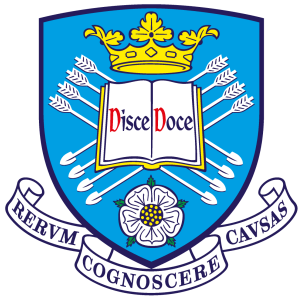



The University Of Sheffield.

# Theoretical Investigations into the Interaction Between Soil Organic Carbon and Soil Minerals

**Aneesa Ahmad**

Supervised by: Dr. N. Martsinovich and Dr. A. Chauvet

A thesis submitted in partial fulfilment of the requirements

for the degree of Doctor of Philosophy

26 November 2021

*"My Lord, increase me in knowledge."*

Holy Quran, 20:114

Dearest Dad, what I would give for you to see this day. I cannot thank you and Mum enough for the love, guidance and sacrifices that you have both made. This thesis is dedicated to you both.

With love.

# Abstract

Organic carbon (OC) is an essential component of soil; its presence in soil is important for growing crops, for maintaining the stability of soils, and more generally for capturing $CO_2$ from the atmosphere. Sorption of OC to oxide mineral surfaces is considered a key process in the soil preservation due to its ability to protect OC from microbial degradation. To understand the sorption of OC in soils, it is important to obtain a quantitative description of the binding of organic molecules to soil minerals and soil metal cations. In this work, we conducted theoretical modelling investigations on the binding of water and small organic molecules, typical building blocks of OC, on $\alpha$-$Al_2O_3$ (alumina), a common soil mineral. Alumina was modelled both bare and hydroxylated using (0001)-oriented periodic slabs, using density functional theory (DFT) calculations and empirical dispersion. Adsorption energies and binding energies were calculated. The amine, amide and carboxylic acid functional groups were found to adsorb more strongly than water during our small molecule adsorption investigations on the bare and hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. These investigations illustrated the key functional groups required for strong OC binding in soils. With knowledge gained from this first stage of investigations, we conducted our biopolymer adsorption investigations. Several biopolymers such as, cellulose and chitin were adsorbed to the hydroxylated alumina surface. Pectin was the most strongly adsorbing of the biopolymers and hydrogen bonding was a major contributor to its stability. Finally, we also investigated the binding of small organic molecules to the hydrated $[Al(H_2O)_6]^{3+}$ and $[Fe(H_2O)_6]^{3+}$ complexes to investigate the binding of cations in the soil solution. In the cation investigations we calculated the Gibbs energy of reaction for the products of the ligand substitution process. Acetate, amine and amide ligands yielded the most energetically favourable substitution products for both the $Al^{3+}$ and $Fe^{3+}$ studies. $Fe^{3+}$ produced more energetically favoured products compared to its $Al^{3+}$ counterpart. Overall, many organic functional groups, particularly, the amine, amide and acid bind strongly to the bare and hydroxylated alumina surface and metal cations. Hence, they are likely to remain adsorbed on alumina and cations in soils.

# Acknowledgements

The past four years have been some of the most intellectually challenging yet enjoyable years of my academic career. To my supervisor Natalia Martsinovich, I am incredibly grateful for the help, support and guidance throughout my project. I am truly grateful to have had a supportive supervisor that has helped develop me as both a theoretician and a person. To my second supervisor Adrien Chauvet, thank you for supporting the project and always being there to lend a kind ear. Thank you to Maria Romero-Gonzalez, for her support and contributions during my first year. To all of my colleagues, particularly the Martsinovich group and the theory cluster thank you for the thought provoking discussions throughout my PhD. Lastly, a big thank you to the Grantham centre for sustainable futures for funding this project.

To my slightly tolerable, yet incredibly supportive siblings thank you all for always reminding me that anything is possible, even when you're the middle child. Your words of encouragement have always resonated with me and kept me going. To Aqib, thank you. Thank you for my morning cups of coffee, reading my work and helping me to deal with the occasional meltdown.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

Soil chemistry originated in the early 1850's with the research of J. T. Way.[17] Often considered as the father of soil science, Way found that soils could adsorb both anions and cations, and that these ions could be exchanged with other ions. His research found that ion exchange was rapid and treating soils with either heat or acid reduced the soils ability to adsorb ions. Much of the research between 1850 and the early 20th century was an extension of Way's work. Modern day soil chemistry has predominantly been motivated by the modern environmental movement, which emphasises reducing pollution and protecting and preserving natural reserves.[2]

Soil chemistry is a sub-discipline of soil science that deals with chemical composition, chemical properties and chemical reactions of soils.[2] Soils are porous media created at the surface of land through weathering processes mediated by geological, biological and hydrological phenomena. In chemistry, soils are chemical systems containing reactive species of solids, liquids and gases. As they are open systems, soils exchange both matter and energy with the surrounding biosphere, atmosphere and hydrosphere. Soils are the predominant characteristic of the life supporting critical zone. The critical zone extends from the top of the vegetation canopy to the bottom of the groundwater aquifer (permeable rock).[1]

## 1.1 Soil Minerals

Approximately one half to two thirds of the soil volume is comprised of solid matter. Of this solid matter, more than 90% represents inorganic compounds, in which humus accounts for $\sim$ 10% of the solid matter and minerals account for the remaining solid matter.[1] Elements present in the largest quantities in soils are oxygen, silicon, aluminium, iron, carbon, calcium, potassium, sodium

and magnesium, There elements are also found in the earth's crust and sediments. Oxygen is the most ubiquitous element in soils and the earth's crust.[2,18]

The elements in soil minerals typically occur as charged ions. These species are arranged in crystal lattices and held together by ionic and covalent bonds. The two most important properties of these ions in soil minerals are their valence and ionic radius. The ionic radius of a single ion cannot be measured. However, the ionic radius can be defined based on the following assumptions: The first assumption states that the radius of the bivalent oxygen ion ($O^{2-}$) is 0.140 nm in all minerals. The second assumption states that the sum of the radii of the cation and anion involved in binding is equal to the measured inter-atomic distance between the two ions. The final assumption states that the element's ionic radius has the same value in all mineral structures containing an ion with a given coordination number (CN). Most chemical bonds have a combination of covalent and ionic character. For example, the Al-O bond is approximately 40 % covalent and 60 % ionic.[1]

Two important parameters can be defined for the atomic properties of metal cations and anions that combine to form crystals. The first parameter is the ionic potential, $IP$

$$IP = \frac{Z}{R} \tag{1.1}$$

where $Z$ is the valence of the metal cation and $R$ is the ionic radius. The significance of the ionic potential (IP) is is that metal cations with an IP $< 30$ nm$^{-1}$ are found in the soil as free cations. Metal cations with IP $< 100$ nm$^{-1}$ have the tendency to hydrolyse and precipitate and those with an IP $> 100$ nm$^{-1}$ tend to be present within soils as free oxyanions. The second important parameter is bond strength, $s$

$$s = \frac{|Z|}{CN}$$

The bond strength assigns an equal proportion of the valence of a cation to each of the chemical bonds it forms with the nearest neighbouring anion.[1]

### 1.1.1 Pauling's Rules

Most of the minerals in soils are ionically bonded, thus their structures can be predicted based on Pauling's rules.[19]

*Rule 1*: A coordinated polyhedron of anions is formed about each cation. The cation-anion dis-

tance is determined by the sum of the cation and anion radii. The CN is therefore determined by the cation to anion radius ratio.

Table 1.1: Minimum radius ratio of the cation-anion inter-atomic distance and their corresponding coordination numbers (CN)

| **Minimum Radius Ratio** | **CN** |
|---|---|
| 1.00 | 12 |
| 0.732 | 8 |
| 0.414 | 6 |
| 0.225 | 4 |

*Rule 2*: In a stable mineral structure the sum of the bond strengths coming from adjacent cations to an anion is equal to the charge of the anion. For example, goethite (FeOOH) illustrates Pauling rule 2. The $Fe^{3+}$ radius is 0.065 nm. The CN of the $Fe^{3+}$ cation according to the Pauling rule 1 is found from

$$Radius\ Ratio\ of\ Fe^{3+} = \frac{Radii\ of\ Fe^{3+}}{Radii\ of\ O^{2-}} = \frac{0.065}{0.140} = 0.464 \tag{1.2}$$

As according to table 1.1, $0.464 > 0.414$, therefore the CN of $Fe^{3+}$ with $O^{2-}$ is 6. Hence,

$$s(Fe^{3+}) = \frac{|3|}{6} = 0.5 \tag{1.3}$$

Following rule 2 four $Fe^{3+}$ should bond to each $O^{2-}$ in the goethite structure. However, the structure reveals that each $O^{2-}$ binds to three $Fe^{3+}$. However, there is a proton in the formula of goethite (FeOOH), this proton can be used to provide a cation, therefore resulting in a fourth bond to $O^{2-}$. As there are twice as many oxygen atoms in goethite as hydrogen atoms, each proton must be shared by two $O^{2-}$. Thus, $s = 0.5$ as required and the fourth bond to $O^{2-}$ (that replaces $Fe^{3+}$) is a hydrogen bond.

*Rule 3*: Pauling rule 3 is a statement of Coulombs law of repulsion for cations. The Coulomb repulsive interaction between cations in a mineral are weakened or screened by the Coulomb field of the anions (negatively charged) surrounding the cations. Alumina, $Al_2O_3$ reflects this property whereby there are both short and long Al-O bonds.

*Rule 4 and Rule 5*: Pauling Rules 4 and 5 are extensions of rule 3. They reflect that stable minerals containing different types of cations do not tolerate sharing of the polyhedra.[1,2]

### 1.1.2 Primary Soil Minerals

The two most abundant chemical elements present within soils are oxygen and silicon. Therefore, these two elements combine to produce silicate minerals.[1,2,20] These minerals are given the name, *primary minerals* as they are sand and silt particles inherited from parent materials such as metamorphic and igneous rocks, Feldspars are examples of such minerals. The potassium feldspars are important as they are a major component of the mineral form of potassium.[2,9,21,22]

Table 1.2: Common primary soil minerals.[1]

| **Mineral Name** | **Chemical Formula** |
|---|---|
| Quartz | $SiO_2$ |
| Feldspar | $(Na,K)AlO_2[SiO_2]_3$, $CaAl_2O_4[SiO_2]_2$ |
| Mica | $K_2Al_2O_5[Si_2O_5]_3Al_4(OH)_4$ (muscovite) |
| | $K_2Al_2O_5[Si_2O_5]_3(Mg,Fe)_6(OH)_4$ (biotite) |
| Pyroxene | $(Ca,Mg,Fe,Ti,Al)_2(Si,Al)_2O_6$ |
| Olivine | $(Mg,Fe)_2SiO_4$ |

### 1.1.3 Secondary Soil Minerals

*Secondary soil minerals* are given their name as they occur as a result of primary mineral weathering transformations.[1,23] Particles formed by secondary minerals are clay or silt size. Clays are inorganic materials that are $< 2$ mm in size, whereas clay minerals are specific minerals that occur in a clay sized fraction.[21] Clay minerals have several different species, the most common are phyllosilicates and aluminosilicates. These clay minerals have a plate like structure. These minerals such as micas, are layered and contain sandwhiches of tetrahedral and octahedral sheets. Classified into three-layer types, they are are distinguished by the number of tetrahedral and octahedral sheets which combine to form a layer. They are further classified into five groups which are differentiated by the location of the isomorphic cation substitutions in the layers.[1] Clay particles are highly reactive due to their large specific surface area (compared to sand and silt particles).[9,24] Amongst the clay minerals, are oxides minerals. The term oxide mineral also refers to metal hydroxides, oxyhydroxides and hydrous oxides (whereby non-stoichiometric water is present within the structure).[21] Aluminium, iron and manganese oxides play an extremely important role in soil chemistry. While they are not found in large quantities (unlike silicates), they have significant effects on many chemical processes that take place in soils, such as sorption and redox processes because of their high reactivity and specific surface areas.[21,25–31]

Table 1.3: Common secondary soil minerals.[1]

| **Mineral Name** | **Chemical Formula** |
|---|---|
| Kaolinite | $Si_4Al_4O_{10}(OH)_8$ |
| Allophane | $Si_xAl_4O_{6+2x} \cdot nH_2O$, $1.6 \leq x \leq 4$, $n \geq 4$ |
| Calcite | $CaCO_3$ |
| Gypsum | $CaSO_4$ |
| Smectite, Illitle or | $M_y(Si,Al)_8(Al,Fe,Mg)_4O_{20}(OH)_4$ |
| Vermicite | M inter-layer cation, $0.4 \leq y \leq 2.0$ |

**Aluminium Oxides**

In nature, a number of crystalline aluminium hydroxides, oxyhydroxides and oxides are present in soils. The hydroxides are the most common crystalline minerals and the oxyhydroxides are the least common. Only two aluminium oxide minerals are naturally occurring within soils, gibbsite ($Al(OH)_3$) a hydroxide and boehmite ($\gamma$-AlOOH) an oxyhydroxide. Diaspore ($\alpha$-AlOOH) and corrundum ($\alpha$-$Al_2O_3$) can also be found. Anhydrous aluminium oxides are often found in igneous and metamorphic rocks.[2,32–34] Gibbsite is the most common aluminium oxide mineral. Gibbsite is often found in highly weathered soils such as oxisols (tropical mineral soils) and utisols (clay-rich soils found in warm climates).[2,3,9,35,36]

**Iron Oxides**

The major iron oxides and oxyhydroxides are listed in table 1.4. The basic structural arrangement for the iron oxides is an octahedron. The iron oxides primary difference occurs in their octahedral structure and how they are linked.[2,37] Goethite ($\alpha$-FeOOH) and hematite ($\alpha$-$Fe_2O_3$) are called the alpha ($\alpha$) phase structures. Their structures are based on hexagonal close packing of anions. On the other hand, lepidocrocite ($\gamma$-FeOOH) and maghemite ($\gamma$-$Fe_2O_3$) are the gamma ($\gamma$) phases, and their structure are based on the cubic close packing of the anions. The $\gamma$ phases are less stable than the $\alpha$ phases. The $Fe^{3+}$ can be replaced by the following metal cations, $Al^{3+}$, $Mn^{3+}$ and $Cr^{3+}$ via isomorphous substitution. Cations such as those of nickel, titanium, cobalt and zinc can also be present in the iron oxide structure.[2,6,38–40] Goethite is the most common iron oxide present within soils. Hematite is the second most common iron oxide in soils. It is commonly found in highly weathered soils. It is the mineral that gives many red soils their strong colour.[6,41–45]

Table 1.4: Common aluminium and iron oxides, hydroxides and oxyhydroxides found in soils.[2–6]

| Mineral Name | Chemical Formula |
|---|---|
| **Aluminium Oxides** | |
| Bayerite | $\alpha$-$Al(OH)_3$ |
| Boehmite | $\gamma$-AlOOH |
| Corundum | $\alpha$-$Al_2O_3$ |
| Diaspore | $\alpha$-AlOOH |
| Gibbsite | $\gamma$-$Al(OH)_3$ |
| **Iron Oxides** | |
| Akaganeite | $\beta$-FeOOH |
| Ferrihydrite | $Fe_5HO_8{\cdot}4H_2O$ |
| Feroxyhyte | $\delta$-FeOOH |
| Goethite | $\alpha$-FeOOH |
| Hematite | $\alpha$-$Fe_2O_3$ |
| Lepidocrocite | $\gamma$-FeOOH |
| Maghemite | $\gamma$-$Fe_2O_3$ |
| Magnetite | $Fe_3O_4$ |

## 1.2 Soil Organic Matter

Similar to the inorganic components of soil, organic components also play a significant role, which affects the chemistry of soils.[46–48] Soil organic matter (SOM) is the intermediate phase of carbon between completely oxidised carbon ($CO_2$) and photosynthetically fixed carbon.[9] SOM is important as it contributes to plant growth by altering chemical, physical and biological properties of soil. SOM also supplies nitrogen, sulphur and phosphorous for plant growth.[49] It also serves as an energy source for soil micro-organisms and promotes the production of good soil structure.[9]

By mass, SOM is approximately 50 % carbon. As a carbon reservoir, SOM is the largest pool of organic carbon on earth. The soil carbon reservoir is four times larger than than the terrestrial biosphere (plants and microbes etc).[50,51,51] Soils vary significantly in their organic matter content.[9] For example, soils on the surface horizons of prairie grassland contain 1 to 6 % organic carbon by mass. Whereas, soils present within sandy deserts may contain as little as 0.1 % organic carbon by mass. The largest global mass of soil carbon is stored in histosols (soils that primarily consist of organic material, often found in Canada and Scandinavia in marshy areas). Histosols have more than 50 % organic carbon my mass.

### 1.2.1 Soil Humus

The main constituents of SOM are carbon (52-58%), oxygen (34-39%), hydrogen (3.3-4.8%) and nitrogen (3.7-4.1%). SOM contains humic and non-humic compounds. Non-humic compounds consist of carbohydrates, peptides, proteins, fats and waxes. These molecules are attacked easily by the microorganisms present within soils. Therefore, they only exist in soils for a brief period of time.[2] SOM is considered to be synonymous with humus, it includes all organic compounds within soils.[2,7,52–54] The carbon containing compounds in humus are a product of degraded plant, animal, microbial molecules, burned carbon and micro-organisms.[9]

Humus consists of humic substances. Humic substances can be defined as naturally occurring heterogeneous biogenic, organic substances.[2] Humic substances are the major constituent of organic molecules in soils and sediments. They arise from the degradation of plant and animal matter, and from activity of microorganisms. Humic substances are large complex biomolecules, as they are predominantly produced from living organisms.[55,56]

The humic products formed have the tendency to associate into more complex and more stable structures, than those of the starting material.[57–60] They can generally be recognised as being either yellow or black in colour.[2,61] Humic substances are also partly aromatic, and polyelectrolytic materials,[62] They can be subdivided into humic acid, fulvic acid and humin.[2]

Several mechanisms have been proposed for the formation of humic substances,

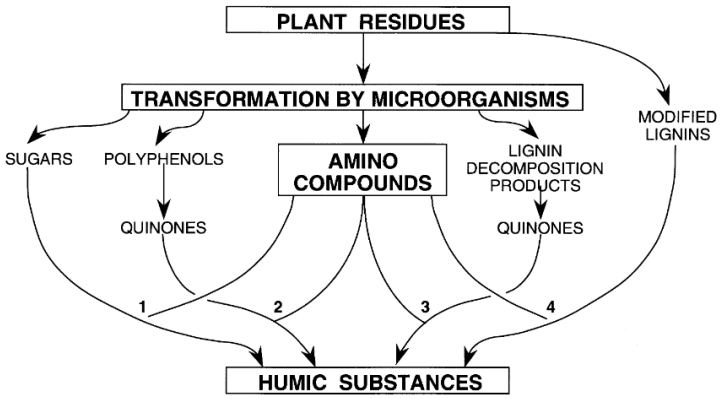



Figure 1.1: Mechanisms for the formation of soil humic substances, amino compound synthesized by microorganisms react with (1) reducing sugar, (2 and 3) quinones and (4) modified lignins.[2,7]

Pathway 1 assumes humic substances come from sugar reduction, it is not considered a significant mechanism, as the quinone pathways are considered to be more significant. The lignin theory (pathway 4) states that humic substances are modified lignins that remain after microbial attack. The modified lignins are characterised by a loss of methoxy ($OCH_3$) groups and the presence

of ortho-hydroxyphenols and the oxidation of aliphatic side chains to form carboxylic (COOH) groups. These lignins undergo further modifications resulting in humic acids and then fulvic acids. The most recent view of humic substance generation (pathway 2 and 3) involve quinones. Phenolic aldehydes and acids are also important as they are released from lignin during microbial attack and are altered to quinones. These quinones polymerise in the absence/presence of amino compounds producing humic macro-molecules. Pathways 2 and 3 are analogous to one another, except that in pathway 2 the polyphenols are microbially synthesised from non-lignin based carbon sources such as, cellulose.[2,7]

**Key Functional groups of SOM**

Table 1.5: Important functional groups of SOM.[7]

| **Functional Group** | **Structure** |
|---|---|
| | **Acdic Groups** |
| Carboxylic acid | R-C=O(-OH) |
| Enol | R-CH=CH-OH |
| Phenol OH | Ar-OH |
| Quinone | Ar=O |
| | **Netrual Groups** |
| Alcohol OH | R-$CH_2$-OH |
| Ether | R-$CH_2$-O-$CH_2$-R |
| Ketone | R-C=O(-R) |
| Aldehyde | R-C=O(-H) |
| Ester | R-C=O(-OR) |
| | **Basic Groups** |
| Amine | R-$CH_2$-$NH_2$ |
| Amide | R-C=O-(NH-R) |

where R is an aliphatic straight or branched carbon chain (saturated and unsaturated) and Ar is an aromatic ring.

There are several functional groups present in humic substances. The main and most important groups are the carboxylic acid and phenolic OH functional groups with the carboxylic acid group as the most important. Fulvic acids are more acidic than humic acids. Small amounts of alcoholic OH, quinonic and ketone groups are also found. Fulvic acids are high in carboxylic groups, and humin has more alcoholic OH groups than humic and fulvic acid.

**Biomolecules in Humus**

Large complex humic substances are the major contributor of organic carbon in soils. However, several different types of biomolecules are also present within soils with well-defined chemical structures. In most soils, weathering (microbially catalysed) of litter and humus under aerobic and anaerobic conditions results in the production of the six key building blocks of biomolecules. The key elements found in biomolecules are hydrogen, nitrogen, carbon, oxygen, sulfur and phosphorous (often released). Biomolecules can range from simple organic acids to large and very complex biopolymers. Organic acids such as formic acid and oxalic acid are amongst some of the best characterised acids.

Table 1.6: Common aliphatic organic acids present within soils.[1]

| **Functional Group** | **Structure** | $pH_{diss}$ |
|---|---|---|
| Formic acid | HCOOH | 3.8 |
| Acetic acid | $CH_3COOH$ | 4.8 |
| Oxalic | HOOCCOOH | 1.3 |
| Tartaric acid | HOOCH(-OH)CH(OH)COOH | 3.0 |

Formic acid is a mono-carboxylic acid produced by bacteria and found in root exudates (root excretions) of maize. Acetic acid is also produced microbially, often under anaerobic conditions. It is found in the root exudates of grass and herbs. The concentration of formic and acetic acid in soils ranges from 2 - 5 mM. Oxalic acid is ubiquitous in soils. Both oxalic and tartaric acid are di-carboxylic acids produced by fungi and then excreted by plant roots. Their soil solution concentrations range from 0.5 - 1 mM. Alongside the aliphatic organic acids, soils also contain aromatic organic acids. Their concentrations range from 0.05 - 3.0 mM.[1]

Soils also contain amino acids.[1,63,64] These acids have a concentration of 0.05 - 0.6 mM in soils. Amino acids can account for as much as a half of the nitrogen in soil humus. The most common amino acids that are present within soils are: glycine, alanine, aspartic acid, glutamic acid, arginine and lysine. Glycine and alanine are an example of *neutral* amino acids, for which the side chain R contains neither a carboxyl group nor the amino group, they contain aliphatic groups. The name *neutral* is apt because the COOH group is dissociated to give the negatively charged $COO^-$ species and the $NH_2$ group gains a proton to become the positively charged $NH_3^+$ species. Neutral amino acids account for two thirds of all soil amino acids. Acidic amino acids on the other hand, the R group includes a carboxyl group, for example aspartic and glutamic acids. Basic amino acids have amino groups as their R groups, for example, arginine and lysine.[1]

Carbohydrates are biopolymers of microbial and plant origin.[1,65–67] They account for up to one

quarter of the organic carbon within soils. They have a non-aromatic ring structure, with OH substituent groups on the ring. Glucose is an example of a monosaccharide that can polymerise to form the polysaccharide cellulose, after the elimination of water. The aromatic counterparts of the polysaccharides are lignin and tannin. These polymers are formed by linking together phenols or other aromatic OH groups.[68] Lignin is the most abundant aromatic biopolymer in terrestrial plants.[1,69] It accounts for up to a third of plant litter biomass.

## 1.3 The Soil Solution

Soils contain both organic and inorganic components, but they also contain water. The amount of water within in soils is known to be temperature dependent.[70,71] Liquid water in soil is a repository for dissolved solutes (gases and solids). For this reason it is therefore given the term the *'soil solution'*.[1] Soil water resides in the pore spaces of soils and it reacts with the soil particles themselves, as well as roots, microbes and gases.[9] Several complex chemical reactions take place within the soils solution[1,2,9] In saturated (or near saturated) soils the water is also known as the bulk solution, chemical reactions take place within the bulk solution.[9]

The chemical reactions that take place in the bulk solution are, (1) protonation and deprotonation of acids, (2) complexation, (3) hydrolysis (under speciation), (4) solid precipitation and dissolution and (5) gas dissolution. The first three reactions are strictly aqueous. Gas dissolution reactions involve reactions between liquids and gases. Precipitation and dissolution reactions deal with the movement of ions between solid and aqueous phases.[9]

### 1.3.1 Protonation and Deprotonation of Acids

In soils some of the molecules liberate $H^+$ ions, thus they are known as Bronsted acids. Strong acids such as HCl and $HNO_3$ completely dissociate into protons and the conjugate bases (anions). Weak acids on the other hand (such as $HCO_3^-$ and $CH_3COOH$) deprotonate at a higher pH (compared to strong acids). For a generic weak acid (HA), the deprotonation reaction has the following equation

$$HA \rightleftharpoons H^+ + A^- \tag{1.4}$$

where $A^-$ is the anion (conjugate base). Therefore, the acidity constant ($K_a$) is

$$K_a = \frac{[H^+][A^-]}{[HA]} \tag{1.5}$$

using $pK_a = -\log K_a$ we will take the negative logarithm of the expression for the acidity constant and separating terms, we obtain the following approximation

$$pK_a = -log[H^+] - log\frac{[A^-]}{[HA]} \tag{1.6}$$

$pH = -\log[H^+]$ therefore,

$$pK_a = pH - log\frac{[A^-]}{[HA]}$$

making pH the subject gives

$$pH = pK_a + log\frac{[A^-]}{[HA]} \tag{1.7}$$

This equation is known as the *Henderson-Hasselbalch equation.*[9,72] From this equation one can calculate the ratio of protonated and deprotonated acids as a function of the pH.

$$log\frac{[A^-]}{[HA]} = pH - pK_a \tag{1.8}$$

The behaviour of acids in soils can be calculated using the above equation, which calculates the ratio of protonated and deprotonated species within soils. The most important factor within protonation and deprotonation reactions within soils is the pH.[9]

### 1.3.2 Complexation

Molecules and ions that are strongly interacting often become complex ions or complexes.[9] Complexes are molecules that are composed of a central ions surrounded by ligands.[1] Typically, ligands are negatively charged and attracted to central ions that are typically positively charged cations, alternatively ligands have electrons that are available for binding to cations. In soils, complexes such as $[Fe(H_2O)_6]^{3+}$ form from the interaction of the free $Fe^{3+}$ cation with $H_2O$. In the $[Fe(H_2O)_6]^{3+}$ complex ligands can replace a single or several $H_2O$ molecules in the primary hydration sphere. Alternatively, other ligands may exist in the second solvation sphere. When a ligand is in the second solvation sphere an *ion pair* is formed. The interactions between an ion pair (electrostatic) are weaker than those of the complexed ions. Complex ions are known as *inner-sphere complexes* and ion pairs are known as *outer-sphere complexes.*[9,73] Many alkaline earth metals occur in soil solutions as either complex ions or ion pairs. Ligands must complete with water in the central

ion's solvation sphere to associate with a central ion.[9]

#### Predicting Complexation

Complexation within soils can be predicted using the hard and soft acid and base (HASB) theory. HASB theory predcits the polarisibility of ions. The HASB theory states that hard Lewis acids tend to interact with hard Lewis bases and soft Lewis acids tend to interact with soft Lewis bases. In aqueous solutions, such as the soil solution, cations such as $Al^{3+}$, $Fe^{3+}$ and $Co^{3+}$ (and more) react strongly with oxyanion ligands, such as $CO_3^{2-}$, $SO_4^{2-}$ and $PO_4^{3-}$. Metal cations such as $Cu^{2+}$ and $Cd^{2+}$ interact with non-oxygen anions such as $S^{2-}$ and $CN^{-}$. Cations that prefer to interact with oxygen containing anions (hard Lewis bases) are hard Lewis acids and are less polarisable. Whereas, cations that interact with non-oxygen containing anions (soft Lewis bases) and soft Lewis acids. Intermediate acids and bases can interact with either hard or soft acids and bases and are more polarisable. The HASB theory allows the prediction of cation and anion complexation. In soils, reduced oxidation states such tend to softer Lewis acids and bases. For example, $Fe^{2+}$ is a softer acid than $Fe^{3+}$, hence they behave differently in soils.

### 1.3.3 Speciation of the Soil Solution

Speciation is a chemical process by which a species has formed. In soils, speciation of free metal ions reduces their toxicity.[73–77] In the soil solution it is paramount as one should know the form of the free and complex species that are present within soils. With this knowledge one gains a better understanding of the reaction of the species present. Ions in the soil solution can form a number of different species. One example of speciation reactions are hydrolysis reactions. Hydrolysis reactions split a $H^+$ ion from a water molecule producing $H_3O^+$ by interacting with neighbouring water molecules. The general equation for the hydrolysis of a hydrated metal ion in solution is

$$[M(H_2O)_x]^{n+} \rightleftharpoons [M(OH)_y(H_2O)_{x-y}]^{(n-y)+} + yH^+ \tag{1.9}$$

where $y$ is a stoichiometric coefficient and n is the charge of the complex. The degree of hydrolysis is pH dependent. Several reactions can occur within soils, another example is the reaction between a hydrated metal complex and a ligand. The general equation between a hydrated metal complex and negatively charge ligand is the as follows

$$a[M(H_2O)_x]^{n+} + bL^- \rightleftharpoons [M(H_2O)_x]_aL_b \tag{1.10}$$

where $a$ and $b$ are stoichiometric coefficients. L is a singly charged ligand.

### 1.3.4 Dissolution and Precipitation

Precipitation is the process of depositing a substance from the soil solution and dissolution is the process of dissolving a solid. These reactions determine the fate of inorganic mineral components within soils.[2] Mineral weathering processes are affected by the dissolution-precipitation equilibria of the solid phase.[1,2] Precipitation of minerals within soils only occurs when supersaturation conditions exist in the soil solution. On the other hand, dissolution only occurs when the soil solution is undersaturated with respect to the minerals within the soil.[2]

Silicate minerals such as anorthite ($CaAl_2Si_2O_8$) undergo dissolution reactions to form various secondary minerals that are important in soils. An example of the weathering reaction of anorthite to produce kaolinite ($Al_2Si_2O_5(OH)_4$) is as follows[2,78]

$$CaAl_2Si_2O_8 + 2CO_2 + 3H_2O \rightleftharpoons Al_2Si_2O_5(OH)_4 + Ca^{2+} + HCO_3^- \tag{1.11}$$

### 1.3.5 Gas Dissolution

Soils also contain gases. The exchange of gas from the soil solution to the atmosphere is a process that controls the aqueous chemistry of ammonia, carbonate, nitrogen, oxygen and sulfur. An example is the gas dissolution reaction for di-nitrogen

$$N_{2(g)} + H_2O_{(l)} \rightleftharpoons N_{2(aq)} \tag{1.12}$$

The equilibrium distribution for gas dissolution is predicted using *Henry's law*

$$K_H = \frac{a_i}{P} \tag{1.13}$$

where $K_H$ is the Henry's law constant. $a_i$ is the aqueous activity of the dissolved species $i$, and $P$ is the partial pressure of the gas. As the partial pressure of a gas increases, the aqueous activity also increases.[9]

## 1.4 Adsorption Phenomena of Soils

We have discussed the important components of soils, particularly the inorganic and organic components. We have also discussed the importance of water and the important chemical reactions that occur within the soil solution. We will now discuss the importance of adsorption phenomena within soils. Adsorption within soils can be defined as the accumulation of a material (adsorbate) at an interface between either a mineral surface or a SOM surface and the aqueous solution.[2,9] Adsorption is considered to be one of the most (if not the most) important chemical process within soils. It determines the the availability and retention of plant nutrients, pesticides and metals within soils. It is one of the primary processes that affects the transport of nutrients and also contaminants within soils. Both physical and chemical forces are involved in adsorption processes within soils. Physical forces include Van der Waals forces and electrostatic outer-sphere complexes. Contrarily, chemical forces result from the interaction of a short range nature. These interactions include covalent bonding, hydrogen bonding and ionic bonding, a process that involves a ligand exchange mechanism.[2]

Adsorption reactions can be classified into *outer-sphere adsorption* and *inner-sphere adsorption*. Within outer-sphere adsorption processes at least one water molecule remains between the surface and the adsorbate. Inner-sphere adsorption processes on the other hand involve the adsorbate directly binding through either a covalent or an ionic bond with the mineral surface. These two adsorption mechanisms are analogous to outer-sphere and inner-sphere complexes as discussed in section 1.3.2.[9,79]

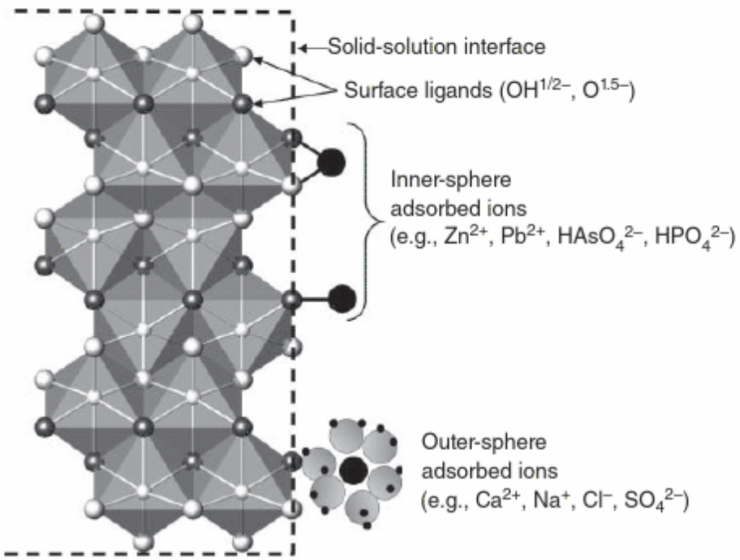



Figure 1.2: Bonding characteristics of inner and outer-sphere adsorption.[9]

### 1.4.1 Outer-sphere Adsorption

Outer-sphere adsorbed ions are located on mineral surfaces. Although they are located on the surface of a mineral they do not form chemical bonds with the functional groups present on the surface. Ions such as $Ca^{2+}$, $Mg^{2+}$, $Na^{+}$, $K^{+}$, $NO_3^-$ and $SO_4^{2-}$ can be adsorbed by outer-sphere ions. They acts as the major reservoir of many micro-nutrients. Cations are adsorbed electrostatically to negatively charged surfaces and anions are adsorbed (electrostatically) onto positively charged surfaces. Several factors that affect outer-sphere electrostatic adsorption reactions are the ion valence, type of surface charge, ion hydration and the concentration of ions in the soil solution. Their reactions are all fast and reversible and they occur on minute to seconds timescales.

On oxide mineral surfaces, outer-sphere adsorbed ions exist as hydrated ions. These hydrated ions are attracted to the charge of the surface functional groups. The charge of the surface functional groups on oxide minerals is not strong enough to dehydrate the hydrated outer-sphere adsorbed ions. Minerals such as phyllosilicates have a delocalised charge, and on these surfaces cations maintain their hydration spheres.[9]

#### Cation Exchange in Soils

Cation exchange in soils is primarily an outer-sphere adsorption reaction. Cation exchange reactions predominate within soils that have a high mineral content. The *Cation exchange capacity* (CEC) is the amount of cations a particular soil can adsorb.[9,80] It is equal to the net negative charge of the soil particles. The majority of a soil's CEC is provided by SOM and soil clays. Exchangeable cations are ions that are replaced by other ions salt solutions in soils. The major exchangeable cations (in order) are $Ca^{2+}$ (80%) > $Mg^{2+}$ (15%) > $K^{+}$ (4%) > $Na^{+}$ (1%). As the acidity of soils increases, $H^{+}$ and $Al^{3+}$ ions also occupy the cation exchange sites on mineral surfaces. As a result of the varying charge, size and polarisability of cations, their adsorptive behaviour also varies. Cations with a low charge and a large hydration radius such as $Ca^{2+}$, $Mg^{2+}$ and $Na^{+}$ tend to adsorb via the formation of outer-sphere bonds on most mineral surfaces.

The attraction of an ion and a charged surface is determined by the strength of electrostatic adsorption. Strongly adsorbed ions ions are always preferentially adsorbed over less strongly adsorbed ions. The magnitude of this interaction in outer-sphere adsorption reactions is often described by a phenomenon known as *cation selectivity*.[9]

#### Outer-sphere Anion Adsorption

Similar to, outer-sphere adsorbed cations, anions also undergo outer-sphere adsorption reaction. Anions are attracted to positively charged sites on clay mineral edges, allophane surfaces and

hydrous oxides. They are repelled by permanent charge clay mineral inter-layers. Soils that contain large amounts of oxide minerals and kaolinite often have a positive charge overall. The positively charged sites electrostatically attract anions. $Cl^-$, $NO_3^-$ and $SO_4^{2-}$ anions are the most important anions within soils. The mobility of the $Cl^-$ anion is similar to that of the $NO_3^-$ anion. Therefore, it is often used as indicator for the mobility of $NO_3^-$ within soils.

### 1.4.2 Inner-sphere Adsorption

We will now discuss inner-sphere adsorption reactions. Inner-sphere adsorption reactions are chemisorption reactions, whereby covalent bonds are formed between an adsorbate and a mineral or SOM surface. Metal cations such as $Al^{3+}$, $Be^{2+}$ and $Pb^{2+}$, alongside transition metal cations (and several others) all form inner-sphere bonds with the functional groups that exist on SOM and mineral surfaces. Amorphous materials have more surface functional groups than crystalline materials. Thus, they have a greater inner-sphere adsorption capacity. For example, ferrihydrite ($Fe_5HO_8 \cdot 4H_2O$) has a greater adsorption capacity than that of goethite ($\alpha$-FeOOH).

Inner-sphere adsorption reactions are dependent upon an ions water solubility and the ability to form covalent bonds. Charge, surface area and the type of ligand are important factors for the surface. Compared to outer-sphere adsorption reactions inner-sphere reactions are initially very rapid, however the rates get slower over long periods of time. The pH is also an important factor in inner-sphere adsorption reactions as the charge of a functional group is pH dependent.[2,9,75,81]

#### Inner-sphere Cation Adsorption

The following metal cations $Cu^{2+}$, $Cd^{2+}$, $Co^{2+}$, $Fe^{3+}$, $Al^{3+}$, $Hg^{2+}$, $Mn^{2+}$, $Ni^{2+}$, $Pb^{2+}$, $Be^{2+}$ and $Zn^{2+}$ are all micro-nutrients essential within soils. These cations can form inner-sphere complexes on mineral surfaces, due to their electron orbital energies favouring the formation of covalent bonds. Aluminium and iron ion concentrations are the largest within soils, the remaining ion concentrations are fairly low.

Inner-sphere metal cations adsorb to the edges of clay minerals and on the surfaces of soil mineral oxides. The general reaction for the deprotonation of a reactive surface functional group (in this case surface hydroxyls) and an $M^{2+}$ cation as an inner-sphere complex is

$$\equiv S - OH + M^{2+} \rightleftharpoons \equiv S - O - M^+ + H^+ \tag{1.14}$$

Generally, the divalent cations $Hg^{2+}$, $Cu^{2+}$ and $Pb^{2+}$ are more strongly adsorbed than $Cd^{2+}$, $Zn^{2+}$, $Mn^{2+}$ and $Ni^{2+}$. Trivalent cations such as $Fe^{3+}$ and $Al^{3+}$ form minerals within soils. The

adsorption of trivalent cations is also stronger than that of divalent cations in soils.[9]

**Inner-sphere Anion Adsorption**

There are several anions that are strongly retained by inner-sphere complexes such as $PO_4^{3-}$, $AsO_4^{3-}$, $MoO_4^{2-}$, $CrO_4^{2-}$, $S^{2-}$, $SeO_3^{2-}$, $B(OH)_4^-$ and $F^-$ and their protonated forms. These anions have a great impact on soils as they are added to soils in fertilisers, agricultural waste, coal combustion and fly ash. Initially, the addition of strongly adsorbed anions increase their bio-availability and the soil solution concentration. However, the anions then adsorb to soil colloids, thus decreasing availability and leaching.

$PO_4^{3-}$ and $F^-$ anions displace $O^{2-}$ ligands that coordinate to $Al^{3+}$ and $Fe^{3+}$ metal cations through a ligand exchange process. Ligand exchange reactions take place on the surfaces of oxide minerals. Unlike outer-sphere ligand exchange reactions which depend on the surface charge to be positive, inner-sphere ligand exchange depends on the approach of the anions towards the surface. This approach of anions is independent of the surface charge. The general equation for the ligand exchange process for anion adsorption is

$$\equiv S - OH + A^- \rightleftharpoons \equiv S - A + OH^- \tag{1.15}$$

### 1.4.3 Adsorption of OC to mineral surfaces

A number of studies have been carried out to determine the OC content and binding in soils using spectroscopic techniques.[82–85] For example McKnight et al.[82] investigated sorption of riverine OC onto Fe and Al hydrous oxides. Their results suggested that carboxyl groups bonded to aromatic molecules are preferentially adsorbed onto the Fe and Al hydrous oxides of soils. They also found that alkyl carbon fractions accumulated within solution, thus suggesting aliphatic groups have weaker binding due to their lack of carboxyl functional groups.[86,87] These comparative studies were conducted on the chemical structure of the OC before and after it was adsorbed onto $\alpha$-FeOOH and $Al(OH)_3$ surfaces. Thus, concluding that aromatic carbonyl groups bind to soil, specifically to iron and aluminium based minerals.[88,89] Kaiser et al.[85] also examined the organic compounds present in forest floor-derived OC by dissolving samples of OC in water and then adsorbing onto the $Al(OH)_3$ surface. They analysed and compared the OC composition in the initial solution and the effluent (portion of OC not adsorbed). The effluent was found to contain slightly less carbonyl and aromatic carbon and more alkyl carbon than the initial solution. This indicates that the OC that remains bound to soil is enriched in carbonyl and aromatic groups, while alkyl groups are less strongly bound and can be removed from the soil by a solvent.

Many studies showed that OC binds to the surfaces of mineral oxides, in particular, aluminium and iron oxides and hydroxides.[82,86,90] Sorption of OC is considered a key process in the stabilization of OC, as it is protects OC from microbial decay.[83–85] The stability of the mineral-attached OC depends upon three factors: (i) the availability of the mineral surfaces during sorption, (ii) structural properties of the OC and (iii) the presence of inorganic ions, such as phosphate, competing with OC for the sorption sites.[83] Sorption of OC onto mineral surfaces also controls the mobility of organic matter within soils. The main sorbents of OC within soils are metal oxides and hydroxides.[83,91] OC sorption onto mineral surfaces occurs via a range of different mechanisms, the most prominent being the ligand exchange mechanism, which is a complex mechanism consisting of both electrostatic attraction and surface complexation; it is considered as the most efficient mechanism to stabilise OC from microbial decay.[83] Thus, Al and Fe oxides and hydroxides have been identified as the minerals that are primarily responsible for the stability of OC in soil. The overall aim of this thesis is to investigate the mechanism of OC binding to these minerals, using calculations of model systems.

## 1.5 Motivation and Aims

Soils are complex systems composed of inorganic minerals, SOM, living organisms, gases and water.[1,2,9] These important soil components result in reactions that preserve the quality of soils. In particular, the presence of OC is very important for the stability and fertility of soils. However, it has received less attention than the adsorption of cations and anions. The adsorption of OC to mineral oxide surfaces is an important process as it protects organic carbon from microbial degradation.[83,84,92] However, the exact nature of the interaction of OC to mineral surfaces and mineral complexes is unknown.

Although there have been some experimental studies into OC interaction within soil, the chemical form of carbon that preferentially adsorbs has not been investigated. Most investigations of OC in soil have also been carried out using experimental techniques, with very few theoretical investigations of mineral-molecule interactions. Hence, we will conduct a series of theoretical investigations to obtain a comprehensive and systematic description of the binding of OC to soil mineral complexes and soil mineral surfaces. We aim to obtain a greater understanding of the complex nature of adsorption phenomena within soils. These investigations are a step towards the production of *carbon fertilisers* a concept based upon improving the soil via the addition of carbon-based molecules. Carbon based fertilisers have a similar purpose as nitrogen, phosphorous and potassium based fertilisers, except they contain carbon. Their purpose is to improve the soil for plant growth by containing strongly adsorbing OC that binds to soil minerals and soil cations.

## 1.6 Outline of Thesis

We will be conducting theoretical modelling investigations into the interaction of OC with mineral complexes and mineral surfaces. We will be the investigating the adsorption of small organic functional groups to alumina a common soil mineral oxide that is believed to be important for OC binding, and specifically the $\alpha$-$Al_2O_3$ (0001) surface, a well characterised surface. We will then hydrate the $\alpha$-$Al_2O_3$ (0001) surface and conduct adsorption studies of organic functional groups to the hydroxylated surface. This will enable us to compare and contrast the binding of organic groups to the bare and hydroxylated alumina surface to determine strongly adsorbing groups and also the nature of binding on each surface. Investigations into the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface are important, as water is naturally present within environmental systems. Therefore, mineral surfaces are naturally hydrated.

We will then present the most realistic modelling investigations into OC interactions with the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. Using knowledge attained from our initial adsorption studies we will deduce which small organic functional groups bind strongly to the mineral oxide surface. Hence, we will research naturally occurring biopolymers that contain these strongly adsorbing carbon containing functional groups. We will then adsorb these biopolymers to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface in several different geometries. From these investigations we will determine the geometries particular biopolymers prefer to adsorb in. We will also determine the strongest adsorbing biopolymer. Our final set of investigations will focus on the ligand substitution reaction between small organic functional groups (previously studied) to the hydrated hexa-aqua $[Al(H_2O)_6]^{3+}$ and $[Fe(H_2O)_6]^{3+}$ complexes. These models will yield a further understanding of the functional groups that bind to aluminium and iron cations in the soil solution.

# Chapter 2

# Theory

## 2.1 Quantum Mechanics

Quantum mechanics is a set of laws that governs the behaviour of particles. These particles include atomic nuclei and electrons. Its predecessor classical mechanics failed to provide an accurate description of matter on this small scale. An example of a failure of classical mechanics is its inability to describe the true nature of electrons. Classically, an electron would either be described as a particle or a wave. However, electrons neither behave exclusively as a particle nor a wave, as they exhibit the phenomenon known as wave-particle duality, whereby the electron does not correspond to point object, but rather it spreads as a wave, with the wavelength.[93]

$$\lambda = \frac{h}{mv} \tag{2.1}$$

where $\lambda$ is the wavelength, $h$ is Plancks constant, $m$ is the mass and $v$ is the velocity. Niels Bohr proposed a theoretical hydrogen model which is well known as the Bohr model. The model expresses the positions of electrons and nuclei, under the consideration of wave-particle duality. Bohr was unable to extend the model to a many electron system.

As a solution the quantum wavefunction, $\Psi$ was proposed. The wavefunction contains a complete description of a quantum system. For a single electron system the wavefunction is represented by $\Psi(\mathbf{r}_1)$, which contains one radial parameter ($\mathbf{r}_1$). The radial parameter is used to represent the coordinates of an electron. In a many-electron system $n$-radial parameters are included. Thus, the wavefunction is $\Psi(\mathbf{r}_1, \mathbf{r}_2 ... \mathbf{r}_n)$, defined for electron 1, electron 2, to electron $n$ respectively.[93]

## 2.2 The Schrödinger Equation

The year 1926 marks the origin of modern quantum mechanics,[94] where Austrian physicist Erwin Schrödinger proposed a non-relativistic form of the wave equation.[95] Schrödinger linked the energy, $E$, to the wavefunction, $\Psi$, of a system via the Hamiltonian operator, $\hat{H}$.

$$\hat{H}\Psi = E\Psi \tag{2.2}$$

Equation 2.2 is the Schrödinger equation in its time-independent form. $\hat{H}$ is known as the Hamiltonian operator, which yields the total energy of the system.[94] The Schrödinger equation is an eigenvalue equation, where the solutions of the equation $E$ and $\Psi$ denote eigenvalue and eigenfunction.

The wavefunction, $\Psi$, and the energy, $E$, are linked together by the Hamiltonian operator, $\hat{H}$. In the Schrödinger equation the Hamiltonian operator acts upon the wavefunction to yield the energy of a system. The Hamiltonian can be separated into its kinetic $\hat{T}$ and potential energy $\hat{V}$ terms.

$$\hat{H} = \hat{T} + \hat{V} \tag{2.3}$$

$$\hat{T} = -\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) = -\frac{\hbar^2}{2m}\hat{\nabla}^2 \tag{2.4}$$

Equation 2.4 is the kinetic energy operator $(\hat{T})$ in a three-dimensional system of coordinates $x$, $y$ and $z$, where $\hat{\nabla}^2$, is the Laplacian. $\hbar$ is the reduced Planck's constant:

$$\hbar = \frac{h}{2\pi} \tag{2.5}$$

The potential energy operator, $\hat{V}$, is just a multiplicative factor. Thus, we can construct the Hamiltonian operator for a system containing electrons and nuclei using the kinetic energy and potential energy operators (equations 2.3 and 2.4).

$$\hat{H} = -\frac{\hbar^2}{2m_e}\sum_i \hat{\nabla}_i^2 - \frac{\hbar^2}{2}\sum_A \frac{1}{M_A}\hat{\nabla}_A^2 + \frac{e^2}{4\pi\varepsilon_0}\left(\sum_{A>B}\frac{Z_A Z_B}{\mathbf{R}_{AB}} - \sum_{A,i}\frac{Z_A}{\mathbf{r}_{Aj}} + \sum_{i>j}\frac{1}{\mathbf{r}_{ij}}\right) \tag{2.6}$$

where $m_e$ is the mass of an electron, $M_A$ is the mass of the nuclei and $A$ and $B$ represent the

nuclei $A$ and $B$. $\varepsilon_0$ is the dielectric permitivity of the vacuum, $e$ is the charge of an electron, $i$ and $j$ are electrons and $\mathbf{r}_{ij}$ is the separation between electrons $i$ and $j$. $\mathbf{R}_{AB}$ represents the separation between nuclei $A$ and nuclei $B$ and $Z_A$ and $Z_B$ are the nuclear charges. The first two terms of the Hamiltonian for a polyatomic system account for the kinetic energy of the electron and the nuclei. The final three terms in equation 2.6 are potential energy terms. The first term in the parentheses represents the repulsive energy between nuclei, the second term represents the attractive energy between the electron and nuclei, as opposite charges attract. The final term in the brackets represents the repulsive energy between electrons, as like charges repel.

The Hamiltonian for a polyatomic system can be simplified using Hartree atomic units. Hartree atomic units define the charge of a proton, mass of an electron, $m_e$, reduced Planck's constant, $\hbar$ and the $4\pi\varepsilon_0$ term to 1. With this knowledge the Hamiltonian can be written in a more simple form.

$$\hat{H} = -\frac{1}{2}\sum_i \hat{\nabla}_i^2 - \frac{1}{2}\sum_A \frac{1}{M_A}\hat{\nabla}_A^2 + (\sum_{A>B}\frac{Z_A Z_B}{\mathbf{R}_{AB}} - \sum_{A,i}\frac{Z_A}{\mathbf{r}_{Ai}} + \sum_{i>j}\frac{1}{\mathbf{r}_{ij}}) \tag{2.7}$$

Hartree atomic units are employed to simplify the complex Hamiltonian operator for a polyatomic system. However, a solution to the Schrödinger equation is not easily obtained for a polyatomic system. Approaches have been developed to simplify it, the first of these is the Born-Oppenheimer approximation, which exploits the mass difference between electrons and nuclei. However, the solution for a polyatomic system still remains a challenge.

## 2.3 The Born-Oppenheimer Approximation

The Born-Oppenheimer approximation exploits the mass difference between electrons and nuclei (mass of nuclei ∼1836 mass of an electron). Since electrons move more than a thousand times faster than nuclei, electrons are able to adjust their positions instantaneously with respect to nuclei. Hence, nuclei can be considered to be stationary, on the timescale of electron movement. Consequently, the wavefunction can be separated into its electronic and nuclear components respectively.[96]

$$\Psi(\mathbf{r},\mathbf{R}) = \Phi(\mathbf{r};\mathbf{R})\chi(\mathbf{R}) \tag{2.8}$$

Where $\Phi$ is the electronic wavefunction and $\chi$ is the nuclear wavefunction. $\mathbf{r}$ and $\mathbf{R}$ are the coordinates of the electrons and the nuclei. The Born-Oppenheimer approximation can then be

applied to the Hamiltonian operator simplifying the nuclear motion, thus generating the clamped-nucleus Hamiltonian. Within this approximation, the nuclear kinetic energy is set to zero and the nuclear-nuclear repulsive potential energy is treated as a constant.

$$\hat{H} = -\frac{1}{2}\sum_{i}\hat{\nabla}_i^2 - \sum_{A,i}\frac{Z_A}{r_{Ai}} + \sum_{i>j}\frac{1}{\mathbf{r}_{ij}} + \sum_{A>B}\frac{Z_A Z_B}{\mathbf{R}_{AB}} \tag{2.9}$$

One can now generate the clamped nucleus Schrödinger equation. This equation computes the electronic energy, $E$ for fixed nuclear positions $\mathbf{R}$ generating what are known as potential energy surfaces.

$$\hat{H}\Phi(\mathbf{r;R})\chi(\mathbf{R}) = E\Phi(\mathbf{r;R})\chi(\mathbf{R}) \tag{2.10}$$

By employing the clamped-nucleus approximation the electronic part of the Schrödinger equation can be solved. For a one-electron system an exact solution to the above Schrödinger equation can be obtained. However, in a many body system an approximate solution must be obtained.[96,97] One of the early and well known methods to solve it is the mean-field approximation, which is also known as the Hartree-Fock method.

## 2.4 Hartree-Fock Theory

The Hartree-Fock method is considered a fundamental starting point in ab initio calculations.[93] It is the based on molecular orbital (MO) theory. MO theory posits that the motion of each electron can be described by an orbital (single particle function).

Hartree-Fock theory is based on the premise of treating electrons as independent and non-interacting. The electronic Schrödinger equation is solvable for hydrogen, if we imagine that we add an electron to the hydrogen atom, we would obtain $H^-$. Treating the two electrons in $H^-$ as non-interacting, would result in a separable Hamiltonian, $\hat{H}$ and the total wavefunction, $\Psi(\mathbf{r}_1,\mathbf{r}_2)$, which describes the motion of electrons 1 and 2. Therefore, the total electronic wavefunction would simply be a product of two hydrogen atom wavefunctions (orbitals), $\Psi_H(\mathbf{r}_1)\Psi_H(\mathbf{r}_2)$. Thus, the *Hartree Product* is born, whereby the wavefunction has the form:

$$\Psi_{HF}(\mathbf{r}_1, \mathbf{r}_2 ... \mathbf{r}_n) = \phi_1(\mathbf{r}_1)\phi_2(\mathbf{r}_2)...\phi_n(\mathbf{r}_n) \tag{2.11}$$

where, $\Psi_{HF}$ is the Hartree wavefunction and $\phi_1(\mathbf{r}_1)$, $\phi_2(\mathbf{r}_2)$ and $\phi_n(\mathbf{r}_n)$ are spatial orbitals corresponding to electrons 1 and 2, all the way through to electron n. Although useful, the *Hartree Product* has a major shortcoming, which is its inability to satisfy the Pauli antisymmetry principle.[98–101] The Pauli Antisymmetry principle[98,100,101] states that the wavefunction describing fermions should by antisymmetric with respect to the interchange of the space-spin coordinates. Space-spin coordinates refer to fermions not only having three spatial degrees of freedom, but also a set of spin coordinates. The spin can either be $\alpha$ or $\beta$. With the full set of coordinates the *Hartree Product* becomes:

$$\Psi_{HF}(\mathbf{x}_1, \mathbf{x}_2 ... \mathbf{x}_n) = \psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)...\psi_n(\mathbf{x}_n) \tag{2.12}$$

where $\psi_1(\mathbf{x}_1)$, $\psi_2(\mathbf{x}_2)$ and $\psi_n(\mathbf{x}_n)$ are the space-spin orbitals corresponding to electrons 1 and 2, all the way through to electron n. However, this does not satisfy the antisymmetry principle as:

$$\Psi_{HF}(\mathbf{x}_1, \mathbf{x}_2) = \psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2) \tag{2.13}$$

If we swap the coordinates of electrons 1 and 2:

$$\Psi_{HF}(\mathbf{x}_2, \mathbf{x}_1) = \psi_1(\mathbf{x}_2)\psi_2(\mathbf{x}_1) \tag{2.14}$$

The only way to obtain a negative wavefunction is as follows:

$$\Psi_{HF}(\mathbf{x}_2, \mathbf{x}_1) = -\psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2) \tag{2.15}$$

which is not true, hence the Hartree product does not contain the required properties.[98] Hence, to satisfy the Pauli antisymmetry principle the following wavefunction must be employed for a two electron system.

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}}[\psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2) - \psi_1(\mathbf{x}_2)\psi_2(\mathbf{x}_1)] \tag{2.16}$$

This wavefunction now satisfies the Pauli antisymmetry principle, but also the Pauli exclusion principle which is a consequence of the antisymmetry principle. In the case of many electrons, or

$N$ electrons, the above wavefunction can be written by using a determinant, known as the Slater determinant.[102]

$$\tilde{\psi}_{HF}(\mathbf{x}_1, \mathbf{x}_2 ... \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \dots & \psi_n(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \dots & \psi_n(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & ... & \psi_n(\mathbf{x}_N) \end{vmatrix} \tag{2.17}$$

Equation 2.17 can be represented by its principal diagonal using Dirac bra-ket notation,[103] which is the shorthand notation for the Slater determinant:

$$\tilde{\Psi}_{HF}(\mathbf{x}_1, \mathbf{x}_2 ... \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \dots & \psi_n(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \dots & \psi_n(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & ... & \psi_n(\mathbf{x}_N) \end{vmatrix} = |\psi_1(\mathbf{x}_1)\psi_2(\mathbf{x}_2)...\psi_n(\mathbf{x}_N)\rangle \tag{2.18}$$

where, $\frac{1}{\sqrt{N!}}$ represents the normalisation factor, $\tilde{\Psi}_{HF}$ is the trial wavefunction and $\psi_n(\mathbf{x}_N)$ represents the spin orbital of the nth electron.[102,104] Equation 2.17 illustrates the Slater determinant for an N-body electron wavefunction which is constructed from one-electron functions via a linear combination of atomic orbitals (LCAO method). Slater determinants are conveniently able to satisfy an important requirement for the wavefunction, which is the need to satisfy the Pauli antisymmetry principle.[?,101]

### 2.4.1 Energy Expression

Using the bra-ket notation we now have simplified notation for the electronic wavefunction, which we know takes the form of a Slater determinant. The expectation value of the energy is given by the following equation in bra-ket notation.

$$E_{el} = \langle \Psi | \Psi \rangle \tag{2.19}$$

where $E_{el}$ is the total electronic energy, $\hat{H}$ is the Hamiltonian operator and $\Psi$ is the wavefunction (antisymmetrized). The *variational theorem*[98,105,106] states that the energy is always the upper bound to the true energy and it depends on the wavefunction. Hence, one can obtain better

wavefunctions by varying their parameters with the given space. Thus, the correct molecular orbitals are the ones that minimise the electron energy.

### 2.4.2 The Hartree-Fock Equations

The Hartree-Fock method seeks to approximately find a solution to the electronic Schrödinger equation. The method assumes that the wavefunction can be approximated by a single Slater determinant made up of one spin orbital per electron. The *Variational theorem* is applied.[98]

The many-body Schrödinger equation is mathematically transformed by conversion into the one-electron Hartree-Fock equation .[93]

$$\hat{F}\chi_i = \varepsilon_i \chi_i \tag{2.20}$$

where $\hat{F}$ denotes the Fock operator (one-electron operator for a spin orbital), $\chi_i$ is the one electron space-spin wavefunction and $\varepsilon_i$ is an eigenvalue which denotes the energy of orbital i. The Fock operator is defined below

$$\hat{F} = h_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\hat{J}_{ij} - \hat{K}_{ij}) \tag{2.21}$$

Where $h_i$ is a one electron operator which includes the kinetic energy and Coulomb (potential) energy between the atomic nucleus and electrons. The final two terms are the two-electron terms, which describe the electron-electron interaction. $\hat{J}_{ij}$ and $\hat{K}_{ij}$ denote the Coulomb and exchange operators.

$$\hat{J_{ij}} = \langle \psi_i \psi_j | \frac{1}{r_{12}} | \psi_i \psi_j \rangle \tag{2.22}$$

$$\hat{K_{ij}} = \langle \psi_i \psi_j | \frac{1}{r_{12}} | \psi_j \psi_i \rangle \tag{2.23}$$

where $r_{12}$ is the inter-electronic distance between electrons one and two. The Coulomb operator, $\hat{J}_{ij}$, accounts for electrostatic (classical) interactions. It essentially states that an electron in orbital $\psi_i$ experiences repulsion from an electron that is in the spin orbital $\psi_j$. On the other hand, the exchange operator, $\hat{K}_{ij}$, is an interaction that is purely quantum mechanical. The term arises due to the antisymmetric nature of the wavefunction.

### 2.4.3 Open and Closed Shell Systems

In a closed shell system, there is a one restriction that $\alpha$-spin and $\beta$-spin electrons are paired in the same spatial orbital ($\psi$). Hartee-Fock in the closed shell system is called restricted Hartree-Fock (RHF). On the other hand, in an open shell system the spin orbital ($\phi$) of the $\alpha$ electron is independent from the $\beta$ electron. Open shell Hartree-Fock is also called unrestricted Hartree-Fock (UHF). The expressions for Hartree-Fock energies are illustrated below:

$$E_0^{HF} = \sum_i^n \langle \psi_i | \hat{h} | \psi_i \rangle + \frac{1}{2} \sum_{ij}^n (\langle \psi_i \psi_j | \frac{1}{r_{12}} | \psi_i \psi_j \rangle - \langle \psi_i \psi_j | \frac{1}{r_{12}} | \psi_j \psi_i \rangle) \tag{2.24}$$

$$E_0^{RHF} = 2 \sum_i^{n/2} \langle \psi_i | \hat{h} | \psi_i \rangle + \sum_{ij}^{n/2} (2 \langle \psi_i \psi_j | \frac{1}{r_{12}} | \psi_i \psi_j \rangle - \langle \psi_i \psi_j | \frac{1}{r_{12}} | \psi_j \psi_i \rangle) \tag{2.25}$$

Where $E_0^{HF}$ represents the Hartee-Fock energy for one-electron spin orbitals and $E_0^{RHF}$ represents the Hartree-Fock energy for restricted closed shell spatial orbitals.[93]

### 2.4.4 Hartree-Fock Matrix Equation

In a closed shell system, due to the orthogonality of the spin functions, the Hartree-Fock equation for the $i$th electron can be written as the following:

$$\hat{F}_i \psi_i(r_i) = \varepsilon_i \psi_i(r_i) \tag{2.26}$$

$\hat{F}_i$ denotes the Fock operator, $\varepsilon_i$ is an eigenvalue, which denotes the orbital energy. $\psi_i$ represents the wavefunction of the $i$th molecular orbital. As a solution to obtain the eigenvalue and wavefunction analytically, a set of functions known as basis functions are introduced.

Let us now consider the Hartree-Fock equations in their spatial form. In another approximation, the spatial functions $\psi_i$ functions are expanded with a finite number of basis functions. Basis functions are a set of functions that can be used to represent the electronic wavefunction. Introducing a basis set transforms the Hartree-Fock equations into the Roothaan equations.

$$\psi_i(r_i) = \sum_{\lambda=1}^{N^\lambda} C_{\lambda_i} \phi_\lambda \tag{2.27}$$

Where $N_\lambda$ is a number of basis functions. The number of basis functions is usually larger than

the number of electrons. $C_{\lambda_i}$ is an expansion coefficient and $\phi_\lambda$ is a defined basis function. By introducing the basis set the Hartree-Fock equation can be rewritten:

$$\hat{F}\sum_{\lambda=1}^{N^\lambda} C_{\lambda_i}\phi_\lambda = \varepsilon_i \sum_{\lambda=1}^{N^\lambda} C_{\lambda_i}\phi_\lambda \tag{2.28}$$

Multiplying both sides of equation 2.28 by $\phi_\gamma^*$ and integrating both sides yields the *Roothaan-Hall* equations.[106, 107]

$$\sum_{\lambda=1}^{N^\lambda} C_{\lambda_i}\langle\phi_\gamma|\hat{F}|\phi_\lambda\rangle = \varepsilon_i \sum_{\lambda=1}^{N^\lambda} C_{\lambda_i}\langle\phi_\gamma|\phi_\lambda\rangle \tag{2.29}$$

The multiplication and integration of equation 2.28 allows us to define the Fock matrix $F_{\gamma\lambda}$ and the overlap matrix $S_{\gamma\lambda}$.

$$F_{\gamma\lambda} = \langle\phi_\gamma|\hat{F}|\phi_\lambda\rangle \tag{2.30}$$

$$S_{\gamma\lambda} = \langle\phi_\gamma|\phi_\lambda\rangle \tag{2.31}$$

By employing the notation used in equations 2.30 and 2.31, equation 2.29 can be rewritten as the following:

$$\sum_{\lambda=1}^{N^\lambda} F_{\gamma\lambda}C_{\lambda_i} = \varepsilon_i \sum_{\lambda=1}^{N^\lambda} S_{\gamma\lambda}C_{\lambda_i} \tag{2.32}$$

Equation 2.32 can be written is a simpler form as a matrix equation, also known as the *Hartree-Fock-Roothaan-Hall* equation.

$$\mathbf{FC} = \mathbf{SC}\varepsilon \tag{2.33}$$

Where $\mathbf{F}$ is the Fock matrix, $\mathbf{S}$ is the overlap matrix, $\mathbf{C}$ contains the expansion coefficients and $\varepsilon$ is the diagonal matrix element which comprises of orbital energies.[93, 107, 108]

### 2.4.5 The Self Consistent Field Method

By varying the expansion coefficients $C_{\lambda i}$ and calculating the orbital energies $\epsilon_i$, the lowest energy solution can be obtained. This is known as the self-consistent field method. The Fock operator, $\hat{F}_i$, contains within it a description of electron electron interactions. Thus, the Fock operator depends on the orbitals i.e. on the coefficients $C_{\lambda i}$. These are calculated using an iterative method starting with an educated guess for $C_{\lambda i}$ (unknowns). Firstly, the initial guess for the unknown $C_{\lambda i}$ is made and then the Fock operator, $\hat{F}_i$ is built. The Fock equations are then solved to yield a new set of orbitals which are then used to construct new Fock operators, $\hat{F}_i$. The cycle is then repeated until there is no significant difference from one step to another. The cycle also employs the use of the variational principle, which yields the lowest total energy.[10,93]

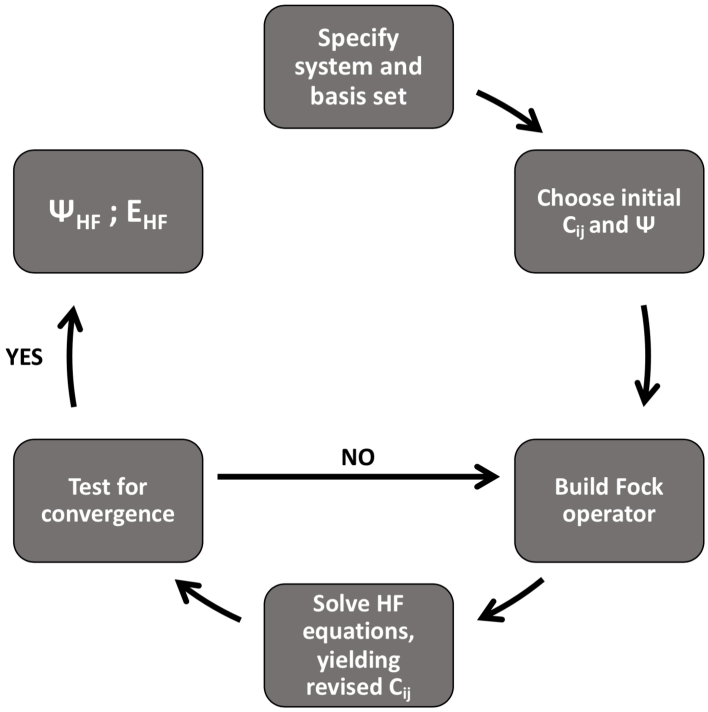



Figure 2.1: Hartree-Fock self-consistent field method schematic adapted from Harvey.[10]

## 2.5 Basis Sets

In order to carry out Hartree-Fock and essentially all other ab-initio quantum mechanical calculations the basis set approximation is required. In practical calculations a linear combination of atomic orbitals is used to express the one electron wavefunction (as described in section 2.4.4). To obtain the highest level of accuracy, one would choose a basis set that is infinite in size (complete basis set). However, a basis set of infinite in size is impossible to obtain in actual calculations. An unknown molecular orbital can be treated as a function, spanned by the whole basis set. When a finite basis set is used, only the components of the molecular orbital along the coordinate axes

corresponding to selected basis functions can be represented. Thus, the smaller the basis the poorer the representation.[10,106]

### 2.5.1 Slater Type Basis Sets

Slater type orbitals (STOs)[109] are functions used as atomic orbitals. They, resemble the wavefunction of the hydrogenic atom. STOs have a cusp at the nucleus and an exponential decay of the from $e^{-\zeta r}$. The Cartesian form of an STO is the following:

$$\chi_{STO} = Nx^i y^j z^k e^{-\zeta r} \tag{2.34}$$

$N$ is the normalisation constant and $x^i y^j z^k$ is the angular component. STOs are primarily used for atomic and diatomic systems. The main advantage of an STO is that it has the correct radial form. However, STOs do not allow us to obtain analytical solutions for two-electron integrals. On the other hand Gaussian type orbitals are able to overcome this.[93,106]

### 2.5.2 Gaussian Type Basis Sets

Gaussian type orbitals (GTOs)[110] were developed as a solution to the limitations of STOs. By using a GTO it is possible to calculate the Coulomb and exchange integrals analytically. GTOs can be written as the following:

$$\chi_{GTO} = Nx^i y^j z^k e^{-\alpha r^2} \tag{2.35}$$

The origin of the coordinates is the nucleus. $N$ is the normalisation constant and $x^i y^j z^k$ is the angular component. $-\alpha r^2$ is the exponent, note the notation has changed from $\zeta$ (for an STO) to $\alpha$ (for a GTO) to avoid confusion. $i$, $j$ and $k$ are non-negative integers that determine the Cartesian nature of the orbital. The main disadvantage of a GTO is that unlike an STO they give a poor representation of radial wavefunction near the atomic nucleus. Gaussian orbitals also decay at a faster rate than Slater orbitals.[93]

In order to improve the radial wavefunction of a GTO, one orbital is expressed by a linear combination of GTOs (basis functions). This process is known as contraction.

$$\sum_{\mu=1}^{L} d_\mu \chi_a^{Gauss} \tag{2.36}$$

### 2.5.3 Classification of Basis Sets

After selecting either an STO or GTO, the most important factor is determining the number of functions to represent an orbital. The smallest number of functions yields the *minimum basis set*.[106] The *minimum basis set* is the smallest basis set that could be used to obtain reasonable molecular orbitals for a particular molecule.[10] For example, for hydrogen only a single *s*-function would be used. For lithium and beryllium only two *s*-functions are required. However, one set of *p*-functions are typically added ($2p_x$, $2p_y$ and $2p_z$).

The next improvement on the basis set is the doubling of all basis functions. Thus, producing a *Double Zeta* (DZ) basis set. The zeta term stems from the exponent term of an STO basis set (denoted $\zeta$). Therefore, for Hydrogen a DZ basis set employs two *s*-functions (1s and 1s'). After a DZ basis set the next step up in the basis set range is the *Triple Zeta* (TZ) basis set. Such a basis set contains three times as many functions as the minimum basis functions.[106]

When a covalent bond is formed between different orbitals electron density distribution is different from isolated atoms. Therefore, basis functions with higher angular momentum (polarisation functions) are added as a correction. Adding a set of polarization functions (*p*-functions on hydrogens and *d*-functions on heavy atoms) results in the DZ basis set becoming the *Double Zeta Polarized* (DZP) basis set.[106]

In the highest spin-state electrons are allocated in the outer shell, which is unoccupied in the lowest spin-state. The extra basis function is known as the *diffuse basis function*. They are added to represent excited state electrons or anions.[93]

### 2.5.4 Basis Set Superposition Error

Using an insufficiently large basis set can result in an incorrect description of molecular orbitals, which can also lead to errors in calculated energies. One way in which this can occur is when calculating the binding energy between two molecules. The basis set superposition error (BSSE) is an artificial increase in the binding energy. This increase arises due to the use of incomplete atom centred basis sets. The binding energy between molecule $A$ and molecule $B$ can be calculated using the following formula:

$$\Delta E_{Bind} = E^{ab}(A-B) - E^{a}(A) - E^{b}(B) \tag{2.37}$$

$A$-$B$ is the complex of molecules $A$ and molecule $B$, $a$ and $b$ are the basis sets associated with molecules $A$ and molecule $B$ respectively. If $a$ and $b$ are finite sets of atom centred functions the

basis set used in the calculation of the complex ($ab$) will be larger than both $a$ and $b$. As a larger basis set provides more variational freedom, the energy of the complex may be artificially lowered. This is relative to the individual molecules themselves.[10,111]

A commonly used method for correcting the BSSE is the counterpoise method. Applying one form of the correction[112,113] to equation 2.37 yields:

$$\Delta E_{Bind}^{CP} = E^{ab}(A-B) - E^{ab}(A) - E^{ab}(B) \tag{2.38}$$

By implementing this form some of the basis functions are present without the corresponding atom. The basis functions of species $B$ are present in the calculation of species $A$, and the basis functions of species $A$ are present in the calculation of species $B$. These basis functions are present without the atoms corresponding to these functions being present.[113]

## 2.6 Density Functional Theory

Density functional theory (DFT) is a popular widely used method in the world of modern quantum mechanics.[114,115] DFT takes a different approach to solve the Schrödinger equation. Instead of using the electronic wavefunction, $\psi$, the Schrödinger equation is solved using the electron density, $\rho(\mathbf{r})$.[10,93,106,116] For a many-electron system the use of the electron density is a much simpler method than the use of the wavefunction.[10] In DFT the electron density is used to compute the properties of a system rather than the wavefunction in principle. A functional is essentially a function that maps a function to a number.

The origins of DFT can be traced back to the Thomas-Fermi method.[117–120] Thomas and Fermi recognised the basic nature of the electron density, which they then applied to atoms.[119,120] The Thomas-Fermi theory stated that the kinetic energy density at any particular point is approximated by that of a uniform electron gas of non-interacting electrons. Dirac expanded on the work of Thomas and Fermi by introducing an exchange correlation term, thus leading to the Thomas-Fermi-Dirac energy functional.[116] A major shortcoming of the above models was their inability to predict chemical bonds between atoms, Hence, the methods were not of much use to chemists.[121] Building upon the work of their predecessors Hohenberg and Kohn were able to formulate the first and second Hohenberg and Kohn theorems. These two theorems are the two main theorems in DFT.

### 2.6.1 The First Hohenberg-Kohn theorem

The first Hohenberg-Kohn theorem states: The external potential, $\nu(\mathbf{r})$ and hence the total energy, is a unique potential of the electron density $\rho(\mathbf{r})$.[122] Hohenberg and Kohn derived their first theorem by considering a collection of an arbitrary number of electrons enclosed in a large box moving under the external potential $\nu(\mathbf{r})$ and the mutual Coulomb repulsion.[122,123] The Hamiltonian has the form

$$H = T + V + U \tag{2.39}$$

where $T$ is the N-electron kinetic energy. $V$ is the N-electron potential energy (from the external field) and $U$ is the electron-electron interaction energy. in the ground state the electron density has the following expression

$$\rho(\mathbf{r}) = (\Psi, \psi^*(\mathbf{r})\psi(\mathbf{r})\Psi) \tag{2.40}$$

Hence from the above equation $\rho(\mathbf{r})$ is a function of $\nu(\mathbf{r})$ because the wavefunction, $\psi(\mathbf{r})$ can be determined by $\nu(\mathbf{r})$. As $\nu(\mathbf{r})$ is a component of $H$[122,123]

This can be proved by assuming the existence of an alternative potential $\nu^{'}(\mathbf{r})$ with the ground-state wavefunction, $\psi^{'}$, that also produces the electron density $\rho^{'}(\mathbf{r})$. $\psi$ and $\psi^{'}$ are distinct since:

$$\nu(\mathbf{r}) - \nu^{'}(\mathbf{r}) \neq 0 \tag{2.41}$$

the above equation holds as $\psi \neq \psi^{'}$. Hence both $\psi$ and $\psi^{'}$ satisfy two different Schrödinger equations. $\psi$ satisfies

$$\hat{H}\psi = E\psi \tag{2.42}$$

and $\psi^{'}$ satisfies:

$$\hat{H}^{'}\psi^{'} = E^{'}\psi^{'} \tag{2.43}$$

Therefore, the above Schrödinger equations have different Hamiltonian's and ground state ener-

gies.[122] Using the variational property the ground state electronic energy we obtain:

$$E^{'} = \langle \psi^{'} | \hat{H}^{'} | \psi^{'} \rangle < \langle \psi | \hat{H} | \psi \rangle + \langle \psi | \hat{H}^{'} - \hat{H} | \psi \rangle \tag{2.44}$$

so that

$$E^{'} < E + \int [\nu^{'}(\mathbf{r}) - \nu(\mathbf{r})] \rho(\mathbf{r}) d\mathbf{r} \tag{2.45}$$

By simply interchanging the primed and unprimed quantities we obtain

$$E < E^{'} + \int [\nu(\mathbf{r}) - \nu^{'}(\mathbf{r}] \rho(\mathbf{r}) d\mathbf{r} \tag{2.46}$$

Addition of the above two equations yields the following

$$E^{'} + E < E + E^{'} \tag{2.47}$$

which cannot hold because if $E = E^{'}$ then you get 0<0. Therefore $\nu^{'}(\mathbf{r})$ does not come from the same $\rho(\mathbf{r})$. Thus, proving that $\nu(\mathbf{r})$ is a unique functional of $\rho(\mathbf{r})$. Hence, the many electron ground state is a unique functional of the electron density $\rho(\mathbf{r})$.[122,123]

### 2.6.2 The Second Hohenberg-Kohn Theorem

The second theorem states that the lowest energy is obtained if the electron density is the true ground-state electron density. Their second theorem further proved that by taking the energy functional as a function of the the electron density allows the ground state energy to be found. The universal energy functional can be defined as

$$F[\rho(\mathbf{r})] = \langle \psi | T + V | \psi \rangle \tag{2.48}$$

where $F[\rho(\mathbf{r})]$ is a universal functional that is valid for any number of particles and any potential.[123] When given the external potential $\nu(\mathbf{r})$ the energy functional can be expressed as

$$E_{\nu}[\rho(\mathbf{r})] = \int \nu(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r} + F[\rho(\mathbf{r})] \tag{2.49}$$

By definition the above equation is equal to the energy functional of the wavefunction. This can be written as

$$\epsilon_\nu[\psi^{'}] = \langle \psi^{'} | V | \psi^{'} \rangle + \langle \psi^{'} | T + V | \psi^{'} \rangle \tag{2.50}$$

where $\epsilon_\nu[\psi^{'}]$ is the energy functional of the wavefunction. $\epsilon_\nu[\psi^{'}]$ attains the ground state electron energy as its minimum when $\psi^{'}$ is the ground state wavefunction for a system of N electrons according to the variational principle. If $\psi^{'}$ denotes the ground state wavefunction that corresponds to the external potential $\nu^{'}(\mathbf{r}) \neq \nu(\mathbf{r})$ then

$$\epsilon_\nu[\psi^{'}] > \epsilon_\nu[\psi] = E \tag{2.51}$$

Hence,

$$\epsilon_\nu[\rho^{'}(\mathbf{r}] > \epsilon_\nu[\rho(\mathbf{r}] = E \tag{2.52}$$

Thus, proving that $\epsilon_\nu[\rho(\mathbf{r}] < \epsilon_\nu[\rho^{'}(\mathbf{r}]$ for all $\rho^{'}(\mathbf{r})$ associated with $\nu^{'}(\mathbf{r})$. Where $\rho(\mathbf{r})$ is the exact electron density associated with the ground state electronic energy. The second Hohenberg and Kohn theorem, therefore provides a variational principle for the ground state directly in terms of the electron density. Thus, allowing the electron density to be used as a basic variable in quantum mechanical calculations.[122,123]

### 2.6.3 Kohn-Sham Theory

A vital flaw of Hohenberg and Kohn's theory was its inability to accurately represent the kinetic energy. Kohn and Sham[120] further developed DFT methods by introducing orbitals into the method. They provided a form for the Hohenberg-Kohn (HK) functional which avoided the explicit use of the kinetic energy. In the Kohn-Sham method the kinetic energy functional is split into two parts, one that can be calculated exactly and one small correction term. The Kohn-Sham method is closely related to the Hartree-Fock method. Thus, sharing the identical form for the kinetic, electron-electron, nuclear-electron and Coulomb energies.

Splitting the kinetic energy into two parts can be justified as follows. If we assume that the Hamiltonian operator takes the form below with $0 \leq \lambda \leq 1$.

$$H_\lambda = T + V_{ext}\lambda + \lambda V_{ee} \tag{2.53}$$

where $\lambda$ is a constant between 0 and 1 and represents the level of interaction in the system. $H$ is the Hamiltonian, $V_{ext}$ is the external potential and $V_{ee}$ is the electron-electron Coulomb interaction energy. The external potential $V_{ext}$ is equal to $V_{ne}$ when $\lambda = 1$. $V_{ne}$ is the nuclear-electron Coulomb interaction energy. When $\lambda = 0$, the electrons are non-interacting. In this case, the exact solution to the Schrödinger equation is given as a Slater determinant composed of molecular orbitals ($\psi_i$). The exact kinetic energy is given as follows[106]

$$T_s = \sum_{i=1}^{Nelec} \langle \psi_i | -\frac{1}{2}\nabla^2 | \psi_i \rangle \tag{2.54}$$

Where $T_s$ is the kinetic energy calculated from a Slater determinant. At $\lambda = 1$ the system is contains interacting electrons. This equation can be modified to obtain the exact kinetic energy of the system, by calculating using natural orbitals ($NO$), which arise from the exact density matrix. Thus, the exact kinetic energy can be written as follows

$$T[\rho_{exact}] = \sum_{i=1}^{\infty} \langle \psi_i^{NO} | -\frac{1}{2}\nabla^2 | \psi_i^{NO} \rangle \tag{2.55}$$

$$\rho_{exact} = \sum_{i=1}^{\infty} n_i |\psi_i^{NO}|^2 \tag{2.56}$$

where

$$N_{elec} = \sum_{i=1}^{\infty} n_i \tag{2.57}$$

where $\rho_{exact}$ is the exact density and $n_i$ is an eigenvalue of the density matrix that has values between 0 and 1 and corresponds to the number of electrons in the spin orbital. Representing the exact density requires an infinite set of natural orbitals. An approximation is made as the exact matrix density remains unknown. Therefore, the novel approximate density matrix is now written as one electron functions.

$$\rho_{approx} = \sum_{i=1}^{Nelec} |\psi_i|^2 \tag{2.58}$$

$\rho_{approx}$ is the approximated density matrix and $\psi_i$ are the one-electron functions. this equation directly relates to equations 2.55 - 2.57 with occupation numbers 0 or 1.[106]

The key to Kohn-Sham theory is to calculate the kinetic energy, under the assumption that the electrons in the system are non-interacting. Realistically electrons are interacting. However, the difference between exact kinetic energy and the kinetic energy calculated using the non-interacting orbital approximation is small. The remaining kinetic energy is added into a term known as the exchange correlation term. As well as the kinetic energy, the exchange correlation term contains the non-classical electron-electron interaction energy. Hence, we can now specify a general expression for the DFT energy as follows

$$E_{DFT}[\rho] = T_s[\rho] + E_{ne}[\rho] + J[\rho] + E_{xc}[\rho] \tag{2.59}$$

$E_{DFT}$ is the DFT energy. $E_{ne}[\rho]$ and $J[\rho]$ are potential energy terms, with $J[\rho]$ defined as the Coulomb energy. $E_{xc}[\rho]$ is the exchange correlation term. Equating $E_{DFT}$ to the exact energy enables the $E_{xc}$ term to be defined. The $E_{xc}$ term is the result of subtracting the kinetic energy (non-interacting) and the $E_{ne}[\rho]$ and $J[\rho]$ potential terms from the total DFT energy. This can be written as the following[106]

$$E_{xc}[\rho] = (T[\rho] - T_s[\rho]) + (E_{ee}[\rho] - J[\rho]) \tag{2.60}$$

The first set of parenthesis in equation 2.60 are considered as the kinetic correlation energy and the second set contains both potential correlation and exchange energy. $E_{ee}$ denotes all of the electron-electron interaction energy.

In principle, DFT is able to provide an exact solution of the Schrödinger equation. However, the universal exchange-correlation energy is not defined at present, therefore an exact solution is difficult to obtain. During the introduction of basis sets into the Kohn-Sham equation the expansion coefficients are obtained and orbital energies are calculated using SCF process.[93] Several exchange correlation functionals have been developed over the last fifty years in an attempt to produce a representative solution to the many-body problem.

### 2.6.4 The Exchange-Correlation Functional

Kohn-Sham theory introduced the exchange-correlation functional. The main difference between different DFT methods is the functional form that the exchange-correlation functional takes. An

explicit form of the functional is not known except for the case of a uniform electron gas. Nevertheless, one can derive properties that the exchange-correlation functional should have.[106,124]

One of the first properties is that the exchange-correlation functional should be self-interaction free. For example, in the case of a one electron hydrogen atom the Coulomb electron-electron energy should cancel and the correlation energy should be zero. Another property is that the coordinate scaling of the exchange energy should be linear. So multiplying by a constant should yield a similar linear scaling of the exchange energy.[106] However, no direct scaling applies to the correlation energy, but scaling the electron coordinates by a factor greater than one should increase the correlations magnitude.[106,125] Another property of the exchange-correlation functional is that as the scaling parameter tends to infinity the correlation energy for a finite system approaches a negative constant. The final properties relate to asymptotic behaviour. The exchange potential should exhibit asymptotic behaviour.[106,126] The correlation potential should also show asymptotic behaviour.[106]

The exchange-correlation functional can be separated into its exchange and correlation parts, as shown below and expressions for the exchange and correlation parts derived separately based on the principles listed above

$$E_{xc}[\rho] = E_x[\rho] + E_c[\rho] \tag{2.61}$$

### 2.6.5 Local Density Approximation

The *local density approximation* (LDA) method is amongst the oldest forms of the exchange-correlation functional.[106] The LDA functional was proposed by Hohenberg and Kohn in 1964.[122] The LDA comprises of locally approximating the true exchange-correlation energy of a system by the exchange-correlation energy associated with a uniform electron gas. The uniform electron gas is the only system for which the exchange-correlation energy is known exactly. The exchange-correlation energy for the electron gas of a uniform density can be presented as:

$$E_{XC}^{LDA}[\rho(\mathbf{r})] = \int \rho(\mathbf{r})\varepsilon_{XC}^{hom}[\rho(\mathbf{r})]d\mathbf{r} \tag{2.62}$$

Where $\varepsilon_{XC}^{hom}$ denotes the exchange correlation energy density and $\rho(\mathbf{r})$ is the electron density. Equation 2.62 shows the general principle of how the exchange-correlation functional depends on the exchange-correlation energy density and electron density in LDA.

LDA functionals have been found to give accurate computational results for solid-state physics. However, they severely overbind chemical bonds and overestimate barrier heights. Therefore, they are not a useful model within chemistry.[93]

### 2.6.6 Generalised Gradient Approximation

The LDA method was improved through the development of a method that considered a non-uniform electron gas. Hohenberg and Kohn presumed that the LDA functional was too simple for modelling atoms and molecules, so they proposed the *generalised expansion approximation* (GEA).[122] The GEA is a series expansion of higher order density gradient terms. The first order density gradient was tested for atoms and molecules, this procedure failed. Despite the failure of the GEA method, it became the basis for the *Generalised gradient approximation* (GGA), which is currently the most popular exchange-correlation functional in condensed matter physics. Unlike the LDA method, in the GGA method the exchange and correlation energies depend on not only the electron density, but also on the derivatives of the electron density. The first derivative of the electron density is included as a variable.[106] The GGA method improves on the LDA description of atoms, molecules and solids.[127] Unlike LDA functionals, GGA functionals tend to improve the description of the energy,[127,128] atomisation energies[128,129] and energy barriers.[130–132] GGA functionals also expand and soften bonds,[127] this is an effect that is able to correct[133] and over-correct the LDA prediction.[134]

The principles of the GGA functionals were developed by Perdew and his coworkers.[135] They designed a cutoff procedure. This cutoff procedure terminates the GEA exchange correlation hole in real space. This is achieved by using delta functions. As a result, the GGA exhange correlation functional can be written in terms of a function known as the enhancement factor, $F_{XC}[\rho(\mathbf{r}), \nabla\rho(\mathbf{r})]$, that modifies the LDA energy density.

$$E_{XC}^{GGA}[\rho(\mathbf{r})] = \int (\mathbf{r})\varepsilon_{XC}^{hom}[\rho(\mathbf{r})]F_{XC}[\rho(\mathbf{r}), \nabla\rho(\mathbf{r})]d\mathbf{r} \tag{2.63}$$

Perdew and his coworkers developed several exchange correlation functionals. The exchange correlation functionals they developed have the acronyms PW86 (Perdew-Wang),[136] PW91 (Perdew-Wang 1991)[137,138] and PBE (Perdew-Burke-Erzerhof).[127] These three functionals are considered to be refinements of the same model.

PW91[137] is an important GGA functional within solid-state DFT. PW91 is a non-empirical functional, as it does not contain any free parameters, rather it is determined from exact quantum

mechanical relations. In PW91 the enhancement factor, $F_X^{PW91}(s)$ has the form:

$$F_X^{PW91}(s) = \frac{1 + 0.19645s\ sinh^{-1}(7.7956s) + (0.2743 - 0.15084e^{100s^2})s^2}{1 + 0.19645s\ sinh^{-1}(7.7956s) + 0.004s^4} \tag{2.64}$$

Equation 2.64 is an extension of the work conducted by Becke[139] who developed the B88 functional. For some time, PW91 was the main functional used in quantum physics. However, it was superseded by a modified functional known as the PBE functional, which was devised by Perdew, Burke and Erzerhof.[127] The exchange part of the functional ($\epsilon_{XC}^{PBE}$) is written as an enhancement factor multiplied by the LDA functional.[106]

$$E_{XC}^{PBE} = E_{XC}^{hom} F_{XC}[\rho(\mathbf{r}), \nabla\rho(\mathbf{r})] \tag{2.65}$$

The PBE functional uses a simplified enhancement factor of the form:

$$F_X^{PBE} = 1 + \kappa - \frac{\kappa}{1 + \mu s^2/\kappa'} \tag{2.66}$$

Where $\mu$=0.21951 and $\kappa$=0.804. PBE is a simple functional which retains only the most energetically important conditions satisfied by PW91.[127]

### 2.6.7 Meta-Generalised Gradient Approximation

GGA functionals have been modified further to improve them. Including higher order density functionals improves the accuracy of calculations. These improve on traditional GGA functionals by allowing the exchange and correlation functionals to depend on higher order derivatives of the electron density, where the Laplacian is the second order term.[106,140,141] Alternatively the functional can be taken to depend on the orbital kinetic energy $\tau$.[106] When employing the use of meta-GGA functionals numerical stability is improved via the introduction of a dependence on the kinetic energy, $\tau$.

$$\tau(\mathbf{r}) = \frac{1}{2} \sum_{i}^{occupied} |\nabla\psi_i^2(\mathbf{r})|^2 \tag{2.67}$$

The meta-GGA functionals may be written in the general form:

$$E_{XC}^{meta-GGA}[n(\mathbf{r})] = \int f[\rho(\mathbf{r}), \nabla\rho(\mathbf{r}), \nabla^2\rho(\mathbf{r}), \tau(\mathbf{r}), \mu(\mathbf{r}), ...\gamma(\mathbf{r})]d(\mathbf{r}) \tag{2.68}$$

Where $\mu(\mathbf{r}), ...\gamma(\mathbf{r})$ are other quantities that can be used in the construction of meta-GGA's. PKZB[142] and TPSS[143] are examples of meta-GGA functionals.

### 2.6.8 Hybrid Functionals

The final class of functionals to consider are hybrid functionals.[144] Hybrid functionals combine Hartree-Fock like exchange with conventional GGA exchange and correlation. The development of hybrid functionals came from Becke's introduction of the *Half and Half* method.[106, 144]

$$E_{XC}^{H+H} = \frac{1}{2}E_X^{exact} + \frac{1}{2}(E_X^{LDA} + E_C^{LDA}) \tag{2.69}$$

where $E_{XC}^{H+H}$ is the half and half exchange correlation energy, $E_X^{exact}$ is the exact DFT energy. The $E_X^{LDA}$ and $E_C^{LDA}$ terms denote the exchange and correlation energy terms for the LDA functional. Since GGA methods yield more accurate results than LDA methods, a generalised version of the *Half and Half* method may be defined. It may be defined by writing the exchange energy as a combination of the exact LDA exchange energy and a gradient correction term.[106]

The general form of the exchange correlation functional for hybrid functionals has the form:

$$E_{XC}^{hybrid} = \alpha(E_X^{HF} - E_X^{GGA}) + E_{XC}^{GGA} \tag{2.70}$$

Where $\alpha$ determines the amount of exact exchange mixing. B3LYP[129, 144–146] and B97[147] are both examples of hybrid functionals. The Becke, Lee, Yang and Parr B3LYP[129, 144–146] functional is amongst one of the most popular hybrid functionals and it has an exchange correlation functional of the form[144, 148]

$$E_{XC}^{B3LYP} = (1-a)(E_X^{LDA} + aE_X^{HF} + b\Delta E_X^B + (1-c)E_c^{GGA}) + cE_c^{LYP} \tag{2.71}$$

Where a = 0.2, b = 0.72 and c = 0.81 were fitted to a training set of experimental atomisation energies, ionisation potentials, proton affinities and total atomic energies. The B3LYP functional proposed in 1993 is still one of the most successful functionals due to its overall performance.[106]

### 2.6.9 Dispersion

Recent investigations have found that many functionals lead to incorrect energies and structure for large molecules.[10] This is because of the poor treatment of the physically and chemically important London dispersion forces.[149,150] Inclusion, of London dispersion forces in theoretical simulations of weakly bonded Van der Waals systems[151] is important to reach chemical accuracy.[149] Inclusion of these interactions also influences the accuracy of theoretical thermodynamics.[152–154] London dispersion forces are the interaction forces between an instantaneous dipole that arises as a random fluctuation in one part of a system with an instantaneous dipole in another part of the system. In small molecules, these interactions are negligible. However, they are no longer negligible in large molecular systems, as their effects become larger.[10] A solution is to therefore add an *empirical* dispersion correction term, DFT with added dispersion correction is often labelled as DFT-D. The general form of the dispersion energy is[149]

$$E_{disp}^{DFT-D} = -\frac{1}{2}\sum_{A\neq B}\sum_{6,8,10,\ldots} S_n \frac{C_n^{AB}}{R_{AB}^n} f_{damp}(\mathbf{R}_{AB}) \tag{2.72}$$

The sum is over all atom pairs in the system. AB represents the atom pair AB and $\mathbf{R}_{AB}$ in the inter-nuclear distance between atom A and atom B. $C_n^{AB}$ denotes the averaged (isotropic) $n^{th}$ order dispersion coefficient of orders $n = 6, 8, 10, ....$ $S_n$ denotes the global scaling factors, they can be adjusted to adjust the correction to the repulsive behaviour of the exchange correlation functional in use.[155] $f_{damp}$ is the damping function. It determines the short range behaviour of of the correction, it is needed to avoid near-singularities at small $\mathbf{R}_{AB}$ values. It is also needed to remove double counting effects of electron correlation at intermediate $\mathbf{R}_{AB}$ distances.[149] Typical expressions for $f_{damp}$ are[156]

$$f_{damp}(\mathbf{R}_{AB}) = \frac{1}{1 + 6(\mathbf{R}_{AB}/(S_{r,n}R_0^{AB}))^{-\gamma}} \tag{2.73}$$

or[157,158]

$$f_{damp}(\mathbf{R}_{AB})' = \frac{1}{1 + e^{-\gamma(\mathbf{R}_{AB}/S_{r,n}R_0^{AB}-1)}} \tag{2.74}$$

where $R_0^{AB}$ is a cut off radius for the atom pair AB and $S_{r,n}$ is a functional dependent (global) radii scaling factor.[159] $\gamma$ is a parameter that determines the steepness of the function for small $\mathbf{R}_{AB}$ distances.[149]

There are several different types of dispersion correction D1,[155] D2,[158] and D3.[160] The total dispersion-corrected DFT energy is given by

$$E_{DFT-D3} = E_{KS-DFT} + E_{disp} \tag{2.75}$$

$E_{Ks-DFT}$ is the usual self-consistent Kohn-Sham energy as obtained for the chosen density functional. $E_{disp}$ is the dispersion correction as a sum of two and three body energies (for D3),

$$E_{disp} = E_2 + E_3 \tag{2.76}$$

the most important two-body term is given by

$$E^2 = \sum_{AB} \sum_{6,8,10,\ldots} S_n \frac{C_n^{AB}}{R_{AB}^n} f_{damp}(\mathbf{R}_{AB}) \tag{2.77}$$

Global scaling factors, $S_n$ are adjusted for n>6 to ensure asymptotic exactness which is fulfilled when the $C_6^{AB}$ are exact. This is the fundamental difference between the DFT-D3 approach and the DFT-D1 and DFT-D2 approaches, whereby $S_6$ was not equal to unity and only a scaled asymptotic value is obtained.[160] The DFT-D3 method will be utilised in our investigations.

## 2.7 Periodic systems

### 2.7.1 Structures of Crystals

Real systems, such as mineral particles, are macroscopic objects, constituted by a large number of atoms. They can effectively be viewed as infinite.[161] In a crystalline solid a small number of atoms, also known as a *basis* are repeated periodically in space.[161,162] There are many ways to choose a basis. When considering the whole symmetry of the crystalline solid the minimum unit of the whole system, is given the term *unit cell* or *Wigner-Seitz cell.* combining the unit cell with lattice vectors enables reproduction of the crystalline structure (infinitely). The vectors that produce the infinite crystal are termed *unit* or *primitive* vectors.[161] The set of translations corresponding to combinations of primitive vectors is known as the Bravais lattice.[162] The translation, $\mathbf{T}$, can be written as multiples of primitive vectors.

$$\mathbf{T} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3 \tag{2.78}$$

where $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ are the primitive lattice vectors of the direct unit cell. $n_1$, $n_2$ and $n_3$ are integers. Each lattice has a *reciprocal lattice*. The reciprocal lattice has lattice vectors $\mathbf{a}_1^*$, $\mathbf{a}_2^*$ and $\mathbf{a}_3^*$ which follow orthogonality rules relative to the direct lattice vectors. The primitive vectors in the reciprocal space are defined by the following relation

$$\mathbf{a}_i \cdot \mathbf{a}_j^* = 2\pi\delta_{ij}$$

where $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$.

In one dimension[162]

$$\mathbf{a}_i^* = \frac{2\pi}{\mathbf{a}_i} \tag{2.79}$$

In three-dimensions

$$\mathbf{a}_i^* = 2\pi \frac{\mathbf{a}_j \times \mathbf{a}_k}{\mathbf{a}_i \cdot \mathbf{a}_j \times \mathbf{a}_k} \tag{2.80}$$

Amongst the possible choices of the unit cell in a reciprocal lattice that are equivalent, the preferred choice is the *Brillouin zone*.[163] This can be obtained by connecting a point in the reciprocal lattice to neighbouring points and passing orthogonal planes through their midpoints. The volume between the planes is known as the *Brillouin zone*.[163]

### 2.7.2 Bloch's Theorem

Bloch's theorem[164] connects the properties of the electrons in the unit cell with those of the electrons in an infinite periodic system. The potential energy of a crystal must be a periodic function with the same periodicity as the lattice. So, the translation by any lattice vector for example $\mathrm{a}_i$, does not change the potential energy

$$V(\mathbf{r}) = V(\mathbf{r} + \mathbf{a}_i)$$

hence, the Schrödinger equation has the form

$$\hat{H}(\mathbf{r})\Psi(\mathbf{r}) = E\Psi(\mathbf{r}) \tag{2.81}$$

which is equivalent to

$$\hat{H}(\mathbf{r}+\mathbf{a}_i)\Psi(\mathbf{r}+\mathbf{a}_i) = E\Psi(\mathbf{r}+\mathbf{a}_i) \tag{2.82}$$

Therefore, the wavefunction can be written as a product of a function with the same periodicity of the potential, $U_{\mathbf{k}}(\mathbf{r})$ and the phase factor (or a plane wave), $e^{i\mathbf{k}(\mathbf{r})}$.

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{kr}}u_{\mathbf{k}}(\mathbf{r}) \tag{2.83}$$

where $u_{\mathbf{k}(\mathbf{r})}$ is periodic and equals

$$u_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r}+\mathbf{a}_i) \tag{2.84}$$

Hence, the periodic wavefunction has the form

$$\psi_{\mathbf{k}}(\mathbf{r}+\mathbf{a}_i) = e^{i\mathbf{ka}_i}\psi_{\mathbf{k}}(\mathbf{r}) \tag{2.85}$$

which means that the neighbouring cells differ only by the phase factor $e^{i\mathbf{k}(\mathbf{r})}$. The electron density by definition is equal to the probability density of the periodic wavefunction, when the complex conjugate is taken

$$\rho_{\mathbf{k}}(\mathbf{r}) = 2|\psi_{\mathbf{k}}(\mathbf{r})|^2 \tag{2.86}$$

Consequently, the electron density is also periodic.[161–163]

$$\rho_{\mathbf{k}}(\mathbf{r}+\mathbf{a}_i) = \rho_{\mathbf{k}}(\mathbf{r}) \tag{2.87}$$

Therefore, all information about the system's electron density is contained in a single unit cell (Brillouin zone). This allows us to use a single unit cell to model infinite periodic crystals using DFT methods as described in section 2.6.

## 2.8 Molecular Dynamics

In this section, we will discuss molecular dynamics (MD) simulations a class of computer simulations that aim to calculate the structure and the physical movements of atoms in a system.[165–169] MD simulations predict how every atom in a system will move. These simulations are often used to capture a variety of processes such as; ligand binding, change in geometry and in bio-molecular processes such as protein folding.[167] The first MD simulations were carried out on simple gases in the 1950s.[170] Over the years MD has gained popularity and visibility, particularly in biomolecule research.[167,171]

The basic principle behind MD simulations is fairly simple. Given the positions of atoms in a system (e.g. a hydrated surface surrounded by an organic molecule), one can calculate the force exerted on each atom by all of the other atoms in the system.[167] Therefore, one can use Newton's laws of motion to predict the position (spatial) of each individual atom as a function of time.[166,167,170,172,173]

### 2.8.1 Molecular Dynamics Algorithms

Nuclei are heavy enough, such that to a good approximation, they are believed to behave as classical particles. Thus, the dynamics can be simulated by solving Newton's second law:

$$\mathbf{F} = m\mathbf{a} \tag{2.88}$$

where $\mathbf{F}$ is the force, $m$ is the mass and $\mathbf{a}$ is the acceleration. This equation can be written in the following differential form

$$-\frac{dV}{d\mathbf{r}} = m\frac{d^2\mathbf{r}}{dt^2} \tag{2.89}$$

$V$ is the potential energy at position $\mathbf{r}$, where $\mathbf{r}$ contains the coordinates for all particles. $\mathbf{r}$ has Cartesian coordinates and it is a vector of length $3N_{atom}$. The left hand of the equation is the negative of the energy gradient, also known as the force, $\mathbf{F}$.[106]

At a small time step $\Delta$t the positions of a set of particles that were initially at position $\mathbf{r}_i$ at time $t$ can be given by the following Taylor expansion

$$\mathbf{r}_{i+1} = \mathbf{r}_i + \frac{\partial \mathbf{r}}{\partial t}(\Delta t) + \frac{1}{2}\frac{\partial^2 \mathbf{r}}{\partial t^2}(\Delta t)^2 + \frac{1}{6}\frac{\partial^3 \mathbf{r}}{\partial t^3}(\Delta t)^3 + ...$$

$$\mathbf{r}_{i+1} = \mathbf{r}_i + \mathbf{v}_i + \frac{1}{2}\mathbf{a}_i(\Delta t)^2 + \frac{1}{6}\mathbf{b}_i(\Delta t)^3 \tag{2.90}$$

$\mathbf{v}_i$ are velocities which are the first derivatives of the position with respect to time ($\frac{\partial \mathbf{r}}{\partial t}$) at time $\mathbf{t}_i$. The accelerations, $\mathbf{a}_i$, are the second derivatives ($\frac{\partial^2 \mathbf{r}^2}{\partial t}$) at time $\mathbf{t}_i$. The $\mathbf{b}_i$ terms are the hyper-accelerations and are the third derivatives and so forth. By substituting $\Delta$t with -$\Delta$t we obtain

$$\mathbf{r}_{i-1} = \mathbf{r}_i - \mathbf{v}_i + \frac{1}{2}\mathbf{a}_i(\Delta t))^2 - \frac{1}{6}\mathbf{b}_i(\Delta t))^3 \tag{2.91}$$

Addition of equations 2.90 and 2.91 give

$$\mathbf{r}_{i+1} = (2\mathbf{r}_i - \mathbf{r}_{i-1}) + \mathbf{a}_i(\Delta t))^2 + ... \tag{2.92}$$

$$\mathbf{a}_i = \frac{\mathbf{F}_i}{m_i} = -\frac{1}{m_i}\frac{dV}{d\mathbf{r}_i} \tag{2.93}$$

where the acceleration $\mathbf{a}_i$ can be derived from equations 2.88 and 2.89. This algorithm is known as the *Verlet algorithm*.[174–176] It solves Newton's equations numerically to obtain $\mathbf{r}_{i-1}$ at time $t + \Delta t$, if we know the coordinates $\mathbf{r}_i$ and $\mathbf{r}_{i-1}$ and accelerations assuming third order ($\mathbf{b}$) and higher terms are negligible. At the initial point, previous positions are not available, however they can be approximated from a first order approximation of equation 2.90

$$\mathbf{r}_{-1} = \mathbf{r}_0 - \mathbf{v}_0\Delta t \tag{2.94}$$

At each time step the acceleration must be evaluated from the forces (equation 2.93). This then allows the atomic positions to be propagated in time, thus, generating a trajectory. As the time-step size $\Delta$t decreases, the trajectory approximation becomes a better approximation of what is known as the "true trajectory". This occurs until practical problems arise, such as the issue of finite numerical accuracy. A small time step does however mean that more steps are needed for propagating the system. Hence, the computational expense increases with the size of the time step.

A key disadvantage of the Verlet algorithm is that the new positions are obtained by the addition

of a term proportional to $\Delta t^2$ to a difference in the positions $(2\mathbf{r}_i - \mathbf{r}_{i-1})$. As $\Delta t^2$ is a small number and $(2\mathbf{r}_i - \mathbf{r}_{i-1})$ is a large number as it is the difference between two large numbers, it can often lead to truncation errors due to finite precision. Further disadvantages of the Verlet algorithm are its inability to display velocities explicitly. This is often a problem when generating an ensemble, ensembles will be discussed further in the this section.

The introduction of the *Leap-frog*[177] algorithm deals with the numerical aspect of the Verlet algorithm and the inability to display velocities explicitly. By performing expansions analogous to those of the Taylor expansion (equation 2.90), with a half time step followed by subtraction we get

$$\mathbf{r}_{i+1} = \mathbf{r}_i + \mathbf{v}_{i+1/2}\Delta t \tag{2.95}$$

Similarly, the velocity is also obtained

$$\mathbf{v}_{i+1/2} = \mathbf{v}_{i-1/2} + \mathbf{a}_i\Delta t \tag{2.96}$$

Both equations 2.95 and 2.96 define the leap-frog algorithm and it can be seen that the velocity and position updates are out of phase. The numerical accuracy of the leap-frog algorithm is considered better than that of the Verlet algorithm. However, the disadvantage of the leap-frog algorithm is that the velocities and positions are not known at the same time. This drawback is however removed by the *velocity Verlet algorithm.*[174] The equations used to propagate atoms are given by

$$\mathbf{r}_{i+1} = \mathbf{r}_i + \mathbf{v}_i\Delta t + \frac{1}{2}\mathbf{a}_i\Delta t^2$$

$$\mathbf{v}_{i+1} = \mathbf{v}_i + \frac{1}{2}(\mathbf{a}_i + \mathbf{a}_{i+1})\Delta t \tag{2.97}$$

In this approach, both the velocities and position are known at the same time.

### 2.8.2 Generating Ensembles

Standard MD simulations generate what is known as an $NVE$ ensemble. In this ensemble temperature and pressure fluctuate while the total energy remains constant. The total energy is the sum of kinetic and potential energies, it can be calculated from the positions and velocities

$$E_{tot} = \sum_{i=1}^{N} \frac{1}{2}\mathbf{m}_i\mathbf{v}_i^2 + V(\mathbf{r}) \tag{2.98}$$

In practice the total energy may not be constant, however preservation of the energy within a given threshold may be achieved by defining the maximum (permissible) time-step.

Although the $NVE$ ensemble is the simplest ensemble, it is also possible to generate $NVT$ and $NPT$ ensembles. These ensembles are generated by modifying the positions or velocities in each time step to maintain constant temperature or pressure. Our work will utilise the $NVT$ ensemble, whereby the temperature of the system is fixed. The temperature of the system is is known to be proportional to the the average kinetic energy

$$E_{kinetic} \geq \frac{1}{2}(3N_{atom} - N_{constraint})kT \tag{2.99}$$

where $N_{atom}$ is the number of atoms and $N_{constraint}$ is the number of constraints. The number of constraints, typically three, correspond to the conservation of linear momentum. As the kinetic energy of the system is the difference between the total energy which is almost constant and the potential energy, which is dependent upon positions, the kinetic energy fluctuates. Hence, the temperature is constant, but has an average value with an associated fluctuation. Similarly, if the volume of a system is fixed, the pressure will be the property that fluctuates.

In the $NVT$ ensemble the system is coupled to a "heat bath". This heat bath gradually removes energy to or from the system with a suitable constant. This procedure is known as a *thermostat*. The kinetic energy of the system is modified (in this case) by scaling the velocities. The rate of heat transfer is controlled by a coupling constant $\tau$

$$\frac{dT}{t} = \frac{1}{\tau}(T_{desired} - T_{actual}) \tag{2.100}$$

$$velocity\ scaling\ factor = \sqrt{1 + \frac{\Delta t}{\tau}(\frac{T_{desired}}{T_{actual}} - 1)} \tag{2.101}$$

In the Nosé-Hoover methods[178,179] the heat bath is an integral part of the system.[106,180] The NVT ensemble corresponds to a closed system which can exchange energy, with the environment, which is a realistic situation in chemical systems. In this work, DFT-based MD calculations were carried out, where all forces, and therefore positions and accelerations were obtained by DFT calculations.

# Chapter 3

# The Bare $\alpha$-$Al_2O_3$ (0001) Surface

## 3.1 Adsorption to the $Al_2O_3$ (0001) Surface

The structure and properties of $\alpha$-$Al_2O_3$ surfaces have been studied both experimentally.[82–85] and computationally.[87,181] Studies show that its most stable surface is the (0001)-oriented surface. Among the possible terminations of the clean (0001) surface, the single-Al termination is the most favourable.[87] Moreover, computational investigations of thermodynamics under vacuum and ambient conditions showed that the preferred surface termination depends on the temperature and the partial pressure of water vapour: the most stable termination varies from the bare surface to the fully hydroxylated surface, as the partial pressure of the water vapour increases.[182] Multiple computational studies have investigated the adsorption of water on this surface.[183–185] Both dissociative and molecular adsorption structures were identified, with dissociative adsorption being the preferred binding mode at low coverage, and complex combinations of dissociated and molecularly adsorbed water predicted at high coverages. The adsorption of water had noticeable effects on both the geometry and electronic structure of this surface: upon adsorption the distance between the terminating aluminium atoms and the oxygen plane increased.[183,186] Surface hydroxyls formed by dissociative adsorption of water were observed experimentally both for monolayers of water and single water molecules on the $\alpha$-$Al_2O_3$ (0001) surface. There have been fewer computational studies of adsorption of organic molecules onto $\alpha$-$Al_2O_3$: examples include methanol and ethanol[187] phenols,[188,189] carboxylic acids[190] and nitrogen containing molecules, such as; methylnitrile,[191] methylamine,[192] nitroamine,[192] nitroguanidine[193] and nitromethane.[194]

Johnston[187] investigated the adsorption of ethanol and found that ethanol bound slightly more strongly to the (0001) surface in a dissociative configuration over a molecular configuration by 0.07

– 0.14 eV. The adsorption of non-dissociated structures was dominated by electrostatic interactions and the adsorption energy was found to be influenced by the proximity of the ethanol molecule's oxygen atom and the surface aluminium atom.[92] The adsorption of phenols was also investigated[189] Yeh et al.[188,189] studied the adsorption of catechol (phenolic alcohol) onto the $\alpha$-$Al_2O_3$ surface in both hydrous and anhydrous conditions. In anhydrous conditions, catechol adsorbed onto hydroxylated and non-hydroxylated surfaces. In aqueous conditions, catechol displaced water molecules and was strongly attracted to the non-hydroxylated alumina surface. Ruan et al.[195] investigated the adsorption of formic acid onto the $\alpha$-$Al_2O_3$ surface and found that the 1,2-dissociated structure where the carboxyl oxygen and hydroxyl hydrogen atoms of formic acid were adsorbed onto the $\alpha$-$Al_2O_3$ surface was the most stable. A very strong interaction was found between β-D-glucopyranose and the $\alpha$-$Al_2O_3$ (0001) surface, which has been attributed to the formation of bridging aluminium-oxygen bonds and hydrogen bonds.[190] Most of the computational studies that investigated the adsorption of molecules onto the $\alpha$-$Al_2O_3$ surface used DFT,[185,187–189,191–194] (or, for water adsorption, MP2 methods) except the investigation conducted by Yeh et al.[189] which used force fields.

A recent investigation into the adsorption of OC to the alumina and hematite (bare and hydroxylated) surfaces was carried out by Blanck et al.[196] Using DFT at the PBE level of theory they calculated the binding energies for the adsorption of several organic functional groups and phosphate and nitrogen containing groups to the $\gamma$-$Al_2O_3$ (100) and the $\alpha$-$Fe_2O_3$ (0001) surfaces. They found that the distance between the adsorption sites of the $\gamma$-$Al_2O_3$ (100) surface were smaller (2.6 Å and 2.9 Å) than those of the $\alpha$-$Fe_2O_3$ (0001) surface (5.1 Å). Hence, bidentate adsorption was preferred on the hematite surface . Hematite was the more reactive surface for all molecules adsorbed. Methylphosphate and dimethylphosphate were the most strongly dissociatively adsorbed molecules. Whereas, benzene and acetonitrile were the least adsorbed (molecular adsorption). Acetic acid yielded very similar adsorption energies on both alumina and hematite. On alumina acetic acid adsorbed in a bidentate configuration, whereas on hematite it adsorbed molecularly. Thus, the bidentate configuration compensated for the strong reactivity of the hematite surface.

In this work, we aim to obtain a quantitative description of the binding of organic molecules containing a variety of functional groups, in order to understand how OC binds to soil minerals. We will be investigating the binding of water and small organic molecules onto $\alpha$-$Al_2O_3$ (corundum), a common soil mineral. $\alpha$-$Al_2O_3$ is the simplest representative of the family of aluminium oxides, hydroxides and oxyhydroxides, and it has been shown to bind OC strongly.[87]

## 3.2 Computational Method

### 3.2.1 Adsorption to the Bare $\alpha$-$Al_2O_3$ (0001) Surface

DFT calculations of $\alpha$-$Al_2O_3$, its (0001) surface and adsorption structures were performed using the CP2K software package.[197–199] Suitability of several DFT functionals was tested by optimising the lattice parameter of the rhombohedral $\alpha$-$Al_2O_3$. The methods considered were based on the generalised gradient approximation (GGA): PBE,[200] PBEsol[201] and revPBE[202] functionals, and meta-GGA: TPSS functional.[143] Grimme's D3[149] empirical dispersion correction was used with these functionals. The calculations utilised double-$\zeta$ basis sets with diffuse and polarization functions (DZVP) optimised for use in CP2K,[203] and Goedecker-Teter-Hutter (GTH) pseudopotentials.[203] All calculations were done at the $\Gamma$ k-point.[198] 2x2x2, 3x3x3 and 4x4x4 extended rhombohedral unit cells of $\alpha$-$Al_2O_3$ were used for bulk calculations. In these calculations, the angles were fixed at their experimental values $\alpha = \beta = \gamma \neq 90°$, while the lattice constant (a = b = c) was allowed to optimise. The calculated lattice constants were: 5.159 Å (PBE functional with the D3 correction), 5.145 Å (PBEsol+D3), 5.166 Å (revPBE+D3) and 4.720 Å (meta-GGA TPSS+D3). This compares to the experimental value of 5.150 Å.[204] Thus, all of the tested GGA functionals were found to give very good agreement of the $\alpha$-$Al_2O_3$ bulk lattice parameter with the experimental value, therefore the GGA PBE functional was used in all the following calculations of the $\alpha$-$Al_2O_3$ (0001) surface and adsorption. The $\alpha$-$Al_2O_3$ (0001) surface was modelled using periodic slabs, where the lattice parameters were fixed at their optimised bulk values, while the positions of all atoms were fully optimised. Slabs were separated by 15 Å of vacuum in the vertical direction. 2x2 and 3x3 extended cells were used for surface calculations. Convergence of surface energies with respect to slab thickness was tested. A 2x2 12 layer converged slab was chosen and then used for calculations of adsorption of water and small organic molecules. The slab consisted of 80 atoms and a length of approximately 14.01 Å. The calculated binding adsorption geometries' energies were corrected for the basis set superposition error (BSSE) using the counterpoise method,[112] the correction range from 0.07 - 0.14 eV. All structures were visualised with the UCSF Chimera[205] software package.

## 3.3 Results and Discussion

### 3.3.1 Surface Energies of $\alpha$-$Al_2O_3$ (0001) Slabs

First, we investigated the convergence of the surface energy of the $\alpha$-$Al_2O_3$ (0001) surface with respect to the thickness of the slab. The 2x2x1 and the 3x3x1 extended unit cell were tested, and the surface energy was calculated with the equation below.

$$E_{surface} = \frac{E_{slab} - E_{bulk} \times Number\ of\ Layers \times Extension}{2A} \tag{3.1}$$

where $E_{surface}$, $E_{slab}$ and $E_{bulk}$ are the energies of the $\alpha$-$Al_2O_3$ (0001) slabs and the $\alpha$-$Al_2O_3$ bulk unit cell. $A$ represents the surface area of the $\alpha$-$Al_2O_3$ (0001) slabs. Surface cell extension was 4 and 6 for the 2x2x1 and the 3x3x1 extended unit cells, respectively.

Table 3.1: The surface energies in $Jm^{-2}$ for various thicknesses of the $\alpha$-$Al_2O_3$ (0001)-orientated 2x2x1 and 3x3x1 extended slabs.

| **No. of Layers** | **Surface Energy/ J m$^{-2}$** **2x2x1 cell extension** | **Surface Energy/ J m$^{-2}$** **3x3x1 cell extension** |
|---|---|---|
| 6 | 2.10 | 2.07 |
| 9 | 1.90 | 1.88 |
| 12 | 2.01 | 1.98 |
| 15 | 2.03 | 1.99 |
| 18 | 2.03 | 1.98 |
| 21 | 2.03 | - |
| 24 | 2.04 | - |
| 27 | 2.04 | - |

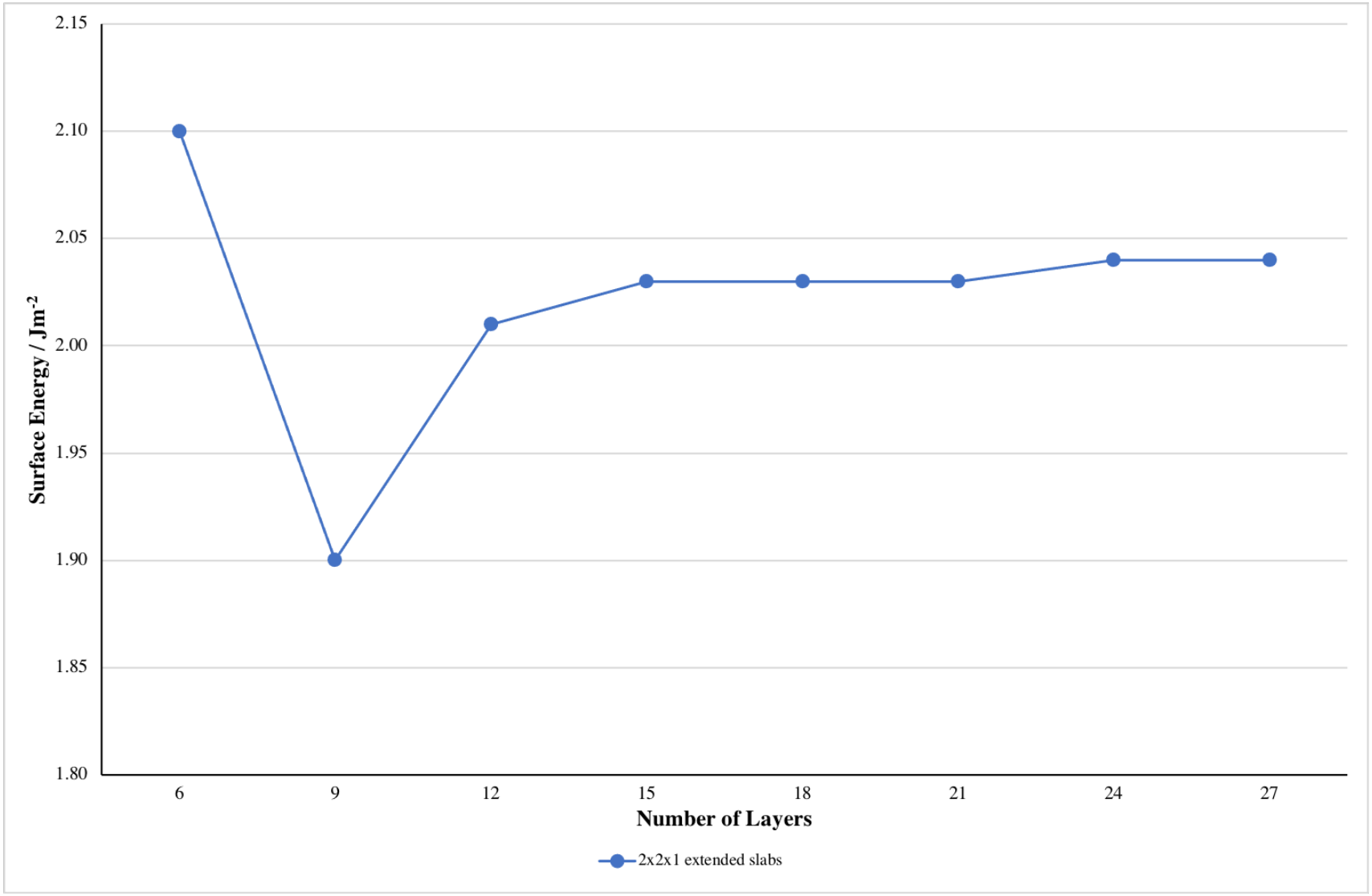



Figure 3.1: Dependence of the surface energy for the $\alpha$-$Al_2O_3$ (0001)-oriented 2x2x1 extended slabs.

The surface energies of the $\alpha$-$Al_2O_3$ (0001)-orientated 2x2x1 and the 3x3x1 extended slabs of different thicknesses were compared, and the results can be seen in table 3.1. The converged surface energy for the thickest considered 18-layer 3x3-extended slab is 1.98 J $m^{-2}$. These results are similar to the literature values for the surface energy of the $\alpha$-$Al_2O_3$ surface which ranged between 0.11 – 0.13 eV/Å$^2$ for 9 to 18 layer slabs,[191–193] The 2x2-extended slabs follow the trend of the 3x3-extended slabs with 0.04 J $m^{-2}$ higher energies. Therefore, we further tested the convergences of 21, 24 an 27 layer slabs for the 2x2-extended slab. The 12-layer slabs can be considered converged. The 2x2x1 extended 12-layer slab was chosen for modelling adsorption of small organic molecules onto the $\alpha$-$Al_2O_3$ (0001) surface, as the thinnest converged slab that represents a good balance between accuracy and system size.

### 3.3.2 Adsorption of Small Molecules onto the $\alpha$-$Al_2O_3$ (0001) Surface

Table 3.2: BSSE corrected Binding energies for adsorbates on the $\alpha$-$Al_2O_3$ (0001) surface.

| **Adsorbate** | **Adsorption Configuration** | **Binding Energy / eV** |
|---|---|---|
| Water | molecular | -1.32 |
| | dissociated | -1.62 |
| Methanol | molecular | -1.53 |
| | dissociated | -1.63 |
| Methanethiol | molecular | -1.36 |
| | dissociated | -1.47 |
| Methylamine | molecular | -2.03 |
| | dissociated | -1.12 |
| Diemthyl ether | molecular | -1.66 |
| Formic acid | molecular | -1.42 |
| | dissociated | -2.33 |
| Acetic acid | chelate | -2.04 |
| Formamide | molecular O-bound | -1.64 |
| | dissociated N-bound | -1.85 |
| Methyl formate | molecular C-O-C-bound | -1.54 |
| Methyl formate | molecular C=O-bound | -1.06 |
| Methane | molecular | -0.38 |
| Cyclohexane | molecular | -0.76 |
| Benzene | molecular | -1.08 |

The adsorption of several small organic molecules was investigated. To evaluate the strength of binding of various organic functional groups to the $\alpha$-$Al_2O_3$ (0001) surface. We applied the BSSE correction to all of our calculated binding energies: the correction was small and ranged from 0.08 – 0.12 eV. The binding energies were calculated using the following equation

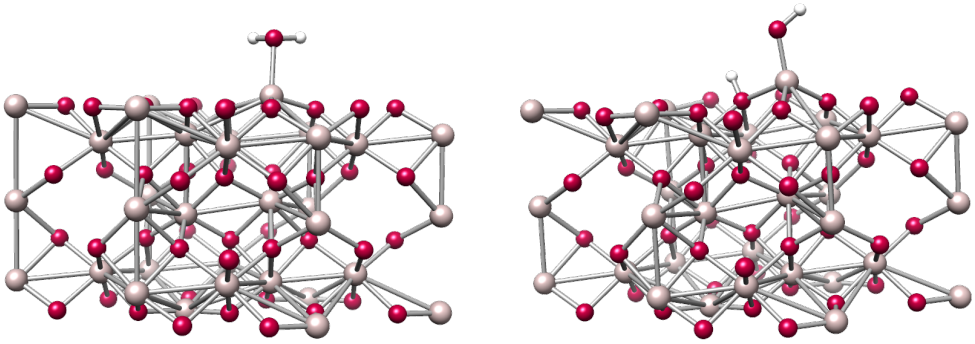
$$E_{binding} = E_{surface+Adsorbate} - (E_{surface} + E_{Adsorbate}) \tag{3.2}$$

Table 3.3: Bond lengths in Å for adsorbates on the $\alpha$-$Al_2O_3$ (0001) surface.

| **Adsorbate** | **Adsorption Configuration** | **Al-O** | **Al-S** | **Al-N** | **Al-C** | **O-H** |
|---|---|---|---|---|---|---|
| Water | molecular | 1.95 | | | | |
| | dissociated | 1.72 | | | | 0.98 |
| Methanol | molecular | 1.93 | | | | |
| | dissociated | 1.71 | | | | 0.98 |
| Methanethiol | molecular | | 2.40 | | | |
| | dissociated | | 2.20 | | | 0.98 |
| Methylamine | molecular | | | 1.98 | | 0.98 |
| | dissociated | | | 1.79 | | 0.98 |
| Diemthyl ether | molecular | 2.40 | | | | |
| Formic acid | molecular | 2.00 | | | | |
| | dissociated | 1.81 | | | | 1.04 |
| | | | | | | 1.58 |
| | | | | | | (O...H) |
| Acetic acid | chelate | 1.98,2.14 | | | | 1.03 |
| Formamide | molecular O-bound | 1.87 | | | | |
| | dissociated N-bound | | | 1.85 | | 0.98 |
| Methyl formate | molecular C-O-C-bound | 2.00 | | | | |
| Methyl formate | molecular C=O-bound | 1.91 | | | | |
| Methane | molecular | | | | 2.48 | |
| | | | | | VDW | |
| Cyclohexane | molecular | | | | | |
| Benzene | molecular | | | | 2.27 | |

### Adsorption of Water onto the $\alpha$-$Al_2O_3$ (0001) Surface

The binding of water to the $\alpha$-$Al_2O_3$ (0001) surface was investigated. Water was adsorbed both molecularly and dissociatively. The calculated binding energy for molecular adsorption of water was -1.32 eV and the chemisorption bond length of the Al-O bond was 1.95 Å. For dissociatively adsorbed water, the calculated binding energy was -1.62 eV and the chemisorption bond lengths for the Al-O and O-H bonds were 1.72 Å and 0.98 Å. Our calculated binding energies are very similar to those calculated by Shapovalov et al.[183] For molecularly adsorbed water Shapovalov et al.[183] calculated a binding energy of -1.22 eV and for dissociatively adsorbed water, it was -1.60 eV. Our chemisorption bond lengths are also fairly similar to those determined within the literature, molecularly adsorbed water had a chemisorption Al-O bond length of 1.99 Å and for dissociatively adsorbed water, the Al-O bond length was 1.75 Å.20-21,45–48[82,91,183,184,186,206–210] The agreement with the literature of the binding energies and bond lengths allows us to conclude that of the two types of structures water adsorption is more stable when it adsorbs to the $\alpha$-$Al_2O_3$ (0001) surface in a dissociative configuration.



(a) -1.32 eV (b) -1.62 eV

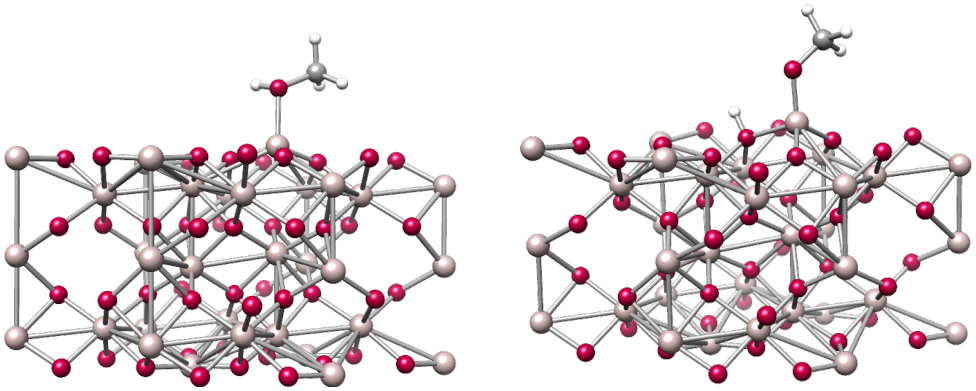
Figure 3.2: Optimised adsorption geometries of (a) molecularly adsorbed water (b) dissociatively adsorbed water onto the $\alpha$-$Al_2O_3$ (0001) surface.

### Adsorption of Alcohol, Amines, Thiols and Ethers onto the $\alpha$-$Al_2O_3$ (0001) Surface

Since our results regarding water adsorption agree with previous experimental and computational studies, we then progressed onto modelling carbon containing functional groups, which are important to carbon-mineral interactions. First, the adsorption of alcohol, amine and thiol molecules was considered, because like water, they contain a hydrogen-heteroatom bond capable of dissociation; ether group was also considered because it is structurally similar to alcohols, but they are without the possibility of dissociation.

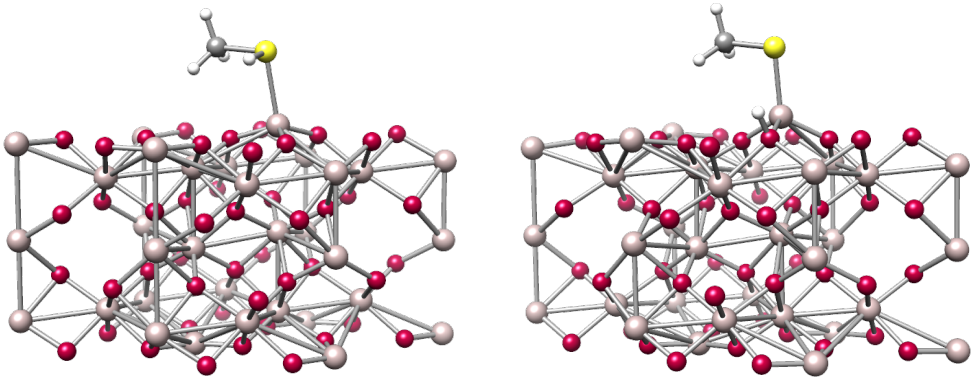
The adsorption of methanol as a model alcohol onto the $\alpha$-$Al_2O_3$ (0001) surface was investigated. The calculated binding energy for molecular adsorption of methanol was -1.53 eV (figure 3.3 (a)) and the chemisorption bond length of the Al-O bond was 1.93 Å. For dissociatively adsorbed

methanol (figure 3.3 (b)) the calculated binding energy was -1.63 eV and the chemisorption bond lengths for the Al-O and O-H bonds were 1.71 Å and 0.98 Å respectively. Methanethiol was the next molecule under investigation, similar to methanol a molecular and dissociative configuration were considered. The calculated binding energy for molecular adsorption (figure 3.4 (a)) of methanethiol was -1.36 eV and the chemisorption bond length of the Al-S bond was 2.40 Å. For dissociatively adsorbed methanethiol (figure 3.4 (b)) the calculated binding energy was -1.47 eV and the chemisorption bond lengths for the Al-S and O-H bonds were 2.20 Å and 0.98 Å.



(a) -1.53 eV (b) -1.63 eV

Figure 3.3: Optimised adsorption geometries of (a) molecularly adsorbed methanol and (b) dissociatively adsorbed methanol onto the $\alpha$-$Al_2O_3$ (0001) surface.

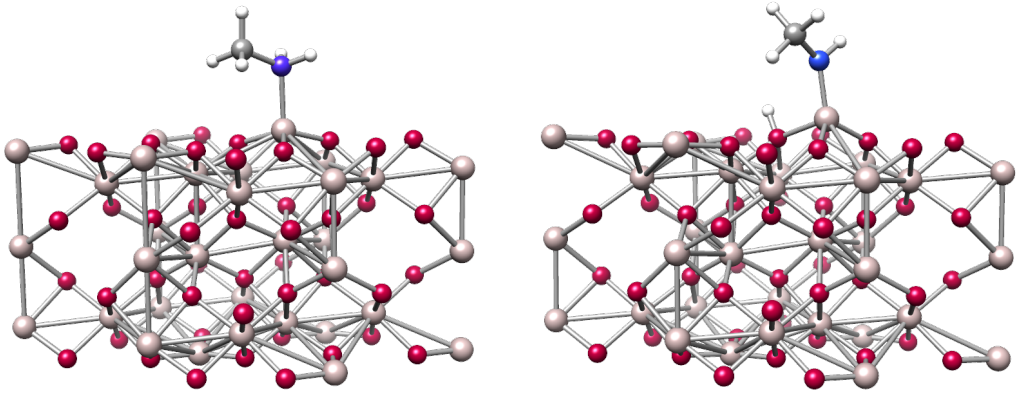


(a) -1.36 eV (b) -1.47 eV

Figure 3.4: Optimised adsorption geometries of (a) molecularly adsorbed methanethiol (b) dissociatively adsorbed methanethiol onto the $\alpha$-$Al_2O_3$ (0001) surface.

For the adsorption of amines, methylamine was the model amine considered for binding onto the $\alpha$-$Al_2O_3$ (0001) surface. Several adsorption configurations of methylamine were investigated. The calculated binding energy for the most stable molecular adsorption geometry (figure 3.5 (a)) of methylamine was -2.03 eV and the chemisorption bond length of the Al-N bond was 1.98 Å. For dissociatively adsorbed methylamine (figure 3.5 (b)) the calculated binding energy was -1.12 eV and the chemisorption bond lengths for the Al-N and O-H bonds were 1.79 Å and 0.98 Å. A doubly dissociated methylamine geometry was also considered, upon optimisation and similar to water this

structure optimised to a geometry where only a single N-H bond was dissociated. To compare to the alcohol functional group, dimethyl ether was adsorbed onto the $\alpha$-$Al_2O_3$ (0001) surface in a molecular configuration (figure 3.6). A dissociated structure was considered to be unlikely as it would require cleavage of the strong O-C bond. Its calculated binding energy was -1.66 eV and its adsorption bond length was 2.40 Å.



(a) -2.03 eV (b) -1.12 eV

Figure 3.5: Optimised adsorption geometries of (a) molecularly adsorbed methylamine (b) dissociatively adsorbed methylamine onto the $\alpha$-$Al_2O_3$ (0001) surface.

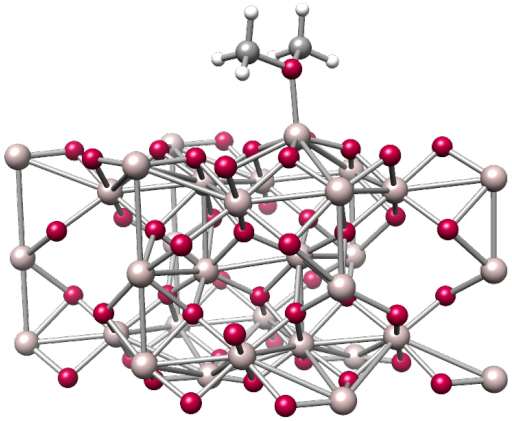


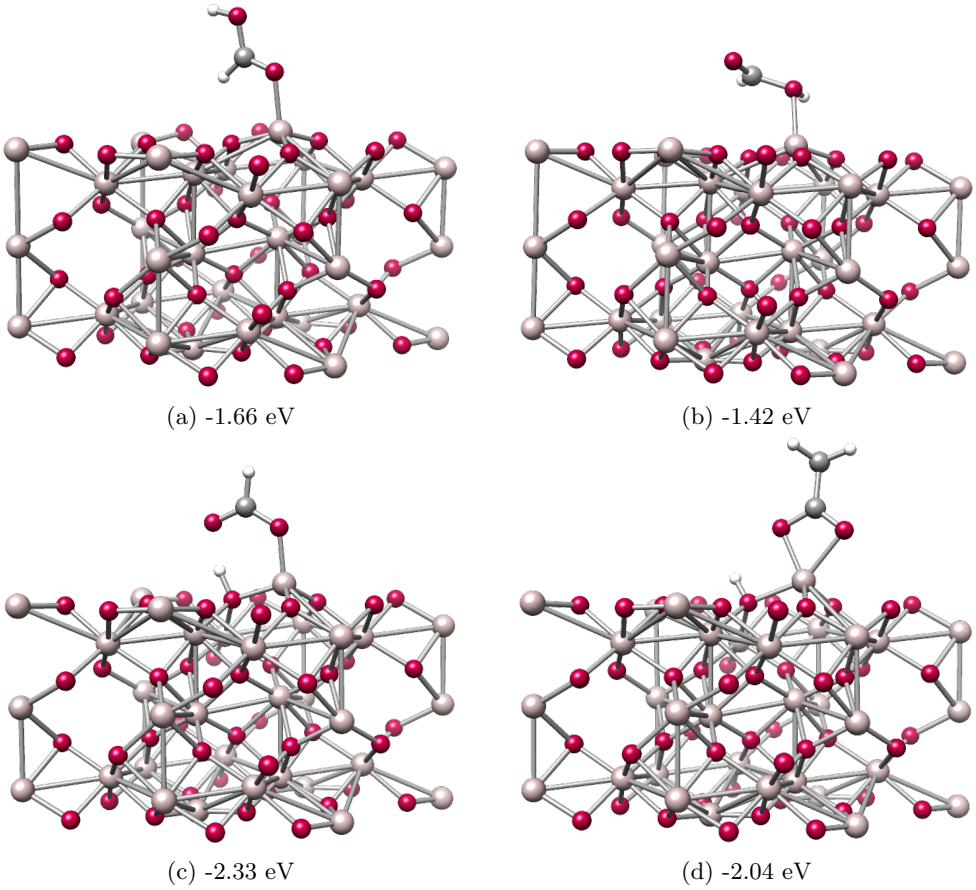
Figure 3.6: Optimised adsorption geometry of molecularly adsorbed dimethyl ether (-1.66 eV) onto the $\alpha$-$Al_2O_3$ (0001) surface.

The calculated binding energies and respective bond lengths between the surface Al and adsorbate O atoms for methanol, methanethiol, similar to water adsorption show that these molecules prefer to adsorb in a dissociative configuration. Dimethyl ether could only bind in one configuration and that was the molecular configuration. Methanol and methanethiol are iso-structural and the observed adsorption geometries illustrate that the two molecules are very similar. Methanethiol adsorbs less strongly by approximately 0.16 eV. The Al-S bond is also longer than the Al-O bond and this can be explained by the larger radius of S. The stronger Al-O bond can also be explained by the greater electronegativity of O compared to S and therefore a greater negative charge of O. Methanol, methylamine and dimethyl ether all have binding energies greater than that of

water, therefore they would displace water when binding to the $\alpha$-$Al_2O_3$ (0001) surface in soils. methylamine is the only surface adsorbate that preferred to bind in a molecular configuration when it could bind in both a molecular and dissociated configuration. This can be explained by the pKa value of a primary amine which is approximately 35. Therefore, the proton will not leave as the pKa value is very high and the $CH_3NH^-$ species is not as stable. Methanethiol had a binding energy smaller than that of water, thiols bind less strongly to the $\alpha$-$Al_2O_3$ (0001) surface due to less ionic character within the bond. Therefore, methanethiol will not compete with water when binding to soil aluminium oxide minerals as water binds more strongly. Hence water would displace the thiol adsorbate. Thus far, alcohols, amines and ethers all bind strongly to the $\alpha$-$Al_2O_3$ (0001) surface and thiols are weaker adsorbates.

### Adsorption of Acids, Amides and Esters onto the $\alpha$-$Al_2O_3$ (0001) Surface

Next, we consider the adsorption of carbonyl compounds: carboxylic acids, amides and esters, which are able to bind to the mineral surface via two atoms: the carbonyl oxygen and the second oxygen or nitrogen atom. We are therefore able to present a large variety of possible adsorption structures. The adsorption of formic acid as the simplest carboxylic acid was considered. The adsorption of formic acid to the $\alpha$-$Al_2O_3$ (0001) surface enabled several different structures to be built. Two molecular configurations were possible for formic acid as adsorption to the surface could occur through the hydroxyl oxygen atom or the carbonyl oxygen atom of formic acid. The calculated binding energy for the structure whereby formic acid bound through the hydroxyl group of formic acid (figure 3.7 (b)) was -1.42 eV and its Al-O chemisorption bond length was 2.00 Å. For the formic acid bound through the carbonyl oxygen (figure 3.7 (a)) the calculated binding energy was -0.95 eV and its corresponding Al-O chemisorption bond length was 1.92 Å. These findings tell us that formic acid prefers to bind through the hydroxyl end. Formic acid was also adsorbed in a dissociative configuration (figure 3.7 (c)) to the $\alpha$-$Al_2O_3$ (0001) surface where the OH bond of the formic acid was dissociated. The calculated binding energy for dissociatively adsorbed formic acid was -1.80 eV and the chemisorption bond length of the Al-O bond was 1.76 Å. Chelate configurations (figure 3.7 (d)) were also considered. Formic acid did not remain in a chelate geometry upon optimisation, however it did optimise to a stable dissociated structure with some hydrogen bonding. The chemisorption bond lengths for this dissociated adsorption were 1.81 Å (Al-O), 1.04 Å (O-H) and 1.58 Å (O...H). The calculated binding energy for this configuration was -2.33 eV, this is significantly larger than the -1.80 eV binding energy for the dissociated configuration. This is most likely due to the additional stabilisation produced by the formation of the O...H hydrogen bond.

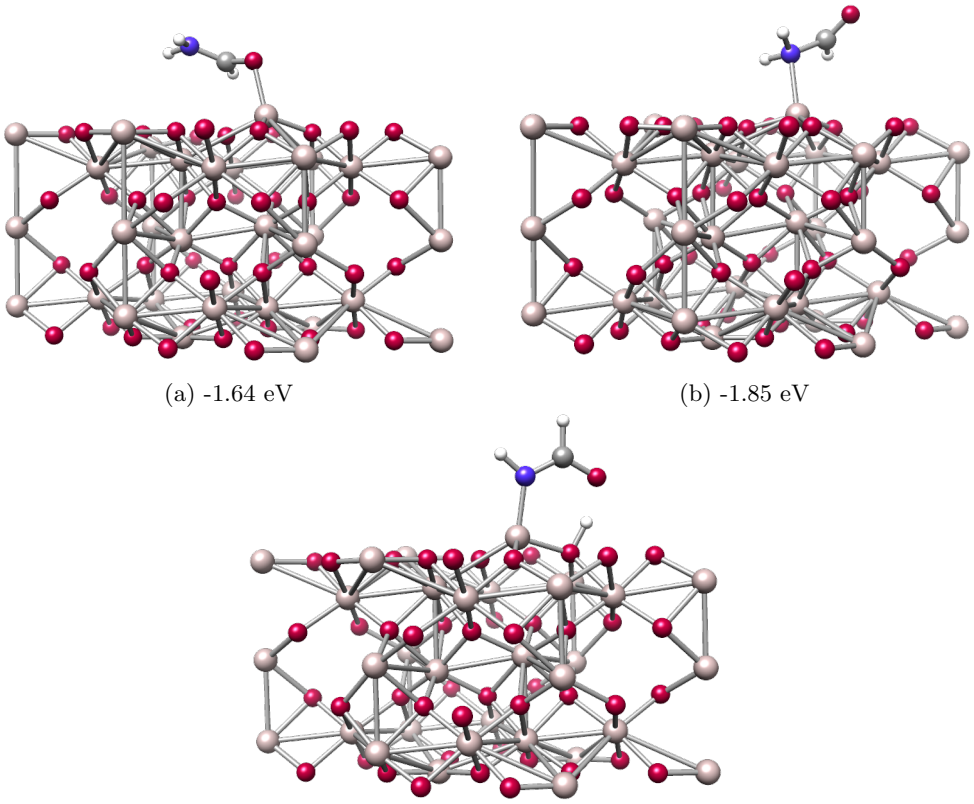


(a) -1.66 eV (b) -1.42 eV

(c) -2.33 eV (d) -2.04 eV

Figure 3.7: Optimised adsorption geometries of (a) molecularly adsorbed formic acid through its C=O oxygen (b) molecularly adsorbed formic acid through its O-H oxygen (c) dissociatively adsorbed formic acid and (d) chelating acetic acid onto the $\alpha$-$Al_2O_3$ (0001) surface.

The adsorption of amides onto the $\alpha$-$Al_2O_3$ (0001) surface was investigated using formamide. Formamide was found to bind to the $\alpha$-$Al_2O_3$ (0001) surface through either its oxygen or nitrogen atoms in a molecular configuration. The oxygen bound formamide (figure 3.8 (a)) had a calculated binding energy of -1.64 eV. The chemisorption bond length is 1.87 Å for the Al-O bond. The Nitrogen bound formamide (figure 3.8 (b)) had a calculated binding energy of -1.16 eV and an Al-N chemisorption bond length of 2.37 Å. The binding energies of molecularly bound formamide show that the oxygen bound formamide is more stable. Oxygen's lone pair is more available than nitrogen's lone pair. Hence, oxygen's lone pairs bind more strongly to the $Al^{3+}$ surface ions, producing a stronger bond to the surface. Another adsorption configuration considered for formamide adsorption onto the $\alpha$-$Al_2O_3$ surface was the dissociated configuration (figure 3.8 (c)). The calculated binding energy of formamide adsorption to the $\alpha$-$Al_2O_3$ surface was -1.22 eV and the chemisorption bond length was 1.85 Å for the Al-O bond and was 0.98 Å for the O-H bond.



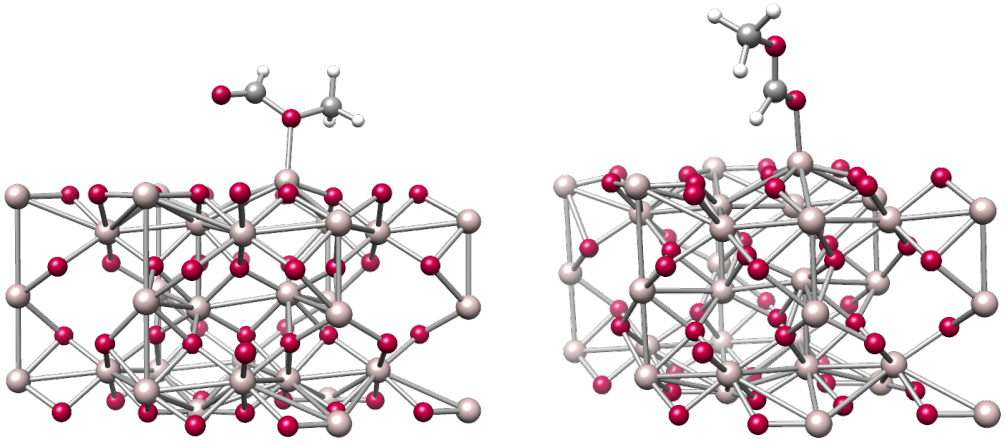
(a) -1.64 eV

(b) -1.85 eV

(c) -1.85 eV

Figure 3.8: Optimised adsorption geometries of (a) molecularly adsorbed formamide through its C=O oxygen (b) molecularly adsorbed formamide through its nitrogen atom (c) dissociatively adsorbed formamide onto the $\alpha$-$Al_2O_3$ (0001) surface.

Methyl formate was the ester considered for adsorption to the $\alpha$-$Al_2O_3$ (0001) surface. Fewer configurations can be expected for esters, compared to those of acids and amides. Only molecularly adsorbed configurations are possible for esters. Methyl formate was adsorbed molecularly onto the $\alpha$-$Al_2O_3$ (0001) surface. For binding through the C-O-C oxygen (figure 3.9 (a)) the calculated binding energy was -1.58 eV and the chemisorption bond length for the Al-O bond was 1.91 Å. For the configuration where binding occurred through the carbonyl group of methyl formate (figure 3.9 (b)) the calculated binding energy was -1.10 eV and the chemisorption bond length for the Al-O bond was 2.00 Å. Configuration (b) is more stable, therefore esters are more likely to bind to through their carbonyl oxygen atom.



(a) -1.54 eV (b) -1.06 eV

Figure 3.9: Optimised adsorption geometries of (a) molecularly adsorbed methyl formate through its C-O-C oxygen atom (b) molecularly adsorbed methyl formate through its C=O oxygen onto the $\alpha$-$Al_2O_3$ (0001) surface.

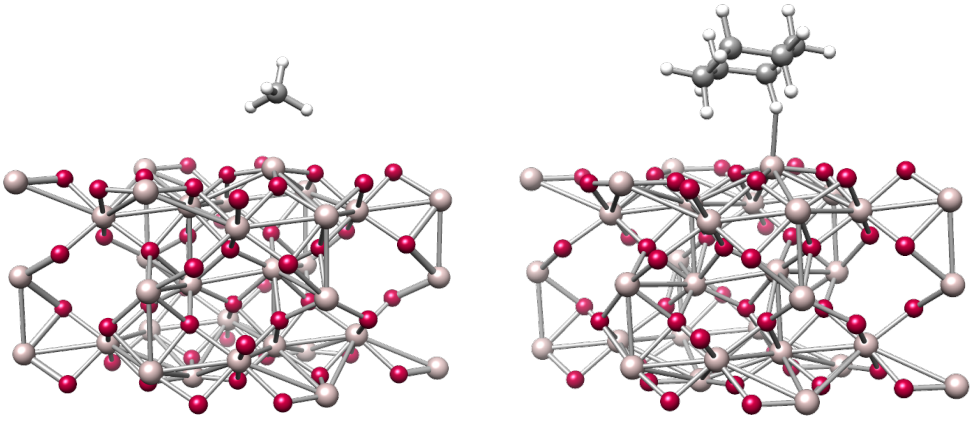
The calculated binding energies and respective bond lengths between the surface aluminium and adsorbate oxygen atoms for formic acid and formamide adsorption show that these molecules prefer to adsorb in a dissociative configuration. The most stable adsorption configuration for formic acid was the dissociated configuration, which had additional hydrogen bond stability. Similar to the acid functional group amides had several modes of adsorption, but preferred to adsorb in a dissociative configuration. Both the acid and amide functional groups had binding energies larger than those of water. Thus, we can conclude that both functional groups would displace water when binding to the $\alpha$-$Al_2O_3$ (0001) surface in soils as they bind more strongly, specifically in their dissociative configurations. The ester functional group was only adsorbed onto the surface in a molecular configuration. The largest binding energy for the ester was -1.54 eV. Therefore, it is expected that the ester will not compete nor displace water when binding to the $\alpha$-$Al_2O_3$ (0001) surface.

### Adsorption of Hydrocarbons onto the $\alpha$-$Al_2O_3$ (0001) Surface

Organic molecules that enter soil from plants (e.g. cellulose) typically contain $sp^3$ carbons as well as functional groups. Sorption studies also indicated the presence of aromatic groups in OC sorbed in soils.[86] Therefore, adsorption of aliphatic and aromatic hydrocarbons was modelled, to compare to the adsorption of functional groups.

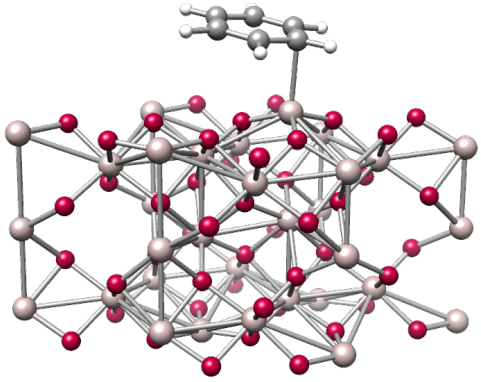
Adsorption of methane on $\alpha$-$Al_2O_3$ (0001) was modelled first, as the simplest example of a hydrocarbon molecule. Methane adsorbed molecularly (figure 3.10 (a)), with a binding energy of -0.38 eV and an Al-C distance of 2.48 Å; this binding is clearly much weaker than water or any organic functional group and is attributed to Van der Waals dispersion interactions. Next, physisorption of cyclohexane and benzene was considered, as comparable examples (with the same number of

carbon atoms) of aliphatic and aromatic hydrocarbons. Cyclohexane can also be viewed as an analogue of tetrahydropyran (with an oxygen atom replaced by a $CH_2$ group), the 6-membered ring that is the building block of cellulose and chitin. Cyclohexane adsorbed to the surface in its molecular form (figure 3.10 (b)), without dissociation, and had a binding energy of -0.76 eV. The most stable structure was that whereby the cyclohexane molecule was adsorbed above an oxygen atom on the $\alpha$-$Al_2O_3$ (0001) surface. Finally, benzene was adsorbed on the surface. Its adsorption is again molecular (figure 3.10 (c)), and the most stable adsorption structure has benzene bound to a surface aluminium atom with an Al-C bond length of 2.27 Å. The binding energy was -1.08 eV, larger than of cyclohexane, indicating that aromatic hydrocarbons adsorb more strongly than aliphatic hydrocarbons on the $\alpha$-$Al_2O_3$ (0001) surface. In both cases, adsorption can be attributed to fairly strong Van der Waals dispersion interactions.



(a) -0.38 eV

(b) -0.76 eV



(c) -1.08 eV

Figure 3.10: Optimised adsorption geometries of (a) molecularly adsorbed methane (b) molecularly adsorbed cyclohexane and (c) molecularly adsorbed benzene onto the $\alpha$-$Al_2O_3$ (0001) surface.

The adsorption energy of methane, cyclohexane and benzene are all much smaller than that of water and of organic functional groups. We can therefore conclude that they will not displace water when binding to the $\alpha$-$Al_2O_3$ (0001) surface in soils because water binds more strongly. Moreover, since organic functional groups bind more strongly than hydrocarbons, these functional groups are likely to be the major factors controlling the binding of OC to soil minerals. However, our calculations show that the binding of hydrocarbons is non-negligible and will contribute significantly to the overall binding of OC to soils; notably, the binding of aromatic hydrocarbons is more favourable than that of aliphatic hydrocarbons, suggesting that aliphatic hydrocarbons are more likely to be preserved in soil.

## 3.4 Conclusion

Our first set of investigations focused on yielding an initial understanding of small organic functional group binding to the bare $\alpha$-$Al_2O_3$ (0001) surface. The binding energies of the interaction between water and small organics with the alumina surface were calculated. The first of these investigations studied the adsorption of water, which was found to adsorb dissociatively to the surface. The binding of small organics followed and several organic adsorbates particularly, acids, alcohols, amines and amides were all found to bind strongly to the surface. Amongst these adsorbates, acids were the most strongly adsorbing carbon containing functional group. Acids bound most strongly when they dissociatively adsorbed to the surface, due to additional stability provided by hydrogen bonding. These results show us which organic functional groups are strongly binding. However, the limitation is that we have considered the bare $\alpha$-$Al_2O_3$ (0001) surface, while in realistic environmental conditions it is expected to be hydrated. Therefore, in the next chapter, we will further investigate the interaction of OC with soil minerals via the investigation of the binding of small organics (from this study) with the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

# Chapter 4

# The Hydroxylated $\alpha$-$Al_2O_3$ (0001) Surface

## 4.1 Hydroxylation of the $\alpha$-$Al_2O_3$ (0001) surface

In regular environmental conditions the $\alpha$-$Al_2O_3$ (0001) surface is likely to be hydroxylated.[196] The hydroxylated alumina surface produced via the interaction of water with alumina is extensively used, similar to alumina, it is also used as a catalyst support as well as a catalyst. It also has a role in environmental chemistry.[206]

It is known that the bare $\alpha$-$Al_2O_3$ (0001) surface readily reacts with water to form surface hydroxyls,[11] As discussed in the previous chapter, theoretical ab-initio calculations have revealed that molecularly adsorbed water on the aluminum terminated $\alpha$-$Al_2O_3$ (0001) surface is meta-stable. The dissociation of water produces two different types of surface hydroxyl groups. The $O_{ads}H$ surface hydroxyl oxygen atom comes from water and the $O_sH$ oxygen atom is from the surface (figure 4.1). Thermal desorption[211] and X-ray photo-electron spectroscopy (XPS)[212–217] have confirmed the existence of OH-terminated $\alpha$-$Al_2O_3$ (0001) surfaces. Initially, at low exposures to water hydroxyl formation is attributed to reactions occurring at surface defect sites. On the other hand, higher water exposure results in hydroxylations occurring at the basal plane,[211, 212, 218–221] DFT simulations have described the first step of the basal plane hydroxylation as water dissociation across one of three equivalent aluminium oxygen $Al_s$-$O_s$ bonds.[196, 222]

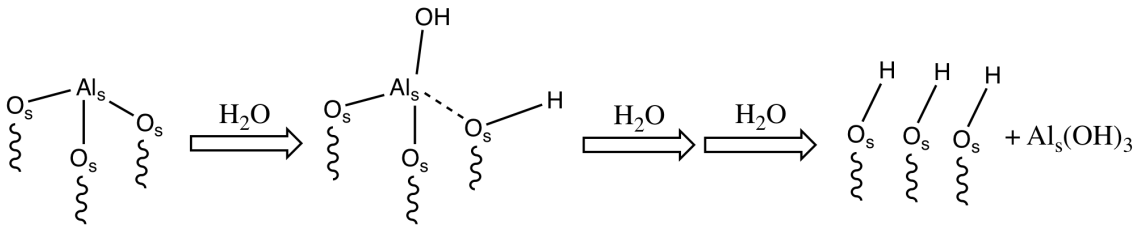



Figure 4.1: Hydroxylation of the $\alpha$-$Al_2O_3$ surface.[11]

The first step in figure 4.1 in the hydroxylation of the $\alpha$-$Al_2O_3$ (0001) surface results in the $H_2O$ molecule adsorbing to the surface to form an adsorbed $OH^-$ and a $H^+$ ion. The next step illustrates the full hydroxylation of the $\alpha$-$Al_2O_3$ (0001) surface resulting in the production of a 'gibbsite' like structure.[11] Car-Parinello molecular dynamics (CPMD) DFT simulations show that the first step in the hydroxylation process has a small activation barrier and occurs rapidly at ambient temperatures.[223,224] Thus, the $\alpha$-$Al_2O_3$ (0001) surface readily exists as the surface produced in the first hydroxylation step in realistic ambient environments, such as hydrated naturally occurring soils.

CMPD simulations have also revealed that the fully hydroxylated surface possess a dynamic layer of hydroxyl groups, consistent with vibrational spectroscopy.[87,210] This surface composition is also consistent with first principles thermodynamic modelling, which demonstrates that the fully hydroxylated $\alpha$-$Al_2O_3$ (0001) surface minimizes the alumina surface free energy under most conditions, with the exception of dry conditions.[208,209,225–228]

## 4.2 Adsorption to the hydroxylated $Al_2O_3$ (0001) surface

Although the initial water dissociation step and composition of the $\alpha$-$Al_2O_3$ (0001) surface in the presence of water have received a great deal of attention, little work has been carried out on the adsorption of molecules to the fully hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.[11] Yeh et al.[189] investigated the adsorption behaviour of catechol on alumina surfaces using molecular dynamics (MD) simulations. In order to investigate the hydrophillic and hydrophobic behaviour of catechol and related phenolic compounds, adsorption was investigated on both the bare and fully hydroxylated $\alpha$-$Al_2O_3$ (0001) surfaces. Hydrogen bonding occurred between neighbouring hydroxyl groups and surface hydroxyl groups of the $\alpha$-$Al_2O_3$ (0001) surface.[189]

Blanck et al.[196] also investigated the adsorption of small organic molecules to related materials, such as the hydroxylated $\gamma$-$Al_2O_3$ (100) and $\alpha$-$Fe_2O_3$ (0001) surfaces. The effects of hydrating the surfaces yielded lower adsorption energies compared to the chemisorption adsorption process that occurred on the bare surface. The stability for the surface adsorbate interactions was provided by

hydrogen bonding.

In this chapter we will be investigating the adsorption behaviour of acids and other organic functional groups on the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface, to determine how strongly they adsorb to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. This follows on from our earlier study of adsorption of organic molecules on the bare $\alpha$-$Al_2O_3$ (0001) surface. The hydroxylated $\alpha$-$Al_2O_3$ (0001) surface is a more realistic model of the alumina surface in soils. We will employ DFT to determine the nature and strength of the binding between the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface and small organic molecules.

## 4.3 Computational Method

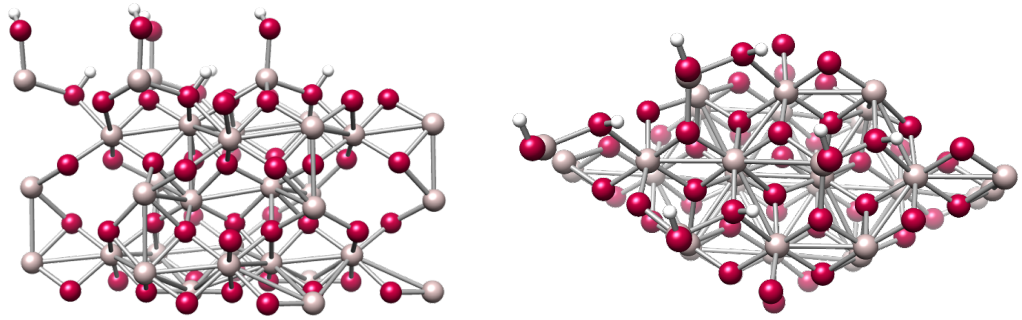
### 4.3.1 Adsorption to the Hydroxylated $\alpha$-$Al_2O_3$ (0001) Surface

DFT calculations of the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface and adsorption structures were performed using the CP2K software package.[197–199] All DFT calculations used the GGA PBE exchange-correlation functional[200] with Grimme's D3[149] empirical dispersion correction. The calculations utilised double-$\zeta$ basis sets with diffuse and polarization functions (DZVP) optimised for use in CP2K,[203] and Goedecker-Teter-Hutter (GTH) pseudopotentials.[203] All calculations were done at the $\Gamma$ k-point.[198] The hydroxylated $\alpha$-$Al_2O_3$ (0001) surface was modelled using periodic slabs containing 12 atomic layers, where the lattice parameters were fixed at their optimised bulk values, while the positions of all atoms were fully optimised. Slabs were separated by 15 Å of vacuum in the vertical direction and were 2x2 extended in the horizontal directions to accommodate the adsorbate. Hydroxylated slabs consisted of 92 atoms and a length of approximately 14.01 Å. The calculated binding energies reported in this work were corrected for the basis set superposition error (BSSE) using the counterpoise method,[112] the correction ranged from 0.01 - 0.10 eV. Molecular dynamics (MD) simulations[229] were performed within the cannonical ensemble (NVT).[178,230] All structures were visualised with the UCSF Chimera[205] software package.

## 4.4 Results and Discussion

### 4.4.1 Adsorption of Small Organic Molecules onto the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface

Under ambient conditions the bare $\alpha$-$Al_2O_3$ (0001) surface is partially hydroxylated as shown in figure 4.2. Therefore, to investigate the binding of OC in soil we have adsorbed several small organic molecules and water onto the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. Multiple adsorption configurations were investigated to get a range of different structures and determine the most stable structure per molecule. The adsorption geometries and the strength of binding of the organic functional groups were also investigated. All structures were adsorbed molecularly (without dissociation) through hydrogen bonding. Hydrogen bond analyses were carried out on all optimised structures to determine the number of hydrogen bonds present in each geometry and the length of each hydrogen bond.



(a) (b)

Figure 4.2: The hydroxylated $\alpha$-$Al_2O_3$ (0001) surface (a) side view and (b) top view.

Binding energies were calculated using equation 3.1 (chapter 3).

Table 4.1: Binding energies for small molecules on the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

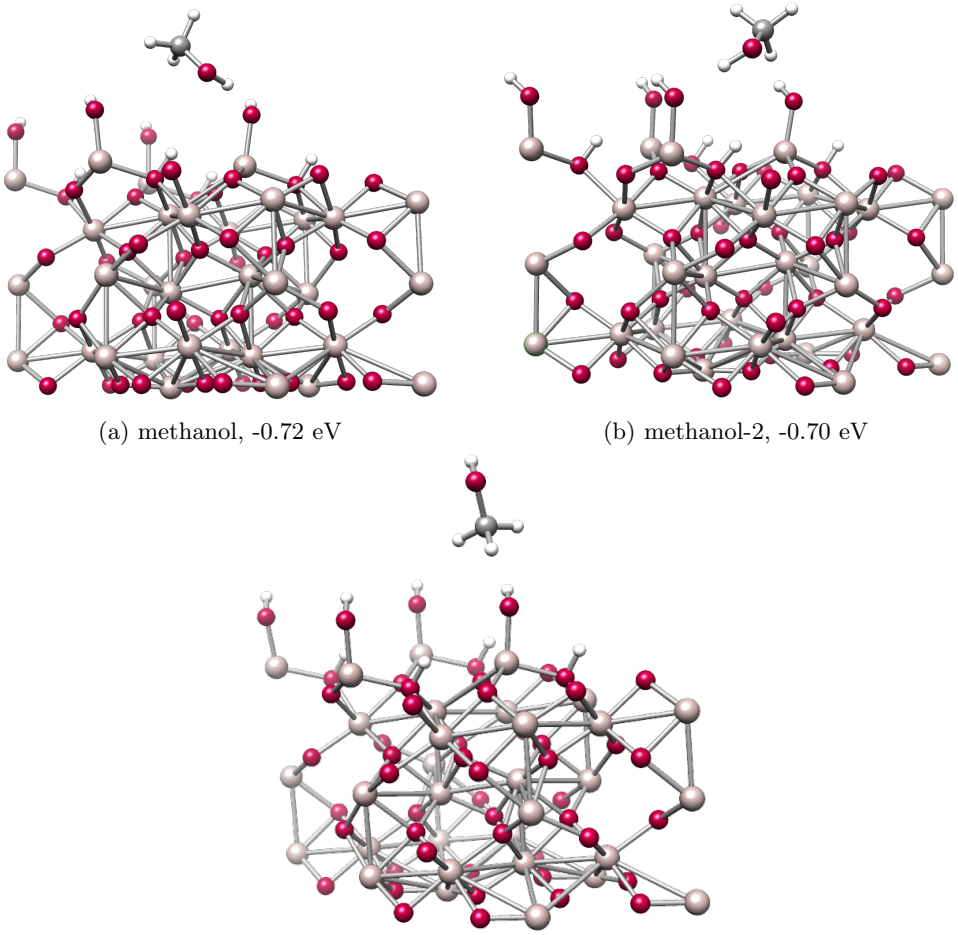
| **Adsorbate** | **Binding Energy** /eV |
|---|---|
| methanol | -0.72 |
| methanol-2 | -0.70 |
| methanol-3 | -0.10 |
| methylamine | -0.79 |
| methylamine-2 | -0.59 |
| methylamine-3 | -0.58 |
| acetic acid | -1.11 |
| acetic acid-2 | -0.98 |
| acetic acid-3 | -0.95 |
| acetic acid-4 | -0.55 |
| acetic acid-5 | -0.55 |
| oxalic acid | -0.78 |
| oxalic acid-2 | -0.75 |
| oxalic acid-3 | -0.30 |
| formamide | -0.70 |
| formamide-2 | -0.61 |
| formamide-3 | -0.46 |
| acetamide | -0.73 |
| acetamide-2 | -0.72 |
| acetamide-3 | -0.69 |
| methyl acetate | -0.57 |
| methyl acetate-2 | -0.32 |
| methyl acetate-3 | -0.19 |
| water | -1.40 |
| water-2 | -1.17 |
| water-3 | -0.47 |

Table 4.2: Hydrogen bond distances for organic molecules on the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface; where $OH_s$ denotes the surface OH groups.

| **Adsorbate** | **Hydrogen Bond** | **Hydrogen Bond Distance / Å** |
|---|---|---|
| methanol | $CH_3OH\cdots OH_s$ | 1.75 |
| methanol-2 | $CH_3OH\cdots OH_s$ | 1.85 |
| methanol-3 | no H-bond | - |
| methylamine | $CH_3NH_2\cdots OH_s$ | 2.16 |
| methylamine-2 | $OH_s\cdots NHCH_3H\cdots OH_s$ | 1.76, 1.95 |
| methylamine-3 | $OH_s\cdots NHCH_3H\cdots OH_s$ | 1.79, 2.04 |
| acetic acid | $OH_s\cdots OCCH_3OH\cdots OH_s$ | 1.55, 1.49 |
| acetic acid-2 | $CH_3COOH\cdots OH_s$ | 1.61 |
| acetic acid-3 | $OH_s\cdots OCCH_3HO\cdots HO_s$ | 1.80, 2.00 |
| acetic acid-4 | $HO_s\cdots OCCH_3OH$ | 1.84 |
| acetic acid-5 | - | 2.34, 1.79 (not H-bonds) |
| oxalic acid | $OH_s\cdots OCOHCOOH$ | 1.66 |
| oxalic acid-2 | $HO_s\cdots HOCOCOHO\cdots HO_s$ | 1.57, 1.86 |
| oxalic acid-3 | $OH_s\cdots OCOHCOOH$ | 1.71 |
| formamide | $OH_s\cdots OCHNH_2\cdots OH_s$ | 1.66, 1.84 |
| formamide-2 | $HO_s\cdots HNHCHO\cdots HO_s$ | 1.87, 1.75 |
| formamide-3 | $HO_s\cdots HNHCHO\cdots HO_s$ | 1.94, 2.27 |
| acetamide | $OH_s\cdots OCCH_3NH_2\cdots OH_s$ | 1.85, 1.74 |
| acetamide-2 | $CH_3CNH_2O\cdots HO_s$ | 1.77 |
| acetamide-3 | $OH_s\cdots NH_2CCH_3O\cdots HO_s$ | 2.32, 1.77 |
| methyl acetate | $OH_s\cdots OCCH_3OCH_3$ | 1.71 |
| methyl acetate-3 | $OH_s\cdots OCH_3COCH_3$ | 2.4 |
| methyl acetate-2 | no H-bond | - |
| water | $HO_s\cdots H\text{-}O\text{-}H\cdots OH_s$, $H_2O\cdots HO_s$ | 1.83, 1.76, 1.55 |
| water-2 | $HO_s\cdots H\text{-}O\text{-}H\cdots OH_s$, $H_2O\cdots HO_s$ | 1.86, 1.79, 1.61 |
| water-3 | $HO_s\cdots H\text{-}O\text{-}H\cdots OH_s$ | 1.93, 1.91 |

**Adsorption of Alcohols and Amines onto the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface**

The strength of binding was evaluated with respect to the number of hydrogen bonds in the adsorbed structures. Small organic molecules with multiple binding sites were bound through all possible sites for example, amides were bound through their nitrogen and carbonyl oxygen atoms. Informed by our adsorption study on the bare surface, where alcohols and amines were found to be strongly binding, Alcohols and amines were the first small organic molecules adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.



(a) methanol, -0.72 eV

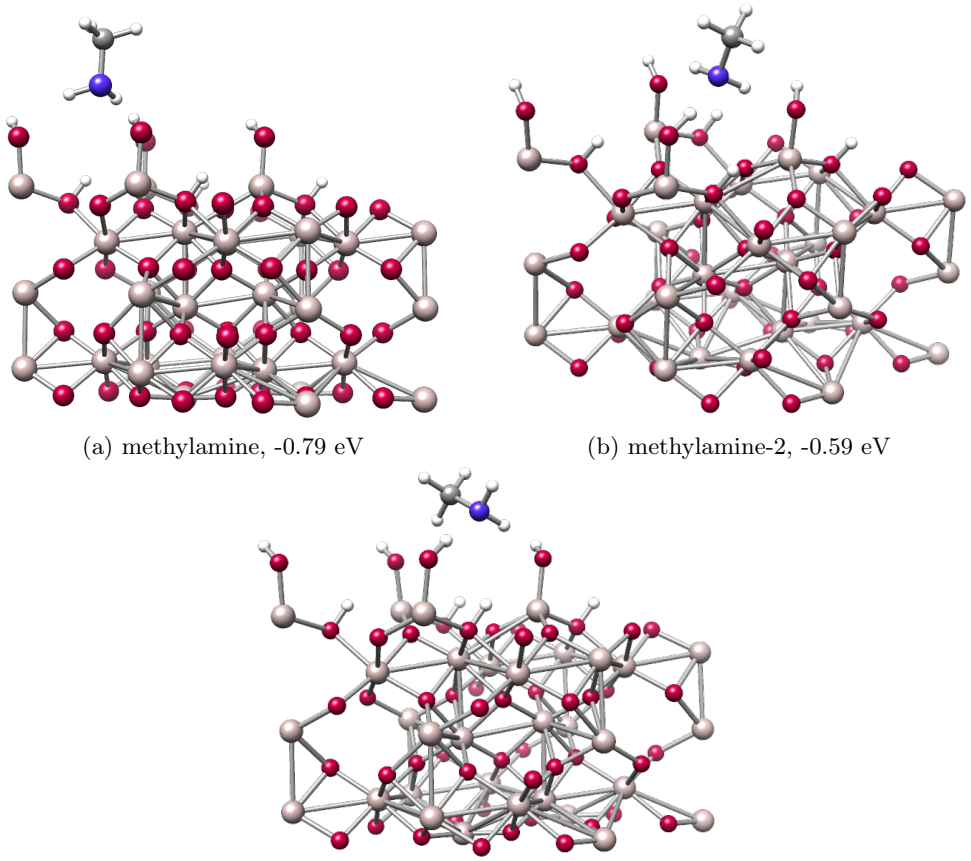
(b) methanol-2, -0.70 eV

(c) methanol-3, -0.10 eV

Figure 4.3: Optimised adsorption geometries and binding energies of methanol adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

Methanol was the model alcohol chosen to be adsorbed onto the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. Methanol was adsorbed in a molecular adsorption configuration. Several adsorption configurations were constructed to investigate the role of hydrogen bonding in the adsorption of methanol. Structures labelled methanol and methanol-2 (figures 4.3 (a) and (b)) involve direct hydrogen bonding between the alcohol hydrogen and the surface oxygen. These two structures are relatively strongly adsorbed (-0.72 eV and -0.70 eV). For comparison the methanol-3 (figure 4.3 (c)) structure does not involve any hydrogen bonding, and the methanol molecule moved

upright during optimisation. Therefore, methanol-3 only interacts with the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface via dispersion interactions. The calculated adsorption energy was -0.10 eV for this geometry. Hence, these results confirm that hydrogen bonding is essential for strong binding of adsorbates to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.



(a) methylamine, -0.79 eV

(b) methylamine-2, -0.59 eV

(c) methylamine-3, -0.58 eV

Figure 4.4: Optimised adsorption geometries and binding energies of methylamine adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

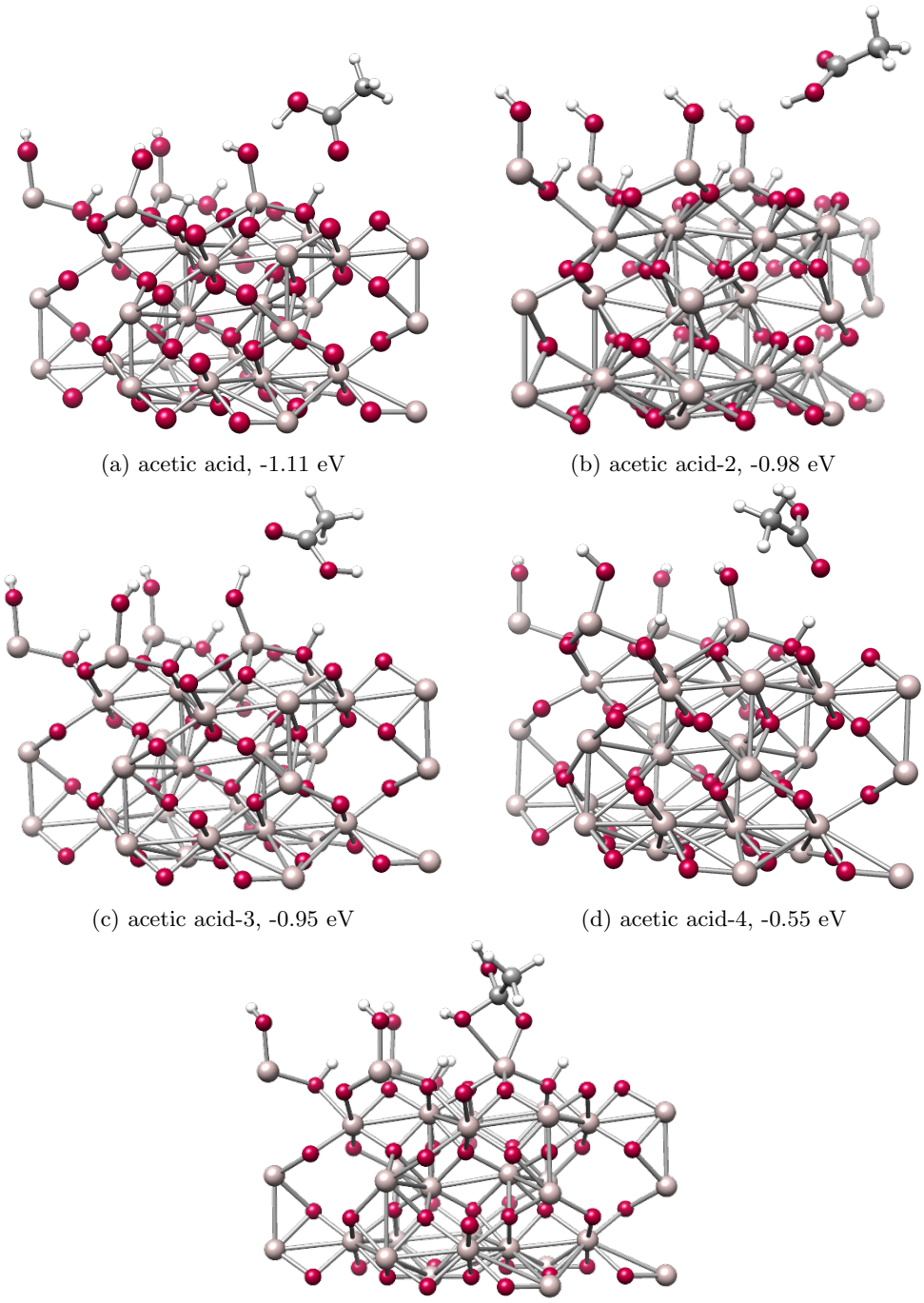
The next molecule adsorbed to the surface was methylamine. Similar to methanol, methylamine was also adsorbed in a molecular configuration. Again, we adsorbed several different structures for methylamine to get the best understanding of the nature of binding between the amine and the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. The calculated binding energy for the most stable molecularly adsorbed methylamine (figure 4.4 (a)) molecule to the hydroxylated alumina surface was -0.79 eV. In this particular geometry a hydrogen bond is produced between the amine hydrogen and the surface oxygen. Although a single hydrogen bond was detected for the first methylamine structure, it was more stable than methylamine-2 and methylamine-3. This stability was the result of an additional hydrogen bond between a hydrogen atom hanging off the surface fragment and the rest of the surface. Methylamine-2 (figure 4.4 (b)) was the second most stable structure, with a calculated binding energy of -0.59 eV, which is similar to the binding energy of methylamine-

3 (figure 4.4 (c)), (-0.58 eV). The structures of methylamine-2 and methylamine-3 illustrate two hydrogen bonds between the surface and the methylamine molecule. Methylamine-2 is more stable, as it has formed shorter stronger hydrogen bonds to the surface, than those of methylamine-3 (1.76 Å, 1.95 Å compared to 1.79 Å, 2.04 Å).

The calculated binding energies for methanol and methylamine illustrate that the binding energies for the alcohol and amine functional groups are similar. Both the alcohol and amine functional groups binding to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface was facilitated by hydrogen bonding. This is as expected with a hydroxylated surface as the surface hydroxyls prevent chemisorption interactions between the small organic molecules and surface aluminium ions. The difference in bond strengths between methanol-1 and 2 and methylamine-2 and 3 can be explained using electronegativity arguments. Oxygen is more electronegative than nitrogen, therefore the OH bond is more polar than the NH bond, thus enabling the alcohol to form stronger hydrogen bonds. The enhanced stability of the amines can be attributed to their ability to form two hydrogen bonds with the hydroxylated surface. Whereby, the alcohol typically only forms a single hydrogen bond between its hydroxyl hydrogen and a surface oxygen.

**Adsorption of Acids, Amides and Esters onto the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface**

Our study on adsorption to the bare surface found that acids were the most strongly adsorbing molecules. Amides and esters had intermediate binding strengths, with the former binding more strongly. Therefore, we considered the adsorption of carbonyl compounds: carboxylic acids, amides and esters for comparison. For the acid and amide functional groups a small model molecule was first investigated (acetic acid and formamide) and then more complex molecules were investigated (oxalic acid, formally known as ethanedioic acid, and acetamide). These molecules were investigated in order to gain an understanding as to how the addition of an extra functional group in a molecule such as an additional carboxylic acid or an additional methyl group may affect binding to the hydroxylated surface. Thus, allowing us to gain some insight into how large biomolecules with multiple functional groups will bind to minerals.



(a) acetic acid, -1.11 eV

(b) acetic acid-2, -0.98 eV

(c) acetic acid-3, -0.95 eV

(d) acetic acid-4, -0.55 eV

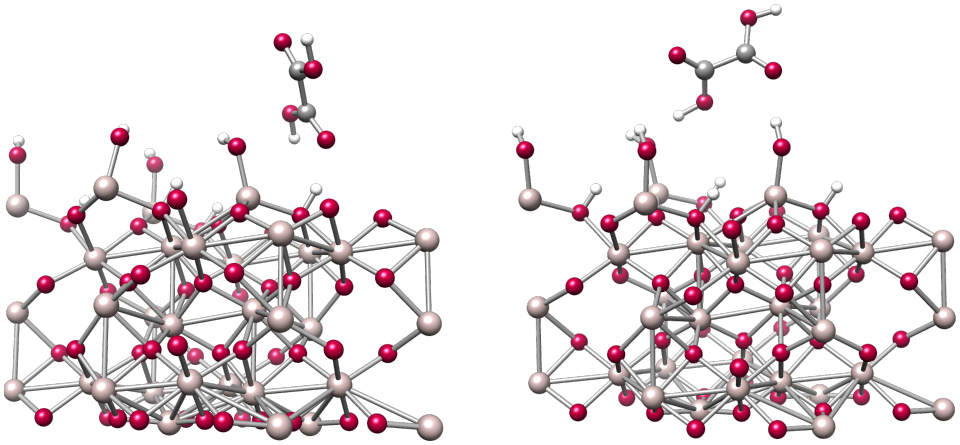
(e) acetic acid-5, -0.55 eV

Figure 4.5: Optimised adsorption geometries and binding energies of acetic acid adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

In our study of acetic acid adsorption on the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface several adsorption geometries were investigated. Acetic acid's most stable adsorption configuration (figure 4.5 (a)) yielded a binding energy of -1.11 eV. In this geometry the acid was able to adsorb to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface via two hydrogen bonds. Upon analysis of the hydrogen bonds, the bonds were short and strong (1.55 and 1.49 Å), thus resulting in a stable adsorption configuration. The second most stable acetic acid adsorption geometry, acetic acid-2 (figure 4.5

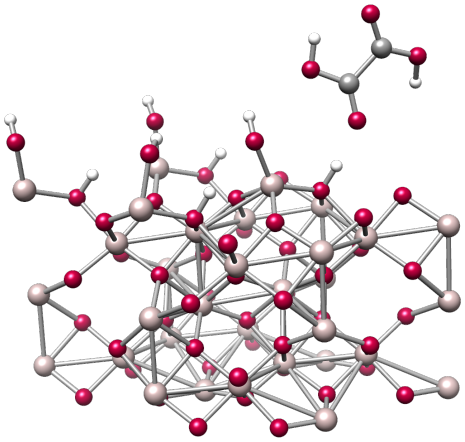
(b)) has a single, but strong hydrogen bond between the surface and the OH of the acid (1.61 Å). We have also found another structure, acetic acid-3 (figure 4.5 (c)), an adsorption geometry with two hydrogen bonds between the acid and the hydroxylated surface, similar to the first acetic acid geometry. However, the hydrogen bonds formed between the surface and acid are longer in this case (1.80 and 2.00 Å). Thus, resulting in a less stable structure with a binding energy of -0.95 eV. Acetic acid was also adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface through the carbonyl oxygen. When binding through the carbonyl oxygen of the acid the calculated binding energy was -0.55 eV. Thus, acids can bind through their hydroxyl and carbonyl regions. Acetic acid-4 (figure 4.5 (d)) illustrates a more weakly bound structure with a binding energy of -0.55 eV. A chelate geometry was also considered (acetic acid-5) (figure 4.5 (e)), its calculated binding energy was -0.55 eV and its bond lengths were 2.34 Å and 1.79 Å respectively. Typically chelate structures are very stable, however this structure is not stabilised by additional hydrogen bonding. The Al-O covalent bond lengths are long, particularly for the Al-O bond measuring 2.34 Å, whereby the surface aluminium has covalently bonded to the carbonyl end of the acid. This structure is interesting to discuss as a possible transition state in a reaction of adsorbed OH being replaced with acetic acid.

The second acid to be adsorbed onto the surface was oxalic acid, It is an acid with two carboxylic acid functional groups. The binding energy for the most stable oxalic acid adsorption configuration (figure 4.6 (a)) was -0.78 eV. In this geometry oxalic acid formed a single strong hydrogen bond (1.66 Å) to the surface. Oxalic acid-2 (figure 4.6 (b)) is the second most stable adsorption geometry, in this adsorption configuration oxalic acid was able to form two hydrogen bonds with the surface. The bond lengths were 1.57 Å and 1.86 Å. Despite the structure containing two fairly strong hydrogen bonds, the calculated binding energy was -0.75 eV. Thus, the addition of an extra acid functional group does not strengthen binding further. The weakest of the adsorption configurations was oxalic acid-3 (figure 4.6 (c)), in which oxalic acid formed only a single hydrogen bond to the surface and an adsorption energy of -0.30 eV was calculated. The weak binding of oxalic acid-3 can be explained by repulsion arguments, as the hydroxyl oxygen of the acid will repel surface hydroxyl oxygen's. Thus, resulting in weaker adsorption.



(a) oxalic acid, -0.78 eV

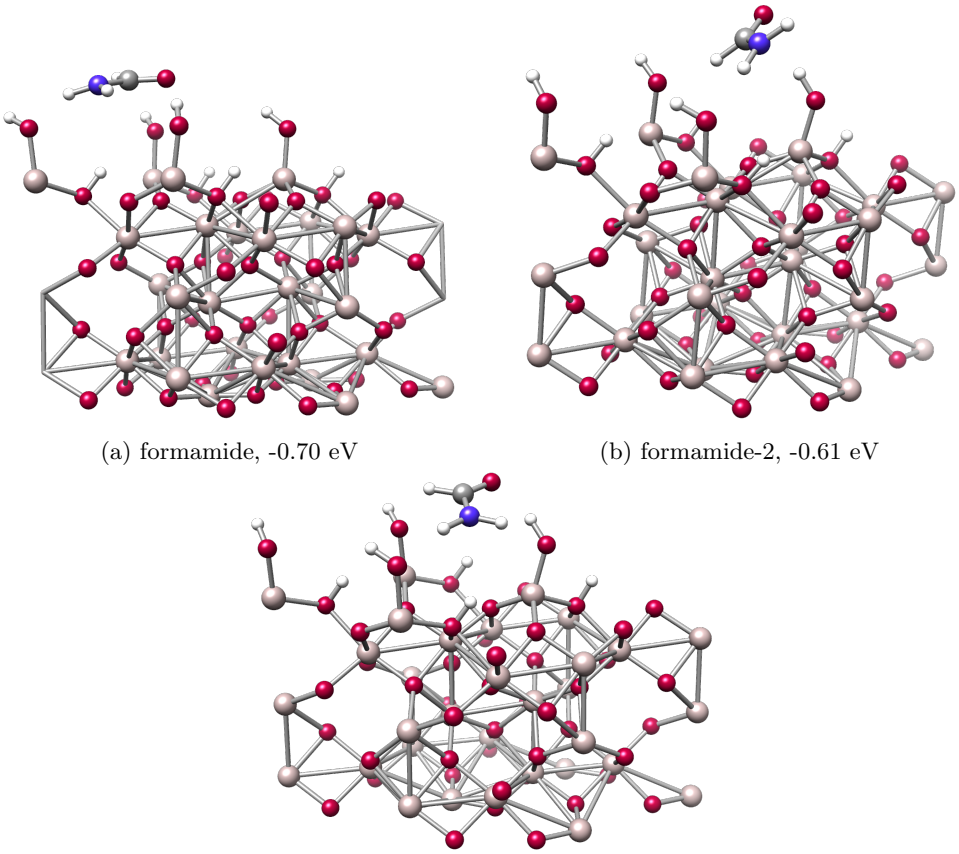
(b) oxalic acid-2, -0.75 eV



(c) oxalic acid-3, -0.30 eV

Figure 4.6: Optimised adsorption geometries and binding energies of oxalic acid adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

Amides were the next class of molecules to be adsorbed onto the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface, first we adsorbed formamide the simplest amide. All of the adsorption geometries obtained for formamide adsorption onto the surface form two hydrogen bonds (figures 4.7 (a), (b) and (c)). The strongest adsorption geometry has an adsorption energy of -0.70 eV, which is less stable than acetic acid, but comparable with oxalic acid. All hydrogen bonds in the optimised geometries were formed between a surface hydrogen and the oxygen of the amide (1.66 Å) and the $NH_2$ hydrogen and surface oxygen (1.84 Å). Formamide-2 (figure 4.7 (b)) was the second most stable amide adsorption geometry with a calculated binding energy of -0.61 eV and hydrogen bond lengths of 1.87 Å and 1.75 Å. The final formamide geometry, formamide-3 has a calculated adsorption energy of -0.46 eV and hydrogen bond lengths of 1.94 Å and 2.27 Å.



(a) formamide, -0.70 eV

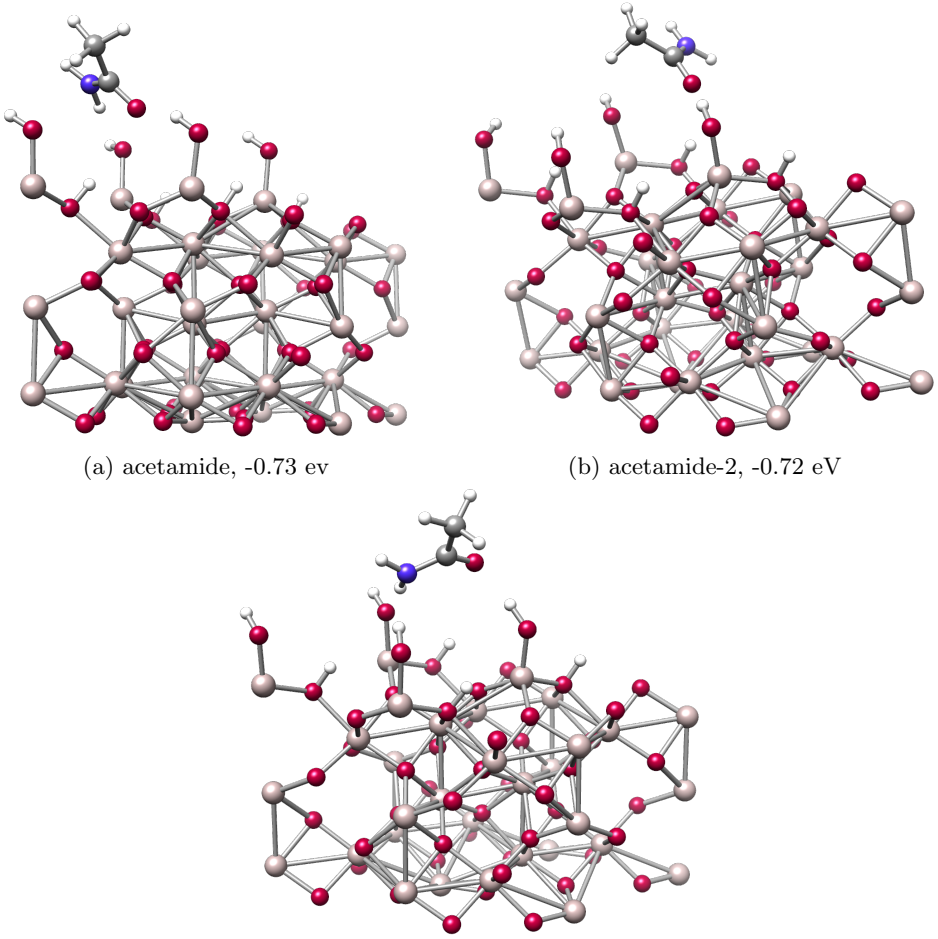
(b) formamide-2, -0.61 eV

(c) formamide-3, -0.46 eV

Figure 4.7: Optimised adsorption geometries and binding energies of formamide adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

For comparison with formamide, the acetamide molecule which contains a donor (methyl) group was also adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. Similar to formamide and the other small organic molecules we adsorbed we considered several adsorption configurations for acetamide. Acetamide (figure 4.8 (a)) and acetamide-3 (figure 4.8 (c)) adsorbed to the hydroxylated surface via two hydrogen bonds, Acetamide-2 adsorbed via a single hydrogen bond. Acetamide-1 geometry was the most stable geometry, with an adsorption energy of -0.73 eV, the structure is similar to that of acetic acid-1. The hydrogen bonds in this structure were formed between the surface hydrogen and the amide oxygen (1.85 Å) and the $NH_2$ hydrogen and a surface oxygen (1.74 Å). Acetamide-2 (figure 4.8 (b)) was the second most stable, with an adsorption energy of -0.72 eV and a hydrogen bond length of 1.77 Å. Acetamide-3 (figure 4.8 (c)) was the least stable of the geometries with a binding energy of -0.69 eV. Its hydrogen bonds have a length of 2.32 Å and 1.77 Å. The hydrogen bonds in this case are longer than those of the first acetamide structure. The bonds were also formed between a surface hydrogen and an acetamide nitrogen and an acetamide oxygen and a surface OH hydrogen. Thus, the hydrogen bonds are weaker, as hydrogen bonds formed between a N-H are weaker than those that form between O-H, as oxygen is more electronegative. The addition of a methyl group to the amide functional group did not affect the strength of binding, as the binding

energies obtained for acetamide were very similar to those obtained for formamide (approximately -0.70 eV) However, there is a difference in adsorption geometries. formamide is small and adsorbs almost flat, whereas acetamide adsorbs tilted or upright. Both acetic acid and acetamide can be considered better representations of biomolecules, because biomolecules themselves are bulky in nature.



(a) acetamide, -0.73 ev

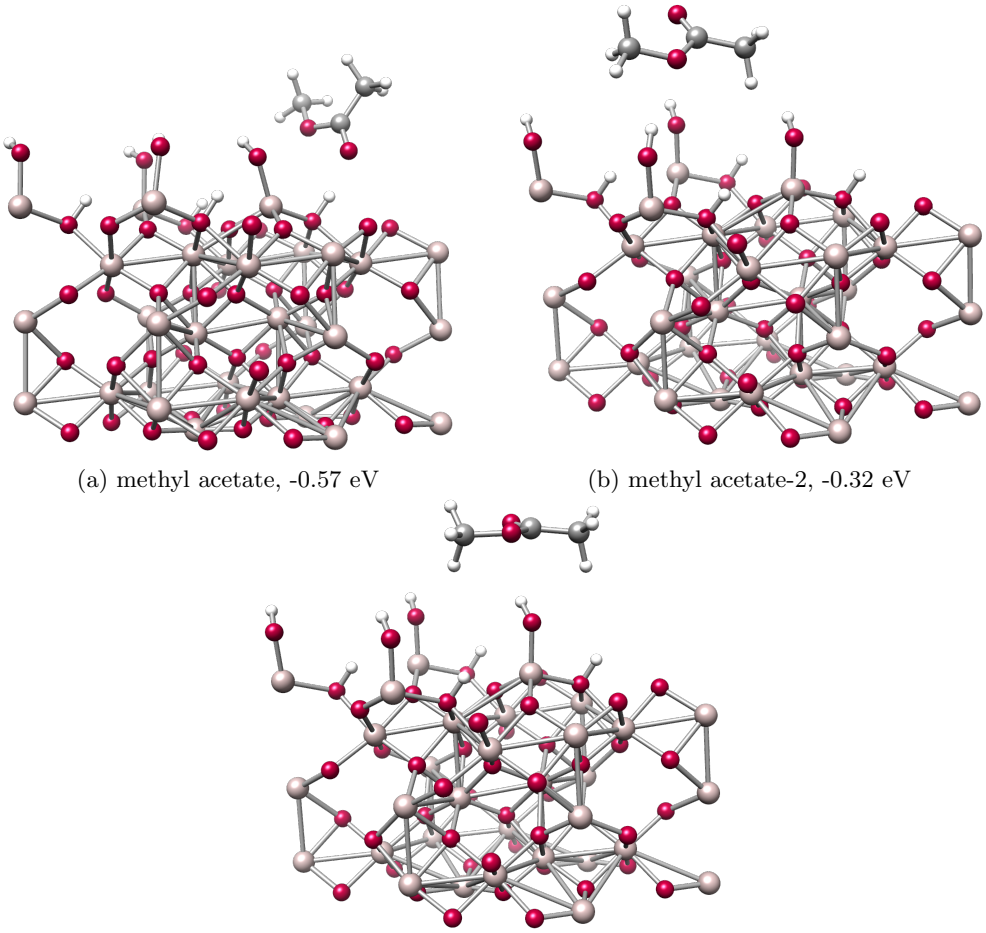
(b) acetamide-2, -0.72 eV

(c) acetamide-3, -0.69 eV

Figure 4.8: Optimised adsorption geometries and binding energies of acetamide adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

Lastly, we adsorbed methyl acetate to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. The most stable geometry has a calculated binding energy of -0.57 eV and a hydrogen bond length of 1.77 Å. The first methyl acetate geometry (figure 4.9 (a)) illustrates why the ester functional group is not as stably adsorbed as some of the other organic functional groups. This is due to the O-O repulsion between the surface OH oxygen and an oxygen on the ester. Methyl acetate-2 (figure 4.9 (b)) which has a hydrogen bond to the ester oxygen is the second most stable ester geometry with a calculated binding energy of -0.32 eV. The least stable geometry was that of methyl acetate-3 (figure 4.9 (c)) where the molecule was lying flat. The binding energy is -0.19 eV and no hydrogen bonds were detected for this structure. Therefore, the ester adsorbed to the hydroxylated $\alpha$-$Al_2O_3$

(0001) surface via VDW dispersion interactions.



(a) methyl acetate, -0.57 eV

(b) methyl acetate-2, -0.32 eV

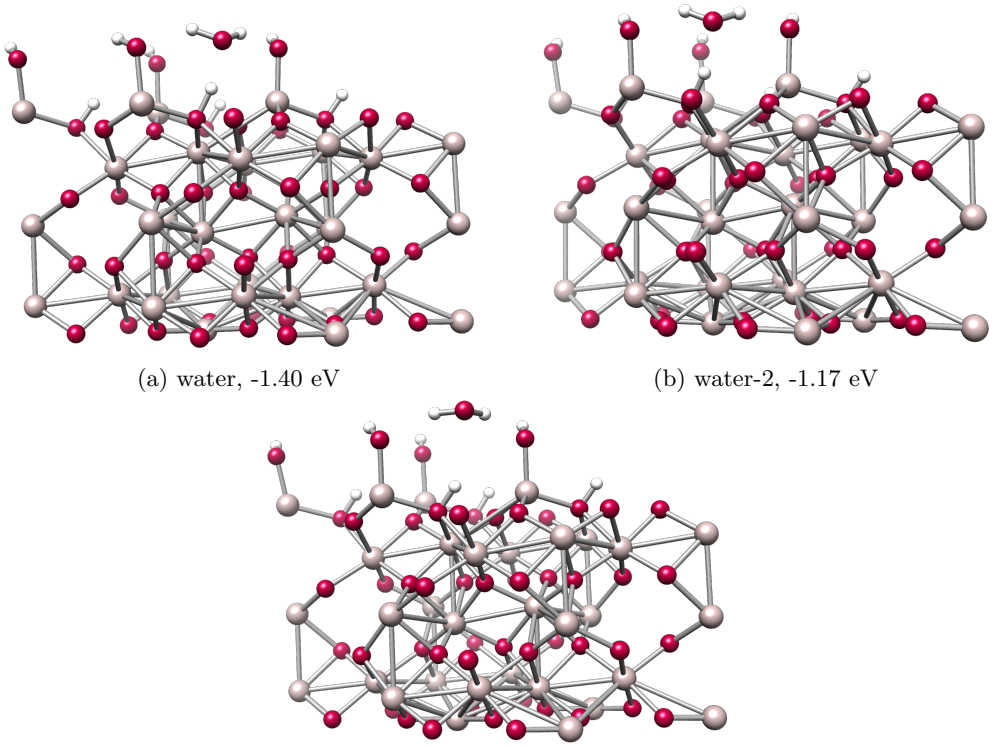
(c) methyl acetate-3, -0.19 eV

Figure 4.9: Optimised adsorption geometries and binding energies of acetamide adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

### 4.4.2 Adsorption of Water on the Hydroxylated $\alpha$-$Al_2O_3$ (0001) Surface

Our final adsorbate was water, it was adsorbed to the hydroxylated surface as it is naturally present in soils. Therefore, the binding energy was of particular interest. Amongst all adsorbates water (water-1 figure 4.10 (a)) adsorbed most strongly to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface with a binding energy of -1.40 eV. Upon analysis of this structure, three hydrogen bonds have been identified between the water molecule and the surface. The first of the these hydrogen bonds is between a surface OH oxygen and the H of the water molecule, this is the longest of the three hydrogen bonds with a bond length of 1.83 Å. The second hydrogen bond occurred between the hydrogen on water and another surface OH oxygen, this had a bond length of 1.76 Å. The final hydrogen bond identified was the shortest and strongest of the three bonds with a length of 1.55 Å. This bond formed between the water molecule's oxygen and a surface hydrogen. Similar results can be seen in table 4.2 for the adsorbate water-2 (figure 4.10 (b)), which also formed three hydrogen

bonds upon water adsorption. Water-2 has a binding energy of -1.17 eV and hydrogen bond lengths of 1.86, 1.79 Å and 1.61 Å respectively. This adsorption geometry is similar in structure to the first water structure (figure 4.10 (a)). However, the hydrogen bonds are longer and thus the adsorbed molecule is slightly less stable, as water is less tightly held. Upon analysis of water-3, only two hydrogen bonds were detected (1.93 Å and 1.91 Å). It is significantly less stable than water and water-2 with a binding energy of -0.47 eV. The structure of water-3 (figure 4.10 (c)) illustrates why water-3 was unable to form three hydrogen bonds with the surface. Figure 4.10 (c) illustrates that the water molecule' oxygen atom is pointing upwards. Thus, a hydrogen bond cannot form between that oxygen atom and a surface hydrogen.



(a) water, -1.40 eV

(b) water-2, -1.17 eV

(c) water-3, -0.47 eV

Figure 4.10: Optimised adsorption geometries and binding energies of water adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

Although water is most strongly adsorbing of all the adsorbates, it has adsorbed in the same plane as the surface hydroxyl groups. Whereas, the small organic adsorbates all adsorbed above the plane of the surface hydroxyl groups. Therefore, the organic molecules would still be able adsorb to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface after water adsorption. The organic adsorbates could also potentially push water molecules into the free spaces between surface hydroxyl groups upon adsorption. However, we have not tested this hypothesis. The most strongly adsorbing organic molecule in these investigations was the acid, followed by the amine, amide and alcohol groups that had similar binding energies.

### 4.4.3 Molecular Dynamics

During the adsorption of acetic acid to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface, acetic acid-MD (figure 4.11) was obtained. This geometry is distinct, as acetic acid chemisorbed to the surface, unlike the other structures whereby binding was facilitated by hydrogen bonding. Acetic acid-MD had a binding energy of -0.64 eV and chemisorption bond lengths of 1.93 Å ($H_2O$-Al) and 2.08 Å (Al-OH of $CH_3COOH$). Acetic acid-MD adsorbed via the OH oxygen (similar to the bare surface). Interestingly, upon adsorption of acetic acid the surface OH bonded to the same Al atom acquired a surface H, thus, resulting in the formation of a water molecule. Therefore, we employed MD calculations to determine whether water or the acid would prefer to desorb. We ran molecular dynamics calculations at 298K for 7.5 ps starting from the acetic acid-MD structure using the NVT ensemble to determine the movements of atoms. Several calculations were run for a sufficiently long period of time. The simulations also considered two temperatures, 298K and 373K. We have reported the results for the 298K calculation results, as the 373K calculation did not yield significant results.

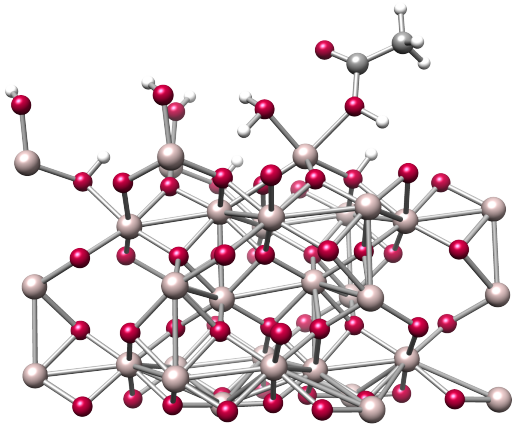


Figure 4.11: Acetic acid adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface (for referencing purposes this structure will be called acetic acid-MD

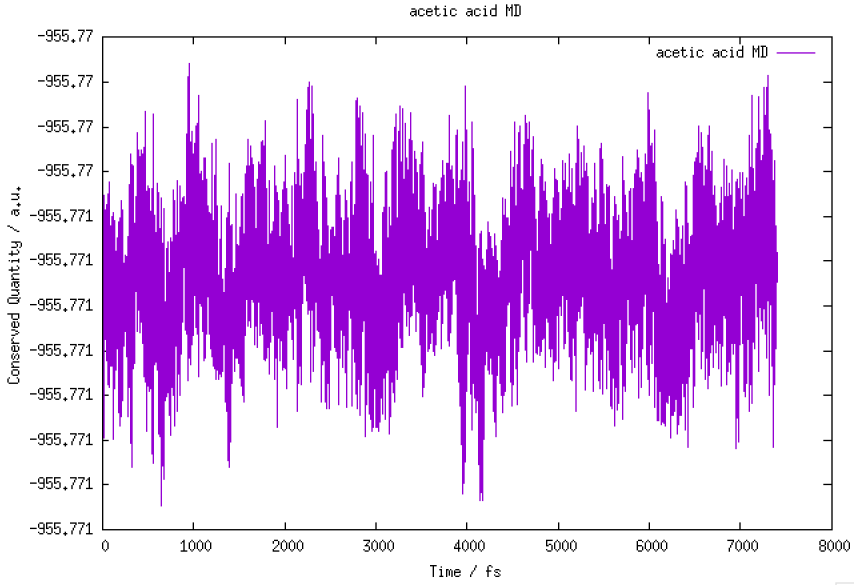



Figure 4.12: A graph of the total energy in a.u. against time in fs for the MD calculation of acetic acid-MD.

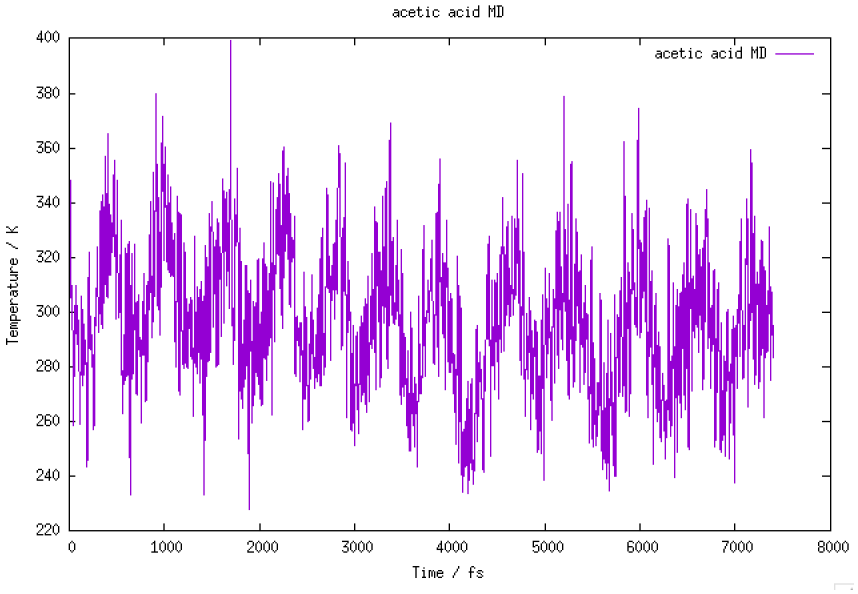



Figure 4.13: A graph of temperature in K against time in fs for the MD calculation of acetic acid-MD.

Figure 4.12 is an illustration of how the energy of acetic acid-MD changed during the MD calculation. It can be seen that the energy remains relatively constant during the entire calculation, as it ranges between -955.771 a.u. and -955.770 a.u. which is a change of <0.03 eV. Similarly, the temperature also remained constant, as illustrated in figure 4.13. Figure 4.13 illustrates some temperature oscillations between 220 and 400 K. However, the average temperature remained at around 298K. Inspection of the MD trajectory in figure 4.14 illustrates that neither the acid or water would desorb. Figure 4.14 shows that there has been some movement of both the acid and the water molecule on the surface. Upon inspection of the bonds, the $H_2O$-Al bond has extended

to a length of 2.08 Å and the Al-O bond of the acid to the surface has shortened to 1.86 Å. Figure 4.14 also illustrates that the acid OH bond has extended, from 1.02 Å (in figure 4.11) to 1.31 Å. The extension of the $H_2O$-Al bond could suggest that water would be the leaving group. Interestingly the acid is now adsorbed through the carbonyl, whereas it was adsorbed via the hydroxyl at the start of the simulation (figure 4.11). On the timescales of our MD calculations (7.5 ps) desorption of either water or the acid did not take place. Therefore, longer simulations would be required to observe desorption of these groups

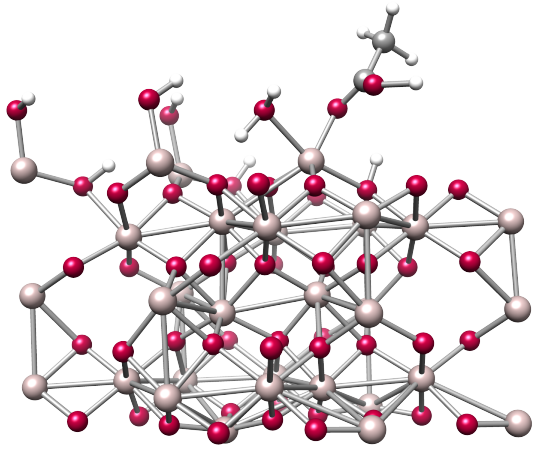


Figure 4.14: Geometry of acetic acid-MD at the end of the MD calculation

## 4.5 Conclusion

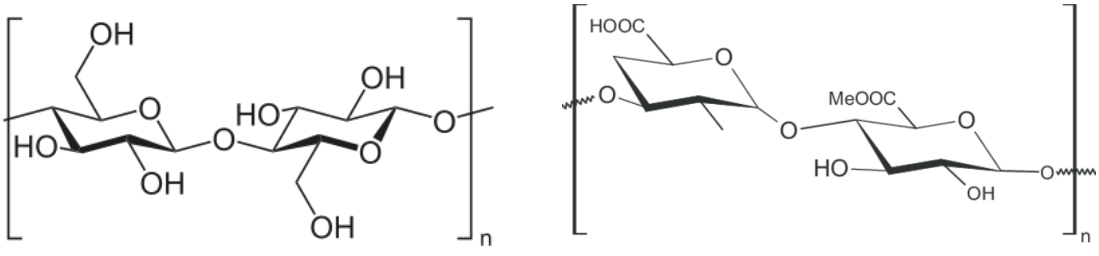
In this chapter we investigated the binding of small organic molecules with the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. This investigation gave a more realistic representation of binding to soil minerals as the alumina surface is hydrated in natural soil environments, due to the presence of water. Binding to this surface was facilitated by hydrogen bonding between the adsorbate and the surface hydroxyl layer. Acids were the strongest adsorbates amongst the small organic functional groups. All of the organic adsorbates were able to form two hydrogen bonds and sometimes three with the surface, yielding stable geometries. We also investigated the leaving group ability of water and the acid functional group using MD simulations. However, conclusive results were not obtained from these investigations.

A follow up investigation will now be described, where we will bind complex naturally occurring biopolymers to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface instead of small organic molecules. These investigations will produce a more comprehensive understanding of binding between OC and soil minerals.

# Chapter 5

# Biopolymers

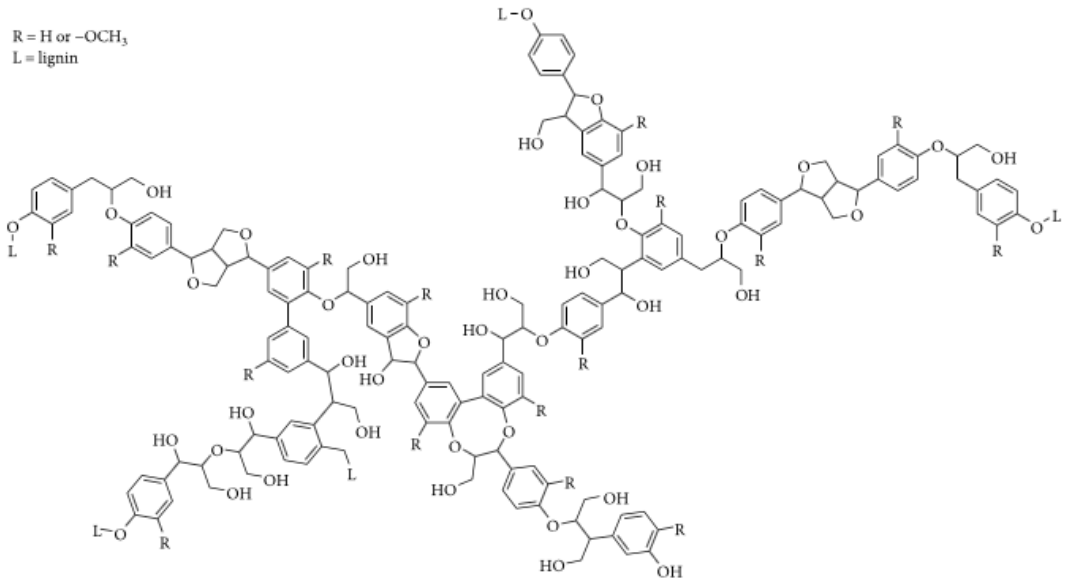
Biomolecules are compounds that can form in nature as part of the life cycles of organisms, they range in structure from simple organic acids, to complex biopolymers.[1] Biopolymers are typically classified into three groups, these groups are based on their origins. Biopolymers can be plant-based, animal-derived or produced by microorganisms.[1,231,232] Plant-based biopolymers are typically produced from plants (dead and alive) and agricultural scraps. Cellulose, pectin and lignin are all examples of plant-based biopolymers. Cellulose is the most abundant biopolymer on earth and the major carbohydrate found in plants. It is is a polysaccharide formed from the polymerisation reaction of monosaccharides. Two units of the glucose monosaccharide undergo a polymerisation reaction to produce cellulose, through the elimination a water molecule. The glucose units are linked together by an oxygen molecule and this link is known as the 1-4 glycosidic bond.[1] Pectin is predominantly a polymer of D-galacturonic acid, it includes oxygen containing functional groups such as aldehyde, carboxylic acid and alcohol groups, with a high exchange capacity.[233,234] Pectin is present within the cell walls of plants, particularly in root tips.[235] It can also be extracted from agricultural by-products such as apple pomace and citrus peel.[236] After cellulose, lignin is the most abundant biopolymer on earth.[231] Lignin is a relatively hydrophobic aromatic polymer, that contributes to the formation of cell walls.[237] Lignin accounts for approximately one-third of plant litter biomass. Alongside its alcohol functional groups, the phenol monomers in lignin can also have aliphatic alcohols and one or two methoxy (O-$CH_3$) substituents on the benzene ring.[1]



(a) cellulose



(b) pectin




(c) lignin

Figure 5.1: Chemical structures of cellulose,[12] pectin[13] and lignin.[14]

Animal-derived biopolymers are the biopolymers that originate from animal sources. Chitin and chitosan are common animal-derived biopolymers.[231] Chitin is the most abundant amino-polysaccharide in nature, it is the building material that gives strength to the exo-skeletons of insects, crustaceans, and the cell walls of fungi.[238] Through enzymatic or chemical deacetylation, chitin can be converted to its well known derivative chitosan. Chitin can be completely deacetylated or partially deacetylated, to yield chitosan. Chemically, it can be extracted by treatment with sodium hydroxide.[238–243] Chitosan is a water soluble biopolymer, that tends to bind to negatively charged surfaces.[231,244]

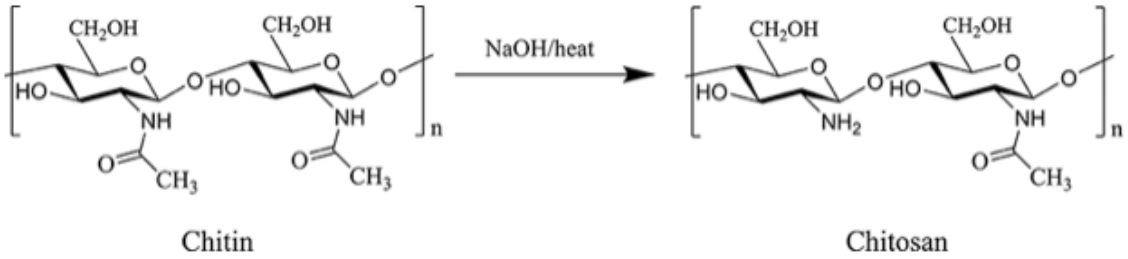



Figure 5.2: Chemical reaction between chitin and NaOH to produce partially deacetylated chitosan.[15]

Biopolymers such as xantham gum and dextran are the products of bacterial fermentation processes. These are considered as microorganism based biopolymers. Xantham gum is an anionic polysaccharide, produced through the fermentation of glucose and sucrose by the *Xanthomonas campestris* bacterium.[231] Dextran is an insoluble biopolymer. It is produced through the bacterium known as *Leuconostoc mesenteroides*.[245] These biopolymers are all commonly present within soils.[1]

## 5.1 Adsorption of Biopolymers

Interactions between biopolymers and inorganic surfaces play an important role in soils. The retention of biomolecules via their interaction with soil surfaces is also of great importance.[246] The interaction of large biomolecules and soil mineral surfaces is believed to improve the soil strengths.[247] Large biomolecules are believed to be selectively retained by mineral surfaces within soils. Thus, protecting soils from microbial degradation[248–250] Theoretical and experimental investigations on the adsorption of large biomolecules onto mineral surfaces have suggested that strong adsorption occurs across a range of conditions.[248,251–259] Inter-facial contact between the surface and adsorbate, may result in the conformational rearrangement between the adsorbed biomolecule and mineral surface. This is an adsorption requirement as it maximises the interaction between the adsorbed biomolecule and the mineral surface[253,255,259,260] Surface adsorption of biopolymers has been demonstrated the release large amounts of free energy, thus favouring inter-facial binding of biopolymers at mineral surface interfaces in experimental investigations. This is because the adsorption geometry (biopolymer + surface) is more stable than the biomolecule and free surface alone.[248,250,261]

For example, Hoogendam et al.[262] conducted adsorption and desorption studies of carboxymethyl cellulose (CMC) on $TiO_2$ and $\alpha$-$Fe_2O_3$ surfaces experimentally. Long and short chains of CMC were found to adsorb to the $\alpha$-$Fe_2O_3$ surface at low biopolymer concentrations in the poly-disperse system. However, with an increase of biopolymer concentration the number of long chains adsorbing increased, thereby displacing the shorter chains. The desorption of CMC on the $\alpha$-$Fe_2O_3$ surface was found to be a very slow process. The effects of the pH on adsorption was also investigated. The pH investigations showed that the amount of CMC adsorbed to both the $TiO_2$ and $\alpha$-$Fe_2O_3$ surfaces decreased with increasing pH. At a high pH CMC will be deprotonated. Therefore, increasing the pH results in electrostatic repulsion between CMC chains and $TiO_2$ and $\alpha$-$Fe_2O_3$ surfaces, particularly between negatively charged oxygen anions.

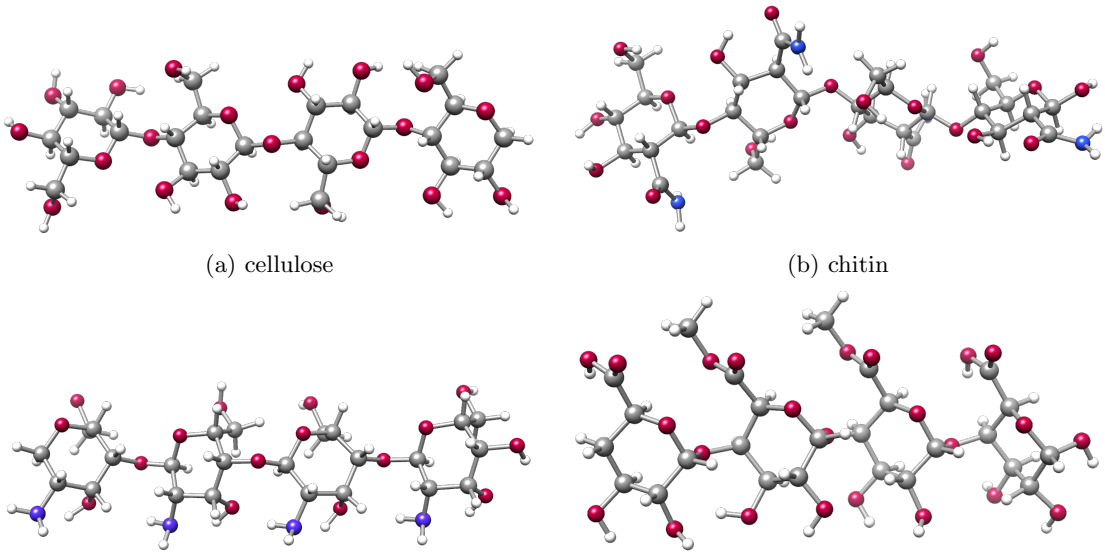
The adsorption of pectin was also investigated. Pectin was adsorbed on to the amorphous iron and aluminium hydroxide surfaces by Wang et al.[233] They conducted batch experiments on the ad-

sorption of pectin to amorphous iron and aluminium hydroxide surfaces, and examined the results using X-ray diffraction, scanning electron microscopy and FTIR spectroscopy. The adsorption of pectin on the aluminium and iron surfaces was determined at different pH values. The amount of pectin adsorbed by amorphous aluminium hydroxide was much greater than the amount of pectin adsorbed by amorphous iron hydroxide in the pH range of 3.5 - 6.5. This was attributed to the amount of positive charge on the amorphous iron and aluminium hydroxide surfaces. At the same pH the positive charge on amorphous aluminium hydroxide was greater than that of amorphous iron hydroxide. Thus, the electrostatic interaction between pectin and the amorphous aluminium hydroxide surface was larger than that of the iron hydroxide. With an increasing pH the adsorption of pectin to both hydroxides decreased. The decrease in positive charge on both Fe and Al hydroxides led to a decrease in the electrostatic interaction between pectin and the hydroxides, thus leading to decrease in the amount of pectin adsorbed.

The adsorption of Nucleic acids (deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) onto mineral surfaces has also been investigated.[248,263–265] Nucleic acids are macro-molecular biopolymers and are produced by bio-active macro-molecules. They are present in soil through excretion from microorgasnisms, plants and animals or by lysis of dying cells.[263,266,267] Schmidt et al.[248] investigated the adsorption of DNA/BSA (bovine serum albumin protein) complex onto the goethite (FeOOH) surface. Spectroscopic studies of DNA/BSA complex adsorption showed that complexation of DNA to the protein BSA hindered the binding of DNA's phosphodiester groups with goethite. However, complexation of DNA with BSA reduced DNA-DNA electrostatic repulsion. Increasing the concentration of the BSA protein in the DNA/BSA complex enhanced the binding of DNA to goethite, due to the reduced repulsion between the helices of DNA. These studies were carried out as the formation of molecule rich supra-molecular complexes are important for the development of mineral stabilised organic matter.[248] These investigations illustrate that the binding of large biomolecules has been investigated. However, there are few literature studies and the nature of the interactions between other important biomolecules is yet to be understood.

In this research we aim to investigate the binding of several biopolymers to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. We will be investigating the strength of binding to the hydroxylated alumina surface, which will yield a greater understanding of the mechanism by which these molecules bind. We will be binding plant-derived biopolymers (cellulose and pectin) and animal-derived biopolymers (chitin and chitosan), that are commonly present in soils.[1] We will determine which biopolymer yields the most strongly bound adsorption structures. Four biopolymers with different functional groups have been selected to be adsorbed to the surface. We aim to achieve the most comprehensive understanding of OC interaction with oxide minerals in soils through our biopolymer

adsorption investigations.



(a) cellulose

(b) chitin

(c) chitosan

(d) pectin

Figure 5.3: Structures of optimised biopolymers to be adsorbed onto the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

## 5.2 Computational Method

### 5.2.1 Biopolymer Adsorption to the Hydroxylated $\alpha$-$Al_2O_3$ (0001) Surface

DFT calculations of the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface and adsorption structures were performed using the CP2K software package.[197–199] All DFT calculations used the GGA PBE exchange-correlation functional[200] with Grimme's D3[149] empirical dispersion correction. The calculations utilised double-$\zeta$ basis sets with diffuse and polarization functions (DZVP) optimised for use in CP2K,[203] and Goedecker-Teter-Hutter (GTH) pseudopotentials.[203] All calculations were done at the $\Gamma$ k-point.[198] The hydroxylated $\alpha$-$Al_2O_3$ (0001) surface was modelled using periodic slabs containing 12 atomic layers, where the lattice parameters were fixed at their optimised bulk values, while the positions of all atoms were fully optimised. Slabs were separated by 15 Å of vacuum in the vertical direction and were 6x6 extended to accommodate the adsorbate. The slab had a length of 47.16 Å and 828 atoms in total. The calculated binding energies reported in this work were corrected for the basis set superposition error (BSSE) using the counterpoise method,[112] the correction ranged from 0.05 - 0.11 eV. All hydrogen bonds in the calculated structures were analysed using the vmd[268] software package with an angle cut off of 35. All structures were visualised with the UCSF Chimera[205] software package.

## 5.3 Results and Discussion

### 5.3.1 Adsorption of Biopolymers to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

The first step in our adsorption calculations of biopolymers to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface was extending the 2x2 the hydroxylated surface (the surface in chapter 4). The size of the surface was chosen based upon the size of the biopolymers. As the average biopolymer length was approximately 18 Å, the surface was 6x6 extended.

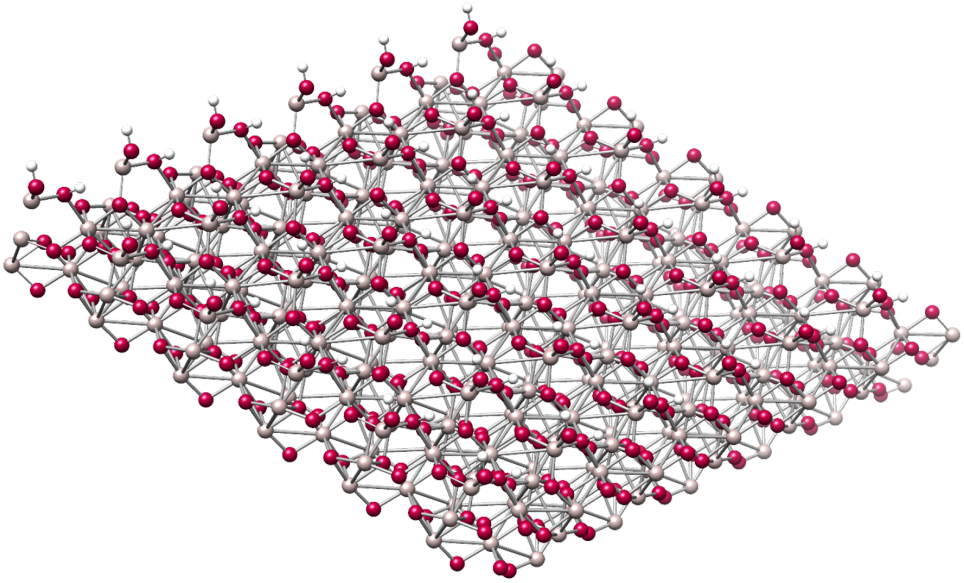


Figure 5.4: Optimised geometry of the 6x6 extended hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

The adsorption of four common biopolymers to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface was investigated. These biopolymers were chosen due to their natural presence within soils as naturally occurring biopolymers. We also selected these biopolymers, based on our investigations on the adsorption of small molecules to the bare and hydroxylated alumina surfaces. These investigations determined that the acid, amine, and amide groups bind to the alumina surface very strongly and the alcohol group adsorbs with a similar strength to water. Therefore, we selected cellulose (R-OH groups), chitin (R-OH and R-$NH_2$), chitosan (R-OH and R-$CONH_2$) and pectin (R-OH and R-COOH). Our model biopolymers contained four monosaccharide units. We employed a systematic procedure for adsorbing our biopolymers to the surface. For each biopolymer several adsorption geometries were considered. Biopolymers were adsorbed both parallel and perpendicular to the x-axis along the lengths of the surface OH rows. We also adsorbed the biopolymers flat and parallel, as well as in an upright position where they were adsorbed on the edge.

We calculated the binding energies and deformation energies in eV for all biopolymers that were

adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. The binding energies were calculated using the following equation

$$E_{binding} = E_{surface+Adsorbate} - (E_{surface} + E_{Adsorbate}) \tag{5.1}$$

The deformation energies were calculated using the following equation with values generated from BSSE calculations

$$E_{deformation} = E_{deformation\ surface} + E_{deformation\ adsorbate} \tag{5.2}$$

where $E_{deformation\ surface}$ is the deformation energy of the surface and $E_{deformation\ adsorbate}$ is the deformation of the adsorbate, in this case the biopolymer.

The deformation energy of the $E_{deformation\ surface}$ subsystem can be calculated using the following

$$E_{deformation\ surface} = E_{surface+adsorbate} - E_{surface} \tag{5.3}$$

where $E_{surface+adsorbate}$ is the energy of the surface and the adsorbate and $E_{surface}$ is the energy of the surface. Similarly, the $E_{deformation\ adsorbate}$ can be calculated using the following equation

$$E_{deformation\ adsorbate} = E_{surface+adsorbate} - E_{adsorbate} \tag{5.4}$$

where $E_{adsorbate}$ is the energy the isolated adsorbate.

We have investigated five different geometries illustrating a variation of different binding abilities for each polymer. The results have been illustrated in a graph (figure 5.5) which shows a correlation between the binding energy to the total number of hydrogen bonds formed between the biopolymer and the surface upon adsorption.

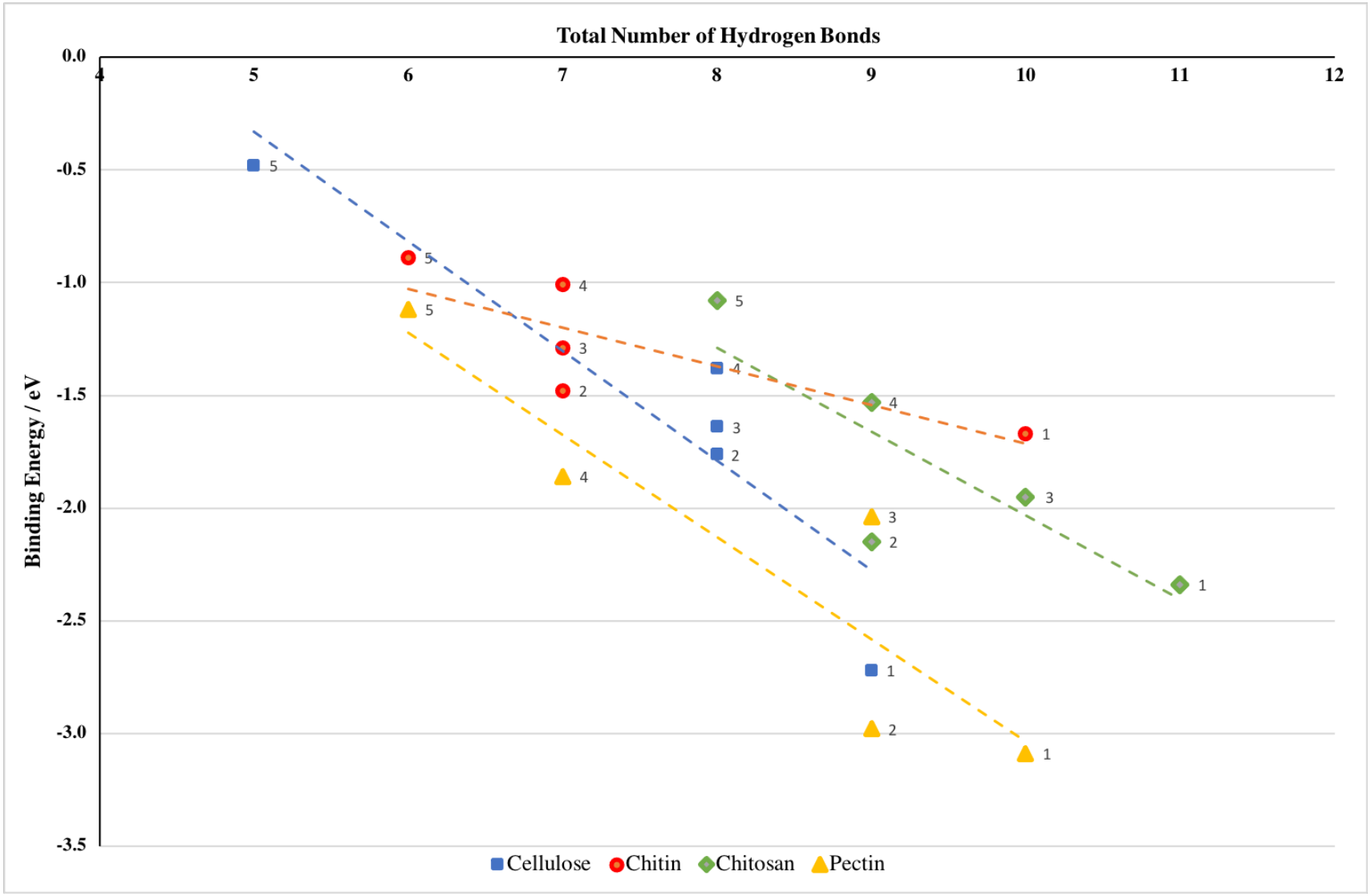



Figure 5.5: Number of hydrogen bonds against the binding energy in eV of different adsorption geometries for several biopolymers adsorbed onto the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

### Cellulose

Cellulose was the first biopolymer to undergo adsorption to the hydroxylated $\alpha$-$Al_2O_3$ surface. It was adsorbed in the centre of the surface for visual clarity. We adsorbed several different geometries of cellulose along different surface OH rows, as well as flat and parallel and on the edge.

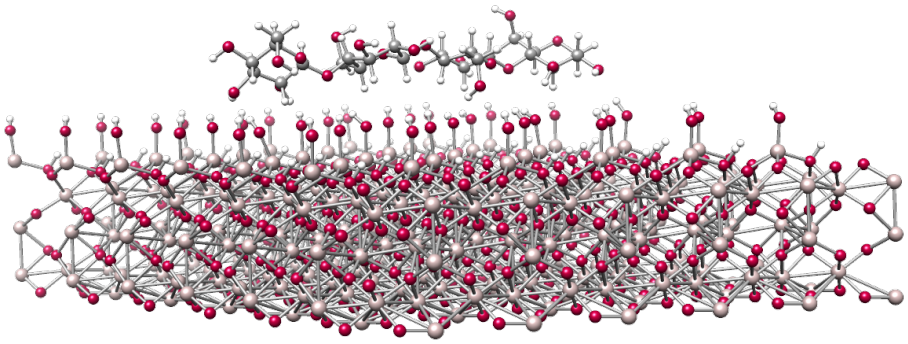


Figure 5.6: Adsorption geometry of cellulose-1, during the adsorption of cellulose to the hydroxylated $\alpha$-$Al_2O_3$ surface.

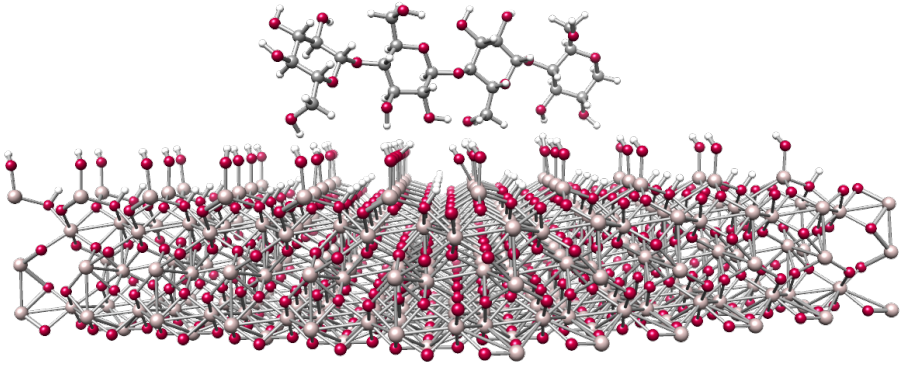


Figure 5.7: Adsorption geometry of cellulose-2, during the adsorption of cellulose to the hydroxylated $\alpha$-$Al_2O_3$ surface.

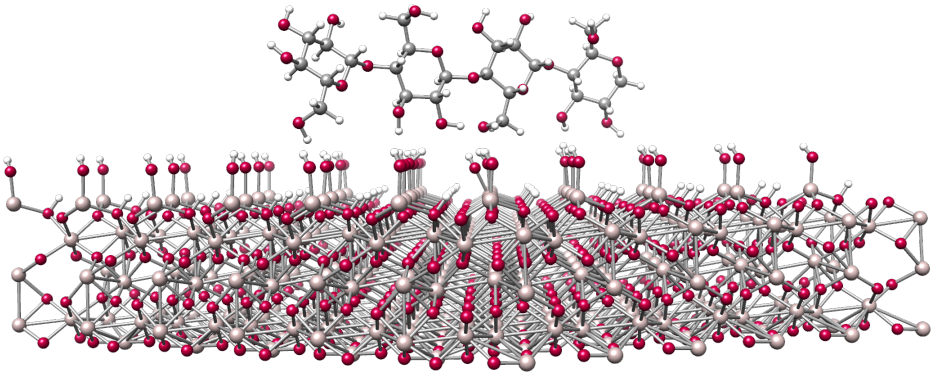


Figure 5.8: Adsorption geometry of cellulose-3, during the adsorption of cellulose to the hydroxylated $\alpha$-$Al_2O_3$ surface.

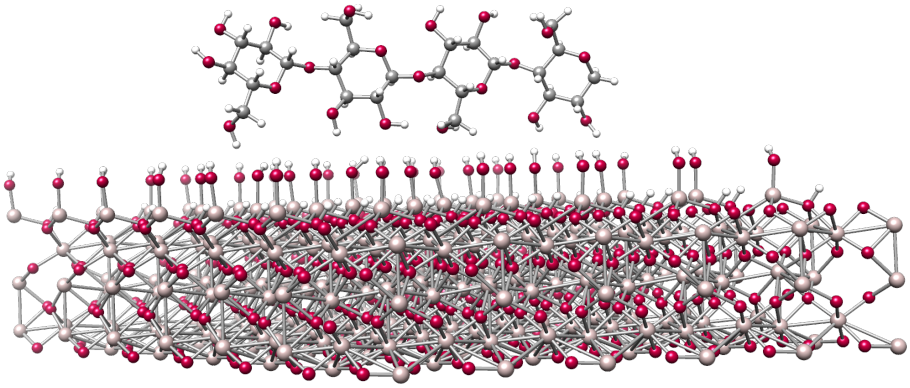


Figure 5.9: Adsorption geometry of cellulose-4, during the adsorption of cellulose to the hydroxylated $\alpha$-$Al_2O_3$ surface.

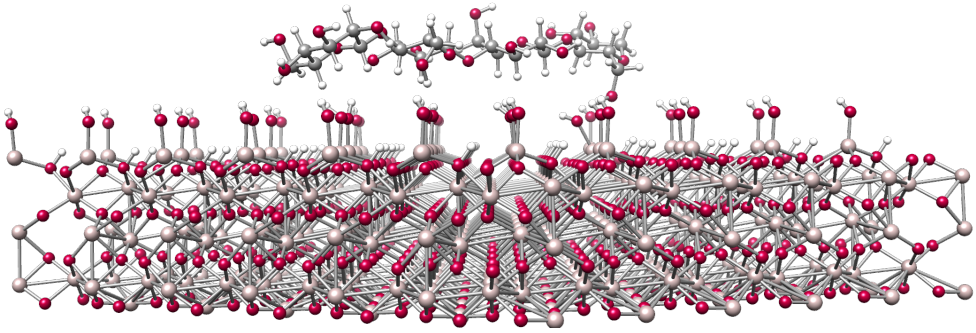


Figure 5.10: Adsorption geometry of cellulose-5, during the adsorption of cellulose to the hydroxylated $\alpha$-$Al_2O_3$ surface.

Table 5.1: Calculated binding energies, deformation energies and total number of hydrogen bonds formed between cellulose and the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

| **Biopolymer** | **Binding Energy / eV** | **Deformation Energy / eV** | **No. of H-Bonds** |
|---|---|---|---|
| Cellulose-1 | -2.72 | 0.58 | 9 |
| Cellulose-2 | -1.76 | 0.55 | 8 |
| Cellulose-3 | -1.64 | 0.59 | 8 |
| Cellulose-4 | -1.38 | 0.40 | 8 |
| Cellulose-5 | -0.48 | 11.45 | 5 |

Table 5.2: Hydrogen bond analysis of the hydrogen bonds that formed during the adsorption of cellulose to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface, all bond lengths are given in Å.

| **H-bond** | $\mathbf{O}_s\cdots\mathbf{H}_m$ | $\mathbf{O}_m\cdots\mathbf{H}_s$ | $\mathbf{O}_m\cdots\mathbf{H}_m$ |
|---|---|---|---|
| **Biopolymer** | | | |
| Cellulose-1 | 1.78, 2.05, 2.35 | 1.83, 2.02, 2.36 | 2.30, 2.33, 2.40 |
| Cellulose-2 | 1.85, 1.87, 2.23, 2.34 | 1.83, 1.92 | 2.25, 2.43 |
| Cellulose-3 | 1.72, 1.88, 1.92 | 1.78, 1.92, 2.04 | 2.30, 2.39 |
| Cellulose-4 | 1.86, 2.09, 2.19 | 1.86, 2.03, 2.10 | 2.31, 2.39 |
| Cellulose-5 | 1.81 | 1.71, 2.44 | 2.26, 2.48 |

Where the subscripts $_s$ and $_m$ represent the surface and the biopolymer.

The adsorption structures of cellulose have been labelled cellulose 1-5, to illustrate the different geometries upon adsorption of cellulose. The most stable geometry obtained during the cellulose investigations is illustrated in figure 5.6. The calculated binding energy of cellulose-1 is -2.72 eV. Figure 5.6 illustrates that cellulose adsorbs most strongly to the surface when it is lying flat and parallel to the surface. In this geometry it is able to form six H-bonds with the surface, which are additionally stabilised by three internal H-bonds within cellulose. Cellulose-2 is another stable geometry (as illustrated in figure 5.7). Figure 5.7 illustrates that cellulose has adsorbed in an vertical position on its edge, whereby the OH groups of cellulose are aligned with the OH groups of the surface. This alignment maximises hydrogen bonding between the surface and the adsorbate. Eight H-bonds formed between the surface and cellulose molecule in this geometry, of which six bound the adsorbate to the surface (similar to cellulose-1) and two were internal cellulose H-bonds. However, the binding energy is less negative than that of cellulose-1, indicating cellulose-2 has a less stable than cellulose-1. Hence, cellulose adsorbs more strongly when it adopts a flat and parallel geometry, than when it adsorbed on its edge. The larger number of hydrogen bonds in cellulose-1 can be attributed to its structure, as it is flat. Therefore, the dispersion interactions are greater, resulting in a more stable geometry.[269–271]

Cellulose-3 (figure 5.8) has a binding energy (-1.64 eV) which is similar to that of cellulose-2 (-1.76 eV), however it is sightly less stable. This geometry is less stable, as it was adsorbed off centre relative to the row of OH groups on the surface. The final cellulose geometry (cellulose-5) represents the weakest calculated adsorption energy. In the structure of Cellulose-5 (figure 5.10), three H-bonds were detected between the adsorbate and the surface, unlike its counterparts that form 6 H-bonds. Hence, it has a significantly less stable binding energy of -0.48 eV, as fewer H-bonds were formed between cellulose and the surface. The reduced stability of cellulose-5 can also be attributed to its deformation energy, as seen in table 5.1. Cellulose 1 to 4 have deformations

energies ranging from 0.40 to 0.58 eV. These values suggest there is little deformation of the biopolymer and surface upon adsorption. However, the deformation energy of cellulose-5 is 11.45 eV. Thus, suggesting that the cellulose biopolymer and surface significantly deformed upon the adsorption of cellulose, to maximise interactions of cellulose with the surface. Upon inspection of figure 5.10, the deformation of the surface is visible via an out of plane surface OH group. Upon further inspection, four OH groups were found to have deformed alongside an OH on cellulose. We were unable to illustrate all of the deformities, however a clear surface OH deformity can be seen in figure 5.10.

### Chitin

The second biopolymer to undergo adsorption to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface was chitin. Chitin is an animal-derived biopolymer that contains, alcohol and amide functional groups. Unlike its biopolymer counterparts, only half of the chitin molecule can be adsorbed at any given time. This is because adjacent chains in chitin alternate, such that they are anti-parallel resulting in a twisted configuration.[272–274]

Table 5.3: Calculated binding energies, deformation energies and total number of hydrogen bonds formed between several chitin and the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

| **Biopolymer** | **Binding Energy / eV** | **Deformation Energy / eV** | **No. of H-Bonds** |
|---|---|---|---|
| Chitin-1 | -1.67 | 1.10 | 10 |
| Chitin-2 | -1.48 | 0.66 | 7 |
| Chitin-3 | -1.29 | 0.50 | 7 |
| Chitin-4 | -1.01 | 0.15 | 7 |
| Chitin-5 | -0.89 | 0.49 | 6 |

Table 5.4: Hydrogen bond analysis of the hydrogen bonds that formed during the adsorption of chitin to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface, all bond lengths are given in Å.

| **H-bond** | $\mathbf{O}_s\cdots\mathbf{H}_m$ | $\mathbf{O}_m\cdots\mathbf{H}_s$ | $\mathbf{O}_m\cdots\mathbf{H}_m$ |
|---|---|---|---|
| **Biopolymer** | | | |
| Chitin-1 | 1.65, 1.77, 1.89 | 1.79, 1.83, 2.01 | 1.79, 1.99, 2.22, 2.37 |
| Chitin-2 | 2.24, 2.39 | 2.49 | 1.76, 2.00, 2.45, 2.50 |
| Chitin-3 | 2.23, 2.34 | 1.80, 1.97, 2.03, 2.46 | |
| | | 2.47 | |
| Chitin-4 | 1.91 | 1.80, 2.02, 2.06, 2.20 | |
| | | 2.35, 2.45 | |
| Chitin-5 | 1.74, 2.48 | 1.80, 2.00, 2.27, 2.44 | |

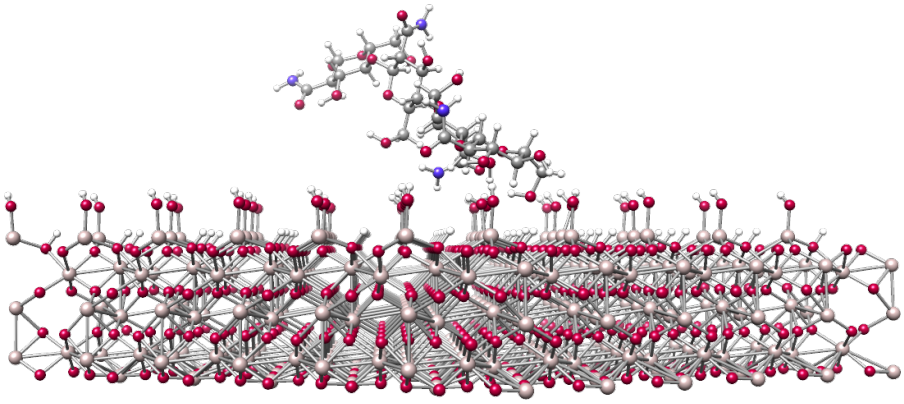


Figure 5.11: Adsorption geometry of chitin-1, during the adsorption of chitin to the hydroxylated $\alpha$-$Al_2O_3$ surface.

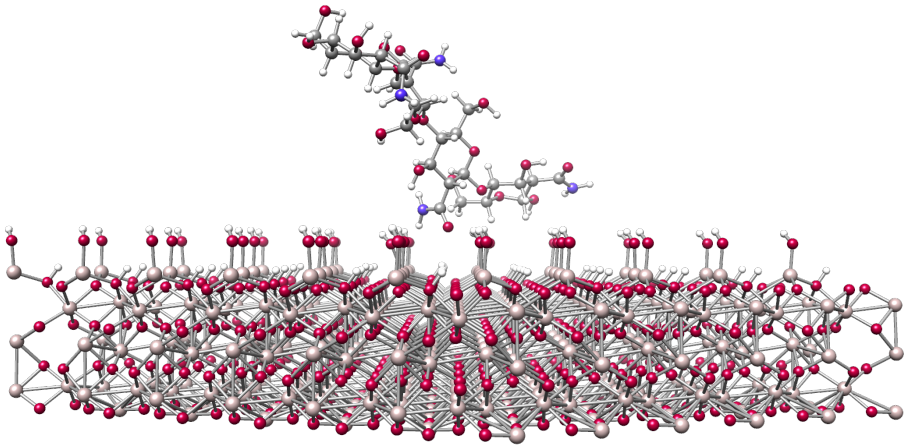


Figure 5.12: Adsorption geometry of chitin-2, during the adsorption of chitin to the hydroxylated $\alpha$-$Al_2O_3$ surface.

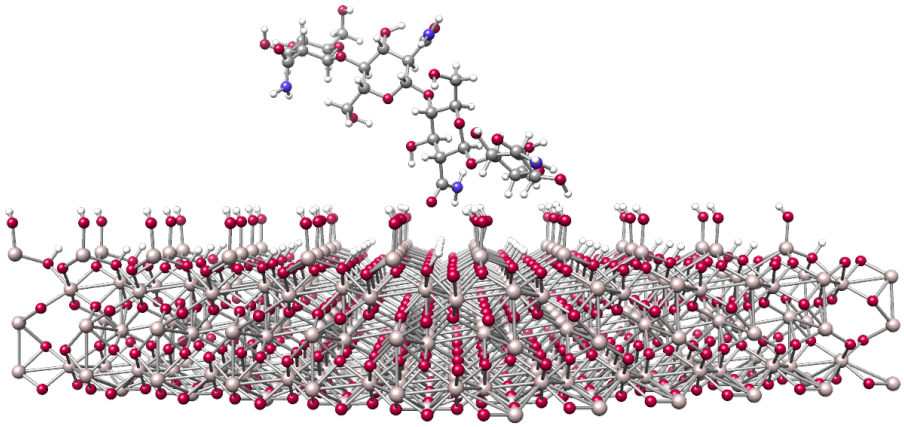


Figure 5.13: Adsorption geometry of chitin-3, during the adsorption of chitin to the hydroxylated $\alpha$-$Al_2O_3$ surface.

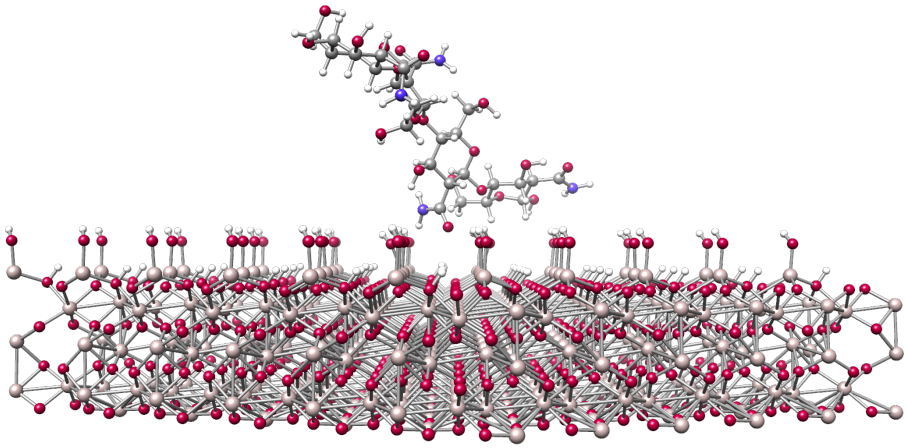


Figure 5.14: Adsorption geometry of chitin-4, during the adsorption of chitin to the hydroxylated $\alpha$-$Al_2O_3$ surface.

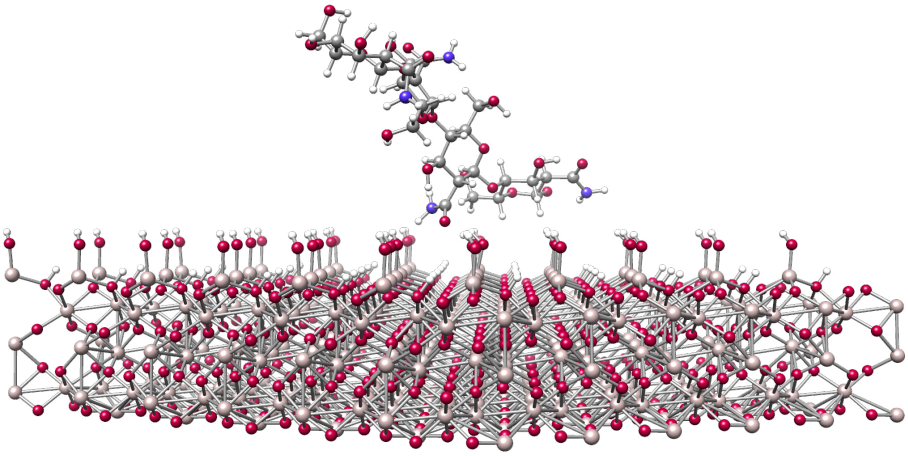


Figure 5.15: Adsorption geometry of chitin-5, during the adsorption of chitin to the hydroxylated $\alpha$-$Al_2O_3$ surface.

The most stable geometry of chitin (chitin-1) is illustrated in figure 5.11. With the structure of chitin in mind, one would expect the adsorption of chitin to be weaker than that of cellulose, as more of the cellulose interacts with the surface. This is indeed the case, as table 5.3 illustrates

that the most stable chitin adsorption geometry has a binding energy of -1.67 eV which is -1.05 eV less stable than cellulose-1. Chitin-1 was able to form 10 hydrogen bonds upon adsorption to the surface, with six O-H H-bonds between the surface and the adsorbate. Chitin-2 is illustrated in figure 5.12 and it is the second most strongly adsorbing chitin configuration. The weaker adsorption energy of chitin-2 can be explained by table 5.4 which highlights that, only three H-bonds formed between chitin and the surface during adsorption. The additional stability was therefore provided by the four internal hydrogen bonds within the chitin molecule itself.

Chitin-3 (figure 5.13) illustrates chitin adsorbing in a different geometry which is similar to chitin-1. However, this geometry is less stable than chitin-1, with a calculated binding energy of -1.29 eV. Despite seven H-bonds between the surface and adsorbate in chitin-3, there are no internal H-bonds within the molecule itself. Upon further analysis of the seven H-bonds, the bonds are longer than those in chitin-1. On the other hand, the bonds in chitin-3 are shorter than those of chitin-2, so we would expect it to be more stable. Nonetheless, both chitin-1 and chitin-2 have additional internal H-bonds, providing them with additional stability. Chitin-4 and chitin-5 (figures 5.14 and 5.15) have similar geometries. Both have one less H-bond than chitin-3 resulting in weaker adsorption to the surface. The calculated deformation energies range from 0.49 - 1.10 eV, suggesting that there was no significant deformation of chitin and the surface upon adsorption of chitin to the surface. This feature can be attributed to the anti-parallel structure of chitin.

### Chitosan

Chitosan, the deacetylation product of chitin was the next biopolymer to be adsorbed to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. Similar to chitin, chitosan is another animal-derived biopolymer found in the shells of crustaceans. The functional groups present in chitosan are alcohol and amine groups.

Table 5.5: Calculated binding energies, deformation energies and total number of hydrogen bonds formed between several chitosan and the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

| **Biopolymer** | **Binding Energy / eV** | **Deformation Energy / eV** | **No. of H-Bonds** |
|---|---|---|---|
| Chitosan-1 | -2.34 | 1.66 | 11 |
| Chitosan-2 | -2.15 | 1.40 | 9 |
| Chitosan-3 | -1.95 | 0.15 | 9 |
| Chitosan-4 | -1.53 | 0.66 | 9 |
| Chitosan-5 | -1.08 | 0.49 | 8 |

Table 5.6: Hydrogen bond analysis of the hydrogen bonds that formed during the adsorption of chitosan to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface, all bond lengths are given in Å.

| **H-bond** | $\mathbf{O}_s\cdots\mathbf{H}_m$ | $\mathbf{O}_m\cdots\mathbf{H}_s$ | $\mathbf{O}_m\cdots\mathbf{H}_m$ | $\mathbf{N}_m\cdots\mathbf{H}_s$ | $\mathbf{N}_m\cdots\mathbf{H}_m$ |
|---|---|---|---|---|---|
| **Biopolymer** | | | | | |
| Chitosan-1 | 1.95, 2.15 | 1.85, 1.86, 2.06 | 2.02, 2.08, 2.12 | | 1.93, 1.98 |
| | | 2.38 | | | |
| Chitosan-2 | 1.56, 1.84 | 1.89, 2.02, 2.04 | 1.85 | 1.62, 1.78 | 1.75 |
| Chitosan-3 | 2.18, 2.41 | 1.86, 2.12, 2.33 | 1.94, 2.10 | | 1.89, 1.94 |
| Chitosan-4 | 2.35, 2.38 | 2.22 | 1.69, 2.19, 2.47 | | 1.90, 2.24, 2.42 |
| Chitosan-5 | 2.05, 2.06 | 2.49 | 2.49 | | 1.97, 2.02, 2.03 |
| | | | | | 2.08 |

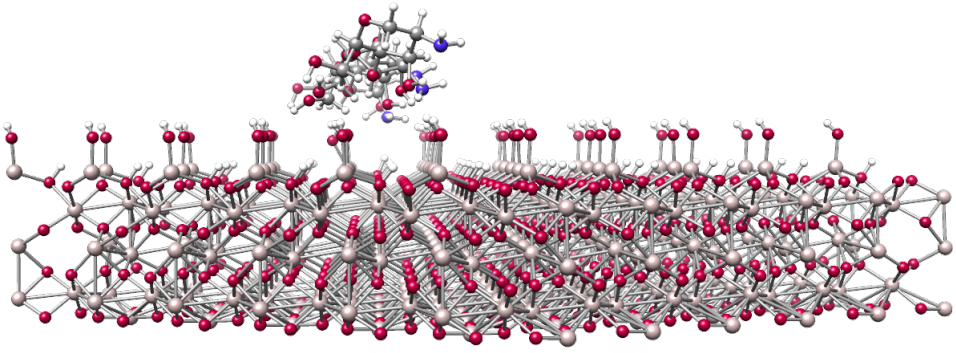


Figure 5.16: Adsorption geometry of chitosan-1, during the adsorption of chitosan to the hydroxylated $\alpha$-$Al_2O_3$ surface.

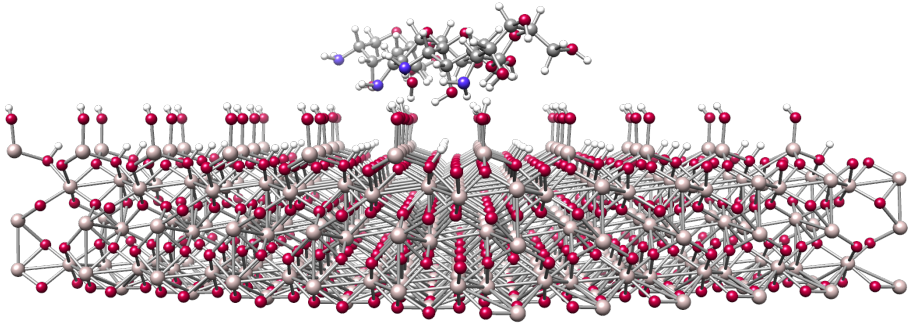


Figure 5.17: Adsorption geometry of chitosan-2, during the adsorption of chitosan to the hydroxylated $\alpha$-$Al_2O_3$ surface.

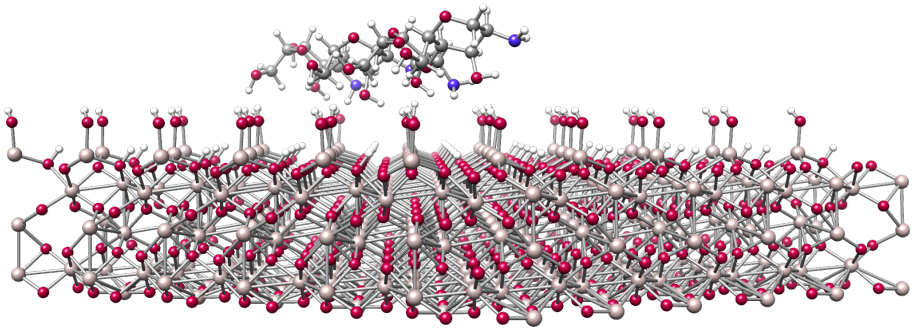


Figure 5.18: Adsorption geometry of chitosan-3, during the adsorption of chitosan to the hydroxylated $\alpha$-$Al_2O_3$ surface.

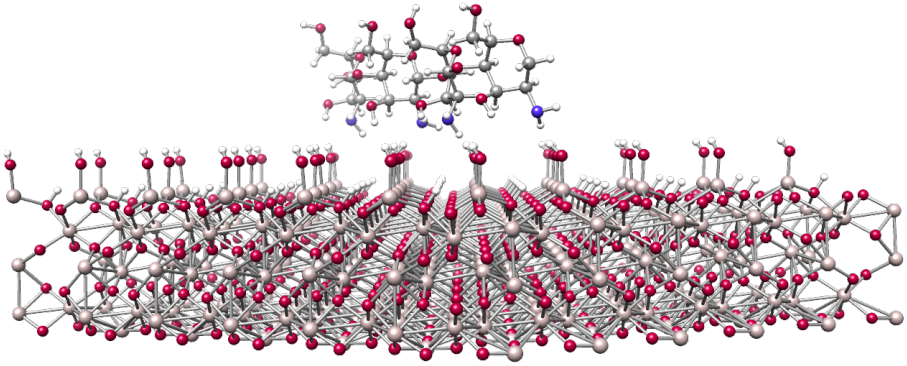


Figure 5.19: Adsorption geometry of chitosan-4, during the adsorption of chitosan to the hydroxylated $\alpha$-$Al_2O_3$ surface.

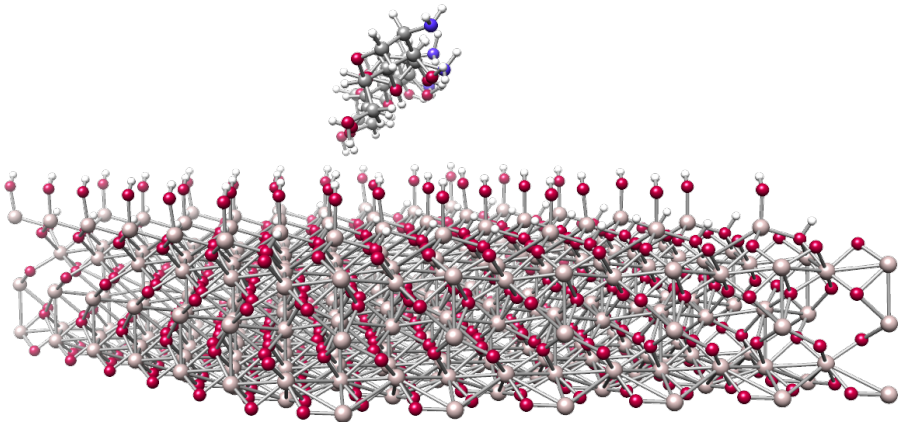


Figure 5.20: Adsorption geometry of chitosan-5, during the adsorption of chitosan to the hydroxylated $\alpha$-$Al_2O_3$ surface.

Several structures of chitosan were modelled from parallel structures to structures with chitosan on its edge. We will now discuss these structures further. Chitosan-1 is the most stable chitosan geometry (figure 5.16), with a binding energy of -2.34 eV, which is closer to the stability of cellulose-1. In this geometry chitosan is lying flat and parallel to the surface. This adsorption geometry is the most stable amongst the chitosan geometries as nine H-bonds formed between the chitosan adsorbate and the surface. Additional stability for this geometry was provided by two short and strong internal hydrogen bonds within chitosan. Chitosan-2 (figure 5.17) was the second most

stable chitosan configuration with a binding energy of -2.15 eV. Chitosan-2 is the only geometry amongst the nitrogen containing biopolymers with two N-H atom H-bonds. Overall, chitosan-2 has seven direct H-bonds between the surface and adsorbate. Hence, chitosan-1 is the most stable as it has more H-bonds between chitosan and the surface. The H-bonds of chitosan-1 are also between O and H atoms (unlike two N-H hydrogen bonds of chitosan-2). Oxygen is more electronegative than nitrogen, therefore the electronegativity difference is greater between O and H. Hence, the resultant H-bond is stronger. Much like chitosan-1 and 2, chitosan-3 (figure 5.18) also adsorbed parallel and flat to the surface. Chitosan was able to form nine H-bonds with the surface in chitosan-3, which is the same as chitosan-2. However, upon inspection of the H-bond lengths (table 5.6) the H-bonds in chitosan-3 are longer than those in chitosan-2. Therefore, chitosan is more strongly held by the surface in chitosan-2, resulting in a more stable adsorption geometry.

The final two chitosan geometries are illustrated in figures 5.19 and 5.20. Chitosan has adsorbed on its edge in these configurations. Chitosan-4 (figure 5.19) illustrates that chitosan has adsorbed with its amine groups pointing downwards towards the surface. Despite the amine groups pointing downwards, no N-H hydrogen bonds were detected between the hydrogen's of the surface and the nitrogen's of the chitosan adsorbate. Conversely, the structure of chitosan-5 illustrates that the chitosan molecules OH groups are pointing downwards towards the surface. At first glance, one would assume that chitosan-5 would have a stronger adsorption geometry due to the interaction of its OH groups with the surface OH groups. However, upon closer inspection and analysis of the H-bonds in table 5.6, it can be seen that chitosan-4 has six H-bonds between chitin and the surface and chitosan-5 only has four. Hence, chitosan-4 has the more stable adsorption geometry. The data for all chitosan geometries illustrates that chitosan prefers to bind to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface in a flat configuration similar to cellulose. Similar to chitin, the deformation energies range from 0.49 - 1.66 eV, suggesting that there was no significant deformation of chitin and the surface upon adsorption of chitin to the surface, with the chitosan-1 as the most deformed.

**Pectin**

Table 5.7: Calculated binding energies, deformation energies and total number of hydrogen bonds formed between pectin and the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface.

| Biopolymer | Binding Energy / eV | Deformation Energy / eV | No. of H-Bonds |
|---|---|---|---|
| Pectin-1 | -3.09 | -0.50 | 10 |
| Pectin-2 | -2.98 | 0.88 | 9 |
| Pectin-3 | -2.04 | 0.77 | 9 |
| Pectin-4 | -1.86 | 0.63 | 6 |
| Pectin-5 | -1.12 | 0.45 | 6 |

Table 5.8: Hydrogen bond analysis of the hydrogen bonds that formed during the adsorption of pectin to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface, all bond lengths are given in Å.

| **H-bond** | $\mathbf{O}_s\cdots\mathbf{H}_m$ | $\mathbf{O}_m\cdots\mathbf{H}_s$ | $\mathbf{O}_m\cdots\mathbf{H}_m$ |
|---|---|---|---|
| **Biopolymer** | | | |
| Pectin-1 | 1.66, 1.81 | 1.84, 1.97, 1.99, 2.11 | 1.73, 1.80, 2.20, 2.20 |
| Pectin-2 | 1.73, 1.96 | 1.81, 1.92 1.96, 2.29 | 1.70, 2.05, 2.44 |
| Pectin-3 | 1.66, 2.15 | 1.71, 1.80, 2.27 | 1.80, 2.15, 2.23, 2.36 |
| Pectin-4 | 1.73 | 1.94, 2.30, 2.37 | 2.28, 2.35, 2.37 |
| Pectin-5 | | 1.91, 1.99, 2.11 | 1.81, 2.17, 2.29 |

Lastly, we modelled the adsorption of pectin onto the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. Pectin contains the three different functional groups: carboxlyic acid, ester and alcohol functional groups.

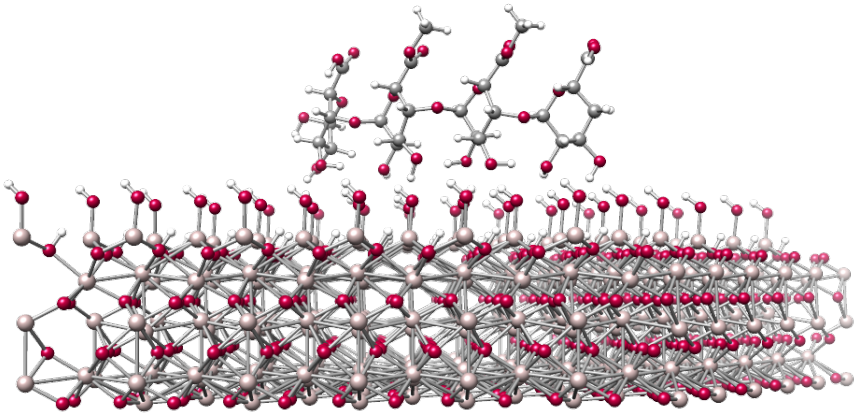


Figure 5.21: Adsorption geometry of pectin-1, during the adsorption of pectin to the hydroxylated $\alpha$-$Al_2O_3$ surface.

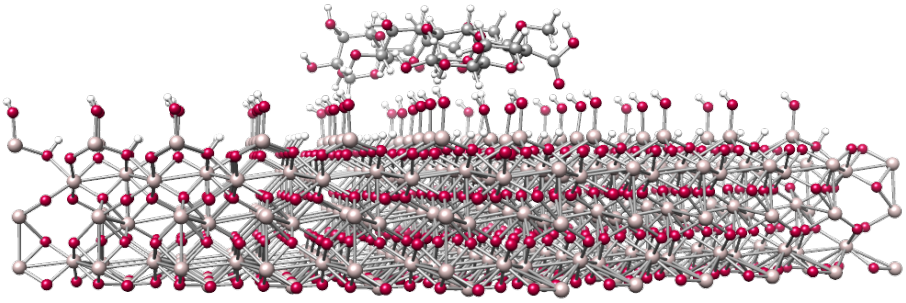


Figure 5.22: Adsorption geometry of pectin-2, during the adsorption of pectin to the hydroxylated $\alpha$-$Al_2O_3$ surface.

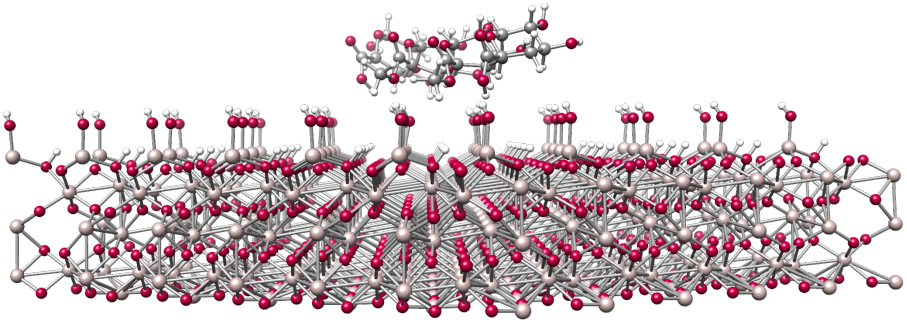


Figure 5.23: Adsorption geometry of pectin-3, during the adsorption of pectin to the hydroxylated $\alpha$-$Al_2O_3$ surface.

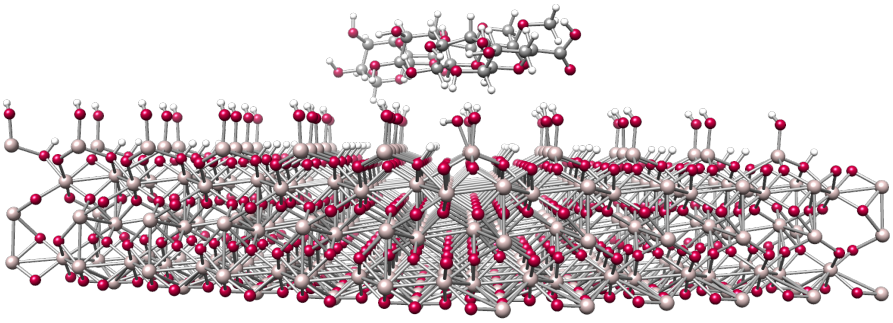


Figure 5.24: Adsorption geometry of pectin-4, during the adsorption of pectin to the hydroxylated $\alpha$-$Al_2O_3$ surface.

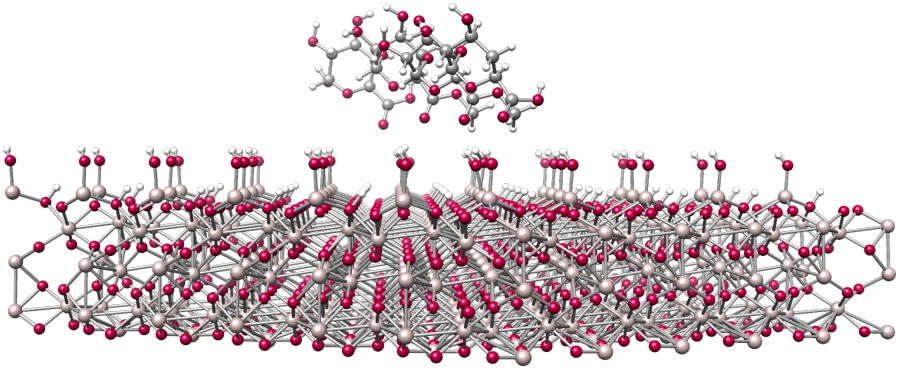


Figure 5.25: Adsorption geometry of pectin-5, during the adsorption of pectin to the hydroxylated $\alpha$-$Al_2O_3$ surface.

The most stable adsorption geometry obtained during our calculation was pectin-1 (figure 5.21), with a binding energy of -3.09 eV which is -0.37 eV more stable than cellulose-1 (second most stable configuration). Pectin-1 is the most stable geometry of all the biopolymers considered in this work. In this configuration a total of six H-bonds were detected between the pectin biopolymer and the surface. The additional stability was provided by the internal H-bonding within pectin itself. Figure 5.21 illustrates that pectin is in an upright position on its edge with its alcohol groups pointing downwards towards the surface. Interestingly, our previous investigations on small molecule adsorption to the bare and hydroxylated $\alpha$-$Al_2O_3$ (0001) surface determined that acids were strong surface adsorbates. Therefore, it is interesting that the acid groups have not directly

contributed to binding through their OH oxygen atom or their carbonyl oxygen atom. However, the ester (groups in between the acids) contain a bulky methyl constituent, therefore it is not likely that a configuration whereby the acid functional groups point downwards towards the surface would be favoured. We modelled this configuration whereby the acid and ester substituents of pectin were pointing downwards towards the surface to test this hypothesis. Upon optimisation the geometry did not remain in the configuration. This is because of the resultant steric clash that occurs between the surface and pectin's methyl constituents. However, during adsorption of pectin-2, the acid functional groups were able to contribute to binding. Figure 5.22 illustrates the geometry of pectin-2, the second most stable pectin geometry, with a binding energy of -2.98 eV. Pectin-2 formed six direct H-bonds with the surface, of those six, four of the H-bonds occurred between carbonyl oxygen atoms of pectin and surface hydrogen atoms. pectin-2 has one less internal H-bond than pectin-1, resulting in a more stable geometry for pectin-1. The carboxylic acid and ester functional groups in pectin-1 are also providing an inductive effect towards the OH groups that are involved with binding, hence the resultant H-bonds are stronger.

Pectin-3 (figure 5.23) was the geometry that was obtained with an intermediate binding strength (-2.04 eV). Five direct H-bonds formed between pectin and the surface in pectin-3. Hence, there are fewer H-bonds that bind directly to the surface in pectin-3 than pectin 1 and 2. Therefore, the resultant pectin-3 geometry is less stable than its predecessors. Pectin-4 and pectin-5 (figures 5.24 and 5.25) were the least stable pectin adsorption configurations with binding energies of -1.86 eV and -1.12 eV. Fewer direct hydrogen bonds were formed upon the adsorption of pectin to the surface in these two geometries (4 and 3). Therefore, the pectin molecules are less tightly held in these configurations and have lower calculated binding energies. The deformation energies of pectin 2-5 illustrate that little deformation occurred during the adsorption of pectin to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. However, pectin-1 has an unusual deformation energy of -0.50 eV, suggesting that a deeper minimum of the adsorbate (pectin) configuration was obtained upon optimisation.

**Comparing the Binding Energies of Biopolymers**

Here, we will compare the results obtained from our investigations on the adsorption of biopolymers to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. The binding hierarchy for biopolymers is as follows (most to least stable),

$$\text{Pectin} > \text{Cellulose} > \text{Chitosan} > \text{Chitin} \tag{5.5}$$

Pectin was the most strongly adsorbing biopolymer, with the largest most negative binding energy (-3.09 eV). Cellulose is most similar in structure to pectin. The key difference between the two is the addition of carboxylic acid and ester functional groups along one side of the pectin molecule. Therefore, one may expect their most stable geometries to be similar. However, in actuality the most stable adsorption geometry for pectin illustrated that pectin adopts an on-edge position with its acid and ester functional groups pointing upwards and its alcohol groups pointing down towards the surface. On the other hand, the most favoured geometry for cellulose, was that which resulted in cellulose adsorbing flat and parallel to the surface. This is in fact similar to chitosan which also favoured a flat adsorption configuration where it adsorbed parallel to the surface. The flat and parallel configuration for cellulose and chitosan allows for maximum interaction between the surface and biopolymer. Upon the adsorption of pectin in a flat and parallel configuration (pectin-2), binding was not as strong as on-edge adsorption. Pectin also has methyl constituents resulting in steric clash with the surface. Therefore, despite the acid groups (in pectin-1) not directly resulting in binding, their presence still increased the stability of the overall geometry. The carboxyl groups in the acid and ester groups also have an inductive effect. The calculations have illustrated (particularly for cellulose, chitosan and pectin), there are two favourable binding modes of biopolymer adsorption to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. These two modes are flat and parallel adsorption and on-edge adsorption.

Chitin on the other hand adsorbed in a different manner to its biopolymer counterparts. As chitin is a non-linear twisted molecule, only half of the molecule is able to adsorb to the surface at any given time. Hence, there is less interaction overall between chitin and the surface. Therefore, chitin is the weakest adsorption biopolymer despite containing strongly adsorbing amide groups.

The alcohol functional group was the common site for binding in all of our biopolymer investigations. A H-bond either formed between the $O_s\cdots H_m$ or $O_m\cdots Hs$ (OH groups on the surface and biopolymer) for most of the biopolymers. Pectin had some interactions to the surface via its carboxylic acid groups, for example, in pectin-2, whereby the C=O oxygen atom bound to the

surface OH hydrogen atoms. Both chitin and chitosan contain nitrogen based functional groups (amide and amine). Thus, the nitrogen atoms had a possibility to form hydrogen bonds with the surface. However, no N-H hydrogen bonds were detected for chitin. Chitosan was able to form some N-H hydrogen bonds (chitosan-2). This suggests that the amine group would be favoured over the amide group during adsorption, as the amine group was directly involved in the binding of chitosan to the surface. Thus far, we have determined that the alcohol amine and acid groups all contribute to the binding of biopolymers to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. Pectin is the most favoured adsorption biopolymer as it contains the acid group providing additional stability to the overall geometry.

We have summarised the results from our biopolymer investigations on a graph (figure 5.5) to illustrate the strongest adsorbing biopolymers with respect to the total number of hydrogen bonds. We plotted the binding energy against the number of hydrogen bonds. The graph illustrates a correlation. The relationship illustrated by the graph suggests that as the number of hydrogen bonds increases the binding energy becomes larger (more negative). Several stable adsorption geometries have been found for all of the biopolymers investigated and amongst all geometries, the flat and parallel and on-edge geometries were favoured.

## 5.4 Conclusion

Our investigations into the adsorption of biopolymers to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface aimed to obtain a realistic picture of OC binding to soil mineral surfaces. We also aimed to identify strongly adsorbing biopolymers. All biopolymers adsorbed to the surface via hydrogen bonding. As the number of hydrogen bonds formed between the surface and biopolymers increased, as did the strength of binding. The flat and parallel and on-edge configurations produced the most stable structures. Pectin, cellulose and chitosan were the stronger adsorbing biopolymers and chitin the weakest adsorbate. We find that the presence of carboxylic acid, hydroxyl and amine groups provide strong binding of biopolymers, while amide groups result in a twisted structure that cannot adsorb effectively. Overall, pectin adsorbed most strongly above its biopolymer counterparts. Therefore, pectin is a prime candidate for *'carbon based'* fertilisers as it is a naturally abundant biopolymer that can be added to soil that will remain strongly bound to soil minerals. Cellulose also emerges as a good candidate, which provides strong binding and it is also the most abundant naturally occurring biopolymer.

# Chapter 6

# Metal Ions

Hydrated Metal (III) species, particularly Al(III) and Fe(III) species are important and of great interest in nature. For example, in clays and hydroxide phases.[16,275] They are also considered to be the reactive species involved in sorption processes within soils.[83,84,91,276–278] The Fe(III) species has been subject to a vast array of classical computer simulations (molecular dynamics). The literature has greatly emphasised the solvation shell structure of $Fe^{3+}$,[279–285] ion and water diffusion,[285] hydrolysis,[286] charge transfer[287–291] and solvent isotope effects.[283,291]

Similarly, the Al(III) species has also undergone a vast array of scientific investigations. Investigations have found that the $Al^{3+}$ species has a rich variety of solute structures[16,292–294] and has a high toxicity to plants and humans.[16,295] Two thirds of the soils (acidic) in the world face potential Al toxicity problems with respect to the growth of crops. In soil chemistry terms, the low pH of acidic soils is accompanied by a proton attack on Al-bearing minerals. This leads to the production of soluble Al in the soil and adsorbed Al on particle surfaces. When the concentration of free $Al^{3+}$ ions is in the macro-molecular range, toxicity to plant roots occurs, therefore inhibiting plant growth.[1,296,297] Aluminium solubility in soils is influenced by a variety of minerals such as gibbiste, kaolinite and beidellite to name a few. Dissolution reactions for these minerals become essential, such as the following dissolution reaction for gibbsite below[1]

$$Al(OH)_3(s) + 3H^+ \rightarrow Al^{3+} + 3H_2O(l) \tag{6.1}$$

The $Al^{3+}$ and $Fe^{3+}$ metal cations are similar in charge and size (54pm and 65pm) for $Al^{3+}$ and high-spin $Fe^{3+}$.[298,299] It has been determined that $Al^{3+}$ and $Fe^{3+}$ cations have a six coordinate, octahedral first hydration shell[8,298–302]

## 6.1 Metal(III) Ion-Water Interactions

The nature of metal(III) cation binding to water is of paramount interest. Natural soil environments contain water, therefore one can expect some interaction between metal cations and water molecules within soils. The polarising ability of metal ions is often quite strong, such that a proton from a water molecule in the primary hydration shell is repelled by the metal ion in aqueous complexes.[8,303–305] Under this condition the aqueous metal ion acts as an acid (proton donor). Therefore, the strength of this proton donation is measured using the pKa. Chang et al.[8] measured the pKa values of the first hydrolysis step of several metal 2+ and 3+ cations. They also calculated binding energies using density functional methods,[306–308] particularly molecular dynamics simulations.

Table 6.1: Binding energies[8] of hexa-aqua metal ions with VWN, PB and BLYP functionals in eV

| **Metal ion** | **VWN** | **BP** | **BLYP** |
|---|---|---|---|
| $Al^{3+}$ | -1.11 | -1.12 | -1.12 |
| $Fe^{3+}$ | -2.73 | -2.74 | -2.74 |

The binding energies of $Al^{3+}$ and $Fe^{3+}$ metal ions are of particular interest to us. These literature values are of importance as they are comparing the adsorption energies of different functionals. Chang et al.[8] found that aqueous metal cations with inert electron configurations are more stable than aqueous cations with partially filled non-inert electron configurations due to the symmetry of electron alignment and the exchange energy.

$$\text{aqua} - Mg^{2+}, Ca^{2+} > \text{aqua} - Mn^{2+}, Al^{3+}, Sc^{3+} > \text{aqua} - Cr^{3+}, Fe^{3+} \tag{6.2}$$

They also found that aqueous cations with a larger ionic radius are more stable than aqueous cations with a smaller one.

$$\text{aqua} - Ca^{2+} > \text{aqua} - Mg^{2+} > \text{aqua} - Sc^{3+} > \text{aqua} - Al^{3+} \tag{6.3}$$

### 6.1.1 Mechanism of Metal-Water Binding

As the nature of metal (III) cation binding to water is of paramount interest, one should therefore understand the mechanism by which the binding of water occurs. The binding process to solvated (III) cations is essentially a ligand substitution process. Many biological and catalytic reactions are

controlled by the lability of coordinated solvent molecules. Hence, a lot of effort and work has been devoted to the elucidation of the reaction mechanisms. This is not only to gain an understanding of the reaction mechanisms themselves, but to also develop knowledge that will enable modification and systematic control of such processes. Theoretical calculations[309–311] alongside sophisticated high-pressure kinetic techniques such as NMR line broadening[312,313] have introduced the nature of the transition state involved in the substitution process.[16,314]

One important aspect to consider is the influence of the pH on water exchange reactions. A multitude of studies[16,315–318] have been devoted to investigating the influence of pH and it is now well documented that the deprotonation of hydrated metal ions such as $[Fe(H_2O)_6]^{3+}$ to form the corresponding mono-hydroxyo $[Fe(H_2O)_5OH]^{2+}$ complex leads to a drastic change in the water exchange rate constant. It also leads to a change in the underlying reaction mechanism from an associative (A) reaction mechanism to a dissociative (D) mechanism. The D mechanism is induced by the trans effect of the coordinated hydroxide ion.[16,319] Hanauer et al.[16] investigated solvent exchange reactions on the $[Al(H_2O)_6]^{3+}$ and $[Al(H_2O)_5OH]^{2+}$ complexes. For the former complex, their calculations determined that the $[Al(H_2O)_6]^{3+}$ complex exchanges water via a dissociative interchange ($I_d$) mechanism.[315,320] Deprotonation of $[Al(H_2O)_6]^{3+}$ has been reported to increase the water exchange rate.[321]

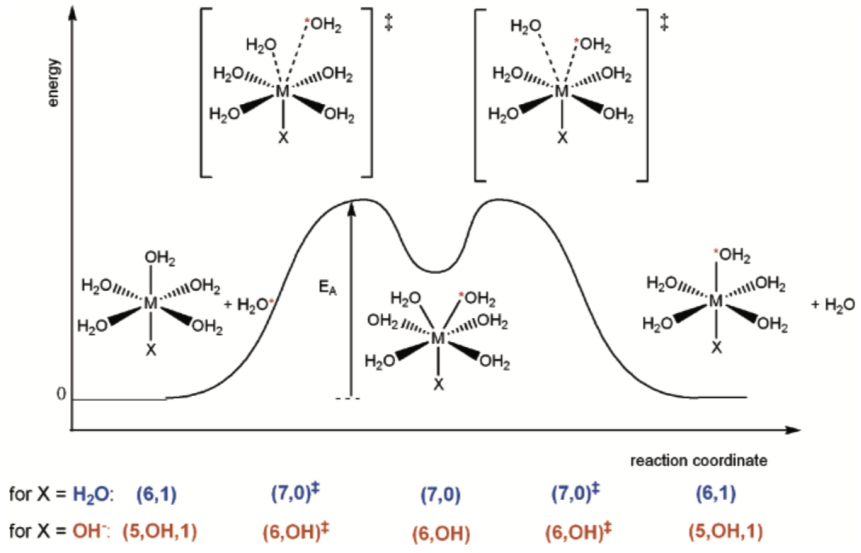



Figure 6.1: Schematic energy profile and structures for the associative (A) water exchange mechanism.[16]

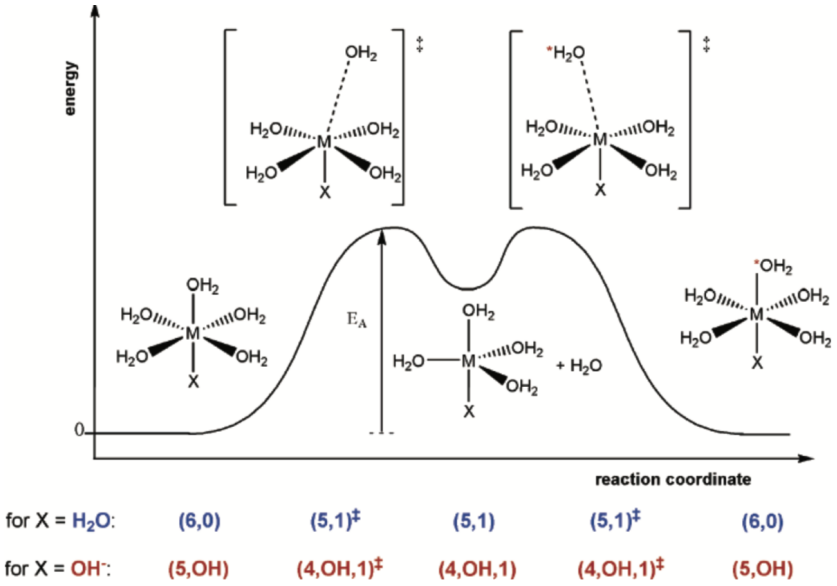



Figure 6.2: Schematic energy profile and structures for the dissociative (D) water exchange mechanism.[16]

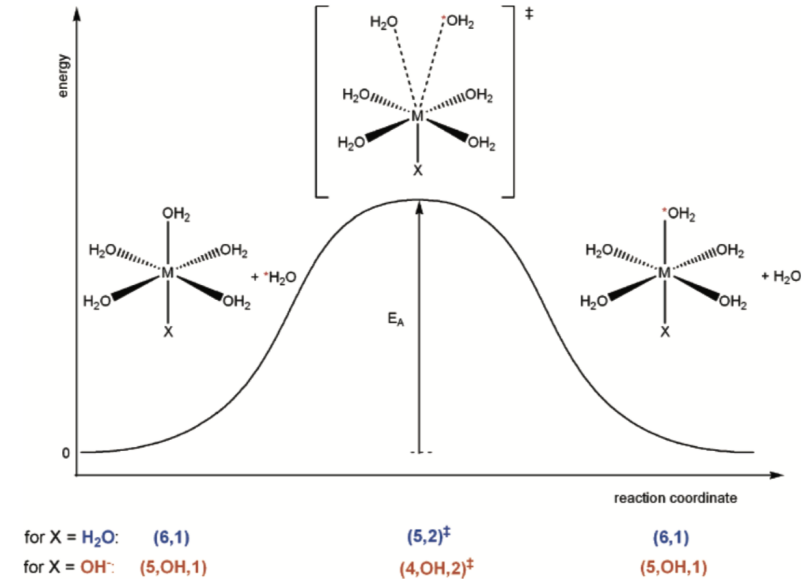



Figure 6.3: Schematic energy profile and structures for the Interchange ($I_d$) water exchange mechanism.[16]

Figures 6.1, 6.2 and 6.3 illustrate the associative (A), dissociative (D) and Interchange ($I_d$) water exchange mechanisms.[16] The A water exchange mechanism was investigated between the hexa-aqua Al(III) complex and a seventh water molecule and the penta-aqua Al(III) complex and a sixth water molecule. During the A mechanism the 7-fold transition state of the A mechanism yielded a 6-fold complex. The D mechanism was favoured for the gas phase. In this mechanism two transition states were found and they are illustrated in figure 6.2. Lastly, in the $I_d$ mechanism a water molecule was exchanged for another water molecule. Figure 6.3 illustrates this, via a transition state with weak coordination of two exchanging water molecules.

## 6.2 Interaction of Aquated Metal(III) Complexes with organic molecules

Now that we have discussed the nature of binding between metal(III) ions and water, we will discuss work in the literature on the interaction of organic compounds with the hydrated $Al^{3+}$ and $Fe^{3+}$ complexes. $M^{3+}$-organic reactions are significant, as they are used in water purification systems to drive coagulation and precipitation reactions between dissolved organic matter and minerals.[322,323] Complexation of $Al^{3+}$ and other metals such as $Fe^{3+}$ by organic matter has been shown to dramatically reduce the toxicity of metals.[296,322,324] Literature studies on the binding between carboxylic acids and the hexa-aqua $[Al(H_2O)_6]^{3+}$ complex have been conducted. Carboxylic acids are thought to represent metal-complexing functional groups present within the organic matter in soils.[322,325] Thus, allowing the complex nature of $Al^{3+}$-organic matter to be discovered. In addition, complexation between simple carboxylic acids can increase the rates of solubility and dissolution of $Al^{3+}$ bearing minerals.[322,326–328] Hence, the presence of carboxylic acids can affect the transport of other metals and $Al^{3+}$ within soils, groundwater and sediments.[322,329,330]

Kubicki et al.[322] calculated the reaction energetics ($\Delta E$) of $Al^{3+}$-carboxylic acid complex formation in aqueous solution using DFT methods. They calculated complexation energies of several $Al^{3+}$-carboxylic acid complexes to determine which resulted in a more energetically stable complex. The first carboxlyic acids that underwent complexation were the deprotonated acetate and lactate carboxylate anions. The research found that when the acetate anion replaced a water molecule the reaction resulted in a negative $\Delta E$, but when the acetate anion replaced an $OH^-$, $\Delta E$ was positive.

$$[Al(H_2O)_6]^{3+} + Ac^- \bullet 8H_2O \rightarrow [Al(H_2O)_5(Ac)]^{2+} + H_2O \bullet 8H_2O \tag{6.4}$$

where $Ac^-$ represents the acetate anion and $\Delta E$ = -168 kJ $mol^{-1}$.

$$[Al(H_2O)_5(OH)]^{2+} + Ac^- \bullet 8H_2O \rightarrow [Al(H_2O)_5(Ac)]^{2+} + OH^- \bullet 8H_2O \tag{6.5}$$

where $\Delta E = +13$ kJ $mol^{-1}$.

The second conclusion that was drawn from their investigations was that the formation of bidentate species was less energetically favourable than the formation of a monodentate species in regards to the acetate anion. For the lactate anion bidentate formation was more energetically favourable than acetate.

$$[Al(H_2O)_6]^{3+} + Ac^- \bullet 8H_2O \rightarrow [Al(H_2O)_4(Ac)]^{2+} + 10H_2O \quad (6.6)$$

where $\Delta E$ = -119 kJ $mol^{-1}$.

$$[Al(H_2O)_6]^{3+} + Lact^- \bullet 4H_2O \rightarrow [Al(H_2O)_4(Lact)]^{2+} + 6H_2O$$

where $Lact^-$ is the lactate anion and $\Delta E$ = -145 kJ $mol^{-1}$. They also calculated the $\Delta E$ values for the complexation of the oxalate, malonate and malate anions. When replacing a water molecule, very large negative $\Delta E$ values were obtained (-177, -123 and -88 kJ $mol^{-1}$) for the formation of the substitution products. Again, positive $\Delta E$ values were obtained when replacing an $OH^-$ for all anions. Although monodentate speciation is exothermic, bidentate replacement of water molecules by oxalate and malate is more strongly favoured.

The aim of our investigations is to understand the nature of binding between $Al^{3+}$ and $Fe^{3+}$ metal cations and organic ligands within soils. Firstly, we will model the $[Al(H_2O)_6]^{3+}$ and $[Fe(H_2O)_6]^{3+}$ complexes and compare them to literature studies. These initial investigations will enable us to gain an understanding of the interaction between $Al^{3+}$ and $Fe^{3+}$ cations and water ligands. Next, we will move on to investigating the reactions of the hexa-aqua $[Al(H_2O)_6]^{3+}$ and $[Fe(H_2O)_6]^{3+}$ complexes with small organic ligands. Thus far in the literature, carboxylic acids have undergone the ligand substitution reaction with the $[Al(H_2O)_6]^{3+}$ complex. Therefore, we would like to extend our research by investigating the substitution reaction between carboxylic acids and the $[Fe(H_2O)_6]^{3+}$ complex, due the likeness of the $Al^{3+}$ and $Fe^{3+}$ metal cations. We will also subject several organic ligands to the ligand substitution reaction. Investigating several organic functional groups will yield a greater understanding of the functional groups that bind strongly to the $Al^{3+}$ and $Fe^{3+}$ metal cations, enabling us to determine which of the substitution reaction products are most stable (energetically favourable). Finally, we will compare the results of the binding between the organics and the $[Al(H_2O)_6]^{3+}$ and $[Fe(H_2O)_6]^{3+}$ complexes to determine which of the metal cations produces more stable products.

## 6.3 Computational Method

### 6.3.1 Metal Ions

DFT calculations of metal ions and metal-ligand complexes were performed using the Gaussian 09 software package.[331] All calculations on the hexa-aqua $[Al(H_2O)_6]^{3+}$ and $[Fe(H_2O)_6]^{3+}$ and their corresponding metal-ligand complexes were performed using the B3LYP[144] functional at the 6-31G++(d,p)[144,145,332] level of theory. This level of theory was chosen due to the B3LYP functionals reliability and extensive use in the literature.[160,333,334] The PBE[335] functional was also used in our calculations due to its heightened popularity as a GGA functional,[336] with Grimme's D3[149] empirical dispersion correction. The PBE-D3 functional was chosen for comparison to our results from other investigations, again at the 6-31G++(d,p)[144,145,332] level of theory. The bulk aqueous solution environment was stimulated by the polarizable continuum model (PCM model). The model treats the bulk solvent molecules as a homogeneous isotropic dielectric medium.[337] With water as the bulk solvent the Gibbs free energy of reaction, $\Delta_r G$ was calculated for all optimised complexes at a temperature of 298K. All structures were visualised with the UCSF Chimera[205] software package.

## 6.4 Results and Discussion

### 6.4.1 Hexa-aqua Metal(III) Complex

#### The $[Al(H_2O)_6]^{3+}$ Complex

In order to mimic the presence of metal ions, specifically $Al^{3+}$ ions, within soils, the $[Al(H_2O)_6]^{3+}$ complex was modelled. $Al^{3+}$ ions are commonly surrounded by water ligands in soils, due to the presence of water in natural soil environments. Hence, the octahedral $[Al(H_2O)_6]^{3+}$ was studied. Initially, we modelled the $[Al(H_2O)_6]^{3+}$ complex to obtain the total energy ($E_e$), these values were then compared to the literature.

Table 6.2: Total energy for the $[Al(H_2O)_6]^{3+}$ complex given in Hartree a.u.

| | **Method** | **Basis Set** | **$E_e$ / a.u.** |
|---|---|---|---|
| **This work** | PBE-D3 | 6-31++G(d,p) | -700.010 |
| | PBE | 6-31G++(d,p) | -700.007 |
| | B3LYP | 6-31G++(d,p) | -700.652 |
| | B3LYP | 6-311G++(d,p) | -700.812 |
| | B3LYP | 6-31G+(d,p) | -700.649 |
| | B3LYP | 6-311G+(d,p) | -700.787 |
| **Literature** | B3LYP/MP2 | SVP +SP | -697.97[338] |
| | MP2/VE | cc-PVDZ | -698.53[339] |
| | MP2 | cc-PvTZ | -699.04[339] |

Next, we modelled the formation of the hexa-aqua $[Al(H_2O)_6]^{3+}$ complex from the reaction between the penta-aqua $[Al(H_2O)_5]^{3+}$ complex and a water molecule to determine the stability of the hexa-aqua $[Al(H_2O)_6]^{3+}$ complex.

$$[Al(H_2O)_5]^{3+} + H_2O \rightarrow [Al(H_2O)_6]^{3+} \tag{6.7}$$

The Gibbs reaction energy, $\Delta_r G$, was calculated using Hess's Law

$$\Delta_r G = \sum \Delta G(Products) - \sum \Delta G(reactants) \tag{6.8}$$

where $\Delta_r G$ is the Gibbs reaction energy and $\Delta G(Products)$ and $\Delta G(Reactants)$ are the Gibbs formation energies of the products and the reactants.

Bench-marking investigations were carried out to investigate which DFT functional and basis set would best model the $[Al(H_2O)_6]^{3+}$ complex in accordance with the literature and computational expense.

Table 6.3: $\Delta_r G$ values for the formation of the $[Al(H_2O)_6]^{3+}$ complex.

| | **Method** | **Basis Set** | $\Delta_r G$ / **eV** |
|---|---|---|---|
| **This work** | PBE-D3 | 6-31++G(d,p) | -1.17 |
| | PBE | 6-31G++(d,p) | -1.00 |
| | B3LYP | 6-31G++(d,p) | -0.95 |
| | B3LYP | 6-311G++(d,p) | -0.94 |
| | B3LYP | 6-31G+(d,p) | -0.92 |
| | B3LYP | 6-311G+(d,p) | -0.24 |
| **Literature** | B3LYP | 6-311G++(d,p) | -0.58[338] |
| | B3LYP | 6-311G+(d,p) | -0.59[16] |
| | MP2 | 6-31G+(d,p) | -0.67[340] |
| | HF | cc-vPVDZ | -0.63[339] |

Table 6.2. illustrates that all methods yield fairly similar values for the $[Al(H_2O)_6]^{3+}$ raw Hartree energy. Table 6.2 also illustrates the use of more precise basis sets and functionals, such as the MP2 methods and triple zeta basis sets. As the results are within agreement of each other, we will not be employing the use of the more precise methods, as they are more computationally expensive. Table 6.3 illustrates the stability of the $[Al(H_2O)_6]^{3+}$ complex which was calculated from the reaction between the $[Al(H_2O)_5]^{3+}$ complex and single water molecule. The PBE functional was chosen, our research following the metal ion investigations will be conducted with the PBE functional as accordance with the literature. Between the PBE and PBE-D3 methods there is a small discrepancy of -0.17 eV, Hence our research for will be carried out at the PBE-D3 6-31G++(d,p) level of theory as the added dispersion makes the method more accurate than PBE alone. We have also chosen to employ the use of the B3LYP functional in our research, as the method is popular amongst the literature. Again, we will use the 6-31G++(d,p) basis set in keeping with the basis set we used for the PBE functional. The optimised geometry of the $[Al(H_2O)_5]^{3+}$ complex optimised under the PBE-D3 functional is illustrated in figure 6.4. The B3LYP and PBE-D3 geometries are similar, hence we will only illustrate the PBE-D3 adsorption geometries for all adsorbates.

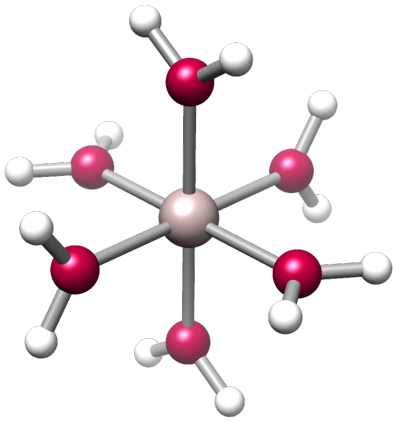


Figure 6.4: Optimised adsorption geometry of the hexa-aqua $[Al(H_2O)_6]^{3+}$ complex optimised under the PBE-D3 functional

### The $[Fe(H_2O)_6]^{3+}$ Complex

The $[Fe(H_2O)_6]^{3+}$ complex was also investigated. Similar to $[Al(H_2O)_6]^{3+}$ studies the optimised $[Fe(H_2O)_6]^{3+}$ complex was also investigated. The binding of the $Fe^{3+}$ ion to six water ligands yields the octahedral hexa-aqua metal complex, for which the Gibbs reaction energies, $\Delta_r G$, are given below.

Table 6.4: $\Delta_r G$ values for the $[Fe(H_2O)_6]^{3+}$ complex in Hartree a.u.

| | **Method** | **Basis Set** | $\Delta_r G$ **/ a.u.** |
|---|---|---|---|
| **This work** | PBE-D3 | 6-31G++(d,p) | -1720.937 |
| | B3LYP | 6-31G++(d,p) | -1720.669 |
| **Literature** | BLYP | - | -1721.23[8] |
| | BP | - | -1721.33[8] |
| | VWN | - | -1715.57[8] |

The calculated Gibbs reaction energies for the formation of the $[Fe(H_2O)_6]^{3+}$ complex are -0.76 eV (PBE-D3 6-31G++(d,p) and 12.101 eV (B3LYP 6-31G++(d,p)). The Gibbs energy values in table 6.4. illustrate that the methods we are employing in our investigation yield energy values very similar to those cited in the literature. Hence, we will model all calculations with the 6-31G++(d,p) basis set and the PBE-D3 and B3LYP functionals. We have illustrated the $[Fe(H_2O)_6]^{3+}$ complex in figure 6.5.

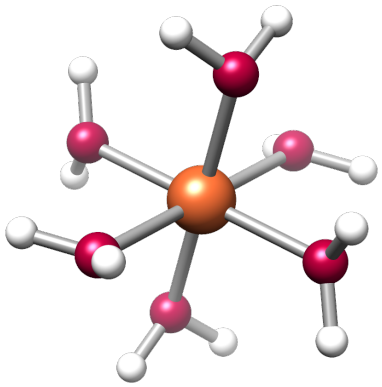


Figure 6.5: Optimised adsorption geometry of the hexa-aqua $[Fe(H_2O)_6]^{3+}$ complex optimised under the PBE-D3 functional

### 6.4.2 Organic Ligand Binding to the $[Al(H_2O)_6]^{3+}$ Complex

The binding of several organic ligands was investigated. Organic ligands were bound to evaluate the strength of binding of various different functional groups to the $[Al(H_2O)_6]^{3+}$ complex. The mechanism for the reaction is essentially a ligand substitution reaction. For the ligand substitution to be favourable (i.e. for the $[Al(H_2O)_6]^{3+}$ complex to lose a water ligand) the small organic ligand must bind more strongly than the water ligand. Thus, resulting in a water molecule being lost in the reaction. The complexes were constructed by systematically replacing a water molecule from the $[M(H_2O)_6]^{3+}$ complex with an organic monodentate ligand. The equation below represents the general case for a ligand binding to a hexa-aqua metal 3+ complex, $[M(H_2O)_6]^{3+}$.

$$[M(H_2O)_6]^{3+} + OL \rightarrow [M(H_2O)_5(OL)]^{3+} + H_2O \tag{6.9}$$

M is a metal with a charge of 3+ and OL is a neutral monodentate organic ligand.

Using equation 6.9 the $\Delta_r G$ values for the PBE-D3 and B3LYP systems were calculated as follows using methanol (with PBE-D3 data) as an example,

$$[Al(H_2O)_6]^{3+} + CH_3OH \rightarrow [Al(H_2O)_5(CH_3OH)]^{3+} + H_2O \tag{6.10}$$

$$\Delta_r G = \sum G([Al(H_2O)_5(CH_3OH)]^{3+} + H_2O)_{products} - \sum G([Al(H_2O)_6]^{3+} + CH_3OH)_{reactants} \tag{6.11}$$

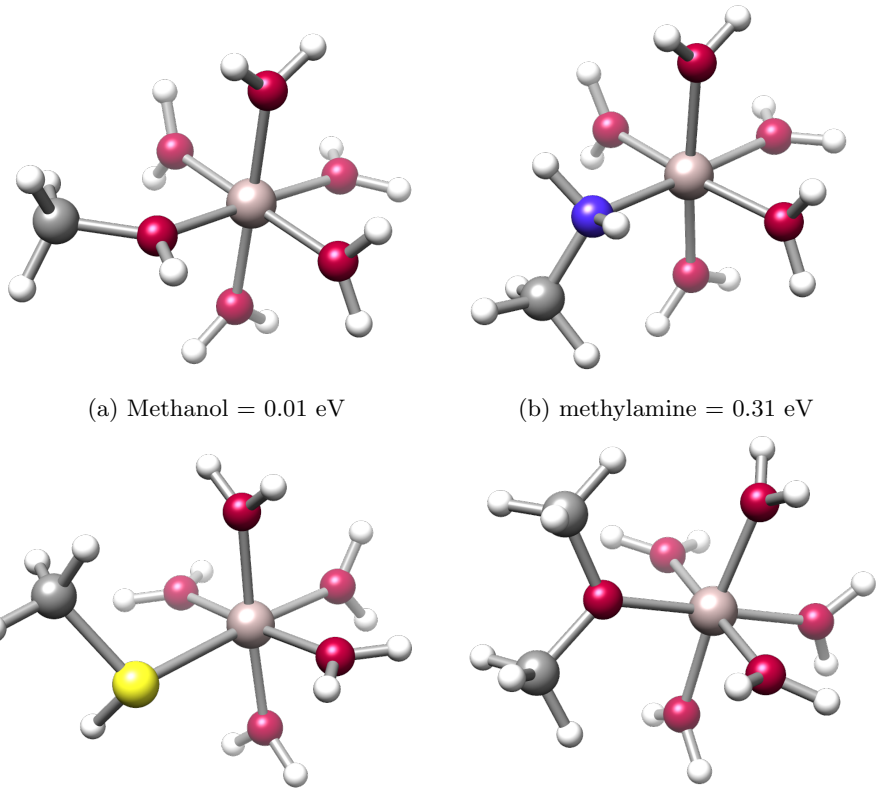
$$\Delta_r G = (-739.22 + -76.35)\ a.u. - (-700.01 + -115.57)\ a.u. = 2.65 \times 10^{-4}\ a.u. \equiv 0.01\ eV \quad (6.12)$$

Table 6.5: $\Delta_r G$ values the ligand substitution reaction between the $[Al(H_2O)_6]^{3+}$ complex and organic ligands (monodentate binding) for the B3LYP and PBE-D3 functionals at the 6-31G++(d,p) level of theory.

| | **B3LYP** | **PBE-D3** |
|---|---|---|
| **Organic Ligand** | $\Delta_r G$ / **eV** | $\Delta_r G$/ **eV** |
| Methanol | 0.05 | 0.01 |
| Methylamine | -0.12 | -0.31 |
| Methylthiol | 0.75 | 0.59 |
| Dimethyl ether | 0.36 | 0.28 |
| Formic acid (O-H bound) | 0.83 | 0.75 |
| Formic acid (C=O bound) | 0.26 | 0.24 |
| Acetic acid (O-H bound) | 0.92 | 0.81 |
| Acetic acid (C=O bound) | 0.30 | 0.23 |
| Acetate | -6.99 | -0.74 |
| Oxalic acid (O-H bound) | 1.04 | 1.17 |
| Oxalic acid (C=O bound) | 0.35 | 0.30 |
| Formamide (O bound) | -0.29 | -0.26 |
| Formamide (N bound) | 0.92 | 0.78 |
| Acetamide (O bound) | -0.23 | -0.35 |
| Acetamide (N bound) | 0.70 | 0.47 |
| Methyl acetate (C=O bound) | 0.08 | 0.07 |
| Methly acetate (O bound) | 0.84 | 6.38 |

**Adsorption of Alcohols, Amines, Thiols and Ethers**

The first set of organic molecules present within soils to undergo ligand substitution reactions with the Al-complex were methanol and its analogues. Similar to water, the alcohol, amine and thiol functional group contain a hydrogen-heteroatom bond that is capable of dissociation. The ether functional group was also considered due to its structural similarities with the alcohol functional group. However, ethers do not have the ability to dissociate.



(a) Methanol = 0.01 eV

(b) methylamine = 0.31 eV

(c) Methyl thiol = 0.59 eV

(d) Dimethyl ether = 0.28 eV

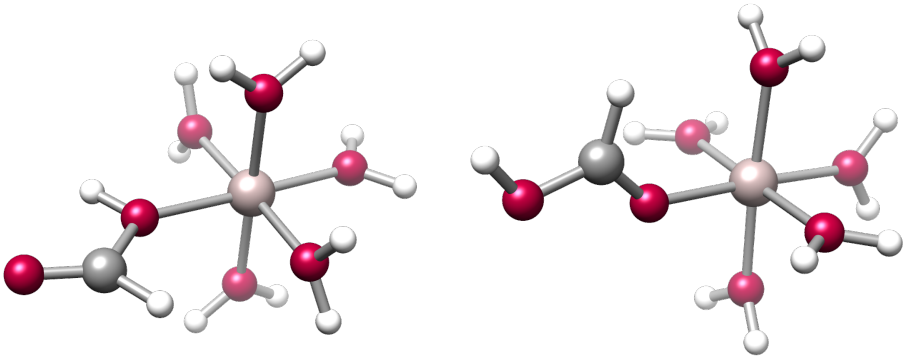
Figure 6.6: Optimised adsorption geometries of the ligand substitution product from the reaction between the $[Al(H_2O)_6]^{3+}$ complex and methanol and its analogues, at the PBE-D3 level of theory.

From our investigations methylamine (figure 6.6 (b)) formed strongly bound ligand substitution products with $\Delta_r G$ values of -0.31 eV (PBE-D3) and -0.12 eV (B3LYP). The stable negative $\Delta_r G$ values can be attributed to the electron donating ability of nitrogen, which forms a very strong bond with the $Al^{3+}$ cation. The remaining $\Delta_r G$ values for the ligand substitution products between the $[Al(H_2O)_6]^{3+}$ complex and methanol, methyl thiol and dimethyl ether were all positive. Methanol yielded $\Delta_r G$ values of 0.01 eV (PBE-D3) and 0.01 eV (B3LYP). Methyl thiol yielded $\Delta_r G$ values of 0.59 eV (PBE-D3) and 0.75 eV (B3LYP) and dimethyl ether yielded $\Delta_r G$ values of 0.28 eV (PBE-D3) and 0.36 eV (B3LYP). As the calculated Gibbs energy of reaction for methanol, methyl thiol and dimethyl ether are positive, the ligand substitution reaction between these organic ligands and the $[Al(H_2O)_6]^{3+}$ complex are endergonic. Therefore, these reactions are unfavourable and would require an additional driving force for the reactions to occur. Hence, for the alcohol functional group and its analogues, methylamine produces the most energetically favourable ligand substitution product.

**Adsorption of Acids, Amides and Esters**

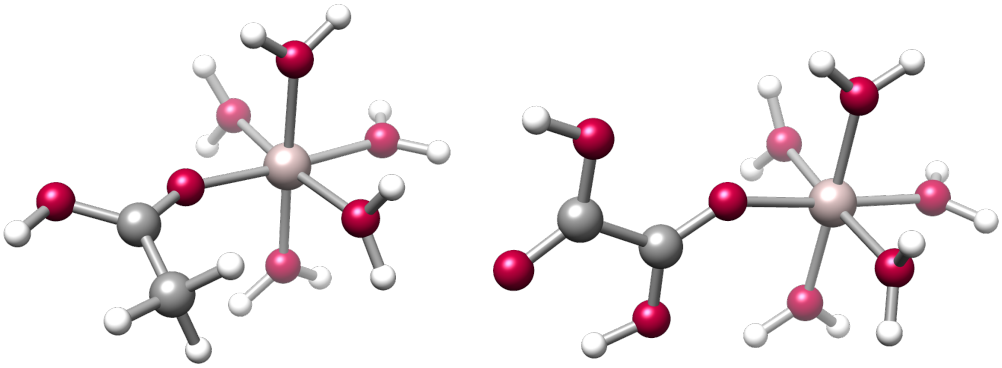
The next set of small organic molecules to undergo the ligand substitution reaction were acids, amides and esters. The smallest acid to undergo the reaction was formic acid. During our investigations we considered several different binding modes of acids. Acids can bind through their O-H oxygen atom, their carbonyl oxygen atom, as a chelating ligand and dissociatively whereby the O-H bond of the acid is broken. After several attempts to optimise a chelating complex, we were unable to optimise such structures. Formic acid was bound through both its O-H oxygen atom (figure 6.7 (a)) and its carbonyl oxygen atom (figure 6.7 (b)). Formic acid also produced endergonic ligand substitution products with calculated $\Delta_r G$ values of 0.75 eV (PBE-D3) and 0.83 eV (B3LYP) for binding through the O-H end of the acid. The calculated $\Delta_r G$ values for carbonyl binding were also endergonic with calculated $\Delta_r G$ values of 0.24 eV (PBE-D3) and 0.26 eV (B3LYP). The calculated $\Delta_r G$ values for complexation through the carbonyl end are less positive than those of the hydroxyl end of the acid. However, both reactions are unfavourable and require an additional driving force to drive the reaction ahead.

Larger acids also underwent the ligand substitution reaction to see the effect they had on binding. We continued to bind larger acids through O-H and carbonyl oxygen atoms. Similar to formic acid acetic acid also yielded $\Delta_r G$ values of 0.81 eV (PBE-D3) and 0.92 eV (B3LYP) for binding through the O-H end of the acid. For binding through the carbonyl end (figure 6.7 (c)) the $\Delta_r G$ values were 0.23 eV (PBE-D3) and 0.30 eV (B3LYP). Therefore, the addition of a methyl group had little effects on binding for the acid functional group. The complexation of oxalic acid was also investigated, to determine the effects of an additional acid functional group on binding. Oxalic acid complexation follows the trend of its smaller acid predecessors with endergonic binding. The calculated $\Delta_r G$ values were 1.17 eV (PBE-D3) and 1.04 eV (B3LYP) for the O-H end of oxalic acid. $\Delta_r G$ values of 0.30 eV (PBE-D3) and 0.35 eV (B3LYP) were calculated for the carbonyl end of oxalic acid (figure 6.7 (d)). Therefore, molecular binding of acids is thermodynamically unfavourable. The addition of a methyl constituent and an additional acid functional group do not strengthen binding. Although binding for all of the molecularly complexed acids yielded less positive $\Delta_r G$ values when bound through the carbonyl region of the acids, neither binding mode was preferred. Hence, we have illustrated both binding modes for formic acid and only carbonyl binding for the larger acids.



(a) Formic acid (O-H oxygen bound) =0.75 eV

(b) Formic acid (C=O oxygen bound) = 0.24 eV



(c) Acetic acid (C=O oxygen bound) = 0.23 eV

(d) Oxalic acid (C=O oxygen bound) = 0.30 eV

Figure 6.7: Optimised adsorption geometries of the ligand substitution product from the reaction between the $[Al(H_2O)_6]^{3+}$ complex and several acids, at the PBE-D3 level of theory.

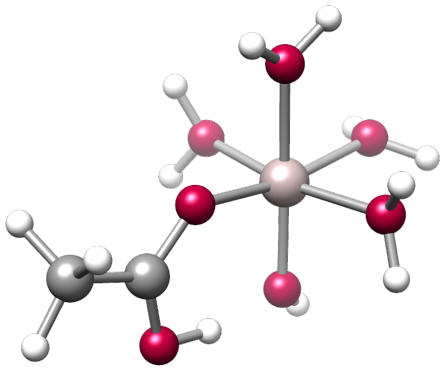


Figure 6.8: Optimised adsorption geometry of the ligand substitution product from the reaction between the $[Al(H_2O)_6]^{3+}$ complex and the acetate anion, at the PBE-D3 level of theory.

The binding of the acetate anion was also investigated to determine the difference in the strength of complexation when a neutral monodentate ligand undergoes complexation and when a negatively charged monodentate ligand undergoes complexation. The ligand substitution reaction for the acetate anion was modelled using the following equation

$$[M(H_2O)_6]^{3+} + CH_3COO^- + H^+ \rightarrow [M(H_2O)_5(CH_3COO)]^{2+} + H_3O^+ \tag{6.13}$$

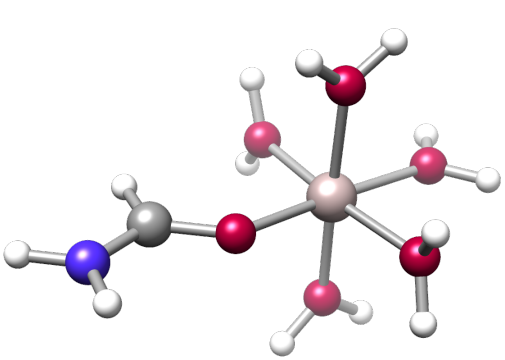
However, upon inspection of the acetate geometry (figure 6.8) it can be seen that the ligand substitution product follows the equation 6.14

$$[M(H_2O)_6]^{3+} + CH_3COO^- + H^+ \rightarrow [M(H_2O)_4(CH_3COOH)(OH)]^{2+} + H_3O^+ \quad (6.14)$$

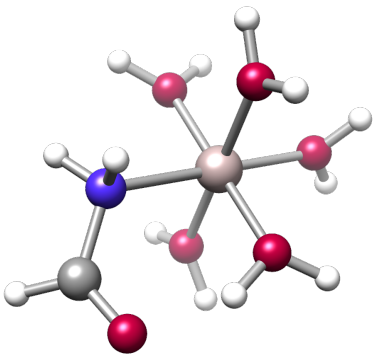
Upon optimisation acetate binding switched from the carbonyl region to the hydroxyl region, whereby the new acid proton was taken from a nearby complexed water molecule. This reaction was facilitated by hydrogen bonding and a clear hydrogen bond can be seen in figure 6.8 between the acid and hydroxyl oxygen. The ligand substitution product was the most stable of all organics investigated with calculated $\Delta_r G$ values of -0.74 eV (PBE-D3) and -6.99 eV (B3LYP). Therefore, the binding of acids is more energetically favourable during dissociative binding. These results are synonymous with our surface investigations whereby acids bound most strongly in dissociative configurations.

Similar, to acids, amides can also bind through two different atoms. Our investigations modelled the binding of formamide through its oxygen and nitrogen atoms. The oxygen bound amide (figure 6.9 (a)) produced a very strongly bound ligand substitution product, with $\Delta_r G$ values of -0.26 eV (PBE-D3) and -0.29 eV (B3LYP). When bound through the nitrogen atom (figure 6.9 (b)) the ligand substitution product had $\Delta_r G$ values of 0.78 eV (PBE-D3) and 0.92 eV (B3LYP). Therefore, binding of the amide is dominated through the oxygen end as binding through the nitrogen is endergonic, hence it is thermodynamically unfavourable. Acetamide was also modelled to determine the effects of a larger amide. When bound through its oxygen atom (figure 6.9 (c)) it too produced a highly stable ligand substitution product, with respective $\Delta_r G$ values of -0.35 eV (PBE-D3) and -0.23 eV (B3LYP). Similar to formamide, the nitrogen bound substitution product was also thermodynamically unfavourable.

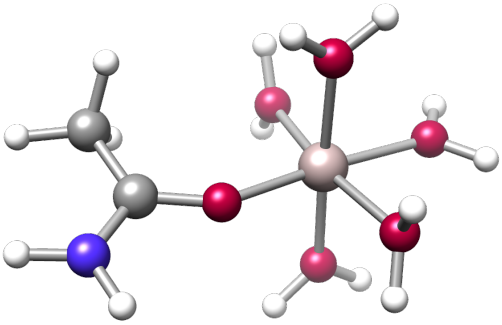
The final small organic ligand to undergo a ligand substitution reaction was methyl acetate. Much alike acids and amides, esters are also able to bind through two regions of the molecule. Binding can occur through the carbonyl oxygen atom (figure 6.9 (d)) or the C-O oxygen atom (figure 6.9 (e)). When bound through the C-O oxygen atom the ligand substitution product had $\Delta_r G$ values of 6.38 eV (PBE-D3) and 0.84 eV (B3LYP). Whereas the respective $\Delta_r G$ values were 0.07 (PBE-D3) and 0.08 eV (B3LYP) when the ester bound through the carbonyl region. Thus, the ligand substitution products of esters are thermodynamically unfavourable and are not likely to form under ambient conditions.



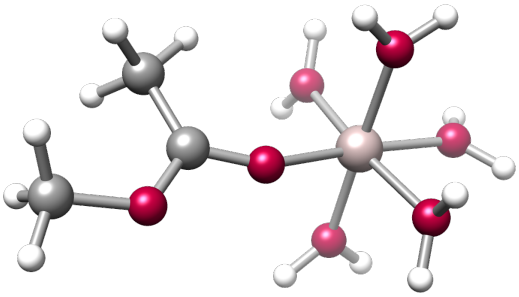
(a) Formamide (Oxygen bound)



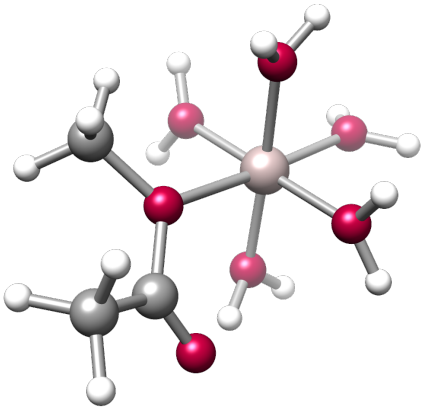
(b) Formamide (Nitrogen bound)



(c) Acetamide (Oxygen bound)



(d) Methyl acetate (C=O bound)



(e) Methyl acetate (Oxygen bound)

Figure 6.9: Optimised adsorption geometries of the ligand substitution product from the reaction between the $[Al(H_2O)_6]^{3+}$ complex and amides and esters, at the PBE-D3 level of theory.

The ligand substitution reaction between the $[Al(H_2O)_6]^{3+}$ complex and several small organic ligands has enabled us to determine which organic functional groups form strongly bound substitution products. Amines, amides and the dissociatively bound acetate anion all form thermodynamically stable substitution products. The remaining organic ligands all produced ligand substitution products with positive $\Delta_r G$ values. Therefore, these reactions were endergonic and the products produced in the reactions were thermodynamically favourable. The scheme below illustrates the binding hierarchy of small organic ligands, starting from the most strongly binding ligands (left hand side). The ligands that will not produce stable substitution products are written after the $\neq$

sign.

$$\text{Acetate} >> \text{Amide} > \text{Amine} \neq \text{Alcohol, Ester, Acid, Ether, Thiol} \tag{6.15}$$

Following the investigations on the $[Al(H_2O)_6]^{3+}$ complex, we will now model the same ligand substitution reactions between the $[Fe(H_2O)_6]^{3+}$ complex and the same small organic ligands. These investigations will enable us to compare the behaviour of $Al^{3+}$ and $Fe^{3+}$ metal ion complexes to determine whether they would behave similarly in soils, or whether one mineral ion would bind carbon in the soil more strongly than the other.

### 6.4.3 Organic Ligand Binding to the $[Fe(H_2O)_6]^{3+}$ Complex

Table 6.6: $\Delta_r G$ values the ligand substitution reaction between the $[Fe(H_2O)_6]^{3+}$ complex and organic ligands using for the B3LYP and PBE-D3 functionals at the 6-31G++(d,p) level of theory.

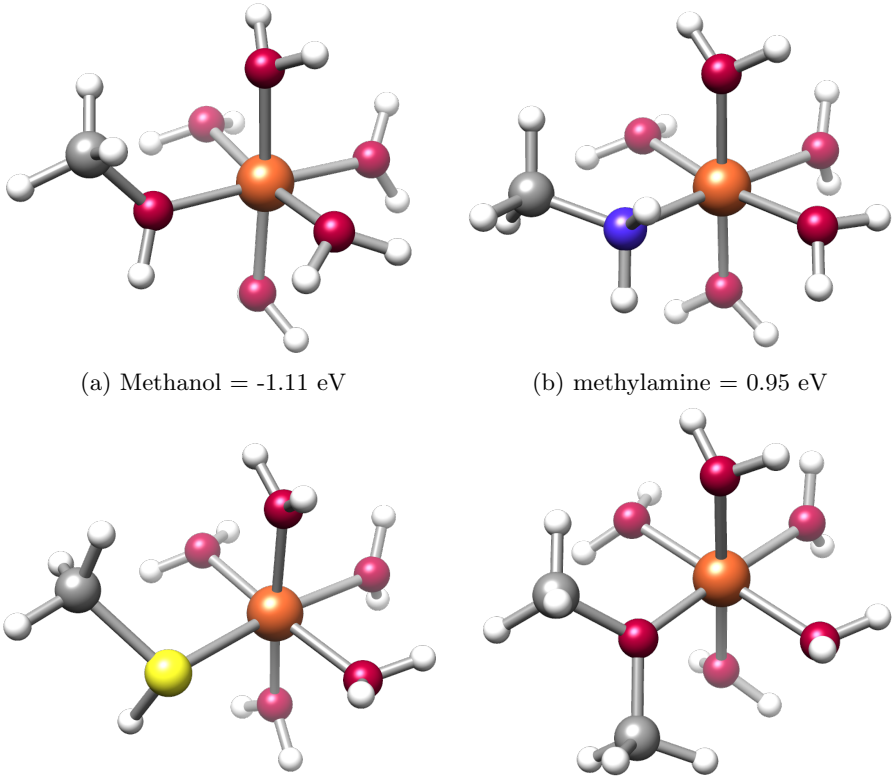
| | **B3LYP** | **PBE-D3** |
|---|---|---|
| **Organic Ligand** | $\Delta_r G$ / **eV** | $\Delta_r G$/ **eV** |
| Methanol | -0.05 | -1.11 |
| Methylamine | -0.77 | -0.95 |
| Methylthiol | 0.03 | -0.21 |
| Dimethyl ether | 0.05 | -0.14 |
| Formic acid (O-H bound) | 0.69 | 0.60 |
| Formic acid (C=O bound) | 0.08 | 0.02 |
| Acetic acid (O-H bound) | 0.72 | 0.58 |
| Acetic acid (C=O bound) | 0.12 | 0.01 |
| Acetate | -7.99 | -8.06 |
| Oxalic acid (O-H bound) | 0.72 | 0.58 |
| Oxalic acid(C=O bound) | 0.23 | 0.07 |
| Formamide (O bound) | -0.48 | -0.56 |
| Formamide (N bound) | 0.46 | 0.27 |
| Acetamide (O bound) | -0.38 | -0.55 |
| Acetamide (N bound) | 0.27 | -0.08 |
| Methyl acetate (C=O bound) | -0.06 | -0.26 |
| Methyl acetate (O bound) | 0.55 | 0.30 |

The ligand substitution reaction between the $[Fe(H_2O)_6]^{3+}$ complex and several small organic ligands was also investigated. The reaction follows equation 6.8 resulting in the formation of the $[Fe(H_2O)_5(OL)]^{3+}$ complex and the loss of a water molecule. For the acetate ligand a metal 2+ complex and $H_3O^+$ are formed.

**Adsorption of Alcohols, Amines, Thiols and Ethers**

Similar to the substitution reactions for the $Al^{3+}$ metal complexes, the amine (figure 6.9 (b)) produced highly stable ligand substitution products for both functionals. The $\Delta_r G$ values of the complexes produced were -0.95 eV (PBE-D3) and -0.77 eV (B3LYP). The ligand substitution product between the $[Fe(H_2O)_6]^{3+}$ complex and methylamine produces $[Fe(H_2O)_6(CH_3NH_2)]^{3+}$ and water. The strength of the $[Fe(H_2O)_6(CH_3NH_2)]^{3+}$ complex can again be attributed to the strong Fe-N bond produced between the highly charged $Fe^{3+}$ (hard Lewis acid) metal cation and the highly electron donating nitrogen in methylamine.

The alcohol also produced stable ligand substitution products. For methanol (figure 6.10 (a)) the respective $\Delta_r G$ values were -1.11 eV (PBE-D3) and -0.05 eV (B3LYP). The stability of the resultant complexes can be attributed to the availability of the alcohol lone pair, which forms a strong bond with the $Fe^{3+}$ metal cation. Both the ether and the thiol groups have conflicting calculated $\Delta_r G$ values. For the PBE-D3 functional the calculated $\Delta_r G$ values were -0.21 eV and -0.14 ev for the thiol and ether functional groups. However, the calculated $\Delta_r G$ values were 0.03 eV and 0.05 eV when using the B3LYP functional. The B3LYP functional is more reliable for transition metals as it represents the electrons better. Therefore, the results for the B3LYP functional are more likely to be more representative than those of the PBE-D3 functional. With this in mind, the ligand substitution products for the thiol and ether groups can be considered as endergonic, as the $\Delta_r G$ values were positive. Hence, no reaction is likely to occur between them and the $[Fe(H_2O)_6]^{3+}$ complex.



(a) Methanol = -1.11 eV

(b) methylamine = 0.95 eV

(c) Methyl thiol = -0.21 eV

(d) Dimethyl ether = -0.14 eV

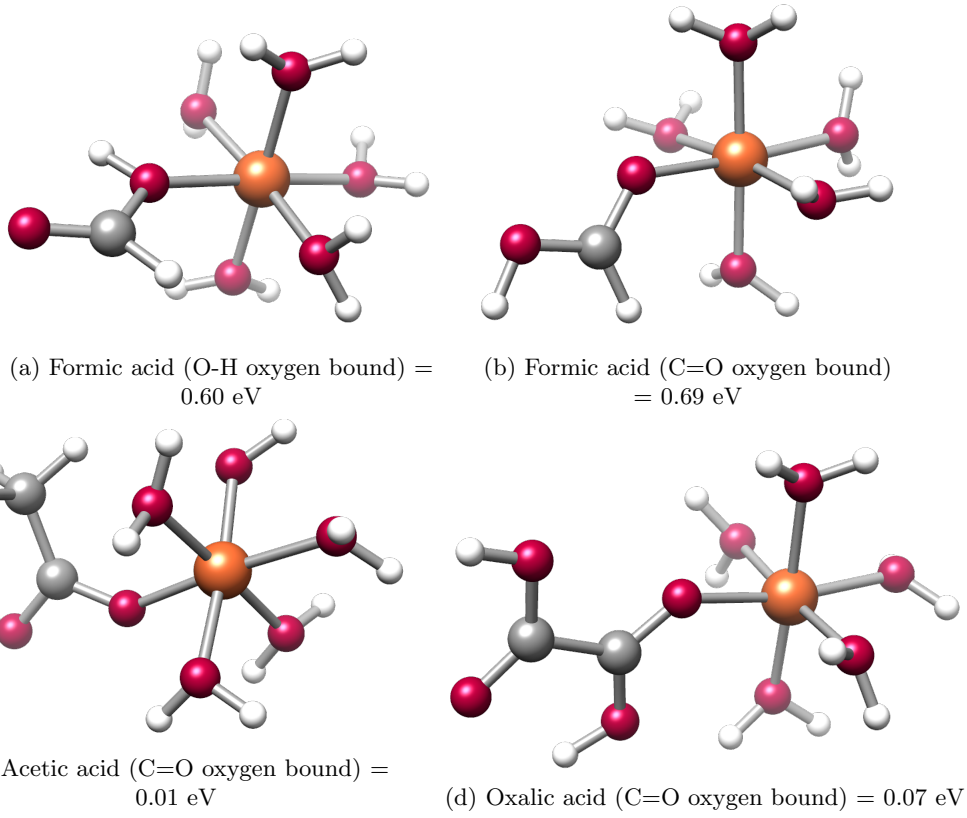
Figure 6.10: Optimised adsorption geometries of the ligand substitution product from the reaction between the $[Fe(H_2O)_6]^{3+}$ complex and methanol and its analogues, at the PBE-D3 level of theory.

### Adsorption of Acids, Amides and Esters

Acids and their analogues were the next small organics to undergo the ligand substitution reaction. Similar to the $Al^{3+}$ investigations, the smallest acid investigated was formic acid. Again, we bound acids through their hydroxyl and carbonyl oxygen atoms. Similar to the $Al^{3+}$ complex results, formic acid, acetic acid and oxalic acid also produced endergonic substitution products. When bound through the O-H oxygen atom the respective $\Delta_r G$ values were 0.60 eV (PBE-D3) (figure 11 (a)) and 0.69 eV (B3LYP). Binding through the carbonyl oxygen yielded substitution products with $\Delta_r G$ values of 0.02 eV (PBE-D3) (figure 11 (b)) and 0.08 eV (B3LYP). Acetic acid was the next acid to undergo the ligand substitution reaction. Again, we modelled binding through the carbonyl and hydroxyl regions of the molecule, we found that binding through the hydroxyl region yielded calculated $\Delta_r G$ values of 0.58 eV (PBE-D3) and 0.72 eV (B3LYP). For the carbonyl region the calculated $\Delta_r G$ values were 0.01 eV (PBE-D3) (figure 11 (c)) and 0.12 eV (B3LYP) for the substitution complexes.

The final acid to be investigated was oxalic acid, to determine the effects of an additional acid functional group on binding. The results once again demonstrate that the ligand substitution

reaction was endergonic, due to positive calculated $\Delta_r G$ values. The $\Delta_r G$ values for oxalic acid when bound via the hydroxyl oxygen atom were 0.58 eV (PBE-D3) and 0.72 eV (B3LYP). The $\Delta_r G$ values for oxalic acid when bound via the carbonyl oxygen atom (figure 11 (d)) were 0.07 eV (PBE-D3) and 0.23 eV (B3LYP). Hence, the reaction for acids and the $[Fe(H_2O)_6]^{3+}$ complex was unfavourable and would require additional driving forces to drive the reaction forward.



(a) Formic acid (O-H oxygen bound) = 0.60 eV

(b) Formic acid (C=O oxygen bound) = 0.69 eV

(c) Acetic acid (C=O oxygen bound) = 0.01 eV

(d) Oxalic acid (C=O oxygen bound) = 0.07 eV

Figure 6.11: Optimised adsorption geometries of the ligand substitution product from the reaction between the $[Al(H_2O)_6]^{3+}$ complex and several acids, at the PBE-D3 level of theory.

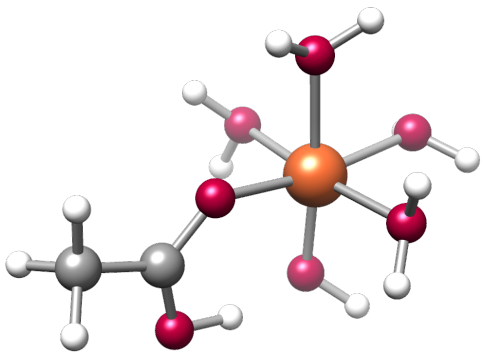
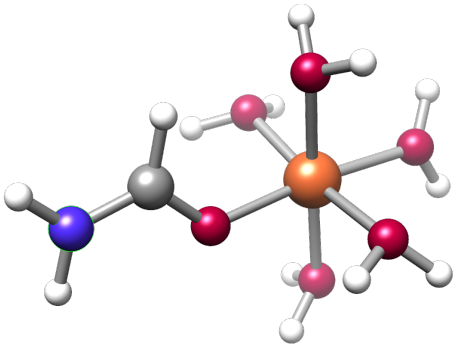
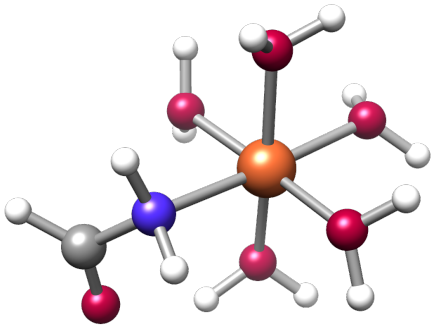


Figure 6.12: Optimised adsorption geometry of the ligand substitution product from the reaction between the $[Fe(H_2O)_6]^{3+}$ complex and the acetate anion, at the PBE-D3 level of theory.

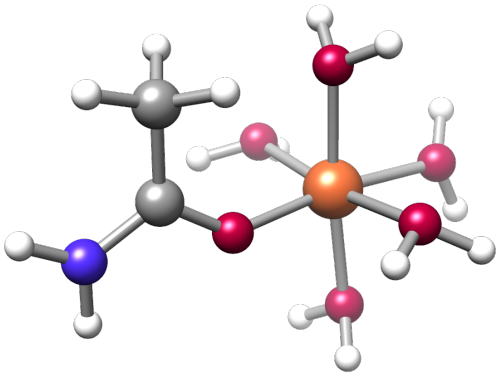
Binding of the acetate ligand was modelled based on equation 6.13. Again, upon binding the and further inspection of the acetate geometry in figure 12, binding of the acetate anion reflected binding of equation 6.14. The calculated $\Delta_r G$ values were -8.06 eV (PBE-D3) and -7.99 eV (B3LYP). These highly negative energies indicate that a very stable energetically favourable substitution product was produced for the reaction between the acetate and the $[Fe(H_2O)_6]^{3+}$ complex. Again, the stability can be attributed to the strong bond that formed between the acetate's O atom and the $Fe^{3+}$ centre and an additional H-bond that can be seen in figure 12.



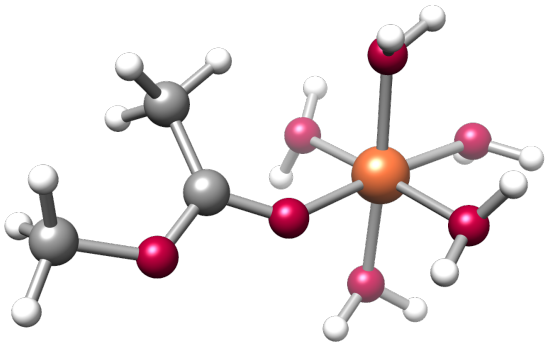
(a) Formamide (Oxygen bound) = -0.56 eV



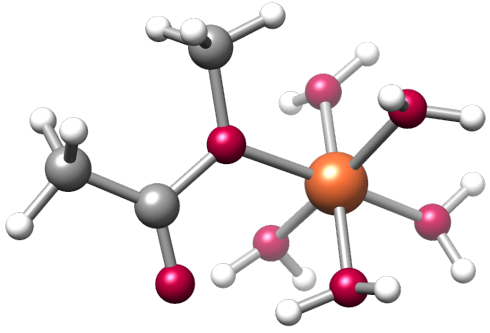
(b) Formamide (Nitrogen bound) = 0.27 eV



(c) Acetamide (Oxygen bound) = -0.55 eV



(d) Methyl acetate (C=O bound) = -0.26 eV



(e) Methyl acetate (Oxygen bound) = 0.30 eV

Figure 6.13: Optimised adsorption geometries of the ligand substitution product from the reaction between the $[Al(H_2O)_6]^{3+}$ complex and amides and esters, at the PBE-D3 level of theory.

Similar, to acids, we bound amides through two different atoms. Our investigations modelled the binding of formamide through its oxygen and nitrogen atoms. Oxygen bound formamide (figure

6.13 (a)) produced a very strongly bound ligand substitution product, with $\Delta_r G$ values of -0.56 eV (PBE-D3) and -0.48 eV (B3LYP). When bound through the nitrogen atom (figure 6.13 (b)) the ligand substitution product had $\Delta_r G$ values of 0.27 eV (PBE-D3) and 0.46 eV (B3LYP). Therefore, formamide binding is preferred through its oxygen atom. Binding through the nitrogen atom leads to an endergonic reaction (positive $\Delta_r G$ values), hence the substitution product will not form. Similar results can be seen for acetamide, a larger amide. Again, binding through the amide oxygen atom (figure 6.13 (c)) is preferred, as the substitution products yielded $\Delta_r G$ values of of -0.55 eV (PBE-D3) and -0.38 eV (B3LYP). Nitrogen-bound ligand substitution products were again, thermodynamically unfavourable for B3LYP calculations (0.27 eV).

Esters were the final small organic ligand to undergo the ligand substitution reaction. Much alike acids and amides, esters were also able bound through two regions of the molecule. Binding occurred through the carbonyl oxygen atom (figure 6.13 (d)) and the C-O oxygen atom (figure 6.13 (e)). When bound through the C-O oxygen atom the ligand substitution product had $\Delta_r G$ values of 0.30 eV (PBE-D3) and 0.55 eV (B3LYP). The respective $\Delta_r G$ values were -0.26 eV (PBE-D3) and -0.06 eV (B3LYP) when methyl acetate bound through the carbonyl oxygen atom. Therefore, the ester functional group produces an energetically favourable substitution product when it binds through the carbonyl oxygen atom, similar to amides. Reactions through the C-O oxygen atom are energetically favourable.

The ligand substitution reaction between the $[Fe(H_2O)_6]^{3+}$ complex and several small organic ligands has enabled us to determine which organic functional groups form strongly bound substitution products. Several organic functional groups produced energetically favourable ligand substitution products. The most stable product was produced by the acetate ligand, followed by the amine and amide. On the other hand, acid, ether and thiol groups all resulted in thermodynamically unfavourable products.

PBE-D3:

$$\text{Acetate} >> \text{Alcohol} > \text{Amine} > \text{Amide} > \text{Ester} > \text{Thiol} \neq \text{Ether}, \text{Acid} \tag{6.16}$$

B3LYP:

$$\text{Acetate} >> \text{Amine} > \text{Amide} > \text{Ester} > \text{Alcohol} >\neq \text{Thiol}, \text{Ether}, \text{Acid} \tag{6.17}$$

The schemes above illustrate the binding hierarchies for the PBE-D3 and B3LYP functionals,

starting from the most strongly binding ligands (left) to the least strongly binding ligands (right).

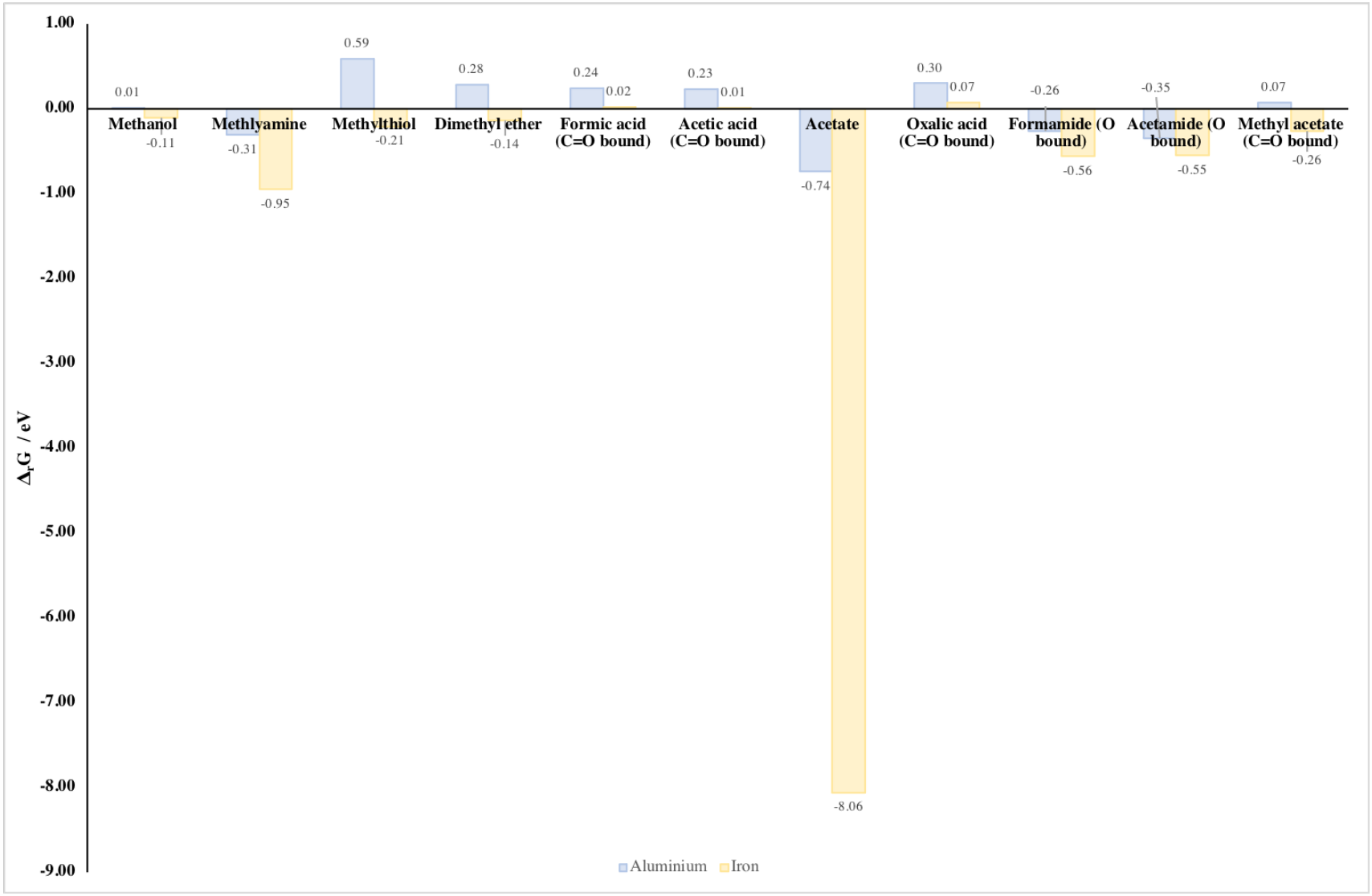



Figure 6.14: Calculated $\Delta_r G$ values for the ligand substitution product produced between the reactions of the $[Al(H_2O)_6]^{3+}$ and $[Fe(H_2O)_6]^{3+}$ complexes with organics at the PBE-D3 631G++(d,P) level of theory.

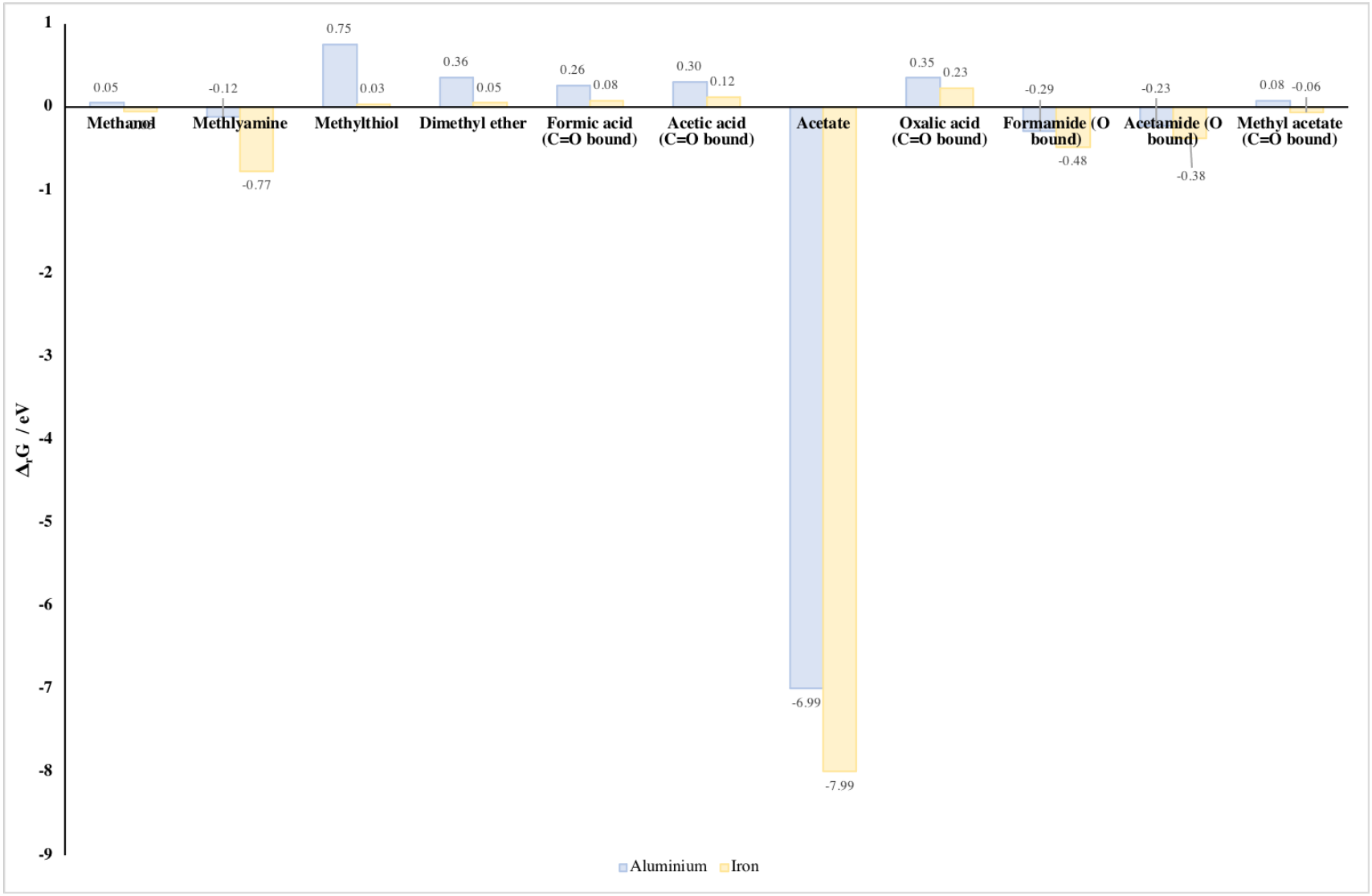



Figure 6.15: Calculated $\Delta_r G$ values for the ligand substitution product produced between the reactions of the $[Al(H_2O)_6]^{3+}$ and $[Fe(H_2O)_6]^{3+}$ complexes with organics at the B3LYP 631G++(d,P) level of theory.

### 6.4.4 Binding of $Al^{3+}$ vs. $Fe^{3+}$

Here, we will compare and contrast the results from our investigations to determine which metal cation produces more energetically favourable products during the ligand substitution reaction between the $[M(H_2O)_6]^{3+}$ complex and small organics. For $Al^{3+}$ the binding hierarchy for both functionals can be summed up as follows,

$$\text{Acetate} >> \text{Amide} > \text{Amine} \neq \text{Alcohol}, \text{Ester}, \text{Acid}, \text{Ether}, \text{Thiol} \tag{6.18}$$

$Fe^{3+}$ can be interpreted as the following for both functionals,

$$\text{Acetate} >> \text{Amine} > \text{Amide} > \text{Alcohol}, \text{Ester} \neq \text{Acid}, \text{Ether}, \text{Thiol} \tag{6.19}$$

For both the $Al^{3+}$ and $Fe^{3+}$ metal cation investigations the acetate ligand produced the most energetically favourable ligand substitution products. The strong binding can be explained by the availability of the oxygen lone pair in acetate, which forms a strong bond with the metal cation. Inspection of the geometries of bound acetate illustrate that the geometry was further stabilised by hydrogen bonding between a surface hydroxyl and an acetate hydrogen atom. These results are in agreement with the surface investigations whereby dissociative acids bound very strongly to the bare and hydroxylated alumina (0001) surface. Due to time constraints we were unable to adsorb other dissociated species such as $CH_3O^-$. However, the acetate investigations have led us to believe binding would be strong and favourable between negatively charged species and positive metal cation centres.

After acetate, both the amine and amide also produced energetically favourable ligand substitution products. The $Al^{3+}$ yielded larger $\Delta_r G$ values when bound to the amide followed by the amine and $Fe^{3+}$ yielded larger values for the amine followed by the amide. Similarly, amines and amides also bound very strongly to the alumina surface. Hence, biopolymers such as chitin and chitosan were chosen to model amine and amide adsorption in complex molecules. For the $Fe^{3+}$ investigations alcohols and esters also produced energetically favourable substitution products. However, $Al^{3+}$ investigations illustrated that the binding of alcohols and esters was endergonic. Acid, ether and thiol groups also led to endergonic reactions, therefore their subsequent substitution products were not energetically favourable and the reaction would require work to be done to drive the reactions forward.

For both the $Al^{3+}$ and $Fe^{3+}$ investigations, particularly during the acid, ester and amide investi-

gations structures were optimised through different regions of the molecule, a product of different binding sites. Acids were bound through their carbonyl oxygen atom and their hydroxyl oxygen atom. They can also bind through the oxygen atom of the dissociated O-H bond to produce the carboxylate anion, and as a chelating ligand. Unfortunately, we were unable to optimise acids as chelating ligands due to time constraints. Weak acid adsorption for both the $Al^{3+}$ and $Fe^{3+}$ investigations is due the lack of dissociation of the O-H bond. The bond produced between the $M^{3+}$ and the carboxylate anion is a lot stronger than the bond between an uncharged acid in a molecular formation. This is because the attraction between the positive charge of the metal and negative charge of the anion is strong, thus the resultant substitution complex would be more energetically favoured. Binding of acids was neither preferred through the carbonyl or hydroxyl oxygen, as reactions were endergonic, so the substitution products would not form. Binding of the amide was energetically favourable when it occurred through the amide oxygen. Lastly, ester binding was preferred through the carbonyl oxygen for our $Fe^{3+}$ investigations, and reactions were endergonic for our $Al^{3+}$ investigations.

Schemes 6.18 and 6.19 both illustrate which ligands produced the most energetically favourable products during the ligand substitution reaction. Our $Al^{3+}$ investigations yielded the same results for the PBE-D3 and B3LYP functionals. However, results for the $Fe^{3+}$ investigations differed slightly. For $Fe^{3+}$ PBE-D3 investigations (scheme 6.16) ethers and acids produced unfavourable products, whereas in the B3LYP investigations thiols, ethers and acids produced unfavourable products. Both functional represent the $Al^{3+}$ well as they yielded similar results. However, $Fe^{3+}$ is a transition metal and the spread of electrons is better represented by a hybrid functional (B3LYP) than a pure DFT functional (PBE).

Figures 6.14 and 6.15 illustrate the most stable complexes for each organic ligand in the PBE-D3 and B3LYP $Al^{3+}$ and $Fe^{3+}$ investigations. For all energetically favourable products, one can see that binding to the $Fe^{3+}$ cation was preferred. The pKa (dissociation constant) values for the $Al^{3+}$ and $Fe^{3+}$ metal cations explain why the $Fe^{3+}$ cation complexes were more energetically favourable. The pKa values for the $Al^{3+}$ and $Fe^{3+}$ metal cations are 5.0 and 2.2 at 298K.[341] Therefore, $Al^{3+}$ is a weaker acid in solution as it has a larger pKa. Consequently, the resultant complexes from the ligand substitution reaction between the $[Al(H_2O)_6]^{3+}$ complex and the organic ligands will result in less energetically favoured substitution products when compared directly to the complexes produced in the reaction between the $[Fe(H_2O)_6]^{3+}$ complex and organic ligands.

## 6.5 Conclusion

The aims of our metal ion investigations was to investigate the nature of metal cation binding within soils to identify strongly binding groups. Our research also aimed to compare the binding of Al and Fe cations and to evaluate the ability of organic groups to capture these metal cations. The ligand substitution reaction between hydrated $Al^{3+}$ and $Fe^{3+}$ cations with small organic molecules was investigated. In the $Al^{3+}$ investigations replacing a water ligand with an organic ligand was unfavourable for ether and acid groups. However, replacing water with acetate, amine or amide ligands was favourable. The substitution of water with an acetate anion was most favourable. On the other hand, $Fe^{3+}$ investigations showed that the substitution of water by organic ligands was more favourable compared to $Al^{3+}$. Again, acetate was the strongest binding ligand, followed by amines and amides. Therefore, the best ligands for binding metal ions in soil solutions are the amine, amide and dissociated acids. Overall, $Fe^{3+}$ complexes were more energetically favoured than $Al^{3+}$ complexes, as the $Fe^{3+}$ cation is more acidic than $Al^{3+}$. Hence, organic molecules bind more strongly to the $Fe^{3+}$ cation.

# Chapter 7

# Conclusions

In this work we investigated the interaction of organic carbon with soil minerals and soil metal cation complexes using DFT methods. We calculated binding energies, Gibbs reaction energies and deformation energies to identify the carbon containing molecules that bound most strongly to the alumina surfaces and $Al^{3+}$ and $Fe^{3+}$ complexes within soils. These investigations enabled us to gain a quantitative description of binding within soils in an effort to preserve OC within soils.

We began our investigations of OC-mineral interactions by investigating the adsorption of small organic molecules to the bare $\alpha$-$Al_2O_3$ (0001) surface (chapter 3) using DFT methods. In chapter 3 we adsorbed water and several small organic functional groups to the $\alpha$-$Al_2O_3$ (0001) surface. Water was first adsorbed as it is an important molecule that is naturally present within in environmental systems. Therefore, organic molecules within soils will have to compete with water when binding to soil mineral surfaces. Alcohols, amines, acids and amides all bound more strongly than water to the surface. The acid was the strongest surface adsorbate and its adsorption was facilitated by hydrogen bonding. All adsorbates chemisorbed to the surface, with the exception of methane that was bound via dispersion interactions. Several molecules could also adsorb molecularly and dissociatively and dissociative adsorption was the preferred binding mode for most adsorbates.

Following our bare surface investigations we investigated the adsorption of small organic molecules to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface (chapter 4), as mineral surfaces are typically hydrated in soils. In chapter 4 we adsorbed water and several small organic functional groups on the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface using DFT methods. Unlike the chemisorption adsorption process occurring in chapter 3, adsorption in chapter 4 was facilitated by hydrogen bonding. Molecules such as the amide and acid were able to form two hydrogen bonds with the surface. Thus, resulting in stronger adsorption. Water and the amine functional group were also strongly adsorbing. Sim-

ilar to chapter 3 the acids were again the strongest carbon containing adsorbate. DFT based MD simulations were employed to investigate the leaving group ability of water and acetic acid when bound to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. Though there was some movement of bonds in these calculations, we were unable to obtain any conclusive evidence of either group leaving. Therefore, further investigation would be required.

We then moved onto our most comprehensive OC-mineral interactions investigations (chapter 5). In chapter 5 we adsorbed four different biopolymers to the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface, again using DFT methods. Biopolymers were selected using the knowledge attained from the investigations in chapters 3 and 4. In chapters 3 and 4 we determined that the alcohol, amine, acid and amide functional groups were all strongly binding. Therefore, we researched naturally occurring biopolymers that can be found within soils that also contain these key functional groups. Hence, we adsorbed cellulose (alcohol groups), chitin (amide and alcohol groups), chitosan (amine and alcohol groups) and pectin (alcohol, acid and ester groups). From these investigations we found that the binding of biopolymers is facilitated by hydrogen bonding similar to binding in chapter 4. However, biopolymers adsorbed more strongly to the hydroxylated surface as they were able to form a greater number of hydrogen bonds upon adsorption. During the biopolymer investigations two adsorption geometries yielded the largest binding energies, the flat and parallel geometry and the on-edge geometry. Pectin was the most strongly adsorbing biopolymer and chitin was the weakest. The weaker adsorption of chitin was a product of its twisted anti-parallel structure which only enabled half of the chitin molecule to be adsorbed at any given time.

Upon completion of our adsorption investigations of OC to the bare and hydroxylated $\alpha$-$Al_2O_3$ (0001) surfaces, we investigated the binding of OC to the hydrated $Al^{3+}$ and $Fe^{3+}$ complexes (chapter 6). The ligand substitution reaction between hydrated $Al^{3+}$ and $Fe^{3+}$ complexes and small organic molecules yielded interesting results. The most strongly adsorbing ligand in both of our metal cation investigations was the acetate ligand. Amides and amines followed suit, as strongly binding ligands. Acids and ethers produced unfavourable ligand substitution products. The acid was amongst one of the weakest adsorbates in our complex investigations, although it was the strongest adsorbate in our surface investigations. In solution acids are typically dissociated, therefore we would expect stronger binding of acids, if dissociated, as was the case for acetate. $Fe^{3+}$ complexes were more energetically favoured in our investigations, due to the acidity of the central $Fe^{3+}$ cation. The $Fe^{3+}$ metal cation investigations also yielded an understanding of the way $Fe^{3+}$ based minerals may behave in soils and which organic molecules would bind strongly to them.

During all investigations we adopted a systematic approach for binding a variety of organic

molecules to the alumina surfaces and $Al^{3+}$ and $Fe^{3+}$ complexes. We considered several structures for our adsorbates during all of our binding investigations. However, other low energy structures are likely to exist, but due to time constraints we were unable to obtain every structure. The investigations carried out gave a representative picture of the interaction between OC with Al-based and Fe-based species in soil. The results we have obtained enable the prediction of which organic species are likely to be stable in soil. Generally, molecules containing the the acid and amine groups and the biopolymers pectin and cellulose are likely to be stable in soils. Based on our investigations in this work, both pectin and cellulose can be recommended to add into soils, to increase the carbon content and the retention of organic carbon in soils.

## 7.1 Future Work

Our investigations on OC adsorption to the bare and hydroxylated alumina surfaces and the hexa-aqua $Al^{3+}$ and $Fe^{3+}$ complexes, has given rise to a greater understanding of the interaction between organic carbon and mineral surfaces present within soils. Our current research focused on the adsorption of small organic molecules to the hexa-aqua metal complexes, via a fairly simple ligand substitution reaction. Our mineral oxide surface investigations focused upon the adsorption of small organics and biopolymers to the bare and hydroxylated $\alpha$-$Al_2O_3$ (0001) surfaces. With our research we have gained a better understanding of OC interaction in soils. However, soil chemistry is vast due to its complex nature.[1] Soils contain a multitude of inorganic mineral surfaces, hence, the interaction of other minerals should also be investigated. A mechanistic understanding of binding would also be beneficial as one would not only have the knowledge of key groups that bind to minerals, but also the mechanism, be it associative, dissociative or an interchange type of mechanism.[1,16] Therefore, as a continuation of this project we would further investigate other species and their role in OC binding. For example, our metal ion investigations determined that the $Fe^{3+}$ complexes produced more energetically favourable reaction products. Thus, we would model adsorption investigations to $Fe^{3+}$ based mineral oxide surfaces such as hematite ($\alpha$-$Fe_2O_3$) and goethite ($\alpha$-FeOOH).

At present, sixteen different types of iron oxides are known. These oxides differ in composition, oxidation state of the Fe ion and most importantly, crystal structure.[342,343] Goethite and hematite are ubiquitous on the Earth's surface, as weathering products from iron bearing sedimentary and metamorphic rocks.[344] Hematite is also the most abundant iron oxide on earth and is also iso-structural with $\alpha$-$Al_2O_3$ (alumina).[196,342,345] Both goethite and hematite are antiferromagentic iron oxides[342,346,347] With respect to Alumina the $Al^{3+}$ binding site is less reactive than the $Fe^{3+}$ binding site in hematite.[196] Hence, knowledge of the interaction between organic carbon and the

more reactive iron oxide surfaces would be beneficial for creating carbon based fertilisers that bind to soil minerals more effectively.

Few theoretical studies[196,342] were conducted in the literature on the adsorption of OC to iron oxide surfaces. The majority of the studies were experimental investigations.[82,84,90,91,348] A very recent theoretical study[196] compared the $\gamma$-$Al_2O_3$ (100) surface and the $\alpha$-$Fe_2O_3$ (0001) surface and demonstrated high reactivity of the hematite surface. These results are consistent with results from our metal ion investigations. It would be important to investigate the binding properties of a variety of minerals, in particular the interaction with polysaccharides, as they are naturally occurring within soils. For example, cellulose is the most abundant plant biopolymer on earth, hence knowledge of its interaction with several mineral surfaces will give us a greater understanding of the binding of complex molecules. Ultimately, we would like to design *'carbon based'* fertilisers, whereby the carbon is of a nature where it strongly binds to mineral oxides within soils. Thus, allowing us to fine tune the carbon containing molecules, to give the maximum improvement in soil quality.

Another avenue of our future work would be to consider the adsorption of amino acids. Amino acids are very important as a half of the nitrogen in soil humus comprises of amino acids. However, their absorption behaviour to mineral surfaces has not been investigated. Amino acids can be neutral, basic or acidic. Therefore, they are pH sensitive molecules, which will yield thought-provoking investigations allowing us to further delve into the complex nature of OC binding in soils.

The interaction of metal cations with soil minerals can also be investigated. Research into the goethite surface[349,350] has revealed the effects of metal cation ($Ca^{2+}$, $Mg^{2+}$ and $Cr^{3+}$) poisoning on the adsorption of acids. Free metal cations are toxic to the soil and inhibit plant growth. Therefore, determining the organic molecules that have a higher affinity for the hematite surface than the toxic metal ions would be beneficial. We would follow on from the experimental work in the literature on goethite[349,350] and also model the hematite surface using theoretical calculations. The outlined directions of the investigation would gain a more comprehensive and complete understanding of the interaction between OC and soil minerals. We hope to achieve this by building upon our previous research on the alumina surface using the hematite and goethite surfaces, to further develop the design of *'carbon based'* fertilisers.

Finally, our investigations into the binding of organics with soil minerals and mineral cations predominately focused on thermodynamics. Binding energies and Gibbs energies of reaction were calculated to deem whether the geometries obtained after adsorption/binding were energetically stable. However, the reaction kinetics of adsorption processes in soils are also important. We

obtained several stable geometries, but bonds were broken to create these structures. Breaking bonds requires energy, therefore these process will have an energy barrier. Hence, the calculation of activation energies for the adsorption of small organics to mineral surfaces/cations would be required to determine the rate of reaction. Thus, we would be able to compare both the kinetics and thermodynamics of adsorption processes within soils, to determine the likeliness of them occurring.

# Bibliography


[1] G Sposito. *The chemistry of soils*. Oxford University Press, Oxford, 2008.

[2] DL Sparks. *Environmental soil chemistry*. Elsevier, Delaware, 2003.

[3] JB Dixon and SB Weed. Aluminum hydroxides and oxyhydroxides. *Minerals in soil environments*, 1:331–378, 1989.

[4] DG Schulze. An introduction to soil mineralogy. *Soil mineralogy with environmental applications*, 7:1–35, 2002.

[5] DG Schulze. An introduction to soil mineralogy. *JB & Weed SB (red.)*, pages 1–34, 1989.

[6] U Schwertmann and RM Cornell. *Iron oxides in the laboratory: preparation and characterization*. John Wiley & Sons, Switzerland, 2008.

[7] FJ Stevenson. *Humus chemistry: genesis, composition, reactions*. John Wiley & Sons, Illinois, 1994.

[8] CM Chang and MK Wang. Linear relationship for acidity and stability in hexaaqua metal ions—density functional studies. *Chemical physics letters*, 286(1-2):46–50, 1998.

[9] DG Strawn, HL Bohn, and O'Connor GA. Redox reactions in soil. *Soil chemistry 4th ed. John Wiley and Sons, Hoboken, NJ*, pages 213–143, 2015.

[10] J Harvey. *Computational chemistry*. Oxford University Press, Oxford, 2018.

[11] VA Ranea, I Carmichael, and WF Schneider. DFT investigation of intermediate steps in the hydrolysis of $\alpha$-al$_2$o$_3$ (0001). *The Journal of Physical Chemistry C*, 113(6):2149–2158, 2009.

[12] Y Nishiyama, P Langan, and H Chanzy. Crystal structure and hydrogen-bonding system in cellulose I$\beta$ from synchrotron x-ray and neutron fiber diffraction. *Journal of the American Chemical Society*, 124(31):9074–9082, 2002.

[13] SD McConaughy. *Examination of an aloe vera galacturonate polysaccharide capable of in situ gelation for the controlled release of protein therapeutics*. PhD thesis, The University of Southern Mississippi, 2008.

[14] Y Lu, Y-C Lu, H-Q Hu, F-J Xie, X-Y Wei, and X Fan. Structural characterization of lignin and its degradation products with spectroscopic methods. *Journal of Spectroscopy*, 2017, 2017.

[15] GM Ruiz and HFZ Corrales. Chitosan, chitosan derivatives and their biomedical applications. *Biological activities and application of marine polysaccharides*, 87, 2017.

[16] H Hanauer, R Puchta, T Clark, and R van Eldik. Searching for stable, five-coordinate aquated Al (iii) species. water exchange mechanism and effect of pH. *Inorganic chemistry*, 46(4):1112–1122, 2007.

[17] JT Way. *On the power of soils to absorb manure*. London, 1850.

[18] LG Berry and BH Mason. Mineralogy, concepts, descriptions, determinations. 1959.

[19] L Pauling. The principles determining the structure of complex ionic crystals. *Journal of the American Chemical Society*, 51(4):1010–1026, 1929.

[20] PM Huang. Feldspars, olivines, pyroxenes, and amphiboles. *Minerals in soil environments*, 1:975–1050, 1989.

[21] DL Sparks. *Potassium dynamics in soils*. Springer, New York, 1987.

[22] S Feigenbaum, A Bar-Tal, and DL Sparks. Dynamics of soil potassium in multicationic systems. *Development of K fertilizer Recommendations*, pages 145–161, 1990.

[23] C Chang, F Li, C Liu, J Gao, H Tong, and M Chen. Fractionation characteristics of rare earth elements (rees) linked with secondary Fe, Mn, and Al minerals in soils. *Acta Geochimica*, 35(4):329–339, 2016.

[24] DM Moore and RC Reynolds Jr. *X-ray Diffraction and the Identification and Analysis of Clay Minerals*. Oxford University Press, Oxford, 1989.

[25] E Paterson, BA Goodman, and VC Farmer. The chemistry of aluminium, iron and manganese oxides in acid soils. *Soil acidity*, pages 97–124, 1991.

[26] AC Birnie and E Paterson. The mineralogy and morphology of iron and manganese oxides in an imperfectly-drained scottish soil. *Geoderma*, 50(3):219–237, 1991.

[27] D Postma. The reactivity of iron oxides in sediments: a kinetic approach. *Geochimica et Cosmochimica Acta*, 57(21-22):5027–5034, 1993.

[28] OK Borggaard. Iron oxides in relation to phosphate adsorption by soils. *Acta Agriculturae Scandinavica*, 36(1):107–118, 1986.

[29] RR Haese. *The reactivity of iron*. Springer, Berlin, 2000.

[30] A Violante, M Ricciardella, S Del Gaudio, and M Pigna. Coprecipitation of arsenate with metal oxides: nature, mineralogy, and reactivity of aluminum precipitates. *Environmental science & technology*, 40(16):4961–4967, 2006.

[31] PM Bertsch and PR Bloom. Aluminum. *Methods of Soil Analysis: Part 3 Chemical Methods*, 5:517–550, 1996.

[32] PP Reichertz and WJ Yost. The crystal structure of synthetic boehmite. *The Journal of chemical physics*, 14(8):495–501, 1946.

[33] NM Miles, GC Scott, and AJ Green. The occurrence of gibbsite in a british columbia soil. *Canadian Journal of Soil Science*, 58(4):529–533, 1978.

[34] C Shen and R Lu. The color origin of gem diaspore: correlation to corundum. *Gems & Gemology*, 54(4):394–403, 2018.

[35] JD Gale, AL Rohl, V Milman, and MC Warren. An ab initio study of the structure and properties of aluminum hydroxide: gibbsite and bayerite. *The Journal of Physical Chemistry B*, 105(42):10236–10242, 2001.

[36] J Frenzel, AF Oliveira, HA Duarte, T Heine, and G Seifert. Structural and electronic properties of bulk gibbsite and gibbsite surfaces. *Zeitschrift für anorganische und allgemeine Chemie*, 631(6-7):1267–1271, 2005.

[37] N Kämpf, AC Scheinost, and DG Schulze. Oxide minerals. *Handbook of soil science*, pages 125–168, 2000.

[38] K Norrish and RM Taylor. The isomorphous replacement of iron by aluminium in soil goethites. *Journal of Soil Science*, 12(2):294–306, 1961.

[39] P Rengasamy, K Murti, and VAK Sarma. Isomorphous substitution of iron for aluminum in some soil kaolinites. *Clays and Clay Minerals*, 23(3):211–14, 1975.

[40] E Bazilevskaya, DD Archibald, M Aryanpour, JD Kubicki, and CE Martínez. Aluminum co-precipitates with Fe (hydr) oxides: Does isomorphous substitution of $Al^{3+}$ for $Fe^{3+}$ in goethite occur? *Geochimica et Cosmochimica Acta*, 75(16):4667–4683, 2011.

[41] RA Berner. Goethite stability and the origin of red beds. *Geochimica et Cosmochimica Acta*, 33(2):267–273, 1969.

[42] J Majzlan, K-D Grevel, and A Navrotsky. Thermodynamics of fe oxides: Part ii. enthalpies of formation and relative stability of goethite ($\alpha$-FeOOH), lepidocrocite ($\gamma$-FeOOH), and maghemite ($\gamma$-$Fe_2O_3$). *American Mineralogist*, 88(5-6):855–859, 2003.

[43] YT He, J Wan, and T Tokunaga. Kinetic stability of hematite nanoparticles: the effect of particle sizes. *Journal of Nanoparticle research*, 10(2):321–332, 2008.

[44] JL Junta-Rosso and MF Hochella Jr. The chemistry of hematite 001 surfaces. *Geochimica et Cosmochimica Acta*, 60(2):305–314, 1996.

[45] Y Tardy and D Nahon. Geochemistry of laterites, stability of Al-goethite, Al-hematite, and Fe3+-kaolinite in bauxites and ferricretes: an approach to the mechanism of concretion formation. *American Journal of Science*, 285(10):865–903, 1985.

[46] S Manahan. *Environmental chemistry*. CRC press, New York, 2017.

[47] M Schnitzer. Soil organic matter—the next 75 years. *Soil science*, 151(1):41–58, 1991.

[48] AV Spain, RF Isbell, and ME Probert. Soil organic matter. *Soils: an Australian viewpoint*, pages 551–563, 1983.

[49] H Tiessen, E Cuevas, and P Chacon. The role of soil organic matter in sustaining soil fertility. *Nature*, 371(6500):783–785, 1994.

[50] J Lehmann and M Kleber. The contentious nature of soil organic matter. *Nature*, 528(7580):60–68, 2015.

[51] JA Baldock and PN Nelson. *Soil organic matter*. CRC press, Boca Raton, 2000.

[52] M Schnitzer and SU Khan. Soil organic matter. *Soil Science*, 151(1):41–48, 1975.

[53] A Piccolo. Humus and soil conservation. *Humic substances in terrestrial ecosystems*, pages 225–264, 1996.

[54] S Nardi, G Concheri, and G Dell'Agnola. Biological activity of humus. In *Humic substances in terrestrial ecosystems*, pages 361–406. Elsevier, 1996.

[55] R Sutton and G Sposito. Molecular structure in soil humic substances: the new view. *Environmental science & technology*, 39(23):9009–9015, 2005.

[56] M Pittarello, JG Busato, P Carletti, LV Zanetti, J da Silva, and LB Dobbss. Effects of different humic substances concentrations on root anatomy and Cd accumulation in seedlings of avicennia germinans (black mangrove). *Marine pollution bulletin*, 130:113–122, 2018.

[57] M Schnitzer. Humic substances: chemistry and reactions. In *Developments in soil science*, volume 8, pages 1–64. Elsevier, 1978.

[58] P MacCarthy. The principles of humic substances. *Soil Science*, 166(11):738–751, 2001.

[59] CEW Steinberg, T Meinelt, MA Timofeyev, M Bittner, and R Menzel. Humic substances. *Environmental Science and Pollution Research*, 15(2):128–135, 2008.

[60] EA Ghabbour, G Davies, and G Davies. *Humic substances: structures, models and functions*, volume 273. Royal Society of Chemistry, 2001.

[61] GR Aiken, DM Mcknight, RL Wershaw, and P Maccarthy. Humic substances in soil, sediment, and water. 1985. *Soil Science*, 142(5):323, 1986.

[62] M Schnitzer and HR Schulten. The analysis of soil organic matter by pyrolysis-field ionization mass spectrometry. *Soil Science Society of America Journal*, 56(6):1811–1817, 1992.

[63] ZN Senwo and MA Tabatabai. Amino acid composition of soil organic matter. *Biology and Fertility of Soils*, 26(3):235–242, 1998.

[64] JM Bremner. The amino-acid composition of the protein material in soil. *Biochemical Journal*, 47(5):538, 1950.

[65] LE Lowe. Carbohydrates in soil. In *Developments in Soil Science*, volume 8, pages 65–93. Elsevier, 1978.

[66] NC Mehta, P Dubach, and H Deuel. Carbohydrates in the soil. In *Advances in carbohydrate chemistry*, volume 16, pages 335–355. Elsevier, 1962.

[67] MV Cheshire et al. Nature and origin of carbohydrates in soils. *Developments in Plants and Soil Sciences*, 16:263–268, 1979.

[68] H Bahri, M-F Dignac, C Rumpel, DP Rasse, C Chenu, and A Mariotti. Lignin turnover kinetics in an agricultural soil is monomer specific. *Soil Biology and Biochemistry*, 38(7):1977–1988, 2006.

[69] R Datta, A Kelkar, D Baraniya, A Molaei, A Moulick, RS Meena, and P Formanek. Enzymatic degradation of lignin in soil: a review. *Sustainability*, 9(7):1163, 2017.

[70] SMO Souvi, M Badawi, F Virot, S Cristol, L Cantrel, and J-F Paul. Influence of water, dihydrogen and dioxygen on the stability of the $Cr_2O_3$ surface: A first-principles investigation. *Surface Science*, 666:44–52, 2017.

[71] M Digne, P Sautet, P Raybaud, P Euzen, and H Toulhoat. Hydroxyl groups on $\gamma$-alumina surfaces: a DFT study. *Journal of Catalysis*, 211(1):1–5, 2002.

[72] HN Po and NM Senozan. The henderson-hasselbalch equation: its history and limitations. *Journal of Chemical Education*, 78(11):1499, 2001.

[73] R Rahnemaie, T Hiemstra, and WH van Riemsdijk. Inner-and outer-sphere complexation of ions at the goethite–solution interface. *Journal of Colloid and Interface Science*, 297(2):379–388, 2006.

[74] JO Nriagu. Lead orthophosphates—iv formation and stability in the environment. *Geochimica et cosmochimica acta*, 38(6):887–898, 1974.

[75] LJ Evans. Chemistry of metal retention by soils. *Environmental Science & Technology*, 23(9):1046–1056, 1989.

[76] A Davis, MV Ruby, and PD Bergstrom. Bioavailability of arsenic and lead in soils from the butte, montana, mining district. *Environmental Science & Technology*, 26(3):461–468, 1992.

[77] A Kabata-Pendias. Behavioural properties of trace metals in soils. *Applied geochemistry*, 8:3–9, 1993.

[78] W Stumm and R Wollast. Coordination chemistry of weathering: Kinetics of the surface-controlled dissolution of oxide minerals. *Reviews of Geophysics*, 28(1):53–69, 1990.

[79] DG Strawn. Sorption mechanisms of chemicals in soils. *Soil Systems*, 5(1):13, 2021.

[80] RL Parfitt, DJ Giltrap, and JS Whitton. Contribution of organic matter and clay minerals to the cation exchange capacity of soils. *Communications in soil science and plant analysis*, 26(9-10):1343–1355, 1995.

[81] W Stumm. The inner-sphere surface complex: A key to understanding surface reactivity. 1995.

[82] DM McKnight, KE Bencala, GW Zellweger, GR Aiken, GL Feder, and KA Thorn. Sorption of dissolved organic carbon by hydrous aluminum and iron oxides occurring at the confluence of deer creek with the snake river, summit county, colorado. *Environmental science & technology*, 26(7):1388–1396, 1992.

[83] MPW Schneider, T Scheel, R Mikutta, P Van Hees, K Kaiser, and K Kalbitz. Sorptive stabilization of organic matter by amorphous Al hydroxide. *Geochimica et Cosmochimica Acta*, 74(5):1606–1619, 2010.

[84] K Lalonde, A Mucci, A Ouellet, and Y Gélinas. Preservation of organic matter in sediments promoted by iron. *Nature*, 483(7388):198–200, 2012.

[85] K Kaiser, G Guggenberger, L Haumaier, and W Zech. Dissolved organic matter sorption on sub soils and minerals studied by 13C-NMR and DRIFT spectroscopy. *European Journal of Soil Science*, 48(2):301–310, 1997.

[86] K Kaiser and G Guggenberger. The role of DOM sorption to mineral surfaces in the preservation of organic matter in soils. *Organic geochemistry*, 31(7-8):711–725, 2000.

[87] KC Hass, WF Schneider, A Curioni, and W Andreoni. The chemistry of water on alumina surfaces: Reaction dynamics from first principles. *science*, 282(5387):265–268, 1998.

[88] EG Jobbágy and RB Jackson. The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecological applications*, 10(2):423–436, 2000.

[89] H Snyder. The composition of humus. *Journal of the American Chemical Society*, 19(9):738–744, 1897.

[90] K Kaiser and G Guggenberger. Sorptive stabilization of organic matter by microporous goethite: sorption into small pores vs. surface complexation. *European Journal of Soil Science*, 58(1):45–59, 2007.

[91] E Tipping. The adsorption of aquatic humic substances by iron oxides. *Geochimica et cosmochimica acta*, 45(2):191–199, 1981.

[92] K Kaiser and W Zech. Natural organic matter sorption on different mineral surfaces studied by DRIFT spectroscopy. *Sciences of Soils*, 2, 1997.

[93] Taku Onishi. *Quantum computational chemistry*. Springer, Singapore, 2018.

[94] JA Pople. Quantum chemical models (nobel lecture). *Angewandte Chemie International Edition*, 38(13-14):1894–1902, 1999.

[95] E. Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev.*, 28:1049–1070, Dec 1926.

[96] M Born and R Oppenheimer. On the quantum theory of molecules. In *Quantum Chemistry: Classic Scientific Papers*, pages 1–24. World Scientific, 2000.

[97] F Wang, T Ziegler, E van Lenthe, S van Gisbergen, and EJ Baerends. The calculation of excitation energies based on the relativistic two-component zeroth-order regular approximation and time-dependent density-functional with full use of symmetry. *The Journal of chemical physics*, 122(20):204103, 2005.

[98] CD Sherrill. An introduction to Hartree-Fock molecular orbital theory. *School of Chemistry and Biochemistry Georgia Institute of Technology*, 2000.

[99] J Schütt and MC Böhm. Influence of the Pauli antisymmetry and exclusion principle on the electronic structure of $\pi$ systems—a combined green function Quantum Monte Carlo and configuration interaction study. *Physics Letters A*, 219(1-2):79–85, 1996.

[100] W Pauli. Über den einfluß der Geschwindigkeitsabhängigkeit der Elektronenmasse auf den Zeemaneffekt. *Zeitschrift für Physik*, 31(1):373–385, 1925.

[101] W Pauli. On the connexion between the completion of electron groups in an atom with the complex structure of spectra. *Zeitschrift für Physik*, 31:765, 1925.

[102] JC Slater. The theory of complex spectra. *Physical Review*, 34(10):1293, 1929.

[103] PAM Dirac. *A new notation for quantum mechanics*, volume 35. Cambridge University Press, Cambridge, UK, 1939.

[104] JC Slater. A simplification of the hartree-fock method. *Physical review*, 81(3):385, 1951.

[105] AV Turbiner. Perturbation theory and variation principle in quantum mechanics. *Zh. Eksp. Teor. Fiz*, 79:1719, 1980.

[106] F Jensen. *Introduction to computational chemistry*. John wiley & sons, Sussex, 2017.

[107] CCJ Roothaan. New developments in molecular orbital theory. *Reviews of modern physics*, 23(2):69, 1951.

[108] GG Hall. The molecular orbital theory of chemical valency viii. a method of calculating ionization potentials. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 205(1083):541–552, 1951.

[109] JC Slater. Atomic shielding constants. *Physical Review*, 36(1):57, 1930.

[110] SF Boys. Electronic wave functions-i. a general method of calculation for the stationary states of any molecular system. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 200(1063):542–554, 1950.

[111] JC Christopher. *Essentials of computational chemistry: theories and models*. John Wiley & Son, Ltd, Sussex, 2004.

[112] SF Boys and FJMP Bernardi. The calculation of small molecular interactions by the differences of separate total energies. some procedures with reduced errors. *Molecular Physics*, 19(4):553–566, 1970.

[113] KN Kirschner, JB Sorensen, and JP Bowen. Calculating interaction energies using first principle theories: Consideration of basis set superposition error and fragment relaxation. *Journal of chemical education*, 84(7):1225, 2007.

[114] P Atkins and J De Paula. *Atkins' physical chemistry*. Oxford University Press, Oxford, UK, 2014.

[115] KA Baseden and JW Tye. Introduction to density functional theory: Calculations by hand on the helium atom. *Journal of Chemical Education*, 91(12):2116–2123, 2014.

[116] K Burke and LO Wagner. Dft in a nutshell. *International Journal of Quantum Chemistry*, 113(2):96–101, 2013.

[117] LH Thomas. The calculation of atomic fields. In *Mathematical proceedings of the Cambridge philosophical society*, volume 23, pages 542–548. Cambridge University Press, 1927.

[118] E Fermi. Eine statistische methode zur bestimmung einiger eigenschaften des atoms und ihre anwendung auf die theorie des periodischen systems der elemente. *Zeitschrift für Physik*, 48(1-2):73–79, 1928.

[119] RO Jones. Density functional theory: Its origins, rise to prominence, and future. *Reviews of modern physics*, 87(3):897, 2015.

[120] W Kohn and LJ Sham. doi: 10.1103/physrev. 140. a1133. *Phys. Rev. A*, 140:113, 1965.

[121] M Finnis. *Interatomic forces in condensed matter*, volume 1. OUP Oxford, 2003.

[122] P Hohenberg and W Kohn. Inhomogeneous electron gas. *Physical review*, 136(3B):B864, 1964.

[123] D Zhou. An introduction of density functional theory and its application. *Physics. Drexel. Edu*, 2007.

[124] JP Perdew, J Tao, VN Staroverov, and GE Scuseria. Meta-generalized gradient approximation: Explanation of a realistic nonempirical density functional. *The Journal of chemical physics*, 120(15):6898–6911, 2004.

[125] M Levy and JP Perdew. Hellmann-feynman, virial, and scaling requisites for the exact universal density functionals. shape of the correlation potential and diamagnetic susceptibility for atoms. *Physical Review A*, 32(4):2010, 1985.

[126] R Van Leeuwen and EJ Baerends. Exchange-correlation potential with correct asymptotic behavior. *Physical Review A*, 49(4):2421, 1994.

[127] JP Perdew, K Burke, and M Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.

[128] JP Perdew, JA Chevary, SH Vosko, KA Jackson, MR Pederson, DJ Singh, and C Fiolhais. Erratum: Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B*, 48(7):4978, 1993.

[129] AD Becke. Density-functional thermochemistry. iii. the role of exact exchange. *The Journal of chemical physics*, 98(7):5648–5652, 1993.

[130] BJKN Hammer and JK Nørskov. Electronic factors determining the reactivity of metal surfaces. *Surface Science*, 343(3):211–220, 1995.

[131] DR Hamann. Generalized gradient theory for silica phase transitions. *Physical Review Letters*, 76(4):660, 1996.

[132] PHT Philipsen, G Te Velde, and EJ Baerends. The effect of density-gradient corrections for a molecule-surface potential energy surface. slab calculations on Cu (100) c (2x2)-CO. *Chemical physics letters*, 226(5-6):583–588, 1994.

[133] V Ozoliņš and M Körling. Full-potential calculations using the generalized gradient approximation: Structural properties of transition metals. *Physical Review B*, 48(24):18304, 1993.

[134] C Filippi, DJ Singh, and CJ Umrigar. All-electron local-density and generalized-gradient calculations of the structural properties of semiconductors. *Physical Review B*, 50(20):14947, 1994.

[135] JP Perdew. Accurate density functional for the energy: Real-space cutoff of the gradient expansion for the exchange hole. *Physical Review Letters*, 55(16):1665, 1985.

[136] JP Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Physical Review B*, 33(12):8822, 1986.

[137] JP Perdew, P Ziesche, and H Eschrig. *Electronic structure of solids' 91*. Akademie Verlag, Berlin, 1991.

[138] JP Perdew, JA Chevary, SH Vosko, KA Jackson, MR Pederson, DJ Singh, and C Fiolhais. Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation. *Physical Review B*, 46(11):6671, 1992.

[139] AD Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical review A*, 38(6):3098, 1988.

[140] AD Becke and MR Roussel. Exchange holes in inhomogeneous systems: A coordinate-space model. *Physical Review A*, 39(8):3761, 1989.

[141] SK Ghosh and RG Parr. Phase-space approach to the exchange-energy functional of density-functional theory. *Physical Review A*, 34(2):785, 1986.

[142] JP Perdew, S Kurth, A Zupan, and P Blaha. Accurate density functional with correct formal properties: A step beyond the generalized gradient approximation. *Physical review letters*, 82(12):2544, 1999.

[143] J Tao, JP Perdew, VN Staroverov, and GE Scuseria. Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids. *Physical Review Letters*, 91(14):146401, 2003.

[144] AD Becke. A new mixing of Hartree-Fock and local density-functional theories. *The Journal of chemical physics*, 98(2):1372–1377, 1993.

[145] C Lee, W Yang, and RG Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Physical review B*, 37(2):785, 1988.

[146] PJ Stephens, FJ Devlin, CFN Chabalowski, and MJ Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *The Journal of Physical Chemistry*, 98(45):11623–11627, 1994.

[147] AD Becke. Density-functional thermochemistry. v. systematic optimization of exchange-correlation functionals. *The Journal of chemical physics*, 107(20):8554–8560, 1997.

[148] AV Arbuznikov. Hybrid exchange correlation functionals and potentials: Concept elaboration. *Journal of Structural Chemistry*, 48(1):S1–S31, 2007.

[149] S Grimme, S Ehrlich, and L Goerigk. Effect of the damping function in dispersion corrected density functional theory. *Journal of computational chemistry*, 32(7):1456–1465, 2011.

[150] IG Kaplan. *Intermolecular interactions: physical picture, computational methods and model potentials*. John Wiley & Sons, Hoboken, NJ, 2006.

[151] S Grimme, J Antony, T Schwabe, and C Mück-Lichtenfeld. Density functional theory with dispersion corrections for supramolecular structures, aggregates, and complexes of (bio) organic molecules. *Organic & Biomolecular Chemistry*, 5(5):741–758, 2007.

[152] T Schwabe and S Grimme. Theoretical thermodynamics for large molecules: walking the thin line between accuracy and computational cost. *Accounts of chemical research*, 41(4):569–579, 2008.

[153] R Huenerbein, B Schirmer, J Moellmann, and S Grimme. Effects of london dispersion on the isomerization reactions of large organic molecules: a density functional benchmark study. *Physical Chemistry Chemical Physics*, 12(26):6940–6948, 2010.

[154] T Schwabe, R Huenerbein, and S Grimme. Large molecules-small energies: Challenges for contemporary quantum chemistry. *Synlett*, 2010(10):1431–1441, 2010.

[155] S Grimme. Accurate description of van der waals complexes by density functional theory including empirical corrections. *Journal of computational chemistry*, 25(12):1463–1473, 2004.

[156] J-D Chai and M Head-Gordon. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Physical Chemistry Chemical Physics*, 10(44):6615–6620, 2008.

[157] Q Wu and W Yang. Empirical correction to density functional theory for van der waals interactions. *The Journal of chemical physics*, 116(2):515–524, 2002.

[158] S Grimme. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *Journal of computational chemistry*, 27(15):1787–1799, 2006.

[159] P Jurečka, J Černỳ, P Hobza, and DR Salahub. Density functional theory augmented with an empirical dispersion term. interaction energies and geometries of 80 noncovalent complexes compared with ab initio quantum mechanics calculations. *Journal of computational chemistry*, 28(2):555–569, 2007.

[160] S Grimme, J Antony, S Ehrlich, and H Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *The Journal of chemical physics*, 132(15):154104, 2010.

[161] J Kohanoff. *Electronic structure calculations for solids and molecules: theory and computational methods*. Cambridge university press, Cambridge, UK, 2006.

[162] RM Martin. *Electronic structure: basic theory and practical methods*. Cambridge university press, Cambridge, UK, 2020.

[163] R Dovesi, B Civalleri, C Roetti, VR Saunders, and R Orlando. Ab initio quantum simulation in solid state chemistry. *Reviews in computational chemistry*, 21:1–125, 2005.

[164] F Bloch. Quantum mechanics of electrons in crystal lattices. *Z. Phys*, 52:555–600, 1928.

[165] K Binder, J Horbach, W Kob, W Paul, and F Varnik. Molecular dynamics simulations. *Journal of Physics: Condensed Matter*, 16(5):S429, 2004.

[166] Kurt Binder and Giovanni Ciccotti. *Monte Carlo and Molecular Dynamics of Condensed Matter Systems: Euroconference on Computer Simulation in Condensed Matter Physics and Chemistry: Como, 3-28 July 1995*, volume 49. Compositori, 1996.

[167] SA Hollingsworth and RO Dror. Molecular dynamics simulation for all. *Neuron*, 99(6):1129–1143, 2018.

[168] J Polanski. Chemoinformatics in: Comprehensive chemometrics. *Polanski, J., Brown, SD, Tauler, R., Walczak, B., Eds*, pages 459–506, 2009.

[169] M Karplus and JA McCammon. Molecular dynamics simulations of biomolecules. *Nature structural biology*, 9(9):646–652, 2002.

[170] BJ Alder and TE Wainwright. Phase transition for a hard sphere system. *The Journal of chemical physics*, 27(5):1208–1209, 1957.

[171] T Hansson, C Oostenbrink, and WF van Gunsteren. Molecular dynamics simulations. *Current opinion in structural biology*, 12(2):190–196, 2002.

[172] A Rahman. Correlations in the motion of atoms in liquid argon. *Physical review*, 136(2A):A405, 1964.

[173] A Rahman. Liquid structure and self-diffusion. *The Journal of Chemical Physics*, 45(7):2585–2592, 1966.

[174] L Verlet. Computer" experiments" on classical fluids. i. thermodynamical properties of Lennard-Jones molecules. *Physical review*, 159(1):98, 1967.

[175] L Verlet. Computer experiments on classical fluids. ii. equilibrium correlation functions. *Physical Review*, 165(1):201, 1968.

[176] J Dummer. A simple time-corrected verlet integration method. *Game Developer*, 2004.

[177] RD Skeel. Variable step size destabilizes the Störmer/Leapfrog/Verlet method. *BIT Numerical Mathematics*, 33(1):172–175, 1993.

[178] S Nosé. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of chemical physics*, 81(1):511–519, 1984.

[179] WG Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Physical review A*, 31(3):1695, 1985.

[180] Alan Hinchliffe. *Molecular modelling for beginners*. John Wiley & Sons, Sussex, 2003.

[181] I Frank, D Marx, and M Parrinello. Structure and electronic properties of quinizarin chemisorbed on alumina. *The Journal of chemical physics*, 104(20):8143–8150, 1996.

[182] Z Łodziana, JK Nørskov, and P Stoltze. The stability of the hydroxylated (0001) surface of $\alpha$-$Al_2O_3$. *The Journal of chemical physics*, 118(24):11179–11188, 2003.

[183] V Shapovalov and TN Truong. Ab initio study of water adsorption on $\alpha$-$Al_2O_3$ (0001) crystal surface. *The Journal of Physical Chemistry B*, 104(42):9859–9863, 2000.

[184] LV Moskaleva, VA Nasluzov, Z-X Chen, and N Rösch. Elastic polarizable environment cluster embedding approach for water adsorption on the $\alpha$-$Al_2O_3$ (0001) surface. a density functional study. *Physical Chemistry Chemical Physics*, 6(18):4505–4513, 2004.

[185] S Alavi, DC Sorescu, and DL Thompson. Adsorption of HCl on single-crystal $\alpha$-$Al_2O_3$ (0001) surface: A DFT study. *The Journal of Physical Chemistry B*, 107(1):186–195, 2003.

[186] JM Wittbrodt, WL Hase, and HB Schlegel. Ab initio study of the interaction of water with cluster models of the aluminum terminated (0001) $\alpha$-aluminum oxide surface. *The Journal of Physical Chemistry B*, 102(34):6539–6548, 1998.

[187] K Johnston. A van der waals density functional study of the adsorption of ethanol on the $\alpha$-alumina (0001) surface. *Surface science*, 621:16–22, 2014.

[188] SD Chakarova-Käck, Ø Borck, E Schröder, and BI Lundqvist. Adsorption of phenol on graphite (0001) and $\alpha$-$Al_2O_3$ (0001): Nature of van der waals bonds from first-principles calculations. *Physical Review B*, 74(15):155402, 2006.

[189] I-C Yeh, JL Lenhart, and BC Rinderspacher. Molecular dynamics simulations of adsorption of catechol and related phenolic compounds to alumina surfaces. *The Journal of Physical Chemistry C*, 119(14):7721–7731, 2015.

[190] H Zhao, N Qi, and Y Li. Interaction between polysaccharide monomer and $SiO_2$/$Al_2O_3$/$CaCO_3$ surfaces: A DFT theoretical study. *Applied Surface Science*, 466:607–614, 2019.

[191] Y-Q Wang, X-F Yan, W Xiao, and Y-X Shao. DFT analysis of the adsorption of methyl nitrate on $Al_2O_3$ surfaces. *Bulletin of the Korean Chemical Society*, 38(6):625–631, 2017.

[192] Ø Borck, P Hyldgaard, and E Schröder. Adsorption of methylamine on $\alpha$-$Al_2O_3$ (0001) and $\alpha$-$Cr_2O_3$ (0001): Density functional theory. *Physical Review B*, 75(3):035403, 2007.

[193] S-Q Zhou, X-H Ju, F-Q Zhao, and S-Y Xu. Periodic DFT study of adsorption of nitroamine molecule on $\alpha$-$Al_2O_3$ (001) surface. *Applied surface science*, 258(19):7334–7342, 2012.

[194] MK Shukla, J Wang, and J Seiter. Understanding the fate of insensitive munitions compounds: Computational study on adsorption of nitroguanidine (NQ) and 1,1-diamino-2,2-dinitroethylene (FOX7) on Pristine and Al-hydroxylated $\alpha$-Alumina surfaces. *The Journal of Physical Chemistry C*, 121(21):11560–11567, 2017.

[195] M Ruan, H Hou, W Li, and B Wang. Theoretical study of the adsorption/dissociation reactions of formic acid on the $\alpha$-$Al_2O_3$ (0001) surface. *The Journal of Physical Chemistry C*, 118(36):20889–20898, 2014.

[196] S Blanck, C Martí, S Loehlé, SN Steinmann, and C Michel. (dis) similarities of adsorption of diverse functional groups over alumina and hematite depending on the surface state. *The Journal of Chemical Physics*, 154(8):084701, 2021.

[197] TD Kühne, M Iannuzzi, M Del Ben, VV Rybkin, P Seewald, F Stein, T Laino, RZ Khaliullin, O Schütt, F Schiffmann, D Golze, J Wilhelm, S Chulkov, MH Bani-Hashemian, V Weber, U Borštnik, M Taillefumier, AS Jakobovits, A Lazzaro, H Pabst, T Müller, R Schade, M Guidon, S Andermatt, N Holmberg, Gk Schenter, A Hehn, A Bussy, F Belleflamme, G Tabacchi, A Glöß, M Lass, I Bethune, CJ. Mundy, C Plessl, M Watkins, J VandeVondele, M Krack, and J Hutter. CP2K: An electronic structure and molecular dynamics software package - quickstep: Efficient and accurate electronic structure calculations. *The Journal of Chemical Physics*, 152(19):194103, 2020.

[198] J VandeVondele, M Krack, F Mohamed, M Parrinello, T Chassaing, and J Hutter. Quickstep: Fast and accurate density functional calculations using a mixed gaussian and plane waves approach. *Computer Physics Communications*, 167(2):103–128, 2005.

[199] J Hutter, M Iannuzzi, F Schiffmann, and J VandeVondele. CP2K: atomistic simulations of condensed matter systems. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4(1):15–25, 2014.

[200] JP Perdew, K Burke, and M Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 77(18):3865, 1996.

[201] JP Perdew, A Ruzsinszky, GI Csonka, OA Vydrov, GE Scuseria, LA Constantin, X Zhou, and K Burke. Restoring the density-gradient expansion for exchange in solids and surfaces. *Physical review letters*, 100(13):136406, 2008.

[202] Y Zhang and W Yang. Comment on "Generalized gradient approximation made simple". *Physical Review Letters*, 80(4):890, 1998.

[203] S Goedecker, M Teter, and J Hutter. Separable dual-space Gaussian pseudopotentials. *Physical Review B*, 54(3):1703, 1996.

[204] R Dovesi, R Orlando, A Erba, CM Zicovich-Wilson, B Civalleri, S Casassa, L Maschio, M Ferrabone, M De La Pierre, and P d'Arco. CRYSTAL14: A program for the ab initio investigation of crystalline solids, 2014.

[205] EF Pettersen, TD Goddard, CC Huang, GS Couch, DM Greenblatt, EC Meng, and TE Ferrin. UCSF Chimera a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 25(13):1605–1612, 2004.

[206] P Thissen, G Grundmeier, S Wippermann, and WG Schmidt. Water adsorption on the $\alpha$-$Al_2O_3$ (0001) surface. *Physical Review B*, 80(24):245403, 2009.

[207] B Wang, H Hou, Y Luo, Y Li, Y Zhao, and X Li. Density functional/all-electron basis set slab model calculations of the adsorption/dissociation mechanisms of water on $\alpha$-$Al_2O_3$ (0001) surface. *The Journal of Physical Chemistry C*, 115(27):13399–13411, 2011.

[208] Y-H Lu, S-Y Wu, and H-T Chen. $H_2O$ adsorption/dissociation and $H_2$ generation by the reaction of $H_2O$ with $Al_2O_3$ materials: a first-principles investigation. *The Journal of Physical Chemistry C*, 120(38):21561–21570, 2016.

[209] S Heiden, Y Yue, H Kirsch, J Wirth, P Saalfrank, and RK Campen. Water dissociative adsorption on $\alpha$-$Al_2O_3$ (1120) is controlled by surface site undercoordination, density, and topology. *The Journal of Physical Chemistry C*, 122(12):6573–6584, 2018.

[210] KC Hass, WF Schneider, A Curioni, and W Andreoni. First-principles molecular dynamics simulations of $H_2O$ on $\alpha$-$Al_2O_3$ (0001). *The Journal of Physical Chemistry B*, 104(23):5527–5540, 2000.

[211] JW Elam, CE Nelson, MA Cameron, MA Tolbert, and SM George. Adsorption of $H_2O$ on a single-crystal $\alpha$-$Al_2O_3$ (0001) surface. *The Journal of Physical Chemistry B*, 102(36):7008–7015, 1998.

[212] P Liu, T Kendelewicz, GE Brown Jr, EJ Nelson, and SA Chambers. Reaction of water vapor with $\alpha$-$Al_2O_3$ (0001) and $\alpha$-$Fe_2O_3$ (0001) surfaces: synchrotron X-ray photoemission studies and thermodynamic calculations. *Surface Science*, 417(1):53–65, 1998.

[213] R Lazzari and J Jupille. Wetting and interfacial chemistry of metallic films on the hydroxylated $\alpha$-$Al_2O_3$ (0001) surface. *Physical Review B*, 71(4):045409, 2005.

[214] R Lazzari and J Jupille. Chemical reaction via hydroxyl groups at the titanium/$\alpha$-$Al_2O_3$ (0001) interface. *Surface science*, 507:683–687, 2002.

[215] SA Chambers, T Droubay, Dr R Jennison, and TR Mattsson. Laminar growth of ultrathin metal films on metal oxides: Co on hydroxylated $\alpha$-$Al_2O_3$ (0001). *Science*, 297(5582):827–831, 2002.

[216] JA Kelber, Chengyu Niu, K Shepherd, DR Jennison, and A Bogicevic. Copper wetting of $\alpha$-$Al_2O_3$ (0001): theory and experiment. *Surface science*, 446(1-2):76–88, 2000.

[217] C Niu, K Shepherd, D Martini, J Tong, JA Kelber, DR Jennison, and A Bogicevic. Cu interactions with $\alpha$-$Al_2O_3$ (0001): effects of surface hydroxyl groups versus dehydroxylation by Ar-ion sputtering. *Surface science*, 465(1-2):163–176, 2000.

[218] Q Fu, T Wagner, and M Rühle. Hydroxylated $\alpha$-$Al_2O_3$ (001) surfaces and metal/$\alpha$-$Al_2O_3$ (0001) interfaces. *Surface science*, 600(21):4870–4877, 2006.

[219] JM McHale, A Auroux, AJ Perrotta, and A Navrotsky. Surface energies and thermodynamic phase stability in nanocrystalline aluminas. *Science*, 277(5327):788–791, 1997.

[220] JM McHale, A Navrotsky, and AJ Perrotta. Effects of increased surface area and chemisorbed $H_2O$ on the relative stability of nanocrystalline $\gamma$-$Al_2O_3$ and $\alpha$-$Al_2O_3$. *The Journal of Physical Chemistry B*, 101(4):603–613, 1997.

[221] HA Al-Abadleh and VH Grassian. FT-IR study of water adsorption on aluminum oxide surfaces. *Langmuir*, 19(2):341–347, 2003.

[222] MC Valero, M Digne, P Sautet, and P Raybaud. DFT study of the interaction of a single Palladium atom with-Alumina surfaces: the role of hydroxylation. *Oil & Gas Science and Technology-Revue de l'IFP*, 61(4):535–545, 2006.

[223] J Janeček, RR Netz, M Flörsheimer, R Klenze, B Schimmelpfennig, and R Polly. Influence of hydrogen bonding on the structure of the (001) corundum–water interface. density functional theory calculations and Monte Carlo simulations. *Langmuir*, 30(10):2722–2728, 2014.

[224] J Lutzenkirchen. Specific ion effects at two single-crystal planes of sapphire. *Langmuir*, 29(25):7726–7734, 2013.

[225] NW Assaf, M Altarawneh, I Oluwoye, M Radny, SM Lomnicki, and BZ Dlugogorski. Formation of environmentally persistent free radicals on $\alpha$-$Al_2O_3$. *Environmental science & technology*, 50(20):11094–11102, 2016.

[226] S-Y Ma, L-M Liu, and S-Q Wang. Water film adsorbed on the $\alpha$-$Al_2O_3$ (0001) surface: structural properties and dynamical behaviors from first-principles molecular dynamics simulations. *The Journal of Physical Chemistry C*, 120(10):5398–5409, 2016.

[227] T Calais, B Playe, J-M Ducéré, J-F Veyan, S Rupich, A Hemeryck, MD R, C Rossi, YJ Chabal, and A Estève. Role of alumina coatings for selective and controlled bonding of DNA on techno-

logically relevant oxide surfaces. *The Journal of Physical Chemistry C*, 119(41):23527–23543, 2015.

[228] X Cheng, Y Liu, and D Chen. Mechanisms of hydrolysis-oligomerization of aluminum alkoxide $Al(OC_3H_7)_3$. *The Journal of Physical Chemistry A*, 115(18):4719–4728, 2011.

[229] J Kolafa. Time-reversible always stable predictor-corrector method for molecular dynamics of polarizable molecules. *Journal of computational chemistry*, 25(3):335–342, 2004.

[230] S Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molecular physics*, 52(2):255–268, 1984.

[231] H Fatehi, DEL Ong, J Yu, and I Chang. Biopolymers as green binders for soil improvement in geotechnical applications: A review. *Geosciences*, 11(7):291, 2021.

[232] M Niaounakis. *Biopolymers: applications and trends*. William Andrew, Oxford, UK, 2015.

[233] R Wang, X Zhu, W Qian, Z Hong, H Tang, R Xu, and Y Yu. Pectin adsorption on amorphous Fe/Al hydroxides and its effect on surface charge properties and Cu (ii) adsorption. *Journal of Soils and Sediments*, 17(10):2481–2489, 2017.

[234] ML Ferreira and ME Gschaider. Theoretical and experimental study of $Pb^{2+}$ and $Hg^{2+}$ adsorption on biopolymers, theoretical study. *Macromolecular Bioscience*, 1(6):233–248, 2001.

[235] JB Wehr, FPC Blamey, PM Kopittke, and NW Menzies. Comparative hydrolysis and sorption of Al and La onto plant cell wall material and pectic materials. *Plant and soil*, 332(1):319–330, 2010.

[236] AGJ Voragen, G-J Coenen, RP Verhoef, and HA Schols. Pectin, a versatile polysaccharide present in plant cell walls. *Structural Chemistry*, 20(2):263–275, 2009.

[237] PT Martone, JM Estevez, F Lu, K Ruel, MW Denny, C Somerville, and J Ralph. Discovery of lignin in seaweed reveals convergent evolution of cell-wall architecture. *Current biology*, 19(2):169–175, 2009.

[238] D Elieh-Ali-Komi and MR Hamblin. Chitin and chitosan: production and application of versatile biomedical nanomaterials. *International journal of advanced research*, 4(3):411, 2016.

[239] Ilhan Chang, Jooyoung Im, and Gye-Chun Cho. Introduction of microbial biopolymers in soil treatment for future environmentally-friendly and sustainable geotechnical engineering. *Sustainability*, 8(3):251, 2016.

[240] D Kafetzopoulos, A Martinou, and V Bouriotis. Bioconversion of chitin to chitosan: purification and characterization of chitin deacetylase from mucor rouxii. *Proceedings of the National Academy of Sciences*, 90(7):2564–2568, 1993.

[241] S-I Aiba. Preparation of n-acetylchitooligosaccharides by hydrolysis of chitosan with chitinase followed by n-acetylation. *Carbohydrate research*, 265(2):323–328, 1994.

[242] K Tokuyasu, M Mitsutomi, I Yamaguchi, K Hayashi, and Y Mori. Recognition of chitooligosaccharides and their n-acetyl groups by putative subsites of chitin deacetylase from a deuteromycete, Colletotrichum lindemuthianum. *Biochemistry*, 39(30):8837–8843, 2000.

[243] I Younes and M Rinaudo. Chitin and chitosan preparation from marine sources. structure, properties and applications. *Marine drugs*, 13(3):1133–1174, 2015.

[244] DW Lee, C Lim, JN Israelachvili, and DS Hwang. Strong adhesion and cohesion of chitosan in aqueous solutions. *Langmuir*, 29(46):14222–14229, 2013.

[245] S-M Ham, I Chang, D-H Noh, T-H Kwon, and B Muhunthan. Improvement of surface erosion resistance of sand by microbial biopolymer formation. *Journal of Geotechnical and Geoenvironmental Engineering*, 144(7):06018004, 2018.

[246] D Costa, P-A Garrain, and M Baaden. Understanding small biomolecule-biomaterial interactions: A review of fundamental theoretical and experimental approaches for biomolecule interactions with inorganic surfaces. *Journal of Biomedical Materials Research Part A*, 101(4):1210–1222, 2013.

[247] A Soldo, M Miletić, and ML Auad. Biopolymers as a sustainable solution for the enhancement of soil mechanical properties. *Scientific reports*, 10(1):1–13, 2020.

[248] Michael P Schmidt and Carmen Enid Martínez. Supramolecular association impacts biomolecule adsorption onto goethite. *Environmental science & technology*, 52(7):4079–4089, 2018.

[249] M Kleber, P Sollins, and R Sutton. A conceptual model of organo-mineral interactions in soils: self-assembly of organic molecular fragments into zonal structures on mineral surfaces. *Biogeochemistry*, 85(1):9–24, 2007.

[250] RM Hazen and DA Sverjensky. Mineral surfaces, geochemical complexities, and the origins of life. *Cold Spring Harbor perspectives in biology*, 2(5):a002162, 2010.

[251] Z Lin, R Hu, J Zhou, Y Ye, Z Xu, and C Lin. A further insight into the adsorption mechanism of protein on hydroxyapatite by FTIR-ATR spectrometry. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 173:527–531, 2017.

[252] CT Johnston, GS Premachandra, T Szabo, J Lok, and RA Schoonheydt. Interaction of biological molecules with clay minerals: a combined spectroscopic and sorption study of lysozyme on saponite. *Langmuir*, 28(1):611–619, 2012.

[253] MP Schmidt and CE Martínez. Ironing out genes in the environment: an experimental study of the DNA-goethite interface. *Langmuir*, 33(34):8525–8532, 2017.

[254] BE Givens, Z Xu, J Fiegel, and VH Grassian. Bovine serum albumin adsorption on $SiO_2$ and $TiO_2$ nanoparticle surfaces at circumneutral and acidic pH: A tale of two nano-bio surface interactions. *Journal of colloid and interface science*, 493:334–341, 2017.

[255] MP Schmidt and CE Martínez. Kinetic and conformational insights of protein adsorption onto montmorillonite revealed using in situ ATR-FTIR/2D-COS. *Langmuir*, 32(31):7719–7729, 2016.

[256] M-A Thyveetil, PV Coveney, HC Greenwell, and JL Suter. Computer simulation study of the structural stability and materials properties of DNA-intercalated layered double hydroxides. *Journal of the American Chemical Society*, 130(14):4742–4756, 2008.

[257] K Kolman, MM Makowski, AA Golriz, M Kappl, J Pigłowski, H-J Butt, and A Kiersnowski. Adsorption, aggregation, and desorption of proteins on smectite particles. *Langmuir*, 30(39):11650–11659, 2014.

[258] X Liu, K Eusterhues, J Thieme, V Ciobota, C Höschen, CW Mueller, K Küsel, I Kögel-Knabner, P Rösch, and J Popp. STXM and NanoSIMS investigations on EPS fractions before and after adsorption to goethite. *Environmental science & technology*, 47(7):3158–3166, 2013.

[259] TH Nguyen and M Elimelech. Plasmid DNA adsorption on silica: kinetics and conformational changes in monovalent and divalent salts. *Biomacromolecules*, 8(1):24–32, 2007.

[260] P Cai, Q-Y Huang, and X-W Zhang. Interactions of dna with clay minerals and soil colloidal particles and protection against degradation by DNase. *Environmental science & technology*, 40(9):2971–2976, 2006.

[261] CA Haynes and W Norde. Globular proteins at solid/liquid interfaces. *Colloids and surfaces B: Biointerfaces*, 2(6):517–566, 1994.

[262] CW Hoogendam, A De Keizer, MA Cohen-Stuart, BH Bijsterbosch, JG Batelaan, and PM Van der Horst. Adsorption mechanisms of carboxymethyl cellulose on mineral surfaces. *Langmuir*, 14(14):3825–3839, 1998.

[263] WH Yu, N Li, DS Tong, CH Zhou, CXC Lin, and CY Xu. Adsorption of proteins and nucleic acids on clay minerals and their interactions: A review. *Applied Clay Science*, 80:443–452, 2013.

[264] P Mignon and M Sodupe. Theoretical study of the adsorption of DNA bases on the acidic external surface of montmorillonite. *Physical Chemistry Chemical Physics*, 14(2):945–954, 2012.

[265] E Paget, M Lebrun, G Freyssinet, and P Simonet. The fate of recombinant plant DNA in soil. *European Journal of Soil Biology*, 34(2):81–88, 1998.

[266] G Recorbet, C Picard, P Normand, and P Simonet. Kinetics of the persistence of chromosomal DNA from genetically engineered Escherichia coli introduced into soil. *Applied and Environmental Microbiology*, 59(12):4289–4294, 1993.

[267] F Widmer, RJ Seidler, and LS Watrud. Sensitive detection of transgenic plant marker gene persistence in soil microcosms. *Molecular Ecology*, 5(5):603–613, 1996.

[268] W Humphrey, A Dalke, and K Schulten. VMD: visual molecular dynamics. *Journal of molecular graphics*, 14(1):33–38, 1996.

[269] VM Rao, R Sanghvi, and HJ Zhu. Solubility of pharmaceutical solids. In *Developing solid oral dosage forms*, pages 1–24. Elsevier, 2009.

[270] R Ludwig. The effect of dispersion forces on the interaction energies and far infrared spectra of protic ionic liquids. *Physical Chemistry Chemical Physics*, 17(21):13790–13793, 2015.

[271] R Ludwigabc. The effect of dispersion forces on the interaction energies and far infrared spectra of protic ionic liquids1. *Phys. Chem. Chem. Phys*, 17:13790–13793, 2015.

[272] RAM Minke and J Blackwell. The structure of $\alpha$-chitin. *Journal of molecular biology*, 120(2):167–181, 1978.

[273] P Binu and M Mohan. Chitosan-based bionanocomposite in regenerative medicine. In *Bionanocomposites in Tissue Engineering and Regenerative Medicine*, pages 169–185. Elsevier, 2021.

[274] J Chen. Synthetic textile fibers: regenerated cellulose fibers. In *Textiles and Fashion*, pages 79–95. Elsevier, 2015.

[275] P Zatta. Recent topics in aluminium chemistry. *Coordination Chemistry Reviews*, 228(2), 2002.

[276] D Adhikari and Y Yang. Selective stabilization of aliphatic organic carbon by iron oxide. *Scientific reports*, 5(1):1–7, 2015.

[277] W Feng, AF Plante, AK Aufdenkampe, and J Six. Soil organic matter stability in organo-mineral complexes as a function of increasing C loading. *Soil Biology and Biochemistry*, 69:398–405, 2014.

[278] SM Henrichs. Sedimentary organic matter preservation: an assessment and speculative synthesis—a comment. *Marine Chemistry*, 49(2-3):127–136, 1995.

[279] S Amira, D Spångberg, V Zelin, M Probst, and K Hermansson. Car-Parrinello Molecular Dynamics simulation of $Fe^{3+}$ (aq). *The Journal of Physical Chemistry B*, 109(29):14235–14242, 2005.

[280] LA Curtiss, JW Halley, J Hautman, and A Rahman. Nonadditivity of abinitio pair potentials for molecular dynamics of multivalent transition metal ions in water. *The Journal of chemical physics*, 86(4):2319–2327, 1987.

[281] F Floris, M Persico, A Tani, and J Tomasi. Ab initio effective pair potentials for simulations of the liquid state, based on the polarizable continuum model of the solvent. *Chemical physics letters*, 199(6):518–524, 1992.

[282] CL Kneifel, HL Friedman, and MD Newton. Calculation of the thermodynamic solvent isotope effect for ferrous and ferric ions in water. *Zeitschrift für Naturforschung A*, 44(5):385–394, 1989.

[283] CL Kneifel, MD Newton, and HL Friedman. Simulation of solvent isotope effects on aqueous ferrous and ferric ions. *Journal of Molecular Liquids*, 60(1-3):107–145, 1994.

[284] L Dégrève and C Quintale Jr. The interfacial structure around ferric and ferrous ions in aqueous solution: the nature of the second hydration shell. *Journal of Electroanalytical Chemistry*, 409(1-2):25–31, 1996.

[285] S Amira, D Spångberg, M Probst, and K Hermansson. Molecular dynamics simulation of $Fe^{2+}$ (aq) and $Fe^{3+}$ (aq). *The Journal of Physical Chemistry B*, 108(1):496–502, 2004.

[286] JR Rustad, BP Hay, and JW Halley. Molecular dynamics simulation of iron (iii) and its hydrolysis products in aqueous solution. *The Journal of chemical physics*, 102(1):427–431, 1995.

[287] A González-Lafont, JM Lluch, A Oliva, and J Bertrán. A monte carlo simulation of free energy relationships for the electron transfer reaction between $Fe^{+}$ and $Fe^{2+}$ in water. *Journal of computational chemistry*, 12(10):1165–1171, 1991.

[288] RA Kuharski, JS Bader, D Chandler, M Sprik, ML Klein, and RW Impey. Molecular model for aqueous ferrous-ferric electron transfer. *The Journal of chemical physics*, 89(5):3248–3257, 1988.

[289] JS Bader and D Chandler. Computer simulation study of the mean forces between ferrous and ferric ions in water. *The Journal of Physical Chemistry*, 96(15):6423–6427, 1992.

[290] PV Kumar and BL Tembe. Solvation structure and dynamics of the $Fe^{2+}$-$Fe^{3}$+ ion pair in water. *The Journal of chemical physics*, 97(6):4356–4367, 1992.

[291] CS Babu, M Madhusoodanan, G Sridhar, and BL Tembe. Orientations of $[Fe(H_2O)_6]^{2+}$ and $[Fe(H_2O)_6]^{3+}$ complexes at a reactive separation in water. *Journal of the American Chemical Society*, 119(24):5679–5681, 1997.

[292] WH Casey. Large aqueous aluminum hydroxide molecules. *Chemical reviews*, 106(1):1–16, 2006.

[293] WH Casey, BL Phillips, and G Furrer. Aqueous aluminum polynuclear complexes and nanoclusters: a review. *Reviews in Mineralogy and Geochemistry*, 44(1):167–190, 2001.

[294] EJ Bylaska, M Valiev, JR Rustad, and JH Weare. Structure and dynamics of the hydration shells of the $Al^{3+}$ ion. *The Journal of chemical physics*, 126(10):104505, 2007.

[295] T Mossor-Pietraszewska. Effect of aluminium on plant growth and metabolism. *Acta Biochimica Polonica*, 48(3):673–686, 2001.

[296] DR Parker and PM Bertsch. Identification and quantification of the" Al13" tridecameric aluminum polycation using ferron. *Environmental Science & Technology*, 26(5):908–914, 1992.

[297] EG Bollard. Involvement of unusual elements in plant growth and nutrition. *Encyclopedia of plant physiology. New series*, 1983.

[298] SA Bogatko, EJ Bylaska, and JH Weare. First principles simulation of the bonding, vibrational, and electronic properties of the hydration shells of the high-spin $Fe^{3+}$ ion in aqueous solutions. *The Journal of Physical Chemistry A*, 114(5):2189–2200, 2010.

[299] DT Richens. *The chemistry of aqua ions: synthesis, structure, and reactivity: a tour through the periodic table of the elements*. Wiley New York, 1997.

[300] H Ohtaki and T Radnai. Structure and dynamics of hydrated ions. *Chemical reviews*, 93(3):1157–1204, 1993.

[301] M Magini, G Licheri, G Paschina, G Piccaluga, and G Pinna. *X-ray diffraction of ions in aqueous solutions: Hydration and complex formation*. CRC press, 2018.

[302] B-K Teo. Extended X-ray absorption fine structure (EXAFS) spectroscopy: techniques and applications. In *EXAFS Spectroscopy*, pages 13–58. Springer, Boston, MA, 1981.

[303] CF Baes and RE Mesmer. *The Hydrolysis of Cations. Iron*. Wiley, New York, 1976.

[304] Edward J King. Qualitative analysis and electrolytic solutions. 1959.

[305] F Basolo and RG Pearson. Mechanisms of inorganic reactions; a study of metal complexes in solution. 1967.

[306] P Hohenberg and WJPR Kohn. Density functional theory (DFT). *Phys. Rev*, 136:B864, 1964.

[307] RG Parr and W Yang. Density-functional theory of atoms and molecules. international series of monographs on chemistry. *Oxford University Press, New York*, 3:14312–14321, 1994.

[308] T Ziegler. Approximate density functional theory as a practical tool in molecular energetics and dynamics. *Chemical Reviews*, 91(5):651–667, 1991.

[309] FP Rotzinger. Treatment of substitution and rearrangement mechanisms of transition metal complexes with quantum chemical methods. *Chemical reviews*, 105(6):2003–2038, 2005.

[310] R Puchta, M Galle, N van Eikema Hommes, E Pasgreta, and R van Eldik. Evidence for associative ligand exchange processes on solvated lithium cations. *Inorganic chemistry*, 43(26):8227–8229, 2004.

[311] H Erras-Hanauer, T Clark, and R van Eldik. Molecular orbital and DFT studies on water exchange mechanisms of metal ions. *Coordination chemistry reviews*, 238:233–253, 2003.

[312] L Helm and AE Merbach. Inorganic and bioinorganic solvent exchange mechanisms. *Chemical reviews*, 105(6):1923–1960, 2005.

[313] DT Richens. Ligand substitution reactions at inorganic centers. *Chemical reviews*, 105(6):1961–2002, 2005.

[314] L Helm and AE Merbach. Water exchange on metal ions: experiments and simulations. *Coordination Chemistry Reviews*, 187(1):151–181, 1999.

[315] Th Kowall, P Caravan, H Bourgeois, L Helm, FP Rotzinger, and AE Merbach. Interpretation of activation volumes for water exchange reactions revisited: Ab i nitio calculations for $Al^{3+}$, $Ga^{3+}$, and $In^{3+}$, and new experimental data. *Journal of the American Chemical Society*, 120(26):6569–6577, 1998.

[316] V Migliorati and P D'Angelo. Unraveling the $Sc^{3+}$ hydration geometry: The strange case of the far-coordinated water molecule. *Inorganic chemistry*, 55(13):6703–6711, 2016.

[317] HVR Annapureddy and LX Dang. Understanding the rates and molecular mechanism of water-exchange around aqueous ions using molecular simulations. *The Journal of Physical Chemistry B*, 118(30):8917–8927, 2014.

[318] M Gupta, EF da Silva, and HF Svendsen. Explicit solvation shell model and continuum solvation models for solvation energy and pKa determination of amino acids. *Journal of chemical theory and computation*, 9(11):5021–5037, 2013.

[319] A Cusanelli, U Frey, DT Richens, and AE Merbach. The slowest water exchange at a homoleptic mononuclear metal center: variable-temperature and variable-pressure 17o NMR study on $[Ir(H_2O)_6]^{3+}$. *Journal of the American Chemical Society*, 118(22):5265–5271, 1996.

[320] D Hugi-Cleary, L Helm, and AE Merbach. Variable-temperature and variable-pressure 17o-NMR study of water exchange of hexaaquaaluminium (iii). *Helvetica chimica acta*, 68(3):545–554, 1985.

[321] JP Nordin, DJ Sullivan, BL Phillips, and WH Casey. An 17O-NMR study of the exchange of water on $[AlOH(H_2O)^{2+}$ (aq). *Inorganic chemistry*, 37(19):4760–4763, 1998.

[322] JD Kubicki, D Sykes, and SE Apitz. Ab initio calculation of aqueous aluminum and aluminum-carboxylate complex energetics and 27Al NMR chemical shifts. *The Journal of Physical Chemistry A*, 103(7):903–915, 1999.

[323] P MacCarthy and IH Suffet. *Aquatic humic substances and their influence on the fate and treatment of pollutants*, volume 219. Washington DC, 1989.

[324] NV Hue, GR Craddock, and Fred Adams. Effect of organic acids on aluminum toxicity in subsoils. *Soil Science Society of America Journal*, 50(1):28–34, 1986.

[325] SE Herbes and L R Schwall. Microbial transformation of polycyclic aromatic hydrocarbons in pristine and petroleum-contaminated sediments. *Applied and Environmental Microbiology*, 35(2):306–316, 1978.

[326] W Furrer, Gand Stumm. The coordination chemistry of weathering: I. dissolution kinetics of $\delta$-$Al_2O_3$ and BeO. *Geochimica et Cosmochimica Acta*, 50(9):1847–1860, 1986.

[327] SA Welch and WJ Ullman. The effect of organic acids on plagioclase dissolution rates and stoichiometry. *Geochimica et Cosmochimica Acta*, 57(12):2725–2736, 1993.

[328] JB Fein and JE Hestrin. Experimental studies of oxalate complexation at 80 c: Gibbsite, amorphous silica, and quartz solubilities in oxalate-bearing fluids. *Geochimica et Cosmochimica Acta*, 58(22):4817–4829, 1994.

[329] DD Eberl. Three zones for illite formation during burial diagenesis and metamorphism. *Clays and Clay Minerals*, 41(1):26–37, 1993.

[330] JB Fein. Porosity enhancement during clastic diagenesis as a result of aqueous metal-carboxylate complexation: Experimental studies. *Chemical Geology*, 115(3-4):263–279, 1994.

[331] MJ Frisch, GW Trucks, HB Schlegel, GE Scuseria, MA Robb, JR Cheeseman, G Scalmani, V Barone, GA Petersson, H Nakatsuji, X Li, M Caricato, AV Marenich, J Bloino, BG Janesko, R Gomperts, B Mennucci, HP Hratchian, JV Ortiz, AF Izmaylov, JL Sonnenberg, D Williams-Young, F Ding, F Lipparini, F Egidi, J Goings, B Peng, A Petrone, T Henderson, D Ranasinghe, VG Zakrzewski, J Gao, N Rega, G Zheng, W Liang, M Hada, M Ehara, K Toyota, R Fukuda, J Hasegawa, M Ishida, T Nakajima, Y Honda, O Kitao, H Nakai, T Vreven, K Throssell, JA Montgomery, Jr., JE Peralta, F Ogliaro, MJ Bearpark, JJ Heyd, EN Brothers, KN Kudin, VN Staroverov, TA Keith, R Kobayashi, J Normand, K Raghavachari, AP Rendell, JC Burant, SS Iyengar, J Tomasi, M Cossi, JM Millam, M Klene, C Adamo, R Cammi, JW Ochterski, RL Martin, K Morokuma, O Farkas, JB Foresman, and DJ Fox. Gaussian 09, revision a.02. *Inc., Wallingford CT*, 121:150–166, 2009.

[332] AD Becke. Density-functional thermochemistry. i. the effect of the exchange-only gradient correction. *The Journal of chemical physics*, 96(3):2155–2160, 1992.

[333] S Dong, J Zhang, X Yuan, and S Bi. Density functional theory studies on the real and apparent water-exchange reaction kinetics of $Al^{3+}$ in aqueous solution. *ACS Earth and Space Chemistry*, 3(10):2315–2322, 2019.

[334] MS Caetano, TC Ramalho, DF Botrel, EFF da Cunha, and WC de Mello. Understanding the inactivation process of organophosphorus herbicides: A DFT study of glyphosate metallic complexes with Zn2+, $Ca^{2+}$, $Mg^{2+}$, $Cu^{2+}$, $Co^{3+}$, $Fe^{3+}$, $Cr^{3+}$, and $Al^{3+}$. *International Journal of Quantum Chemistry*, 112(15):2752–2762, 2012.

[335] JP Perdew, K Burke, and M Ernzerhof. Generalized gradient approximation made simple. *Physical review letters*, 78(7):1396–1396, 1997.

[336] TJ Giese and DM York. Density-functional expansion methods: Evaluation of LDA, GGA, and meta-GGA functionals and different integral approximations. *The Journal of chemical physics*, 133(24):244107, 2010.

[337] J Tomasi, B Mennucci, and R Cammi. Quantum mechanical continuum solvation models. *Chemical reviews*, 105(8):2999–3094, 2005.

[338] AJA Aquino, D Tunega, G Haberhauer, M Gerzabek, and H Lischka. A density functional theoretical study on solvated $Al^{3+}$–oxalate complexes: structures and thermodynamic properties. *Physical Chemistry Chemical Physics*, 2(13):2845–2850, 2000.

[339] E Wasserman, JR Rustad, and SS Xantheas. Interaction potential of $Al^{3+}$ in water from first principles calculations. *The Journal of chemical physics*, 106(23):9769–9780, 1997.

[340] MI Lubin, EJ Bylaska, and JH Weare. Ab initio molecular dynamics simulations of aluminum ion solvation in water clusters. *Chemical Physics Letters*, 322(6):447–453, 2000.

[341] SJ Hawkes. All positive ions give acid solutions in water. *Journal of chemical education*, 73(6):516, 1996.

[342] NY Dzade, A Roldan, and NH De Leeuw. A density functional theory study of the adsorption of benzene on hematite ($\alpha$-$Fe_2O_3$) surfaces. *Minerals*, 4(1):89–115, 2014.

[343] RM Cornell and U Schwertmann. *The iron oxides: structure, properties, reactions, occurrences and uses*. John Wiley & Sons, 2003.

[344] P Yue, N Chen, D Peak, NM Bompoti, M Chrysochoou, A Onnis-Hayden, and P Larese-Casanova. Oxygen atom release during selenium oxyanion adsorption on goethite and hematite. *Applied Geochemistry*, 117:104605, 2020.

[345] G Rollmann, A Rohrbach, P Entel, and J Hafner. First-principles calculation of the structure and magnetic phases of hematite. *Physical Review B*, 69(16):165107, 2004.

[346] R Marino-Fernandez, SH Masunaga, N Fontaina-Troitino, MP Morales, J Rivas, and V Salgueirino. Goethite ($\alpha$-FeOOH) nanorods as suitable antiferromagnetic substrates. *The Journal of Physical Chemistry C*, 115(29):13991–13999, 2011.

[347] J An, P Wanaguru, C Xia, M Tao, and Q Zhang. First-principles study of sulfur atom doping and adsorption on $\alpha$–$Fe_2O_3$ (0001) film. *Physics Letters A*, 380(38):3149–3154, 2016.

[348] B Gu, Ju Schmitt, Z Chen, L Liang, and JF McCarthy. Adsorption and desorption of natural organic matter on iron oxide: mechanisms and models. *Environmental science & technology*, 28(1):38–46, 1994.

[349] K Mesuere and W Fish. Chromate and oxalate adsorption on goethite. Surface complexation modeling of competitive adsorption. *Environmental science & technology*, 26(12):2365–2370, 1992.

[350] E Tipping and D Cooke. The effects of adsorbed humic substances on the surface charge of goethite ($\alpha$-FeOOH) in freshwaters. *Geochimica et Cosmochimica Acta*, 46(1):75–80, 1982.