University of
HUDDERSFIELD

Modelling Magnetite Biomineralisation: The Interactions of Proteins and Fe_3O_4 Surfaces

Amy Monnington
E-mail: Amy.Monnington@hud.ac.uk

May 12, 2014

Abstract

The biosynthesis of magnetite is the earliest known example of biomineralisation; however, much of the detailed atomistic mechanisms by which the process occurs are unknown. Within the bacterial strain *Magnetospirillum magneticum* AMB-1, the formation of magnetite nanoparticles is thought to occur under the influence of the Mms6 protein. The C–terminal of this protein is highly acidic, containing dense carboxyl and hydroxyl groups, and exhibits direct interaction with the magnetite surface. In this thesis, a novel atomistic model of Mms6-driven magnetite formation was developed and the interactions of amino acids, dipeptide, tetrapeptide and pentapeptide sequences, related to the C-terminus of Mms6, with the {100} and {111} magnetite surfaces (both *in vacu* and solvated) have been investigated. Each study was split into two systems; a classic molecular dynamics system and a constrained molecular dynamics system utilising the Potential of Mean Force.

Initially, the attachment of the individual amino acids to magnetite surfaces was considered. From these results, it was established that the {111} surface was the favoured for surface for amino acid attachment and bonding occurred through octahedral iron ions, rather than tetrahedral iron ions. Furthermore, the charged amino acids demonstrated a higher affinity for iron binding and solvation of unconstrained systems diminished the iron binding abilities of all the amino acids.

Secondly, based on a glutamate repeat motif, the attachment of a series of di- and tetrapeptides to the {100} and {111} magnetite surfaces was explored. It was hypothesised that if the negatively charged glutamate was substituted for a charge neutral alanine, the iron binding potential of the sequence would reduce. The results suggested that the substitution of glutamate for alanine significantly reduces the iron binding affinity of the system on the {100} surface, irrespective of sequence length and composition. However, on the {111} surface, the introduction of alanine differentially modulates the iron binding activity of the sequences investigated. Sequential substitution in a two amino acid chain confers inhibition of iron binding, conversely, in a four amino acid chain, iron binding affinity is enhanced.

The final chapter utilised pentapeptides taken from the Cterminal region, thus ensuring the full sequence was explored. The binding behaviour of these pentapeptides and their related mutants, were investigated. It was found that the different sections behaved differently from each other, suggesting that the binding activity of the C-terminal sequence is partly dependent on how the amino acids interact with each other. It was theorised that sequence mutation would decrease iron binding; however, the data suggested that this was not always the case and was sequence dependent. Based on the constrained system data, mutation of the original sequences confirmed the hypothesis for DIESA, LRDAL and EVELR on the {100} surface, and for SRDIE and SDEEV on the {111} surface, whereas, the theory was contradicted for the counterparts surfaces and for both surfaces of ELRDA. This data also suggests that the {111} surface was the preferred surface of attachment, with the exception of LRDAL. For the unconstrained systems, the observations differed dependent on the data analysis technique utilised, as well as on the pentapeptide original sequence, with none of the sequences explored fully confirming the hypothesis. Furthermore, the presence of water in the unconstrained systems was detrimental to the iron binding potential of the pentapeptides. The data from both the unconstrained and constrained systems propose that, there are many factors affecting the iron binding ability other than sequence mutation, such as, surface type, iron type and sequence dependence.

Copyright Statement

Acknowledgements

Undertaking this PhD has been one of the most challenging experiences of my life so far. Without the help, patience and support of the following amazing people this project would not have been completed.

First and foremost, I would like to thank my supervisor, Dr. David Cooke, for his continuous encouragement and guidance. He was always available to give advice and on occasion lend an ear for me to have a rant. I have learnt so much from him and the opportunities that were opened up to me during this time have been life changing. Through this I have grown immensely in personal confidence, for which I am eternally grateful. I have also had the privilege of travelling the world to present my work which has been incredible.

Valuable support was received from Dr Sarah Staniland and her group, by way of experimental data and useful discussion. I would like to acknowledge the University of Huddersfield for funding my studentship and the EPSRC and the University of Huddersfield HPCRC for computer time.

The best and worst moments of my doctoral journey have been shared with my friends at the University of Huddersfield, and I am deeply grateful to have had you there along the way. A special mention is needed for Heidi Joao and Sam Bullock. Thank you so much ladies for providing a safe haven away from work to unwind and have fun. I feel that together we have learnt to just roll with the punches and embrace the madness (and that it is totally normal to break down and cry for no apparent reason!!).

I also wish to thank my family, for their affection and encouragement throughout both this PhD and life in general, in particular, my dad Tony Monnington. You sacrificed so much in life for me and your love and support has been unconditional all of these years. It is your example that I try to emulate in all that I do. I could never ever wish for a better dad and I truly hope I have made you proud. Thank you!

One of the most amazing things to have come out of my PhD is my brilliant fiancé and best friend, Dr Daniel Patten. To you I owe so much. You have been my absolute rock and a true inspiration throughout this project. You believed in me even when I didn't even believe in myself, and made me smile at times when I didn't think it was possible. There are not enough words to express my gratitude to you, so I shall just say, I love you with all of my heart, you are my soul mate and my world, and this thesis is dedicated to you.

Contents

List of Figures

List of Tables

List of Abbreviations

- AMBER Assisted Model Building with Energy Refinement
- BCM Biologically Controlled Mineralisation
- BIM Biologically Induced Mineralisation
- CoM Centre of mass
- CN Coordination number
- DFT Density Functional Theory
- D O Aspartate side chain oxygen
- EM Energy minimisation
- E O Glutamate side chain oxygen
- FEO Octahedral iron bonding
- FET Tetrahedral iron bonding
- GTPase Guanosine triphosphotase
- MD Molecular dynamics
- METADISE Minimum Energy Technique Applied to Dislocations, Interfaces and Surfaces Energies
- MM Magnetosome membrane
- Mms Magnetosome membrane specific
- MNP Magnetic Nanoparticles
- MTB Magnetotactic bacteria
- NBO N-terminal peptide bond oxygen
- OATZ Oxic-anoxic transition zone
- PMF Potential of mean force
- Q O Glutamine side chain oxygen
- RDF Radial distribution function
- RT Residence time
- SD Single–magnetic domains
- S O Serine side chain oxygen
- x-x PBO Inter–sequence peptide bond oxygen

Chapter 1

Introduction

1.1 Biomineralisation

"Biomineralisation: the study of the formation, structure and properties of inorganic solids deposited in biological system"[1].

Biomineralisation occurs throughout the natural world, from unicellular prokaryotes to the skeletal tissue of higher mammals, and refers to the processes whereby organisms deposit minerals in order to form structures such as shells, teeth and bones[2, 3]. These processes are distinguished from abiotic mineralisation by the control exerted over mineral formation by the organisms. Biominerals are complex organic–inorganic composites, where an organic matrix component plays a significant role in the mineral formation. It has been suggested that biominerals habitually have properties quite dissimilar to their inorganically produced counterparts such as shape, size, and trace elemental compositions, due to their formation under controlled conditions.

Biomineralisation processes are divided into two fundamentally different groups, which are classified according to the degree of biological control[4, 5]. Biologically induced mineralisation (BIM) is a passive process whereby microbes sorb solutes onto their cell surface and/or release reactive metabolites. This subsequently alters the saturation state of the solution surrounding the cell surface, changing the chemical equilibrium of the surrounding environment causing BIM mineral formation. These changes can be triggered by events such as the introduction of biologically produced metabolic end–products, the release of particular cations by the cell, or as a by–product of cell surface charge[6]. The biological system has minimal control over the type of minerals precipitated, the composition and properties of the minerals formed are greatly dependent on environmental conditions[7]. The biominerals formed are generally characterized by broad particle–size distribution,

poor crystallinity, lack of specific crystal morphologies and chemical heterogeneousity[8], also appearing to have no specific recognized functions. In contrast, biologically controlled mineralisation (BCM) is an active process. The organisms involved exert an enormous degree of control, both chemical and genetic, over all aspects of nucleation and crystal growth stages[9]. It is presumed this is due to the biominerals produced serving some physiological function. Within BCM geochemical conditions independent from the bulk solution are created by sealing off specific sites within the cytoplasm or on the cell wall from the external environment[6]. Specific ions sequestering by the cell occurs, these ions are subsequently transferred into these intracellular compartments, where the concentration is increased until supersaturation (this refers to a solution that contains more of the dissolved material than could be dissolved by the solvent under normal circumstances). Biomineral nucleation site, crystal growth and morphology are actively controlled by the organism. The biominerals that are formed are characterized by well–ordered crystals with a relatively narrow size distribution, specific consistent morphologies and of high chemical purity[9, 10].

The ability to form minerals within first prokaryotes and then eukaryotes has established over the last 3.5 billion years. Many of the different minerals known to date appeared about 540 million years ago when living systems first began to exert control over biomineral formation[11]. There have been over sixty different minerals, from many different phyla, identified so far. Whilst Table 1.1 represents the names and corresponding chemical formulae of some of these organism produced minerals, this is by no means a complete list as new biogenic minerals continue to be discovered.

Name	Formula
Calcite, Aragonite & Vaterite	$CaCO_3$
Dolomite & Protodolomite	$(Mg_xCa_{1-x})CO_3$ & $CaMg(CO_3)_2$
Hydrocerussite	$Pb_3(CO_3)_2(OH)_2$
Amorphous calcium carbonate	$CaCO_3{\cdot}H_2O$ or $CaCO_3$
Octacalcium phosphate	$Ca_8H_2(PO_4)_6$
Brushite	$CaHPO_4 \cdot 2H_2O$
Hydroxyapatite & Francolite	$Ca_{10}(PO_4)_6(OH)_2$ & $Ca_{10}(PO_4)_6F_2$
Whitlockite	$Ca_{18}H_2(Mg,Fe)_2^{+2}(PO_4)_{14}$
Struvite	$Mg(NH_4)(PO_4){\cdot}6H_2O$
Vivianite	$Fe_3^{+2}(PO_4)_2 \cdot 8H_2O$
Gypsum, Barite & Celestite	$CaSO_4 \cdot 2H_2O$, $BaSO_4$ & $SrSO_4$
Jarosite	$KFe_3^{+3}(SO_4)_2(OH)_6$
Pyrite	FeS_2
Hydrotroilite	$FeS{\cdot}nH_2O$
Sphalerite & Wurtzite	ZnS
Galena	PbS
Greigite	Fe_3S_4
Mackinawite	$(Fe,Ni)_9S_8$
Acanthite	Ag_2S
Amorphous silica	$SiO_2{\cdot}nH_2O$
Atacamite	$Cu_2Cl(OH)_3$
Fluorite	CaF_2
Hieratite	K_2SiF_6
Magnetite	Fe_3O_4
Goethite	α-FeOOH
Lepidocrocite	γ-FeOOH
Ferrihydrite	$5Fe_2O_3 \cdot 9H_2O$
Todorokite	$(Mn^+2CaMg)Mn_3^{+4}O_7{\cdot}H_2O$
Birnessite	$Na_4Mn_{14}O_{27} \cdot 9H_2O$

Table 1.1: Names and chemical compositions of biominerals produced by BCM and BIM processes. Adapted from Weiner and Dove 2003[11].

It has been identified that approximately 50% of known biominerals are calcium–bearing minerals[5], with the calcium carbonate biogenic minerals being the most abundant. They are widespread among many different taxa[5], with eight known polymorphs, three of which are pure calcium carbonate; calcite, aragonite and vaterite. Within molluscs, calcite has been found to form the outer layer of its shell. Calcite has also been identified as an inner–ear gravity receptor in mammals, whereas, aragonite has been found to be a gravity receptor in fish. Protective spicules of sea squirts are formed from vaterite. Of the known biogenic mineral types phosphates comprise around 25%. The majority of phosphate minerals are produced by the BCM method, with the exception of struvite and brushite. Hydroxyapatite is the most abundantly produced of these[5], being the mineral present in the bones and teeth of vertebrates. 25% of the known biominerals are amorphous. Amorphous silica is a commonly formed example, being present in the teeth of limpets and the tips of the stinging hairs on nettles to name but a few functions.

Iron biominerals comprise approximately 40% of all minerals formed by organisms[12, 10]. They are not readily evaluated by mineral class because they have significant occurrences as oxides, hydroxides and sulphides[13, 14]. Some sulphates and phosphates have also been reported[15].

1.1.1 Iron Oxide and Oxyhydroxide Biomineralisation

Iron is one of the most abundant elements on the Earth and many organisms are known to contain minerals of iron. It has two valence states: ferric iron (Fe^{3+}), and ferrous iron (Fe^{2+}). This profusion of iron minerals could perhaps be on account of, the significant role of iron in numerous metabolic processes. The production of iron biominerals allows for the accumulation of iron, for future metabolic needs, within the organisms, at the same time as avoiding high concentrations of intracellular ferrous iron[16].

A key group of iron minerals are the iron oxides and oxyhydroxides, occurring as several phases. This have been reported in geological deposits, environmental water sediments, industrial processes, and as biological system components in the case of some phases[17, 18, 19, 20]. The ferric oxide and oxyhydroxide biominerals are an important class of iron biomineral, with examples exhibited in Table 1.2. These biominerals occur as amorphous, colloidal precipitates, as quasi–crystalline minerals, or as crystalline minerals[16].

Iron oxide and oxyhydroxide biominerals are formed by a wide variety of organisms, in which they serve various functions (see Table 1.3), such as, usage in the strengthening and

Oxides		Oxyhydroxides	
Formula	Mineral	Formula	Mineral
$5Fe_2O_3 \cdot 9H_2O$	ferrihydrite **	α-FeOOH	goethite **
α-Fe_2O_3	hematite	β-FeOOH	akaganeite
γ-Fe_2O_3	maghemite **	γ-FeOOH	lepidocrocite **
Fe_3O_4	magnetite **	δ-FeOOH	feroxyhte

Table 1.2: The major iron oxides and oxyhydroxides[21]. ** Found in biological systems.

hardening of tissues[16, 22]. They are also associated with iron overload diseases and are involved in intracellular iron storage and detoxification. A linkage has also been found between the biominerals of iron oxide and oxyhydroxide and the sensing of magnetic fields for magnetic navigation[23, 24, 25, 26]. Studies have even found crystals of magnetic iron minerals in the human brain[27, 28]. Konhauser[14] discussed at length the subject of bacterial iron biomineralisation.

Mineral	Formula	Organism	Location	Function
magnetite	Fe_3O_4	bacteria	intracellular	magnetotaxis
		chitons	teeth	teeth hardening
		salmon/honeybees	head	magnetic navigation
geothite	α-FeOOH	limpets	teeth	teeth hardening
lepidocrocite	γ-FeOOH	sponges	filaments	unknown
		chitons	teeth	teeth hardening
ferrihydrite	$5Fe_2O_3 \cdot 9H_2O$	plants	phytoferritin	Fe storage
		animals/humans	ferritin	Fe storage
		bacteria	intracellular	metabolic byproduct
		chitons	teeth	precursor phase
		beaver/rat/fish	tooth surface	mechanical strength
	+ phosphate	bacteria	ferritin	Fe storage
		sea cucumber	dermis	mechanical strength

Table 1.3: Iron oxide biominerals and their functions[1, 19, 16].

Much of the research on biomineralisation to date is aimed at BCM in an attempt to elucidate the structure of the organic macromolecules involved in the processes and also discovering the nature of the related organic–inorganic interactions. It has been deduced that iron biomineral formation in eukaryotic organisms is almost exclusively BCM. An

important iron biomineral that can be formed by the BCM method is the iron oxide, magnetite (Fe_3O_4).

1.1.2 Magnetite, Magnetotactic Bacteria and Magnetosomes

Of all known naturally occurring minerals on Earth, magnetite is the most magnetic and is also one of the two most common iron oxides[29]. Magnetite can be thought of as having an inverse spinel crystal structure, which has the general formula AB_2O_4. However, most recent studies of magnetite have suggested that this is an ideal case and instead the structure is based on a face–centred cubic array of oxide ions with A ions, in the case of magnetite Fe^{2+}, occupying octahedral holes and the B ions, Fe^{3+}, are equally distributed between octahedral and tetrahedral holes. This crystal structure is represented in Figure 1.1.

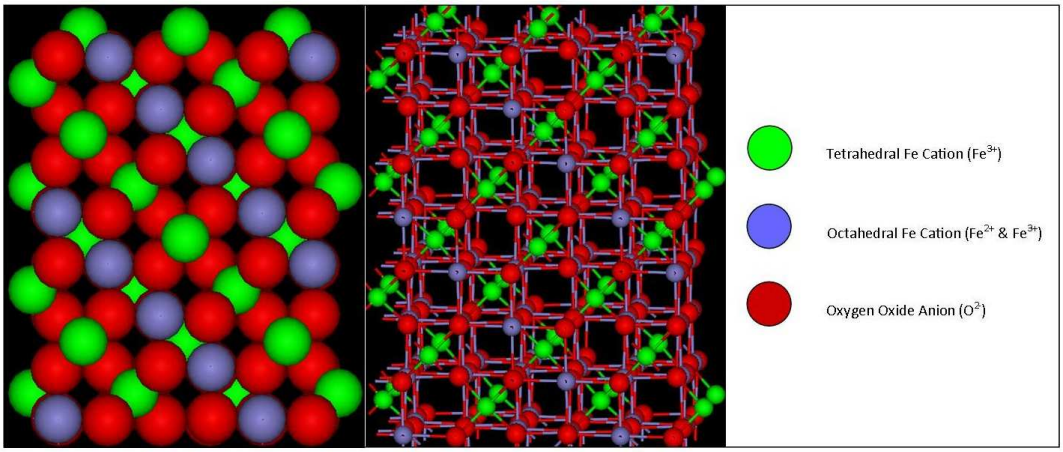

Figure 1.1: The inverse spinel crystal structure of magnetite.

In magnetite, the Fe^{2+} and Fe^{3+} ions have unpaired spins, leading to interesting magnetic properties. Fe^{3+} has a d5 electronic configuration and five unpaired electrons. The spins of the Fe^{3+} ions are cancelled out as half of the ions are on octahedral sites and half on tetrahedral sites. Thus the spins of the Fe^{3+} ions in octahedral and tetrahedral holes are antiparallel (due to the spins of the ions on the octahedral sites interact with those on the tetrahedral sites via the oxide ions) and no net magnetisation results from these ions. The Fe^{2+} ions have a d6 electronic configuration with four unpaired electrons and are responsible for the net magnetisation. These divalent ions tend to align their spins parallel to those of the trivalent ions in adjacent octahedral holes. This alignment confers a net magnetic moment on the crystal. This all leads to magnetite being ferrimagnetic

(see Figure 1.2), whereby the magnetic moments of the atoms on different sublattices are opposed, however, at the same time the opposing moments are unequal and a spontaneous magnetisation remains[30]. Magnetite is known to interact over 10^6 times more strongly with an external magnetic field than any other naturally occurring mineral[27]. Hence, magnetite interacts very strongly with the Earth's magnetic fields.

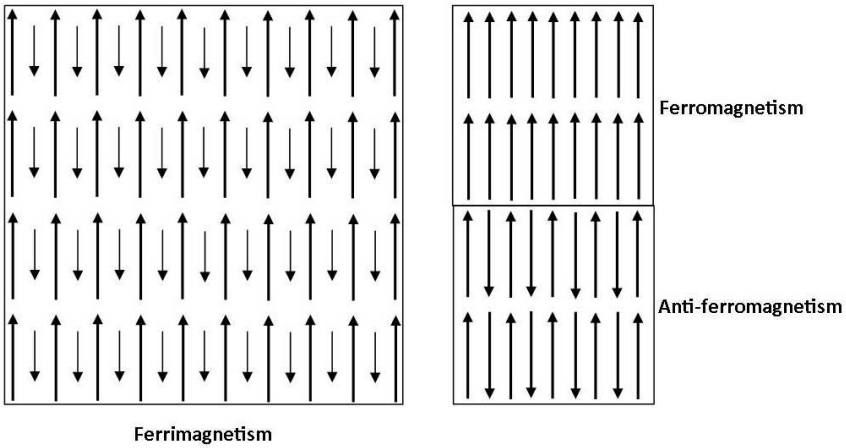

Figure 1.2: Ferrimagnetism; the magnetic moments of the atoms on different sublattices are opposed and unequal conferring a net magnetic moment.

In this work, the {100} and {111} magnetite surfaces were used. The {111} surface can be either iron terminated or oxygen terminated. In this instance, the iron terminated version was utilised as shown in Figure 1.3.

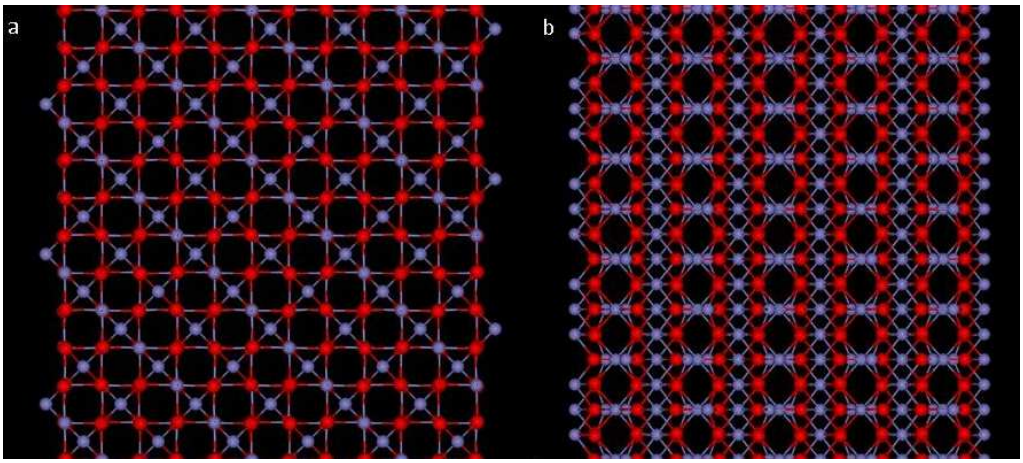

Figure 1.3: Magnetite surface types modelled in this work. a) {100} and b) {111} Fe terminated surface.

Inorganic magnetite is usually found in igneous and metamorphic rock, having an octahedral crystal habit, as shown in Figure 1.4. These crystals often have lattice dislocations and other crystal defects, broad size distributions, and the inclusion of impurities such as;

Mg, Zn, Mn, Ni, Cr, Ti, V and Al. However, as mentioned earlier magnetite can also be mineralised biologically.

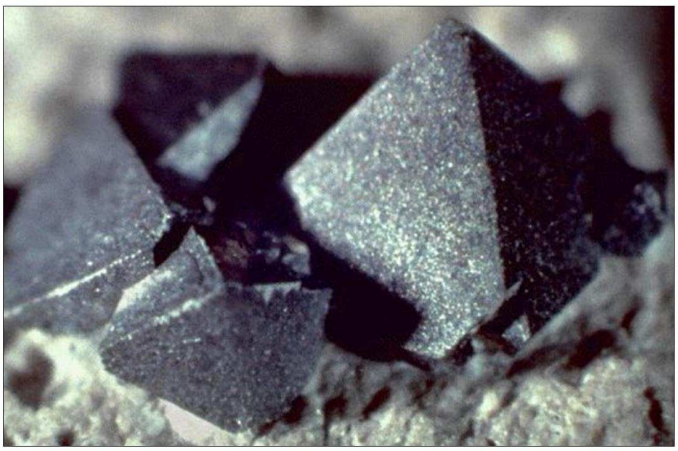

Figure 1.4: An example of inorganic magnetite. Note the octahedral crystal habit[31].

The earliest known example of biomineralisation is that of the biosynthesis of magnetite. Microfossils of magnetite were discovered in the Gunflint Chert, a sequence of banded iron formation rocks that are exposed in the Gunflint Range of northern Minnesota and north western Ontario along the north shore of Lake Superior[32, 33]. These can be traced back approximately two billion years[33], occurring long before biominerals became common within the Cambrian.

Magnetite is produced by both BIM and BCM in the same sedimentary environment; nevertheless, the most extensively researched method of magnetite mineralisation is BCM. Biogenic magnetite was first identified in the Molluscan class *Polyplacophora* as a capping material in the radula teeth of chitons[34, 35]. Prior to this discovery it was thought that magnetite only formed in igneous or metamorphic rock under high temperature and pressure. However, the major step forward in understanding the BCM formation of magnetite occurred due to the report of Blakemore in 1975, detailing the discovery of magnetotactic bacteria (MTB) in Cape Cod Marsh, Massachusetts[36]. When looking at Coccoid bacteria, Blakemore discovered that a large population were migrating in one direction however when an external magnetic field was applied it was noted that this direction could be reversed.

Blakemore[36, 37] discovered that MTB are a diverse group of Gram–negative bacteria (so called because they do not retain crystal violet dye used in the Gram staining protocol) that use a process known as magnetotaxis[38], in which they align and swim along Earth's geomagnetic field lines, using their flagellum, in order to migrate to their ideal habitat. By

taking up iron ions from the surrounding environment; intracellular, membrane–bound, nano–sized crystals of magnetite of high chemical purity[10], referred to as magnetosomes, are biomineralised by MTBs. These magnetosomes are usually aligned in chains, enabling magnetotaxis. Bazylinski[39] suggested that MTB produce these magnetic nano–particles (MNPs) under precise biochemical, chemical and genetic control. Blakemore[40] and Matsunaga[41] discovered that MTB take up approximately 100 times more iron than non–magnetic bacteria to synthesize intracellular MNPs. The molecular mechanism of MTB magnetite formation is hypothesized to be a distinct stepwise method; however the processes by which this occurs are still not clear.

MTB are ubiquitous to freshwater and marine habitats[42, 43, 44, 45, 46], preferring environments that contain low oxygen concentrations due to their microaerophilic or anaerobic tendencies[36, 39]. It has been found that MTB are most abundant at the oxic–anoxic transition zone (OATZ)[39, 47, 48], which, in most sediments, occur at or just below the sediment–water interface. The OATZ is an interface where oxygenated water or sediment, meets oxygen–deficient water or sediment. Bacteria located in environments with higher or lower than optimal oxygen concentration migrate to the ideal living conditions of the OATZ by rotating their flagella in a specific direction to move either forwards (counter clockwise) or backwards (clockwise). Figure 1.5 gives a basic illustration of this system of migration. The MTB have been found to migrate to the magnetic north in the Northern hemisphere[36], to the magnetic south in the Southern hemisphere[49, 50], and on the geomagnetic equator there is migration in both directions[51] (although at the equator there are much lower populations of MTB[32]).

MTBs are a heterogeneous group of prokaryotes known as α-proteobacteria. The morphological types are diverse, with many cocci, rods, vibrios, spirilla and multicellular forms having been derived[37, 53, 54, 55]. Within MTBs, MNPs of uniform size and morphology are formed; however, it has been observed that compositions and morphologies are species- or strain–dependant, thus suggesting an element of high biological control [56, 57]. Three of the most commonly occurring morphologies for MNPs are; elongated prismatic, roughly cuboidal and tooth/bullet shaped[36, 39, 58]. The idealised magnetosome morphologies are based on a combination of;

- {100} faces = cube
- {110} faces = dodecahedron
- {111} faces = octahedron

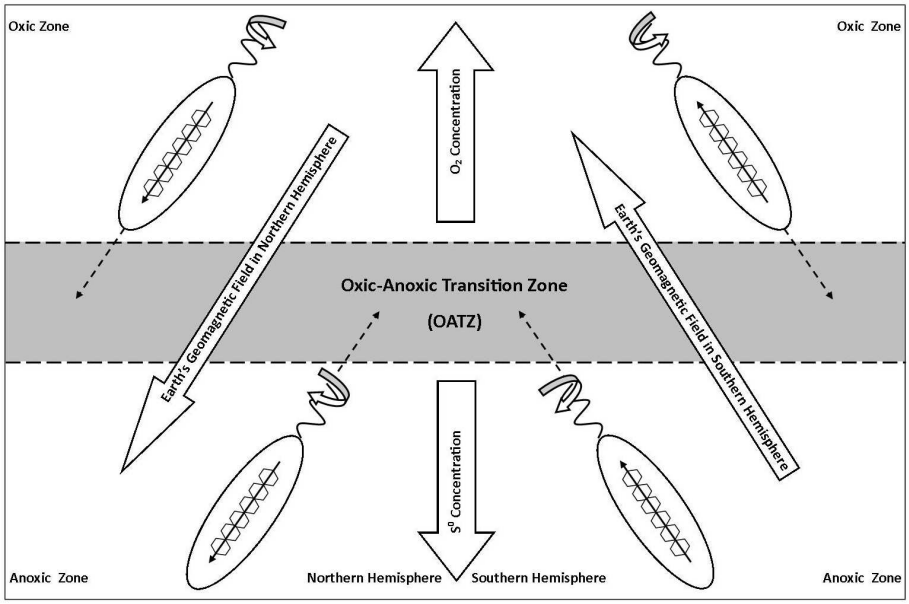

Figure 1.5: Illustration of MTB use of magnetotaxis to facilitate dwelling in the OATZ. Dashed arrows show the bacterium swimming directionality. Solid arrows show the MNP's (hexagons) alignment with the Earths geomagnetic field. The circular arrows show the rotation of the flagellum allows the bacterium to swim backward or forward within a water column (rotating clockwise to swim backward or rotating counter clockwise to swim forward)[52].

And all possible distortions and elongations of these[59].

The magnetosome is a distinct subcellular compartment exhibiting a unique biochemical composition. The structure of the magnetosome membrane (MM) is similar to that of eukaryotic organelles, being composed of a lipid bilayer. The MM represents a third membranous compartment alongside the outer membrane and cytoplasmic membrane[38, 60]. It is believed to originate from the cytoplasmic membrane by invagination. The optimal method of MM visualisation is transition-electron microscopy (TEM).

The MNPs have a crystal size of approximately 35-120 nm[61], a narrow size distribution with an average diameter of 50-100 nm[58], and are permanent single–magnetic domains (SD) at ambient temperature[62]. The crystals are uniformly magnetised within this size range, meaning that their magnetic dipole moment is at its maximum (equal to that of the saturation moment). The crystals smaller than SD at ambient temperature are superparamagnetic, with a remnant magnetisation coming close to zero. Whereas, larger crystals are metastable SD or support formation of domain walls making their magnetic moments smaller than that of SD crystals. Therefore, the SD crystals have the largest possible remnant magnetisation[8]. The individual grain moments align parallel to each other along the magnetosome chain direction, this process is caused by the magnetotactic

interactions between the SD magnetosomes. In doing so, a permanent magnetic dipole is created, roughly equal to the sum of the individual magnetosome magnetic moments[63]. Frankel[64] suggested that the orientation of the dipole results in the orientation of the cell as the dipole is in a fixed cell. The magnetosomes of magnetite are aligned along the {111} direction, parallel to the magnetosome chain, this is believed to be a biological mechanism for maximising the magnetic moment per particle, as this particular direction yields approximately 3% higher saturation magnetisation than the other directions[9, 65, 27]. Due to their unique characteristics, the identification of bacterially precipitated magnetite in the fossil record is possible. Fossilised MTB have even been found in deep sea sediment[66, 67] and Precambrian stromatolites[32].

Interest and research into MTBs since their discovery has been wide–ranging and scientifically interdisciplinary. Various commercial applications for bacterial MNPs have been proposed and fall into two categories: those involving the whole living cells and their magnetotactic behaviour; and those utilising isolated magnetosome particles, allowing for modification with organic molecules. Examples of the former are limited, but include the removal of radionuclides and heavy metals from waste water[68, 69], whilst usages of the latter are more widespread, with potential for use as contrast agents for MRI[70, 71, 72, 73, 74, 75], magnetic antibodies[76, 77], hyperthermia therapeutic agents[78, 79, 80], drug[75, 81, 82, 83, 84, 85] and gene delivery systems[86, 87, 88]. While the biotechnological potentials have been established, thus far such applications are not commercially viable at present. The main problems for commercial scale utilisation being; the expense of Magnetotactic bacteria mass cultivation, and the lack of essential understanding of the genetic and biochemical principles involved in the process of magnetite biomineralisation. In order to produce MNPs more economically the biomineralisation process needs to be further explored.

It has been hypothesised by many, including Arakaki *et al*[89]and Devouard[59], that MNP biomineralisation has three major stages. The first stage is the formation of intracellular vesicles resulting from the invagination of the cytoplasmic membrane. This process is believed to be primed by a membrane specific GTPase[90]. Secondly, ferric iron is transported into the vesicles and is reduced to ferrous iron on the cell surface before being taken into the cytoplasm. These are then transported into the vesicle, via transmembrane iron transporters, and subsequently oxidized to produce magnetite. Finally, the crystallisation of magnetite occurs in the vesicle. This process still remains unclear, however, it is suggested that tightly bound MNP proteins trigger magnetite crystal nucleation and/or

regulate the resulting morphology. It is believed that various MM proteins play functional roles in magnetite synthesis, including, the accumulation of the iron concentrations required for supersaturation, the maintenance of intracellular reductive conditions, the induction of mineralisation via the oxidation of iron, and the transformation of ferrihydrite to magnetite through its partial reduction and dehydration[89].

Cultivation of MTB in a laboratory has proven difficult to achieve, despite their high abundance. These difficulties arise from attempting to reconstruct their natural environment under laboratory conditions. The *Magnetospirilla* family was the first to be cultivated in a laboratory[40], with the first MTB isolated in pure culture being the *Magnetospirillum magnetotacticum* strain MS-1[40]. However, its usefulness in magnetosome formation molecular studies was limited; due to its fastidious nature and incapability for robust laboratory based growth. With the isolation of two other closely related species of *Magnetospirilla*; *Magnetospirillum grysphiswaldense* strain MSR-1 and *Magnetospirillum magneticum* strain AMB-1, this issue was bypassed, with these species subsequently taking centre stage as models for understanding magnetosome formation[91, 92, 60]. To date at least twenty species of magnetite producing MTB have been cultivated in pure culture. Over recent years, the genomes of several MTB have been sequenced; the first of which was produced for *Magnetospirillum magneticum* strain AMB-1[93]. This was undertaken to elucidate the mechanism of magnetosome formation and to provide a template for determining how MTB maintain a species–specific, nano–sized, magnetic SD morphology. Figure 1.6 shows TEM images of an example of *Magnetospirillum magneticum* strain AMB-1 and it MNPs.

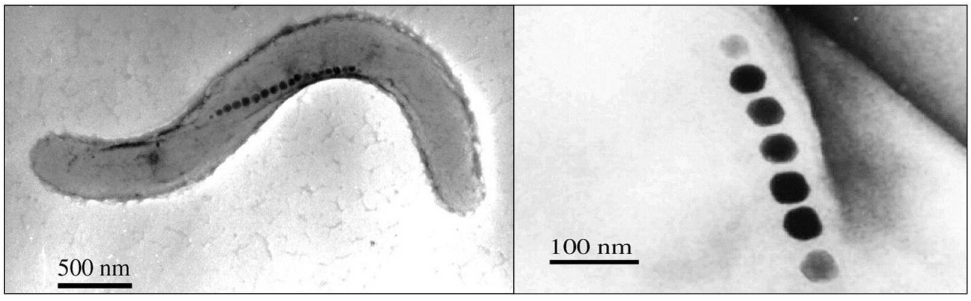

Figure 1.6: TEM images of *Magnetospirillum magneticum* strain AMB-1. A) Full bacterial specimen. B) MNPs. Adapted from Arakaki *et al* 2008[89]

Magnetospirillum magneticum strain AMB-1, a facultative anaerobic magnetotactic *spirilla*, is a Gram–negative, spiral shaped, aquatic, mesophillic bacterium, preferring temperatures between $25 - 40\,^{\circ}\mathrm{C}$. They synthesize MNPs that are cubo–octahedral in shape

consisting of both {100} and {111} crystal faces, elongated in the {111} direction (see Figure 1.7). The {111} axis corresponds to the magnetic easy axis, hence, having the lowest energy. This arises from anisotropy in the magnetocrystalline energy, the interaction of spin magnetic moments within the crystalline matrix, in magnetite this is above 120 K[94]. The {100} direction is the magnetic hard axis, thus, being higher energy. Matsunaga *et al*[41] found that these nitrate–reducing bacteria were capable of growing under both micro–aerobic and aerobic conditions in liquid or solid media, making this strain an ideal candidate for genetic manipulation[91].

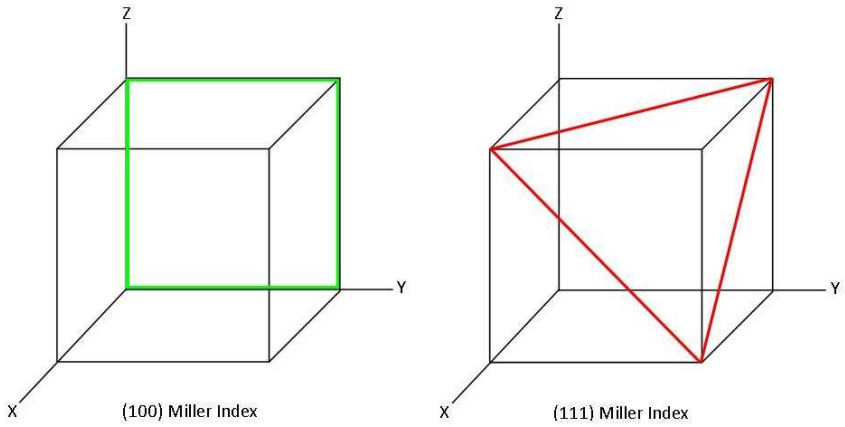

Figure 1.7: An illustrations of the {100} and {111} Miller indices.

As mentioned before elucidating the magnetosome formation mechanism is of great importance. Much attention has been focused on the magnetosome membrane proteins of *Magnetospirillum magneticum* strain AMB-1, in order to analyse this precisely regulated molecular mechanism[43, 38]. Tanaka *et al*[95] postulated that the magnetosome membrane was derived from the cytoplasmic membrane of the cell, this was backed up by the identification of 78 magnetosome membrane proteins also being found in the cytoplasmic membrane, several of which are believed to be related to magnetosome biosynthesis. Most magnetosome–specific proteins are shared by different *Magnetospirilla*[95] and are suggested to function in protein–protein interactions, vesicle formation, magnetosomal iron transport, magnetite crystallisation, intracellular arrangement of magnetite particles and activation of magnetosomes. Proteome analysis of the magnetosome membrane identified several novel proteins with potentially crucial roles in MNP biomineralisation elucidated by a number of molecular studies[90, 95, 96, 97, 98, 99, 100].

1.1.3 Magnetosome Membrane Protein Mms6

From the analysis of the magnetite crystal surface of the MNPs, Arakaki *et al*[100] identified four proteins, within *Magnetospirillum magneticum* strain AMB-1, which were tightly associated with the bacterial magnetite crystal surface; Mms5, Mms6, Mms7 and Mms13. These Mms- (Magnetosome particle Membrane–Specific) proteins have common amphiphilic features; with a hydrophobic N–terminal and a hydrophilic C–terminal but had no sequence similarities to known functional proteins, with the only observed similarity being to the MM proteins MamC and MamD from *Magnetospirillum grysphiswaldense*[101]. The N–terminal in Mms5, 6 and 7, exhibit a common leucine (L) and glycine (G) repetitive amino acid motif, which is reminiscent of self–aggregating proteins of other biomineralisation systems[100, 102], it is believed to do this via hydrophobic interactions. Within Mms6, the C–terminal is a highly acidic region, containing dense carboxyl and hydroxyl groups that show direct interaction with the magnetite surface. The carboxyl groups have a strong affinity for metal ions and bridging ligands, whilst the hydroxyl groups also have metal binding capabilities. Both groups are able to bind iron ions[100, 103].

Much of the studies into these novel proteins have been focused on the Mms6 protein. During in vitro magnetite synthesis via partial oxidation of ferrous hydroxide, Amemiya *et al* [104] found that in the presence of the recombinant Mms6 protein, uniform sized magnetite crystals were formed. These crystals had a cubo–octahedral morphology, consisting of {100} and {111} crystal faces. It was found that the crystals produced by Mms6–mediated synthesis were similar to the magnetite crystals isolated from *Magnetospirillum magneticum* strain AMB-1. The formation of {100} faces was considered to be due to Mms6 face–specific interaction, this being said, they did not achieve identification of the functional site of the molecule. In the absence of the protein, octahedral crystals with {111} faces were detected. Amemiya *et al*[104] postulated that the Mms6 protein could serve many key roles in magnetite biosynthesis such as; promoting the growth of a lower surface energy crystal lattice with the self–assembled aggregate acting as a scaffold, regulating crystal size, controlling the shape of the crystal via the face–specific interaction.

Arakaki *et al*[105] also focused their attention on the magnetosome membrane protein, Mms6, wanting to understand its role in magnetite biosynthesis. They investigated magnetite formation using synthetic peptides (shown in Table 1.4) that mimic the Mms6 protein, again utilising a partial oxidation method as represented in Equation 1.1. Arakaki *et al* focused primarily on the C–terminal acidic region and the hydrophobic GL repeat

region of the N–terminal.

$$\begin{aligned} 2Fe^{2+} + 4OH^- &\longrightarrow 2Fe(OH)_2 \\ 2Fe(OH)_2 + NO_3^- + H_2O &\longrightarrow 2Fe(OH)_3 + NO_2^- \\ Fe(OH)_2 + 2Fe(OH)_3 &\longrightarrow Fe_3O_4 + H_2O \end{aligned} \tag{1.1}$$

Synthetic Peptide	Amino Acid Sequence	kDa	pI
M6A	DIESAQSDEEVE	1.4	3.6
GLM6A	GLGLGLGLGLDIESAQSDEEVE	2.3	3.6
M6B	KIKSAQSKKKVK	1.4	10.7
GLM6B	GLGLGLGLGLKIKSAQSKKKVK	2.3	10.7

Table 1.4: Amino acid sequences of the synthetic peptides mimicking Mms6. kDa = Molecular weight and pI = Isoelectric point

Their findings showed that the crystals synthesized in the presence of peptides containing the C–terminal acidic region (M6A and GLM6A) were uniform–sized cubo–octahedral crystals with a narrow size distribution, similar to biogenic magnetite, conversely, crystals synthesized in the absence of peptides containing the C–terminal acidic region (M6B and GLM6B) were octahedral with sharp corners. Statistical evaluation of the size distribution and circularity was used to confirm the morphological characteristics. Despite these findings, the specific localisation of proteins or peptides on particular crystal faces was again not determined. Arakaki *et al* suggested that the results of this study indicated that the negatively charged C–terminal region is crucial for the formation of cubo–octahedral crystals. However, the observed size and morphology of the crystals synthesized in the presence of Mms6 and the other mimic peptides cannot be attributed to only the isoelectric points or the net negative charge of these molecules, thus, the conformation of the C–terminal region may contribute to controlling the morphology of the magnetite crystals.

Various magnetite synthesis methods have been explored ranging from dry hematite reduction[106, 107] to wet precipitate methods[108] including; simple co–precipitate[109], partial oxidation[110, 111], ferric hydrolysis of urea[112] and controlled colloidal methods[113]. However, the crystals formed are very much method–dependent. With the exception of the biogenic synthesis of magnetite and Mms6–mediated syntheses, the available literature

suggests that, the only other methods to produce cubo–octahedral magnetite involve very high temperatures and organic solvents[114, 115].

Many studies over recent years have attempted to demonstrate the in vitro roles of the Mms6 protein in the control of morphology and/or crystal size, via chemical synthesis using purified proteins or synthetic peptides. However, the morphological regulatory molecular mechanisms remained unclear due to a lack of direct evidence of protein–mediated in vivo control, until Tanaka *et al* in 2011[3]. By constructing and analysing an Mms6 gene deletion mutant strain of *Magnetospirillum magneticum* strain AMB-1, Tanaka *et al* endeavoured to understand this role. The gene deletion strain proved to synthesise magnetite crystals with lower aspect ratios, smaller size and indefinite morphologies with uncommon crystal faces, very unlike the uniform cubo–octahedral crystals displaying {100} and {111} crystal faces produced with the wild type. The faces present in the Mms6 gene deletion strain included the high index crystal faces {210}, {211} and {311}. Also, the higher energy, unstable {110} face, along with the {100} and {111} surfaces present in the wild type. The difference between the results in the absence of the Mms6 protein, for the in vivo work of Tanaka *et al* and in vitro work of Arakaki *et al*, was postulated to be due to the very different reaction conditions used in each study.

Tanaka *et al*[3] also found that gene deletion and complementation had no effect on bacterial growth. The magnetite crystals aligned in a chain, and the average number of magnetite crystals per cell was similar to that of wild type, suggesting that the structure of the vesicles were unaffected by the gene deletion, furthermore, that Mms6 doesn't function in the nucleation step of magnetite biomineralisation. Surprisingly, they discovered that the crystals found in the middle of the chain were elongated toward the {111} direction, as was seen with the wild type, proposing that the crystallographic direction remained unaffected by the deletion of Mms6. This finding suggests that the Mms6 protein regulates the orientation of the crystal growth, however, another protein is responsible for the induction of growth towards the {111} direction. In addition to the role Mms6 plays in the in vivo regulation of the crystal structure during biomineralisation, the results of Tanaka *et al* also suggest that the protein serves alternative functions, these include the surface stabilisation of the magnetite crystal growth, and acting as a scaffold to stabilise the protein complexes, in order to accurately localise them onto the magnetite crystal surfaces. Additionally, the homeostasis of other proteins (Mms5, 7 and 13) on the crystal surface was significantly reduced by the deletion of Mms6, proposing an association between these proteins and the potential of co-localisation on the magnetite crystal sur-

face in the form of protein complexes. These proteins, which are tightly associated with the magnetite crystal surface, are considered to have similar functions, and due to this cooperate in the formation of magnetite with a consistent crystal surface.

Despite the results produced from the work in this area, further work needs to be undertaken in order to elucidate the molecular mechanisms behind magnetite crystal formation and what role exactly Mms6 has in these. A technique that works successfully alongside experiment, that enables individual components of mechanisms to be considered in isolation, is computer modelling at the atomistic level. Atomistic simulations, which have not previously been utilised to establish the role of the Mms6 protein in the biomineralisation of magnetite, offer the advantage of direct observation of the atomic and molecular processes involved, this being the topic of this thesis.

1.2 Computational Studies

Despite the progress made experimentally in the field of biomineralisation, allowing for greater understanding of the roles that biomolecules, such as peptides and proteins, play in biomineralisation mechanisms, understanding of the biomolecule–mineral interactions during biomineralisation is at present far from complete. Computer simulations can be an effective tool to further this understanding, allowing for the attainment of atomistic details at the biomolecule–mineral surface interface. Increasingly, theoretical models are being used, with some success, for processes such as biomineralisation involving the interface between biomolecules and mineral surfaces. A range of computational techniques have been used to study the mechanisms of biomineralisation over the years, utilising theories ranging from quantum mechanics and atomistic simulations at the nanoscale to continuum mechanics and composite theories at the macroscale. Despite the great challenges of simulating the structure and properties of biomaterials, many attempts have been made to do so at atomic and meso length and timescales, which incorporate the range of physical effects important in the formation of biomaterials. These simulation studies generally fall into three groups of techniques;

Electronic Structure Calculations[116] This technique provides a description of the system–specific electron–electron interactions by numerical solutions of Schrödinger's equation. One of the most popular theories utilised in electronic structure calculations is the Density Functional Theory (DFT) which utilises code such as CASTEP[117],

VASP[118] and CRYSTAL[119]. To reduce the amount of numerical analysis involved in the calculation, approximations, such as Local Density Approximation (LDA) and Generalised Gradient Approximation (GGA), are introduced.

Atomistic Simulations[120] This technique involves the use of potential based parameterised functions to describe the short- and long–range interatomic interactions. This technique is generally applied in two ways, molecular dynamics simulations and Monte Carlo simulations. Molecular dynamics simulations[121] calculate atomic and molecular physical movements in the context of N–body simulation, whereby, Newton's equations of motion are numerically solved, in order to determine the trajectories of atoms and molecules for a system of interacting particles. A large number of general purpose codes are available for this technique, such as DL_POLY[122] and LAMMPS[123]. There are also more specific biomolecular codes such as AMBER[124] and NAMD[125]. Monte Carlo[126] is a broad class of computational algorithms, for simulating the behaviour of various physical and mathematical systems, that rely on repeated random sampling. This method is stochastic, using algorithms that are non–deterministic in some manner and using random number or pseudo–random number generators. Monte Carlo codes are written by individual groups for specific applications.

Mesoscale Simulations Mesoscale refers to the scale between atomistic and macroscopic. These methods apply molecular dynamics or Monte Carlo techniques in domains too large for individual molecules and atoms to be included explicitly, instead using "pseudo–atoms" to represent groups of atoms. These methods can tackle the problems of large system size and long timescales by using reduced representations, which coarse–grained techniques are examples of. Coarse–graining[127, 128] consists of replacing an atomistic description of biological molecules, such as proteins, with a lower–resolution coarse–grained model that averages away fine details, for example degrees of freedom.

There are problems with all theoretical approaches to biomineralisation processes, for example, the restriction of system sizes to mesoscale and below necessitate consideration of a wide range of techniques, from ab initio methods to coarse–grained simulation methods. However, the main challenge posed for biomineral system simulation is that of the interfaces between soft and hard matter of the biominerals, as they are often the boundaries between different model types. These components of the biomineralisation system have been, up until recently, considered as two very separate theoretical fields, organic and

inorganic simulation. Much has been published on the individual elements, with both fields having reliable potential models which have been utilised to tackle a wide range of problems. However, very rarely are the fields approached together as combining the two poses immense challenges. Reviews have been published giving excellent discussions of computational techniques for modelling organic–inorganic interfaces; however, much of the information provided favours either hard matter simulation or soft matter simulation dependent on the author[129, 130]. One possible method for combining the two components is by defining cross–term potentials, for example, by using the Lorentz–Berthelot mixing rules. However, from previous studies[131, 132], this method has been found to seriously overestimate interfacial binding energies. This issue could be resolved by systematically refitting the potential sets, however, this approach is extremely computationally expensive. An alternative approach, which has been used by de Leeuw *et al*[133] and Freeman *et al*[134], is the Schröder method[135], which recasts the interatomic cross–term potentials to fit the different charges between the systems.

1.2.1 Mineral-Water Interface

Another problem facing biomaterial simulators is that the systems studied are, by nature, aqueous, which means that the chemistry of the mineral–water interface must also be considered alongside the detailed structural effects the presence of water will have on the biomolecule–mineral surface interactions, making biomineralisation systems highly complex. The behaviour of water within a system is a crucial factor to consider as it may be an important component of biomolecule adsorption at mineral surfaces in aqueous conditions. Many studies[132] have confirmed this importance, suggesting that water may mediate the biomolecule–mineral surface interactions, due to the reports that a solvent monolayer persists between the biomolecule and surface. There are, however, major challenges to face when simulating water[136] and this in itself has resulted in an exhaustive literature leading to the existence of many force fields and models to describe it[137].

Calcium based biominerals have been the predominant focus for biomineral–water interface studies. Calcite surfaces in aqueous conditions have been studied extensively and the results from computer simulations[138, 139] and experiments[140] agree well, with both showing clear evidence of an ordered layering structure to the water near the surface and that the details of the water structure vary with the different surface structures of the calcite[141]. It was also found that the free energy of adsorption was greatly influenced by the layering of the water at the interface. In one study, density functional theory

calculations were employed to explore the interaction between associated and dissociated water molecules and the {10.4} surface of calcite. It was revealed, through static relaxations, that the dissociated water molecules were energetically unfavourable[142]. Again, with this study, the presence of well–defined water layers was highlighted, with the water molecules forming three distinct layers, whereby, the molecules nearest to the surface lay flat on the surface, and the water molecules slightly further out split between interacting with the surface and interacting with the other water molecules present.

In 2003, Kerisit *et al* focused their research on the dissociative adsorption of water on the {10.4} calcite surface. It was found, by the use of atomistic simulation methods, that when water adsorbs onto the surface it loses its hydrogen–bond network, leading to an oscillation of water density in close proximity to the surface. The simulation results also suggested that, the carbonate groups present at some step edges and low–index surfaces form OH groups on the surface by dissociating water and subsequently release carbon dioxide[138]. Most recently, studies into the calcite–water interface are interested in direct comparison between simulation and experimental data in an attempt to ensure the accuracy of the simulated models. Fenter *et al*[143] compared four different modern force fields models(two non–polarizable, one polarizable, and one reactive force field) utilising classical molecular dynamics simulations with high–resolution specular X-ray reflectivity (XR) data in order to give further understanding of the interfacial structure. The comparative results produced suggested that the interfacial structures created through simulation, when taken as a whole, were not in agreement with the precision and accuracy of the XR data, however, there was a higher level of consistency with the XR data for the simulated interfacial water profiles, with the rigid–ion model results showing semi–quantitative agreement.

Hydroxyapatite is another biomineral of importance in mineralwater interface research. Pan *et al*[144], using molecular dynamics simulations, studied the behaviour of water on the {001} and {100} hydroxyapatite surfaces. The results again highlight the presence of an ordered layering structure to the water and that this layered structure varies dependent on surface, with the {001} direction resulting in more structured water layers than the {100} direction. It was also found, when observing the interaction of the water molecules with the calcium and phosphate sites at the interface between the hydroxyapatite surfaces and water, that there are multiple pathways of water adsorption onto the surfaces, with the first layer of water exhibiting specific adsorption sites. The group postulated that the water layers could block ion migration, which could significantly impact hydroxyapatite crystal growth and dissolution. Prior to this, it was largely assumed that biomolecule

adsorption onto the surfaces of hydroxyapatite was related to calcium ion and phosphate ion site distribution on the surface, whilst ignoring the water layer functions[145, 146].

In 2009, Corno *et al*[147] investigated water adsorption on the {001} and {010} hydroxyapatite surfaces using the quantum mechanical code CRYSTAL at the B3LYP level of theory, finding that there was a great affinity for water for both surfaces. Water adsorbed molecularly on the {001} surface, while it dissociated on the {010} surface, which lead to new surface terminations. A reasonable agreement in results was shown on comparison with experimental water adsorption enthalpies. Among the most recent studies into the hydroxyapatite–water interface, the surface stability of the {001} hydroxyapatite surface and the phosphate–exposed, calcium–exposed, and hydroxyl–exposed terminations of the {010} hydroxyapatite surface were investigated, with particular interest into surface resistance to hydrolytic remodelling[148]. It was established that dissociative water adsorption was beneficial for the {001} surface and the phosphate–exposed {010} surface, whilst, only molecular adsorption of water was feasible for the hydroxyl–terminated {010} surface. The study also showed that the hydroxyl–terminated (010) surface was the most stable surface and hence, in experiments, should be the predominant surface exposed.

Kerisit in 2011[149] exhibited, through molecular dynamics simulations utilising four different models (based on interatomic potentials), the atomic–level structure of three hematite-water interfaces({001},{012} and {110}). It was found, with the exception of one termination with one model, that all of the models tested predicted the correct interfacial water molecule arrangement, and that there was good to excellent experimental agreement for water positions, distances, and layer occupancies. The study also showed that, at all surfaces, each of the three surface functional groups present (triply–coordinated oxo, doubly–coordinated hydroxo, and singly–coordinated aquo groups) formed similar hydrogen bond configurations with the adsorbed water molecules. Kerisit postulated that the structure of water at the interface with the mineral surface plays a role in the thermodynamic and kinetics of adsorption[150].

Great progress had been made, working up to this study, in the field of hematite–water interfaces, moving from modelling a simple monolayer of water molecules, to discovering that the major mineral–water interactions are through water oxygen ions with surface iron ions, followed by hydrogen–bonding to surface oxygen ions[151, 152, 153, 154], to more advanced potential model simulations of hydrated hematite surfaces in contact with water layers tens of angstroms thick[155, 156, 157, 158], with much focus on the {001} and {012} surfaces. The earlier work of Kerisit and colleagues[159] showed that, for the {001}

hematite surface, there were water layers at two distinct distances above the surface, and that in the first layer water molecules bonded to the surface hydroxyl groups, whereas, for the second layer the water molecules interact with the surface hydroxyls and the iron ions.

Recent research into goethite[160], another iron oxide biomineral, investigates the proton affinity implications of water structure and hydrogen bonding at goethite–water interfaces. Boily[160] utilised molecular dynamics simulation to explore the {010}, {100}, {110}, and {021} surfaces of goethite in an aqueous environment, in order to determine the structures of interfacial water in close proximity to the (hydr)oxo groups on the goethite surface. The calculations showed that water adopted highly surface–specific configurations, and that, under normal environmental conditions, singly coordinated surface groups are predominantly linked to charge uptake. Yet, revised proton affinity constants showed that, in the presence of strongly binding negatively charged ligands, protonation of doubly–coordinated hydroxo groups and one type of triply–coordinated oxo group may be preferred.

1.2.2 Mineral-Organic Molecule Interface

Research into mineral–organic molecule interfaces is of great interest for furthering understanding of the mechanisms and processes of biomineralisation. It is also of great importance to consider and understand the critical role of structured water layers on mineral–organic molecule interface behaviour. Yu *et al*[161] highlighted how organic molecules react at a mineral surface differed significantly dependent on a presence or absence of water. Their study published in 2003 showed that when the system was in vacu, the molecule maximised its contact area with the mineral surface. Whereas, when the system was solvated, molecular binding to the surface often occurred through only a single functional group, with the remainder of the organic molecules structure being immersed in the water, and, in some instances, direct binding between the organic molecule and the mineral surface was not observed at all, due to a lack of ability for the molecule to displace the strongly bound innermost water layer. In this case, the organic molecule either binds to the mineral through this water layer, disrupting only the second and subsequent layers (as shown experimentally by Magdans *et al*[162]), or it just fails to bind. Several other studies since have also noted these behaviours.

Most recently, Zhu *et al*[163] demonstrated that the particular surface of attachment effects the organic–inorganic interactions along with the significant contribution that water layers

have to interfacial interactions, through a study comparing adsorption of an acrylic acid dimer on the {10.4} calcite surface to its adsorption on the {11.0} calcite surface. The results, which are consistent with experimental observations, showed that on the {10.4} surface two carboxyl groups interact with the crystal but the molecule has to compete with water due to the well–structured hydration layer, whereas, on the {11.0} surface only one carboxyl group interacts, however, the water layer was relatively loose so that the molecule can easily replace water, leading to a stronger interaction with {11.0} surface than with {10.4} surface.

There has been a large body of work on the calcite–organic molecule interfaces, with studies into many different interfacial behaviours including: polysaccharide adsorption behaviours with stepped calcite surfaces, demonstrating that adsorption onto acute–stepped surfaces was more favourable than onto obtusestepped, and the potential usage of monosaccharides to inhibit crystal growth through selective surface coverage[164]. These studies illustrated how biomolecule adsorption can be greatly affected by peptide configuration, with studies such as that of Yang *et al*[165], whereby peptide chain interactions with the calcite {10.4} surface were investigated, revealing that a β-turn peptide configuration has the strongest interaction with calcite surface, while the α-helix configuration exhibits the weakest interaction.

For Freeman *et al*, the question of what role the chicken eggshell protein Ovocleidin-17 (OC-17) (a protein found only in hen ovaries) had in eggshell formation led to the first molecular simulation of spontaneous crystallisation of amorphous calcium carbonate (ACC) in the presence of OC-17[166]. Using metadynamic simulations, Freeman *et al* explored how the onset of mineral crystallisation can be controlled and what role this native protein had in the process, through the simulation of OC-17 adsorption onto ACC in an aqueous environment. It was found that OC-17 bound most readily to the nanoparticles through two clusters of arginine (ARG) residues, and that OC-17 acts as a catalyst in eggshell formation by aiding the transformation of ACC particles into calcite crystals. Freeman *et al* then progressed onto exploring the adsorption of OC-17, in several different configurations, onto the calcite {10.4} surface, using classical molecular dynamics simulations[167]. This study again demonstrates competitive binding between biomolecules and strongly bound surface water, and that arginine residues are the major binding sites of OC-17. The simulation data produced shows that, firstly, the protein configuration most energetically favourable for binding in fact has the fewest residues–surface contact points, as this causes the least disruption of the strongly bound water

layer through water molecule displacement. Secondly they exhibited that ARG residues achieve strong interactions with the surface carbonate–oxygen and calcium ions due to the ability of its long side chain group, to penetrate the water layer with relative ease.

In a paper published in 2012, Freeman *et al*, went on to study OC-17 binding on stepped calcite surfaces (the vicinal {31.16} and {31.8} surfaces), again utilising classical molecular dynamics simulations[168]. The simulations demonstrate that binding was facilitated by the surface water organisation on stepped surfaces, with binding occurring in areas of least water density, and that binding was greater at the obtuse step. The data also suggests that it was unlikely that OC-17 was important in controlling crystal morphology, due to the vicinal surfaces exhibited weaker binding than the planar {10.4} surface, suggesting that OC-17 may not prevent crystal growth in particular directions by binding to steps, which was a finding that agreed with previous crystallisation simulations[166]. Freeman *et al* also found that the structure of OC-17 remains unchanged when it was in contact with the calcite surface and that binding to the surface was usually through basic amino acid residues, in particular ARG.

It is known that hydroxyapatite is grown as nano–sized mineral platelets at nucleation sites on a collagen protein template, however, the exact role of the collagen matrix during biomineralisation remains unknown, although it is postulated that collagen functions to actively control mineralization through growth direction effects. Knowledge of the collagen–hydroxyapatite interface is required to understand the nucleation and directed growth of hydroxyapatite at the collagen matrix. This problem is the focus of much research into hydroxyapatite biomineralisation and the interface between hydroxyapatite and biomolecules. The {00.1} and {01.0} hydroxyapatite surfaces have predominantly been the crystallographic faces considered in relation to simulation studies of the hydroxyapatite–collagen interface, as the {00.1} surface is the most thermodynamically stable hydroxyapatite surface[169]. The {00.1} surface is also the direction in which the apatite mineral is aligned along the collagen fibril. The {01.0} surface, however, has been found to be the dominant surface in biological material morphology[170]. Hydroxyapatite–collagen studies have moved from constituent collagen matrix amino acid attachment[171] to collagen peptide attachment[172, 173], leading through to the simulation of early stage hydroxyapatite nucleation at a collagen template[174], utilising density functional theory techniques and atomistic simulations. In a collagen molecule, approximately one third of the amino acid residue sequence is made up of glycine (GLY) (one GLY residue in every third position), whilst a further quarter consists of proline (PRO) and hydroxyproline (HYP)[175]. Lysine

(LYS)and hydroxylysine (HYL) residues are also present, hence these residues have been the focus of constituent amino acid and peptide studies.

The constituent amino acid density functional theory study by Almora-Barrios and de Leeuw[171] in 2009, showed that GLY, PRO and HYP are able to form multiple surface interactions with the hydroxyapatite, and the strongest amino acid–surface binding was through HYP onto the {01.0} hydroxyapatite surface. When collagen peptides(amphiphilic - PRO-HYP-GLY and HYP-PRO-GLY, hydrophobic - PRO-LYS-GLY and PRO-HYL-GLY) were investigated[172, 173] the results were in agreement with that of experiment and previous simulation studies[171, 176, 177, 178, 179, 180, 181], in that, multiple surface interactions were formed between the peptides and the hydroxyapatite surface species, and as with the collagen amino acid study[171], the peptides more strongly bound to the less stable {01.0} surface than to the thermodynamically stable {00.1} surface. Thus surface stability and surface geometry (e.g. bridging between two or more calcium ions) have a significant effect on the interaction strength of the peptides. Of the collagen peptides, the positively charged NH^{3+} side groups (LYS and HYL) and the OH side groups (HYL and HYP) principally were found to bind strongly to the surface, leading to Almora-Barrios and de Leeuw[172] to propose that the presence of these residues in the collagen protein should promote hydroxyapatite growth in biomineralisation and synthetic mineralisation. When water was present in the system there was competition for adsorption sites, in some cases water molecules can block peptide adsorption at the preferred surface site, however, the presence of water does not change the surface that the peptide most strongly binds to, as only the {01.0} surface interacts strongly with peptides, further helping to elucidate the reason behind the biological hydroxyapatite morphology where, as a result of the collagen matrix growth–directing effect, this surface was expressed preferentially.

In 2012, Almora-Barrios and de Leeuw attempted to simulate early stage hydroxyapatite nucleation[174] by using molecular dynamic simulations in which a collagen template was immersed in a stoichiometric solution of Ca^{2+}, PO_4^{3-} and OH^- ions. Calcium phosphate clusters were found to form at the collagen template due to electrostatic attraction between the calcium ions and the phosphate ions. The calcium ions also reacted electrostatically predominantly with the oxygen atoms of the GLY and HYP residues of the collagen template and some of the phosphate ions interacted with the hydroxy groups of the HYP residues through hydrogen bonding, although the interaction of the phosphate ions with the collagen template was weaker than that of the interaction of the calcium ions with

the collagen template. Due to the short timescale possible with this type of simulation it was too early in the clustering and nucleation process to be able to differentiate between distinct surfaces of hydroxyapatite. In another hydroxyapatite study, Rimola *et al*[182], showed that although adsorption is explained by electrostatic interactions, the main interaction species are dependent on peptide configuration, exhibiting that if the GLU/LYS mutated 12-gly-polypeptide was in a α-helix folded state then the acidic and basic residues interact mostly with the Ca^{2+} surface ion, however, if the peptide was in a coiled state then the interactions with the Ca^{2+} surface ions were mainly through the backbone COO. It was also shown that this trend was not modified when water molecules were added at the adsorption points.

In other examples of calcium based biomineral interfacial interactions, the importance of phosphorylation in mediating interactions between calcium oxalate monohydrate (COM) surfaces and the extracellular structural bone protein osteopontin (OPN) was investigated. The study, by Hug *et al*[183], compared the interaction of aspartic acid (ASP), also a dipeptide formed of ASP and phosphoserine (pSER) with COM, establishing that although ASP formed close contacts with the COM surface, through the α-carboxyl and amine groups, they were only temporary contacts unlike the contacts that the dipeptide made through the carboxyl groups which were permanent, showing that for OPN–COM interactions, carboxyl groups are crucial. The results confirmed the work of earlier OPN–related studies (both experimental and simulation)[184, 185, 186], that the presence of a phosphate group in a peptide or protein sequence has a strong effect on the attachment of carboxyl groups with the COM surface. The behaviour of octacalcium phosphate (OCP), the precursor to hydroxyapatite, at an interface with proteins has also been studied. One such study[187], investigated the adsorption of lysozyme (LSZ) and human serum albumin (HSA) onto different OCP surfaces, finding that the adsorption energy of LSZ was higher than that of HSA, which suggested that LSZ attachment to OCP surfaces would be most favourable, agreeing with experimental work carried out by Zhu *et al*[188]. It was also observed from this investigation that interaction energies of the proteins differed with a change in OCP surface ($\{001\}$ $\{111\}$ $\{110\}$ $\{100\}$); however, the trend in changes was similar for both LSZ and HSA. The surface energy trend also matched that of the interaction energy. The adsorption differences of the two proteins that were observed were proposed to be due acidic/basic residue ratio differences on the protein backbone.

1.2.3 Magnetite Computational Studies

Despite the fact that, of the iron oxides, magnetite has been one of the most studied experimentally, with an increasingly large collection of research data being compiled, computational publications remain limited in comparison to other biominerals. Mazo-Zuluaga *et al*[189, 190] focused their research into magnetite on magnetism, in particular, on elucidating the critical magnetic behaviour of both bulk magnetite and associated with the transition from ferrimagnetism to paramagnetism, and the effect that having different exchange integrals has upon this behaviour. Whilst, Rustad *et al*[191] used molecular dynamics to investigate magnetite {001} surface reconstruction, as this particular surface has a half–occupied tetrahedral layer termination. They proposed a novel mode of reconstruction, whereby, the Fe^{2+} ions in the top monolayer move downward to occupy a vacant half–octahedral site in the plane of the second–layer iron ions, whilst, half of the tetrahedral iron ions in the third iron layer are move upward to occupy an adjacent octahedral vacancy at the level of the second–layer iron ions. The proposed reconstruction was found to be in agreement with experimental data. Soontrapa and Chen[192] used molecular dynamics to model magnetite formation via the oxidation of iron. In doing so they utilised appropriate pair potentials based on energy minimization charge distributions. The publication introduced a novel model for studying magnetite formation with results in satisfactory agreement with experiment, particularly with regard to oxide layer thickness.

To date, very little research has explored magnetite interfaces. The earliest work to be published specifically on the magnetite–water interface was produced by Rustad *et al*[193] in 2003, concentrating on the {001} magnetite surface in the presence of pure water and in a 2.3 M $NaClO_4$ solution. The investigation established that the electrolyte presence in the simulation system had negligible effect on surface functional group protonation states, and that they were preferentially arranged within the defined water molecule layers. The results also showed that extensive hydrolysis of the interfacial water molecules had a significant effect on the protonation states of the surface species. Subsequently, Kundu *et al*[194] performed static energy minimisation techniques to establish the behaviour of five different magnetite surfaces ({001}, {011}, {101}, {110} and {111}) in the presence of molecular and dissociated water and formic acid, focusing on surface structure and adsorption behaviours of water and carboxylic groups on pure surfaces. Kundu *et al* postulated that the magnetite surfaces would be hydroxylated in nature due to the very highly negative reaction energy produced. It was also observed that, with the exception

of the {011} surface, water adsorption is preferred over formic acid adsorption on hydroxylated magnetite surfaces. Research into magnetite–water interfaces has also looked into not only liquid water but water vapour. Tombacz *et al*[195] revealed that several well–defined layers are formed by water physisorbed onto the magnetite surface; and these layers are occur simultaneously, resulting in the adsorption layer's external surface being somewhat irregular.

In 2008 Grillo *et al*[196] produced a DFT+U study investigating the adsorption of water onto the tetrahedral iron–terminated {111} surface of magnetite, establishing that initial water adsorption was dissociative with the iron sites being occupied by OH groups, while the hydrogen atoms bind to surface oxygen, and subsequent water adsorption was molecular, leading to water bridging the OH and H groups forming a hydronium–ion–like structure. Most recently in magnetite–water interface research, the magnetite {110} surface was explored[197], looking at two different surface terminations that were close in surface energy, A and B, where the A termination contained two surface octahedral iron cations and four surface oxygen anions, whilst the B termination contained two surface octahedral iron cations, two surface tetrahedral iron cations and four surface oxygen anions. It was found that for both terminations, adsorption energies were as a result of not only the adsorption of water to the iron ions but also the hydrogen bonding between the water molecules and to the surface oxygen atoms. It was also established that, for the A termination {110} surface, with just one and two water molecules, molecular adsorption of water is favoured, whereas, for three and four water molecules, a combination of molecular and dissociative adsorption is ideal, however, on the B termination, dissociative adsorption of water is favoured for one and two H_2O molecules, whilst, again a combination adsorption modes, also, full dissociative adsorption were possible, for three and four molecules of water. This suggests that water will adsorb in different modes dependent on termination type as well as on surface type.

Very few publications exist that focus on simulation of the magnetite–biomolecule interface, and only one in the context of biomineralisation. The first study was from 2011, concentrating gold/magnetite nanocomposites[198], in particular, the role of an intermediate layer during gold nanoparticle deposition onto the {111} magnetite surface, finding that gold–magnetite interactions could be enhanced by the use of a suitable surfactant, polymer or SiOH terminated amorphous polymer to modify the magnetite surface. The magnetite–intermediate layer interaction strength was found to be dependent on the functional groups present in the intermediate layer, with amine functionalised molecules (e.g.

polyethyleneimines) being preferable. When the interactions of these composites with amino acids were examined it was established that there was a reasonably strong absorption of cysteine, which was believed to be due to the strong sulphur–gold bond formed, reducing molecular diffusivity.

To the best of the author's knowledge, the only publication that explores magnetite–biomolecule interfacial interactions with regard to biomineralisation was produced by Bürger *et al*[199] in 2013. This report focuses on amino acids linked to the magnetotactic bacteria *Magnetospirillum gryphiswaldense* magnetosome membrane proteins, MamJ and MamG, which are dominated by aspartic acid(ASP) and glutamic acid (GLU) in MamJ[200], and leucine (LEU) and glycine (GLY) in MamG[102]. Bürger *et al* utilised force field simulations to dock ASP, GLU, LEU and GLY onto the {111} magnetite surface, finding that electrostatic interactions dominate amino acid physisorption, with the involved species being the Fe^{3+} surface ions and the oxygen atoms of the amino acid carboxyl and carboxylate groups. GLY and LEU interacted with the magnetite through a bridging binding mode to the carboxylate oxygen atom, ASP did so through a bidentate binding mode to the carboxyl oxygen atom, whilst GLU exhibited both binding modes.

1.3 Conclusion

There have obviously been many advances in studying the mechanisms and processes of biomineralisation of magnetite and certainly of the biomineralisation of other biominerals, however, there is a long way to go before we fully understand what is occurring at the magnetite–biomolecule interface, particularly in the field of computational simulation. Key questions remain unanswered such as: how does the magnetite surface structure influence which biomolecule will attach; how does the presence of water influence whether or not a biomolecule will attach to the magnetite surface; what is the mechanism of biomolecule adsorption onto the magnetite surface; can reliable potentials be derived to describe the interactions between biomolecules, magnetite surfaces and water, and ultimately, how does the attachment of the biomolecule influence the growth of the crystal? This thesis attempts to explore these issues, however, before presenting the findings of this work on magnetite–biomolecule interfaces, the potential model and theoretical methods utilised to simulate the systems explored are first described.

Chapter 2

The Potential Model

In order to gain reliable simulation results, knowledge of the interatomic interactions within a system is required. Ideally this could be achieved using *ab initio* techniques, with the potential for giving an exceptionally accurate description of numerous essential processes within the system. However, this is very computationally expensive, and even with the use of modern petascale computers this type of technique is not feasible to simulate systems containing more than a few hundred atoms, making the technique inappropriate for the time and length scales of most real systems. This is particularly true when studying an interface between two phases, as in this work.

The use of parameterised potentials, to evaluate interatomic forces, provides a faster, molecular mechanics, alternative method, known as atomistic simulation, which enables the modelling of larger systems; however, the simulations require careful parameterisation, in order to ensure that the interatomic interactions remain accurate with changing geometry. Such a model was used for the work described in this thesis.

The basic requirement of a potential model is to describe the total energy of the system as a sum over intermolecular and intramolecular interactions, i.e.

$$U_{system} = U_{inter} + U_{intra} \tag{2.1}$$

The intramolecular interactions can be further divided into the sum of the energy required to stretch, bend or twist a number of adjacent atoms or bonds;

$$U_{intra} = U_{bonds} + U_{angles} + U_{torsions} \tag{2.2}$$

Each of which is defined by a series of parameterised equations. Similarly, the intermolecular interactions can be subdivided into coulombic and van der Waals interactions.

$$U_{inter} = U_{coul} + U_{vdw} \tag{2.3}$$

In an ionic solid such as magnetite the intermolecular interactions dominate to the point where the intramolecular terms can be neglected. This is sometimes described as the Born Model of Solids and as the theory equally applies to the intermolecular interactions in the amino acid residues and peptides it is the starting point for our discussions.

2.1 Intermolecular Interactions

Ionic or polar solid interactions are described using the Born Model of Solids[201], and a similar model is used to describe intermolecular interactions in molecular systems. In both models point charge particles are used to represent the atoms of a system that interact by means of short range interactions and long range electrostatic interactions. The interaction energy between two ions or atoms is obtained by

$$U_{inter} = \frac{1}{2}\sum_{i}^{N}\sum_{\substack{j \\ j\neq i}}^{N}\frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}} + \frac{1}{2}\sum_{i}^{N}\sum_{\substack{j \\ j\neq i}}^{N}\Phi_{ij}(r_{ij}) \tag{2.4}$$

Long-range Coulombic interactions are expressed in the first term, ε_0 is the permitivity of vacuum, q_i and q_j are the ionic or atomic charges and r_{ij} is the distance between i and j (or interatomic distance). The short-range interactions between the particles are described in the second term ($\Phi(r_{ij})$). Such short-range interactions include van der Waals attractive forces and the repulsion between the electron charge clouds.

It is not possible to obtain the Coulombic energy of the system by simply summing over all pairwise electrostatic interactions, as shown in Equation 2.4. The reason for this is that the contribution of the point-charges to the electrostatic potential decays as $1/r$, this causes very poor convergence of the Coulombic term. To counteract this, summation methods, are used to calculate the electrostatic interactions. The other terms of Equation 2.4, the short-range interactions, can usually be calculated by a simple summation as they converge much faster.

2.1.1 Long-range Intermolecular Interactions

As mentioned previously, the long-range electrostatic interactions are Coulombic. A particle i will interact with all other particles j in the simulation box. It will also, interact with these particle's periodic images, including those of i. This leads to the Coulombic contribution, U_{COUL}, to the interaction energy being represented by

$$U_{COUL} = \frac{1}{2}\left(\frac{1}{4\pi\varepsilon_0}\right) \sum_{n}{}' \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{q_i q_j}{|\mathbf{r}_{ij} + n\mathbf{L}|} \tag{2.5}$$

where q_i and q_j represent the charges on particles i and j, r_{ij} is the interatomic distance, ε_0 is the permitivity of free space, and $\mathbf{L}$ is the set of simulation cell vectors reflecting the periodicity of the simulation box. The sum of all the periodic images is expressed as the sum over n, where n is the ordered triple of integers that define the periodic images of the simulation cell. The prime on the first summation indicates that $i = j$ is ignored for $n = 0$.

However, as mentioned previously, due to the fact that the sum in Equation 2.5 is very slow to converge as a result of the $1/r$ term, a quicker and more reliable summation scheme, such as the Ewald method[202], is required.

Ewald Summation

Within the Ewald method[202] every particle i, of charge q_i, is assumed to be surrounded by a diffuse charge distribution of opposite sign. The total charge of the cloud is such that it exactly cancels q_i, meaning that the electrostatic potential, due to particle i, is due exclusively to the fraction of q_i not screened. This is usually taken to be a Gaussian distribution such as

$$\rho_i(r) = -q_i\left(\frac{\alpha}{\pi}\right)^{3/2} \exp(-\alpha r^2) \tag{2.6}$$

α, the arbitrary parameter, determines the distribution width, and r is the position relative to the centre of the distribution. Because of this, only the fraction of q_i not screened contributes to the electrostatic potential due to the particle i. This fraction rapidly converges to zero at long distances, making the screened interactions short-ranged. Hence, direct summation in real space can now be used to calculate the electrostatic interactions

between these screened charges.

The total contribution to the interaction energy of the screened Coulombic interactions, U_{REAL}, is then expressed as

$$U_{REAL} = \frac{1}{2}\left(\frac{1}{4\pi\varepsilon_0}\right) \sum_{n}{}' \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{q_i q_j}{|\mathbf{r}_{ij} + n\mathbf{L}|} \mathrm{erfc}(\sqrt{\alpha}|\mathbf{r}_{ij} + n\mathbf{L}|) \tag{2.7}$$

where $\mathrm{erfc}(x)$ is the complementary error function, shown as

$$\mathrm{erfc}(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)\mathrm{d}t \tag{2.8}$$

This tends to zero with increasing x. Hence, the only term that contributes to the sum in real space is that with $n = 0$, if α is large enough.

However, a new charge distribution must be added to each particle i in order to cancel the effect of the first charge distribution, as the electrostatic potential due to interacting screened charges is not the quantity of interest. This new compensating charge distribution has the same sign as the original charge q_i and the same shape as the distribution $\rho_i(r)$. By summing their Fourier transforms in reciprocal space, the contribution of this set of charge distributions to the electrostatic potential can be calculated, as the Gaussian distributions are a set of smoothly changing periodic functions.

Poisson's equation denotes the electrostatic potential at a point r_i, due to a charge distribution $\rho_1(r)$, that consists of a periodic sum of Gaussians. This is represented by

$$-\nabla^2 \phi_1(r) = \frac{1}{\varepsilon_0} \rho_1(r) \tag{2.9}$$

where

$$\rho_1(r) = \sum_{n} \sum_{j=1}^{N} q_j \left(\frac{\alpha}{\pi}\right)^{3/2} \exp\left[-\alpha|\mathbf{r} - (\mathbf{r}_j + n\mathbf{L})|^2\right] \tag{2.10}$$

In Fourier form, Poisson's equation becomes

$$k^2 \phi_1(k) = \frac{1}{\varepsilon_0} \rho_1(k) \tag{2.11}$$

Fourier transforming the charge density ρ_1 gives

$$\rho_1(\mathbf{k}) = \frac{1}{V}\int_V \rho_1(\mathbf{r})\exp(-i\mathbf{k}\cdot\mathbf{r})\mathrm{d}\mathbf{r} \tag{2.12}$$

If $\rho_1(r)$ is replaced by its expression as defined in Equation 2.10, then Equation 2.12 becomes (see reference [203] for derivation)

$$\rho_1(\mathbf{k}) = \frac{1}{V}\sum_{j=1}^{N} q_j \exp(-i\mathbf{k}\cdot\mathbf{r}_j)\exp(-k^2/4\alpha) \tag{2.13}$$

Combining Equations 2.11 and 2.13 yields

$$\phi_1(k) = \frac{1}{k^2\varepsilon_0}\frac{1}{V}\sum_{j=1}^{N} q_j \exp(-i\mathbf{k}\cdot\mathbf{r}_j)\exp(-k^2/4\alpha) \tag{2.14}$$

The electrostatic potential in real space is

$$\begin{aligned}\phi_1(r) &= \sum_{k\neq 0}\phi_1(\mathbf{k})\exp(i\mathbf{k}\cdot\mathbf{r}) \\ &= \frac{1}{V}\sum_{k\neq 0}\sum_{j=1}^{N}\frac{q_j}{k^2\varepsilon_0}\exp\left[i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}_j)\right]\exp(-k^2/4\alpha)\end{aligned} \tag{2.15}$$

Hence the contribution to the interaction energy due to ϕ_1 is

$$\begin{aligned}U_{RECIP.} &= \frac{1}{2}\sum_i q_i\phi_1(r_i) \\ &= \frac{1}{2}\sum_{k\neq 0}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_iq_j}{Vk^2\varepsilon_0}\exp(i\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_j))\exp(-k^2/4\alpha) \\ &= \frac{V}{2}\sum_{k\neq 0}\frac{1}{k^2\varepsilon_0}|\rho(\mathbf{k})|^2\exp(-k^2/4\alpha)\end{aligned} \tag{2.16}$$

where

$$\rho(\mathbf{k}) = \frac{1}{V}\sum_{i=1}^{N} q_i\exp(i\mathbf{k}\cdot\mathbf{r}_i) \tag{2.17}$$

However, the contribution to the interaction energy given in Equation 2.16 includes the interaction between the continuous Gaussian charge cloud, of charge q_i, with itself. This

term, called the self-interaction term, needs to be corrected for. The contribution of this charge distribution to the interaction energy is given by[203]

$$U_{SELF} = \frac{1}{4\pi\varepsilon_0}\sqrt{\frac{\alpha}{\pi}}\sum_{i=1}^{N} q_i^2 \tag{2.18}$$

The U_{SELF} term is independent of the particle's positions and thus is constant throughout the simulation. Therefore, the total contribution to the potential energy is

$$U_{COUL} = U_{REAL} + U_{RECIP.} - U_{SELF} \tag{2.19}$$

or

$$\begin{aligned} U_{COUL} = \frac{1}{2}\left(\frac{1}{4\pi\varepsilon_0}\right)\sum_{n}{}' \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{q_i q_j}{|\mathbf{r}_{ij}+n\mathbf{L}|} * \mathrm{erfc}(\sqrt{\alpha}|\mathbf{r}_{ij}+n\mathbf{L}|) \\ + \frac{V}{2}\sum_{k\neq 0}\frac{1}{k^2\varepsilon_0}|\rho(\mathbf{k})|^2 \exp(-k^2/4\alpha) \\ - \frac{1}{4\pi\varepsilon_0}\sqrt{\frac{\alpha}{\pi}}\sum_{i=1}^{N} q_i^2 \end{aligned} \tag{2.20}$$

As α is present in both the real space and reciprocal space sums it can be chosen to optimise the speed of the simulation.

Figures 2.1 and 2.2 represent the charge distributions for reciprocal and real space respectively.

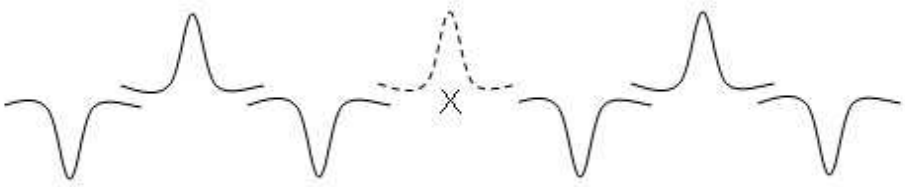

Figure 2.1: Charge distribution for Fourier reciprocal space.

Smooth Particle Mesh Ewald Summation

The smooth particle mesh Ewald (SPME) method is a modification of the Ewald summation. This technique is generally faster than conventional Ewald summation, leading

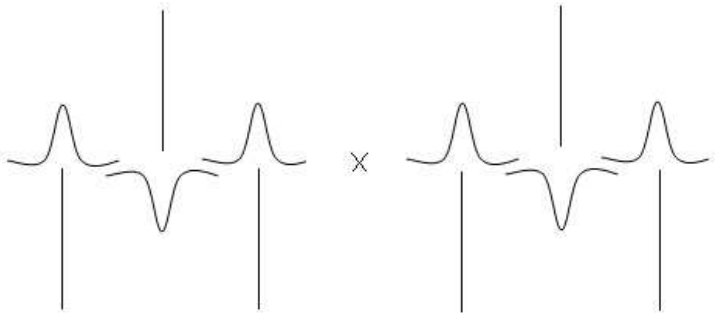

Figure 2.2: Charge distribution for real space.

to it being used throughout this work, in molecular dynamics simulations. Within this method the reciprocal space term is converted into a form suitable for Fast Fourier Transform (FFT). To accomplish this, the Gaussian charge distribution is approximated by a gridded charge distribution and 3D FFT is applied to the grid. The SPME method scales as $N\log(N)$, whereas the Ewald sum usually scales as $N^{3/2}$.

Parry Summation

The Parry summation[204, 205] is a variation of the Ewald summation. It is used for surface simulations where crystals are periodic in two-dimensions. This method suggests that the crystal consists of a series of charged planes of infinite size as opposed to an infinite lattice. When the electrostatic interactions are summed up, the vectors are now separated into in-plane vectors and vectors perpendicular to the plane. Hence meaning the sum of plane charges can now not be presumed to be zero. Evaluation of the reciprocal space term is needed. Heyes *et al*[206] provide a detailed derivation of the reciprocal space term.

2.1.2 Short-range Intermolecular Interactions

The contribution to the lattice energy made by short-range interactions consists of a combination of different contributions. Repulsive interactions occur at small separations due to neighbouring ion electron charge clouds and core-core interactions. At the same time, dipole-dipole interactions occur at these distances, due to fluctuating dipoles on each ion, resulting in attractive van der Waals forces. Other terms can also be included in the model, such as short-range interactions between three ions or more when considering covalent systems, in order to represent directionality in the bonding. Throughout this work,

the short-range attractive and repulsive contributions are described by simple parameterized potential functions. As a result, in order to obtain reliable quantitative results, careful consideration of the parameters used to accurately describe the properties of the lattice is needed. Potential parameter derivation occurs from either using empirical fitting to experimental data, or, fitting to more accurate simulations. The most commonly used short-range potentials, hence those being used in this work, are described below.

Lennard-Jones Potential

The Lennard-Jones potential is a mathematical approximation that describes the complicated nuclear and electronic repulsions which dominate the attractive intermolecular interactions at short separations. The potential takes the form:

$$U(r_{ij}) = \frac{A_{ij}}{r_{ij}^n} - \frac{B_{ij}}{r_{ij}^m} = \frac{E_0}{n-m}\left[m\left(\frac{r_0}{r_{ij}}\right)^n - m\left(\frac{r_0}{r_{ij}}\right)^m\right] \tag{2.21}$$

Where r is the interatomic distance and E_0 is the interaction energy at the equilibrium separation (or bond length) r_0.

However, the form used in this work was

$$U(r_{ij}) = \frac{A}{r_{ij}^{12}} - \frac{B}{r_{ij}^6} = E_0\left[\left(\frac{r_0}{r_{ij}}\right)^{12} - 2\left(\frac{r_0}{r_{ij}}\right)^6\right] \tag{2.22}$$

The first term, denotes the repulsion between electron clouds, which dominate at very short distances. At larger distances the second term dominates, making up the attractive part of the potential, modelling the van der Waals forces. It must be mentioned that Equation 2.22 is known as the Lennard-Jones (12-6) potential (see Figure 2.3 for a diagram of its potential function), the numbers referring to the powers that r_{ij} are raised to. The most common values used for the repulsive and attractive terms are 12 and 6, respectively and are used exclusively in this work. However, values such as 9-3, 9-6 and 12-10 could also be used, dependant on the system being modelled, for example, 12-10 is often used to describe hydrogen bonding, and 9-3 often used to model the interaction between a continuous solid wall and the atoms/molecules of a liquid. The A and B parameters are chosen to fit the materials physical properties.

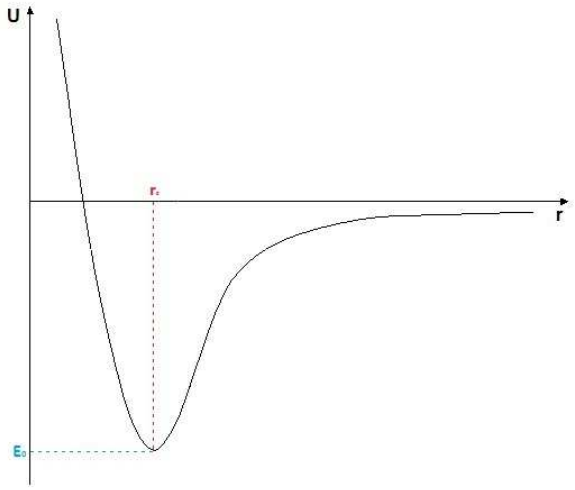

Figure 2.3: 12-6 Lennard-Jones potential diagram.

Buckingham Potential

The Buckingham potential is similar to the Lennard-Jones potential in that as with the Lennard-Jones potential, the first term represents the repulsive forces and the second term represents the attractive forces. The difference being that the repulsive term is replaced by an exponential as shown in Equation 2.23

$$U(r_{ij}) = A_{ij}e^{-r_{ij}/\rho_{ij}} - \frac{C_{ij}}{r_{ij}^6} \tag{2.23}$$

A_{ij}, ρ_{ij} and C_{ij} are parameters which differ for each pair of interacting ions. A_{ij} and ρ_{ij} represent ion size and hardness respectively. The exponential within the repulsive term of the equation, enables it to be more flexible than the Lennard-Jones potential. It is also a better mathematical match for the underlying physics. Due to this property the Buckingham potential is widely used in the successful modelling of polar solids, where the repulsive terms are more significant. However, this thesis has made use of Lennard-Jones potentials to ease mixing organic and inorganic parameters.

2.2 Intramolecular Interactions

When building a model to describe a pure ionic solid e.g. MgO, the intermolecular terms described in the previous section are sufficient to reproduce the fundamental properties of the system. However, as the system gains a more covalent nature, e.g. calcium carbonate,

then additional terms must be included to describe the bonding within the covalent parts of the system. In the case of the organic molecules studied in this thesis, the covalent nature dominates and the intramolecular potentials describing bond vibrations, bond bending and bond twisting must be derived such that this model can be written as

$$U_{intra} = U_{bonds} + U_{angles} + U_{torsions} \tag{2.24}$$

2.2.1 Bonding Interactions

By far the simplest way to describe the oscillations of a bond is to assume the vibrations are purely harmonic and can therefore be described as a simple harmonic oscillator of the form of Equation 2.25:

$$U(r_{ij}) = \frac{1}{2}k\left(r_{ij} - r_0\right)^2 \tag{2.25}$$

In Figure 2.4 the energy is plotted in relation to Equation 2.25, where the energy curve is parabolic. The zero of the curve and subsequently the equation is found at $r = r_0$, where r_0 is the equilibrium distance (or bond length). Any energy in excess of this, for example E1, arises because of extension or compression of the bond, which may be likened to the behaviour of a spring.

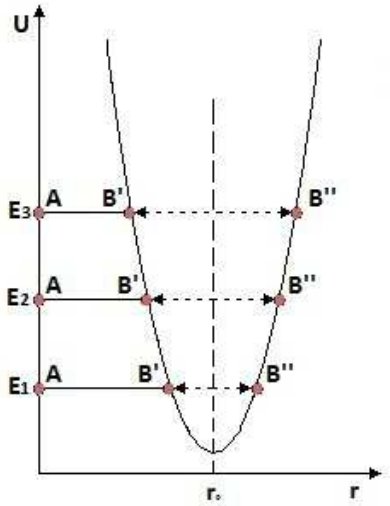

Figure 2.4: Parabolic harmonic oscillator curve of energy against interatomic distance.

The energy of the whole system is at a minimum when the two atoms stay at a mean interatomic distance, so that the attractive and repulsive forces are balanced. Any attempt

to pull the atoms further apart is met by resistance from the attractive forces. Repulsive forces arise quickly as a result of any attempt to squeeze the atoms closer together. In order to distort bonds an input of energy is required, whether it be for extension or compression of bonds. Figure 2.4 shows that if one atom (A) at r=0 is believed to be static, the other atom oscillates between B' and B''. If the energy is increased to E3 for example, there are more vigorous oscillations present, leading to a greater degree of compression or extension. However, as noted in the observed differences between the infra-red spectrum of HCl based on a pure simple harmonic model and experimental data, real molecules do not completely obey the laws of harmonic motion. Every bond will inevitably reach its breaking point, leading to the dissociation of the molecule into atoms. For smaller bond length distortions the bond can be said to be perfectly elastic, making the simple harmonic parabola an acceptable method. However, when the bond distortions get larger (>10 % of bond length), a more complicated behaviour must be assumed.

A simple deviation from the simple harmonic oscillator, often used in molecular modelling, is the Morse potential. It takes the form of

$$U(r_{ij}) = A_{ij}(1 - e^{(-B_{ij}(r_{ij}-r_0))})^2 - A_{ij} \tag{2.26}$$

where A_{ij} is the bond dissociation energy, r_0 is the equilibrium bond distance, and B_{ij} is a function of the slope of the potential energy well.

Figure 2.5 exhibits a comparison of the energy against interatomic distance curves for both the harmonic oscillator and Morse potentials.

In many well tested models for describing organic molecules (AMBER, DREIDING, etc.) it has been shown that the speed up seen from using a simple harmonic motion is more significant than the loss in accuracy compared to the Morse potential and thus is the model of choice for describing the bonding in this work. One advantage of using the simple harmonic model is that atoms initially far from their equilibrium position will quickly equilibrate without the risk of bond breaking.

2.2.2 Three-Body Potentials

This potential describes the directionality of the bonds involved in a system. Hence, it is a further component of the interactions of covalent species is the bond-bending term. It is added to take into account the energy penalty for deviations from the equilibrium value.

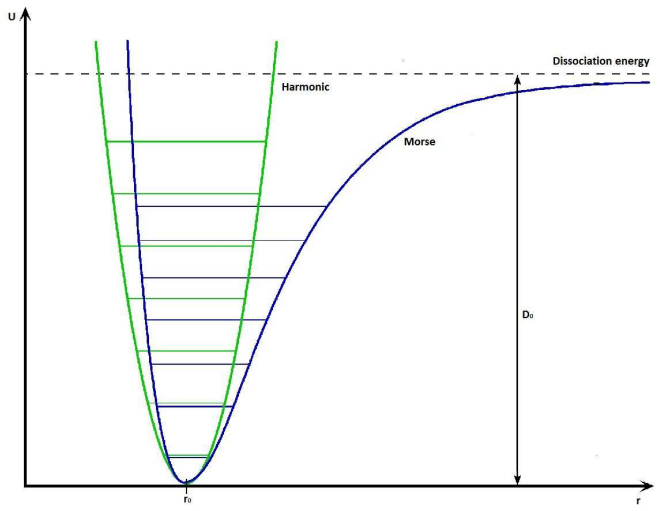

Figure 2.5: A comparison of the harmonic oscillator curve and the Morse curve.

Again in its most basic form a simple harmonic model is used.

$$U(\theta_{ijk}) = \frac{1}{2}k_{ijk}(\theta_{ijk} - \theta_0)^2 \tag{2.27}$$

where k_{ijk} is the three-body force constant and θ_0 the equilibrium angle.

The same weaknesses as discussed in the previous section apply here, and more complex anharmonic models, based on trigometric functions could be used. However, for the purpose of this work, the simple harmonic model three-body potentials, as integrated into AMBER force fields, were utilised.

2.2.3 Four-Body Potentials

A four-body potential is incorporated to denote the effect of deviations of torsion angles from their equilibrium values on the energy, taking the form

$$U(\phi_{ijk}) = k_{ijkl}(1 - \cos(n\phi_{ijkl})) \tag{2.28}$$

where k_{ijkl} is the four-body force constant, n is equal to 2 and ϕ_{ijkl} is the torsion angle.

2.3 Polarisability

As noted in the discussion of short range intermolecular interactions, one important interaction that occurs at short separations is the polarisation of the atoms in the system.

The induction of a dipole in an ion's electron charge cloud when it is brought close to an asymmetric field is known as ionic polarisability. The short-range interactions between ions can be affected by this dipole induction. Simple rigid ion models, such as in Equation 2.4, ignore this polarisability, considering each ion as a formal point charge. A way of including ionic polarisability into the model is to use the shell model, whereas in the organic systems polarisability can be incorporated by means of multipole models.

2.3.1 The Shell Model

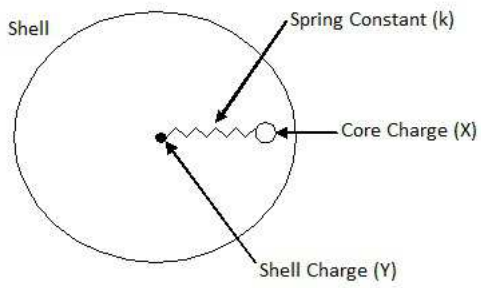

Figure 2.6: Schematic of the shell model.

The shell model was originally developed in 1958 by Dick and Overhauser[207]. It is a simple mechanical model, in which an ion is represented by two components; a core and a shell, which are connected by a harmonic spring. The harmonic core-shell potential models the interactions between the negatively charged electron cloud and the positively charged nucleus, as shown in Equation 2.29.

$$U(r_{c-s}) = \frac{1}{2} k r_{c-s}^2 \tag{2.29}$$

Ion polarisability (α) is related to the shell charge (Y) and the spring constant (k)

$$\alpha = \frac{Y^2}{k} \tag{2.30}$$

Several studies have been involved in the comparison between the rigid ion model and

the shell model[208, 209, 210]. The rigid ion model has been used to produce some very successful investigations for a wide range of systems, despite it being a very simplistic model[211, 212, 213]. However, in the comparison work, the shell model was shown to be superior. Although, the major disadvantage of the shell model is that it significantly increases the particle number, resulting in longer computational times and subsequently higher computational costs. Therefore, molecular dynamics techniques often use the simple rigid ion model, as is the case in this work.

2.4 Amino Acid Interactions

Amino acids can interact with other molecules, including other amino acids, in a variety of ways. The four major non-covalent bonds or forces responsible for intermolecular interactions can be categorised as either electrostatic or hydrophobic in nature. Electrostatic interactions can then be further sub-divided down into hydrogen bonding, charge-charge interactions and van der Waals forces[214].

Hydrogen bonds are amongst the strongest non-covalent forces, being strong enough to confer structural stability but weak enough to be readily broken. They form between hydrogen atoms and more electronegative elements, such as oxygen and nitrogen. The usual length of a hydrogen bond is approximately 2 Å, with the total distance between the two electronegative atoms typically being 2.7 to 3 Å. The hydrogen bond has many of the characteristics of a covalent bond but much weaker. The strength of the hydrogen bond is dependent on the alignment between the hydrogen atom and the two electronegative atoms, with small deviations from a linear alignment being allowed, however, the strength of the resultant hydrogen bond is weakened. Figure 2.7 exhibits some common examples of hydrogen bonds. All of the displayed functional groups can also form hydrogen bonds with water, in fact, when exposed to water they are far more likely to interact with water due to the concentration of water available to them. In order for hydrogen bonds to form within biochemical macromolecules the donor and acceptor groups must be shielded from water, hence, hydrogen bonds most often occur within the hydrophobic interior of a polypeptide where water cannot penetrate[214].

Charge-charge interactions are the strongest non-covalent forces and extend greater distances than other intermolecular forces. The strength of these interactions in an aqueous environment is greatly dependent on the nature of solvent; water significantly weakens

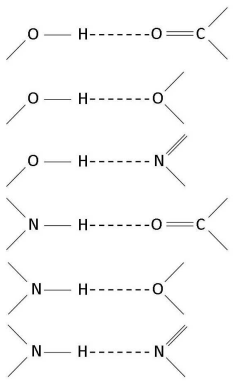

Figure 2.7: Common examples of hydrogen bonds.

these forces. These interactions play a role in the recognition of one molecule by another. As well as being in charge of the attractions between oppositely charged functional groups in a protein, charge-charge interactions are also responsible for the mutual repulsion of like charged groups[214].

The interactions between two permanent dipoles, or the interactions between a permanent dipole and an induced dipole in a neighbouring molecule, are known as van der Waals forces. Although they operate over similar distances to hydrogen bonds van der Waals forces are much weaker. These forces are produced between all neutral atoms by transient electrostatic interactions, occurring when the atoms are very close together, involving both attraction and repulsion. At short internuclear distances van der Waals forces are strongly repulsive and are very weak at long intermolecular distances. The separation of two atoms by the sum of their van der Waals radii is said to be the distance at which attractive forces are maximal. Table 2.1 presents the van der Waals radii of several atoms of importance in biochemical macromolecules. Despite their individual weakness, the clustering of atoms within a protein allows for the establishment of a large number of van der Waals forces, thus, playing a key role in maintaining molecular structures[214].

Atom	Radius (nm)
Hydrogen	0.12
Oxygen	0.14
Nitrogen	0.15
Carbon	0.17
Sulfur	0.18

Table 2.1: Van der Waals radii of some biochemically important atoms[214].

2.5 Potential Parameters Utilised

This thesis focuses on the following four key systems:

1. The organic molecules.
2. The magnetite crystal.
3. The water.
4. The interaction of the above three with each other (cross terms).

In order to study these systems, a model that describes the whole system could, potentially, be defined from first principles; however, this would be a laborious process. Alternatively, to gain an insight into the real systems, rather than just fit the parameters, well-defined literature values were utilised, thus leaving the challenge of fitting to derive reliable cross terms as practiced by Freeman *et al*[134].

2.5.1 Amino Acid and Peptide Potential Parameters

The organic molecules are described using parameters originally derived as part of the AMBER (or Assisted Model Building with Energy Refinement) project, which was originally developed by P. Kollman's group, as a collection of force fields for use in molecular dynamics of biomolecules. AMBER is also the name for the related software package, used to simulate the aforementioned force fields.

The AMBER force field function takes the form[215] below, as described in the previous sections:

$$\begin{aligned} E_{total} &= \sum_{bonds} K_r(r-r_0)^2 \\ &+ \sum_{angles} K_\theta(\theta-\theta_0)^2 \\ &+ \sum_{dihedrals} \frac{V_n}{2}\left[1+\cos(n\phi-\gamma)\right] \\ &+ \sum_{i<j}\left[\frac{A_{ij}}{R^12_{ij}} - \frac{B_{ij}}{R^6_{ij}} + \frac{q_iq_j}{\varepsilon R_{ij}}\right] \end{aligned} \tag{2.31}$$

Equation 2.31 defines the potential energy of a system. Note that despite the term force field, this equation defines the potential energy of the system; the force is the derivative of this potential with respect of position.

The first term represents the energy between covalently bonded atoms. The second term relates to the bond bending energy and can be thought of as being due to electron orbital geometry involved in covalent bonding. The third term expresses the energy of the bond torsions due to bond order and neighbouring bonds or lone pairs of electrons. It must be noted that a single bond can possess more than one of these terms, leading to the total torsional energy being expressed as a sum. The fourth term is linked to the non-bonded energy present between all atom pairs. This can be split into two components: the first being for van der Waals forces and the second being for electrostatic energies.

In order to use the AMBER force field, parameter values of the force field are needed (e.g. force constants, equilibrium bond length and angles, charges). There are a large number of parameter sets in existence, all of which are described in detail in the AMBER software user manual. Each parameter set was derived for a certain type of molecule. GAFF (General AMBER Force Field)[216] provides parameters for small organic molecules for use with, for example, small molecule ligands in conjunction with biomolecules. Peptide, protein and nucleic acid parameters are provided by parameter sets with the prefix ff and containing a two digit year number, e.g. ff99[215]. In 1995 Rob Woods[217] developed the GLYCAM force fields for simulating carbohydrates. Within this work the parameter set ff99SB was used.

The focus of this thesis was the relationship between amino acids and peptides with the {100} and {111} magnetite surfaces. The biological molecules range in size from a monomer (individual amino acids) to sequences of up to 6 monomers. In order to produce the required amino acids and peptides the AMBER software package was used, as mentioned previously. The initial structures of the amino acid sequences were generated using the AMBERTOOLS package TLEAP[218]. The sequences were capped, using the ACE and NME method[218], thus neutralising the C- and N-terminal charges and more accurately replicating conditions within the Mms6 protein, as discussed in Chapter 1. This system was then relaxed *in vacu*, using a combination of Energy Minimisation and NVT (an ensemble with constant particle number, constant volume and constant temperature, which will be discussed further in Chapter 3) molecular dynamics for 1 ns at 300 K in a simulation cell of dimensions 40 Å x 40 Å x 40 Å. Water was then added using the utility distributed with DL_POLY classic and run for a further 1 ns of NVT MD. This

was done to prepare the system before attachment to the magnetite surfaces. All of the intermolecular forces between the different biological molecules were modelled with 12-6 Lennard-Jones potentials, as produced by AMBER, these are shown in Table 2.3.

Ion	AMBER definition
N	sp^2 N in amides
H	H attached to N
CT	any sp^3 C
H1	H attached to aliphatic C with 1 electron-withdrawing substituent
HC	H attached to aliphatic C with no electron-withdrawing substituents
C	any carbonyl sp^2 C
O	sp^2 O in amides
OH	sp^3 O in alcohols, & protonated carboxylic acids
HP	H attached to C directly bonded to formally $+ve$ atoms
S	S in methionine

Table 2.2: Definitions of the different ions used in AMBER[215].

Ion pair (ij)	A_{ij} (eV $\mathring{A}^{12}$)	B_{ij} (eV $\mathring{A}^{6}$)
N-N	40972.883380	34.769428
H-H	0.006074	0.004068
CT-CT	45259.251180	29.314815
H1-H1	141.438220	0.620809
HC-HC	326.122582	0.942682
C-C	35578.570414	23.044553
O-O	16482.836953	24.510403
OH-OH	25244.442106	30.362014
HP-HP	8.757123	0.154474
S-S	181990.850560	88.862720

Table 2.3: Amino acid and peptide potential parameters[215].

2.5.2 Magnetite Potential Parameters

There are several methods for describing the magnetite potential parameters. Lewis and Catlow[219] provide potentials that incorporate a shell model built with full ionic charges,

however, this presents a problem of how to describe the charge distribution on the tetrahedral sites, where half are Fe^{+2} and half are Fe^{+3}. Also, whilst a full charge model gives an excellent reproduction of the material properties, the large difference in charges between those of the organic molecules and those of the magnetite makes fitting cross terms more complicated[134].

More recently, Rustad *et al*[193] proposed a partial charge model for magnetite, whereby, in the bulk magnetite structure above the Verwey transition, each O^{2-} was thought of as being coordinated to three octahedral $Fe^{2.5+}$ ions and one tetrahedral Fe^{3+}. When Rustad *et al*[220] worked with a magnetite-water system, all oxygens, protons, and iron ions were treated on an equivalent basis, in that, the oxygens in the crystal are the same as the oxygens in the water, and the only difference between the Fe^{3+} and $Fe^{2.5+}$ ions was the charge.

Cygan *et al*[221] developed a potential set, designed to be compatible with the organic force fields, which was aimed at being a transferrable potential set for modelling silicates. The data s*et al*so included terms for tetrahedral and octahedral Al and octahedral Fe. Its compatibility with the AMBER force fields makes it an attractive option but for the absence of a tetrahedral Fe term, this issue was later rectified by Kerisit[149] through modification of the original CLAYFF parameters. Hence, the interatomic potentials for magnetite used for this investigation were developed by Kerisit[149] using modified CLAYFF parameters. This modification was due to the results produced from an energy minimisation of a bulk hematite structure using the original CLAYFF parameters[221]. The lattice parameters yielded proved to be much larger than that of the experimental data. In order to combat this, the r_0 parameter of the octahedral species was reduced. This principle was adapted for use with magnetite, yielding results in excellent agreement with experimental data as shown in Table 2.4. This agreement was also proven to be conclusive for a series of other iron (hydr)oxide minerals.

Lattice Parameter	Experimental Å	Calculated Å	Difference %
Fe_3O_4	8.394	8.412	0.2

Table 2.4: Experimental[222] and calculated[149] lattice parameters obtained with the modified CLAYFF model.

As mentioned previously, in this work a rigid iron model is utilised. Within this model, the two cation types (Fe^{2+} and Fe^{3+}) have the same short range potentials as this proved to be the method with the most stable system and the lowest surface energies. A partial

charge model was also decided upon, as the system was more stable and better results were yielded than when a fully charged model was used. Lennard-Jones potentials were used to describe all interactions considered. The potential parameters used within this work are presented in Table 2.5. The {100} and {111} surfaces were chosen for this work, as these particular surfaces are linked with bacterial magnetite production, as mentioned before.

Ion	Charges (e)	Ion pair (ij)	A_{ij} (eV $\mathring{A}^12$)	B_{ij} (eV $\mathring{A}^6$)
FET	1.5750	FET-FET	32.5633	0.0071
FEO	1.3130	FEO-FEO	32.5633	0.0071
Oxygen (OM)	-1.0500	OM-OM	27290.9548	27.1226

Table 2.5: Magnetite potential parameters[149]. Iron in Tetrahedral Sites = FET. Iron in Octahedral Sites = FEO.

2.5.3 Water Potential Parameters

Discussion of the force fields available for modelling water could be the topic of a thesis in its own right. London South Bank University has created a database, documenting over 100 different water models[137]. Here, the water potential used is a simple flexible three-point transferable interaction potential model, known as TIP3P/Fs, as presented in the paper of Wu *et al*[223]. The TIP3P potential represents a non-polarisable model for water. The intramolecular interaction of TIP3P/Fs includes harmonic bond-stretching and bond-bending potentials as shown in Table 2.6. This particular water model was chosen as it is incorporated into both AMBER and CLAYFF, reducing the need to explicitly fit cross terms. The oxygen(OW)-oxygen intermolecular forces were modelled with a 12-6 Lennard-Jones potential. The TIP3P model does not have potential parameters for any hydrogen(HW) intermolecular forces (neither indeed do most of the simple water models). The reason being essentially that all of the hydrogen bonding can be accounted for in the Coulombic interactions.

2.5.4 Deriving Cross Terms

The interactions between organic and inorganic systems create issues when producing force fields for them. These challenges arise predominantly due to the different types of force field used to model the two systems. Thus, because different charge models are used, there

Ion	Charges (e)
OW	-0.8340
HW	0.4170

Bonds $(k(r-r_0)^2$		
Ions	k $(eV\mathring{A})^{-2}$	r_0 $(\mathring{A})$
OW-HW	23.9907	0.9572

Angles $(k(\theta-\theta_0)^2$		
Ions	k (eV rad^2)	θ (°)
HW-OW-HW	4.3383	104.5200

Ion pair (ij)	A_{ij} (eV $\mathring{A}^{12}$)	B_{ij} (eV $\mathring{A}^6$)
OW-OW	25246.0590	25.8052
HW-HW	0.0	0.0

Table 2.6: Water potential parameters. Water Oxygen = OW. Water Hydrogen = HW.

is no unique way for deriving the cross-terms. In this work the Lorentz-Berthelot mixing rules were used in order to generate the required potentials, as these are already employed in the AMBER force field. The rule states that the AB parameters for a Lennard-Jones representation are suitable averages of the Lennard-Jones parameters for AA and BB interactions, as shown in the work of Allen and Tildesley 1989[224], and Leach 2001[225].

The basic principles of the rule relates back to the Lennard-Jones equation in Chapter 2 (Equations 2.21).

The cross terms for R (R_{ij}) and E (E_{ij}) are created by

$$R_{ij} = \frac{1}{2}(R_{ii} + R_{jj}) \tag{2.32}$$

And

$$E_{ij} = \sqrt{E_{ii}E_{jj}} \tag{2.33}$$

All of the intermolecular forces were modelled with 12-6 Lennard-Jones potentials, as shown in Tables 2.7, 2.8, 2.9 and 2.10.

Ion pair (ij)	A_{ij} (eV $\mathring{A}^{12}$)	B_{ij} (eV $\mathring{A}^6$)
FET-OW	1001.2367	0.4498
FEO-OW	1001.2367	0.4498
OM-OW	26249.4643	26.4562

Table 2.7: Magnetite-water potential parameters.

Ion pair (ij)	A_{ij} (eV $\mathring{A}^{12}$)	B_{ij} (eV $\mathring{A}^{6}$)
FET/FEO-N	1247.262137	0.516298166
FET/FEO-H	5.457577171	0.018826465
FET/FEO-CT	1275.380255	0.467609854
FET/FEO-H1	98.41898679	0.079950986
FET/FEO-HC	135.8338467	0.09392657
FET/FEO-C	1130.78556	0.414595152
FET/FEO-O	853.6819681	0.450310798
FET/FEO-OH	1023.386429	0.493276859
FET/FEO-HP	37.09092915	0.049081549
FET/FEO-S	2501.164948	0.805129336

Table 2.8: Magnetite(FET/FEO)-amino acid/peptide potential parameters.

Ion pair (ij)	A_{ij} (eV $\mathring{A}^{12}$)	B_{ij} (eV $\mathring{A}^{6}$)
OM-N	33474.12726	30.72488384
OM-H	69.56586966	0.772112995
OM-CT	35414.33752	28.3052683
OM-H1	2153.351019	4.295915188
OM-HC	3128.233445	5.177831863
OM-C	31399.27979	25.09619275
OM-O	21353.21078	25.87079686
OM-OH	26287.5924	28.71839912
OM-HP	687.8435741	2.427969087
OM-S	71972.76209	49.61265566

Table 2.9: Magnetite(OM)-amino acid/peptide potential parameters.

Ion pair (ij)	A_{ij} (eV $\mathring{A}^{12}$)	B_{ij} (eV $\mathring{A}^{6}$)
OW-N	32208.5821	29.9755
OW-H	65.9863	0.7479
OW-CT	34096.9833	27.6236
OW-H1	2064.0189	4.1831
OW-HC	3001.3595	5.0443
OW-C	30231.2790	24.4918
OW-O	20518.9579	25.2232
OW-OH	25273.1025	28.0065
OW-HP	657.2350	2.3605
OW-S	69341.2416	48.4338

Table 2.10: Water-amino acid/peptide potential parameters.

Chapter 3

Theoretical Methods & Computational Techniques

The potential model used to describe the calculation of the interatomic interactions within a system was detailed in the previous chapter. This chapter moves onto describe the use of the model, within atomistic simulation techniques, to attain important information about the systems in question. Two of such techniques that can be used are discussed in detail.

The first technique is that of Energy Minimisation. This method is used to minimise the total interaction energy of a system in order to obtain its equilibrium configuration. This method as with all methods has its advantages and disadvantages. Energy minimisation is a fast technique that has been used successfully over the years in the study of mineral and molecular structures and energies[226, 219, 227], however, it completely disregards the effects of temperature, due to the technique not taking into account the vibrational properties of the crystal. This essentially means that the system is in effect being run at 0 K, with the zeropoint energy ignored. However, if the potential parameters used in the calculation are fitted to room temperature experimental data, the effect of temperature will appear implicitly in the model, though such effects will be small. Depending on the choice of minimisation regime, energy minimisation can be the least computationally demanding technique, meaning that it can be readily applied to systems with large numbers of atoms. Although when the systems are larger the simulations are expensive in terms of memory and CPU.

The second technique is Molecular Dynamics. This differs from energy minimization because the particles considered are effectively involved in time dependent motion. The kinetic energy of the system is included within this method and, as such, the effects of temperature are considered. Molecular dynamics allows for the kinetic energy of a system to evolve with time, in order for the low energy configurations to be sampled. The major drawback of this technique is that, whilst each iteration is cheap, the number of

iterations is so large that the number of steps required means the simulation is expensive in terms of time. Hence this technique is computationally a lot more expensive than energy minimisation. This is however a vital technique in the modelling of liquids and mineral surfaces that come into contact with liquid[228, 159, 229]

This chapter also details the different approaches available for describing mineral surfaces based on both two-dimensional periodicity and three-dimensional periodicity. Initially though, periodic boundary conditions must be discussed, as they have great importance in atomistic modelling.

3.1 Periodic Boundary Conditions

As the size of a system being studied has great implications on the amount of computer resources a calculation requires, it is important to take advantage of any properties of the system that reduce the work required. In the case of a solid system the unit cell defines a repeat unit, that if expanded will reproduce the system at a macro-level. If it is assumed that any effects on the system are local it is possible to describe the bulk solid simply by studying the unit cell or a supercell containing a small number of unit cells.

This approach works on the principle that a simulation cell is surrounded by an infinite number of images, so that when a particle leaves the simulation box, an image re-enters from the opposite side; meaning that the system becomes periodic, with a periodicity equal to the dimension of the simulation cell. Figure 3.1 gives a basic demonstration of this concept. The highlighted cell represents the simulation cell, whereas the surrounding cells are the images.

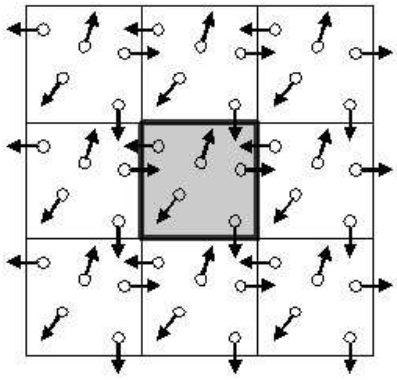

Figure 3.1: Schematic of the concept of periodic boundary conditions, the simulation cell is highlighted.

This method is very effective for crystalline solid simulations provided there is no effect

that exists on a length scale greater than the periodicity. It is also adequate for the modelling of liquids or amorphous solids, provided the simulation cell is large enough, otherwise pseudo periodic effects may appear, which lead to liquid systems appearing glossy or like disordered solids. This method can also be extended to non-bulk systems not periodic in three dimensions.

Two possible strategies can be employed when simulating surfaces and grain boundaries; the first being the use of two-dimensional periodic boundary conditions, where the periodicity of the cell is kept in only two dimensions, so there is no periodicity present perpendicular to the surface or the interface. The second strategy is to employ three-dimensional periodicity, but have the simulation cell of a size such that the interactions between the interface and its images are negligible. Also, making sure that the slab is sufficiently thick in order to ensure no interaction between the two surfaces, and that the centre of the slab behaves as a bulk would. These two strategies require different methods for the summation of Coulombic interactions.

Having developed a potential model and a method for defining the coordinates of the system, we now consider how the energy and related properties can be calculated.

3.2 Energy Minimisation

At equilibrium the energy of a system is at a minimum, therefore, the calculated interaction energy within a system should be the minimum interaction energy of the system, though this is very often not the case when setting up a simulation cell. Thus, energy minimisation is needed in order to remove the residual stresses. When considering a periodic system, there are two ways to achieve this mechanical equilibrium; the first is to perform a constant volume minimisation, where the dimensions of the cell are kept fixed but the atomic positions can change. The second method is to perform a constant pressure minimisation, where both the cell dimensions and the ions are relaxed, so that the forces on the atoms and the cell dimensions are removed. This can be expressed as the atoms being at their minimised positions when all the forces are at zero.

$$\frac{\partial U}{\partial \mathbf{r}} = 0 \tag{3.1}$$

where ∂U is the change in the energy and $\partial \mathbf{r}$ is the change in position of the ion.

There are various methods by which Equation 3.1 can be achieved; all are iterative but differ in their accuracy and computational expense. Two methods are discussed here.

3.2.1 Conjugate Gradients

The conjugate gradient technique[230] is an extension of the steepest descent method, whereby successive steps use the information on the forces from previous steps to direct the minimisation. The new positions in the steepest descent method, are calculated using Equation 3.2.

$$\mathbf{r}_{n+1} = \mathbf{r}_n - \alpha_n \mathbf{S}_n \tag{3.2}$$

where $\mathbf{r}_n$ is the coordinate at time n, α_n is a numerical constant chosen for each iteration to optimise the efficiency of the minimisation, and $\mathbf{s}_n$ is the displacement vector, which is given as

$$\mathbf{S}_n = -\mathbf{g}_n \tag{3.3}$$

With

$$\mathbf{g}_n = \frac{\partial U}{\partial \mathbf{r}_n} \tag{3.4}$$

Within this method, the displacement vector, also known as the search direction, is gained from information on the previous gradient values.

$$\mathbf{S}_n = -\mathbf{g}_n + \gamma_n \mathbf{S}_{n-1} \tag{3.5}$$

Where

$$\gamma_n = \frac{\mathbf{g}_n^{\mathrm{T}} \cdot \mathbf{g}_n}{\mathbf{g}_{n-1}^{\mathrm{T}} \cdot \mathbf{g}_{n-1}} \tag{3.6}$$

and $\mathbf{S}_1 = \mathbf{g}_1$. The superscript T refers to the transpose of the vector. When $\mathbf{S}_n$ is zero or close enough to zero to satisfy an accuracy criteria, the minimisation is complete. Each iteration is very rapid due to the fact that only the first derivative of the energy with

respect to the positions is required. However, the algorithm is much less effective than that of matrix methods such as the Newton-Raphson technique and therefore requires many more steps to reach the energy minimum.

Traditionally the gain in speed meant that conjugate gradients were still much faster even though more steps were required. As computers have got faster this is less the case for molecular mechanics but is still true for quantum mechanics.

3.2.2 Newton-Raphson

In the Newton-Raphson technique[231], a Taylor expansion is used to expand $U(r)$ to second order.

$$U(r_{n+1}) = U(r_n) + \mathbf{g}_n \cdot \delta\mathbf{r}^{\mathrm{T}} + \frac{1}{2}\delta\mathbf{r}^{\mathrm{T}} \cdot \mathbf{W}_n \cdot \delta\mathbf{r} \tag{3.7}$$

Where the displacement of a given ion is $\delta\mathbf{r}$

$$\delta\mathbf{r} = \mathbf{r}_{n+1} - \mathbf{r}_n \tag{3.8}$$

And the second derivative matrix is represented by $\mathbf{W}_n$

$$\mathbf{W}_n = -\frac{\partial^2 U}{\partial \mathbf{r}_n^2} \tag{3.9}$$

Consequently, when the system is at equilibrium (assuming that the energy change with the strain is zero)

$$\frac{\partial U}{\partial \mathbf{r}} = 0 = \mathbf{g}_n + \mathbf{W}_n \cdot \delta\mathbf{r} \tag{3.10}$$

Giving

$$\delta\mathbf{r} = -\mathbf{W}_n^{-1} \cdot \mathbf{g}_n \tag{3.11}$$

Hence

$$\mathbf{r}_{n+1} = \mathbf{r}_n - \mathbf{g}_n\mathbf{H}_n \tag{3.12}$$

Where $\mathbf{H}_n$ is the Hessian matrix and is equivalent to $\mathbf{W}_n^{-1}$.

If the system energy was perfectly harmonic in $\mathbf{r}$ the minimum system energy would be obtained in one step. However, if it is not harmonic but the displacement gives rise to a lower energy configuration, then the minimum energy is found by iteratively repeating this step. As the Newton-Raphson technique requires the calculation of both derivatives of displacement energy, as well as the inversion of the second derivative matrix, it is potentially very expensive computationally for large systems, although this overhead has reduced with faster processors and more RAM being available. This method is however incredibly reliable when compared to other methods, including the conjugate-gradients technique. The problem of computational expense can be reduced by using an approximation of the inverted second derivative matrix and recalculation of the matrix only occurs after a fixed iteration number[232]. This is the so called DFP method METADISE[233] uses, whilst more modern codes use the related BFGS method.

The updated Hessian approximation is given as

$$\mathbf{H}_{n+1} = \mathbf{H}_n + \frac{\delta\mathbf{r} \cdot \delta\mathbf{r}^{\mathrm{T}}}{\delta\mathbf{r}^{\mathrm{T}} \cdot \delta\mathbf{g}} - \frac{\mathbf{H}_n \cdot \delta\mathbf{g} \cdot \delta\mathbf{g}^{\mathrm{T}} \cdot \mathbf{H}_n}{\delta\mathbf{g}^{\mathrm{T}} \cdot \mathbf{H}_n \cdot \delta\mathbf{g}} \tag{3.13}$$

The approaches towards minimising to constant pressure are similar[234]. In this work, energy minimisation has been used for two main purposes; firstly as a way of comparing the quality of the magnetite potentials, and secondly, in the relaxation of the geometry of the organic molecules created in drawing packages to ensure the system is in a stable configuration before molecular dynamics are applied. However, as mentioned before, energy minimisation techniques neglect the temperature effects on a system. For this reasoning, molecular dynamics methodologies were used throughout this work in order to incorporate temperature effects into the investigation of the systems of interest.

3.3 Molecular Dynamics

Molecular dynamics utilises Newton's laws of motion to calculate the energy of a system over a finite period of time for all of the particles within the system. This method, as mentioned earlier, encompasses kinetic energy, doing so explicitly by assigning all particles of the system a position and a velocity, enabling the system to reach a target temperature, allowing the system to evolve with time. This allows atoms and molecules to potentially

jump over energy barriers, to reach a global energy minimum. The problem with this method is that it can only apply to small energy barriers (e.g. in the order of a few k_BT), due to the very short real time accessible to molecular dynamics simulations. The molecular dynamics simulations produced throughout this work were performed using the computer code DL_POLY[235] developed by W. Smith, T.R. Forester and I. Todorov in Daresbury, UK.

All particles of a system within a molecular dynamics simulation are initially given random velocities, enabling the system to start with the required temperature and making sure the simulation cell has no translational momentum.

$$\sum_{i=1}^{N} m_i \cdot v_i^2 = 3Nk_BT \tag{3.14}$$

And

$$\sum_{i=1}^{N} m_i \cdot v_i = 0 \tag{3.15}$$

Where the number of particles is shown as N, k_B is the Boltzmann constant, temperature is T, m_i is the mass of particle i, and velocity of particle i is v_i.

Calculating the force acting on each particle is the second step of a molecular dynamics simulation. These calculations use the same method as energy minimisation. Once the forces, $\mathbf{F}_i$, are established, calculation of the accelerations, $\mathbf{a}_i$, can occur, and the ion positions, $\mathbf{r}_i$, and velocities, $\mathbf{v}_i$, can be updated, for an infinitely small time step, as follows

$$\mathbf{a}_i(t) = \frac{\mathbf{F}_i(t)}{m_i} \tag{3.16}$$

$$\mathbf{v}_i(t + \delta t) = \mathbf{v}_i(t) + \mathbf{a}_i(t)\delta t \tag{3.17}$$

$$\mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\delta t \tag{3.18}$$

Equations 3.16, 3.17 and 3.18 are Newton's equations of motion and can only be applied strictly for an infinitesimal time step. In practice, integration algorithms, such as the Verl*et*

*al*gorithm[224], are used within computer codes to solve Newton's equations, by combating the errors due to time step, δt, size. The time step choice is incredibly important in these equations. If the time step is too large, the molecular vibrations will occur within the time step, producing large errors. On the other hand, if δt is too small, it will require too many iterations for the particles to move a significant distance, hence making the simulation time too long. Run time properties (eg. potential energy, temperature or pressure) of a system can be calculated after each step. Then the process is repeated for the required amount of time for the simulation, which could be several thousand or million steps. The velocities of the particles are scaled for the first few tens of thousands of steps to meet the required temperature. This is known as the equilibration period, where the system reaches equilibrium at a given pressure and temperature, before the collection of data. The particle's velocities are from then on not scaled and the simulation is run for as long as possible, in order to obtain converged averages of the run time properties of interest. Throughout the majority of this work a simulation time of 1 ns was used, however where indicated, the simulation time went up to 5 ns.

3.3.1 Integration Algorithms

Integration algorithms are used for updating particle coordinates by a finite time step. A Taylor expansion in the time can be applied to attain an estimate of the positions, velocities and accelerations.

$$\begin{aligned}
\mathbf{r}(t+\delta t) &= \mathbf{r}(t) + \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2 + \frac{1}{6}\mathbf{b}(t)\delta t^3 + ... \\
\mathbf{v}(t+\delta t) &= \mathbf{v}(t) + \mathbf{a}(t)\delta t + \frac{1}{2}\mathbf{b}(t)\delta t^2 + ... \\
\mathbf{a}(t+\delta t) &= \mathbf{a}(t) + \mathbf{b}(t)\delta t + ... \\
\mathbf{b}(t+\delta t) &= \mathbf{b}(t) + ...
\end{aligned} \tag{3.19}$$

Where the position of the particle is $\mathbf{r}$, $\mathbf{v}$ is its velocity, $\mathbf{a}$ is its acceleration and $\mathbf{b}$ is the third time derivative of $\mathbf{r}$. Using Equation 3.19 it is possible to calculate particle position about a position $\mathbf{r}(t)$, before and after a time step δt

$$\mathbf{r}(t+\delta t) = \mathbf{r}(t) + \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2 + \frac{1}{6}\mathbf{b}(t)\delta t^3 + \vartheta(\delta t^4) \tag{3.20}$$

$$\mathbf{r}(t-\delta t) = \mathbf{r}(t) - \mathbf{v}(t)\delta t + \frac{1}{2}\mathbf{a}(t)\delta t^2 - \frac{1}{6}\mathbf{b}(t)\delta t^3 + \vartheta(\delta t^4) \tag{3.21}$$

Where $\vartheta(x)$ is the order of accuracy. Adding Equation 3.20 and 3.21 gives

$$\mathbf{r}(t+\delta t) + \mathbf{r}(t-\delta t) = 2\mathbf{r}(t) + \mathbf{a}(t)\delta t^2 + \vartheta(\delta t^4) \tag{3.22}$$

Or

$$\mathbf{r}(t+\delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\delta t) + \frac{\delta t^2}{m}\mathbf{f}(t) + \vartheta(\delta t^4) \tag{3.23}$$

This is the basis of the Verl*et al*gorithm[236]. This method is used to calculate the new position of a particle from the previous and current positions, in addition to the current force on the particle. Hence, the velocities are not required to compute trajectories, but are used to estimate the kinetic energy. To calculate the velocities in the system Equation 3.21 can be subtracted from Equation 3.20.

$$\mathbf{r}(t+\delta t) - \mathbf{r}(t-\delta t) = 2\mathbf{v}(t)\delta t + \vartheta(\delta t^3) \tag{3.24}$$

Or

$$\mathbf{v}(t) = \frac{\mathbf{r}(t+\delta t) - \mathbf{r}(t-\delta t)}{2\delta t} + \vartheta(\delta t^2) \tag{3.25}$$

This means that the calculation of velocities can only occur once $\mathbf{r}(t+\delta t)$ is known and they are subject to a δt^2 order of errors. Unfortunately with this algorithm numerical imprecision can be introduced, due to the addition of a small term ($\vartheta(\delta t^2)$) to a difference of large terms ($\vartheta(\delta t)$) in Equation 3.23. The algorithm used throughout this work is a modified version of the Verl*et al*gorithm known as the Verlet leapfrog scheme[237], defined as

$$\mathbf{v}(t+\frac{1}{2}\delta t) = \frac{\mathbf{r}(t+\delta t) - \mathbf{r}(t)}{\delta t} \tag{3.26}$$

And

$$\mathbf{v}(t-\frac{1}{2}\delta t) = \frac{\mathbf{r}(t) - \mathbf{r}(t-\delta t)}{\delta t} \tag{3.27}$$

Thus, Equation 3.23 can be defined as

$$\frac{\mathbf{r}(t+\delta t)-\mathbf{r}(t)}{\delta t}=\frac{\mathbf{r}(t)-\mathbf{r}(t-\delta t)}{\delta t}+\frac{\delta t}{m}\mathbf{f}(t)+\vartheta(\delta t^3) \tag{3.28}$$

Then from Equations 3.26 and 3.27

$$\mathbf{v}(t+\frac{1}{2}\delta t)=\mathbf{v}(t-\frac{1}{2}\delta t)+\frac{\delta t}{m}\mathbf{f}(t)+\vartheta(\delta t^3) \tag{3.29}$$

And

$$\mathbf{r}(t+\delta t)=\mathbf{r}(t)+\mathbf{v}(t+\frac{1}{2}\delta t)\delta t+\vartheta(\delta t^4) \tag{3.30}$$

The values of the positions and forces at time t and the velocities half a time step behind are required for the algorithm. The initial step is to use Equation 3.29 to calculate the new velocities, whereby the velocities leap over the coordinates to produce the subsequent half step values $\mathbf{v}(t+\frac{1}{2}\delta t)$. Current velocities can be calculated during this step as follows

$$\mathbf{v}(t)=\frac{1}{2}\left[\mathbf{v}(t+\frac{1}{2}\delta t)+\mathbf{v}(t-\frac{1}{2}\delta t)\right] \tag{3.31}$$

When the velocities have advanced, the positions can be updated using Equation 3.30. Fortunately, due to the fact that at no point in the calculations is the difference of two large quantities taken to obtain a small one, the numerical precision of the algorithm is improved.

3.3.2 Ensembles

Within molecular dynamic simulations the conditions are known as ensembles. Three of such ensembles are the NVE, NVT and NPT ensembles.

The microcanonical ensemble (NVE) is where the number of particles, the volume and the total energy of the system are kept constant.

$$\mathcal{H}_{NVE}=U+K.E. \tag{3.32}$$

Where U is the potential energy and $K.E.$ the kinetic energy.

The canonical ensemble (NVT) is where the number of particles, the volume and the temperature of the system are kept constant. The temperature of the system is kept constant by using a Nosé-Hoover thermostat[238] as a heat bath. Newton's equations of

motion are modified due to the Nosé-Hoover algorithm by including a friction coefficient, χ

$$\frac{d\mathbf{v}(t)}{dt} = \frac{\mathbf{f}(t)}{m} - \chi(t)\mathbf{v}(t) \tag{3.33}$$

Where the friction coefficient is controlled by the first order differential equation

$$\frac{d\chi(t)}{dt} = \frac{1}{\tau_T^2}\left[\frac{T}{T_e xt} - 1\right] \tag{3.34}$$

Where τ_T is an arbitrary time constant for temperature fluctuations, T_{ext} is the temperature of the heat bath, and T is the instantaneous temperature. Thus, modification of the Verlet leapfrog algorithm occurs as follows

$$\begin{aligned}
\chi(t+\frac{1}{2}\delta t) &= \chi(t-\frac{1}{2}\delta t) + \frac{\delta t}{\tau_T^2}\left[\frac{T}{T_e xt} - 1\right] \\
\chi(t) &= \frac{1}{2}\left[\chi(t-\frac{1}{2}\delta t) + \chi(t+\frac{1}{2}\delta t)\right] \\
\mathbf{v}(t+\frac{1}{2}\delta t) &= \mathbf{v}(t-\frac{1}{2}\delta t) + \left[\frac{\mathbf{f}(t)}{m} - \chi(t)\mathbf{v}(t)\right]\delta t \\
\mathbf{v}(t) &= \frac{1}{2}\left[\mathbf{v}(t-\frac{1}{2}\delta t) + \mathbf{v}(t+\frac{1}{2}\delta t)\right] \\
\mathbf{r}(t+\delta t) &= \mathbf{r}(t) + \mathbf{v}(t+\frac{1}{2}\delta t)\delta t
\end{aligned} \tag{3.35}$$

However, as $\mathbf{v}(t)$ is needed to calculate T and therefore itself, in order to acquire self-consistency several iteration are required. Within the DL_POLY code, the iteration number is set to 3 and the standard Verlet leapfrog algorithm is utilized to obtain the first prediction of $\mathbf{v}(t)$ and T. In NVT, the conserved quantity, which is derived from the extended Hamiltonian for the system is

$$\mathcal{H}_{NVT} = \mathcal{H}_{NVE} + f k_B T_{ext}\left(\frac{\tau_T^2 \chi^2(t)}{2} + \int_0^t \chi(s)\mathrm{d}(s)\right) \tag{3.36}$$

The isobaric-isothermal ensemble (NPT) is where the number of particles, the pressure and temperature of the system are kept constant. These conditions can either be isotropic or anisotropic. Isotropic; meaning that only the dimensions of the cell could vary. Anisotropic; meaning that both the dimensions and shape of the cell may change. A modification of the Hoover algorithm, known as the Melchionna modification can be

used[239], which entails the use of a Nosé-Hoover thermostat and a barostat that follows a similar algorithm, where the velocities are calculated as follows

$$\frac{d\mathbf{v}(t)}{dt} = \frac{\mathbf{f}(t)}{m} - \big[\chi(t) + \eta(t)\big]\mathbf{v}(t) \tag{3.37}$$

Where η is the friction coefficient of the barostat

$$\frac{d\eta(t)}{dt} = \frac{1}{Nk_BT_{ext}\tau_p^2}\mathbf{V}(t)[P - P_{ext}] \tag{3.38}$$

Where P_{ext} is the pressure of the barostat, P is the instantaneous pressure, τ_p is an arbitrary time constant for pressure fluctuations and $\mathbf{V}(t)$ is the volume of the system at time t, with

$$\frac{d\mathbf{V}(t)}{dt} = [3\eta(t)]\mathbf{V}(t) \tag{3.39}$$

Thus modification of the Verlet leapfrog algorithm occurs as follows

$$\begin{aligned}
\eta(t + \frac{1}{2}\delta t) &= \eta(t - \frac{1}{2}\delta t) + \frac{\mathbf{V}(t)\delta t}{Nk_BT_{ext}\tau_p^2}[P - P_{ext}] \\
\eta(t) &= \frac{1}{2}\left[\eta(t + \frac{1}{2}\delta t) + \eta(t - \frac{1}{2}\delta t)\right] \\
\mathbf{v}(t + \frac{1}{2}\delta t) &= \mathbf{v}(t - \frac{1}{2}\delta t) + \left(\frac{\mathbf{f}(t)}{m} - \big[\chi(t) + \eta(t)\big]\mathbf{v}(t)\right]\delta t \\
\mathbf{v}(t) &= \frac{1}{2}\left[\mathbf{v}(t + \frac{1}{2}\delta t) + \mathbf{v}(t - \frac{1}{2}\delta t)\right]
\end{aligned}$$

As

$$\begin{aligned}
\frac{d\mathbf{r}(t)}{dt} &= \mathbf{v}(t) + \eta(t)[r(t) - R-)]
\end{aligned}$$

Where R_0 is the centre of mass of the system, then

$$\mathbf{r}(t + \delta t) = \mathbf{r}(t) + \left(\mathbf{v}(t + \frac{1}{2}\delta t) + \eta(t + \frac{1}{2}\delta t)\big[\mathbf{r}(t + \frac{1}{2}\delta t) - R_0\big]\right)\delta t$$

With

$$\mathbf{r}(t + \frac{1}{2}\delta t) = \frac{1}{2}\left[\mathbf{r}(t) + \mathbf{r}(t + \delta t)\right] \tag{3.40}$$

Again, in order to obtain self-consistency several iterations are required. Within the DL_POLY code, the iteration number is set at 4 and the standard Verlet leapfrog algorithm is used to predict initial estimates for T, P, $\mathbf{v}(t)$ and $\mathbf{r}(t+\frac{1}{2}\delta t)$. The new volume can be derived using

$$\mathbf{V}(t+\delta t) = \mathbf{V}(t)\exp\left[3\delta t\eta(t+\frac{1}{2}\delta t)\right] \tag{3.41}$$

And the new cell vectors from

$$\mathbf{H}(t+\delta t) = \mathbf{H}(t)\exp\left[\eta(t+\frac{1}{2}\delta t)\delta t\right] \tag{3.42}$$

Where $\mathbf{H}$ is the cell matrix whose columns are the three vectors. Conserved quantity in the isotropic conditions is

$$\mathcal{H}_{NPT} = \mathcal{H}_{NVT} + P_{ext}\mathbf{V}(t) + \frac{3Nk_BT_{ext}}{2}\eta(t)^2\tau_p^2 \tag{3.43}$$

In the anisotropic conditions, adjustments of the isotropic algorithm occur, allowing for the cell shape to change by defining η as tensor $\underline{\underline{\eta}}$. Anisotropic conditions were, however, not used in this work.

3.3.3 Molecular Dynamics Properties

A number of properties can be calculated using molecular dynamics simulations. There are two classes these properties can be separated into; static system properties and dynamical system properties.

Static system properties

Within this class further separations can be applied, these are; structural properties and thermodynamic properties. Thermodynamic properties include:

- Kinetic Energy

$$\langle K.E.\rangle = \left\langle \frac{1}{2}\sum_i^N m_i v_i^2 \right\rangle \tag{3.44}$$

- Configuration Energy

$$U = \left\langle \sum_{i}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} U(r_{ij}) \right\rangle \tag{3.45}$$

- Temperature

$$T = \frac{2}{3Nk_B} \langle K.E. \rangle \tag{3.46}$$

- Pressure

$$P = \frac{Nk_BT}{V} - \frac{1}{3V} \left\langle \sum_{i}^{N} \mathbf{r}_i \cdot \mathbf{f}_i \right\rangle \tag{3.47}$$

Structural properties include:

- Radial Distribution Function (RDF) (or pair correlation)

$$G_{A-B}(r) = \frac{\langle n_B(r, \delta r) \rangle}{4\pi \frac{N_B}{V} r^2 \delta r} \tag{3.48}$$

  Where V is the volume of the system, r is the diameter of the shell, $n_B(r, \delta r)$ is the number of particles between shells at $r - \delta r/2$ and $r + \delta r/2$, and N_B is the total number of particle B.

Dynamical system properties

This class consists mainly of correlation functions, including

- Mean Square Displacement (MSD)(Einstein relation)

$$2Dt = \frac{1}{3} \langle |\mathbf{r}_i(t) - \mathbf{r}_i(0)|^2 \rangle \tag{3.49}$$

  Therefore, if $\langle |\mathbf{r}_i(t) - \mathbf{r}_i(0)|^2 \rangle$ is plotted as a function of time, the curve gradient produced is equal to six times the coefficient of self-diffusion of particle i. There is no diffusion in a solid, hence the MSD is flat. However, particles diffuse randomly in a liquid, and so the gradient of the curve is proportional to the diffusion coefficient. Often, the initial curve takes a parabolic form; this is due to the short amount of time for an atom to feel the effect of the presence of other atoms.

- Velocity Correlation Function (VACF)(Green-Kubo relation)

$$D = \frac{1}{3} \int_0^\infty \langle \mathbf{v}_i(t) \cdot \mathbf{v}_i(0) \rangle \mathrm{d}t \tag{3.50}$$

This is another method of establishing diffusion coefficients. Numerous other transport coefficients have had their Green-Kudo relations derived, such as the shear viscosity and the thermal conductivity[240]. In a solid, the Fourier transform of the normalised VACF gives the frequency distribution of phonon states, i.e. the density of states[234, 241].

3.3.4 Free Energy

The main properties that can be derived directly from molecular dynamics simulations were discussed in the previous section. However, certain properties of a system cannot be derived directly, such as the entropy and free energy (Helmholtz and Gibbs). This is due to these properties not being explicit functions of the phase space coordinates of a system. These properties are instead related to the phase space volume available to the system and are known as thermal quantities. Properties such as the pressure and the temperature, which can be expressed as a function of the velocities and coordinates of all of the system particles, are known as mechanical properties.

The thermodynamic integration method is one possible technique that can be used to obtain thermal quantities from molecular dynamics simulations. The method is based on the fact that deriving thermal quantities often generate mechanical properties. For example, assuming we have two states A and B, of which the free energy of state A is known and state B is unknown, at constant volume and temperature. A reversible path links the two states, consequently, by the integration of the energy along said path, the free energy of state B can be found. This path does not have to be a physical path and any parameter can be used as a thermodynamic variable, ϕ, providing that the potential energy, U, depends linearly on ϕ, so that for $\phi = 0$, U corresponds to the potential energy of the reference system and, for $\phi = 1$, U is the potential energy of the system of interest. The partition function of a system with a U that corresponds to a value of ϕ between 0 and 1 is[240]

$$Q(N, V, T, \phi) = \frac{1}{\Lambda^{3N} N!} \int \exp\left[-\frac{1}{k_B T} U(\phi) \right] \mathrm{d}\mathbf{r}^N \tag{3.51}$$

Where Λ is the thermal de Broglie wavelength

$$\Lambda - \left(\frac{h^2}{2\pi m k_B T} \right)^{1/2} \tag{3.52}$$

Where h is Planck's constant and m is the mass of a particle. The derivative of the Helmholtz free energy $A(\phi)$ with respect to ϕ is displayed as

$$\begin{aligned}\left(\frac{\partial A(\phi)}{\partial \phi}\right)_{NVT} &= -k_B T \frac{\partial ln Q(N,V,T,\phi)}{\partial \phi} \\ &= -k_B T \frac{1}{Q(N,V,T,\phi)} \frac{\partial Q(N,V,T,\phi)}{\partial \phi} \\ &= \frac{\int \mathrm{d}\mathbf{r}^N (\partial U(\phi)/\partial \phi) \exp[-\frac{U(\phi)}{k_B T}]}{\int \mathrm{d}\mathbf{r}^N \exp[-\frac{U(\phi)}{k_B T}]} \\ &= \left\langle \frac{\partial U(\phi)}{\partial \phi} \right\rangle_{\phi} \end{aligned} \tag{3.53}$$

Where $\left\langle \frac{\partial U(\phi)}{\partial \phi} \right\rangle_{\phi}$ is the ensemble average of the derivative of the potential energy of the system with respect to the thermodynamic variable. Therefore, by integrating Equation 3.53, the free energy difference between states A and B can be established.

$$A(B) - A(A) = \int_{\phi=0}^{\phi=1} \left\langle \frac{\partial U(\phi)}{\partial \phi} \right\rangle_{\phi} \mathrm{d}\phi \tag{3.54}$$

Molecular dynamic simulations are a way of deriving the free energy difference, as it only depends on the ensemble average of the mechanical quantities. The disadvantage of this is that for large systems, a large amount of simulation time is lost simulating parts of the system that are unimportant or change little as the system moves along the reaction coordinate.

3.3.5 Potential of Mean Force

An alternative approach is to define super atoms by creating one interaction unit by joining together several atoms. Potential of Mean Force (PMF) is used to define the interaction between the atoms on the super atom. The PMF is the system free energy as a function of a reaction coordinate. The reaction coordinate R is dependent on the coordinates of the atoms in the system

$$R(r) = R(r_1, r_2, \cdots, r_N) \tag{3.55}$$

A system confined in physical terms to such a reaction coordinate $R(r)$ is also restricted to a hyperspace $R(r)$ in phase space, and would be characterized by the partition function

$Q(R)$ and a free energy of $A(R)$. The probability of a given system configuration being on the hyperspace $R(r)$ is

$$\pi(p^N, r^N; R') = \frac{\exp[-\beta H(p^N, r^N)]\delta(R' - R(r^N))}{\int dp^N dr^N \exp[-\beta H(p^N, r^N)]} \tag{3.56}$$

Where the system is characterized by the sum of the kinetic and potential energy at position r and momenta p, also known as Hamiltonian $H(r, p)$. The probability of the system being at the specified reaction coordinate can be found by the integration of equation 3.56 over phase space

$$\begin{aligned} P(R') &= \int dp^N dr^N \pi(p^N, r^N; R') \\ &= \frac{\int dp^N dr^N exp[-\beta H(p^N, r^N)]\delta(R' - R(r^N))}{\int dp^N dr^N \exp[-\beta H(p^N, r^N)]} \\ &= \frac{Q(R')}{Q} \end{aligned} \tag{3.57}$$

Which in terms of free energy $A(R')$ reduces to

$$\begin{aligned} A(R) &= -k_\beta T \ln P(R') - k_\beta T \ln Q \\ &= k_\beta T \ln P(R') + constant \end{aligned} \tag{3.58}$$

The determination of $P(R\prime)$ is dependent on the application but in general there are three approaches; Boltzmann sampling, Umbrella sampling and the use of constraints. In this work constraints were used to calculate the probability of the system being at a specified reaction coordinate. The system in question is constrained at a particular value of reaction coordinate R_c

$$\sigma(r; R_c) = R(r) - R_c \tag{3.59}$$

The constraints allow the atoms Cartesian coordinates to become dependent on each other, consequently, it is necessary to transform to a set of generalized coordinates (q, r), which equate one of the coordinates to the reaction coordinate R. In order to obtain the free

energy, $A(R)$, associated with a particular reaction coordinate, R_c, the force is integrated to keep it at the reaction coordinate

$$A(R_c) = \int_{\infty}^{R_c} A(R(r))dR \tag{3.60}$$

$$A(R_c) = \frac{\left\langle (\partial r/\partial R) \cdot [-\partial U/\partial r + k_\beta T \partial \ln(|J|)/\partial r] \delta(R' - R(r^N)) \right\rangle}{\left\langle \delta(R' - R(r^N)) \right\rangle} \tag{3.61}$$

Where $|J| = |\partial r/\partial q|$. If $R(r)$ is linearly dependent on the Cartesian coordinates the Jacobian J is zero. This term may be neglected when the range of R is limited and this term is essentially constant. This approach was used throughout this to calculate the change in free energy of system involving magnetite surfaces and the attachment of amino acids and peptides at different ranges from the surface.

3.4 Mineral Surface Simulation

The main focus of this thesis is the magnetite mineral surface and its interactions. For the simulation of mineral surfaces there are two common approaches; a two-dimensional approach and a three-dimensional approach. Both methods will be discussed in detail, but first different surface types need to be examined.

3.4.1 Surface Type

Crystal surfaces can be considered as a stack of planes periodic in two-dimensions. The cleaving of mineral crystals in a particular direction, as specified by the Miller index, produced the surfaces used in this work. Within a simulation, not only the direction of the surface cut is important, but also describing the location of the surface is key. The different surface types that can be generated by different cut locations have been described by Tasker[242] (as shown in Figure 3.2). A type I surface is where each plane is charge neutral as they consist of both anions and cations in a stoichiometric ratio. There is no dipole moment perpendicular to the surface. Type II surfaces consist of a stacking sequence of charged planes but the repeat unit consists of several planes in a symmetrical configuration. This means there is no dipole moment perpendicular to the surface with

this type of surface either. Type III surfaces are made up of an alternately charged stack of planes, hence a dipole moment perpendicular to the surface is produced.

The simulation cells containing surfaces must be charge neutral and must not have a dipole moment perpendicular to the surface for the electrostatic energy to converge. Bertaut[243] showed that if within a unit cell, a dipole moment is present perpendicular to the surface; the surface energy diverges and is infinite, making Type III surfaces naturally unstable. Therefore, in order to simulate Type III surfaces, the removal of the dipole is required. Oliver *et al*[244] suggested a method by which this could be achieved. It entailed the removal of half of the ions on the top layer and transferring them to the bottom of the unit cell, as depicted in Figure 3.3. This process is known in nature as the faceting of polar surfaces into neutral surfaces. This method reconstructs the surface so that an unstable surface is built from very stable surfaces. Figure 3.4 depicts an example of faceting in MgO.

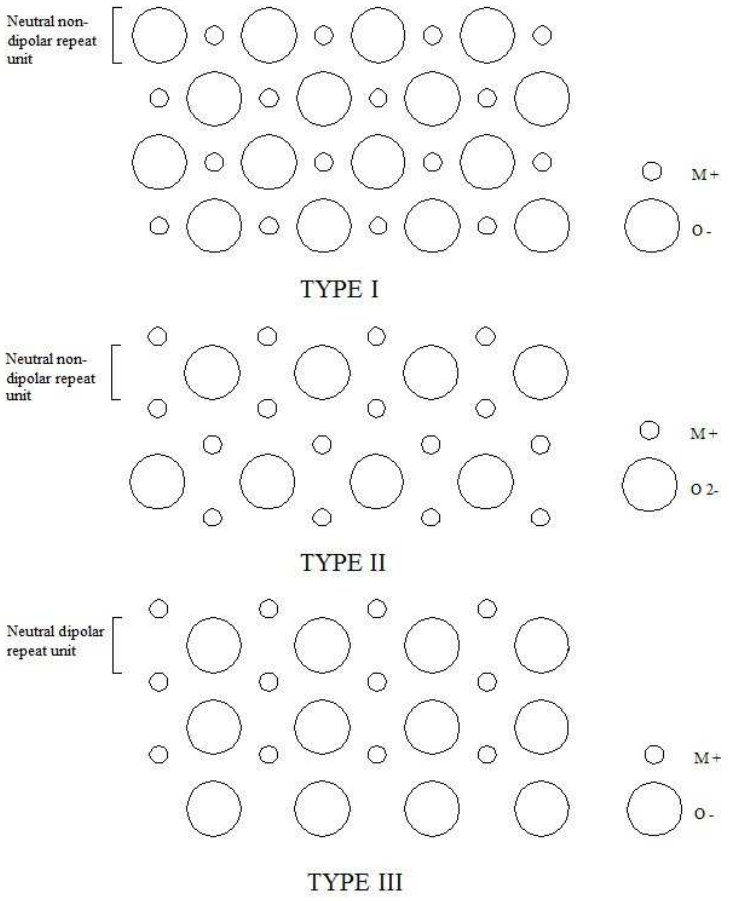

Figure 3.2: The three types of stacking surface[242].

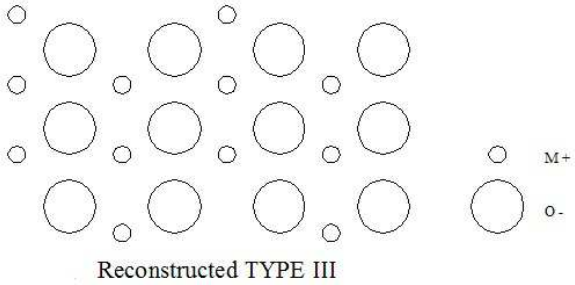

Figure 3.3: The reconstructed type III stacking sequence[242].

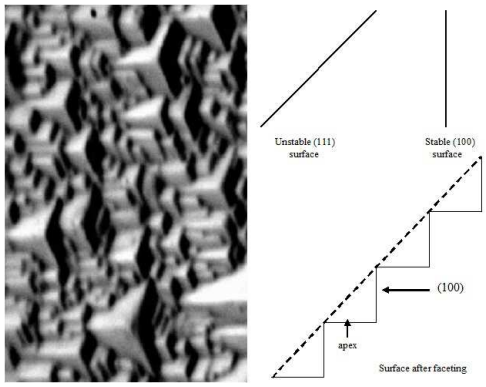

Figure 3.4: Example of faceting in MgO.

3.4.2 Two-dimensional Approach

As mentioned before, surface simulations regard a crystal as being made up of a series of charged planes parallel to the surface and periodic in two dimensions. The METADISE[233] (Minimum Energy Techniques Applied to Dislocation, Interface & Surface Energies) computer code uses the two-region method developed by Tasker[242], to model the bulk and surfaces of minerals in this work, alongside the potential model. Tasker's method suggests that a simulated crystal consists of two blocks, each of which is separated into two regions, periodic in two-dimensions. First region atoms are those near the surface and are able to relax mechanically. Those atoms in region II, however, are held fixed at their bulk equilibrium positions, representing the remainder of the surface. A surface is formed when two blocks separate. Figure 3.5 shows a schematic of Tasker's approach.

The energy of a block consists of two parts; E_1, the energy of region I and E_2, the energy of region II. Region I energy can be further separated into two components; interaction energy between region I ions and interaction energy between region I and region II ions. This further separation can also apply to region II, although due to the fact that region

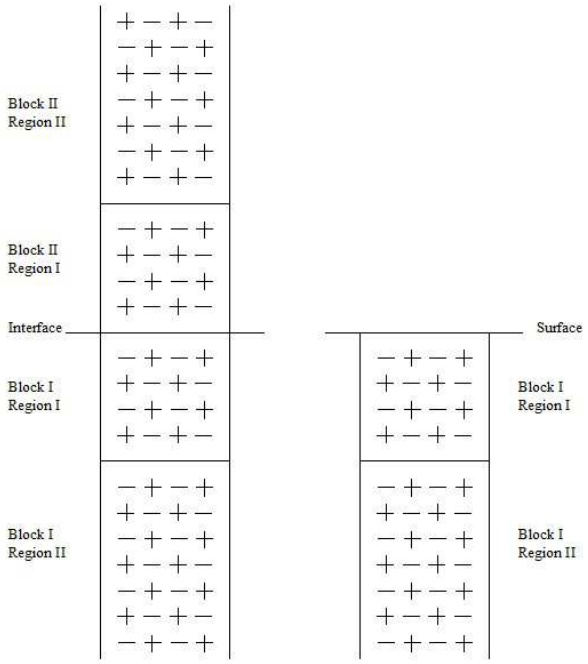

Figure 3.5: Schematic of the two region approach put forward by Tasker [242].

II ions are kept fixed throughout the simulation, the interaction energy between region II ions does not change and so is taken to be zero. Hence total crystal energy is

$$E_{TOT} = \sum_{\substack{i\epsilon I \\ j\epsilon I}}^{N} U(r_{ij}) + \frac{1}{2}\sum_{\substack{i\epsilon I \\ j\epsilon II}}^{N} U(r_{ij}) + \frac{1}{2}\sum_{\substack{i\epsilon II \\ j\epsilon I}}^{N} U(r_{ij}) \tag{3.62}$$

Electrostatic forces are calculated using the Parry technique, while short-range interactions are described by parameterised analytical functions.

3.4.3 Three-dimensional Approach

One of the main advantages that the three-dimensional Ewald method has over the two-dimensional Parry method is that it is exceptionally fast and efficient. Therefore, a special application of periodic three-dimensional boundary conditions is used when carrying out surface molecular dynamics calculations using DL_POLY. First the system is relaxed to the bulk structure and oriented so that two out of three lattice vectors are parallel to the surface. The third vector, which is perpendicular to the surface, is then increased in order to introduce a void into the crystal, creating repeating slabs of crystal, having the chosen

surface on opposite sides. This created void must be of a size that ensures no interactions between atoms on opposite sides. At the same time, the crystal slab must be of a thickness that will eliminate surface interactions between the two surfaces on opposite sides.

The Coulombic energy summation approach was chosen, due to the high efficiency, for a three-dimensional simulation cell. Once the void has been produced, minimisation is run on the system again to relax the newly formed surfaces. The simulations run throughout this work were almost exclusively run at 300K, with the ensembles available limited to NVT, due to the problem that if the volume was not kept fixed, the slabs would reform bulk structures, as this configuration will continually be more thermodynamically favourable.

3.4.4 Interfacial Energy

In this work we are interested in the interfacial or adsorption energy of a biomolecule being adsorbed on to the magnetite surfaces. The interfacial energy E_{int} of these systems is defined as

$$E_{int} = (E_{s/aa/w} - E_{s/w})(E_{aa/w} - E_w) \tag{3.63}$$

Where E_{int} is the magnetite crystal/amino acid residue/water interfacial energy, $E_{s/aa/w}$ is the average potential energy of the magnetite slab/amino acid residue/water MD simulation, $E_{s/w}$ is the average potential energy of the magnetite slab/water MD simulation, $E_{aa/w}$ is the average potential energy of a solvated amino acid residue MD simulation, and E_w is the average potential energy of a box of water MD simulation. This method has been successfully applied to calculating the interfacial energies of polysaccharides on calcite in the study of the biological control polysaccharides have on the crystallization of calcite[164].

Chapter 4

The Interaction of Amino Acids with Magnetite

4.1 Introduction

The focus of this investigation is the C–terminal region of the Mms6 protein, native to the magnetotactic bacteria Magnetospirillum magneticum strain AMB-1, and its involvement in the control of nucleation and growth of magnetite nanocrystals. Moreover, particular emphasis is given to the regions within the C–terminal which are responsible for magnetite binding. This chapter explores the interfacial relationship between the magnetite {100} and {111} crystal surfaces and the amino acids present in the C–terminal sequence (Figure) 4.1.

…MKSRDIESAQSDEEVELRDALA

Figure 4.1: Amino acid sequence of the C–terminal region of the Mms6 protein.

Amino acids consist of a central α-carbon atom surrounded by an amino group, a carboxylic acid group, a hydrogen atom and a unique side chain group, known as the R group (see Figure 4.2). In nature, there are 20 amino acids known to be involved in protein synthesis, however, the C–terminal of Mms6 consists of just 11 of these. Amino acids can be divided into subgroups based on their side chain properties: non–polar (or hydrophobic) and uncharged; polar (or hydrophilic) and uncharged; acidic (polar and charged) and basic (polar and charged). Figure 4.3 shows the structures of the 11 different amino acids present in the C–terminal of the Mms6 protein. The properties of the side chain groups are

significant as they not only affect the amino acids individually, but also greatly influence the overall three dimensional (tertiary) conformation of the resultant protein [214].

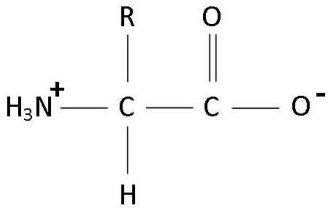

Figure 4.2: Basic structure of an amino acid. Note that it appears in zwitterion form, as it would under most biological conditions. A zwitterion is created by the transfer of a hydrogen ion from the acid portion to the base portion forming a carboxylate group (COO^-) and an ammonium group ($-NH_3^+$).

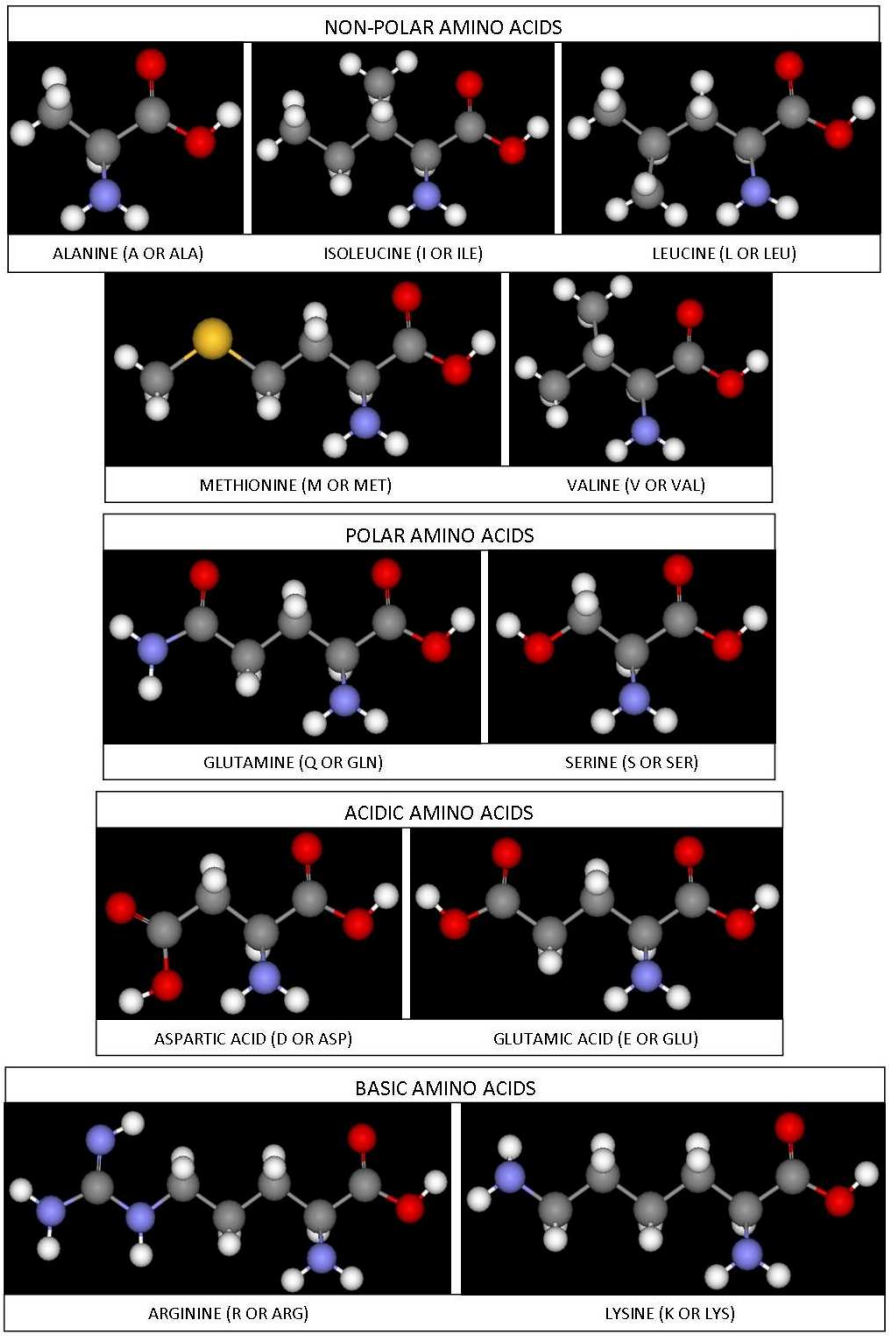

Figure 4.3: Structures of the C–terminal Mms6 protein amino acids.

Under normal physiological conditions, in the pH range of 6.8 to 7.4, the amino group exists in a protonated state (-NH_3^+ as opposed to -NH_2) due to the amino group possessing a pK_a value (also known as its isoelectric point)close to 9. Conversely, the carboxylic acid group subsists in an ionized state (COO^- as opposed to COOH), as this group possesses a pK_a value below 3. Thus, under most biological conditions, amino acids exist as zwitterions. the typical pK_a values for the amino acids involved in the C–terminal of the Mms6 protein are shown below (Table 4.1).

Amino Acid	pK_a COO^-	pK_a NH_3^+	pK_a R
Alanine	2.4	9.9	
Arginine	1.8	9.0	12.5
Aspartic Acid	2.0	9.9	3.9
Glutamine	2.2	9.1	
Glutamic Acid	2.1	9.5	4.1
Isoleucine	2.3	9.6	
Leucine	2.3	9.7	
Lysine	2.2	9.1	10.5
Methionine	2.1	9.3	
Serine	2.2	9.2	
Valine	2.3	9.7	

Table 4.1: pK_a values for the Mms6 amino acids at 25 °C[214].

Non–polar amino acids are hydrophobic, with the hydrophobicity increasing with increasing number of carbons in the hydrocarbon chains. Saturated aliphatic side chains play an important role in establishing and maintaining the tertiary structure of proteins because of their tendency to aggregate away from water. This phenomenon results from the inclination of non–polar compounds to associate with each other rather than with water molecules. Isoleucine, leucine and valine are all particularly hydrophobic due to the branching of the hydrocarbons on their side chains. Additionally, methionine contains a non–polar methyl thioether group, making it one of the most hydrophobic amino acids and resulting in its inclusion as the first amino acid in a polypeptide chain. On the other hand, alanine is an ambivalent amino acid, meaning that it can be inside or outside of the protein molecule[214].

Polar amino acids are not ionisable and, as a consequence, are charge neutral. Serine has a β-hydroxyl group in is side chain, giving a hydrophilic nature to the aliphatic side

chain. This hydroxymethyl group possesses the weak ionization properties of a primary alcohol and, consequently, does not significantly ionize in aqueous solution. Glutamine is the amide derivative of glutamic acid. Despite the neutral nature of its side chain group, glutamine is highly polar and can often be found on protein surfaces, allowing for interaction with water molecules. Glutamine can also form hydrogen bonds from its polar amide group to atoms in the side chains of other polar amino acids[214].

Aspartic acid and glutamic acid, also known as aspartate and glutamate respectively, are dicarboxylic amino acids and maintain negatively charged hydrophilic side chains at physiological pH (approximately pH7). The secondary carboxylic acid groups (β- and γ-carboxyl groups for aspartate and glutamate respectively) present in their side chain, is a weaker acid than the primary group, hence, having a higher pK_a. Both acidic amino acids confer negative charges on proteins because their side chains are ionized at physiological pH. This extra carboxylate group plays an important role in many protein–metal ion interactions and in ionic interactions[214].

Basic amino acids are polar and positively charged at pH values below their pK_a, and are very hydrophilic. Both lysine and arginine have hydrophilic side chains that are nitrogenous, allowing for the acceptance of a hydrogen ion. Lysine is a diamino acid, having both α- and ϵ-amino groups. At neutral pH the ϵ- amino group exists as an alkylammonium ion. The guanidinium ion on the side chain of arginine is protonated under all biological conditions, making it the most basic of the 20 protein based amino acids. The side chains of the basic amino acids impart a positive charge on proteins[214].

Amino acids are joined together by peptide bonds. This is achieved, when the amino group of one amino acid and the carboxylate group of another undergo a condensation reaction, during protein synthesis forming an amide linkage. Unlike the carboxyl and amino groups of free amino acids in solution, the groups involved in the peptide bond carry no ionic charges. The end to end joining of many amino acids forms a linear polypeptide. These linked amino acid moieties are known as amino acid residues[214]. The effect of peptide bonds is incorporated into this body of work by capping the amino acids to remove the effects of their α-amino and α-carboxyl groups.

The peptide bond is a partial double bond; hence, rotation about this bond is restricted. This rotational restriction of the peptide bond has major significance in determining the polypeptide chain conformation. The partial double bond character is a consequence of the electronic configuration of the nitrogen atom and of the π bonding in the carbonyl

group. The side–side merging of the p–orbitals of the nitrogen atom, and the carbon and oxygen atoms of the carbonyl group, forms π covalent bonds, causing the delocalisation of electrons. There is a variable amount (about 40 %) of π bonding between the N and C, sufficient to restrict bond rotation. Hence, each of the peptide bonds exhibits no free rotation about the carbon–nitrogen bond because of resonance form contribution, leading to a planar structure (see Figure 4.4). Many important properties of the peptide bond are utilised during protein synthesis. For instance, during polypeptide folding, the peptide bond rigidity reduces the degrees of freedom of the polypeptide. Also, due to the double bond character, the six atoms involved in the peptide bond group are always planar, as can be seen in Figure 4.5. This leads to the, rotation about the C-N and C-C bonds, by angles of ϕ and ψ respectively, defining the shape of the polypeptide (see Figure 4.5[214].

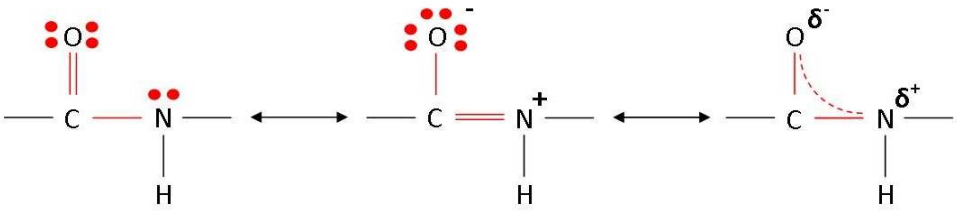

Figure 4.4: The resonance structure of the peptide bond. Adapted from Horton *et al*[214].

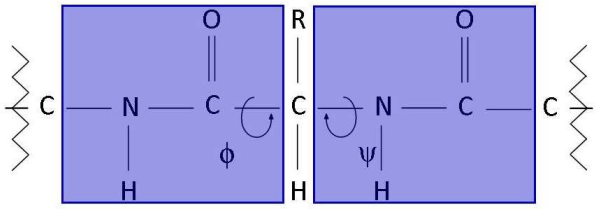

Figure 4.5: The planar conformation of a peptide bond. Note the phi and psi angles of rotation are indicated. Adapted from Horton *et al*[214].

Previously, there has been several experimental studies involving amino acids and magnetite[245, 246, 247, 248, 249, 250]. The work of these groups focused on the synthesis of magnetite nanoparticles in the presence of amino acids in solution or magnetite nanoparticles coated in, and functionalised with, amino acids for use in biomedical applications (see Chapter 1). These studies have mostly utilised the charged polar amino acids, but do not consider the use of biogenic magnetite and its formation through biomineralisation. However, Arakaki *et al* and Amemiya *et al* focused their work on the biosynthesis of magnetite using the tightly bound magnetosome membrane protein, Mms6, as it was shown to mediate the formation of cubo–octahedral magnetite nanocrystals consisting of {100} and {111} crys-

tal faces [105, 104]. Arakaki *et al* synthesised peptides that mimicked partial sequences of the Mms6 protein, with results suggesting that the C–terminal acidic region of the protein having significant control over the morphology of magnetite crystals[105].

This chapter explores the attachment of the individual C–terminal sequence amino acids, both *in vacu* and solvated, to the {100} and {111} crystal surfaces. This study was split into two sections; the first focused on a classic molecular dynamics system and the second experimented with a constrained molecular dynamic system using the Potential of Mean Force (PMF).

4.2 Unconstrained system

4.2.1 Computational Methods

As mentioned in Chapter 2, TLEAP was utilised to create the capped amino acid residues. The structures were relaxed with AMBER and subsequently using DL_POLY. TIP3P/fs water was added and the simulation was run for 1 ns. The relaxed *in vacu* amino acids were then placed in the vacuum gap above the magnetite slab surface and run for an additional 5 ns. Water was then re-added to the system and another 1 ns of molecular dynamics was run. All systems were run at 300 K. The ensembles, potentials and parameters used are described in Chapters 2 and 3.

The evolution of potential energy of all amino acid residues as a function of time were plotted, an example of which is shown for solvated alanine in Figure 4.6. A full set of potential energy vs time plots can be found in the Appendix. The equilibration period for all simulations was 20 ps, after which point the potential energy of all the amino acid residue systems converge.

4.2.2 Results and Discussion

Non–polar Amino Acid Attachment

The magnetite crystal consists of both tetrahedral and octahedral iron ions which have positive charges leading to an attraction to the negatively charged oxygen ions in the amino acid residues. It is difficult to gauge the distance between an oxygen atom within the amino acid residues, and it's next nearest neighbour tetrahedral or octahedral iron.

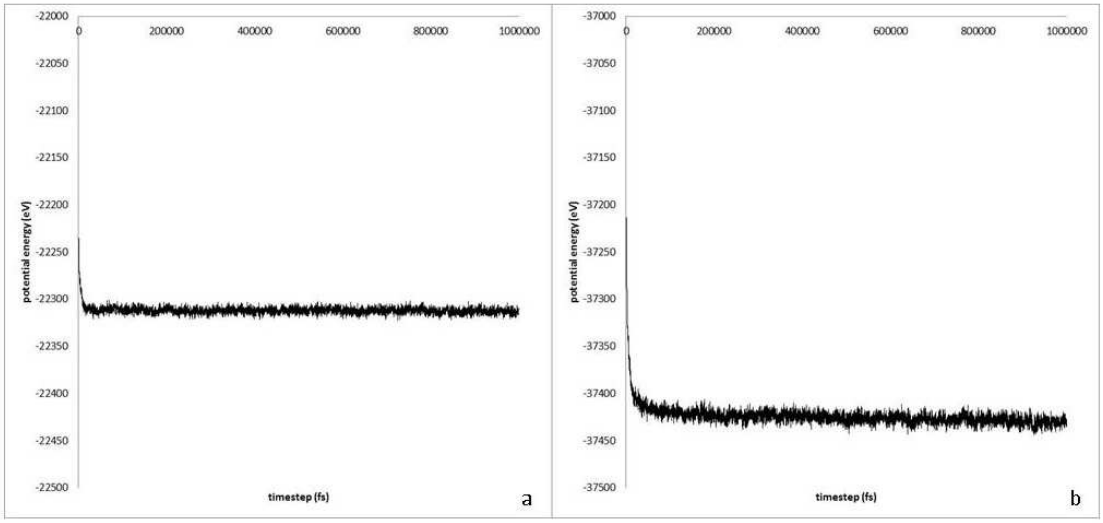

Figure 4.6: Evolution of potential energy as a function of time plots for the alanine residue. a) {100} solvated surface. b) {111} solvated surface.

This is, however, possible to measure by studying the radial distribution function (RDF) of a tetrahedral (FET) or octahedral (FEO) iron and the oxygen of the amino acid residues, providing a graphical representation of how the density of residue oxygen ions varies as a function of distance from a reference particle (FET or FEO), averaged over all Fe ions. A distance of 1.5 Å to 2.5 Å from an amino acid oxygen to the next nearest neighbouring tetrahedral or octahedral iron of magnetite is classified as a bond. Amino acid oxygen-iron distances or r values for all non–polar amino acids can be found in Tables 4.2 and 4.3.

Figure 4.7 exhibits the RDF profiles for alanine. The RDF profiles of the remaining non–polar amino acids can be found in the Appendix. The non–polar amino acid RDF data shows that for the {100} surface, only tetrahedral iron bonding was present for alanine and valine, whereas, with the {111} surface only octahedral iron bonding was exhibited. For isoleucine, octahedral iron bonding was exhibited for both systems on the {100} surface, whilst tetrahedral iron bonding is only present for the *in vacu* system. In contrast, only octahedral iron bonding is existent on the {111} surface. When the bonding behaviours of leucine and methionine were investigated, both octahedral and tetrahedral iron bonding was exhibited for the {100} surface. However, again only octahedral iron bonding was available on the {111} surface. These results suggest that when the non–polar amino acids are attached to the magnetite surfaces, different iron attachment sites are favoured, with octahedral iron attachment being the preferred for the {111} surface.

Additionally, it was shown that the RDF first peak intensity was greater in the in vacu sys-

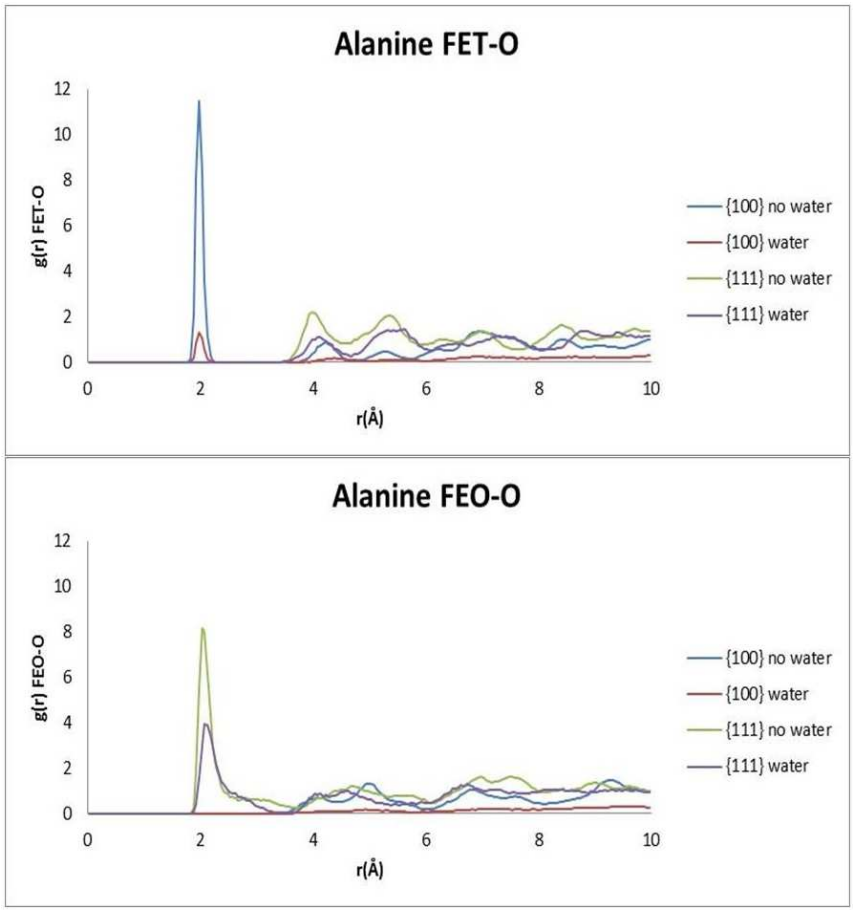

Figure 4.7: RDF plots for alanine. FET is tetrahedral iron, FEO is octahedral iron. Blue is {100} *in vacu*. Red is {100} solvated. Green is {111} *in vacu*. Purple is {111} solvated.

tem than in the equivalent solvated system, with the exception of methionine–tetrahedral iron bonding on the {100} surface, which shows similar values for both system types. This data suggests upon system solvation it is more difficult for amino acid oxygen to have an iron as a next nearest neighbour and for this to remain for the simulation duration. For the solvated systems, the most intense peak was shown for valine on the {111} surface having octahedral iron as its next nearest neighbour, suggests that the oxygen present in this amino acid is more freely accessible to the next nearest neighbouring irons within the bond distance of 1.5 Å to 2.5 Å throughout the simulation duration.

The RDF data can be examined alongside the $FE - O_{aminoacid}$ bond distance data extracted from the final simulation coordinates (Table 4.4).

Only N–terminal peptide bond oxygen bonding was possible due the lack of oxygen in the non–polar amino acid residue side chains. The non–polar amino acid bond distance data shows that, there was no bonding present for alanine, methionine and valine in the *in vacu* {100} surface system, however, in the equivalent isoleucine and leucine systems,

N–terminal peptide bond oxygen bonding was exhibited. For both surfaces, no bonding between magnetite iron and amino acid oxygen occurred when the systems were solvated, suggesting more favourable bonding between magnetite iron and the oxygen in water bonds, than the oxygen in isoleucine and leucine, leading to the amino acids being pushed away from the magnetite surface, allowing for the movement of water closer to the surface and binding (Figure 4.8). The RDF data reflects this finding, with the solvated systems having less intense $g(r)$ peaks than the corresponding *in vacu* systems, with the slight exception of methionine–tetrahedral iron bonding on the {100} surface. A greater amount of bonding was present for the {111} surfaces suggesting preferable attachment, however, the presence of water in the system is detrimental to iron bonding.

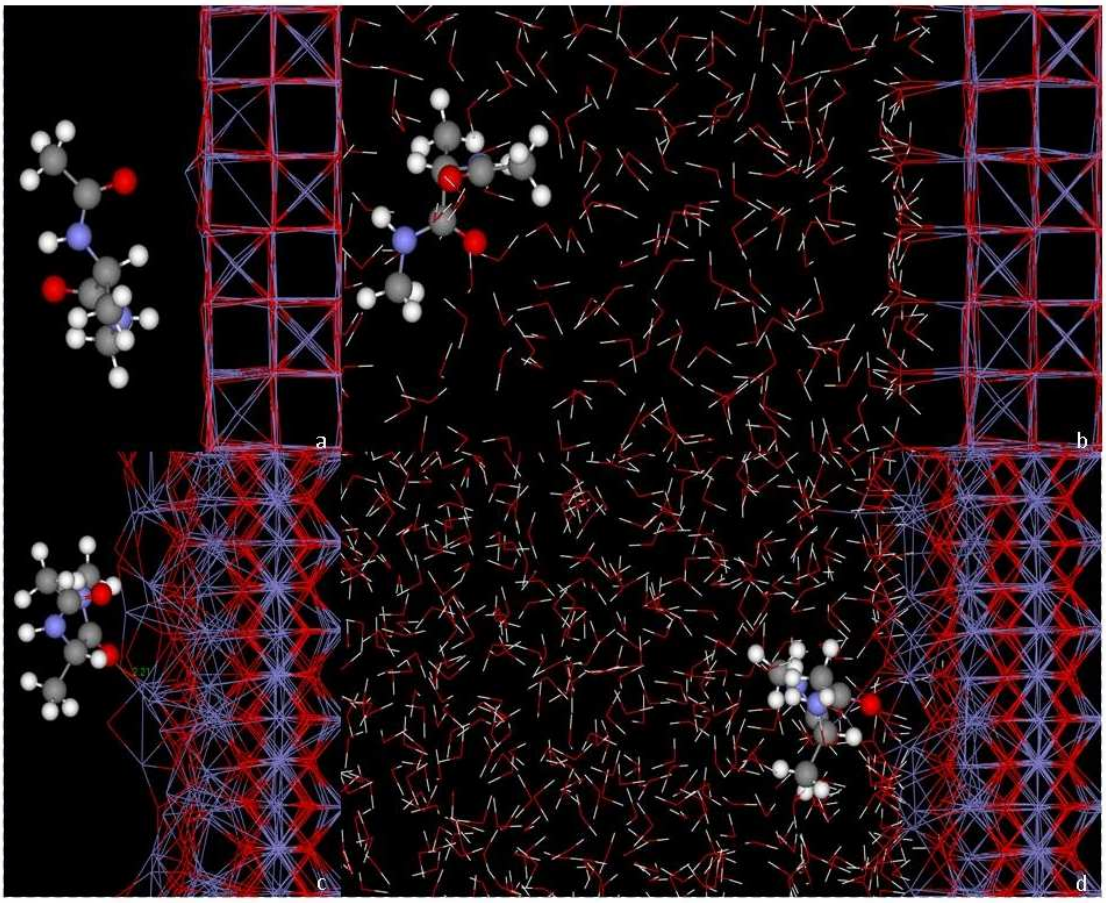

Figure 4.8: Bond distance images for the *in vacu* and solvated systems of the alanine residue. a){100} *in vacu*, b){100} solvated, c){111} *in vacu* and d){111} solvated.

In order to get a more accurate representation of bonding behaviour, throughout the simulation, residence times and coordination numbers where considered. Residence time is the average amount of time that a particle spends in a particular system, in this case, how long (ps) a bond between an FET or FEO ion and an oxygen ion from the amino acid

residues remains between 0 and 2.5 Å. The coordination number is the total number of points of attachment to a central atom or ion. Table 4.5, represents the average residence times and coordination numbers for all oxygen ions present in the non–polar amino acids residues, attached to either/both FET and FEO ions, on the {100} and {111} magnetite surfaces.

The non–polar amino acid residence time data shows that, for alanine there was no bonding lasting over 15% of the simulation length. With isoleucine attached to the {100} surface, there was no bonding lasting over 36% of the simulation length. Conversely, for the {111} surface *in vacu* system, no tetrahedral iron bonding was present, although, octahedral iron bonding to the N–terminal peptide bond oxygen was displayed lasting over 99% of the simulation length. With the addition of water, the residence time of this bond drops to 215.78 ps, suggesting that attachment at this site is no longer favourable.

A very similar trend was exhibited for leucine as for isoleucine, whereby, there was no bonding lasting over 21% of the simulation length for the {100} surface, and no tetrahedral iron bonding present for the {111} surface. There was again octahedral iron bonding to the N–terminal peptide bond oxygen lasting over 99% of the simulation length for the *in vacu* system. However, unlike with isoleucine, when this system is solvated this bond continues to last over 99% of the simulation length, suggesting that the presence of water has little to no effect. With methionine there was no bonding lasting over 2% of the simulation length.

Valine exhibits the same type of bonding as isoleucine. When attached to the {100} surface, there was no bonding lasting over 0.2% of the simulation length. Whilst, for the {111} surface *in vacu* system, there was no tetrahedral iron bonding exhibited, however, octahedral iron bonding to the N–terminal peptide bond oxygen was shown lasting over 96% of the simulation length. Upon solvation, the residence time of this bond drops to 306.52 ps, suggesting that attachment at this site is no longer favourable.

These results show that there is very minimal bonding present between magnetite and the non–polar amino acid residues, with methionine exhibiting almost no bonding to magnetite; this could be due to presence of sulphur. Leucine exhibits the most bonding, a possible reason being the effect of its branched aliphatic side chain; it is also a major component in ferritin sub-units, suggesting a link with iron.

Amino Acid	Fe type	Surface	System	r (Å)
ALA	FET	100	*in vacu*	1.98
			solvated	1.98
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	-
			solvated	-
		111	*in vacu*	2.03
			solvated	2.08
ILE	FET	100	*in vacu*	2.08
			solvated	-
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.08
			solvated	2.23
		111	*in vacu*	2.03
			solvated	2.03
LEU	FET	100	*in vacu*	2.13
			solvated	2.18
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.13
			solvated	2.18
		111	*in vacu*	2.08
			solvated	2.03

Table 4.2: Non–polar amino acid r (Å) values from RDF data.

Amino Acid	Fe type	Surface	System	r (Å)
MET	FET	100	*in vacu*	2.08
			solvated	2.28
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.13
			solvated	2.28
		111	*in vacu*	1.98
			solvated	2.33
VAL	FET	100	*in vacu*	1.98
			solvated	1.98
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	-
			solvated	-
		111	*in vacu*	2.03
			solvated	2.03

Table 4.3: Continued. Non–polar amino acid r (Å) values from RDF data.

{100} surface attachment			
Amino Acid	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
ALA	NBO	NB	NB
ILE	NBO	2.00	NB
LEU	NBO	2.27	NB
MET	NBO	NB	NB
VAL	NBO	NB	NB
{111} surface attachment			
Amino Acid	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
ALA	NBO	2.21	NB
ILE	NBO	2.05	NB
LEU	NBO	2.18	2.04
MET	NBO	NB	NB
VAL	NBO	1.98	2.13

Table 4.4: Non–polar amino acid $FE - O_{aminoacid}$ bond distance data. BL is bond length, NBO is N–terminal peptide bond oxygen, and NB is no bonding.

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
ALA {100} surface	NBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ALA {111} surface	NBO	0.00	0.00	518.39	1.09	0.00	0.00	145.83	0.29
ILE {100} surface	NBO	1765.08	0.66	1104.97	0.33	0.00	0.00	14.07	0.07
ILE {111} surface	NBO	0.00	0.00	4977.00	1.00	0.00	0.00	215.78	0.55
LEU {100} surface	NBO	7.83	0.00	1033.73	0.95	0.00	0.00	11.27	0.17
LEU {111} surface	NBO	0.00	0.00	4993.00	1.00	0.00	0.00	999.00	1.00
MET {100} surface	NBO	0.00	0.00	5.86	0.00	0.00	0.00	4.89	0.01
MET {111} surface	NBO	0.00	0.00	0.00	0.00	0.00	0.00	16.37	0.03
VAL {100} surface	NBO	0.00	0.00	8.07	0.00	0.00	0.00	0.00	0.00
VAL {111} surface	NBO	0.00	0.00	4832.84	1.00	0.00	0.00	306.52	1.00

Table 4.5: Residence times (RT) and average coordination numbers (av. CN) of the non–polar amino acid residues. (NBO = N–terminal peptide bond oxygen).

The magnetite-amino acid interfacial energy is another property that can be estimated from these simulations, giving an indication of how easily the amino acid residue can be adsorbed. The interfacial energies were calculated using Equation 4.1.

$$E_{int} = (E_{s/aa/w} - E_{s/w})(E_{aa/w} - E_w) \tag{4.1}$$

Where E_{int} is the magnetite crystal/amino acid residue/water interfacial energy, $E_{s/aa/w}$ is the average potential energy of the magnetite slab/amino acid residue/water MD simulation, $E_{s/w}$ is the average potential energy of the magnetite slab/water MD simulation, $E_{aa/w}$ is the average potential energy of a solvated amino acid residue MD simulation, and E_w is the average potential energy of a box of water MD simulation. A schematic representation is shown in Figure 4.9. This method has been successfully applied to calculating the interfacial energies of polysaccharides on calcite in the study of the biological control polysaccharides have on the crystallization of calcite[164].

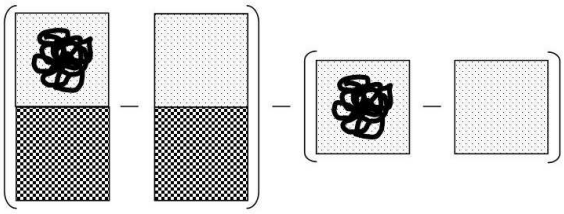

Figure 4.9: Schematic representation of the interfacial energy of the magnetite crystal–amino acid residue–water system. The checked box represents the magnetite crystal, the dotted box represents the water and the black line indicates the amino acid residue in solution.

All related energy data from the interfacial energy calculations can be found in Table 4.6.

The solvated system interfacial energy data shows that the {100} surface produces lower interfacial energy values (between 4.98 eV and 6.21 eV) than the {111} surface (between 11.03 eV and 14.45 eV). On the {100} surface, methionine has the highest interfacial energy value and valine has the lowest, whilst on the {111} surface, again methionine has the highest interfacial energy however alanine now exhibited the lowest value. As methionine consistently present with the highest energy, it suggests that the presence of the side chain sulphur in the system was detrimental to iron binding. Additionally, this data shows that the addition of water to the system has a detrimental effect on the interfacial energy, as the *in vacu* system energies are much lower (in most cases of a negative value). The solvation of the system also affects the way that the amino acid residues react at the interface, as the interfacial energy trends for the *in vacu* systems are

{100} surface attachment				
	E_{int} of s-a (eV)	E_{diff} of s-a (eV)	E_{int} of s-a-w (eV)	E_{diff} of s-a-w (eV)
ALA	-2.25	2.73	5.70	0.50
ILE	-1.90	2.38	5.63	0.58
LEU	-3.61	4.09	5.86	0.34
MET	0.48	0.00	6.21	0.00
VAL	-3.66	4.13	4.98	1.22
{111} surface attachment				
	E_{int} of s-a (eV)	E_{diff} of s-a (eV)	E_{int} of s-a-w (eV)	E_{diff} of s-a-w (eV)
ALA	-3.65	6.31	11.03	3.42
ILE	2.66	0.00	13.54	0.91
LEU	-4.54	7.20	13.05	1.40
MET	-1.48	4.15	14.45	0.00
VAL	-9.64	12.31	13.77	0.68

Table 4.6: Interfacial energies for all non–polar amino acid residues, using the Yang *et al* method[164]. E_{int} of s-a is the interfacial energy of the slab–amino acid residue system. E_{diff} of s-a is the difference in interfacial energy from the highest interfacial energy of the slab–amino acid residue system. E_{int} of s-a-w is the interfacial energy of the slab–amino acid residue–water system. E_{diff} of s-a-w is the difference in interfacial energy from the highest interfacial energy of the slab–amino acid residue–water system.

very different to the solvated systems. For the *in vacu* systems, the lowest energies were exhibited for valine, however, the highest solvated interfacial energies were presented for methionine and isoleucine, on the {100} and {111} surfaces, respectively. This suggests that valine is the preferred amino acid residue for attachment in the *in vacu* systems. The increase in interfacial energy with the addition of water could be the result of a preferential attachment of the magnetite surface to the water, leading to energy needing to be put into the system to allow for the amino acid residue to get close to the magnetite surface.

Whilst the absolute values for the adsorption energy are clearly an order of magnitude larger than expected, the relative energies within a sequence may be more reliable for amino acid binding potential comparison. The trends shown when water is present in the system are counter-intuitive and contradictive of experimental studies; however, it is clear that this is an artefact of the energy calculation method utilised, as the expected trend is revealed in the relative residence times of various surface interactions. As it can be demonstrated that:

$$\mathrm{R}T \times ln(\tau^1/\tau^2) = (\Delta H^2_{ads} - \Delta H^1_{ads}) = (E^1_{des} - E^2_{des}) \tag{4.2}$$

then this method is perhaps more reliable for determining the change in adsorption energy between the various amino acids. This interfacial energy method also requires the measurement of small differences between large energies, which leads to large uncertainties in the results. Another possible error incorporated into the calculation of the interfacial energies for the hydrated systems, is the assumption that all the water molecules within the systems containing the slab and water are perfectly hydrated throughout the slab, i.e. the bulk density is that of the bulk water simulation. However, as the simulations are run within an NVT ensemble, this is not necessarily the case. Therefore, if insufficient water is present at the start of the simulation, some of the water molecules will eventually become under coordinated, thus increasing the energy of the simulation cells. Furthermore, rather than directly calculating the adsorption energy, via this interfacial energy method, a further alternative is to use an indirect method, such as those based on the Potential of Mean Force, and instead, which has been considered in Section 3.3.5.

Polar Amino Acid Attachment

R values for the polar amino acids can be found in Table 4.7.

Figure 4.10 exhibits the RDF profiles for glutamine. The RDF profiles of the remaining polar amino acids can be found in the Appendix. The RDF data for the polar amino acids shows that glutamine exhibits both tetrahedral and octahedral iron bonding on the {100} surface. While, for the {111} surface only octahedral iron bonding is present. A similar trend is shown for serine, however, there is also no octahedral iron bonding present for the {100} solvated system. This data suggests that different iron attachment sites are favoured dependent on which surface is used. The lack of octahedral iron bonding for the {100} solvated system of serine, suggests that the addition of water is detrimental to the iron binding potential.

The polar amino acid RDF data also demonstrates that, the solvated systems produce a first peak with weaker intensity than their corresponding *in vacu* systems, proposing that with the introduction of water into the system, it is more difficult for iron to be the next nearest neighbour of an amino acid oxygen and for this to remain for the simulation duration. For the solvated systems, serine on the {111} surface having octahedral iron as its next nearest neighbour, was shown to exhibit the most intense $g(r)$, although this is

Amino Acid	Fe type	Surface	System	r (Å)
GLN	FET	100	*in vacu*	1.98
			solvated	1.98
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	-
			solvated	-
		111	*in vacu*	2.03
			solvated	2.08
SER	FET	100	*in vacu*	2.08
			solvated	-
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.08
			solvated	2.23
		111	*in vacu*	2.03
			solvated	2.03

Table 4.7: Polar amino acid r (Å) values from RDF data.

very close in intensity to that of glutamine on the {111} surface having octahedral iron as its next nearest neighbour. These results suggest that in the {111} surface systems the oxygen present in the polar amino acids is more freely accessible to the next nearest neighbouring irons.

On comparison of the polar amino acid RDF data with that of the non–polar it was shown that the same type of bonding was seen for glutamine as was exhibited for leucine and methionine suggesting that the presence of the two charged groups in glutamine has no effect on iron binding type, which may be due to the oppositely charged regions of the side chain cancelling each other out through ionic bonding. A similar trend is shown between serine and alanine and valine, with the additional octahedral iron bonding present for the {100} *in vacu* serine system, suggesting that the side chain oxygen presence in serine does to some extent increase its iron binding potential.

The $FE-O_{aminoacid}$ bond distance data for the polar amino acids is represented in Table 4.8.

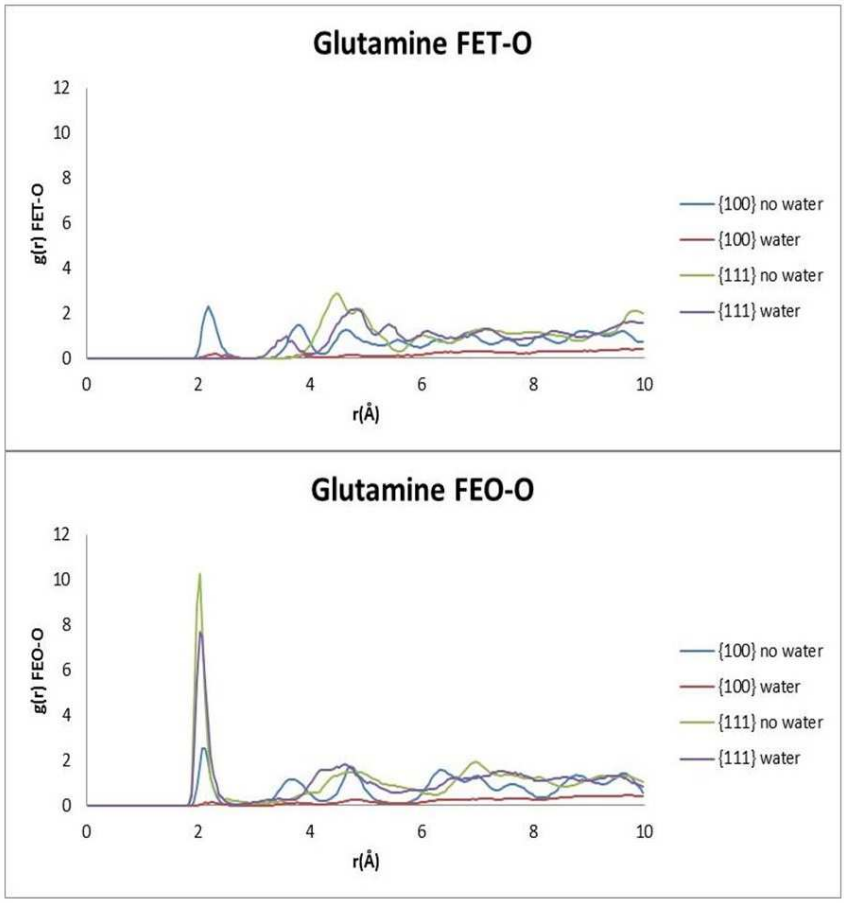

Figure 4.10: RDF plots for glutamine. FET is tetrahedral iron, FEO is octahedral iron. Blue is {100} *in vacu*. Red is {100} solvated. Green is {111} *in vacu*. Purple is {111} solvated.

The polar amino acid bond distance data shows that for the {100} *in vacu* system, serine displayed no bonding present and glutamine exhibited only N–terminal peptide bond oxygen bonding. Upon solvation, neither amino acid exhibited any form of bonding, suggesting more favourable bonding between magnetite iron and water oxygen (Figure 4.11). For the {111} *in vacu* system, much more bonding was present signifying that this is the preferred attachment surface. When water was added to the system all of the bonds present in the *in vacu* system were again present, suggesting that the addition of water has no effect on iron binding.

On comparison of the polar amino acid bond distance data with that of the non–polar it was shown that oxygen being present in the side chains has minimal effect on {100} surface iron binding, as no bonding was exhibited for serine, as with alanine, methionine and valine. Glutamine exhibited a similar bonding to isoleucine and leucine, however the bond was to the side chain oxygen not the N–terminal peptide bond oxygen, as with isoleucine and leucine. With the {111} surface, no increased iron binding effect was seen

{100} surface attachment			
Amino Acid	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
GLN	NBO	NB	NB
	Q O	2.05	NB
SER	NBO	NB	NB
	S O	NB	NB
{111} surface attachment			
Amino Acid	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
GLN	NBO	1.95	2.19
	Q O	1.88	2.11
SER	NBO	2.20	1.97
	S O	NB	NB

Table 4.8: Polar amino acid $FE-O_{aminoacid}$ bond distance data. BL is bond length, NBO is N–terminal peptide bond oxygen, Q O is glutamine side chain oxygen, S O is serine side chain oxygen is and NB is no bonding.

by the presence of oxygen in serine as the same amount of bonding was demonstrated as with leucine and valine. However the presence of the charged oxygen in glutamine has a positive effect on iron binding as it exhibits more bonding than the non–polar amino acids systems.

Table 4.9 represents the residence times and related average coordination numbers for bonding between all oxygen present in the polar amino acid residues and tetrahedral and/or octahedral iron on both the {100} and {111} magnetite surfaces.

The polar amino acid residence time data shows that, for glutamine there was no side chain oxygen bonding lasting over 42% and 26% of the simulation length for the {100} and the {111} surface, respectively. However, with the N–terminal peptide bond oxygen, there was no bonding seen for the {100} surface, but the {111} surface exhibited octahedral iron bonding lasting over 99% of the simulation length for both the *in vacu* and solvated systems, suggesting favourable attachment.

With serine attached to the {100} surface, there was no N–terminal peptide bond oxygen bonding present and no side chain oxygen bonding lasting over 1% of the simulation length. For the {111} surface, there was no tetrahedral iron bonding existent, nevertheless, side chain oxygen–octahedral iron bonding was seen lasting over 99% of the simulation length

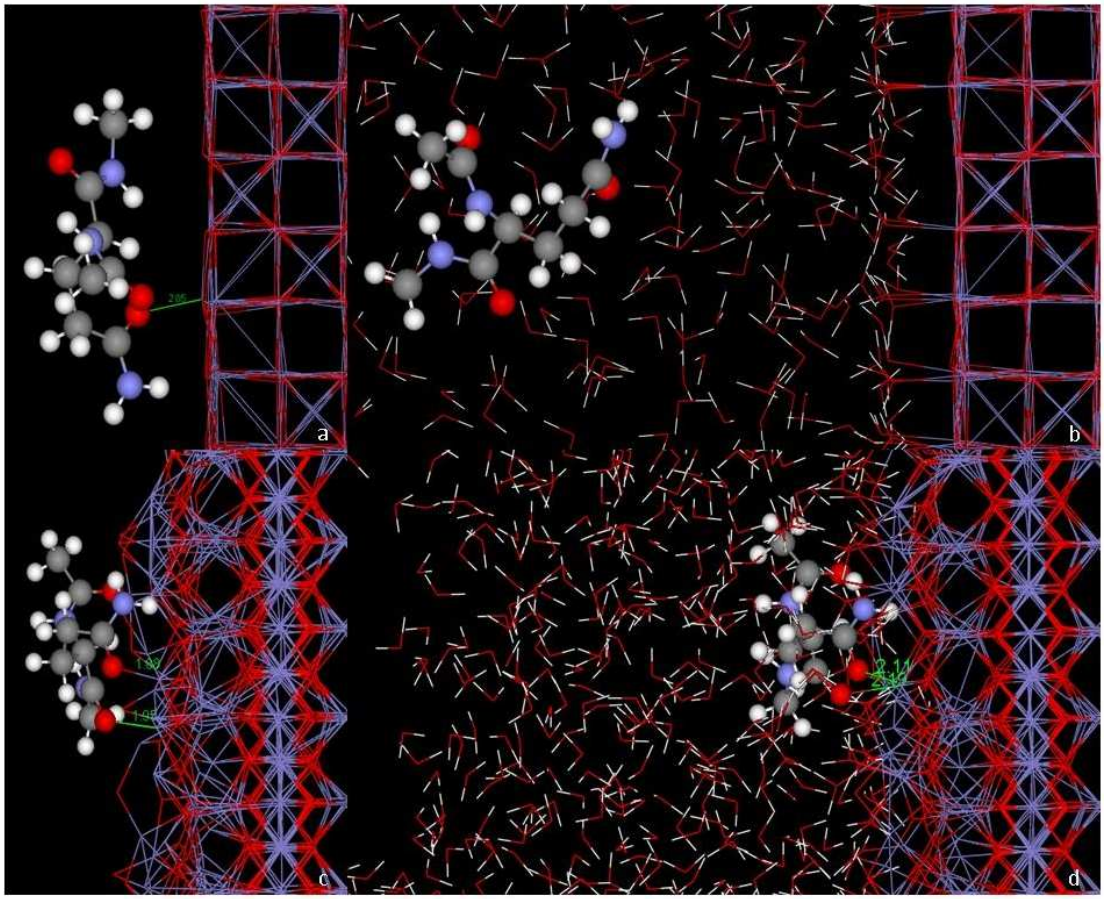

Figure 4.11: Bond distance images for the *in vacu* and solvated systems of the glutamine residue. a){100} *in vacu*, b){100} solvated, c){111} *in vacu* and d){111} solvated.

for the *in vacu* system. Upon solvation, the residence time of this bond drops to 407.28 ps, suggesting that this attachment is no longer beneficial. Octahedral iron bonding was also present to the N–terminal peptide bond oxygen lasting only 1100.01 ps in the *in vacu* system, but increasing in residence time to lasting over 99% of the simulation length for the solvated system. This data suggests that the addition of water effects how the amino acid oxygen bonds to magnetite, with side chain oxygen bonding being less favourable than water binding but N–terminal peptide bond oxygen bonding being more favourable.

The polar residence time data suggests that the amino acids favour binding to the {111} surface as small or no residence times are exhibited on the {100} surface, with the exception glutamine side chain oxygen–octahedral iron bonding in the *in vacu* system. Glutamine consistently shows a preference for N–terminal peptide bond oxygen bonding, whilst serine shows no preference.

On comparison of the polar amino acid residence time data with that of the non–polar it was shown that, the addition of oxygen to the residue side chains has minimal effect on the

iron binding, as the bonding lasting over 99% of the simulation length was through the N–terminal peptide bond oxygen as opposed to side chain oxygen, with the exception of the serine {111} *in vacu* system. For both amino acid groups, the {111} surface is the preferred for attachment. Additionally, little or no tetrahedral iron bonding was present, with the exception of the isoleucine {100} *in vacu* system, showing a preference for octahedral iron bonding.

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
GLN {100} surface	Q O	0.00	0.00	2087.49	0.99	0.00	0.00	21.99	0.10
	NBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
GLN {111} surface	Q O	0.00	0.00	1155.17	1.16	0.00	0.00	252.16	0.96
	NBO	0.00	0.00	4995.00	1.00	0.00	0.00	999.00	1.00
SER {100} surface	S O	0.00	0.00	7.10	0.10	0.00	0.00	8.62	0.07
	NBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
SER {111} surface	S O	0.00	0.00	4994.00	1.00	0.00	0.00	407.28	0.98
	NBO	0.00	0.00	1100.01	1.07	0.00	0.00	994.67	1.01

Table 4.9: Residence times (RT) and average coordination numbers (av. CN) of the polar amino acid residues. NBO is N–terminal peptide bond oxygen, Q O is glutamine side chain oxygen and S O is serine side chain oxygen.

The interfacial energies of the polar amino acid residues on the magnetite surfaces can be found in Table 4.10.

{100} surface attachment				
	E_{int} of s-a (eV)	E_{diff} of s-a (eV)	E_{int} of s-a-w (eV)	E_{diff} of s-a-w (eV)
GLN	-7.40	4.44	5.40	1.06
SER	-2.97	0.00	6.46	0.00
{111} surface attachment				
	E_{int} of s-a (eV)	E_{diff} of s-a (eV)	E_{int} of s-a-w (eV)	E_{diff} of s-a-w (eV)
GLN	-7.63	4.89	12.72	0.00
SER	-2.75	0.00	11.01	1.71

Table 4.10: Interfacial energies for all polar amino acid residues, using the Yang *et al* method[164]. E_{int} of s-a is the interfacial energy of the slab–amino acid residue system. E_{diff} of s-a is the difference in interfacial energy from the highest interfacial energy of the slab–amino acid residue system. E_{int} of s-a-w is the interfacial energy of the slab–amino acid residue–water system. E_{diff} of s-a-w is the difference in interfacial energy from the highest interfacial energy of the slab–amino acid residue–water system.

From the solvated polar interfacial energy data, it can be observed that the {100} surface produces the lowest interfacial energies (5.40 eV and 6.46 eV as opposed to 12.72 eV and 11.01 eV). On the {100} surface glutamine has the lowest interfacial energy value, whereas, on the {111} surface serine has the lowest, suggesting that the preferred residue of attachment is surface dependent. The addition of water to the system has a detrimental effect on the interfacial energy, as the *in vacu* system produces much lower energies. Before solvation very similar interfacial energy values were exhibited for the different surfaces (-7.40 eV and -2.97 eV for the {100}, and -7.63 eV and -2.75 eV for the {111}), however, very differing energy values were seen in the presence of water. This data suggests that the magnetite–water interactions are surface dependent and the presence of water in the {111} surface system is much more detrimental than for the {100}. Preferential water attachment could explain the increase in interfacial energy witnessed. Furthermore, system solvation affects the amino acid interfacial behaviours, as the solvated interfacial energy trend differs from the *in vacu* trend. Glutamine exhibits lower interfacial energy values than serine for the *in vacu* system, yet when water is present, glutamine continues to exhibit the lowest interfacial energy on the {100} surface, but the opposite is seen for the {111} surface.

On comparison of the polar amino acid interfacial energy data with that of the non–polar it was shown that, for the {100} surface solvated systems, serine produces the highest interfacial energy value, whereas glutamine produces the second lowest value (after valine),

suggesting that the presence of a hydroxyl oxygen is more detrimental to iron binding than having no oxygen, although its reaction with water may play a part in this. Conversely, the opposite is seen for serine in the solvated {111} system, as it exhibited the lowest interfacial energy value, suggesting that the hydroxyl oxygen presence is beneficial to iron binding.

Again, there were issues with this method of directly calculating the adsorption energy, with larger than expected energy values produced and expected trends not exhibited. Potential reasons for the problems displayed were discussed previously for the non–polar amino acids. Indirect Potential of Mean Force based methods of adsorption energy calculation are an alternative, and are considered in Section 3.3.5.

Acidic Amino Acid Attachment

R values for the acidic amino acids can be found in Table 4.11.

Amino Acid	Fe type	Surface	System	r (Å)
ASP	FET	100	*in vacu*	1.98
			solvated	1.98
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	-
			solvated	-
		111	*in vacu*	2.03
			solvated	2.08
GLU	FET	100	*in vacu*	2.08
			solvated	-
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.08
			solvated	2.23
		111	*in vacu*	2.03
			solvated	2.03

Table 4.11: Acidic amino acid r (Å) values from RDF data.

Figure 4.12 exhibits the RDF profiles for aspartate. The RDF profiles of the remaining

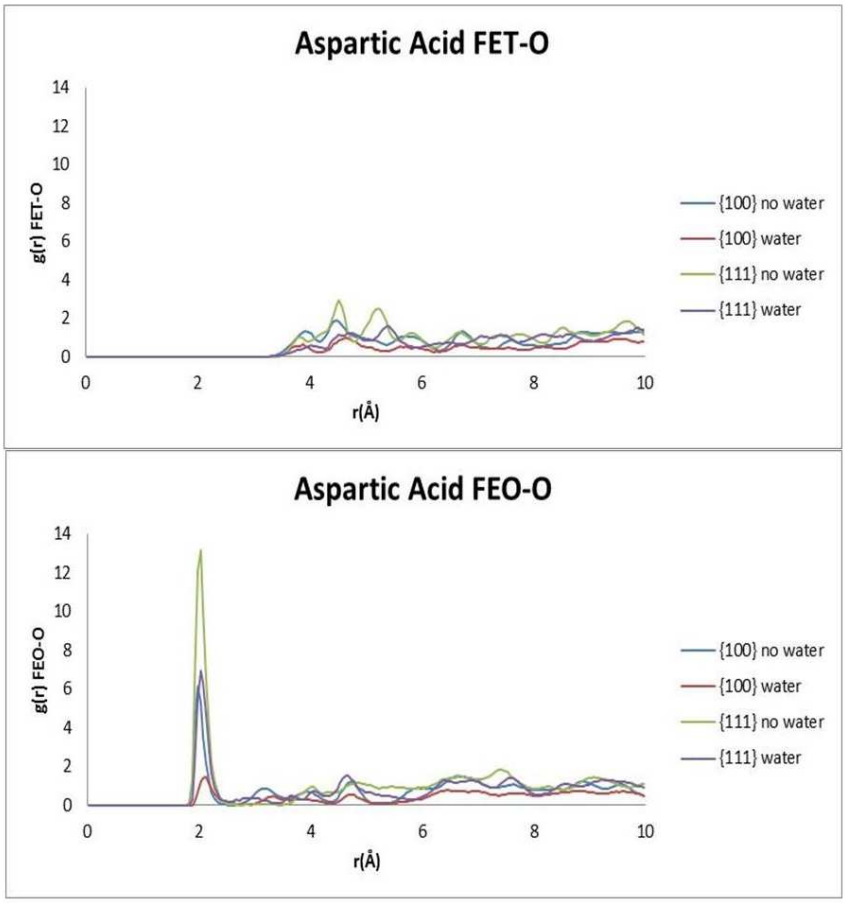

Figure 4.12: RDF plots for aspartate. FET is tetrahedral iron, FEO is octahedral iron. Blue is {100} *in vacu*. Red is {100} solvated. Green is {111} *in vacu*. Purple is {111} solvated.

acidic amino acids can be found in the Appendix. The acidic amino acid RDF data shows that for aspartate, all bonding is through octahedral iron, suggesting a definitive preference for octahedral iron binding. Whilst, for glutamate, the {111} surface exhibits only octahedral iron bonding, whereas, the {100} surface exhibits both tetrahedral and octahedral iron bonding suggesting no binding preference.

As with the non–polar and polar amino acids, the RDF data also shows that, the *in vacu* system produced a more intense $g(r)$ peak than the equivalent solvated system, suggesting that the presence of water was unfavourable for iron binding. Glutamate on the {111} solvated surface, having octahedral iron as its next nearest neighbour, demonstrates the greatest first peak intensity, indicating that the oxygen in glutamate are more freely accessible to the next nearest neighbouring octahedral iron, than in aspartate. This phenomenon is also true of the {100} versus the {111} surface, as the solvated {111} surfaces, bonding through octahedral iron, produced more intense $g(r)$ peaks than the corresponding {100} surfaces.

When the RDF data of the acidic amino acids was compared to that of the non–polar and polar amino acids it was shown that, the same bonding types were seen for glutamate as was exhibited for glutamine, leucine and methionine, indicating that the presence of the two oxygens in the side chain have no effect on the type of iron being bound. No other amino acid compared exhibited the same bonding type trend as aspartate.

The $FE-O_{aminoacid}$ bond distance data for the acidic amino acids is represented in Table 4.12.

{100} surface attachment			
Amino Acid	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
ASP	NBO	NB	NB
	D O	1.93	NB
	D O	1.99	NB
GLU	NBO	NB	NB
	E O	2.05	2.51
	E O	2.07	2.01
{111} surface attachment			
Amino Acid	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
ASP	NBO	1.97	NB
	D O	2.22	2.12
		2.03	2.13
	D O	1.94	1.93
		2.09	NB
GLU	NBO	1.99	NB
	E O	2.06	2.04
		NB	2.03
	E O	2.09	NB
		2.10	NB

Table 4.12: Acidic amino acid $FE-O_{aminoacid}$ bond distance data. BL is bond length, NBO is N–terminal peptide bond oxygen, D O is aspartate side chain oxygen, E O is glutamate side chain oxygen and NB is no bonding.

The acidic amino acid bond distance data shows more bonding present on the {111} *in vacu* surface than the {100}, as each amino acid exhibited two side chain oxygen bonds in the {100} systems, whereas, in the {111} systems, aspartate exhibited four side chain

oxygen bonds and glutamate exhibited three side chain oxygen bonds. N–terminal peptide bond oxygen bonding was only present for the {111} systems. Reduced bonding was seen for the solvated systems than the *in vacu*, as for aspartate, the {111} surface showed three bonds and the {100} showed no bonding, whilst glutamate exhibited the same amount of bonding on both surfaces (two bonds).

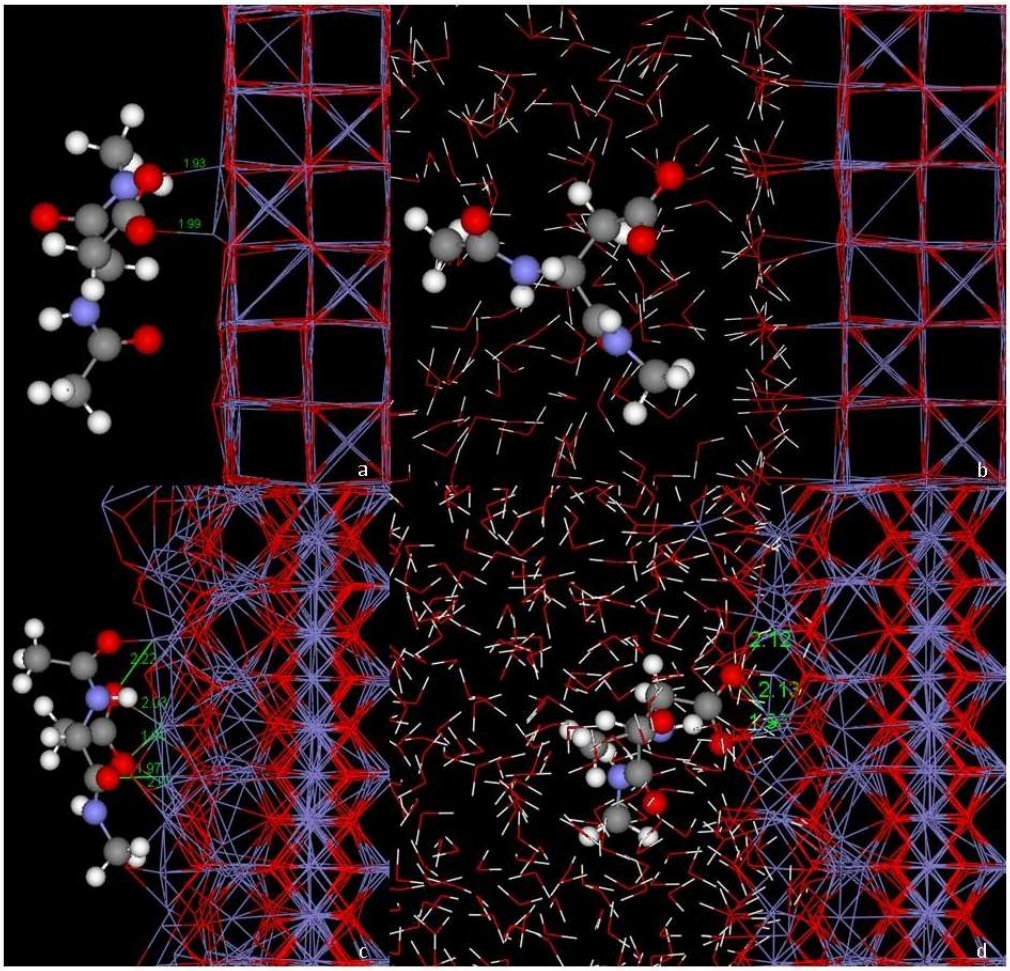

Figure 4.13: Bond distance images for the *in vacu* and solvated systems of the aspartate residue. a){100} *in vacu*, b){100} solvated, c){111} *in vacu* and d){111} solvated.

When the acidic amino acid bond distance data was compared to that of the non–polar and polar amino acids it was shown that, the presence of two side chain oxygen increases the amount of iron bonding, as for the {100} surface, the non–polar and polar amino acids exhibited either no bonding present or only one bond, and for the {111} surface, the non–polar and polar amino acids show no more than two bonds present. Additionally, this data shows that acidic amino acid bonding is predominantly through the side chain oxygen, whilst for the non–polar and polar it is primarily through the N–terminal peptide bond oxygen, with the exception of glutamine which appears to behave in a similar way

to the acidic amino acid residues, although this could be expected due it being the amide form of glutamate.

Table 4.13 represents the residence times and related average coordination numbers for bonding between all oxygen present in the acidic amino acid residues and tetrahedral and/or octahedral iron on both the {100} and {111} magnetite surfaces.

The acidic amino acid residence time data shows that, for aspartate, there was no tetrahedral iron bonding present. There was also no bonding lasting over 0.05% of the simulation length to the N–terminal peptide bond oxygen from the {100} surface. Aspartate in the {100} *in vacu* system, presented with octahedral iron bonding to both side chain oxygen lasting over 99% of the simulation length. On addition of water, the residence times of these bonds drop (74.22 ps and 66.06 ps) suggesting that attachment at this site is no longer favourable and water bonding is preferential. In the {111} surface *in vacu* system, octahedral iron bonding to one of the side chain oxygen of aspartate, and the N–terminal peptide bond oxygen was exhibited lasting over 98% of the simulation length. The other side chain oxygen of aspartate also showed octahedral iron bonding lasting 3455.92 ps. Upon solvation, the residence times of these bonds drop (766.44 ps, 29.34 ps and 870.81 ps, respectively) suggesting that water addition is detrimental to iron binding but to different extents dependent on which oxygen type is bonded.

For glutamate in the {100} *in vacu* system, bonding lasting over 99% of the simulation length was exhibited for both side chain oxygen (one to tetrahedral iron and one to octahedral iron). Upon solvation, the residence times of these bonds drop (181.56 ps and 143.98 ps). No N–terminal peptide bond oxygen bonding lasting over 2% of the simulation length was displayed. For the {111} surface *in vacu* system, no bonding lasting over 47% of the simulation length was revealed, however, with the addition of water into the system, again no bonding lasting nearly the full length of the simulation was exhibited, however, bonding between octahedral iron and one of the side chain oxygen lasting over 69% of the simulation length was seen, suggesting that water addition somewhat improves iron binding for glutamate on the {111} surface.

On comparison of the acidic amino acids, aspartate appears to be the preferred residue of attachment, particularly on the {111} surface. {111} surface attachment is also the preferred for attachment in the solvated systems. Aspartate exhibited only octahedral iron bonding, whereas glutamate, the {100} surface displayed both iron bonding types for the {100} surface, and only octahedral iron bonding for the {111} surface.

When the acidic amino acid residence time data was compared to that of the non–polar and polar amino acids it was shown that, the presence of two side chain oxygens improves iron binding, in that the bonding lasting over 99% of the simulation length was predominantly through the side chain oxygen. For the solvated systems, the {111} surface is the preferred surface of attachment. Also, there is little or no bonding through tetrahedral iron, showing a preference for octahedral iron bonding.

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
ASP {100} surface	1st D O	0.00	0.00	4989.00	1.00	0.00	0.00	74.22	0.48
	2nd D O	0.00	0.00	4984.15	1.00	0.00	0.00	66.06	0.53
	NBO	0.00	0.00	2.20	0.00	0.00	0.00	0.00	0.00
ASP {111} surface	1st D O	0.00	0.00	3455.92	1.91	0.00	0.00	870.81	1.85
	2nd D O	0.00	0.00	4905.18	1.94	0.00	0.00	766.44	1.22
	NBO	0.00	0.00	4998.00	1.00	0.00	0.00	29.34	0.18
GLU {100} surface	1st E O	4995.00	1.00	0.00	0.00	181.56	0.94	0.00	0.00
	2nd E O	0.00	0.00	4996.00	1.00	0.00	0.00	143.98	0.92
	NBO	0.00	0.00	60.88	0.04	0.00	0.00	0.00	0.00
GLU {111} surface	1st E O	0.00	0.00	891.48	1.40	0.00	0.00	298.41	1.77
	2nd E O	0.00	0.00	881.56	1.55	0.00	0.00	694.73	1.28
	NBO	0.00	0.00	2312.48	0.46	0.00	0.00	77.93	0.23

Table 4.13: Residence times (RT) and average coordination numbers (av. CN) of the acidic amino acid residues. (NBO = N–terminal peptide bond oxygen, D O = aspartate side chain oxygen and E O = glutamate side chain oxygen).

The interfacial energies of the acidic amino acid residues on the magnetite surfaces can be found in Table 4.14.

{100} surface attachment				
	E_{int} of s-a (eV)	E_{diff} of s-a (eV)	E_{int} of s-a-w (eV)	E_{diff} of s-a-w (eV)
ASP	-8.94	3.16	6.05	0.07
GLU	-5.78	0.00	6.12	0.00
{111} surface attachment				
	E_{int} of s-a (eV)	E_{diff} of s-a (eV)	E_{int} of s-a-w (eV)	E_{diff} of s-a-w (eV)
ASP	-6.00	0.06	21.19	0.00
GLU	-5.94	0.00	11.22	9.97

Table 4.14: Interfacial energies for all acidic amino acid residues, using the Yang *et al* method[164]. E_{int} of s-a is the interfacial energy of the slab–amino acid residue system. E_{diff} of s-a is the difference in interfacial energy from the highest interfacial energy of the slab–amino acid residue system. E_{int} of s-a-w is the interfacial energy of the slab–amino acid residue–water system. E_{diff} of s-a-w is the difference in interfacial energy from the highest interfacial energy of the slab–amino acid residue–water system.

From the solvated interfacial energy data it can be observed that, the {100} surface produces the lowest interfacial energies (6.05 eV and 6.12 eV as opposed to 21.19 eV and 11.22 eV). On the {100} surface, the acidic residues have similar interfacial energy values, although aspartate was marginally lower, whereas, on the {111} surface glutamate has the lowest. Again, the presence of water has a disadvantageous effect on the interfacial energy, as the energies are much lower for the *in vacu* system. Before solvation the interfacial energy values of glutamate are similar on the different surfaces, however, when the water is added, the different surfaces exhibit very different energy values, with the presence of water having a greater detrimental on the {111} surface (-5.94 eV *in vacu* to 11.22 eV solvated). This phenomenon is also apparent for aspartate, although there is an even greater difference in energy values between the solvated surfaces (15.14 eV), and preferential water attachment may be to blame. The solvation of the system also affects the way that the amino acid residues react at the interface, as the solvated systems interfacial energy trends are different to the *in vacu* trends. For the {100} surface *in vacu* system, there was a difference of over 3 eV between the acidic residue values, however when solvated the difference became only 0.7 eV, with aspartate displaying the lowest values. Whereas, with the {111} surface *in vacu* system, there was a difference of only 0.6 eV between the acidic residue values, with aspartate showing the lowest, then upon solvation, the difference between the values was significantly greater with a variance of

9.97 eV with glutamate now presenting the lowest energy value. Thus, from this data, aspartate appears to be the preferred acidic amino acid residue of attachment.

On comparison of the acidic amino acid interfacial energy data with that of the non–polar and polar amino acids it can be seen that, for the {100} solvated systems, the acidic residues produce midrange interfacial energy values, this suggesting that the presence of the two oxygen on each of their side chains does not improve iron binding, and in some instances, is more detrimental to iron binding than having a purely hydrocarbon side chain, although the reaction of water with the residues or the magnetite surfaces may play a part in this. For the {111} solvated systems, the acidic amino acids show very different trends in the interfacial energy value, as glutamate exhibits the third lowest energy value whilst aspartate exhibits the highest energy value with a gap of 6.74 eV to the next energy value. Thus, for glutamate the presence of the two side chain oxygen is beneficial to iron binding, however, for aspartate it is extremely detrimental, although this effect could be due to a combination of factors involving the interactions of the individual system components.

There are, as with the non–polar and polar amino acids, concerns with this direct method of calculating the adsorption energy, possible reasons for these were discussed previously for the non–polar amino acids. An alternative indirect Potential of Mean Force based method of adsorption energy calculation is considered in Section 3.3.5.

Basic Amino Acid Attachment

R values for the basic amino acids can be found in Table 4.15.

Figure 4.14 exhibits the RDF profiles for lysine. The RDF profiles of the remaining basic amino acids can be found in the Appendix. From the basic amino acid RDF data it can be seen that, for arginine, only tetrahedral iron bonding is displayed on the {100} surface and only octahedral iron bonding is displayed on the {111} surface. With lysine, the {111} surface, again, exhibits only octahedral iron bonding, conversely, the {100} surface exhibits both tetrahedral and octahedral iron bonding. These data suggest that for arginine and the lysine {100} surface there is a definite iron binding type preference, whereas, the lysine {111} surface shows no such preference.

Additionally, from the basic amino acid RDF data it can be derived that, as with the other amino acid groups, the *in vacu* system produces a stronger $g(r)$ peak than the complementary solvated system, and that the largest $g(r)$ peak is displayed for lysine on

Amino Acid	Fe type	Surface	System	r (Å)
ARG	FET	100	*in vacu*	1.98
			solvated	1.98
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	-
			solvated	-
		111	*in vacu*	2.03
			solvated	2.08
LYS	FET	100	*in vacu*	2.08
			solvated	-
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.08
			solvated	2.23
		111	*in vacu*	2.03
			solvated	2.03

Table 4.15: Basic amino acid r (Å) values from RDF data.

the {111} surface having octahedral iron as its next nearest neighbour. This reason for this may be that the water is being attracted to the positively charged group of the side chain rather than the magnetite surface, freeing up the interaction between the magnetite surface and the amino acid oxygen.

When the basic amino acid RDF data was compared to that of the non–polar, polar and acidic amino acids it was shown that, the same type of bonding was seen for lysine as was exhibited for glutamate, glutamine, leucine and methionine. This suggests that the presence of the ϵ-amino group in the side chain has no effect on the type of iron being bound. No other amino acid compared exhibited the same bonding type trend as aspartate.

The $FE - O_{aminoacid}$ bond distance data for the basic amino acids is represented in Table 4.16.

When the basic amino acid bond distance data was examined it was shown that, more bonding was present in the {111} systems than the {100}. No bonding was exhibited on

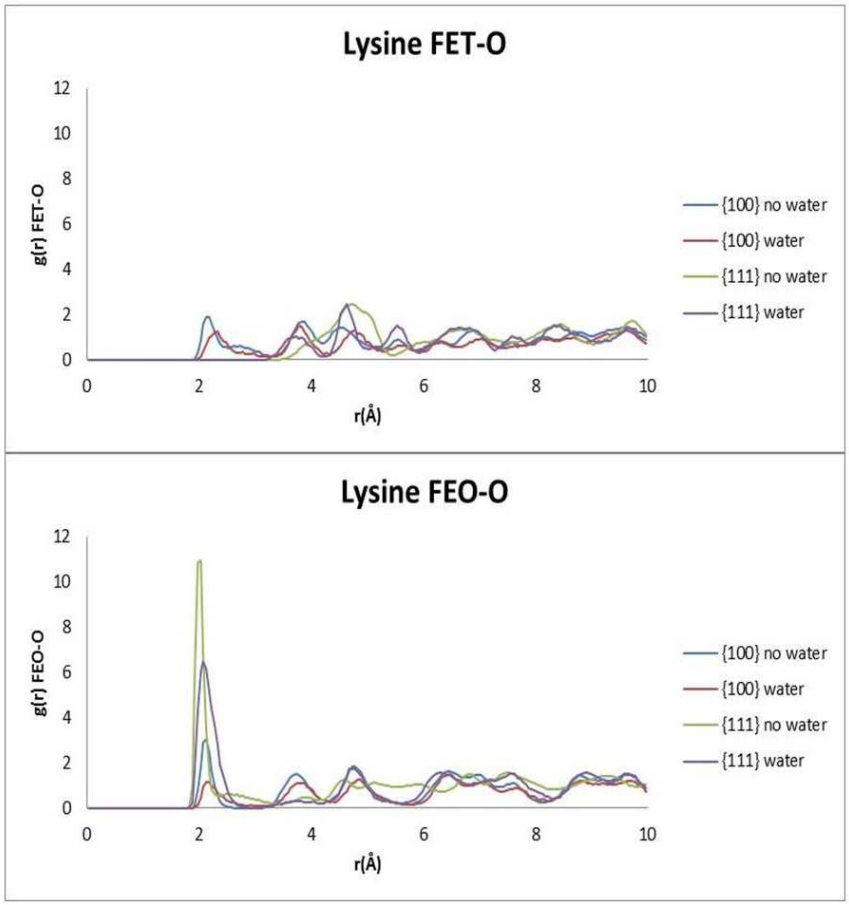

Figure 4.14: RDF plots for lysine. FET is tetrahedral iron, FEO is octahedral iron. Blue is {100} *in vacu*. Red is {100} solvated. Green is {111} *in vacu*. Purple is {111} solvated.

the {100} surface. Bonding was displayed to the N–terminal peptide bond oxygens in the {111} surface *in vacu* systems, however upon solvation, no bonding was present for lysine, thus the addition of water is detrimental for iron binding in lysine {111} system.

When the basic amino acid bond distance data was compared to that of the non–polar, polar and acidic amino acids it was shown that, the presence of the positively charged groups in the side chains has minimal effect on {100} surface iron binding, as there was no bonding present, also displayed for alanine, methionine, valine and serine. With the {111} surface, no increased iron binding effect was seen by the presence of the guanidinium group in arginine as, like with leucine, valine and serine, there is one bond present. {111} surface lysine, however, exhibits one bond for the *in vacu* system and no bonding for the solvated system, as see for alanine and isoleucine.

Table 4.17 represents the residence times and related average coordination numbers for bonding between all oxygen present in the basic amino acid residues and tetrahedral and/or octahedral iron on both the {100} and {111} magnetite surfaces.

{100} surface attachment			
Amino Acid	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
ARG	NBO	NB	NB
LYS	NBO	NB	NB
{111} surface attachment			
Amino Acid	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
ARG	NBO	2.02	2.06
LYS	NBO	2.01	NB

Table 4.16: Basic amino acid $FE-O_{aminoacid}$ bond distance data. BL is bond length, NBO is N–terminal peptide bond oxygen, and NB is no bonding.

The basic amino acid residence time data shows that, for the {100} surface, arginine exhibits no bonding lasting over 0.03% of the simulation length and lysine exhibits no bonding lasting over 4% of the simulation length. With the {111} surface, there was no tetrahedral bonding present, however, arginine showed octahedral iron bonding lasting over 86% of the *in vacu* simulation length, and over 99% of the solvated simulation length. Whilst, with lysine in the {111} *in vacu* system, octahedral iron bonding lasting over 99% of the solvated simulation length was demonstrated, however, upon solvation, the residence times of this bond drops 333.43 ps, suggesting that water addition is detrimental to iron binding.

On comparison of the basic amino acid residues, the same amount of bonding lasting over 86% of the simulation length was present for the *in vacu* systems, however, when solvated, arginine shows the same amount of this length of bonding as the *in vacu* system, whilst lysine shows none, suggesting that arginine {111} surface attachment is not effected by the addition of water, whilst it has an unfavourable effect on lysine binding.

When the basic amino acid residence time data was compared to that of the non–polar, polar and acidic amino acids it was shown that, the {111} surface is the preferred surface of attachment for all amino acid groups, particularly for the solvated systems. Also, there is little or no tetrahedral iron bonding for any of the amino acid groups, showing a preference for octahedral iron bonding. Arginine exhibits the same amount and type of long term bonding as leucine and glutamine, whilst lysine exhibits the same amount and type of long term bonding as isoleucine and valine.

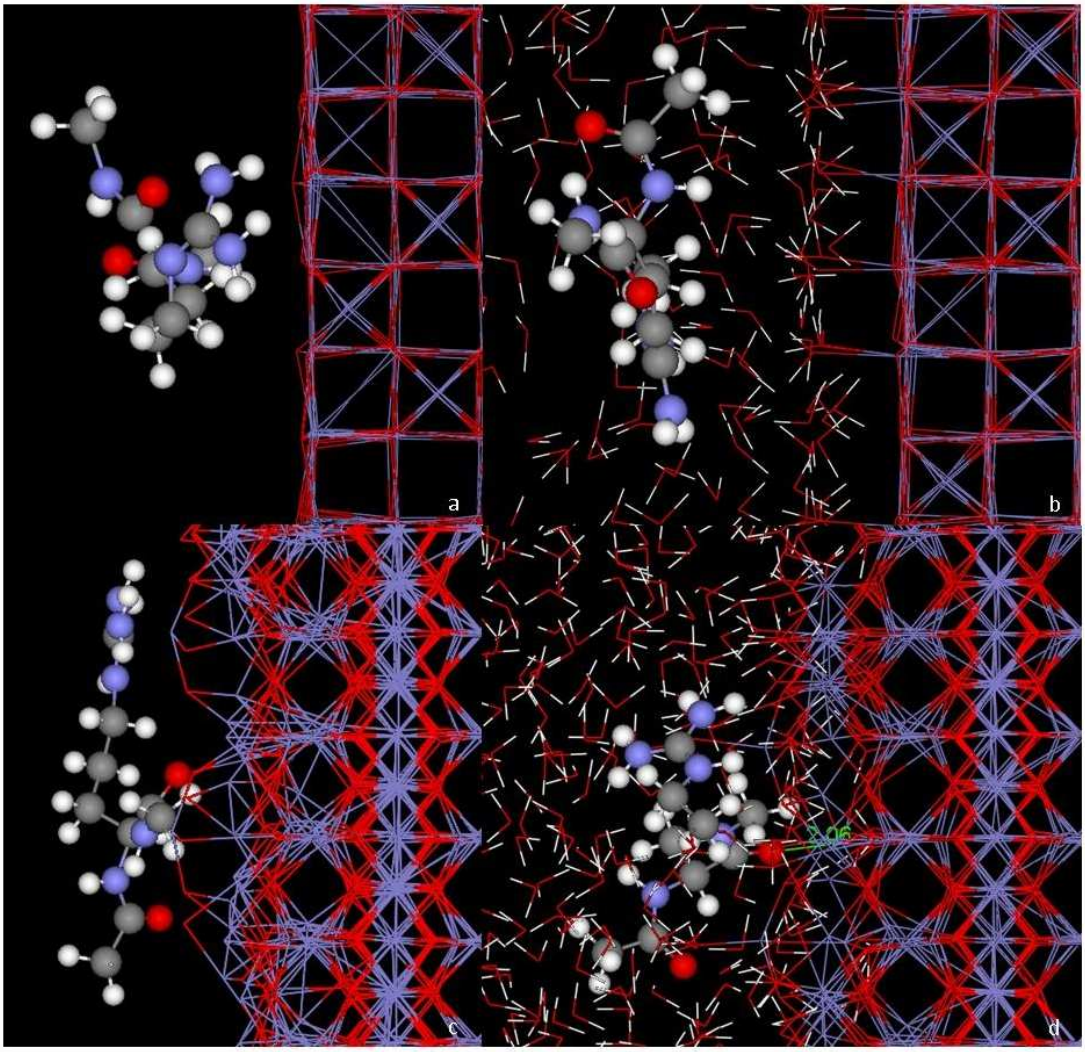

Figure 4.15: Bond distance images for the *in vacu* and solvated systems of the arginine residue. a){100} *in vacu*, b){100} solvated, c){111} *in vacu* and d){111} solvated.

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
ARG {100} surface	NBO	0.00	0.00	1.16	0.01	0.00	0.00	0.00	0.00
ARG {111} surface	NBO	0.00	0.00	4330.20	0.98	0.00	0.00	999.00	1.00
LYS {100} surface	NBO	22.54	0.43	0.00	0.00	35.27	0.32	9.88	0.05
LYS {111} surface	NBO	0.00	0.00	4995.00	1.00	0.00	0.00	333.43	0.64

Table 4.17: Residence times (RT) and average coordination numbers (av. CN) of the basic amino acid residues. (NBO = N–terminal peptide bond oxygen).

The interfacial energies of the acidic amino acid residues on the magnetite surfaces can be found in Table 4.18.

{100} surface attachment				
	E_{int} of s-a (eV)	E_{diff} of s-a (eV)	E_{int} of s-a-w (eV)	E_{diff} of s-a-w (eV)
ARG	-16.20	14.81	6.23	0.00
LYS	-1.39	0.00	4.18	2.05
{111} surface attachment				
	E_{int} of s-a (eV)	E_{diff} of s-a (eV)	E_{int} of s-a-w (eV)	E_{diff} of s-a-w (eV)
ARG	-14.11	11.62	9.81	0.00
LYS	-2.49	0.00	8.59	1.22

Table 4.18: Interfacial energies for all basic amino acid residues, using the Yang *et al* method[164]. E_{int} of s-a is the interfacial energy of the slab–amino acid residue system. E_{diff} of s-a is the difference in interfacial energy from the highest interfacial energy of the slab–amino acid residue system. E_{int} of s-a-w is the interfacial energy of the slab–amino acid residue–water system. E_{diff} of s-a-w is the difference in interfacial energy from the highest interfacial energy of the slab–amino acid residue–water system.

From the solvated basic amino acid interfacial energy data, it can be seen that, as with the amino acid groups, the {100} surface produces the lowest interfacial energy values, with the {100} surface showing values of 6.23 eV and 4.18 eV for arginine and lysine respectively, whilst the {111} surface showed values of 9.81 eV and 8.59 eV for arginine and lysine respectively. Lysine exhibits the lowest interfacial energy values, suggesting this is the preferred amino acid of attachment. This data also shows that, the addition of water to the system has a detrimental effect on the interfacial energy, as the energies calculated are much lower for the *in vacu* system. The solvation of the system also affects the interfacial behaviour of the amino acids, as the interfacial energy trends for the *in vacu* systems are very different to the solvated systems. For the *in vacu* systems, the lowest energy was exhibited for arginine; however, with the addition of water the opposite is true.

On comparison of the basic amino acid interfacial energy data with the non–polar, polar and acidic amino acid residue data it can be seen that, for the {100} surface solvated systems, the basic residues produce the two extremes of interfacial energy values, with arginine producing one of the highest interfacial energy values and lysine producing the lowest interfacial energy value; this suggests that the presence of a primary amine group in the side chain of lysine is beneficial to iron binding whilst the presence of the guanidinium group in arginine is greatly detrimental to iron binding for this surface. For the {111} surface solvated system, a very different basic amino acid trend was shown, as arginine

and lysine exhibit the second lowest and lowest interfacial energy values, respectively; this suggests that the presence positively charged groups in the side chains is beneficial to iron binding.

As with the other amino acid groups, there are concerns with this direct method of calculating the adsorption energy, possible reasons for these were discussed previously for the non–polar amino acids. An alternative indirect method of adsorption energy calculation is explored in the following section. This method utilises a constrained molecular dynamics version of the systems, focusing on the Potential of Mean Force and free energy evaluation.

4.3 Constrained system

4.3.1 Computational Methods

As with the unconstrained system, the initial amino acid structures, including terminal end capping, were produced with AMBERTOOLS TLEAP. The DL_POLY code was used for all calculations, and the TIP3P/fs potential model for water was utilised. The ensembles, potentials and parameters used are described in Chapters 2 and 3.

The iron binding affinity of the peptide chains was determined by running a series of Potential of Mean Force (PMF)[251] simulations, with the distance between the amino acids centre of mass (CoM) and the magnetite surface constrained in the direction perpendicular to the surface to distances between 0 Å and 10 Å, but free to move parallel to the surface. The amino acid chiral carbon is, for the purposes of this study, the centre of mass. The additional force exerted on the simulation due to this constraint was monitored. Each PMF calculation was run for 1 ns NVT at 300 K and the average force was integrated with respect to the constraint distance to produce the free energy of binding[139].

4.3.2 Results and Discussion

Free energy profiles were created for the aforementioned systems as shown in Figures 4.16, 4.20, 4.22 and 4.24. Within the free energy profiles, 0 Å on the x–axis is an approximate representation of where the magnetite surface lies based on the first strongest peak for an FEO iron ion in the related density profile. The free energy of the simulation systems as a function of the centre of mass distance from the magnetite surface are shown in the free energy profiles. Peaks in the water density profiles indicate areas within the system

containing high water molecule concentrations.

Non–polar Amino Acid Attachment

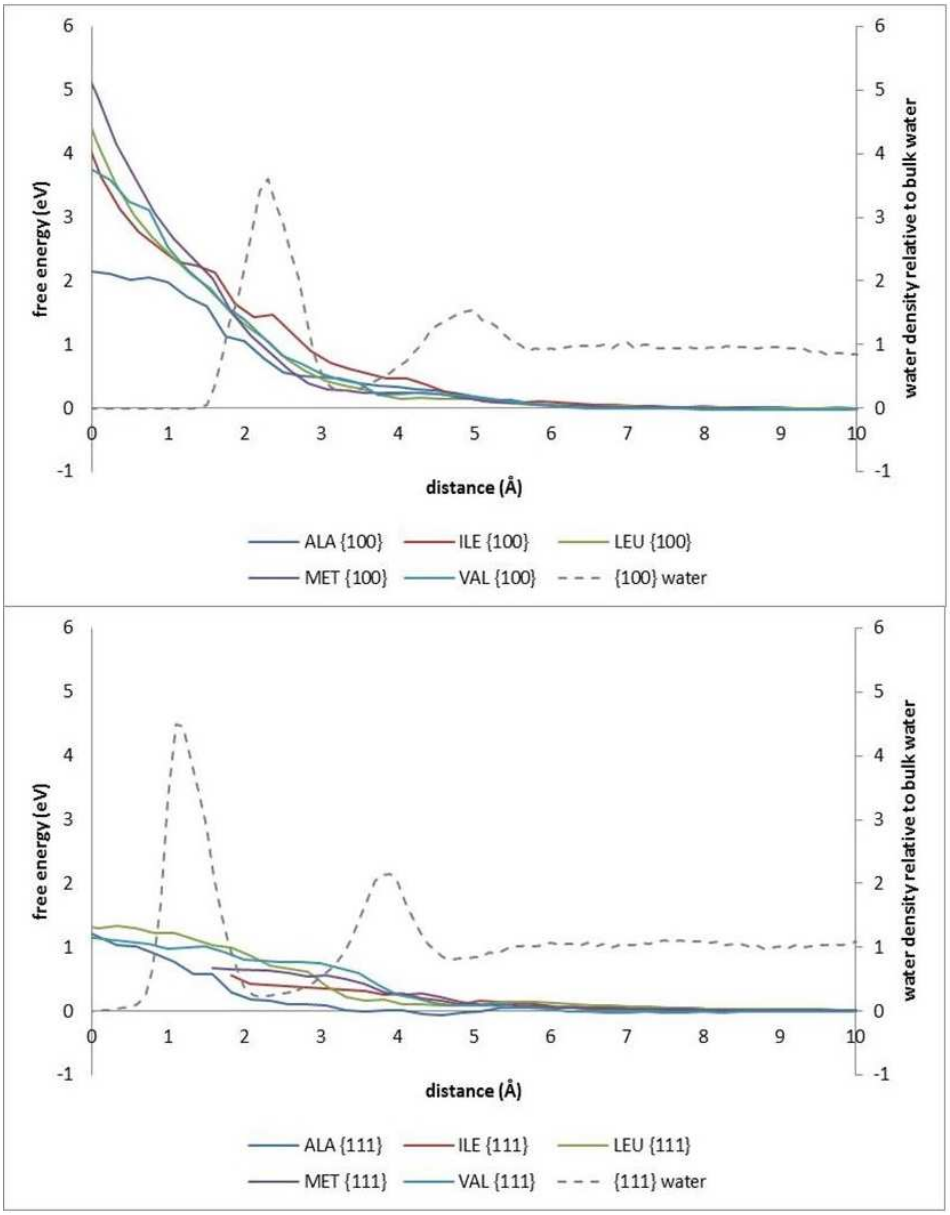

Figure 4.16: A comparison of the free energy profiles of the non–polar amino acid residues. Distance refers to distance between the peptide CoM and the magnetite surface. The dashed line represents the water density profile for the system.

From the non–polar amino acid {100} surface free energy profile results (Figure 4.16) it is evident that, alanine has the lowest free energy value of 2.15 eV when the centre of mass was at the surface, suggesting preferential iron binding over the other longer hydrocarbon chain and branched hydrocarbon chain amino acid side groups. Methionine exhibited the highest free energy value of 4.92 eV when the centre of mass was at the surface, with a

difference of 0.7 eV to the next amino acid (leucine) in energy, suggesting that the presence of sulphur in the side chain is detrimental to iron binding. When the amino acid residues centre of mass is between 10 Å and 5.5 Å the amino acid energies are comparable, however, as the amino acid centre of masses get closer to the surface the free energies increase, with the increases in free energy corresponding with the distance at which strong water adsorption layers occur. The free energy trend is similar for the non–polar amino acids, with the exception of alanine, which produced slightly lower free energy values, thus indicating a greater iron binding affinity. Optimum iron binding was seen between 10 Å and 6 Å. We propose that the successive increase in free energy is attributable to a requirement for more energy within the system, for the amino acids to pass through the strong water barriers, and to enable amino acid attachment as they move closer towards the surface.

From the non–polar amino acid {111} surface data, it was found that the centre of mass could not be put any closer than 2 Å and 1.5 Å for isoleucine and methionine respectively, when they were moved closer amino acid and magnetite surface bond breaking occurred (Figures 4.18 and 4.19). As can be seen from the isoleucine images, the amino acid entered the magnetite surface causing large scale surface dislocation, amino acid disintegration and water vacuum creation. This behaviour could be due to a combination of factors including: the configuration of the side chain; the reaction of the hydrophobic hydrocarbon side chain with the water; or the reaction of the water with the magnetite (competitive binding). The same was true of methionine, as seen in the images, although an additional factor that may cause this behaviour could be the reaction of sulphur within the system.

A comparison of the non–polar amino acid {111} surface free energy profiles show that, alanine, leucine and valine have very similar energies when their CoM reached the surface (1.17 eV, 1.31 eV and 1.19 eV respectively). Between 10 Å and 5.5 Å, the same trend is exhibited as for the {100} surface, and as the distance gets closer to the surface a gradual increase in free energies, with the exception of alanine. For alanine, when the centre of mass reaches 5.5 Å away from the surface its free energy begins to decrease, with its lowest energy reaching -0.06 eV, after which the free energy increases again. The increase in free energy corresponds with the distance at which the weaker water absorption layer occurs, with the stronger layer appearing to exert minimal influence on the free energy profiles. Again, the highest free energies were found to be at the distance closest to the {100} surface, and equally, the lowest energies were exhibited when the constrained coordinate is moved away from the magnetite surface, further suggesting that amino acid sequence

attachment is more favourable at greater distances, particularly for alanine which had the lowest non–polar amino acid free energy (CoM at 4.6 Å).

Upon examination of the energy minima for the non-polar amino acids (Figure 4.17) it can be observed that, on the {100} surface, all of the non-polar amino acid residues except for alanine exhibited a negative value for their energy minima, with valine producing the lowest value at -0.0312 eV. The energy minima for these residues fall between 8.7 Å and 9.6 Å, suggesting that attachment of the residue centre of mass at this distance would be preferable. On the {111} surface, fewer of the non polar amino acid residues exhibited a negative value for their energy minima, with alanine showing the lowest value at -0.0614 eV. The trend of favourable attachment distance is also different from the {100} surface, with alanine, isoleucine and valine exhibiting centre of mass distances of 4.6 Å , 7.5 Å and 10 Å respectively.

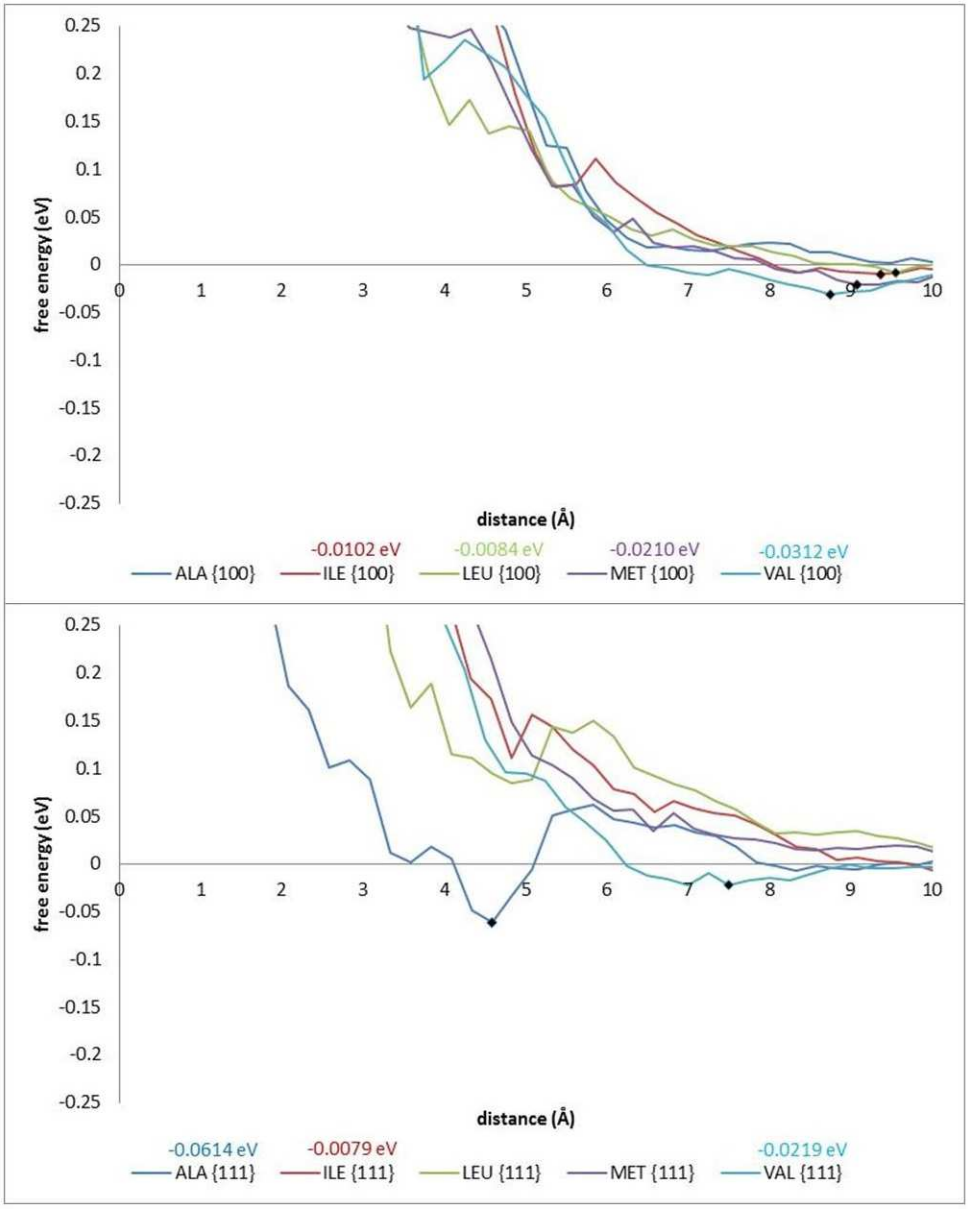

Figure 4.17: A close-up comparison of the free energy profiles of the non–polar amino acid residues between -0.25 eV and 0.25 eV, exhibiting the free energy minima.

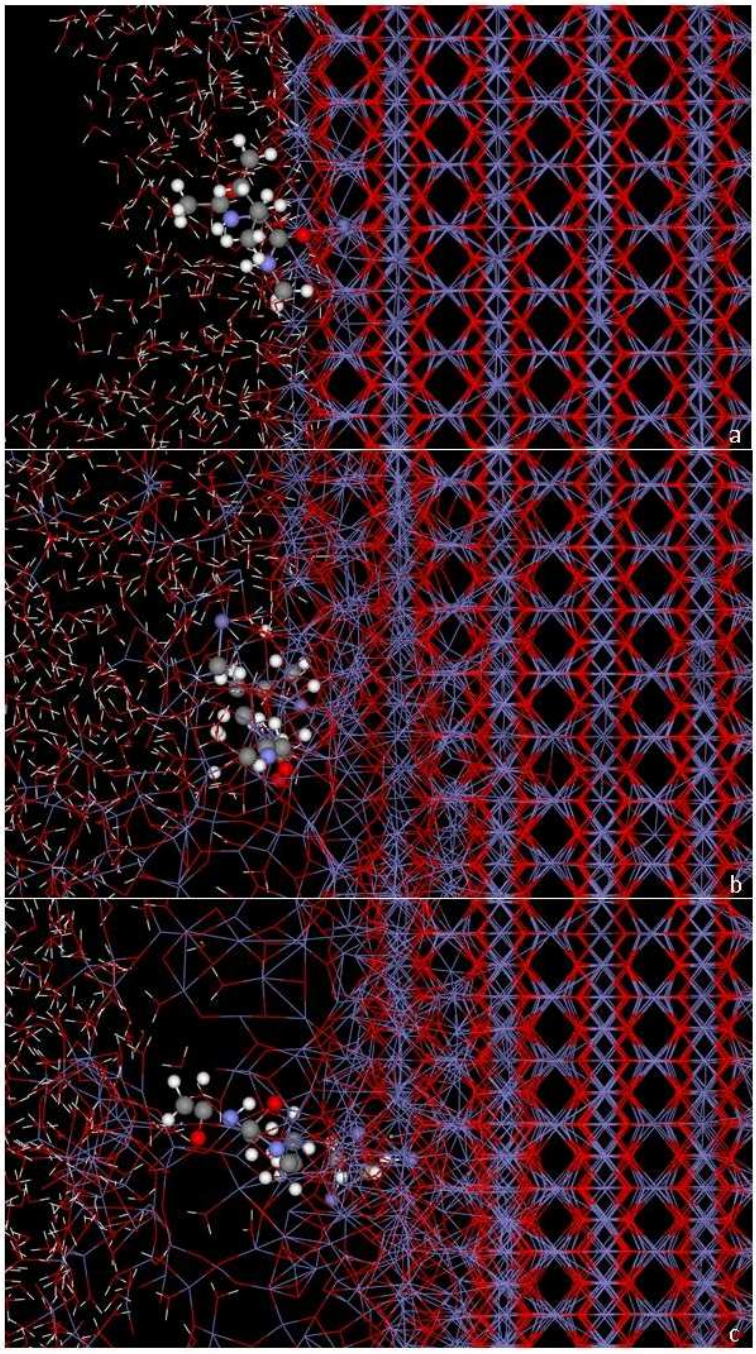

Figure 4.18: A comparison of the final simulation coordinates for ILE CoM at (a) 2 Å from the magnetite surface and then (b) 1.75 Å and (c) 1.5 Å exhibiting bond breaking.

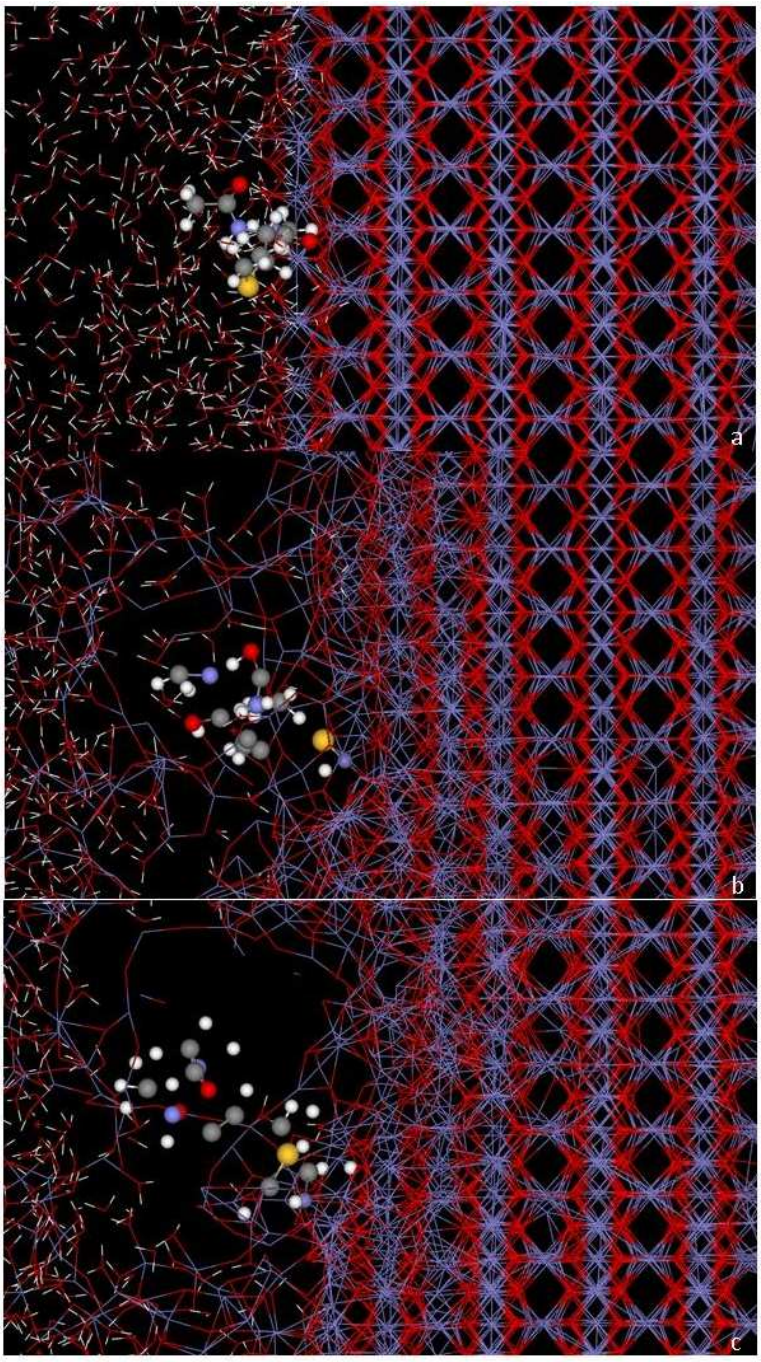

Figure 4.19: A comparison of the final simulation coordinates for MET CoM at (a) 1.5 Å from the magnetite surface and then (b) 1.25 Å and (c) 1 Å exhibiting bond breaking.

Polar Amino Acid Attachment

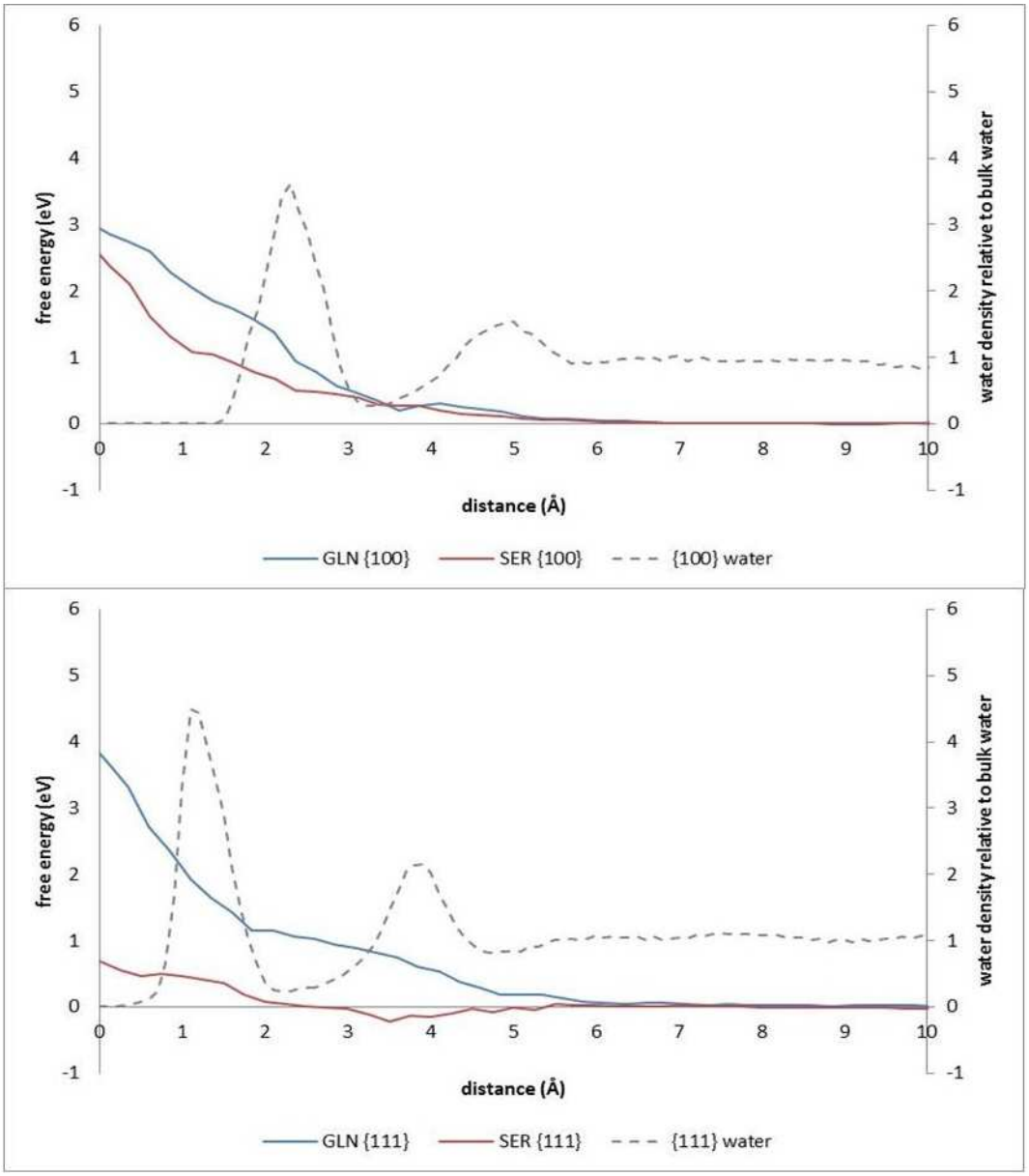

Figure 4.20: A comparison of the free energy profiles of the polar amino acid residues. Distance refers to distance between the peptide CoM and the magnetite surface. The dashed line represents the water density profile for the system.

The polar amino acid {100} surface free energy profile results (Figure 4.20) show that, glutamine and serine exhibit comparable iron binding affinity, between 10 Å and 3 Å, with optimal binding seen between 10 Å and 6.5 Å. An initial increase in free energy is observed at 6 Å, which corresponds with the distance at which the weaker water absorption layer occurs. A further increase in energy is observed at 3 Å, which corresponds with the stronger water absorption layer occurs. However, at distances less than 3 Å, serine exhibits slightly lower free energy values, thus indicating a higher iron binding activity. The increase in

free energies identified could be due to either the need for more energy to allow the amino acid sequences to pass through the water barrier or for them to overcome the unfavourable attachment as they approach the magnetite surface, or a combination of the two.

Conversely, on the {111} surface, a large difference in free energy at the surface is observed. Serine has the lowest free energy value of 0.70 eV, whilst glutamine has an energy value of 3.69 eV. Comparable iron binding affinity is exhibited between 10 Å and 6 Å, however, as the distance to the surface decreases, the amino acid behaviours differ. Serine has a slight decrease in free energy until 3.5 Å reaching a minimum of -0.22 eV, subsequently, the free energy increases steadily until it reaches the surface. This initial decrease in free energy indicates a distance for preferable attachment. As the glutamine centre of mass moves in towards the surface from 6 Å, its free energy increases.

Upon examination of the energy minima for the polar amino acids (Figure 4.21) it can be observed that, on the {100} surface, only glutamine exhibited a negative value for its energy minima, producing a value of -0.0078 eV at 9.4 Å, suggesting that attachment of the residue centre of mass at this distance would be preferable. On the {111} surface, only serine exhibited a negative value for its energy minima, producing a value of -0.2227 eV at 3.5 Å, suggesting that attachment of the residue centre of mass at this distance would be preferable.

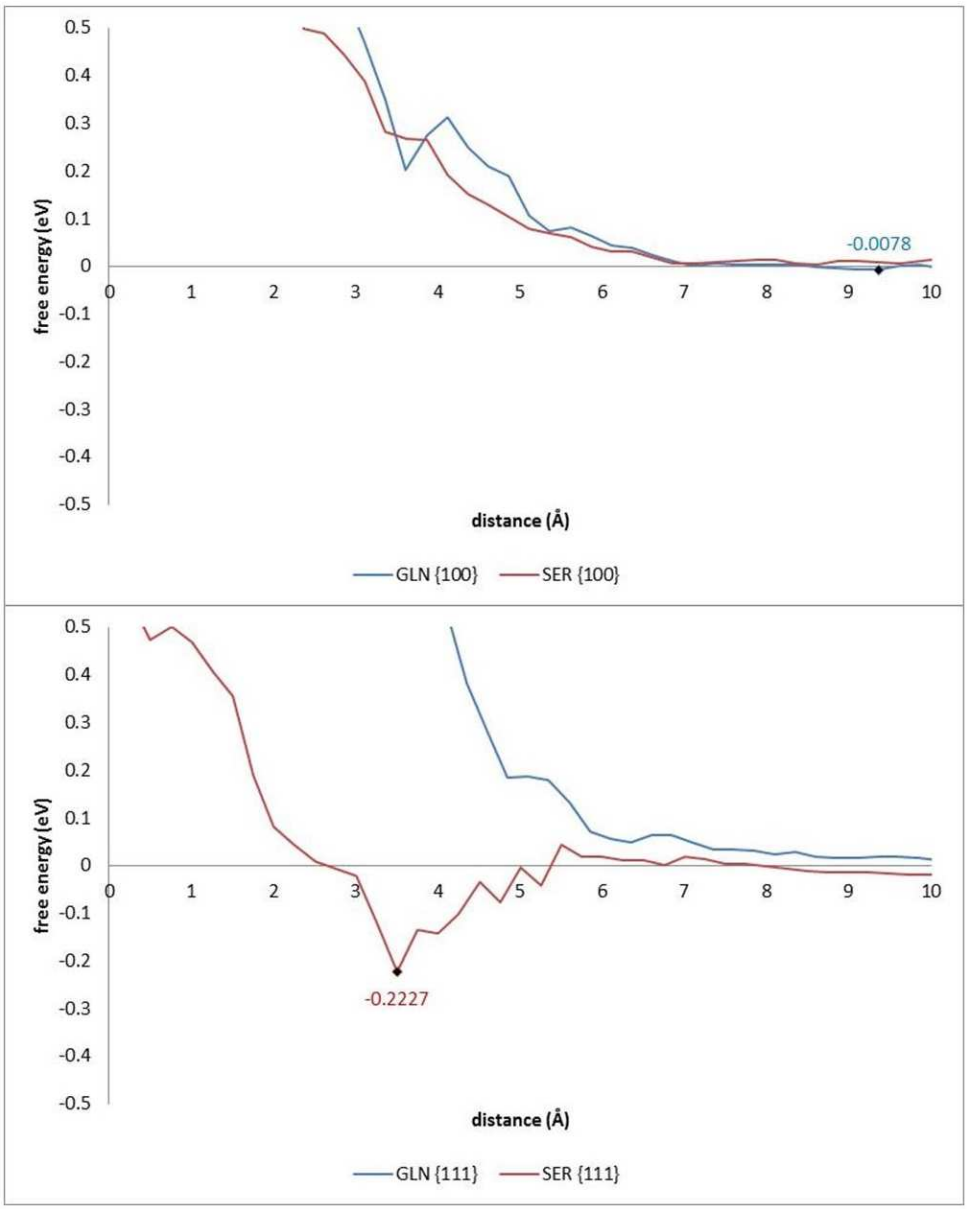

Figure 4.21: A close-up comparison of the free energy profiles of the polar amino acid residues between -0.5 eV and 0.5 eV, exhibiting the free energy minima.

Acidic Amino Acid Attachment

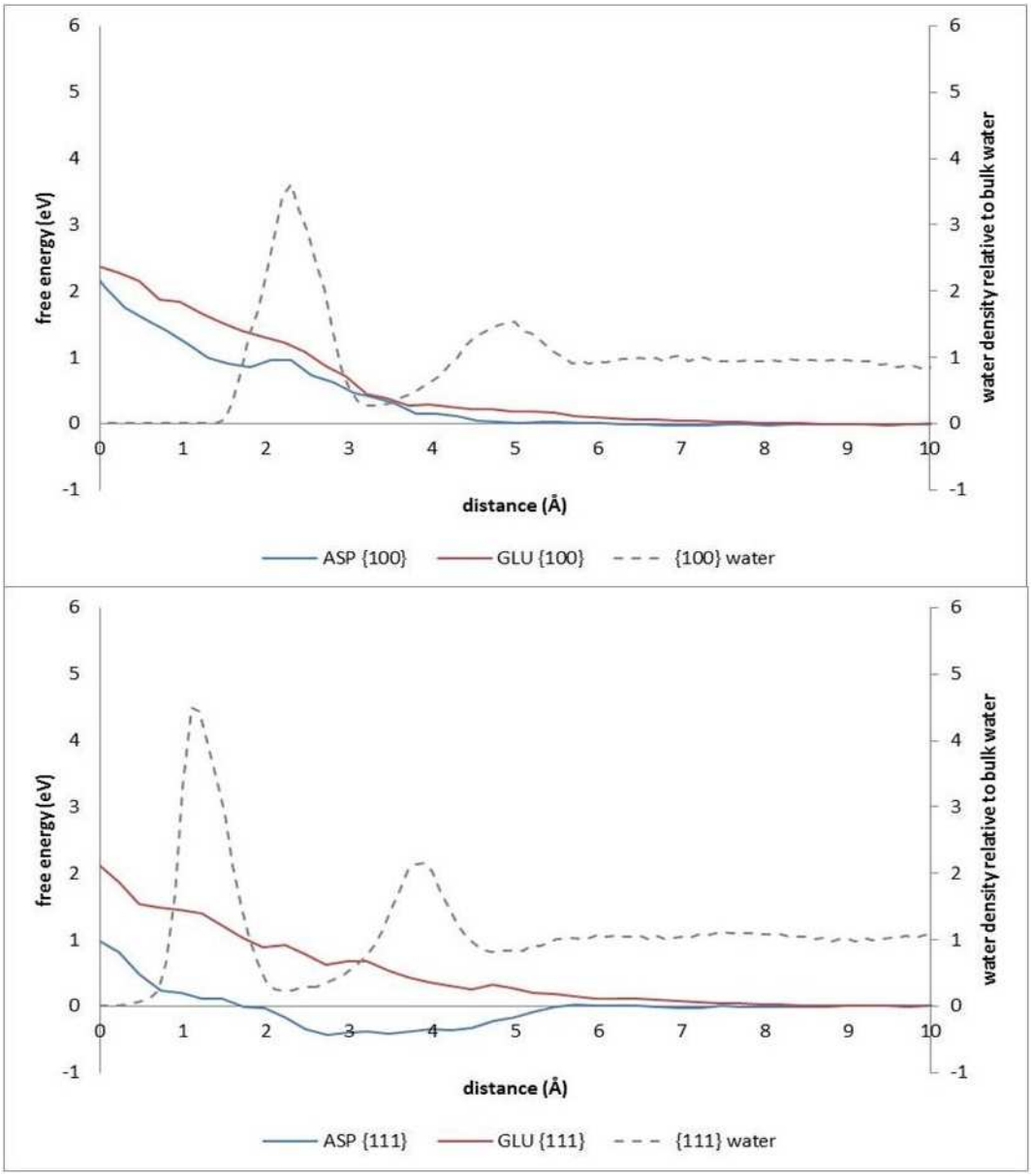

Figure 4.22: A comparison of the free energy profiles of the acidic amino acid residues. Distance refers to distance between the peptide CoM and the magnetite surface. The dashed line represents the water density profile for the system.

On comparison of the acidic amino acid free energy profile data for the {100} surface (Figure 4.22) it was evident that, the profiles were similar throughout, with aspartate having a slightly lower free energy values. The free energy profile of aspartate remained at approximately 0 eV from 10 Å to 4.5 Å, whereupon, it showed an initial increase in free energy, peaking at the distance corresponding with the stronger of the water barriers. Subsequently, a slight decrease in energy was seen until 1.5 Å, after which, the energy increases again. This data suggests that energy was needed in the system for aspartate

to pass through the stronger water barrier, after which point it became slightly easier for attachment until the magnetite slab began to have an influence on the profile. For glutamate, the free energy remains consistently at approximately 0 eV from 10 Å to 7.5 Å, at which point, an initial increase in free energy is shown. A further, greater increase in energy is observed at 1.85 Å, which corresponds with the stronger water absorption layer, suggesting that the presence of water layering has a great influence on amino acid iron binding.

When the acidic amino acids {111} surface free energy profiles were examined, a large difference in the free energy profiles was witnessed. As with the {100} surface, aspartate exhibited the lowest free energy profile. When the amino acid residues centre of mass was between 10 Å and 6 Å the energies stayed at around 0 eV, however, as the CoM got closer to the surface, glutamate presented with a gradual increase in free energy as its centre of mass moved towards the surface. Whilst, aspartate exhibited a decrease in free energy until 2.7 Å , reaching its energy minimum of -0.44 eV, whereupon, the free energy again increased steadily.

Upon examination of the energy minima for the acidic amino acids (Figure 4.23) it can be observed that, on the {100} surface, both acidic amino acid residues exhibited a negative value for their energy minima, with aspartate producing the lowest value at -0.0312 eV. The energy minima for these residues fall 2.5 Å apart, suggesting that preferable attachment of the residue centre of mass is amino acid dependent. On the {111} surface, again, both acidic amino acid residues exhibited a negative value for their energy minima, with aspartate producing the lowest value at -0.4361 eV. The energy minima for these residues fall 7 Å apart, suggesting that preferable attachment of the residue centre of mass is again amino acid dependent. This data suggests that independent of surface type, aspartate consistently exhibited the lowest energy minima value, however, the {111} surface produced a considerable lower value.

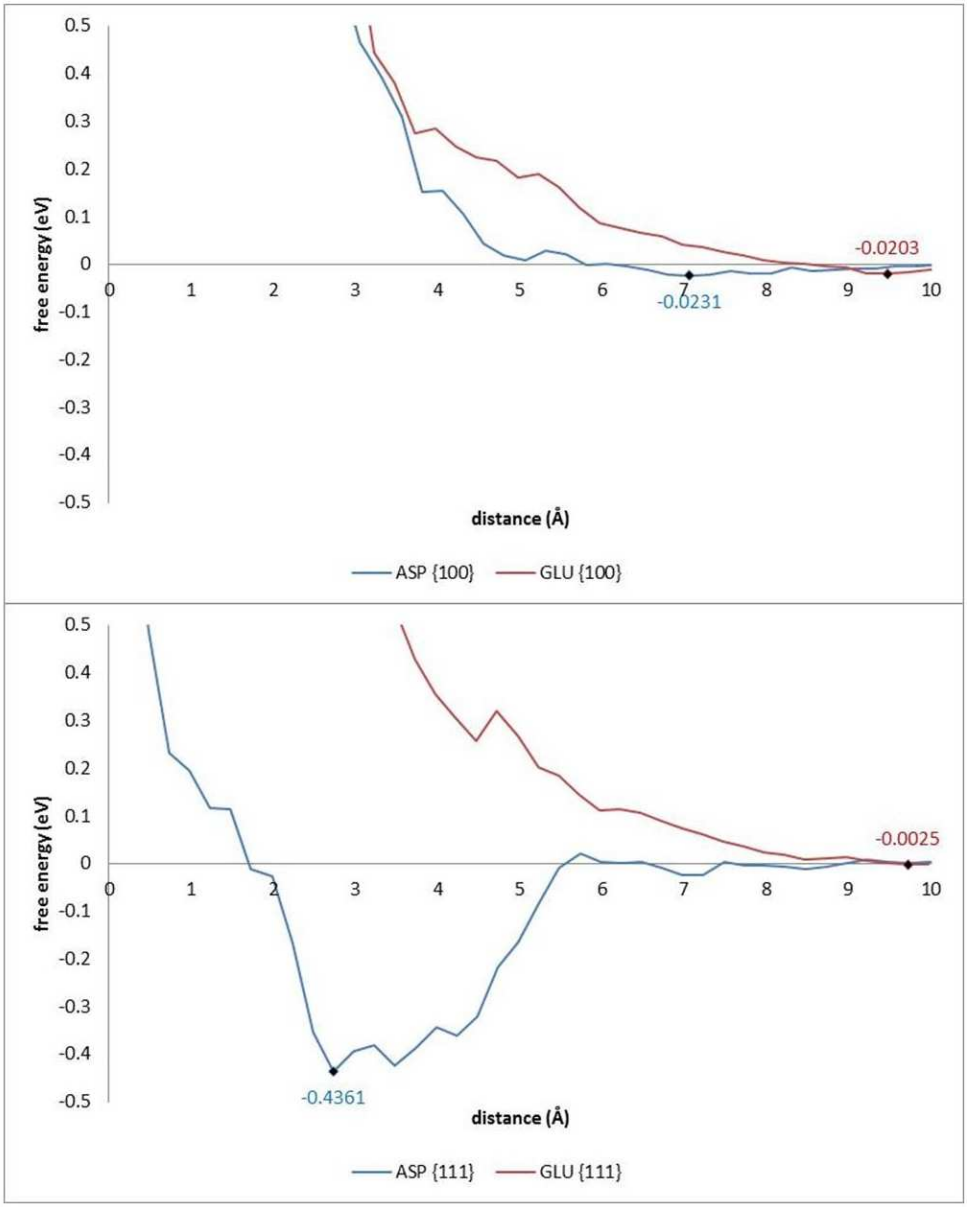

Figure 4.23: A close-up comparison of the free energy profiles of the acidic amino acid residues between -0.5 eV and 0.5 eV, exhibiting the free energy minima.

Basic Amino Acid Attachment

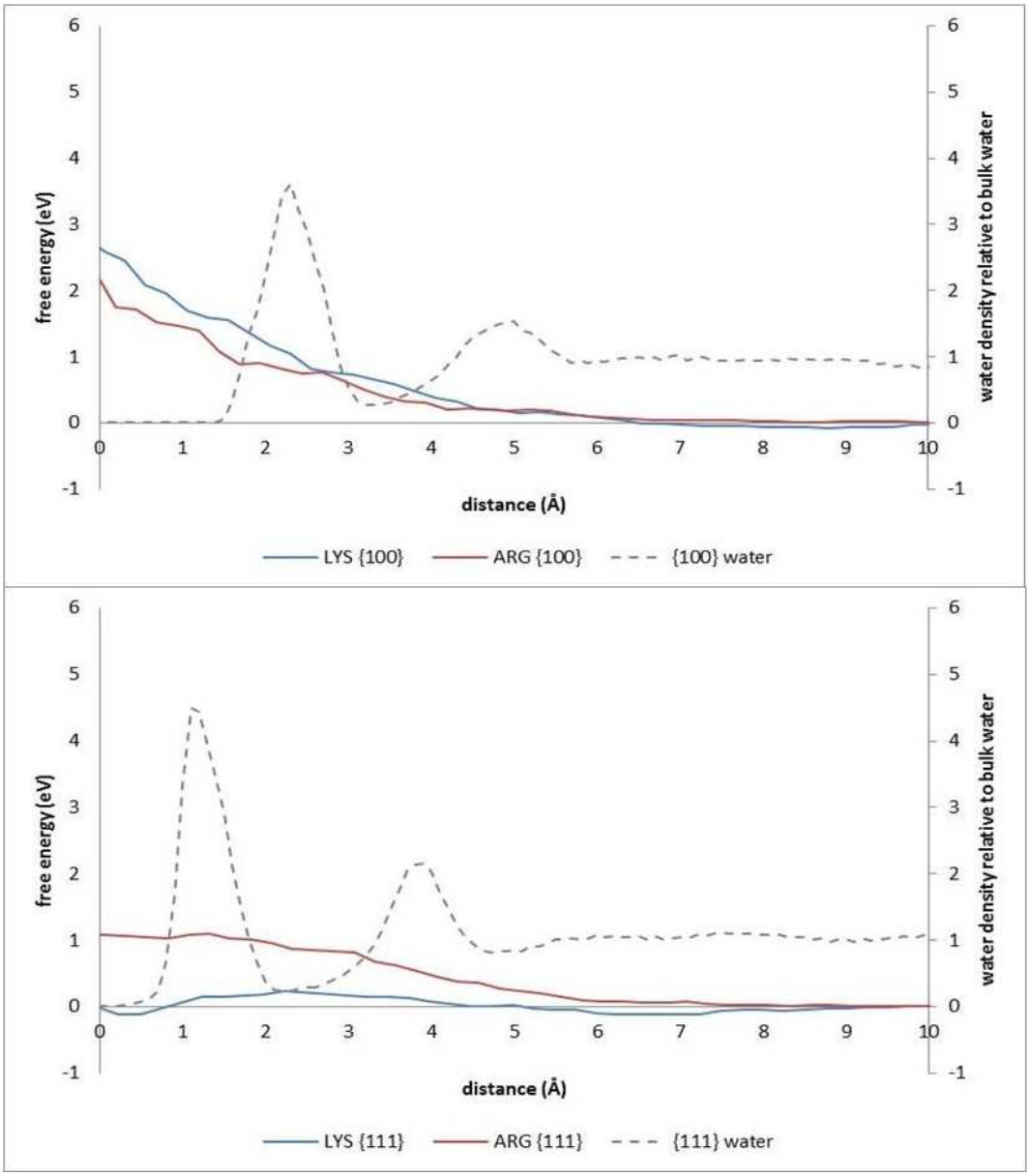

Figure 4.24: A comparison of the free energy profiles of the basic amino acid residues. Distance refers to distance between the peptide CoM and the magnetite surface. The dashed line represents the water density profile for the system.

When the {100} surface basic amino acid free energy profiles (Figures 4.24) were studied it was shown that, the profiles were similar, with arginine having a slightly lower free energy. Between 10 Å and 6.5 Å is optimal for attachment, after which point, a gradual increase in free energy was detected, corresponding with the presence of the intense water layers.

Whereas, with the {111} surface energy profile, a different free energy trend is seen, with lysine exhibiting the lowest free energy values. When the amino acid CoM is between 10

Å and 6 Å the energies stay at approximately 0 eV, however, as the CoM gets closer to the surface the amino acid residue behaviours differ. Arginine presents a gradual rate of free energy increase, whilst, lysine also presents an increase, however this is only small, with its maxima at 0.23 eV at a distance of 2.23 Å. From this distance the free energy decreases to -0.12 eV at its minima, suggesting an area of preferential attachment, subsequently increasing again up to the surface.

Upon examination of the energy minima for the basic amino acids (Figure 4.25) it can be observed that, on the {100} surface, only lysine exhibited a negative value for its energy minima, producing a value of -0.0746 eV at 8.8 Å, suggesting that attachment of the residue centre of mass at this distance would be preferable. On the {111} surface, again, only lysine exhibited negative values for its energy minima, producing a value of -0.1224 eV at 0.2 Åand -0.1226 eV at 6.5 Å, suggesting that attachment of the residue centre of mass at either of these distances would be preferable.

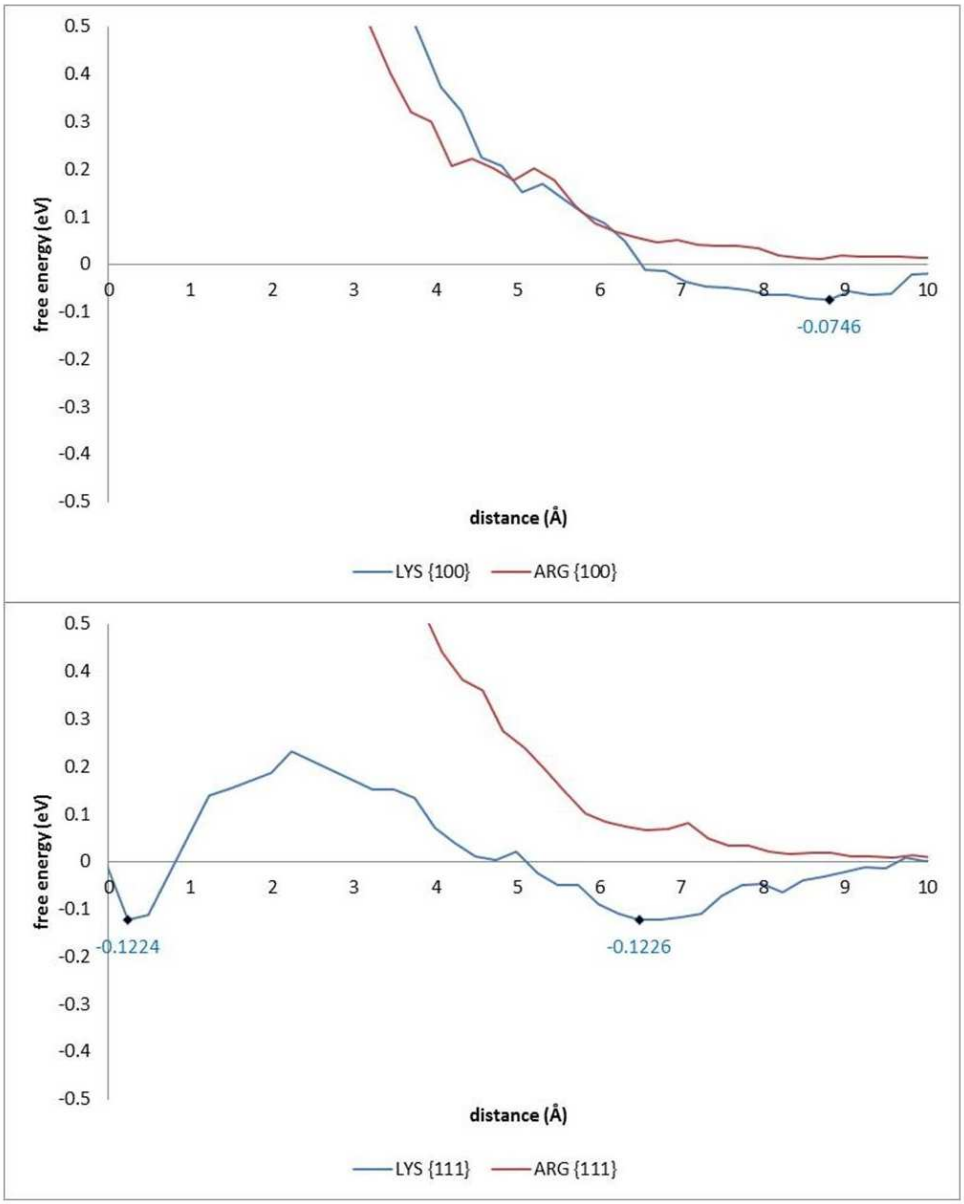

Figure 4.25: A close-up comparison of the free energy profiles of the basic amino acid residues between -0.5 eV and 0.5 eV, exhibiting the free energy minima.

Surface Free Energy

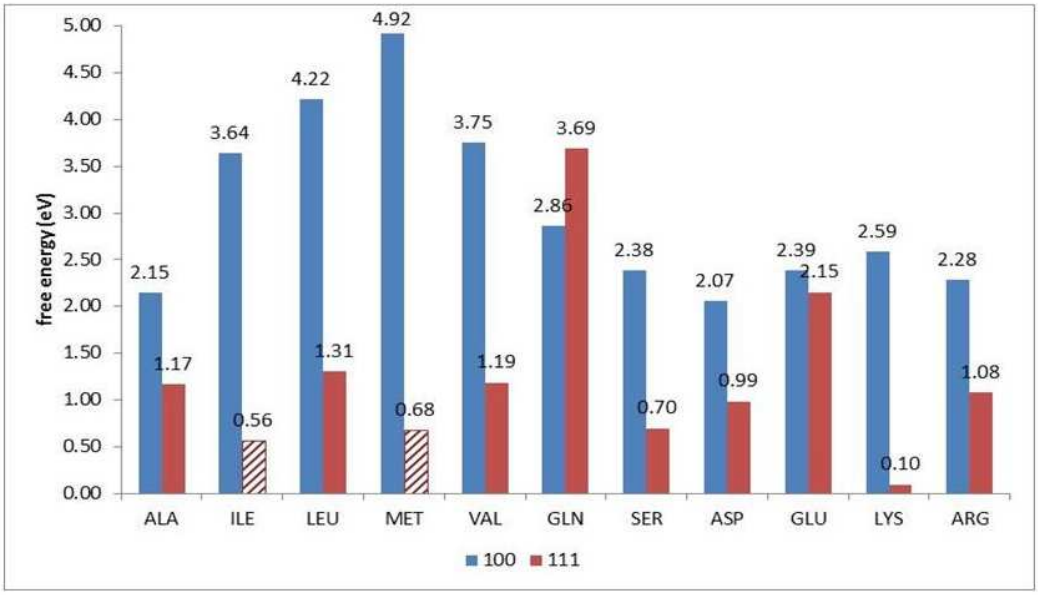

Figure 4.26: A comparison of the free energy values for all Mms6 C–terminal amino acid residue CoM at the magnetite surface for both the {100} and {111} surfaces.

The free energy values for all of the amino acid residue centre of masses at the magnetite surface were examined for both the {100} and {111} surfaces (Figure 4.26).

The non–polar amino acids demonstrated a higher free energy value on the {100} surface, with isoleucine and methionine exhibiting the largest differences. However, it is worth noting that these data are not representative of the true centre of mass at surface values. Nonetheless, there were still large differences between the non–polar amino acid surface values, particularly for leucine and valine. When comparing the polar amino acid residues differing behaviours were seen. For glutamine the {111} surface has the largest free energy value but for serine the {100} surface has the larger free energy value, with serine having the greatest energy difference between the surfaces. With the acidic amino acid residues, aspartate and glutamate both show that the {100} surface produces the larger free energy values, although glutamate showed minimal between the two surface. When comparing the basic amino acid residues, again both show that the {100} surface produces the larger free energy values.

These data show that the {111} surface has significantly lower free energy values than the {100} surface for the Mms6 C–terminal amino acid residues, representing greater iron binding, with the exception of glutamine, suggesting that for the majority the {111} surface is the preferential surface of attachment. It also showed that the amino acid residue

with the lowest free energy value was lysine with a free energy of 0.10 eV, suggesting that the presence of a ammonium group in the side chain as opposed to any form of oxygen in the residue side chain may be more beneficial to iron binding.

4.4 Summary

4.4.1 Unconstrained

From the RDF data, octahedral iron attachment is preferred for the {111} surface, whilst the {100} surface exhibits no particular preference, with the exception of alanine, valine and aspartate which exhibits only octahedral iron binding. For the solvated systems, glutamate on the {100} surface produces the most intense $g(r)$ peak for the tetrahedral iron attachment, whereas, valine on the {111} surface produces the highest intensity $g(r)$ value for the octahedral iron attachment, leading to the idea that attachment of the amino acid residues to the magnetite surface is dependent on many things such as surface type and iron type. For the bond distance data, more bonding was present in the {111} surface system, suggesting that this is the preferred surface of attachment. The acidic amino acid residues exhibit the most bonding, suggesting that the presence of carboxyl oxygen in the residue is beneficial to iron binding. The residence time data indicates that for the *in vacu* systems, on the {100} surface, the acidic amino acid residues exhibit the most bonding lasting over 99% of the simulation length. On the {111} surface, aspartate exhibits the most bonding lasting over 99% of the simulation length, with all bonding being through octahedral iron. For the solvated systems, no {100} surface bonding was exhibited lasting over 20% of the simulation length. On the {111} surface, only leucine, glutamine, serine and arginine, exhibit bonding lasting over 99% of the simulation length. The interfacial energy data suggests that the lowest solvated energies were found on the {100} surface, however, this data looks at average energy values of the full magnetite/amino acid residue/water interface over a simulation, whereas, other forms of data analysis focus on bonding present within the systems throughout the simulations. The {100} surface may exhibit lower interfacial energy due to the surfaces behaviour with water. The *in vacu* interfacial energies are considerably lower, further enforcing the detrimental effect of system solvation. Lysine has the lowest interfacial energy value on both surfaces, whilst, serine and aspartate have the highest on the {100} and {111} surfaces respectively.

4.4.2 Constrained

The Potential of Mean Force data indicates that the {111} surface is the preferred surface of attachment, as this surface produces the lowest free energy values throughout the simulations. Also, the presence of charged groups within the residue side chains does not always lead to improved iron binding ability, as surface type and water influence also have an effect.

4.4.3 General

This chapter demonstrates, for the first time, the interaction of the individual C–terminal Mms6 amino acid residues with *in vacu* and solvated magnetite crystal {100} and {111} surfaces, in both unconstrained and constrained systems. The results from the unconstrained system suggest that acidic amino acids exhibit a higher affinity for iron binding, particularly through octahedral iron, on the {111} surface. The PMF data obtained from the constrained system further supported this, with the majority of amino acid investigated demonstrating significantly lower free energy values in simulations of the {111} surface. Furthermore, the iron binding abilities of all the amino acids, in both systems, were shown to be diminished by the addition of water molecules.

Chapter 5

The Interaction of Di- and Tetrapeptides with Magnetite

5.1 Introduction

Within magnetotactic bacteria, magnetite nanoparticles of uniform size and morphology are formed. In *Magnetospirillum magneticum* strain AMB-1, cubo–octahedral crystals, based on combinations of the {100} and {111} crystallographic faces, are produced, thus suggesting an element of high biological control [56, 57]. In order to replicate this phenomenon for commercial production of magnetic nanoparticles, the biological processes need to be understood. Several proteins have been identified as being involved in magnetite biocrystal nucleation. Such proteins commonly possess acidic amino acid repeat motifs which have a strong affinity for metal ions. The residue sequence of the magnetosome membrane protein, Mms6, is amphiphilic, possessing a hydrophobic N–terminus, and a highly acidic hydrophilic C–terminal region which contains dense carboxyl and hydroxyl groups that are able to bind iron ions[100, 103, 105].

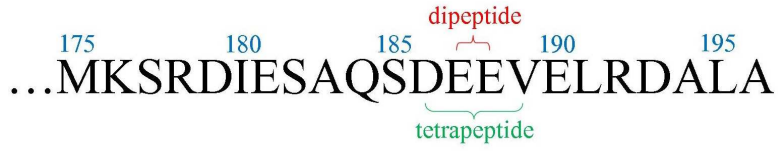

Figure 5.1: Sequence of the C–terminal region of the Mms6 protein, exhibiting residue number. The dipeptide and tetrapeptide regions have been highlighted.

Preliminary experimental investigation[252] identified a glutamate repeat motif within the C–terminal Mms6 sequence, as shown in Figure 5.1, that could potentially contribute to iron binding. It was hypothesised that if the negatively charged carboxylate side chain of glutamate (E) was substituted for a charge neutral side chain, such as alanine (A), the

iron binding potential of the C–terminal sequence would reduce. Based on this repeat motif, the original sequence (EE) and the substitution of one or both of the glutamate with alanine (AE, AA) and its effect on iron binding to the {100} and {111} magnetite surfaces were investigated. Subsequently, the model was extended to include the adjacent aspartic acid (D) and valine (V) moieties in order to more closely replicate the wild–type C–terminal sequence. The effects of the additional amino acids on the iron binding of the original repeat motif (DEEV) were explored, along with the alanine substitutions DEAV, DAEV and DAAV. The hypothesis that carboxylate substitution causes a reduction in iron binding would be supported by an increase in free energy.

In this chapter, molecular dynamics (based on classical atomistic potentials) are used to study the attachment of dipeptides and tetrapeptides created from the C–terminal sequence residues 186-189 (as shown in Figures 5.2 and 5.3), both *in vacu* and solvated, to the {100} and {111} crystal surfaces, and how this effects iron binding. This investigation also examines the interfacial relationships and how these can vary dependent on the size of the peptide chain, the surface utilised and which amino acid residues are involved in the peptide chain. It is also worth noting that this study was divided into two systems; a classic molecular dynamics system and a constrained molecular dynamics system utilising the Potential of Mean Force (see Chapter 3.3.5).

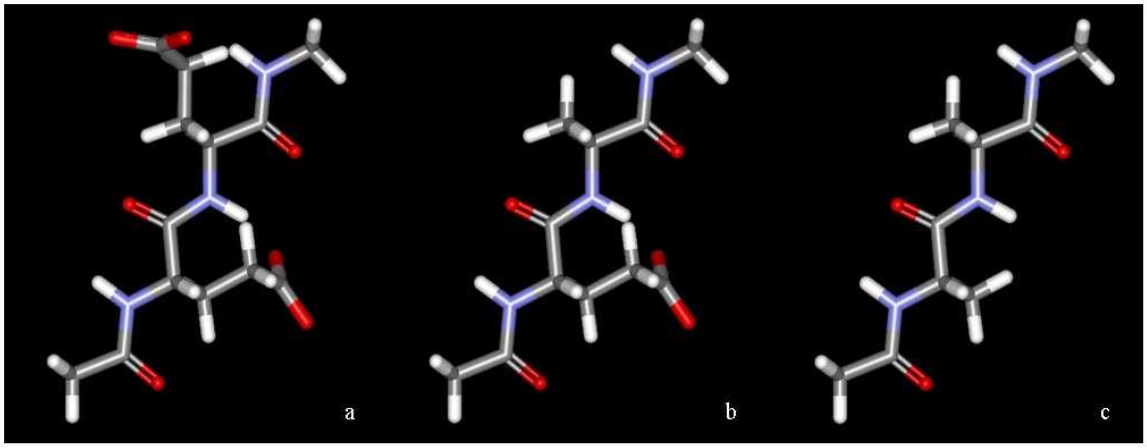

Figure 5.2: Structures of the dipeptides investigated. a)EE, b)AE and c)AA.

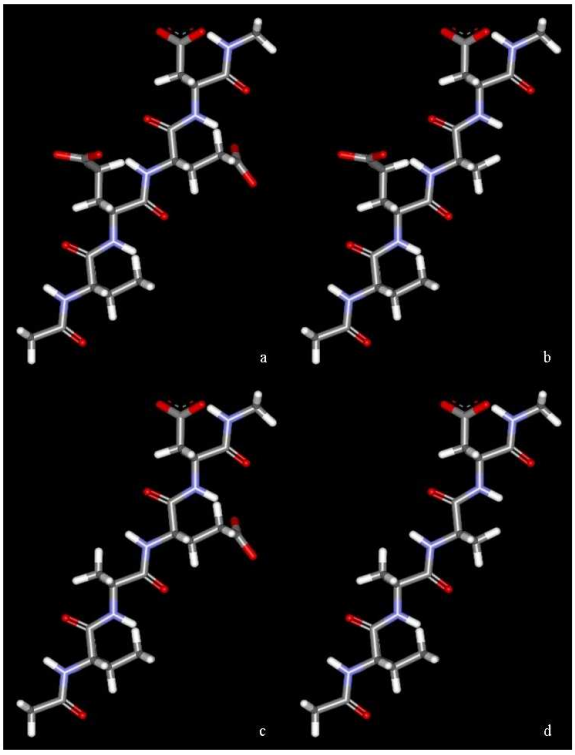

Figure 5.3: Structures of the tetrapeptides investigated. a)DEEV, b)DAEV, c)DEAV and d)DAAV.

5.2 Unconstrained system

5.2.1 Computational Methods

As with the systems in Chapter 4, the starting structures of the peptide chains were generated and capped using TLEAP, then relaxed using SANDER and PMEMD. The systems were then relaxed *in vacu* using DL_POLY, with TIP3P/fs water subsequently added and the system was simulated for a further 1 ns. The water was then removed and the peptides were then placed in the vacuum gap above the magnetite slab surface and run further *in vacu* and solvated for 5 ns and 1ns, respectively. All systems were simulated at 300 K. The ensembles, potentials and parameters used are described in Chapters 2 and 3.

The evolution of potential energy of the di- and tetrapeptides as a function of time were plotted, an example of which is shown for the original dipeptide, EE, in Figure 5.4. A full set of potential energy plots can be found in the Appendix. The equilibration period for all simulations was 20 ps, after which point all of the di- and tetrapeptide energies converge.

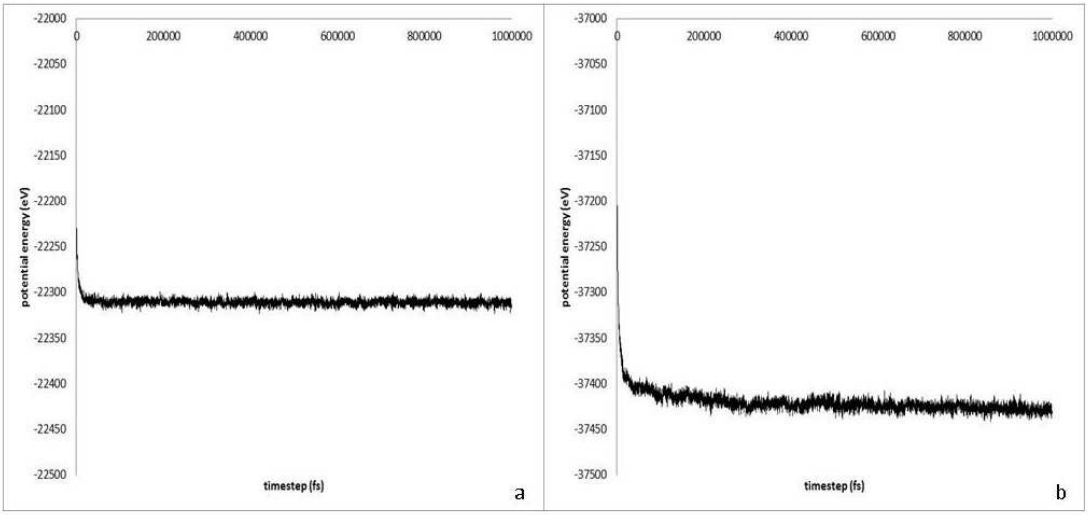

Figure 5.4: Evolution of potential energy as a function of time plots for the original dipeptide, EE. a) {100} solvated surface. b) {111} solvated surface.

5.2.2 Results and Discussion

Dipeptide Attachment

Radial distribution function (RDF) data was examined for the dipeptide attachment to the {100} and {111} magnetite surfaces. As with the approach used for the amino acids attachments of Chapter 4, a distance of 1.5 Åto 2.5 Åfrom dipeptide oxygen to the next nearest neighbouring tetrahedral or octahedral iron of the magnetite is classified as a bond. Dipeptide oxygen-iron distances or r values for all dipeptides can be found in Table 5.1.

Figure 5.5 exhibits the RDF profiles for EE. The RDF profiles of the remaining dipeptides can be found in the Appendix. When the RDF data for the dipeptides was examined, it can be inferred that oxygen atoms of EE bond to both octahedral and tetrahedral iron of the {100}, in both the *in vacu* and solvated systems. Conversely, on the {111} surface, only octahedral iron bonding is present. A similar trend is also presented for both mutated sequences (AE and AA); however, in these instances, tetrahedral iron bonding is not present for the solvated system on the {100} surface. This suggests that when the dipeptides are attached to either surface, different iron attachment sites are favoured, with octahedral iron attachment being the preferred for both surfaces, with the exception of the original sequence, where tetrahedral iron attachment is the preferred on the {100} surface. These data suggest that a single or double alanine substitution, along with the addition of water, is detrimental to the iron binding potential as it appears to have inhibited

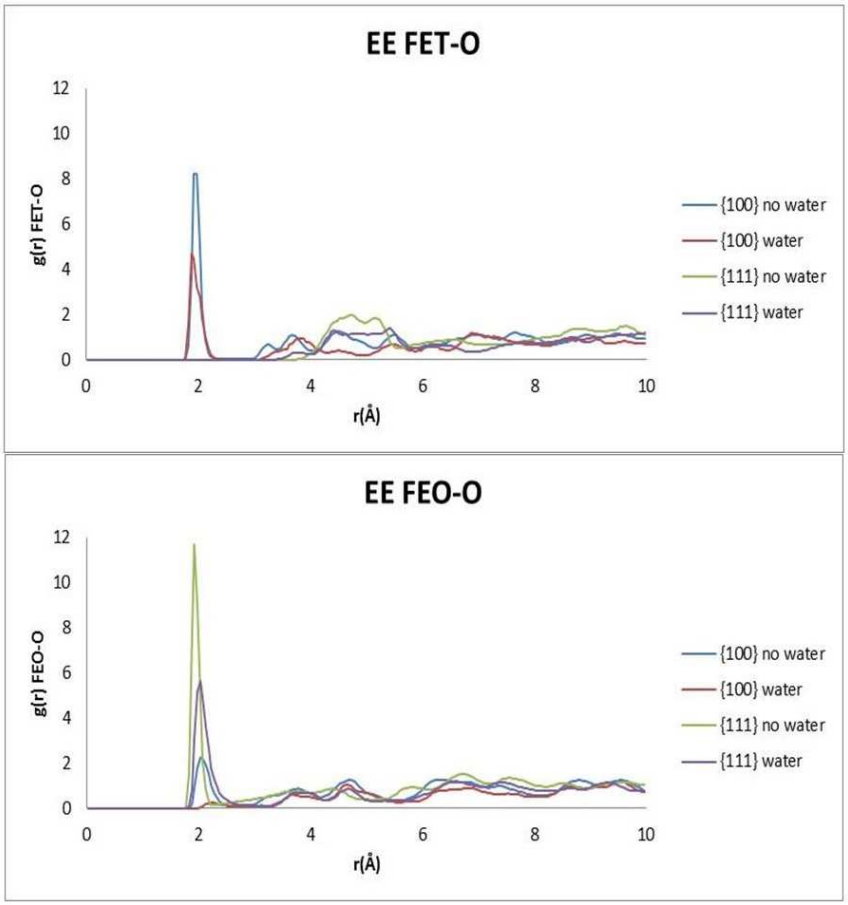

Figure 5.5: RDF plots for EE. FET is tetrahedral iron, FEO is octahedral iron. Blue is {100} *in vacu*. Red is {100} solvated. Green is {111} *in vacu*. Purple is {111} solvated.

tetrahedral iron bonding on the {100} surface.

These data also show that, the intensity of the first peak of the RDF is greater in the *in vacu* systems, suggesting that once water is introduced into the system, it is more difficult for dipeptide oxygen to have an iron atom as a next nearest neighbour and for this to remain for the simulation duration, as water will bond preferably to the magnetite surface. For the solvated systems, the most intense peak is shown for EE on the {111} surface having octahedral iron as its next nearest neighbour, proposing that in these systems the oxygen present is more freely accessible to the next nearest neighbouring irons.

The RDF data can be examined alongside the $FE - O_{aminoacid}$ bond distance data extracted from the final simulation coordinates (Table 5.2).

Dipeptide	Fe type	Surface	System	r (Å)
EE	FET	100	*in vacu*	1.98
			solvated	1.88
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.03
			solvated	2.28
		111	*in vacu*	1.93
			solvated	2.03
AE	FET	100	*in vacu*	2.08
			solvated	-
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	1.98
			solvated	2.08
		111	*in vacu*	1.93
			solvated	2.03
AA	FET	100	*in vacu*	2.18
			solvated	-
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.13
			solvated	2.23
		111	*in vacu*	2.03
			solvated	2.08

Table 5.1: Dipeptide r (Å) values from RDF data.

	{100} surface attachment			{111} surface attachment		
Dipeptide	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
EE	1st E O	1.96	NB	1st E O	1.91	2.20
		1.97	1.90		1.92	2.03
	2nd E O	2.02	2.11	E-E PBO	2.15	NB
		2.03	NB	2nd E O	2.09	2.12
	NBO	2.17	NB		2.27	2.15
					1.95	NB
AE	NBO	2.08	NB	NBO	2.01	NB
	E O	2.01	NB	E O	1.88	2.18
		1.91	NB		-	2.23
					1.94	NB
AA	NBO	2.24	NB	A-A PBO	1.96	1.97
				NBO	1.92	2.07

Table 5.2: Dipeptide $FE-O_{dipeptide}$ bond distance data. BL is bond length, E O is glutamate side chain oxygen, x-x PBO is inter–residue peptide bond oxygen and NBO is N–terminal peptide bond oxygen.

The dipeptide bond distance data shows that, for EE *in vacu*, bonding through both side chain oxygens of glutamate occurred, and even, two bonds from one oxygen of the side chain oxygen of the C–terminal glutamate in the {111} surface system. This suggests strong bonding of the dipeptide to the magnetite, which was evident with the original dipeptide showing a more intense $g(r)$ peak in the RDF data. Less bonding occurred when the systems were solvated, suggesting that the oxygen in water bonds more favourably to the magnetite iron compared to oxygen in the dipeptides, leading to the dipeptide being pushed away from the magnetite surface, or remaining outside of the water adsorption layer, to allow the water to move closer to the surface and bind (Figure 5.6). This was reflected by the solvated systems having less intense $g(r)$ peaks than the corresponding *in vacu* systems. The average bond lengths for the original sequence are; 2.03 Å and 2.01 Å on the {100} surface for *in vacu* and solvated systems respectively, and 2.07 Å and 2.13 Å on the {111} surface for *in vacu* and solvated systems respectively.

With a single substitution of glutamate to alanine, both oxygens on the side chain of glutamate bonded to magnetite iron. Again, the addition of water to the system was detrimental to this bonding; however, with mutation this effect was worsened, as there was a loss of potential binding sites due to the substitution of glutamate side chain oxygen for a methyl group. This is reflected in that there is no bonding present for AE on the {100} solvated surface. This effect is less severe on the {111} solvated surface as there are still two bonds from one of the side chain oxygen of glutamate to the magnetite present. When this data is compared to that of the RDF it can be seen that the removal of the potential binding sites of the second glutamate reduces the intensity of the $g(r)$ compared to the original dipeptide values. The average bond lengths in AE are; 2.00 Å for the *in vacu* system on the {100} surface, and 1.94 Å and 2.21 Å on the {111} surface for the *in vacu* and solvated systems respectively.

For the double substitution mutant, AA, only bonding to peptide bond oxygen was possible. On the {100} surface, only an N–terminal peptide bond oxygen bond was present, however, upon system solvation it is no longer present. With the {111} surface, there was an additional bond present (A-A PBO), and this bond is also present when the system is solvated. The average bond lengths in AA is; 2.24 Å for the *in vacu* system on the {100} surface, and 1.94 Å and 2.02 Å on the {111} surface for the *in vacu* and solvated systems respectively.

As with the amino acid residues, residence times and average coordination numbers were considered in order to gain a more accurate representation of bonding behaviour through-

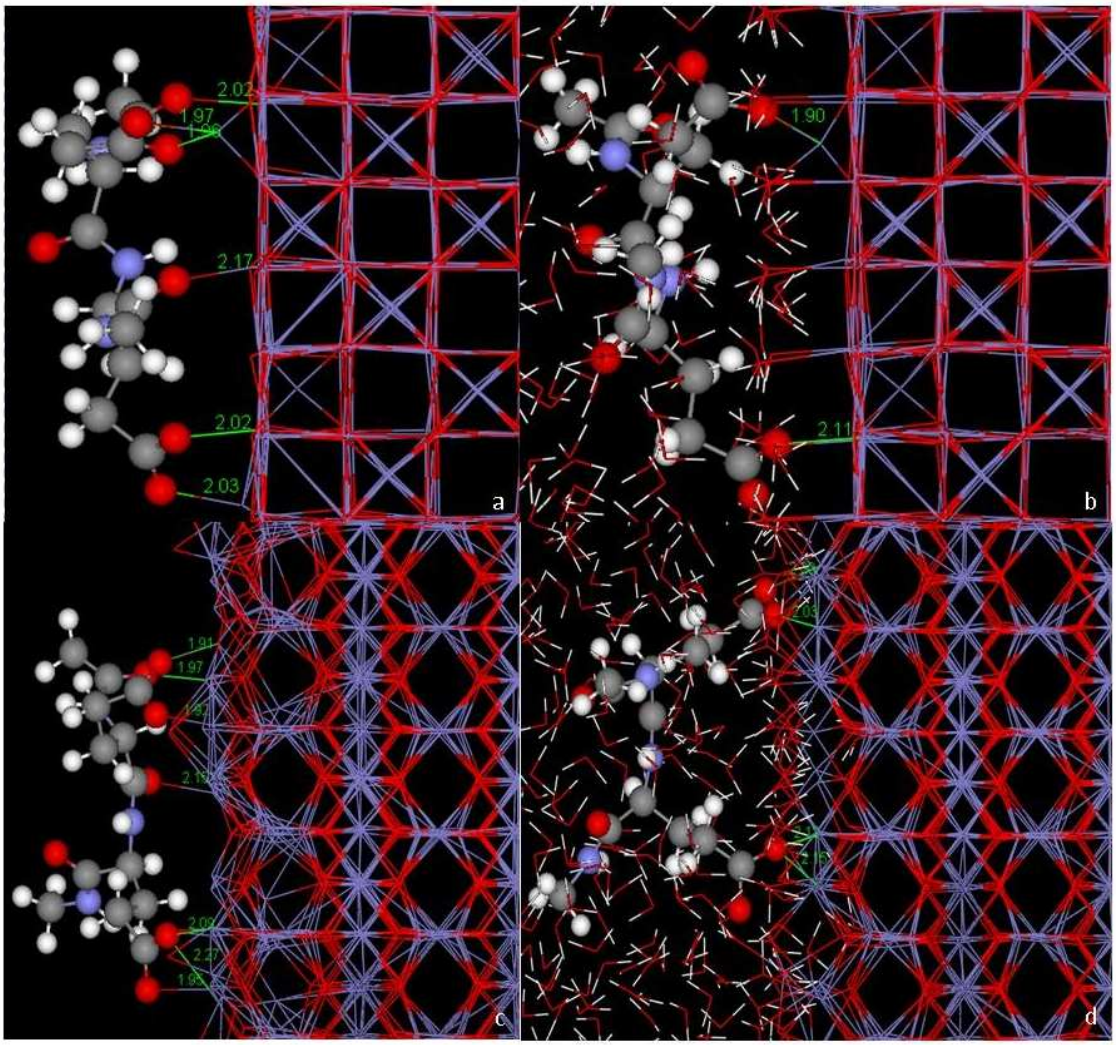

Figure 5.6: Bond distance images for the *in vacu* and solvated systems of EE. a){100} *in vacu*, b){100} solvated, c){111} *in vacu* and d){111} solvated.

out the simulations.

Tables 5.3 and 5.4 represent the residence times and average coordination numbers for bonding between all oxygen present in the dipeptides and tetrahedral and/or octahedral iron on both the {100} and {111} magnetite surfaces. Figure 5.7 describes the different possible oxygen binding sites using the original dipeptide, EE, as an example.

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
EE {100} surface	E O	98.63	0.95	3.33	0.09	999.00	1.00	0.00	0.00
	E O	119.26	0.96	2.41	0.08	8.03	0.04	0.00	0.00
	E-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	4998.00	1.00	0.00	0.00	2.74	0.09
	E O	4998.00	1.00	0.00	0.00	999.00	1.00	0.00	0.00
	NBO	0.00	0.00	172.13	0.96	0.00	0.00	31.31	0.25
EE {111} surface	E O	0.00	0.00	4952.86	1.01	0.00	0.00	896.89	1.82
	E O	0.00	0.00	4951.75	1.00	0.00	0.00	374.47	0.98
	E-E PBO	0.00	0.00	4958.03	1.00	0.00	0.00	81.23	0.19
	E O	0.00	0.00	4991.19	1.00	0.00	0.00	329.54	1.12
	E O	0.00	0.00	4951.17	1.20	0.00	0.00	359.07	1.33
	NBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 5.3: Residence times (RT) and average coordination numbers (av. CN) of the dipeptide sequences. (E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
AE {100} surface	E O	0.00	0.00	4984.13	1.01	0.00	0.00	169.36	0.88
	E O	0.00	0.00	4990.86	1.00	0.00	0.00	235.98	0.79
	A-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NBO	485.20	0.55	204.30	0.42	0.00	0.00	12.64	0.04
AE {111} surface	E O	0.00	0.00	4975.04	1.01	0.00	0.00	178.65	1.84
	E O	0.00	0.00	4977.32	1.01	0.00	0.00	399.66	0.73
	A-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NBO	0.00	0.00	4939.19	1.00	0.00	0.00	125.98	0.25
AA {100} surface	A-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NBO	185.33	0.17	875.15	0.79	0.00	0.00	52.99	0.10
AA {111} surface	A-A PBO	0.00	0.00	4993.47	1.00	0.00	0.00	70.21	1.26
	NBO	0.00	0.00	4989.00	1.00	0.00	0.00	723.60	0.98

Table 5.4: Continued. Residence times (RT) and average coordination numbers (av. CN) of the dipeptide sequences. (E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

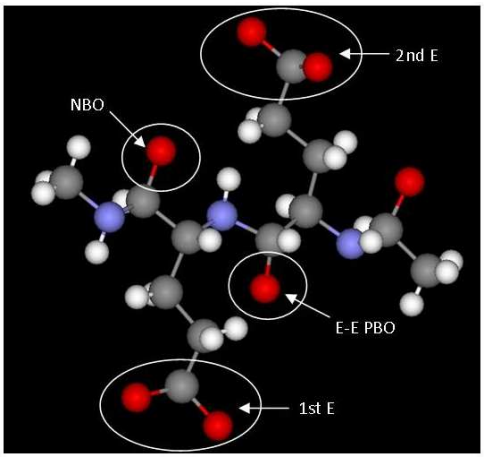

Figure 5.7: Different possible oxygen binding sites of EE. NBO is N–terminal peptide bond oxygen. E-E PBO is glutamate glutamate peptide bond oxygen. 1st E relates to the side chain oxygen of the first glutamate.

The dipeptide residence time data shows that, for EE *in vacu*, when attached to the {100} surface, both oxygens of the N–terminal glutamate bound to iron for over 99% of the simulation length (4998.00 ps), however, each oxygen attached to a different iron type. When this system is solvated the $FET - O_{dipeptide}$ bond continues to be present for over 99% of the simulation length (999.00 ps), however, the residence time of the $FEO - O_{dipeptide}$ bond drops to 2.74 ps, suggesting that attachment at this site is no longer favourable. Another bond is, however, formed lasting for over 99% of the simulation length (999.00 ps) for one of the oxygen of the C–terminal glutamate to tetrahedral iron, proposing that the low value for the equivalent *in vacu* residence time is due to the formation of this bond toward the latter stages of this system simulation. The residence time data for the {111} surface shows a very different trend, as there is no tetrahedral iron bonding present at all. For the *in vacu* system all oxygen showed octahedral iron bonding lasting for over 99% of the simulation length, with the exception of the N–terminal peptide bond oxygen showing no bonding (0.00 ps). Upon solvation bonding reduces greatly, with one of the C–terminal glutamate oxygen having the longest residence time at 896.89 ps, whilst the remaining possible bonds do not last longer than 38% of the simulation length. This data suggests that the addition of water provides competitive binding to the magnetite, and the magnetite in many cases has a stronger affinity for the water than the original dipeptide.

With a single alanine substitution, no tetrahedral iron bonding was exhibited, except for to the N–terminal peptide bond oxygen (485.20 ps) on the {100} *in vacu* surface. However, for {100} surface octahedral iron bonding, bonds formed to each of the glutamate oxygen

lasting for over 99% of the simulation length (4984.13 ps and 4990.86 ps). Upon solvation, no tetrahedral iron bonding was present and no octahedral iron bonding lasted longer than 24% of the simulation length, again suggesting competitive water binding, with the substitution of glutamate for alanine increasing the magnetite–water affinity. For the AE–{111} surface, octahedral iron bonding was existent for all oxygen present, with the exception of the A-E peptide bond oxygen, lasting over 98% of the simulation length. On addition of water, bonding greatly diminishes, with the bond between octahedral iron and one of the glutamate oxygen exhibiting the longest residence time at 399.66 ps, whilst the remaining possible bonds do not last longer than 18% of the simulation length, suggesting that single substitution of a glutamate for an alanine clearly has a detrimental effect on the potential iron binding.

With the mutant consisting of a double alanine substitution, the N–terminal peptide bond oxygen and the A-A peptide bond oxygen are the only possible bonding areas. The {100} *in vacu* surface showed no bonding for the A-A peptide bond oxygen, and no bonding lasting over 18% of the simulation length to the N–terminal peptide bond oxygen. When the system was solvated, only one bond was exhibited (NBO-FEO), which lasted just 52.99 ps. For the {111} surface, as with EE and AE, no tetrahedral iron bonding was present. With octahedral iron, bonding to both of the peptide bond oxygen was seen lasting over 99% of the simulation length for the *in vacu* system. With the addition of water to this system the residence times of these bonds drop (70.21 ps and 723.60 ps), suggesting that attachment at this site is no longer favourable.

On comparison of all dipeptides, it is shown that the amount of bonding lasting over 99% of the simulation length reduces more with the removal of each glutamate from the sequence (e.g. *in vacu* {111} systems from 5 to 3 to 2 bonds), particularly when water is present and providing competitive binding, suggesting that the oxygen present in the glutamate are areas of favourable iron binding.

As with the amino acid residues of Chapter 4, the interfacial energies of the peptide sequences on the {100} and {111} magnetite surfaces were calculated using the Yang *et al* method[164], giving an indication of how easily the peptide sequences can be adsorbed onto the different magnetite surfaces. All related energy data from the interfacial energy calculations of the dipeptides can be found in Table 5.5.

The solvated dipeptide interfacial energy data shows that, the {100} surface produces lower interfacial energy values (between 5.98 eV and 7.23 eV) than the {111} surface

{100} surface attachment				
	E_{int} of s-p (eV)	E_{diff} of s-p (eV)	E_{int} of s-p-w (eV)	E_{diff} of s-p-w (eV)
EE	-6.33	7.86	7.23	0.00
AE	-4.42	5.95	6.22	1.01
AA	1.53	0.00	5.98	1.24
{111} surface attachment				
	E_{int} of s-p (eV)	E_{diff} of s-p (eV)	E_{int} of s-p-w (eV)	E_{diff} of s-p-w (eV)
EE	-8.73	6.99	7.39	4.38
AE	-7.80	6.06	11.78	0.00
AA	-1.74	0.00	7.60	4.18

Table 5.5: Interfacial energies for all dipeptides on the {100} and {111} surface, using the Yang *et al* method[164]. E_{int} of s-p is the interfacial energy of the slab–peptide system. E_{diff} of s-p is the difference in interfacial energy from the highest interfacial energy of the slab–peptide system. E_{int} of s-p-w is the interfacial energy of the slabpeptide–water system. E_{diff} of s-p-w is the difference in interfacial energy from the highest interfacial energy of the slab–peptide–water system.

(between 7.39 eV and 11.78 eV). On the {100} surface, EE had the highest interfacial energy value and AA was shown to have the lowest, whilst on the {111} surface, AE had the highest interfacial energy and EE now exhibited the lowest value. This data suggests that the interfacial energy trends are surface dependent and that, for the {100} surface, iron binding appears to improve with each glutamate substitution. Furthermore, this data shows that the presence of water in the system has an unfavourable effect on the interfacial energy, as much lower energies (predominantly of a negative value) are presented for the *in vacu* system. Water also affects the interfacial behaviour of the dipeptides, as the *in vacu* system interfacial energy trends differ greatly from the solvated systems, as for both surfaces, the lowest energy was exhibited for the original dipeptide sequence and the interfacial energy increases with each subsequent alanine substitution. For the {100} surface, this is the complete opposite of what is seen for the solvated surface, whereas, with the {111} surface, EE still exhibits the lowest energy, however, AE exhibits the highest energy value, which is considerably higher than both EE and AA (4.39 eV than EE as opposed to 0.21 eV higher as seen for AA. The increase in interfacial energy with the addition of water could be explained by the need for more energy in the system to allow for the dipeptide to get close to the magnetite surface through the structured water layers.

As discussed in more detail with the amino acid interfacial energies (Section 4.2.2 Non–polar Amino Acid Attachment), there are many problems with this method of calculating

adsorption energy. The absolute values themselves are noticeably an order of magnitude larger than anticipated and the solvated trends are not as expected from experimental studies. The relative residence times of various surface interactions, however, reveal the expected trend, and as it can be shown that $RT \times ln(\tau^1/\tau^2) = (\Delta H^2_{ads} - \Delta H^1_{ads}) = (E^1_{des} - E^2_{des})$, then this method for determining the change in adsorption energy is perhaps more reliable. Furthermore, rather than directly calculating the adsorption energy, via this interfacial energy method, an alternative would be to use an indirect method, such as those based on the Potential of Mean Force, which has been considered in Section 3.3.5.

Tetrapeptide Attachment

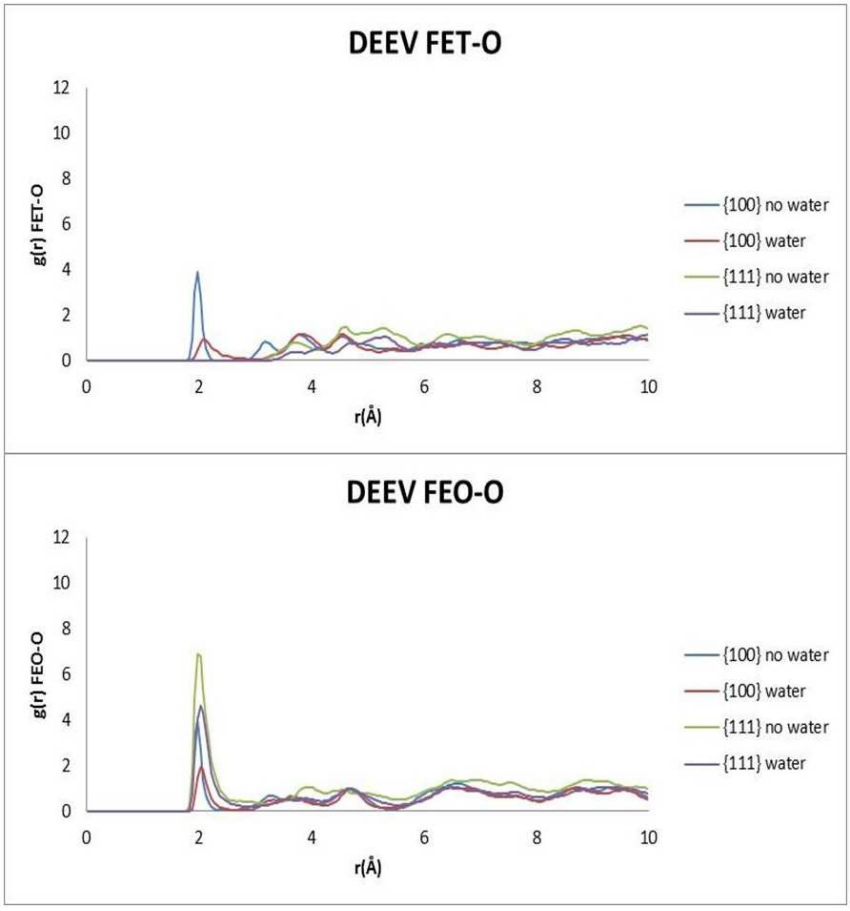

Figure 5.8: RDF plots for DEEV. FET is tetrahedral iron, FEO is octahedral iron. Blue is {100} *in vacu*. Red is {100} solvated. Green is {111} *in vacu*. Purple is {111} solvated.

Figure 5.8 exhibits the RDF profiles for DEEV. The RDF profiles of the remaining tetrapeptides can be found in the Appendix. The tetrapeptide RDF data shows that, as with the dipeptides, only octahedral iron bonding is present for the {111} surface attachments. With the mutated tetrapeptides, tetrahedral iron bonding is present in the

{100} solvated systems, unlike in the dipeptide mutated systems, suggesting that the additional amino acid residues to the sequences encourages tetrahedral iron attachment, which could be due to the presence of further potential oxygen binding sites. The tetrapeptides show the same type of bonding suggesting that mutation of the original sequence does not affect iron bonding type, unlike with the dipeptides.

From the tetrapeptide RDF data it can also be seen that, the *in vacu* systems had higher intensity first peaks than the corresponding solvated systems. For each tetrapeptide, the most intense $g(r)$ peak belonged to the oxygen bonding through an octahedral iron on the {111} surface, suggesting that this is the preferred means of attachment for the tetrapeptides. It was unexpected to see that the two single substituted mutants presented with more intense $g(r)$ peaks than the original sequence, in particular DAEV, leading to the belief that single substitution improves octahedral iron binding on the {111} surface. When the solvated system data was examined, again octahedral iron bonding on the {111} surface produce the most intense $g(r)$ peaks, with DAEV having the largest.

When the tetrapeptide solvated values are compared to the dipeptide solvated values, for tetrahedral iron bonding on the {100} surface, EE, has the largest $g(r)$ value, however, no other dipeptide presents with this type of bonding, whilst all of the tetrapeptides exhibit this bonding type. This suggests that the addition of the extra amino acids affects the binding potentials of the peptides in this system, improving the bonding through tetrahedral iron but not to the extent of the original dipeptide sequence. When the octahedral iron bonding on the {100} surface was compared for the peptides, it was shown that the single mutated dipeptide and the original tetrapeptide had the highest intensity first peaks, and that, the tetrapeptides produced more intense $g(r)$ peaks than the dipeptides, with the exception of AE, again proposing that the addition of the extra amino acids improves the potential octahedral iron bonding but not to the extent of the single mutated dipeptide sequence. No tetrahedral iron bonding was present on the {111} surface for any of the peptides, suggesting that the addition of the extra amino acids has no effect on this type of bonding. On comparison of the {111} surface octahedral iron bonding, it can be seen that the single mutated tetrapeptide DAEV produces the most intense $g(r)$ peak, whilst the equivalent dipeptide, AE, produces the smallest. This suggests that the addition of the extra amino acids has a significant effect on potential octahedral iron bonding for this particular sequence. Tetrapeptide oxygen-iron distances or r values for all tetrapeptides can be found in Tables 5.6 and 5.7.

The $FE - O_{tetrapeptide}$ bond distance data for all tetrapeptides are represented in Tables

Tetrapeptide	Fe type	Surface	System	r (Å)
DEEV	FET	100	*in vacu*	1.98
			solvated	2.08
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	1.98
			solvated	2.03
		111	*in vacu*	1.98
			solvated	2.03
DAEV	FET	100	*in vacu*	1.98
			solvated	1.88
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	1.98
			solvated	2.08
		111	*in vacu*	2.03
			solvated	2.03

Table 5.6: Tetrapeptide r (Å) values from RDF data.

5.8, 5.9 and 5.10.

Tetrapeptide	Fe type	Surface	System	r (Å)
DEAV	FET	100	*in vacu*	1.88
			solvated	1.93
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	1.98
			solvated	2.13
		111	*in vacu*	1.93
			solvated	2.03
DAAV	FET	100	*in vacu*	2.03
			solvated	2.23
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	1.98
			solvated	2.03
		111	*in vacu*	1.98
			solvated	2.03

Table 5.7: Continued. Tetrapeptide r (Å) values from RDF data.

	{100} surface attachment			{111} surface attachment		
Tetrapeptide	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
DEEV	D O	1.94	NB	D-E PBO	2.16	NB
		2.02	2.11	1st E O	2.28	2.08
	1st E O	2.11	2.05		2.23	NB
		2.11	2.55		1.97	2.04
	2nd E O	2.00	NB		2.14	NB
		1.97	NB	E-E PBO	1.91	2.03
				2nd E O	2.05	2.04
					2.18	2.19
					2.01	2.06
					2.08	NB
				E-V PBO	2.14	NB

Table 5.8: Tetrapeptide $FE - O_{tetrapeptide}$ bond distance data. BL is bond length, D O is aspartate side chain oxygen, E O is glutamate side chain oxygen, x-x PBO is inter–residue peptide bond oxygen and NBO is N–terminal peptide bond oxygen.

	{100} surface attachment			{111} surface attachment		
Tetrapeptide	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
DAEV	NBO	1.95	NB	NBO	2.07	2.01
	D O	1.95	NB	D O	2.03	2.08
		1.99	2.00	D-A PBO	2.15	2.20
	D-A PBO	2.18	NB	A-E PBO	2.08	1.86
	E O	1.96	NB	E O	2.09	2.12
		1.95	NB		1.97	2.13
					2.05	2.03
					2.03	NB
				E-V PBO	2.02	2.24

Table 5.9: Continued. Tetrapeptide $FE - O_{tetrapeptide}$ bond distance data. BL is bond length, D O is aspartate side chain oxygen, E O is glutamate side chain oxygen, x-x PBO is inter–residue peptide bond oxygen and NBO is N–terminal peptide bond oxygen.

	{100} surface attachment			{111} surface attachment		
Tetrapeptide	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
DEAV	D O	1.98	NB	NBO	2.07	NB
		1.94	2.13	D O	1.94	2.18
	D-E PBO	2.13	2.18		-	2.11
	E O	1.87	2.07		1.91	2.20
		1.97	1.99	E O	1.98	2.00
	E-A PBO	2.21	2.26		1.93	2.02
				E-A PBO	2.21	NB
DAAV	NBO	2.22	NB	NBO	2.05	NB
	D O	1.99	2.00	D O	1.95	2.16
		2.00	NB		-	2.08
	D-A PBO	1.98	NB		2.11	2.13
					-	2.00
				A-V PBO	1.94	2.28

Table 5.10: Continued. Tetrapeptide $FE - O_{tetrapeptide}$ bond distance data. BL is bond length, D O is aspartate side chain oxygen, E O is glutamate side chain oxygen, x-x PBO is inter–residue peptide bond oxygen and NBO is N–terminal peptide bond oxygen.

The tetrapeptide bond distance data (Figure 5.9) shows that the DEEV tetrapeptide exhibits increased bonding on the {111} surface than the {100} surface. However, on the {100} surface, only iron bonding through aspartate and glutamate side chain oxygen was seen, whereas, on the {111} surface, bonding from both the side chain oxygen and the peptide bond oxygen of glutamate (D-E, E-E and E-V) was evident. This suggests that different surfaces have different preferences for oxygen type.

This data also suggests that, iron binding ability was reduced with the addition of water to the systems, due to more favourable bonding between the magnetite slab and water molecules. The trends shown are represented in the RDF data by the more intense $g(r)$ peaks for the {111} surface rather than the {100} surface, and for the *in vacu* systems rather than the solvated systems. The average bond lengths for an iron to a tetrapeptide oxygen in DEEV was; 2.03 Å and 2.24 Å on the {100} surface for the *in vacu* and solvated systems respectively, and 2.10 Å and 2.07 Å on the {111} surface for the *in vacu* and solvated systems respectively.

As would be expected, when the bond distance data of DEEV was compared to that of EE, the addition of extra amino acids resulted in increased bonding due to the greater number of available oxygen binding sites. This suggests that as sequence length increases as does iron binding potential. There was also similar types of bonding on each surface seen, i.e. the {100} surface is predominantly side chain oxygen attachment whereas with the {111} surface there is a mix of side chain oxygen and peptide bond oxygen attachments.

With a single substitution of glutamate to alanine, there is the same amount of bonding with similar bonding types found for the DAEV and DEAV sequences in the {100} *in vacu* system, as both show bonding to each of the aspartate and glutamate side chain oxygens and to two types of peptide bond oxygen (N–terminal and D-A for DAEV, and D-E and E-A for DEAV). Whereas, with the {111} *in vacu* system, more bonding was present with DAEV than DEAV, suggesting that sequence order effects potential iron binding and that two charged amino acids together in a sequence reduces iron binding potential, however, more bonding was present for DEEV than DAEV suggests proposing that three charged residues together actually improves the potential binding. Iron binding decreases with the substitution of glutamate for alanine.

For DAEV, solvation has a greater detrimental effect on the {100} iron binding than {111} surface as fewer bonds that were present in the *in vacu* systems remain for the solvated. For the solvated DEAV systems, the same amount of bonding was exhibited

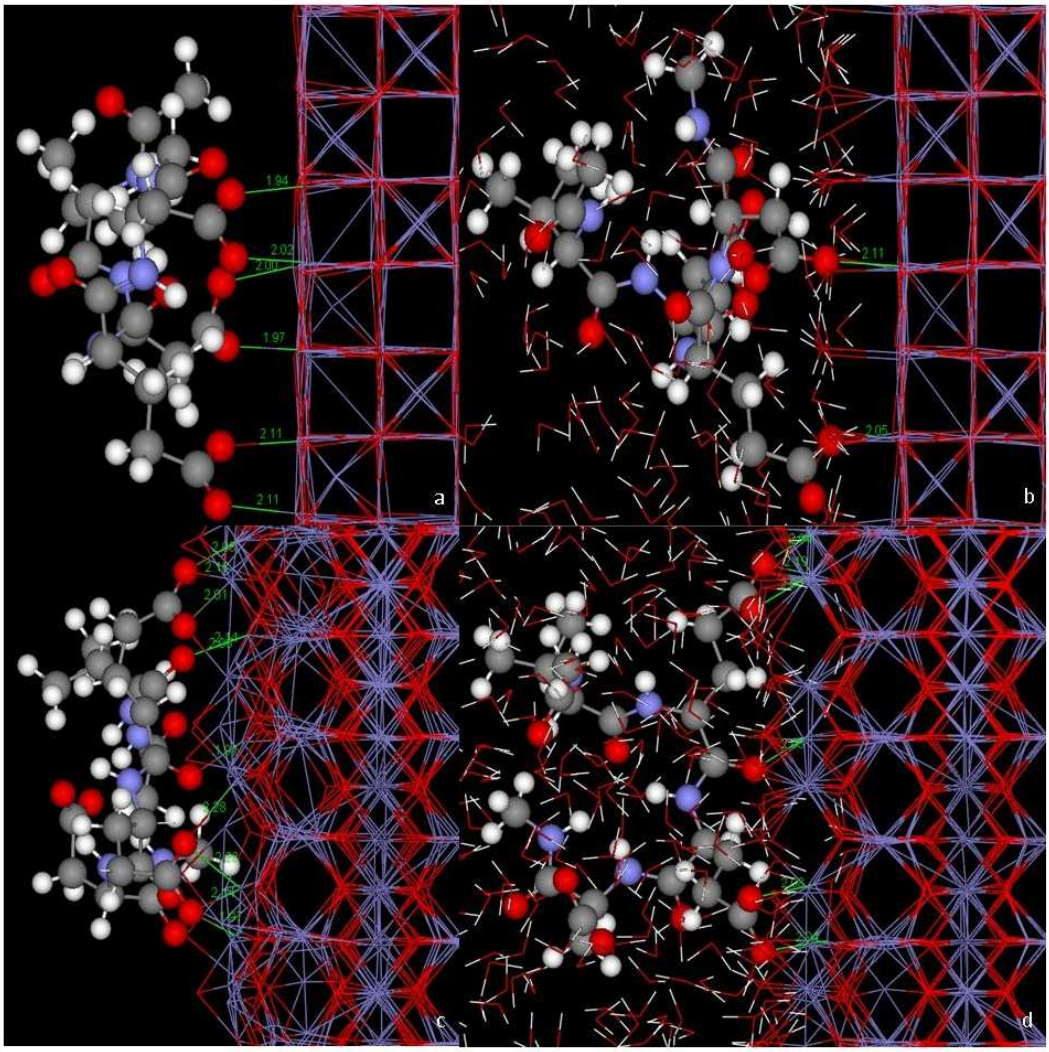

Figure 5.9: Bond distance images for the *in vacu* and solvated systems of DEEV. a){100} *in vacu*, b){100} solvated, c){111} *in vacu* and d){111} solvated.

for both surface types, however, on the {100} surface there was a mixture of side chain oxygen and peptide bond oxygen binding exhibited, whilst on the {111} surface, only side chain oxygen bonding was present. The average bond lengths for an iron to a tetrapeptide oxygen in DAEV were; 2.00 Å and 2.00 Å on the {100} surface for the *in vacu* and solvated systems respectively, and 2.05 Å and 2.08 Å on the {111} surface for the *in vacu* and solvated systems respectively. The average bond lengths for an iron to a tetrapeptide oxygen in DEAV were; 2.02 Å and 2.13 Å on the {100} surface for the *in vacu* and solvated systems respectively, and 2.01 Å and 2.10 Å on the {111} surface for the *in vacu* and solvated systems respectively.

On comparison of the single substitution sequences, as expected, with the addition of extra amino acid residues more bonding was present. In the dipeptide, an N–terminal

peptide bond oxygen bond and 2/3 glutamate side chain oxygen bonds were present for the *in vacu* systems, whereas, for the tetrapeptides, DAEV also exhibited this bonding but also presented with bonding from the aspartate side chain oxygens and the inter–residue peptide bond oxygens. DEAV also exhibits these bonding types, with the exception of the N–terminal peptide bond oxygen, as only the {111} surface exhibited with this.

When considering the double substitution mutant, more bonding is present in the {111} surface systems than the {100}. Both surfaces exhibit aspartate side chain oxygen bonding but to differing extents, particularly for the solvated systems, with more bonding in the solvated {111} system than in the {111} *in vacu* system. Peptide bond oxygen bonding was also present for both (NBO and D-A PBO for the {100} surface, and NBO and A-V PBO for the {111}). As expected, bonding decreased greatly with the double substitution. The average bond lengths for an iron to a tetrapeptide oxygen in DAAV is; 2.03 Å and 2.24 Å on the *in vacu* surface for the *in vacu* and solvated systems respectively, and 2.10 Å and 2.07 Å on the {111} surface for the *in vacu* and solvated systems respectively.

On comparison of the fully mutated tetrapeptide bond distance data with its dipeptide counterpart, more bonding is apparent with the addition of extra amino acid residues, due to the presence of further potential oxygen binding sites, with substantial aspartate side chain oxygen bonding being seen, particularly on the {111} surface, suggesting that as sequence length increases as does iron binding potential.

Tables 5.11, 5.12, 5.13, 5.14 and 5.15 represent the residence times and related average coordination numbers for bonding between all oxygen present in the tetrapeptides and tetrahedral and/or octahedral iron on both the {100} and {111} magnetite surfaces.

		FET *in vacu*		FEO *in vacu*		FET solvated		FEO solvated	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
DEEV {100} surface	E-V PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	4998.00	1.00	0.00	0.00	144.57	0.57	0.00	0.00
	E O	1.00	0.00	4997.00	1.00	0.00	0.00	14.17	0.05
	E-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	1835.57	1.01	0.00	0.00	999.00	0.99
	E O	0.00	0.00	1852.14	1.00	0.00	0.00	316.35	0.90
	D-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D O	4993.00	1.00	0.00	0.00	20.03	0.60	0.00	0.00
	D O	0.00	0.00	4993.00	1.00	0.00	0.00	999.00	1.00
	NBO	3.50	0.01	57.18	0.43	0.00	0.00	34.02	0.42

Table 5.11: Residence times (RT) and average coordination numbers (av. CN) of the DEEV sequence. (D O = side chain oxygen of aspartate E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		FET *in vacu*		FEO *in vacu*		FET solvated		FEO solvated	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
DEEV {111} surface	E-V PBO	0.00	0.00	4929.61	1.00	0.00	0.00	19.09	0.12
	E O	0.00	0.00	3021.27	1.86	0.00	0.00	664.65	1.98
	E O	0.00	0.00	3022.16	1.86	0.00	0.00	221.90	1.74
	E-E PBO	0.00	0.00	4906.94	1.06	0.00	0.00	707.02	1.53
	E O	0.00	0.00	1931.81	1.14	0.00	0.00	975.07	1.03
	E O	0.00	0.00	1843.62	1.13	0.00	0.00	981.22	1.01
	D-E PBO	0.00	0.00	4983.00	1.00	0.00	0.00	236.62	0.47
	D O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NBO	0.00	0.00	1979.52	1.00	0.00	0.00	33.96	0.11

Table 5.12: Continued. Residence times (RT) and average coordination numbers (av. CN) of the DEEV sequence. (D O = side chain oxygen of aspartate E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		FET *in vacu*		FEO *in vacu*		FET solvated		FEO solvated	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
DAEV {100} surface	E-V PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	2027.11	1.00	0.00	0.00	93.93	0.19
	E O	0.00	0.00	4992.47	1.00	0.00	0.00	72.70	0.14
	A-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D-A PBO	173.20	0.83	0.00	0.00	112.64	0.51	0.00	0.00
	D O	4625.86	0.97	31.70	0.02	0.00	0.00	0.00	0.00
	D O	4216.05	0.99	16.75	0.01	999.00	1.00	0.00	0.00
	NBO	0.00	0.00	4781.00	0.96	0.00	0.00	22.00	0.49
DAEV {111} surface	E-V PBO	0.00	0.00	4924.33	1.01	0.00	0.00	139.58	0.95
	E O	0.00	0.00	2781.10	1.99	0.00	0.00	992.02	1.00
	E O	0.00	0.00	3181.39	2.00	0.00	0.00	999.00	2.00
	A-E PBO	0.00	0.00	4987.00	1.00	0.00	0.00	999.00	1.00
	D-A PBO	0.00	0.00	4334.18	1.00	0.00	0.00	999.00	1.00
	D O	0.00	0.00	4639.00	0.93	0.00	0.00	999.00	1.00
	D O	0.00	0.00	0.00	0.00	0.00	0.00	4.82	0.04
	NBO	0.00	0.00	4324.62	1.00	0.00	0.00	999.00	1.00

Table 5.13: Residence times (RT) and average coordination numbers (av. CN) of the DAEV sequence. (D O = side chain oxygen of aspartate E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		FET *in vacu*		FEO *in vacu*		FET solvated		FEO solvated	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
DEAV {100} surface	A-V PBO	6.48	0.04	1.20	0.08	0.00	0.00	2.07	0.05
	E-A PBO	74.66	0.87	1.00	0.00	255.36	0.94	0.00	0.00
	E O	4998.00	1.00	0.00	0.00	999.00	1.00	0.00	0.00
	E O	1.19	0.07	4998.00	1.00	2.86	0.02	95.83	0.94
	D-E PBO	0.00	0.00	4994.00	1.00	0.00	0.00	999.00	0.98
	D O	0.00	0.00	4969.25	1.00	48.99	0.61	11.50	0.02
	D O	4528.00	0.91	233.61	0.09	11.50	0.02	6.70	0.04
	NBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DEAV {111} surface	A-V PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E-A PBO	0.00	0.00	4427.27	0.95	0.00	0.00	35.10	0.62
	E O	0.00	0.00	4644.14	0.99	0.00	0.00	665.46	1.54
	E O	0.00	0.00	3824.66	1.00	0.00	0.00	310.16	1.00
	D-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D O	0.00	0.00	4512.05	1.07	0.00	0.00	604.82	1.89
	D O	0.00	0.00	4166.74	1.19	0.00	0.00	247.75	0.90
	NBO	0.00	0.00	2009.25	0.99	0.00	0.00	118.17	0.23

Table 5.14: Residence times (RT) and average coordination numbers (av. CN) of the DEAV sequence. (D O = side chain oxygen of aspartate E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		FET *in vacu*		FEO *in vacu*		FET solvated		FEO solvated	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
DAAV {100} surface	A-V PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	A-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D-A PBO	4654.95	0.93	0.00	0.00	16.26	0.48	6.20	0.06
	D O	0.00	0.00	4998.00	1.00	0.00	0.00	329.38	0.99
	D O	4998.00	1.00	0.00	0.00	9.90	0.04	0.00	0.00
	NBO	1215.60	0.92	154.46	0.06	20.51	0.22	0.00	0.00
DAAV {111} surface	A-V PBO	0.00	0.00	825.42	0.91	0.00	0.00	145.77	1.32
	A-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D O	0.00	0.00	4979.66	1.00	0.00	0.00	904.91	1.95
	D O	0.00	0.00	4999.00	1.00	0.00	0.00	517.61	1.65
	NBO	0.00	0.00	2652.98	1.04	0.00	0.00	49.14	0.32

Table 5.15: Residence times (RT) and average coordination numbers (av. CN) of the DAAV sequence. (D O = side chain oxygen of aspartate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

The DEEV residence time data shows that, when the {100} surface was *in vacu*, bonding between tetrahedral iron and the side chain oxygen of aspartate and the C–terminal glutamate was exhibited, lasting over 99% of the simulation length. Bonding for this length of time was also seen between octahedral iron and the remaining side chain oxygens of aspartate and the C–terminal glutamate. No bonding existed for any inter–residue peptide bond oxygen and only octahedral bonding lasting under 1.2% of the simulation length was shown for the N–terminal glutamate side chain oxygens. When this system was solvated, only the octahedral iron–aspartate side chain oxygen bond remained lasting over 99% of the simulation length. The residence times of all other *in vacu* bonds lasting over 99% of the simulation length drop significantly, showing no bonding lasting over 32% of the simulation length, with the exception of an octahedral iron bond to one of the oxygen of the N–terminal glutamate, which, upon solvation, lasted over 99% of the simulation length, suggesting that the bond formed in the latter stages of the *in vacu* simulation and remained when the system was solvated.

When this data was compared to the equivalent dipeptide data, for the tetrapeptides, more bonding was seen, in particular more bonding lasting over 99% of the simulation length was exhibited, this is to be expected with the addition of the extra amino acids, in particular the addition of aspartate. The trend of no peptide bond oxygen bonding, except for NBO, was also apparent for the dipeptide on the {100} surface. Different trends in bonding lasting over 99% of the simulation length for the solvated systems are exhibited, for EE, the bonds were to tetrahedral iron, whereas, for DEEV, they were to octahedral iron.

When examining the {111} surface systems for DEEV, it can be seen that no tetrahedral iron bonding exists. Octahedral iron bonding lasting over 98% of the simulation length was exhibited to the D-E, E-E and E-V peptide bond oxygens when the system was *in vacu*, however, upon solvation, the residence times of these bonds drop significantly, with the longest retention time being less than 71% of the simulation length. Solvation of the system also leads to bonds between octahedral iron and the N–terminal glutamate side chain oxygens lasting over 97% of the simulation length. The equivalent *in vacu* bonds last approximately 38% of the simulation length, suggestive of them forming in the latter stages of the simulation and remaining once solvated.

When this data was compared to the equivalent dipeptide data, no tetrahedral iron bonding was present for either peptide type. The octahedral iron bonding trends differ greatly. For the *in vacu* systems, EE exhibits bonding lasting over 99% of the simulation length

to all potential binding oxygen, with the exception to the N–terminal peptide bond oxygen, whereas, DEEV shows only inter–residue peptide bond oxygen bonds lasting for this length of time. On addition of water, EE shows no bonding lasting over 90% of the simulation length, whilst DEEV exhibits two bonds lasting over 97% of the simulation length. This data suggests that the addition of more potential binding sites to the sequence has a greater beneficial effect on the iron binding affinity of the solvated system as opposed to the *in vacu*.

The residence time data for the single substituted tetrapeptide mutants on the {100} surface shows that, for the DAEV *in vacu* system, tetrahedral iron bonding lasting over 84% of the simulation length was exhibited for both side chain oxygen of aspartate, and octahedral iron bonding lasting over 95% of the simulation length were seen to a side chain oxygen of glutamate and the N–terminal peptide bond oxygen. When solvated, the residence times of these bonds drop (the biggest drop being a 93% drop in length), with the exception of the tetrahedral iron bond to the aspartate side chain oxygen, which is present for over 99% of the simulation length. For DEAV on the {100} surface, more bonding was present lasting over 90% of the simulation length than with DAEV, suggesting that the iron ions bind preferentially to DEAV. For the *in vacu* system, tetrahedral iron bonding lasting over 90% of the simulation length was exhibited to one of the side chain oxygen of aspartate and one of the side chain oxygen of glutamate, and octahedral iron bonding lasting over 99% of the simulation length were shown to the remaining side chain oxygens of aspartate and glutamate and to the D-E peptide bond oxygen. When this system is solvated, the majority of the residence times of the *in vacu* bonds drop (the biggest drop being a 98% drop in length), with the exception of the tetrahedral iron bond to a glutamate side chain oxygen and the octahedral iron bond to the D-E peptide bond, which lasted for over 99% of the simulation length.

On comparison of AE against DAEV and DEAV on the {100} surface, more bonding was present, with an increase in bonding lasting over 99% of the simulation length also being seen for the tetrapeptides, in particular DEAV, which was to be expected with the addition of extra potential iron binding sites. The lack of tetrahedral iron bonding to glutamate oxygen and A-E peptide bond oxygen exhibited in AE is also seen in DAEV, as the majority of the tetrahedral iron bonding present involves the aspartate residue, however, this is not so for DEAV, as a tetrahedral iron bond to a glutamate side chain oxygen is exhibited lasting over 99% of the simulation length of both the *in vacu* and solvated simulations, suggesting that this type of bonding is sequence dependent. With

the octahedral iron bonding for AE, each of the oxygen of glutamate has a bond that lasts over 99% of the simulation length for the *in vacu* system, however, with the addition of the extra amino acids, this strength of bonding of glutamate diminishes, as for both tetrapeptides only one of the two glutamate oxygen have bonds lasting over 99% of the simulation length. Also, for the solvated systems, no octahedral iron bonding lasting over 24% and 9% of the simulation length was present for either AE or DAEV, respectively; yet, DEAV exhibits a bond which last over 99% of the simulation length to the D-E peptide bond oxygen.

When the single substituted tetrapeptide {111} surface systems were examined it was found that, no tetrahedral iron bonding was present for either DAEV or DEAV. For the octahedral iron *in vacu* system, with DAEV, bonds lasting over 86% of the simulation length were found to all peptide bond oxygen and to one side chain oxygen of aspartate. Upon solvation, the residence time of the bond to E-V peptide bond oxygen dropped by 85%, however, the other bonds that lasted over 86% of the simulation length in the *in vacu* system lasted over 99% of the simulation length when solvated. Additionally, bonds between octahedral iron and both oxygen of glutamate lasting over 99% of the simulation length were present for the solvated system. For the {111} surface, more bonding was present lasting over 83% of the simulation length for DAEV than DEAV, suggesting that the iron ions bind preferentially to DAEV. For the DEAV *in vacu* system, bonds were present between octahedral iron and both side chain oxygen of aspartate, the one side chain oxygen of glutamate, and E-A peptide bond oxygen, lasting over 83% of the simulation length. However, upon solvation of the system, the residence times of these bonds drop significantly (the biggest drop being an 85% drop in length).

When the tetrapeptide {111} surface data was compared to its dipeptide counterpart data it was shown that, all single alanine substituted sequences exhibit no tetrahedral iron bonding. With octahedral iron bonding, the addition of extra amino acids leads to increased bonding for the tetrapeptides, as would be expected. The DAEV residence time data did not follow the trend of AE for the *in vacu* system, with neither of the glutamate side chain oxygen of DAEV showing bonding lasting over 64% of the simulation length, whilst, for AE these bonds lasted over 99% of the simulation length, and the A-E peptide bond oxygen bond in DAEV lasted over 99% of the simulation length, whereas, for AE this bond was not present. For the corresponding solvated systems, the trend of AE was again not followed, as all but two possible bonds on DAEV lasted over 99% of the simulation length, however, there was no bonding lasting over 40% of the simulation length in the AE

solvated system. The DEAV residence time data also did not follow the trend of AE for the *in vacu* system, as only one bond was present in both systems lasting over 92% of the simulation length (between octahedral iron and one of the side chain oxygen of glutamate). When the DEAV system was solvated it showed a similar trend to AE, showing no long lasting bonding (the longest residence times were 665.46 ps and 399.66 ps for DEAV and AE, respectively). This data suggests that the addition of extra amino acids, changes the way in which the sequence binds to iron ions.

The residence time data for the alanine double substitution on the {100} surface it was shown that, there was no A-A or A-V peptide bond oxygen bonding present. There was also no long term iron bonding for the solvated systems, with the longest residence time being 329.38 ps. For the *in vacu* systems, bonding lasting over 93% of the simulation length was exhibited between tetrahedral iron and one of the side chain oxygen of aspartate and D-A peptide bond oxygen, and between octahedral iron and the other side chain oxygen of aspartate.

On comparison with the {100} surface data of the double substituted dipeptide it was seen that, there was an increase in bonding for DAAV with the addition of extra amino acid residues. Neither DAAV nor AA exhibited any A-A peptide bond oxygen bonding, and the N–terminal peptide bond oxygen bonding is minimal for all systems.

With the {111} surface there was no D-A or A-A peptide bond oxygen bonding present. As with the other tetrapeptides, there was also no tetrahedral iron bonding is present. With the octahedral iron bonding, for the *in vacu* system, both side chain oxygen of aspartate have bonds lasting over 99% of the simulation length, however, upon solvation, only the bond to the second oxygen of aspartate continues to last over 90% of the simulation length.

On comparison of the {111} surface data of DAAV with that of AA, it was shown that, very different trends were exhibited, as for the *in vacu* octahedral bonding of AA, bonds lasting over 99% of the simulation length were present for the N–terminal peptide bond oxygen and the A-A peptide bond oxygen, however, with the addition of extra amino acid residues, there was no A-A peptide bond oxygen bonding present and the N–terminal peptide bond oxygen bond last only 2652.98 ps. This data suggests that the addition of the extra amino acid residues, effects how the residues from the dipeptide sequence react with the magnetite iron.

The interfacial energies of the tetrapeptides on the magnetite surfaces can be found in Table 5.16.

{100} surface attachment				
	E_{int} of s-p (eV)	E_{diff} of s-p (eV)	E_{int} of s-p-w (eV)	E_{diff} of s-p-w (eV)
DEEV	-6.16	2.65	8.21	0.00
DAEV	-8.25	4.74	7.76	0.45
DEAV	-5.24	1.73	8.09	0.12
DAAV	-3.51	0.00	6.19	2.02
{111} surface attachment				
	E_{int} of s-p (eV)	E_{diff} of s-p (eV)	E_{int} of s-p-w (eV)	E_{diff} of s-p-w (eV)
DEEV	-10.13	2.94	15.21	0.32
DAEV	-13.23	6.04	15.52	0.00
DEAV	-9.70	2.51	10.57	4.96
DAAV	-7.19	0.00	15.14	0.38

Table 5.16: Interfacial energies for all tetrapeptides on the {100} and {111} surface, using the Yang *et al* method[164].E_{int} of s-p is the interfacial energy of the slab–peptide system. E_{diff} of s-p is the difference in interfacial energy from the highest interfacial energy of the slab–peptide system. E_{int} of s-p-w is the interfacial energy of the slab–peptide–water system. E_{diff} of s-p-w is the difference in interfacial energy from the highest interfacial energy of the slab–peptide–water system.

The solvated tetrapeptide interfacial energy data shows that, the {100} surface produces lower interfacial energy values (between 6.19 eV and 8.21 eV) than the {111} surface (between 10.57 eV and 15.52 eV). For the {100} solvated surface, the DEEV sequence produces the highest interfacial energy value, whereas the fully mutated sequence produces the lowest interfacial energy value, suggestive of the presence of glutamate in the sequence leading to a reduced iron binding potential, particularly if the glutamate is next to aspartate in the sequence. For the solvated {111} system, the values for the DEEV, DAEV and DAAV are similar, with a much smaller value exhibited for DEAV, suggesting that this particular mutation improves the iron binding potential compare to the original sequence.

This data also shows that, solvating the system has an unfavourable effect on the interfacial energy, as the observed energies are again considerably lower for the *in vacu* system. The solvation of the system also affects the interfacial behaviour of the tetrapeptides, as the trends seen for the *in vacu* systems are very dissimilar to the solvated systems. DAEV consistently produces the lowest energy for the *in vacu* systems, followed by DEEV, and then interfacial energy increases with each subsequent alanine substitution. The increase in interfacial energy suggests potential competitive binding of the magnetite iron surface

to the water oxygen.

On comparison of the dipeptide interfacial energy data with the tetrapeptide data, it can be seen that for the solvated systems the tetrapeptides produce higher interfacial energy values than their dipeptide counterparts. These results suggest that the addition of the extra amino acid residues to the sequence reduce the iron binding potential of the sequences. However, for the *in vacu* systems, the opposite is seen as the tetrapeptides produce lower interfacial energies than their dipeptide counterparts, again proposing that the presence of water in the systems is detrimental to the iron binding, providing competitive oxygen for the magnetite iron to bind with. This data also suggests that in the *in vacu* system, the addition of the extra amino acids actually enhances the iron binding potential of the sequence.

As with the dipeptides, there are concerns with this direct method of calculating the adsorption energy, possible reasons for these were discussed previously for the dipeptides. An alternative indirect method of adsorption energy calculation is explored in the following section. This method utilises a constrained molecular dynamics version of the systems, focusing on the Potential of Mean Force and free energy evaluation.

5.3 Constrained system

5.3.1 Computational Methods

The methods used for this section of work are as detailed in Chapter 4.3.1. A series of Potential of Mean Force simulations were run to determine the iron binding affinity of the capped peptide chains, previously relaxed with SANDER, PMEMD and DL_POLY, then attached to the magnetite {100} or {111} surface and solvated. Each PMF calculation was run for 1 ns NVT at 300 K and the average force was integrated with respect to the constraint distance to produce the free energy of binding[139].

5.3.2 Results and Discussion

Free energy profiles were created for the aforementioned systems. Within the free energy profiles, 0 Å on the x-axis is an approximate representation of where the magnetite surface lies based on the first strongest peak for an FEO iron ion in the related density profile. The free energy of the simulation system as a function of peptide chain centre of mass (CoM)

distance from the magnetite surface for the dipeptides (EE, AE, AA) and tetrapeptides (DEEV, DAEV, DEAV and DAAV) attached to the {100} and {111} surfaces are shown in the free energy profiles. Peaks in the water density profiles indicate areas within the system which contain a high concentration of water molecules.

Dipeptide Attachment

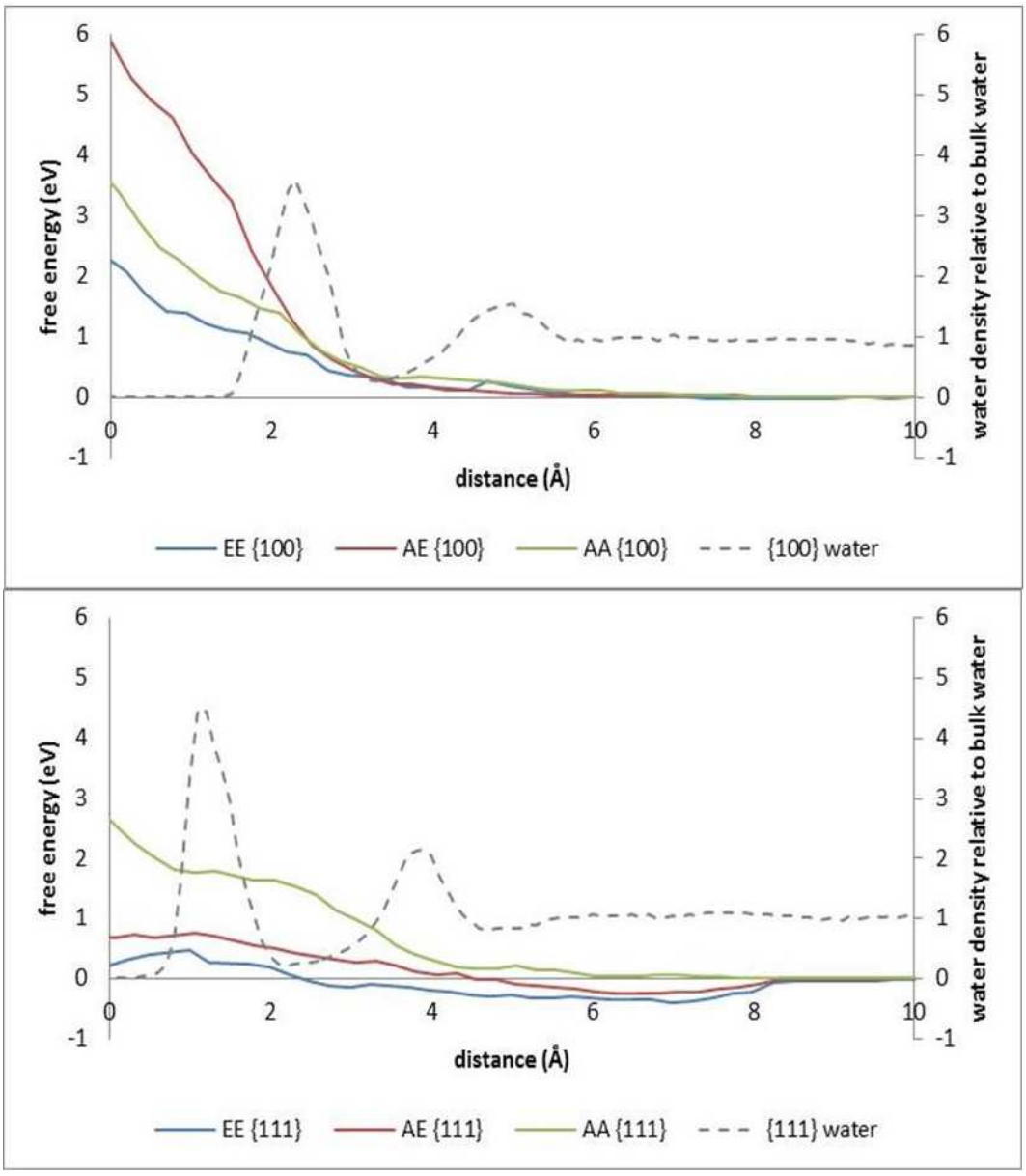

Figure 5.10: A comparison of the free energy profiles of the dipeptides. Distance refers to distance between the peptide CoM and the magnetite surface. The dashed line represents the water density profile for the system.

From the free energy profile results for the different dipeptides (Figure 5.10) it was evident that, on the {100} surface, the double glutamic acid to alanine substitution does appear to

reduce the iron binding activity of the sequence. Nevertheless, the AE sequence produces a significantly higher free energy profile, suggesting that a single E to A substitution confers even greater reductive effect on iron binding. Between 10 Å and 5.5 Å is optimal for attachment, after which point, an increase in free energy was detected, corresponding with the presence of the intense water layers.

Whereas, with the {111} surface energy profile, a different free energy trend is seen, with AA exhibiting the highest free energy values, proposing that iron binding affinity is surface dependent. For AA, between 10 Å and 5.5 Å is optimal for attachment, after which point, an increase in free energy was detected, corresponding with the presence of the intense water layers. For EE and AE, the maximal free energies correspond to the distance at which the larger water peak occurs, suggesting that the water layers affect the free energy profiles, however, as the maximal free energies are low (0.49 eV and 0.76 eV for EE and AE, respectively) this water influence is minimal. Between 8 Å and 2.5 Å and 8 Å and 5 Å are optimal for attachment for EE and AE, respectively, as the profiles exhibit negative values within this region.

Upon examination of the energy minima for the dipeptide sequences (Figure 5.11) it can be observed that, on the {100} surface, only the original sequence, EE, exhibited a negative value for its energy minima, producing a value of -0.0310 eV at 7.9 Å, suggesting a preferable attachment distance. On the {111} surface, all of the dipeptide sequences except for the double substituted sequence, AA, exhibited a negative value for their energy minima, with EE producing the lowest value at -0.3831 eV. The energy minima for these residues fall between 6.6 Å and 7 Å, suggesting that attachment of the residue centre of mass at this distance would be preferable.

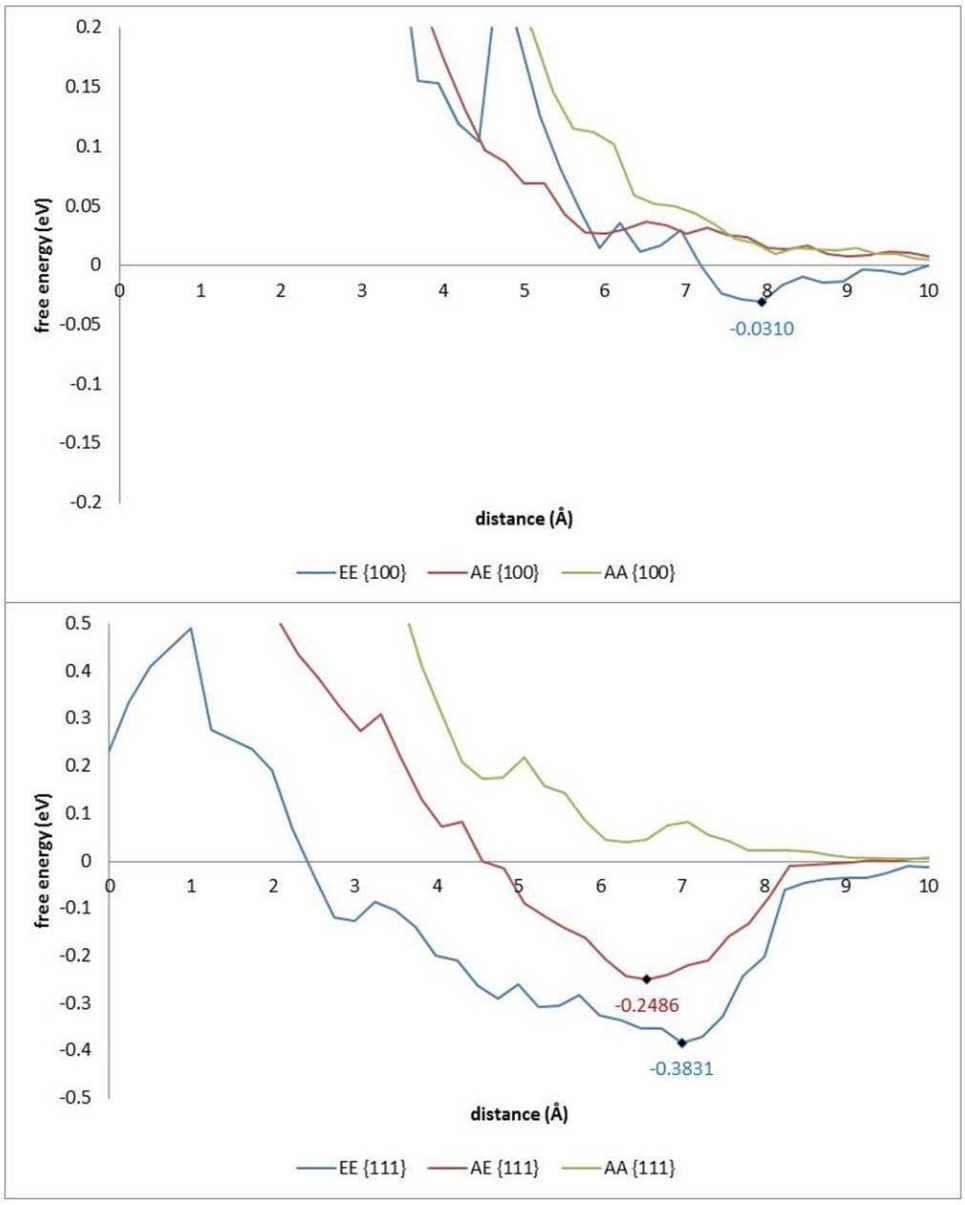

Figure 5.11: A close-up comparison of the free energy profiles of the dipeptides between -0.5 eV and 0.5 eV, exhibiting the free energy minima.

Tetrapeptide Attachment

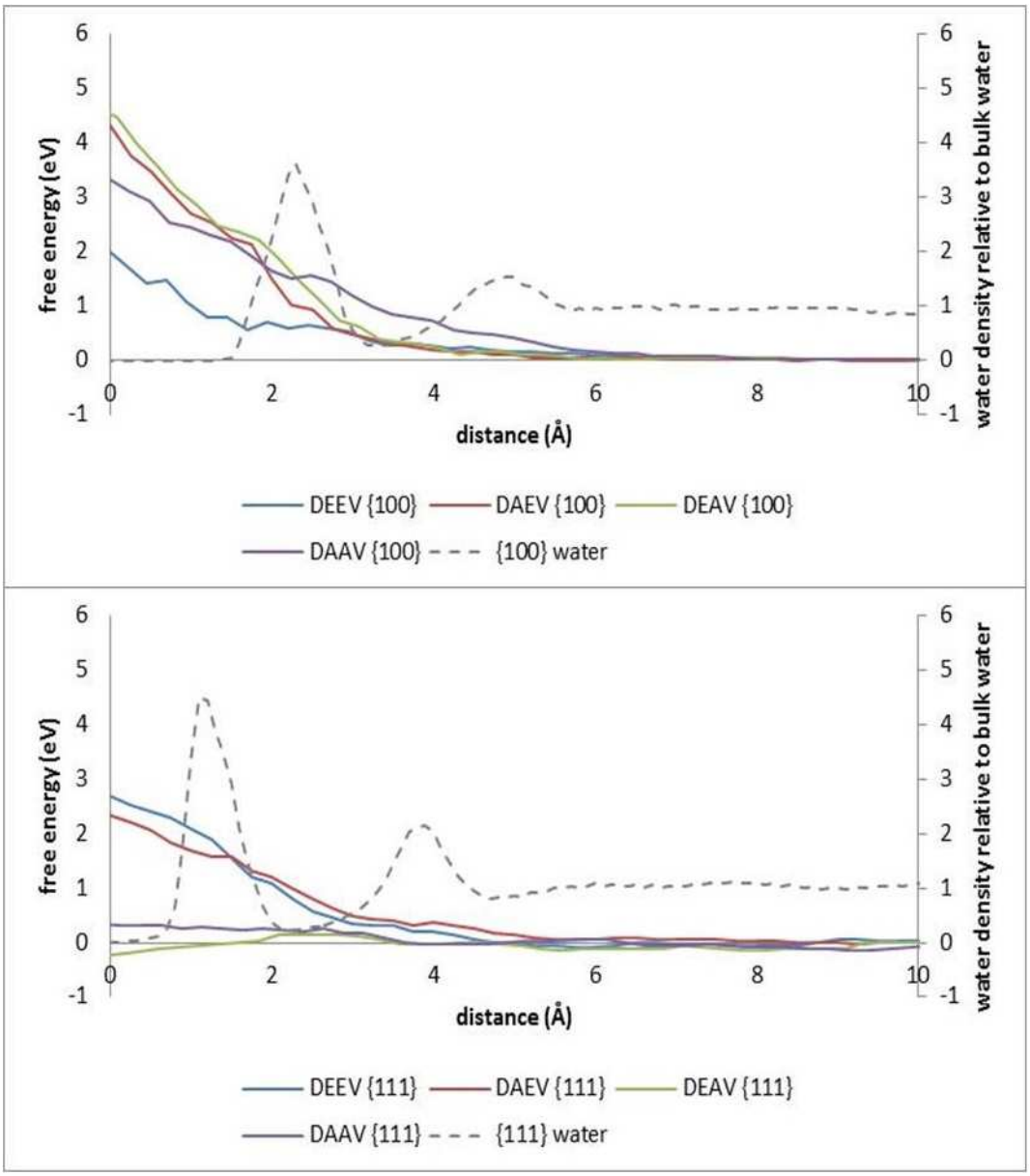

Figure 5.12: A comparison of the free energy profiles of the tetrapeptides. Distance refers to distance between the peptide CoM and the magnetite surface. The dashed line represents the water density profile for the system.

From the free energy profile results for the different tetrapeptides (Figure 5.12) it was shown that, on the {100} surface, a single E to A substitution (both DEEV to DAEV and DEEV to DEAV) produces a considerable increase in the maximal free energy compared to that of DEEV, suggesting that irrespective of its position within the sequence, a single E to A mutation has a significant reductive effect on iron binding. In addition, this reductive effect was also observed when a double substitution occurs (DEEV to DAAV), however, to a lesser extent. Between 10 Å and 6 Å is optimal for attachment, after which point,

an increase in free energy was detected, corresponding with the presence of the intense water layers. The initial increase in free energy corresponds with the distance at which the smaller of the two water peaks occurs. The increase in free energy at this distance was greater for DAAV than for the other tetrapeptides, which exhibit a gradual increase in energy, however, the increase for this sequence remains at this rate as the CoM gets closer to the surface. A further increase in free energy, corresponding with the distance at which the larger water peak occurs, was displayed for the single substituted sequences.

This data relates to that of the dipeptides on the {100} surface, suggesting that for this particular surface a single substitution has a greater reductive effect on iron binding ability than a double substitution. The EE and DEEV, and the AA and DAAV sequences show similar free energy profiles on the {100} surface, therefore, the addition of D and V to EE and AA appears to have had little, if any, effect. The addition of D and V to AE, to form either DAEV or DEAV, shows a reduction in free energy from the AE sequence for both variations, suggesting that the extra amino acids increase the iron binding ability of the AE amino acid sequence, irrespective of the sequence configuration.

Whereas, with the {111} surface energy profile, a different free energy trend is seen, with DEEV and DAEV producing similar free energy profiles, whilst, DEAV and DAAV exhibit greatly decreased free energy values, with DEAV actually producing a negative free energy value when the CoM was closest to the surface. This data suggests that the DEAV and DAAV mutations have much higher iron binding affinities than the original DEEV sequence. For DEEV and DAEV, between 10 Å and 5.5 Å is optimal for attachment, after which point, an increase in free energy was detected, corresponding with the presence of the intense water layers. This is similar for DAAV, however, the optimal attachment region is extended to 3.5 Å. For DEAV, between 1.5 Å and 0 Å is optimal for attachment, as the profiles exhibit negative values within this region, proposing that this is the preferential sequence for surface attachment. For DEAV and DAAV sequences, the water density of the system appears to have little or no effect on the free energy profiles. It can be deduced from the results that a single alanine substitution has differing effects on the iron binding affinity dependent on where the substitution is within the sequence, with DAEV exhibiting a minimal effect on free energy when compared to DEEV, whilst DEAV appears to significantly enhance iron binding ability. From this data it can be deduced that the E adjacent to the V has the greatest influence on the iron binding as when this was substituted, in both DAAV and DEAV, a very large decrease in free energy was displayed.

On comparison of the dipeptide and tetrapeptide {111} surface data it was shown that,

the EE and DEEV sequences show differing free energy profiles in that the addition of extra amino acids increases the effect that the water layers have on sequence attachment, with EE exhibiting minimal influence on the free energy profiles from the water layers, whilst DEEV displayed an increase in free energy corresponding with the presence of the intense water layers. This suggests that the addition of D and V to EE on the {111} surface has had a detrimental effect on iron binding ability. This occurrence is also true of the AE and DAEV sequences. For the AE and DEAV sequences, the addition of D and V to AE reduces the free energy on the {111} surface, with the profile exhibiting no obvious influence from the water layers, suggesting again that the effects seen are greatly sequence dependent. The AA and DAAV sequences produce significantly different free energy profiles; with the AA profile demonstrating that sequence attachment is more favourable as the CoM distance from the surface increases, whereas the DAAV profile shows sequence affinity is independent of distance from the surface. This implies that the addition of D and V to AA on the {111} surface has increased the iron binding ability of the sequence.

Upon examination of the energy minima for the tetrapeptide sequences (Figure 5.13) it can be observed that, on the {100} surface, only the original sequence, DEEV, and the single substituted sequence, DAEV, exhibited a negative value for their energy minima, with DEEV producing the lowest value at -0.0164 eV. The energy minima for these residues fall at approximately 9.5 Å , suggesting that attachment of the residue centre of mass at this distance would be preferable. On the {111} surface, all of the tetrapeptide sequences exhibited a negative value for their energy minima, with DEAV producing the lowest value at -0.2690 eV. The trend of favourable attachment distance is very different from the {100} surface, with DEEV, DAEV, DEAV and DAAV exhibiting centre of mass distances of 6 Å , 9.3 Å , 0 Å and 9.1 Å respectively.

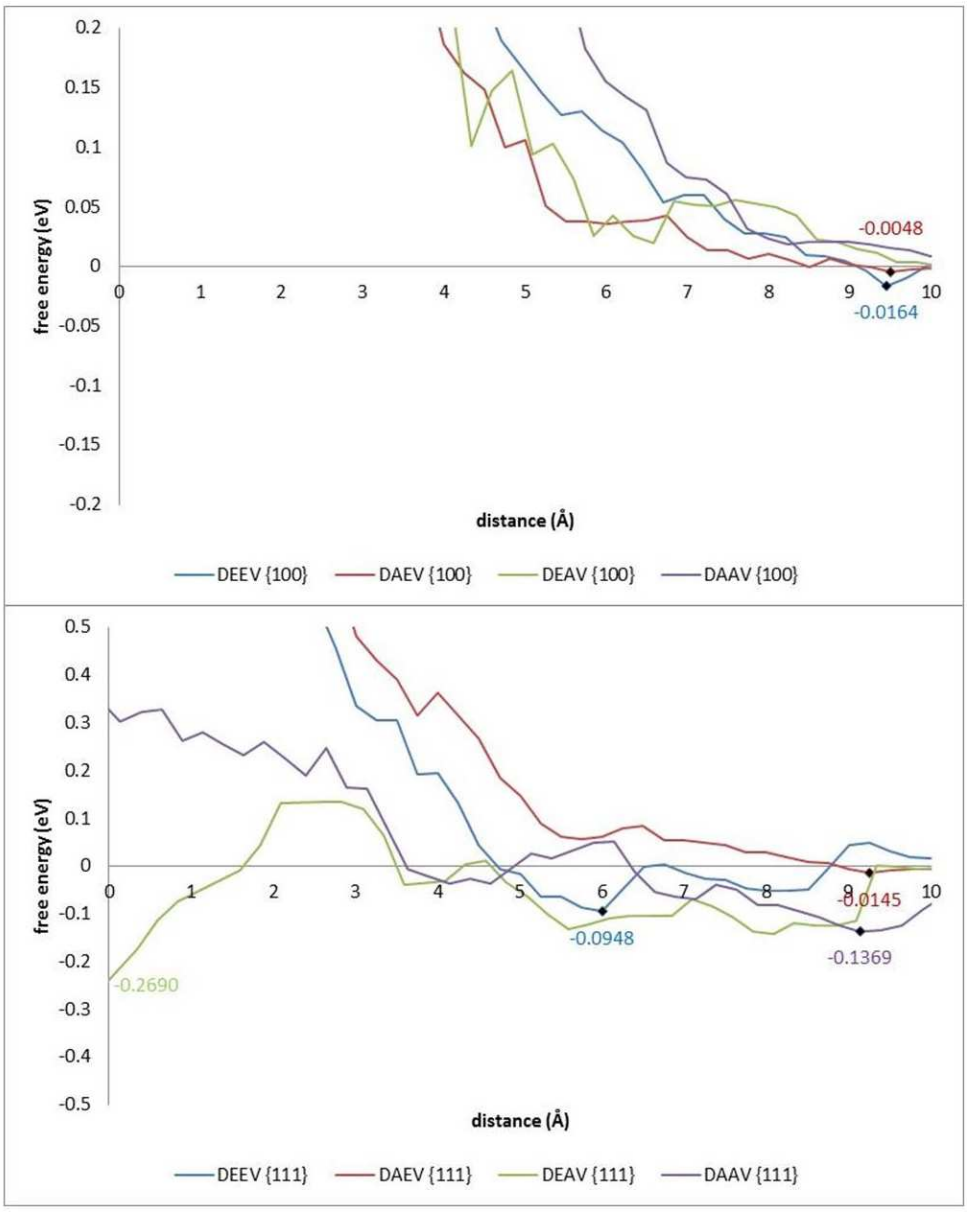

Figure 5.13: A close-up comparison of the free energy profiles of the tetrapeptides between -0.5 eV and 0.5 eV, exhibiting the free energy minima.

Surface Free Energy

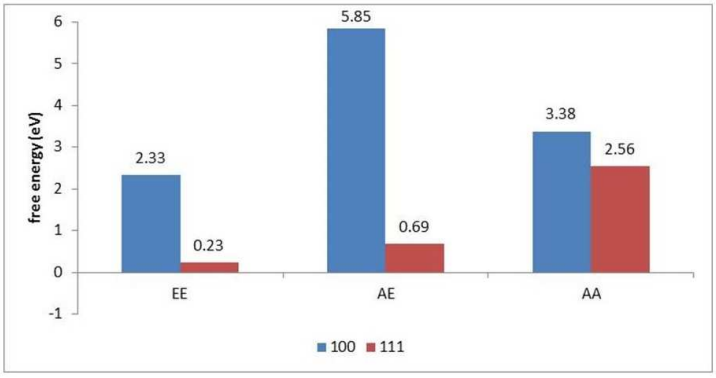

Figure 5.14: A comparison of the free energy values for the dipeptide CoM at the magnetite surface for both the {100} and {111} surfaces.

When the free energy values for the dipeptides centre of mass at the magnetite surface were examined (Figure 5.14) it was shown that, on the {100} surface, the EE sequence exhibits a free energy value of 2.33 eV, whereas, on the {111} surface, its value was significantly lower (0.23 eV). Similarly, the AE and AA sequences demonstrated free energy values of 5.85 eV and 3.38 eV, respectively, on the {100} surface, but much lower values on the {111} surface (0.69 eV and 2.56 eV, respectively). These data show that the {111} surface has lower free energy values, representing greater iron binding, suggesting that this is the preferential surface of attachment.

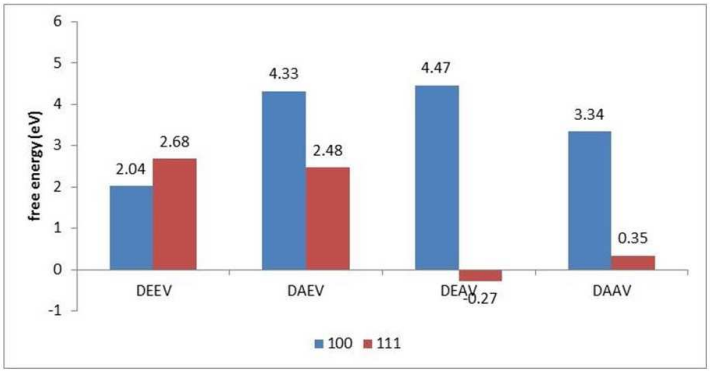

Figure 5.15: A comparison of the free energy values for the four amino acid chains at the magnetite surface for both the {100} and {111} surfaces.

The free energy values for the tetrapeptides centre of mass at the magnetite surfaces are shown in Figure 5.15. On the {100} surface, the DEEV sequence exhibits a free energy value of 2.04 eV, whereas, on the {111} surface, its value is higher, 2.68 eV. Conversely, the single substitution mutants, DAEV and DEAV, show free energy values of 4.33 eV

and 4.47 eV, respectively, on the {100} surface, but exhibit largely decreased values on the {111} surface (2.48 eV and -0.27 eV, respectively). Similarly, on the {100} surface, the double mutation (DAAV) gives a free energy value of 3.34 eV, but a reduced value of 0.35 eV on the {111} surface. These data show that, with the exception of DEEV, the tetrapeptides have a higher iron binding affinity for the {111} surface.

On the {100} surface, the addition of D and V to AE (giving both DAEV and DEAV) shows an increase in iron binding ability, suggesting that the extra amino acids increase the affinity of the sequences for the {100} surface. Equally, on the {111} surface, a significant decrease in free energy was seen for the DEAV sequence from the AE sequence value, suggesting again that the extra amino acids improve iron binding ability. However, for the DAEV sequence, an increase in free energy from the value for AE is exhibited, proposing that for this sequence the addition of D and V is not beneficial to iron binding. The addition of D and V to AA shows no effect on binding affinity to the {100} surface, whereas on the {111} surface binding affinity increases thus confirming that the addition of these particular extra amino acids improves the affinity for iron binding. In the formation of DEEV, on the {100} surface, an increase in the attraction between surface and amino acid sequence was displayed. Yet, the opposite was true for the {111} surface, consequently, the {100} surface shows more favourable sequence attachment when the extra amino acids were present.

5.4 Summary

This chapter demonstrates, for the first time, the interaction of *in vacu* and solvated magnetite crystal {100} and {111} surfaces with a glutamic acid repeat motif, found in the wild–type Mms6 protein C–terminus, and its subsequent sequential mutation with alanine moieties, in both unconstrained and constrained systems.

5.4.1 Unconstrained

Our results for the unconstrained systems indicate that, octahedral iron attachment is preferred for the {111} surface, whilst the {100} surface exhibits no particular preference, with the exception of the mutated dipeptides which exhibits only octahedral iron binding for the solvated systems. More bonding was present in the {111} surface systems than the {100} surface systems, with the exception of DEAV which exhibits the same amount

of bonding on both surfaces, suggesting that this is the preferred surface of attachment. Less bonding present in the solvated systems than the *in vacu* systems, proposing that the addition of water to the systems is detrimental to the iron binding potential of the peptides, providing competitive binding to the magnetite. The interfacial energy data suggests that the lowest solvated energies were found on the {100} surface, however, this data looks at average energy values of the full magnetite/amino acid residue/water interface over a simulation, whereas, other forms of data analysis focus on bonding present within the systems throughout the simulations. The {100} surface may exhibit lower interfacial energy due to the surfaces behaviour with water. The *in vacu* interfacial energies are considerably lower, further enforcing the detrimental effect of system solvation.

The hypothesis of this study was that iron binding would decrease with sequential mutations of the original glutamate repeat motif with alanine moieties. The data show that, the substitution of glutamate for alanine clearly has a detrimental effect on the potential iron binding of the dipeptides, as the bonding present reduces more with the removal of each glutamate from the sequence. The exception to this is the interfacial energy data of solvated systems, in particular for the {100} surface, where the opposite of this hypothesis is seen, with EE exhibiting the highest interfacial energy and the AA displaying the lowest interfacial energy. This suggests that the behaviour of the interface between the magnetite surface and the water has a strong effect on the overall interfacial energy. When the dipeptides sequences were extended to tetrapeptides, with the addition of the aspartate and valine residues, more bonding was present. This is believed to be due to the greater availability of potential oxygen binding sites. This was also demonstrated with the *in vacu* interfacial energies, where the lowest interfacial energies were exhibited for the tetrapeptides. Proposing that, the addition of these particular amino acid residues to the sequences leads to preferential tetrapeptide sequence attachment. However, when the systems were solvated, the interfacial behaviours exhibited differed greatly, with the lowest interfacial energy sequences being AA on the {100} surface and EE on the {111} surface again proposing that the presence of water in the system greatly effects the system interfacial behaviours.

5.4.2 Constrained

The Potential of Mean Force data indicates that the {111} surface is the preferred surface of attachment, with the lone exception of DEEV, as this surface produces the lowest free energy values throughout the simulations. This is in concordance with the {111} surface

also conforming to the magnetic easy axis[94], having the lowest energy, and thus representing the more favourable attachment site. The constrained system data also shows that the substitution of glutamic acid for alanine has surface–dependent differential modulatory effects on the iron binding ability of peptide sequences to magnetite. An increase in free energy with both single and double E to A mutations from the original EE sequence was demonstrated, thus indicating a reduction in iron binding affinity, on the {100} and {111} surfaces. When the EE sequence was extended DEEV, and was subsequently mutated, sequence attachment to the {100} surface again showed increases in free energy, thus further indicating a reduction in iron binding affinity. However, on the {111} surface the substitution of E for A appears to enhance iron binding ability, as shown by a decrease in free energy. The E adjacent to the V on the {111} surface appears to have the greatest influence on the iron binding, as its substitution produces the most significant decrease in free energy.

Chapter 6

The Interaction of Pentapeptides with Magnetite

6.1 Introduction

As mentioned previously, the C–terminal of the Mms6 is a hydrophilic and acid rich region[100], which is believed to play an important role in binding ferric ions with a high affinity [253]. To test this hypothesis, a range of mutants were designed by experimentalists [254], based on the C–terminal of the Mms6 protein, as shown in Figures 6.1, 6.2, 6.3, 6.4, 6.5 and Table 6.1.

Figure 6.1: Sequence of the C–terminal region of the Mms6 protein, exhibiting the designed mutants. The pentapeptide and tetrapeptide regions have been highlighted.

Acidic amino acids were targeted by the experimentalists for mutagenesis as they are likely to have ability for iron binding due to their negative charge at neutral to basic pH. This is a required property for magnetite formation and stability, allowing for the coordination of the positively charged iron ions. Thus, all aspartic acids residues and glutamic acids residues present in the C–terminal sequence are substituted for alanine producing mutants with designations such as D179A. A double glutamic acid mutant, EE187AA, was also created. The GinsV mutant was created in order to investigate the spacing between clusters of acidic amino acids. The arginine substitution mutant (R192A) was created to investigate its role in iron binding, as the arginine residue is conspicuous due to it being a lone residue of the opposing charge state within a highly acidic region.

In this chapter, molecular dynamics (based on classical atomistic potentials) are used to

Pentapeptide	Mutation Name	Mutation Explanation
SRDIE	D179A	Substitution of D at position 179 for A.
DIESA	E181A	Substitution of E at position 181 for A.
SDEEV	D186A	Substitution of D at position 186 for A.
	E187A	Substitution of E at position 187 for A.
	E188A	Substitution of E at position 188 for A.
	EE187AA	Double substitution of E from position 187 for A.
EVELR	E190A	Substitution of E at position 190 for A.
	GinsV	Insertion of G after V at position 189.
ELRDA	R192A	Substitution of R at position 192 for A.
LRDAL	D193A	Substitution of D at position 193 for A.

Table 6.1: Explanation of the pentapeptides and related mutations investigated in this study. (D is aspartate, A is alanine, E is glutamate, G is glycine, V is valine and R is arginine.)

study the attachment of pentapeptides created from the C–terminal sequence and their related mutations, both *in vacu* and solvated, to the {100} and {111} crystal surfaces, and the effects these mutations have on iron binding. The investigation examines the interfacial relationships and how these can vary dependent on the surface utilised, which amino acid residues are involved in the peptide chain and whether the sequence is mutated. This study was divided into two; a classic molecular dynamics system and a constrained molecular dynamics system using the Potential of Mean Force.

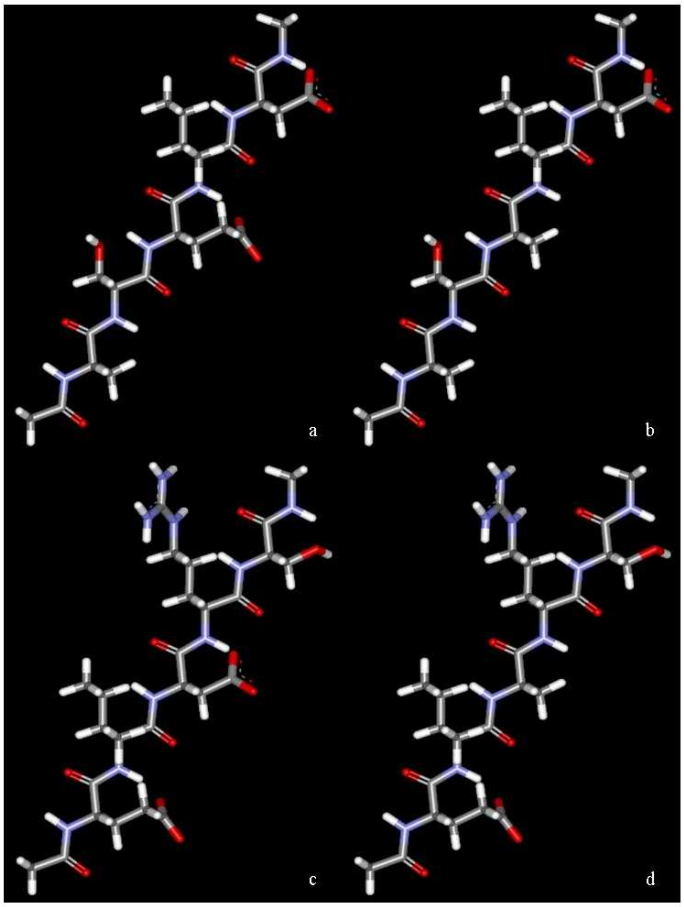

Figure 6.2: Structures of the SRDIESA sequences investigated. a) DIESA, b) DIASA c) SRDIE and d) SRAIE.

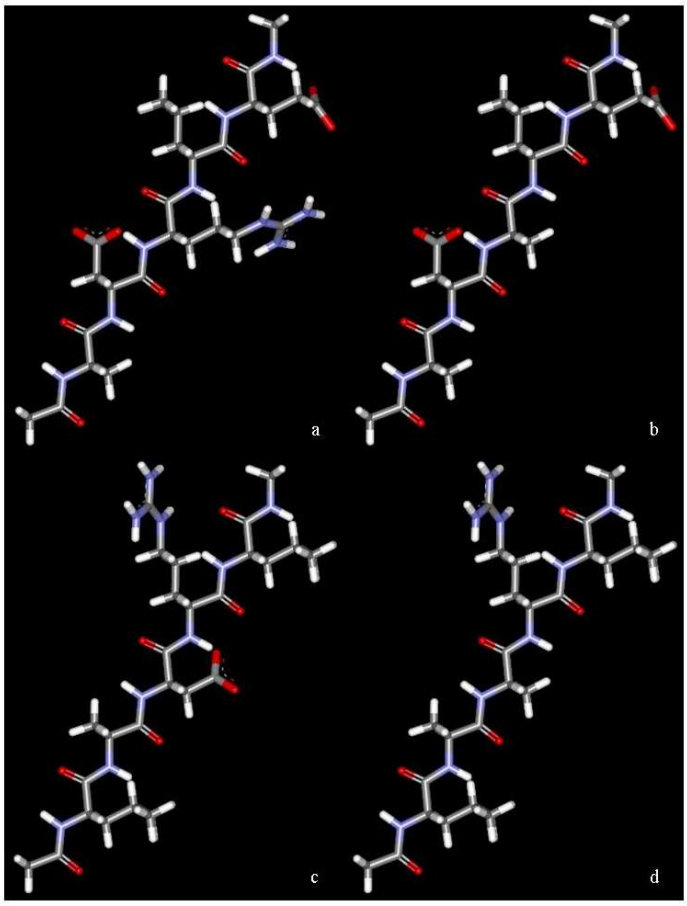

Figure 6.3: Structures of the ELRDAL sequences investigated. a) ELRDA, b) ELADA c) LRDAL and d) LRAAL.

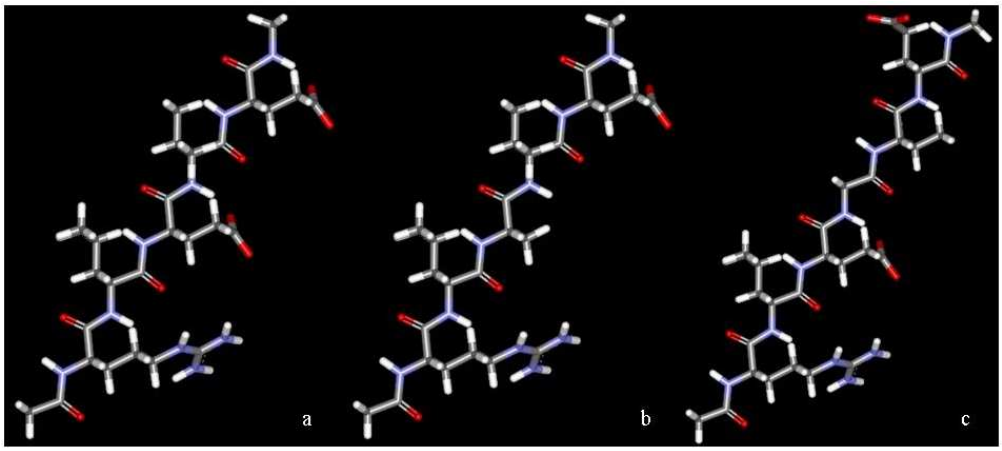

Figure 6.4: Structures of the EVELR sequences investigated. a) EVELR, b) EVALR and c) EVGELR.

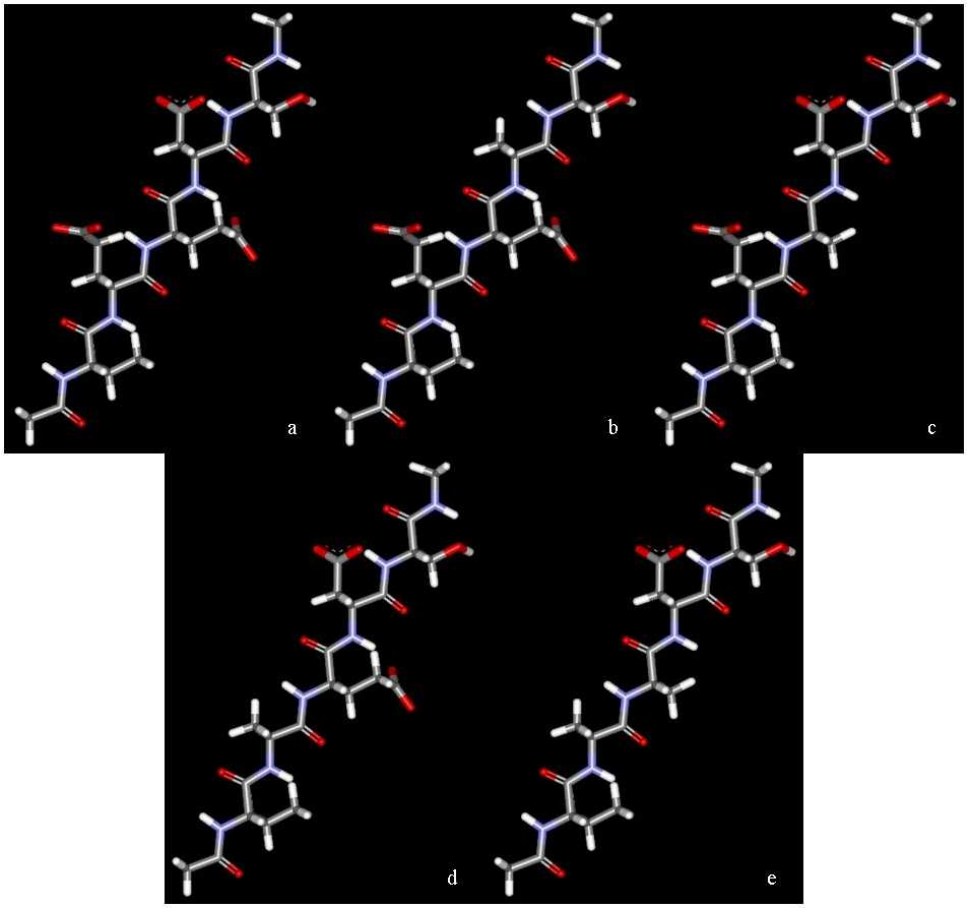

Figure 6.5: Structures of the SDEEV sequences investigated. a) SDEEV, b) SAEEV c) SDAEV d) SDEAV and e) SDAAV.

6.2 Computational Methods

6.2.1 Unconstrained System

As with Chapters 4 and 5, the starting structures of the peptide chains were generated and capped using TLEAP. SANDER, PMEMD and DL_POLY were used for structural and system relaxation. The TIP3P/fs potential was used for any water present in the simulations. Upon placement of the relaxed peptide chains in the vacuum gap vacuum gap above the magnetite slab surface and run further *in vacu* and solvated for 5 ns and 1ns, respectively. All systems were run at 300 K. The ensembles, potentials and parameters used are described in Chapters 2 and 3.

The evolution of potential energy of all of the pentapeptides as a function of time were plotted, an example of which is shown for one of the original sequences, DIESA (Figure 6.6). A full set of potential energy plots can be found in the Appendix. The equilibration period for all simulations was 20 ps, after which point all of the pentapeptide energies converge.

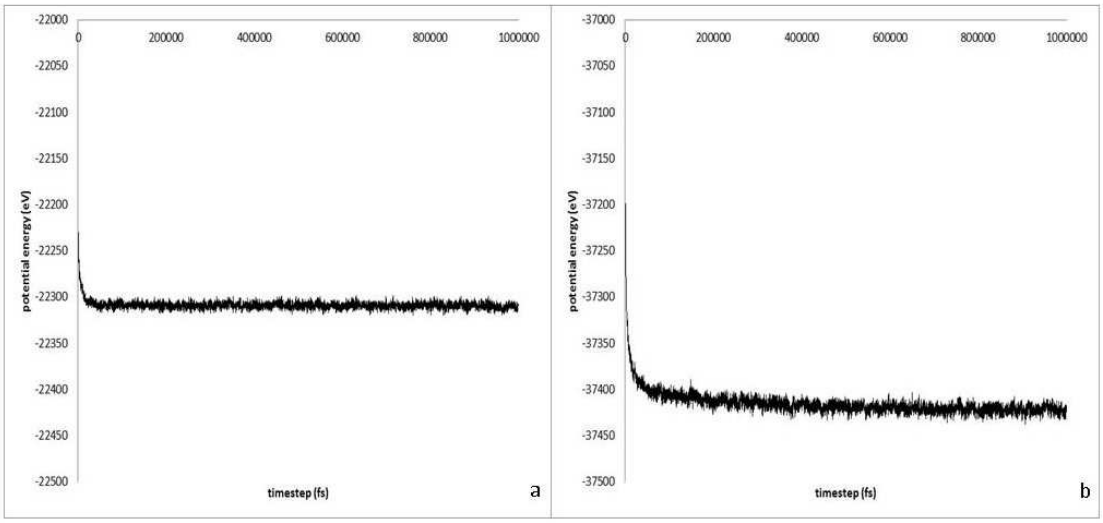

Figure 6.6: Evolution of potential energy as a function of time plots for DIESA. a) {100} solvated surface. b) {111} solvated surface.

6.2.2 Constrained System

As with the unconstrained system, AMBERTOOLS TLEAP was used to produce the initial amino acid structures, including terminal end capping. AMBER's in built software

was used to relax the peptide structures in preparation for conversion into DL_POLY format, whereby, they were relaxed further, both *in vacu* and solvated, using molecular dynamics. Subsequently, the water was removed and the peptides were included into the slab surface system, where they were run further *in vacu* then solvated. As before, the systems were at 300 K throughout. A series of Potential of Mean Force simulations were run to determine the iron binding affinity of the capped peptide chains. Each PMF calculation was run for 1 ns NVT at 300 K and the average force was integrated with respect to the constraint distance to produce the free energy of binding[139].

6.3 Results and Discussion

6.3.1 Residues 177-183 (DIESA & SRDIE)

Unconstrained System

As with Chapters 4 and 5, radial distribution function (RDF) data was examined for the 177-183 pentapeptides attachment to the {100} and {111} magnetite surface, using a distance of 1.5 Å to 2.5 Å . R values for pentapeptides in the region of 177-183 can be found in Tables 6.2 and 6.3.

Figure 6.7 exhibits the RDF profiles for DIESA. The RDF profiles of the remaining 177-183 pentapeptides can be found in the Appendix. The RDF data for the 177-183 pentapeptides shows that, both octahedral and tetrahedral iron bonding was exhibited for the {100} surface, with the exception of SRAIE, which exhibited no octahedral iron bonding for the solvated {100} system. Whereas, with the {111} surface, only octahedral iron bonding was present. When the 177-183 pentapeptides were attached to the {111} surface, octahedral iron attachment was favoured. The iron bonding trend was the same for all 177-183 pentapeptides, suggesting that sequence alteration had no effect on the resultant iron bonding type.

From the 177-183 pentapeptide RDF data it can also be seen that, the *in vacu* system produced a more intense first peak than the corresponding solvated system, with the exception of tetrahedral iron bonding in the D179A {100} system, proposing that, upon solvation, it was more difficult for a pentapeptide oxygen to have an iron as a next nearest neighbour and for this to remain for the simulation duration. For the solvated systems, the largest $g(r)$ peak was shown for octahedral iron bonding in the D179A {111} system,

Pentapeptide	Fe type	Surface	System	r (Å)
DIESA	FET	100	*in vacu*	1.98
			solvated	1.88
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.08
			solvated	2.23
		111	*in vacu*	1.98
			solvated	2.03
DIASA	FET	100	*in vacu*	1.98
			solvated	1.88
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.13
			solvated	2.23
		111	*in vacu*	1.93
			solvated	1.98

Table 6.2: R (Å) values from RDF data for pentapeptides in the region of 177-183.

suggesting that, the oxygen present in this pentapeptide was more freely accessible to the next nearest neighbouring irons.

By examining the $FE - O_{pentapeptide}$ bond distance data, the results of the RDF data can be some way explained. This data for the 177-183 pentapeptides is represented in Tables 6.4 and 6.5.

Pentapeptide	Fe type	Surface	System	r (Å)
SRDIE	FET	100	*in vacu*	1.93
			solvated	1.93
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.03
			solvated	2.08
		111	*in vacu*	1.98
			solvated	2.03
SRAIE	FET	100	*in vacu*	2.13
			solvated	2.03
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.03
			solvated	-
		111	*in vacu*	1.93
			solvated	2.03

Table 6.3: R (Å) values from RDF data for pentapeptides in the region of 177-183.

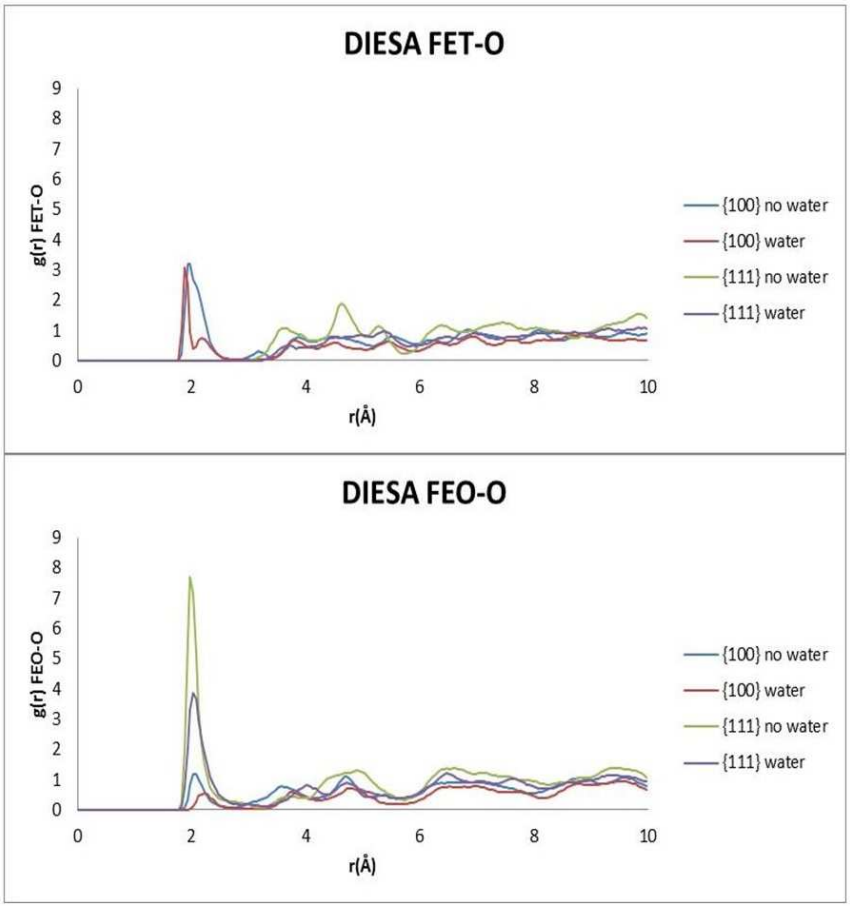

Figure 6.7: RDF plots for DIESA. FET is tetrahedral iron, FEO is octahedral iron. Blue is {100} *in vacu*. Red is {100} solvated. Green is {111} *in vacu*. Purple is {111} solvated.

	{100} surface attachment			{111} surface attachment		
Pentapeptide	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
DIESA	NBO	2.05	NB	NBO	2.09	NB
	D O	1.95	1.92	D O	1.92	2.00
		1.97	NB		2.10	2.33
	D-I	2.02	2.18	D-I	2.10	NB
	I-E	2.10	2.48	I-E	2.20	NB
	S-A	2.16	NB	E O	1.96	1.98
					2.11	1.95
					1.98	NB
				E-S	2.09	2.13
DIASA	D O	1.94	NB	NBO	2.21	NB
		2.06	1.84	D O	2.02	1.91
	D-I	2.02	2.11		2.06	1.93
	A-S	2.21	2.10	I-A	2.10	NB
	S-A	2.17	NB	S-A	2.04	NB

Table 6.4: 177-183 pentapeptide $FE-O_{pentapeptide}$ bond distance data. BL is bond length, E O is glutamate side chain oxygen, D O is aspartate side chain oxygen, x-x PBO is inter–residue peptide bond oxygen and NBO is N–terminal peptide bond oxygen.

	{100} surface attachment			{111} surface attachment		
Pentapeptide	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
SRDIE	NBO	2.18	NB	NBO	1.95	2.11
	D O	1.95	NB	S-R	2.02	2.15
	E O	1.93	1.89	D O	2.27	2.01
		1.97	NB		1.94	2.09
				I-E	2.01	NB
				E O	1.95	NB
					1.87	NB
SRAIE	NBO	2.17	NB	NBO	1.99	NB
	S-R	2.18	NB	S-R	2.25	NB
	E O	1.94	NB	E O	1.90	2.03
		-	2.03		1.89	2.05

Table 6.5: Continued. 177-183 pentapeptide $FE - O_{pentapeptide}$ bond distance data. BL is bond length, E O is glutamate side chain oxygen, D O is aspartate side chain oxygen, x-x PBO is inter–residue peptide bond oxygen and NBO is N–terminal peptide bond oxygen. NB is no bonding.

The 177-183 pentapeptides bond distance data (Figure 6.8) shows that, for DIESA *in vacu*, bonding differed dependant on the surface of attachment. The {111} surface showed bonding through both side chain oxygen of glutamate, however, no glutamate bonding was apparent for the {100} surface. The {100} surface showed S-A peptide bond oxygen bonding, whereas, the {111} surface exhibited E-S peptide bond oxygen bonding. There were some commonalities in bonding type between the two surfaces, with both surfaces exhibiting N–terminal, D-I and I-E peptide bond oxygen bonding and bonding through both side chain oxygen of aspartate. This data suggests that, the DIESA sequence had more of an affinity to the {111} surface, particularly in the case of glutamate. Less bonding occurred when the systems were solvated, suggesting that the iron ions may be preferentially binding to the oxygen of water as opposed to the oxygens of the pentapeptides. The average bond lengths for a magnetite iron to a pentapeptide oxygen in the DIESA sequence was; 2.04 Å and 2.19 Å on the {100} surface for *in vacu* and solvated systems respectively, and 2.06 Å and 2.08 Å on the {111} surface for *in vacu* and solvated systems respectively.

With the mutation of the DIESA sequence to DIASA, less bonding was present than for the original sequence, proposed to be due to the removal of potential oxygen binding sites from the sequence. Both surfaces exhibited S-A peptide bond oxygen bonding and bonding through both side chain oxygen of aspartate, however, other bonding types present differed dependent on surface type, with D-I and A-S peptide bond oxygen bonding being exhibited for the {100} surface, whilst, the {111} surface displayed N–terminal and S-A peptide bond oxygen bonding. Again, the addition of water to the system was detrimental to this bonding, as a reduced amount of bonding was seen, with all of the peptide bond oxygen bonding removed from the {111} surface. The average bond lengths for an iron to a pentapeptide oxygen in DIASA was; 2.08 Å and 2.02 Å on the {100} surface for the *in vacu* and solvated systems respectively, and 2.09 Å and 1.92 Å on the {111} surface for the *in vacu* and solvated systems respectively.

For the SRDIE sequence, less bonding was present than for the DIESA sequence, suggesting that the iron binding potential is sequence dependent. For SRDIE *in vacu*, both surfaces displayed N–terminal peptide bond oxygen bonding, bonding through both side chain oxygen of glutamate and bonding through side chain oxygen of aspartate (one bond for the {100} surface and two for the {111} surface). The {100} surface presented with no other bonding, whilst, the {111} surface exhibited S-R and I-E peptide bond oxygen bonding. This data suggests that, the SRDIE sequence has more of an affinity to the {111}

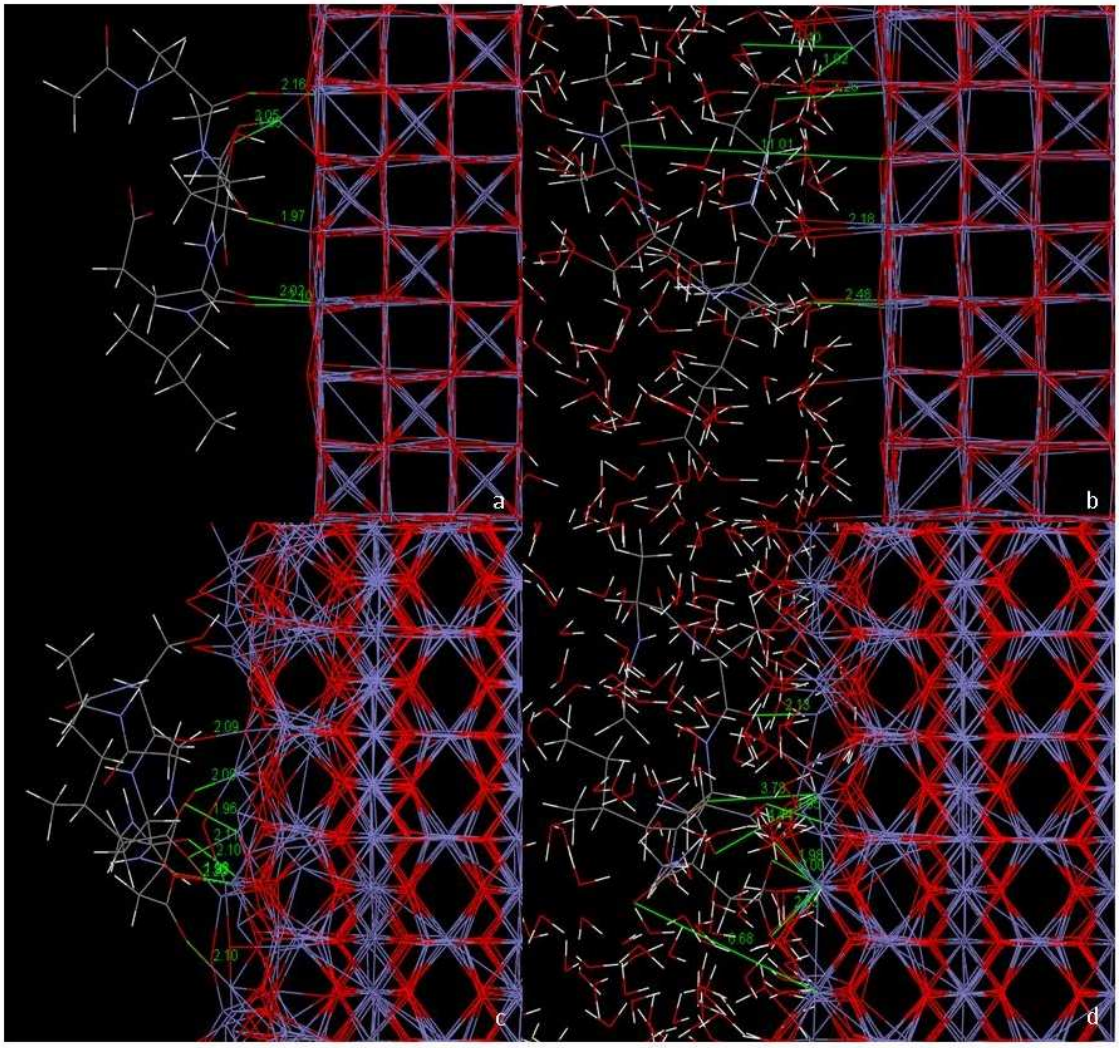

Figure 6.8: Bond distance images for the *in vacu* and solvated systems of DIESA. a){100} *in vacu*, b){100} solvated, c){111} *in vacu* and d){111} solvated.

surface, particularly with peptide bond oxygen bonding. As previously seen, less bonding arose when the systems were solvated, suggesting more favourable bonding between the iron ions and the oxygen of water, which is detrimental to peptide–magnetite binding. The average bond lengths for a magnetite iron to a pentapeptide oxygen in the SRDIE sequence was; 2.01 Å and 1.89 Å on the {100} surface for *in vacu* and solvated systems respectively, and 2.00 Å and 2.09 Å on the {111} surface for *in vacu* and solvated systems respectively.

The SRAIE sequence (D179A), presented with less bonding than the original sequence, particularly for the {111} surface, again, assumed to be due to glutamate substitution and thus removal of potential oxygen binding sites. Both surfaces exhibited the same types of bonding; N–terminal and S-R peptide bond oxygen bonding and bonding through side

chain oxygen of glutamate (one bond for the {100} surface and two for the {111} surface), suggesting that attachment to different surface types had minimal effect on iron binding potential. Again, solvation was detrimental to bonding, with fewer bonds present in the solvated systems. For the {111} surface, no bonding present in the *in vacu* system was exhibited, however, bonding was displayed through the other side chain oxygen of glutamate. The average bond lengths for an iron to a pentapeptide oxygen in SRAIE was; 2.10 Å and 2.03 Å on the {100} surface for the *in vacu* and solvated systems respectively, and 2.01 Å and 2.04 Å on the {111} surface for the *in vacu* and solvated systems respectively.

As with Chapters 4 and 5, residence times were considered in order to gain a more accurate representation of bonding behaviour throughout the simulations. Tables 6.6, 6.7, 6.8, 6.9, 6.10 and 6.11 represent the residence times and average coordination numbers for bonding between all oxygen present in the 177-183 pentapeptides and tetrahedral and/or octahedral iron on both the {100} and {111} magnetite surfaces.

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
DIESA {100} surface	S-A PBO	59.21	0.49	52.59	0.02	16.00	0.03	0.00	0.00
	S O	9.85	0.01	1.49	0.04	0.00	0.00	0.00	0.00
	E-S PBO	141.66	0.13	310.68	0.44	0.00	0.00	39.10	0.34
	E O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	I-E PBO	1446.84	0.70	216.74	0.14	72.16	0.82	9.90	0.05
	D-I PBO	0.00	0.00	494.21	0.91	0.00	0.00	77.20	0.89
	D O	3203.39	1.00	2.27	0.00	999.00	1.00	0.00	0.00
	D O	913.00	0.37	3158.36	0.63	1.50	0.00	13.80	0.04
	NBO	3166.00	0.63	515.18	0.24	5.67	0.01	1.33	0.00

Table 6.6: Residence times (RT) and average coordination numbers (av. CN) of the DIESA sequence. (E O = side chain oxygen of glutamate, D O = side chain oxygen of aspartate, S O = side chain oxygen of serine, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
DIESA {111} surface	S-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	S O	0.00	0.00	2705.51	1.09	0.00	0.00	150.66	0.99
	E-S PBO	0.00	0.00	4998.00	1.00	0.00	0.00	998.26	1.00
	E O	0.00	0.00	4990.40	1.00	0.00	0.00	968.51	1.25
	E O	0.00	0.00	3046.57	1.85	0.00	0.00	205.96	0.52
	I-E PBO	0.00	0.00	4999.00	1.00	0.00	0.00	113.91	0.27
	D-I PBO	0.00	0.00	2323.78	1.68	0.00	0.00	33.66	0.87
	D O	0.00	0.00	4799.67	1.06	0.00	0.00	128.89	1.74
	D O	1.50	0.00	4963.95	1.00	0.00	0.00	353.68	0.99
	NBO	0.00	0.00	4650.00	0.93	0.00	0.00	74.97	0.15

Table 6.7: Continued. Residence times (RT) and average coordination numbers (av. CN) of the DIESA sequence. (E O = side chain oxygen of glutamate, D O = side chain oxygen of aspartate, S O = side chain oxygen of serine, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
DIASA {100} surface	S-A PBO	140.04	0.16	751.32	0.80	0.00	0.00	17.62	0.09
	S O	0.00	0.00	61.84	0.12	0.00	0.00	1.00	0.00
	A-S PBO	240.03	0.73	0.00	0.00	42.38	0.64	0.00	0.00
	I-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D-I PBO	0.00	0.00	2155.81	0.99	0.00	0.00	29.96	0.67
	D O	4810.22	0.97	79.20	0.03	999.00	1.00	0.00	0.00
	D O	4998.00	1.00	1.00	0.00	33.74	0.07	0.00	0.00
	NBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DIASA {111} surface	S-A PBO	0.00	0.00	4964.29	1.00	0.00	0.00	38.51	1.16
	S O	0.00	0.00	1.33	0.00	0.00	0.0	0.00	0.00
	A-S PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	I-A PBO	0.00	0.00	4960.95	1.00	0.00	0.00	25.68	0.06
	D-I PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D O	0.00	0.00	4981.81	1.00	0.00	0.00	730.68	1.14
	D O	0.00	0.00	4978.42	1.00	0.00	0.00	996.43	1.02
	NBO	0.00	0.00	43.33	1.36	0.00	0.00	20.97	0.08

Table 6.8: Residence times (RT) and average coordination numbers (av. CN) of the DIASA sequence. (E O = side chain oxygen of glutamate, D O = side chain oxygen of aspartate, S O = side chain oxygen of serine, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
SRDIE {100} surface	E O	4997.00	1.00	0.00	0.00	333.96	0.98	0.00	0.00
	E O	3.00	0.00	4992.00	1.00	0.00	0.00	327.92	0.87
	I-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D-I PBO	0.00	0.00	186.38	0.88	0.00	0.00	4.76	0.10
	D O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D O	2660.36	0.85	503.57	0.15	30.30	0.11	0.00	0.00
	R-D PBO	2.50	0.00	28.97	0.74	0.00	0.00	1.00	0.00
	S-R PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	S O	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
	NBO	1330.00	0.27	830.64	0.71	3.50	0.01	0.00	0.00

Table 6.9: Residence times (RT) and average coordination numbers (av. CN) of the SRDIE sequence. (E O = side chain oxygen of glutamate, D O = side chain oxygen of aspartate, S O = side chain oxygen of serine, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
SRDIE {111} surface	E O	0.00	0.00	4996.00	1.00	0.00	0.00	231.90	0.46
	E O	0.00	0.00	4856.87	1.05	0.00	0.00	3.17	0.01
	I-E PBO	0.00	0.00	4997.00	1.00	0.00	0.00	52.33	0.18
	D-I PBO	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
	D O	0.00	0.00	1975.98	1.95	0.00	0.00	735.14	1.97
	D O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	R-D PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	S-R PBO	0.00	0.00	4979.81	1.00	0.00	0.00	847.96	1.00
	S O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NBO	0.00	0.00	4987.53	1.00	0.00	0.00	998.38	1.00

Table 6.10: Continued. Residence times (RT) and average coordination numbers (av. CN) of the SRDIE sequence. (E O = side chain oxygen of glutamate, D O = side chain oxygen of aspartate, S O = side chain oxygen of serine, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
SRAIE {100} surface	E O	0.00	0.00	4330.00	0.87	0.00	0.00	1.50	0.00
	E O	0.00	0.00	0.00	0.00	996.00	0.99	0.00	0.00
	I-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	A-I PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	R-A PBO	7.00	0.00	7.25	0.00	0.00	0.00	0.00	0.00
	S-R PBO	320.06	0.83	323.99	0.13	4.44	0.02	0.00	0.00
	S O	1.75	0.02	2.12	0.02	1.00	0.00	27.84	0.15
	NBO	320.68	0.13	103.19	0.80	21.98	0.04	12.91	0.05
SRAIE {111} surface	E O	0.00	0.00	4940.54	1.04	0.00	0.00	466.77	1.07
	E O	0.00	0.00	4911.88	1.01	0.00	0.00	435.43	1.82
	I-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	A-I PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	R-A PBO	0.00	0.00	241.19	0.86	0.00	0.00	695.43	0.98
	S-R PBO	0.00	0.00	2864.34	0.99	0.00	0.00	48.58	0.20
	S O	0.00	0.00	23.12	0.42	0.00	0.00	10.23	0.08
	NBO	0.00	0.00	71.87	0.52	0.00	0.00	83.47	0.19

Table 6.11: Residence times (RT) and average coordination numbers (av. CN) of the SRAIE sequence. (E O = side chain oxygen of glutamate, D O = side chain oxygen of aspartate, S O = side chain oxygen of serine, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

The 177-183 pentapeptides residence time data shows that, for DIESA {100} *in vacu* system, there was no bonding present that lasted over 65% of the simulation length. Bonding lasting between 63% and 64% of the simulation length was exhibited between tetrahedral iron and the N–terminal peptide bond oxygen and one of the side chain oxygen of aspartate, and also, between octahedral iron and the other side chain oxygen of aspartate. When this system was solvated, the bond between tetrahedral iron and one of the side chain oxygen of aspartate was present for over 99% of the simulation length. This suggests that the lower value for the equivalent *in vacu* residence time was due to the formation of this bond toward the latter stages of this system simulation. No other bonding lasting over 8% of the simulation length was present for this system, suggesting preferential magnetite–water bonding. There was no glutamate bonding for either DIESA {100} system, proposing that this acidic residue is not of interest for iron binding within this sequence.

The DIESA {111} surface residence time data showed a very different trend, with no tetrahedral iron bonding present for either system. For the *in vacu* system, bonding lasting over 93% of the simulation length was seen between octahedral iron and N–terminal, I-E, and E-S peptide bond oxygen, both side chain oxygen of aspartate, and one of the side chain oxygen of glutamate. All other available oxygen showed bonding lasting between 46.5% and 61% of the simulation length, with the exception of S-A peptide bond oxygen, which showed no bonding. However, upon solvation, only the bonds between octahedral iron and one of the side chain oxygen of glutamate and E-S peptide bond oxygen lasted over 97% of the simulation length. No other bonding was present lasting over 36% of the simulation length, suggesting that the addition of water had a detrimental effect on iron binding, as water provides competitive binding to the magnetite.

With the single alanine substitution of DIESA to DIASA, in the {100} *in vacu* system, bonding lasting over 96% of the simulation length was exhibited between tetrahedral iron and both side chain oxygen of aspartate. No other bonding lasting over 44% of the simulation length was seen. When this system was solvated, only one of the bonds between tetrahedral iron and the side chain oxygen of aspartate continued to last over 99% of the simulation length. There was no other bonding lasting over 4.5% of the simulation length present for the solvated system, suggesting that, again, water was providing competitive binding and the glutamate substitution reduced the number of potential binding sites, leading to a stronger affinity of the magnetite for water than for the pentapeptide.

On examination of the corresponding DIASA {111} surface data, no tetrahedral iron bonding was present. Bonding was exhibited to both side chain oxygen of aspartate, and

to I-A and S-A peptide bond oxygen through octahedral iron that lasted over 99% of the simulation length, within the *in vacu* system. No other bonding lasted over 1% of the simulation length. Again, with the addition of water, only the bond between octahedral iron and one of the side chain oxygen of aspartate lasted over 99% of the simulation length. The other side chain oxygen of aspartate showed bonding that lasted for 73% of the simulation length. No other bonding lasted over 4% of the simulation length. The substitution of a glutamate for an alanine had an adverse effect on the potential iron binding of the pentapeptide, as hypothesised, as there was a considerable drop in the residence times exhibited when compared to that of the DIESA sequence.

When the residence times of the SRDIE sequence were considered it was shown that, in the {100} *in vacu* system, only two bonds present lasted over 99% of the simulation length, these were between tetrahedral iron and one of the side chain oxygen of glutamate, and octahedral iron and the other side chain oxygen of glutamate. Bonding lasting for 53% of the simulation length was also exhibited between tetrahedral iron and one of the side chain oxygen of aspartate. No other bonding lasting over 27% of the simulation length was present. Upon solvation, no bonding lasting over 34% of the simulation length was displayed, suggesting that the addition of water had an extremely detrimental effect on the sequences iron binding potential.

When the residence time data was examined for SRDIE on the {111} surface it was shown that, as with DIESA and DIASA, no tetrahedral iron bonding was present. For the *in vacu* system, bonding lasting over 97% of the simulation length was seen between octahedral iron and N–terminal, S-R and I-E peptide bond oxygen, and both side chain oxygen of glutamate. A bond lasting for 40% of the simulation length was also displayed to one of the side chain oxygen of aspartate. On addition of water, only the N–terminal peptide bond oxygen bond continued to last over 99% of the simulation length, however, bonding lasting between 74% and 85% of the simulation length was present to S-R peptide bond oxygen and to one of the side chain oxygen of aspartate. This reduction in residence times suggests that solvation leads to water providing competitive magnetite binding.

On comparison of the residence time data for the two original sequences (DIESA and SRDIE) it can be seen that, for the {100} *in vacu* system, only SRDIE exhibited bonding lasting over 99% of the simulation length, whilst the longest residence time displayed for DIESA was 3203.39 ps (64% of the simulation length). However, DIESA shows more bonding in general, but most of these bonds did not last over 29% of the simulation length. When these systems were solvated, DIESA exhibited one bond that lasted over 99% of the

simulation length, whereas, SRDIE exhibited no such bonding. When the {111} surface *in vacu* system was examined it was shown that, DIESA presented with six bonds lasting over 93% of the simulation length, however, SRDIE showed only five bonds lasting over this length of time. Upon solvation, DIESA presented with two bonds lasting over 97% of the simulation length, whereas, SRDIE exhibited just one bond that lasted over this length of time. From the solvated systems data it can be seen that, DIESA is the preferred sequence of attachment, with the {111} surface being the favoured for this sequence.

The residence time data for the single alanine substitution to SRAIE, in the {100} *in vacu* system, showed one bond lasting over 86% of the simulation length, between octahedral iron and one of the side chain oxygen of glutamate, with no other bonding lasting over 6.5% of the simulation length present. When this system was solvated, a bond was formed between tetrahedral iron and one of the side chain oxygen of glutamate lasting over 99% of the simulation length, whilst, no other bonding lasted over 2.8% of the simulation length.

When the {111} surface simulations were examined, as with the other 177-183 pentapeptides, no tetrahedral iron bonding was present. For the *in vacu* system, bonding lasting over 98% of the simulation length was seen between octahedral iron and both side chain oxygen of glutamate. A bond lasting over 57% of the simulation length was also exhibited to S-R peptide bond oxygen. Upon solvation, no bonding was present that lasted lasting over 70% of the simulation length, suggesting that the presence of water was unfavourable for the iron binding potential of the sequence.

On comparison of the residence time data of the two mutated sequences (DIASA and SRAIE) it can be seen that, for the {100} *in vacu* system, DIASA presented with more bonding lasting over 86% of the simulation length than SRAIE (two bonds and one bond, respectively). This suggests that DIASA was the preferred pentapeptide of attachment. On addition of water, both DIASA and SRAIE exhibited one bond lasting over 99% of the simulation length, suggesting no favoured sequence of attachment when the systems were solvated. From the {111} *in vacu* system data it was shown that, DIASA exhibited more bonding lasting over 98% of the simulation length than SRAIE (four bonds and two bonds, respectively), suggesting that DIASA is the preferred sequence of attachment. Upon solvation, this trend was exhibited again, as DIASA presented with one bond lasting over 99% of the simulation length, whilst, SRAIE exhibited no bonding lasting over this length of time. It is worth noting that, for the mutated sequences, the most prominent bonding was seen to the remaining acidic residues. In the original counterparts, these acidic residues were also the areas of prominent iron binding, with the substituted acidic

residues having exhibited far less bonding. The retention time data of the natural solvated systems shows that, there was no preference for either sequence with the {100} surface, however, for the {111} surface, DIASA is the preferred sequence of attachment.

As with Chapters 4 and 5, the interfacial energies of the 177-183 pentapeptide sequences on the {100} and {111} magnetite surfaces were calculated. All related energy data from the interfacial energy calculations of the pentapeptides can be found in Table 6.12.

{100} surface attachment				
	E_{int} of s-p (eV)	E_{diff} of s-p (eV)	E_{int} of s-p-w (eV)	E_{diff} of s-p-w (eV)
DIESA	-4.34	0.00	7.72	0.00
DIASA	-5.09	0.75	7.23	0.49
SRDIE	-28.87	24.53	7.63	0.09
SRAIE	-22.52	18.18	7.39	0.33
{111} surface attachment				
	E_{int} of s-p (eV)	E_{diff} of s-p (eV)	E_{int} of s-p-w (eV)	E_{diff} of s-p-w (eV)
DIESA	-13.61	6.10	10.73	7.50
DIASA	-7.51	0.00	18.23	0.00
SRDIE	-27.42	19.91	13.61	4.62
SRAIE	-22.01	14.50	11.66	6.56

Table 6.12: Interfacial energies for all 177-183 pentapeptides on the {100} and {111} surface, using the Yang, Stipp and Harding method[164]. E_{int} of s-p is the interfacial energy of the slab–peptide system. E_{diff} of s-p is the difference in interfacial energy from the highest interfacial energy of the slab–peptide system. E_{int} of s-p-w is the interfacial energy of the slabpeptide–water system. E_{diff} of s-p-w is the difference in interfacial energy from the highest interfacial energy of the slab–peptide–water system.

From the solvated 177-183 pentapeptides interfacial energy data it can be seen that, the {100} surface produced lower interfacial energy values (between 7.23 eV and 7.72 eV) than the {111} surface (between 10.73 eV and 18.23 eV). For the {100} surface, the mutated sequences produced lower interfacial energies than their original counterparts, suggesting that mutation improves sequence iron binding potential. For the {111} surface, DIASA produced a higher interfacial energy than DIESA, suggesting that the substitution was detrimental to the iron binding affinity of the sequence (as hypothesised), whereas, with SRDIE, the mutated sequence (SRAIE) produced a lower interfacial energy, proposing that the substitution improved the sequence iron binding potential.

This data also shows that, the presence of water in the system had a detrimental effect

on the interfacial energies, as the energies exhibited were much lower for the *in vacu* systems, in particular for SRDIE and its mutation, for example, on the {111} surface, solvation significantly increased the interfacial energy of SRDIE from -27.42 eV to 13.61 eV. Solvation of the systems also affected how SRDIE and its mutation reacted at the interface, as the interfacial energy trends differed dependent on whether the systems were solvated or not. The lowest interfacial energy was exhibited for the original sequence, when *in vacu*, however when solvated the opposite was observed. The increase in interfacial energy with the addition of water could be resultant from preferential attachment of water to the magnetite surface.

Whilst the absolute values for the adsorption energy are clearly an order of magnitude larger than expected, the relative energies within a sequence may be more reliable for amino acid binding potential comparison. The trends shown upon solvation are counter-intuitive and contradictive of experimental studies; however, it is clear that this is an artefact of the energy calculation method utilised, as the expected trend is revealed in the relative residence times of various surface interactions. As it can be demonstrated that, $\mathrm{R}T \times ln(\tau^1/\tau^2) = (\Delta H^2_{ads} - \Delta H^1_{ads}) = (E^1_{des} - E^2_{des})$, then this method is perhaps more reliable for determining the change in adsorption energy between the various pentapeptide sequences. This interfacial energy method also requires the measurement of small differences between large energies, which leads to large uncertainties in the results. Another possible error incorporated into the calculation of the interfacial energies for the hydrated systems, is the assumption that all the water molecules within the systems containing the slab and water are perfectly hydrated throughout the slab, i.e. the bulk density is that of the bulk water simulation. However, as the simulations are run within an NVT ensemble, this is not necessarily the case. Therefore, if insufficient water is present at the start of the simulation, some of the water molecules will eventually become under coordinated, thus increasing the energy of the simulation cells. Furthermore, rather than directly calculating the adsorption energy, via this interfacial energy method, an alternative indirect method of adsorption energy calculation is explored in the following section. This method utilises a constrained molecular dynamics version of the systems, focusing on the Potential of Mean Force and free energy evaluation.

6.4 Constrained system

Free energy profiles were created for 177-183 pentapeptide sequences. As previously, 0 Å on the x-axis is an approximate representation of where the magnetite surface lies. The profiles again exhibit the free energy of the simulation system as a function of peptide chain centre of mass (CoM) distance from the magnetite surface.

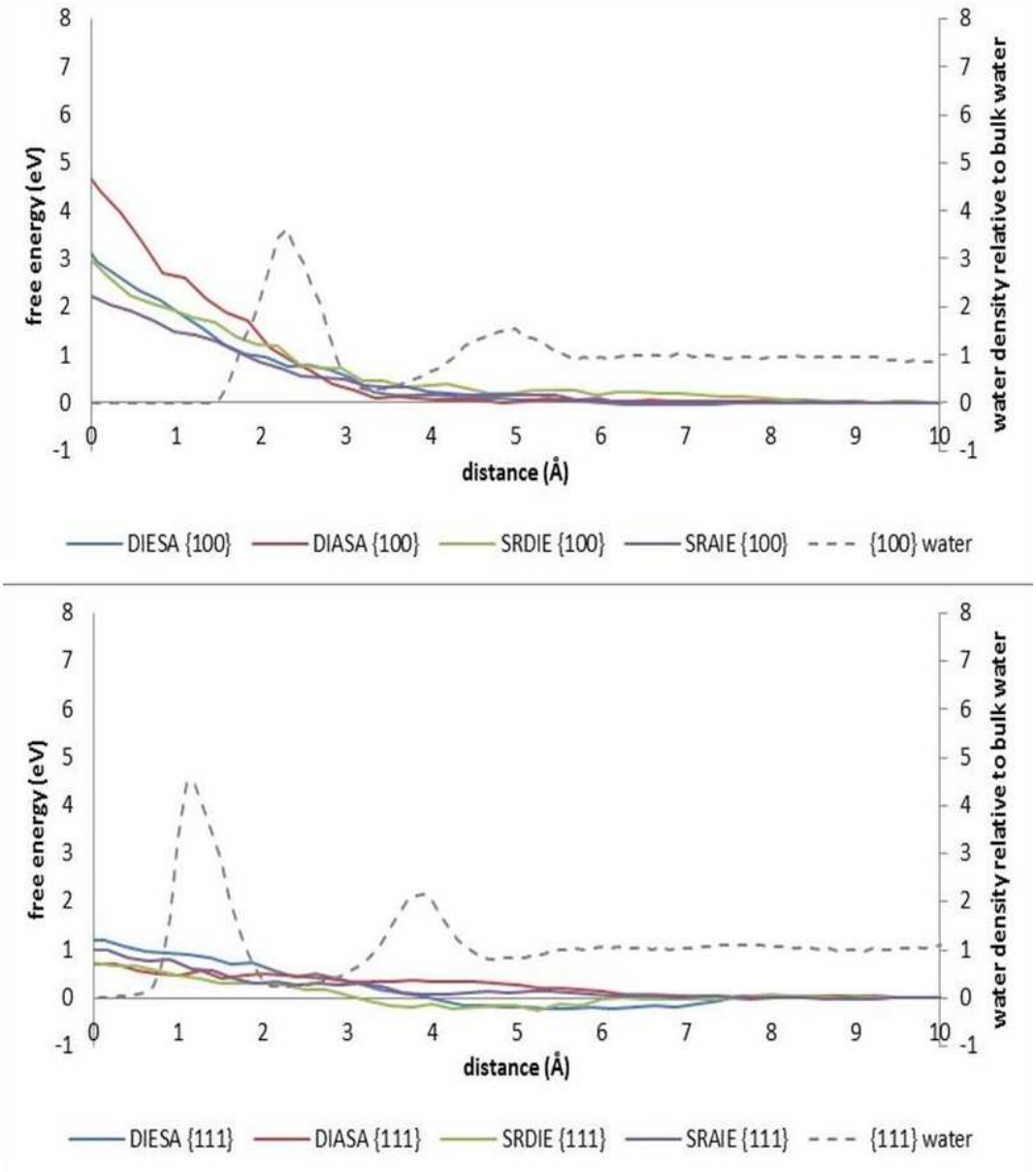

Figure 6.9: A comparison of the free energy profiles of the 177-183 pentapeptides. Distance refers to distance between the peptide CoM and the magnetite surface. The dashed line represents the water density profile for the system.

From the free energy profile results for the different 177-183 pentapeptides (Figure 6.9) it was evident that, on the {100} surface, the mutation of DIESA to DIASA reduced the

iron binding activity of the sequence as hypothesised, whereas, the mutation of SRDIE to SRAIE increases the iron binding activity by reducing the free energy. The mutant DIASA produced the highest free energy of all 177-183 pentapeptide sequences. Between 10 Å and 4 Å is optimal for attachment, after which point, an increase in free energy was detected, corresponding with the presence of the more intense water peak. It is proposed that the successive increase in free energy was attributable to a requirement for more energy within the system for the sequences to pass through the water barrier in order to enable sequence attachment to the {100} surface.

Whereas, with the {111} surface energy profile, there was very little difference between the free energy profiles of the four sequences. It can also be seen that, the mutation of DIESA to DIASA appears to slightly increase the iron binding activity of the sequence, whereas the mutation of SRDIE to SRAIE appears to slightly reduce the iron binding activity of the sequence. The original sequence DIESA produces the highest free energy of all 177-183 pentapeptide sequences. For DIASA and SRAIE, the optimal distance for attachment was between 10 Å and 6 Å , and between 10 Å and 4 Å , respectively, after which point, an increase in free energy was detected, corresponding with the presence of the intense water layers. Between 7.5 Å and 4 Å and 6 Å and 3 Å are optimal for attachment for DIESA and SRDIE, respectively, as the profiles exhibit negative values within this region. The small changes in free energy seen for the sequence profiles suggests that the water absorption layers have minimal influence on the free energy profiles.

Upon examination of the energy minima for the 177-183 pentapeptides (Figure 6.10) it can be observed that, on the {100} surface, all of the 177-183 pentapeptides exhibited a negative value for their energy minima, with DIESA producing the lowest value at -0.0468 eV. The trend of favourable attachment distance varies with sequence, with DIESA, DIASA, SRDIE and SRAIE exhibiting centre of mass distances of 6.6 Å , 9.6 Å , 10 Å and 6.7 Å respectively. On the {111} surface, again, all of the 177-183 pentapeptides exhibited a negative value for their energy minima, with SRDIE producing the lowest value at -0.2429 eV. The trend of favourable attachment distance showed DIESA, DIASA, SRDIE and SRAIE exhibiting centre of mass distances of 5.6 Å , 7.8Å , 5.3 Å and 8.7 Å respectively.

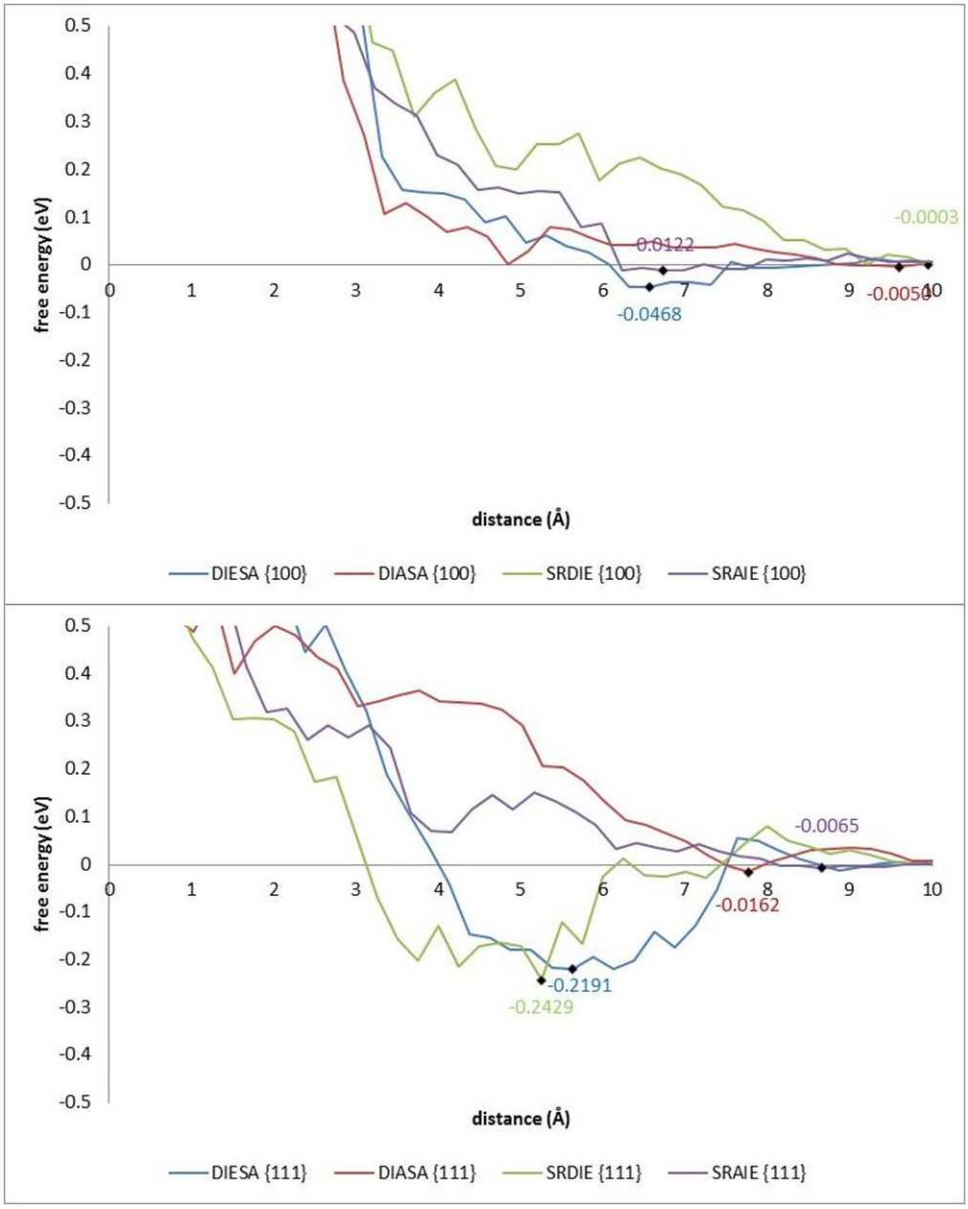

Figure 6.10: A close-up comparison of the free energy profiles of the 177-183 pentapeptides between -0.5 eV and 0.5 eV, exhibiting the free energy minima.

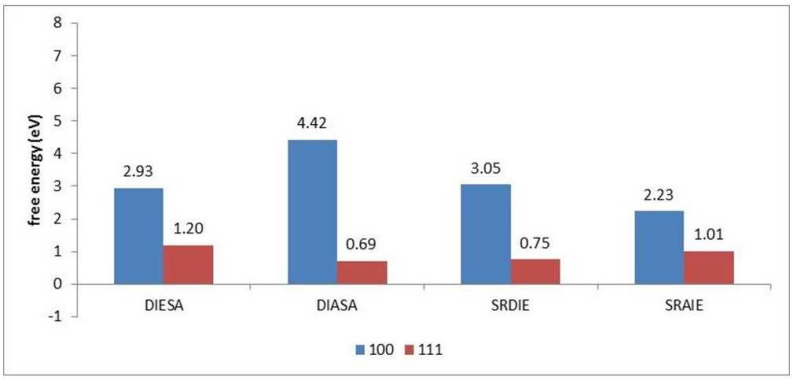

Figure 6.11: A comparison of the free energy values for the 177-183 pentapeptides CoM at the magnetite surface for both the {100} and {111} surfaces.

The free energy values for the 177-183 pentapeptides centre of mass at the magnetite surfaces are shown in Figure 6.11. On the {100} surface, the DIESA sequence exhibits a free energy value of 2.93 eV, whereas, on the {111} surface, its value is lower, 1.20 eV. Similarly, the DIASA, SRDIE and SRAIE sequences demonstrated much higher free energy values (4.42 eV, 3.05 eV and 2.23 eV, respectively) on the {100} surface, than exhibited on the {111} surface (0.69 eV, 0.75 eV and 1.01 eV, respectively). These data show that, the {111} surface has significantly lower free energy values, representing greater iron binding, suggesting that this was the preferential surface of attachment for the 177-183 pentapeptide sequences. It also shows that, the sequence with the lowest free energy when the CoM is at the surface was SRAIE for the {100} surface and DIESA for the {111} surface.

6.4.1 Residues 190-195 (ELRDA & LRDAL)

Unconstrained System

R values for pentapeptides in the region of 190-195 can be found in Tables 6.13 and 6.14.

Pentapeptide	Fe type	Surface	System	r (Å)
ELRDA	FET	100	*in vacu*	2.03
			solvated	2.08
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.08
			solvated	2.13
		111	*in vacu*	1.93
			solvated	2.03
ELADA	FET	100	*in vacu*	1.98
			solvated	2.03
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.03
			solvated	2.13
		111	*in vacu*	1.98
			solvated	2.03

Table 6.13: R (Å) values from RDF data for pentapeptides in the region of 190-195.

Figure 6.12 exhibits the RDF profiles for ELRDA. The RDF profiles of the remaining 190-195 pentapeptides can be found in the Appendix. The 190-195 pentapeptides RDF data shows that, for ELRDA and its mutant ELADA, both octahedral and tetrahedral iron bonding was exhibited for the {100} surface. Whereas, with the {111} surface only octahedral iron bonding was present. However, this trend differed for LRDAL and its mutant LRAAL. With LRDAL, both octahedral and tetrahedral iron bonding was exhibited for the {111} surface. Whereas, with the {100} surface, only octahedral iron bonding was present. LRAAL exhibited a similar trend to both ELRDA and ELADA, however, there was also no tetrahedral iron bonding for the solvated LRAAL {100} system. When the 190-195 pentapeptides were attached to the {111} surface, octahedral iron attachment was the preferred attachment type.

Pentapeptide	Fe type	Surface	System	r (Å)
LRDAL	FET	100	*in vacu*	-
			solvated	-
		111	*in vacu*	2.03
			solvated	1.98
	FEO	100	*in vacu*	2.08
			solvated	2.13
		111	*in vacu*	2.03
			solvated	2.03
LRAAL	FET	100	*in vacu*	2.18
			solvated	-
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.13
			solvated	2.18
		111	*in vacu*	2.03
			solvated	2.13

Table 6.14: Continued. R (Å) values from RDF data for pentapeptides in the region of 190-195.

The iron bonding trend was the same for all 190-195 pentapeptides, with the exception of LRDAL, suggesting that the sequence mutation had no effect on the resultant iron bonding type. The change in trend for LRDAL suggests that, the presence of two oppositely charged residues adjacent to each other effects the iron bonding types possible. The presence of the positively charged residue in LRAAL may also explain the removal of tetrahedral iron bonding in the {100} solvated system.

The 190-195 pentapeptide RDF data also shows that, the *in vacu* systems produced a higher intensity first peak than the corresponding solvated systems, suggesting that, upon system solvation, it was more difficult pentapeptide oxygens to have an iron ion as a next nearest neighbour. For the solvated systems, the strongest g(r) peak was shown for octahedral iron bonding in the ELADA {111} system, proposing that in this system the pentapeptide oxygen present was more freely accessible to the next nearest neighbouring irons.

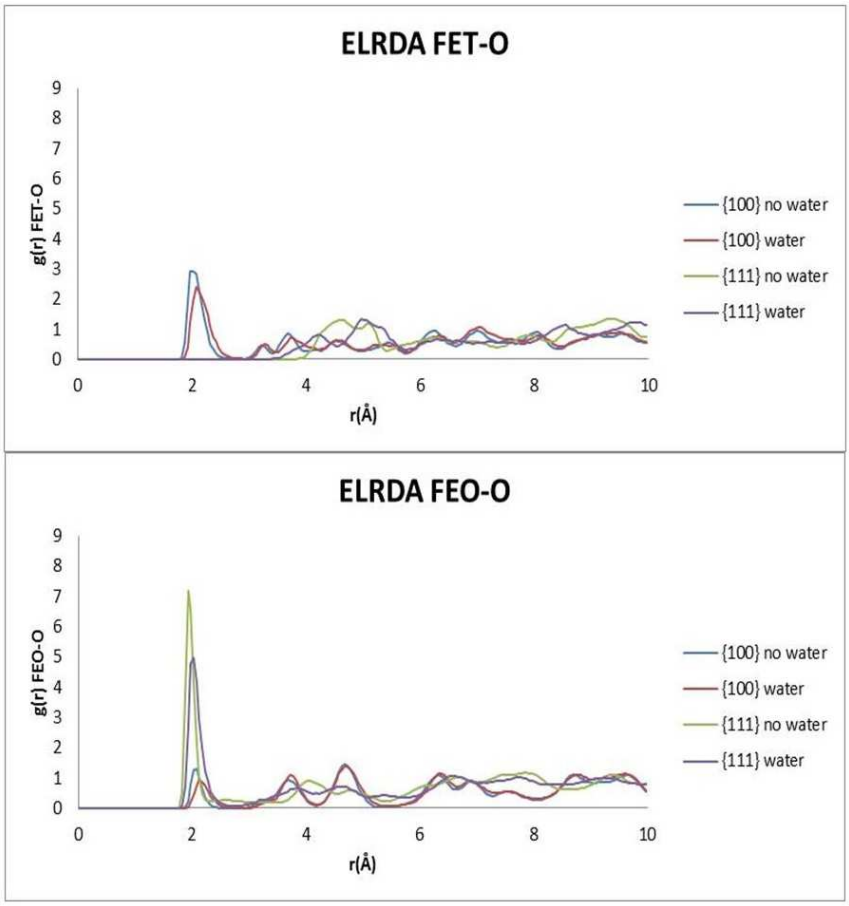

Figure 6.12: RDF plots for ELRDA. FET is tetrahedral iron, FEO is octahedral iron. Blue is {100} *in vacu*. Red is {100} solvated. Green is {111} *in vacu*. Purple is {111} solvated.

	{100} surface attachment			{111} surface attachment		
Pentapeptide	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
ELRDA	NBO	2.32	2.25	E O	1.92	2.27
	E-L	2.43	2.19		1.98	1.97
	L-R	2.14	2.36	D O	1.93	NB
	D O	1.97	2.11		1.88	1.99
		1.96	1.97	D-A	1.88	2.25
ELADA	NBO	2.00	NB	NBO	1.99	2.08
	E-L	2.09	NB	E O	1.92	1.97
	A-D	2.08	2.15		2.02	2.22
	D O	1.93	2.02	E-L	2.10	2.01
		2.12	NB	A-D	1.97	2.25
				D O	1.96	2.19
					1.96	NB
				D-A	2.24	NB

Table 6.15: 190-195 pentapeptide $FE - O_{pentapeptide}$ bond distance data. BL is bond length, E O is glutamate side chain oxygen, D O is aspartate side chain oxygen, x-x PBO is inter–residue peptide bond oxygen and NBO is N–terminal peptide bond oxygen. NB is no bonding.

	{100} surface attachment			{111} surface attachment		
Pentapeptide	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
LRDAL	NBO	2.21	2.14	NBO	2.08	NB
	L-R	2.01	2.35	D O	2.05	1.96
	R-D	1.97	NB		1.89	NB
	D-A	2.21	NB		2.00	1.94
	A-L	2.10	NB	D-A	2.19	NB
LRAAL	A-A	2.10	NB	NBO	1.99	NB
				L-R	2.07	NB
				R-A	1.95	2.23

Table 6.16: Continued. 190-195 pentapeptide $FE-O_{pentapeptide}$ bond distance data. BL is bond length, D O is aspartate side chain oxygen, x-x PBO is inter–residue peptide bond oxygen and NBO is N–terminal peptide bond oxygen. NB is no bonding.

The bond distance data is represented in Tables 6.15 and 6.16. The 190-195 pentapeptides bond distance data (Figure 6.13) shows that, for the ELRDA *in vacu* systems, bonding differed dependant on the attachment surface. On the {111} surface, bonding was present through both side chain oxygens of glutamate, however, no glutamate bonding was apparent for the {100} surface. The {100} surface showed N–terminal, E-L and L-R peptide bond oxygen bonding, whereas, the {111} surface exhibited D-A peptide bond oxygen bonding. The only commonality in bonding type between the two surfaces was the bonding through both side chain oxygens of aspartate. This data suggests that, for the ELRDA sequence, there was no surface preference, however, there was more of an affinity for particular oxygen types dependent on surface type (acidic side chain oxygen on the {111} surface and peptide bond oxygen on the {100} surface). On the {100} surface, the amount of bonding present does not change upon solvation, suggesting that the addition of water had minimal effect on the iron bonding potential of the sequence. However, upon solvation of the {111} surface, one of the bonds to one of the side chain oxygens of aspartate was no longer present, proposing that the presence of water in this system was detrimental to iron binding. The average bond lengths for a magnetite iron to a pentapeptide oxygen in the ELRDA sequence was; 2.16 Å and 2.18 Å on the {100} surface for *in vacu* and solvated systems respectively, and 1.92 Å and 2.12 Å on the {111} surface for *in vacu* and solvated systems respectively.

With the mutation of the ELRDA sequence to ELADA, the same amount of bonding was seen as for ELRDA in the {100} *in vacu* system, although, the type of bonding differed, with bonding through A-D peptide bond oxygen for ELADA, as opposed to L-R peptide bond oxygen bonding. On the {111} surface, however, more bonding was exhibited for ELADA, with the same bonding type exhibited as for ELRDA but with additional bonding to N–terminal, E-L and A-D peptide bond oxygen bonding. This could be due to the removal of the positively charged arginine residue, which may cause repulsive ionic interactions with the positively charged iron ions, leading to potential system stabilisation upon sequence mutation. The addition of water to the system was detrimental to bonding, in particular for the {100} surface, reducing the amount of bonding present. The average bond lengths for an iron to a pentapeptide oxygen in ELADA was; 2.04 Å and 2.09 Å on the {100} surface for the *in vacu* and solvated systems respectively, and 2.02 Å and 2.12 Å on the {111} surface for the *in vacu* and solvated systems respectively.

The bond distance data for the LRDAL *in vacu* systems shows that, there was the same amount of bonding present as in the ELRDA sequence, however, the bonding type differed,

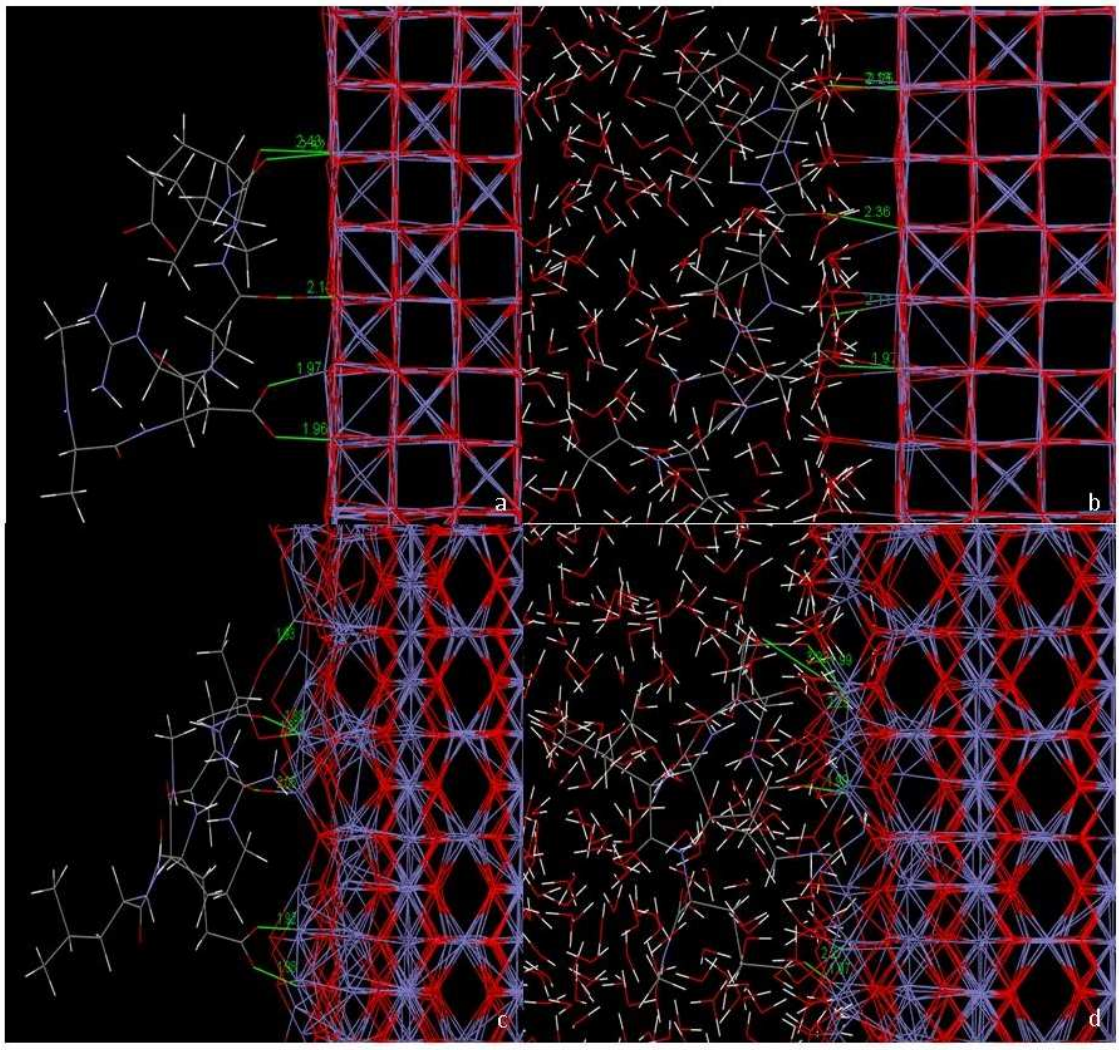

Figure 6.13: Bond distance images for the *in vacu* and solvated systems of ELRDA. a){100} *in vacu*, b){100} solvated, c){111} *in vacu* and d){111} solvated.

as for the {100} surface there was only peptide bond oxygen bonding present in LRDAL, and for the {111} surface there was more peptide bond oxygen bonding and obviously no glutamate bonding displayed for LRDAL. This data suggests that, although the amount of potential iron binding was not sequence dependent, the type of bonding was. For the LRDAL *in vacu* systems, both surfaces exhibited N–terminal and D-A peptide bond oxygen bonding. The {100} surface also displayed L-R, R-D and A-L peptide bond oxygen bonding, whereas, the {111} surface presented with three bonds through the side chain oxygens of aspartate. This data suggests that, there was no surface preference for the LRDAL sequence, however, there was more of an affinity for particular oxygen types dependent on the surface involved (acidic side chain oxygen on the {111} surface and peptide bond oxygen on the {100} surface, as with ELRDA). Both surfaces exhibited less bonding when the systems were solvated, suggesting more favourable bonding between

iron ions and the oxygen of water. The average bond lengths for a magnetite iron to a pentapeptide oxygen in the LRDAL sequence was; 2.10 Å and 2.25 Å on the {100} surface for *in vacu* and solvated systems respectively, and 2.04 Å and 1.95 Å on the {111} surface for *in vacu* and solvated systems respectively.

The LRAAL sequence displayed less bonding than the LRDAL sequence, particularly for the {100} surface. This was hypothesised to be due to the removal potential oxygen binding sites, with the substitution of aspartate. The two surfaces exhibited different bonding, with the {100} surface showing only A-A peptide bond oxygen bonding, whilst, the {111} surface presented with N–terminal, L-R and R-A peptide bond oxygen bonding, proposing that different surface types affect the amount of potential iron binding possible. Once again, solvation was detrimental to bonding, as the {100} surface displayed no bonding after solvation, and the {111} surface showed only R-A peptide bond oxygen bonding. The severe lack of bonding present for the LRAAL systems suggests that, due to an absence of acidic residues, there were very limited areas along the sequence that could potentially provide iron binding sites. The average bond lengths for an iron to a pentapeptide oxygen in LRAAL was; 2.10 Å for the *in vacu* system on the {100} surface, and 2.01 Å and 2.23 Å on the {111} surface for the *in vacu* and solvated systems respectively.

Tables 6.17, 6.18, 6.19 and 6.20 represent the residence times and average coordination numbers for bonding between all oxygen present in the 190-195 pentapeptides and tetrahedral and/or octahedral iron on both the {100} and {111} magnetite surfaces.

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
ELRDA {100} surface	D-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D O	0.00	0.00	4477.00	0.89	0.00	0.00	174.90	0.95
	D O	4990.00	1.00	0.00	0.00	999.00	1.00	0.00	0.00
	R-D PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	L-R PBO	4991.00	1.00	0.00	0.00	175.25	0.92	0.00	0.00
	E-L PBO	0.00	0.00	4991.00	0.99	0.00	0.00	355.70	0.95
	E O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NBO	585.47	0.95	0.00	0.00	46.24	0.85	0.00	0.00
ELRDA {111} surface	D-A PBO	0.00	0.00	4978.90	1.00	0.00	0.00	999.00	0.99
	D O	0.00	0.00	4026.94	1.34	0.00	0.00	185.12	0.75
	D O	0.00	0.00	4989.00	1.00	0.00	0.00	998.54	1.12
	R-D PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	L-R PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E-L PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	4988.50	1.00	0.00	0.00	625.40	1.75
	E O	0.00	0.00	4791.13	1.08	0.00	0.00	230.86	0.97
	NBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.17: Residence times (RT) and average coordination numbers (av. CN) of ELRDA. (E O = side chain oxygen of glutamate, D O = side chain oxygen of aspartate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
ELADA {100} surface	D-A PBO	1.00	0.00	1.11	0.01	0.00	0.00	1.73	0.03
	D O	0.00	0.00	4994.00	1.00	0.00	0.00	3.46	0.16
	D O	4997.00	1.00	0.00	0.00	999.00	1.00	0.00	0.00
	A-D PBO	0.00	0.00	4974.82	1.00	0.00	0.00	253.35	0.91
	L-A PBO	32.00	0.07	0.00	0.00	1.00	0.00	8.72	0.03
	E-L PBO	1643.70	0.93	0.00	0.00	56.48	0.37	0.00	0.00
	E O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NBO	0.00	0.00	765.62	0.97	0.00	0.00	5.36	0.23
ELADA {111} surface	D-A PBO	0.00	0.00	1503.89	0.99	0.00	0.00	1.50	0.00
	D O	2.16	0.01	2993.42	0.60	0.00	0.00	7.00	0.01
	D O	0.00	0.00	1635.68	1.00	0.00	0.00	462.69	1.79
	A-D PBO	0.00	0.00	4994.00	1.00	0.00	0.00	999.00	1.00
	L-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E-L PBO	0.00	0.00	4995.00	1.00	0.00	0.00	358.15	0.91
	E O	0.00	0.00	1532.57	1.59	0.00	0.00	999.00	1.00
	E O	0.00	0.00	3435.03	1.31	0.00	0.00	999.00	1.99
	NBO	0.00	0.00	4979.81	1.00	0.00	0.00	307.53	0.99

Table 6.18: Residence times (RT) and average coordination numbers (av. CN) of ELADA. (E O = side chain oxygen of glutamate, D O = side chain oxygen of aspartate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
LRDAL {100} surface	A-L PBO	0.00	0.00	4992.00	0.99	0.00	0.00	38.22	0.80
	D-A PBO	0.00	0.00	1548.02	0.97	0.00	0.00	1.67	0.01
	D O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	R-D PBO	0.00	0.00	70.33	0.88	0.00	0.00	7.89	0.40
	L-R PBO	0.00	0.00	4994.00	1.00	0.00	0.00	172.99	0.84
	NBO	0.00	0.00	4997.00	0.98	0.00	0.00	80.78	0.76
LRDAL {111} surface	A-L PBO	0.00	0.00	13.54	0.10	0.00	0.00	0.00	0.00
	D-A PBO	0.00	0.00	3848.00	0.77	0.00	0.00	3.36	0.03
	D O	1.00	0.00	1681.99	0.99	0.00	0.00	643.68	0.97
	D O	4724.91	0.95	4990.00	1.00	292.24	0.58	350.65	0.85
	R-D PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	L-R PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NBO	0.00	0.00	4982.63	1.00	0.00	0.00	37.18	0.11

Table 6.19: Residence times (RT) and average coordination numbers (av. CN) of LRDAL. (D O = side chain oxygen of aspartate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
LRAAL {100} surface	A-L PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	A-A PBO	0.00	0.00	1677.08	0.98	0.00	0.00	7.84	0.03
	R-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	L-R PBO	293.90	0.12	1.00	0.01	0.00	0.00	0.00	0.00
	NBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
LRAAL {111} surface	A-L PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	A-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	R-A PBO	0.00	0.00	4990.18	1.00	0.00	0.00	205.17	0.99
	L-R PBO	0.00	0.00	1400.99	0.99	0.00	0.00	74.40	0.32
	NBO	0.00	0.00	4992.00	1.00	0.00	0.00	56.92	0.25

Table 6.20: Residence times (RT) and average coordination numbers (av. CN) of LRAAL. (x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

The 190-195 pentapeptides residence time data shows that, for the ELRDA *in vacu* systems, when attached to the {100} surface, there were four bonds present that lasted for over 90% of the simulation length, (tetrahedral iron bonding to L-R peptide bond oxygen and to one of the side chain oxygens of aspartate and octahedral iron bonding to E-L peptide bond oxygen and to the other side chain oxygen of aspartate). No other bonding was exhibited lasting for over 12% of the full simulation for this system. Upon solvation, only one bond was present for over 99% of the simulation length, between tetrahedral iron and one of the side chain oxygens of aspartate. Again, no other bonding lasting over 36% of the simulation was present for this system, suggesting preferential magnetite–water bonding.

When the residence time data was examined for ELRDA on the {111} surface, a different trend was shown, with no tetrahedral iron bonding present for either system. For the *in vacu* system, bonding for over 81% of the full length of the simulation was displayed between octahedral iron and both side chain oxygens of glutamate, both side chain oxygens of aspartate and D-A peptide bond oxygen. No other bonding was seen for this system. Upon solvation, only the bonds between octahedral iron and one of the side chain oxygens of aspartate and D-A peptide bond oxygen continued to last over 99% of the simulation length and no other bonding present lasted over 63% of the full length of the simulation. This data suggests that, the addition of water, provides competitive binding to the magnetite.

The residence time data for the single alanine substitution to ELADA, in the {100} *in vacu* system, showed bonding lasting over 99% of the simulation length, exhibited between tetrahedral iron and one of the side chain oxygens of aspartate, and between octahedral iron and A-D peptide bond oxygen and the other side chain oxygen of aspartate. No other bonding lasting over 33% of the simulation was seen. When this system was solvated, only the bonding between tetrahedral iron and one of the side chain oxygens of aspartate continued lasting over 99% of the full length of the simulation, whilst, no other bonding lasted over 26% of the simulation length. This data suggests that, again, water was providing competitive binding.

On examination of the corresponding {111} surface data, no tetrahedral iron bonding was present, with the exception of the tetrahedral iron bond to one of the side chain oxygens of aspartate lasting for only 0.04% of the simulation length in the *in vacu* system. With the octahedral iron, bonding was exhibited to N–terminal, E-L and A-D peptide bond oxygen that lasted over 99% of the full length of the simulation. Bonding lasting between

60% and 69% of the simulation length was also seen to one of the side chain oxygens of glutamate and one of the side chain oxygens of aspartate. No other bonding lasting over 35% of the simulation length was exhibited for this system. With the addition of water, the bond to A-D peptide bond oxygen continued to last over 99% of the full length of the simulation, however, the two other bonds that lasted over this length of time in the *in vacu* simulation no longer did, now presenting with bonds lasting between 31% and 36% of the simulation length. Both side chain oxygens of glutamate exhibited bonding lasting over 99% of the simulation length, suggesting that the lower values for the equivalent *in vacu* residence times are due to the formation of this bond toward the latter stages of this system simulation. No other bonding lasting over 47% of the simulation length was exhibited.

When the residence times of LRDAL {100} systems were considered it was shown that, no tetrahedral iron bonding was present, which was a very different trend to that seen for ELRDA (and ELADA). In the *in vacu* system, three bonds present lasted over 99% of the full length of the simulation; between octahedral iron and N–terminal, L-R, and A-L peptide bond oxygen. No other bonding lasting over 31% was exhibited. When this system was solvated, no bonding was present lasting over 18% of the full length of the simulation, suggesting that the addition of water to the system was severely detrimental to the LRDAL {100} surface iron binding potential.

When the residence time data was examined for LRDAL on the {111} surface, a tetrahedral iron bond to one of the side chain oxygen of aspartate was present lasting over 95% of the simulation length, this, again, was different to the trend seen for ELRDA and ELADA. Bonding lasting for over 99% of the full length of the simulation was also seen between octahedral iron and N–terminal peptide bond oxygen, and one of the side chain oxygens of aspartate. Bonding was also exhibited between octahedral iron and D-A peptide bond oxygen lasting for 77% of the simulation length. No other bonding lasting over 34% of the full simulation length was displayed. Upon solvation, there was no bonding present lasting over 65% of the simulation length. This reduction in residence times suggests, again, preferential magnetite–water bonding.

On comparison of the residence time data for the two original sequences (ELRDA and LRDAL) it can be seen that, for the {100} *in vacu* system, ELRDA had more bonding lasting over 90% of the simulation length (four bonds for ELRDA and three bonds for LRDAL). Upon solvation, ELRDA exhibited more bonding lasting over 99% of the simulation length than LRDAL (one bond for ELRDA and no bonds for LRDAL). When the

{111} surface *in vacu* system was examined it was shown that, ELRDA, again, exhibited more bonding lasting over 81% of the simulation length (five bonds for ELRDA and three bonds for LRDAL). Upon addition of water, this trend was continued with ELRDA presenting with two bonds that lasted over 99% of the full length of the simulation, whilst, LRDAL exhibited no bonding lasting for this length of time. From the solvated systems data it can be seen that, ELRDA was the preferred sequence of attachment, with the {111} surface being the favoured for this sequence.

The residence time data for LRAAL attached to the {100} surface showed that, there was no bonding exhibited lasting over 34% of the simulation, with one bond present for only 0.02% of the simulation length, suggesting that this sequence was very unfavourable for attachment to magnetite.

When residence time data for the {111} surface showed that, no tetrahedral iron bonding was present for system. For the *in vacu* system, bonding lasting for over 99% of the full length of the simulation was seen between octahedral iron and N–terminal and R-A peptide bond oxygen. No other bonding lasting over 29% of the simulation length was exhibited. Upon solvation, no bonding was present that lasted over 21% of the simulation length. The {111} surface data, again, suggests that LRAAL attachment on this surface was not favourable.

On comparison of the residence time data for ELADA and LRAAL it can be seen that, for the {100} surface *in vacu* system, ELADA had more bonding lasting over 99% of the simulation length than LRAAL (three bonds for ELADA and no bonds for LRAAL), suggesting that ELADA was the preferred sequence of attachment. When these systems were solvated, this trend was seen again, with ELADA exhibiting one bond lasting over 99% of the full length of the simulation, whilst, LRAAL presented with no bonds lasting for this length of time. For the {111} *in vacu* system it was shown that, ELADA again, had more bonding lasting over 99% of the simulation length, exhibiting three bonds of this time length, whilst, LRAAL presented with two bonds of this length scale, again proposing that ELADA was favoured for attachment to magnetite. Upon solvation, ELADA presented with three bonds that lasted over 99% of the full length of the simulation, whereas, LRAAL exhibited no bonding for this time scale. The solvated system data showed that, ELADA was the preferred sequence of attachment, with the {111} surface displaying the strongest attachment. The retention time data for the 190-195 pentapeptides also showed that, the 190-194 sequence and its mutation (R192A) were preferred over the 191-195 sequence and its mutation (D193A).

All related energy data from the interfacial energy calculations of the 190-195 pentapeptides can be found in Table 6.21.

{100} surface attachment				
	E_{int} of s-p (eV)	E_{diff} of s-p (eV)	E_{int} of s-p-w (eV)	E_{diff} of s-p-w (eV)
ELRDA	-27.44	22.25	6.42	2.43
ELADA	-5.19	0.00	8.85	0.00
LRDAL	-25.68	20.49	8.71	0.13
LRAAL	-16.55	11.37	5.98	2.86
{111} surface attachment				
	E_{int} of s-p (eV)	E_{diff} of s-p (eV)	E_{int} of s-p-w (eV)	E_{diff} of s-p-w (eV)
ELRDA	-42.49	33.42	3.58	15.97
ELADA	-9.06	0.00	11.22	8.33
LRDAL	-24.22	15.16	19.55	0.00
LRAAL	-16.88	7.82	16.58	2.97

Table 6.21: Interfacial energies for all 190-195 pentapeptides on the {100} and {111} surface, using the Yang, Stipp and Harding method[164]. E_{int} of s-p is the interfacial energy of the slab–peptide system. E_{diff} of s-p is the difference in interfacial energy from the highest interfacial energy of the slab–peptide system. E_{int} of s-p-w is the interfacial energy of the slab–peptide–water system. E_{diff} of s-p-w is the difference in interfacial energy from the highest interfacial energy of the slab–peptide–water system.

From the solvated 190-195 pentapeptides interfacial energy data it can be seen that, the {100} surface produced lower interfacial energy values (between 5.98 eV and 8.85 eV) than the {111} surface, with the exception of ELRDA (between 11.22 eV and 19.55 eV (excluding ELRDA at 3.58 eV)). For the {100} surface, ELADA produced a larger interfacial energy than ELRDA, suggesting that the mutation reduced the iron binding potential of the sequence. However, for LRAAL, a smaller interfacial energy was produced than LRDAL, suggesting that the substitution of the acidic residue improved the iron binding potential of the sequence. For the {111} surface, ELADA produced a much larger interfacial energy than ELRDA, suggesting that the substitution is detrimental to the interfacial energy of the system, whereas, as with the {100} surface, LRAAL produced a lower interfacial energy, suggesting that mutation improved the interfacial energy of the system.

This data also shows that, solvation had a detrimental effect on the iron binding potential of the sequences, as the *in vacu* energies produced were much lower than the solvated energies, for example, on the {111} surface, ELRDA exhibited a substantial increase in

interfacial energy from -42.49 eV to 3.58 eV. The addition of water to the system also affected the pentapeptides interfacial behaviour, as the interfacial energy trends differed for the *in vacu* and solvated systems. For example, the LRDAL {100} *in vacu* showed one of the lowest interfacial energies, however, upon solvation, it exhibited one of the highest. Preferential attachment of water to the magnetite surface could explain the increase in interfacial energy with the addition of water.

Again, there are concerns with this direct method of calculating the adsorption energy, with larger than expected energy values produced and expected trends not exhibited, possible reasons for which were discussed for the 177-183 pentapeptides. An alternative indirect method of adsorption energy calculation is explored in the following section. This method utilises a constrained molecular dynamics version of the systems, focusing on the Potential of Mean Force and free energy evaluation.

6.5 Constrained system

Free energy profiles were created for 190-195 pentapeptide sequences, as shown in Figure 6.14.

The free energy profiles for the different 190-195 pentapeptides (Figure 6.14) showed that, on the {100} surface, the mutation of ELRDA to ELADA increased the iron binding activity, with the exception of between 2 Å and 4 Å , whereby, the ELADA sequence slightly increased the free energy of the system, compared to that seen for ELRDA. From this data it was also evident that, the mutation of LRDAL to LRAAL greatly reduced the iron binding activity of the sequence, as hypothesised, with LRDAL exhibiting energies as low as -0.47 eV at its minima and LRAAL exhibiting energies as high as 7.38 eV at its maxima. The mutant LRAAL produced the highest free energy of all 190-195 pentapeptide sequences. For ELRDA, ELADA and LRAAL, between 10 Å and 5.5 Å was optimal for attachment, after which point, an increase in free energy was detected (a substantial increase for LRDAL), corresponding with the presence of the more intense water peak. It is proposed that the successive increase in free energy was attributable to a requirement for more energy within the system for the sequences to pass through the water barrier in order to enable sequence attachment to the {100} surface. The much larger energy increase for LRDAL, suggests that it was considerably harder for this sequence to attach within close range of the magnetite surface, for this system, compared to ELRDA and its mutation. However, for LRDAL, the optimal attachment region was between 8.5 Å and

1.5 Å as the profile exhibited negative values within this region. From this distance toward the surface an increase in free energy was exhibited, proposing that it was more difficult for the sequence to attach closer to the magnetite surface.

Whereas, with the {111} surface energy profile, there was little difference between the free energy profiles of the four sequences, with the exception of LRDAL. It can also be seen that, mutation of the original sequences appears to slightly increase the iron binding affinity of the sequences. The original sequence LRDAL produced the highest free energy of all 190-195 pentapeptide sequences. For ELRDA, LRDAL and LRAAL, between 10 Å and 3.75 Å was optimal for attachment, after which point, an increase in free energy was detected (a more substantial increase for LRDAL), corresponding with the presence of the intense water layers. ELADA exhibited a negative free energy throughout its profile, thus, attachment was favourable at any distance explored, however, the free energy of this sequence was at its lowest when the CoM was at 1 Å from the magnetite surface, suggesting that this was the preferred distance for attachment. The small changes in free energy seen for the sequence profiles, with the exception of LRDAL, suggest that the water absorption layers have minimal influence on the free energy profiles.

Upon examination of the energy minima for the 190-195 pentapeptides (Figure 6.15) it can be observed that, on the {100} surface, only ELADA and LRDAL exhibited a negative value for their energy minima, with LRDAL producing the lowest value at -0.4681 eV. The energy minima for these residues fall between 4.1 Å and 5.7 Å, suggesting that attachment of the residue centre of mass at this distance would be preferable. On the {111} surface, all of the 190-195 pentapeptides except for LRAAL exhibited a negative value for their energy minima, with ELADA producing the lowest value at -0.3559 eV. The trend of favourable attachment distance differs from the {100} surface, with ELRDA, ELADA and LRDAL exhibiting centre of mass distances of 6.6 Å , 1 Å and 6.3Å respectively.

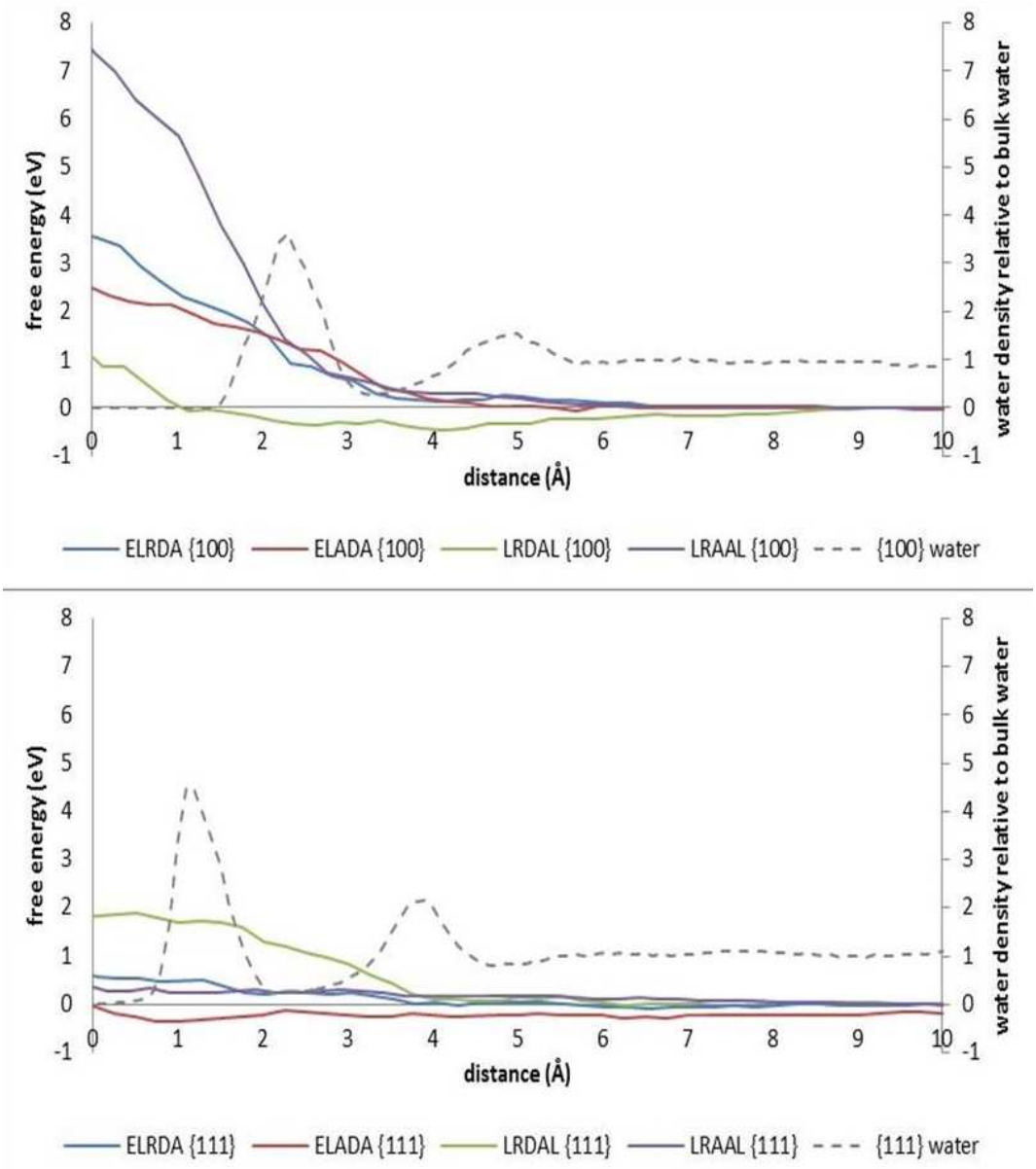

Figure 6.14: A comparison of the free energy profiles of the 190-195 pentapeptides. Distance refers to distance between the peptide CoM and the magnetite surface. The dashed line represents the water density profile for the system.

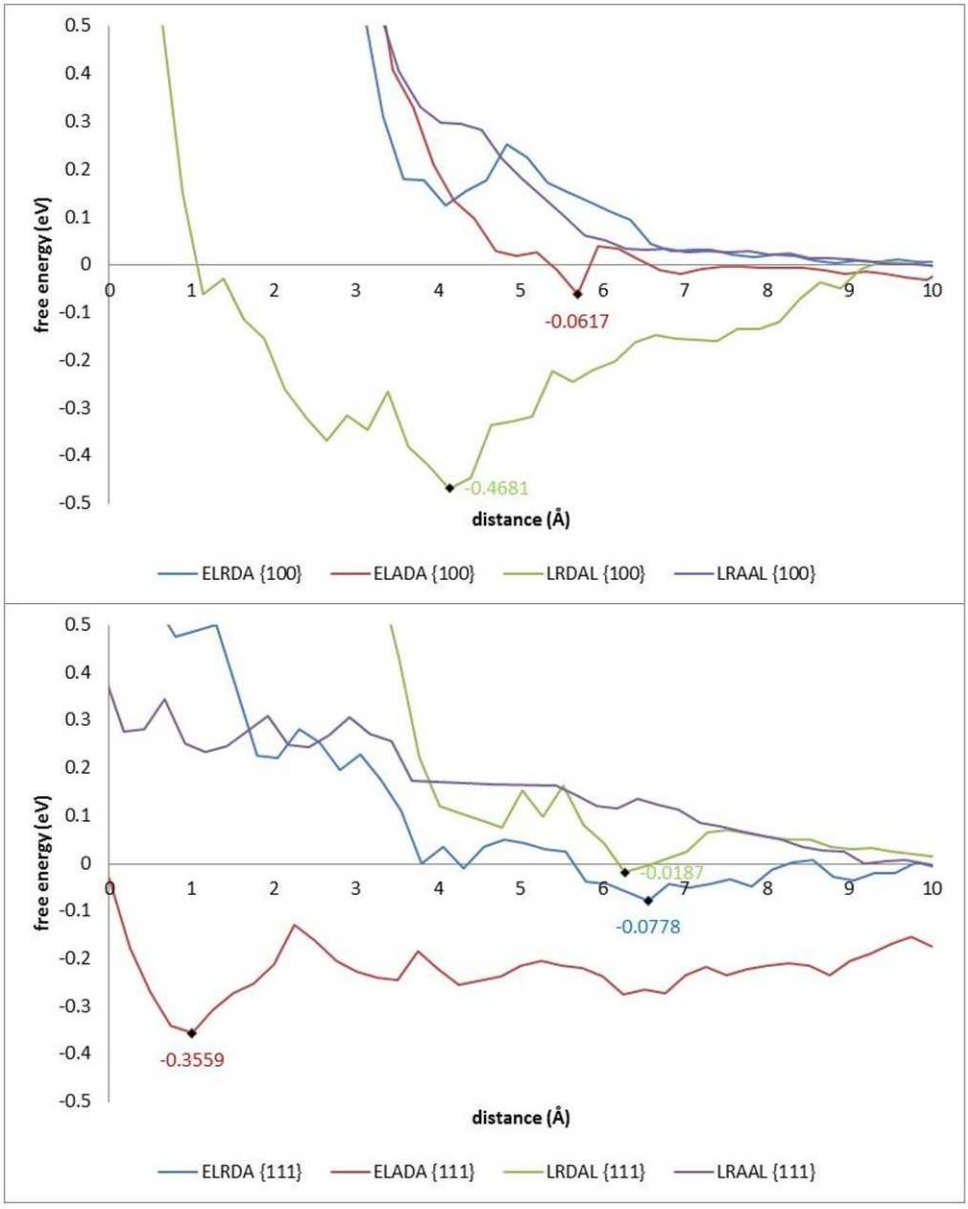

Figure 6.15: A close-up comparison of the free energy profiles of the 190-195 pentapeptides between -0.5 eV and 0.5 eV, exhibiting the free energy minima.

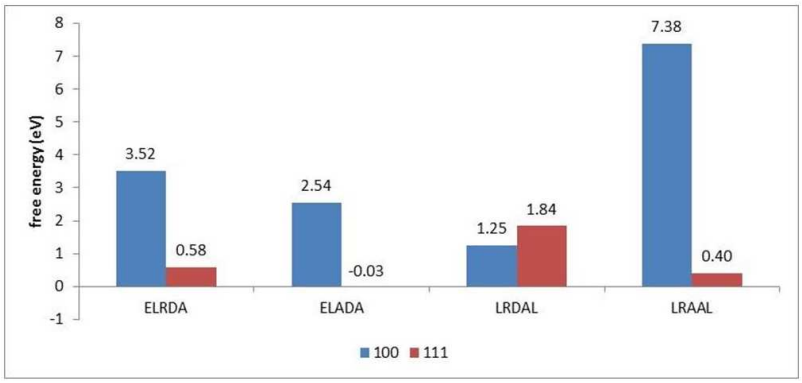

Figure 6.16: A comparison of the free energy values for the 190-195 pentapeptide CoM at the magnetite surface for both the {100} and {111} surfaces.

The free energy values for the 190-195 pentapeptides centre of mass at the magnetite surfaces are shown in Figure 6.16. On the {100} surface, the ELRDA sequence exhibits a free energy value of 3.52 eV, whereas, on the {111} surface, its value was much lower, 0.58 eV. Similarly, the ELADA and LRAAL sequences demonstrated much higher free energy values (2.54 eV and 7.38 eV, respectively) on the {100} surface, than exhibited on the {111} surface (-0.03 eV and 0.40 eV, respectively), particularly for LRAAL. However, for LRDAL, the {111} surface exhibits a higher free energy value than the {100} surface, although the difference in energy was minimal (0.59 eV). This data showed that, with exception of LRDAL, the {111} surface exhibited considerably lower free energy values, representing greater iron binding, proposing that this was the preferential surface of attachment for the 190-195 pentapeptide sequences. The free energy profiles also showed that, the sequence with the lowest free energy when the CoM is at the surface was LRDAL for the {100} surface and ELADA for the {111} surface.

6.5.1 Residues 188-192 (EVELR)

Unconstrained System

R values for pentapeptides in the region of 188-192 can be found in Table 6.22.

Pentapeptide	Fe type	Surface	System	r (Å)
EVELR	FET	100	*in vacu*	2.18
			solvated	2.28
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.03
			solvated	2.03
		111	*in vacu*	2.03
			solvated	1.98
EVALR	FET	100	*in vacu*	2.33
			solvated	2.33
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.03
			solvated	2.03
		111	*in vacu*	2.03
			solvated	2.03
EVGELR	FET	100	*in vacu*	1.93
			solvated	1.93
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	1.98
			solvated	2.08
		111	*in vacu*	2.03
			solvated	2.08

Table 6.22: R (Å) values from RDF data for pentapeptides in the region of 188-192.

Figure 6.17 exhibits the RDF profiles for EVELR. The RDF profiles of the remaining 188-192 pentapeptides can be found in the Appendix. The RDF data for the 188-192 pentapeptides showed that, for all pentapeptides, both octahedral and tetrahedral iron

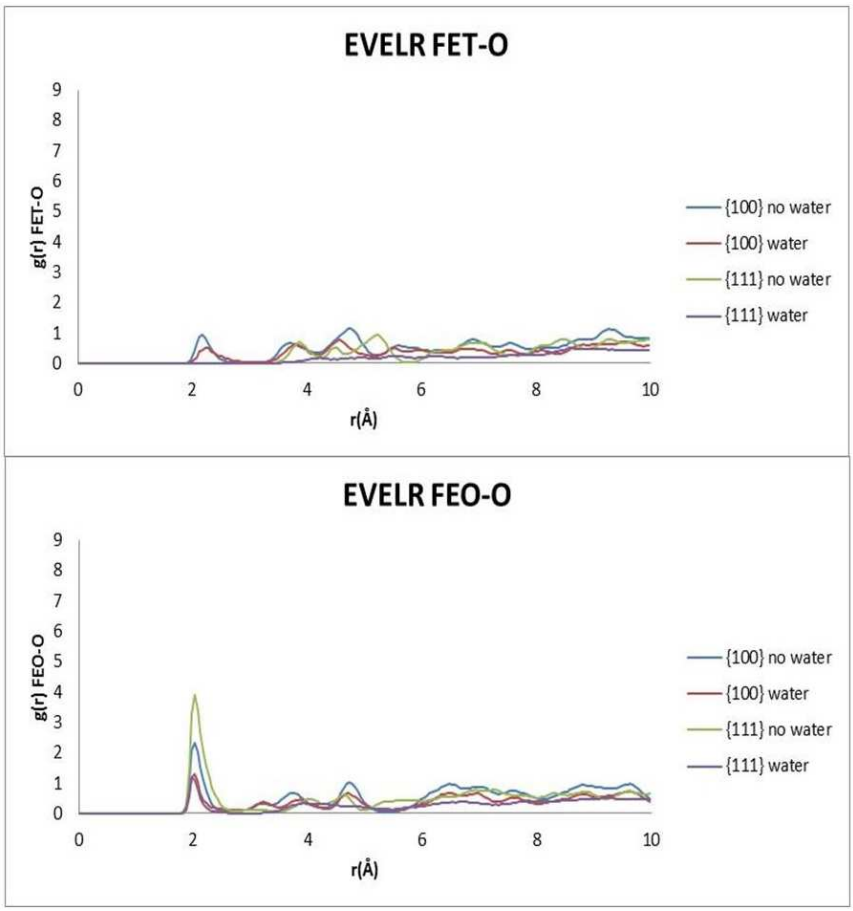

Figure 6.17: RDF plots for EVELR. FET is tetrahedral iron, FEO is octahedral iron. Blue is {100} *in vacu*. Red is {100} solvated. Green is {111} *in vacu*. Purple is {111} solvated.

bonding was exhibited for the {100} surface. Whereas, with the {111} surface, only octahedral iron bonding was present. When the 188-192 pentapeptides were attached to the {111} surface, octahedral iron attachment was the preferred attachment type. The iron bonding trend was the same for all 188-192 pentapeptides proposing that sequence alteration had no effect on the resultant type of iron bonding.

From the 188-192 pentapeptide RDF data it can also be seen that, the *in vacu* system produced a more intense first peak than the corresponding solvated system, suggesting that, upon addition of water to the systems, it was more difficult for a pentapeptide oxygen to have an iron as a next nearest neighbour. For the solvated systems, the strongest $g(r)$ peak was shown for through octahedral iron bonding in the EVALR {111} system, proposing that the oxygen present in this sequence was more freely accessible to the next nearest neighbouring irons.

The $FE - O_{pentapeptide}$ bond distance data for the 188-192 pentapeptides is represented in Tables 6.23 and 6.24.

	{100} surface attachment			{111} surface attachment		
Pentapeptide	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
EVELR	NBO	2.11	2.34	E-L	1.99	NB
	E O	2.01	1.98	E O	2.28	NB
		2.04	2.20		1.96	NB
	E-V	2.27	2.39		2.14	NB
EVALR	NBO	2.07	NB	NBO	2.09	2.22
	E O	1.96	2.12	E O	2.17	2.01
		2.06	1.94		1.92	2.26
	E-V	2.26	NB		1.93	2.09
					-	1.99
				V-A	1.99	2.09
				L-R	1.96	2.07

Table 6.23: 188-192 pentapeptide $FE - O_{pentapeptide}$ bond distance data. BL is bond length, E O is glutamate side chain oxygen, x-x PBO is inter–residue peptide bond oxygen and NBO is N–terminal peptide bond oxygen. NB is no bonding.

	{100} surface attachment			{111} surface attachment		
Pentapeptide	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
EVGELR	E O	2.06	2.34	NBO	2.13	NB
		2.05	1.84	E O	1.92	2.00
	E-V	2.20	NB		2.11	2.03
					2.03	2.03
					2.00	2.15
				G-E	1.97	NB
				E O	1.94	2.25
					2.11	1.97
				E-L	2.14	NB

Table 6.24: Continued. 188-192 pentapeptide $FE-O_{pentapeptide}$ bond distance data. BL is bond length, E O is glutamate side chain oxygen, x-x PBO is inter–residue peptide bond oxygen and NBO is N–terminal peptide bond oxygen. NB is no bonding.

The 188-192 pentapeptides bond distance data (Figure 6.18) shows that, bonding differed dependant on surface of attachment. For the EVELR *in vacu* system, the two surfaces showed bonding through both side chain oxygens of the glutamate (N–terminal glutamate for the {100} surface and C–terminal glutamate for the {111} surface). There were no other commonalities present, as the {100} surface showed N–terminal and E-V peptide bond oxygen bonding, whereas, the {111} surface exhibited only E-L peptide bond oxygen bonding. This data suggests that, as the EVELR sequence exhibited the same amount of bonding on each surface there was no surface preference, however, there was more of an affinity for side chain oxygen on the {111} surface, whereas, the {100} surface had more of an affinity to peptide bond oxygen, proposing that surface type had no effect on the bonding amount but did affect the bonding type. On the {111} surface, no bonding between magnetite iron and pentapeptide oxygen occurred when the system was solvated, suggesting more favourable bonding between magnetite iron and the oxygen of water. The average bond lengths for a magnetite iron to a pentapeptide oxygen in the EVELR sequence was; 2.11 Å and 2.23 Å on the {100} surface for *in vacu* and solvated systems respectively, and 2.09 Å in the solvated {111} surface system.

With the mutation to EVALR, on the {100} surface, the same amount and type of bonding was seen for the *in vacu* system, as was exhibited for EVELR, suggesting that, the mutation of the sequence in the *in vacu* system had no effect on the sequence iron binding affinity. For the {111} surface, three bonds to glutamate side chain oxygens were exhibited, as with EVELR, however, the mutation lead to more bonding being present than for the original sequence, with the addition of N–terminal, V-A and L-R peptide bond oxygen bonding being present in this system. The E-L peptide bond oxygen bond was not exhibited as it does not exist in the mutated sequence. These results suggest that substitution was beneficial to the iron binding affinity of the sequence. With the addition of water, for the {100} surface, the peptide bond oxygen bonds were no longer present, suggesting that bonding between magnetite iron and the oxygen of water was more favourable than to peptide bond oxygen. For the {111} surface, more bonding was seen when the system was solvated than when it was *in vacu*, again, suggesting that mutation increased the iron binding affinity of the sequence. The average bond lengths for an iron to a pentapeptide oxygen in EVALR was; 2.09 Å and 2.03 Å on the {100} surface for the *in vacu* and solvated systems respectively, and 2.01 Å and 2.10 Å on the {111} surface for the *in vacu* and solvated systems respectively.

For the glycine insertion mutant, EVGELR, less bonding was present for the {100} surface

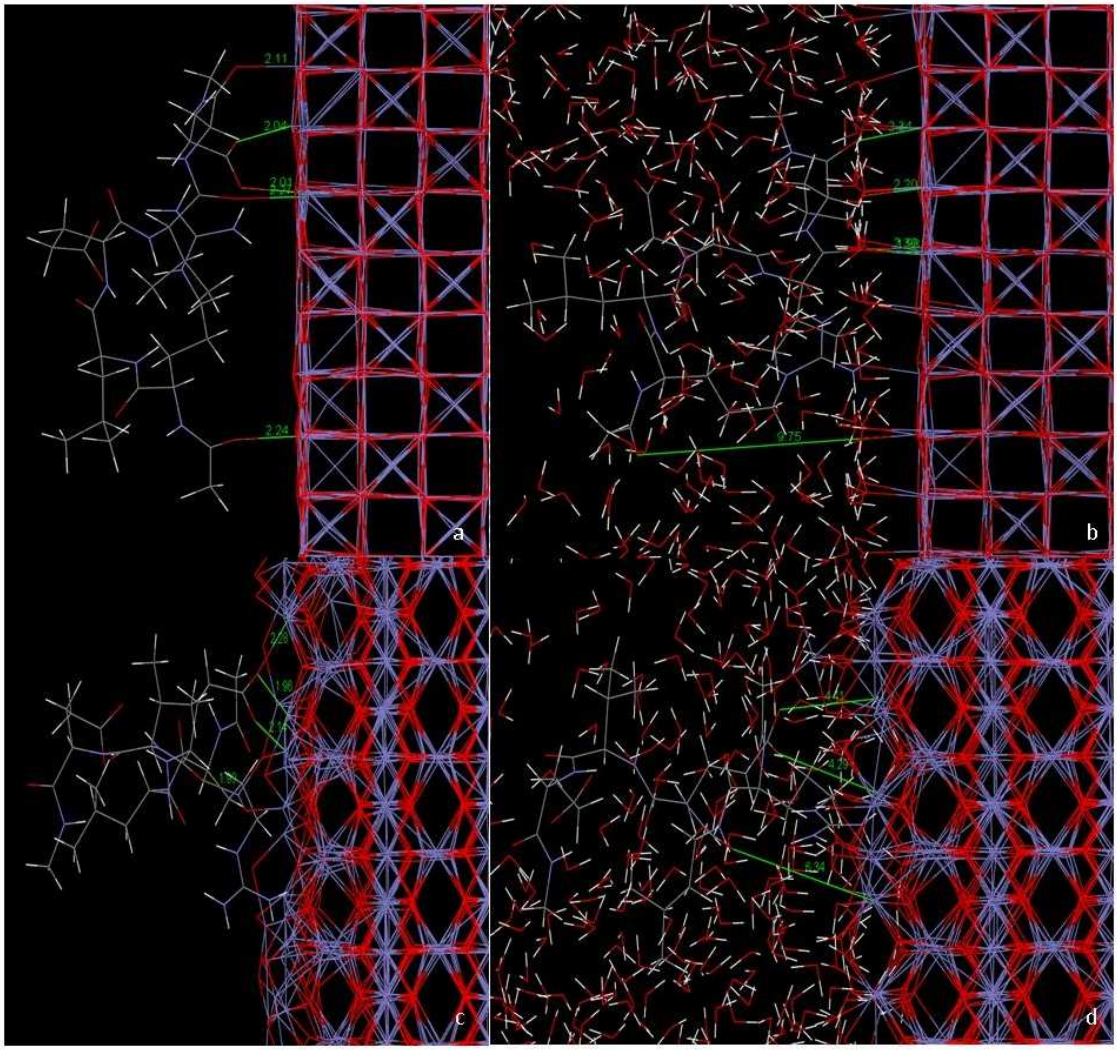

Figure 6.18: Bond distance images for the *in vacu* and solvated systems of EVELR. a){100} *in vacu*, b){100} solvated, c){111} *in vacu* and d){111} solvated.

than for EVELR and EVALR, with the lack of N–terminal peptide bond oxygen bonding, which was present in EVELR and EVALR. Hence, the addition of glycine into this position of the sequence was detrimental to the iron binding affinity. More bonding was present for EVGELR on the {111} surface than for EVELR and EVALR, as EVGELR exhibited four bonds to magnetite iron from the side chain oxygens of glutamate, whilst EVELR and EVALR displayed only three of such bonds. EVGELR also exhibited E-L peptide bond oxygen bonding (as with EVELR), and N–terminal peptide bond oxygen bonding (as with EVALR). Unlike the other 188-192 pentapeptide sequences, the glutamate side chain oxygen bonding for EVGELR was to both glutamate in the sequence. This sequence also exhibited G-E peptide bond oxygen bonding. This data suggests that the insertion of glycine greatly improved the iron binding affinity of the sequence. With the addition of water, no peptide bond oxygen bonding was present, again, suggesting preferential

bonding of magnetite to the oxygen of water than to peptide bond oxygen. This residence time data suggests that within the solvated systems, acidic residue bonding is the preferred method of attachment. The average bond lengths for an iron to a pentapeptide oxygen in EVGELR was; 2.10 Å and 2.09 Å on the {100} surface for the *in vacu* and solvated systems respectively, and 2.04 Å and 2.07 Å on the {111} surface for the *in vacu* and solvated systems respectively.

Tables 6.25, 6.26, 6.27 and 6.28 represent the residence times and average coordination numbers for the bonding present in the 188-192 pentapeptides.

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
EVELR {100} surface	L-R PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E-L PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	V-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E-V PBO	212.41	0.92	0.00	0.00	36.46	0.61	2.33	0.01
	E O	0.00	0.00	4946.96	1.01	0.00	0.00	999.00	1.00
	E O	0.00	0.00	4997.00	1.00	0.00	0.00	92.03	0.87
	NBO	0.00	0.00	4987.32	0.99	0.00	0.00	2.51	0.20
EVELR {111} surface	L-R PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E-L PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	1042.77	1.90	0.00	0.00	168.69	0.72
	E O	0.00	0.00	4974.47	1.99	0.00	0.00	169.49	0.34
	V-E PBO	0.00	0.00	4978.00	1.00	0.00	0.00	161.51	0.32
	E-V PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.25: Residence times (RT) and average coordination numbers (av. CN) of the EVELR. (E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
EVALR {100} surface	L-R PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	A-L PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	V-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E-V PBO	22.04	0.61	0.00	0.00	17.11	0.22	0.00	0.00
	E O	0.00	0.00	4973.34	1.00	0.00	0.00	159.03	0.98
	E O	0.00	0.00	4974.00	0.99	0.00	0.00	999.00	1.00
	NBO	0.00	0.00	4947.23	0.99	0.00	0.00	7.95	0.13
EVALR {111} surface	L-R PBO	0.00	0.00	1476.96	0.99	0.00	0.00	163.57	1.30
	A-L PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	V-A PBO	0.00	0.00	938.18	0.88	0.00	0.00	999.00	1.00
	E-V PBO	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	4070.40	1.80	0.00	0.00	999.00	2.00
	E O	0.00	0.00	3683.83	1.29	0.00	0.00	333.20	1.82
	NBO	0.00	0.00	61.44	1.23	0.00	0.00	32.09	1.00

Table 6.26: Residence times (RT) and average coordination numbers (av. CN) of the EVALR. (E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
EVGELR {100} surface	L-R PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E-L PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	G-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	V-G PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E-V PBO	1106.00	0.98	0.00	0.00	3.53	0.02	0.00	0.00
	E O	1.00	0.00	4891.61	1.00	20.20	0.05	479.00	0.89
	E O	4999.00	1.00	0.00	0.00	999.00	1.00	0.00	0.00
	NBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.27: Residence times (RT) and average coordination numbers (av. CN) of the EVGELR. (E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
EVGELR {111} surface	L-R PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E-L PBO	0.00	0.00	369.35	1.09	0.00	0.00	18.30	0.06
	E O	0.00	0.00	2206.00	0.43	0.00	0.00	333.45	0.98
	E O	0.00	0.00	4993.00	1.00	0.00	0.00	101.24	0.91
	G-E PBO	0.00	0.00	2000.66	0.71	0.00	0.00	113.50	0.23
	V-G PBO	0.00	0.00	982.27	0.40	0.00	0.00	0.00	0.00
	E-V PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	3406.64	1.98	0.00	0.00	703.60	1.98
	E O	0.00	0.00	4965.95	1.99	0.00	0.00	999.00	2.00
	NBO	0.00	0.00	4996.00	1.00	0.00	0.00	8.49	0.05

Table 6.28: Continued. Residence times (RT) and average coordination numbers (av. CN) of the EVGELR. (E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

The 188-192 pentapeptides residence time data shows that, for the EVELR *in vacu* system, when attached to the {100} surface, there were three bonds present lasting over 99% of the simulation length, all of which were through octahedral iron. This bonding was to N–terminal peptide bond oxygen and to both of the side chain oxygens of the N–terminal glutamate. No other bonding exhibited lasted over 5% of the simulation length. When this system was solvated, only one bond was present for over 99% of the full length of the simulation, between octahedral iron and one of the side chain oxygen of the N–terminal glutamate. No other bonding lasted over 10% of the simulation length for this system, suggesting that the presence of water was disadvantageous to the iron binding potential of this sequence.

When the equivalent {111} surface residence time data was examined, no tetrahedral iron bonding was found to be present for either system. For the *in vacu* system, bonding that lasted over 99% of the full length of the simulation was seen between octahedral iron and the V-E peptide bond oxygen and one of the side chain oxygen of the C–terminal glutamate. No other bonding exhibited for this system lasted over 21% of the simulation length. Upon solvation, no bonds were present that lasted over 17% of the simulation length, proposing that, the addition of water provides competitive binding to the magnetite.

The single alanine substitution to EVALR, for the {100} *in vacu* system, showed no bonding lasting over 0.5% of the simulation length exhibited through tetrahedral iron. Through octahedral iron, however, three bonds were present lasting over 99% of the full length of the simulation. These were to N–terminal peptide bond oxygen and to both of the side chain oxygens of glutamate. No other bonding was displayed for this system. Upon system solvation, only one bond was present for over 99% of the simulation length, between octahedral iron and one of the side chain oxygens of glutamate. No other bonding was present for this system. This data suggests that, water provides competition for iron binding.

The corresponding {111} surface data showed no tetrahedral iron bonding. With the *in vacu* system octahedral iron bonding, a bond was exhibited to one of the side chain oxygens of glutamate that lasted for 81% of the simulation length. A bond that lasted for 74% of the simulation length was also seen to the other side chain oxygen of glutamate. No other bonding present lasted over 30% of the simulation length. With the addition of water, two bonds were displayed which lasted over 99% of the simulation length. One of the bonds was to one of the side chain oxygen of glutamate, which continued from

the *in vacu* system. The other bond was to V-A peptide bond oxygen, suggesting that the shorter equivalent *in vacu* V-A peptide bond oxygen residence time was due to the formation of this bond toward the latter stages of this system simulation. For this system no other bonding lasting over 34% of the simulation length was exhibited.

On comparison of the residence time data for EVALR and EVELR, it can be seen that for the {100} *in vacu* system, EVALR had the same amount of bonding that lasted over 99% of the full length of the simulation as EVELR (three bonds for each sequence), and these bonds were also of the same type (octahedral iron to N–terminal peptide bond oxygen and to both of the side chain oxygens of the N–terminal glutamate). This data suggests that, the sequence mutation had no effect on the iron binding potential of the {100} *in vacu* system. When these systems were solvated, again, EVALR and EVELR presented with the same amount of bonding lasting over 99% of the full length of the simulation (one bond to the side chain oxygen of glutamate for both). When the {111} surface *in vacu* system was examined, EVELR showed more bonding lasting over 82% of the simulation length (two bonds for EVELR and one bond for EVALR). These results suggest that, mutation of EVELR reduced the potential iron binding of the sequence. Upon solvation, EVALR exhibited the most bonding lasting over 99% of the simulation length (two bonds for EVALR and no bonds for EVELR). Based on the solvated systems data, for the {100} surface, sequence mutation had no effect on iron binding, hence, attachment was not sequence dependent. EVALR was the preferred sequence of attachment for the {111} surface.

The residence time data for EVGELR showed that, for the {100} *in vacu* system, bonding was exhibited through tetrahedral iron to one of the side chain oxygens of the N–terminal glutamate and through octahedral iron to the other side chain oxygen of the N–terminal glutamate, lasting over 98% of the simulation length. No other bonding was exhibited lasting over 23% of the simulation length. When this system was solvated, only the tetrahedral iron bond to one of the side chain oxygens of the N–terminal glutamate continued to last for over 99% of the simulation length. The octahedral iron bond to the other side chain oxygen of the N–terminal glutamate exhibited bonding lasting for 48% of the simulation length. No other bonding lasting over 2% of the simulation length was observed. When the residence time data for EVGELR on the {111} surface was examined, again, no tetrahedral iron bonding was exhibited. Octahedral iron bonding was exhibited to N–terminal peptide bond oxygen, one of the side chain oxygen of the N–terminal glutamate, and one of the side chain oxygen of the C–terminal glutamate that lasted over 99% of the

simulation length, within the *in vacu* system. Bonding lasting between 40% ad 69% of the simulation length was also seen to the other side chain oxygens of both glutamate, and to G-E peptide bond oxygen. No other bonding lasted over 20% of the simulation length. Upon solvation, the bond to one of the side chain oxygen of the N–terminal glutamate continued to last over 99% of the simulation length, however the bond to one of the side chain oxygens of the C–terminal glutamate only lasted for 10% of the simulation length with the presence of water. Bonding was seen to the other side chain oxygen of the N–terminal glutamate lasting for 70% of the simulation length, however, no other bonding presented lasting over 34% of the simulation length.

On comparison of the residence time data for EVGELR and EVELR it can be seen that, for the {100} surface *in vacu* system, EVELR had more bonding lasting over 99% of the simulation length (three bonds for EVELR and two bonds for EVGELR), which suggests that, the sequence mutation reduced the iron binding potential of the sequence, as hypothesised. However, when these systems were solvated, EVGELR exhibited the same amount of bonding lasting over 99% of the full length of the simulation as EVELR (one bond for each to the side chain oxygen of the N–terminal glutamate), suggesting no sequence preference. When the {111} surface *in vacu* system was examined it was shown that, EVGELR exhibited more bonding lasting over 99% of the full length of the simulation (three bonds for EVGELR and two bonds for EVELR), proposing that, the mutation improved the sequence iron binding potential. Upon solvation, this trend remained the same, with EVGELR displaying one bond that lasted over 99% of the simulation length, whilst, EVELR exhibited no bonding that lasted for this length of time. From the solvated systems data it was shown that, for the {100} surface, mutation of the original sequence had no effect on the iron binding potential, hence, attachment was not sequence dependent. EVGELR was the preferred sequence of attachment for the {111} surface, proposing that, as with EVALR, the mutation improved iron binding affinity.

All related energy data from the interfacial energy calculations of the 188-192 pentapeptides can be found in Table 6.29.

The solvated 188-192 pentapeptides interfacial energy data showed that, the {100} surface produces lower interfacial energy values (between 6.84 eV and 7.50 eV) than the {111} surface, with the exception of EVGELR (between 8.28 eV and 13.65 eV (excluding EVGELR at 5.72 eV)). For the {100} surface, both mutations exhibited larger interfacial energy values than EVELR, with EVALR producing the largest, suggesting that mutation reduced the iron binding potential of the sequence, as hypothesised, but the extent of this effect

{100} surface attachment				
	E_{int} of s-p (eV)	E_{diff} of s-p (eV)	E_{int} of s-p-w (eV)	E_{diff} of s-p-w (eV)
EVELR	-26.40	6.69	6.84	0.66
EVALR	-19.71	0.00	7.50	0.00
EVGELR	-26.60	6.89	7.34	0.16
{111} surface attachment				
	E_{int} of s-p (eV)	E_{diff} of s-p (eV)	E_{int} of s-p-w (eV)	E_{diff} of s-p-w (eV)
EVELR	-23.91	5.52	13.65	0.00
EVALR	-18.39	0.00	8.28	5.37
EVGELR	-25.56	7.17	5.72	7.93

Table 6.29: Interfacial energies for all pentapeptides on the {100} and {111} surface, using the Yang, Stipp and Harding method[164]. E_{int} of s-p is the interfacial energy of the slab–peptide system. E_{diff} of s-p is the difference in interfacial energy from the highest interfacial energy of the slab–peptide system. E_{int} of s-p-w is the interfacial energy of the slab–peptide–water system. E_{diff} of s-p-w is the difference in interfacial energy from the highest interfacial energy of the slab–peptide–water system.

was dependent on the type of mutation. For the {111} surface, both mutations produced smaller interfacial energy values than EVELR, with EVGELR producing the smallest value, suggesting that the sequence mutation improved the iron binding potential, and again, the extent of this influence was dependent on mutation type.

This data also indicated that, again, the addition of water was damaging to the systems interfacial energy, as the energies exhibited were much lower for the *in vacu* system, for example, on the {111} surface, EVELR presented with a significant increase in interfacial energy values upon solvation, from -23.91 eV to 13.65 eV. Preferential attachment of water to the magnetite surface could explain the increase in interfacial energy with the addition of water. Solvation of the system also affected the pentapeptides interfacial behaviour, as the energy trends for the *in vacu* systems are very different to the related solvated systems, with the exception of EVALR on the {100} surface, and EVGELR on the {111} surface, both of which exhibited the same trend whether *in vacu* or solvated.

As with the other pentapeptides, there are concerns with this direct method of calculating the adsorption energy, which are discussed for the 177-183 pentapeptides along with possible reasons for these. An alternative indirect method is explored in the following section which utilises a constrained molecular dynamics version of the systems, focusing on the Potential of Mean Force and free energy evaluation.

6.6 Constrained system

Free energy profiles were created for the aforementioned systems as shown in Figure 6.19.

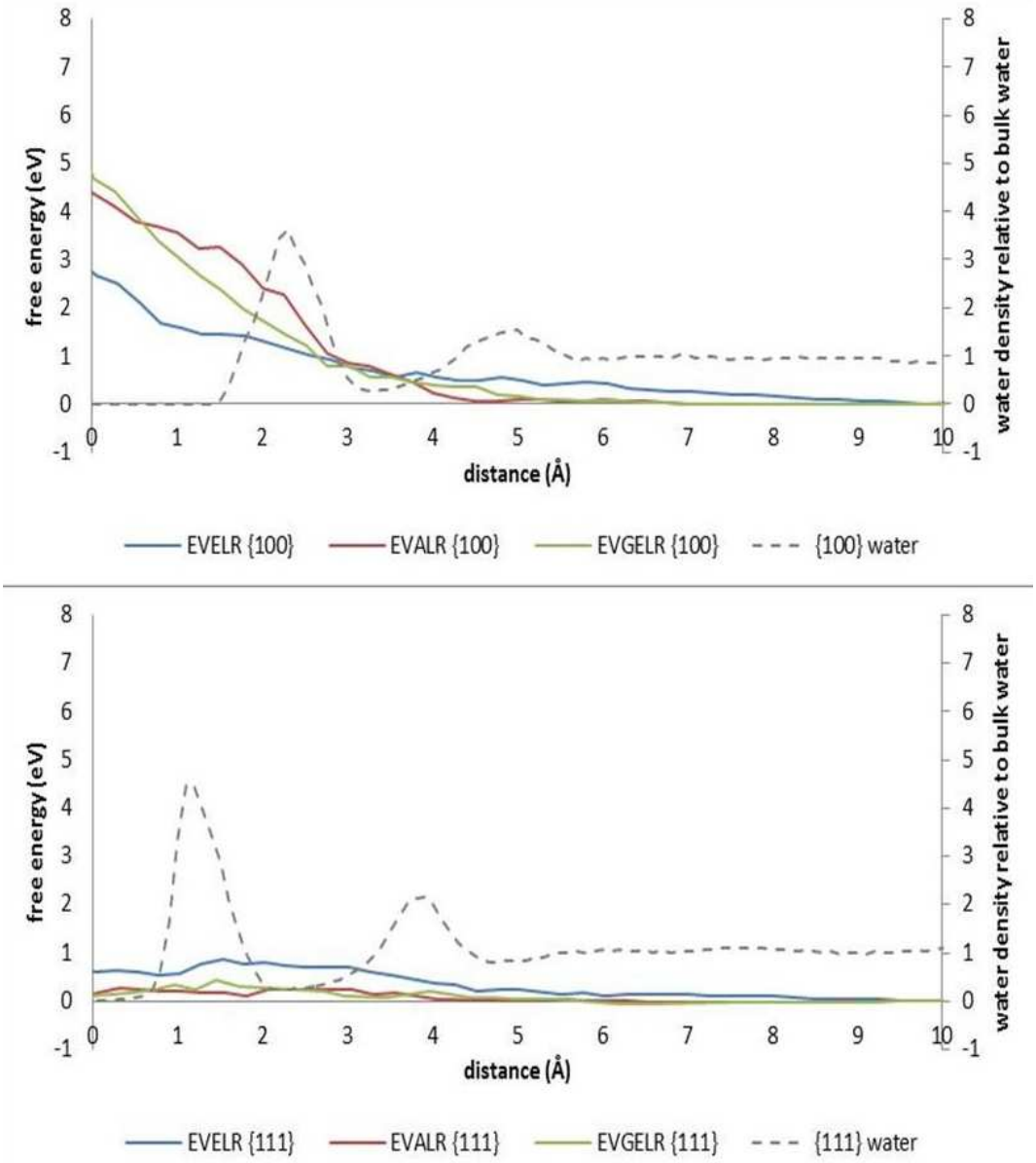

Figure 6.19: A comparison of the free energy profiles of the 188-192 pentapeptides. Distance refers to distance between the peptide CoM and the magnetite surface. The dashed line represents the water density profile for the system.

The free energy profiles for the different 188-192 pentapeptides (Figure 6.19) showed that, on the {100} surface, both mutations of EVELR reduced the iron binding activity between 0 Å and 3.5 Å , however, between 3.5 Å and 9 Å the mutation increase the iron binding activity. Between 0.75 Å and 3.5 Å EVALR produced the highest mutant free energy, however, from this distance until the sequence CoM is at the surface, EVGELR produced

the higher free energy value. For EVELR, between 10 Å and 9 Å was optimal for attachment, after which point, a gradual increase in free energy was detected, which appears to be independent of any water barrier influence. Thus, it is put forward that this successive increase in free energy is attributable to a requirement for more energy within the system to enable sequence attachment to the magnetite crystal as it moves closer towards the {100} surface. For EVALR and EVGELR, between 10 Å and 4.5 Å and between 10 Å and 5 Å , respectively, was optimal for attachment, after which point, an increase in free energy was detected, corresponding with the presence of the smaller water absorption layer. It is proposed that the successive increase in free energy was attributable to a requirement for more energy within the system for the sequences to pass through the water barriers in order to enable sequence attachment to the {100} surface.

Whereas, with the {111} surface energy profile, there was little difference between the free energy profiles of the three sequences, with EVELR showing a marginally larger change in free energy than EVALR and EVGELR. It can also be seen that, EVELR mutation appears to slightly increase the iron binding activity of the sequence and minimal difference in the free energy profiles of the two types of mutation was displayed. For EVALR and EVGELR, between 10 Å and 4 Å and between 10 Å and 4.5 Å , respectively, was optimal for attachment, after which point, a slight increase in free energy was detected, corresponding with the presence of the smaller water absorption layer. It is proposed that the successive increase in free energy was attributable to a requirement for more energy within the system for the sequences to pass through the water barriers in order to enable sequence attachment to the {100} surface. For EVELR, between 10 Å and 8.5 Å was optimal for attachment, after which point, a slight increase in free energy was detected, which appears to be independent of any water barrier influence. Thus, the successive increase in free energy was proposed to be attributable to the need for more system energy, to enable sequence attachment. The small changes in free energy seen for the sequence profiles suggest that the water absorption layers have minimal influence on the free energy profiles.

Upon examination of the energy minima for the 188-192 pentapeptides (Figure 6.20) it can be observed that, on the {100} surface, only EVALR and EVGELR exhibited a negative value for their energy minima, with EVALR producing the lowest value at -0.0178 eV. The energy minima for these residues fall between 7.3 Å and 7.5 Å, suggesting that attachment of the residue centre of mass at this distance would be preferable. On the {111} surface, again, only EVALR and EVGELR exhibited a negative value for their energy minima,

with EVGELR producing the lowest value at -0.0442 eV. The energy minima for these residues fall between 6.5 Å and 7.1 Å, suggesting that attachment of the residue centre of mass at this distance would be preferable.

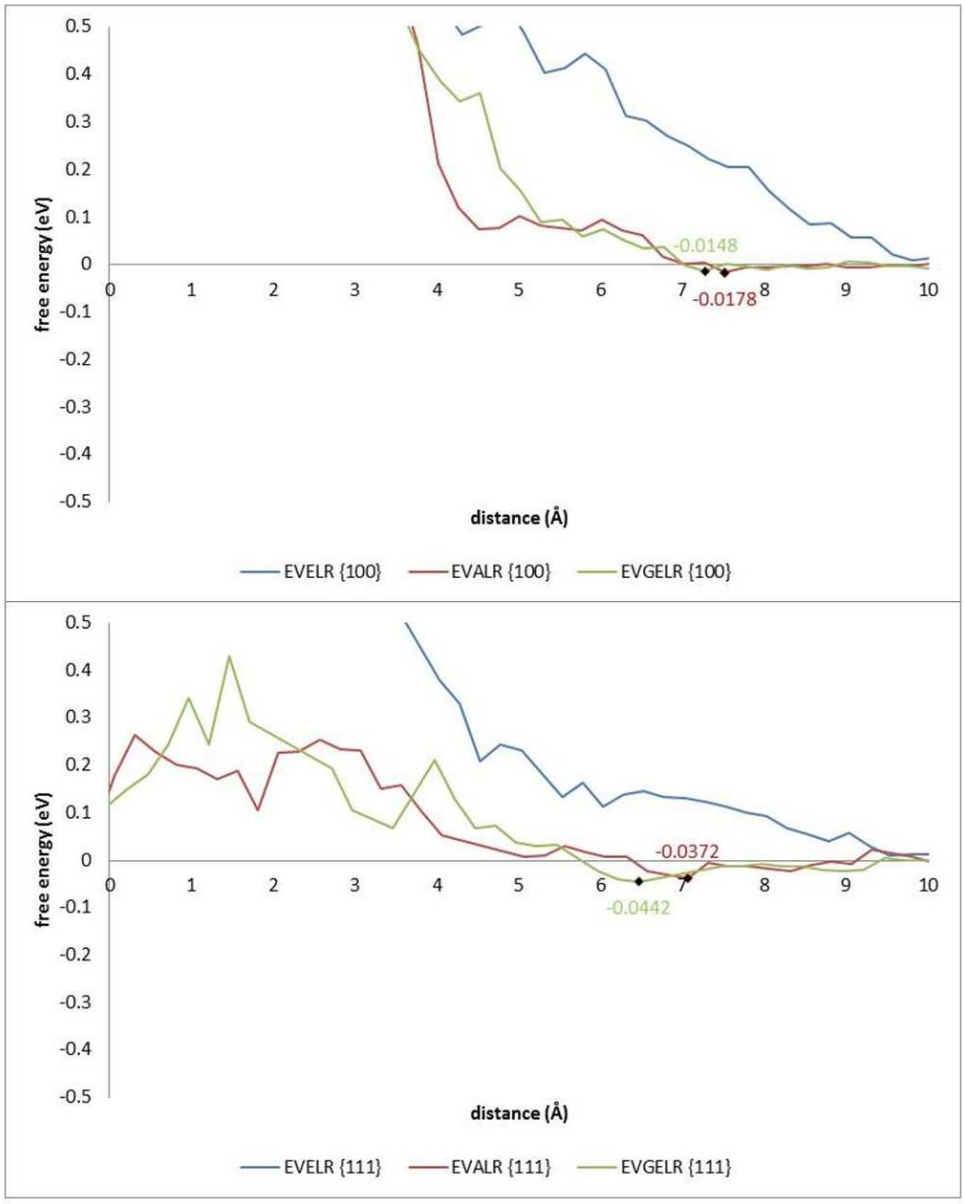

Figure 6.20: A close-up comparison of the free energy profiles of the 188-192 pentapeptides between -0.5 eV and 0.5 eV, exhibiting the free energy minima.

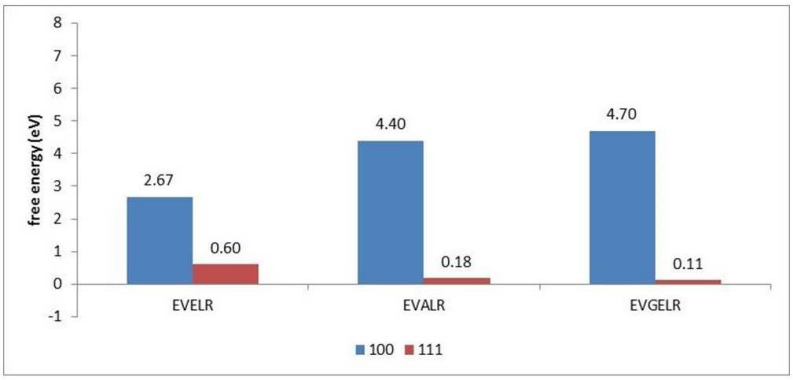

Figure 6.21: A comparison of the free energy values for the 188-192 pentapeptide CoM at the magnetite surface for both the {100} and {111} surfaces.

The free energy values for the 188-192 pentapeptides centre of mass at the magnetite surfaces are shown in Figure 6.21. On the {100} surface, the EVELR sequence exhibits a free energy value of 2.67 eV, whereas, on the {111} surface its value was much lower, 0.60 eV. Likewise, the EVALR and EVGELR sequences demonstrated much higher free energy values (4.40 eV and 4.70 eV, respectively) on the {100} surface, than exhibited on the {111} surface (0.18 eV and 0.11 eV, respectively). With both sets of mutation energy values being similar, it suggests that similar effects on the systems are seen despite very different sequence mutations being present. This data showed that, the {111} surface presented with significantly lower free energy values than the {100} surface, representing superior iron binding, proposing that this was the preferential surface of attachment. The free energy profiles also showed that, the sequence with the lowest free energy when the CoM is at the surface was EVELR for the {100} surface and EVGELR for the {111} surface.

6.6.1 Residues 185-189 (SDEEV)

Unconstrained System

R values for pentapeptides in the region of 185-189 can be found in Tables 6.30 and 6.31.

Pentapeptide	Fe type	Surface	System	r (Å)
SDEEV	FET	100	*in vacu*	1.98
			solvated	1.98
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	1.98
			solvated	2.08
		111	*in vacu*	2.03
			solvated	2.08
SAEEV	FET	100	*in vacu*	1.93
			solvated	1.93
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	1.98
			solvated	2.03
		111	*in vacu*	1.98
			solvated	2.03
SDAEV	FET	100	*in vacu*	2.03
			solvated	2.03
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	1.98
			solvated	2.08
		111	*in vacu*	1.98
			solvated	2.03

Table 6.30: R (Å) values from RDF data for pentapeptides in the region of 185-189.

Figure 6.22 exhibits the RDF profiles for SDEEV. The RDF profiles of the remaining 185-189 pentapeptides can be found in the Appendix. The RDF data for the 185-189 pentapeptides shows that, for all pentapeptides, both octahedral and tetrahedral iron bonding

Pentapeptide	Fe type	Surface	System	r (Å)
SDEAV	FET	100	*in vacu*	1.98
			solvated	2.03
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	1.98
			solvated	2.13
		111	*in vacu*	2.03
			solvated	2.03
SDAAV	FET	100	*in vacu*	2.08
			solvated	2.18
		111	*in vacu*	-
			solvated	-
	FEO	100	*in vacu*	2.03
			solvated	2.03
		111	*in vacu*	1.98
			solvated	2.03

Table 6.31: Continued. R (Å) values from RDF data for pentapeptides in the region of 185-189.

was displayed for the {100} surface. Whereas, with the {111} surface only octahedral iron bonding was exhibited. As the iron bonding trend was the same for all 185-189 pentapeptides, the data suggests that mutation of the sequence had no effect on the iron bonding type that resulted.

The RDF data for the 185-189 pentapeptide also showed that, the most intense first peak was exhibited for the *in vacu* systems rather than the corresponding solvated system, suggesting that, with the addition of water, it was more difficult for the pentapeptide oxygens to have an iron ion as a next nearest neighbour. For the solvated systems, the strongest $g(r)$ peak was shown for octahedral iron bonding in the {111} surface SDEAV system, suggesting that, the oxygen present was more freely accessible to the next nearest neighbouring irons.

The $FE - O_{pentapeptide}$ bond distance data for the 185-189 pentapeptides is represented in Tables 6.32, 6.33 and 6.34.

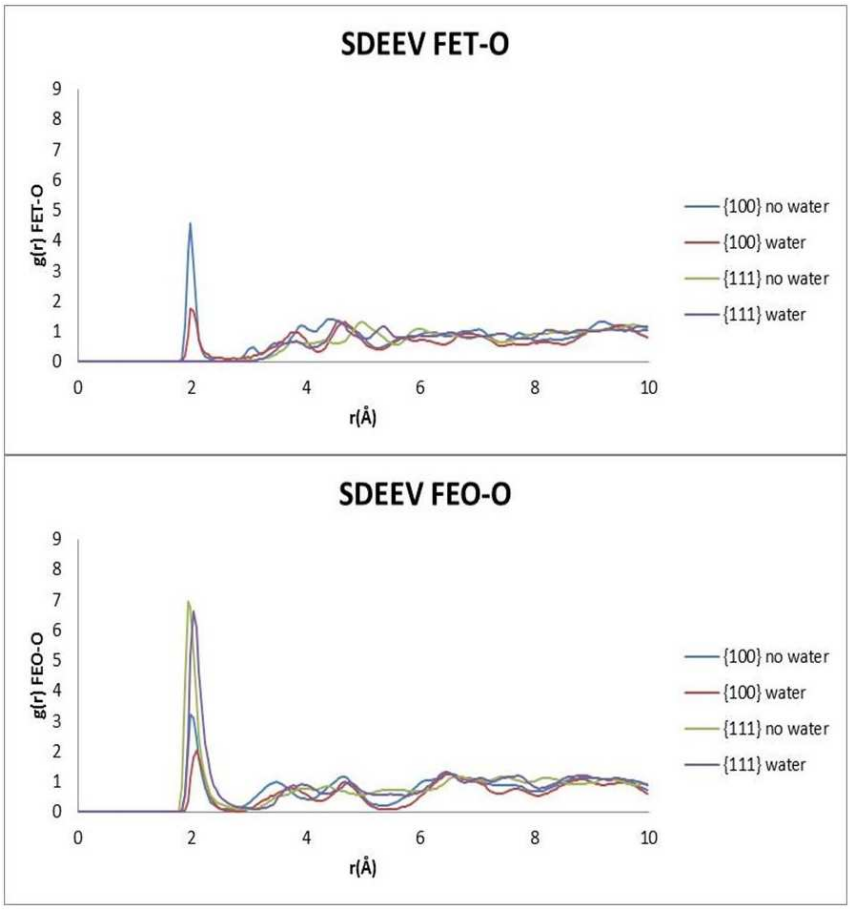

Figure 6.22: RDF plots for SDEEV. FET is tetrahedral iron, FEO is octahedral iron. Blue is {100} *in vacu*. Red is {100} solvated. Green is {111} *in vacu*. Purple is {111} solvated.

	{100} surface attachment			{111} surface attachment		
Pentapeptide	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
SDEEV	S-D	1.97	NB	NBO	1.96	1.98
	D O	1.95	2.06	S-D	2.21	2.11
		2.07	2.18	D O	2.17	2.13
	E O	2.03	2.07		2.03	2.04
		1.95	2.09		2.00	NB
	E-E	2.13	2.26		2.19	NB
	E O	2.04	2.13	D-E	2.05	2.51
		2.00	NB	E O	1.92	2.17
	E-V	2.19	NB		-	2.14
					2.26	NB
					2.02	NB
				E O	2.02	2.24
					-	2.16
					1.97	2.05

Table 6.32: 185-189 pentapeptide $FE - O_{pentapeptide}$ bond distance data. BL is bond length, E O is glutamate side chain oxygen, D O is aspartate side chain oxygen, x-x PBO is inter–residue peptide bond oxygen and NBO is N–terminal peptide bond oxygen. NB is no bonding.

	{100} surface attachment			{111} surface attachment		
Pentapeptide	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
SAEEV	E O	2.08	NB	NBO	1.97	NB
		1.93	1.90	S-A	1.99	NB
	E O	2.02	1.99	E O	1.93	2.12
		1.90	1.92		-	2.05
					1.95	1.97
				E O	1.93	2.10
					2.26	NB
					2.07	2.06
					1.98	2.16
				E-V	1.98	NB
SDAEV	D O	2.02	NB	NBO	2.03	1.96
		2.15	2.04	S-D	1.96	2.49
	E O	1.96	NB	D O	2.17	2.06
		1.93	NB		1.96	2.00
				A-E	2.15	2.24
				E O	2.02	2.14
					-	2.05
					2.08	2.06
					2.13	2.08

Table 6.33: Continued. 185-189 pentapeptide $FE - O_{pentapeptide}$ bond distance data. BL is bond length, E O is glutamate side chain oxygen, D O is aspartate side chain oxygen, x-x PBO is inter–residue peptide bond oxygen and NBO is N–terminal peptide bond oxygen. NB is no bonding.

	{100} surface attachment			{111} surface attachment		
Pentapeptide	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)	Oxygen type	BL *In Vacu* (Å)	BL Solvated (Å)
SDEAV	S-D	-	2.13	NBO	2.04	1.99
	D O	2.08	1.98	S-D	2.22	2.14
		1.98	NB	D O	2.09	2.08
	E O	2.14	NB		2.21	NB
		1.96	1.93		1.86	1.94
				D-E	2.12	2.03
				E O	1.92	2.15
					-	2.02
					2.03	1.96
					2.15	2.11
				A-V	2.06	NB
SDAAV	S-D	2.05	2.19	NBO	1.94	2.28
	D O	1.93	NB	S-D	1.98	2.09
		2.06	2.15	D O	1.94	2.39
	A-A	2.17	2.42		1.90	2.21
					-	2.00
				A-A	2.00	2.39

Table 6.34: Continued. 185-189 pentapeptide $FE - O_{pentapeptide}$ bond distance data. BL is bond length, E O is glutamate side chain oxygen, D O is aspartate side chain oxygen, x-x PBO is inter–residue peptide bond oxygen and NBO is N–terminal peptide bond oxygen. NB is no bonding.

The 185-189 pentapeptides bond distance data (Figure 6.23) showed that, for the SDEEV *in vacu* system, bonding differed dependant on surface of attachment. For the {100} surface, bonding was present through the two side chain oxygens of both glutamate, and both side chain oxygens of aspartate. This bonding was also observed for the {111} surface, however, more acidic residue side chain oxygen bonding was present for this surface. Both surfaces also exhibited S-D peptide bond oxygen bonding. The {100} surface displayed E-E and E-V peptide bond oxygen bonding, whilst, the {111} surface showed N–terminal and D-E peptide bond oxygen bonding. This data suggests that, the {111} surface was slightly more favourable for iron binding. Upon solvation, the amount of bonding present was reduced, suggesting that, the presence of water was disadvantageous to the iron bonding potential of the sequence, however, the {111} surface remained the preferred surface of attachment, as the amount of bonding was greater than that presented for the {100} systems. The average bond lengths for a magnetite iron to a pentapeptide oxygen in the SDEEV sequence was; 2.04 Å and 2.13 Å on the {100} surface for *in vacu* and solvated systems respectively, and 2.07 Å and 2.15 Å on the {111} surface for *in vacu* and solvated systems respectively.

With the mutation of SDEEV to SAEEV, less bonding was displayed for the mutant sequence, particularly for the {100} surface. This phenomenon was proposed to be due to the removal of the potential oxygen binding sites of aspartate. Bonding to the two side chain oxygens of both glutamate was exhibited, although, more of this bonding type was present on the {111} surface than the {100} surface (four bonds to the C–terminal glutamate for the {111} surface and two bonds for the {100} surface). The {100} surface exhibited no other bonding, whereas, the {111} surface also presented with N–terminal, S-A and E-V peptide bond oxygen bonding, suggesting that different surface types affect the iron binding potential of the sequence, with {111} surface attachment again favoured. The addition of water to the systems had an adverse effect on bonding, particularly on the {111} surface, as all peptide bong oxygen bonding was removed for this surface. The average bond lengths for an iron to a pentapeptide oxygen in SAEEV was; 1.98 Å and 1.94 Å on the {100} surface for *in vacu* and solvated systems respectively, and 2.01 Å and 2.08 Å on the {111} surface for *in vacu* and solvated systems respectively.

For the SDAEV sequence {100} *in vacu* system, the same amount of bonding was present as in the equivalent SAEEV system, thus, less bonding was present than for the original sequence. In this system, only acidic side chain oxygen bonding was present, (to both side chain oxygens of aspartate and to both side chain oxygens of glutamate). In the {111}

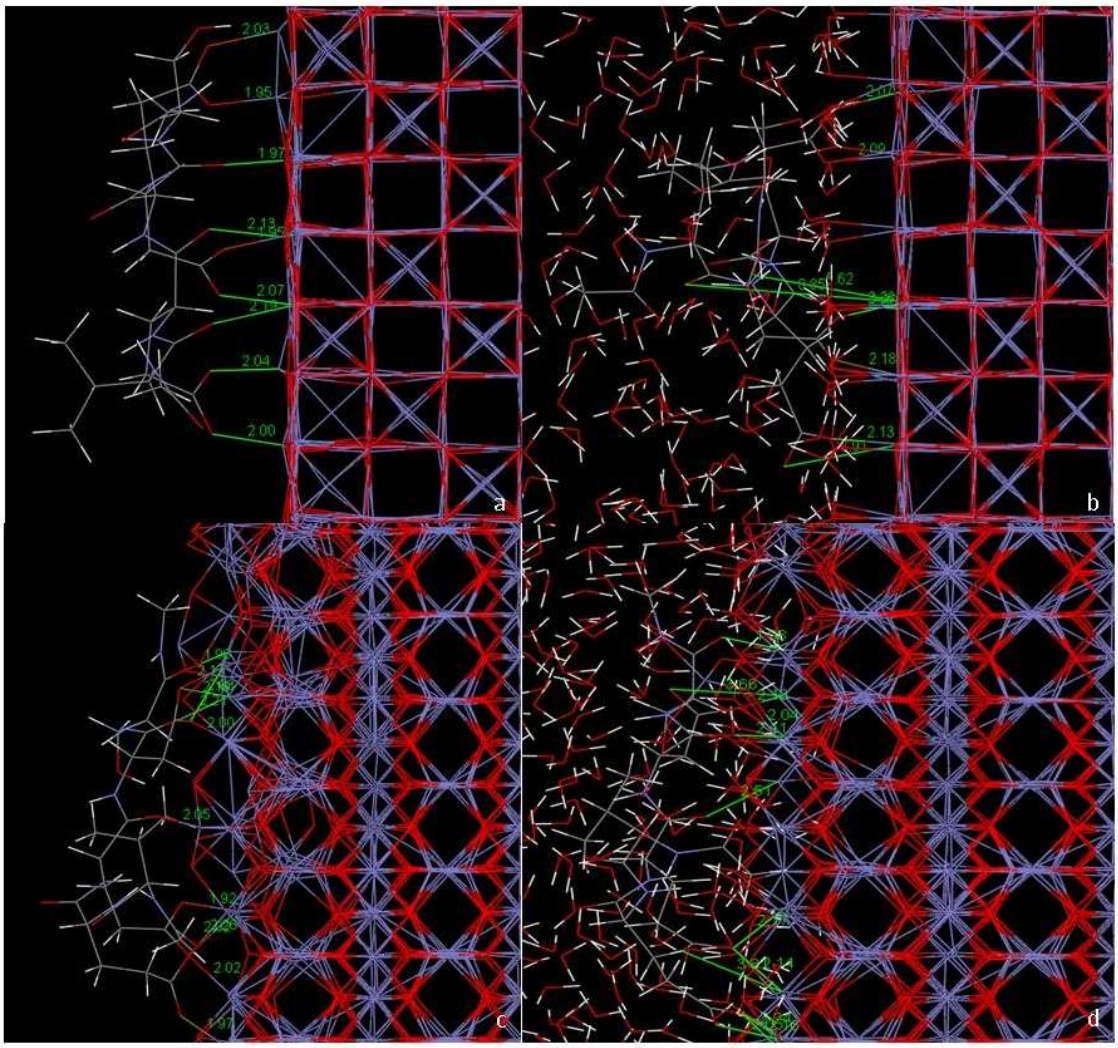

Figure 6.23: Bond distance images for the *in vacu* and solvated systems of SDEEV. a){100} *in vacu*, b){100} solvated, c){111} *in vacu* and d){111} solvated.

surface, SDAEV exhibited one more bond than SAEEV, presenting with N–terminal, S-D, A-E peptide bond oxygen bonding, and bonding to both side chain oxygens of aspartate (one bond per each oxygen) and bonding to both side chain oxygens of glutamate (one bond to one oxygen and two bonds to the other oxygen), suggesting that potential iron binding was surface dependent. Less bonding occurred when the {100} surface system was solvated, as only one bond was displayed, suggesting more favourable bonding between the iron ions and the oxygen of water. For the {111} surface, more bonding was seen upon solvation, with both side chain oxygens of glutamate presenting with two bonds each. This data postulates that, the addition of water improved the sequence iron binding potential. The average bond lengths for a magnetite iron to a pentapeptide oxygen in the SDAEV sequence was; 2.02 Å and 2.04 Å on the {100} surface for *in vacu* and solvated systems respectively, and 2.06 Å and 2.12 Å on the {111} surface for *in vacu* and solvated systems

respectively.

For the SDEAV sequence {100} *in vacu* system, the same amount and type of bonding were seen, as in the equivalent SDAEV system (bonding to both side chain oxygens of aspartate and to both side chain oxygens of glutamate). However, the {111} *in vacu* system, exhibited more bonding than the other mutated sequences, though, there was still less bonding than was seen for SDEEV. SDAEV exhibited bonding to N–terminal, S-D, D-E and A-V peptide bond oxygen bonding, bonding to both side chain oxygens of glutamate, and bonding to both side chain oxygens of aspartate. These results suggested that the {111} surface had more of a preference for peptide bond oxygen bonding than the {100} surface. As previously seen, less bonding was displayed upon solvation, suggesting more favourable bonding between the magnetite iron and the oxygen of water. For the {100} surface, a bond to one of the side chain oxygens of aspartate and a bond to one of the side chain oxygens of glutamate were no longer present, however, a new bond existed to S-D peptide bond oxygen. For the {111} surface, the A-V peptide bond oxygen bond and one of the side chain oxygen aspartate bonds were no longer present, however, a new bond exists to the other glutamate side chain oxygen, proposing that, along with being detrimental to the amount of bonding, solvation also changed the type of bonding present. The average bond lengths for a magnetite iron to a pentapeptide oxygen in the SDEAV sequence was; 2.04 Å and 2.01 Å on the {100} surface for *in vacu* and solvated systems respectively, and 2.07 Å and 2.05 Å on the {111} surface for *in vacu* and solvated systems respectively.

For the double substitution mutant, SDAAV, less bonding was present than for any other 185-189 pentapeptide sequence. This phenomenon was postulated to be due to the removal of the two acidic residues from the original sequence, thus, removal of more potential oxygen binding sites than the other mutant sequences. Both surfaces exhibited bonding to both side chain oxygens of aspartate, and to S-D and A-A peptide bond oxygen. Whilst, the {111} surface, also showed N–terminal peptide bond oxygen bonding. This data suggests that, the {111} surface was the preferred surface for attachment. Upon solvation, less bonding occurred for the {100} surface system, with the loss of an aspartate side chain oxygen bond. For the {111} surface, a new aspartate side chain oxygen bond was formed with the addition of water, suggesting that the presence of water improved the iron binding potential of this sequence. The average bond lengths for a magnetite iron to a pentapeptide oxygen in the SDAAV sequence was; 2.05 Å and 2.25 Å on the {100} surface for *in vacu* and solvated systems respectively, and 1.95 Å and 2.23 Å on the {111} surface for *in*

vacu and solvated systems respectively.

Tables 6.35, 6.36, 6.37, 6.38, 6.39, 6.40, 6.41, 6.42 and 6.43 represent the residence times and average coordination numbers for bonding between all oxygen present in the 185-189 pentapeptides and tetrahedral and/or octahedral iron on both the {100} and {111} magnetite surfaces.

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
SDEEV {100} surface	E-V PBO	0.00	0.00	224.46	0.74	0.00	0.00	4.71	0.03
	E O	0.00	0.00	581.95	0.99	0.00	0.00	5.63	0.01
	E O	169.03	0.14	857.35	0.87	0.00	0.00	495.11	0.97
	E-E PBO	0.00	0.00	675.62	0.98	0.00	0.00	90.46	0.87
	E O	0.00	0.00	4967.57	1.00	0.00	0.00	999.00	1.00
	E O	395.01	0.47	547.59	0.53	0.00	0.00	973.90	0.98
	D-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D O	4998.00	1.00	0.00	0.00	999.00	1.00	0.00	0.00
	D O	4998.00	1.00	0.00	0.00	42.78	0.50	0.00	0.00
	S-D PBO	0.00	0.00	494.16	0.92	0.00	0.00	398.61	0.71
	S O	0.00	0.00	1.00	0.00	7.22	0.11	0.00	0.00
	NBO	41.13	0.12	1.00	0.00	0.00	0.00	0.00	0.00

Table 6.35: Residence times (RT) and average coordination numbers (av. CN) of SDEEV. (S O = side chain oxygen of serine, D O = side chain oxygen of aspartate, E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
SDEEV {111} surface	E-V PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	1499.32	1.19	0.00	0.00	507.50	0.95
	E O	0.00	0.00	1306.17	1.24	0.00	0.00	628.78	1.77
	E-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	1929.87	1.10	0.00	0.00	883.25	1.66
	E O	0.00	0.00	1872.59	1.24	0.00	0.00	105.00	0.94
	D-E PBO	0.00	0.00	1543.15	0.44	0.00	0.00	41.43	0.68
	D O	0.00	0.00	1990.82	0.79	0.00	0.00	41.46	0.25
	D O	0.00	0.00	4824.94	1.39	0.00	0.00	631.75	1.98
	S-D PBO	0.00	0.00	924.52	0.97	0.00	0.00	781.02	1.69
	S O	0.00	0.00	6.28	0.33	0.00	0.00	0.00	0.00
	NBO	0.00	0.00	4995.00	1.00	0.00	0.00	999.00	1.00

Table 6.36: Continued. Residence times (RT) and average coordination numbers (av. CN) of SDEEV. (S O = side chain oxygen of serine, D O = side chain oxygen of aspartate, E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
SAEEV {100} surface	E-V PBO	19.50	0.55	0.00	0.00	18.66	0.32	0.00	0.00
	E O	4999.00	1.00	0.00	0.00	999.00	1.00	0.00	0.00
	E O	1.02	0.01	4999.00	1.00	0.00	0.00	999.00	1.00
	E-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	3479.70	0.70	735.04	0.30	999.00	1.00	0.00	0.00
	E O	1548.71	1.00	2.50	0.00	9.00	0.02	0.00	0.00
	A-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	S-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	S O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 6.37: Residence times (RT) and average coordination numbers (av. CN) of SAEEV. (S O = side chain oxygen of serine, E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
SAEEV {111} surface	E-V PBO	0.00	0.00	4996.00	1.00	0.00	0.00	230.03	0.45
	E O	0.00	0.00	3292.41	1.76	0.00	0.00	279.13	1.63
	E O	0.00	0.00	3302.07	1.81	0.00	0.00	293.26	1.76
	E-E PBO	0.00	0.00	148.60	0.26	0.00	0.00	0.00	0.00
	E O	0.00	0.00	1157.13	1.26	0.00	0.00	239.88	1.81
	E O	0.00	0.00	4895.34	1.01	0.00	0.00	999.00	1.00
	A-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	S-A PBO	0.00	0.00	4991.00	1.00	0.00	0.00	14.61	0.08
	S O	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
	NBO	0.00	0.00	4363.03	1.00	0.00	0.00	33.95	0.06

Table 6.38: Continued. Residence times (RT) and average coordination numbers (av. CN) of SAEEV. (S O = side chain oxygen of serine, E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
SDAEV {100} surface	E-V PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	4269.00	0.85	0.00	0.00	3.04	0.02
	E O	0.00	0.00	4272.00	0.85	0.00	0.00	37.47	0.14
	A-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D-A PBO	234.60	0.24	1.00	0.00	0.00	0.00	0.00	0.00
	D O	1199.75	0.75	0.00	0.00	377.62	1.00	0.00	0.00
	D O	4995.00	1.00	0.00	0.00	4.63	0.02	0.00	0.00
	S-D PBO	0.00	0.00	256.53	0.45	0.00	0.00	0.00	0.00
	S O	1.00	0.00	2.80	0.12	0.00	0.00	2.57	0.01
	NBO	0.00	0.00	1.91	0.00	0.00	0.00	0.00	0.00

Table 6.39: Residence times (RT) and average coordination numbers (av. CN) of SDAEV. (S O = side chain oxygen of serine, D O = side chain oxygen of aspartate, E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
SDAEV {111} surface	E-V PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	4956.77	1.99	0.00	0.00	665.20	1.97
	E O	0.00	0.00	4424.68	1.38	0.00	0.00	993.59	1.96
	A-E PBO	0.00	0.00	4997.53	0.97	0.00	0.00	18.62	0.67
	D-A PBO	0.00	0.00	9.38	0.00	0.00	0.00	0.00	0.00
	D O	0.00	0.00	3647.00	0.73	0.00	0.00	999.00	1.00
	D O	0.00	0.00	1829.00	1.19	0.00	0.00	202.77	1.04
	S-D PBO	0.00	0.00	4997.00	1.00	0.00	0.00	100.64	0.94
	S O	0.00	0.00	7.70	0.00	0.00	0.00	1.00	0.00
	NBO	0.00	0.00	4996.16	1.00	0.00	0.00	999.00	1.00

Table 6.40: Continued. Residence times (RT) and average coordination numbers (av. CN) of SDAEV. (S O = side chain oxygen of serine, D O = side chain oxygen of aspartate, E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
SDEAV {100} surface	A-V PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E-A PBO	0.00	0.00	2.00	0.00	0.00	0.00	0.00	0.00
	E O	4998.00	1.00	0.00	0.00	999.00	1.00	0.00	0.00
	E O	0.00	0.00	4998.00	1.00	0.00	0.00	62.58	0.44
	D-E PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D O	4997.00	1.00	0.00	0.00	806.42	0.99	0.00	0.00
	D O	4997.00	1.00	0.00	0.00	17.30	0.06	0.00	0.00
	S-D PBO	0.00	0.00	48.28	0.63	0.00	0.00	267.01	0.95
	S O	11.27	0.33	69.72	0.39	24.08	0.77	0.00	0.00
	NBO	83.24	0.08	1.33	0.00	0.00	0.00	0.00	0.00

Table 6.41: Residence times (RT) and average coordination numbers (av. CN) of SDEAV. (S O = side chain oxygen of serine, D O = side chain oxygen of aspartate, E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
SDEAV {111} surface	A-V PBO	0.00	0.00	4989.40	1.00	0.00	0.00	33.32	0.08
	E-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	E O	0.00	0.00	4994.47	1.99	0.00	0.00	971.96	1.93
	E O	0.00	0.00	4981.72	1.00	0.00	0.00	793.79	1.84
	D-E PBO	0.00	0.00	4999.00	1.00	0.00	0.00	999.00	1.00
	D O	0.00	0.00	4731.90	1.18	0.00	0.00	987.34	1.04
	D O	0.00	0.00	4214.23	1.99	0.00	0.00	636.54	1.90
	S-D PBO	0.00	0.00	49.57	0.71	0.00	0.00	377.27	1.00
	S O	0.00	0.00	4.23	0.59	0.00	0.00	45.90	1.11
	NBO	0.00	0.00	4995.40	1.00	0.00	0.00	999.00	1.00

Table 6.42: Continued. Residence times (RT) and average coordination numbers (av. CN) of SDEAV. (S O = side chain oxygen of serine, D O = side chain oxygen of aspartate, E O = side chain oxygen of glutamate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

		in vacu system				solvated system			
		FET		FEO		FET		FEO	
		RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN	RT (ps)	Av. CN
SDAAV {100} surface	A-V PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	A-A PBO	0.00	0.00	656.48	0.95	0.00	0.00	5.62	0.35
	D-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D O	0.00	0.00	4995.00	1.00	0.00	0.00	999.00	1.00
	D O	0.00	0.00	3566.87	0.98	0.00	0.00	9.69	0.03
	S-D PBO	1557.50	0.99	0.00	0.00	246.40	0.97	0.00	0.00
	S O	1.83	0.00	17.38	0.23	0.00	0.00	25.08	0.57
	NBO	1.17	0.03	222.50	0.09	0.00	0.00	0.00	0.00
SDAAV {111} surface	A-V PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	A-A PBO	0.00	0.00	1731.89	1.34	0.00	0.00	57.02	1.65
	D-A PBO	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	D O	0.00	0.00	4991.00	1.00	0.00	0.00	999.00	0.99
	D O	0.00	0.00	3227.39	1.77	0.00	0.00	999.00	1.98
	S-D PBO	0.00	0.00	4990.00	1.00	0.00	0.00	999.00	0.98
	S O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	NBO	0.00	0.00	4992.00	1.00	0.00	0.00	999.00	1.00

Table 6.43: Residence times (RT) and average coordination numbers (av. CN) of SDAAV. (S O = side chain oxygen of serine, D O = side chain oxygen of aspartate, x-x PBO = inter–residue peptide bond oxygen and NBO = N–terminal peptide bond oxygen).

The 185-189 pentapeptides residence time data showed that, for the SDEEV {100}*in vacu* system, there were three bonds present that lasted over 99% of the full length of the simulation (tetrahedral iron bonding to both side chain oxygens of aspartate and octahedral iron bonding to one of the side chain oxygen of the N–terminal glutamate). No other bonding exhibited lasted over 18% of the simulation length. When this system was solvated, three bonds were present for over 97% of the full length of the simulation, however, the bonding type differed to that shown for the *in vacu* system, with a tetrahedral iron bond to one of the side chain oxygens of aspartate, and a bond to each of the side chain oxygens of the N–terminal glutamate through octahedral iron being present, suggesting that the addition of water makes glutamate attachment more favourable than aspartate attachment. No other bonding was present lasting over 50% of the simulation length. As the amount of bonding lasting over 97% of the simulation length did not alter upon solvation, it was proposed that, water had a minimal effect on iron binding potential.

The residence time data for SDEEV on the {111} surface showed that, no tetrahedral iron bonding was present for either system. For the *in vacu* system, bonding lasting over 96% of the full length of the simulation was seen between octahedral iron and N–terminal peptide bond oxygen, and one of the side chain oxygens of aspartate. No other bonding lasting over 40% of the simulation length was seen for this system. Upon solvation, only the bond between octahedral iron and N–terminal peptide bond oxygen continued to last over 99% of the full length of the simulation. This data suggests that the addition of water provides competitive binding to the magnetite. Bonding was also present, lasting between 51% and 88% of the simulation length, between octahedral iron and all of the side chains oxygens of the glutamates and the S-D peptide bond oxygen.

With a single alanine substitution of SDEEV to SAEEV, for the {100} *in vacu* system, bonding lasting over 99% of the full length of the simulation was exhibited between tetrahedral iron and one of the side chain oxygens of the C–terminal glutamate, and between octahedral iron and the other side chain oxygen of the C–terminal glutamate. Bonding lasting for 70% of the simulation length was also seen between tetrahedral iron and one of the side chain oxygens of the N–terminal glutamate. No other bonding present for this system lasted over 31% of the simulation length. Upon solvation, the bonding which lasted over 99% of the simulation length in the *in vacu* system, continued to last for this length of time, however, the bond between tetrahedral iron and one of the side chain oxygens of the N–terminal glutamate also lasted over 99% of the full length of the simulation with the addition of water, proposing that, the shorter equivalent *in vacu* residence time was due

to the formation of this bond towards the latter stages of this simulation. This systems residence time data suggests that, the addition of water improved the iron binding affinity of the sequence. No other bonding present lasted over 2% of the simulation length.

Again, no tetrahedral iron bonding was exhibited when the corresponding {111} surface data was examined. With the octahedral iron, bonding was exhibited to N–terminal, S-A and E-V peptide bond oxygen, and one of the side chain oxygens of the N–terminal glutamate that lasted over 87% of the full length of the simulation. Bonding lasting for 66% of the simulation length was also seen to both side chain oxygens of the C–terminal glutamate. No other bonding present lasted over 24% of the simulation length. With the addition of water, only the bond between octahedral iron and one of the side chain oxygens of the N–terminal glutamate lasted over 99% of the simulation length, suggesting that, waters presence in the system was unfavourable for the iron binding ability of the sequence. No other bonding was exhibited lasting over 30% of the simulation length.

On comparison of the residence time data of SAEEV and SDEEV, it was seen that, for the {100} surface *in vacu* system, SDEEV had more bonding lasting over 99% of the simulation length than SAEEV (three bonds for SDEEV and two bonds for SAEEV), thus, the SDEEV sequence was the preferred for attachment in this system. Upon solvation, both sequences exhibited three bonds that lasted over 97% of the full length of the simulation, suggesting no preferential sequence for this system. The {111} *in vacu* system data showed that, SAEEV exhibited more bonding than SDEEV (four bonds for SAEEV and two bonds for SDEEV), proposing that the SAEEV sequence was preferred for attachment. When these systems were solvated, both sequences exhibited one bond that lasted over 99% of the simulation length, proposing that there was no favoured sequence. Based on the natural solvated systems, the {100} surface was the preferred surface of attachment, however, there was no preferred sequence of attachment.

For the SDAEV {100} *in vacu* system, there were three bonds present lasting over 85% of the full length of the simulation(octahedral iron bonding to both side chain oxygens of glutamate and tetrahedral iron bonding to one of the side chain oxygens of the aspartate). No other bonding was exhibited for this system lasting over 24% of the simulation length. Upon addition of water, no bonding was present lasting over 38% of the simulation length, suggesting that the addition of water to the system is severely detrimental to the iron binding potential of this sequence.

When the corresponding {111} surface residence time data, there was no tetrahedral iron

bonding displayed. With the octahedral iron, N–terminal, S-D and A-E peptide bond oxygen bonding, and bonding to both side chain oxygens of glutamate was exhibited lasting over 88% of the simulation length. Bonding lasting for 73% of the simulation length was also revealed between octahedral iron and one of the side chain oxygens of aspartate. No other bonding was exhibited lasting over 37% of the simulation length. Upon solvation, the bonds between octahedral iron and N–terminal peptide bond oxygen and one of the side chain oxygens of glutamate continued to last over 99% of the full length of the simulation. Also, the bond to one of the side chain oxygens of aspartate lasted over 99% of the simulation length, suggesting that the lower value for the equivalent *in vacu* residence time is due to the formation of this bond toward the latter stages of the simulation. No other bonding present lasted over 67% of the full length of the simulation. This reduction in residence times suggests that when this system was solvated, the water provides competitive binding to the magnetite.

On comparison of the residence time data for SDAEV and SDEEV, it was seen that, for the {100} *in vacu* system, both sequences exhibited three bonds that lasted over 85% of the full length of the simulation, suggesting no preferential sequence attachment for this system type. When water was added to these systems, SDEEV exhibited more bonding lasting over 97% of the simulation length (three bonds for SDEEV and no bonds for SDAEV), suggesting that SDEEV was the preferred attachment sequence. When the {111} *in vacu* system was examined it was shown that, SDAEV exhibited more bonding lasting over 88% of the simulation length than SDEEV (five bonds for SDAEV and two bonds for SDEEV), proposing that SDAEV was the preferred sequence of attachment, hence, the mutation improved the iron binding affinity of the sequence. Upon solvation, SDAEV presented with more bonding lasting over 99% of the simulation length than SDEEV (three bonds for SDAEV and one bond for SDEEV). Based on the natural solvated systems, each surface had a different sequence preference. For the {100} surface the preferred sequence of attachment was SDEEV, whereas, SDAEV was the preferred sequence of attachment for the {111} surface.

For the SDEAV {100} *in vacu* system it was shown that, bonding lasting over 99% of the full length of the simulation was exhibited between tetrahedral iron and both side chain oxygens of aspartate and one of the side chain oxygens of glutamate, and between octahedral iron and one of the side chain oxygens of glutamate. No other bonding lasting over 2% of the simulation length was displayed for this system. Upon solvation, only the bond between tetrahedral iron and one of the side chain oxygen of glutamate continued

to last over 99% of the simulation length. Bonding was also observed between tetrahedral iron and one of the side chain oxygens of aspartate lasting for 81% of the simulation length. This data suggests that, an adverse effect was seen for the sequence iron binding potential with the addition of water to the system.

When the residence time data for the SDAEV {111} system was examined, there was no tetrahedral iron bonding present. With the octahedral iron, bonding was exhibited to N–terminal, D-E and A-V peptide bond oxygen, both side chain oxygens of aspartate, and both side chain oxygens of glutamate that lasted over 84% of the full length of the simulation. No other bonding lasting over 1% of the simulation length was present for this system. Upon solvation, the bonds between octahedral iron and N–terminal and D-E peptide bond oxygen, one of the side chain oxygens of aspartate and one of the side chain oxygens of glutamate lasted over 97% the full length of the simulation. Bonds lasting between 64% and 79% of the simulation length were also displayed between octahedral iron and the other side chain oxygens of aspartate and glutamate. No other bonding lasted over 38% of the simulation length. This reduction in residence times suggests that, when water was present in the system, competitive binding to the magnetite was provided.

On comparison of the residence time data for SDEAV and SDEEV it can be seen that, for the {100} *in vacu* system, SDEAV had more bonding lasting over 99% of the simulation length (four bonds for SDEAV and three bonds for SDEEV). This occurrence suggests that, the mutated sequence is the preferred for attachment. When these systems were solvated, SDEEV exhibited more bonding lasting over 97% of the simulation length than SDEAV (three bonds for SDEEV and one bond for SDEAV). When the {111} *in vacu* system was examined it was shown that, SDEAV had more bonding lasting over 84% of the simulation length than SDEEV (seven bonds for SDEAV and two bonds for SDEEV), again suggesting that, SDEAV was the preferred sequence of attachment. Upon solvation, SDEAV again presents as the sequence of preferred attachment, as it presents with more bonding lasting over 97% of the simulation length than SDEEV (four bonds for SDEAV and one bond for SDEEV). Based on the solvated systems residence time data, a different sequence preference was seen for each surface, as the {100} surface showed a preference for SDEEV, whereas, SDAEV was the preferred sequence of attachment for the {111} surface.

The residence time data for the double substitution mutant, SDAAV, shows that, for the {100} *in vacu* system, there was no bonding present lasting over 32% of the simulation length through tetrahedral iron.

With the octahedral iron, bonding was exhibited to one of the side chain oxygens of aspartate that lasted over 99% of the full length of the simulation. Bonding lasting for 71% of the simulation length was also seen between octahedral iron and the other side chain oxygen of aspartate. No other octahedral iron bonding present lasted over 14% of the simulation length. Upon solvation, the bond between octahedral iron and one of the side chain oxygens of aspartate continued to last over 99% of the full length of the simulation. No other bonding was exhibited lasting over 25% of the simulation length, suggesting that competitive binding of water caused the retention times of water to reduce.

When the {111} surface simulations were examined, as with the other 185-189 pentapeptide sequences, no tetrahedral iron bonding was present. For the *in vacu* system, bonding lasting over 99% of the full length of the simulation was seen between octahedral iron and N–terminal and S-D peptide bond oxygen, and one of the side chain oxygens of aspartate. Bonding was also exhibited between octahedral iron and the other side chain oxygen of aspartate lasting for 65% of the simulation length. No other bonding lasting over 35% of the simulation length was displayed. When this system was solvated, the bonding between octahedral iron and N–terminal and S-D peptide bond oxygen, and one of the side chain oxygens of aspartate continued to last over 99% of the full length of the simulation. Additionally, the bond between octahedral iron and the other side chain oxygen of aspartate lasted over 99% of the simulation length. This data suggests that the addition of water improved the iron binding potential of the sequence.

On comparison of the residence time data for SDAAV and SDEEV it was seen that, for the {100} *in vacu* system, SDEEV had more bonding lasting over 99% of the simulation length than SDAAV (three bonds for SDEEV and one bond for SDAAV), suggesting that SDEEV was the pentapeptide of preferred attachment. Upon solvation, again SDEEV had more bonding lasting over 97% of the simulation length than SDAAV (three bonds for SDEEV and one bond for SDAAV). On examination of the {111} surface *in vacu* system residence time data it was shown that, SDAAV had more bonding lasting over 96% of the simulation length than SDEEV (three bonds for SDAAV and two bonds for SDEEV), suggesting that the mutant sequence was the preferred for attachment. Upon addition of water, again SDAAV had more bonding lasting over 99% of the simulation length than SDEEV (four bonds for SDAAV and one bond for SDEEV). From the natural solvated system data, each surface exhibited a different sequence preference, with SDEEV being favoured for attachment on the {100} surface the preferred sequence of attachment is SDEEV, whereas, for the {111} surface SDAEV was the preferred sequence of attachment.

All related energy data from the interfacial energy calculations of the 185-189 pentapeptides can be found in Table 6.44.

{100} surface attachment				
	E_{int} of s-p (eV)	E_{diff} of s-p (eV)	E_{int} of s-p-w (eV)	E_{diff} of s-p-w (eV)
SDEEV	-9.13	2.98	8.03	0.45
SAEEV	-6.15	0.00	7.54	0.94
SDAEV	-8.55	2.40	7.18	1.30
SDEAV	-7.56	1.41	7.72	0.76
SDAAV	-9.44	3.29	8.48	0.00
{111} surface attachment				
	E_{int} of s-p (eV)	E_{diff} of s-p (eV)	E_{int} of s-p-w (eV)	E_{diff} of s-p-w (eV)
SDEEV	-11.81	5.07	11.11	7.04
SAEEV	-15.91	9.17	6.76	11.38
SDAEV	-13.45	6.71	12.08	6.07
SDEAV	-9.80	3.05	18.14	0.00
SDAAV	-6.74	0.00	6.97	11.17

Table 6.44: Interfacial energies for all 185-189 pentapeptides on the {100} and {111} surface, using the Yang, Stipp and Harding method[164]. E_{int} of s-p is the interfacial energy of the slab–peptide system. E_{diff} of s-p is the difference in interfacial energy from the highest interfacial energy of the slab–peptide system. E_{int} of s-p-w is the interfacial energy of the slab–peptide–water system. E_{diff} of s-p-w is the difference in interfacial energy from the highest interfacial energy of the slab–peptide–water system.

The solvated 185-189 pentapeptide interfacial energy data showed that, the {100} surface produced lower interfacial energy values (between 7.18 eV and 8.48 eV) than the {111} surface, with the exception of SAEEV and SDAAV (between 11.11 eV and 18.14 eV (excluding SAEEV at 6.76 eV and SDAAV at 6.97 eV)). For the {100} surface, the SAEEV, SDAEV and SDEAV mutant sequences produced smaller interfacial energy values than SDEEV (with SDAEV producing the smallest), whereas, SDAAV produced a larger interfacial energy value than SDEEV (as hypothesised). This data suggests that, only the mutation to SDAAV causes a reduction in iron binding potential, with the single substitutions of acidic residues to alanine leading to increased iron binding affinity. For the {111} surface, the SDAEV and SDEAV mutations produced larger interfacial energies than SDEEV, whilst, SAEEV and SDAAV produced smaller interfacial energy values than SDEEV (with SAEEV producing the smallest), proposing that, the effect of mutation on iron binding potential was sequence dependent, as single glutamate mutations reduced iron

binding affinity, whereas, aspartate mutation and double glutamate mutation improved iron binding potential.

This data also showed that, solvation of the systems had an adverse effect on the interfacial energies, as the energies displayed were much lower for the *in vacu* system, with values increasing greatly, for example, on the {111} surface, SDEEV presented with an increase in interfacial energy from -11.81 eV to 11.11 eV upon solvation. The increase in interfacial energy can be explained by preferential attachment of water to the magnetite surface. Solvation of the system also affected the way that the sequences reacted at the interface, as the energy trends for the *in vacu* systems were very different to the related solvated systems. An example being, SDAAV on the {100} surface, in which, *in vacu*, the sequence had the lowest interfacial energy of the 185-189 pentapeptides, however, upon solvation it exhibited the highest interfacial energy.

There are, as with the other pentapeptides, issues with this method of directly calculating the adsorption energy, with larger than expected energy values produced and expected trends not exhibited. Potential reasons for the problems displayed were discussed for the 177-183 pentapeptides. Indirect Potential of Mean Force based methods of adsorption energy calculation are an alternative, and are considered in the following section, which utilise a constrained molecular dynamics version of the systems.

6.7 Constrained system

Free energy profiles were created for the aforementioned systems as shown in Figure 6.24.

The free energy profiles for the different 185-189 pentapeptides (Figure 6.24) showed that, on the {100} surface, all sequences exhibited very similar profiles. The mutants SDAEV and SDEAV slightly reduce the iron binding activity (with the free energy profile for SDAEV following very closely that of SDEEV), whereas, the SAEEV and SDAAV mutants slightly increase the iron binding activity by exhibiting lower free energy values. The mutant SDEAV produced the highest free energy of all 185-189 pentapeptide sequences. For SDEEV, SDAEV, SDEAV and SDAAV, between 10 Å and 5.5 Å was optimal for attachment, after which point, an increase in free energy was detected, corresponding with the presence of the smaller water absorption layer. It is proposed that the successive increase in free energy was attributable to a requirement for more energy within the system for the sequences to pass through the water barriers in order to enable sequence attach-

ment. Although, SDAAV exhibits a larger rate of increase in free energy corresponding with the stronger of the two water peaks, suggesting that, this particular water peak had more of an effect on iron binding ability of the sequence. However, for SAEEV, the optimal attachment region was between 8.5 Å and 3.5 Å as the profile exhibits negative values within this region. From this distance toward the surface an increase in free energy was exhibited, proposing that it was more difficult for the sequence to attach closer to the magnetite surface.

Whereas, the {111} surface energy profile showed that, SDAEV, SDEAV and SDAAV appeared to display similar free energy trends, whilst SDEEV and SAEEV exhibited a similar trend in profile. SDEEV presented with the lowest free energy profile, suggesting that the mutations are doing as hypothesised, whilst SDAEV produced the highest free energy of all 190-195 pentapeptide sequences. For SDAEV, SDEAV and SDAAV, between 10 Å and 5 Å, between 10 Å and 3 Å and between 10 Å and 7.5 Å, respectively, was optimal for attachment, after which point, a slight increase in free energy was detected, corresponding with the presence of the smaller water absorption layer. For SDAEV and SDEAV, the stronger of the water layers appears to have a greater effect on the free energy as a larger rate of increase was exhibited corresponding with this peak for these sequences, suggesting that, this particular water peak had more of an effect on iron binding ability of the sequence. For SDAAV, however, a greater rate of increase in free energy was exhibited when the sequence centre of mass had passed through the water layers and was getting closer to the magnetite surface, suggesting possible configurational effects. SDEEV exhibited a negative free energy throughout its profile until 0.5 Å, thus, attachment was favourable at any distance explored until then. The small changes in free energy seen for the SDEEV and SAEEV sequences suggest that the water absorption layers have minimal influence on the free energy profiles.

Upon examination of the energy minima for the 185-189 pentapeptides (Figure 6.25) it can be observed that, on the {100} surface, all of the 185-189 pentapeptides except for SDAAV exhibited a negative value for their energy minima, with SAEEV producing the lowest value at -0.4354 eV. The energy minima for these residues fall between 6.7 Å and 9.3 Å, suggesting that attachment of the residue centre of mass at this distance would be preferable. On the {111} surface, all of the 185-189 pentapeptides exhibited a negative value for their energy minima, with SDEEV producing the lowest value at -0.3578 eV. The trend of favourable attachment distance is different from the {100} surface, with SDEEV, SAEEV, SDAEV, SDEAV and SDAAV exhibiting centre of mass distances of 2.5 Å , 8.1

Å , 6.5 Å , 4.2 Å and 9.5 Å respectively.

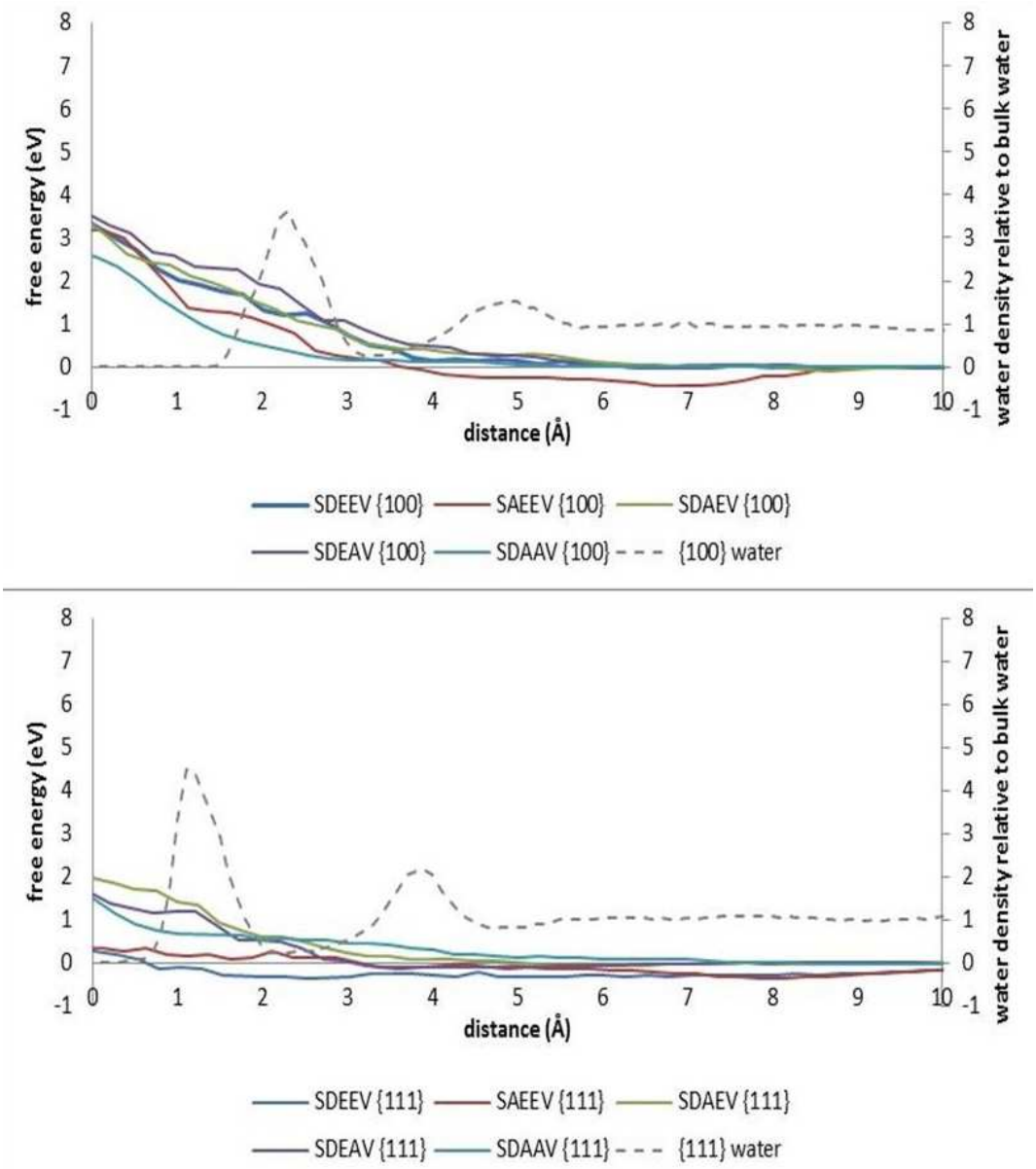

Figure 6.24: A comparison of the free energy profiles of the 185-189 pentapeptides. Distance refers to distance between the peptide CoM and the magnetite surface. The dashed line represents the water density profile for the system.

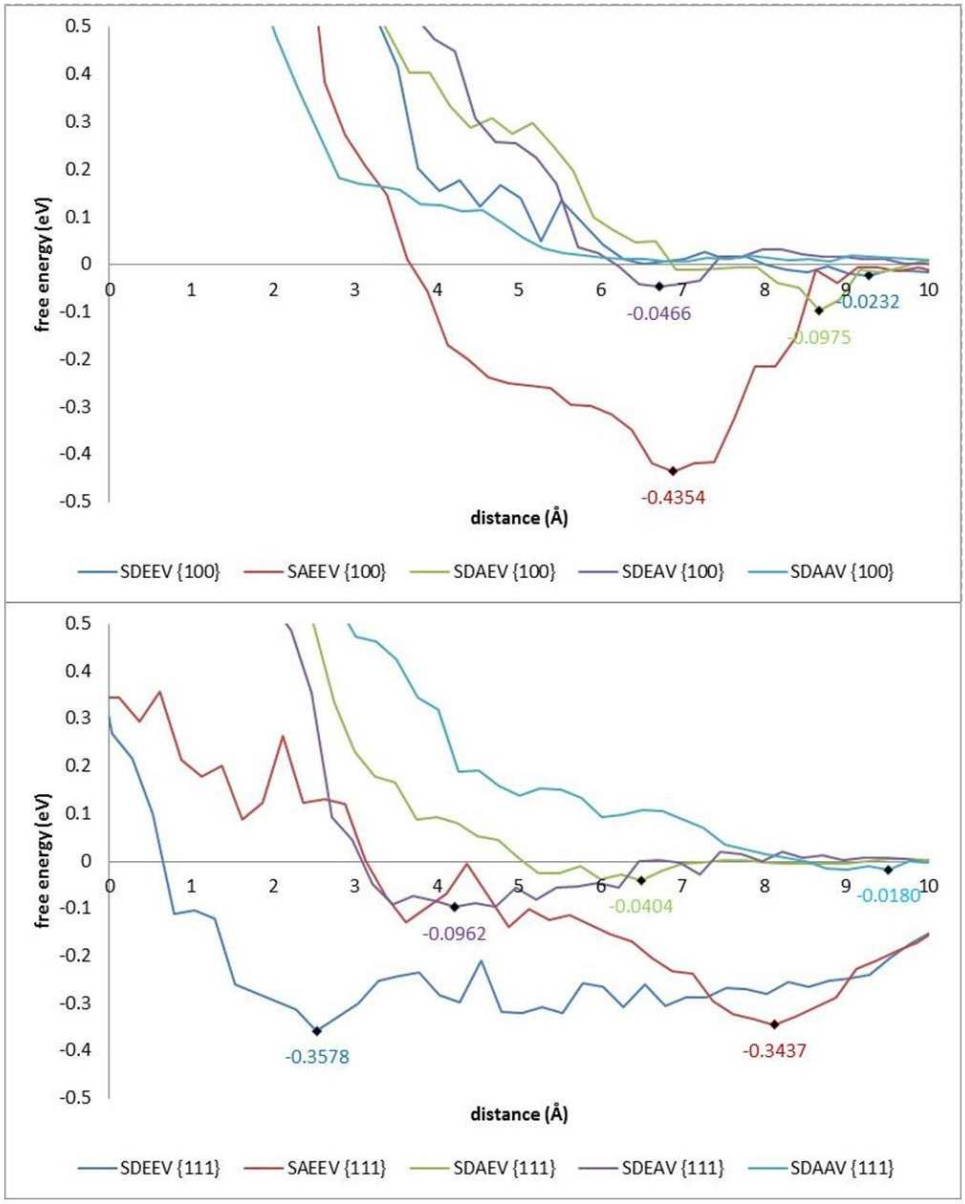

Figure 6.25: A close-up comparison of the free energy profiles of the 185-189 pentapeptides between -0.5 eV and 0.5 eV, exhibiting the free energy minima.

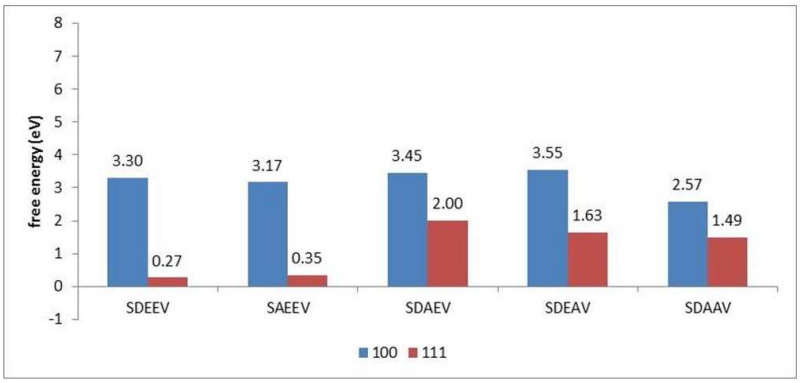

Figure 6.26: A comparison of the free energy values for the pentapeptide CoM at the magnetite surface for both the {100} and {111} surfaces.

The free energy values for the 185-189 pentapeptide centre of mass at the magnetite surfaces are shown in Figure 6.26. On the {100} surface, the SDEEV sequence exhibits a free energy value of 3.30 eV, whereas, on the {111} surface its value was much lower (0.27 eV). Similarly, the mutation sequences demonstrated higher free energy values (3.17 eV, 3.45 eV, 3.55 eV and 2.57 eV, respectively) on the {100} surface, than exhibited on the {111} surface (0.35 eV, 2.00 eV, 1.63 eV and 1.49 eV, respectively). The SDEEV and SAEEV sequences showed very similar free energy values for both surfaces, proposing that the aspartate substitution had little effect on iron binding ability. The mutation SDEAV had the greatest reductive effect on iron binding for the {100} surface, whereas, for the {111} surface, SDAEV had the greatest reductive effect on iron binding. On the {100} surface, the single glutamate substitution mutants reduced the iron binding affinity of the sequence, whilst, the double glutamate substitution mutant increased the iron binding ability. On the {111} surface, all mutants increased the free energy of the sequence attachment. This data showed that, the {111} surface presented with significantly lower free energy values than the {100} surface, representing superior iron binding, suggesting preferential attachment. The free energy profiles also showed that, the sequence with the lowest free energy when the CoM is at the surface was SDAAV for the {100} surface and SDEEV for the {111} surface.

6.8 Summary

6.8.1 Residues 177-183 (SRDIESA)

Unconstrained

Our results for the unconstrained 177-183 pentapeptide systems indicate that from the RDF profiles, tetrahedral iron attachment is preferred for the {100} surface and octahedral iron attachment is preferred for the {111} surface. For the solvated systems, DIASA on the {100} surface produces the most intense $g(r)$ peak for the tetrahedral iron attachment, whereas, the mutation SRAIE on the {111} surface produces the highest intensity $g(r)$ value for the octahedral iron attachment, leading to the idea that attachment of the pentapeptides to the magnetite surface is dependent on many things such as surface type, iron type and sequence.

For the bond distance data, more bonding was present in the {111} surface systems, suggesting that this is the preferred surface of attachment, with the exception of DIASA which exhibits the same amount of bonding on both surfaces for the *in vacu* system and shows a preference for the {100} surface in the solvated system. Also, in the solvated systems DIESA and DIASA exhibit the most bonding for the {100} surface, whereas, DIESA displayed the most on the {111} surface. Additionally, at least one bond is present to each acidic residue present when solvated, with the exception of SRDIE which exhibits only glutamate bonding on the {100} surface and only aspartate bonding on the {111} surface.

The residence time data indicates that for the *in vacu* systems, on the {100} surface, the most bonding lasting over 90% of the simulation length was exhibited for DIASA and SRDIE (bonds through acidic residues), whilst, on the {111} surface, DIESA exhibited the most, suggesting that attachment was surface and sequence dependent. For the solvated systems, on the {100} surface, the most bonding lasting over 90% of the simulation length was exhibited for DIESA, DIASA and SRAIE (all are bonds between tetrahedral iron and acidic residues), whilst, on the {111} surface, again, DIESA exhibited the most, suggesting that this sequence was the favoured for the {111} surface.

The interfacial energy data suggests that, the lowest solvated interfacial energies were found on the {100} surface, however, this data looks at average energy values of the full magnetite/peptide/water interface over a simulation, whereas other forms of data analysis

focus on bonding present between the peptides and the magnetite surfaces throughout the simulations. The {100} surface may exhibit lower interfacial energy due to the surfaces behaviour with water. The *in vacu* interfacial energies are considerably lower, further enforcing the detrimental effect of system solvation. For the solvated systems, on the {100} surface the mutations have lower interfacial energy values than their original sequence counterparts, which was contrary to the expected behaviour. On the {111} surface, DIESA has a lower interfacial energy than its mutated counterpart, however, SRDIE has a higher interfacial energy than its mutated counterpart, again reinforcing that attachment is surface and sequence dependent.

Constrained

The Potential of Mean Force data indicates that, the {111} surface was the preferred surface of attachment, as this surface produced much lower free energy values, thus much improved iron binding. Also, on the {100} surface, the mutation of DIESA to DIASA reduced iron binding, whilst on the {111} surface the mutation improved iron binding ability. With SRDIE, the opposite was displayed, with the mutation improving iron binding ability on the {100} surface, whilst, on the {111} surface the mutation reduces iron binding. This suggested that, sequence mutation alone does not always lead to reduced iron binding ability, as sequence, surface type and water influence also have an effect.

6.8.2 Residues 190-195 (ELRDAL)

Unconstrained

From the RDF data, for ELRDA and ELADA, for the {111} surface octahedral iron attachment was preferred, whilst, the {100} surface exhibits no particular preference. With LRDAL, octahedral iron attachment was preferred for the {100} surface, whilst the {111} surface exhibited no iron type preference. With LRAAL, octahedral iron attachment was the preferred for both surfaces. For the solvated systems, ELRDA on the {100} surface produced the most intense $g(r)$ peak for the tetrahedral iron attachment, whereas, ELADA on the {111} surface produced the highest intensity $g(r)$ value for the octahedral iron attachment, leading to the idea that attachment of the pentapeptides to the magnetite surface was dependent on many things such as surface type, iron type and sequence.

For the bond distance data, more bonding was present for the mutated peptide sequences

in the {111} surface systems, suggesting that this was the preferred surface of attachment. However, *in vacu*, the original sequences exhibited the same amount of bonding on both surfaces, suggesting that there was no surface of preference. Yet, when solvated, LRDAL exhibited the same amount of bonding on both surfaces, whilst, ELRDA exhibited more bonding on the {100} surface. Additionally, ELRDA exhibited the most bonding on the {100} surface, whereas, on the {111} surface, ELADA displayed the most bonding. Furthermore, at least one bond is present to each acidic residue present when solvated, with the exception of LRDAL on the {100} surface (exhibiting no acidic residue bonding) and LRAAL (which contains no acidic residues).

The residence time data indicates that, for the *in vacu* systems, ELRDA exhibits the most bonding lasting over 89% of the simulation length, for both surfaces, suggesting that attachment of this pentapeptide to the magnetite surface is favoured. For the solvated systems, ELRDA and ELADA exhibit the most bonding lasting over 99% of the simulation length on the {100} surface, whereas, on the {111} surface, ELADA exhibits the most bonding for this time length, suggesting that, attachment was surface and sequence dependent.

The interfacial energy data suggests that the lowest solvated energies were found on the {100} surface, with the exception of ELRDA whose lowest energy was on the {111} surface, however, this data looks at average energy values of the full magnetite/peptide/water interface over a simulation, whereas, other forms of data analysis focus on bonding present within the systems throughout the simulations. The *in vacu* interfacial energies are considerably lower, particularly for ELRDA, further establishing the unfavourable effect of water on the system. For the solvated systems, for both surfaces, ELRDA had a lower interfacial energy than ELADA, whilst, LRDAL had a higher interfacial energy than LRAAL, again emphasising that attachment was surface and sequence dependent.

Constrained

The Potential of Mean Force data indicates that the {111} surface is the preferred surface of attachment, as this surface produces the lowest free energy values throughout the simulations, with the exception of LRDAL. Also, for the free energy values when the CoM was at the magnetite surface, on both surfaces, the mutation of ELRDA (ELADA) improved iron binding. With LRDAL, on the {100} surface, the mutation reduced iron binding ability, whilst, on the {111} surface the mutation improved iron binding. Again, show-

ing that, surface type and pentapeptide sequence affect iron binding ability, along with sequence mutation.

6.8.3 Residues 188-192 (EVELR)

Unconstrained

From the RDF data, octahedral iron attachment is preferred for the {111} surface, whilst the {100} surface exhibits no particular preference. For the solvated systems, EVGELR on the {100} surface produced the highest intensity $g(r)$ value for the tetrahedral iron attachment, whereas, EVALR on the {111} surface produced the most intense $g(r)$ peak for the octahedral iron attachment, proposing that attachment of the pentapeptide to the magnetite surface was dependent on many things including surface type, iron type and sequence.

For the bond distance data, more bonding was present in the {111} surface system than the {100} surface system, with the exception of EVELR, which exhibited the same amount of bonding on both surfaces for the *in vacu* system, and displayed more bonding on the solvated {100} surface, suggesting that the {111} surface was the preferred surface of attachment. Additionally, EVELR exhibited the most bonding on the {100} surface, whereas, on the {111} surface, EVALR displayed the most bonding.

The residence time data indicates that for the *in vacu* systems, on the {100} surface, EVELR and EVALR exhibited the most bonding lasting over 99% of the simulation length, whereas, on the {111} surface, EVGELR exhibited the most bonding for this time length. For the solvated systems, on the {100} surface, all of the pentapeptides exhibit the same amount of bonding lasting over 99% of the simulation length, thus attachment is independent of sequence. Whilst, on the {111} surface, EVALR exhibited the most bonding lasting over 99% of the simulation length.

The interfacial energy data suggests that the lowest solvated energies were found on the {100} surface, with the exception of EVGELR, however, this data looks at average energy values of the full magnetite/amino acid residue/water interface over a simulation, whereas, other forms of data analysis focus on bonding present within the systems throughout the simulations. The *in vacu* interfacial energies were considerably lower than their equivalent solvated values. For the solvated systems, on the {100} surface, EVELR had a lower interfacial energy than its mutated counterparts (as hypothesised), however, on the {111}

surface, EVELR had a higher interfacial energy than EVALR and EVGELR.

Constrained

The Potential of Mean Force data indicates that the {111} surface is the preferred surface of attachment, as this surface produces the lowest free energy values throughout the simulations. From the pentapeptide CoM at the magnetite surface free energy data, on the {100} surface, both mutations reduced iron binding ability, whilst, on the {111} surface both mutations improved iron binding ability.

6.8.4 Residues 185-189 (SDEEV)

Unconstrained

From the RDF data, octahedral iron attachment is preferred for the {111} surface, whilst the {100} surface exhibits no particular preference. For the solvated systems, SAEEV on the {100} surface produced the highest intensity $g(r)$ value for the tetrahedral iron attachment, whereas, SDEAV on the {111} surface produced the most intense $g(r)$ peak for the octahedral iron attachment, again proposing that there are many factors affecting attachment other than sequence mutation.

For the bond distance data, more bonding was present in the {111} surface system, suggesting that this is the preferred surface of attachment. Also, SDEEV exhibited the most bonding for both surfaces. Additionally, at least one bond is present to each acidic residue present in the solvated system, with the exception of SDAEV on the {100} surface which exhibited no bonding to glutamate.

The residence time data indicates that for the *in vacu* systems, on both surfaces, SDEAV exhibited the most bonding that lasting over 90% of the simulation length, suggesting that attachment of this pentapeptide is preferential. For the solvated systems, on the {100} surface, SDEEV and SAEEV exhibit the most bonding lasting over 99% of the simulation length, whereas, on the {111} surface, SDEAV and SDAAV exhibit the most bonding lasting for this length of time.

The interfacial energy data suggests that, the lowest solvated energies were found on the {100} surface, with the exception of SAEEV and SDAAV. The {100} surface may exhibit lower interfacial energy due to the surfaces behaviour with water. The *in vacu* interfacial

energies are, again, considerably lower than the corresponding solvated system values. For the solvated systems, on the {100} surface, SDAEV had the lowest interfacial energy, however, SAEEV had the lowest interfacial energy on the {111} surface, again, reinforcing that attachment is surface and sequence dependent.

Constrained

The Potential of Mean Force data indicates that the {111} surface is the preferred surface of attachment for all peptide sequences explored, as this surface produces much lower free energy values, thus much improved iron binding. From the pentapeptide CoM at the magnetite surface free energy data it was seen that, on the {100} surface the sequences have very similar free energy values, suggesting that mutation has very minimal effect on the iron binding affinity with this surface, with the exception of SDAAV which improved iron binding ability. On the {111} surface, SAEEV and SDEEV exhibited similar free energy values, whilst, the other mutations reduced iron binding affinity, with SDAEV having the greatest effect.

6.8.5 General

This chapter demonstrates, for the first time, the interaction of experimentally designed pentapeptides and their related mutants based on the C–terminal Mms6 sequence with *in vacu* and solvated magnetite crystal {100} and {111} surfaces, in both unconstrained and constrained systems. The hypothesis of this study was that iron binding would decrease with mutation of the original Mms6 C–terminal sequences; however, the data showed that this was not always the case. For the unconstrained 177-183 pentapeptides (SRDIESA) natural solvated system data, the hypothesis was confirmed for the {111} surface, with the exception of the interfacial energy for SRDIE. However, for the {100} surface, it was deduced that the mutation either had no effect on how the sequence bound or the mutation improved iron binding. For the unconstrained 190-195 pentapeptides (ELRDAL) natural solvated system data, the hypothesis was only confirmed with the bond distance data for LRDAL on both surfaces, and ELRDA on the {100} surface, and using the interfacial energy data for ELRDA on both surfaces. The hypothesis was contradicted using the residence time data, as for both ELRDA and LRDAL the mutants either exhibited the same amount of bonding as the original sequence lasting over 89% of the simulation length or displayed more bonding lasting for this length of time. Thus, no obvious trend in the

effect of mutation was exhibited for the 190-195 sequences. For the unconstrained 188-192 pentapeptides (EVELR) natural solvated system data, the hypothesis was confirmed for the {100} surface, with the exception of the residence time data, whereby, the mutations had the same amount of bonding lasting over 99% of the simulation length as EVELR. However, for the {111} solvated system, it was deduced that one or other of the mutations, dependent on the data analysis method, improved iron binding (EVALR from the bond distance data and the residence time data and EVGELR from the interfacial energy data). For the unconstrained 185-189 pentapeptides (SDEEV) natural solvated system data, the hypothesis was only confirmed using the bond distance data on both surfaces, and the residence time data on the {100} surface. The hypothesis was contradicted using the interfacial energy data for both surfaces, as in both cases a mutant produces the lowest interfacial energy value (SDAEV for the {100} surface and SAEEV for the {111} surface), and using the residence time data on the {111} surface, as the mutant sequences SDEAV and SDAAV exhibit the most bonding lasting over 99% of the simulation length. Thus, no obvious trend in the effect of mutation was exhibited for the 185-189 sequences. The unconstrained data proposes that, there are many factors affecting the iron binding ability of all pentapeptides explored other than sequence mutation, such as, surface type, iron type and sequence dependence.

The constrained system data also showed this to be true, as the Potential of Mean Force data indicated that, for the 177-183 pentapeptides, on the {100} surface, the mutation of DIESA to DIASA confirmed the hypothesis, whilst, on the {111} surface the mutation improved iron binding ability. However, the opposite was seen for the mutation of SRDIE, with the {111} surface confirming the hypothesis. For the 190-195 pentapeptides, on both surfaces, the mutation of ELRDA disproved the hypothesis, improving iron binding. The same phenomenon was exhibited for LRDAL on the {111} surface, however, on the {100} surface the hypothesis was proved correct. For the 188-192 pentapeptides, the hypothesis was confirmed for the {100} surface, however, on the {111} surface the mutation improved iron binding affinity. For the 185-189 pentapeptides, on the {100} surface, it was deduced that the mutation had minimal effect on how the sequence bound, as the mutants produced similar free energy values to SDEEV, with the exception of SDAAV which improved iron binding ability. On the {111} surface, the hypothesis was confirmed for the mutants with the exception of SAEEV which exhibited a similar free energy value to SDEEV. The constrained data proposes, again, that sequence mutation was not the only factor affecting iron binding potential. The {111} surface was the preferred surface of attachment, as this surface produces the lowest free energy values throughout the simulations, with the

exception of LRDAL.

Chapter 7

Conclusion

Atomistic simulations were used in this thesis to model the magnetite surface interactions at the atomic level. The aim of this work was to attempt to further the current understanding of the biomineralisation of magnetite in magnetotactic bacteria through computational simulations. The results in this thesis cover a range of organic molecule attachments on the {100} and {111} magnetite surfaces.

The aim of Chapter 4 was to reveal, for the first time, the interaction of the individual C–terminal Mms6 amino acid residues with magnetite surfaces. The {100} and {111} surfaces were considered both *in vacu* and solvated. It was established from the unconstrained system data that, the {111} surface was particularly favourable for iron binding to occur and that octahedral iron attachment was preferred over tetrahedral iron attachment. Additionally, it was apparent that a higher affinity for iron binding was displayed by the acidic amino acids. Furthermore, the addition of water molecules to the unconstrained systems diminished the iron binding abilities of all the amino acids. The Potential of Mean Force data obtained further supported that, the 111 surface was the preferred surface of attachment, as this surface produced significantly lower free energy values for the majority of amino acid investigated, the exception being glutamine.

Chapter 5 focused on the interaction of *in vacu* and solvated magnetite surfaces with a glutamic acid repeat motif and its corresponding sequence mutations, in both unconstrained and constrained systems. The results of data analysis for the unconstrained systems showed that, as with the amino acids, octahedral iron bonding was more favourable and that the {111} surface was the preferred for dipeptide and tetrapeptide attachment. The hypothesis of this study was that iron binding ability would decrease with sequential mutations. This was proven to be true for the dipeptides, with less bonding present with each subsequent glutamate substitution. The addition of extra amino acid residues to the sequences lead to preferential tetrapeptide sequence attachment. Again, solvation

was detrimental to the iron binding potential of the peptides. The constrained system data, again, confirmed the unconstrained findings, indicating that the {111} surface is the preferred surface of attachment, with the exception of DEEV. The Potential of Mean Force data also shows that, mutation has surface–dependent differential modulatory effects on the iron binding ability of peptide sequences, as for the dipeptides on both surfaces and the tetrapeptides on the {100} surface, confirmed the hypothesis whereas, for the tetrapeptides on the {111} surface the hypothesis was contradicted.

In Chapter 6, for the first time, the interactions of the C–terminal Mms6 pentapeptides and their related mutants with the magnetite {100} and {111} surfaces, in both unconstrained and constrained systems, were investigated. The hypothesis of this study was that iron binding would decrease with mutation of the original Mms6 C–terminal sequences; however, the data showed that this was not always the case. For the unconstrained solvated 177-183 pentapeptides (SRDIESA), the hypothesis was proven correct for the {111} surface, with the exception of the interfacial energy for SRDIE. However, for the {100} surface, mutation either improved iron binding or had no effect on sequence binding. For the solvated unconstrained 190-195 pentapeptides (ELRDAL), the hypothesis was only confirmed with the bond distance data for LRDAL (both surfaces) and ELRDA ({100} surface), and using the interfacial energy data for ELRDA (both surfaces). The hypothesis was disproved using the residence time data for both ELRDA and LRDAL, as the mutants either exhibit the same amount of bonding as, or more bonding than, the original sequences lasting over 89% of the simulation length. For the solvated unconstrained 188-192 pentapeptides (EVELR), the hypothesis was confirmed for the {100} surface, with the exception of the residence time data. Conversely, for the {111} surface, one or other of the mutations improved the iron binding potential of the sequence (EVALR with bond distance and residence time data and EVGELR with the interfacial energy data). For the solvated unconstrained 185-189 pentapeptides (SDEEV), the hypothesis was only confirmed using the bond distance data (both surfaces) and the residence time data ({100} surface). The hypothesis was refuted using the interfacial energy data (both surfaces), as a mutant produced the lowest interfacial energy (SDAEV for the {100} surface and SAEEV for the {111} surface), and using the residence time data ({111} surface), as the mutant sequences SDEAV and SDAAV exhibit the most bonding lasting over 99% of the simulation length. Additionally, the presence of water in the unconstrained systems was detrimental to the iron binding potential of the pentapeptides.

The constrained pentapeptide system data also showed that, iron binding did not always

decrease with mutation of the original sequences, as the free energy profiles showed that, for the 177-183 pentapeptides, the mutation of DIESA confirmed the hypothesis on the {100} surface, but, improved iron binding ability on the {111} surface, whereas, for the SRDIE mutation the reverse was seen. For the 190-195 pentapeptides, the mutation of ELRDA disproved the hypothesis, improving iron binding, which was also the case for LRDAL on the {111} surface, however, the hypothesis was proven on the {100} surface. For the 188-192 pentapeptides, for the {100} surface the hypothesis was confirmed, however, mutation improved iron binding affinity on the {111} surface. For the 185-189 pentapeptides, on the {111} surface, the hypothesis was verified, with the exception of SAEEV, however, on the {100} surface, as mutation either had minimal effect on sequence binding or improved it (SDAAV). The Potential of Mean Force data obtained further supported that, the 111 surface was the preferred surface of attachment, as this surface produces the lowest free energies, with the exception of LRDAL.

For all studies undertaken, the interfacial energy method of data analysis encountered problems, with the absolute adsorption energy values appearing to be an order of magnitude larger than expected, and the solvated behavioural trends shown being counter-intuitive and contradictive of experimental studies. Possible reasons for these issues could be that the interfacial energy method entails the measurement of small variances between large energies, leading to large errors in the results. Also, for the hydrated systems, if insufficient water was present at the start of the simulation, some of the water molecules will eventually become under coordinated, thus increasing the energy of the simulation cells. Future work would validate this method of data analysis or establish a more reliable method of direct adsorption energy calculation.

This work has, for the first time, investigated the interaction of the C–terminal Mms6 protein with the magnetite {100} and {111} surfaces, in both unconstrained and constrained systems, through the use of amino acids and, di-, tetra- and pentapeptides and related mutants. These studies have highlighted the strong influence that water has over the systems and that for the majority of biomolecules explored the 111 surface was the preferred surface of attachment. However, this work can be taken much further. Future work to advance the current investigations would explore ever increasing polypeptide chain lengths in order to explore the influence of extra amino acids on the iron binding potential, and whether particular residues within the sequence affect how others behaviour, also how this influences iron binding. This would naturally lead to the attachment of the full Mms6 C-terminal sequence to the magnetite surfaces, which could subsequently be mutated.

Another possible direction for future work would be to investigate different starting positions and configurations for the biomolecules, establishing whether the binding behaviours would significantly differ from what has already been revealed. Furthermore, individual iron ions could be utilised to study iron binding, as opposed magnetite surfaces, in order to get a better idea of where the preferred areas of iron attachment are on the original sequences. This could then lead to the addition of extra iron ions into the biomolecule-iron ion system, so as to examine the influence this addition has on the biomolecule configuration and the free energy of the system.

References

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

[2] E. Crabb and E. Moore. *Metals and Life*. Royal Society of Chemistry in association with the Open University, 2010.

[3] M. Tanaka, E. Mazuyama, A. Arakaki, and T. Matsunaga. Mms6 protein regulates crystal morphology during nano-sized magnetite biomineralization in vivo. *Journal of Biological Chemistry*, 286(8):6386–6392, 2011.

[4] H.A. Lowenstam. Minerals formed by organisms. *Science*, 211(4487):1126–1131, 1981.

[5] H.A. Lowenstam and S. Weiner. *On biomineralization*. Oxford University Press, 1989.

[6] D.A. Bazylinski and S. Schübbe. Controlled biomineralization by and applications of magnetotactic bacteria. *Advances in Applied Microbiology*, 62:21–62, 2007.

[7] PE Bell, AL Mills, and JS Herman. Biogeochemical conditions favoring magnetite formation during anaerobic iron reduction. *Applied and Environmental Microbiology*, 53(11):2610–2616, 1987.

[8] R.B. Frankel. Biological permanent magnets. *Hyperfine Interactions*, 151(1):145–153, 2003.

[9] S. Mann, N.H. Sparks, M.M. Walker, and J.L. Kirschvink. Ultrastructure, morphology and organization of biogenic magnetite from sockeye salmon, oncorhynchus nerka: implications for magnetoreception. *Journal of Experimental Biology*, 140(1):35–49, 1988.

[10] D.A. Bazylinski and R.B. Frankel. Biologically controlled mineralization in prokaryotes. *Reviews in Mineralogy and Geochemistry*, 54(1):217, 2003.

[11] S. Weiner and P.M. Dove. An overview of biomineralization processes and the problem of the vital effect. *Reviews in Mineralogy and Geochemistry*, 54(1):1–29, 2003.

[12] H.A. Lowenstam. Mineralization processes in monerans and protoctists. *Biomineralization in lower plants and animals*, 30:1–17, 1986.

[13] D.A. Bazylinski and B.M. Moskowitz. Microbial biomineralization of magnetic iron minerals; microbiology, magnetism and environmental significance. *Reviews in Mineralogy and Geochemistry*, 35(1):181–223, 1997.

[14] K.O. Konhauser. Bacterial iron biomineralisation in nature. *FEMS Microbiology Reviews*, 20(3-4):315–326, 1997.

[15] K.O. Konhauser. Diversity of bacterial iron mineralization. *Earth Science Reviews*, 43(3-4):91–121, 1998.

[16] R.B. Frankel and R.P. Blakemore. *Iron biominerals*. Plenum Press, 1991.

[17] U. Schwertmann and R.M. Cornell. The iron oxides: Structure, properties, occurrences, and uses, 1991.

[18] R.W. Fitzpatrick, R. Naidu, and P.G. Self. Iron deposits and microorganisms in saline sulfidic soils with altered soil water regimes in south australia. *Catena, Supplement*, 21:263–286, 1992.

[19] S. Mann, J. Webb, and R.J.P. Williams. *Biomineralization: chemical and biochemical perspectives*. VCH, 1989.

[20] K. Simkiss and K.M. Wilbur. *Biomineralization*. Elsevier, 1989.

[21] J. Webb, D.J. Macey, W. Chua-Anusorn, T.G. St Pierre, L.R. Brooker, I. Rahman, and B. Noller. Iron biominerals in medicine and the environment. *Coordination Chemistry Reviews*, 190:1199–1215, 1999.

[22] H.A. Lowenstam. Lepidocrocite, an apatite mineral, and magnetite in teeth of chitons (polyplacophora). *Science*, 156(3780):1373–1375, 1967.

[23] K.J. Lohmann and C.M.F. Lohmann. Detection of magnetic field intensity by sea turtles. *Nature*, 380:5–61, 1996.

[24] K.J. Lohmann, S.D. Cain, S.A. Dodge, and C.M.F. Lohmann. Regional magnetic fields as navigational markers for sea turtles. *Science*, 294(5541):364–366, 2001.

[25] C.V. Mora, M. Davison, J.M. Wild, and M.M. Walker. Magnetoreception and its trigeminal mediation in the homing pigeon. *Nature*, 432(7016):508–511, 2004.

[26] R.A. Holland, K. Thorup, M.J. Vonhof, W.W. Cochran, and M. Wikelski. Navigation: Bat orientation using earth's magnetic field. *Nature*, 444(7120):702–702, 2006.

[27] J.L. Kirschvink, A. Kobayashi-Kirschvink, and B.J. Woodford. Magnetite biomineralization in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 89(16):7683, 1992.

[28] F. Brem, A.M. Hirt, M. Winklhofer, K. Frei, Y. Yonekawa, H-G. Wieser, and J. Dobson. Magnetic iron compounds in the human brain: a comparison of tumour and hippocampal tissue. *Journal of The Royal Society Interface*, 3(11):833–841, 2006.

[29] R.J. Harrison, R.E. Dunin-Borkowski, and A. Putnis. Direct imaging of nanoscale magnetic interactions in minerals. *Proceedings of the National Academy of Sciences*, 99(26):16556–16561, 2002.

[30] N.A. Spaldin. *Magnetic Materials: Fundamentals and Applications*. Magnetic Materials: Fundamentals and Applications. Cambridge University Press, 2010.

[31] L. Perloff. Black opaque octahedral crystals of magnetite on matrix. http://www.test.org/doe/http://www.webmineral.com/data/Magnetite.shtml, 2013. Accessed: 20/02/2013.

[32] S-B.R. Chang, J.F. Stolz, J.L. Kirschvink, and S.M. Awramik. Biogenic magnetite in stromatolites. ii. occurrence in ancient sedimentary environments. *Precambrian Research*, 43(4):305–315, 1989.

[33] J.W. Schopf, E.S. Barghoorn, M.D. Maser, and R.O. Gordon. Electron microscopy of fossil bacteria two billion years old. *Science*, 149(3690):1365–1367, 1965.

[34] H.A. Lowenstam. Magnetite in denticle capping in recent chitons (polyplacophora). *Geological Society of America Bulletin*, 73(4):435–438, 1962.

[35] J.L. Kirschvink and H.A. Lowenstam. Mineralization and magnetization of chiton teeth: Paleomagnetic, sedimentologic, and biologic implications of organic magnetite. *Earth and Planetary Science Letters*, 44(2):193–204, 1979.

[36] R.P. Blakemore. Magnetotactic bacteria. *Science*, 190(4212):377–379, 1975.

[37] R.P. Blakemore. Magnetotactic bacteria. *Annual Reviews in Microbiology*, 36(1):217–238, 1982.

[38] Y.A. Gorby, T.J. Beveridge, and R.P. Blakemore. Characterization of the bacterial magnetosome membrane. *Journal of Bacteriology*, 170(2):834–841, 1988.

[39] D.A. Bazylinski, R.B. Frankel, B.R. Heywood, S. Mann, J.W. King, P.L. Donaghay, and A.K. Hanson. Controlled biomineralization of magnetite (fe (inf3) o (inf4)) and greigite (fe (inf3) s (inf4)) in a magnetotactic bacterium. *Applied and Environmental Microbiology*, 61(9):3232–3239, 1995.

[40] R.P. Blakemore, D. Maratea, and R.S. Wolfe. Isolation and pure culture of a freshwater magnetic spirillum in chemically defined medium. *Journal of Bacteriology*, 140(2):720–729, 1979.

[41] T. Matsunaga, T. Sakaguchi, and F. Tadakoro. Magnetite formation by a magnetic bacterium capable of growing aerobically. *Applied Microbiology and Biotechnology*, 35(5):651–655, 1991.

[42] D.A. Bazylinski, R.B. Frankel, and H.W. Jannasch. Anaerobic magnetite production by a marine, magnetotactic bacterium. *Nature*, 334:518, 1988.

[43] R.B. Frankel, R.P. Blakemore, and R.S. Wolfe. Magnetite in freshwater magnetotactic bacteria. *Science*, 203(4387):1355, 1979.

[44] J.W.E. Fassbinder, H. Stanjekt, and H. Vali. Occurrence of magnetic bacteria in soil. *Nature*, 343(6254):161–163, 1990.

[45] J.F. Stolz, S-B.R. Chang, and J.L. Kirschvink. Magnetotactic bacteria and single-domain magnetite in hemipelagic sediments. *Nature*, 321:849–851, 1986.

[46] D.R. Lovley, J.F. Stolz, G.L. Nord, and E.J.P. Phillips. Anaerobic production of magnetite by a dissimilatory iron-reducing microorganism. *Nature*, 330(6145):252–254, 1987.

[47] R.P. Blakemore and R.B. Frankel. Biomineralization by magnetogenic bacteria. *Metal-microbe Interactions. IRL Press, New York*, pages 85–98, 1989.

[48] J.F. Stolz. Magnetotactic bacteria: biomineralization, ecology, sediment magnetism, environmental indicator. *Catena Supplement*, 21:133–133, 1993.

[49] J.L. Kirschvink. South-seeking magnetic bacteria: Short communications. *Journal of Experimental Biology*, 86(1):345–347, 1980.

[50] R.P. Blakemore, R.B. Frankel, and A.J. Kalmijn. South-seeking magnetotactic bacteria in the southern hemisphere. *Physics*, page 159, 1980.

[51] R.B. Frankel, R.P. Blakemore, F.F. T. De Araujo, D.M.S. Esquivel, and J. Danon. Magnetotactic bacteria at the geomagnetic equator. *Science*, 212(4500):1269–1270, 1981.

[52] L. Chen, D.A. Bazylinski, and B.H. Lower. Bacteria that synthesize nano-sized compasses to navigate using earth's geomagnetic field. *Nature Education Knowledge*, 1(10):14, 2010.

[53] D.A. Bazylinski, A.J. Garratt-Reed, and R.B. Frankel. Electron microscopic studies of magnetosomes in magnetotactic bacteria. *Microscopy research and technique*, 27(5):389–401, 1994.

[54] S. Spring and K-H. Schleifer. Diversity of magnetotactic bacteria. *Systematic and Applied Microbiology*, 18(2):147–153, 1995.

[55] D. Schüler. Formation of magnetosomes in magnetotactic bacteria. *Journal of Molecular Microbiology and Biotechnology*, 1(1):79–86, 1999.

[56] F.C. Meldrum, S. Mann, B.R. Heywood, R.B. Frankel, and D.A. Bazylinski. Electron microscopy study of magnetosomes in a cultured coccoid magnetotactic bacterium. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 251(1332):231–236, 1993.

[57] F.C. Meldrum, S. Mann, B.R. Heywood, R.B. Frankel, and D.A. Bazylinski. Electron microscopy study of magnetosomes in two cultured vibrioid magnetotactic bacteria. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 251(1332):237–242, 1993.

[58] S. Mann, N.C.H. Sparks, R.B. Frankel, D.A. Bazylinski, and H.W. Jannasch. Biomineralization of ferrimagnetic greigite (fe3s4) and iron pyrite (fes2) in a magnetotactic bacterium. *Nature*, 343:258, 1990.

[59] B. Devouard, M. Posfai, X. Hua, D.A. Bazylinski, R.B. Frankel, and P.R. Buseck. Magnetite from magnetotactic bacteria; size distributions and twinning. *American Mineralogist*, 83(11-12 Part 2):1387, 1998.

[60] A. Komeili. Molecular mechanisms of magnetosome formation. *Biochemistry*, 76(1):351, 2007.

[61] D.A. Bazylinski and R.B. Frankel. Biologically controlled mineralization of magnetic iron minerals by magnetotactic bacteria. *Environmental microbe-metal interactions*, pages 109–143, 2000.

[62] R.F. Butler and S.K. Banerjee. Theoretical single-domain grain size range in magnetite and titanomagnetite. *Journal of Geophysical Research*, 80(29):4049–4058, 1975.

[63] D.A. Bazylinski and R.B. Frankel. Magnetosome formation in prokaryotes. *Nature Reviews Microbiology*, 2(3):217–230, 2004.

[64] R.B. Frankel, D.A. Bazylinski, M.S. Johnson, and B.L. Taylor. Magneto-aerotaxis in marine coccoid bacteria. *Biophysical Journal*, 73(2):994–1000, 1997.

[65] H. Vali and J.L. Kirschvink. Magnetofossil dissolution in a palaeomagnetically unstable deep-sea sediment. *Nature*, 339(6221):203–206, 1989.

[66] J.L. Kirschvink and S.B.R. Chang. Ultrafine-grained magnetite in deep-sea sediments: Possible bacterial magnetofossils. *Geology*, 12(9):559, 1984.

[67] N. Petersen, T. von Dobeneck, and H. Vali. Fossil bacterial magnetite in deep-sea sediments from the south atlantic ocean. *Nature*, 320(6063):611–615, 1986.

[68] A.S. Bahaj, I.W. Croudace, and P.A.B. James. Treatment of heavy metal contaminants using magnetotactic bacteria. In *IEEE Transactions on Magnetics. Joint MMM-Intermag Conference*, pages 4707–4709, 1994.

[69] A.S. Bahaj, I.W. Croudace, P.A.B. James, F.D. Moeschler, and P.E. Warwick. Continuous radionuclide recovery from wastewater using magnetotactic bacteria. *Journal of Magnetism and Magnetic Materials*, 184(2):241–244, 1998.

[70] J.W.M. Bulte and R.A. Brooks. Magnetic nanoparticles as contrast agents for mr imaging. In *Scientific and clinical applications of magnetic carriers*, pages 527–543. Springer, 1997.

[71] J.W.M. Bulte, T. Douglas, S. Mann, R.B. Frankel, B.M. Moskowitz, R.A. Brooks, C.D. Baumgarner, J. Vymazal, and J.A. Frank. Magnetoferritin: biomineralization as a novel molecular approach in the design of iron-oxide-based magnetic resonance contrast agents. *Investigative Radiology*, 29:S214–216, 1994.

[72] J.W.M. Bulte, T. Douglas, S. Mann, R.B. Frankel, B.M. Moskowitz, R.A. Brooks, C.D. Baumgarner, J. Vymazal, M.P. Strub, and J.A. Frank. Magnetoferritin: characterization of a novel superparamagnetic mr contrast agent. *Journal of Magnetic Resonance Imaging*, 4(3):497–505, 1994.

[73] R. Reszka. Applications for magnetosomes in medical research. *Biomineralization. Wiley-VCH, Weinheim, Germany*, pages 81–92, 2000.

[74] E. Bäuerlein, D. Schüler, R. Reszka, and S. Päuser. Specific magnetosome, method for the production and use thereof, 2001. US Patent 6,251,365.

[75] C. Sun, J.S.H. Lee, and M. Zhang. Magnetic nanoparticles in mr imaging and drug delivery. *Advanced Drug Delivery Reviews*, 60(11):1252–1265, 2008.

[76] N. Nakamura, J.G. Burgess, K. Yagiuda, S. Kudo, T. Sakaguchi, and T. Matsunaga. Detection and removal of escherichia coli using fluorescein isothiocyanate conjugated monoclonal antibody immobilized on bacterial magnetic particles. *Analytical Chemistry*, 65(15):2036–2039, 1993.

[77] N. Nakamura and T. Matsunaga. Highly sensitive detection of allergen using bacterial magnetic particles. *Analytica Chimica Acta*, 281(3):585–589, 1993.

[78] R. Hergt, R. Hiergeist, M. Zeisberger, D. Schüler, U. Heyen, I. Hilger, and W.A. Kaiser. Magnetic properties of bacterial magnetosomes as potential diagnostic and therapeutic tools. *Journal of Magnetism and Magnetic Materials*, 293(1):80–86, 2005.

[79] R. Hergt and S. Dutz. Magnetic particle hyperthermiabiophysical limitations of a visionary tumour therapy. *Journal of Magnetism and Magnetic Materials*, 311(1):187–192, 2007.

[80] S. Mornet, S. Vasseur, F. Grasset, and E. Duguet. Magnetic nanoparticle design for medical diagnosis and therapy. *Journal of Materials Chemistry*, 14(14):2161–2175, 2004.

[81] J. Dobson. Magnetic nanoparticles for drug delivery. *Drug Development Research*, 67(1):55–60, 2006.

[82] A.S. Lübbe, C. Alexiou, and C. Bergemann. Clinical applications of magnetic drug targeting. *Journal of Surgical Research*, 95(2):200–206, 2001.

[83] N. Kohler, C. Sun, J. Wang, and M. Zhang. Methotrexate-modified superparamagnetic nanoparticles and their intracellular uptake into human cancer cells. *Langmuir*, 21(19):8858–8864, 2005.

[84] N. Kohler, C. Sun, A. Fichtenholtz, J. Gunn, C. Fang, and M. Zhang. Methotrexate-immobilized poly (ethylene glycol) magnetic nanoparticles for mr imaging and drug delivery. *Small*, 2(6):785–792, 2006.

[85] T.K. Jain, M.A. Morales, S.K. Sahoo, D.L. Leslie-Pelecky, and V. Labhasetwar. Iron oxide nanoparticles for sustained delivery of anticancer agents. *Molecular pharmaceutics*, 2(3):194–205, 2005.

[86] F. Scherer, M. Anton, U. Schillinger, J. Henke, C. Bergemann, A. Krüger, B. Gänsbacher, and C. Plank. Magnetofection: enhancing and targeting gene delivery by magnetic force in vitro and in vivo. *Gene therapy*, 9(2), 2002.

[87] F. Krötz, H-Y. Sohn, T. Gloe, C. Plank, and U. Pohl. Magnetofection potentiates gene delivery to cultured endothelial cells. *Journal of vascular research*, 40(5):425–434, 2003.

[88] J. Dobson. Gene therapy progress and prospects: magnetic nanoparticle-based gene delivery. *Gene Therapy*, 13(4):283–287, 2006.

[89] A. Arakaki, H. Nakazawa, M. Nemoto, T. Mori, and T. Matsunaga. Formation of magnetite by bacteria and its application. *Journal of the Royal Society interface*, 5(26):977–999, 2008.

[90] Y. Okamura, H. Takeyama, and T. Matsunaga. A magnetosome-specific gtpase from the magnetic bacteriummagnetospirillum magneticum amb-1. *Journal of Biological Chemistry*, 276(51):48183–48188, 2001.

[91] T. Matsunaga, C. Nakamura, J.G. Burgess, and K. Sode. Gene transfer in magnetic bacteria: transposon mutagenesis and cloning of genomic dna fragments required for magnetosome synthesis. *Journal of bacteriology*, 174(9):2748–2753, 1992.

[92] D. Schultheiss and D. Schüler. Development of a genetic system for magnetospirillum gryphiswaldense. *Archives of microbiology*, 179(2):89–94, 2003.

[93] T. Matsunaga, Y. Okamura, Y. Fukuda, A.T. Wahyudi, Y. Murase, and H. Takeyama. Complete genome sequence of the facultative anaerobic magnetotactic bacterium magnetospirillum sp. strain amb-1. *DNA research*, 12(3):157–166, 2005.

[94] S.K. Banerjee and B.M. Moskowitz. Ferrimagnetic properties of magnetite. In *Magnetite biomineralization and magnetoreception in organisms*, pages 17–41. Springer, 1985.

[95] M. Tanaka, Y. Okamura, A. Arakaki, T. Tanaka, H. Takeyama, and T. Matsunaga. Origin of magnetosome membrane: proteomic analysis of magnetosome membrane and comparison with cytoplasmic membrane. *Proteomics*, 6(19):5234–5247, 2006.

[96] A. Komeili, H. Vali, T.J. Beveridge, and D.K. Newman. Magnetosome vesicles are present before magnetite formation, and mama is required for their activation. *Proceedings of the National Academy of Sciences of the United States of America*, 101(11):3839–3844, 2004.

[97] C. Nakamura, T. Kikuchi, J.G. Burgess, and T. Matsunaga. Iron-regulated expression and membrane localization of the maga protein in magnetospirillum sp. strain amb-1. *Journal of Biochemistry*, 118(1):23–27, 1995.

[98] A. Scheffel, M. Gruska, D. Faivre, A. Linaroudis, J.M. Plitzko, and D. Schüler. An acidic protein aligns magnetosomes along a filamentous structure in magnetotactic bacteria. *Nature*, 440(7080):110–114, 2006.

[99] T. Matsunaga, T. Sakaguchi, and Y. Okamura. Molecular and biotechnological aspects of bacterial magnetite. *Biomineralization: Progress in Biology, Molecular Biology and Application*, 2006.

[100] A. Arakaki, J. Webb, and T. Matsunaga. A novel protein tightly bound to bacterial magnetic particles in magnetospirillum magneticum strain amb-1. *Journal of Biological Chemistry*, 278(10):8745–8750, 2003.

[101] K. Grünberg, C. Wawer, B.M. Tebo, and D. Schüler. A large gene cluster encoding several magnetosome proteins is conserved in different species of magnetotactic bacteria. *Applied and Environmental Microbiology*, 67(10):4573–4582, 2001.

[102] D. Schüler. Molecular analysis of a subcellular compartment: the magnetosome membrane in magnetospirillum gryphiswaldense. *Archives of Microbiology*, 181(1):1–7, 2004.

[103] T. Matsunaga and Y. Okamura. Genes and proteins involved in bacterial magnetic particle formation. *Trends in Microbiology*, 11(11):536–541, 2003.

[104] Y. Amemiya, A. Arakaki, S.S. Staniland, T. Tanaka, and T. Matsunaga. Controlled formation of magnetite crystal by partial oxidation of ferrous hydroxide in the presence of recombinant magnetotactic bacterial protein mms6. *Biomaterials*, 28(35):5381–5389, 2007.

[105] A. Arakaki, F. Masuda, Y. Amemiya, T. Tanaka, and T. Matsunaga. Control of the morphology and size of magnetite particles with peptides mimicking the mms6 protein from magnetotactic bacteria. *Journal of Colloid and Interface Science*, 343(1):65–70, 2010.

[106] M.D. Alcala, J.M. Criado, C. Real, T. Grygar, M. Nejezchleba, J. Subrt, and E. Petrovsky. Synthesis of nanocrystalline magnetite by mechanical alloying of iron and hematite. *Journal of Materials Science*, 39(7):2365–2370, 2004.

[107] I. David and A.J.E. Welch. The oxidation of magnetite and related spinels. constitution of gamma ferric oxide. *Transactions of the Faraday Society*, 52:1642–1650, 1956.

[108] A.E. Regazzoni, G.A. Urrutia, M.A. Blesa, and A.J.G. Maroto. Some observations on the composition and morphology of synthetic magnetites obtained by different routes. *Journal of Inorganic and Nuclear Chemistry*, 43(7):1489–1493, 1981.

[109] P.S. Sidhu, R.J. Gilkes, and A.M. Posner. The synthesis and some properties of co, ni, zn, cu, mn and cd substituted magnetites. *Journal of Inorganic and Nuclear Chemistry*, 40(3):429–435, 1978.

[110] T. Sugimoto and E. Matijevic. Formation of uniform spherical magnetite particles by crystallization from ferrous hydroxide gels. *Journal of Colloid and Interface Science*, 74:227–243, 1980.

[111] Y. Tamaura, T. Yoshida, and T. Katsura. The synthesis of green rust ii (feiiiu1-feiiu2) and its spontaneous transformation into fe3o4. *Bulletin of the Chemical Society of Japan*, 57(9):2411–2416, 1984.

[112] K. Matsuda, M. Sumida, K. Fujita, and S. Mitsuzawa. The control of the particles of magnetite. *Bulletin of the Chemical Society of Japan*, 60(12):4441–4442, 1987.

[113] M Hilgendorff and M Giersig. Synthesis of colloidal magnetic nanoparticles. In *Low-Dimensional Systems: Theory, Preparation, and Some Applications*, pages 151–161. Springer, 2003.

[114] M.F. Casula, Y.W. Jun, D.J. Zaziski, E.M. Chan, A. Corrias, and A.P. Alivisatos. The concept of delayed nucleation in nanocrystal growth demonstrated for the case of iron oxide nanodisks. *Journal of the American Chemical Society*, 128:1675–1682, 2006.

[115] D.C. Golden, D.W. Ming, R.V. Morris, A.J. Brearley, H.V. Lauer, A.H. Treiman, M.E. Zolensky, C.S. Schwandt, G.E. Lofgren, and G.A. McKay. Evidence for exclusively inorganic formation of magnetite in martian meteorite alh84001. *American Mineralogist*, 89(5-6):681–695, 2004.

[116] C. Zhang and P.J.D. Lindan. Towards a first-principles picture of the oxide–water interface. *The Journal of Chemical Physics*, 119(17):9183–9190, 2003.

[117] M.D. Segall, P.J.D. Lindan, M.J. Probert, C.J. Pickard, P.J. Hasnip, S.J. Clark, and M.C. Payne. First-principles simulation: ideas, illustrations and the castep code. *Journal of Physics: Condensed Matter*, 14(11):2717, 2002.

[118] G. Sun, J. Kürti, P. Rajczy, M. Kertesz, J. Hafner, and G. Kresse. Performance of the vienna ab initio simulation package (vasp) in chemical applications. *Journal of Molecular Structure: THEOCHEM*, 624(1):37–45, 2003.

[119] R. Dovesi, B. Civalleri, R. Orlando, C. Roetti, and V.R. Saunders. Ab initio quantum simulation in solid state chemistry. *Reviews in Computational Chemistry*, 21:1, 2005.

[120] S.A. Ojo, L. Whitmore, B. Slater, and C. R.A. Catlow. Understanding nucleation and growth using computer simulation. *Solid state sciences*, 3(7):821–826, 2001.

[121] B.J. Alder and T.E. Wainwright. Studies in molecular dynamics. i. general method. *The Journal of Chemical Physics*, 31(2):459–466, 1959.

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

[123] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1–19, 1995.

[124] D.A. Case, T.E. Cheatham, T. Darden, H. Gohlke, R. Luo, K.M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R.J. Woods. The amber biomolecular simulation programs. *Journal of Computational Chemistry*, 26(16):1668–1688, 2005.

[125] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kal, and K. Schulten. Scalable molecular dynamics with namd. *Journal of Computational Chemistry*, 26(16):1781–1802, 2005.

[126] I.T. Todorov, N.L. Allan, M.Y. Lavrentiev, C.L. Freeman, C.E. Mohn, and J.A. Purton. Simulation of mineral solid solutions at zero and high pressure using lattice statics, lattice dynamics and monte carlo methods. *Journal of Physics: Condensed Matter*, 16(27):S2751–2770, 2004.

[127] A.Y. Shih, A. Arkhipov, P.L. Freddolino, and K. Schulten. Coarse grained protein-lipid model with application to lipoprotein particles. *The Journal of Physical Chemistry B*, 110(8):3674–3684, 2006.

[128] A. Arkhipov, P.L. Freddolino, and K. Schulten. Stability and dynamics of virus capsids described by coarse-grained modeling. *Structure*, 14(12):1767–1777, 2006.

[129] A. Redondo and R. LeSar. Modeling and simulation of biomaterials. *Annual Review of Materials Research*, 34:279–314, 2004.

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

[131] Dorothy M Duffy and John H Harding. Simulation of organic monolayers as templates for the nucleation of calcite crystals. *Langmuir*, 20(18):7630–7636, 2004.

[132] A.N. Cormack, R.J. Lewis, and A.H. Goldstein. Computer simulation of protein adsorption to a material surface in aqueous solution: Biomaterials modeling of a ternary system. *The Journal of Physical Chemistry B*, 108(52):20408–20418, 2004.

[133] N.H. De Leeuw, S.C. Parker, and K.H. Rao. Modeling the competitive adsorption of water and methanoic acid on calcite and fluorite surfaces. *Langmuir*, 14(20):5900–5906, 1998.

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

[135] K-P. Schröder, J. Sauer, M. Leslie, C.R.A. Catlow, and J.M. Thomas. Bridging hydrodyl groups in zeolitic catalysts: a computer simulation of their structure, vibrational properties and acidity in protonated faujasites (h-y zeolites). *Chemical Physics Letters*, 188(3):320–325, 1992.

[136] B. Guillot. A reappraisal of what we have learnt during three decades of computer simulations on water. *Journal of Molecular Liquids*, 101(1):219–260, 2002.

[137] Martin Chaplin. Water models. http://www.lsbu.ac.uk/water/models.html, 2014. Accessed: 12/02/2014.

[138] S. Kerisit, S.C. Parker, and J.H. Harding. Atomistic simulation of the dissociative adsorption of water on calcite surfaces. *The Journal of Physical Chemistry B*, 107(31):7676–7682, 2003.

[139] S. Kerisit and S.C. Parker. Free energy of adsorption of water and calcium on the {10 1 [combining macron] 4} calcite surface. *Chemical Communications*, pages 52–53, 2004.

[140] P. Geissbühler, P. Fenter, E. DiMasi, G. Srajer, L.B. Sorensen, and N.C. Sturchio. Three-dimensional structure of the calcite–water interface by surface x-ray scattering. *Surface science*, 573(2):191–203, 2004.

[141] D. Spagnoli, S. Kerisit, and S.C. Parker. Atomistic simulation of the free energies of dissolution of ions from flat and stepped calcite surfaces. *Journal of Crystal Growth*, 294(1):103–110, 2006.

[142] J.S. Lardge, D.M. Duffy, and M.J. Gillan. Investigation of the interaction of water with the calcite (10.4) surface using ab initio simulation. *The Journal of Physical Chemistry C*, 113(17):7207–7212, 2009.

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

[144] H. Pan, J. Tao, T. Wu, and R. Tang. Molecular simulation of water behaviors on crystal faces of hydroxyapatite. *Frontiers of Chemistry in China*, 2(2):156–163, 2007.

[145] E.C. Moreno, M. Kresak, and D.I. Hay. Adsorption of molecules of biological interest onto hydroxyapatite. *Calcified Tissue International*, 36(1):48–59, 1984.

[146] N.V. Nikolenko and E.E. Esajenko. Surface properties of synthetic calcium hydroxyapatite. *Adsorption Science & Technology*, 23(7):543–553, 2005.

[147] M. Corno, C. Busco, V. Bolis, S. Tosoni, and P. Ugliengo. Water adsorption on the stoichiometric (001) and (010) surfaces of hydroxyapatite: a periodic b3lyp study. *Langmuir*, 25(4):2188–2198, 2009.

[148] J. Zeglinski, M. Nolan, D. Thompson, and S.A.M. Tofail. Reassigning the most stable surface of hydroxyapatite to the water resistant hydroxyl terminated (010) surface. *Surface Science*, 2014.

[149] S. Kerisit. Water structure at hematite–water interfaces. *Geochimica et Cosmochimica Acta*, 75(8):2043–2061, 2011.

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

[151] S.C. Parker, N.H. de Leeuw, and S.E. Redfern. Atomistic simulation of oxide surfaces and their reactivity with water. *Journal of the Chemical Society, Faraday Transactions*, 114:381–393, 1999.

[152] F. Jones, A.L. Rohl, J.B. Farrow, and W. Van Bronswijk. Molecular modeling of water adsorption on hematite. *Physical Chemistry Chemical Physics*, 2(14):3209–3216, 2000.

[153] J. Wang and J.R. Rustad. A simple model for the effect of hydration on the distribution of ferrous iron at reduced hematite (0 1 2) surfaces. *Geochimica et Cosmochimica Acta*, 70(21):5285–5292, 2006.

[154] N.H. de Leeuw and T.G. Cooper. Surface simulation studies of the hydration of white rust fe(oh)2, goethite -feo(oh) and hematite -fe2o3. *Geochimica et Cosmochimica Acta*, 71(7):1655–1673, 2007.

[155] S. Kerisit, E.S. Ilton, and S.C. Parker. Molecular dynamics simulations of electrolyte solutions at the (100) goethite surface. *Journal of Physical Chemistry B*, 110(41):20491–20501, 2006.

[156] S. Kerisit and K.M. Rosso. Computer simulation of electron transfer at hematite surfaces. *Geochimica et Cosmochimica Acta*, 70(8):1888–1903, 2006.

[157] D. Spagnoli, D.J. Cooke, S. Kerisit, and S.C. Parker. Molecular dynamics simulations of the interaction between the surfaces of polar solids and aqueous solutions. *Journal of Materials Chemistry*, 16(20):1997–2006, 2006.

[158] D. Spagnoli, B. Gilbert, G.A. Waychunas, and J.F. Banfield. Prediction of the effects of size and morphology on the structure of water around hematite nanoparticles. *Geochimica et Cosmochimica Acta*, 73(14):4023–4033, 2009.

[159] S. Kerisit, D.J. Cooke, D. Spagnoli, and S.C. Parker. Molecular dynamics simulations of the interactions between water and inorganic solids. *Journal of Materials Chemistry*, 15(14):1454–1462, 2005.

[160] J-F. Boily. Water structure and hydrogen bonding at goethite/water interfaces: implications for proton affinities. *The Journal of Physical Chemistry C*, 116(7):4714–4724, 2012.

[161] C-H. Yu, S.Q. Newton, M.A. Norman, L. Schäfer, and D.M. Miller. Molecular dynamics simulations of adsorption of organic compounds at the clay mineral/aqueous solution interface. *Structural Chemistry*, 14(2):175–185, 2003.

[162] U. Magdans, X. Torrelles, K. Angermund, H. Gies, and J. Rius. Crystalline order of a water/glycine film coadsorbed on the (104) calcite surface. *Langmuir*, 23(9):4999–5004, 2007.

[163] B. Zhu, X. Xu, and R. Tang. Hydration layer structures on calcite facets and their roles in selective adsorptions of biomolecules: A molecular dynamics study. *The Journal of chemical physics*, 139(23):234705, 2013.

[164] M. Yang, J.H. Harding, and S.L.S. Stipp. Simulations of monosaccharide on calcite surfaces. *Mineralogical Magazine*, 72(1):295–299, 2008.

[165] M. Yang, P.M. Rodger, J.H. Harding, and S.L.S. Stipp. Molecular dynamics simulations of peptides on calcite surface. *Molecular Simulation*, 35(7):547–553, 2009.

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

[167] C.L. Freeman, J.H. Harding, D. Quigley, and P.M. Rodger. Simulations of ovocleidin-17 binding to calcite surfaces and its implications for eggshell formation. *The Journal of Physical Chemistry C*, 115(16):8175–8183, 2011.

[168] C.L. Freeman, J.H. Harding, D. Quigley, and P.M. Rodger. Protein binding on stepped calcite surfaces: simulations of ovocleidin-17 on calcite {31.16} and {31.8}. *Physical Chemistry Chemical Physics*, 14(20):7287–7295, 2012.

[169] D. Mkhonto and N.H. de Leeuw. A computer modelling study of the effect of water on the surface structure and morphology of fluorapatite: introducing a ca10 (po4) 6f2 potential model. *Journal of Materials Chemistry*, 12(9):2633–2642, 2002.

[170] R. Rohanizadeh, M. Trecant-Viana, and G. Daculsi. Ultrastructural study of apatite precipitation in implanted calcium phosphate ceramic: influence of the implantation site. *Calcified Tissue International*, 64(5):430–436, 1999.

[171] N. Almora-Barrios, K.F. Austen, and N.H. de Leeuw. Density functional theory study of the binding of glycine, proline, and hydroxyproline to the hydroxyapatite (0001) and (0110) surfaces. *Langmuir*, 25(9):5018–5025, 2009.

[172] N. Almora-Barrios and N.H. de Leeuw. A density functional theory study of the interaction of collagen peptides with hydroxyapatite surfaces. *Langmuir*, 26(18):14535–14542, 2010.

[173] N. Almora-Barrios and N.H. de Leeuw. Modelling the interaction of a hyp-pro-gly peptide with hydroxyapatite surfaces in aqueous environment. *CrystEngComm*, 12(3):960–967, 2010.

[174] Neyvis Almora-Barrios and Nora H De Leeuw. Molecular dynamics simulation of the early stages of nucleation of hydroxyapatite at a collagen template. *Crystal Growth & Design*, 12(2):756–763, 2012.

[175] J.D. Prockop. Collagens: molecular biology, diseases, and potentials for therapy. *Annual Review of Biochemistry*, 64(1):403–434, 1995.

[176] N. Ikawa, T. Kimura, Y. Oumi, and T. Sano. Amino acid containing amorphous calcium phosphates and the rapid transformation into apatite. *Journal of Materials Chemistry*, 19(28):4906–4913, 2009.

[177] S. Koutsopoulos and E. Dalas. Inhibition of hydroxyapatite formation in aqueous solutions by amino acids with hydrophobic side groups. *Langmuir*, 16(16):6739–6744, 2000.

[178] S. Koutsopoulos and E. Dalas. Hydroxyapatite crystallization in the presence of serine, tyrosine and hydroxyproline amino acids with polar side groups. *Journal of crystal growth*, 216(1):443–449, 2000.

[179] S. Koutsopoulos and E. Dalas. Hydroxyapatite crystallization in the presence of amino acids with uncharged polar side groups: Glycine, cysteine, cystine, and glutamine. *Langmuir*, 17(4):1074–1079, 2001.

[180] A. Rimola, M. Corno, C.M. Zicovich-Wilson, and P. Ugliengo. Ab initio modeling of protein/biomaterial interactions: glycine adsorption at hydroxyapatite surfaces. *Journal of the American Chemical Society*, 130(48):16181–16183, 2008.

[181] D.K. Dubey and V. Tomar. Role of the nanoscale interfacial arrangement in mechanical strength of tropocollagen-hydroxyapatite-based hard biomaterials. *Acta biomaterialia*, 5(7):2704–2716, 2009.

[182] A. Rimola, M. Aschi, R. Orlando, and P. Ugliengo. Does adsorption at hydroxyapatite surfaces induce peptide folding? insights from large-scale b3lyp calculations. *Journal of the American Chemical Society*, 134(26):10899–10910, 2012.

[183] S. Hug, G.K. Hunter, H.A. Goldberg, and M. Karttunen. ¡ i¿ ab initio¡/i¿ simulations of peptide-mineral interactions. *Physics Procedia*, 4:51–60, 2010.

[184] B. Grohe, J. O'Young, D.A. Ionescu, G. Lajoie, K.A. Rogers, M. Karttunen, H.A. Goldberg, and G.K. Hunter. Control of calcium oxalate crystal growth by face-specific adsorption of an osteopontin phosphopeptide. *Journal of the American Chemical Society*, 129(48):14946–14951, 2007.

[185] J. OYoung, S. Chirico, N. Al Tarhuni, B. Grohe, M. Karttunen, H.A. Goldberg, and G.K. Hunter. Phosphorylation of osteopontin peptides mediates adsorption to and incorporation into calcium oxalate crystals. *Cells Tissues Organs*, 189(1-4):51–55, 2008.

[186] A. Langdon, G.R. Wignall, K. Rogers, E.S. Sorensen, J. Denstedt, B. Grohe, H.A. Goldberg, and G.K. Hunter. Kinetics of calcium oxalate crystal growth in the presence of osteopontin isoforms: An analysis by scanning confocal interference microcopy. *Calcified Tissue International*, 84(3):240–248, 2009.

[187] K. Wang, Y. Leng, X. Lu, F. Ren, and X. Ge. Study of protein adsorption on octacalcium phosphate surfaces by molecular dynamics simulations. *Journal of Materials Science: Materials in Medicine*, 23(4):1045–1053, 2012.

[188] X.D. Zhu, H.S. Fan, C.Y. Zhao, J. Lu, T. Ikoma, J. Tanaka, and X.D. Zhang. Competitive adsorption of bovine serum albumin and lysozyme on characterized calcium phosphates by polyacrylamide gel electrophoresis method. *Journal of Materials Science: Materials in Medicine*, 18(11):2243–2249, 2007.

[189] J. Mazo-Zuluaga and J. Restrepo. Monte carlo study of the bulk magnetic properties of magnetite. *Physica B: Condensed Matter*, 354(1):20–26, 2004.

[190] J.M. Florez, J. Mazo-Zuluaga, and J. Restrepo. Ferrimagnetic to paramagnetic transition in magnetite: Mössbauer versus monte carlo. *Hyperfine interactions*, 161(1-4):161–169, 2005.

[191] J.R. Rustad, E. Wasserman, and A.R. Felmy. A molecular dynamics investigation of surface reconstruction on magnetite (001). *Surface Science*, 432(1):L583–L588, 1999.

[192] C. Soontrapa and Y. Chen. Molecular dynamics potentials in magnetite (fe_3o_4) modeling. *Computational Materials Science*, 50(12):3271–3281, 2011.

[193] J.R. Rustad, A.R. Felmy, and E.J. Bylaska. Molecular simulation of the magnetite-water interface. *Geochimica et cosmochimica acta*, 67(5):1001–1016, 2003.

[194] T.K. Kundu, K. H. Rao, and S.C. Parker. Atomistic simulation studies of magnetite surface structures and adsorption behavior in the presence of molecular and dissociated water and formic acid. *Journal of colloid and interface science*, 295(2):364–373, 2006.

[195] E. Tombacz, A. Hajdu, E. Illes, K. Laszlo, G. Garberoglio, and P. Jedlovszky. Water in contact with magnetite nanoparticles, as seen from experiments and computer simulations. *Langmuir*, 25(22):13007–13014, 2009.

[196] M.E. Grillo, M.W. Finnis, and W. Ranke. Surface structure and water adsorption on fe 3 o 4 (111): Spin-density functional theory and on-site coulomb interactions. *Physical Review B*, 77(7):075407, 2008.

[197] X. Yu, Y. Li, Y-W. Li, J. Wang, and H. Jiao. Dft+u study of molecular and dissociative water adsorptions on the fe3o4 (110) surface. *The Journal of Physical Chemistry C*, 117(15):7648–7655, 2013.

[198] J. Yue, X. Jiang, and A. Yu. Molecular dynamics study on au/fe3o4 nanocomposites and their surface function toward amino acids. *The Journal of Physical Chemistry B*, 115(40):11693–11699, 2011.

[199] A. Bürger, U Magdans, and H. Gies. Adsorption of aminoacids on the magnetite-(111)-surface: a force field study. *Journal of Molecular Modeling*, 19(2):851–857, 2013.

[200] K. Grünberg, E-C. Müller, A. Otto, R. Reszka, D. Linder, M. Kube, R. Reinhardt, and D. Schüler. Biochemical and proteomic analysis of the magnetosome membrane in magnetospirillum gryphiswaldense. *Applied and environmental microbiology*, 70(2):1040–1050, 2004.

[201] M. Born and K. Huang. *Dynamical theory of crystal lattices.* Oxford University Press, 1954.

[202] P. P. Ewald. Die berechnung optisher und elektostatischer gitterpotentiale. *Annalen der Physik*, 64:253–287, 1921.

[203] D. Frenkel and B. Smit. *Understanding Molecular Simulation: From Algorithms to Applications.* Computational science. Elsevier Science, 2001.

[204] D.E. Parry. The electrostatic potential in the surface region of an ionic crystal. *Surface Science*, 49(2):433–440, 1975.

[205] D.E. Parry. Errata. *Surface Science*, 54:195, 1976.

[206] D.M. Heyes, M. Barber, and J.H.R. Clarke. Molecular dynamics computer simulation of surface properties of crystalline potassium chloride. *Journal of the Chemical Society, Faraday Transactions 2*, 73(7):1485–1496, 1977.

[207] B.G. Dick and A.W. Overhauser. Theory of the dielectric constants of alkali halide crystals. *Physical Review Letters*, 112(1):90–103, 1958.

[208] G.W. Watson, S.C. Parker, and A. Wall. Molecular dynamics simulation of fluoride-perovskites. *Journal of Physics: Condensed Matter*, 4:2097–2108, 1992.

[209] N.L. Allan, A.L. Rohl, D.H. Gay, C.R.A. Catlow, R.J. Davey, and W.C. Mackrodt. Calculated bulk and surface properties of sulfates. *Faraday Discussions*, 95:273–280, 1993.

[210] P.M. Oliver, S.C. Parker, J. Purton, and D.W. Bullett. Atomistic simulations and electronic structure of tio2 (100) surfaces. *Surface Science*, 307:1200–1205, 1994.

[211] G.V. Paolini, P.J.D. Lindan, and J.H. Harding. The thermal conductivity of defective crystals. *Journal of Chemical Physics*, 106(9):3681–3687, 1997.

[212] R. Takagi, F. Hutchinson, P.A. Madden, A.K. Adya, and M. Gaune-Escard. The structure of molten dycl3 and dyna3cl6 simulated with polarizable- and rigid-ion models. *Journal of Physics: Condensed Matter*, 11(3):645–658, 1999.

[213] D.C. Sayle, S.A. Maicaneanu, and G.W. Watson. Synthesis of a bao thin film supported on mgo(001) using a shell-model amorphisation and recrystallisation strategy. *Physical Chemistry Chemical Physics*, 4(20):5189–5198, 2002.

[214] H.R. Horton. *Principles of biochemistry.* Pearson Prentice Hall, 2006.

[215] W.D. Cornell, P. Cieplak, C.I. Bayly, I.R. Gould, K.M. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, and P.A. Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117(19):5179–5197, 1995.

[216] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, and D.A. Case. Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9):1157–1174, 2004.

[217] R.J. Woods, R.A. Dwek, C.J. Edge, and B. Fraser-Reid. Molecular mechanical and molecular dynamic simulations of glycoproteins and oligosaccharides. 1. *glycam*$_9$3 parameter development. *Journal of Physical Chemistry*, 99(11):3832–3846, 1995.

[218] D.A. Case, T.A. Darden, I.T.E. Cheatham, C.L. Simmerling, J. Wang, R.E. Duke, R. Luo, R.C. Walker, W. Zhang, and K.M. Merz. Ambertools version 1.4, 2010. *University of California, San Francisco, USA*, 2010.

[219] G.V. Lewis and C.R.A. Catlow. Potential models for ionic oxides. *Journal of Physics C: Solid State Physics*, 18:1149, 1985.

[220] J.R. Rustad. *Molecular Modeling Theory: Applications in the Geosciences.*, chapter 6, pages 169–198. Mineralogical Society of America, 2001.

[221] R.T. Cygan, J-J. Liang, and A.G. Kalinichev. Molecular models of hydroxide, oxyhydroxide, and clay phases and the development of a general force field. *Journal of Physical Chemistry B*, 108(4):1255–1266, 2004.

[222] M.E. Fleet. The structure of magnetite: Symmetry of cubic spinels. *Journal of Solid State Chemistry*, 62(1):75–82, 1986.

[223] Y. Wu, H.L. Tepper, and G.A. Voth. Flexible simple point-charge water model with improved liquid-state properties. *The Journal of Chemical Physics*, 124:024503, 2006.

[224] M.P. Allen and D.J. Tildesley. *Computer simulation of liquids.* Oxford science publications. Clarendon Press, 1989.

[225] A.R. Leach. *Molecular modelling: principles and applications.* Pearson Education. Pearson, Prentice Hall, 2001.

[226] C.R.A. Catlow, K.M. Diller, and M.J. Norgett. Interionic potentials for alkali halides. *Journal of Physics C: Solid State Physics*, 10:1395, 1977.

[227] D.J. Harris, J.H. Harding, and S.C. Parker. Simulations of surfaces and interfaces in mgo. *Radiation Effects and Defects in Solids*, 151(1-4):299–304, 1999.

[228] R.W. Impey, P.A. Madden, and I.R. McDonald. Hydration and mobility of ions in solution. *The Journal of Physical Chemistry*, 87(25):5071–5083, 1983.

[229] R.J. Kirkpatrick, A.G. Kalinichev, and J. Wang. Molecular dynamics modelling of hydrated mineral interlayers and surfaces: structure and dynamics. *Mineralogical Magazine*, 69(3):289, 2005.

[230] R. Fletcher and C.M. Reeves. Function minimization by conjugate gradients. *The Computer Journal*, 7(2):149–154, 1964.

[231] M.J Norgett and R. Fletcher. Fast matrix methods for calculating the relaxation about defects in crystals. *Journal of Physics C: Solid State Physics*, 3(11):L190, 1970.

[232] R. Fletcher and M.J.D. Powell. A rapidly convergent descent method for minimization. *The Computer Journal*, 6(2):163–168, 1963.

[233] G.W. Watson, E.T. Kelsey, N.H. de Leeuw, D.J. Harris, and S.C. Parker. Atomistic simulation of dislocations, surfaces and interfaces in mgo. *Journal of the Chemical Society, Faraday Transactions*, 92(3):433, 1996.

[234] C.R.A. Catlow. *Computer modelling in inorganic crystallography*. Academic Press, 1997.

[235] W. Smith and T.R. Forester. Dl_poly is a package of molecular simulation routines, copyright the council for the central laboratory of the research councils, daresbury laboratory at daresbury, nr. *Nr. Warrington*, 1996.

[236] L. Verlet. Computer "experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Physical Review Letters*, 159:98–103, 1967.

[237] R.W. Hockney. *Methods in Computational Physics:The potential calculation and some applications*, volume 9. Academic Press, 1970.

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

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

[240] D. Frenkel and B. Smit. *Understanding molecular simulation: from algorithms to applications*. Academic Press, 2nd edition, 2002.

[241] D.J. Cooke. *Atomistic simulation of segregation to the surfaces of hematite and goethite*. PhD thesis, University of Bath, 2003.

[242] P.W. Tasker. The stability of ionic crystal surfaces. *Journal of Physics C: Solid State Physics*, 12(22):4977, 1979.

[243] F. Bertaut. Le terme electrostatique de lenergie de surface. *Comptes Rendus Hebdomadaires Des Seances De L' Acadamie Des Sciences*, 246(25):3447–3450, 1958.

[244] P.M. Oliver, S.C. Parker, and W.C. Mackrodt. Computer simulation of the crystal morphology of nio. *Modelling and Simulation in Materials Science and Engineering*, 1(5):755, 1993.

[245] M. Ma, Y. Zhang, W. Yu, H-Y Shen, H-Q Zhang, and N. Gu. Preparation and characterization of magnetite nanoparticles coated by amino silane. *Colloids and Surfaces A: physicochemical and engineering aspects*, 212(2):219–226, 2003.

[246] S-L. Tie, Y-Q. Lin, H-C. Lee, Y-S. Bae, and C-H. Lee. Amino acid-coated nano-sized magnetite particles prepared by two-step transformation. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 273(1):75–83, 2006.

[247] G. Marinescu, L. Patron, D.C. Culita, C. Neagoe, C.I. Lepadatu, I. Balint, L. Bessais, and C.B. Cizmas. Synthesis of magnetite nanoparticles in the presence of aminoacids. *Journal of Nanoparticle Research*, 8(6):1045–1051, 2006.

[248] K. Aurich, M. Schwalbe, J.H. Clement, W. Weitschies, and N. Buske. Polyaspartate coated magnetite nanoparticles for biomedical applications. *Journal of Magnetism and Magnetic Materials*, 311(1):1–5, 2007.

[249] J.Y. Park, E.S. Choi, M.J. Baek, and G.H. Lee. Colloidal stability of amino acid coated magnetite nanoparticles in physiological fluid. *Materials Letters*, 63(3):379–381, 2009.

[250] J.L. Viota, F.J. Arroyo, A.V. Delgado, and J. Horno. Electrokinetic characterization of magnetite nanoparticles functionalized with amino acids. *Journal of Colloid and Interface Science*, 344(1):144–149, 2010.

[251] T.L. Hill. *Statistical Mechanics: Principles and Selected Applications*. McGraw-Hill series in advanced chemistry. McGraw-Hill, 1956.

[252] S. A. Staniland, A. Rawlings, and J. Bramble. private communication, 2011.

[253] Lijun Wang, Tanya Prozorov, Pierre E Palo, Xunpei Liu, David Vaknin, Ruslan Prozorov, Surya Mallapragada, and Marit Nilsen-Hamilton. Self-assembly and biphasic iron-binding characteristics of mms6, a bacterial protein that promotes the formation of superparamagnetic magnetite nanoparticles of uniform size and shape. *Biomacromolecules*, 13(1):98–105, 2011.

[254] A. Rawlings, S. A. Staniland, and J. Bramble. private communication, 2011.

Appendix A

Appendix

A.1 Chapter 4

Evolution of the potential energy of the amino acid residues as a function of time. The x-axis is the timestep in ps and the y-axis is the potential energy in eV. The red line indicates the *in vacu* system and the blue line represents the solvated system.

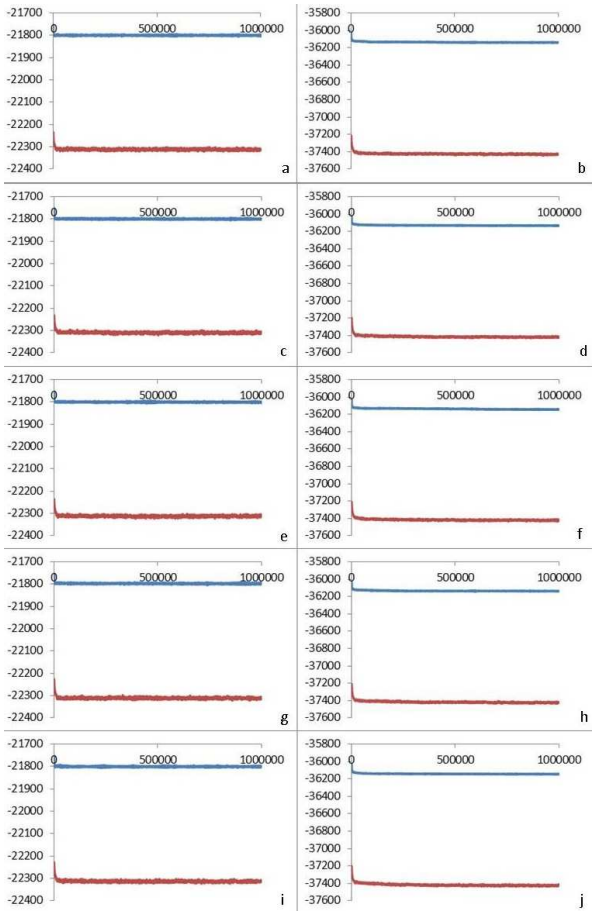

Figure A.1: Non-polar amino acid residues. a) ALA {100}, b) ALA {111}, c) ILE {100}, d) ILE {111}, e) LEU {100}, f) LEU {111}, g) MET {100}, h) MET {111}, i)VAL {100} and j) VAL {111}.

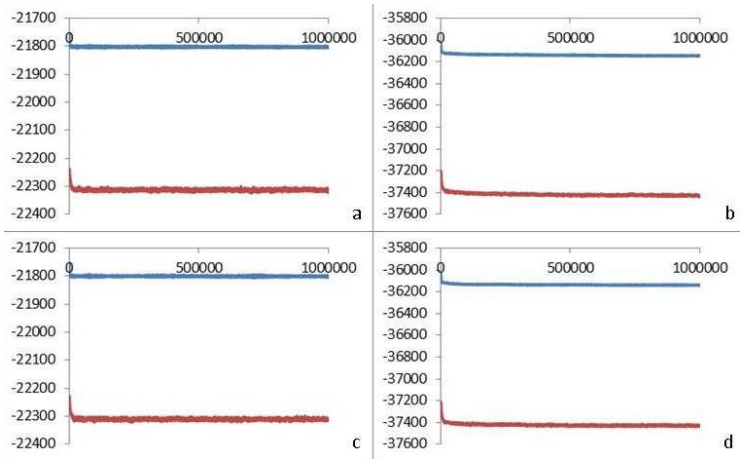

Figure A.2: Polar amino acid residues. a) GLN {100}, b) GLN {111}, c) SER {100} and d) SER {111}.

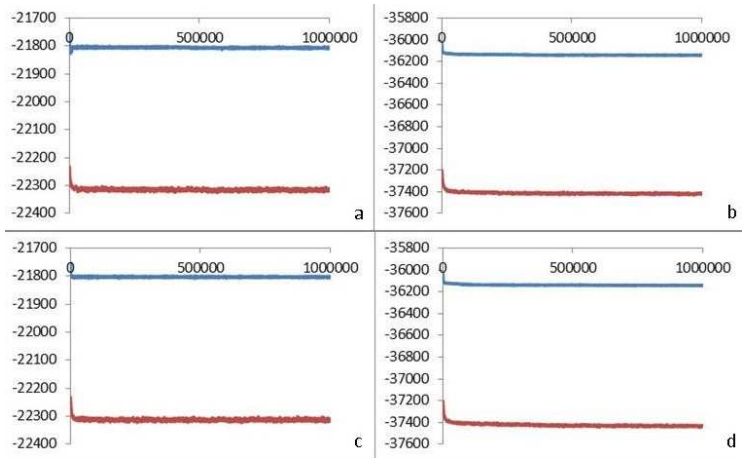

Figure A.3: Acidic amino acid residues. a) ASP {100}, b) ASP {111}, c) GLU {100} and d) GLU {111}.

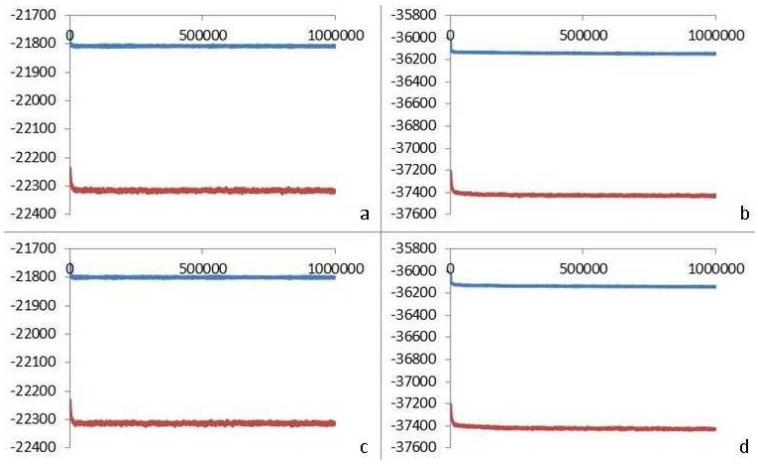

Figure A.4: Basic amino acid residues. a) ARG {100}, b) ARG {111}, c) LYS {100} and d) LYS {111}.

Radial distribution function profiles for the amino acid attachment to the {100} and {111} surfaces both *in vacu* and solvated. FET is tetrahedral iron, FEO is octahedral iron. Blue is {100} *in vacu*. Red is {100} solvated. Green is {111} *in vacu*. Purple is {111} solvated. The x-axis is the r value (or distance) in Å and the y-axis is the g(r) value.

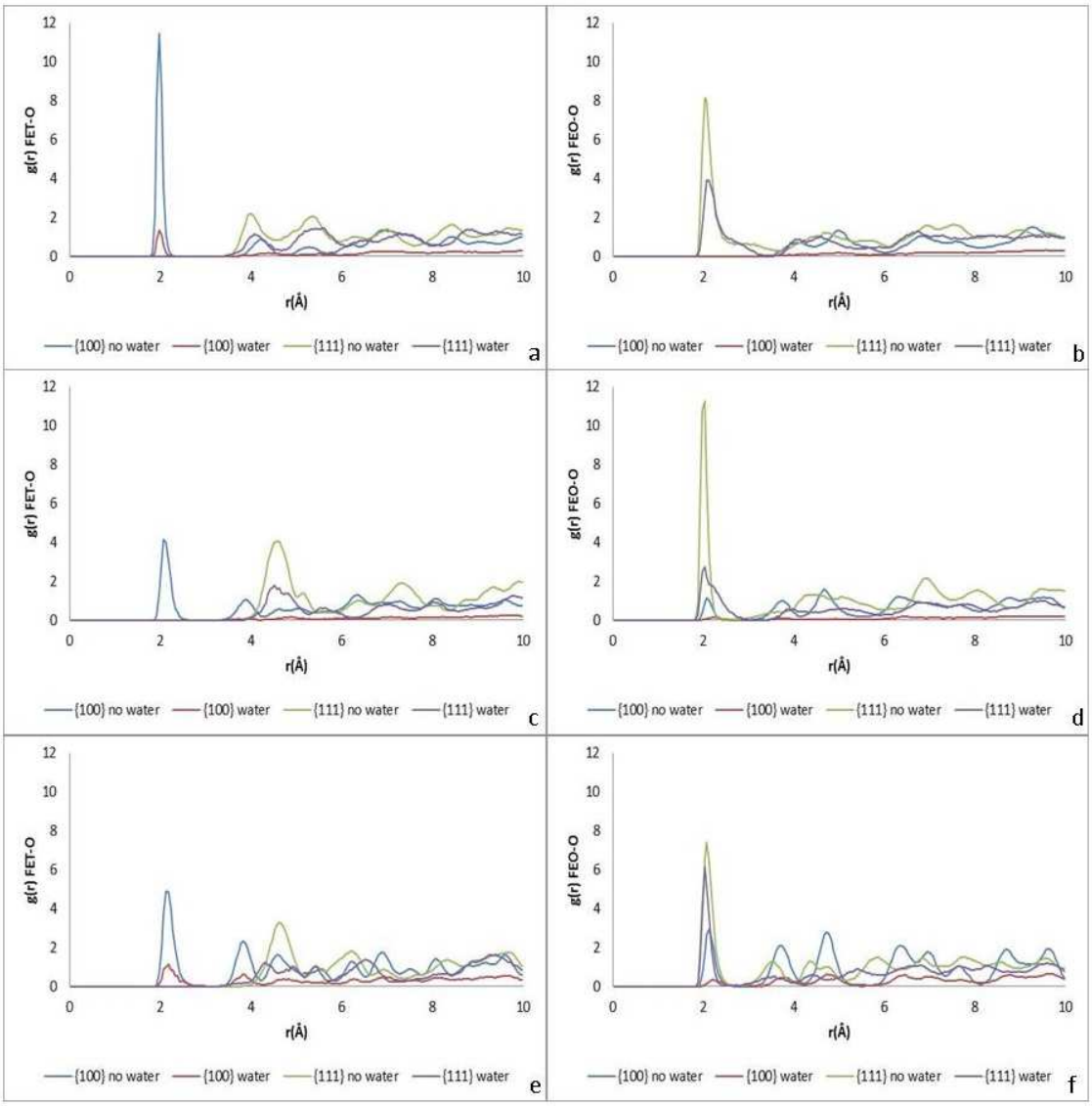

Figure A.5: RDF plots for the non-polar amino acids; a) ALA O-FET, b) ALA O-FEO, c) ILE O-FET, d) ILE O-FEO, e) LEU O-FET and f) LEU O-FEO.

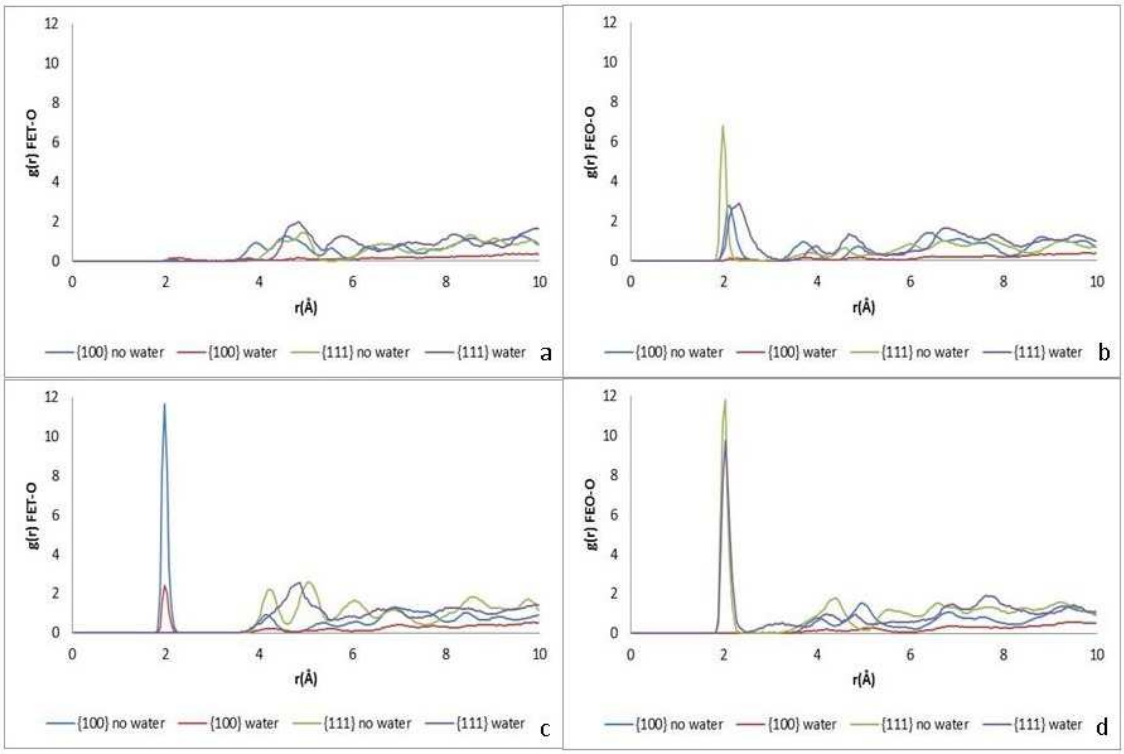

Figure A.6: RDF plots for the non-polar amino acids; a) MET O-FET, b) MET O-FEO, c) VAL O-FET and d) VAL O-FEO iron.

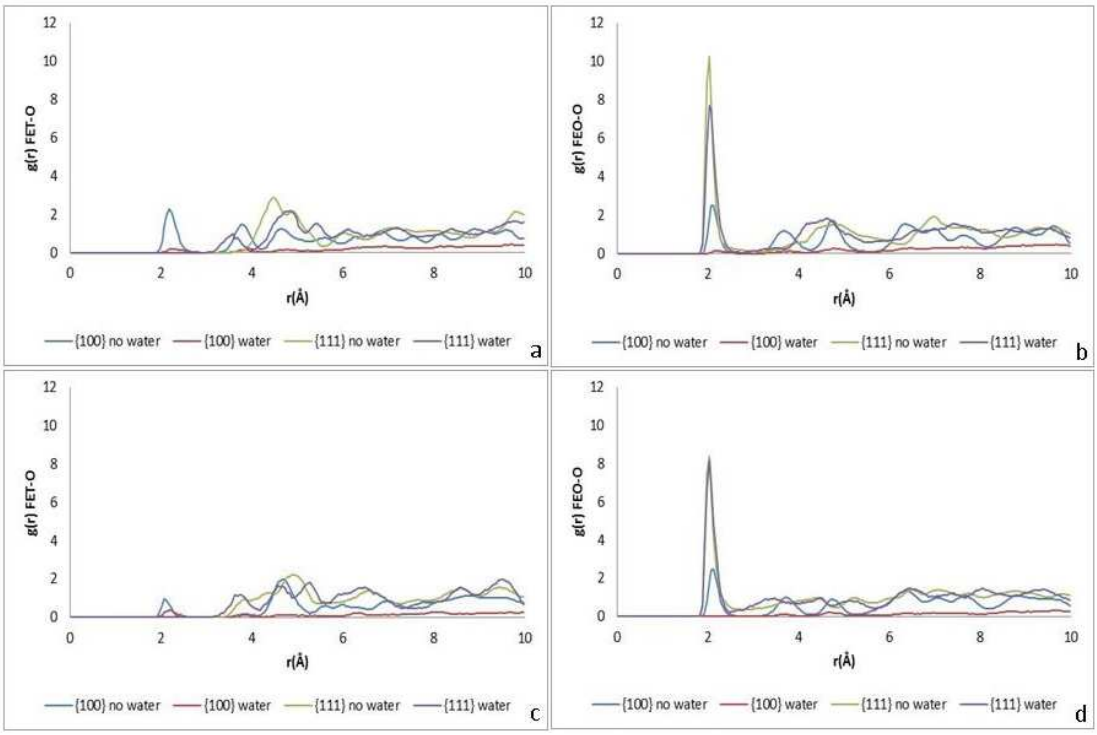

Figure A.7: RDF plots for the polar amino acids; a) GLN O-FET, b) GLN O-FEO, c) SER O-FET, and d) SER O-FEO iron.

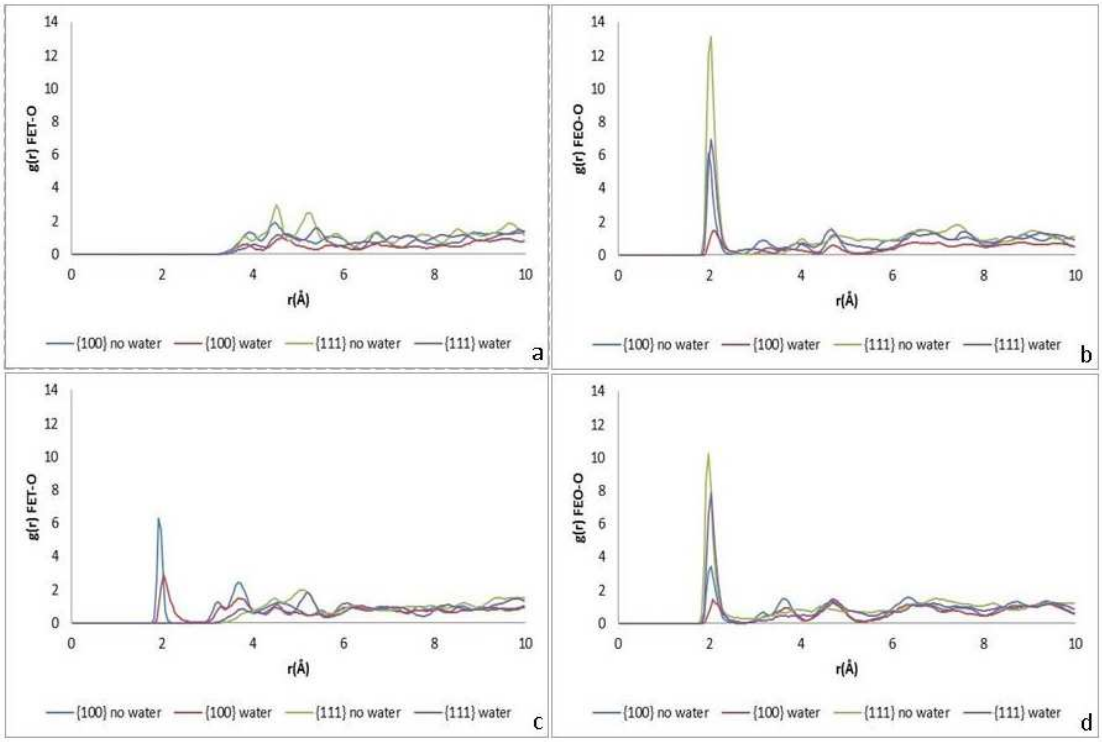

Figure A.8: RDF plots for the acidic amino acids; a) ASP O-FET, b) ASP O-FEO, c) GLU O-FET, and d) GLU O-FEO iron.

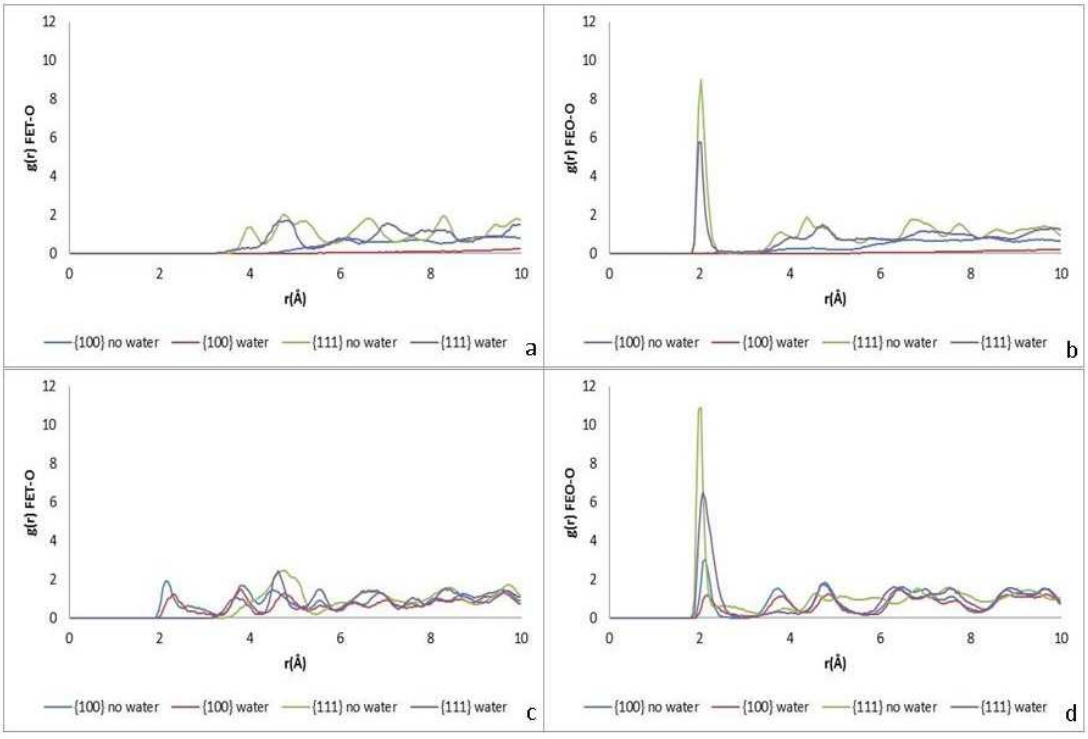

Figure A.9: RDF plots for the basic amino acids; a) ARG O-FET, b) ARG O-FEO, c) LYS O-FET, and d) LYS O-FEO iron.

A.2 Chapter 5

Evolution of the potential energy of the di- and tetrapeptides as a function of time. The x-axis is the timestep in ps and the y-axis is the potential energy in eV. The red line indicates the *in vacu* system and the blue line represents the solvated system.

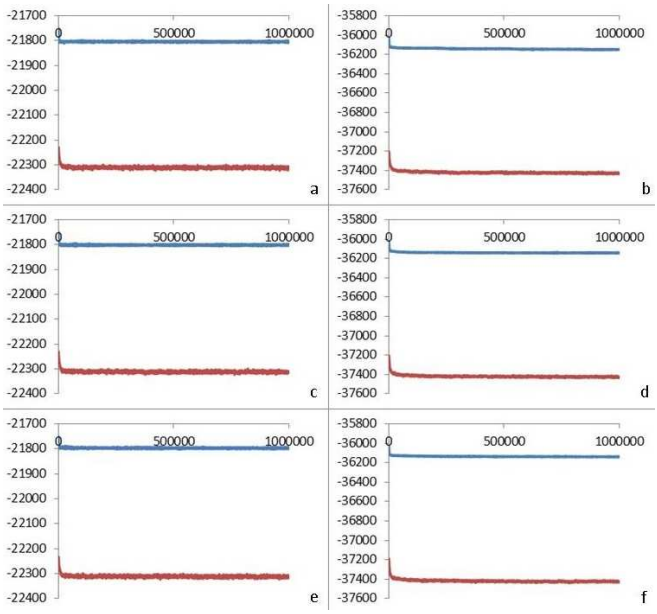

Figure A.10: Dipeptides. a) EE {100}, b) EE {111}, c) AE {100}, d) AE {111}, e) AA {100}, f) AA {111}.

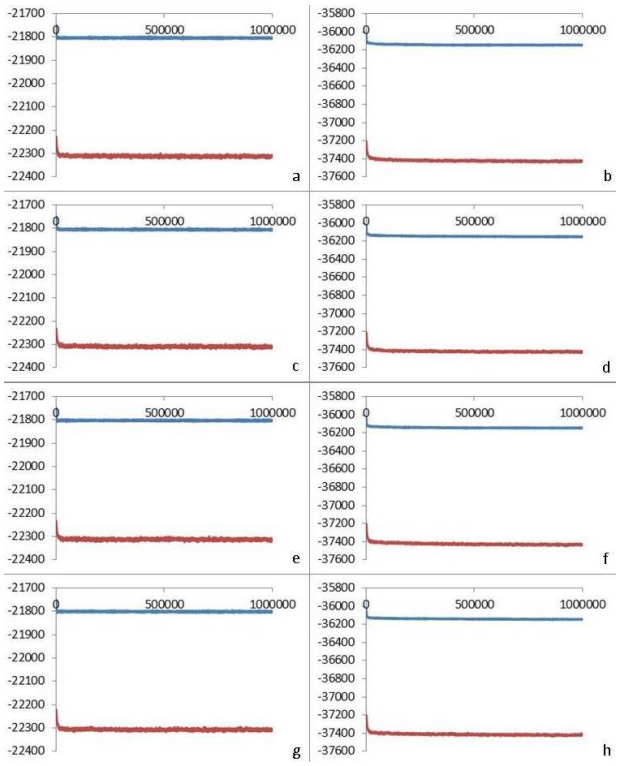

Figure A.11: Tetrapeptides. a) DEEV {100}, b) DEEV {111}, c) DAEV {100}, d) DAEV {111}, e) DEAV {100}, f) DEAV {111}, g) DAAV {100}, h) DAAV {111}.

Radial distribution function profiles for the di- and tetrapeptide attachment to the {100} and {111} surfaces both *in vacu* and solvated. FET is tetrahedral iron, FEO is octahedral iron. Blue is {100} *in vacu*. Red is {100} solvated. Green is {111} *in vacu*. Purple is {111} solvated. The x-axis is the r value (or distance) in Å and the y-axis is the g(r) value.

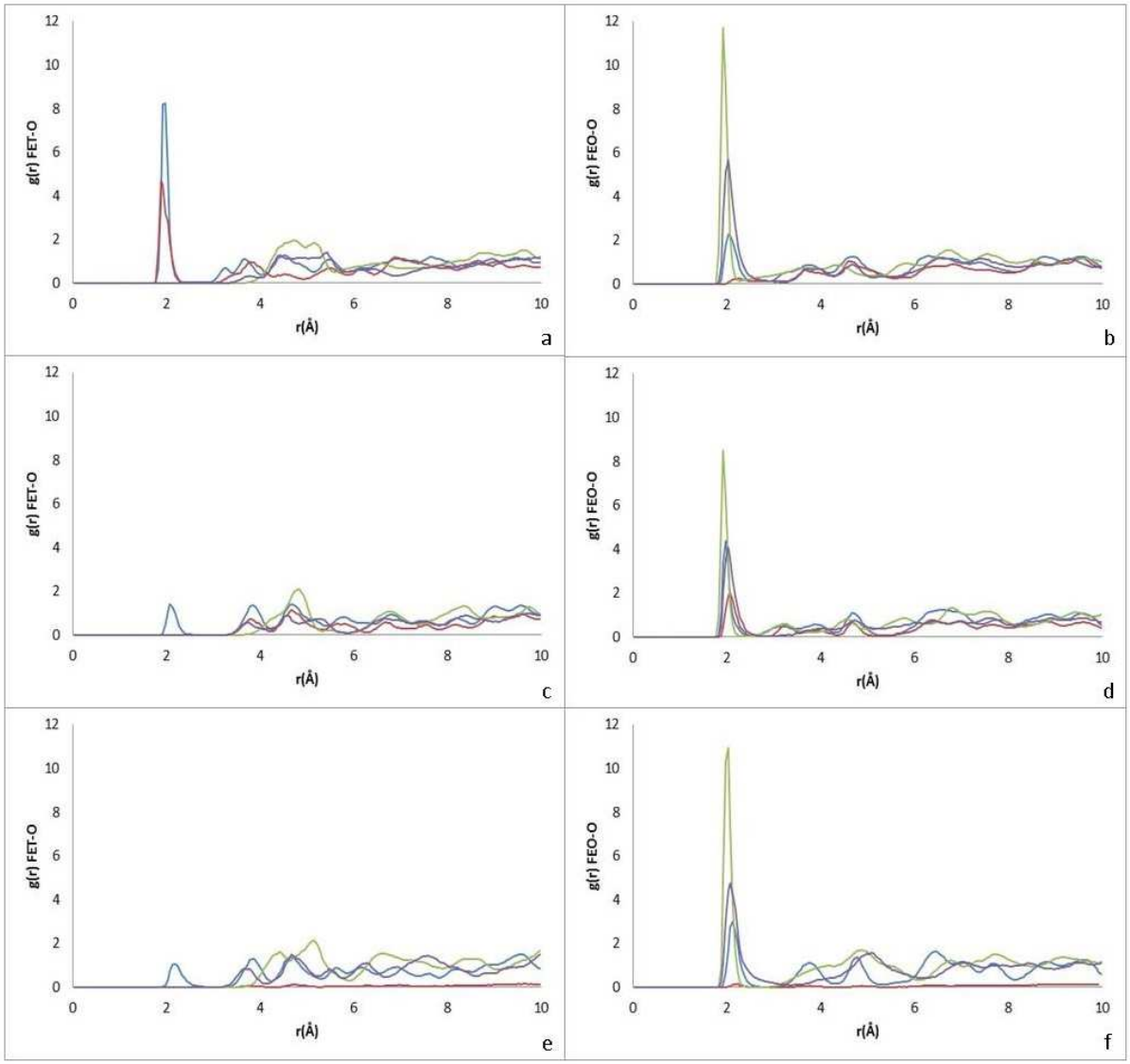

Figure A.12: Radial distribution function plots for the dipeptides; a) EE O-FET, b) EE O-FEO, c) AE O-FET, d) AE O-FEO, e) AA O-FET and f) AA O-FEO iron.

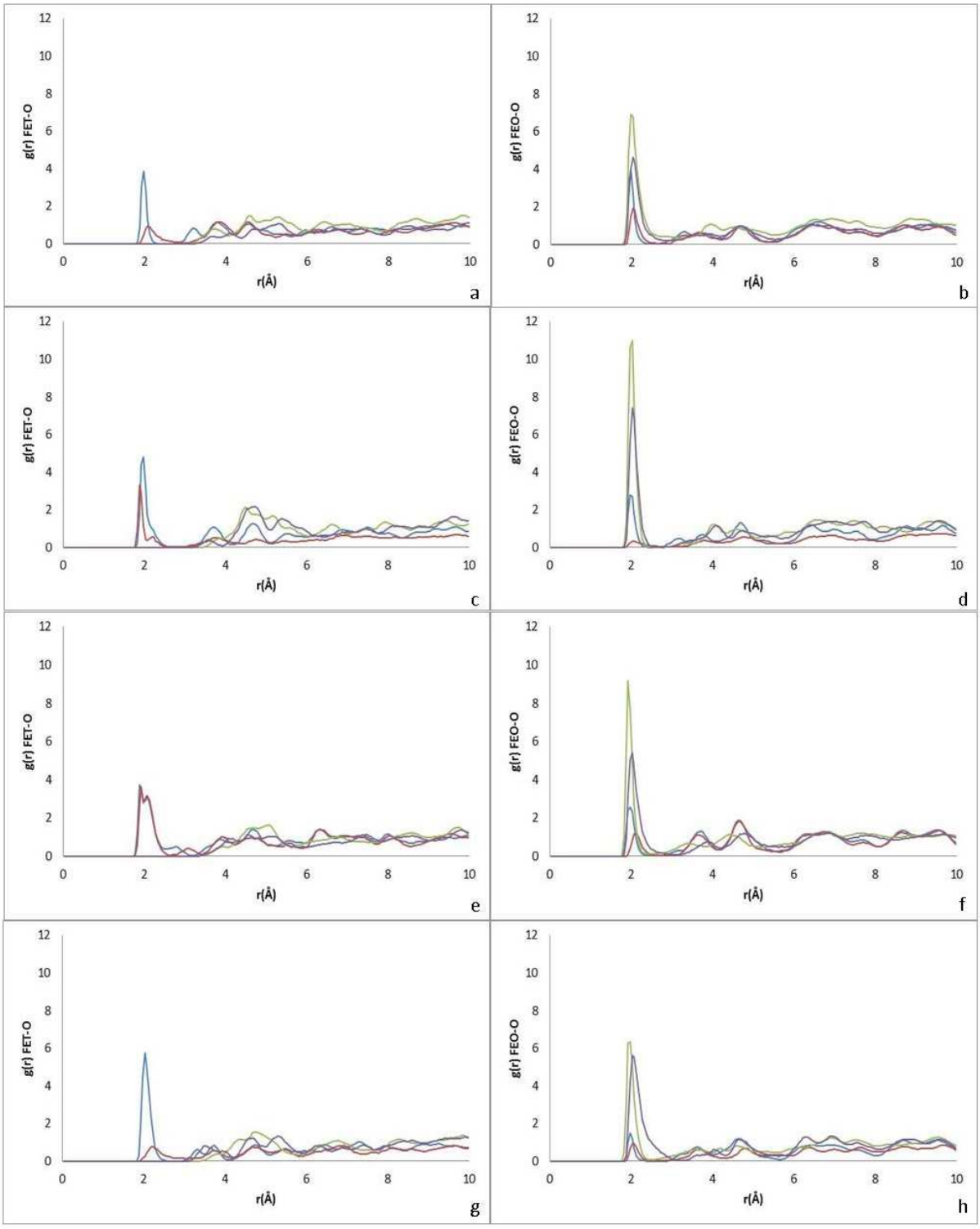

Figure A.13: Radial distribution function plots for the tetrapeptides; a) DEEV O-FET, b) DEEV O-FEO, c) DAEV O-FET, d) DAEV O-FEO, e) DEAV O-FET, f) DEAV O-FEO, g) DAAV O-FET and h) DAAV O-FEO iron.

A.3 Chapter 6

Evolution of the potential energy of the pentapeptides as a function of time. The x-axis is the timestep in ps and the y-axis is the potential energy in eV. The red line indicates the *in vacu* system and the blue line represents the solvated system.

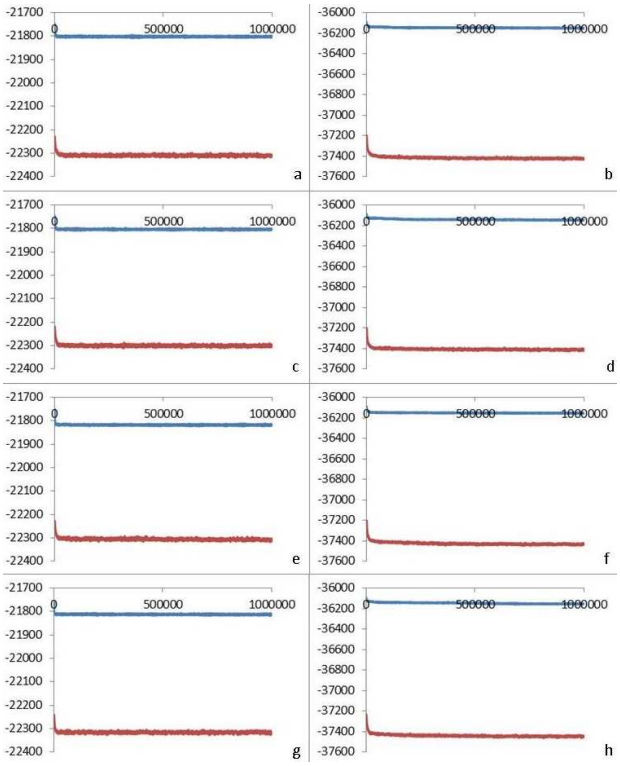

Figure A.14: 177-183 pentapeptides. a) DIESA {100}, b) DIESA {111}, c) DIASA {100}, d) DIASA {111}, e) SRDIE {100}, f) SRDIE {111}, g) SRAIE {100} and h) SRAIE {111}.

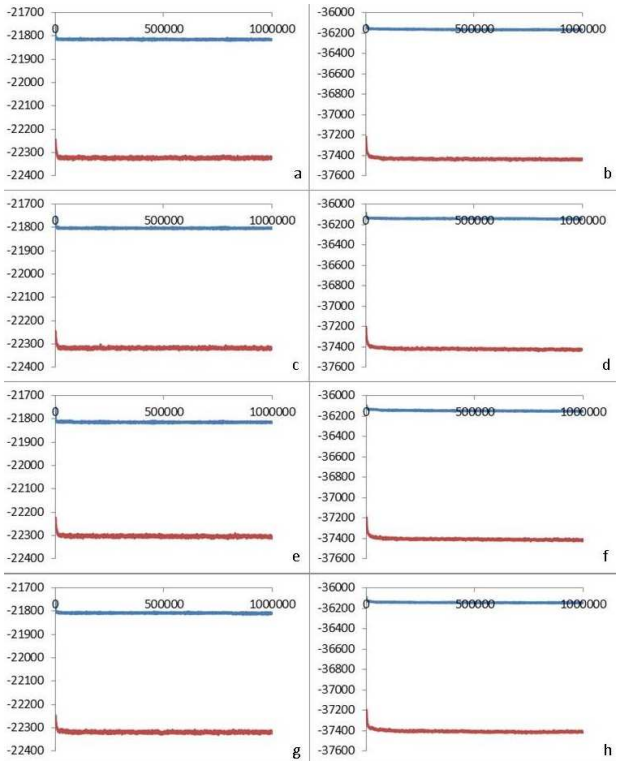

Figure A.15: 190-195 pentapeptides. a) ELRDA {100}, b) ELRDA {111}, c) ELADA {100}, d) ELADA {111}, e) LRDAL {100}, f) LRDAL {111}, g) LRAAL {100} and h) LRAAL {111}.

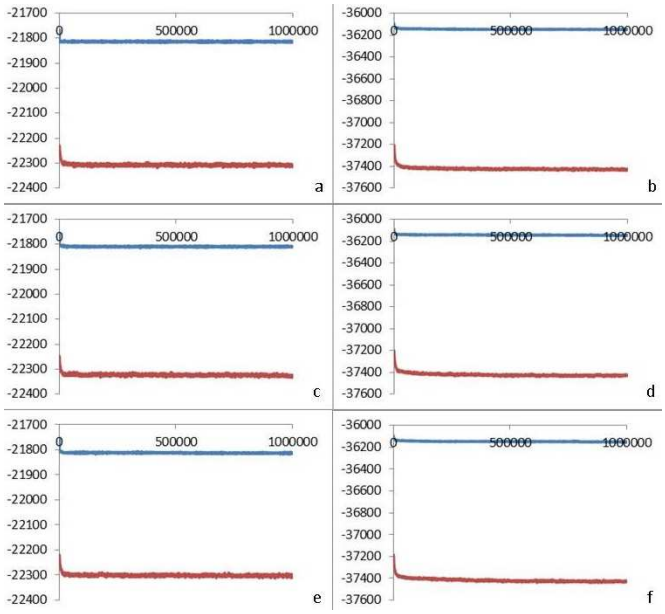

Figure A.16: 188-192 pentapeptides. a) EVELR {100}, b) EVELR {111}, c) EVALR {100}, d) EVALR {111}, e) EVGELR {100} and f) EVGELR {111}.

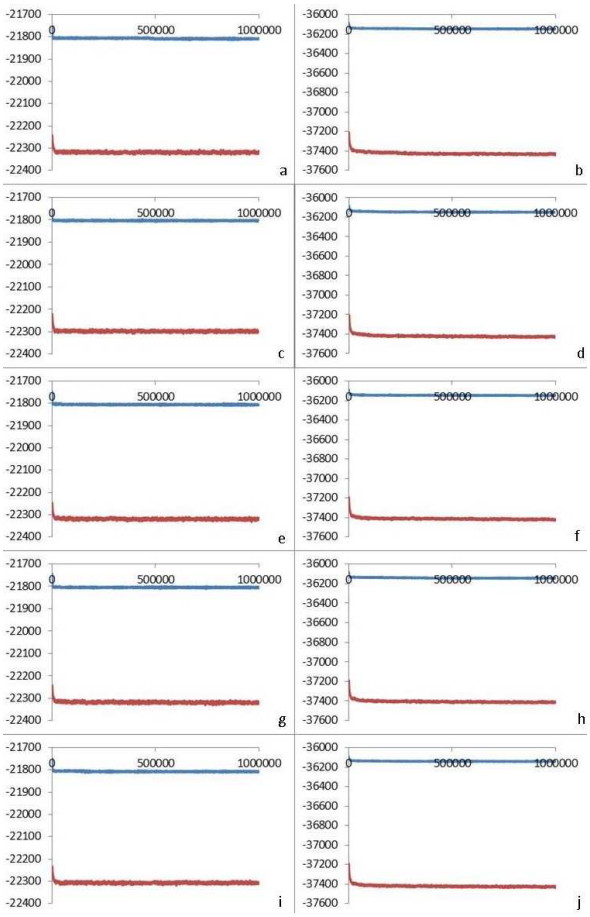

Figure A.17: 185-189 pentapeptides. a) SDEEV {100}, b) SDEEV {111}, c) SAEEV {100}, d) SAEEV {111}, e) SDAEV {100}, f) SDAEV {111}, g) SDEAV {100}, h) SDEAV {111}, i) SDAAV {100} and j) SDAAV {111}.

Radial distribution function profiles for the pentapeptide attachment to the {100} and {111} surfaces both *in vacu* and solvated. FET is tetrahedral iron, FEO is octahedral iron. Blue is {100} *in vacu*. Red is {100} solvated. Green is {111} *in vacu*. Purple is {111} solvated. The x-axis is the r value (or distance) in Å and the y-axis is the g(r) value.

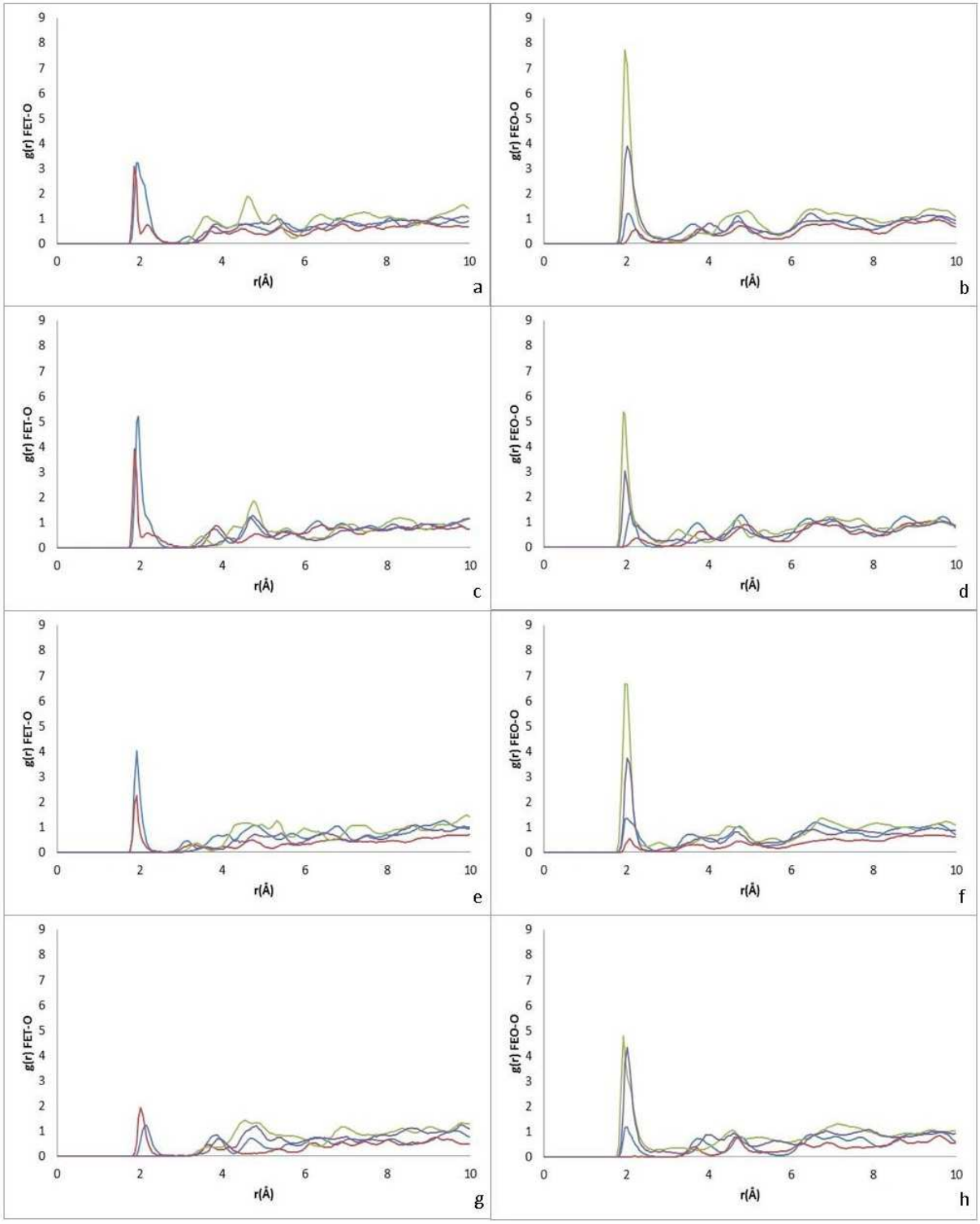

Figure A.18: Radial distribution function plots for the 177-183 pentapeptides; a) DIESA O-FET, b) DIESA O-FEO, c) DIASA O-FET, d) DIASA O-FEO, e) SRDIE O-FET, f) SRDIE O-FEO, g) SRAIE O-FET and h) SRAIE O-FEO iron.

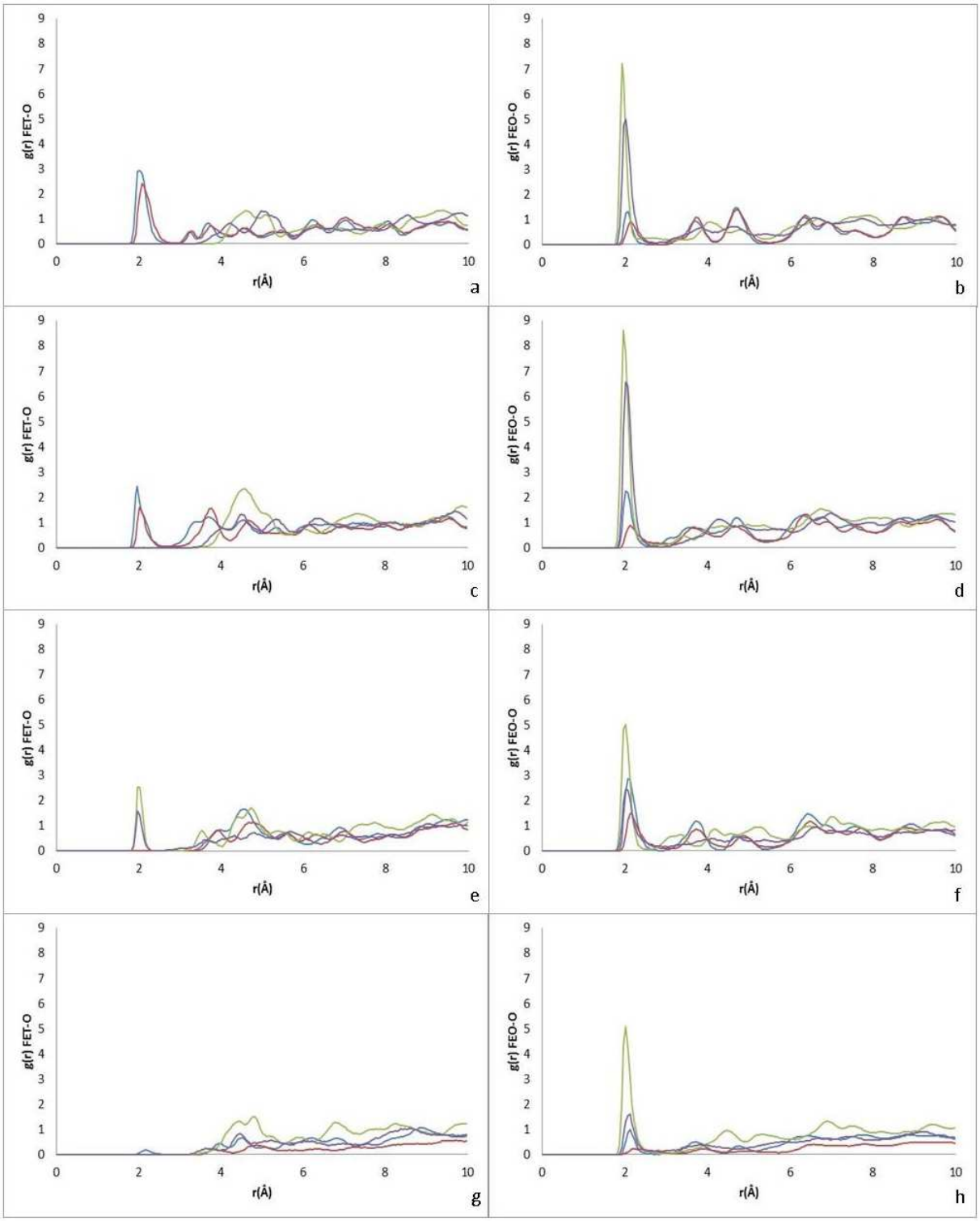

Figure A.19: Radial distribution function plots for the 190-195 pentapeptides; a) ELRDA O-FET, b) ELRDA O-FEO, c) ELADA O-FET, d) ELADA O-FEO, e) LRDAL O-FET, f) LRDAL O-FEO, g) LRAAL O-FET and h) LRAAL O-FEO iron.

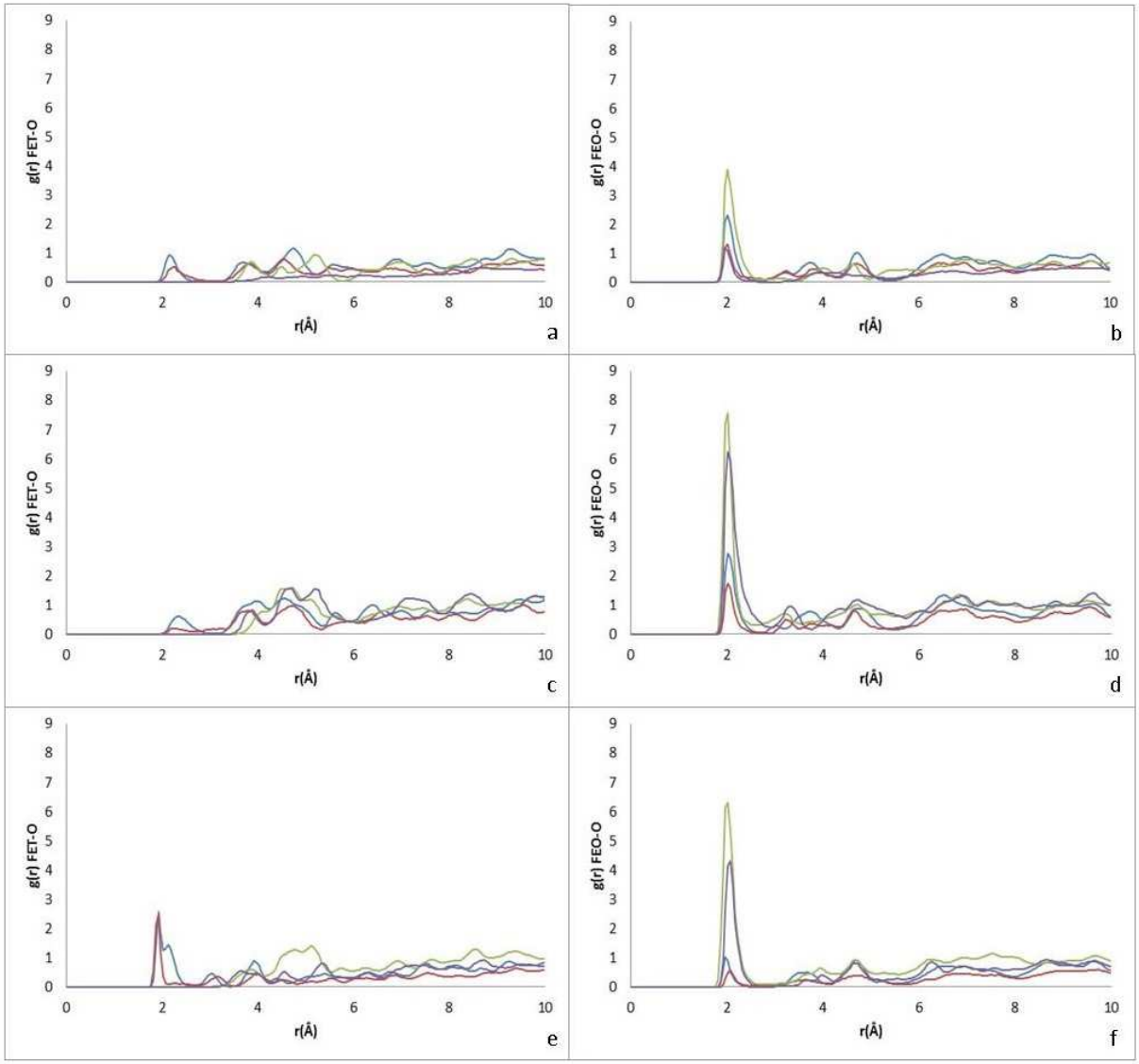

Figure A.20: Radial distribution function plots for the 188-192 pentapeptides; a) EVELR O-FET, b) EVELR O-FEO, c) EVALR O-FET, d) EVALR O-FEO, e) EVGELR O-FET and f) EVGELR O-FEO iron.

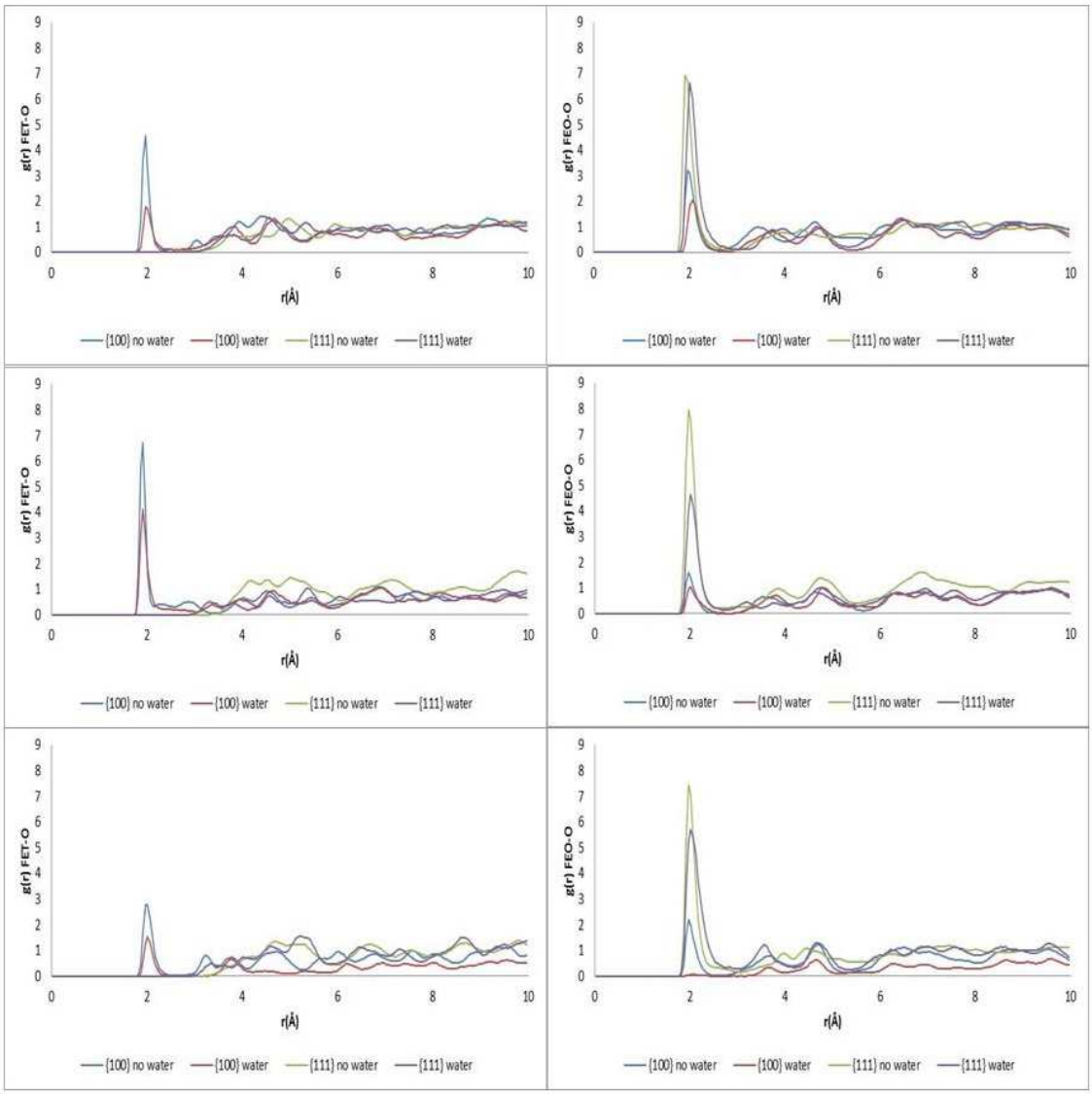

Figure A.21: Radial distribution function plots for the 185-189 pentapeptides; a) SDEEV O-FET, b) SDEEV O-FEO, c) SAEEV O-FET, d) SAEEV O-FEO, e) SDAEV O-FET and f) SDAEV O-FEO.

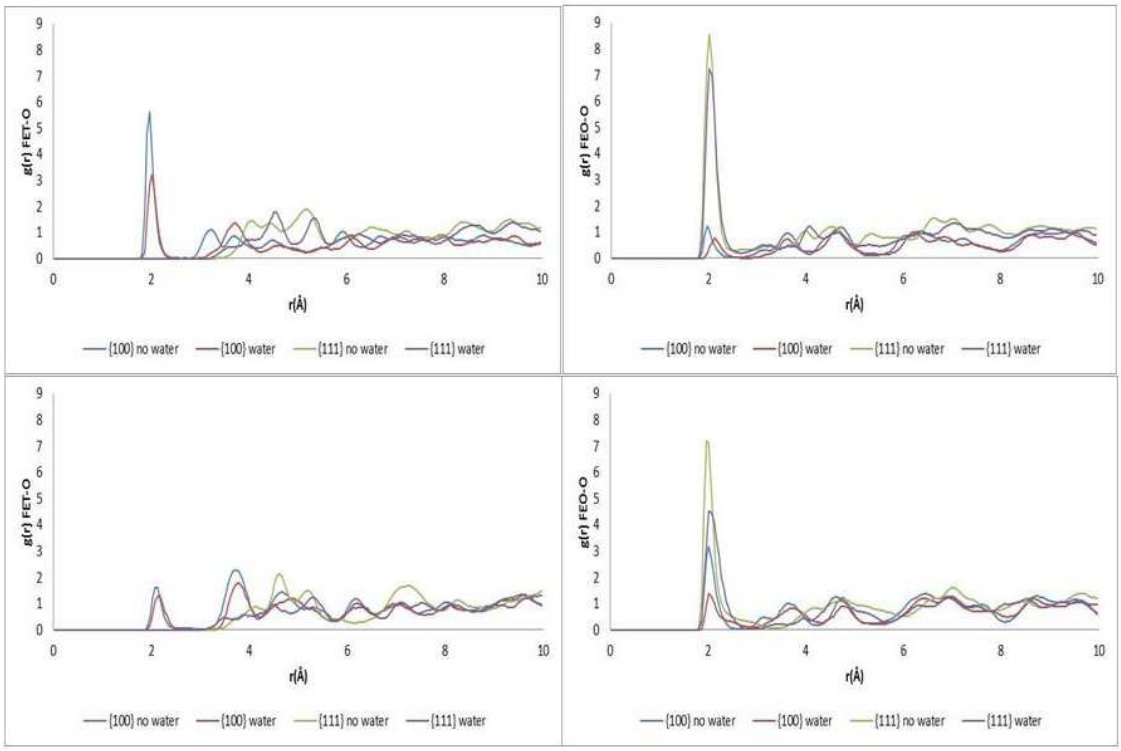

Figure A.22: Radial distribution function plots for the 185-189 pentapeptides; a) SDEAV O-FET, b) SDEAV O-FEO, c) SDAAV O-FET and d) SDAAV O-FEO iron.