Posner-like clusters formed with formula $Ca_9(PO_4)_6$; (c) Distances between the central Ca1 and the neighbouring oxygens. Colour key Ca:light blue, P:green, O:red, Atoms not forming Posner's clusters:grey.

In the presence of sodium with a Na:Ca ratio of 2:1, the aggregates are similar in shape to those obtained without sodium (see Figure 4.15(a)) but a smaller number of different

Posner-like clusters forms (see Figure 4.16). The clusters have a lower level of symmetry when compared to those obtained without sodium, where C2 is the point group with the highest symmetry obtained considering Ca and P sites; the C2 cluster is shown in Figure 4.15(b). Other clusters obtained in the same trajectory, however, belong to the point group C_s, or they present no elements of symmetry at all. In general, the calcium ions have to be aligned with the central one (*i.e.* Ca1 in HA) in order to confer some sort of symmetry.

(a) **(b)**

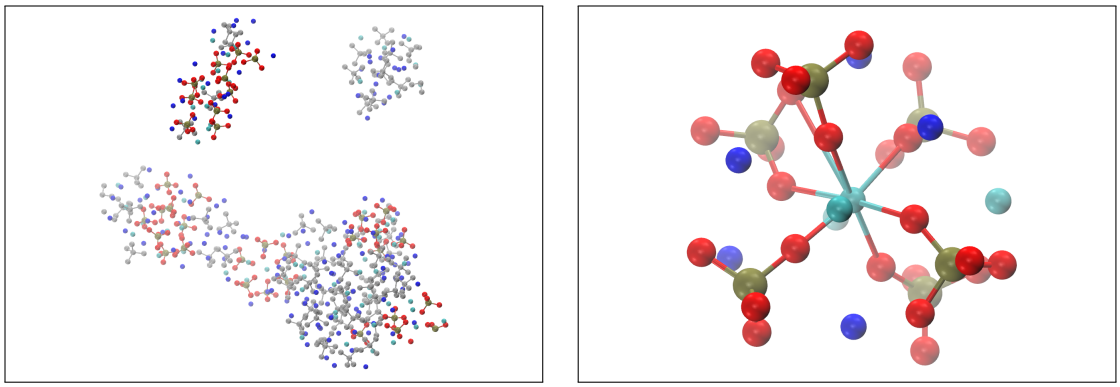


Figure 4.15 Solution (IIIa): (a) Aggregates after 5 ns; (b) Detail of one of the Posner-like clusters formed with formula $[Ca_4Na_5(PO_4)_6]^{-5}$. Colour key Ca:light blue, Na:dark blue, P:green, O:red, Atoms not forming Posner's clusters:grey.

During the aggregation of Ca^{2+} and $PO_4{}^{3-}$, the effect of sodium is opposite to what we observe in solutions I and Ia: there are more different Posner-like clusters without sodium. Furthermore, under these conditions sodium shows itself to be a valid substitute for calcium in the external layer of the Posner-like clusters, as we experience various degrees of Na/Ca substitutions. We observe sharing of some of the phosphate ions among the clusters, as also found in the solution Ia.

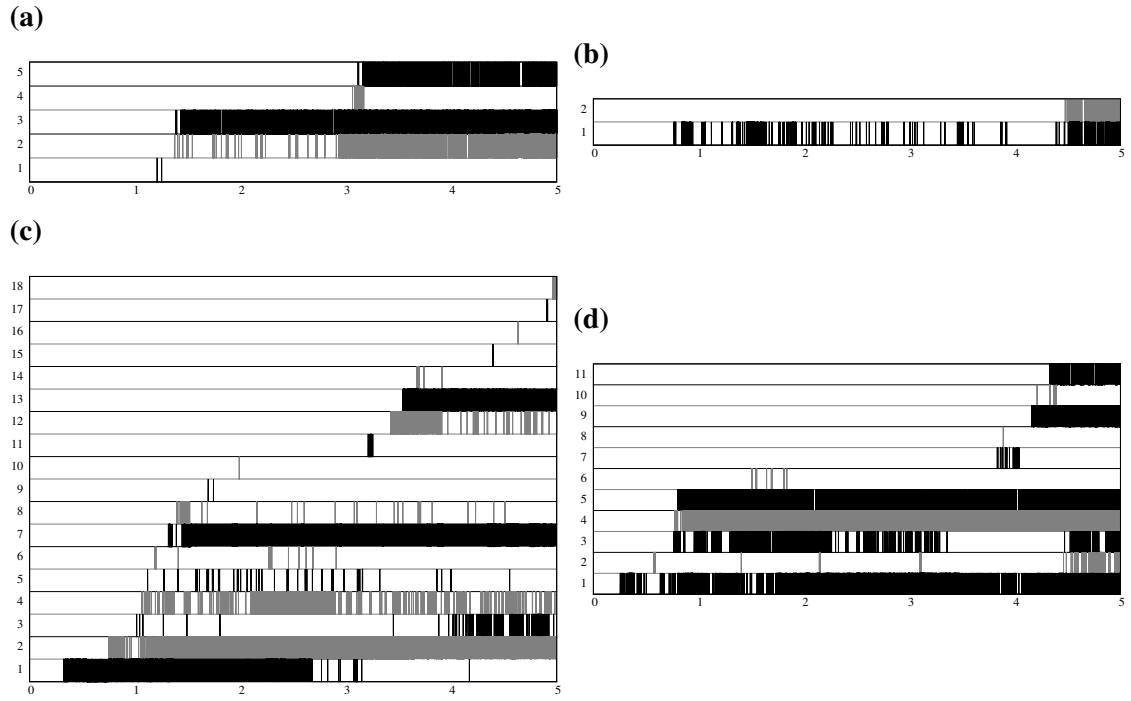


Figure 4.16 Number of different clusters and their appearance during 5 ns of trajectory, (a) Solution (II); (b) Solution (IIa); (c) Solution (III); (d) Solution (IIIa). Each stripe is shaded in grey or black for the time intervals when the cluster is formed. For a better view, clusters are depicted alternatively in black or grey colour. The number of stripes corresponds to the total number of observed clusters in the trajectory.

4.4 Discussion

Under the solution conditions considered in the present study, the aggregation of calcium phosphate species takes place spontaneously during the first ns of the simulations and the coordination environment around the calcium ions does not show significant variations thereafter (see Figure 4.8). The aggregates are spherical in the presence of protonated phosphate species, whereas they tend to be more fragmented and needle-like when constituted by PO_4^{3-} ions. In general, they are flexible and may resemble the liquid-like ionic polymers observed for homogeneous nucleation of calcium carbonate by Demichelis *et al.*[86]

The position of the first peaks in the Ca_c-Ca, Ca_c-P, Ca_c-O_{ph} and Ca_c-Na RDFs of the Posner-like clusters detected during the simulations are in very good agreement with the RDF of the single Posner cluster in water (see the dashed line in Figures 4.17, 4.18, 4.19, 4.20, 4.21 and 4.22). The Ca_c-P first peak positions also match in the presence

of protonated phosphates (solution I, Figure 4.17) and the introduction of Na^+ does not influence the Ca_c-Ca, Ca_c-P or Ca_c-O_{ph} distances for the Posner-like clusters (solution Ia, Figure 4.20, solution IIa, Figure 4.21 and solution IIIa, Figure 4.22).

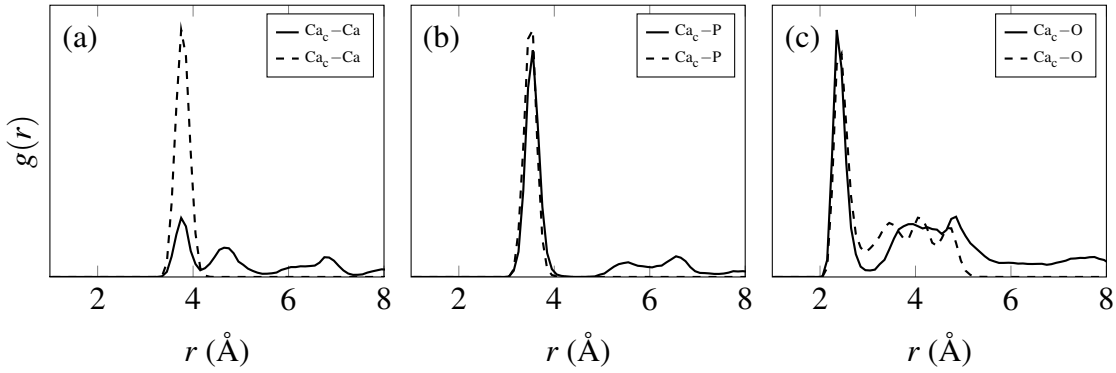


Figure 4.17 Solution I: RDF of (a) Ca_c-Ca, (b) Ca_c-P, (c) Ca_c-O. The RDF of the single Posner's cluster in water are reported as a dashed line for comparison.

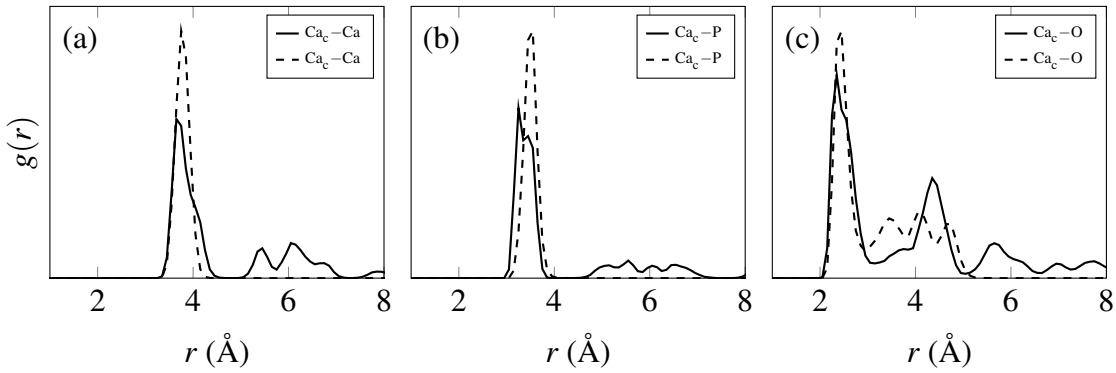


Figure 4.18 Solution II: RDF of (a) Ca_c-Ca, (b) Ca_c-P, (c) Ca_c-O. The RDF of the single Posner's cluster in water are reported as a dashed line for comparison.

To form HA, each phosphate group needs to bridge four Ca1, resulting in a $CN(P-Ca_c)$ equal to 4, where Ca_c is equivalent to Ca1 in the crystal structure of HA (all the Ca1 atoms are the centres of Posner clusters in a stoichiometric crystal) and represents the centre of a Posner cluster. The $CN(P-Ca_c)$ is, therefore, an important measure of the progression from the early aggregates of calcium phosphate in solution to the formation of the solid phase. Table 4.9 reports the average $CN(P-Ca_c)$ for each simulated solution considered in the present study, which have been calculated applying Equation 4.1 to the frames containing two or more clusters and some degree of phosphate bridging (*i.e.* $CN(P-Ca_c) > 1$). The $CN(P-Ca_c)$ do not present significant differences between solutions I and Ia.

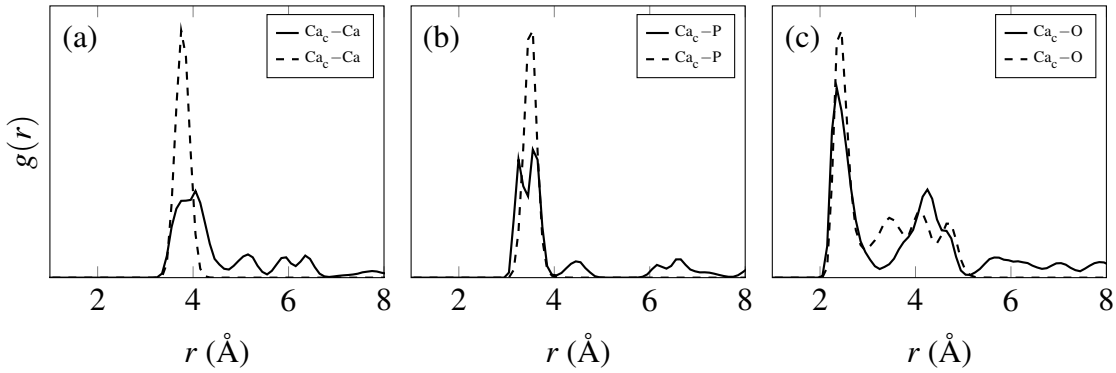


Figure 4.19 Solution III: RDF of (a) Ca_c-Ca, (b) Ca_c-P, (c) Ca_c-O. The RDF of the single Posner's cluster in water are reported as a dashed line for comparison.

Considering solutions II and IIa, the $CN(P-Ca_c)$ is higher in the absence of sodium. Comparing the highly concentrated solutions III and IIIa, the $CN(P-Ca_c)$ is much larger in the presence of sodium, which is probably due to the inability of sodium to efficiently compensate the highly negative charges of the orthophosphate anions; here the clusters need to share part of the phosphate groups in order to maximise the interaction with the fewer Ca^{2+}ions. Contrary to what could be expected, the $CN(P-Ca_c)$ decreases in more concentrated solutions ($CN(P-Ca_c)$ in solution II and IIa is larger than in solutions III and IIIa), because a higher concentration leads to the formation of more clusters, which are not necessarily bridged, as they are diluted in a larger aggregate.

In the neutral solutions I and Ia, which simulate the process taking place in the early stages of calcium phosphate aggregation in body fluids, sodium can efficiently substitute calcium to generate a larger number of Posner-like clusters. Interestingly, the Ca-deficient clusters and the spherical aggregates obtained from the simulations fit into the post-nucleation stage discussed by Habraken *et al.* who described spherical aggregates with a Ca/P ratio equal to 0.67 and post-nucleation clusters with formula $[Ca_2(HPO_4)_3]^{2-}$.[23] Analysis of the simulations indeed reveals that (1) the central calcium may coordinate fewer than 8 outer calcium ions, (2) the Ca/P ratio during the nucleation stage can be lower than 1.5 (typical of ACP), (3) other ions normally present in body fluids can substitute calcium and phosphates (in particular we verified the ability of sodium to replace calcium) and (4) some phosphate ions act as bridging ligands connecting different Posner-like clusters. In fact, the Posner-like clusters obtained in our simulations validate the assumptions of Du *et al.* that non-idealised clusters exist in solution.[87] Lastly, the formation of clusters

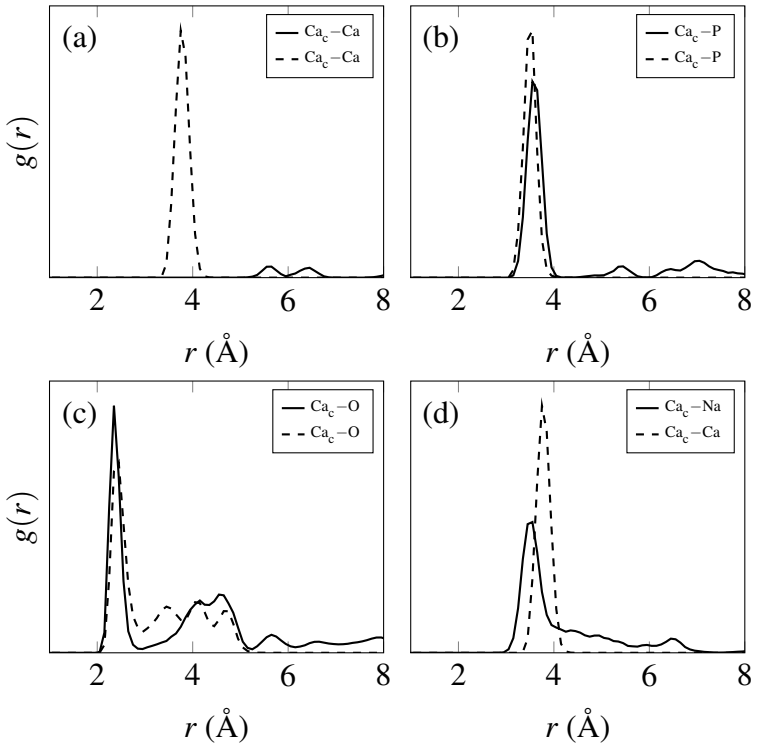


Figure 4.20 Solution Ia: RDF of (a) Ca_c-Ca, (b) Ca_c-P, (c) Ca_c-O, (d) Ca_c-Na. The RDF of the single Posner's cluster in water are reported as a dashed line for comparison.

containing both $H_2PO_4^-$ and HPO_4^{2-} is in agreement with experiment, because NMR studies show that HPO_4^- ions are present at the bone surface,[193] and a certain degree of phosphate protonation (5-15%) is typical of ACP originating from a neutral solution.[194] Figure 4.23 reports the percentage of trajectory frames with zero, one or more Posner-like clusters. This graph suggests that the increase in the Ca/P ratio from 0.8 (solution I) to 1.5 (solution II) leads to the formation of more Posner-like clusters. This is evidenced by the increase in the number of MD frames with one or more clusters from 10.6% to 16.3%.
The aggregates originated in solutions I and Ia resemble the apatite-like structure of superficial bone tissue in contact with body fluids, where some of the phosphate ions are protonated and the sodium ions occupy a number of Ca2 sites.[189] We observed that upon addition of Na ions, the proportion of frames with one or more Posner-like clusters significantly increases from 10.6% to 40% and up to five clusters form simultaneously in solution Ia, with Na effectively occupying Ca2 positions. At the same time, the aggregates from solution II are more like the deep bone mineral phase, which is poor in hydrogen phosphate and sodium. The analysis of solutions II and IIa shows a clear decrease in the formation of Posner-like clusters in the presence of sodium ions: the percentage of

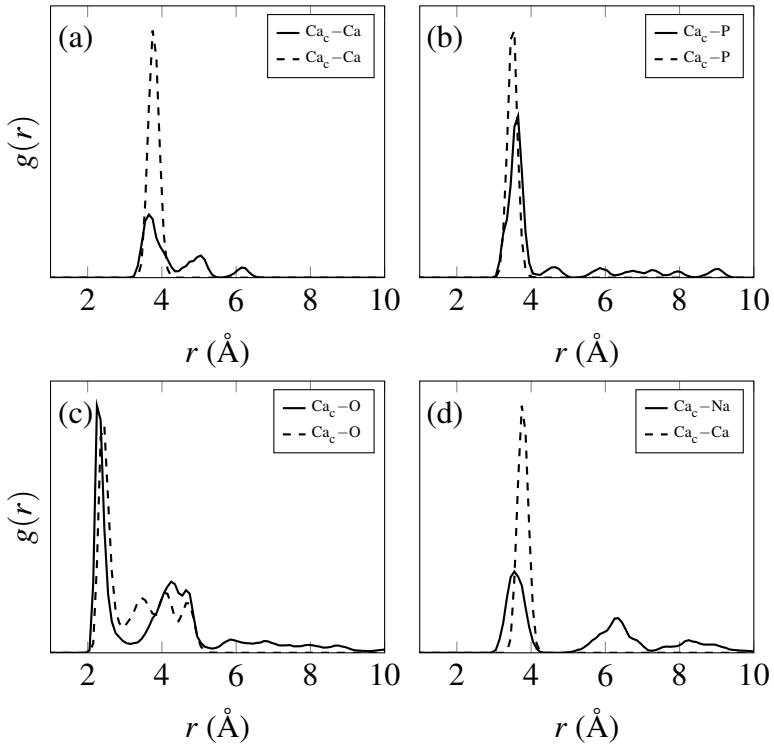


Figure 4.21 Solution IIa: RDF of (a) Ca_c-Ca, (b) Ca_c-P, (c) Ca_c-O, (d) Ca_c-Na. The RDF of the single Posner's cluster in water are reported as a dashed line for comparison.

frames with one or more Posner-like clusters decreases from 16.3% (solution II) to 7.4% (solution IIa). Our data suggest that the effect of sodium on the formation of the mineral phase depends on the counter ions present in solution: sodium is enhancing the clustering in the presence of HPO_4^{2-} and $H_2PO_4^-$, while it is inhibiting clustering in presence of PO_4^{3-}. This trend is supported by experimental findings,[189] which show the presence of Ca, Na and protonated phosphates at the surface of bones and of Ca and PO_4^{3-} in the deep tissue. We explained this phenomenon considering the charges of the ions: the 3- negative charge of PO_4^{3-} is more efficiently counterbalanced by Ca^{2+} than Na^+. Therefore, the Na^+ ions, that help to form Posner-like clusters during the early stage of nucleation, should be replaced by Ca^{2+} in order to promote the formation of Posner-like clusters as the hydrogenated phosphates start to release the protons, explaining why sodium is found only in traces in bone tissue.[195] The expulsion of a proton from an $H_2PO_4^-$ or an HPO_4^{2-} leads to a more negatively charged phosphate ion, which is able to coordinate Ca^{2+} in solution, raising the Ca/P ratio of the ion association. Provided that the newly incorporated Ca^{2+} was already part of a calcium phosphate aggregate, the phosphate acts as a bridging ligand between two calciums, increasing its P-Ca coordination towards the value of 4 as

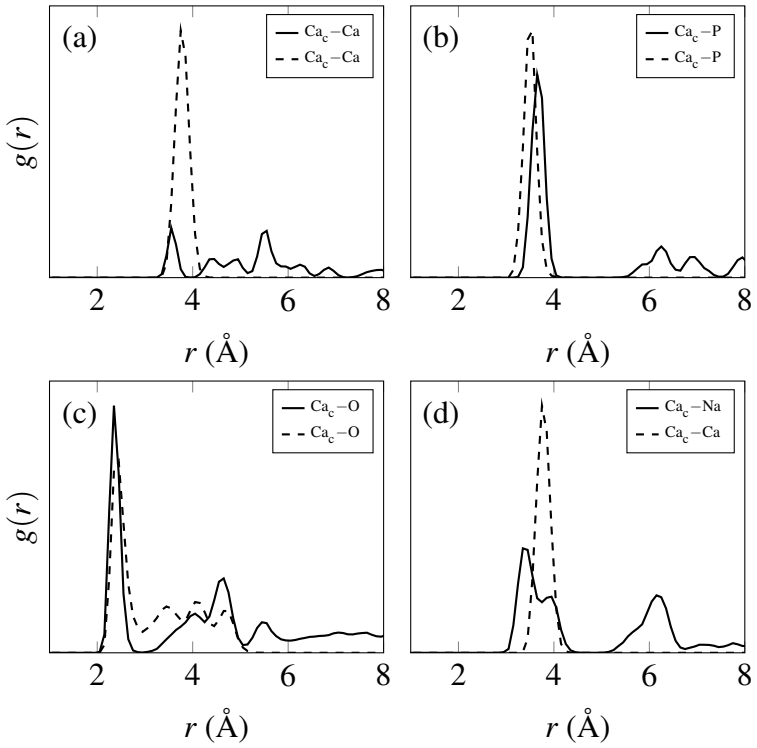


Figure 4.22 Solution IIIa: RDF of (a) Ca_c-Ca, (b) Ca_c-P, (c) Ca_c-O, (d) Ca_c-Na. The RDF of the single Posner's cluster in water are reported as a dashed line for comparison.

found in HA.

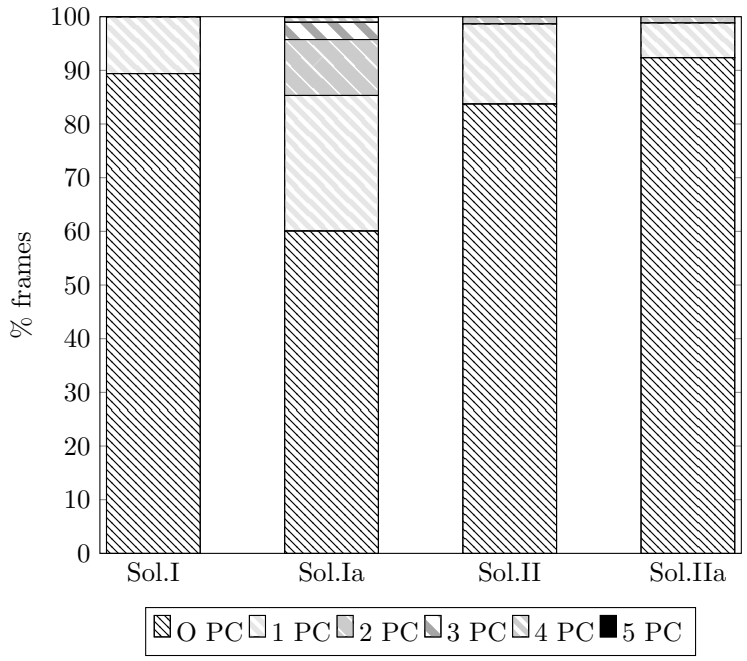


Figure 4.23 Percentage of trajectory frames in which 0, 1, 2, 3, 4 or 5 Posner-like clusters (PC) are detected at the same time. The bars refer to the solutions I, Ia, II and IIa, which have a $[Ca^{2+} + \frac{1}{2}Na^{+}] \approx 0.65$ mol/l

Table 4.9 Average CN(P$-$Ca$_c$) calculated over all the frames presenting bridging phosphates

Solution	**CN(P$-$Ca$_c$)**
Solution (I)	1.20 ± 0.00^a
Solution (II)	1.35 ± 0.06
Solution (III)	1.06 ± 0.01
Solution (Ia)	1.17 ± 0.07
Solution (IIa)	1.20 ± 0.00^a
Solution (IIIa)	1.14 ± 0.04

[a]Zero standard deviation means that we obtained the same CN(P$-$Ca$_c$) for all the frames presenting bridging phosphates.

4.5 Conclusions

In the present work, the aggregation of calcium and phosphate ions in water at body temperature and physiological pH was studied by means of classical Molecular Dynamics simulations. The aggregation of Ca^{2+} and PO_4^{3-} in solution was also considered, in order to reproduce the Ca/P ratio and composition found in solid amorphous calcium phosphate. Under neutral pH conditions, Posner-like clusters, deficient in calcium ions and composed of protonated phosphate groups form spontaneously during the simulation after a few ns. Sodium ions in solution, which are normally present in body fluids, were also considered and it was found that they can replace calcium in the outer layer of the Posner's cluster. The symmetry of the clusters originated at neutral pH is C1 (as it is in vacuum[14]), although the alignment of two calcium ions with the central one and the presence of deprotonated phosphate groups confers a higher symmetry to the cluster. Two or more clusters can share part of their ions, since one or more of their phosphate groups bridge two central calciums, as already observed in experiment[87] and during MD simulations of Ca^{2+}, PO_4^{3-} and OH^- clustering at a collagen template.[168]

The simulations here reported represent the first theoretical investigation of the structure of the Posner's cluster in solution and confirm the formation of these species in the early stages of calcium phosphate crystallisation from solution, supporting previous experimental findings.

Chapter 5

Calcium phosphate deposition on planar and stepped (101) surfaces of anatase TiO_2

5.1 Introduction

The first application of titanium in orthopaedic surgery dates back to the early fifties.[196] Nowadays, due to its high corrosion- and wear-resistance, lightness, the absence of allergic reactions and its mechanical reliability in aqueous environment, titanium and its alloys are commonly employed for load-bearing prostheses, such as hip joint and bone replacements and artificial teeth.[35,68]

Unfortunately, titanium does not bond to the living tissue. For this reason, to promote its integration and minimise the risks related to the release of metallic ions into the body, titanium is usually coated by HA, i.e. the calcium phosphate CaP which constitutes the main inorganic component of bone and tooth tissues. HA is commonly deposited as a 40-200 μm thick layer onto the metallic surface using the plasma spraying technique.[61] The coating can also be prepared through electrophoretic deposition,[62] sol-gel deposition,[63] electrochemical deposition,[64] sputtering,[65] hydrothermal synthesis[197] and biomimetic deposition.[66] The last technique consists of soaking the metallic implant in simulated body fluids (SBFs) at body temperature, to obtain a ceramic coating which is very similar to the bone mineral.[198]

The high corrosion-resistance of the material stems from the passivation of titanium by a thin layer of its oxide. The most common polymorphs of TiO_2 are rutile and anatase.[70] Although rutile is the thermodynamically most stable bulk phase, at a nanometric scale the surface energy for anatase particles is lower than that of rutile particles of the same size,[72,199] and this increased stability of nanoparticulate anatase is the reason behind our choice of the anatase polymorph for this study.

Classical MD simulations employing IP can provide access to sizeable systems over time scales in the order of nanoseconds at a relatively low computational cost. Contrary to the much more expensive *ab initio* methods, the classical procedure relies on parameters calculated in advance, which reduce complex interatomic interactions to simple analytic equations. The derivation of these parameters is often the most significant effort in obtaining a classical description of the chemical processes.

In this work, we have consistently merged a force field for phosphate-based glasses[5] with the Matsui-Akaogi potential for anatase and rutile[6] as well as a set of parameters describing the interaction of their surfaces with water.[17] We found that this combination describes both the Ca^{2+} and $PO_4{}^{3-}$ interactions with the most stable (101) surface of anatase to an acceptable degree compared to density functional theory calculations (DFT). Moreover, we have used the derived force field to obtain significant insight into the aggregation of Ca^{2+} and $PO_4{}^{3-}$ ions on anatase. In summary, we propose the first classical description of the deposition processes occurring in a CaP-containing aqueous solution in contact with a TiO_2 surface.

5.2 Methods

The results reported in this work have been performed using both DFT and IP methods with periodic boundary conditions.

5.2.1 Density functional theory calculations

The DFT code Vienna Ab initio Simulation Package (VASP), version 5.4.1,[103–106] was employed to derive the first-principles data needed to benchmark the proposed force field. We have used the generalized gradient approximation as derived by Perdew, Burke and

Ernzerhof (PBE)[100] to account for the exchange-correlation energy. The PBE functional was combined with the Grimme correction to include the contribution of long-range dispersion interactions to the total energy of the system. We have used the latest modification of the Grimme's method, which incorporates geometry information into the *ab initio* parametrisation of the dispersion coefficients and three-body terms to correct the overbinding of previous methods with only two-body contributions.[200–203] Energies were converged with Monkhorst-Pack *k*-points meshes of $8\times8\times4$ for the unit cell of anatase and $4\times4\times1$ for the slab representing the (101) surface of anatase.[204] Only the valence electrons were treated explicitly using a plane wave basis set with a cut-off energy of 400 eV. We have used the projected-augmented-wave method (PAW) to describe the nodal features of the valence electrons and their interaction with the inner core of the atom.[124,125] The convergence criterion for the electronic self-consistent iterations was set at 10^{-5} eV, whereas the threshold for the ionic forces was set at 0.03 eV/Å. To improve the electronic convergence, we have used the Gaussian smearing method with a band width of 0.1 eV.[205,206]

5.2.2 Interatomic potentials simulations

To the best of our knowledge, a force field to study the deposition of CaP on a TiO_2 surface has not yet been reported. We overcame this limitation by combining a number of IP already in the literature. The CaP was modelled by the force field developed by Ainsworth *et al.*[5], whilst the shell-model for water developed by De Leeuw and Parker[7] with the modification introduced by Kerisit and Parker[8] allowed us to introduce the aqueous medium.[207] For the description of the TiO_2, we selected the force field of Matsui and Akaogi,[6] which has been successfully employed by many authors[17,199,208–212] and gives unit cell parameters and the bulk modulus of anatase in excellent agreement with the experimental and DFT values (see Table 5.1).

Table 5.1 Anatase bulk parameters, unit cell volume and bulk modulus obtained (i) after optimisation with VASP at the GGA (PBE+D3) level of theory, (ii) optimisation with GULP using the Matsui-Akaogi force field[6] and (iii) experimental values.[15,16] Ti−O distances and angles in the optimised bulk are also compared.

	GGA (PBE+D3)	**IP**[6]	**Experiment**
a=b (Å)	3.792	3.770	3.785[16]
c (Å)	9.560	9.568	9.519[16]
Volume (Å^3)	137.49	136.00	136.37[16]
Bulk mod. (GPa)	175	176[a]	178±1[15]
Ti−O[b] (Å)	1.94	1.93	1.966[214]
Ti−O (Å)	1.99	1.99	1.937[214]
O−Ti−O	92.42°	92.49°	92.604°[214]
O−Ti−O	101.85°	102.02°	102.308°[214]

[a] Value obtained using the Hill's approximation.[213]

[b] Anatase has two non-equivalent Ti−O bonds.

In both CaP and TiO_2 force fields, the atoms are subjected to short-range two-body interactions described by Buckingham potentials. Similarly to the procedure employed by other authors,[215] we chose to use the same potential form for the missing cross terms, scaling the A parameter for the interactions between attractive (repulsive) atoms proportionally (inversely proportionally) to the ratio between the oxygen charges of TiO_2 and $PO_4{}^{3-}$, labelled respectively as O_{Ti} and O_{ph}:

$$A(\mathrm{Ti{-}O_{ph}}) = A(\mathrm{Ti{-}O_{Ti}}) \times \frac{q(\mathrm{O_{ph}})}{q(\mathrm{O_{Ti}})} \tag{5.1}$$

$$A(\mathrm{O_{Ti}{-}O_{ph}}) = A(\mathrm{O_{Ti}{-}O_{Ti}}) \times \frac{q(\mathrm{O_{Ti}})}{q(\mathrm{O_{ph}})} \tag{5.2}$$

$$A(\mathrm{Ca{-}O_{Ti}}) = A(\mathrm{Ca{-}O_{ph}}) \times \frac{q(\mathrm{O_{Ti}})}{q(\mathrm{O_{ph}})} \tag{5.3}$$

$$A(\mathrm{P{-}O_{Ti}}) = A(\mathrm{P{-}O_{ph}}) \times \frac{q(\mathrm{O_{Ti}})}{q(\mathrm{O_{ph}})} \tag{5.4}$$

The parameters for the water-TiO_2 interaction were taken from Alimohammadi *et al.*[17], who derived the Lennard-Jones parameters for the Extended Single Point Charge

(SPC/E)[216] water molecules interacting with TiO_2 surfaces described by the Matsui-Akaogi potentials.[6] The oxygen charge of water in the SPC/E and the shell-model differ only by 0.0476 e and for this reason, we did not scale the interactions between the oxygens of water and anatase. All the parameters used in this work are reported in Table 5.2.

Table 5.2 Force field parameters from refs.[5,6,8,17,18] The shadowed lines report the parameters we obtained applying equations 1-4.

Potential parameters				
Atom		**Mass (u)**		**Charge (e)**
Ti		47.867		+2.196
O_{Ti}		16.0		-1.098
Ca		40.078		+2.0
P		30.9738		+5.0
$O_{ph}c$		15.8		+0.84819
$O_{ph}s$		0.2		-2.84819
Ow_c		15.8		+1.25
Ow_s		0.2		-2.05
Hw		1.0		+0.4
Core-Shell Spring $\frac{1}{2}k_{c-s}r^2$				
				k_{c-s} ($\frac{eV}{A^2}$)
$O_{ph}c$	$O_{ph}s$			74.92038
Ow_c	Ow_s			209.45
Buckingham				$Ae^{-r/\rho} - Cr^{-6}$
		A (eV)	ρ(Å)	C (eV/A^6)
Ti	Ti	31090.38	0.154	5.24
Ti	O_{Ti}	16941.28	0.194	12.57
Ti	$O_{ph}s$	30858.43	0.194	12.57
Ti	Ow_s	1238.72	0.276	6.41
O_{Ti}	O_{Ti}	11771.46	0.234	30.19
O_{Ti}	$O_{ph}s$	6462.53	0.234	30.19

Ca	O_{Ti}	1181.64	0.309227	0.09944
Ca	O_{phs}	2152.3566	0.309227	0.09944
Ca	Ow_s	1186.6	0.297	0.0
P	O_{Ti}	559.98	0.34322	0.03
P	O_{phs}	1020.0	0.34322	0.03
P	Ow_s	465.25	0.34322	0.03
O_{phs}	O_{phs}	22764.3	0.149	27.88
O_{phs}	Ow_s	23987.77	0.213	12.09
O_{phs}	Hw	758.468	0.23	0.0
LJ				$4\varepsilon_{ij}[(\frac{\sigma_{ij}}{r_{ij}})^{12}-(\frac{\sigma_{ij}}{r_{ij}})^{6}]$
		σ (Å)	ε (eV)	
O_{Ti}	Ow_s	3.446	0.0067	
12-6				$(\frac{A}{r_{ij}})^{12}-(\frac{B}{r_{ij}})^{6}$
		A (eV)	B (eV)	
Ow_s	Ow_s	39344.98	42.15	
nm				$\frac{E_0}{(n-m)}[m(\frac{r_0}{r_{ij}})^{n}-n(\frac{r_0}{r_{ij}})^{m}]$
		E_0 (eV)	n m	r_0(Å)
Ow_s	Hw	0.0555555	9 6	1.817121
schrm				$\frac{k}{2}(\theta_{jik}-\theta_0)^2 exp[-\frac{r_{ij}}{\rho_1}+\frac{r_{ik}}{\rho_2}]$
	k	θ_0	ρ_1	ρ_2
O_{phs} P O_{phs}	3.3588	109.47	10000000.0	10000000.0

We ran the MD simulations using the DL_POLY package, version 4.07,[102] which easily allows the use of a customised force field. Each simulation consisted of an equilibration run of 25 ps in the NVE ensemble followed by 25 ps in the NVT ensemble, during which the Ca^{2+} and PO_4^{3-} ions and the TiO_2 slab were kept frozen to allow the water to relax and a subsequent 1 ns production run in the NPT ensemble, where all species were free to move.[207] The radial distribution functions (RDFs) were collected between 0.75 and 1 ns. We chose a temperature of 310 K (body temperature) and a pressure of 1 bar. The timestep was set to 0.05 fs, which is compatible with the frequency of vibration of the core-shell units. We employed a cutoff of 8 Å and the Nosé-Hoover algorithm,[131,132]

with a relaxation time of 0.1 ps for both the thermostat and the barostat.

5.2.3 Anatase (101) planar and stepped

The (101) plane is the most stable surface of anatase (Figure 5.1(a)). It presents two- and three-coordinated oxygen atoms (O_{2c} and O_{3c}, respectively) and five- and six-coordinated titanium atoms (Ti_{5c} and Ti_{6c}, respectively). The O_{3c} and Ti_{6c} sites have the same coordination environment of the bulk phase, whereas the O_{2c} and Ti_{5c} are unsaturated and therefore more reactive.

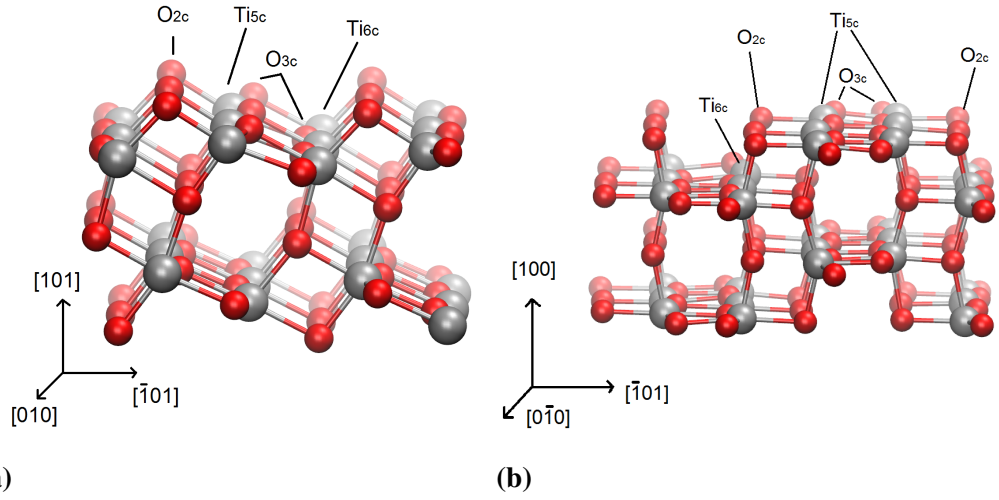


Figure 5.1 Anatase (a) (101) and (b) (100) surfaces. Titanium and oxygen atoms are labelled according to their coordination number. Colour key: Ti:grey, O:red

The reactivity of metal oxides is influenced by the presence of defects at the surface.[70] Monoatomic step-edges on anatase (101) are common defects which have a characteristic trapezoidal shape.[75] The parallel sides of the trapezoidal islands are oriented along the [010] direction, whereas the other two sides along the [$\bar{1}$11] and [11$\bar{1}$]. Among all the possible step-edges, we considered the long parallel side of the trapezoidal islands, *i.e.* step B from the study of Gong *et al.*[4] (see Figure 1.7). Since the side of the step can itself be considered as a tiny slice of a surface, and follows its reactivity, we have also simulated the CaP deposition on the (100) surface of anatase (Figure 5.1(b)), which is the extended surface corresponding to the side of step B.[4]

The (100) surface in Figure 5.1(b) exposes saturated Ti_{6c} and O_{3c} sites and under-coordinated Ti_{5c} and O_{2c} sites. Here, the Ti_{6c} are even less accessible than in the (101)

surface because they are positioned in the recess of the saw-tooth orientation of the surface. When ordering the relaxed anatase surfaces in terms of their relative surface energies, the (100) surface follows after the (101). Surface energies are usually correlated to their reactivity and here depend on the density of exposed uncoordinated Ti atoms; thus we expect the (100) surface to be less stable but more reactive than the (101).[77]

5.2.4 Preparation of the TiO_2 slabs

The optimisation of bulk anatase was performed employing both VASP and the IP-based General Utility Lattice Program (GULP) code, version 4.4.[107–109] The starting coordinates were taken from reference.[16]

Compared to experiment, the DFT and IP methods yielded values with errors of no more than 3% for the cell volume and vectors, bulk modulus and Ti−O bond distances and angles (see Table 5.1 and Figure 5.2). We then used the DFT and IP optimised lattice parameters to construct the surface unit cells.

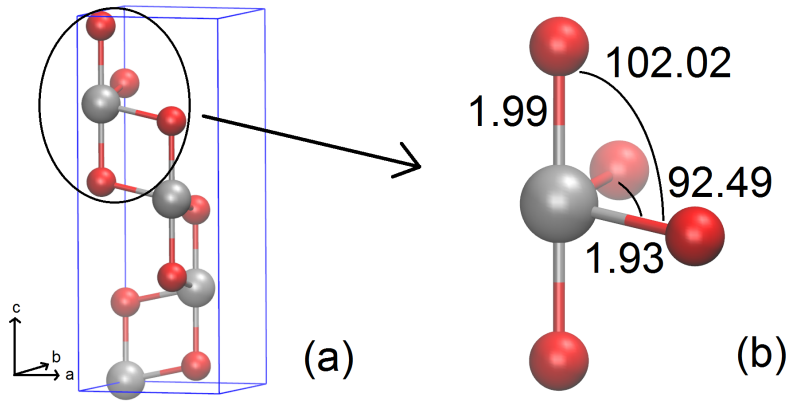


Figure 5.2 Geometry optimisation of the TiO_2 anatase bulk with the interatomic potentials of Matsui and Akaogi.[6](a) Anatase unit cell, (b) Ti−O distances in Angstrom and O−Ti−O angles. Colour key: Ti:grey, O:red

The anatase (101) and (100) surface slabs were created using the METADISE code, version 5.64,[110] which guarantees that no net dipole moment occurs normal to the surface.[139,145] For the DFT calculations, which were performed on the most stable (101) surface, the optimised (101) surface unit cell (Figure 5.3) was expanded 3 times along the [010] direction and 4 times along the [100] direction, resulting in a thickness of 12.9 Å (4 layers). We added a vacuum gap of 20 Å along [101] to avoid spurious interactions

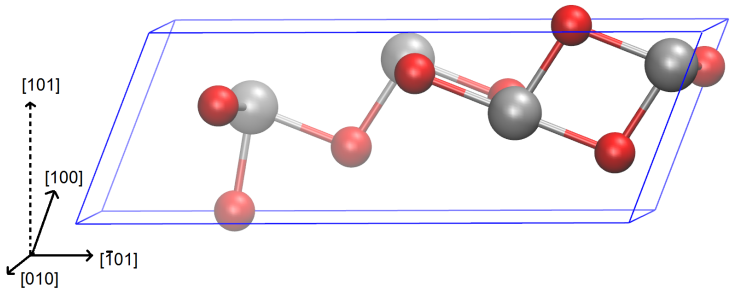


Figure 5.3 Anatase (101) unit cell. Colour key: Ti:grey, O:red

between the slab and its periodic images. The atoms in the two central layers of the slab were kept frozen at their bulk positions, whereas the top and the bottom layers were left free to rearrange. For the IP simulations, the GULP optimised structure was expanded 8 times along the [010] direction, 3 times along the [$\bar{1}$01] direction and 6 times along the [100] direction, resulting in a thickness of 17 Å (6 layers). The slab was centred in an orthorhombic box of approximately $30 \times 30 \times 77$ Å^3 and the space above and below the anatase slab was filled with a $30 \times 30 \times 30$ Å^3 cube of water.

Similarly, the (100) surface unit cell obtained with METADISE was expanded 8 times along the [0$\bar{1}$0] direction, 3 times along [$\bar{1}$01] and 6 times along [100], resulting in a thickness of 21 Å (6 layers). The slab was centred in an orthorhombic box with approximate dimension of $30 \times 29 \times 78$ Å^3.

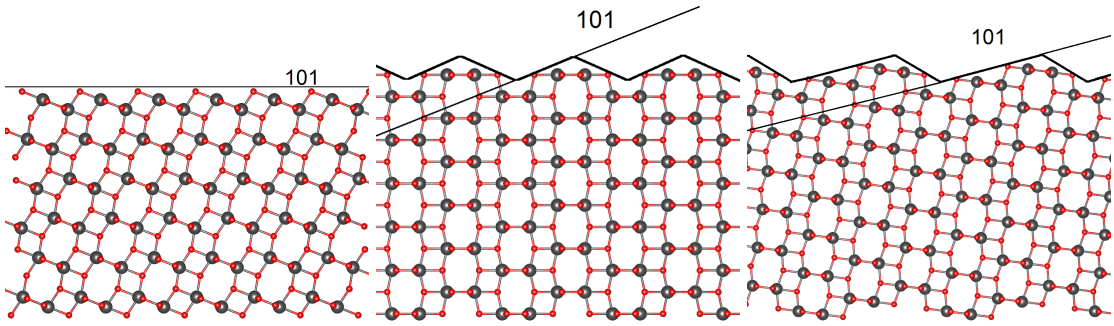


Figure 5.4 (a) (101), (b) (100) and (b) $(100)_{rot}$ surfaces and their relationship with the (101). Colour key: Ti:grey, O:red

A third slab was obtained rotating 14° clockwise the [$\bar{1}$01] axis of the (100) so that the new surface, labelled as $(100)_{rot}$, has mixed features between the (101) and the (100). Figure 5.4 shows all the slabs used for IP simulations and their relationship with the (101)

surface.

5.3 Results

5.3.1 Anatase (101) surface

5.3.1.1 Water adsorption

Our first DFT calculations focused on the molecular adsorption of water, which previously has been shown to be favoured over dissociative adsorption on the (101) surface of anatase.[71,217] In order to evaluate the sites for water adsorption on the surface, we placed a water molecule at the top and at the bottom of the slab on equivalent positions (see Figure 5.5 for the starting adsorption sites). After optimisation, the water molecule preferentially adsorbed on three different sites (see Figures 5.6 and 5.8(a-c)).

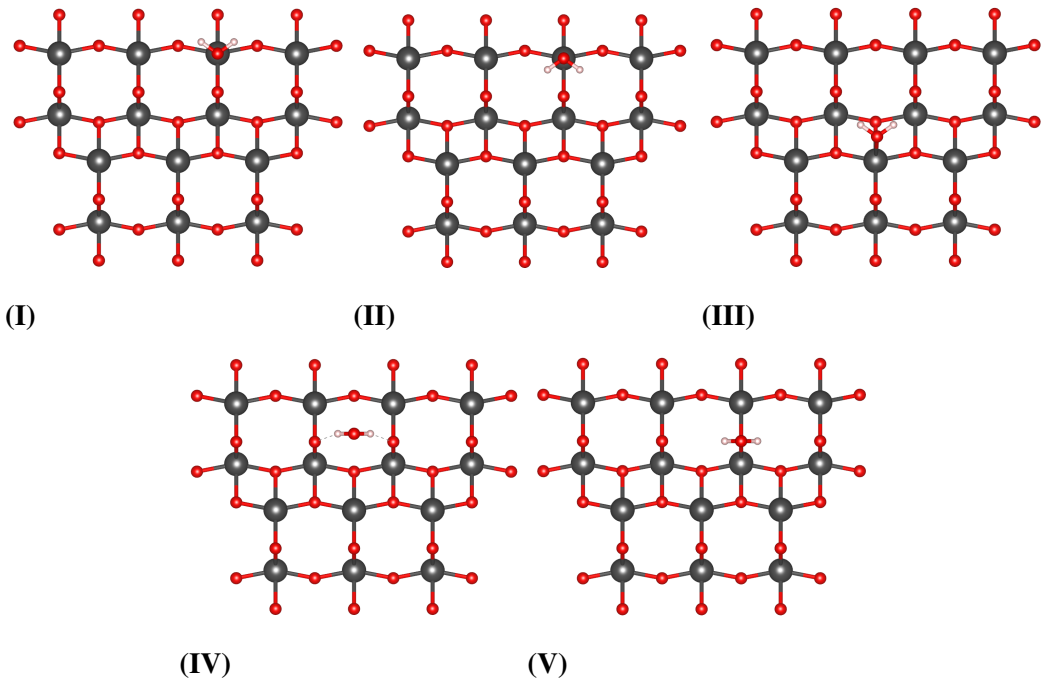


Figure 5.5 Adsorption of a single water molecule on a TiO_2 anatase [101] surface, different starting configurations. Colour key: Ti:grey, O:red, H:white.

In the most stable position, which matches that reported by previous studies in the literature,[76,218,219] the water molecule simultaneously interacts with a Ti_{5c} centre via its O atom (hereafter referred as O_w) and with nearby O_{2c} sites through the formation of two hydrogen-bonds, as shown in Figure 5.8(a). The second most stable adsorption retains

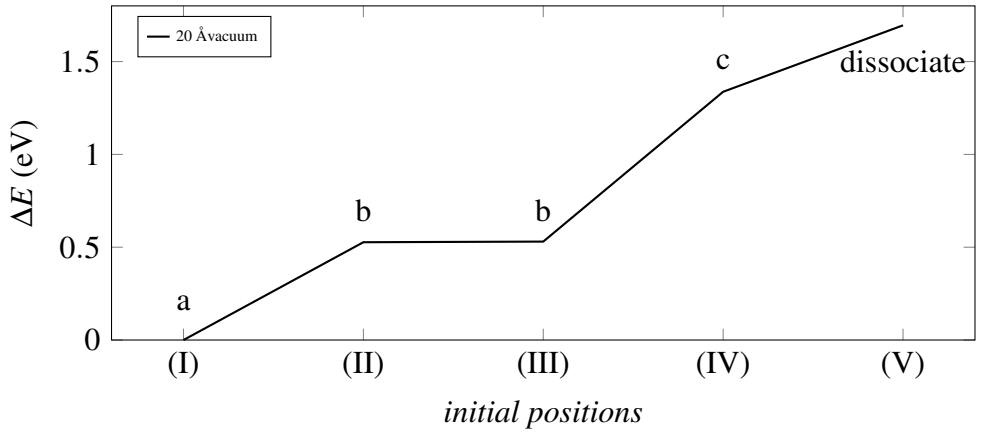


Figure 5.6 Relative energies after adsorption of two water molecules, one at the top and one at the bottom of the slab, in equivalent positions and a vacuum of 20 Å. The roman notation refers to the initial adsorption sites in Figure 5.5

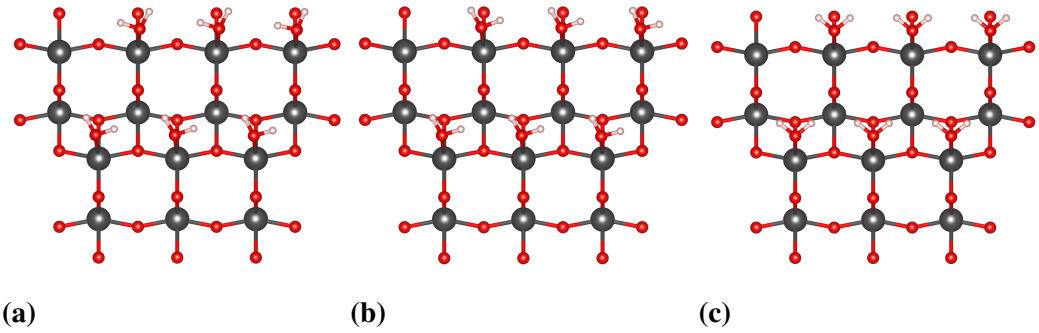
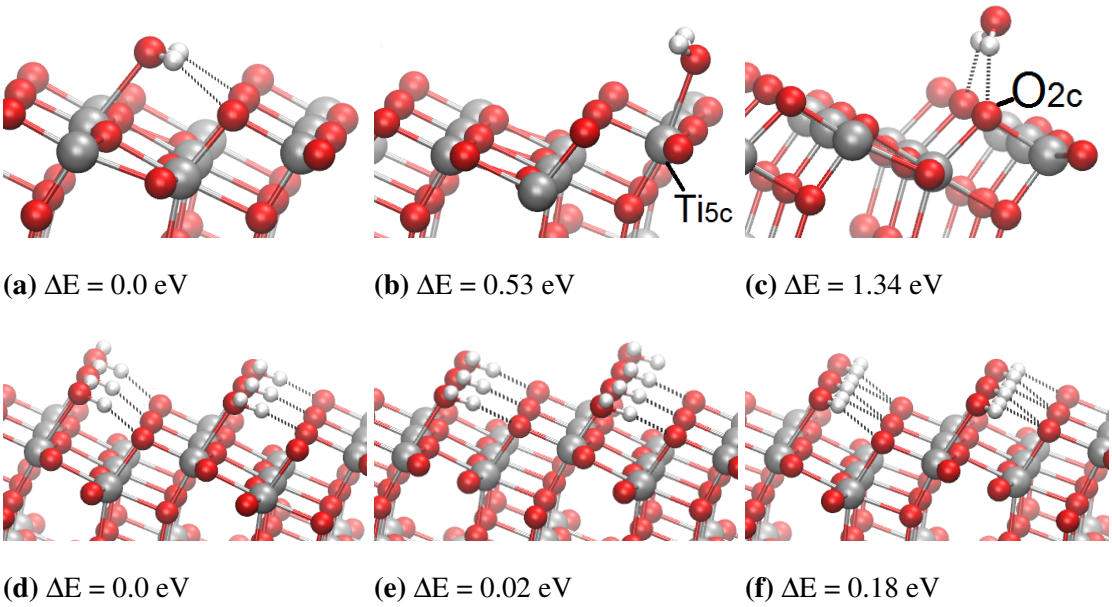


Figure 5.7 Adsorption of water on a TiO_2 anatase [101] surface, optimised geometries for full coverage. Colour key: Ti:grey, O:red, H:white

the $Ti_{5c}-O_w$ link but cleaves the hydrogen-bonds, leading to an energy penalty of 0.53 eV (Figure 5.8(b)). Finally, in the third position (Figure 5.8(c)), water binds the surface only via two hydrogen-bonds with the exposed O_{2c} sites, increasing the system energy by 1.34 eV compared to the most stable configuration. All the relative adsorption energies are plotted in Figure 5.6.

We next studied the change in the geometry of adsorbed water when the coverage is increased to a full monolayer. We tested several orientations for the water molecules using the structural information from the single molecule adsorption (see Figure 5.7), obtaining the full-coverage configurations reported in Figure 5.8(d), 5.8(e) and 5.8(f). In the lowest energy geometry, all water molecules adsorbed on Ti_{5c} sites are slightly rotated and form a single short hydrogen bond with the O_{2c} atoms, while the water molecules'



(a) ΔE = 0.0 eV **(b)** ΔE = 0.53 eV **(c)** ΔE = 1.34 eV

(d) ΔE = 0.0 eV **(e)** ΔE = 0.02 eV **(f)** ΔE = 0.18 eV

Figure 5.8 (a-c) DFT optimisation of a single water adsorption on the anatase (101) surface. (d-f) DFT optimisation of a monolayer of water on anatase (101). The energies of the adsorption configurations per water molecule are reported relative to the most stable system at each coverage. Colour key: Ti:grey, O:red, H:white

dipole moments are oriented in a zig-zag fashion along the [$\bar{1}$01] direction (Figure 5.8(d)). The alternative configuration with all dipole moments oriented in the same direction is only 0.02 eV higher in energy (see Figure 5.8(e)). In the least stable configuration, the number of hydrogen bonds is maximised at the expense of their lengths. From the DFT distances reported in Table 5.3, it is possible to see that, in general, the adsorption of a monolayer of water on Ti_{5c} sites results in the elongation of the $Ti_{5c}-O_w$ bond and a shortening of the hydrogen bond $O_{2c}-H_w$. Similar to the single molecule adsorption, the

Table 5.3 Interatomic distances between water and the anatase (101) surface. The DFT columns use the same letter labelling of Figure 5.8. The IP column reports the values obtained from the first peak of the RDFs with the corresponding coordination numbers within parentheses.

	DFT						IP
Distances (Å)	(a)	(b)	(c)	(d)	(e)	(f)	
$Ti_{5c}-O_w$	2.28	2.26	/	2.31	2.32	2.31	2.15 (0.250)
$O_{2c}-H_w$	2.23	/	1.99	2.04-2.07	2.04	2.12	2.25 (0.213)

full coverage results agree qualitatively with those in the literature.[218–220]

The input for the MD simulations was prepared as described in section 5.2.4. The resulting RDFs are reported in Figures 5.10-5.11 and Table 5.4. We obtained average $Ti_{5c}-O_w$ and $O_{2c}-H_w$ distances of 2.15 and 2.25 Å, respectively, which is in agreement with the DFT results (Table 5.3).

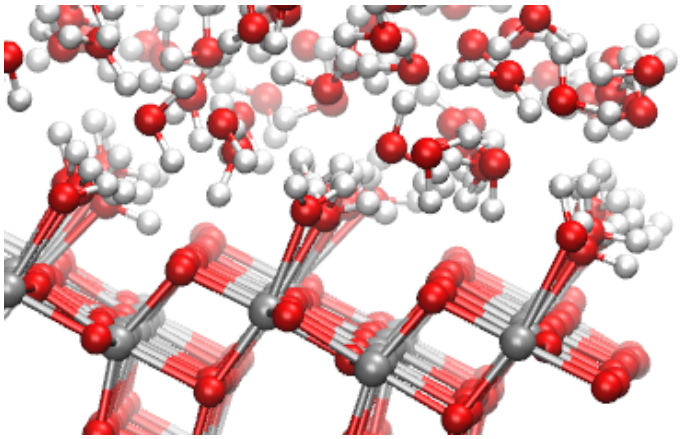


Figure 5.9 Classical MD snapshot after 1 ns in an NPT ensemble showing water adsorption on the anatase (101) surface. Colour key: Ti:grey, O:red, H:white

Figure 5.9 shows a snapshot after 1 ns of NPT simulation. The first monolayer of adsorbed water interacts with the anatase surface in configurations that resemble the DFT orientations in Figures 5.8(d) and 5.8(e). As expected, these water molecules do not only establish hydrogen-bonds with the O_{2c} sites, but also with other water molecules from the aqueous medium. In addition to the water that binds directly to the Ti_{5c} centres, other water molecules adsorb more loosely via hydrogen-bonds with the O_{2c} sites, forming semi-ordered rows along the [010] direction. A similar arrangement of water has been found previously by DFT methods for a 2 ML coverage.[218] These results highlight the good transferability of the force field to describe the water-anatase interface using water models different from the original SPC/E.

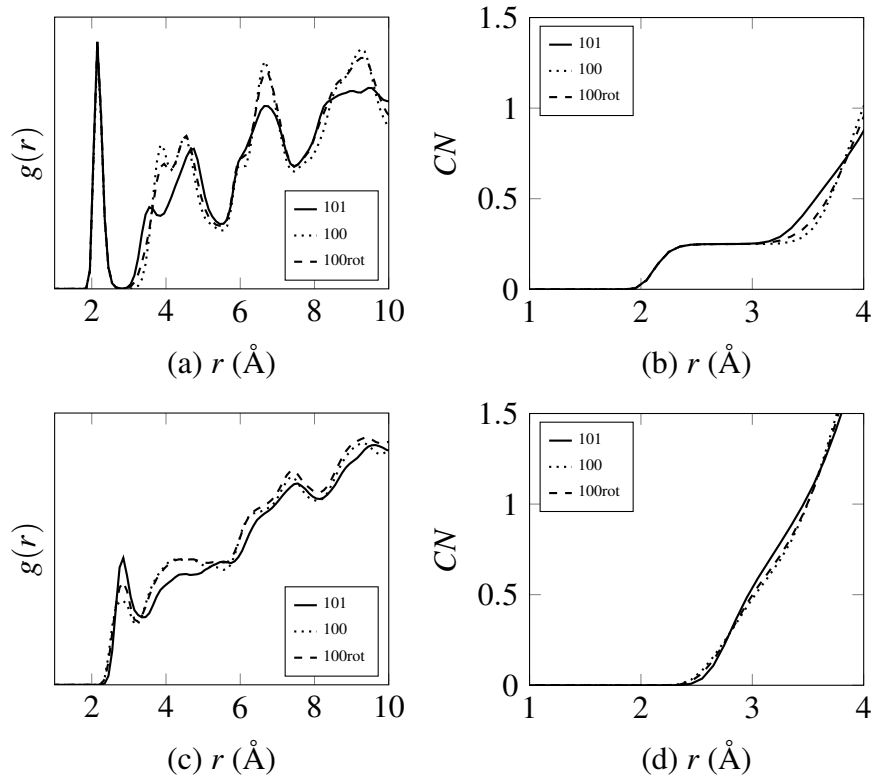


Figure 5.10 Radial distribution functions and coordination number of (a-b) Ti−O_w and (c-d) Ti−H_w

Table 5.4 Interatomic distances with corresponding coordination numbers within parentheses between water and the (100) and $(100)_{rot}$ surfaces of anatase. The DFT distances[4] for molecular adsorption of water on the anatase (100) surface are also reported.

	IP			DFT[4]
Distances (Å)	(101)	(100)	$(100)_{rot}$	
$Ti_{5c}-O_w$	2.15 (0.250)	2.15 (0.250)	2.15 (0.250)	2.282
$O_{2c}-H_w$	2.25 (0.213)	2.25 (0.220)	2.25 (0.215)	2.141

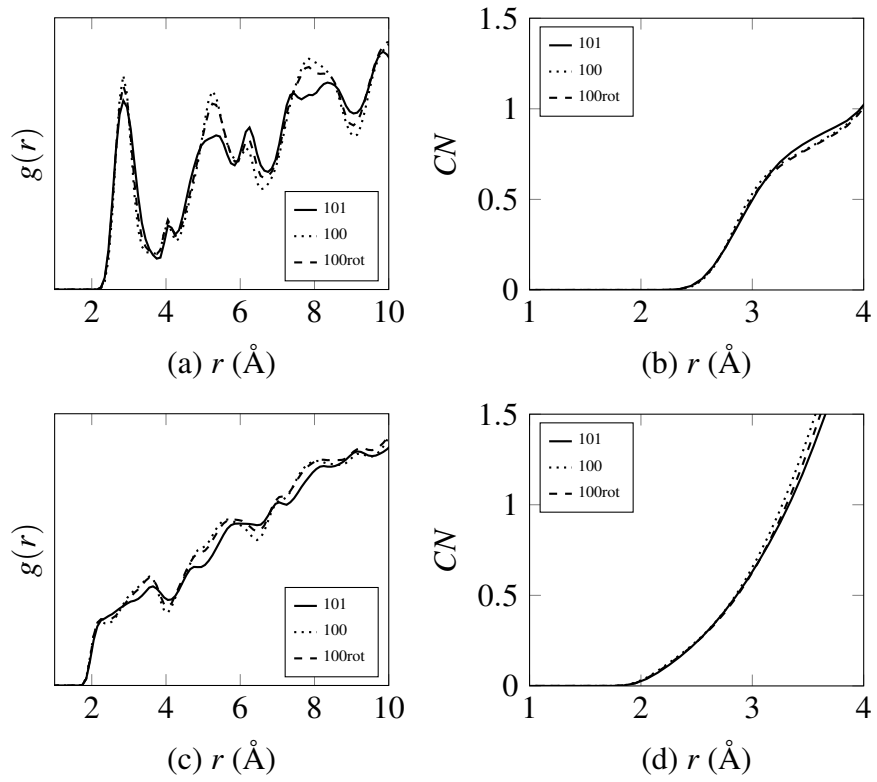


Figure 5.11 Radial distribution functions and coordination number of (a-b) $O_{Ti}-O_w$ and (c-d) $O_{Ti}-H_w$

5.3.1.2 Calcium phosphate adsorption

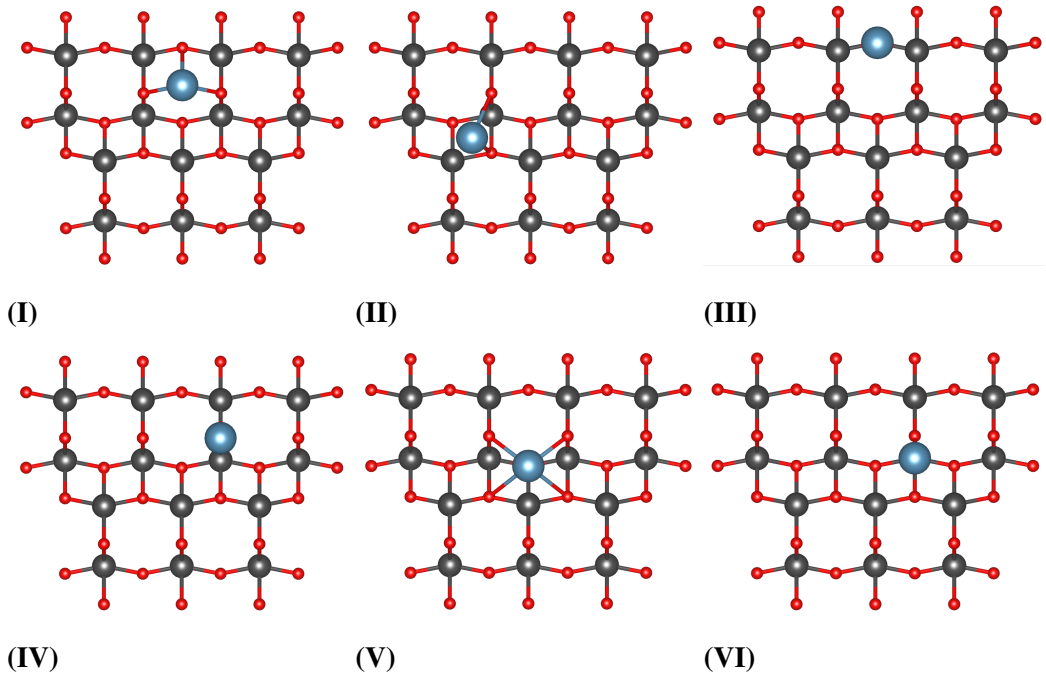


Figure 5.12 Adsorption of Ca^{2+} on a TiO_2 anatase [101] surface, different starting configurations. Colour key: Ti:grey, O:red, Ca:cyan

Table 5.5 Interatomic distances between calcium and the anatase (101) surface. The DFT columns use the same letter labelling of Figure 5.14.

	DFT		**IP**
Distances (Å)	(a)	(b)	
O_{2c}−Ca	2.21	2.12	2.25
O_{2c}−Ca	/	2.29	/
O_{3c}−Ca	2.54	/	3.35

We started by adsorbing two Ca^{2+} ions, one at the top and one at the bottom of the anatase slab at equivalent positions. The optimisations led to only two low-energy adsorption sites for the Ca^{2+} ion: one where it can coordinate two O_{2c} and one O_{3c} (Figure 5.14(a)) and another on top of a Ti_{6c} coordinating two O_{2c} (Figure 5.14(b)). The first position is 1.23 eV more stable than the second. We verified that the energetic stability of the different adsorption sites does not depend either on the vacuum size or the charge of the system by performing single point calculations increasing the vacuum from 20 to 30 Å and using

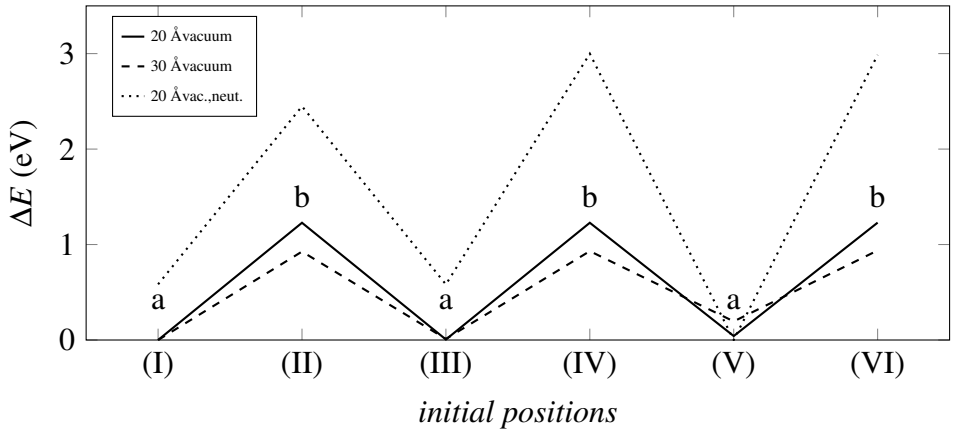


Figure 5.13 Relative energies after adsorption of (continuous line) two Ca^{2+}, one at the top and one at the bottom of the slab, in equivalent positions and a vacuum of 20 Å, (dashed line) two Ca^{2+}, one at the top and one at the bottom of the slab, in equivalent positions and a vacuum of 30 Å and (dotted line) two Ca^{2+}, one at the top and one at the bottom of the slab, in equivalent positions and removal of a Ti^{4+} ion in the frozen core of the slab to make the system neutral. Removing a Ti^{4+} from the slab eliminates the symmetry between the structures a(I) (and a(III)) and the structure a(V).

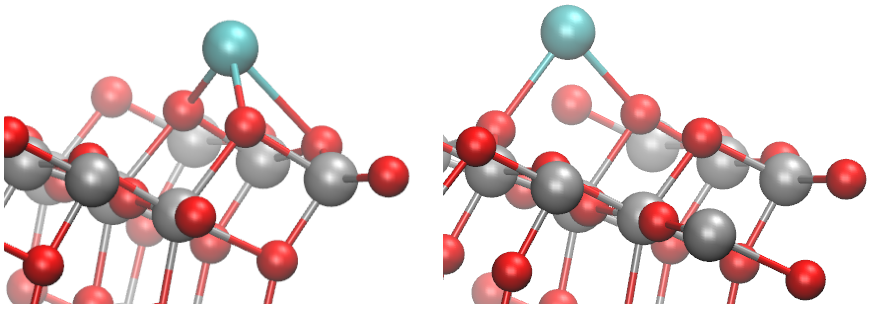
Table 5.6 Interatomic distances between $PO_4{}^{3-}$ and the anatase (101) surface. The DFT columns use the same letter labelling of Figure 5.18.

	DFT		**IP**
Distance (Å)	(a)	(b)	
$Ti_{5c}-O_{ph}$	1.88	1.94	1.85

a neutral slab (see Figures 5.12-5.13). We report in Table 5.5 the distances between the Ca^{2+} ion and the atoms of the surface.

During the MD simulations the Ca^{2+} ions are in a position very similar to that of the lowest energy structure obtained by DFT. In particular, they adsorb exactly on top of two consecutive O_{2c} and their partial solvation results in longer $O_{2c}-Ca$ bonds than in the anhydrous DFT optimisation, moreover Ca^{2+} ions are able to displace the water that is loosely interacting with the O_{2c} sites through hydrogen-bonds (see Figure 5.15 and Table 5.5).

For $PO_4{}^{3-}$, seven different starting configurations were considered (see Figure 5.16), which reduced to four after DFT geometry optimisation. The corresponding structures



(a) ΔE = 0.0 eV **(b)** ΔE = 1.23 eV

Figure 5.14 DFT optimisation of Ca^{2+} adsorption on the anatase (101) surface. The adsorption energies of the configurations are reported relative to the most stable position. Colour key: Ti:grey, O:red, Ca:cyan.

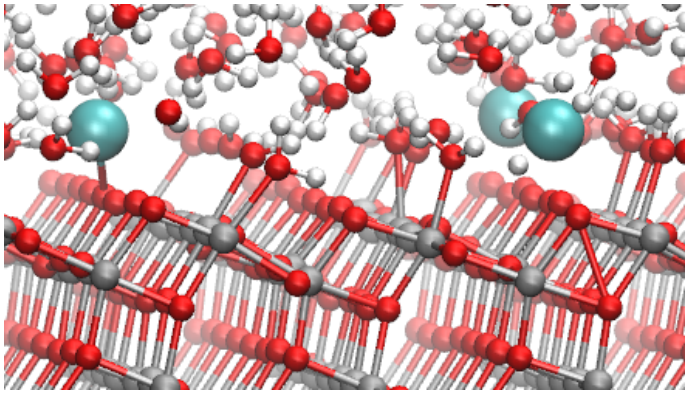


Figure 5.15 Classical MD simulation of Ca^{2+} adsorption on the anatase (101) surface. Colour key: Ti:grey, O:red, H:white, Ca:cyan.

and their relative energies are reported in Figures 5.17 and 5.18, which shows that, as in the case of the Ca^{2+} ion, the vacuum size and the charge of the system do not influence the trend in energy differences among the structures. The two lowest energy configurations and the relative $Ti_{5c}-O_{ph}$ distances are reported in Figure 5.18(a-b) and Table 5.6.

During the MD simulations, we found that the PO_4^{3-} ions adsorb on the anatase surface only in the presence of Ca^{2+} ions. This behaviour reflects the experimental CaP nucleation on TiO_2, where the initial adsorption of Ca^{2+} ions on the oxide is followed by the deposition of PO_4^{3-} ions onto the positively charged layer.[221,222] When a PO_4^{3-} ion adsorbs on the surface, its geometry agrees with that predicted by DFT (compare Figures 5.18(a) and 5.19(a)), *i.e.* two O_{ph} bridge two consecutive Ti_{5c} along the [010] direction. The $Ti_{5c}-O_{ph}$ distance obtained from the RDF is only 0.03 Å shorter than the one obtained

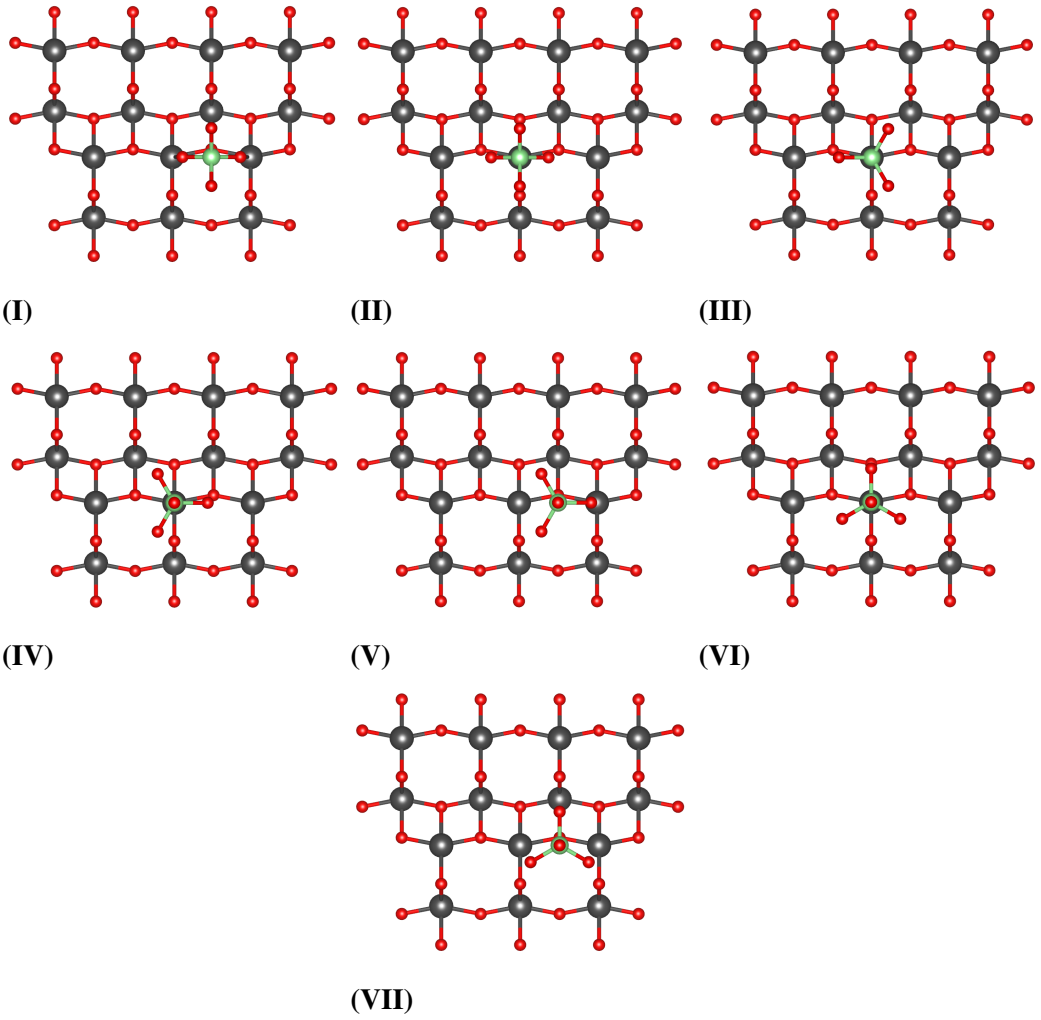


Figure 5.16 Adsorption of PO_4^{3-} on a TiO_2 anatase [101] surface, different starting configurations.

by DFT optimisation (see Table 5.6). Another possibility for the PO_4^{3-} ion is to coordinate a Ca^{2+} ion already adsorbed on the oxide surface (see Figure 5.19(b)). The average $Ca-O_{ph}$ distance for the structures in Figures 5.19(a) and 5.19(b) are respectively 2.25 and 2.35 Å. These distances are shorter than 2.45 Å, which is the value obtained using classical[5,207] and *ab initio* MD[223] to simulate CaP species in solution. The shortening is likely to be due to the geometric restraint imposed by the proximity of the surface active sites.

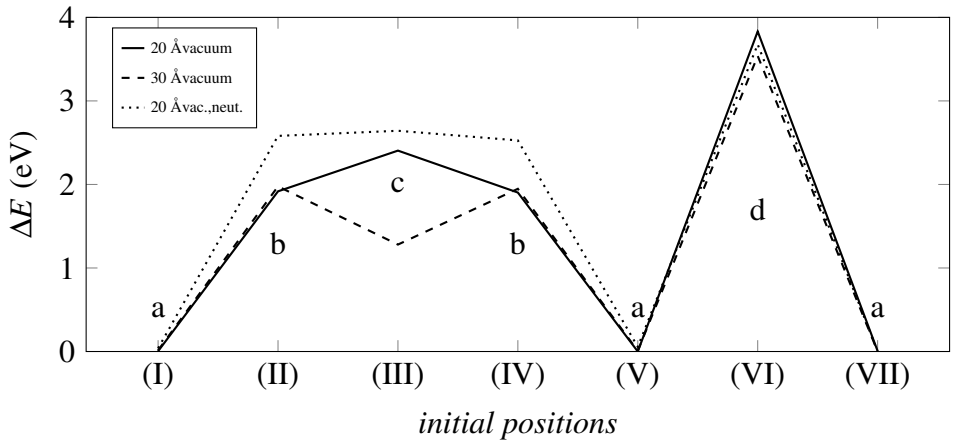
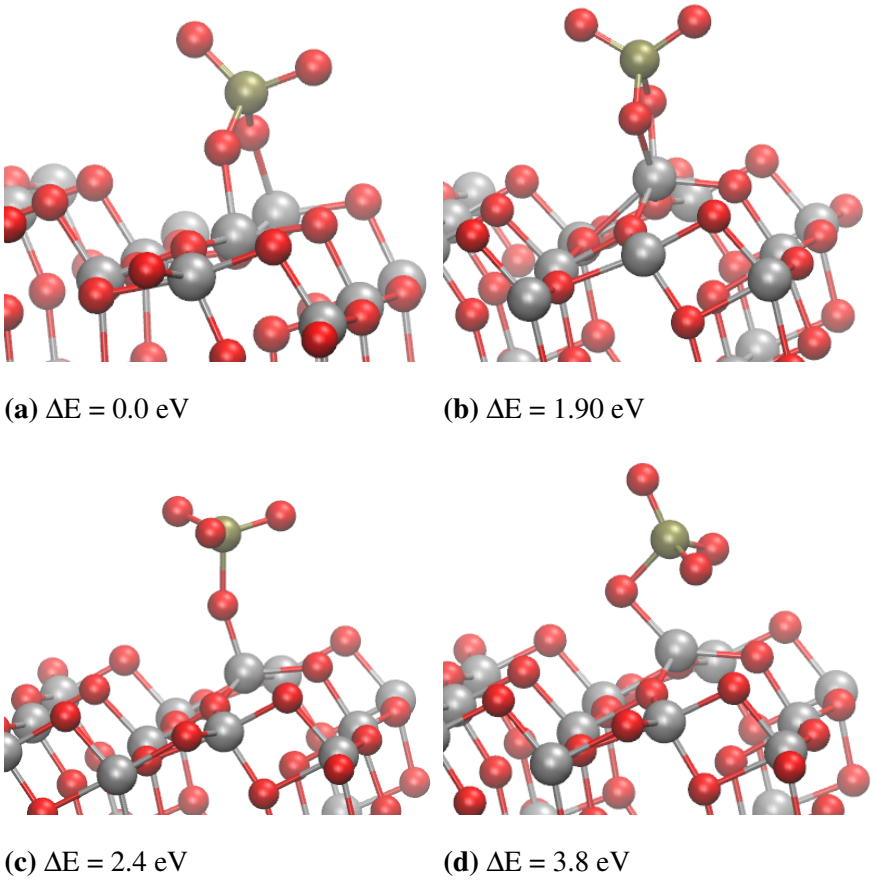


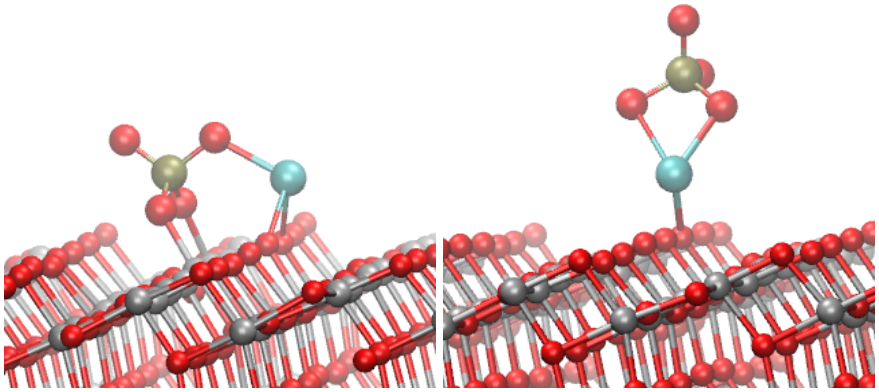
Figure 5.17 Relative energies after adsorption of (continuous line) two PO_4^{3-}, one at the top and one at the bottom of the slab, in equivalent positions and a vacuum of 20 Å (dashed line) two PO_4^{3-}, one at the top and one at the bottom of the slab, in equivalent positions and a vacuum of 30 Å and (dotted line) two PO_4^{3-}, one at the top and one at the bottom of the slab, in equivalent positions and removal of 3 oxygen ions in the frozen core of the slab to make the system neutral.



(a) ΔE = 0.0 eV **(b)** ΔE = 1.90 eV

(c) ΔE = 2.4 eV **(d)** ΔE = 3.8 eV

Figure 5.18 DFT optimised structures of PO_4^{3-} adsorption on the anatase (101) surface. The energy of the configurations are reported relative to the most stable position with a vacuum of 20 Å. Colour key: Ti:grey, O:red, P:green.



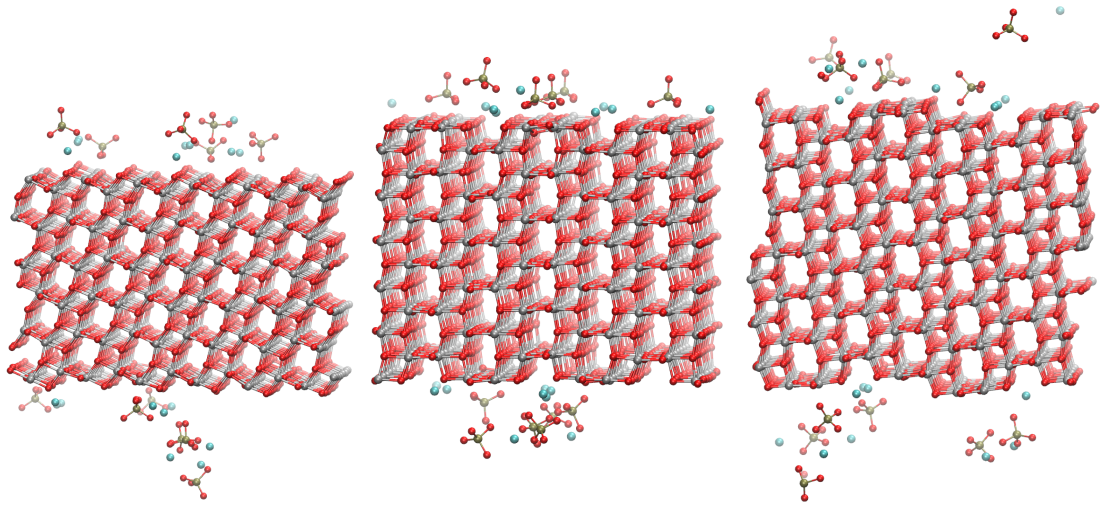
(a) **(b)**

Figure 5.19 Classical MD simulation snapshots of PO_4^{3-} adsorption on the anatase (101) surface. Water molecules are removed for clarity. Colour key: Ti:grey, O:red, P:green.

5.3.2 Calcium phosphate deposition on planar and stepped surfaces

Finally, we have employed the validated force field introduced in section 5.2.2 and reported in Table 5.2 to model the deposition of Ca^{2+} and $PO_4{}^{3-}$ ions in aqueous solution on the anatase (101), (100) and $(100)_{rot}$ surfaces. We placed 9 Ca^{2+} ions above and 9 Ca^{2+} ions below the slab, at $\sim$ 2 Å from the surfaces. The space between the slabs was filled with water and the necessary number of $PO_4{}^{3-}$ ions to neutralise the positive charge given by the Ca^{2+} ions.

Figure 5.20 shows snapshots of the three different surfaces taken after 1 ns. The RDFs and the average coordination numbers are plotted in Figures 5.22-5.23 and listed in Table 5.7. Our simulations show that aggregation and deposition take place concurrently on the (101) surface (Figure 5.20(a)), whereas the Ca^{2+} and $PO_4{}^{3-}$ ions prefer to bind to the (100) surface rather than aggregate in solution (Figure 5.20(b)). The anatase (100) facet is known to be more reactive than the (101) towards the adsorption of water, methanol and formic acid.[76] This tendency is confirmed by an increase in the coordination number for the $Ca-O_{Ti}$, Ca−Ti, $O_{ph}-O_{Ti}$ and O_{ph}−Ti pairings on the (100) surface with respect to the (101). The $Ca-O_{ph}$ pairing, which measures CaP aggregation, exhibits the opposite trend (see Table 5.7).



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

Figure 5.20 Classical MD snapshots of CaP deposition on (a) anatase (101), (b) (100) and (c) $(100)_{rot}$ surface. Water molecules are removed for clarity. Colour key: Ti:grey, O:red, Ca:cyan, P:green.

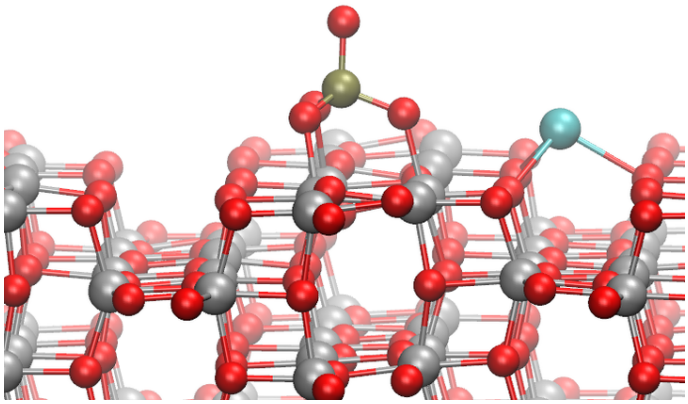


Figure 5.21 Classical MD snapshot of CaP deposition on anatase (100). The preferred adsorption sites for the Ca^{2+} and $PO_4{}^{3-}$ ions are visible. Colour key: Ti:grey, O:red, Ca:cyan, P:green.

Table 5.7 Interatomic distances for the adsorption of Ca^{2+} and $PO_4{}^{3-}$ ions on the anatase (101), (100) and $(100)_{rot}$ surfaces. The corresponding coordination numbers are reported in parentheses.

Distance (Å)	**(101)**	**(100)**	**$(100)_{rot}$**
$Ca-O_{2c}$	2.25 (1.1)	2.35 (2.5)	2.35 (1.6)
$Ca-Ti_{5c}$	3.45 (2.5)	3.75 (4.9)	3.75 (3.3)
$Ca-O_{ph}$	2.35 (2.8)	2.25 (1.8)	2.35 (2.1)
$O_{ph}-O_{2c}$	2.65 (0.6)	2.65 (1.0)	2.75 (0.5)
$O_{ph}-Ti_{5c}$	1.85 (0.2)	1.95 (0.4)	1.95 (0.2)

On the (100) surface, Ca^{2+} ions tend to coordinate to the exposed O_{2c} sites at the edge of the gulleys (see Figure 5.21). The PO_4 adsorb preferentially on the relatively flat regions of the (100) surface, which present two parallel rows of unsaturated positive Ti_{5c} ions that are capable of coordinating the three O_{ph} atoms simultaneously (Figure 5.21). As shown in Figure 5.4, the $(100)_{rot}$ surface has mixed features of both the (101) and (100) surfaces, and indeed it exhibits intermediate behaviour towards CaP deposition (Figure 5.20(c)). Moving from anatase (101) to the (100) surface, we observe that the coordination numbers $Ca-O_{2c}$ and $O_{ph}-Ti_{5c}$ increase by 130% and 72%, respectively, causing a stretching of 0.1 Å in the distances $Ca-O_{2c}$ and $O_{ph}-Ti_{5c}$ (see Table 5.7). Furthermore, during the deposition of CaP, the displacement of water was more pronounced on (100) than on (101), as reported in Table 5.8 and Figures 5.24, 5.25 and 5.26. In fact, upon adsorption

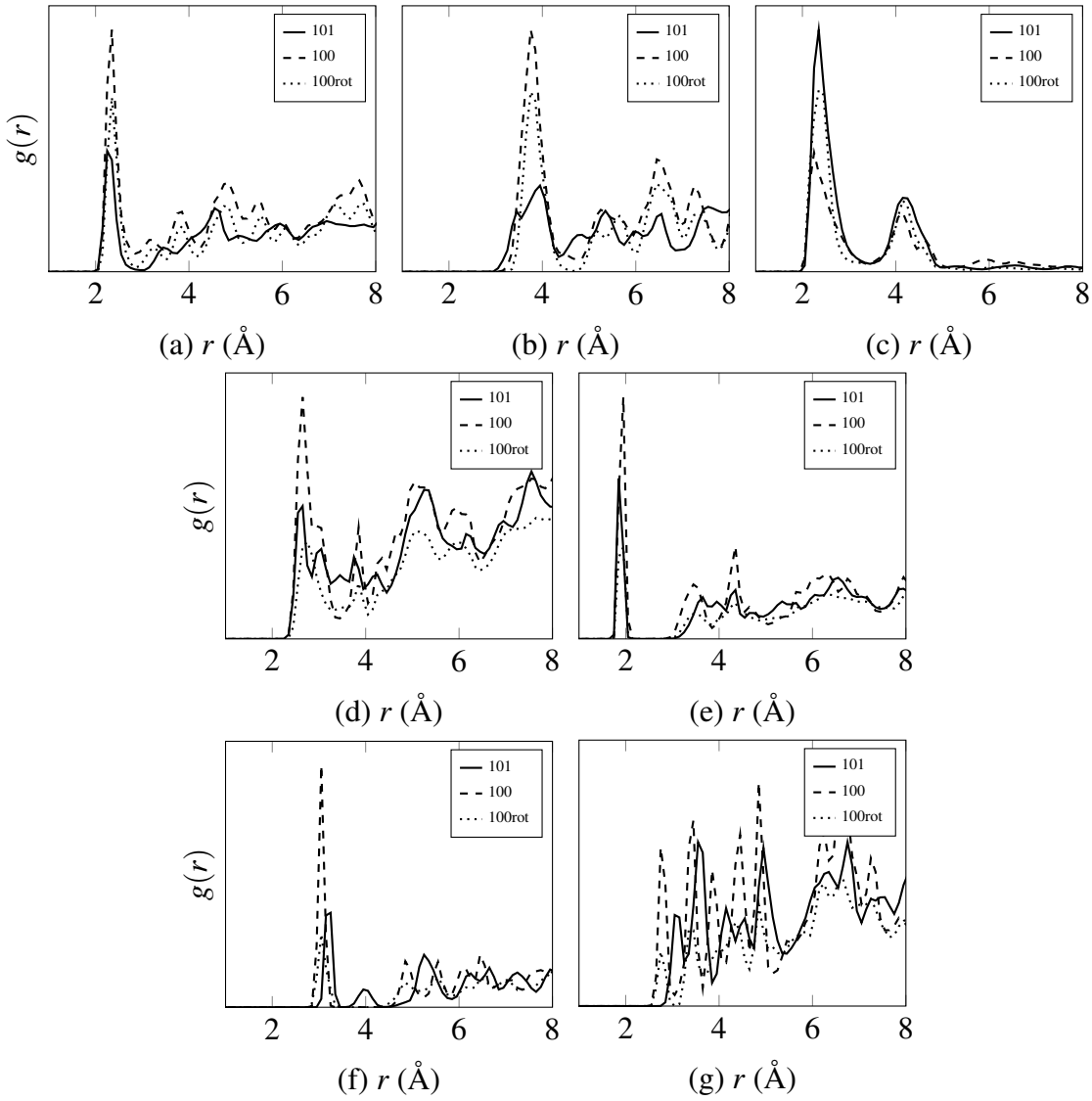


Figure 5.22 Radial distribution functions of (a) Ca−O_{Ti}, (b) Ca−Ti, (c) Ca−O_{ph}, (d) O_{ph}−O_{Ti}, (e) O_{ph}−Ti, (f) P−Ti and (g) P−O_{Ti}.

of the ions, the Ti_{5c}−O_w coordination number decreases by ≈ 12% on (101) and 20% on (100). These modifications indicate that anatase (100) is more reactive than (101) towards the adsorption of CaP. We explain the higher reactivity of (100) over (101) after considering the effect of the distribution of active sites on both surfaces, namely Ti_{5c} and O_{2c}, on the adsorption strength of water and CaP ions.

The (101) and (100) surfaces differ in their arrangements of the active sites, *i.e.* the species O_{2c} and Ti_{5c} which are undercoordinated with respect to the bulk. The (101) surface presents an alternation of positive and negative sites along the [$\bar{1}$01] direction ($\cdots O_{2c}$-Ti_{5c}-O_{2c}-$Ti_{5c}\cdots$) with the O_{2c} sites at the tip of the sawtooth-like corrugations (see Figure

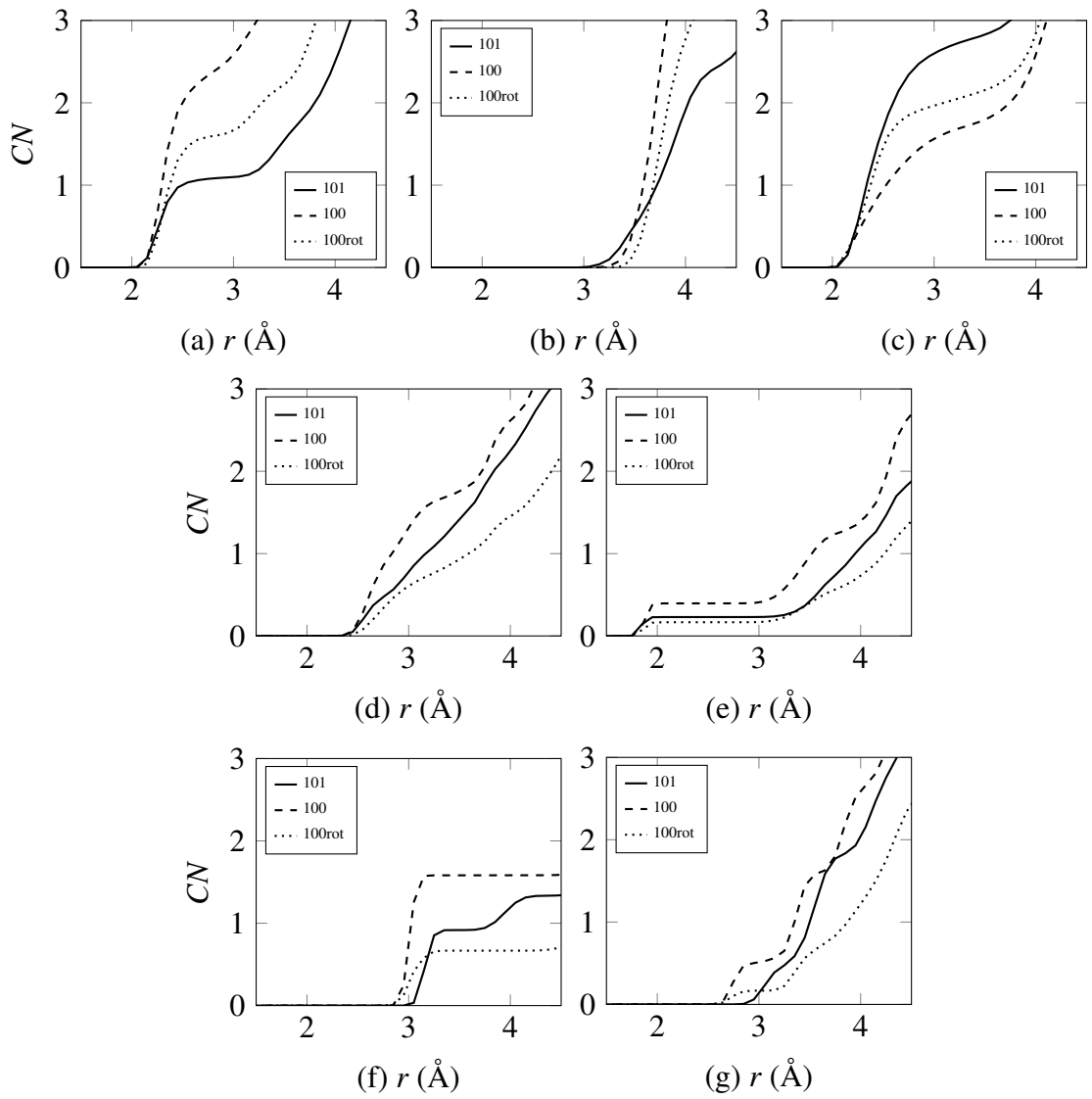


Figure 5.23 Coordination numbers for (a) Ca−O_{Ti}, (b) Ca−Ti, (c) Ca−O_{ph}, (d) O_{ph}−O_{Ti}, (e) O_{ph}−Ti, (f) P−Ti and (g) P−O_{Ti}.

5.1(a)). This arrangement allows water to interact simultaneously with the Ti_{5c} sites, via direct Ti_{5c}−O_w links, and with the O_{2c} sites through hydrogen-bonds, as shown in Figure 5.8(d) and Figure 5.9. In contrast, the (100) surface shows a distribution that places the Ti_{5c} and O_{2c} sites on the same plane (···gulley-O_{2c}-Ti_{5c}-Ti_{5c}-O_{2c}-gulley···), see Figure 5.1(b), thus reducing the possibility for the water molecules to interact simultaneously with both centres. Consequently, the water molecules have a stronger adsorption on (101) than on (100), and hence they are harder to displace by the incoming CaP ions. Note that, although the discussed O_w−H···O_{2c} hydrogen-bonds on (101) are partially disrupted by the presence of bulk water, they are still observable during the MD simulations. This

finding is in agreement with the DFT calculations by Gong and Selloni, who have shown that adsorption is energetically favoured on (101) over (100) by 0.18 eV (or 0.13 eV if the dissociative adsorption of water on (100) is considered).[76] Water molecules adsorb molecularly on (101), but can dissociate on (100). However, our MD simulations could only consider molecular adsorption on both surfaces, whereas the dissociative adsorption on (100) could cause a greater hindrance to the displacement of water by the deposition of CaP, because the water molecules have to assemble again before desorbing from the surface, which requires additional energy.

The arrangement of active sites on (101) and (100) has an opposite effect on the adsorption of CaP ions compared to water. On the (101) surface, the Ca^{2+} ion can only bind two O_{2c} sites simultaneously, placing itself on top of the sawtooth-like corrugations (see Figure 5.14(a)). In comparison, on the (100) surface, the Ca^{2+} ions can bind three O_{2c} sites at the gulleys; the short spacing between the O_{2c} centres, and their zig-zag distribution along the $[0\bar{1}0]$ direction allows this favourable adsorption. Furthermore, whereas on (101) Ca^{2+} needs to compete with the water that is interacting simultaneously with the Ti_{5c} and O_{2c} centres plus the more loosely adsorbed water that only interacts via hydrogen-bonds with O_{2c} sites, on (100) Ca^{2+} only needs to displace the weakly hydrogen-bounded water at the gulleys. Regarding the PO_4 ions, a similar reasoning can be applied. On (101), the PO_4 ion, which has four negative centres at each corner of the tetrahedron available for interaction with the Ti_{5c} sites, can adsorb on no more than two Ti_{5c} centres (Figure 5.18(a) and Figure 5.19(a)). On (100), the flat regions between gulleys provide a more favourable Ti_{5c} zig-zag arrangement, with Ti_{5c}-Ti_{5c} spacings shorter than 4 Å, where the PO_4 can adsorb via three O_{ph}-Ti_{5c} links (Figure 5.21). Therefore, the stronger CaP interaction with the (100) surface produces a higher amount of displaced water compared to (101). Again, it is important to note that these conclusions are derived from MD simulations, where only molecular adsorption was considered. Based on the results above, we expect a corresponding increase in CaP deposition at the steps B compared to the clean substrate. When considering the slightly rotated $(100)_{rot}$ surface, which has features of both the (100) and the (101) surfaces, the reactivity towards Ca^{2+} deposition lies about halfway between the (100) and the (101) facets, suggesting that the plain (100) surface may actually exhibit higher reactivity than step B. Instead, we found that the reactivity of the TiO_2

surfaces towards the adsorption of phosphate does not exhibit the same trend: the RDFs in Figure 5.22 (d-g) and the corresponding coordination numbers in Figure 5.23 (d-g) show that the $PO_4{}^{3-}$– surface reactivity follows the sequence (100) > (101) > $(100)_{rot}$. We may suppose that the observed behaviour derives from the limited number of simulations and the small size of the system; more MD simulations with different concentrations of ions in solution are required to confirm the trends in TiO_2 surfaces reactivity.

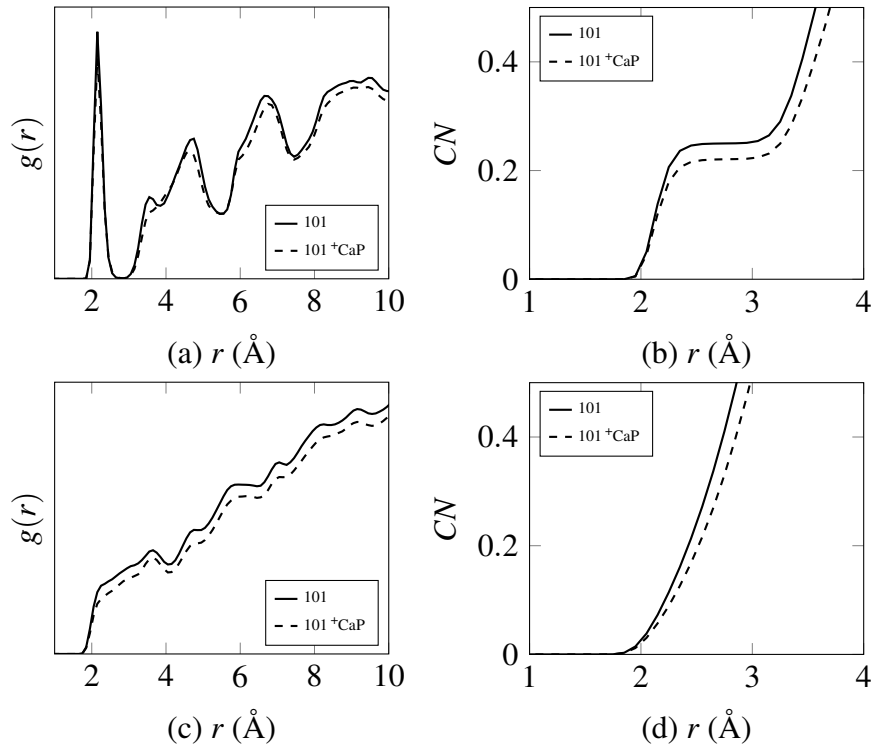


Figure 5.24 (101) surface: RDFs and coordination number of (a-b) $Ti-O_w$ and (c-d) $O_{Ti}-H_w$

Table 5.8 Difference in coordination number of $Ti_{5c}-O_w$ with and without calcium and phosphate ions.

	(101)	**(100)**	**$(100)_{rot}$**
CN $Ti_{5c}-O_w$	1.000	1.000	1.000
CN $Ti_{5c}-O_w$ with ions	0.884	0.804	0.916
Difference	-0.116	-0.196	-0.084

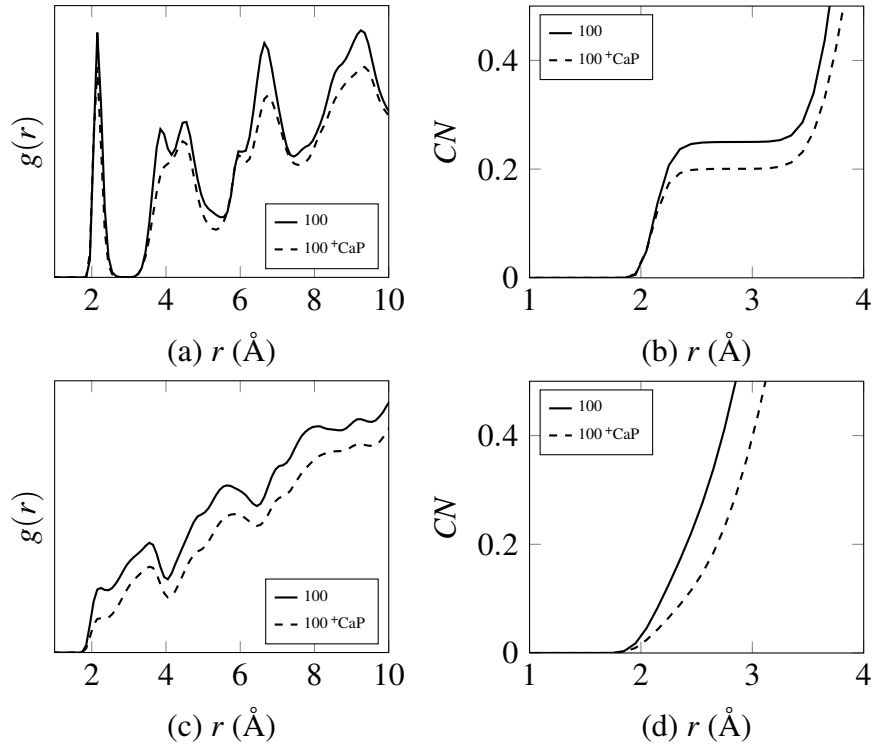


Figure 5.25 (100) surface: RDFs and coordination number of (a-b) $Ti-O_w$ and (c-d) $O_{Ti}-H_w$

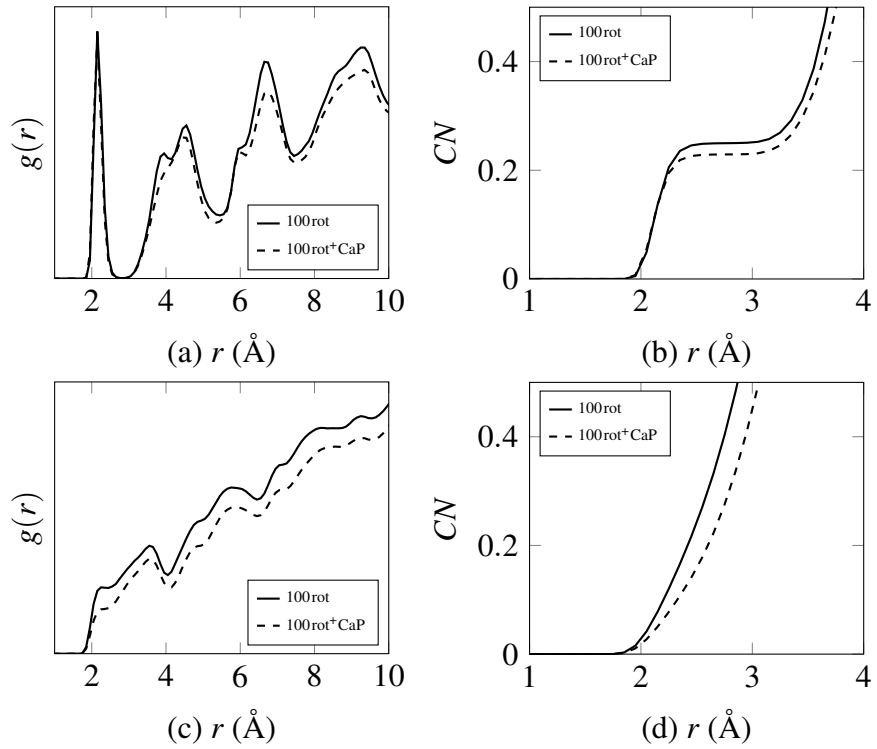


Figure 5.26 $(100)_{rot}$ surface: RDFs and coordination number of (a-b) $Ti-O_w$ and (c-d) $O_{Ti}-H_w$

5.4 Conclusions

In the present work, we have combined existing force fields for CaP, TiO_2 and water, and added missing two-body interactions between CaP and TiO_2, to perform classical MD simulations of CaP deposition on anatase in aqueous solution. The comparison with DFT shows very good performance of the new set of parameters in describing the adsorption of CaP on the most stable (101) surface of anatase. The (100) surface of anatase, which has the same features of one of the steps present in the trapezoidal islands observed experimentally on the (101) surface, was shown to be more reactive than the (101) towards CaP deposition. This finding was explained considering the different distribution of active sites on the two surfaces, which affects both water adsorption as well as calcium and phosphate deposition in opposing ways.

Chapter 6

Conclusions and future work

The nucleation of calcium phosphate in aqueous solution has been investigated using computational tools. Several studies reported in the scientific literature stress the role of *prenucleation clusters* during the early stages of calcium phosphate formation. Two structures for CaP PNCs are proposed: the calcium triphosphate complex $[Ca(HPO_4)_3]^{4-}$ and the Posner's cluster $Ca_9(PO_4)_6$.

In the present work, both PNC structures have been investigated using different levels of theory, depending on the size of the system. The structures and stabilities of $[Ca(HPO_4)_3]^{4-}$ prenucleation complexes in water have been studied by means of *ab initio* molecular dynamics simulations and umbrella sampling techniques. The simulations have shown that, in the most stable configuration, one phosphate group bound to the calcium is monodentate, while the other two are bidentate and placed at the opposite sides of the calcium. In addition, two water molecules also bind the calcium ion, which prefers to stay in a $Ca(\eta^2-{HPO_4}^{2-})_2(\eta^1-{HPO_4}^{2-})(H_2O)_2$ arrangement. The calcium triphosphate PNC is more stable than the isolated ions, as evidenced by the free energy profiles simulating the dissociation process. Moreover, the data suggest that a Ca/P ratio of 1:3 is thermodynamically favoured over other ratios, supporting the experimental findings in the literature.

Classical Molecular Dynamics simulations of the aggregation of calcium and phosphate ions in water revealed that Posner-like clusters, deficient in calcium ions and composed of protonated phosphate groups, form spontaneously during the simulation after a few ns. Sodium ions in solution, which are normally present in body fluids, can replace calcium

in the outer layer of the Posner's cluster. The Posner-like entities do not have symmetry, although the alignment of two calcium ions with the central one and the presence of deprotonated phosphate groups confers a higher symmetry to the cluster. Two or more clusters can share part of their ions, since one or more of their phosphate groups bridge two central calciums, as already observed in experiment and during MD simulations of Ca^{2+}, PO_4^{3-} and OH^- clustering at a collagen template. The simulations here reported represent the first theoretical investigation of the structure of the Posner's cluster in solution and confirm the formation of these species in the early stages of calcium phosphate crystallisation from solution, supporting previous experimental findings. The possibility that calcium triphosphate complexes originate in solution and aggregate to form Posner-like clusters has not been directly investigated in this work, but the dynamic character of the $[Ca(HPO_4)_3]^{4-}$ complexes, which change phosphate denticity and calcium coordination in the timescale of the *ab initio* MD (60 ps) makes them very likely to interact with other ions in solution (*e.g.:* Na^+), forming larger assemblies. Although in chapter 4 we have chosen to detect clusters made of 6 phosphate groups and 7 to 9 cations, this does not exclude the formation of calcium triphosphate complexed during the classical MD simulations: in the future, it would be interesting to re-analyse the same trajectories to reveal $[Ca(HPO_4)_3]^{4-}$. Besides, the starting structure for the PNC studied in chapter 3 is the one given by Habraken *et al.*, who proposed that these units form polymeric assemblies before taking up calcium ions from solution in a post-nucleation stage; however, the authors do not clarify how these negatively charged complexes can aggregate. In chapter 4, we have seen that Na^+ has a pivotal role during the early stages of CaP aggregation, especially in neutral solutions, where it is able to compensate for the negative charge of phosphate ions, allowing the formation of large assemblies. It can be assumed that calcium triphosphate complexes aggregate into larger clusters which contain also other cations normally present in solution, such as sodium, leading to the formation of Posner-like clusters. In a second stage, the protons of hydrogenphosphate ions and most of the cations other than Ca^{2+} should be released increasing the Ca/P ratio of the aggregate towards the value of 1.5, typical of the first solid phase emerging from solution (ACP).

Titanium is commonly employed for replacing hard tissues in the human body, where it is passivated and covered by a CaP hydroxyapatite layer to enhance its biocompatibility.

Therefore the formation of the CaP layer on titanium oxide is the object of the last chapter of this thesis. Unfortunately, a CaP-TiO_2 force field for MD simulations does not appear in the literature. To overcome this limitation, existing force fields for CaP, TiO_2 and water were combined and the missing two-body interactions between CaP and TiO_2 were calculated. The comparison with DFT showed very good performance of the new set of parameters in describing the adsorption of CaP on the most stable (101) surface of anatase. The (100) surface of anatase, which has the same features of one of the steps present in the trapezoidal islands observed experimentally on the (101) surface, resulted to be more reactive than the (101) towards deposition.

Future development may focus on performing more *ab initio* MD simulations of a calcium triphosphate complex in water to validate the structure of the PNC in solution presented in chapter 3 and to give an estimate of the errors on calcium coordination number, denticity of the ligands and number of water molecules in the first solvation sphere. The inclusion of cations (K^+, Mg^{2+}) and anions (Cl^-, $HCO_3{}^-$, $SO_4{}^{2-}$) normally present in body fluids in the calcium phosphate force field used in chapter 4 will make possible a more realistic representation of the nucleation process *in vivo*. More computationally demanding classical MD simulation of larger systems can be performed to observe further steps of calcium phosphate aggregation, up to the ACP nucleation; alternatively, enhanced sampling techniques such as umbrella sampling and metadynamics, may provide interesting insights into the ACP formation at a reasonable computational effort. Additional benchmarking against DFT calculations is required in order to use the CaP-water-TiO_2 combined force field to study the CaP deposition on common titanium oxide defects, such as oxygen vacancies and calcium impurities and on different TiO_2 polymorphs and their facets. It would also be interesting to expand this set of interatomic potential including an atom type for oxygen which describes the OH groups formed on the TiO_2 surface after water dissociation.

Bibliography

[1] W. Humphrey, A. Dalke and K. Schulten, *Journal of Molecular Graphics*, 1996, **14**, 33–38.

[2] L. Siperko and W. Landis, *Journal of Structural Biology*, 2001, **135**, 313–320.

[3] S. V. Dorozhkin, *J Mater Science*, 2007, **42**, 1061–1095.

[4] X.-Q. Gong, A. Selloni, M. Batzill and U. Diebold, *Nature Materials*, 2006, **5**, 665–670.

[5] R. I. Ainsworth, D. Di Tommaso, J. K. Christie and N. H. de Leeuw, *The Journal of chemical physics*, 2012, **137**, 234502.

[6] M. Matsui and M. Akaogi, *Molecular Simulation*, 1991, **6**, 239–244.

[7] N. H. de Leeuw and S. C. Parker, *Physical Review B*, 1998, **58**, 13901–13908.

[8] S. Kerisit and S. C. Parker, *J. Am. Chem. Soc.*, 2004, **126**, 10152–10161.

[9] J. M. Sorenson, G. Hura, R. M. Glaeser and T. Head-Gordon, *Journal of Chemical Physics*, 2000, **113**, 9149–9161.

[10] N. A. Hewish, J. E. Enderby and W. S. Howells, *J. Phys. C: Solid State Phys.*, 1983, **16**, 1777–1791.

[11] H. J. Kulik, N. Marzari, A. A. Correa, D. Prendergast, E. Schwegler and G. Galli, *The journal of physical chemistry. B*, 2010, **114**, 9594–9601.

[12] F. Betts and A. Posner, *Materials Research Bulletin*, 1974, **9**, 353–360.

[13] L. Veselinović, L. Karanović, Z. Stojanović, I. Bračko, S. Marković, N. Ignjatović and D. Uskoković, *Journal of Applied Crystallography*, 2010, **43**, 320–327.

[14] X. Yin and M. J. Stott, *Journal of Chemical Physics*, 2003, **118**, 3717–3723.

[15] V. Swamy and L. S. Dubrovinsky, *Journal of Physics and Chemistry of Solids*, 2001, **62**, 673–675.

[16] M. Rezaee, S. M. Mousavi Khoie and K. H. Liu, *CrystEngComm*, 2011, **13**, 5055–5061.

[17] M. Alimohammadi and K. A. Fichthorn, *The Journal of Physical Chemistry C*, 2011, **115**, 24206–24214.

[18] C. L. Freeman, J. H. Harding, D. J. Cooke, J. A. Elliott, J. S. Lardge and D. M. Duffy, *Journal of Physical Chemistry C*, 2007, **111**, 11943–11951.

[19] M. Hawkesford, W. Horst, T. Kichey, H. Lambers, J. Schjoerring, I. S. Moller and P. White, in *Marschner's Mineral Nutrition of Higher Plants: Third Edition*, Elsevier Ltd, 2012, pp. 135–189.

[20] H. Lowenstam and S. Weiner, *On biomineralization*, Oxford University Press, New York, 1989, pp. 3–6.

[21] S. V. Dorozhkin and M. Epple, *Angewandte Chemie - International Edition*, 2002, **41**, 3130–3146.

[22] A. Dey, P. H. H. Bomans, F. A. Müller, J. Will, P. M. Frederik, G. de With and N. A. J. M. Sommerdijk, *Nature materials*, 2010, **9**, 1010–1014.

[23] W. J. E. M. Habraken, J. Tao, L. J. Brylka, H. Friedrich, L. Bertinetti, A. S. Schenk, A. Verch, V. Dmitrovic, P. H. H. Bomans, P. M. Frederik, J. Laven, P. van der Schoot, B. Aichmayer, G. de With, J. J. DeYoreo and N. A. J. M. Sommerdijk, *Nat. Commun.*, 2013, **4**, 1507.

[24] A. Posner and F. Betts, *Accounts of Chemical Research*, 1975, **8**, 273–281.

[25] C. Y. Pak, *The Lancet*, 1998, **351**, 1797–1801.

[26] P. Koutsoukos, Z. Amjad, M. B. Tomson and G. H. Nancollas, *Journal of the American Chemical Society*, 1980, **102**, 1553–1557.

[27] P. Becker, *Fertilizer Science and Technology Series*, Marcel Dekker, New York, 1989, pp. 6–20.

[28] A. Hesse and D. Heimbach, *World journal of urology*, 1999, **17**, 308–315.

[29] J. J. Crall and J. M. Bjerga, *Journal of oral pathology*, 1987, **16**, 488–491.

[30] S. Takagi, L. C. Chow and K. Ishikawa, *Biomaterials*, 1998, **19**, 1593–1599.

[31] L. C. Chow and E. D. Eanes, *Octacalcium Phosphate*, Karger, Basel, 2001.

[32] A. A. Mirtchi, J. Lemaitre and N. Terao, *Biomaterials*, 1989, **10**, 475–480.

[33] R. LeGeros, *Calcium Phosphates in Oral Biology and Medicine*, Karger, Basel, 1991.

[34] A. S. Posner, A. Perloff and A. F. Diorio, *Acta Crystallographica*, 1958, **11**, 308–309.

[35] W. Suchanek and M. Yoshimura, *Journal of Materials Research*, 1998, **13**, 94–117.

[36] S. Mann, *Biomineralization: Principles and Concepts in Bioinorganic Materials Chemistry*, Oxford University Press, Oxford, 2001.

[37] L. Fisher, D. Torchia, B. Fohr, M. Young and N. Fedarko, *Biochemical and Biophysical Research Communications*, 2001, **280**, 460–465.

[38] G. K. Hunter, P. V. Hauschka, A. R. Poole, L. C. Rosenberg and H. A. Goldberg, *The Biochemical journal*, 1996, **317**, 59–64.

[39] M. Glimcher, A. Hodge and F. Schmitt, *Proc Natl Acad Sci*, 1957, **43**, 860–967.

[40] M. Vallet-Regí and J. M. González-Calbet, *Progress in Solid State Chemistry*, 2004, **32**, 1–31.

[41] K. Young, J. Wise, P. DeSaix, D. H. Kruse, B. Poe, E. Johnson, J. Johnson, O. Korol, G. Betts and M. Womble, *Anatomy \& Physiology*, 2013.

[42] S. Weiner and H. D. Wagner, *Annual Review of Materials Science*, 1998, **28**, 271–298.

[43] A. Sheu and J. Center, *MedicineToday*, 2017, **18**, 29–38.

[44] K. Kawasaki, A. V. Buchanan and K. M. Weiss, *Annual Review of Genetics*, 2009, **43**, 119–142.

[45] M. J. Duer, T. Friščić, D. Proudfoot, D. G. Reid, M. Schoppet, C. M. Shanahan, J. N. Skepper and E. R. Wise, *Arteriosclerosis, Thrombosis, and Vascular Biology*, 2008, **28**, 2030–2034.

[46] M. P. Ginebra, T. Traykova and J. A. Planell, *Journal of Controlled Release*, 2006, **113**, 102–110.

[47] B. R. Constantz, I. C. Ison, M. T. Fulmer, R. D. Poser, S. T. Smith, M. VanWagoner, J. Ross, S. A. Goldstein, J. B. Jupiter and I. Rosenthal, *Science*, 1995, **267**, 1796–1799.

[48] M. Bohner, J. Lemaître, P. Van Landuyt, P. Y. Zambelli, H. P. Merkle and B. Gander, *Journal of Pharmaceutical Sciences*, 1997, **86**, 565–572.

[49] G. Daculsi, *Biomaterials*, 1998, **19**, 1473–1478.

[50] G. Grimandi, P. Weiss, F. Millot and G. Daculsi, *Journal of Biomedical Materials Research*, 1998, **39**, 660–666.

[51] P. Korkusuz, U. O, F. Korkusuz and V. Hasirci, *Biomaterials*, 1999, **20**, 711–719.

[52] R. Zhang and P. X. Ma, *Macromolecular Bioscience*, 2004, **4**, 100–111.

[53] S. E. Ruiz Hernandez, R. I. Ainsworth and N. H. de Leeuw, *Journal of Non-Crystalline Solids*, 2016, **451**, 131–137.

[54] L. L. Hench, *Journal of Materials Science: Materials in Medicine*, 2006, **17**, 967–978.

[55] R. Z. LeGeros, *Chemical Reviews*, 2008, **108**, 4742–4753.

[56] S. J. Kalita, A. Bhardwaj and H. A. Bhatt, *Materials Science and Engineering C*, 2007, **27**, 441–449.

[57] R. G. T. Geesink, *Clinical orthopaedics and related research*, 2002, **395**, 53–65.

[58] M. Semlitsch, *Clinical Materials*, 1987, **2**, 1–13.

[59] H. Ji, C. B. Ponton and P. M. Marquis, *Journal of Materials Science: Materials in Medicine*, 1992, **3**, 283–287.

[60] L. Sun, C. C. Berndt, K. A. Gross and A. Kucuk, *Journal of Biomedical Materials Research*, 2001, **58**, 570–592.

[61] K. de Groot, J. G. Wolke and J. A. Jansen, *Journal of engineering in medicine*, 1998, **212**, 137–147.

[62] P. Ducheyne, W. Van Raemdonck, J. C. Heughebaert and M. Heughebaert, *Biomaterials*, 1986, **7**, 97–103.

[63] H. W. Kim, Y. H. Koh, L. H. Li, S. Lee and H. E. Kim, *Biomaterials*, 2004, **25**, 2533–2538.

[64] M. Shirkhanzadeh, *Materials Letters*, 1992, **14**, 27–30.

[65] J. L. Ong, L. C. Lucas, W. R. Lacefield and E. D. Rigney, *Biomaterials*, 1992, **13**, 249–254.

[66] K. Hata, T. Kokubo, T. Nakamura and T. Yamamuro, *Journal of the American Ceramic Society*, 1995, **78**, 1049–1053.

[67] B. Stadlinger, S. J. Ferguson, U. Eckelt, R. Mai, A. T. Lode, R. Loukota and F. Schlottig, *British Journal of Oral and Maxillofacial Surgery*, 2012, **50**, 74–79.

[68] A. Chandra, J. J. Ryu, P. Karra, P. Shrotriya, V. Tvergaard, M. Gaisser and T. Weik, *Journal of the Mechanical Behavior of Biomedical Materials*, 2011, **4**, 1990–2001.

[69] F. A. Grant, *Reviews of Modern Physics*, 1959, **31**, 646–674.

[70] U. Diebold, *Surface Science Reports*, 2003, **48**, 53–229.

[71] A. Vittadini, A. Selloni, F. P. Rotzinger and M. Grätzel, *Physical Review Letters*, 1998, **81**, 2954–2957.

[72] H. Z. Zhang and J. F. Banfield, *Journal of Materials Chemistry*, 1998, **8**, 2073–2076.

[73] B. Yang, M. Uchida, H. M. Kim, X. Zhang and T. Kokubo, *Biomaterials*, 2004, **25**, 1003–1010.

[74] X. Zhu, J. Chen, L. Scheideler, R. Reichl and J. Geis-Gerstorfer, *Biomaterials*, 2004, **25**, 4087–4103.

[75] W. Hebenstreit, N. Ruzycki, G. S. Herman, Y. Gao and U. Diebold, *Physical Review B*, 2000, **62**, R16334–R16336.

[76] X. Q. Gong and A. Selloni, *Journal of Catalysis*, 2007, **249**, 134–139.

[77] M. Lazzeri, A. Vittadini and A. Selloni, *Physical Review B*, 2001, **63**, 155409.

[78] D. Erdemir, A. Y. Lee and A. S. Myerson, *Accounts of Chemical Research*, 2009, **42**, 621–629.

[79] D. Gebauer, M. Kellermeier, J. D. Gale, L. Bergström and H. Cölfen, *Chemical Society reviews*, 2014, **43**, 2348–2371.

[80] K. Onuma and A. Ito, *Chemistry of Materials*, 1998, **10**, 3346–3351.

[81] G. Treboux, P. Layrolle, N. Kanzaki, K. Onuma and A. Ito, *Journal of the American Chemical Society*, 2000, **122**, 8323–8324.

[82] D. Laurencin, N. Almora-Barrios, N. H. de Leeuw, C. Gervais, C. Bonhomme, F. Mauri, W. Chrzanowski, J. C. Knowles, R. J. Newport, A. Wong, Z. Gan and M. E. Smith, *Biomaterials*, 2011, **32**, 1826–1837.

[83] Q. Hu, M. H. Nielsen, C. L. Freeman, L. M. Hamm, J. Tao, J. R. I. Lee, T. Y. J. Han, U. Becker, J. H. Harding, P. M. Dove and J. J. De Yoreo, *Faraday Discussions*, 2012, **159**, 509–523.

[84] D. Gebauer, A. Völkel and H. Cölfen, *Science*, 2008, **322**, 1819–1822.

[85] J. Baumgartner, A. Dey, P. H. H. Bomans, C. Le Coadou, P. Fratzl, N. A. J. M. Sommerdijk and D. Faivre, *Nature materials*, 2013, **12**, 310–314.

[86] R. Demichelis, P. Raiteri, J. D. Gale, D. Quigley and D. Gebauer, *Nature Communications*, 2011, **2**, 590–598.

[87] L.-W. Du, S. Bian, B.-D. Gou, Y. Jiang, J. Huang, Y.-X. Gao, Y.-D. Zhao, W. Wen, T.-L. Zhang and K. Wang, *Crystal Growth & Design*, 2013, **13**, 3103–3109.

[88] H. Zhang, G. A. Waychunas and J. F. Banfield, *The Journal of Physical Chemistry B*, 2015, **119**, 10630–10642.

[89] J. P. Jolivet, E. Tronc and C. Chanéac, *Comptes Rendus - Geoscience*, 2006, **338**, 488–497.

[90] R. Iler, *The Chemistry of Silica: Solubility, Polymerization, Colloid and Surface Properties and Biochemistry of Silica*, Wiley, New York, 1979.

[91] J. Hutter, M. Iannuzzi, F. Schiffmann and J. VandeVondele, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2014, **4**, 15–25.

[92] U. Borštnik, J. Vandevondele, V. Weber and J. Hutter, *Parallel Computing*, 2014, **40**, 47–58.

[93] M. Krack, *Theoretical Chemistry Accounts*, 2005, **114**, 145–152.

[94] J. Vandevondele, M. Krack, F. Mohamed, M. Parrinello, T. Chassaing and J. Hutter, *Computer Physics Communications*, 2005, **167**, 103–128.

[95] M. Frigo and S. G. Johnson, *Proceedings of the IEEE*, 2005, **93**, 216–231.

[96] J. VandeVondele and J. Hutter, *Journal of Chemical Physics*, 2003, **118**, 4365–4369.

[97] G. J. Martyna and M. E. Tuckerman, *The Journal of Chemical Physics*, 1999, **110**, 2810–2821.

[98] C. Hartwigsen, S. Goedecker and J. Hutter, *Physical Review B*, 1998, **58**, 3641–3662.

[99] G. Lippert, J. Hutter and M. Parrinello, *Molecular Physics*, 1997, **92**, 477–488.

[100] J. P. Perdew, K. Burke and M. Ernzerhof, *Physical Review Letters*, 1996, **77**, 3865–3868.

[101] S. Goedecker, M. Teter and J. Hutter, *Physical Review B*, 1996, **54**, 1703–1710.

[102] I. T. Todorov, W. Smith, K. Trachenko and M. T. Dove, *Journal of Materials Chemistry*, 2006, **16**, 1911–1918.

[103] G. Kresse and J. Hafner, *Physical Review B*, 1993, **47**, 558–561.

[104] G. Kresse and J. Hafner, *Physical Review B*, 1994, **49**, 14251–14269.

[105] G. Kresse and J. Furthmüller, *Computational Materials Science*, 1996, **6**, 15–50.

[106] G. Kresse and J. Furthmüller, *Physical Review B*, 1996, **54**, 11169–11186.

[107] J. D. Gale, *J.Chem. Soc., Faraday Trans.*, 1997, **93**, 629–637.

[108] J. D. Gale and A. L. Rohl, *Molecular Simulation*, 2003, **29**, 291–341.

[109] J. D. Gale, *Zeitschrift fur Kristallographie*, 2005, **220**, 552–554.

[110] G. W. Watson, E. T. Kelsey, N. H. de Leeuw, D. J. Harris and S. C. Parker, *Journal of the Chemical Society, Faraday Transactions*, 1996, **92**, 433–438.

[111] E. Schrödinger, *Physical Review*, 1926, **28**, 1049–1070.

[112] C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, John Wiley & Sons Ltd, Chichester, UK, 2nd edn., 2004.

[113] M. Born and R. Oppenheimer, *Annalen der Physik*, 1927, **389**, 457–484.

[114] P. Hohenberg and W. Kohn, *Physical Review B*, 1964, **136**, 864–871.

[115] I. Levine, *Quantum Chemistry*, Pearson Prentice Hall, Upper Saddle River, NJ, 6th edn., 2009.

[116] W. Kohn and L. J. Sham, *Physical Review*, 1965, **140**, 1133–1138.

[117] J. C. Slater, *Physical Review*, 1930, **36**, 57–64.

[118] S. Boys, *Proceedings of the Royal Society of Chemistry*, 1950, **200**, 542–554.

[119] F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons Ltd, Chichester, UK, 3rd edn., 2017.

[120] F. Bloch, *Zeitschrift Fur Physik*, 1929, **52**, 555–600.

[121] M. Krack and M. Parrinello, *Phys. Chem. Chem. Phys.*, 2000, **2**, 2105–2112.

[122] D. Hamann, M. Schlüter and C. Chiang, *Physical Review Letters*, 1979, **43**, 1494–1497.

[123] D. Vanderbilt, *Physical Review B*, 1990, **41**, 7892–7895.

[124] P. Blochl, *Physical Review B*, 1994, **50**, 17953–17979.

[125] G. Kresse, *Physical Review B*, 1999, **59**, 1758–1775.

[126] R. Car and M. Parrinello, *Physical Review Letters*, 1985, **55**, 2471–2474.

[127] A. Leach, *Molecular Modelling Principles and Applications*, Pearson Prentice Hall, Harlow, UK, 2nd edn., 2001.

[128] L. Verlet, *Physical Review*, 1967, **159**, 98–103.

[129] W. C. Swope, *Journal of Chemical Physics*, 1982, **76**, 637–649.

[130] J. Gibbs, *Elementary principles in Statistical mechanics*, Scribner's sons, New York, 1902.

[131] S. Nosé, *Journal of Chemical Physics*, 1984, **81**, 511–519.

[132] W. G. Hoover, *Physical Review A*, 1985, **31**, 1695–1697.

[133] S. Melchionna, G. Ciccotti and B. Lee, *Molecular Physics*, 1993, **78**, 533–544.

[134] G. J. Martyna, D. J. Tobias and M. L. Klein, *The Journal of Chemical Physics*, 1994, **101**, 4177–4189.

[135] D. Adams, *CCP5 Quarterly Reports*, 1983, **10**, 30–36.

[136] P. P. Ewald, *Annalen der Physik*, 1921, **369**, 253–287.

[137] S. Lifson and A. Warshel, *The Journal of Chemical Physics*, 1968, **49**, 5116–5129.

[138] J. M. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

[139] R. Buckingham, *Proceedings of the Royal Society of London*, 1938, **168**, 264–283.

[140] G. Mie, *Annalen der Physik*, 1903, **316**, 657–697.

[141] J. H. R. Clarke, W. Smith and L. V. Woodcock, *J. Chem. Phys.*, 1986, **84**, 2290–2294.

[142] B. G. Dick and A. W. Overhauser, *Physical Review*, 1958, **112**, 90–103.

[143] C. R. A. Catlow and W. Mackrodt, *Computer Simulation of Solids*, Springer-Verlag, 1982.

[144] P. J. Mitchell and D. Fincham, *Journal of Physics: Condensed Matter*, 1993, **5**, 1031–1038.

[145] M. J. Sanders, M. Leslie and C. R. A. Catlow, *Journal of the Chemical Society, Chemical Communications*, 1984, **4**, 1271–1273.

[146] J. D. Bernal and R. H. Fowler, *The Journal of Chemical Physics*, 1933, **1**, 515–548.

[147] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *The Journal of Chemical Physics*, 1983, **79**, 926–935.

[148] M. W. Mahoney and W. L. Jorgensen, *Journal of Chemical Physics*, 2001, **115**, 10758–10768.

[149] B. Guillot, *Journal of Molecular Liquids*, 2002, **101**, 219–260.

[150] M. V. Fernández-Serra and E. Artacho, *Journal of Chemical Physics*, 2004, **121**, 11136–11144.

[151] J. C. Grossman, E. Schwegler, E. W. Draeger, F. Gygi and G. Galli, *The Journal of Chemical Physics*, 2004, **120**, 300–11.

[152] T. Todorova, A. P. Seitsonen, J. Hutter, I.-F. W. Kuo and C. J. Mundy, *The journal of physical chemistry. B*, 2006, **110**, 3685–91.

[153] G. Torrie and J. Valleau, *Journal of Computational Physics*, 1977, **23**, 187–199.

[154] A. Laio and M. Parrinello, *Proceedings of the National Academy of Sciences of the United States of America*, 2002, **99**, 12562–12566.

[155] R. Radhakrishnan and B. L. Trout, *Journal of the American Chemical Society*, 2003, **125**, 7743–7747.

[156] S. Csajka, D. Chandler, C. Dellago and P. G. Bolhuis, *J. Chem. Phys.*, 1998, **108**, 1964–1977.

[157] A. Kawska, J. Brickmann, O. Hochrein and D. Zahn, *Zeitschrift fur Anorganische und Allgemeine Chemie*, 2005, **631**, 1172–1176.

[158] A. Kawska, J. Brickmann, R. Kniep, O. Hochrein and D. Zahn, *Journal of Chemical Physics*, 2006, **124**, 0–7.

[159] J. Kästner, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2011, **1**, 932–942.

[160] A. Grossfield, *An implementation of WHAM: the weighted histogram analysis method, version 2.0.9*, 2004.

[161] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 1011–1021.

[162] P. A. Kollman, *J. Comput. Chem.*, 1994, **16**, 1339–1350.

[163] B. Roux, *Comput. Phys. Commun.*, 1995, **91**, 275–282.

[164] J. Kästner and W. Thiel, *Journal of Chemical Physics*, 2005, **123**, 144104.

[165] M. Corno, F. Chiatti, A. Pedone and P. Ugliengo, in *Biomaterials-Physics and Chemistry*, ed. R. Pignatello, InTech, 2011, ch. 14, pp. 275–298.

[166] I. Lin, A. P. Seitsonen, M. D. Coutinho-Neto, I. Tavernelli and U. Rothlisberger, *The journal of physical chemistry. B*, 2009, **113**, 1127–1131.

[167] P. H. L. Sit and N. Marzari, *Journal of Chemical Physics*, 2004, **122**, 204510.

[168] N. Almora-Barrios and N. H. de Leeuw, *Crystal Growth and Design*, 2012, **12**, 756–763.

[169] P. Wernet, D. Nordlund, U. Bergmann, M. Cavalleri, M. Odelius, H. Ogasawara, L. A. Näslund, T. K. Hirsch, L. Ojamäe, P. Glatzel, L. G. M. Pettersson and A. Nilsson, *Science*, 2004, **304**, 995–999.

[170] D. Di Tommaso and N. H. de Leeuw, *J. Phys. Chem B*, 2008, **112**, 6965–6975.

[171] T. Dudev and C. Lim, *J. Phys. Chem. B*, 2004, **108**, 4546–4557.

[172] F. Sowrey, L. Skipper, D. Pickup, K. Drake, Z. Lin, M. Smith and R. Newport, *Phys. Chem. Chem. Phys.*, 2004, **6**, 188–192.

[173] J. De Yoreo, *Nature Materials*, 2013, **12**, 284–285.

[174] W. A. House, *Environmental Technology*, 1999, **20**, 727–733.

[175] S. Jiang, H. Pan, Y. Chen, X. Xu and R. Tang, *Faraday Discuss.*, 2015, **179**, 451–461.

[176] M. J. J. M. van Kemenade and P. L. de Bruyn, *Journal of Colloid And Interface Science*, 1987, **118**, 564–585.

[177] Q. Zhang, Y. Jiang, B. D. Gou, J. Huang, Y. X. Gao, J. T. Zhao, L. Zheng, Y. D. Zhao, T. L. Zhang and K. Wang, *Crystal Growth & Design*, 2015, **15**, 2204–2210.

[178] A. Oyane, K. Onuma, T. Kokubo and A. Ito, *The Journal of Physical Chemistry B*, 1999, **103**, 8230–8235.

[179] M. Corno, A. Rimola, V. Bolis and P. Ugliengo, *Phys. Chem. Chem. Phys.*, 2010, **12**, 6309–6329.

[180] L. Wang, S. Li, E. Ruiz-Agudo, C. V. Putnis and A. Putnis, *CrystEngComm*, 2012, **14**, 6252–6256.

[181] G. Treboux, P. Layrolle, N. Kanzaki, K. Onuma and A. Ito, *The Journal of Physical Chemistry A*, 2000, **104**, 5111–5114.

[182] J. Ma, *Journal of Materials Science*, 2014, **49**, 3099–3106.

[183] M. Tanaka, G. Girard, R. Davis, A. Peuto and N. Bignell, *Metrologia*, 2001, **38**, 301–309.

[184] D. Kashchiev and G. M. van Rosmalen, *Crystal Research and Technology*, 2003, **38**, 555–574.

[185] J. Anwar and D. Zahn, *Angewandte Chemie - International Edition*, 2011, **50**, 1996–2013.

[186] J. H. Weaver and P. R. Frederikse, *CRC Handbook of Chemistry and Physics (Internet Version)*, CRC Press, 2016, pp. 5–92 5–93.

[187] A. Oyane, H.-M. Kim, T. Furuya, T. Kokubo, T. Miyazaki and T. Nakamura, *J Biomed Mater Res*, 2003, **65**, 188–195.

[188] B. Xie, T. J. Halter, B. M. Borah and G. H. Nancollas, *Crystal Growth and Design*, 2014, **14**, 1659–1665.

[189] D. Laurencin, A. Wong, W. Chrzanowski, J. C. Knowles, D. Qiu, D. M. Pickup, R. J. Newport, Z. Gan, M. J. Duer and M. E. Smith, *Phys. Chem. Chem. Phys.*, 2010, **12**, 1081–1091.

[190] R. P. Heaney, *Journal of the American College of Nutrition*, 2006, **25**, 271S–276S.

[191] H. El Feki, J. M. Savariault, A. B. Salah and M. Jemal, *Solid State Sciences*, 2000, **2**, 577–586.

[192] Y. Tang, H. F. Chappell, M. T. Dove, R. J. Reeder and Y. J. Lee, *Biomaterials*, 2009, **30**, 2864–2872.

[193] S. Maltsev, M. J. Duer, R. C. Murray and C. Jaeger, *Journal of Materials Science*, 2007, **42**, 8804–8810.

[194] E. D. Eanes, in *Calcium Phosphates in Biological and Industrial Systems*, ed. Z. Amjad, Springer US, Boston, MA, 1998, pp. 21–39.

[195] K. Matsunaga, *Journal of the American Ceramic Society*, 2010, **93**, 1–14.

[196] P. Laing, *Clinical Experience with Prosthetic Materials: Historical Perspectives, Current Problems, and Future Directions*, American Society for Testing and Materials, 1979, vol. 684, pp. 199–211.

[197] M. Lorenzetti, O. Dakischew, K. Trinkaus, K. S. Lips, R. Schnettler, S. Kobe and S. Novak, *Journal of biomaterials applications*, 2015, **30**, 71–84.

[198] F. Barrère, P. Layrolle, C. A. Van Blitterswijk and K. De Groot, *Bone*, 1999, **25**, 107S–111S.

[199] P. K. Naicker, P. T. Cummings, H. Zhang and J. F. Banfield, *J. Phys. Chem. B*, 2005, **109**, 15243–15249.

[200] S. Grimme, *Journal of computational chemistry*, 2006, **27**, 1787–1799.

[201] S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *Journal of Chemical Physics*, 2010, **132**, 154104.

[202] A. Ambrosetti, D. Alfè, R. A. Distasio and A. Tkatchenko, *Journal of Physical Chemistry Letters*, 2014, **5**, 849–855.

[203] S. Grimme, S. Ehrlich and L. Goerigk, *Journal of computational chemistry*, 2011, **32**, 1456–1465.

[204] H. J. Monkhorst and J. D. Pack, *Physical Review B*, 1976, **13**, 5188–5192.

[205] K. Ho, C. Fu, B. Harmon, W. Weber and D. Hamann, *Physical Review Letters*, 1982, **49**, 673–676.

[206] C. L. Fu and K. M. Ho, *Physical Review B*, 1983, **28**, 5480–5486.

[207] G. Mancardi, C. E. Hernandez Tamargo, D. Di Tommaso and N. H. de Leeuw, *J. Mater. Chem. B*, 2017, **5**, 7274–7284.

[208] P. Oliver, G. Watson, E. Kelsey and S. Parker, *Journal of Materials Chemistry*, 1997, **7**, 563–568.

[209] V. Swamy, J. Gale and L. Dubrovinsky, *Journal of Physics and Chemistry of Solids*, 2001, **62**, 887–895.

[210] M. Predota, A. V. Bandura, P. T. Cummings, J. D. Kubicki, D. J. Wesolowski, A. A. Chialvo and M. L. Machesky, *J. Phys. Chem. B*, 2004, 12049–12060.

[211] V. N. Koparde and P. T. Cummings, *Journal of Physical Chemistry B*, 2005, **109**, 24280–24287.

[212] M. Alimohammadi and K. A. Fichthorn, *Nano Letters*, 2009, **9**, 4198–4203.

[213] R. Hill, *Proceedings of the Physical Society. Section A*, 1952, **65**, 349–354.

[214] G. Samsonov, *The Oxide Handbook*, IVI/Plenum Press, New York, 1982.

[215] K. P. Schröder, J. Sauer, M. Leslie, C. Richard, A. Catlow and J. M. Thomas, *Chemical Physics Letters*, 1992, **188**, 320–325.

[216] H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma, *Journal of Physical Chemistry*, 1987, **91**, 6269–6271.

[217] A. S. Barnard, P. Zapol and L. A. Curtiss, *Journal of Chemical Theory and Computation*, 2005, **1**, 107–116.

[218] U. Terranova and D. R. Bowler, *Journal of Materials Chemistry*, 2011, **21**, 4197–4203.

[219] C. Arrouvel, M. Digne, M. Breysse, H. Toulhoat and P. Raybaud, *Journal of Catalysis*, 2004, **222**, 152–166.

[220] A. Tilocca and A. Selloni, *Langmuir*, 2004, **20**, 8379–8384.

[221] T. Kokubo, *Acta Materialia*, 1998, **46**, 2519–2527.

[222] H. Takadama, H. M. Kim, T. Kokubo and T. Nakamura, *Journal of biomedical materials research*, 2001, **55**, 185–193.

[223] G. Mancardi, U. Terranova and N. H. de Leeuw, *Crystal Growth & Design*, 2016, **16**, 3353–3358.