TOWARD ACCURATE SIMULATION OF ELECTROCATALYZED WATER SPLITTING

Enhanced Quantum Chemical Dynamics Simulations of Proton and Electron Transfer Reactions

Ambuj Tiwari
Toward accurate simulation of electrocatalyzed water splitting
Enhanced quantum chemical dynamics simulations of proton and electron transfer reactions

Academisch Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus prof. dr. ir. K.I.J. Maex
ten overstaan van een door het College voor Promoties ingestelde commissie, in het openbaar te verdedigen in de Agnietenkapel

op maandag 3 december 2019, te 16.00 uur

doors

Ambuj Tiwari

geboren te Mainpuri
Promotiecommissie

Promotor:
Dr. ir. B. Ensing
Prof. dr. P.G. Bolhuis

Overige leden:
Prof. dr. E.J. Meijer
Prof. dr. S. Woutersen
Prof. dr. P.D. Iedema
Prof. dr. H.M. Cuppen
Dr. M. Ceriotti
Dr. M. Sulpizi
Dr. F. Buda

Universiteit van Amsterdam
Universiteit van Amsterdam
Universiteit van Amsterdam
Universiteit van Amsterdam
École polytechnique fédérale de Lausanne
Johannes Gutenberg-Universität Mainz
Universiteit Leiden

Faculteit der Natuurwetenschappen, Wiskunde en Informatica
This work is part of the Industrial Partnership Program (IPP) ‘Computational Science for Energy Research’ (Project 14CSER014) which is financially supported by the Netherlands Organization for Scientific Research (NWO). This research program is co-financed by Shell Global Solutions International B.V. This work was carried out at the Van ’t Hoff Institute for Molecular Sciences at the University of Amsterdam. SURFSARA, the Dutch national e-infrastructure is acknowledged for the use of supercomputing facilities.

Cover illustration: The cover takes inspiration from the famous fresco 'The Creation of Adam' by Michelangelo which illustrated God giving life to Adam, the first man. One interpretation of the cover is to imagine the hand of god as an anode, providing energy to split water. The proton hops through the solvent to create clean fuel H₂ at the cathode. This comparison aims to emphasize on the urgency of finding the cleaner sources of fuel for the continual survival of mankind.

Cover Design: Ambuj Tiwari, Thor van Heezch and Yuliia Orlova

Copyright © 2019 by Ambuj Tiwari
ISBN 978-94-6323-948-6
If, in some cataclysm, all of scientific knowledge were to be destroyed, and only one sentence passed on to the next generation of creatures, what statement would contain the most information in the fewest words? I believe it is the atomic hypothesis that all things are made of atoms, little particles that move around in perpetual motion, attracting each other when they are a little distance apart, but repelling upon being squeezed into one another. In that one sentence, you will see, there is an enormous amount of information about the world, if just a little imagination and thinking are applied.

Richard Feynman
Chapter 1

Introduction

With the ever-increasing world population and improved standard of living, especially in developing countries, the demand for energy is rising at an alarming pace. In the 20th century, coal and natural gas remained the primary sources to fulfill this energy demand. More and more people had electricity in their homes and cars that run on fossil fuels, but this led to unforeseen challenges, climate change and global warming. Gladly, albeit a bit late, the consequences were understood, and as a result, in the 21st century, almost every school science textbook emphasizes on the importance of renewable energy. Significant sources of renewable energy are solar energy, wind energy, and hydroelectric power. However, all of these sources are intermittent i.e., their production depends on several uncontrollable natural factors. Solar power needs daylight and a clear sky, wind energy depends on wind strength, and hydroelectricity depends on the seasonal nature of rivers.

The current strategy to deal with this irregular supply from renewables is to tune the production from the coal plants accordingly. However, this makes these coal plants run inefficiently and also keeps our dependency on them during a shortage of renewable energy. One possible solution is to store the excess electricity generated by renewable sources during peak production in the form of chemical bonds and use it whenever required at a later point of time. One of the most promising approaches in this category is to split water using photo/electrolysis to produce hydrogen. This hydrogen can be used as a clean fuel afterward.

1.1 Water splitting

Decomposing water to produce O$_2$ and H$_2$ is known as water splitting:

$$2\text{H}_2\text{O} \rightarrow 2\text{H}_2 + \text{O}_2$$  \hspace{1cm} (1.1)

It is an endothermic process and requires external energy which can be in the form
1. Introduction

of light, heat, electricity, etcetera. Water splitting consists of two half-reactions: the oxygen evolution reaction (OER) at the anode:

\[ 2H_2O(l) \rightarrow O_2(g) + 4H^+(aq) + 4e^- \]  \hspace{1cm} (1.2)

and the hydrogen evolution reaction (HER) at the cathode:

\[ 2H^+(aq) + 2e^- \rightarrow H_2(g) \]  \hspace{1cm} (1.3)

An illustration for such an electrolyzer is shown here:

\[ \text{Figure 1.1: Schematic illustration of an electrolytic water splitting cell. OER occurs at the anode, generated protons diffuses to cathode to participate in the HER through a proton exchange membrane} \]

1.1.1 Basic principle

The Minimum amount of energy required for the water splitting reaction can be calculated using the Gibbs free energy equation:

\[ \Delta G^o = \Delta H^o - T \Delta S^o \]  \hspace{1cm} (1.4)
1.1. Water splitting

where $\Delta H^o$ (286 kJ mol$^{-1}$) and $\Delta S^o$ (0.16 kJ mol$^{-1}$K$^{-1}$) are standard enthalpy of formation and ideal gas entropy of gaseous water respectively at 298 K. Substituting these values in Eqn. 1.4 gives $\Delta G^o = 237$ kJ mol$^{-1}$ or roughly 1.23 eV. It is the minimum thermodynamic energy required to split water to be supplied in the form of a potential difference between cathode and anode of the electrochemical cell.

However, since it is an activated process, there are kinetic barriers. Hence, some excess potential difference is required between the cathode and anode to cross these free energy barriers. This excess potential is known as overpotential and causes a loss of energy. A catalyst can lower kinetic barriers, however simple in theory, achieving a low overpotential catalyst is a formidable task spanning constant research efforts for more than four decades. Several plausible reasons are identified for the overpotential such as charge carrier recombination, side-reactions, the electrical double layer, production of high-energy intermediates, non-linear response effects in the inner-sphere or solvent reorganization, and so forth.

Without applying the overpotential, even though thermodynamically feasible, the water splitting reaction usually has a prohibitively slow reaction rate.

1.1.2 Mechanism of water splitting on RuO$_2$

A catalyst reduces the aforementioned overpotential by accessing different reaction pathways with lower activation barriers. Since a catalyst does not take part in the overall reaction, it usually influences the rate by stabilizing the temporary intermediates. Usually, the catalyst can be categorized as homogeneous or heterogeneous. As the name suggests, homogeneous catalysts are in the same phase as the reactants, whereas heterogeneous catalysts are in a different phase. Most heterogeneous catalysts are in the solid phase acting on liquid or gaseous reactants and thus often are more stable than the homogeneous catalysts.

For electrolytic water splitting the role of the catalyst can be combined with that of an electrode, thus making it electrocatalytic water splitting. For the water splitting reaction, RuO$_2$ has been proposed to be one of the most active electrode materials. However, there is still significant energy loss which is associated with the OER at the RuO$_2$ (110) surface:

The OER is proposed to undergo via a four step mechanism:

$$\ast + H_2O \rightarrow \ast OH + H^+ + e^-$$ (1.5)
$$\ast OH \rightarrow \ast O + H^+ + e^-$$ (1.6)
$$\ast O + H_2O \rightarrow \ast OOH + H^+ + e^-$$ (1.7)
$$\ast OOH \rightarrow \ast + O_2 + H^+ + e^-$$ (1.8)

where $\ast$ represents an active site on the RuO$_2$ surface. Each of these four steps involves a proton coupled electron transfer (PCET) reaction. Even though the equations for OER look straightforward, some chemical restrictions make it challenging to optimize the catalyst. The most notorious of these is the scaling relationship:
Figure 1.2: Gibbs free energies of assumed reaction intermediates of the oxygen evolution reaction (OER) on a RuO$_2$ (110) surface. The Gibbs free energies are depicted for three different electrode potentials: 0 V, 1.23 V (equilibrium potential of the overall reaction) and 1.60 V vs standard hydrogen electrode (SHE). At 1.60 V vs SHE the OER becomes thermodynamically feasible. Image reproduced from Rossmeisl et al. [Copyright 2007, Elsevier].

The scaling relationship$^3$ dictates that the adsorption energy of the intermediates shown in Fig. 1.2 scales with one another. This is rather easy to understand, if an intermediate binds too strongly to the surface then it will need a lot of energy to desorb in the next step. The scaling relationship makes it necessary to find an ideal material that optimizes the water splitting scheme as shown in Fig. 1.2. However, to do so, it is vital to identify the intermediate species and calculate their relative stabilities in realistic conditions. In the next section, we will have a detailed look at the current challenges of the theory that tackles the PCET reactions.

1.2 Proton coupled electron transfer reactions

Proton coupled electron transfer (PCET) processes are ubiquitous. The most characteristic of which is the photosynthesis process which converts light energy into chemical energy as follows:

$$\text{CO}_2 + 2\text{H}_2\text{O} \xrightarrow{\text{hv}} \text{CH}_2\text{O} + \text{O}_2 + \text{H}_2\text{O} \quad \text{(1.9)}$$

and the corresponding half reactions are:

$$2\text{H}_2\text{O} \longrightarrow \text{O}_2 + 4\text{H}^* \quad \text{(1.10)}$$

$$4\text{H}^* + \text{CO}_2 \longrightarrow \text{CH}_2\text{O} + \text{H}_2\text{O} \quad \text{(1.11)}$$
1.2. Proton coupled electron transfer reactions

This simple process not only provides plant energy for their functions but also acts as the lungs of our planet. We have long learned and mimicked from nature to solve technological challenges. The chemical equation 1.10 also served as the motivation for storing the excess electricity into chemical bonds as a clean fuel and has been the basis of decades of research for energy storage and carbon capture.

Although initially specific to the concerted proton and electron transfer, nowadays, PCET stands for any chemical reaction which involves electron and proton transfer in their half-reactions, irrespective of their mechanism. This is especially the case in electrochemical water splitting where the nature of proton and electron transfer has yet to be concluded. However, the reaction rates are shown to depend on the mechanism of the proton and electron transfer, e.g. in case of concerted proton-electron transfer (CPET) by avoiding charge separation. Hence, it is crucial and remains an open challenge to develop new models to understand these fundamental electrochemical processes.

1.2.1 Electron transfer reactions

Electron transfer involves the rearrangement of electrons from a donor to an acceptor species. It is arguably one of the most fundamental processes, occurring all around us in nature, such as photosynthesis, respiration, metal corrosion as well as in human conceived technologies in batteries, solar cells, semiconductors, etc. The electron transfer reaction can be further divided in two sub-categories: inner sphere and outer sphere electron transfer.

**Inner sphere**

Inner sphere electron transfer occurs via covalently bonded bridging atoms/ligands between a donor and an acceptor. A typical example of such a reaction is:

\[
[\text{CoCl(NH}_3\text{)}_5]^{2+} + [\text{Cr(H}_2\text{O)}_6]^{2+} \rightarrow [\text{Co(NH}_3\text{)}_5(\text{H}_2\text{O})]^{2+} + [\text{CrCl(H}_2\text{O)}_5]^{2+}
\] (1.12)

By using radioactive \(\text{Cl}^-\), it was shown that all the \(\text{Cl}^-\) transferred to \(\text{Cr(III)}\) came from \(\text{Co}\) through a formation of an intermediate bimetallic complex.\(^4\) However, most of the biological electron reaction occurs via the second mechanism called outer sphere electron transfer.
Outer sphere

Outer sphere electron transfer, on the other hand, does not involve a rearrangement of bonds and is more prevalent in nature. Therefore, it has been used as a toy system for understanding electron transfer reactions. One such extensively studied model is the self-exchange reaction between two metal ions in aqueous solution:

\[ MN^+ + M^{(n+1)+} \rightarrow M^{(n+1)+} + M^{n+} \]  

(1.13)

where an electron is transferred from one metal ion to another, exchanging their oxidation states as well. This simple reaction presented a lot of questions: What is the driving force behind the electron transfer? What is the mechanism of this electron transfer? How does it depend on the nature of the metal ion?

In 1952, Rudolph A. Marcus proposed a theory to understand this electron transfer reaction.\(^5\) He suggested that the electron transfer is caused by the solvent reorganization and proposed the diabolic Marcus plot, which was confirmed by several molecular dynamics (MD) and Monte Carlo simulations. However, due to an activation barrier, direct sampling of this electron transfer process cannot be achieved during the accessible timescales by the current electronic structure methods, thus making it an open research question for computational or theoretical investigations to unravel the mechanistic details of electron transfer.

1.2.2 Proton transfer reactions

Proton transfer occurs in a myriad of chemical and biological reactions essential to our survival. The mechanism of proton transfer was first proposed by Grotthuss in 1806, while studying the decomposition of water by passing an electrical current.\(^6\) In analogy with positive and negative electrodes proposed by his mentor Volta,\(^7\) he suggested that water also has a negative and a positive pole and is attracted by the electrodes. He further suggested that these fragments move along chains of water molecules until they form hydrogen and oxygen gases at the cathode and anode respectively. Interestingly, when Grotthuss came up with this idea, the exact composition of water was not discovered so according to his theory, both oxygens and protons were hopping. Now, two centuries later, we know it is only protons that hop around in the solvent by constantly making and breaking new chemical bonds with the neighboring molecules leading to proton conduction as shown by Fig. 1.3
Proton conduction has been studied extensively, both experimentally and theoretically, confirming a Grotthuss-like mechanism. However, studying water dissociation is much more difficult as it involves kinetic barriers and reaction free energies that are much higher than those for proton conduction in water. The ease or difficulty of proton-donor dissociation, e.g. in an acid, is quantified by the acidity constant, or $pK_a$, of the donor species.

Modeling proton dissociation and calculating $pK_a$ is again challenging because the chemical bond breaking and making require an accurate quantum mechanical electronic structure approach, while at the same time, an explicit treatment of the solvent involves the sampling of many molecular configurations. This makes direct sampling of proton dissociation rather improbable in the short timescale accessible during quantum chemistry based MD simulations.

1.3 Computational chemistry

The urge to make theoretical models to understand the physical world dates back at least two millennia when several philosophers around the globe devoted their lives to understand the nature of matter. The idea that matter is made up of individual particles was proposed in several cultures such as in Greece by Leucippus together with his student Democritus and in India by Kanada who suggested every object of creation is made of atoms (parmanu) which in turn connect with each other to form molecules (anu). It took more than two millennia to improve
this idea further by a series of development by Faraday (discovery of cathode rays), Kirchhoff (black body radiation), Boltzmann (statistical mechanics), Hertz (photoelectric effect), Max Planck (discrete energy states) and finally Albert Einstein (quantization of energy) leading to the advent of the quantum mechanics which started a new era in theoretical chemistry. Computational chemistry implements these fundamental theoretical models into efficient algorithms to study various chemical and physical aspects of the world.

Despite being constructed to explain the abstract world of atoms, the impact of quantum mechanics on daily life is revolutionary. The fierce development of quantum mechanics was supplemented by the genius of Kohn, Sham, Hohenberg and numerous others who contributed to the development of one of the most used quantum chemistry models, density functional theory (DFT). Finally, John Pople developed the first general-purpose quantum chemistry package, Gaussian, that can accurately estimate properties of molecules and predict chemical reactions. Gaussian package survives to this day and is also used to perform chemical calculations in this thesis. In the last 50 years, two out of the five most cited physicists are Perdew and Becke who are actively contributing to further development of DFT and will be discussed in detail in the methods section. Overall, DFT has emerged as the decisive tool to understand molecular reactions.

A similar revolution is underway now in materials discovery where comparisons with experiments and systematic predictions are made across a broad range of applications. During chemical reactions, atoms move around and have access to multiple chemical pathways which are difficult to capture by single point DFT calculations. This problem is solved by combining DFT with molecular dynamics (DFT-MD) together with the assumption that electrons move much faster than nuclei (Born-Oppenheimer approximation).

Our model

We employed the DFT-MD method which efficiently propagates the system in time by solving classical Newtonian equations of motion. The forces used by these equations are calculated on-the-fly with quantum mechanical calculations of the electronic structure. The method of choice for electronic structure calculation is DFT for this thesis because of the computational efficiency of the method.

Despite being computationally expensive, DFT-MD has numerous advantages: First, the atomic forces are calculated using electronic structure calculations, instead of empirical forcefields, thus DFT-MD can predict chemical reactions within molecular dynamics. Second, along with the dynamics of the system, information about the electronic structure of the atoms is also present, making it ideal for electron transfer reactions. Third, reactions that require the participation of the solvent molecules can also be tackled, in our case specifically, water molecules thus

---

1Walter Kohn and John Pople were awarded the 1998 Nobel Prize in chemistry for their contributions in the development of theory and implementation respectively
1.4 Thesis outline

Efficient water splitting is an important step towards making a sustainable hydrogen based economy. Despite several breakthroughs, there is still a lot of energy loss while conducting water splitting due to the overpotential. To unravel the molecular origins of this overpotential, we need to employ atomistic simulations that can also guide the experiments and predict novel materials. The main scope of this thesis is to develop a robust framework that can be utilized to study proton coupled electron transfer reactions occurring during the water splitting reaction. Furthermore, we apply these models to the RuO$_2$ (110) surface to unravel the mechanism of the water splitting reaction. This is achieved by using DFT-MD simulations along with various enhanced sampling techniques.

In the next chapter, we introduce the theoretical background behind all the methods and techniques used throughout this thesis, including DFT, ab-initio molecular dynamics, a brief review on statistical mechanics, transition path sampling, metadynamics, and path-metadynamics.

In chapter 3, we perform an extensive benchmark simulation of the liquid water for DFT-MD simulations. We show that the properties of liquid water calculated...
1. Introduction

from DFT-MD simulations are quite sensitive to the choice of parameters such as the XC functional, basis set, Van der Waals corrections, smoothing functions and SCF convergence criteria. Moreover, we suggest a new method to efficiently approximate the simulation box size of the system to perform the NVT simulations, when the experimental density is unknown.

In chapter 4, we develop a new framework to simulate electron transfer reactions by combining DFT-MD simulations with transition path sampling. We show that the simple self-exchange reaction between two metal ions is coupled to a proton transfer. We further show that this electron transfer reaction is caused by solvent reorganization as initially predicted by Marcus.

In chapter 5, we develop a new proton transfer collective variable (PTCV) to study the proton dissociation process. We show that the PTCV can be used to study the acid dissociation process, accurately predicting the \( pK_a \) across the whole acid strength spectrum. Moreover, we shed light on the different acid dissociation mechanisms. We propose two different types of mechanism: a slow stepwise mechanism, and a concerted long range mechanism. We elucidate the molecular origin of these two different mechanisms by combining the \( pK_a \) results with DFT-MD equilibration runs. We show that the different degrees of solvent reorganization play a decisive role in selecting one mechanism or the other.

The free energy calculations performed using enhanced sampling methods do not include nuclear quantum effects such as zero point energy (ZPE). In the case of water, ZPE changes the \( pK_a \) of water from 7 to the \( pK_a \) of 7.43 for D\(_2\)O. Hence, in chapter 6, we use the 2PT method to calculate the zero point energy of various liquids and their mixtures by using molecular dynamics trajectories to calculate the density of states and from that the ZPE. We also perform a benchmark calculation to assess the speed and convergence of the 2PT method. We show that systems with 1000 molecules converge within a few ps, but single molecule systems of interest such as one acid molecules in the water solvent need at least 1000 ps to get a converged result. Unfortunately, in the current state of the art, this makes the method much more expensive than originally advertised, but still more economic than full blown path integral MD calculations.

In chapter 7, we study the water splitting reaction on the highly active RuO\(_2\) (110) surface. We start with benchmarking the DFT parameters to get good agreement with the literature adsorption energy. In the next step, we study various proton transfer processes occurring on the metal oxide-water interface. We show that the system reaches dynamical equilibrium and the water molecules dissociate spontaneously upon adsorption on the coordinatively unsaturated (CUS) Ru sites. We proceed by calculating the proton dissociation barrier for each step of the oxygen evolution reaction. We show that by considering the explicit solvent, the barrier for the rate determining step (O-O bond formation) of the OER decreases from 0.74 to 0.4 eV. Finally, this thesis ends with a summary of this work in English and in Dutch.
Quantum mechanics has played a pivotal role to study chemical reactions in a variety of systems. The challenge is always to get the electronic structure by solving the (time-independent) Schrödinger equation:

\[ \hat{H} \Psi(r_1, r_2, ..., r_N, R_1, R_2, R_M) = E \Psi(r_1, r_2, ..., r_N, R_1, R_2, R_M) \quad (2.1) \]

where \( \hat{H} \) is the hamiltonian, \( \Psi(r_1, r_2, ..., r_N, R_1, R_2, R_M) \) is the many-body wavefunction of a system having \( N \) electrons, \( M \) nuclei and \( E \) is the energy. According to the Born-Oppenheimer approximation, the nuclei are stationary with respect to the motion of the electrons; thus, a simplified \( \hat{H} \) can be expressed as:

\[ \hat{H} = -\frac{\hbar^2}{2m_e} \sum_i^N \nabla_i^2 - \sum_{i,l}^{N,M} Z_I e \frac{1}{|r_i - R_I|} + \sum_{i \neq j}^N \frac{e^2}{|r_i - r_j|} \quad (2.2) \]

where \( \hbar \) is the reduced Planck’s constant, \( m_e \) is the mass of an electron, \( \nabla \) is the gradient operator, \( Z_I \) and \( R_I \) are the atomic number and spatial coordinates of nucleus \( I \), \( r_i \) is the spatial coordinate of the electron \( i \) and \( e \) is the electronic charge. The first term of Eq. 2.2 is the electronic kinetic energy, the second term is the electron-nuclei Coulombic interaction energy, the third term is the electron-electron Coulombic interaction energy. Unfortunately, there are only very few systems where we can have the exact solution of Eq. 2.2, such as the particle in a box and the quantum harmonic oscillator. For most realistic systems consisting of multiple interacting nuclei and electrons, we rely on various forms of approximations. One of the most frequently used and computationally efficient solutions is Density Functional Theory (DFT).
2. Theory and Methods

2.1 Density functional theory

There are several theories out there fighting for the name “Density functional theory”, but with more than 10000 citations, \(^{16}\) it is safe to say that the title goes to the one we are going to discuss here. It is only fair to start by pointing out the complexity of the wavefunction for even a system as small as a single water molecule requires solving for a 30-dimensional function and would be computationally expensive. To circumvent the expensive calculation of wavefunctions, DFT was introduced. In 1964, Hohenberg and Kohn proposed two theorems:\(^{17}\)

**Theorem 1** The ground-state electron density \(\rho(r)\) is a unique functional of the external potential \(V_{\text{ext}}(r)\).

The first H-K Theorem implies that there is a one to one mapping between \(V_{\text{ext}}(r)\) and the electron density \(\rho(r)\) proposing a universal H-K functional. By introducing electron density, we reduced the \(4N\) dimensional problem (\(3N\) spatial and \(N\) spin) of finding the wavefunction of the \(N\) electron system to just three dimensions of the electronic density. Unfortunately, the exact form of the universal H-K functional is unknown but the second Hohenberg-Kohn theorem gives significant clues about finding one:

**Theorem 2** The H-K functional delivers the ground state energy if and only if the input density is the true ground state density.

This simple theorem proposes that a variational principle can be used to find the electron density, which gives the lowest ground state energy. One important thing to note is that collectively the two HK theorem reformulate the many-body electron problem to three dimensional electron density but do not actually solve it. This is where the Kohn-Sham ansatz comes to the rescue:

*The exact ground state electron density can be approximated as the ground state density of an imaginary system of non-interacting particles that have their own ground state energy functional.*

The Kohn-Sham ansatz leads to the formulation of the Kohn-Sham equation\(^ {18}\) which can be solved iteratively starting from an initial guess \(\rho(r)\)

\[
\left( \frac{-\hbar^2}{2m_e} \nabla^2 + v_{\text{ext}}(r) + e^2 \int \frac{\rho(r')}{|r-r'|} d^3r' + V_{\text{xc}}(r) \right) \psi_j = \epsilon_j \psi_j \tag{2.3}
\]

where \(\psi_j\) are the single electron Kohn-Sham orbitals and \(\epsilon_j\) is the corresponding eigenvalue. This equation looks similar to Eq. 2.2 except that the sum signs are gone as the Kohn-Sham equation is defined for each electron. The first term is the single-electron kinetic energy (of the non-interacting system), \(v_{\text{ext}}(r)\) is the external potential due to the nuclei (in the simplest case), the third term is
the Coulombic interaction, and $V_{xc}$ is the derivative of the exchange-correlation potential.

The $V_{xc}$ accounts for the missing exchange interaction and the electron correlation interaction. Starting from an initial guess density $\rho_{\text{guess}}(r)$ for the Coulombic electron-electron interaction term, solving the eigenvalue problem of equation 2.3 provides an improved set of orbitals and density $\rho(r)$. By reinserting the improved density in an iterative manner, convergence is reached when the input and output $\rho(r)$ is self-consistent.

The biggest open challenge in the field of DFT development is to find the exact exchange-correlation ($XC$) functional. Several approaches have been proposed to develop working approximate $XC$ functionals, including:

- **The local density approximation (LDA):** The functional depends on the local density at each point, most common examples are Cole-Perdew\textsuperscript{19} and PW92.\textsuperscript{20}

- **The Generalized Gradient Approximation (GGA):** The functional depends on the local density as well as its gradient, popular examples are BLYP\textsuperscript{21,22} and PBE.\textsuperscript{23}

- **Meta-GGA:** The functional depends on the local density, its gradient, and the second derivative. Recent versions may also include the electron kinetic energy density as an input. A popular example in material science is the SCAN functional\textsuperscript{24}.

- **Hybrid:** Includes a percentage of exact Hartree-Fock exchange, popular examples are B3LYP\textsuperscript{25,26} and PBE0.\textsuperscript{27}

LDA has performed surprisingly well for solid-state physics despite the very primitive uniform gas approximation. However, it generally fails to predict accurate energies and geometries for molecules due to the larger variations in the electronic density over a small region of space.\textsuperscript{28} By including these spatial density fluctuations via the density gradients, the introduction of GGAs provided a paramount improvement over the LDA functionals, which started the DFT revolution for chemical applications.

The search for the exact $XC$ functional has led to two key viewpoints. The first approach relies on developing an approximate form and fitting its parameters to experimental data of which BLYP and M06 functionals are the most popular example. The second school of thought focuses on incorporating exact constraints to improve the description of chemical systems. The PBE functional proposed by Perdew, Burke, and Ernzerhof falls in this category. The PBE functional (and its derivatives) has been consistently predictive for solid-state systems,\textsuperscript{29} whereas BLYP has been consistent in condensed phase systems.\textsuperscript{30}

One other issue that most of the GGA functionals face is the underestimation of the long-range dispersion interactions. Two empirical but efficient approaches to incorporate these interactions are based on employing interatomic dispersion
corrections. Grimme et al.\textsuperscript{31–33} proposed the first method that pre-calculates correction coefficients for various pairs of elements. Tkatchenko and Scheffler et al.\textsuperscript{34} proposed the second approach that performs Hirshfeld partitioning of the electron density to compare effective atomic volume and scale the coefficient, which was calculated for the free atom.

However, as shown in Chapter 3, proper benchmarking is crucial as the choice of \( \text{XC} \) functional as well as other DFT parameters have been shown to alter the results of a calculation significantly. Dependence of various DFT results on some of these parameters are compared in Chapter 3 and 7. Furthermore, going high in the so-called ”Jacob’s ladder” to achieve the heaven of chemical accuracy may not necessarily improve the accuracy despite being 10 to 100 times more expensive.

Despite being sensitive to the choice of parameters, DFT has been extensively useful in understanding chemical reactions and has also been the workhorse of this thesis. Hopefully, by the end of this thesis, it will be apparent that DFT is accurate enough to make powerful predictions about the electronic as well as structural properties of complex many-particle systems. It becomes even more useful when combined with Newton’s equation of motion to follow the time evolution of the atoms and molecules, as explained in the next section.

### 2.2 Ab-initio molecular dynamics

With the increased computational power, it has become almost trivial to study the time evolution of the system with Molecular Dynamics simulations where Newton’s equation of motion is solved on-the-fly to propagate the system in time. Most of these simulations use pre-defined empirical forcefields that employ simple mathematical functions to describe intramolecular interactions defining bonds, angles and torsions. These functions are parametrized to reproduce experimental data or high-level electronic structure calculations. This simple approach has been an enormous success and is able to describe many molecular processes, including involved transformations, such as protein folding, phase separation, nucleation etcetera.

However, most forcefields are not able to predict chemical properties accurately, mostly because of critical many-body interactions involved but also because of the fixed charges in these forcefield simulations. Moreover, most of the forcefields developed assume a fixed structure and bond order, thus making it impossible to study chemical reactions. With the fast developments of efficient, parallel algorithms and increased computational power, it is now possible to accurately and efficiently calculate these interactions for a large number of atoms using DFT calculations on-the-fly to calculate energies and forces on the atoms.

Let us consider a system with \( N \) atoms in a box of volume \( V \). The state of the system can be defined by \( 3N \) positions, \( (r_1, r_2, \ldots, r_{3N}) \), and \( 3N \) velocities, \( (v_1, v_2, \ldots, v_{3N}) \). The total kinetic energy of the system can be defined as:


\[ K = \frac{1}{2} \sum_{i=1}^{3N} m_i v_i^2 \]  

(2.4)

where \( m_i \) is the mass of the atom associated with index \( i \), and the total potential energy of the system can be defined as:

\[ U = U(r_1, r_2, ..., r_{3N}). \]  

(2.5)

Here, the total potential energy is a function of the position of all the atoms, \( (r_1, r_2, ..., r_{3N}) \). As discussed earlier, we can use DFT-MD to calculate the total potential energy on-the-fly. Also, we can obtain the force \( F_i \) on the atom \( i \),

\[ F_i = -\frac{\partial U}{\partial r_i}. \]  

(2.6)

Since we assume the motion of the atoms to be classical, Newton’s equations hold and allow us to calculate the acceleration for each particle,

\[ F_i = m_i a_i = m_i \frac{dv_i}{dt} \]  

(2.7)

with \( m_i \) being the mass of the atom \( i \) and \( a_i \) is the acceleration. Eq. 2.6 and 2.7 together form a system of \( 6N \) first-order differential equations. Solving these equations will provide us a time evolution of the system in space. Although for practical application, these equations of motion are too complicated to be exactly solved, fortunately, there are a lot of fast algorithms designed to solve such system of linear differential equations. A very robust and popular one used for MD simulation is called the Verlet algorithm, which can be derived easily from a Taylor expansion of the position as a function of time and substituting \( \Delta t \) with \( -\Delta t \) in Eq. 2.8,

\[ r_i(t + \Delta t) = r_i(t) + \frac{dr_i(t)}{dt} \Delta t + \frac{1}{2} \frac{d^2 r_i(t)}{dt^2} \Delta t^2 + \frac{1}{6} \frac{d^3 r_i(t)}{dt^3} \Delta t^3 + O(\Delta t^4) \]  

(2.8)

\[ r_i(t - \Delta t) = r_i(t) - \frac{dr_i(t)}{dt} \Delta t + \frac{1}{2} \frac{d^2 r_i(t)}{dt^2} \Delta t^2 - \frac{1}{6} \frac{d^3 r_i(t)}{dt^3} \Delta t^3 + O(\Delta t^4) \]  

(2.9)

Finally, adding Eq. 2.8 and 2.9 gives,

\[ r_i(t + \Delta t) = 2r_i(t) - r_i(t - \Delta t) + \frac{F_i(t)}{m_i} \Delta t^2. \]  

(2.10)

Note that if we have the spatial coordinates for the time \( t \) and \( t - \Delta t \), we can propagate the system in time using Eq. 2.10. As long as we chose a small enough timestep \( \Delta t \), Eq. 2.10 should be numerically accurate. If we are going to compare the simulation results with experiments, we need a consistent approach to obtain
macroscopic properties from atomistic DFT-MD simulations. One theory which is particularly useful in connecting the microscopic electronic description of the system to the macroscopic experimental observations is statistical mechanics, which is discussed in brief in the next section.

2.3 Statistical mechanics

Ab initio molecular dynamics is a very effective tool to study the microscopic dynamics of atoms, unravel reaction mechanisms, and provide simultaneously insight into the electronic structure of the system. However, most of the resulting DFT-MD properties, such as the instantaneous electron density, and the atomic positions and velocities cannot be compared directly to experimental observables. Moreover, most of the quantities measured experimentally are averaged over the number of particles as well as over time. If we want to compare our calculations with the experimental observables, we need some systematic approach of getting correct time and particle averages.

Statistical mechanics is the tool that bridges the gap between microscopic simulations and macroscopic observables. Let us again consider the system with $N$ atoms, defined by the positions $r^N$ and momenta $p^N$ of the atoms. Together, position and momenta define a $6N$ dimensional space, commonly called phase space. According to statistical mechanics, a macroscopic observable, $A(p^N, r^N)$, can be calculated from the ensemble average:

$$
\langle A \rangle_{\text{ensemble}} = \frac{\int \int dp^N dr^N A(p^N, r^N) \exp[-\beta \hat{H}(p^N, r^N)]}{\int \int dp^N dr^N \exp[-\beta \hat{H}(p^N, r^N)]}.
$$

(2.11)

The probability of finding the system in a state, $P(p^N, r^N)$, is

$$
P(p^N, r^N) = Q^{-1} \exp[-\beta \hat{H}(p^N, r^N)]
$$

(2.12)

where $Q$ is the partition function:

$$
Q = \int \int dp^N dr^N \exp[-\beta \hat{H}(p^N, r^N)].
$$

(2.13)

One of the most fundamental axioms of statistical mechanics, the ergodic hypothesis, suggests that this ensemble average is also equal to the time average leading to:

$$
\langle A \rangle_{\text{ensemble}} = \langle A \rangle_{\text{time}} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A(p^N, r^N) dt
$$

(2.14)

In principle, this allows us to employ DFT-MD to calculate various macroscopic properties. However, calculating experimentally relevant information, involves sampling a transition state configuration with high energy. The sampling of high
2.4. Rare event sampling methods

Many of the interesting physical and chemical transformations occur via high energy transition states. DFT-MD alone cannot be used to study these chemical transformations where two stable states are separated by energy barriers that are significantly higher than the thermal fluctuations ($k_B T$). The probability of sampling the high energy transition states decreases exponentially with the height of the energy barriers. To alleviate the exponential time scale issue some techniques are required to enhance the sampling of the phase space. These rare event sampling methods can be roughly divided into three categories: 1) methods that improve the sampling by increasing the temperature 2) methods that employ an external bias onto one or more selected order parameters or collective variables (CVs) and 3) methods that finds the transition paths between two stable states. However, not all methods work well in every situation thus choosing the correct one is quite crucial.

As a general guideline, methods based on increasing the temperature work well to explore a rough free energy surface with many degenerate saddle points separated by low energy barriers. An example of this would be study of surface diffusion. Collective variable based methods are the most useful when approximate reaction coordinates that describes the transition are known. Finally, transition path based approaches work best for transformations which are rather complex in nature and involves multiple collective variables which changes in a particular order, such as large protein conformational changes, protein folding, and protein (un)binding. Recently, many methods have been introduced that combine some of these classes to obtain smarter algorithms to tackle the rare event sampling problem. Elaborating on all the methods is outside the scope of this thesis and the focus will be on the three methods that are used in this thesis: transition path sampling, metadynamics, and path-metadynamics.

2.4.1 Transition path sampling

For some chemical processes, defining a collective variable is extremely difficult. One such example is, proton or electron transfer which involves solvent reorganization where many the degrees of freedom of solvent molecules have to be considered to bias the reaction. Most of the time, this collective variable is not well known. An important method to deal with these kind of transformations is Transition Path Sampling (TPS). In TPS, only the definition of stable states is required and the perfect collective variable describing the chemical reaction is usually unknown.
TPS generates an ensemble of reactive trajectories connecting the stable states, starting from an initial reactive trajectory, new trajectories are generated using a Monte Carlo scheme. Technical details about the TPS method can be found in excellent papers by Bolhuis, Dellago, Geisler and Chandler.\textsuperscript{36,37} The collective variables that define the stable states A and B need to fulfil some basic criteria: 1) the value of the collective variable for two states should not overlap resulting in well separated state definitions, 2) the states should have a high probability to be visited (i.e. have a low free energy) beyond the barrier region.

![Figure 2.1](image)

\textbf{Figure 2.1: Schematic representation of the free energy landscape with two stable, attractive wells separated by a barrier. The dotted line represents a new trajectory that was branched of at point $P$ from an old trajectory (bold line) and surpasses the TS barrier at a lower point. Figure is reproduced from the thesis 'Chemistry in Water' by Bernd Ensing.}

A schematic of 'two way shooting' is shown in Figure 2.1, where there are two stable state basins and an initial path (solid line) connecting them. This initial path need not be physical and can be obtained by biasing using any suboptimal reaction coordinate. A point $x_i$ is chosen on this path and then the equations of motion are integrated forward and backward in time after randomly perturbing the atomic momenta. This is termed as a 'two way shooting' and leads to generation of a new path (dashed line) if the forward and backward ends connect the two stable states the path is accepted and then the same process is repeated. If both the paths end up in the same state then a new point is chosen on the previous path. Acceptance of the new path usually depends on the following factors:
2.4. Rare event sampling methods

- the smaller the random perturbation in momenta, the more probable is the acceptance but the new path will closely follow the old one.

- the closer the random point $x_i$ is located to the top of the barrier, the larger the probability of acceptance but the downside is that the decorrelation of the paths will be slower.

Clearly the random perturbation should be optimized to equilibrate to the physical path quickly. After extensive analysis of a simple model, Dellago et al. suggested to tune the random perturbations to obtain an acceptance probability in the range of 30-60% for fast equilibration.

It is important to point out that in principle, the equilibrated paths represents the true dynamics of the system and thus a transition state ensemble can be obtained from these equilibrated paths which can be further analyzed to find the reaction coordinate and understand the mechanism of the chemical process.

2.4.2 Metadynamics

Metadynamics\(^{38}\) discourages the resampling of already visited parts of the phase point in the CV space. This is achieved by adding a history dependent bias $V_{bias}(\vec{s})$ to the underlying potential energy surface, where $\vec{s}$ is the chosen CV. In the simplest formulation, this bias is a small repulsive Gaussian potential which is added at every time interval $\tau$.

This external bias acting on the system at a time $t$ is:

$$V_{bias}(\vec{s}, t) = H \sum_{t' = \tau, 2\tau, \ldots} \exp\left(-\frac{(\vec{s}(t) - \vec{s}(t'))^2}{2w^2}\right)$$ (2.15)

where $H$ and $w$ are the height and width of the Gaussian respectively. As the simulation proceeds, more and more Gaussian are summed up, thus pushing the system to different regions in the CV space. Once the whole free energy landscape is explored, it will lead to the free diffusion of the CVs. It has been shown previously that this bias converges to the negative of the free energy\(^{39}\) (subject to the Langevin dynamics evolution of CV). The resolution of the free energy curve depends on the choice of the Gaussian height and width, the smaller values will result in higher resolution but increasing the computational cost. Also, depositing Gaussian less frequently will improved the accuracy of the free energy profile but slow down the exploration thus increasing the total convergence time. Furthermore, multiple parallel simulations can be used for faster exploration of the free energy landscape.\(^{40}\)

2.4.3 Path-metadynamics

Metadynamics is a powerful tool to calculate the free energy of a chemical reaction but is limited by the number of CVs as the convergence time increases.
exponentially with the number of CVs making it inefficient for many complex processes. On the other hand, TPS gives important mechanistic insights into complex chemical process but does not provide the free energy of the reaction. Path-metadynamics\textsuperscript{41,42} (PMD) combines the best of both worlds. It adaptively finds the average transition path defined in an N-dimensional collective variable space while calculating the free energy on-the-fly along this path using metadynamics.

In PMD, a progress parameter is $\sigma(\vec{z})$ is defined for any point $\vec{z}$ in an N-dimensional CV space by projecting it onto the nearest point on the path $\vec{s}(\sigma)$ as shown in Figure 2.2. This path $\vec{s}(\sigma)$ is then updated to follow the average CV density,

$$ s(\sigma) = \int_{s_\sigma} dS_\sigma \vec{z} p_\sigma(z'_1, \ldots, z'_n), \text{ with} $$

$$ p(z'_1, \ldots, z'_n) = \frac{1}{Z} \int d\vec{q} e^{-\beta U(\vec{q})} \delta(z_1 - z'_1) \ldots \delta(z_n - z'_n) $$

with $\delta$ being the Dirac delta function and $Z$ is the partition function. The integration is over the plane $S_\sigma$ perpendicular to the path $s(\sigma)$. The ensemble average of all the sampled $\vec{z}$ is replaced by a time average,

$$ \langle \vec{z} \rangle_{\sigma_g} = \lim_{t \to \infty} \int_0^t \int_{S_{\sigma_g}} \vec{z}(t') dS_{\sigma_g} dt' $$

Numerically, the transition path is adapted using a set of $M$ sequential equidistant nodes in N-dimensional CV-space. The projection of $\vec{z}$ onto the path is then done using the two closest nodes, and the $\sigma$ is obtained by interpolation. Further details of the PMD and some benchmark systems can be found here.\textsuperscript{41,42}
Chapter 3

Ab-initio molecular dynamics simulations of liquid water at ambient pressure

Density functional theory combined with molecular dynamics (DFT-MD) in Car-Parrinello or Born-Oppenheimer MD simulations has been and still is an immensely powerful instrument to obtain microscopic insight in water and aqueous solutions. The quality of the DFT-MD water model is however limited by the approximate nature of standard GGA functionals, neglect of nuclear quantum effects, and finite size effects. Here, we have performed a comprehensive study of the effects of various parameters such as the choice of XC functional, Van der Waals corrections and smoothing operators on various structural and dynamical properties of bulk water. We show that for condensed phase simulations, the PBE functional even in conjunction with Grimme’s D3 correction overstructures water and results in glassy dynamics. We also propose a framework to estimate the box size for ambient pressure simulation using NVT ensemble. We utilize the new framework to study the structural and dynamical properties of bulk water at ambient conditions. Our results can serve as a benchmark of the current state-of-the-art of DFT-MD simulation of aqueous systems.
3. DFT-MD simulation of liquid water at ambient pressure

3.1 Introduction

Since the introduction of the Car-Parrinello Molecular Dynamics method in 1985 and Born-Oppenheimer Molecular Dynamics soon after, ab initio molecular dynamics (AIMD) simulation has become an invaluable instrument for the study of the structure and dynamics of matter at finite temperatures. AIMD allows for the computation of classical trajectories of atoms and molecules, using, most often, Density Functional Theory (DFT) for the quantum mechanical calculation of the inter-atomic forces from the underlying electronic structure. Arguably the most important contribution of AIMD simulation has been its role in our understanding of water and chemical processes in aqueous solution.

It is a challenge to accurately model from first principles the countless remarkable properties of water, which are ultimately the result of the intricate balance of the molecular interactions arising from e.g. hydrogen-bonding, the dipole moment, the polarizability and dispersion. DFT-MD simulations, using exchange-correlation (XC) functionals based on the generalized gradient approximation (GGA), are efficient enough to sample tens to hundreds of water molecules for tens to hundreds of picoseconds. However, early simulations of water using the BP86, BLYP and PBE functionals showed too much structure and too slow dynamics of the liquid. More recent AIMD simulations using GGA functionals augmented with corrections for Van der Waals interactions or using hybrid functionals showed much improvement. Nevertheless, AIMD simulations are still often performed at a somewhat higher temperature to compensate for the above deficiencies, but also to compensate, albeit in a crude manner, for the neglect of nuclear quantum effects.

A possible culprit that could severely amplify subtle shortcomings in the DFT description is the common use of the canonical (NVT) ensemble in DFT-MD simulations, without prior relaxation of the pressure using a barostat. Equilibration of the system in the isobaric-isothermal (NpT) ensemble, as routinely done for classical forcefield MD simulations, is usually avoided because DFT-MD simulations in the NpT ensemble exhibit much larger pressure fluctuations due to the relatively small system sizes. They further require much larger (planewave) basis sets (or, equivalently, finer electron density grids) to reduce quality dependence on the volume, and they generally show slow converge of the volume with respect to the affordable DFT-MD simulation times. Instead, DFT-MD simulations are typically setup at a density obtained from experiment, a forcefield MD simulation, or an educated guess, which could cause an undesirably large positive or negative internal pressure during the simulation.

Pioneering DFT-based Monte Carlo (MC) and DFT-MD simulations
in the NpT and NpH ensembles revealed that the density of liquid water is severely underestimated at the GGA level of theory; for instance with BLYP by about 20% by McGrath et al.\textsuperscript{67} and by 8% by Ma et al. Inclusion of Grimme’s D2 empirical Van der Waals correction resulted in the densities ranging from 0.992 kg/l\textsuperscript{68} to 1.07 kg/l.\textsuperscript{56} Even better densities were obtained with the hybrid PBE0-D3 functional, and the truly first principles MP2-MD simulation,\textsuperscript{58} although these two methods are still respectively one and two orders of magnitude more computationally demanding for a system of 64 water molecules respectively, compared to the BLYP-D3 level of theory. Nevertheless, these NpT simulations showed that it is now, in principle, feasible to compute the average pressure with enough accuracy to perform AIMD simulations at a desired pressure.

Here, we assess the performance of the two popular GGA functionals, BLYP and PBE, for the modeling of liquid water in the NVT ensemble, in which we focus our attention on the dependence of structural and dynamical properties of the density and the temperature. In particular, we compare NVT simulations performed at the experimental density with those in which the pressure is relaxed towards 1 atm.

### 3.2 Computational details

The DFT-MD simulations were performed using the CP2K package.\textsuperscript{70} The electronic structure part of CP2K, called QUICKSTEP, uses the combined Gaussian and planewave (GPW) method for the calculation of forces and energies.\textsuperscript{71} The GPW method is based on the Kohn-Sham formulation of density functional theory and employs a hybrid scheme of Gaussian and planewave functions. Here the Kohn-Sham orbitals are expanded using a linear combination of atom-centered Gaussian type orbital functions, and an auxiliary basis set of planewaves describes the electronic charge density. This combination is chemically more intuitive and computationally more efficient than the sole use of a planewave basis set. First principles simulations with CP2K sample directly the Born-Oppenheimer surface.

In the DFT-MD simulations, the norm-conserving pseudopotentials of Goedecker et al. (GTH)\textsuperscript{72} were applied to replace the core electrons. We employed the BLYP\textsuperscript{21,73} and PBE\textsuperscript{23} exchange-correlation functionals, augmented with Grimme’s D3 dispersion corrections\textsuperscript{33} to include Van der Waals interactions. For comparison with the previous results we also performed simulations with Grimme’s D2 dispersion corrections\textsuperscript{31,32}. According to Lin et al. dispersion corrections to the BLYP functional have shown to consistently improve the energetics and structure for liquid water. The Gaussian basis set consisted of a triple-Z valence basis set, augmented with two sets of d-type or p-type polarization functions (TZV2P)\textsuperscript{74}
optimized for the use with the GTH pseudopotentials. A charge density cutoff of 400 Ry is used for the auxiliary planewave basis set. For the planewave grid, we applied the nearest neighbour density smoothing operator NN10 and NN50 for the xc calculation of all reported properties, except the pressure tensor. The purpose and effects of these smoothing operators are discussed in details in the section 4.3. In the current version (5.1) of CP2K, the pressure tensor calculation does not work together with the smoothing operator; instead we employed the USE_FINER_GRID option for an accurate (albeit computationally more demanding) estimation of the pressure tensor. A CSVR thermostat\(^75\) with a time constant of 500 fs is used to generate an NVT ensemble. In addition to the D3 correction, the temperature is set higher than ambient, as it has been shown to yield good agreements with structural and dynamical properties measured from experiments of bulk water under ambient conditions. The runs at ambient pressure conditions presented here were carried out on state-of-the-art parallel computers on the Dutch national supercomputer Cartesius using between 24 and 72 processors.

We have performed AIMD simulations of two types of pure bulk water systems consisting of 64 water molecules, firstly a system at the experimental density of 1.0 g/cm\(^{-3}\) (i.e. a box size of 12.43 Å) and the other one at the density corresponding to at the average density corresponding to 1 atm, thus 12.12 Å for BLYP-D3 and 12.43 Å for PBE-D3. The bulk water systems were constructed using the packmol program,\(^76\) pre-equilibrated using the TIP3P forcefield with the molecular dynamics simulation package LAMMPS,\(^77\) and further equilibrated at the DFT-MD level of theory with CP2K for approximately 50 ps. Subsequent sampling is performed for 50 ps. The MD timestep is 0.5 fs.

### 3.2.1 Setting up ambient pressure simulations with NpT or NVT

The straightforward procedure for setting up a DFT-MD simulation in the NVT ensemble at a desired average pressure is to first equilibrate the system in the isobaric-isothermal (NpT) ensemble to determine the appropriate volume. However, for a relatively small system, the convergence of the average volume is rather slow as shown in Figure 3.1a. Moreover, the quality of a finite planewave basis set depends on the volume, which may require a large basis set to avoid a bias on the volume. Further details on DFT-MD simulations in the NpT ensemble can be found here.\(^56,67-69\) As an alternative to performing an initial NpT simulation, we can also estimate the appropriate volume from the previous NVT simulation at the experimental box size with the following formula,

\[
V = V' \exp \left[ \beta_T(p' - p) \right],
\]
which is obtained by integrating the definition of the isothermal compressibility, \( \beta_T = -V^{-1}(\partial V/\partial p)_T \). Here, \( V' \) and \( p' \) are the volume and pressure respectively from a previous simulation. We show that \( p' \) converges relatively fast in a short simulation time of approximately 5 to 10 ps as shown in Figure 3.1b. The fast convergence of \( p \) is in agreement with previous simulations.\(^{56}\) Finally, \( p \) is the target pressure, which is here 1 atm. The experimental isothermal compressibility is \( \beta_T = 45.0 \) Mbar\(^{-1}\).\(^{78}\) However, since the compressibility is somewhat pressure dependent and, more important, the GGA water may have a compressibility that deviates from the experimental value,\(^{79}\) additional NVT simulations may be required to obtain the volume corresponding to the target pressure in an iterative manner (see e.g. Ref.\(^{55}\)). Here however, we used only Eqn. 3.1 to estimate the volume at ambient pressure, which led to a box size of \( L = 12.12 \) Å at the BLYP-D3/TZV2P/400Ry level of theory and a box size of \( L = 12.43 \) Å for PBE-D3/TZV2P/400Ry.

**Fig. 3.1:** The left plot shows the box size fluctuations and the bottom panel of the right plot shows the pressure fluctuations in an NPT simulation of liquid water. The top panel of the right figure shows pressure fluctuations in an NVT simulation of liquid water. All these simulations use BLYP-D3 and their running average values are also shown.
3.3 Methods of analysis

3.3.1 Radial distribution function
The local structure around the water molecules in the liquid phase is adequately measured by the pair-correlation function, or radial distribution function (RDF). The RDF gives the density of atoms around the central atom type, relative to the average density. Integration of the RDF gives information on the coordination numbers around the central species. The RDF can also be obtained from scattering experiments.

3.3.2 Mean square displacements
The translational motion of water molecules is probed by measuring their mean square displacements (MSD), \( \langle \Delta r^2 \rangle \), defined as

\[
\langle \Delta r^2 \rangle = \langle |r_i(t) - r_i(0)|^2 \rangle,
\]

(3.2)

where \( r_i(t) \) and \( r_i(0) \) are the position vectors of the oxygen atom of the \( i \)-th water molecule at time \( t \) and \( t = 0 \), respectively, the averaging is taken over all water molecules and different time origins. The slopes of these MSD curves are used to calculate the diffusion coefficient (\( D_E \)) of water molecules using Einstein’s relation

\[
D_E = \lim_{\Delta t \to \infty} \frac{\langle \Delta r^2 \rangle}{2d\Delta t}.
\]

(3.3)

3.3.3 Orientational tetrahedral order \( q \)
A plethora of measurables have been defined for characterizing the local structure of liquid water, but many are somewhat ambiguous and ad hoc. A convenient and much employed description focuses on the distortion with respect to the structure of crystalline ice, where water molecules are regularly positioned on a well-defined lattice and where the nearest neighbors form a regular tetrahedron. The first solvation shell of each water molecule forms an approximate tetrahedron, distorted by the frequent exchanges between the first and second shells and by the increased probability to find water molecules in an interstitial position between these two shells. How much the local structure of liquid water deviates from an ideal tetrahedron is influenced by a number of factors. For example, decreasing the temperature enhances the local order. Here, we will employ the popular angular orientational tetrahedral order \( q \). This is probably the most widely used parameter to quantify the local structure of water. It was originally proposed by Chau and Hardwick,\(^80\) and subsequently rescaled by Errington and Debenedetti\(^81\) so that
the average value of $q$ varies from 0 for an ideal gas to 1 for a regular tetrahedron. It focuses on the four nearest water oxygen neighbors and is defined as

$$q = 1 - \frac{3}{8} \sum_{j=1}^{3} \sum_{k=j+1}^{4} \left( \cos \psi_{jk} + \frac{1}{3} \right)^2$$  \hspace{1cm} (3.4)

where $\psi_{jk}$ is the angle formed by the lines joining the oxygen atom of the water molecule under consideration and its nearest neighbor oxygen atoms $j$ and $k$. $q$ has been used for example, to study the structure of supercooled water$^{82,83}$ and to understand the changes in the local water structure around various solutes$^{84,85}$ and surfaces.$^{86,87}$

### 3.3.4 Orientational autocorrelation

The rotational motion of water molecules is an important parameter as it can give an estimate of the solute interaction. For instance, water present in the first hydration shell around a protein molecule has a much slower decay time as compared to bulk water. Moreover, the orientational decorrelation can be compared to the anisotropic decay measured with time-resolved femtosecond infrared spectroscopy.

It is important to examine the reproducibility of this orientational motion using the current exchange correlation functionals as well as its dependency on various other parameters such as VdW correction and the basis set used. The time evolution of the vector $\vec{u}$ along a water O-H bond can be measured by calculating the autocorrelation function $C_2(t)$, defined as

$$C_2(t) = \langle P_2[\vec{u}(0) \cdot \vec{u}(t)] \rangle$$  \hspace{1cm} (3.5)

where $P_2$ is the second order Legendre polynomial, and $\vec{u}(t)$ is the unit vector along each of the OH bonds of the water molecules.

### 3.4 Results

#### 3.4.1 GGA water at the experimental density

To set the stage for our analysis of DFT-MD simulation of liquid water at ambient pressure, we first review the behavior of water as it is most commonly modeled from first principles: in the NVT ensemble, with the volume set to match the experimental density at $T = 300$ K. For a cubic box holding 64 water molecules, the edge has then a length of 12.43 Å. To compensate for over-structuring and too slow diffusion, the simulation temperature is typically set somewhat higher
than ambient temperature. Here, we show the temperature dependence of the radial distribution function (RDF) in Figure 3.2a, the tetrahedral order parameter in Figure 3.2b, the mean square displacement (MSD) in Figure 3.2c and the anisotropy decay in Figure 3.2d, computed using the BLYP functional, augmented with Grimme’s D3 parameterization of the Van der Waals correction and the TZV2P/400Ry basis sets. In the top panel in Table 3.1, we list the temperature dependence of the coordination number, the diffusion coefficient, and the anisotropy decay time.

The oxygen–oxygen RDF in Figure 3.2a and the tetrahedral order parameter in Figure 3.2b show that the structure of the liquid becomes less pronounced with increasing temperature, as expected. Comparison of the RDFs with the experimental result\(^8\) (dashed line) shows a very good match of the simulation result at \(T = 303\) K with the X-ray result at 298 K. Our results also compare well with previous BLYP-D3 simulations at \(T = 298\) K,\(^6\) at \(T = 323\) K,\(^5\) and at \(T = 330\) K.\(^6\)

The dynamics of the solvent molecules at several temperature values is shown by the MSD in Figure 3.2c and the anisotropy decay in Figure 3.2d. The translational self-diffusion coefficient, \(D\), of water is obtained from the slope of the MSD, and the characteristic decorrelation time of the water orientation \(\tau_{\text{mid}}\) is computed by fitting the anisotropy decay with a single exponential. Both numbers are compiled in Table 3.1, together with the experimental numbers. As the diffusion coefficient is known to be sensitive to the finite size effects, we do not compare our result directly to the experimental number of bulk water at \(T = 298\) K, which is 0.186 Å\(^2\)/ps for D\(_2\)O,\(^8\) but rather to a first order scaling of the experimental number to our small periodic system, using the equation:

\[
D_{\text{pbc}} = D_\infty - \frac{k_B T \xi}{6\pi \eta L}. \tag{3.6}
\]

Here, \(k_B\) is Boltzmann’s constant, \(\xi = 2.83727\) is a constant, and \(\eta = 1.095\) mPa·s is the viscosity of water at 298 K.\(^9\) This gives \(D_{\text{exp,pbc}} = 0.129\) Å\(^2\)/ps, which is in between the BLYP-D3 diffusion at \(T = 303\) K (\(D = 0.074\) Å\(^2\)/ps) and at \(T = 325\) K (\(D = 0.166\) Å\(^2\)/ps). Since, we compare to the (scaled) experimental number of heavy water, we expect the effects of nuclear quantum effects, which are not included in our simulations, to have a only very small influence. Our simulation result at \(T = 325\) K is a perfect match with the results of \(D = 0.17\) Å\(^2\)/ps at \(T = 323\) K by Jonchiere \textit{et al}.,\(^5\) and of \(D = 0.181\) Å\(^2\)/ps at \(T = 330\) K by Bankura \textit{et al}.,\(^6\)

The anisotropy decay, quantified by the relaxation time, \(\tau_{\text{mid}}\), of the orientational time correlation function, is likely to be affected by finite size effects,
3.4. Results

analogous to the translational diffusion constant. The experimental value of \( \tau_{\text{mid}} = 2.3 \) ps for bulk D\(_2\)O at \( T = 298 \) K is comparable to the BLYP-D3 result at \( T = 345 \) K, similar to the unscaled experimental self-diffusion coefficient.

![Graphs](image)

Fig. 3.2: Oxygen-oxygen radial distribution functions, tetrahedral order parameter, mean square displacement and anisotropic decay of liquid water for a range of temperatures at the BLYP+D3/TZV2P/400Ry level of theory and the experimental water density.

The PBE exchange correlation functional is generally thought to be of comparable quality as the BLYP functional and is probably the second most used functional for DFT-MD simulations of water. In Figures 3.4a to 3.4d, we compare the structural and dynamical properties of water described with the BLYP-D3 and the PBE-D3 functionals with the TZV2P/400Ry basis sets at \( T = 325 \) K. Clearly, there are large differences between the performance of these two GGAs. While the BLYP-D3 functional shows good agreement with experiments, the PBE-D3 functional results in highly over-structured water, as seen both in the RDF and in the tetrahedral order, and a slow translational diffusion and orientational dynamics.
Table 3.1: Computed coordination number, \( n_{cc} \), diffusion coefficient, \( D \), and orientational relaxation time \( \tau_{mid} \) for selected DFT-MD simulation parameterizations. The effect of the technical parameters in the lower panels can be seen by comparison to the second line, printed in bold, which we take as the reference simulation.

<table>
<thead>
<tr>
<th>XC</th>
<th>Basis</th>
<th>grid</th>
<th>T [K]</th>
<th>Lbox [Å]</th>
<th>( n_{cc} )</th>
<th>( D ) [Å²/ps]</th>
<th>( \tau_{mid} ) [ps]</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLYP</td>
<td>D3</td>
<td>TZV2P/400</td>
<td>300</td>
<td>12.43</td>
<td>3.442</td>
<td>5.477</td>
<td>0.111</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>300</td>
<td>12.43</td>
<td>3.273</td>
<td>4.512</td>
<td>0.066</td>
</tr>
<tr>
<td>PBE</td>
<td>D3</td>
<td>TZV2P/400</td>
<td>300</td>
<td>12.43</td>
<td>3.255</td>
<td>4.031</td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>300</td>
<td>12.43</td>
<td>3.188</td>
<td>4.081</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>D3</td>
<td>TZV2P/400</td>
<td>300</td>
<td>12.43</td>
<td>3.488</td>
<td>5.200</td>
<td>0.306</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>300</td>
<td>12.43</td>
<td>3.488</td>
<td>5.228</td>
<td>0.261</td>
</tr>
<tr>
<td></td>
<td>D3</td>
<td>TZV2P/400</td>
<td>300</td>
<td>12.43</td>
<td>3.412</td>
<td>4.811</td>
<td>0.166</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>300</td>
<td>12.43</td>
<td>3.488</td>
<td>5.196</td>
<td>0.182</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>300</td>
<td>12.43</td>
<td>3.488</td>
<td>5.118</td>
<td>0.208</td>
</tr>
</tbody>
</table>

...
3.4. Results

(a) Oxygen-Oxygen RDF

(b) Mean square displacement

Fig. 3.3: Oxygen-oxygen radial distribution functions and mean square displacement of liquid water for a range of temperatures at the PBE+D3/TZV2P/400Ry level of theory and the experimental water density.

Similar results were found for PBE-D3 by Michaelides et al.\textsuperscript{30} regarding the over-structuring of water at room temperature. It is concluded that PBE-D3 dynamics is extremely slow and the major emphasis of the present work has been on getting the proper dynamics for PBE functional. One straightforward way to achieve this is to run the simulations at higher temperature as shown in Figure 3.3b. It can be concluded from here that at a temperature of around 385 K, PBE-D3 has a reasonable dynamics. Also the radial distribution function of bulk water at this temperature as shown in Figure 3.3a matches well with that of bulk water at 300 K.

3.4.2 GGA water at ambient pressure

After calculating the ambient pressure box size for PBE and BLYP for both D2 and D3 Van der Waals corrections using the method described in section 3.2.1, we compared the structural and dynamical properties in all these cases. Interestingly the box size for ambient pressure simulations using BLYP-D2 and PBE-D3 is found to be the same as the experimental box size which is $L = 12.43$ Å. The average pressure for BLYP-D3 ($L = 12.12$ Å) and BLYP-D2 ($L = 12.43$ Å) is 10 bar and 30 bar at 300 K, respectively. For the corresponding PBE-D3 ($L = 12.43$ Å) and PBE-D2 ($L = 12.53$ Å) simulations the average pressure is 40 bar and 70 bar. Figure 3.6a shows the effect of the reduction of the volume on the radial distribution function. Surprisingly, the BLYP-D3 water is less structured at ambient pressure than the experimental curve and that from the simulation at the experimental density. On the other hand, the BLYP-D2 result
3. DFT-MD simulation of liquid water at ambient pressure

![Graphs showing](image)

(a) O-O RDF  
(b) Tetrahedral order parameter

(c) Mean square displacement  
(d) Anisotropic decay

**Fig. 3.4:** Oxygen-oxygen radial distribution functions, tetrahedral order parameter, mean square displacement and anisotropic decay of liquid water for BLYP and PBE at the D3/TZV2P/400Ry level of theory, experimental water density and 325 K

is highly over-structured. For the PBE curves in the lower panel, softening of the structure is seen upon relaxing the pressure, but the RDF still shows much more structure than the experimental curve.

### 3.4.3 MD parameters

Structural and dynamical properties calculated using ab initio molecular dynamics simulations are reported to be highly dependent on the choice of a variety of technical specifications. Many of these have been studied in detail such as basis set dependence and the effect of dispersion corrections. In this section we revisit the effect of various dispersion corrections and basis sets on water properties. We also investigate the effects of various widely used smoothing parameters on the structure and dynamics of liquid water.
3.4. Results

(a) Oxygen-Oxygen RDF

(b) Tetrahedral order parameter

Fig. 3.5: Oxygen-oxygen radial distribution functions and tetrahedral order parameter of liquid water at ambient conditions for PBE and BLYP and its comparison with experimental density at 300 K using two different Van der Waals corrections.

3.4.3.1 Smoothing

The GPW method implemented in CP2K uses GTH pseudopotentials which by construction have a zero pseudocharge density at the core for all elements apart from H. Thus, for these pseudopotentials the core region is a difficult part of the exchange and correlation potential calculation. The pronounced spike of $v^{xc}$ at the core due to zero pseudocharge density gives rise to small variations in the total energy as atoms move relative to the grid. These variations or ripples not only make it difficult to conserve the energy but sometimes can also give rise to strange artifacts such as immobile atoms as found in our simulation of bulk water where in a few simulations an oxygen atom was frozen in space. To overcome these issues a slightly different scheme has been introduced in CP2K: the XC density is smoothed using $S^z$, a nearest neighbour smoothing operator which is defined as
3. DFT-MD simulation of liquid water at ambient pressure

Fig. 3.6: Mean square displacement and anisotropic decay of liquid water at ambient conditions for PBE and BLYP and it’s comparison with experimental density at 300 K using two different Van der Waals corrections.

\[(S^z f)_{i,j,k} = \frac{z^3}{z^3 + 6z^2 + 12z + 8} \times \sum_{l=-1}^{1} \sum_{m=-1}^{1} \sum_{n=-1}^{1} q^{-|l|-|m|-|n|} f_{i+l,j+m,k+n}\]

(3.7)
3.4. Results

Fig. 3.8: Diffusion coefficients for different temperatures computed with BLYP+D3 and the corresponding experimental values.

where function $f$ represents the electron density function over the grid points $i,j,k$. Typical values for $z$ are 10 or 50 depending on the required amount of smoothing. Such a smoothing is equivalent with a redefinition of $E^{XC}$ that reduces to the identity as the planwave cutoff is increased. Here we used $XC_{SMOOTH\_RHO\_NN10}$ for all the GGA calculations at the experimental box size, however the pressure tensor fluctuations for such a smoothing are very large and gives highly negative numbers for a box of water as shown in Figure 3.9. Therefore, instead of using the usual smoothing option we use instead the $USE\_FINER\_GRID$ option available in the CP2K package to deal with these large fluctuations. This alternate method uses a finer grid to calculate the exchange and correlation energy instead of smoothing the density but it is three times more computationally expensive for a 64 water molecule box as compared to the normal smoothing operator. Here we compare the structure and dynamics of water at 325 K using these two alternate methods of calculating the $XC$ energy. As seen in Figure 3.10b the dynamics is slower in case of NN10 as compared to both NN50 and UFG whereas the rdf is quite similar for all three as shown in Figure 3.10a.

3.4.3.2 Van der Waals at $T = 325$ K

To test the effect of dispersion correction on the structure and dynamics of liquid water, we have performed simulations using three popular dispersion corrections namely Grimme’s D2, D3 and D3 including a C9 term. The most important feature introduced in the D3 term is that now atom-specific pairwise dispersion...
3. DFT-MD simulation of liquid water at ambient pressure

Fig. 3.9: Pressure fluctuations in NVT simulations with their running average values while using NN10 smoothing operator.

(a) Oxygen-Oxygen RDF  (b) Mean square displacement

Fig. 3.10: Oxygen-oxygen radial distribution functions and mean square displacement of liquid water for BLYP at 325 K using NN10, NN50 and finer grid.

coefficients and cutoff radii are both computed from first principles. However, the long-range part of the interaction between three ground-state atoms is not exactly equal to the interaction energies taken in pairs. The total three-body contribution is typically $5 - 10\%$ of $E_{\text{disp}}$ therefore three-body nonadditivity terms are taken into account by the C9 term:

$$C_{9}^{ABC} \approx -\sqrt{C_{6}^{AB}C_{6}^{BC}C_{6}^{AC}}$$ (3.8)
Here again we focus on the dynamics and the structure of liquid water. It is already well established that the dispersion corrections leads to the softening of the structure which is in agreement with our results in figure 3.11a. However, in contrast to the improvements in the interaction energy of molecular crystals by including the C9 term,\(^9\) it seems to have a negative effect on the mean square displacement of water. The resulting water is slowed down by 15\% as compared to the original D3 implementation.

\[\text{Fig. 3.11: Oxygen-oxygen radial distribution functions and mean square displacement of liquid water for BLYP at 325 K for the three most commonly used Van der Waals corrections.}\]

### 3.4.3.3 Basis set

Here, we probe a major technical aspect of \textit{ab initio} molecular dynamics simulations namely, the basis set which is used to solve the Kohn-Sham equations. Contrary to the common consensus of TZV2P being the best basis set, the structure and dynamics of water is very similar to TZVP basis set as shown in Figure 3.12a for the structure and in Figure 3.12b for the dynamics. It is important to mention that TZVP is computationally more efficient as compared to TZV2P and hence will be the choice for most of the simulations performed in the later part of the thesis.
3. DFT-MD simulation of liquid water at ambient pressure

Fig. 3.12: Oxygen-oxygen radial distribution functions and mean square displacement of liquid water for BLYP at 325 K for DZVP, TZVP and TZV2P basis set.

3.5 Concluding remarks

DFT-MD simulations of aqueous systems are typically performed at a somewhat elevated temperature to compensate for the observed over-structuring and glassy dynamics of these systems when modeled at the GGA level of theory, even when the DFT functional is augmented with corrections for Van der Waals interactions. Here, we have performed a careful assessment of the structure and the dynamics of liquid water at different temperatures and pressures employing BLYP and PBE, two of the most popular XC-functionals for water simulations, in combination with Grimme’s D2 and D3 Van der Waals corrections. For simulations at the experimental density of water at $T = 300 \text{ K}$, we find that the BLYP-D3 functional shows good agreement with experimental results for the oxygen-oxygen radial distribution function, but that the dynamics is somewhat slow, especially the orientational diffusion. Instead, the PBE-D3 functional shows a strong over-structuring, which only disappears when the simulation is performed at a temperature of about $T = 385 \text{ K}$. Also the self-diffusion coefficient is much too low ($0.002 \text{ Å/ps}^2$) at $T = 300 \text{ K}$ which is almost 2 order of magnitude smaller than the experimental value at 300 K. This indicates glassy behavior in case of PBE and makes it prohibitively inaccurate for the DFT-MD simulations of liquid water.

The average pressure of a BLYP-D3 simulation at $T = 325 \text{ K}$ is $-1.3 \cdot 10^2 \text{ atm}$, while that of PBE-D3 is 70 atm. We note that the convergence of the pressure in an NVT simulation, or the density in an NPT simulation, is system dependent and much slower for the glassy PBE-D3 water model than for the BLYP-D3 water model. Using the experimental isothermal compressibility of water, we have estim-
ated the cell volume at ambient pressure. Interestingly, relaxation of the pressure results in an improved water structure in the case of PBE-D3 and BLYP-D2, but not for BLYP-D3. The radial distribution function in the latter case does not show almost any structure beyond the first peak. Instead, the PBE-D3 structure is seen to improve, although the RDF still shows some over-structuring. From the tetrahedral order parameter it can be seen that at ambient pressure the shoulder at $\theta = 0.5$ is gone. This indicates more water exchange in the first and the second coordination shells. This also explains the over softening of the structure as seen in the RDF for ambient pressure simulations of BLYP. It is intriguing to see how the onset of the first peak of radial distribution function shifts with the change of XC functional. In case of the PBE functional the onset matches well with the experimental curve whereas for BLYP it is slightly shifted towards the right.

While the D2 and D3 parameterisations of the Van der Waals corrections lead to similar RDFs, the effects on the dynamics of liquid water are substantial. The self-diffusion coefficient is underestimated (0.058 Å/ps$^2$) for BLYP-D2 whereas for BLYP-D3 it improves by a factor of 3 (0.166 Å/ps$^2$) at $T = 325$ K. The D3 parametrization contains an additional C9 term that accounts for three-body interactions. Invoking this C9 term results in a slight softening of the structure but further slowdown of the dynamics resulting in a self-diffusion coefficient of 0.112 Å/ps$^2$. Thus, among the three examined choices, BLYP-D3 works the best.

In addition to the water structure and dynamics using these popular DFT models, we have tested several technical parameters of the DFT-MD simulation of water. In particular we determined the effect of constraining the center of mass of the system and the effect of smoothing the XC density on coarse grids. Density smoothing is recommended because fine grids are significantly more computationally demanding (but currently unavoidable to compute the pressure tensor accurately in CP2K), and coarse grids without smoothing may result in spurious dynamics. We find that the effect of the different available smoothing options on the structure and the dynamics of water is not negligible. We compared three different type of smoothing possible NN10, NN50 and USE_FINER_GRID. It was found that the dynamics was best for NN50 whereas the structure was best for NN10. USE_FINER_GRID although three times computationally more expensive gives good values for both the dynamics and structure.

This analysis shows that technicalities are very important for correct simulations of liquid water, which also makes it crucial that users report the XC functional, Van der Waals correction terms, basis set and type of smoothing with their scientific results to be comparable among different DFT-MD studies. We can expect further improvements of the description of water with functionals that are higher up on the scale of Jacob’s ladder, such as hybrid functionals and even...
post HF descriptions, however, this is still much more costly than semi-local DFT functionals.

**Acknowledgements**

This work is part of the Industrial Partnership Programme (IPP) 'Computational sciences for energy research' of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO). This research programme is co-financed by Shell Global Solutions International B.V. The calculations were carried out on the Dutch national e-infrastructure with the support of the SURF Cooperative.
Chapter 4

Direct simulation of self-exchange electron transfer reaction

Outer sphere electron transfer between two ions in aqueous solution is a rare event on the time scale of first principles molecular dynamics simulations. We have used transition path sampling to generate an ensemble of reactive trajectories of the self-exchange reaction between a pair of Ru$^{2+}$ and Ru$^{3+}$ ions in water. To distinguish between the reactant and product states, we use as an order parameter the position of the maximally localised Wannier center associated with the transferring electron. This allows us to align the trajectories with respect to the moment of barrier crossing and compute statistical averages over the path ensemble. We compare our order parameter with two typical reaction coordinates used in applications of Marcus theory of electron transfer: the vertical gap energy and the solvent electrostatic potential at the ions.

4. Direct simulation of self-exchange electron transfer reaction

4.1 Introduction

Electron transfer is the fundamental process in reduction and oxidation (i.e. redox) chemistry, taking place for example in metal corrosion, fuel cell reactions, and photo-synthesis. Early pioneering studies of the prototypical self-exchange reaction between two metal ions in aqueous solution, as illustrated by equation 4.1,

\[ \text{M}^{n+} + \text{M}^{*(n+1)+} \rightarrow \text{M}^{(n+1)+} + \text{M}^{*n+} \quad (4.1) \]

have been pivotal for our understanding of electron transfer processes and the development of Marcus’ theory of electron transfer. Molecular dynamics (MD) and Monte Carlo simulations confirmed that the free energy profiles of electron transfer reactions consists of two intersecting parabolas, as proposed by Marcus’ theory, confirming that the solvent reorganisation during electron transfer is indeed well-described by the linear response approximation.

As with most chemical reactions, the self-exchange reaction is an activated process, in which the system has to surpass a transition state of high (free) energy, which renders direct sampling of an electron transfer reaction in an MD simulation highly improbable. This so-called rare event problem was tackled in the early work by using enhanced sampling techniques, such as umbrella sampling or free energy perturbation, which bias the simulation along an appropriate reaction coordinate, for example the position of the electron.

Unfortunately, it is not so easy to bias the electron transfer process in combination with a quantum-chemical many-electron description of the system, such as with density functional theory (DFT). The reason is that the position of the excess electron, in the reactant state, the product state, or somewhere in between, is governed by the configurational state of the polar solvent environment, which is difficult to capture with a reaction coordinate. The method of constrained-DFT offers alternatively the possibility to use the electron position as the reaction coordinate, although this constraint also confines the orbital to a certain shape.

Instead, most DFT studies on electron transfer focus on half-reactions. With the half-reaction approach, introduced by Warshell, and further developed for DFT-MD simulation by Sprik and co-workers, the redox potential and reorganisation free energy is computed from the average energy needed to add (or remove) an electron to (or from) the system containing only the solvated donor or acceptor species. This approach has been extensively applied to investigate the redox properties of transition metals, organic molecules, and proteins in explicit solvent. However, the half-reaction approach does not provide direct information on the electron transfer between a donor and ac-
We have recently returned to the problem of direct simulation of electron transfer between donor and acceptor species with DFT-MD. Using transition path sampling (TPS) simulations, we can generate an ensemble of reactive trajectories that sample the unbiased dynamics of an electron transfer process. Analysis of the trajectories provides unique information on the transition state ensemble, such as the donor-acceptor distance, the solvent structure in the transition state and the dynamics of the process. Particularly interesting would be an analysis of relevant order parameters or collective variables that correlate with the vertical energy gap, $\Delta E$, which is the reaction coordinate used in Marcus theory. However, direct computation of $\Delta E$ is difficult within the DFT-MD simulations, as it requires constraining the position of the electron.

Here, we analyse a DFT-MD/TPS ensemble of reactive trajectories of the self-exchange reaction between a pair of Ru$^{2+}$ and Ru$^{3+}$ ions in water solvent. Rather than computing $\Delta E$, we obtain statistics of the energy to insert or delete an electron ($\Delta E_{\text{ins}}$, $\Delta E_{\text{del}}$) to the system along the reactive trajectory. We first sample these two vertical gap energies for the pair of Ru$^{2+}$ and Ru$^{3+}$ ions in an equilibrium simulation to reconstruct the free energy landscape, which we compare to that obtained using the half-reaction approach. We then sample $\Delta E_{\text{ins}}$ and $\Delta E_{\text{del}}$ along the reactive trajectories, and compare the sum as a reaction coordinate to the solvent electrostatic potential and an order parameter based on the position of the transferring electron.

In the following, we first briefly summarise the theory of electron transfer as applied in half-reaction simulations, we discuss the issues associated with the use of periodic boundary conditions, and we present the computational details of our combined transition path sampling and DFT-MD simulations. The results are presented in three subsections covering the redox properties obtained with the half-reaction approach, our equilibrium DFT-MD simulations of the pair of Ru$^{2+}$ and Ru$^{3+}$ ions in water, and analysis of the reactive trajectories of the electron transfer obtained with TPS. This is followed by the conclusions.

4.2 Methods

4.2.1 Theory of electron transfer

At the DFT level of theory, redox properties of ions and molecules in explicit solvent are most conveniently computed using the half-reaction approach, which is based on Marcus’ theory of electron transfer and can be connected to the half-reactions
that take place in electrochemical cell experiments. The central quantity to be measured is the vertical energy gap, $\Delta E$, which is the energy needed to add (or delete) an electron to (or from) the solute in a given nuclear configuration. This gap energy plays the role of the reaction coordinate and it quantifies the state of the polarised medium when the solute is an ion. In the case of a molecular solute, $\Delta E$ also captures the state of the molecule itself. In remarkably many cases (but not always\cite{118}), the response of both the polarised medium and the molecular solute, as measured by $\Delta E$, is largely linear with respect to the amount of charge loaded to or from the solute.

Consider the reduction of an oxidant (O) to a reductant (R),

\begin{equation}
O + e^- \longrightarrow R,
\end{equation}

and the vertical energy gap

\begin{equation}
\Delta E = E_R(r^N) - E_O(r^N),
\end{equation}

with $E_x(r^N)$, $x = (R, O)$, the internal energy of the system at a nuclear configuration $r^N$, in the reduced or oxidised state, respectively. Linear response of the system entails that the distribution $P(\Delta E)$ is Gaussian, centered at an (ensemble) average $\langle \Delta E \rangle_x$ in either the reduced or oxidised state:

\begin{equation}
P_x(\Delta E) = \frac{1}{\sigma_x\sqrt{2\pi}} \exp \left[ \frac{(\Delta E - \langle \Delta E \rangle_x)^2}{2\sigma_x^2} \right],
\end{equation}

with $\sigma^2$ the variance. The Landau free energies, as a function of the $\Delta E$ reaction coordinate,

\begin{equation}
\Delta A_x(\Delta E) = -k_BT\ln[P(\Delta E)],
\end{equation}

must then be parabolic and give rise to Marcus’ well-known diabatic free energy landscape, which is illustrated in Figure 4.1 by the dashed curves. Here, $T$ is the temperature and $k_B$ is Boltzmann’s constant.

The overall reaction free energy is a particularly simple function of the average gap energies when linear response holds:

\begin{equation}
\Delta A_r = \frac{1}{2}(\langle \Delta E \rangle_O + \langle \Delta E \rangle_R),
\end{equation}

and a similar relation holds for the third key quantity in Marcus’ theory, the reorganisation free energy:

\begin{equation}
\lambda = \frac{1}{2}(\langle \Delta E \rangle_O - \langle \Delta E \rangle_R),
\end{equation}
4.2. Methods

Fig. 4.1: Schematic illustration of a parabolic free energy landscape of electron transfer. The dashed lines depict Marcus’ diabatic free energy curves, which cross at the diabatic free energy barrier $\Delta A^\ddagger_d$. The solid line shows the adiabatic free energy profile, with a somewhat lower barrier $\Delta A^\ddagger_a$, due to the coupling of the electronic reactant and product states. The difference between $\Delta A^\ddagger_d$ and $\Delta A^\ddagger_a$ is generally much smaller than depicted here. For a half-reaction written in the conventional reduction form (Eqn. 4.2) and the definition of the reaction coordinate, $\Delta E$, from Eqn. 4.3, the reaction thus takes place from right to left.

This $\lambda$ is the free energy associated with the relaxation of the polarised medium into the product state after a vertical excitation, i.e. after adding an electron to the solute at a fixed configuration. The reorganisation free energy is also associated to the curvature of the free energy curves and to the magnitude of the fluctuations of $\Delta E$ in either oxidation state:

$$\lambda_x = \frac{\sigma_x^2}{2k_B T} \quad (4.8)$$

Finally, the diabatic free energy barrier, denoted in Figure 4.1 at the crossing point of the parabolas, is given as a function of the overall reaction free energy (which is zero for a full self-exchange reaction, but not for the half-reaction) and the reorganisation free energy:

$$\Delta A^\ddagger_d = \frac{(\lambda + \Delta A_r)^2}{4\lambda} \quad (4.9)$$

In practice, the diabatic free energy curves are computed by sampling $\Delta E$ along two equilibrium DFT-MD simulations; one in the oxidised state where the energy needed to add an electron is computed and one in the reduced state in which the energy to delete an electron is sampled. In section 4.3.1, we will set the stage by
applying this approach to compute the redox properties of the Ru$^{2+}$/Ru$^{3+}$ couple, which can be compared to the extensive studies on this system by Blumberger and Sprik.110–112,115,123

As pointed out in the introduction, the half-reaction framework cannot be so easily applied to study directly the full reaction, i.e. the electron transfer reaction between a donor molecule and an acceptor molecule. In the full reaction case, $\Delta E$ would be the energy to bring the electron from the donor to the acceptor at a fixed configuration. This $\Delta E$ can be decomposed into two terms: (1) the energy needed to delete the electron from the donor, $\Delta E^{\text{del}}$, and (2) the energy needed to insert the electron at the acceptor, $\Delta E^{\text{ins}}$. The first term is easy to compute as long as the transferring electron resides at the highest occupied orbital (HOMO) of the system. But the second term cannot be computed with ground-state DFT without additional constraints, since the electron would return to the donor in the electronic ground-state. Of course, we can compute the energy needed to insert an electron to the acceptor before removing the electron from the donor, $\Delta E^{\text{ins, before}}$, but that energy would contain an additional interaction energy of the inserted electron with the other electron still at the donor, plus higher order terms in a polarisable environment.

If we consider this additional electron interaction as a correction to the $\Delta E^{\text{ins, before}}$ term, the $\Delta E$ of moving an electron from donor to acceptor at a given nuclear configuration can thus be computed using the half-reaction technique of inserting an electron and deleting an electron:

$$\Delta E = \Delta E^{\text{del}} + \Delta E^{\text{ins}} + U^{\text{corr}}.$$  \hspace{1cm} (4.10)

Here, and hereafter, we omit the “before” superscript; that is, the $\Delta E^{\text{ins}}$ refers to the energy needed to insert an electron into the original system instead of the system after having deleted an electron from the donor. As an initial approximation, the correction term could be taken as the Coulomb energy between an electron centered at the donor and an electron centered at the acceptor:

$$U^{\text{corr}} = \frac{1}{4\pi\varepsilon_0} \frac{1}{|\mathbf{r}_{\text{RuD}} - \mathbf{r}_{\text{RuA}}|}.$$  \hspace{1cm} (4.11)

However, due to the periodic boundary conditions applied in these calculations, the correction is somewhat more complicated, as further discussed next.

4.2.2 Periodic boundary conditions and finite size effects

One of the complications of DFT-MD simulations of molecular processes in condensed materials, and in particular of half-reactions in aqueous solution, is related
4.2. Methods

to the rather small size of the systems and the periodic boundary conditions to mimic the extended phase. The long-range electrostatic interactions are computed using Ewald summation, which is of course almost always much better than using a cutoff, but is nevertheless not without its own issues. The problem is that the half-reaction approach is typically applied to systems that are not neutral, and, moreover, the method requires the calculation of energy differences between systems that carry different charges.

Although in principle a charged system embedded in infinite arrays of periodic copies would have an infinite energy, Ewald summation removes the divergent term in a manner that is commonly interpreted as adding a uniformly smeared out background charge. The density of the background charge is \(-q/L^3\) and thus depends on the box size, \(L\). The charged system interacts with the background charge so that the energy of the system contains a “self-interaction” term that is a function of the box size and shape. For a charged particle in a cubic cell, this (Wigner) self-interaction is:

\[
\Phi = -\frac{\xi_M q^2}{2L} \tag{4.12}
\]

with \(q\) the charge of the particle, \(L\) is the cell dimension, and \(\xi_M = 2.837297\) the Madelung constant of a cubic periodic system. The finite size effects due to the Ewald potential have been extensively investigated, for example for the interaction energy between two ions,\(^{124}\) for ionic hydration free energies,\(^{125-127}\) for the reaction and reorganisation free energies in half-reaction calculations by Sprik,\(^{112,117}\) and by Blumberger\(^{115}\) on the reorganisation free energy in (full reaction) electron transfer.

For a charged particle dissolved in a polar medium such as water, it turns out that the self-interaction energy is largely compensated by the screening of the solvent, effectively reducing this \(L^{-1}\) dependent term by a factor of \(1/\epsilon\), with \(\epsilon\) the dielectric constant, which is approximately 80 for ambient water. The \(L^{-1}\)-dependence of the solvent comes from the repulsion of the polarised solvent in the unit cell with that in the periodic images, which solvate the periodic image of the charged particle in their own cell, rather than the original ion.

Although the \(L^{-1}\)-dependent self-interaction is thus largely canceled in a high-dielectric solvent, higher order terms in \(L\) remain that arise from the finite size of the ion, which creates a cavity with a radius \(R\) in the medium in all periodic images. The total energy difference due to (Ewald) periodic boundary conditions (PBC) with respect to that of a continuum Bohr model\(^{112,126,127}\) (B) is:

\[
E_{\text{PBC}} - E_{\text{B}} = \frac{\xi_M q^2}{2L\epsilon} - \left(1 - \frac{1}{\epsilon}\right) \frac{2\pi q^2 R^2}{3L^3} + O(L^{-5}) \tag{4.13}
\]

In the half-reaction approach presented above, the reaction free energy, \(\Delta A_r\)
4. Direct simulation of self-exchange electron transfer reaction

(Eqn. 4.6), and the reorganisation free energy, $\lambda$ (Eqn. 4.7), are computed from sampled energy differences, $\Delta E$, of the system with and without an extra charge. Sprik et al found that for $\Delta A$, the box size dependence is dominated by the $L^{-3}$ term, whereas $\lambda$ scales as $L^{-1}$. Blumberger and Lamoureux showed that for the full electron transfer reaction, in which the electron moves from donor to acceptor, the box size dependence of the reorganisation free energy is much less severe. This is perhaps not surprising, since in that case only the dipole is changed, but not the total charge.

![Fig. 4.2: The apparent dependence on the distance between the Ru$^{2+}$ and Ru$^{3+}$ ions of the energy gap probability (top) and the free energy curves (bottom) computed with forcefield MD simulations. The curves obtained from $\Delta E^{\text{del+ins}}$ (in black) are not distance dependent. The lines are Gaussian (top) and parabolic (bottom) fits through the measured data (circles). Dashes lines in the bottom panel are computed by adding to $\Delta E^{\text{del+ins}}$ the $U^{\text{corr}}$ term of equation 4.10 as explained in the text.](image)

The dipole term depends on the distance between the donor and the acceptor. In Figure 4.2, we show the $P(\Delta E^{\text{ET}})$ (top panel) and the free energy curve computed with Eqn. 4.5 (bottom panel) computed with simple forcefield MD simulations of a pair of Ru$^{2+}$ and Ru$^{3+}$ aqua complexes, constrained at different donor–acceptor distances, $d$, ranging from 6 to 14.5 Å, solvated in 1000 water
molecules in a cubic box with $L = 31.04 \, \text{Å}$. As expected, the curves show a clear dependence on the donor-acceptor distance.

The black curves in Figure 4.2 were computed using the insertion-deletion scheme, with $\Delta E^{\text{del+ins}} = \Delta E^{\text{del}} + \Delta E^{\text{ins}}$, i.e. the first two terms of Eqn. 4.10, which are, within the statistical accuracy, independent of the Ru-Ru distance. The missing $U^{\text{corr}}$ term of Eqn 4.10 can be computed as the difference in $\Delta E^{\text{del+ins}}$ between a system containing the two ruthenium point charges (without solvent) with, and without, periodic boundary conditions. Adding this to the $\Delta E^{\text{del+ins}}$ indeed recovers perfectly the $\Delta E^{\text{ET}}$ curves, as shown by the dotted lines in the bottom panel of Figure 4.2.

For the DFT-MD simulations that we discuss hereafter, the box size dependence and the $U^{\text{corr}}$ term also contain contributions due to the instantaneous electronic polarisation. The aim of this work, however, is not to obtain quantitative numbers for the redox properties, but rather, to connect the reaction coordinates used in Marcus theory and the half-reaction method to our equilibrium and TPS simulations of the full electron transfer reaction.

4.2.3 Transition path sampling

We use an adapted 2-way TPS algorithm of shooting and shifting moves to generate reactive trajectories, starting from an initial path obtained from a biased MD simulations as explained hereafter. Here, reactive trajectory refers to a trajectory that starts in the stable reactant state and ends in the stable product state, or vice versa. The shooting move proceeds in the usual way. That is, each new trajectory starts from a randomly chosen configuration from the previous trajectory, by adding small perturbations to all atomic momenta, and performing two MD simulations, one forward and one backward (by reversing all velocities) in time (hence, “2-way TPS”). The shooting move is accepted if the new trajectory connects the reactant and product states, or rejected otherwise. The perturbations are done by adding random momenta from a 5 Kelvin Maxwell-Boltzmann distribution to the original momenta, removing any total momentum, and rescaling to the original temperature. This resulted in an average acceptance ratio of the shooting move of 0.46. Instead of the original shifting move, we locate the configuration on the previous trajectory that is closest to halfway the electron transfer process. The shooting configuration was subsequently chosen from a fixed number of saved restart files ranging from 1000 MD steps (i.e. 500 fs) before and after this central time frame. By using this fixed number of restart files to randomly choose the shooting configuration from, the TPS algorithm obeys detailed balance. The actual trajectory length was not fixed however, and ranged typically between 2 and 3 ps.
4. Direct simulation of self-exchange electron transfer reaction

The initial reactive trajectory was obtained from a 1 ps constrained MD simulation that started from an equilibrated system, in which the six ruthenium-oxygen distances of the hexaaqua Ru$^{2+}$ complex were forced to decrease by 0.3 Å. By the time that the constrained coordination shell becomes smaller than that of the (free) Ru$^{3+}$ complex, an electron is expelled from the Ru$^{2+}$ complex and taken up by the Ru$^{3+}$ complex soon after. A second constrained simulation was initiated from a frame just after the forced electron transfer took place, in which the six Ru$-$O distances were kept constant. This second trajectory was used to generate initial unconstrained reactive trajectories, by applying the TPS shooting move but without perturbing the nuclear momenta.

The stable reactant and product states are defined by the numbers of $d$-electrons at each of the ruthenium ions; six on one and five on the other, or vice versa. Counting the number of electrons is conveniently done by transforming the occupied orbitals into maximally localised Wannier functions$^{128,129}$ (MLWFs) and computing the distances of the MLWF centers to each of the Ru ions. The MLWFs are computed every 10 MD steps on-the-fly along each TPS trajectory. We define as an order parameter:

$$\xi_{ET} = (d_{Ru-X} - d_{Ru'-X})/d_{Ru-Ru'},$$

(4.14)

in which Ru and Ru’ denote the two Ru ions, $d$ is the distance and X is the MLWF that is, of the eleven MLWF centers within a cutoff distance of the Ru nuclei, farthest away. This order parameter is practically equal to minus one in the reactant state with the first Ru ion being 2+ and the second Ru’ ion is 3+; in the product state with the charges reversed, $\xi_{ET} \approx 1$; and during electron transfer, $\xi_{ET}$ has a value in between. Now, we define the reactant state as the set of configurations for which $\xi_{ET} < -0.9$ and the product state as the set of configurations for which $\xi_{ET} > 0.9$.

The TPS algorithm of generating reactive trajectories described here was implemented using a bash script that would launch the forward and backward DFT-MD simulations (further detailed below), accept or reject the trajectories, choose a new starting configuration, perturb the momenta, and launch the next simulations, and so forth.

4.2.4 Density functional theory-based molecular dynamics

The DFT-MD simulations were performed using the CP2K software package. The electronic structure part of CP2K, called QUICKSTEP, uses the combined Gaussian and plane-wave (GPW) method for the calculation of forces and energies. The GPW method is based on the Kohn-Sham formulation of density func-
4.3 Results

In the DFT-MD simulations, the norm-conserving pseudopotentials of Goedecker et al. (GTH) were applied to replace the core electrons. We employed the BLYP exchange-correlation functional, augmented with Grimme’s D2 dispersion correction to include van der Waals interactions. The BLYP-D level of theory has been extensively used and tested by the scientific community for the description of various structural and dynamical properties of liquid water and aqueous solutions (see e.g. Refs. 30,31), including the Ru$^{2+}$/Ru$^{3+}$ redox couple. The Gaussian basis set consisted of a double-zeta valence basis set with a single set of polarization functions (DZVP) optimised for the use with the GTH pseudopotentials. A charge density cutoff of 280 Ry was used for the auxiliary plane-wave basis set. For the plane-wave grid, we applied the nearest neighbour smoothing operator NN10. A CSVR thermostat with a time constant of 500 fs was used to generate an NVT ensemble. The temperature was set to 300 K. Periodic boundary conditions were applied to a cubic box with an edge length of 12.4138 Å for the simulations with two Ru ions and 64 water molecules. For the single ion with 32 water molecules, the box dimension was 9.86 Å. The simulations here were carried out on the Dutch national supercomputer Cartesius using 24 processors in parallel.

4.3 Results

We first compute the redox properties of the Ru$^{2+}$/Ru$^{3+}$ couple using the half-reaction approach. Next, we investigate the pair of aqueous Ru$^{2+}$ and Ru$^{3+}$ ions in equilibrium. Thirdly, we analyse the reactive trajectories of the self-exchange reaction harvested with TPS.

4.3.1 Redox properties using the half-reaction approach

The Ru$^{2+}$/Ru$^{3+}$ redox couple was one of the first systems to which Sprik and co-workers applied their half-reaction approach. Here we present the same redox properties computed with the CP2K program at the BLYP+D2/DZVP/280Ry level of theory for a single ion with 32 water molecules in a cubic box with an edge of $L = 9.86$ Å subject to periodic boundary conditions. Two equilibrium DFT-MD simulations were performed, one of Ru$^{2+}$ and one of Ru$^{3+}$, with a length of 10 ps, of which the last 5 ps were used to sample $\Delta E$.

The $P(\Delta E)$ distributions are shown in the top-panel of Figure 4.3. The histograms of $\Delta E$, shown by the circles are fitted very well by Gaussian functions...
(solid lines), as expected, however, the widths of the two distributions, measured by $\sigma$, are not exactly the same. The Ru$^{2+}$/Ru$^{3+}$ system is known to obey Marcus linear response theory rather well,\textsuperscript{110–112,115,123} so the deviation seen here must be due to statistical errors.

![Figure 4.3](image)

**Fig. 4.3:** Top panel: vertical energy gap distributions for the Ru$^{3+} + e^- \rightarrow$ Ru$^{2+}$ half-reaction. The circles show the histogram values obtained from the simulation data; the solid lines are Gaussian fit functions. Bottom panel: free energy curves obtained from the Gaussian functions using Eqn 4.5.

The derived redox properties are compiled in Table 4.1. The uneven standard deviations of $\Delta E$ in the oxidised and reduced states leads to a significant discrepancy between $\lambda_O$ and $\lambda_R$. We have previously seen that the fluctuations in $\Delta E$ are easily underestimated in the rather short DFT-MD simulations, and that $\lambda$ computed from the averages (Eqn 4.7) is the safer estimate for the reorganisation free energy.\textsuperscript{118} Our results are in reasonable agreement with the early work of Blumberger and Sprik, although in that work an external chemical potential was applied to enforce alignment of the minima of the parabolic curves, and the parabola were fitted to the combined $O$ and $R$ data-sets resulting in a more symmetric free energy landscape.

### 4.3.2 Free energy curves of the combined Ru$^{2+}$ and Ru$^{3+}$ ions

Four independent DFT-MD simulation were performed of a pair of Ru$^{2+}$ and Ru$^{3+}$ ions solvated by 64 water molecules in a cubic box with $L = 12.41$ Å subject to...
4.3. Results

Table 4.1: Average and standard deviations of the computed $P(\Delta E)$ distributions using the half-reaction approach, together with the derived redox properties using the indicated equations.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>value [eV]</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\langle \Delta E_R \rangle$</td>
<td>-1.23</td>
<td>4.4</td>
</tr>
<tr>
<td>$\langle \Delta E_O \rangle$</td>
<td>0.74</td>
<td>4.4</td>
</tr>
<tr>
<td>$\sigma_R$</td>
<td>0.18</td>
<td>4.4</td>
</tr>
<tr>
<td>$\sigma_O$</td>
<td>0.22</td>
<td>4.4</td>
</tr>
<tr>
<td>$\Delta A_R$</td>
<td>-0.24</td>
<td>4.6</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.98</td>
<td>4.7</td>
</tr>
<tr>
<td>$\lambda_R$</td>
<td>0.64</td>
<td>4.8</td>
</tr>
<tr>
<td>$\lambda_O$</td>
<td>0.93</td>
<td>4.8</td>
</tr>
</tbody>
</table>

periodic boundaries. These simulations had different starting configurations, but they had in common that in all cases the Ru$^{3+}$ complex was deprotonated. From two of the simulations, we removed at the start the excess proton from the solvent, yielding a total charge of +4 of the system. The other two systems had a charge of +5. In the latter, the extra proton diffused through the solvent via the Grothuss mechanisms, and was not seen to jump back onto the deprotonated ruthenium complex during the simulation. Rather than adding explicit counter ions, which would make the sampling more cumbersome, the system has a neutralising uniform background counter-charge via the Ewald summation, as further detailed in the method section. The simulations had a length of 50 ps, of which the last 30 ps were used for analysis.

Figure 4.4 shows in the top panel the distributions, $P(\Delta E_{\text{del}})$ (red curve), $P(\Delta E_{\text{ins}})$ (blue), and $P(\Delta E_{\text{del+ins}})$ (black). The solid lines are Gaussian functions fitted to the histograms of the data, which are shown in circles for the system without the solvated proton. The fitted Gaussian functions are also shown for the system with the excess proton, with dashed lines, to illustrate that the computed distributions are remarkably independent from the total charge of the system; apart from a small deviation between the blue curves, the distributions obtained from the two systems are, considering the statistical uncertainty, equal to each other.

The distribution of the energy needed to insert an electron at the Ru$^{3+}$ ion is somewhat broader than that of the energy needed to delete an electron at the Ru$^{2+}$ ion. This trend is in agreement with what we found earlier for the $P_O(\Delta E)$ and $P_R(\Delta E)$ distributions with the half-reaction approach (see Figure 4.3), although this is probably a coincidence. The more plausible cause for the different widths is the different coordination shells of the Ru$^{2+}$ and Ru$^{3+}$ ions here, the former being coordinated by six water ligands whereas the latter contains five water ligands and
one hydroxo ligand. This difference between the oxidised and reduced Ru ions could indeed cause such a non-linear effect.

The first and second moments of the distributions, together with the derived reorganisation free energies are listed in Table 4.2. The average energy to delete an electron from the system is in perfect agreement with $-\langle \Delta E_R \rangle$ of the half-reaction (Table 4.1) and also $\sigma$ matches very well. However, this must be somewhat fortuitous, considering the different box sizes and the effect that the nearby Ru$^{3+}$ complex must have in the current case of the full reaction. Also $\sigma_{\text{ins}}$ is in excellent agreement with $\sigma_O$ of the half-reaction, however the average $\langle \Delta E_{\text{ins}} \rangle$ is significantly shifted with respect to $\langle \Delta E_O \rangle$. This discrepancy is mainly
due to the different coordination shells between the Ru\(^{3+}\) complexes, which contains a hydroxo ligand in the current full reaction case, while for the half-reaction none of the water ligands were deprotonated.

The lower panel of Figure 4.4 shows the parabolic free energy profile obtained from the \(P(\Delta E^{\text{del+ins}})\) distribution, using Eqn. 4.6 (solid black line). Note however, that this profile is shifted with respect to that of the actual electron transfer from Ru\(^{2+}\) to Ru\(^{3+}\), as we did not include the correction term, \(U^\text{corr}\), from Eqn. 4.10. But since \(\Delta A_r = 0\) for the self-exchange reaction, the average \(\langle \Delta E \rangle\) must equal \(\lambda_{\text{del+ins}}\). Here we neglect the fluctuation part of the correction term, which is expected to be small. In Figure 4.4, the shifted free energy curve is shown in red, together with its counterpart of the reverse electron transfer reaction (red dashed). The obtained reorganisation free energy, \(\lambda_{\text{del+ins}} = 1.73\) eV is in reasonable agreement with earlier estimates using constrain-DFT (1.62 eV),\(^{123}\) static quantum-chemistry methods (1.95 eV),\(^{132}\) and experimental measurement (2.0 eV).\(^{133}\) Finally, the diabatic free energy barrier for the Ru\(^{2+}\)/Ru\(^{3+}\) self-exchange reaction is computed with Eqn. 4.9 to be \(\Delta A_d^{\text{f}} = 0.43\) eV.

**Table 4.2:** Average and standard deviations of the computed \(P(\Delta E^{\text{del}})\), \(P(\Delta E^{\text{ins}})\), and \(P(\Delta E^{\text{del+ins}})\) distributions, together with the derived redox properties using the indicated equations.

<table>
<thead>
<tr>
<th>Quantity</th>
<th>value [eV]</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\langle \Delta E^{\text{del}} \rangle)</td>
<td>1.23</td>
<td>4.4</td>
</tr>
<tr>
<td>(\langle \Delta E^{\text{ins}} \rangle)</td>
<td>1.19</td>
<td>4.4</td>
</tr>
<tr>
<td>(\langle \Delta E^{\text{del+ins}} \rangle)</td>
<td>2.42</td>
<td>4.4</td>
</tr>
<tr>
<td>(\sigma_{\text{del}})</td>
<td>0.19</td>
<td>4.4</td>
</tr>
<tr>
<td>(\sigma_{\text{ins}})</td>
<td>0.23</td>
<td>4.4</td>
</tr>
<tr>
<td>(\sigma_{\text{del+ins}})</td>
<td>0.30</td>
<td>4.4</td>
</tr>
<tr>
<td>(\Delta A_r)</td>
<td>0.0</td>
<td>4.6</td>
</tr>
<tr>
<td>(\lambda)</td>
<td>2.42</td>
<td>4.7</td>
</tr>
<tr>
<td>(\lambda_{\text{del}})</td>
<td>0.73</td>
<td>4.8</td>
</tr>
<tr>
<td>(\lambda_{\text{ins}})</td>
<td>1.05</td>
<td>4.8</td>
</tr>
<tr>
<td>(\lambda_{\text{del+ins}})</td>
<td>1.73</td>
<td>4.8</td>
</tr>
</tbody>
</table>

**4.3.3 Transition path sampling**

We employed the TPS technique in combination with DFT-MD to generate in total eight sequences of reactive trajectories of the self-exchange reaction between a Ru\(^{2+}\) ion and a Ru\(^{3+}\) ion in water. During the initial constrained MD simulations to generate an initial path (as explained in the Method section), not only an electron transferred, but also a proton was donated by one of the aqua ligands of the Ru\(^{3+}\) complex to the solvent. The first four TPS sequences were generated
4. Direct simulation of self-exchange electron transfer reaction

with this proton in the water solvent. The second four TPS sequences were generated after removal of the solvated proton. The first seven sequences contain 50 reactive trajectories. A typical “path tree” of such a sequence is shown in Figure 4.5. The eighth sequence contained 180 (accepted) paths. The acceptance ratio over all paths was 0.46. The simulation length of each forward or backward path ranged from 0.5 to 1.5 ps, as seen from the path tree.

![Figure 4.5](image1)

**Fig. 4.5:** One of the eight TPS “path trees”, showing the length of each path and the shooting time where the next path branches off a previous one. Only the accepted paths are shown.

Figure 4.6 shows a cartoon of snapshots from a representative reactive trajectory. The octahedrally coordinated ruthenium complexes are shown in ball-stick representation (Ru is blue, O is red, and H is white), while the other solvent water molecules are shown as grey sticks. In the first panel the Wannier center is seen, as a yellow sphere, to depart from the Ru$^{2+}$ ion. The Ru$^{3+}$ complex on the left side has five water ligands and one hydroxo ligand, which points toward the electron donor complex and is hydrated by a solvent water molecule also drawn in ball-stick representation. In the second panel, the Wannier center is halfway
the donor and acceptor complexes, and 45 fs later the electron is taken up by the acceptor species in panel 3. Panels 4 and 5 show the subsequent proton transfer from the (hitherto) Ru$^{2+}$ complex to the acceptor complex, via a bridging solvent water molecule.

![Fig. 4.6: Five snapshots from a typical reactive trajectory showing the electron transfer represented by the departure of its Wannier center (yellow sphere) from the Ru$^{2+}$ ion (right-hand-side blue sphere) in the left-most panel. In the second panel, the Wannier center is approximately in the middle ($t = 0$), and 45 fs later it arrives at the other Ru ion. A proton from the right-hand-side hexaaqua complex is transferred via an intermediate solvent water (fourth panel) to the other complex. Ligands and intermediate H$_2$O molecule are shown in red and white ball-and-stick representation; other solvent molecules are drawn as grey sticks.](image)

The distance between the Ru ions fluctuates around 7 Å. The electron transfer is in all reactive trajectories accompanied by the proton transfer, in no preferential order (as should be expected by time-reversibility). Further details of this proton-coupled electron transfer mechanism are discussed in a separate publication; instead here we will focus on how we can connect the reactive trajectories of the adiabatic electron transfer to the redox quantities from Marcus theory.

In order to perform a statistical analysis over the TPS sequences of reactive trajectories, we have to align the trajectories by time. We take the moment of electron transfer in each path as the zero of the time scale, which we compute by fitting the reaction coordinate values $\xi_{ET}$ (see Eqn. 4.14) along each trajectory.
by the switch function \( f(t) = \tanh[a * (t - t_0)] \), with the parameter \( t_0 \) defining the moment of electron transfer. Figure 4.7 shows \( \xi_{\text{ET}} \), which quantifies the position of the Wannier center, for three trajectories. Also the average \( \xi_{\text{ET}} \) over the sequence of 180 paths is shown (black line). The blue line shows a reactive event in which the electron recrossed back toward the donor and again to the acceptor moiety. However, such barrier recrossing are rather scarce in this electron transfer process. Seen from the average \( \xi_{\text{ET}} \), the actual electron transfer, that is, passing from the reactant state definition to the product state definition takes about \( 10^2 \) fs.

Having aligned the trajectories, statistics of other order parameters can be obtained from the TPS ensemble. Panels B and C of Figure 4.7 shows \( \Delta E^\text{del} \), \( \Delta E^\text{ins} \), and \( \Delta E^\text{del+ins} \), computed from \( t = -0.5 \) to \( t = 0.5 \) ps. The fluctuations seen in the three individual path traces are very large just before and after the electron transfer event, however around \( t = 0 \) they all show a clear spike toward zero. Only for a few trajectories the \( \Delta E \) values are zero at \( t = 0 \), which probably means that our time resolution was not fine enough to see this in the other trajectories. The horizontal dashed line in panel C denotes the average \( \langle \Delta E^{\text{del+ins}} \rangle = 2.42 \) eV from the equilibrium simulation (see also Table 4.2), which is an indication that our relatively short 2-3 ps trajectories indeed connect the stable reactant and product states, and do not sample only the top region of a free energy barrier.

Another interesting property, already suggested by Marcus, is the solvent electrostatic potential, \( \phi_x \), \( x = (D, A) \) acting on the donor (D) and acceptor (A) ions respectively. These potentials are shown in panel D of Figure 4.7. Here, the potentials are approximated by summing the classical Coulomb interaction over all solvent nuclei and associated Wannier centers using the minimum image convention and neglecting the long-range part. The electrostatic potentials are remarkably symmetric with respect to their switching at \( t = 0 \). Note that all these order parameters show clearly the barrier recrossing event in path 50 (blue lines) in agreement with the Wannier center position.

Both the \( \Delta E^{\text{del+ins}} \) and \( \phi_x \) quantities can be used as reaction coordinates for the electron transfer process. The \( \Delta E^{\text{del+ins}} \) parameter does not by itself distinguish between the reactant and product states, as it does not change sign unlike the original \( \Delta E \). This can be remedied by keeping track of which of the Ru ions is involved in the electron deletion and insertion processes. Alternatively, the \( \Delta E^{\text{del+ins}} \) can be multiplied with \( \xi_{\text{ET}} \), which is shown in panel E in Figure 4.7 (black line, left-hand-side axis) for the averaged quantities. For the electrostatic potential, we take the difference between \( \phi_A \) and \( \phi_D \), which is also shown in panel E (red line, right-hand-side axis). Both these combined order parameters switch smoothly from the reactant state to the product state, crossing zero at \( t = 0 \).
4.3. Results

In Figure 4.8, we show the correlation between the three order parameters, by projecting 18 reactive trajectories, taken from the long sequence at an interval of 10, as green points and lines in pairs of the order parameters. The gap energy $\Delta E^{\text{del}+\text{ins}}$, multiplied with $\xi_{\text{ET}}$, is shown versus the Wannier center position, $\xi_{\text{ET}}$, in the top panel, and versus the solvent potential difference, $\phi_A - \phi_D$, in the middle panel. The bottom panel shows $\phi_A - \phi_D$ versus $\xi_{\text{ET}}$. Starting with the top panel, we notice that there is a strong correlation between the gap energy and the Wannier center position, and that their relation is not linear. Since $\Delta E$ captures the linear response of the solvent polarisation to the amount of charge displaced, the latter being quantified here by $\xi_{\text{ET}}$, this is somewhat surprising. Note however, that $\xi_{\text{ET}}$ is obtained from the sampling of adiabatic electron transfer events only (see also Figure 4.1, in which the adiabatic profile deviates from the parabolic curves near the barrier). If diabatic electron transfer events could have been included in the statistics of $\xi_{\text{ET}}$ at gap energies left and right from the center at $\Delta E = 0$, the curve would have been more straight.

Note also that the fluctuations seen in $\Delta E$ in the reactant and product states (i.e. at $\xi_{\text{ET}} = -1$ and $\xi_{\text{ET}} = 1$ respectively) disappear at $\xi_{\text{ET}} = 0$. In other words, not only the average gap energy is zero at barrier crossing, but the gap energy for all reactive trajectories is zero at barrier crossing. However, since $\Delta E$ (i.e. the solvent polarisation) governs the electron position, and not the other way around, this means that a simulation in which the electron position is fixed in the middle, $\Delta E$ would exhibit the same fluctuations as it would in the reactant or product state, whereas, vice versa, a simulation at fixed $\Delta E = 0$, would show very little fluctuations in the electron position.

Panels B in Figure 4.8 shows that there is a good, almost linear, correlation between $\Delta E$ and the difference in solvent electrostatic potential. However, the correlation with $\xi_{\text{ET}}$ in panel C clearly shows that this potential difference does not uniquely determine the electron position. This could be caused by the approximate nature of the calculation of the electrostatic potential order parameters here, using pairwise sums over nuclei and Wannier center distances, but it could also mean that other interactions play a role. For example, we found that the correlation of $\Delta E$ with the solvent electrostatic potential computed using the atomic Mulliken charges would significantly improve when this electrostatic potential was multiplied by a factor depending on the amount of charge transfer from the Ru ion to the aqua ligands in each configuration (data not shown).
4. Direct simulation of self-exchange electron transfer reaction

Fig. 4.7: Panel A: the Wannier center coordinate, $\xi_{\text{ET}}$ showing the fast switch from -1 to 1 during electron transfer. Curves for three reactive trajectories are shown, together with the path average. In path 50, a barrier recrossing is seen (blue line). Panel B and C shows $\Delta E^{\text{del}}$, $\Delta E^{\text{ins}}$, and $\Delta E^{\text{del}+\text{ins}}$. Panel D: the solvent electrostatic potential at the Ru ion positions. Panel E: $\Delta E^{\text{del}+\text{ins}}$ multiplied by $\xi_{\text{ET}}$ (black line, left-hand-side axis) and the difference of the electrostatic potentials between the acceptor and donor ions, $\phi_A - \phi_D$ (red line, right-hand-side axis).
4.3. Results

Fig. 4.8: Correlation between the three reaction coordinations used here to describe the electron transfer reaction. Panel A: the vertical gap energy $\Delta E^{\text{del}+\text{ins}}$ versus the Wannier center position $\xi_{\text{ET}}$; panel B: $\Delta E^{\text{del}+\text{ins}}$ versus the difference between the electrostatic potentials at the donor and acceptor ions $\phi_A - \phi_D$; and panel C: $\phi_A - \phi_D$ versus $\xi_{\text{ET}}$. Green crosses and lines denote the points visited along the reactive trajectories; data from 18 paths with an interval of 10 of the longest sequence is used. The black lines show the average over all accepted paths.
4. Direct simulation of self-exchange electron transfer reaction

4.4 Conclusions

We have used transition path sampling combined with first principles DFT-MD simulations to investigate the proton-coupled electron transfer reaction taking place between a pair of ruthenium(II/III) ions in aqueous solution. Until now, the main approach to study this prototypical self-exchange reaction at the DFT level of theory was by means of the half-reaction approach. Here, we first applied this approach to construct the diabatic free energy landscape and compute the overall reaction and reorganisation free energies, based on Marcus’ linear response theory of electron transfer. Next, we performed equilibrium DFT-MD simulations of a pair of donor and acceptor ions in solution. As it is not possible for the full reaction to sample the vertical gap energy, $\Delta E$, of transferring the electron from the donor ion to the acceptor ion, the gap energy was probed indirectly as the sum of the energy needed to delete an electron from the donor ion, $\Delta E_{\text{del}}$, and the energy required to insert an electron at the acceptor ion, $\Delta E_{\text{ins}}$. This allowed us to compute the free energy profiles of the full diabatic electron transfer reaction.

The reactive trajectories generated with DFT-MD/TPS remain always in the electronic ground-state, and thus sample the adiabatic electron transfer landscape. To define the stable reactant and product states, we used as the order parameter, the position of the center of a maximally localised Wannier function associated with the transferring electron. The moment of barrier crossing along each path was set as the reference of the time scale to align the paths and compute averages over the paths. In particular, we have computed the $\Delta E_{\text{del+ins}}$ along the electron transfer reaction and also the solvent electrostatic potential at the ruthenium ions, both of which are important ingredients in Marcus’ theory of electron transfer.

Correlating the Wannier center position with $\Delta E_{\text{del+ins}}$ shows that there is a one-to-one mapping between the order parameters in the neighbourhood of the transition state, where both order parameters pass through zero. This means that the electron position is strictly governed by the value of $\Delta E_{\text{del+ins}}$ when it is close to zero. The relation between these two order parameters was surprisingly non-linear however, which we believe to be due to the sampling of only adiabatic electron transfer events. Instead the correlation of the difference between the electrostatic potential at the ruthenium ions shows fluctuations in the electrostatic potential that are not different during electron transfer with respect to that in the stable states. This suggests that the solvent electrostatic potential difference is not a very good reaction coordinate for electron transfer, as it does not determine strictly the amount of electron transfer.
APPENDIX

Comparison with Ag$^+$/Ag$^{2+}$ electron transfer

To understand how the electron transfer dynamics depends on the nature of the metal ions, we simulated a similar self exchange electron transfer between Ag$^+$/Ag$^{2+}$ aqua complexes. For Ru$^{2+}$/Ru$^{3+}$ reaction the coordination number for both Ru$^{2+}$ and Ru$^{3+}$ is six and the only difference is the average Ru-O distance in the first coordination shell. In case of silver, not only does the average Ag-O distance change but also the coordination number changes from four, for Ag$^+$, to five, for Ag$^{2+}$ as shown below:

\[
[\text{Ag(H}_2\text{O)}_4]^+ + [\text{Ag}^*(\text{H}_2\text{O})_5]^{2+} \rightarrow [\text{Ag(H}_2\text{O)}_5]^{2+} + [\text{Ag}^*(\text{H}_2\text{O})_4]^+ \quad (4.15)
\]

First, we performed equilibration runs of each of the Ag$^+$ and Ag$^{2+}$ aqua complexes separately, including full solvent to characterize the structure of the complexes. As shown by the radial distribution functions in Figure 4.9, Ag$^+$ has on average four water ligands, Ag$^{2+}$ has five. The average Ag-O distance for Ag$^+$ is 2.32 Å, which is close to the experimental value of 2.35 Å found by Fulton et al.\textsuperscript{134} After this peak, there is no real separation between the first and second solvation shell, which is due to the frequent structural changes for Ag$^+$. For Ag$^{2+}$, the first peak lies at 2.21 Å followed by another small peak, which can be assigned to the fifth water ligand.

![RDF for Ag$^+$ and Ag$^{2+}$](image)

Fig. 4.9: The RDFs and their integrals of Ag$^+$ and Ag$^{2+}$ ion in water solvent. The left y-axis shows the RDF values and the right y-axis shows the coordination number of the silver ion. The x-axis shows the Ag-O distance.

The Ag$^+$ aqua complex, while highly flexible, is approximately tetrahedral.
The structure fluctuates predominantly between a tetrahedral and a square planer configuration. However, both three and five coordinated structures were also encountered as shown in Figure 4.10. The Ag\(_{2}^{+}\) complex, on the other hand, remains mostly a square pyramidal structure. The apex of the pyramid fluctuated between two possible sides of the square base as shown in Figure 4.11.

![Fig. 4.10: Structures of Ag\(^{+}\) encountered in the single ion equilibration calculation.](image1)

![Fig. 4.11: Structures of Ag\(^{2+}\) encountered in the single ion equilibration calculation.](image2)

To probe the mechanism of the electron transfer in this case, we employed the same framework as for the Ru\(^{2+}/Ru^{3+}\) electron transfer. Two silver ions, Ag\(^{+}\) and Ag\(^{2+}\), are put in a box with 64 water molecules, with a temperature of 325 K and a time step of 0.5 fs. After performing an NPT simulation to calculate the box size, followed by an NVT equilibration run we used steered MD to find an initial guess path to start the TPS. A total of 350 paths were generated to elucidate the mechanism of electron transfer. As anticipated from the difference in the equilibration structures of the two silver aqua complexes, there is a significant change in the coordination shell of the complexes during ET. Initially, Ag\(^{+}\) has four ligands in a square planar structure, whereas Ag\(^{2+}\) has five ligands in a square pyramidal structure as shown in Figure 4.12. During the electron transfer, both complexes show a square pyramidal structure and finally, when the electron has settled on the other silver atom, the structures have reversed as compared to the
first frame. In the actual path the fifth ligand comes in and goes out frequently, because the structures need more time to settle into the product state.

(a) $t = 0.275$ ps  
(b) $t = 0.3$ ps  
(c) $t = 0.325$ ps

(d) $t = 0.33$ ps  
(e) $t = 0.335$ ps  
(f) $t = 0.35$ ps

Fig. 4.12: Snapshots from a TPS trajectory of the ET between two silver ions. The green sphere depicts the Wannier center position associated with the transferring electron. The silver ion on the left is the donating ion.

**Acknowledgements**

This work is part of the Industrial Partnership Programme (IPP) ’Computational sciences for energy research’ of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO). This research programme is co-financed by Shell Global Solutions International B.V. The calculations were carried out on the Dutch national e-infrastructure with the support of the SURF Cooperative.
4. Direct simulation of self-exchange electron transfer reaction
Dissociation mechanisms and acidity constants of acids in water from path-metadynamics simulations

Acid-base reactions are an important class of chemical processes. Acid dissociation and proton transfer play key roles in countless biochemical, catalytic, and conformational transitions. At the molecular level however, acid dissociation is still poorly understood and difficult to probe both experimentally and computationally. Here, we show with rare-event first principles simulations that acid dissociation can proceed via two distinct mechanisms, one of which involving a meta-stable contact-ion pair intermediate. The free energy profiles obtained with our approach provide associated $pK_a$ values in very good agreement with experiment, but also barrier heights that can be related to spectroscopic rate estimations. The mechanistic insight may ultimately be exploited to control acid-base reactivity, for example in catalysis.

5. Dissociation mechanisms of proton transfer reactions

5.1 Introduction

Acid dissociation and proton transfer in aqueous solution are fundamental processes in uncountably many chemical and biochemical reactions. Modeling the dissociation mechanism and computing the acidity or $pK_a$ of chemical species have therefore been topics of many theoretical studies. Modeling acid dissociation is challenging because the chemical bond breaking and making requires an accurate quantum mechanical electronic structure approach, while at the same time, an explicit treatment of the solvent involves the sampling of many molecular configurations. Density functional theory based molecule dynamics (DFT-MD) simulation is ideally suited for this purpose when combined with an enhanced sampling technique to overcome the transition state barrier of the dissociation process, which is for weak acids a rare event on the picosecond time scale of a straightforward DFT-MD simulation.

Pioneering constrained DFT-MD simulations in 1998 by Trout and Parrinello on the dissociation of water into $H^+$ and $OH^-$ resulted in a first estimate of the free energy profile. Using the dissociating O-H distance as the reaction coordinate (or collective variable, CV), the free energy could be probed from the reactant minimum up to the transition state maximum, at which point the formed hydronium ion would loose another proton that immediately reacts with the $OH^-$ ion, forming again the reactant state. Soon after, using a coordination number constraint as the CV to avoid this quenching of the products, Sprik estimated the $pK_a$ of water and Ivanov and Klein achieved the same for a histidine molecule, although the free energy profiles were still only captured up to the transition state maximum of the first proton transfer step from the acid molecule to the nearest water molecule. To estimate the dissociation free energy of fully separated product ions, Sprik et al. developed an insertion–deletion scheme, which has since then been employed to compute the $pK_a$ of various systems. However, by deleting the proton from the acid and inserting it into the bulk solvent, no information is obtained on the actual dissociation kinetics and the mechanism. Such mechanistic insight was obtained in 2006 by Park et al. for acetic acid using transition path sampling and metadynamics simulations. The former unbiased simulations revealed a mechanism in which the proton was rapidly expelled into the solvent, whereas the metadynamics simulations using 3 CVs showed a step-wise mechanism via an intermediate contact ion pair state. The question, why a different mechanism is observed depending on the applied enhanced sampling method, remains unanswered. Moreover, how general are these findings to other, weaker or stronger, acids, than acetic acid?

Here, we investigate the dissociation mechanisms and calculate the free ener-
gies profiles of four organic acids in explicit water over a wide spectrum of acidities (see Figure 5.1). We employ DFT-MD simulations with path-metadynamics (PMD),\textsuperscript{41,42} which allows us to find the reaction mechanism from the most likely transition path in the space of relevant CVs. The free energy along this path is obtained simultaneously as the negative of the metadynamics bias potential, which is gradually built-up during the simulation to escape the attractive reactant and product states and enhance the sampling of the barrier crossing. In metadynamics, a repulsive bias potential is constructed on-the-fly by adding small Gaussian-shaped potentials that work on two or three CVs to push the dynamics out of the stable reactant and product states. Instead in PMD, the growing bias is one-dimensional and works along a flexible path variable, which itself is a function of several (possibly many) CVs. Starting from a linear guess path connecting the reactant and product states in this CV-space, the average transition path is found by displacing the path at regular time intervals to the sampled CV-density until convergence.\textsuperscript{41,42}

(a) 2-Chloropropanoic acid (2.8) \hspace{1cm} (b) Acetic acid (4.8)

(c) Imidazolium cation (7.0) \hspace{1cm} (d) Phenol (10.0)

\textbf{Fig. 5.1:} \textit{Snapshots from the PMD/DFT-MD simulations of the dissociation reactions of the four acids, with the conjugate base, the hydronium ion, and selected hydrating solvent molecules shown in ball-stick representation. Dotted lines indicating the local H-bond network serve as a guide to the eye. The experimental acidity constants\textsuperscript{158} are shown above in parentheses.}
5. Dissociation mechanisms of proton transfer reactions

5.2 Computational details

5.2.1 Density functional theory based molecular dynamics

The density functional theory (DFT) based molecular dynamics (MD) simulations performed in this study employed the electronic structure calculation module QUICKSTEP\textsuperscript{71} of the CP2K package,\textsuperscript{70} which combines a Gaussian and plane-wave method to calculate the forces on the nuclei used for the Newtonian dynamics. We used a TZVP basis set\textsuperscript{74} and a 300 Ry cutoff for the plane waves. The SCF energy cutoff was $5 \cdot 10^{-7}$ to converge the wave function to good accuracy. We used the BLYP functional\textsuperscript{21,73} in conjunction with Grimme’s D3 parameterized van der Waals correction.\textsuperscript{33} Core electrons are represented by the norm conserving pseudopotentials of Goedecker et al. (GTH).\textsuperscript{72} To circumvent very fine grids, we used the NN50 smoothing of the exchange-correlation density. The time step was 0.5 fs and the CSVR thermostat\textsuperscript{75} was used with a coupling constant of 250 fs to maintain a constant temperature of 325 K. Each system consisted of a single acid or conjugate base molecule and about 60 water molecules in a cubic box subject to periodic boundary conditions and a volume set to match the experimental density. The number of water molecules was 61 with 2-chloropropanoic acid, 62 for acetic acid, 61 for imidazolium, and 59 for phenol. The positions were saved every 5 fs for further analysis and visualisation. After initial equilibration for 20 ps, production runs were computed for about 120 ps.

5.2.2 Path-metadynamics

We augmented the DFT-MD simulations with the path-metadynamics (PMD) method\textsuperscript{41,42} as implemented in PLUMED\textsuperscript{159} to enhance the sampling of the acid dissociation reaction, obtain the reaction free energy profiles, and unravel the reaction mechanisms. PMD enforces the dynamics away from the stable reactant and product states by gradually constructing a repulsive bias along a path variable that is a function of geometric order parameters, or collective variables (CVs), and connects the reactant and product states. For the initial guess of the path, we chose a linear interpolation between the average CV values in the stable states obtained from equilibrium DFT-MD simulations of the acid (reactant state) and conjugate base (product state) in water. The path was a function of the following two CVs.

The first CV quantifies the distance, $r_{xp}$, between the proton and the acid molecule. Here, $x$ is the position of the carboxylic C atom in the cases of acetic acid and 2-chloropropanoic acid, the hydroxylic O in the case of phenol, and the N atom in the case of the conjugate acid of imidazole. The proton position, $q_p$, or

74
rather, the center of the hydronium ion when the proton is in the water solvent, is computed by taking a weighted sum over all positions, \( q \), of the oxygen and nitrogen atoms in the system that can bind the proton:

\[
q_p = \sum_i w(i)q(i)
\]  

(5.1)

with the weights,

\[
w(i) = \frac{\exp \left[ \gamma(n(i) - n_0(i)) \right]}{\sum_i \exp \left[ \gamma(n(i) - n_0(i)) \right]}
\]  

(5.2)

The position \( q_p \) is thus obtained as an exponentially weighted average of all proton acceptor positions, which are here the water and acid oxygen or nitrogen atoms. With a positive number for the parameter \( \gamma \), the weighting selects the oxygen(s) with the highest coordination number with respect to the H-atoms, \( n_{OH} \). Here, \( n_0(i) \) is the default number of hydrogens bonded to atom \( i \) (without the proton), which is 2 for a water oxygen and 0 for the acid atom. The coordination number is computed with:

\[
n_{xy} = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \frac{1 - \left( \frac{r_{ij}}{r_0} \right)^{nn}}{1 - \left( \frac{r_{ij}}{r_0} \right)^{mm}}
\]  

(5.3)

This proton transfer collective variable (PTCV) is not very sensitive to the exact values of the switching function and the same parameters of the switching function work across the whole \( pK_a \) spectrum. The following parameter values were used: \( nn = 12 \), \( mm = 24 \), \( r_0 = 1.3 \text{ Å} \), and \( \gamma = 8 \).

The second CV, \( n_{OO} \), is the coordination number (Eqn. 7.10) of the acid oxygen(s), or nitrogens in the case of the imidazolium cation, with respect to all water oxygen atoms, with \( nn = 14 \), \( mm = 28 \), and \( r_0 = 2.5 \text{ Å} \).

The path is comprised of 30 nodes with fixed ends (i.e. without trailing nodes) and is updated every 0.5 ps, starting from an initial straight guess path in the 2D CV space. The half-life parameter was 0.5 ps. The metadynamics Gaussian potentials were added with a period of 50 fs, a height of 0.05 kcal/mol and a width of 0.045 in normalised path units.

5.3 Results

Figure 5.2 shows the free energy profiles obtained from the PMD simulations of the four acid dissociation reactions, with on the horizontal axis the progress parameter along the average transition path, \( \sigma \), which is zero in the protonated acid state
and one in the fully dissociated state. The free energy minimum of the protonated acid is taken as the zero for the free energy scale. The profiles show for each acid dissociation reaction a clear free energy well at $\sigma = 0$ that is separated by one or two transition state maxima from a rather flat plateau region where $\sigma$ reaches 1 and that is associated to the fully solvated hydronium ion. Here, fully solvated means that our proton transfer CV has reached its maximum distance of about half the box length and the hydronium ion is hopping over solvent molecules in the third and fourth coordination shells away from the acid molecule.

![Graph](image.png)

**Fig. 5.2:** Acid dissociation free energy profiles along the optimised transition path, quantified by the reaction progress parameter, $\sigma$, which is zero in the reactant state and one in the fully dissociated state. An intermediate minimum can be seen in the profiles for acetic acid and phenol at $\sigma \approx 0.5$, which is associated with a meta-stable contact ion pair.

The acidity constants are calculated from the reaction free energies, as follows:

$$pK_a = \frac{F(\sigma = 1) - F(\sigma = 0)}{k_BT \ln[10]}.$$  \hspace{1cm} (5.4)

The agreement with the experimental values, seen in Table 5.1, is within 1 $pK_a$ unit, which is very good considering the accuracy of the DFT BLYP+D3 level of theory for the electronic structure description, the finite PMD sampling length, the finite system size effects, and the neglect of nuclear quantum effects.

Being able to compute the acid dissociation free energy profile all the way to the product state of fully separated hydronium and conjugate base is here possible because of the non-local nature of the first CV, which captures the distance between the proton and the acid molecule without explicitly specifying which
5.3. Results

Table 5.1: Computed acid dissociation free energies (in kcal/mol) and resulting acidity constants compared to the experimental values.

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\Delta F$</th>
<th>$pK_a$</th>
<th>exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-chloropropanoic acid</td>
<td>2.5</td>
<td>1.8</td>
<td>2.8</td>
</tr>
<tr>
<td>acetic acid</td>
<td>6.7</td>
<td>4.9</td>
<td>4.8</td>
</tr>
<tr>
<td>imidazolium cation</td>
<td>10.5</td>
<td>7.6</td>
<td>7.0</td>
</tr>
<tr>
<td>phenol</td>
<td>14.2</td>
<td>10.3</td>
<td>10.0</td>
</tr>
</tbody>
</table>

Solvent molecules or H-bond wires are involved in the proton transfer. It is now clear from Figure 5.2 that computing the free energy profile only up to the first transition state maximum, as was often done when using a single O-H distance or coordination number CV, leads to a significant systematic error in the $pK_a$ estimation. Moreover, this error is different for each acid due to the long-range undulations in the free energy profile after this first transition state, which turn out to be different for each acid.

The most striking feature in the dissociation free energy profiles, after the initial climb from the reactant state minimum, is the second minimum at $\sigma \approx 0.5$, which is very shallow in the cases of 2-chloropropanoic acid and the imidazolium cation, but clearly visible for acetic acid and phenol. For the latter two, an intermediate meta-stable contact-ion pair is formed by the hydronium ion and the conjugate base after the first proton transfer. Interestingly, these differences between the free energy profiles can be associated with two different dissociation mechanisms, a stepwise mechanism and a concerted long-range mechanism, which were also observed in the work by Park et al.:\textsuperscript{157}

For 2-chloropropanoic acid, the free energy profile is rather flat after the 3.8 kcal/mol low transition state barrier for the initial proton transfer from the acid molecule. And the reaction mechanism observed in the simulations is predominantly the long-range mechanism, in which the proton immediately transfers via a first coordination shell water molecule to the second (or even third) coordination shell with two (or three) concerted Grotthuss mechanistic jumps. Instead for acetic acid, also a carboxylic acid, the free energy profile exhibits an intermediate contact ion pair state and two transition states, with the first maximum at $\sigma \approx 0.3$ being higher than the second maximum at $\sigma \approx 0.7$. The dissociation mechanism for acetic acid is mostly stepwise, but also the concerted long-range mechanism was observed occasionally. The imidazolium free energy profile does not show the contact-ion pair minimum, and, similar as for 2-chloropropanoic acid, the dissociation mechanism is concerted long-range. And finally for phenol, we see a clear contact-ion pair intermediate state and two transition states with the second
5. Dissociation mechanisms of proton transfer reactions

Table 5.2: Coordination numbers of the acid and conjugate base atoms (X=(O, N)) by water oxygen and hydrogen atoms obtained from integration of the RDFs up to 3.0 Å for \( n_c(XO) \) and up to 2.0 Å for \( n_c(XH) \). The solvent reorganization is largest for phenol deprotonation, followed by acetic acid.

<table>
<thead>
<tr>
<th>Compound</th>
<th>( n_c(XO) )</th>
<th>( n_c(XH) )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>acid</td>
<td>base</td>
</tr>
<tr>
<td>2-chloropropionic acid</td>
<td>1.7</td>
<td>2.2</td>
</tr>
<tr>
<td>acetic acid</td>
<td>1.6</td>
<td>2.4</td>
</tr>
<tr>
<td>imidazolium cation</td>
<td>1.1</td>
<td>1.2</td>
</tr>
<tr>
<td>phenol</td>
<td>2.0</td>
<td>3.1</td>
</tr>
</tbody>
</table>

transition state higher than the first. The phenol dissociation follows mainly the stepwise mechanism. In both mechanisms, the proton transfer is concerted with changes in the hydration structure of the donating and accepting molecules in agreement with previous pioneering work on the hydronium ion in water\(^9\) (see also Figure 5.6 in the Appendix).

Why some acids dissociate predominantly in a stepwise mechanism via a metastable contact-ion pair before complete separation, whereas other acids display a fast separation via a concerted long-range mechanism can be understood, at least for these four acids, from the solvent reorganization that takes place during the dissociation process. This solvent reorganization is quite well captured by the radial distribution functions (RDFs) of the solvent atoms around the acid oxygen or nitrogen atoms. We computed the RDFs from equilibrium DFT-MD simulations of each acid in water solvent and of each conjugate base in water solvent (see Figures 5.9 and 5.10 in the Appendix). Comparison of these RDFs before and after dissociation shows that the deprotonation reaction of 2-chloropropionic acid and of imidazolium, which both follow the long-range mechanism, is accompanied with only a small reorganization of the solvent around the proton donating acid atom. Instead for acetic acid, and even more so for phenol, the solvent structure changes significantly upon deprotonation. In particular, the number of water molecules that donate a hydrogen bond to the acid atom is larger for the deprotonated base than for the acid, as shown by the acid atom coordination numbers compiled in Table 5.2. It is this solvent reorganization, which takes place concerted with the contact ion pair breakup, that underlies the second transition state barrier and the resulting stepwise mechanism in the cases of acetic acid and phenol dissociation.
5.4_summary

We have shown here that the entire free energy profile of acid dissociation can now be computed at the DFT-MD level of theory with explicit water solvent using two descriptive CVs combined in a path-metadynamics simulation, from which reliable acidity constants can be calculated. Applying our simulation approach to four common acids spanning \( pK_a \) values from 2.8 to 10.0, we observe two different reaction mechanisms. Acid dissociation can proceed through a two-step mechanism via an intermediate contact-ion pair or via a fast long-range mechanism in which the proton jumps quickly to the second coordination shell without forming the contact-ion pair. The propensity for one mechanism or the other can be understood from the amount of solvent reorganization, concerted with the acid dissociation, that is needed to hydrate and stabilize the conjugate base. Our approach provides unique mechanistic insight and can be particularly useful for systems that are difficult to model with implicit solvent models, for example at metal-oxide surfaces, clays, inside zeolites and MOFs, and in enzymes and biomolecular assemblies.

5.5 Appendix

Figure 5.3 shows the evolution and convergence of the path towards the average transition pathway in the CV space spanned by the PTCV, \( r_{po} \), (x-axis) and the coordination number, \( n_{OO} \), (y-axis) starting from a straight line (green) at \( t = 0 \). During the 120 ps simulation, many reactive transitions are sampled (see also Figure 5.4). Instead of gradually decreasing the Gaussian size or using the well-tempered metadynamics method, we converge the free energy estimate by taking the time average of the metadynamics bias potential starting from the moment that the path-CV progress variable, \( \sigma \), has visited the whole range along the path from zero to one and back.

The dynamics in the space of the two CVs is shown for all four acids in Figure 5.5. Note that for each acid in the reactant state (at \( r_{XP} = 0 \)), the coordination number value, \( n_{XO} \), of the protonated acid atom with respect to the water oxygen atoms is close to zero, and must first increase, by drawing the closest water molecule nearer to the acid, before dissociation occurs and the proton can transfer away from the acid.

Once the excess proton reaches the second hydration shell, the O-OW coordination number does not change much, indicating the stable hydration of the conjugate base whereas in case of conjugate acid of imidazole this O-OW coordination number drops down due to poor hydration of the neutral N atom. As
5. Dissociation mechanisms of proton transfer reactions

Fig. 5.3: Time evolution of path for the acetic acid dissociation shown in the two-dimensional CV-space, starting from a linear initial guess path (green). The path converges in the first 30 ps. Grey dots show the sampled configurations in the CV space.

expected, imidazolium has a very different solvation as compared to the other acids. This is because of the net +1 charge of the molecule which makes it relatively hydrophobic and inhibits strong H bonding with the solvent water molecules as shown in Figure 5.5. After proton dissociation, the conjugate base of all other acids is negatively charged and hence have strong interactions with the solvent water, whereas the conjugate base of imidazolium is neutral and has weaker solvation which is evident from the low values of \( n_{NO} \) in Figure 5.5.

Plumed input examples for the PMD simulation with the proton distance CV are available on the Plumed nest website (www.plumed-nest.org), including the source code for the PTCV under Plumed Nest Project ID: 19.034. The source code for our PMD implementation is found in Plumed Nest Project ID: 19.033.

As outlined in the main text, the PMD simulations revealed two mechanisms for acid dissociation: a stepwise mechanism and a concerted long-range mechanism. Figure 5.6 illustrates the two mechanisms with a scheme of the proton transfer from the acid molecule along an H-bond wire formed by nearby solvent molecules.

The stepwise mechanism occurs via formation of a meta-stable intermediate contact ion pair. The water molecule closest to the acid proton (shown in black) accepts an H-bond from the acid molecule and accept another H-bond from a nearby solvent molecule (shown in red). It also donates two H-bonds. After proton transfer, the formed \( \text{H}_3\text{O}^+ \) molecule does not accept any H-bonds, but donates three H-bonds. The first proton transfer is thus concerted with the H-bond donating solvent molecule (red) breaking its H-bond, therewith stabilising
(a) 2-Chloropropanoic acid

(b) Acetic acid

(c) Imidazolium cation

(d) Phenol

Fig. 5.4: Time evolution of the path progress parameter $\sigma$ for all the four acids during the PMD simulations.

The concerted long-range mechanism takes place in the cases that the acid dissociation requires only little solvent reorganisation to stabilise the conjugate base, as is shown in the manuscript for 2-chloropropanoic acid and imidazolium. In those cases, fluctuations in the H-bond network in the second or third solvation shell can result in a proton transfer as depicted on the right-hand-side in Figure 5.6. Here the H-bond wire, along which the proton transfers, is shown in black. In blue, two second shell solvent molecules are shown that donate and accept H-bonds to the first water molecule in the H-bond wire (starting from the acid). In red, a water molecule is shown that donates an H-bond to the second water in the wire.
5. Dissociation mechanisms of proton transfer reactions

(a) 2-Chloropropanoic acid  (b) Acetic acid

(c) Imidazolium cation  (d) Phenol

Fig. 5.5: Sampled two dimensional collective variable space from the PMD simulations.

Fig. 5.6: Schematic diagrams of the two acid dissociation mechanisms observed in the PMD simulations: the stepwise mechanism (left) and the concerted long-range mechanism (right).

When this H-bond breaks, the proton can jump directly via the first water to the second water in the chain, forming a threefold coordinated hydronium ion. Figure 5.8 shows an example of the concerted long-range mechanism taking place.
Fig. 5.7: Eight PMD simulation snapshots showing phenol dissociation taking place via the stepwise mechanism. Phenol and three solvent molecules forming the H-bond wire along which the proton jumps are shown in ball-and-stick; the rest of the solvent is represented by sticks. The semi-transparent blue sphere shows the proton position as estimated by the PTCV. The contact ion pair is seen in snapshot four at $t = 19.660$ ps.

in the PMD simulation of 2-chloropropanoic acid.

Fig. 5.8: Four PMD simulation snapshots showing the concerted long distance mechanism taking place for dissociation of 2-chloropropanoic acid. The acid and three involved water molecules are shown in ball-and-stick representation; the other solvent is shown as sticks. The semi-transparent blue sphere indicates the proton position as quantified by the PTCV.

We capture the solvent reorganization around acid molecules by comparing the radial distribution functions (RDFs) of the solvent atoms around the acid oxygen or nitrogen atoms before and after dissociation. We computed the RDFs from
equilibrium DFT- MD simulations of each acid in water solvent and of each conjugate base in water solvent (see Figures 5.9 and 5.10). Comparison of the RDFs before and after dissociation shows that dissociation of 2-chloropropanoic acid and of imidazolium, which both follow the long-range mechanism, is accompanied with only a small reorganization of the solvent around the proton donating acid atom. Instead for acetic acid, there is no O-O peak seen before dissociation (black line in Figure 5.9b), in contrast to the case of 2-chloropropanoic acid, whereas after dissociation (blue line), the O-O peak is larger than that for 2-chloropropanoic acid, indicating a significantly larger solvent reorganization. Note also that the pre-dissociation peak seen for 2-chloropropanoic acid at $r \approx 2.5 \text{ Å}$ shows that the acid forms a relatively strong hydrogen bond to the nearest water molecule, whereas for acetic acid the nearest water molecule is on average further away. This led us to employ the second CV in the PMD simulations to pull this water closer and make the first proton transfer possible. From the O-H RDFs (Figure 5.10), we see that the 2-chloropropanoic acid oxygen accepts on average 1.8 hydrogen bonds after dissociation (dashed blue line), which are already partly formed (1.4 hydrogen bonds) when the acid is still intact. Instead for acetic acid these H-bond donors are practically absent before dissociation and are formed to stabilise the negative charge of the conjugate base only after the proton transfers from the first shell to the second shell. This latter solvent reorganization underlies the second transition state barrier. For phenol, for which we only observed the step-wise mechanism, these changes in the hydration of the acid during dissociation are even more pronounced. Before dissociation the acid oxygens accepts one weakly bonded H-bond, then, when forming the contact ion pair the proton accepting water molecule turns from an H-bond acceptor into a second H-bond donor, and finally, crossing the second barrier involves recruitment of a third H-bond donating solvent molecule to stabilise the remaining negatively charged phenol oxygen.
5.5. Appendix

(a) 2-Chloropropanoic acid

(b) Acetic acid

(c) Imidazolium cation

(d) Phenol

Fig. 5.9: Radial distribution functions of water oxygen atoms around the acid oxygen atom, or nitrogen atom in the case of imidazolium, in the protonated and deprotonated states. For the carboxylic acids in the reactant state, \(O_p\) and \(O_d\) refer to the protonated and deprotonated oxygens, respectively, whereas in the product state the distribution is averaged over the two oxygens. Similarly, for imidazolium in the product state, \(N_p\) and \(N_d\) refer to the protonated and deprotonated nitrogens, and in the reactant state, the average over the two nitrogens is taken. The dashed lines show the cumulative coordination number (using the right-hand-side y-axis) computed from integration of the RDFs.
5. Dissociation mechanisms of proton transfer reactions

(a) 2-Chloropropanoic acid
(b) Acetic acid
(c) Imidazolium cation
(d) Phenol

Fig. 5.10: Radial distribution functions of water hydrogen atoms around the acid oxygen atom, or nitrogen atom in the case of imidazolium, in the protonated and deprotonated states. For the carboxylic acids in the reactant state, $O_p$ and $O_d$ refer to the protonated and deprotonated oxygens, respectively, whereas in the product state the distribution is averaged over the two oxygens. Similarly, for imidazolium in the product state, $N_p$ and $N_d$ refer to the protonated and deprotonated nitrogens, and in the reactant state, the average over the two nitrogens is taken. The dashed lines show the cumulative coordination number (using the right-hand-side y-axis) computed from integration of the RDFs.
5.5. Appendix

Acknowledgements

This work is part of the Industrial Partnership Programme (IPP) 'Computational sciences for energy research' of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO). This research programme is co-financed by Shell Global Solutions International B.V. The calculations were carried out on the Dutch national e-infrastructure with the support of the SURF Cooperative.
5. Dissociation mechanisms of proton transfer reactions
Chapter 6

Accurate calculation of zero point energy from molecular dynamics simulations of liquids and their mixtures

The two-phase thermodynamic (2PT) method is used to compute the zero point energy (ZPE) of several liquids and their mixtures. The 2PT method uses the density of states (DoS), which is computed from the velocity autocorrelation function (VAC) obtained from a short classical molecular dynamics trajectory. By partitioning the VAC and the DoS of a fluid into solid and gas-like components, quantum mechanical corrections to thermodynamical properties can be computed. The ZPE is obtained by combining the partition function of the quantum harmonic oscillator with the vibrational part of the solid-like DoS. The resulting ZPE is found to be in excellent agreement with both experimental and ab initio results. Solvent effects such as hydrogen bonding and polarization can be included by the utilization of ab initio molecular dynamics (DFT-MD) simulations. It is found that these effects significantly influence the DoS of water molecules. The obtained results demonstrate that the 2PT model is a powerful method for efficient ZPE calculations, in particular to account for solvent effects and polarization.
6. Accurate calculation of zero point energy from molecular dynamics simulations

6.1 Introduction

Quantum chemical calculation of molecular energies has become a standard tool for various research applications, including thermochemistry, catalysis and reaction kinetics. With constant improvement of commonly used DFT calculations, the error in molecular energy has reduced to only a few kcal/mol requiring other sources of errors to improve as well. One such source of error in many cases has been shown to be quantum nuclear effects, in particular, the zero point energy (ZPE). The ZPE represents the energy of a system at absolute zero, corresponding to the lowest vibrational energy level. Perhaps one of the most important consequences of ZPE is the effect on bond strength as measured by the kinetic isotope effect.\textsuperscript{160–162} The strength of a chemical bond is related to the vibrational energy of that bond. A regular C-H stretch vibration has a frequency of approximately 2900 cm\(^{-1}\), corresponding to a ZPE of 17 kJ/mol. When hydrogen is substituted with deuterium the mass of the atoms doubles, decreasing the C-D stretch vibration to 2100 cm\(^{-1}\), corresponding to a ZPE of approximately 13 kJ/mol. This decrease of 4 kJ/mol means that more energy is required to break a C-D bond compared to the more energetic C-H bond. Chemical reactions on deuterated substrates are thus generally slower than the protonated variant, as the higher energy barrier leads to a lower rate constant.

This effect is even more important in water and aqueous systems. The O-H stretch vibration in water has a frequency of approximately 3700 cm\(^{-1}\), corresponding to a ZPE of 21 kJ/mol which is quite significant.\textsuperscript{163} D\(_2\)O and T\(_2\)O have lower stretch frequencies, corresponding to a ZPE of 15 and 12 kJ/mol respectively. The higher bond strength of D\(_2\)O means that it is less likely to self-ionize in comparison to regular H\(_2\)O, resulting in a higher pH of 7.43 for D\(_2\)O compared to 7.0 for H\(_2\)O. Heavy water also forms stronger hydrogen bonds, as is observed by the increase of the melting point from 273 K for regular water to 277 K for heavy water.\textsuperscript{164}

It has also been demonstrated that hydrogen bonding itself affects the zero point energy of solvated molecules.\textsuperscript{163} When a hydrogen bond is formed between two water molecules the intramolecular vibrations are perturbed, leading to an increased contribution of the bending mode. This consequently leads to an overall decrease of ZPE in the molecule as the vibrational contribution is decreased.\textsuperscript{165} Hydrogen bonding organic molecules are also affected by this ZPE contribution. It has been shown that ring substitutions in anisole change the structure of the hydrogen bond, this in turn affects the zero point energy of the molecule as the H-bond structures perturb the bond vibrations to a varying degree.\textsuperscript{166} All of these effects suggests the need of a thorough study of various liquid systems and analyze
6.2 Theoretical Background

6.2.1 Density of States

The DoS of a system can be obtained from the mass-weighed velocity autocorrelation function (VAC):
6. Accurate calculation of zero point energy from molecular dynamics simulations

\[ C(t) = \sum_{j=1}^{N} \sum_{k=1}^{3} m_j \langle v^k_j(t)v^k_j(t+\tau) \rangle \tau \]  

(6.1)

in which \( m_j \) is the mass of atom \( j \), \( v^k_j(t) \) is the velocity of atom \( j \) along axis \( k \), \( v^k_j(t+\tau) \) is the velocity of the same atom at a time delay \( \tau \), averaged over the different time delay \( \tau \) in a trajectory to get a time series. \( N \) is the number of particles in the system. A typical velocity autocorrelation function is shown in Figure 6.1. The function is 1 when the velocity is identical to the initial velocity, such as at zero time lag. At larger time lags the function quickly decays to zero as the function loses its memory due to the random collisions resulting in a diffusive motion. Because the VAC represents periodicity of a system in the time domain, a Fourier transform can be used to obtain these periodic normal modes in the frequency domain; the density of states, \( S(\nu) \), as follows:

\[ S(\nu) = \frac{2}{k_B T} \lim_{\tau \to \infty} \int_{-\tau}^{+\tau} C(t) \exp(-i2\pi\nu t) \, dt \]  

(6.2)

where \( T \) is the temperature of the system. The physical definition of the density of states at a specific frequency \( \nu \) is that it represents the density of normal modes at that frequency. Specific domains of the DoS can thus be related to chemical and physical properties of the molecule in question. The high frequency range is dominated by vibrational motion, the DoS in this region can thus be compared to the infrared absorption spectrum of a molecule. Lower frequencies correspond

Fig. 6.1: Velocity autocorrelation function of TIP3P water, normalized to one time lag \( (\tau) = 0 \)

92
to other types of motions such as rotation and translation. A typical density of states function is shown in Figure 6.2. The characteristic IR vibrations can clearly be seen, such as the C-H stretch at 3000 cm\(^{-1}\). Rotational and translational normal modes are contained in the bump below 100 cm\(^{-1}\), which does not occur in a normal IR spectrum. Because the DoS is a measure of the density of normal modes, integration over the entire function gives the total amount of normal modes of the system; the degrees of freedom.

\[
\int_0^\infty S(\nu) \, d\nu = 3N \tag{6.3}
\]

The density of states at the origin, \(S(\nu = 0)\), corresponds to the diffusive modes of the system. It can be used to calculate the self-diffusion coefficient of a liquid as follows:

\[
D = \frac{S(0)k_BT}{12Nm} \tag{6.4}
\]

wherein \(N\) is the number of particles with mass \(m\). With the obtained density of states function, and the assumption that the system is fully harmonic, the partition function of a system can be obtained using various models, such as the Debye theory of crystals.\(^{170,182}\) However, the application of such simple models are not straightforward to fluids since they exhibit anharmonicity and diffusion.

\[\text{Fig. 6.2: The density of states of liquid ethanol.}\]

\[\int_0^\infty S(\nu) \, d\nu = 3N \tag{6.3}\]

\[D = \frac{S(0)k_BT}{12Nm} \tag{6.4}\]
6.2.2 The 2PT method

The main limitation of the quantum harmonic oscillator approximation discussed in the previous section is the assumption that the system is completely harmonic. This is not a problem for highly crystalline solids, but presents issues when applied to amorphous liquids and gases. One of these issues is the non-zero density of states at $\nu = 0$ due to diffusion and anharmonicity. To account for the anharmonic motions in liquids a refined method was proposed by Lin et al., which splits the density of states in the solid and gaseous components.\(^{171}\) This was based on the observation that the liquid DoS resembles a superposition of the solid and gaseous states. By decomposing the DoS it is possible to obtain a fully harmonic component (the solid DoS) and an anharmonic component (the gaseous DoS), onto which the appropriate weighing functions can be used to further calculate thermodynamic properties.

$$S_{\text{liquid}}(\nu) = S_{\text{gas}}(\nu) + S_{\text{solid}}(\nu) \quad (6.5)$$

In order to decompose the DoS however, one of the components has to be computed analytically. This is accomplished by modeling the gas-like component as a hard sphere fluid. The velocity autocorrelation function of a hard sphere gas can be calculated analytically as follows:

$$C_{\text{HS}}(t) = \frac{3kT}{m} \exp(-\alpha t) \quad (6.6)$$

in which $\alpha$ is the Enskog friction constant pertaining in the hard sphere model. By applying the Fourier transforms as described in the previous section, an analytic form of the hard sphere DoS, and therefore the gas-like DoS can be obtained as follows:

$$S_{\text{HS}}(\nu) = S_{\text{gas}}(\nu) = \frac{S_0}{1 + \left[\frac{S_0 \pi \nu}{6fN}\right]^2} \quad (6.7)$$

where $S_0$ is the total density of states at $\nu = 0$. Here, $f$ is the fluidity of the system, i.e. the fraction of particles $N$ corresponding to the diffusive hard sphere fluid. The fluidity is proportional to the diffusivity $\Delta$ of a system, which can be calculated from the self-diffusion.

$$\Delta(T, \rho, m, S_0) = \frac{2S_0}{9N} \left(\frac{\pi k_b T}{m}\right)^{1/2} \rho^{1/3} \left(\frac{6}{\pi}\right)^{2/3} \quad (6.8)$$

An expression for the fluidity $f$ can then be obtained by relating the diffusivity of
6.2. Theoretical Background

the system to the pure hard sphere diffusivity with the Carnahan-Starling equation of state:

$$2 \left( \frac{f_5}{\Delta^3} \right)^{3/2} - 6 \left( \frac{f_5}{\Delta^3} \right) - \left( \frac{f_7}{\Delta^3} \right)^{1/2} + 6 \left( \frac{f_5}{\Delta^3} \right)^{1/2} + 2f = 2$$ (6.9)

Because the 2PT method depends on the motion of atoms, molecules with more than one atom require decomposition into the primary degrees of motion. This is accomplished by decomposing the velocity into translational, rotational, and vibrational components as follows:

$$v(t) = v_{trans}(t) + v_{rot}(t) + v_{vib}(t)$$ (6.10)

The translational component of the velocity is obtained by calculating the mass weighed velocity for each molecule

$$v_{trans}(t) = \frac{1}{M_i} \sum_{k=1}^{3} \sum_j m_j v_{j}^k(t)$$ (6.11)

where $m_j$ is the mass of atom $j$ and $M_i$ is the total mass of all atoms $j$ in molecule $i$. The rotational component can be computed from the angular velocity $\omega$ of each molecule

$$v_{rot}(t) = \omega(t) \times v(t)$$ (6.12)

The angular velocity is obtained from the inverse moment of inertia tensor $I_i^{-1}$ and the angular momentum $L(t)$ of each molecule

$$\omega(t) = I_i^{-1} \times L(t)$$

$$L(t) = \sum_j m_j \left( r_j \times v_j(t) \right)$$ (6.13)

where $r_j$ is the distance of an atom to the center of mass of the molecule. The vibrational velocity is then obtained by subtracting the rotational and translational components from the total velocity.

$$v_{vib}(t) = v(t) - v_{trans}(t) - v_{rot}(t)$$ (6.14)

The resulting density of states after decomposition into a solid and diffusive component are shown in Figure 6.3a, diffusive effects are only relevant at extremely low frequencies whereas the remaining major part of the DoS is considered as solid (and therefore harmonic). Figure 6.3b depicts the density of states of the various
6. Accurate calculation of zero point energy from molecular dynamics simulations

(a) The solid and diffusive states
(b) The different velocity components

Fig. 6.3: The decomposed density of states of liquid ethanol

molecular motions. Rotation and translation dominate the lower frequencies, while anything above 500 cm\(^{-1}\) is the result of vibrations.

6.2.3 Zero point vibrational energy

The zero-point energy results from the uncertainty principle. If the lowest possible energy state would have zero energy, both its potential and kinetic energies would have to be zero. Zero kinetic energy would mean that the momentum would be exactly zero, therefore the uncertainty in the momentum \((\Delta p)\) as well as in the position \((\Delta x)\) would be zero, which is not allowed according to the Heisenberg uncertainty principle\(^{183}\) as depicted by Eq. 6.15. Hence the need for a nonzero ground-state energy.

\[
\Delta x \Delta p \geq \frac{1}{2} \hbar
\]  

(6.15)

In the 2PT method, molecular vibrations are represented as quantum harmonic oscillators, which means that the zero point energy of molecules can be calculated from the lowest energy level of the corresponding harmonic oscillator. The permitted energy levels can be calculated from the Schrödinger equation of a quantum harmonic oscillator:

\[
E_n = (n + \frac{1}{2})\hbar \omega = (n + \frac{1}{2})\hbar \nu
\]

\[
\omega = \left( \frac{k_f}{m} \right)^{1/2}, \nu = \frac{\omega}{2\pi}
\]  

(6.16)
where \( \omega \) is the oscillation frequency defined in terms of the force constant \( k_f \) and mass \( m \). When a system is considered with only classical mechanics the energy \( E \) of a linear harmonic oscillator can have any value. However, a more precise quantum mechanical description would lead to a non-zero ground state energy, at energy level \( n = 0 \), which is:

\[
E_{\text{ZPE}} = \frac{1}{2} \hbar \nu
\]  

(6.17)

For a group of atoms that constitutes a set of normal modes, zero-point energy of each mode can be determined independently thus the total zero point energy can be obtained using:

\[
E_{\text{ZPE}} = \sum_{i=1}^{3N-6} \frac{1}{2} \hbar \nu_i
\]  

(6.18)

The density of states represents the mass-weighed distribution of normal modes of a molecular system, the vibrational component, \( S_{\text{vib}}(\nu) \) of the DoS can thus be used to calculate the ZPE of a system.

\[
E_{\text{ZPE}} = \int_{0}^{\infty} S_{\text{vib}}(\nu) \frac{\hbar \nu}{2} d\nu
\]  

(6.19)

Because the ZPE represents vibrations at absolute zero, it is problematic to obtain experimental values. Instead, reported experimental values are extrapolated results from spectroscopic constants. This makes accurate and fast theoretical predictions desirable, and which serves the motivation of this article.

### 6.3 Computational method

#### 6.3.1 Forcefield MD simulations

Molecular dynamics simulations were carried out with the open source LAMMPS software package. Initial configurations of 1000 solvent molecules were generated using the packmol software package. The atomic interactions were modeled with the General Amber Force Field (GAFF). The electrostatic interactions were cutoff at a distance of 10 Å and augmented with the pppm long-range solver. The partial charges of the solvent molecules were fitted using the restrained electrostatic potential (RESP) model, which uses Hartree Fock (HF) calculations to obtain accurate charge distributions for molecules. RESP charges were calculated with the Gaussian software package at the HF/6-31G* level of detail. The initial box configurations were warmed-up to a temperature of 300 K in the NVT
ensemble using a Nosé–Hoover thermostat over the course of 0.1 ns with a 1 fs timestep for the outer time step. A four layer RESPA multi-timestep algorithm was used to evaluate middle range interactions (5.25-9.0 Å) twice per time step, short range interactions (0-5.25 Å) four times per time step, and intra-molecular interactions eight times per time step. The systems were then further equilibrated in both the NPT ensemble at 1 bar, and NVT at 300 K for 1 ns each.

### 6.3.2 DFT-MD simulations

ZPE calculations were also tested using density functional theory based molecular dynamics (DFT-MD) simulations, where the electronic structure and hence the energy and forces are computed at every timestep using DFT, instead of a force field potential. Despite the large increase of computational power required, DFT-MD offers several advantages compared to forcefield simulations. Chemical processes involving bond breaking or forming, such as proton transfer, are not possible in standard forcefield simulations, thus preventing accurate analysis of hydronium ions and other reactive species. Polarization effects are also not included in the forcefield simulations, limiting the effects of solvents and hydrogen-bonding on molecular vibrations. The open source software package CP2K was used to perform DFT-MD simulations. Due to the increased cost of the simulation, a box with 128 molecules was used. The exchange correlation (XC) was computed with the BLYP functional using the TZV2P basis set, with DFT-D3 dispersion correction. The norm-conserving pseudopotentials of Goedecker et al. (GTH) were applied to replace the core electrons. The simulations were performed in the NVT ensemble at a temperature of 330 K with a 0.5 fs timestep. A higher temperature 330 K was used to counter the slow dynamics of BLYP at 300 K.

As discussed in the theoretical background, the 2PT method obtains results from the density of states, which is in turn calculated from the velocity autocorrelation function. To obtain accurate results for the density of states, velocities should be sampled with at least double the frequency of the fastest molecular vibration. For liquids a limit can be set at 4000 cm$^{-1}$ as this accounts for nearly all IR absorption bands. This frequency corresponds to a period of approximately 8 fs, thus positions and velocity’s must be sampled at least every 4 fs.

### 6.4 Results

#### 6.4.1 Bulk solvents

A total of six solvents were analyzed on their performance in reproducing accurate results for the zero point energy with the 2PT method. These results were then
compared to the experimental values, obtained from the NIST Standard Reference Database. As described in the theoretical background, the zero point energy cannot be measured directly experimentally. The extrapolations used in the NIST database are of very high quality however, and can thus be viewed as accurate.

Table 6.1: Zero point energies (in kJ/mol) computed with the 2PT method for six liquids using the GAFF forcefield and for DFT/BLYP+D3 water (bottom row). Our MD results are compared to gas-phase DFT calculations and experimental numbers.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>2PT</th>
<th>Gas phase</th>
<th>Experimental</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethyl acetate</td>
<td>310.40</td>
<td>298.35</td>
<td>308.67</td>
</tr>
<tr>
<td>DMF</td>
<td>258.75</td>
<td>259.10</td>
<td>259.83</td>
</tr>
<tr>
<td>Ethanol</td>
<td>202.15</td>
<td>210.40</td>
<td>202.98</td>
</tr>
<tr>
<td>Acetonitrile</td>
<td>114.29</td>
<td>114.60</td>
<td>115.11</td>
</tr>
<tr>
<td>Chloroform</td>
<td>51.96</td>
<td>50.03</td>
<td>51.36</td>
</tr>
<tr>
<td>Water (TIP3P)</td>
<td>48.86</td>
<td>54.43</td>
<td>53.88</td>
</tr>
<tr>
<td>Water (BLYP)</td>
<td>49.96</td>
<td>54.43</td>
<td>53.88</td>
</tr>
</tbody>
</table>

As demonstrated in Table 6.1, the 2PT method can produce the zero point energy of pure liquids with very high accuracy, deviating from the experimental values by only 2 kJ/mol at most. The exception being water, which performs significantly worse than the other solvents. To assess the effect of bulk solvent on the ZPE, DFT gas phase calculations of a single molecule were also performed with the Gaussian 09 software package at the BLYP/6-311G* level of theory. Zero point energy values obtained with the Gaussian package perform slightly worse compared to the experimental values. This in part can be attributed to the inclusion of the solvent effects in the forcefield simulations that are used in the 2PT method.

Including polarization and higher order many body effects using DFT-MD does not influence much the value of the ZPE as calculated by the 2PT method as shown in Table 6.1. The similarity between the TIP3P and BLYP results might initially indicate that polarization effects do not significantly influence the zero point energy. However, closer examination of the DoS of the classical TIP3P water and BLYP water in Figure 6.4 shows that BLYP water exhibits significant peak
broadening whereas TIP3P has sharp and defined peaks. This difference in DoS is caused by the way Coulombic interactions are calculated in both models. Classical TIP3P water uses a constant point charge for each atom whereas DFT-MD computes the electron density at each simulation step. Polarization effects significantly influence molecular properties such as bond strength, hydrogen bonding strengths and therefore the vibrational frequency. This allows BLYP water to occupy a much broader range of vibrational frequencies compared to the forcefield models. The remarkable resemblance of the ZPE between the TIP3P water and the BLYP-D3 water seems therefore somewhat fortuitous.

\[ E_{ZP}^{i} = \sum x_i E_{ZP}^{i} \]  

(6.20)

**Fig. 6.4:** The vibrational density of states of liquid water, showing the much broader distributions of the DFT BLYP-D3 model compared to the TIP3P forcefield model.

### 6.4.2 Mixtures of solvents

Earlier studies utilizing the 2PT method have concluded that the method can effectively capture the excess Gibbs free energy in mixtures of classical solvents. It is therefore worth investigating whether similar observations can be made for the ZPE of solvent mixtures. To this end, ethanol-water mixtures with varying concentrations were run with the same conditions as the bulk solvents (1000 molecules, GAFF force field, 300 K, NVT ensemble). Obtained zero point energy values were then compared to the ideal mixing (using Eqn. 6.20) of the experimental numbers for the pure liquids, where the ZPE of the each molecular species is weighed by the molar fraction:
Table 6.2: Zero point energies (in kJ/mol) computed with the 2PT method for five ethanol-water mixtures where \( x \) is the mole fraction of ethanol. Our MD results are compared to the ideally mixed experimental numbers, as explained in the text.

<table>
<thead>
<tr>
<th>( x )</th>
<th>2PT</th>
<th>Experimental</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.96</td>
<td>193.32</td>
<td>196.09</td>
<td>2.77</td>
</tr>
<tr>
<td>0.75</td>
<td>159.66</td>
<td>164.25</td>
<td>4.59</td>
</tr>
<tr>
<td>0.50</td>
<td>120.68</td>
<td>126.36</td>
<td>5.68</td>
</tr>
<tr>
<td>0.25</td>
<td>83.42</td>
<td>88.46</td>
<td>5.01</td>
</tr>
<tr>
<td>0.10</td>
<td>60.32</td>
<td>65.72</td>
<td>5.40</td>
</tr>
</tbody>
</table>

The results presented in Table 6.2 demonstrate that mixing water and ethanol decreases the overall zero point energy of the system. IR spectroscopy studies on ethanol water mixtures have demonstrated that hydrogen bonding between the two molecules results in an overall decrease in intensity of the O-H stretch vibration.\(^{194}\) This is a result of electronic interactions between the molecules, which are not captured accurately as the polarization effects are ignored in the forcefield simulations performed. Ethanol water mixtures are also generally known to reduce in volume due to the non-ideal behavior of mixtures; these effects have previously been captured by the 2PT method and shown to influence the entropy and free energy.\(^{193}\)

### 6.4.3 Performance and efficiency of the 2PT method

The advantage of the 2PT method is the efficiency with which it is able to compute various thermodynamic properties such as free energy, entropy and heat capacity.\(^{195-197}\) Trajectories of only 20 ps are required to obtain full convergence of these properties although no benchmarks of ZPE convergence with the 2PT approach have been published to the best of our knowledge.

Figure 6.5 shows that for large molecular systems, the ZPE converges very quickly, after a simulation time of only 1 ps a stable value is obtained. This quick convergence is in part caused by the size of the system; the average of 1000 molecules is calculated. Convergence for individual molecules requires longer simulations lengths as demonstrated in Figure 6.5b, after 1000 ps a stable result is obtained.
6. Accurate calculation of zero point energy from molecular dynamics simulations

(a) Systems containing 1000 molecules
(b) A single water molecule (in a system of 1000 molecules)

Fig. 6.5: Convergence of the ZPE computed with the 2PT methods for systems of different sizes.

The primary reason to decompose the liquid DoS in a solid and diffusive component is to deal with the anharmonicities in the liquid. This decomposition is based on a hard-sphere approximation of the liquid, which depends on the fluidity of the system following equations 6.8 and 6.9. To analyze the performance of the decomposition, the ZPE and the DoS were calculated for ethanol boxes at different temperatures, where it is expected that the gas-like fraction increases at higher temperatures. Because the ZPE represents the energy of a molecule in its lowest energy state, it should be independent of temperature.

Table 6.3: ZPE and fluidity of ethanol at various temperatures.

<table>
<thead>
<tr>
<th>Temperature [K]</th>
<th>ZPE [kJ/mol]</th>
<th>Fluidity factor</th>
<th>Diffusion coefficient $[10^{-5} \text{ cm}^2/\text{s}]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>201.15</td>
<td>0.033</td>
<td>0.013</td>
</tr>
<tr>
<td>100</td>
<td>202.73</td>
<td>0.046</td>
<td>0.036</td>
</tr>
<tr>
<td>300</td>
<td>202.15</td>
<td>0.25</td>
<td>2.32</td>
</tr>
<tr>
<td>500</td>
<td>202.89</td>
<td>0.63</td>
<td>20.15</td>
</tr>
<tr>
<td>1000</td>
<td>200.26</td>
<td>0.86</td>
<td>105.90</td>
</tr>
</tbody>
</table>

102
6.5. Conclusions

Table 6.3 demonstrates that the 2PT method can accurately compute the ZPE across a wide range of temperatures, deviating by only a few kJ/mol even at extreme temperatures. The increased gas-like nature at high temperature is accounted for by the fluidity factor, which describes the fraction of the system corresponding to diffusive behavior. Figures 6.6a and 6.6b show the solid ethanol contributions, which have near negligible diffusive components. Figures 6.6c and 6.6d show significant diffusive contributions below a frequency of 25 cm$^{-1}$, arising from the more anharmonic and fluidic nature of the system.

6.5 Conclusions

We have used the 2PT method to accurately calculate the zero point energy of several pure liquids and mixtures from short MD trajectories. Simulations of bulk solvents require only a picosecond of simulation time as the ZPE can be averaged.
over all the molecules in the box. However, calculating the ZPE for individual molecules requires a much longer simulation time of 1000 ps to fully converge. For most pure liquids investigated here, the ZPE is not affected very strongly by solvent effects, as seen from the already good performance of the ZPE calculation of the isolated (gas phase) molecule compared to the experimental number. The exceptions are ethanol and ethyl acetate, for which the gas phase estimates are respectively 7.4 kJ/mol too high and 10.3 kJ/mol too low, whereas the PT2 full solvent result is within 2 kJ/mol. Remarkably, for liquid water the opposite is observed: the gas phase number is in good agreement with experiment, whereas the 2PT method applied to liquid TIP3P water results in an almost 10% too low number. Moreover, the 2PT estimate from DFT BLYP-D3 water agrees with the TIP3P result. Considering the excellent performance of the PT2 approach for the other liquids, it may be worthwhile to reexamine the experimental estimate for liquid water. For various mixtures of ethanol and water, the ZPE estimate is seen to be somewhat too low, which can be partly attributed to the too low ZPE of water and partly to the neglect of non-ideal mixing in the calculation of the experimental number. The 2PT approach to ZPE calculations can be particularly powerful to estimate these nuclear quantum effects onto the thermodynamics of complex liquids or chemical reactions, such as acid/base reactions. Although within the context of DFT-MD simulation, the calculation of the ZPE of a single acid or conjugate base molecule in water solvent might still require prohibitively long trajectories.

**Acknowledgements**

This work is part of the Industrial Partnership Programme (IPP) ‘Computational sciences for energy research’ of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO). This research programme is co-financed by Shell Global Solutions International B.V. The calculations were carried out on the Dutch national e-infrastructure with the support of the SURF Cooperative.
Free energy barriers for the oxygen evolution reaction at the highly active RuO$_2$ surface

RuO$_2$ has been established as one of the most active catalysts for electro-chemical water splitting experimentally. However, recent theoretical studies are not able to reproduce this high reactivity. Most of these studies were aimed at computing the overpotential based on the adsorption energies of the reaction intermediates. However, the influence of the solvent phase on the surface chemistry has largely been ignored. Solvent reorganization can be particularly important in reactions that involve charge separation. In this study, we use first-principle molecular dynamics simulations combined with state-of-the-art enhanced sampling methods to capture the effects of the solvent reorganization on the RuO$_2$ (110) catalyzed water splitting. We are able to elucidate the mechanism of water splitting and estimate the kinetic barrier of the rate determining step to be 7.2 kcal/mol, which is approximately half of the previously reported value. This approach is straightforward to extend to other catalytic systems involving proton dissociation.
7. Probing the dynamics of water splitting reaction

7.1 Introduction

Electrochemical water splitting is one of the most promising ways to store energy into sustainable fuels, for example in molecular hydrogen. However, water splitting is yet to be widely adopted because of the required high overpotential, making the process inefficient. The water splitting reaction consist of two half reactions: the oxygen evolution reaction (OER) at the anode and the hydrogen evolution reaction (HER) at the cathode. Most of the energy loss is associated with the OER, which is proposed to take place via a four step mechanism as depicted by equations 7.1 to 7.4:

\[
* + H_2O \rightarrow *OH + H^+ + e^- \tag{7.1}
\]
\[
*OH \rightarrow *O + H^+ + e^- \tag{7.2}
\]
\[
*O + H_2O \rightarrow *OOH + H^+ + e^- \tag{7.3}
\]
\[
*OOH \rightarrow * + O_2 + H^+ + e^- \tag{7.4}
\]

RuO₂ has been proposed to be one of the best electrode materials for water splitting as early as 1970 by Trasatti et al. Nevertheless, to further lower the still significant energy losses, a more fundamental understanding of the underlying electrochemical processes is needed. Quantum chemical calculations and molecular simulations are ideally suited to provide such microscopic insight.

There have been numerous theoretical studies of the OER catalyzed by the RuO₂ (110) surface to rationalize its high activity. However, these density functional theory (DFT) based calculations have been unable to reproduce the experimentally observed high activity. Many attempts have been made focusing on various aspects of the simulations, such as improving the DFT model by including Van der Waals corrections, adding a Hubbard potential, by improving the physical description of the system by considering a few layers of solvent molecules explicitly on the RuO₂ slab to better describe the interface, probing other surface terminations of RuO₂ and recently by including the effect of various surface defects, such as point defects, kinks and steps. However, thus far none of these improvements have been able to reproduce the experimentally observed overpotential. These studies have provided critical insight into the OER by focusing on the thermodynamics of the process. However, also kinetic factors have been shown to be important for the water splitting and should be calculated along with the reaction free energies.

Calculating the pKₐ's of the adsorbed intermediates provides insight into the thermodynamics of the water splitting steps. However, probing the intrinsic pKₐ
of surface species is a daunting task both experimentally as well as theoretically. Many models have been developed over the years to theoretically predict the acidity constants of the surface functional groups. They can be broadly classified in empirical models and first principles models. The MultiSite Complexation (MUSIC) model\textsuperscript{209} and bond valence based methods from Bickmore et al\textsuperscript{210} are examples of empirical models. These methods vary in their formulation and use different types of experimental data for their parameterization. Due to the different underlying assumptions made, their predictions for the same metal oxide acidities can differ even as much as 10 $pK_a$ units in certain cases,\textsuperscript{211,212} which illustrates the need for further development of this class of methods.

An example of the first principles approach is to use DFT based molecular dynamics (DFT-MD) simulations combined with free energy perturbation methods as developed by Sprik et al\textsuperscript{117,213} and recently used by Blumberger et al\textsuperscript{156} for a hematite-water interface. This approach relies less on empirical data and gives good experimental agreement mainly because of the accurate description of the H-bonding network and the proper description of the metal oxide-water interactions. A drawback of Spriks proton insertion and deletion approach is that it does not provide information on the mechanism and the kinetics of the proton transfer process.

Here, we combine DFT-MD simulation with constrained MD\textsuperscript{214} (CMD) to enhance the sampling and probe the free energy landscape of the water splitting reaction steps taking place at the interface of the RuO$_2$(110) surface and explicit water solvent.

Hereafter, we provide the computational details and describe the setup of our model, including the equilibration and characterization of the RuO$_2$/water interface. Interestingly, we observe spontaneous hydration and hydroxylation reactions taking place during this equilibration. Next, we simulate the deprotonation reactions of the different water splitting intermediates at zero surface potential as well as the O-O bond formation. To probe the changes in the electronic structure, we include a Wannier center analysis.

### 7.2 Computational details

To allow for solvent mediated reaction mechanisms and solvent reorganization effects, it is essential to properly describe the RuO$_2$ water interface, including explicit water solvent. The solvent molecules are thus simulated at the same accurate quantum chemical level as the electro-catalyst and the reactants, allowing for bond breaking and forming not only at the interface, but also in the solvent.
7. Probing the dynamics of water splitting reaction

7.2.1 System setup

The metal oxide slab is modeled using three layers, consisting of 48 Ru atoms and 96 O atoms for the pristine RuO$_2$ (110) surface. There are two interfaces, each consisting of 8 coordinatively unsaturated sites (CUS) exposing 8 CUS ruthenium (Ru$_{CUS}$) and 8 bridging oxygen atoms (O$_{br}$) as shown in Figure 7.1. For the oxygen terminated system, all the CUS sites were capped with O atoms. The box size was obtained by taking the optimized bulk RuO$_2$ values in the two directions parallel to the slab, and letting the third lattice parameter, orthogonal to the surface, relax in an NPT run.

7.2.2 Ab-initio molecular dynamics

The DFT-MD simulations performed here employ the electronic structure calculation module QUICKSTEP$^{71}$ of the CP2K package,$^{70}$ which combines a Gaussian and plane-wave method to calculate the forces on the nuclei used for the Newtonian dynamics. We used a DZVP-MOLOPT-GTH basis set$^{74}$ and a 300 Ry cutoff for the plane waves. The SCF energy cutoff was $5 \cdot 10^{-7}$ to converge the wave function to good accuracy. We used the BLYP functional$^{21,73}$ in conjunction with Grimme’s D3 parameterized Van der Waals correction.$^{33}$ Core electrons are represented by the norm conserving pseudopotentials of Goedecker et al. (GTH).$^{72}$ To circumvent very fine grids, we used the NN50 smoothing of the exchange-correlation density. The time step was 0.5 fs and the CSVR thermostat$^{75}$ was used with a coupling constant of 250 fs to maintain a constant temperature of 325 K. An orthorhombic simulation box was used to which periodic boundary conditions were applied and the volume for each system was calculated using NPT simulations. The positions were saved every 5 fs for further analysis and visualization. All systems considered here were in an electronic closed shell state. Using above parameters gives excellent agreement with previous literature values of the binding energies of hydrogen within 0.05 eV, which is often used to predict the overpotential. To analyze electron transfers on RuO$_2$, maximally localized Wannier Functions (MLWF’s) were utilized to locate electrons and assign them to the atoms.$^{128}$ The RuO$_2$-water system was equilibrated for 10 ps with DFT-MD simulation before starting with the production runs and free energy calculations.

7.2.3 Constrained molecular dynamics

We use constrained molecular dynamics$^{214}$ to compute the free energy profiles of the activated water splitting reaction steps. However, rather than using holonomic constraints, we employ stiff harmonic springs that restrain the system at a certain value along a reaction coordinate or collective variable (CV).
7.2. Computational details

Fig. 7.1: Snapshot of the system in dynamic equilibrium, showing the periodic supercell, the three layer RuO$_2$ slab with hydrating water molecules, in ball and stick representation. Solvent water is shown as semi-transparent sticks. Note that all surface CUS Ru atoms (black frames) are hydrated or hydroxylated, whereas several surface bridging oxygens, in the in-between rows, are protonated. The same is seen at the bottom interface. See for comparison the initial pristine RuO$_2$ surface in Figure 7.8 on page 117.

$$f_c = -k(X_0 - X)$$ (7.5)

By accumulating the force of constraint, $f_c$, in a series of constrained MD simulations (Eq. 7.5), with the CV constrained at different values ranging from the reactant state to the product state, the Landau free energy profile of the reaction is obtained by integration.
7. Probing the dynamics of water splitting reaction

\[ \Delta G = \int \langle f_c \rangle dX \]  \tag{7.6}

For each restraint simulation, the system was first equilibrated for at least 5000 steps, after which the average restraint force obtained from a restraint NVT run of 4000 steps.

**Fig. 7.2:** Representation of the first PCET showing the $H_2O$-terminated CUS and a solvent water molecule with collective variable ($X = d_1 - d_2$).

7.3 Results and Discussion

7.3.1 Surface characterization

During the 10 ps NVT equilibration simulation, starting from the pristine RuO$_2$(110) surface (shown in Figure 7.8) brought in contact with water solvent, we observe that solvent water molecules adsorb onto the surface and bind to the 5-fold coordinated Ru atoms. These chemisorbed water molecules dissociate spontaneously and donate a proton to a nearby O$_{br}$. This is in agreement with previous theoretical and experimental studies,\textsuperscript{201,215–217} which also reported dissociation of water that adsorbs on the RuO$_2$ surface. While a DFT study by Fang et al.\textsuperscript{205} suggested a fully OH-terminated surface at low potential, we find in our simulations that on
average 50 to 60% of the adsorbed water molecules are dissociated without an external electrode potential in agreement with Rentao et al.\textsuperscript{216,218} The DFT-MD simulations show a very dynamic surface, with a multitude of processes happening on the picosecond time scale. The proton transfer processes occurring can be assigned into the following two categories:

1. Surface proton transfer: in this process a proton is transferred between a water molecule adsorbed on the CUS Ru site and an adjacent surface O\textsubscript{br}, forming two surface hydroxyl groups. This proton transfer can also occur via intermediate solvent molecules as shown in Figure 7.3a. Upon accepting the proton, the lateral distance between the O\textsubscript{br} and the neighboring six fold saturated Ru atom changes on average from 1.85 Å to 2.05 Å to compensate for the newly formed covalent O\textsubscript{br}H bond.

2. Adlayer proton transfer: in this process a proton transfer occurs between a water molecule adsorbed on the CUS Ru site and a hydroxyl group formed by the dissociation of a water molecule adsorbed on the neighboring CUS Ru site as shown in Figure 7.3b.

During the entire DFT-MD run, a proton is never transferred spontaneously...
7. Probing the dynamics of water splitting reaction

beyond the interface to the bulk water. This suggests that proton transport into
the solvent is hampered by a free energy barrier, which is in agreement with previ-
ous theoretical and experimental studies that proposed that the first deprotonation
takes place at an intermediate external electrode potential of 1.0 to 1.5V. In the next section, we will compute the free energy profile of proton transfer to
the solvent with constrained MD simulations.

7.3.2 First proton transfer

As described above, a variety of hydration and hydroxylation reactions were ob-
served during equilibration of the RuO₂ water solvent interface, resulting in a dy-
namic aquo and hydroxo terminated surface. To estimate the \( pK_a \) of this surface,
we compute the free energy profile of transferring a proton from the surface to a
solvent water molecule with a series of constrained MD simulations as explained
in section 7.2.3.

As the constrained collective variable, we choose the difference of the O-H
distance from the surface donor oxygen to the proton and the H-O distance from
the proton to the accepting water solvent oxygen (see also Figure 7.2). The
selected surface oxygen is from an adsorbed hydroxyl group on a Ru\( _{CUS} \) site.

A series of 13 constrained MD simulations was performed, with the CV ranging
from -1.2 to 1.2 Å. The stiff restraint force constant, \( k \), was 200 kcal·mol\(^{-1}\)Å\(^{-2}\). Additional restraints were applied to the OH distances of the receiving water
molecule, with a \( k \) of 500 kcal·mol\(^{-1}\)Å\(^{-2}\). Each run started with an additional
1 ps equilibration, followed by 2 ps of collecting the average restraint force.

The resulting free energy profile, shown in Figure 7.4, is steeply increasing to
15 kcal/mol without showing a product minimum or a separating transition state
barrier. Using equation 7.7, this results in a \( pK_a \) of 12.

\[
pK_a = \frac{\Delta G}{k_BT\ln[10]} \tag{7.7}
\]

As the rather simple CV can only describe the first transfer and not the fol-
lowing Grotthuss-like proton jumps, we estimate that the free energy profile is
reliable up to \( r_2 - r_1 = 0.5 \) Å. A more reliable estimate of the surface \( pK_a \) can be
obtained, in principle, with a path-metadynamics (PMD) simulation in combina-
tion with the proton transfer collective variable (PTCV), as shown in Chapter 5.
Unfortunately, the required optimization of the path-CV together with sampling
of multiple Grotthuss jumps for this system is currently still too computationally
demanding (see section 7.4 in the Appendix for a PMD attempt of the second
proton transfer). Another technical issue is that our current implementation of
the PTCV cannot yet deal with multiple excess protons in the system, which is
7.3. Results and Discussion

the case for this surface.

Fig. 7.4: Free energy curve for the first proton transfer

7.3.3 Second proton transfer

Next, we estimate the free energy barrier for the proton transfer in the second water splitting reaction step. Starting from the previous equilibrated system, for both surface sides eight protons and electrons are removed, leaving all RuCUS OH-terminated and all Odehyde deprotonated. Using a $k$ of 200 kcal·mol$^{-1}$·Å$^{-2}$, the collective variable $X$ was constrained on 10 values each 0.25 Å apart. The obtained free-energy curve (Fig. 7.5) shows an activation barrier of 4.5 kcal/mol, which is lower than the first proton transfer and indicates that the second proton transfer is not the rate determining step. Furthermore, it suggests a self-catalyzing process where subsequent proton transfer processes have lower free energy barriers, which supports the earlier hypothesis that at high external potential the RuO$_2$ water interface is oxygen terminated.

7.3.4 OER kinetics at O-terminated surface

To calculate the free energy barrier of the O-O bond formation (equation 7.3), we start with an oxygen terminated system as it is the most stable phase at the OER conditions. We again utilize constrained molecular dynamics to calculate the free energy.

Here, we use for the CV, the distance between the terminal O and the incoming water oxygen atoms, denoted by $r_{O-O}$. Finally, a free-energy profile is constructed
7. Probing the dynamics of water splitting reaction

Fig. 7.5: Free energy curve of the second proton transfer.

from the constrained molecular dynamics simulations as shown in Figure 7.6.

Fig. 7.6: Free energy profile for the rate determining step involving O → OOH

We find that the kinetic free energy barrier for the rate determining O-O bond formation step is 7.2 kcal/mol. This barrier has been previously estimated to be 17.1 kcal/mol by Liu et al.\textsuperscript{205} The drastic reduction in the barrier height might be associated with the inclusion of a full explicit solvation considered in the current study. As shown in Figure 7.7, there is a second water molecule actively participating in the transition state, which is not considered in the previous gas phase studies. Unlike the commonly proposed mechanism where a water molecule attacks the oxygen terminated surface site, we find that the solvent
7.3. Results and Discussion

Fig. 7.7: Snapshots from four different constrained molecular dynamics runs, at the indicated values of the constrained O-O distance CV. Plot (a) shows a water molecule approaching an oxygen atom adsorbed at Ru\textsubscript{CUS}. Plot (b) shows the transition state, where the first proton is transferred to the solvent. Plot (c) shows the spontaneous second proton transfer to the neighboring O\textsubscript{hr} and plot (d) shows the final O-O bond formation.

water molecules assist the proton dissociation of the attacking water molecule. The reaction thus proceeds through a less kinetically hindered hetero-oxo coupling reaction instead of a high barrier water dissociation reaction.

The hetero-oxo coupling is possible due to the generation of an electrophilic terminal oxygen at the end of the third proton transfer reaction. The electrophilic CUS oxygen can couple with a more nucleophilic OH\textsuperscript{−} group reducing the kinetic barrier of the overall reaction. The reported onset overpotential for the OER on RuO\textsubscript{2} is 6 – 10 kcal/mol, which matches reasonably well with our kinetic barrier calculation.
7.4 Conclusion

We have performed DFT-MD simulations combined with the constraint MD technique to study the oxygen evolution reaction steps catalyzed by the RuO$_2$ (110) surface in contact with explicit water solvent. First, we perform unbiased DFT-MD simulations to characterize the surface structure under zero external potential. This gives us information about the surface coverage by the various oxo-species and on the different proton transfer processes taking place. In particular, our simulations show that at zero external bias, water adsorbs and dissociates spontaneously. In particular, we find that solvent water adsorbs at the Ru$_{CUS}$ site, and that 50-60\% of the adsorbed water dissociates by donating a proton to a nearby surface oxygen. The system reaches a dynamic equilibrium after 5 ps, maintaining the surface coverage unchanged. These simulations suggest that the first proton transfer occurs to the neighboring $O_{br}$ spontaneously. The dissociative adsorption of water molecules on Ru$_{CUS}$ is in agreement with previous studies.\textsuperscript{217,220} Second, we performed DFT-MD simulations in conjunction with constrained DFT-MD simulations to calculate the free energy barrier of the second proton transfer reaction. We found a barrier height of 15 and 4 kcal/mol for the first and the second surface proton respectively. We show with the help of PMD simulations as well as Wannier center analysis that this oxygen terminated system does not accept hydrogen bonds indicating the electrophilic nature. This observation explains the approach of a new water molecule with oxygen pointing towards the surface oxo-species paving the way for the O-O bond formation. Finally, we calculate the kinetic barrier of the rate determining step of the oxygen evolution reaction. We found a value of 6 kcal/mol as opposed to 17 kcal/mol\textsuperscript{205} reported earlier. The improvement can be attributed to inclusion of full solvent effects.

This works extends the understanding of the water splitting reaction on the highly active RuO$_2$ (110) surface. These atomistic simulations not only estimate the free energy barriers but also provide atomistic mechanistic insights on the water dissociation as well as O-O bond formation step. We identify several adsorbed species, different active sites and elucidate their role in the overall catalysis of the OER. This work emphasizes on the importance of explicit solvent in the calculation of the kinetic barriers and proposes a generic framework that can be employed to study other metals or metal oxides for a myriad of reactions. Furthermore, this framework can be easily extended to study the effects of various surface defects.
APPENDIX

Benchmark results

Ruthenium oxide (110) consists of two octahedral ruthenium oxide complexes alternating horizontally and vertically as shown in Fig. 7.8. At the RuO₂ water interface there are two active sites, of which the Coordinately Unsaturated Site (CUS) is the most important. It is a five-fold coordinated ruthenium atom that is capable of binding a sixth ligand. Adjacent to the CUS exists the bridging oxygen (\(O_{br}\)) which 'bridges' two horizontal octahedral RuO₂ structures and as discussed later and shown experimentally before, plays a role in stabilizing CUS-bound intermediates.²¹⁷

![Fig. 7.8: Representation of rutile RuO₂ (110), visible is the three-layered structure with two surface layers.](image)

In order to find suitable system settings, the absolute energy difference between two intermediate structures [Fig. 7.9], where the CUS are \(O^*\) and \(HO^*\) terminated, was calculated and compared to benchmark results acquired from earlier studies, where it was observed that the energy difference between the two intermediate states \((E_{O^*} - E_{HO^*})\) is 1.69 eV² and 1.72 eV.³

To calculate the energy of these intermediate structures a clean RuO₂ surface was taken where \(O^*\), \(HO^*\) or \(HOO^*\) (the asterisk signifying the atom bound to RuO₂) were adsorbed to one of sixteen available CUS, after which gas-phase geometry optimizations were performed to find their absolute energies. These simulations were performed without considering solvation effects.

To test the system settings gas phase geometry optimization simulations were
7. Probing the dynamics of water splitting reaction

**Fig. 7.9:** Simplified visualization of the active surface sites of the RuO$_2$ (110) in different stages of the OER-mechanism. The pink and red spheres represent ruthenium and oxygen respectively. From left to right; Clean CUS; H$_2$O-terminated CUS; HO-terminated CUS; O-terminated CUS.

**Table 7.1:** The energy difference ($E_{O^*} - E_{HO^*}$) obtained from benchmark calculations performed using different functionals and the comparison with the literature values.

<table>
<thead>
<tr>
<th>Functional</th>
<th>BLYP</th>
<th>PBE</th>
<th>revPBE</th>
<th>Rossmeisl et al$^2$</th>
<th>Man et. al.$^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{O^<em>} - E_{HO^</em>}$ (eV)</td>
<td>1.6834</td>
<td>1.7654</td>
<td>1.6333</td>
<td>1.69</td>
<td>1.72</td>
</tr>
</tbody>
</table>

performed using different XC functionals to calculate the energy difference between systems containing $O^*$ and $HO^*$ terminated CUS. The results were compared to benchmark results observed in previous studies$^{2,3}$ and are shown in Table 7.1.

Both the Perdew-Burcke-Ernerhof (PBE) as well as the BLYP functional obtained comparable $E_{O^*} - E_{HO^*}$ values to the benchmark values. This is not surprising and it has been shown that XC functional does not drastically change the binding energy. While PBE is often used for solid state simulations, BLYP has proven to describe water solvation effects more accurately. For this reason the BLYP DFT-D3 functional was chosen to perform further simulations.

**Surface characterization**

During the simulations where the CUS was H$_2$O-terminated it was observed that the RuO$_2$ surface was very dynamic and intramolecular hydrogen transfer occurred frequently between the CUS and the O$_{br}$. This suggests that spontaneous dissociation of the H$_2$O occurs on the CUS leading to a surface containing various adsorbents. Furthermore, the role of the bridging oxygen plays an important role in the first proton dissociation step.

To obtain a good estimate of the surface composition at zero external potential, we performed two DFT-MD simulations. In the first simulation, we start with a system where all the CUS sites are H$_2$O terminated. Within the first picosecond, almost half of the water molecules undergo dissociation and donate their proton to
the neighboring $O_{br}$ as shown in Fig. 7.12a. After simulating for 3 ps, the system achieves a dynamic equilibrium.

In the second simulation, we started with the fully dissociated system as shown in Fig. 7.12b. This was done to check if there are big kinetic barriers for the first proton transfer which might have caused the system to be stuck in a local minimum. Similar to the first simulation, significant proton rearrangement was observed. Again, the system achieved dynamic equilibrium within 5 ps resulting in half dissociated surface indicating that 50-60 % of water molecules dissociates on adsorption on the CUS sites.

As discussed in the main text, the RuO$_2$ water interface is very dynamic in nature and there are a lot of intra-surface proton transfer processes happening. However, during the equilibration simulation, second proton does not happen spontaneously. This compels us to perform enhanced sampling calculation to observe the second proton transfer to the bulk water.

**Path-metadynamics**

To track the proton beyond the first layer, a novel framework is used where DFT-MD simulations were augmented with the path-metadynamics (PMD) method$^{41}$ as implemented in PLUMED$^{159}$ to enhance the sampling of the proton dissociation reaction, obtain the reaction free energy profiles, and unravel the reaction mechanisms. PMD enforces the dynamics away from the stable reactant and product states by gradually constructing a repulsive bias along a path variable that is a function of geometric order parameters, or collective variables (CVs), and connects the reactant and product states. For the initial guess of the path, we chose a linear interpolation between the average CV values in the stable states obtained from equilibrium DFT-MD simulations of the reactant state and the product state. The PathCV is a function of the two CVs explained hereafter.

The first CV quantifies the distance, $r_{xp}$, between the proton and the adsorbed species. Here, $x$ is the position of the adsorbed donor atom at the interface of water and metal oxide. The proton position, $q_p$, or rather, the center of the hydronium ion when the proton is in the water solvent, is computed by taking a weighted sum over all positions, $q$, of the oxygen atoms in the system that can bind the proton:

$$q_p = \sum_i w(i)q(i) \quad (7.8)$$

with the weights,
7. Probing the dynamics of water splitting reaction

\[ w(i) = \frac{\exp\left[\gamma(n(i) - n_0(i))\right]}{\sum_i \exp\left[\gamma(n(i) - n_0(i))\right]} \]  

(7.9)

The position \( q_p \) is thus obtained as an exponentially weighted average of all proton acceptor positions, which are here the water and surface oxygen atoms. With a positive number for the parameter \( \gamma \), the weighting selects the oxygen(s) with the highest coordination number with respect to the H-atoms, \( n_{OH} \). Here, \( n_0(i) \) is the default number of hydrogens bonded to atom \( i \) (without the proton), which is 2 for a water oxygen and 0 for the donating surface oxygen atom. The coordination number is computed with:

\[ n_{xy} = \frac{1 - (r_j/r_o)^{nn}}{1 - (r_j/r_o)^{mm}} \]  

(7.10)

This proton transfer collective variable (PTCV) is not very sensitive to the exact values of the switching function and the same parameters of the switching function work across the whole pK\(_a\) spectrum. The following parameter values were used: \( nn = 12, mm = 24, r_o = 1.3 \ \text{Å} \), and \( \gamma = 8 \).

The second CV, \( n_{OO} \), is the coordination number (equation 7.10) of the surface oxygen(s) with respect to all water oxygen atoms, with \( nn = 14, mm = 28, \) and \( r_o = 2.5 \ \text{Å} \).

The path is comprised of 30 nodes with fixed ends (i.e. without trailing nodes) and is updated every 0.5 ps, starting from an initial straight guess path in the 2D CV space. The half-life parameter was 0.5 ps. The metadynamics Gaussian potentials were added with a period of 50 fs, a height of 0.05 kcal/mol and a width of 0.045 in normalized path units.

Plumed input examples for the PMD simulation with the proton distance CV are available on the Plumed nest website (www.plumed-nest.org), including the source code for the PTCV under Plumed Nest Project ID: 19.034. The source code for our PMD implementation is found in Plumed Nest Project ID: 19.033.

Our results indicate that for an oxygen terminated surface, the adsorbed OH at the CUS site is acidic in nature. The pK\(_a\) is estimated using equation 7.7 and found to be 1.2. As shown in Figure 7.10, the meta-stable contact ion pair (CIP) formation is negligible due to the lack of electrostatic interaction between the newly formed hydronium ion and the residual conjugate base. Furthermore, the O-O coordination number drops to 0 once the proton is transferred to the solvent (Figure 7.11). This drop in the O-O coordination number is attributed to the electrophilic nature of the residual surface CUS oxygen as it does not make stable H-bonds with the solvent water molecules. This is further confirmed by calculating
the position of the Wannier centers that changes upon the second proton transfer from a tetrahedral to double bond arrangement as shown in Figure 7.12.

![Free energy curve for the second proton dissociation using PMD](image)

**Fig. 7.10:** Free energy curve for the second proton dissociation using PMD

**Comparison with acid dissociation**

To understand the second proton dissociation better, we compared it with the acid dissociation in phenol. After proton dissociation, the remaining phenol molecule (phenoxide) is negatively charged and thus makes very strong H-bonds. This is in contrast to what we observed in case of surface oxygen which does not make strong hydrogen bonds after proton dissociation as depicted in Fig. 7.11 by the drastic drop in O-O coordination number.

The lack of hydrogen bonding of the surface oxygen atom with the solvent water molecule in turn can be explained by the immediate transfer of the electron into the bulk surface making the surface oxygen atom rather electrophilic in nature.

This is further confirmed by plotting the Wannier center positions as shown in Fig. 7.12, where one Wannier center changes its position and is now shared between Ru and the surface oxygen atom. In fact, this process also forms the premises for the O-O bond formation (rate determining step) of the OER reaction.
7. Probing the dynamics of water splitting reaction

Fig. 7.11: Comparison of the second proton dissociation in RuO$_2$ with proton dissociation in phenol.

(a) Wannier centers before the second proton transfer

(b) Wannier centers after the second proton transfer

Fig. 7.12: Wannier center rearrangement around the interface oxygen after the second proton transfer reaction. Oxygen, Ruthenium and wannier centers are represented by red, silver and blue balls respectively.
7.4. Conclusion

Rate determining step

Apart from calculating the free energy barrier using the constrained molecular dynamics simulations, we also looked into the Wannier center rearrangement before, during and after the O-O bond formation as shown in Fig. 7.12b and 7.13.

![Wannier centers during the transition state of O-O bond formation](image1)

![Wannier centers after the O-O bond formation](image2)

**Fig. 7.13:** Wannier center rearrangement around the interface oxygen after the second proton transfer reaction. Oxygen, Ruthenium and wannier centers are represented by red, silver and blue balls respectively.

In the oxygen terminated surface, there are 3 wannier centers pointing towards the RuO$_2$ slab whereas, the $Ru_{CUS}$ has 5 wannier center. When another water molecule approaches the surface oxygen, the wannier center rearranges due to a new bond formation as shown in Fig. 7.13a. When there is a complete O-O bond formation, there are 2 wannier centers now between the oxygens involved in O-O bond. Moreover, another wannier center is transferred to the $Ru_{CUS}$, which now has total of 6 wannier centers.

Acknowledgements

This work is part of the Industrial Partnership Programme (IPP) 'Computational sciences for energy research' of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organisation for Scientific Research (NWO). This research programme is co-financed by Shell Global Solutions International B.V. The calculations were carried out on the Dutch national e-infrastructure with the support of the SURF Cooperative.
Summary

With the imminent threat of global warming and climate change that can render millions of people homeless and the earth inhospitable, it has never been more crucial to look for cleaner sources of energy. However, most renewable sources of energy are intermittent in nature and thus storing the excess energy efficiently for use at a later time is important. One of the most promising ways to do so is by storing excess energy in the form of chemical bonds, such as by splitting water to produce Hydrogen. Despite several breakthroughs, there is still a lot of energy loss while conducting water splitting due to the overpotential. We will show how molecular modeling tools developed in this thesis can be used to unravel the molecular origins of this overpotential.

In this thesis, we start by providing a brief overview of the proton and electron transfer reactions involved in the water splitting reaction followed by a historic commentary on the progress of the computational chemistry field and a summary of our model. In the second chapter, we discuss concisely theoretical background behind all the methods and techniques used throughout the thesis, including DFT, ab-initio molecular dynamics, a brief review on statistical mechanics, transition path sampling, metadynamics, and path-metadynamics.

Most of the chemical reactions discussed in this thesis occur in aqueous solvent. To get reliable information about the reaction mechanism and energy barriers, explicit solvent description is required. DFT-MD simulations have been and still is an immensely powerful instrument to obtain microscopic insight in water and aqueous solutions. The quality of the DFT-MD water model is however limited by the approximate nature of standard GGA functionals, neglect of nuclear quantum effects, and finite size effects. In chapter 3, we have performed a comprehensive study of the effects of various parameters such as the choice of XC functional, Van der Waals corrections and smoothing operators on various structural and dynamical properties of bulk water. We show that for condensed phase simulations, the PBE functional even in conjunction with Grimme’s D3 correction overstructures water and results in glassy dynamics. We also propose a framework to estimate the box size for ambient pressure simulation using NVT ensemble. We utilize the new framework to study the structural and dynamical properties of bulk water at ambient conditions.

In chapter 4, we developed a new framework to simulate electron transfer reactions by combining DFT-MD simulations with transition path sampling. We generate an ensemble of reactive trajectories of the self-exchange reaction between a pair of Ru\(^{2+}\) and Ru\(^{3+}\) ions in water. To distinguish between the reactant and
product states, we use the position of the maximally localised Wannier center associated with the transferring electron as an order parameter. We show that the simple self-exchange reaction between two metal ions is coupled to a proton transfer.

The second important process in a water splitting reaction is a proton transfer. To successfully simulate a proton transfer, a robust framework is needed to track and bias the excess proton. In chapter 5, we developed a new proton transfer collective variable (PTCV) to study the proton dissociation process. We show that the PTCV can be used to study the acid dissociation process, accurately predicting the $pK_a$ across a broad acid strength spectrum. Moreover, we shed light on the different acid dissociation mechanisms in acids. We propose two different types of mechanism: a slow step-wise mechanism, and a concerted long range mechanism. We elucidate the molecular origin of these two different mechanisms by combining the $pK_a$ results with DFT-MD equilibration runs. We show that the different degrees of solvent reorganization play a decisive role in selecting one mechanism or the other.

The free energy calculations performed using enhanced sampling methods do not include nuclear quantum effects such as zero point energy (ZPE). In the case of water, ZPE changes the $pK_a$ of water from 7 to the pKa of 7.43 for D$_2$O. Hence, in chapter 6, we use the 2PT method to calculate the zero point energy of various liquids and their mixtures. The 2PT method uses the density of states (DoS), which is computed from the velocity autocorrelation function (VAC) obtained from a short classical molecular dynamics trajectory. We also perform a benchmark calculation to assess the speed and convergence of the 2PT method. We show that systems with 1000 molecules converge within a few ps, but single molecule systems of interest such as one acid molecules in the water solvent need at least 1000 ps to get a converged result. Unfortunately, in the current state of the art, this makes the method much more expensive than originally advertised, but still more economic than full blown path integral MD calculations.

In the final chapter, we apply the newly developed frameworks to study the water splitting reaction on the RuO$_2$ (110) surface. RuO$_2$ has been established as one of the most active catalysts for electro-chemical water splitting experimentally. However, recent theoretical studies are not able to reproduce this high reactivity. Most of these studies were aimed at computing the overpotential based on the adsorption energies of the reaction intermediates. However, the influence of the solvent phase on the surface chemistry has largely been ignored. We use first-principle molecular dynamics simulations combined with state-of-the-art enhanced sampling methods to capture the effects of the solvent reorganization on the RuO$_2$ (110) catalyzed water splitting. We study various proton transfer processes.
occurring on the metal oxide-water interface. We show that the system reaches dynamical equilibrium and that the water molecules dissociate spontaneously upon adsorption on the unsaturated (CUS) Ru sites. We proceed by calculating the proton dissociation barrier for each step of the oxygen evolution reaction. We show that by considering the explicit solvent, the barrier for the rate determining step (O-O bond formation) of the OER decreases from 0.74 to 0.4 eV. Our framework, in principle, is straightforward to extend to other catalytic systems involving proton dissociation and can be used to gain mechanistic insights in a lot of other applications such as homogenous catalysis, hydrogen evolution reaction and meta-organic frameworks.
Samenvatting
Samenvatting

Met de steeds groter wordende dreiging van de opwarming van de aarde en de klimaatverandering die miljoenen mensen dakloos en de aarde onleefbaar kan maken, is het nog nooit zo belangrijk geweest om op zoek te gaan naar schoner energiebronnen. De meeste hernieuwbare energiebronnen kunnen echter vaak niet 24 uur per dag energie leveren (zonne-energie) en daarom is het belangrijk om overtollige energie efficient op te slaan voor gebruik op een later tijdstip. Een van de meest veelbelovende manieren om dit te doen is door overtollige energie op te slaan in de vorm van chemische bindingen, bijvoorbeeld door water te splitsen om waterstof te produceren. Sinds 1970 wordt er al onderzoek gedaan naar een efficiente manier van watersplitsing. Ondanks een aantal doorbraken is er nog steeds veel energieverlies bij het splitsen van het water door het overpotentiaal. We zullen laten zien hoe moleculaire modellerings technieken die in dit proefschrift zijn ontwikkeld kunnen worden gebruikt om de moleculaire oorsprong van dit overpotentiaal te ontrafelen.

We hebben een nieuwe methode ontwikkeld dat gebruikt kan worden om proton en electron transfer en de proton coupled electron transfer reacties die optreden tijdens de splitsing van water te bestuderen. We hebben deze methode toegepast op het RuO$_2$ (110) oppervlak om het mechanisme van de watersplitsing te ontrafelen. Dit is gelukt door gebruik te maken van DFT-MD simulaties samen met verschillende zogenaamde enhanced sampling technieken.

In dit proefschrift beginnen we met een kort overzicht van de proton en electron transfer reacties die betrokken zijn bij de watersplitsing, gevolgd door een beschrijving van de totstandkoming van de huidige computationele chemie en een samenvatting van ons model. In het tweede hoofdstuk geven we een beknopte toelichting van de theoretische achtergrond achter alle methoden en technieken die gebruikt zijn in de loop van het proefschrift, waaronder DFT, ab-initio moleculaire dynamiek, een kort overzicht van de statistische mechanica, transition path sampling, metadynamics en path-metadynamics.

De meeste van de chemische reacties besproken in dit proefschrift komen voor in waterige oplossing. Om betrouwbare informatie te krijgen over het reactiemechanisme en de daarbij behorende energie barrières is een expliciete beschrijving van het oplosmiddel nodig. DFT-MD simulaties waren en zijn nog steeds een enorm krachtig instrument om microscopisch inzicht te krijgen in water en waterige oplossingen. De kwaliteit van het DFT-MD watermodel wordt echter beperkt door de aard van de standaard GGA-functionalen, het weglaten van nucleaire kwantumeffecten en de eindige grootte van het systeem. In hoofdstuk 3
hebben we een uitgebreide studie uitgevoerd naar de effecten van verschillende parameters zoals de keuze van XC-functionaliteit, Van der Waals-correcties en zogenaamde smoothing operators op verschillende structurele en dynamische eigenschappen van bulkwater. We laten zien dat bij gecondenseerde fase simulaties de PBE-functie, zelfs in combinatie met Grimme’s D3-correctie, het water over structureert en resulteert in trage dynamica. We stellen ook een methode voor om de doos grootte te schatten voor simulatie van omgevingsdruk met behulp van een NVT ensemble. We gebruiken de nieuwe methode om de structurele en dynamische eigenschappen van bulkwater bij omgevingscondities te bestuderen.

In hoofdstuk 4 hebben we een nieuwe methode ontwikkeld om electron transfer reacties te simuleren door DFT-MD-simulaties te combineren met transition path sampling. We genereren een ensemble van de self-exchange reactie tussen een Ru$^{2+}$ ion en een Ru$^{3+}$ ion in water. Om onderscheid te maken tussen de begin- en eindtoestand, gebruiken we de positie van het maximaal gelokaliseerde Wannier-centrum die hoort bij het uitgewisselde elektron als orderparameter. Dit stelt ons in staat om nieuwe paden te genereren vanuit het moment van de barriere overgang en statistische gemiddelden over deze paden te berekenen. We tonen aan dat de eenvoudige self-exchange reactie tussen twee metaalionen gekoppeld is aan een proton transfer. We bevestigen en kwantificeren verder dat deze electron transfer reactie wordt veroorzaakt door een verandering in het oplosmiddel zoals Marcus aanvankelijk voorspeld had.

Het tweede belangrijke proces in de splitsing van water is een proton transfer. Om met succes een proton transfer te simuleren, is een betrouwbare methode nodig om het overtollige proton op te sporen en te sturen. In hoofdstuk 5 hebben we een nieuwe proton transfer collective variable (PTCV) ontwikkeld om het proton dissociatie proces te bestuderen. We laten zien dat de PTCV gebruikt kan worden om het zuur dissociatie proces te bestuderen, waarbij de $pK_a$ nauwkeurig wordt voorspeld over een breed zuurgraad spectrum. Bovendien geven we meer inzicht in de verschillende zuur dissociatie mechanismen voor verschillende moleculen. We stellen twee soorten mechanismen voor: een langzaam stapsgewijs mechanisme, en een gecordeerd lange-afstands mechanisme. We verhelderen de oorsprong op molecular van deze twee verschillende mechanismen door de $pK_a$ resultaten te combineren met DFT-MD equilibration runs. We tonen aan dat de verschillende gradaties van herstructurering in het oplosmiddel een beslissende rol spelen bij de selectie van het ene of het andere mechanisme.

De berekeningen van de vrije energie, uitgevoerd met behulp van enhanced sampling methoden, houden geen rekening met nucleaire kwantumeffecten zoals nulpunt-energie (ZPE). In het geval van water verandert ZPE de $pK_a$ van water van 7 naar de $pK_a$ van 7,43 voor D$_2$O. Daarom gebruiken we in hoofdstuk 6 de
2PT-methode om de nulpunt-energie te berekenen van verschillende vloeistoffen en hun mengsels. De 2PT-methode maakt gebruik van de density of states (DoS), die wordt berekend op basis van de velocity autocorrelation function (VAC) die wordt verkregen uit een kort classical MD simulatie. Door de VAC en de DoS van een vloeistof op te delen in vaste en gasvormige componenten, kunnen kwantum-mechanische correcties van thermodynamische eigenschappen worden berekend. We voeren ook een ijkpunt berekening uit om de snelheid en convergentie van de 2PT-methode te beoordelen. We laten zien dat systemen met 1000 moleculen binnen een paar ps convergeren, maar systemen met bijvoorbeeld een zuur molecule in water hebben ten minste 1000 ps nodig om een geconvergeerd resultaat te krijgen. Helaas, met de huidige techniek is de methode dus veel computatieduurder dan oorspronkelijk gedacht, maar nog wel goedkoper dan de volledige path integral MD berekeningen.

In het laatste hoofdstuk passen we de nieuw ontwikkelde methodes toe om de splitsing van water op het oppervlak van RuO$_2$ (110) te bestuderen. RuO$_2$ is een van de meest actieve katalysatoren voor electrochemische watersplitsing, die experimenteel is vastgesteld. Recente theoretische studies zijn echter niet in staat om deze hoge reaktiviteit te reproduceren. De meeste van deze studies waren gericht op het berekenen van het overpotentiaal op basis van de adsorptie energien van de tussenproducten. De invloed van de oplossingsmiddel op de oppervlakte reacties is echter grotendeels genegeerd. We gebruiken ab initio MD simulaties in combinatie met de nieuwste enhanced sampling methoden om de effecten van de herstructurering van het oplosmiddel op de oppervlakte reacties te verduidelijken. We beginnen met het ijken van de DFT-parameters om goede overeenstemming te krijgen met de literatuur over adsorptie-energie. In de volgende stap bestuderen we verschillende proton transfer processen die zich voordoen op het metaaloxide-water raakvlak. We tonen aan dat het systeem een dynamisch evenwicht bereikt en dat de watermoleculen spontaan uiteenvallen bij adsorptie op de onverzadigde (CUS) Ru-sites. We gaan verder door de proton dissociatie barriere voor elke stap van de zuurstof evolutie reactie te berekenen. We tonen aan dat door rekening te houden met het oplosmiddel, de barriere voor de snelheidsbepalende stap (vorming van O-O binding) van de OER afneemt van 0,74 tot 0,4 eV. Onze methode is in principe eenvoudig uit te breiden naar andere katalytische systemen met proton dissociatie en kan gebruikt worden om mechanistische inzichten te verkrijgen in vele andere toepassingen zoals homogene katalyse, waterstof evolutie reacties en meta-organische raamwerken.


136


137


190 Johnson, R. D. *NIST computational chemistry comparison and benchmark database*; 2001; p 261.


<table>
<thead>
<tr>
<th>Reference</th>
<th>Authors</th>
<th>Journal</th>
<th>Year</th>
<th>Volume</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>204</td>
<td>Exner, K. S.; Sohrabnejad-Eskan, I.; Over, H.</td>
<td><em>ACS Catal.</em></td>
<td>2018</td>
<td>8</td>
<td>1864–1879</td>
</tr>
</tbody>
</table>
List of publications

Part of this thesis


5. Tiwari, A., & Ensing, B. In preparation, Free energy barriers for the oxygen evolution reaction at the highly active RuO$_2$ surface.

Not part of this thesis


8. Tiwari, A., Papadopoulos, S., & Ensing, B. In preparation, Mechanistic insights from combined theoretical and experimental studies into shuttling mechanism in Rotaxanes.
Acknowledgement

Four years have gone in a breath! So much learnt, so much explored, so many new experiences, and none of this would have been possible without the help of some very special people. First of all, I want to take this opportunity to thank my parents for all the sacrifices they made for my education and to show me the directions unknown to most in our part of the world. Their constant encouragement and unconditional love kept me motivated in hard times. Next, I want to thank my brother for always being there for me in every situation. Talking for hours about a movie or a new algorithm never gets boring with you!

After my family, if there is another person I want to sincerely thank, it is my supervisor, Bernd. He guided me in all the scientific endeavors, sometimes for hours in a row (Sorry, Francesca). It always felt so easy to approach him for even the simplest of questions. The freedom he extended to choose new projects is simply praiseworthy. Moreover, his help was not limited to just science but with anything under the sun, such as swimming, diving, hiking, repairing a bike and many more. I am particularly very grateful to him for his continued support even after the unfortunate accident to help me finish on time. I also want to thank my second supervisor, Peter for always being ready to help and for keep giving me salary raise letters. Furthermore, the discussions with him about my work were always very insightful.

Coming to a foreign country leaving everyone you know behind can be tough but this was made a peace of cake by a host of people, first of whom was my roommate Julian. Thanks for introducing me to all the people and the awesome parties. In my second year, many awesome people arrived in the group. First of all, I want to start by thanking my paranymphs, Verena and Yuliia for the last three years of craziness and all the fun. I learnt a lot from you two about life. Special thanks to Verena for teaching me driving, swimming and skiing and Yuliia for helping me with the thesis cover and numerous art and music suggestions in general. Another awesome group of friends I want to thank are from my college, Poonam, Rajat, Utkarsh, Arjun and Ayush. You guys have been a constant source of joy and happiness in my life for almost 10 years now and anything I write here would not do justice to our friendship. Also thanks to Shinde, koondi, Vatsa and goel for keep coming to Amsterdam and reminding me of good old kgp times!!

Next, I want to thank Rakesh, Jurn, Alberto, Adithya, Sander, Nitish, Lou, Faidon, Wagner, Ariana, Arthur, Nazila, Chunhua, Mojgan, Hannah, Becca, David, Chiara, Yinying, Ferry, Thor, Tuen, Ivan and Andrii for being amazing colleagues and even better friends, especially Rakesh for being an awesome roommate and
Acknowledgement

Thor for drawing the hands of god and the molecules for the thesis cover. Also to Sander and Jurn for being always ready to help, gruesome coffee corner discussions and introducing me to the Dutch culture. I also want to thank people from the other side of the coffee corner, Max, Benjamin, Dorina, Sheetal, Dongdong and Jan for all the coffee chats and HIMS borrel fun. Further, I want to thank Andreea, Maria, Chandan, Pim, Lucas, Sandra, Rhea and Nicole for all the fun conversations at Friday drinks. Thanks to Herma, Evert Jan, Joceylene, David for all the help, especially Herma for helping me with the projects and thesis corrections when Bernd was injured. I also want to thank Marcel, Renate and Ineke for their constant help and advice throughout my PhD.

I further want to thank Roberto, Huib and Sander for many fruitful collaborations, Pablo and Zeila for kindly hosting me in Barcelona and helping me with TranSiesta, and finally Nitish and Evert Jan for several discussions and quantum meetings.

The most important people in the office are all the nice students, I had the opportunity to work with. I want to thank Philip, Tamika, Stan, Laura, Dylan, Yol and Jim for all the fun, especially Philip for the Dutch translation of my thesis summary. You guys were really fast to learn and always open to new ideas and challenges.

The next group of people I want to thank is the Bollywood Amsterdam, which made my four years pass in a blink. I met the group in the very first week of my arrival and the dance still continues! I want to start by thanking Danitia, Lalitha, Anchal and Sneha for the awesome parties, travel escapades and all the fun. I also want to thank Peter, Faiza, Ankit, Martina, Gaurav, Fabio, Ruchi, Ravi, Abhas, Freddy, Boris, Divya, Delia, Vir and Sohail for all the amazing dancing nights and backstage madness.

During my stay in Amsterdam, I undertook many journeys across the continent and beyond. I was lucky to have amazing friends to go with. It all began with a 24 hours long journey by bus to Barcelona with Vivek, that was so much fun! Then the next trip was with Lalitha and Vishnu to October festival followed by many more, which will end up being another thesis if I write about each of them. Thanks for giving me company in my wanderlust. Next two years flew by exploring countless places with Verena, Yuliia, Andreea, Tony and Vlad. I had so much fun travelling with Mangal to Finland, Norway, Belgium along with Arjun, Hannah, Mohsin, Marloes and Neha, thank you guys for being such sweethearts! I want to thank the Squash group particularly Mark, Gwylan, Allen, Jeroen, Zita, Onno, Chiara and Hans for teaching me squash, organizing a cool trip to Belgium and several other events. The latest travel buddies are also amazing, for which I want to thank Ayush, Souloke and Arno, especially the trip to Poland will go deep down
in the memory lane.

Now is the time to introduce the Temple gang, because of whom I never had to think about Sunday plans. I realized that going to Gurudwara is not awesome just because of the food, but also because of all the after-fun. For this, I want to thank Neeraj, Prasanna and Saikat. I also want to thank fellow Shellers Vishnu, Sangeetika, Anwesha, Soumen, Nitin, Priyanka, Nishant, Harshil, Anirudha, Monika, Vikas, Deepak and Stephen. All the Shell meetings were so much fun because of you guys.

I am going to stop for now and say a collective thanks to everyone who helped me during my PhD and I hope to see some of you in my future endeavors!