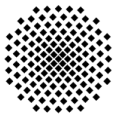
Surface Adsorption on Interstellar Ice I_h

Thomas Bissinger



University of Stuttgart
Germany

Master Thesis supervised by

Prof. Dr. rer. nat. Johannes Kästner

Dr. Jens Smiatek

M. Sc. Jan Meisner

Institute for Theoretical Chemistry, University of Stuttgart, January 2016

Declaration of Authorship

I, Thomas Bissinger, declare that this Master's thesis, titled "Surface Adsorption on Interstellar Ice I_h", and the work presented in it are my own. I have documented all sources and material used. Neither this thesis nor parts of it have been previously presented to another examination board nor have they been published elsewhere.
I confirm that the electronic copy handed in is identical to this printed version.

Signed:

Date:

Abstract

The formation of many molecules in the interstellar medium, especially water with its many formation routes, is thought to be assisted by surface reactions on ice mantles of interstellar dust grains. We introduce a surface adsorption model for a water I_h Fletcher surface within a QM/MM description. For the MM region, we use the TIP3P force field, for the QM region we present a hybrid basis set in combination with a selection of functionals. The performance of DFT functionals have been evaluated in a gas-phase interaction energy benchmark. We give formation and reaction energies for gas-phase molecules as well as adsorption and Eley–Rideal type reaction energies for molecules on the surface. For the latter, we give energy data by combining DFT adsorption energies and coupled-cluster gas-phase reaction energies. We also show that it is sufficient to compute the ZPE correction with only a small selection of the QM molecules instead of the full QM region. This greatly reduces the computational cost. An outlook on binding site analysis of triplet O and a transition state search for the reaction $OH + H_2 \longrightarrow H_2O + H$ shows that the model is fit for further application.

Zusammenfassung

Noch immer besteht Unklarheit über die Details der Entstehung einiger Moleküle im interstellaren Raum, insbesondere die von Wasser, das über verschiedene Reaktionspfade entstehen kann. Eine prominente Theorie nimmt an, dass viele relevante Reaktionen durch Adsorption auf Oberflächen vermittelt werden. Dabei handelt es sich z.T. um Eisoberflächen auf interstellaren Staubpartikeln. Wir stellen ein Modell für Oberflächenadsorption auf einer Eis I_h Fletcher Surface vor, die mit einem QM/MM-Kopplungsverfahren beschrieben wird. Die Wassermoleküle des MM-Gebiets interagieren untereinander mit dem TIP3P-Kraftfeld. Für das QM-Gebiet empfehlen wir eine Hybridbasis und geben eine Auswahl potentieller DFT Funktionale an, die zuvor in einem Benchmark für Interaktionsenergien getestet wurden. Mit diesen Funktionalen berechnen wir sowohl Bildungs- und Reaktionsenergien in der Gasphase als auch Adsorptionsenergien und Eley–Rideal-Reaktionsenergien auf der Oberfläche. Für die Oberflächenreaktionen können wir die Coupled-Cluster Ergebnisse für Gasphasenreaktionsenergien mit den Adsorptionsenergien auf DFT-Ebene kombinieren. Wir zeigen außerdem, dass zur Nullpunktsenergiekorrektur nicht alle Atome des QM-Gebiets berücksichtigt werden müssen, was Rechenzeiten stark reduziert. Wir schließen mit zwei kurzen Studien zur Analyse von Bindungsstellen beim Triplett O und einer Untersuchung des Übergangszustand der Reaktion $OH + H_2 \longrightarrow H_2O + H$ auf der Oberfläche.

Contents

Surface Adsorption on Interstellar Ice I_h

Thomas Bissinger

February 21, 2016

1 Introduction

Interstellar chemistry is the key ingredient to understanding the molecular abundances in our universe. While the formation of atoms takes place in stars, their further reaction and therefore the formation of larger molecules in space largely occurs in interstellar clouds. With modern telescopes it is possible to measure many molecular abundances in the interstellar medium (ISM) to ever increasing levels of accuracy. Over the past few decades, it became evident that the reaction rates governing the formation of molecules can not properly be explained by gas-phase chemistry alone.[1]

A prominent theory to mend this discrepancy is to consider the contribution of surface reactions on interstellar dust grains.[2] The existence of interstellar dust was first inferred by Trumpler,[3] who attributed extinction of starlight in the ISM to the presence of dust. Following this first observation, many investigations confirmed Trumpler's assumption and nowadays the presence of dust grains in many regions of the ISM as well as in circumstellar disks is considered proven.[4,5] Following Draine *et al.*, the vast majority of interstellar dust grains is composed of silicates or carbonaceous material.[6]

From the viewpoint of molecular chemistry, such interstellar dust grains can serve as catalysts for reactions that would not usually take place in the gas phase. The molecules formed by these surface reactions may remain on the dust grains. If many molecules are accreted on a grain, ices may form which have an additional impact on the surface reaction mechanism.

One of the molecular components of such ices is water. It was first observed in the Becklin–Neugebauer infrared point source in the Orion nebula by Gillett and Forrest in 1973.[7,8] In the laboratory, Hiraoka *et al.* found evidence for the formation of water at 12 K by the reaction of D and O in an N_2O matrix in 1998.[9] Today, water is considered the main component of interstellar ices.[10]

At high temperatures the water will not freeze on the grains, but at temperatures as low as in molecular clouds (between 10 and 20 K)[11] some of the grains are believed to have icy mantles.

Among the first to observe water ice in a molecular cloud were Léger *et al.* in 1979.[12] Concluding from the observed abundances, the grain mantle material is mostly composed of H_2O, CO and CO_2. In the case of water, one faces different species of *amorphous solid water* (ASW).[1,13]

The formation of water itself can be described by various reaction channels. In 1982, Tielens and Hagen published an influential paper proposing a network of water formation reactions by hydrogenation of O, O_2 and O_3.[14] The network was refined by later studies, an overview of which was recently presented by Dishoeck *et al.*.[1] The modern version gives an account of the most relevant reactions that lead to water formation. Depending on temperature, density and radiation field, different reactions of the network are assigned varying importance dependent on the interstellar region. For example within a continuous-time random-walk Monte Carlo simulation by Cuppen and Herbst, the reaction $H_2 + OH \longrightarrow H_2O + H$ is most likely dominant in dark molecular clouds, while in diffuse clouds the solid-state reaction $H + OH \longrightarrow H_2O$ is dominant.[15]

The surface can serve as a catalyst to reactions. There are two main mechanisms in molecular clouds to describe that, see Figure 1. The first is the Langmuir-Hinshelwood (LH) mechanism.[16] Here, molecules X and Y both adsorb on the surface. They move by diffusive processes and when they meet, they have a chance to react to a molecule Z (or a set of molecules Z, Z', Z", ...). In the Eley–Rideal (ER) reaction mechanism, only molecule X is adsorbed.[17] Molecule Y approaches X from the gas phase and the two can react to Z (or more molecules). In any case, if the reaction is exothermic, the excess energy has to be compensated by rotational or vibrational excitation of the product(s) – or by the surface atoms. If this is possible, the Z stays intact, otherwise it will react further or undergo desorption. For example, Papoular found in a theoretical study that the ER reaction s-H $+ OH_{(g)} \longrightarrow$ s-H_2O, where s-H means adsorbed hydrogen, leads to immediate

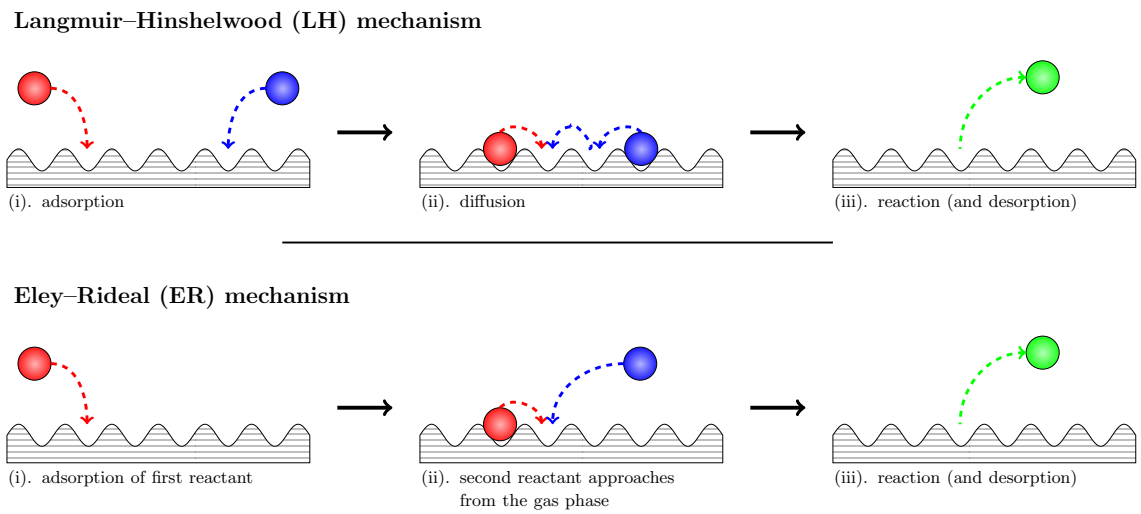


Figure 1 LH and ER reaction mechanisms schematically represented. For more explanation, see text.

desorption of the resulting H_2O,[18] while another theoretical study by Bergeron *et al.* predicted that the product of the LH reaction s-H + s-O $\longrightarrow$ s-OH remains mostly adsorbed on a graphite surface.[19] If the reaction is endothermic or has a barrier, the energy difference has to be supplied by the surface if it can not be supplied by the kinetic energies of the molecules (and neglecting quantum tunneling).

While experimental research in the field of surface astrochemistry has grown over the past years, theoretical data is still sparse. Especially *ab initio* calculations have so far been challenging due to the large system size and the resulting difficulties to find methods that can still give results with acceptable accuracy. Still, there is some noteable research which we want to name here.

Ice surface structure was studied by Cabrera Sanfelix *et al.* in 2003.[20] They analyzed the early accretion of water on a bare graphite grain using density functional theory (DFT) with the PW91 functional.[21,22] They found that the water is physisorbed on the surface and that for high coverages of the surface, ice-like layers are formed. A 2005 study by Lin *et al.* investigated adsorption of small water clusters (not more than six molecules) on a graphite surface with a density functional tight-binding method and a dispersion correction. They found that the ring structure for a water hexamer is not the most stable structure when physisorbed onto the surface.[23]

DFT methods were also employed by Goumans *et al.* to study the hydrogenation of CO on a silica surface with a QM/MM (Quantum Mechanics/Molecular Mechanics) coupling method. They were able to calculate binding energies and the catalytic effect of the surface on the formation of CH_3OH.[24] With similar methods, Goumans *et al.* studied the formation of water on an interstellar dust grain.[25] They gained transition states with the climbing image nudged-elastic band method[26].

For water surfaces, theoretical studies by Woon[27] and Xie *et al.*[28] studied the effect of a water ice surface on reaction barriers. Woon used MP2[29] alongside the QCISD(T)[30] method with double-ζ precision, while Xie *et al.* employed DFT with B3LYP with the 6-31G(d) basis set. Both concluded that H_2O adsorption may significantly alter reaction barriers.

While not directly employing electronic structure theory, Karssemeijer fitted parameters for a classical force field to experimental and *ab initio* data and calculated binding energies and diffusion coefficients with these.[31]

Other computational approaches to ice adsorption did so far not employ an electronic structure theory directly nor indirectly by force fields but used simulation models which had diffusion coefficients and adsorption energies as input values. There is no fully reliable data available for many of these, therefore the simulation parameters were tuned, starting from educated guesses of some kind, to fit experimental data. While this approach can be justified and yields indeed not only good approximations to experimental data but also gives rise to new conclusions, the resulting

parameters still have little physical meaning.

Our approach is therefore to create a model for water I_h ice surface adsorption from scratch. We want to describe a crystalline water surface with enough ice layers to ignore effects of the underlying dust grain, which is assumed to be large enough to allow for a regular planar ice surface. These idealizations will not be met by the majority of ice mantles on interstellar dust grains, therefore we will not put too much stress on the geometric properties of the surface but rather on computational details. The surface will be divided in a quantum mechanical and molecular mechanical region coupled via a QM/MM scheme. The MM scheme is described by a classical force field and we present a selection of DFT functionals that yield reasonable results for adsorption energies. Finally, we give a basis set recommendation.

We will describe the theory underlying the model in the next section. Section 3.1 then describes the benchmarking we performed on smaller test systems to determine the best functionals and basis sets in describing intermolecular interactions and therefore adsorption. With this DFT framework, we analyze gas-phase formation and reaction energies in Section 3.2. Subsequently, we present the geometry and basic properties of the model and give our results for adsorption energies and surface reaction energies in Section 3.3. As a suggestion of further applications, we give an idea of how one can analyze binding sites and transition states for adsorabtes in a small section following that. Finally, there will be concluding remarks and an outlook on possible further application for our findings in Section 4.

2 Theoretical Background

This section focuses on the theoretical framework of the ice surface model. We introduce the main chemical nomenclature in Section 2.1 and then proceed to the physical and mathematical ideas behind DFT in Section 2.2. After that, Section 2.3 will explain how we describe the MM interaction of the system and Section 2.4 explains how QM and MM are coupled by the QM/MM procedure. Finally, Section 2.5 describes how to find energy minima of the potential energy surface.

2.1 Different Types of Energy

We will consider the *interaction energy* between two molecular species X and Y. We call the system that combines both molecules $X-Y$. We also consider the *adsorption energy* of a molecule X on the ice surface. We call this system s-X.

We describe the interaction energy first. Consider a system of two molecules X and Y. We can calculate the energy of the isolated molecule X to be E_X and the energy of the isolated molecule Y to be E_Y. We can also calculate the energy of the full system $X-Y$, which will in general depend

on the distance and the orientation of the two molecules, to be $E_{\mathrm{X-Y}}$. Then, the interaction energy between the two molecules is the energy given by

$$E^{\mathrm{int}}_{\mathrm{X-Y}} := E_{\mathrm{X-Y}} - E_{\mathrm{X}} - E_{\mathrm{Y}}. \tag{2.1}$$

We did not include spatial dependence of $E_{\mathrm{X-Y}}$ into the above definition. A map

$$(\mathbf{R}_{\mathrm{X-Y}}, \Omega_{\mathrm{X}}, \Omega_{\mathrm{Y}}) \mapsto E^{\mathrm{int}}_{\mathrm{X-Y}} \tag{2.2}$$

with the center of mass separation $\mathbf{R}_{\mathrm{X-Y}}$ and the molecular orientation Ω_{X} and Ω_{Y} is called the *potential energy surface* (PES) of the intermolecular interaction. It may also contain internal deformations of the molecule. For that case, one can either include the deformations in E_{X} and E_{Y} or leave them out, depending on the problem to be analyzed.

However, one does often speak of the interaction energy of two molecules without further specification of a point on the PES. This is usually a reference to the *optimum geometry* of $\mathrm{X-Y}$, that is the global minimum of the PES $E_{\mathrm{X-Y}}$ and therefore $E^{\mathrm{int}}_{\mathrm{X-Y}}$. If the interaction between X and Y were purely repulsive, that is $E^{\mathrm{int}}_{\mathrm{X-Y}} > 0$ for all geometries, the global minimum would not be well-defined since it requires inifinite separation of X and Y in an arbitrary direction. However, the algorithms we use will converge to a local minimum, by which we will then classify the strength of the repulsion. But we are not able to determine whether the potential energy minimum we find is the global minimum or within what error its energy is to the global minimum. One must also keep in mind that the optimum geometries to which the system converges depends on the initial geometry from which the search algorithm starts, since there is generally more than one local minimum.

The definition of the adsorption energy is mostly similar to the one for the interaction energy. There, we have the energy of the isolated species E_{X} and the energy minimum of the surface E_{S}. If we denote the system of the surface with the adsorbed molecule X by s-X and its energy minimum by $E_{\text{s-X}}$, we define the adsorption energy to be

$$E^{\mathrm{ads}}_{\mathrm{X}} := E_{\text{s-X}} - E_{\mathrm{X}} - E_{\mathrm{S}}. \tag{2.3}$$

Again, we did not include the dependence of $E_{\text{s-X}}$ and E_{S} on the respective geometries. We even specified that we will consider the individual geometry of minimum energy here. This makes sense because the surface geometry of the system s-X (surface + molecule) may be different from the system S of the surface alone when comparing energy minima. For a fixed value of E_{S}, we could again consider a PES of the type $(\mathbf{R}_i)_i \mapsto E^{\mathrm{ads}}_{\mathrm{X}}$, where the vector $\mathbf{R}_i$ is the coordinate of the

i-th atom in s-X, $1 \leq i \leq N$ for some N. The different energy minima of this map to $E_{\mathrm{X}}^{\mathrm{ads}}$ are called *binding geometries*, and the position and orientation of the molecule X in a binding geometry is called a *binding site*. Exploring binding sites and the strength of the binding $E_{\mathrm{X}}^{\mathrm{ads}}$ may have further use to determining parameters for simulations.

We will also give reaction energies. In this work, we distinguish two kinds of reactions: Gas-phase reactions, in which isolated molecules come in contact, and ER type surface reactions as described in the introduction. We will only consider gas-phase reactions according to the scheme

$$\mathrm{X}_{(\mathrm{g})} + \mathrm{Y}_{(\mathrm{g})} \longrightarrow \mathrm{Z}_{(\mathrm{g})}, \tag{2.4}$$

where the subscript (g) indicates gas-phase molecules. The initially isolated reactants $\mathrm{X}_{(\mathrm{g})}$ and $\mathrm{Y}_{(\mathrm{g})}$ combine to form a product $\mathrm{Z}_{(\mathrm{g})}$. Reaction (2.4) is actually ill-defined since it does not contain information about the vibrational, rotational or electronic state of the reactants and the product. There is also no indication of radiative association, that is a photon of energy $h\nu$ on the left or the right side of the reaction that may also be necessary to maintain conservation of energy. Therefore, equation (2.4) is only a formal representation which we use to stay in analogy to a surface reaction (see below).

The reaction energy consumed or released in (2.4) is

$$E_{\mathrm{gas}}^{\mathrm{react}} := E_{\mathrm{Z}_{(\mathrm{g})}} - E_{\mathrm{X}_{(\mathrm{g})}} - E_{\mathrm{Y}_{(\mathrm{g})}}. \tag{2.5}$$

All energies on the right hand side are ground state energies, possibly corrected for zero-point energy (see below). Positive $E_{\mathrm{gas}}^{\mathrm{react}}$ means an *endothermic*, negative $E_{\mathrm{gas}}^{\mathrm{react}}$ an *exothermic* reaction.

It is also possible to calculate *reaction energies* $E_{\mathrm{ER}}^{\mathrm{react}}$ for the ER mechanism. For that, species X is adsorbed on the surface to s-X and in equilibrium, that is at optimum geometry. From the surrounding gas, a molecule of species Y approaches. The two react to form s-Z. This follows the scheme

$$\text{s-X} + \mathrm{Y}_{(\mathrm{g})} \longrightarrow \text{s-Z}, \tag{2.6}$$

much similar to (2.4).

The $E_{\mathrm{ER}}^{\mathrm{react}}$ reaction energy is

$$E_{\mathrm{ER}}^{\mathrm{react}} := E_{\text{s-Z}} - E_{\text{s-X}} - E_{\mathrm{Y}_{(\mathrm{g})}}. \tag{2.7}$$

Again, all energies are ground-state energies with possible zero-point energy correction.

ER reactions, adsorption energies and gas-phase reactions can be related. We can see that when

starting from (2.7) and inserting (2.3), we find

$$\begin{aligned} E_{\mathrm{ER}}^{\mathrm{react}} :=& E_{\text{s-Z}} - E_{\text{s-X}} - E_{\mathrm{Y_{(g)}}} \\ =& E_{\mathrm{Z}}^{\mathrm{ads}} + E_{\mathrm{S}} + E_{\mathrm{Z_{(g)}}} - \left(E_{\mathrm{X}}^{\mathrm{ads}} + E_{\mathrm{S}} + E_{\mathrm{X_{(g)}}}\right) - E_{\mathrm{Y_{(g)}}} \\ =& E_{\mathrm{Z}}^{\mathrm{ads}} - E_{\mathrm{X}}^{\mathrm{ads}} + E_{\mathrm{Z_{(g)}}} - E_{\mathrm{X_{(g)}}} - E_{\mathrm{Y_{(g)}}} \\ =& E_{\mathrm{Z}}^{\mathrm{ads}} - E_{\mathrm{X}}^{\mathrm{ads}} + E_{\mathrm{gas}}^{\mathrm{react}}. \end{aligned} \tag{2.8}$$

This result is important. It means that gas-phase reaction energies and adsorption energies can be calculated separately to yield ER reaction energies. The two components require very different system sizes – small gas-phase molecules and large surface+molecule systems. But equation (2.8) allows us to independently choose methods adequate to the requirements of these two separate calculations. We will use this result when discussing ER reaction energies in Section 3.3.5.

The ER reaction scheme can basically occur for all molecules X and Y present in the interstellar medium. But astrochemically, only reactions with H and H_2 are likely to participate in an ER reaction scheme due to the high abundances of these two in space.

The superscripts and subscripts on E^{int}, E^{ads} and E^{react} may be ignored if it is clear which energy is meant.

We also want to introduce an energy correction, the *zero-point (vibrational) energy* (ZPE). We describe it for some general system that may contain any arrangement of atoms. For all calculations we perform, we will work with fixed values for the atomic coordinates $\mathbf{r}_i$, $1 \leq i \leq N$ for some $N \in \mathbb{N}$. That description can only be accurate if the atoms were classical particles. However, if we want to allow for them to be quantum objects, we need to include uncertainty into their position. We do this by a harmonic approximation to the energy, which yields the the zero-point vibrational energy of he atoms. It is computed by

$$E^{\mathrm{ZPE}} = E + \Delta E^{\mathrm{ZPE}}, \tag{2.9}$$

where the correction term ΔE^{ZPE} is the *zero-point (vibrational energy) correction*. It depends on the eigenvalues of the *Hessian* matrix $\mathbf{H}$ of the system,

$$\mathbf{H}(\mathbf{r}_1, \ldots, \mathbf{r}_N) := \left(\frac{\partial^2 E}{\partial \mathbf{r}_i \partial \mathbf{r}_j} (\mathbf{r}_1, ..., \mathbf{r}_N) \right)_{1 \leq i,j \leq N}. \tag{2.10}$$

We will use the superscript in E^{ZPE} if we want to denote energies that are corrected with ΔE^{ZPE} as in (2.9), unless it is clear from the context that we mean energies with ZPE correction.

2.2 Methods of Quantum Chemistry

We already saw a few different energy expressions so far. The accurate calculation of these is naturally vital to anything we want to do in this work. We now want to focus on the methods of quantum chemistry which will be used to describe the quantum mechanical part of our system.

2.2.1 Basics of Density Functional Theory

For a system with a time-independent potential V, one is usually interested in a solution of the *time-independent Schrödinger equation*

$$\hat{H}\,|\Psi\rangle = E\,|\Psi\rangle\,. \tag{2.11}$$

This equation holds for all non-relativistic quantum mechanical particles. Within the *Born–Oppenheimer* approximation, one can separate the dynamics of the atomic nuclei from the dynamics of the electrons. We will treat the nuclei in a classical way and only later incorporate the ZPE correction to reduce that error. Therefore, we will focus on solving the Schrödinger equation for $N \in \mathbb{N}$ electrons, that is the Hamiltonian of our system is

$$\hat{H} = -\sum_{i=1}^{N} \frac{\hbar^2}{2m_\mathrm{e}} \nabla_i^2 + \frac{e^2}{4\pi\varepsilon_0} \sum_{1\le i<j\le N} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + V(\mathbf{r}^N). \tag{2.12}$$

$\mathbf{r}_i$ is the spatial coordinate of electron i and ∇_i^2 is the *Laplace operator* applied to the three coordinates contained in $\mathbf{r}_i$. The electrons move in an external potential V given by the geometry of the nuclei and their movement, where $\mathbf{r}^N$ is the vector containing all electron coordinates. We will have for $K \in \mathbb{N}$ atomic nuclei

$$V(\mathbf{r}^N) = \frac{e^2}{4\pi\varepsilon_0} \sum_{1\le A<B\le K} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} - \frac{e^2}{4\pi\varepsilon_0} \sum_{A=1}^{K} \sum_{i=1}^{N} \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|}. \tag{2.13}$$

Here, Z_A is the atomic number of atom A and $\mathbf{R}_A$ is the coordinate of it with corresponding ∇_A^2. We can separate V into a core-core (cc) and an core-electron (ce) potential

$$V(\mathbf{r}^N) = V_\mathrm{cc} + V_\mathrm{ce}(\mathbf{r}^N) \tag{2.14}$$

with

$$V_\mathrm{ce}(\mathbf{r}^N) = \sum_{i=1}^{N} \tilde{V}(\mathbf{r}_i) = -\frac{e^2}{4\pi\varepsilon_0} \sum_{A=1}^{K} \sum_{i=1}^{N} \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|}. \tag{2.15}$$

There will be more than one solution to (2.11), so one can construct the set of all solutions $\{|\Psi_i\rangle\,|i \in \mathbb{N}_0\}$ with corresponding energy eigenvalues E_i. $\{|\Psi_i\rangle\}$ is always the complete basis

of some $\mathbb{C}$-vector space $\mathcal{H}_a$, where the subscript a denotes antisymmetry according to the *Pauli principle*

$$\begin{aligned} &\langle \mathbf{x}_1, \ldots, \mathbf{x}_l, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_N | \Psi_i \rangle \\ = \quad &\Psi_i(\mathbf{x}_1, \ldots, \mathbf{x}_l, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_N) \\ = &-\Psi_i(\mathbf{x}_1, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_l, \ldots, \mathbf{x}_N). \end{aligned} \tag{2.16}$$

We use $\mathbf{x}_i = (\mathbf{r}_i, s_i)$ for orbital coordinates $\mathbf{r}_i$ and the spin coordinate s_i.

If $\mathcal{H}_a \subseteq \mathcal{L}^2$, which is typically the case, the set of solutions can be chosen to be orthonormal $\langle \Psi_i | \Psi_j \rangle = \delta_{ij}$. The set of solutions is usually not finite and the set of eigenvalues (energies) E_i of $\hat{H}$ does not necessarily have an upper bound. But there is always a minimum energy, denoted by E_0, which we call the *(electronic) ground state energy*. The corresponding eigenvector $|\Psi_0\rangle$ is the *(electronic) ground state*. They can both be obtained by the variational ansatz

$$\begin{aligned} E_0 &= \min_{|\Psi\rangle} \left\{ \langle \Psi | \hat{H} | \Psi \rangle \right\}, \\ |\Psi_0\rangle &= \arg\min_{|\Psi\rangle} \left\{ \langle \Psi | \hat{H} | \Psi \rangle \right\}. \end{aligned} \tag{2.17}$$

Note that while E_0 is unique, there may be multiple possibilities for $|\Psi_0\rangle$, although we will not consider that case.

Now, equation (2.11) has a variety of equivalent counterparts. One of them is the key to the approach of DFT. When multiplying (2.11) by the bra $\langle \Psi |$, one can interpret the resulting energy as a (non-linear) functional of the wavefunction by

$$E : \mathcal{H}_a \to \mathbb{R}, \qquad E\left[|\Psi\rangle\right] = \langle \Psi | \hat{H} | \Psi \rangle, \tag{2.18}$$

which would mean that the ground state energy can be found by minimizing the functional $E\left[|\Psi\rangle\right]$ according to (2.17). But so far, the minimization of said functional only differs in semantics from the task of minimizing the energy expectation value.

A truly new task arises from considering the *electron density* ρ instead of the wave function Ψ. The two approaches are related, since the ground state electron density is given by

$$\rho(\mathbf{x}_1, ..., \mathbf{x}_N) = |\Psi_0(\mathbf{x}_1, ..., \mathbf{x}_N)|^2, \tag{2.19}$$

This N-electron density describes the probability of finding the system in a state within a small volume of $\mathrm{d}\mathbf{x}_1 \cdots \mathrm{d}\mathbf{x}_N$ around $(\mathbf{x}_1, \ldots, \mathbf{x}_N)$. The N-electron density can be reduced to the one-electron density by

$$\rho(\mathbf{r}_1, s_1) = \int \mathrm{d}\mathbf{x}_2 \cdots \int \mathrm{d}\mathbf{x}_N \, |\Psi_0(\mathbf{x}_1, ..., \mathbf{x}_N)|^2, \tag{2.20}$$

where the integrals run over the full spin-orbit space for all particles but the first.

For a system of N electrons, Hohenberg and Kohn were able to show that the ground state one-electron density uniquely determines the Hamiltonian except for the addition of a constant, and that conversely there is a functional of the density that has its minimal value at the one-electron ground state density, and for which the minimum value is the ground state energy.[32] Therefore, the task of solving the Schrödinger equation (2.11) is reduced to the task of finding the minimum of this density functional.

The problem here is that not much is known about the nature of this density functional. Especially the kinetic energy, which has a clear and simple form for orbitals, has no accurate representation in orbital-free DFT.

The popular ansatz by Kohn and Sham reintroduces orbitals into DFT.[33] It is called Kohn–Sham (KS-)DFT, and the energy expression is given by

$$E[\rho] = V_{\mathrm{cc}} + \int \mathrm{d}\mathbf{r}\, \tilde{V}(\mathbf{r})\rho(\mathbf{r}) + \frac{e^2}{8\pi\varepsilon_0} \int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + T_s[\rho] + E_{\mathrm{xc}}[\rho] \tag{2.21}$$

The two problematic terms that remain are the *kinetic energy functional* $T_s[\rho]$ and the *exchange correlation functional* $E_{\mathrm{xc}}[\rho]$. The former is treated in the Kohn–Sham approach by introducing *Kohn–Sham orbitals* ϕ_i that solve the equations

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V_{\mathrm{eff}}(\mathbf{r}) \right) \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}) \tag{2.22}$$

with the *effective potential*

$$V_{\mathrm{eff}}(\mathbf{r}) = V_{\mathrm{cc}} + \tilde{V}(\mathbf{r}) + \frac{e^2}{2} \int \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho(\mathbf{r})}. \tag{2.23}$$

The latter term is often expressed by the *exchange correlation potential*

$$V_{\mathrm{xc}} = \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho(\mathbf{r})}. \tag{2.24}$$

With this orbital approach, the full wavefunction can be expressed as a Slater determinant of the ϕ_i, and reducing the corresponding density operator yields a density according to

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2 . \tag{2.25}$$

With that, the kinetic energy functional is given by

$$T_s[\rho] = \sum_{i=1}^{N} \int \mathrm{d}\mathbf{r}\, \phi_i^*(\mathbf{r}) \left(-\frac{\hbar^2}{2m} \nabla^2 \right) \phi_i(\mathbf{r}). \tag{2.26}$$

The last remaining task is to find expressions for $E_{\mathrm{xc}}[\rho]$ (see below).

After one decides on some functional for E_{xc}, one needs to find *self-consistent* solution to (2.22). Since the KS orbitals ϕ_i define the potential V_{eff} in (2.23), these solutions are called the *self-consistent field* (SCF). (2.22) is usually solved iteratively, where one chooses some initial guess for the ϕ_i, then solves (2.22) for the resulting effective potential and uses the new ϕ_i to establish a new potential until convergence is reached. There are numerous techniques like damping and orbital shifts to support the convergence, however there is no mathematical guarantee for it.

2.2.2 DFT Functionals

Over the years, many different expressions for E_{xc} were proposed and used, each defining a separate density functional with different advantages and disadvantages. There are three main classes of KS-DFT functionals: *local density approximation* (LDA) functionals are functionals where $V_{\mathrm{xc}}(\mathbf{r})$ depends only on the value of $\rho(\mathbf{r})$, which may be sensible for slowly varying densities. The more universal *generalized gradient approximation* (GGA) additionally incorporates dependencies on $\nabla\rho(\mathbf{r})$ into $V_{\mathrm{xc}}(\mathbf{r})$. Thirdly, there are approximations that use linear combinations of LDA and GGA calculations and the Hartree-Fock exact exchange energy to obtain more accurate results.[34] These functionals are called *hybrid* functionals. Also, there is a variety of *meta-hybrid GGA* functionals that promise to be potentially even more accurate.

We chose functionals from the GGA, meta-GGA and (meta-)hybrid GGA class.

GGA: BP86,[35,36], BLYP,[35,37] PBE[38] and B97-D.[39]

Meta-GGA: TPSS[40] and PW6B95.[41]

Hybrid: B3LYP,[34,42] BHLYP,[43] TPSSh[40,44] and PBE0.[38,45]

At the beginning of our research, we also considered the meta-GGA functional M06.[46] But in our tests it had the strange deficiency of convergence failures when unbound H-radicals were present in the system, starting with systems as small as $H_2O + H$. The calculations still converged for geometries close to the optimum geometry, but even just slightly longer separations of H and H_2O lead to failure of the SCF convergence. This problem seemed to be mostly independent of the choice of basis set and the implementation of M06 (the problem occured both in TURBOMOLE[47] and NWCHEM[48] calculations), and it could not be solved by standard approaches to facilitate convergence. For that reason, M06 is not included in our further considerations despite it being among the most accurate functionals.

2.2.3 Basis Sets

Now, the ansatz (2.25) is not yet fit for computational application since the functions ϕ_i may pertain to the infinite-dimensional Hilbert space $\mathcal{L}^2(\mathbb{R}^3)$. As usual, we have to restrict the approximation to a finite subspace $\{\chi_\mu \mid 1 \leq \mu \leq M\} = X \subset \mathcal{L}^2(\mathbb{R}^3)$. The standard approach is then to find the functions ϕ_i by *linear combination of atomic orbitals* (LCAO). An atomic orbital is a function $\chi_\mu(\mathbf{r} - \mathbf{R}_{A_\mu})$ for some core A_μ. In the LCAO approximation, one chooses X to be a set of M (approximated) atomic orbital functions, then determines the uniform transformation matrix to diagonalize (2.22) and chooses the N linear combinations $\psi_i = \sum_\mu C_{i\mu}\chi_\mu$ that correspond to the N minimum eigenenergies ε_i. The individual atomic orbital functions vary with atom species. In principle, one could also include basis functions that are not centered around an atom, but we will not consider any such basis sets.

For the atomic orbitals, one often uses *Gaussian type orbitals* (GTO). They resolve the angular dependence by usual *spherical harmonics* Y_{l_μ,m_μ} and the radial dependence by a Gaussian bell curve. The main advantage of bell curves is that one can integrate products of them by simple rules. A main disadvantage is that to get close to the more accurate 1-electron atomic orbitals, the *Slater type orbitals* (STO), one must use linear combinations of multiple GTO per STO. In many cases, the computational efficiency, however, compensates the increase in the number of basis functions. Basically, a larger basis set increases computational accuracy while also demanding more computational resources, which means that some compromise between these two has to be found. This is also done by benchmarking. In our benchmark, we used the def2-SVP,[49] def2-TZVP[50] and def2-QZVP[51] basis sets. We also used the def2-SVPD, def2-TZVPD and def2-QZVPD basis sets with additional diffuse functions,[52] the def2-TZVPP[50] basis set with additional polarization functions and finally the def2-TZVPPD[52] basis set with both additional polarization and diffuse functions.

Also, due to the truncation of the basis set (that is the operation in the finite subspace X), the *basis set superposition error* BSSE arises. That is the basis set of molecule Y plays a part in describing the energy of an electron at the molecule X, thus lowering the energy minimum. If one calculates intermolecular interaction, the difference (2.1) can not be calculated by the difference between the system $\mathrm{X}-\mathrm{Y}$ and the two isolated systems X and Y since the latter do not have the same contribution from the other basis set. This error is normally corrected by considering instead of the system X the system X with the basis functions for X and the additional basis functions of Y. This *dummy atom* Y does not have a mass nor are its electrons included. Analogously, one replaces the system Y. For more than two interacting molecules, this *counterpoise* (CP) correction gets more complicated, see for example Anacker and Friedrich.[53]

We will not include a CP correction in our calculations. The reason is that we will mainly work

with def2-TZVP and def2-TZVPD, so basis sets large enough to consider the BSSE negligible. However, the comparison between def2-TZVP and for example def2-SVP results is then somewhat unfair since the def2-SVP basis set without the CP correction may be significantly less accurate than def2-SVP with CP correction. In the next section's benchmark, we will compare results for different methods to CP corrected reference energies, so the BSSE is at least indirectly accounted for.

2.2.4 Dispersion Corrections

It is well known that contemporary DFT can still not account for dispersion, i.e. *Van der Waals* (VdW) interaction. The problem is mainly explained by a missing long-range self-interaction of the density in the functionals.[54] Since water-water interaction is strongly affected by non-covalent bonds, a correct account of dispersion is desirable for our approach to the system. There were a variety of correction terms proposed,[55,56] of which we will use Grimme's D3 correction.[57] Its most basic idea is to add a dispersion correction to the energy E_{KS} obtained from minimizing (2.21). This yields

$$E_{\text{DFT-D3}} = E_{\mathrm{KS}} + E_{\mathrm{disp}}, \tag{2.27}$$

with E_{disp} including two-body and three-body terms. The two-body terms are given by

$$E^{(2)} = -\sum_{AB}\sum_{n=6,8} s_n \frac{C_n^{AB}}{r_{AB}^n} f_{d,n}(r_{AB}) \tag{2.28}$$

with scaling factors s_n, damping functions $f_{d,n}$ and the average isotropic nth order *dispersion coefficients* C_n^{AB}. The sum runs over all pairs of atoms AB contained in the system. For stability reasons, only the terms for $n = 6, 8$ are included. Grimme chooses $f_{d,n}$ in the form

$$f_{d,n}(r_{AB}) = \frac{1}{1 + 6(r_{AB}/s_{r,n}R_0^{AB})^{-\alpha_n}}. \tag{2.29}$$

R_0^{AB} is a cutoff-radius and α_n is called a "steepness"-parameter. They are both chosen independently of the DFT functional used, as are the coefficients C_n^{AB}. The parameters s_6 and $s_{r,8}$ are both fixed to unity, so only the parameters s_8 and $s_{r,6}$ can be chosen according to the different functionals. All parameters were chosen to minimize the *mean absolute deviation* (MAD) for a large benchmark set.[57]

We will denote dispersion-corrected functionals by METHOD-D3, so BLYP-D3 will be the BLYP functional plus the corresponding D3 dispersion correction.

Note that even for functionals that are already meant to include dispersion there are still D3 corrected versions, like B97-D-D3. For any functional, the correction may sometimes lead to over-

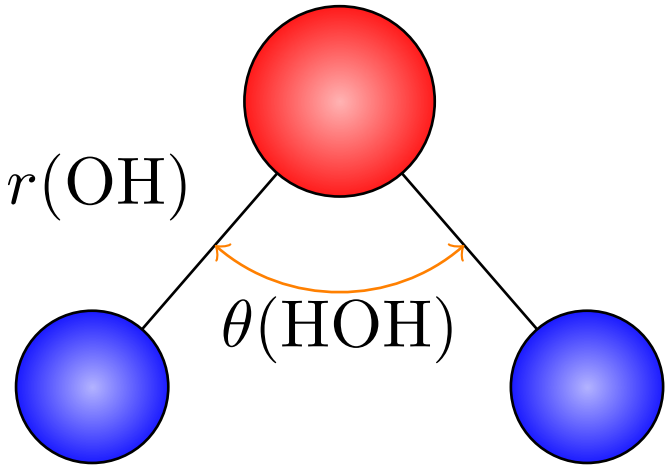


Figure 2 The geometry of TIP3P water. O is red, H is blue. Each atom coincides with a site with charge q and LJ-paramteres ε and σ.

estimation of dispersion or still underestimate it at some circumstances. The effect of the D3 correction will also be tested in next section's benchmark.

2.3 Molecular Mechanics

When there is no bond formation or breaking to be expected, the framework of classical force fields is often also suited to describe a system. That means that all particles move in a classical (effective) potential. We will only treat water molecules as classical particles, so we need a good classical model of water. We decided on the quite popular TIP3P potential by Jorgensen *et al.* [58] The geometry can be seen in Figure 2. The three sites of the TIP3P potential coincide with the atoms forming H_2O. Each site has a charge q_α with *Coulomb interaction*

$$V_{\alpha\beta}^{C}(r_{\alpha\beta}) = \frac{1}{4\pi\varepsilon_0}\frac{q_\alpha q_\beta}{r_{\alpha\beta}}, \tag{2.30}$$

where α,β are the site types and $r_{\alpha\beta}$ is the distance between the sites α and β located at two different water molecules. Sites also interact by Van der Waals forces which are described by a *Lennard–Jones* (LJ) potential

$$V_{\alpha\beta}^{LJ}(r) = 4\varepsilon_{\alpha\beta}\left[\left(\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}}\right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}}\right)^{6}\right]. \tag{2.31}$$

The original TIP3P model only had LJ interaction between the oxygen atoms. In the CHARMM implementation which we use, the H atoms are LJ sites as well. [59] Therefore, a TIP3P water

molecule i moves in the potential of all other TIP3P waters by

$$V_i(\mathbf{r}_i) = \sum_{j \neq i} \sum_{\alpha,\beta} \left[V^{\mathrm{C}}_{\alpha\beta}(r_{\alpha\beta}(\mathbf{r}_i, \mathbf{r}_j)) + V^{\mathrm{LJ}}_{\alpha\beta}(r_{\alpha\beta}(\mathbf{r}_i, \mathbf{r}_j)) \right]. \tag{2.32}$$

The first sum runs over all other water molecules and the second over all pairs of sites. (2.32) contains only intermolecular interaction. There would also be the possibility to include intramolecular degrees of freedom with harmonic potentials for bond and angle stretching. This is not the case for TIP3P, where all sites are fixed within the molecule, reducing the 9-dimensional coordinate $\mathbf{r}_i = \mathbf{r}_i(\mathbf{r}_{\mathrm{O}}, \mathbf{r}_{\mathrm{H1}}, \mathbf{r}_{\mathrm{H2}})$ to a 6-dimensional one, $\mathbf{r}_i = \mathbf{r}_i(\mathbf{r}_{\mathrm{O}}, \Omega)$ with the orientational dependence described by $\Omega \in \mathbb{R}^3$.

The parameters $\varepsilon_{\alpha\beta}$, $\sigma_{\alpha\beta}$ and q_α, q_β as well as the geometric parameters $\mathbf{r}(\mathrm{OH})$ and $\theta(\mathrm{HOH})$ define the TIP3P potential. They are chosen to reproduce (macroscopic) thermodynamic properties of a pure water system. The CHARMM TIP3P potential is fitted to yield good specific heats.[60]

2.4 QM/MM

We already motivated using both QM and MM calculations simultaneously due to the speed of MM and the accuracy of QM results. The simultaneous use demands some notion of coupling between the two systems. The original QM/MM coupling scheme proposed by Warshel and Levitt[61] is nowadays only one among several schemes. We chose an additive QM/MM electrostatic coupling in which the QM part is subject to the external potential defined by the MM charges and the MM part interacts with the QM atoms by the LJ potential (2.31).

2.5 Energy Minima and Transition States

Finding energy minima of the PES for QM or QM/MM systems is basically following the path defined by the gradient (force)

$$\mathbf{F}(\mathbf{R}^M) = -\nabla V(\mathbf{R}^M) \tag{2.33}$$

of the total system's energy $V(\mathbf{R}^M)$, where $\mathbf{R}^M$ is the vector containing all atomic coordinates. Again, there is no knowing whether this leads to a global or a local minimum.

When working on a computer, (2.33) must be integrated in a discretized form. This is best done with step sizes depending on the value of $\mathbf{F}$ and of course one needs tolerances that define convergence. We used the DL-FIND package with the default tolerances to follow the minimum energy path.[62] For optimizations, DL-FIND uses the L-BFGS algorithm.[63]

Evaluating the gradient in (2.33) is fortunately not expensive in comparison to a single energy evaluation of the QM part. The gradient of the MM region can be calculated straighforwardly

from (2.32). For the QM region, there are schemes to obtain analytical results for the gradient for a given wave function.

We will also calculate a *transition state* (TS). This is the state of maximum energy along a *reaction path*, that is a path that starts at some initial optimum geometry with initial *reaction coordinate* ξ_1 and then follows the direction corresponding to the lowest eigenvalue of the Hessian matrix to a saddle point (commonly called the transition state) and after that reaches a new minimum by continuing on its route along the gradient. In Figure 3, an example of such a path between reaction coordinates ξ_1 and ξ_2 is plotted. The energy difference

$$E^{\text{act}} = E_{\text{TS}} - E_1 \tag{2.34}$$

is called the *activation energy*, which is the barrier a reaction would have to overcome on a classical reaction path.

In this work, we obtain transition states with the *dimer method* .[64] It uses two images of the system with a slight offset, then rotates them to find the eigenvector to the minimum eigenvalue of the Hessian matrix (2.10). It follows these minimum eigenvalue directions by an iterative scheme until a saddle point is reached. This is the transition state.

Transition states are important for calculating activation energies as well as reaction and tunneling rates. The former govern reaction rates while the latter are also interesting for the analysis of molecule formation on ice surfaces since at temperatures as low as in the interstellar medium,

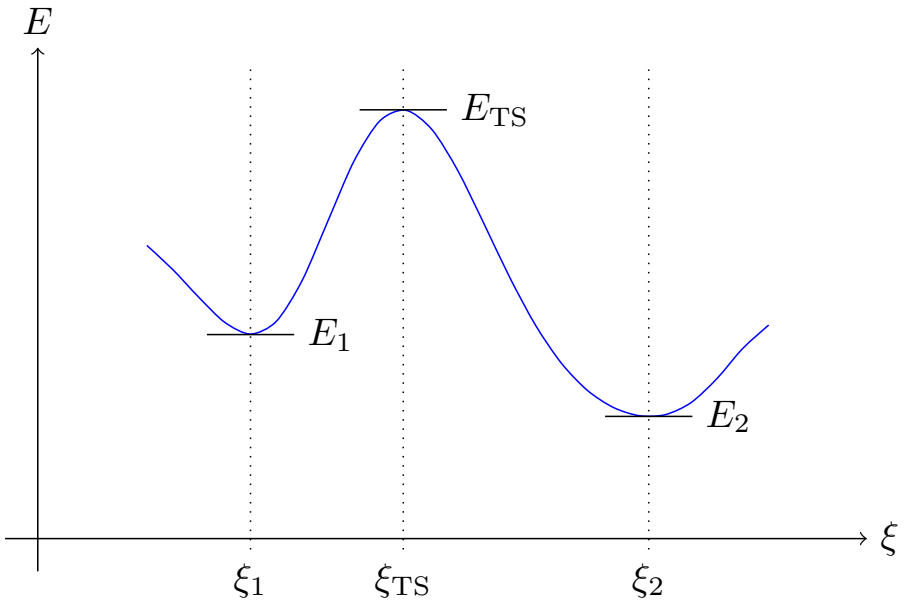


Figure 3 A cut of a PES along reaction coordinate ξ. The reaction coordinate starts at some initial minimum ξ_1 with energy E_1 and follows a path of minimum energy to a neighboring minimum. The second minimum with E_2 is located at ξ_2. The transition state between the two with E_{TS} is a maximum at ξ_{TS}.

tunneling has a significant contribution to the dynamics of these systems.

2.6 Technical Details

Before proceeding to actual calculations, we want to mention the programs and tools used for these. As we already mentioned, the optimizations are carried out with DL-FIND,[62] as are the dimer method and the calculation of Hessian matrices. DFT calculations are performed with TURBOMOLE version 7.0,[47] some preliminary computations used NWCHEM version 6.5.[48] Force field calculations are performed by DL_POLY.[65] These programs were interfaced via ChemShell,[66,67] which also provided the QM/MM coupling.

For the post-HF calculations in the benchmark, we used the MOLPRO package, version 2012.1.[68] All molecular visualizations are created with VMD.[69]

The calculations were performed on the bwForCluster JUSTUS of the German federal country of Baden-Württemberg.[70] Shared memory parallel calculations with up to 16 CPUs were performed. The main limit was the maximal computation time of six days.

3 Results

3.1 Benchmarking

In the previous section, we introduced the basic notions of the theoretical structure underlying our computations. We mentioned already that the choice of DFT functionals and basis sets is of great importance to the accuracy of our calculations. Unfortunately, not much can be said a priori about which functional combined with which basis set should be used to describe a certain system. There are recommendations as some functionals were fitted to best represent a certain group of elements or reactions, but in order to determine the reliability of future results, there is no way around literature research and own benchmark tests.

We already discussed the functionals we want to test in Section 2.2.2, and we have reason to hope that at least the PBE0, PW6B95-D3 and B97-D-D3 functionals will be sufficiently accurate to describe water-water interaction, as they are recommended by Anacker and Friedrichs.[53] The basis sets we want to compare alongside the functionals are those mentioned in Section 2.2.3.

The goal here is to find the best functional/basis set combination to describe adsorption energies for the systems water surface + H or water surface + ^{3}O.* We use the model systems H_2O-H, H_2O-H_2O and $H_2O-{}^3O$.

*Note that we write ^{3}O for the triplet oxygen. The spin multiplicity of a molecule will only be given if the choice is not obvious. In all that follows, this is only the case for the pure oxygen species ^{1}O, ^{3}O, $^1O_2$ and $^3O_2$. All other species have singlet or doublet spin.

In order to perform a benchmark study we use the three interaction energies of the H_2O-H, H_2O-H_2O and $H_2O-{}^3O$ systems. The reference calculations were in all cases performed with the MOLPRO[68] package at the CCSD(T)-F12a[71] level of theory with the VTZ-F12[72] basis set, for which we found energy minima. For $H_2O-{}^3O$, the CCSD(T)-F12a calculations were preceded by *multi-reference* calculation at the MRCISD-F12/VTZ-F12[73] level of theory including Davidson correction[74] to find an optimum geometry, which turned out to be suitable for single-reference treatment.

At each optimum geometry, an additional calculation was carried out to find a CP corrected energy. For H_2O-H and H_2O-H_2O, this calculations was also at the CCSD(T)-F12a/VTZ-F12 level of theory, while for the $H_2O-{}^3O$ interaction again an MRCISD-F12/VTZ-F12 calculation was required due to the necessity of a multireference treatment of the case where the ^3O radical is set as a dummy atom. In all cases, the VTZ-F12 basis set is large enough to lead to only small CP corrections of $0.05-0.18$ kJ/mol.

We performed energy minimizations at the DFT level of theory with different basis sets. The initial geometry was always the reference geometry provided by the CCSD(T)-F12a calculation. The results can be compared to the coupled-cluster data with respect to optimum geometry and energy deviations. We will discuss these for the three test systems separately.

The deviation in interaction energy is calculated by

$$\Delta E^{\text{int}} = E^{\text{int}}_{\text{DFT}} - E^{\text{int}}_{\text{CC}}. \tag{3.1}$$

$E^{\text{int}}_{\text{DFT}}$ stands for the interaction energy (2.1) of the system computed with a certain DFT functional and basis set, and $E^{\text{int}}_{\text{CC}}$ is the CP corrected interaction energy with CCSD(T)-F12a/VTZ-F12.

The deviation of the DFT optimum geometry $\mathbf{R}^M_{\text{DFT}} = (\mathbf{r}^1_{\text{DFT}}, \ldots, \mathbf{r}^M_{\text{DFT}})$ from the reference optimum geometry $\mathbf{R}^M_{\text{CC}} = (\mathbf{r}^1_{\text{CC}}, \ldots, \mathbf{r}^M_{\text{CC}})$ can be quantified by the *root mean square deviation* (RMSD). It is given by

$$\text{RMSD}(\mathbf{R}^M_{\text{DFT}}; \mathbf{R}^M_{\text{CC}}) = \sqrt{\frac{1}{M}\sum_{k=1}^{M} \left|\bar{\mathbf{r}}^k_{\text{DFT}} - \mathbf{r}^k_{\text{CC}}\right|^2}. \tag{3.2}$$

The coordinates $\bar{\mathbf{R}}^M_{\text{DFT}} = (\bar{\mathbf{r}}^1_{\text{DFT}}, \ldots, \bar{\mathbf{r}}^M_{\text{DFT}})$ are calculated by $\bar{\mathbf{R}}^M_{\text{DFT}} = \mathbf{Q}\mathbf{R}^M_{\text{DFT}} + \mathbf{t}$, where $\mathbf{Q}$ is a rotation matrix and $\mathbf{t}$ is a translation vector. $\mathbf{Q}$ and $\mathbf{t}$ are chosen to minimize (3.2), which means that they overlay the coordinates of the DFT optimum geometry and the reference optimum geometry as good as possible. $\mathbf{Q}$ and $\mathbf{t}$ are obtained with the Kabsch algorithm.[75] We will give RMSD in Å.

Note that, obviously, the coupled cluster results are approximations just as the DFT results, however there is good reason to assume that they are more accurate. For the following comparisons, we will therefore consider the coupled cluster results as if they were the correct values for this

interaction.

Another remark before starting the discussion is the accuracy with which numerical integration was carried out. In the TURBOMOLE code, we used an m4 integration grid with SCF iterations converged to 10^{-8} hartree (keyword "$scfconv 8"). The results have been compared to the more accurate m5 grid and convergence to 10^{-9} hartree, which did not result in a change in energy. Therefore, we recommend the less accurate technique due to it being computationally faster while being very close to the more accurate alternative.

3.1.1 H_2O-H Interaction

Reference energy. For H_2O-H, we found $E_{CC}^{int} = -0.40$ kJ/mol, so we have an only weakly bonded system. The CP correction contributes $+0.05$ kJ/mol to the energy.

Basis set. For basis set comparison, we focus on Figure 4. Grouped from most attractive to least attractive interaction energies, the basis sets are def2-SVPD, def2-TZVP, def2-QZVP and def2-TZVPD. For most functionals, def2-TZVP is the most accurate basis set.

Figure 4 does not contain our results for the def2-SVP basis set since they are much worse than all other results – up to -20 kJ/mol and usually more than -5 kJ/mol of deviation. However, the addition of diffuse functions effects quite a remarkable improvement to the def2-SVP results, such that def2-SVPD calculations are well comparable to results obtained by basis sets of more than

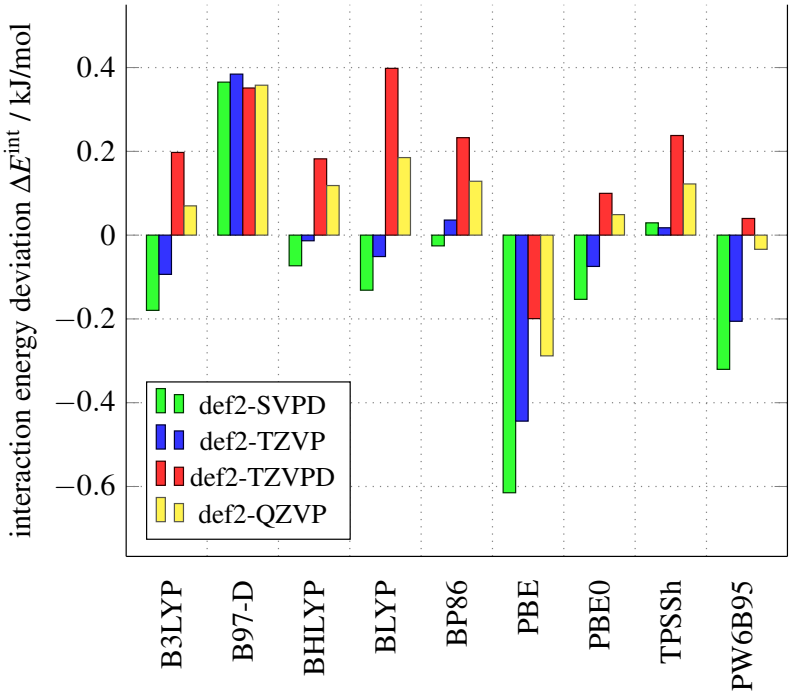


Figure 4 The H_2O-H benchmark results for different basis sets without dispersion correction. Energies are plotted as differences to the reference energy from CCSD(T)-F12a/VTZ-F12 calculations, see (3.1).

double-ζ order.

It is also noteworthy that def2-QZVP does not seem to offer a clear advantage over the other three basis sets in Figure 4. It tends to yield similar accuracy as def2-TZVPD, while usually being inferior to def2-TZVP for most functionals. def2-QZVP takes on average 1.44 times as long to complete an energy and gradient calculation as def2-TZVPD, which in turn takes on average only 1.02 times as long as an energy and gradient calculation with def2-TZVP.

Functionals and dispersion correction. Again, we focus on the data in Figure 4. Comparing functionals with the def2-TZVP basis set, the most accurate results are obtained by BHLYP, TPSSh and BP86, mediocre results are obtained by BLYP, PBE0 and B3LYP. The PW6B95 functional has its best result at the def2-TZVPD basis set.

The results here seem to be very accurate in absolute deviation, while relative deviations are often greater than 50%. Therefore, we can generally say that the description of H_2O-H seems to be tricky with the functionals and basis sets we used in this test. A few functionals need special remarks, but rather for their deficiencies. The first one are the TPSS and TPSS-D3 functionals. They predict much too attractive energies for this benchmark, beyond -7 kJ/mol and -9 kJ/mol, respectively. They predict a wrong optimum geometry, where the H radical is at a distance of around 2.20 Å from the O atom in the water molecule, which is much less than the reference value

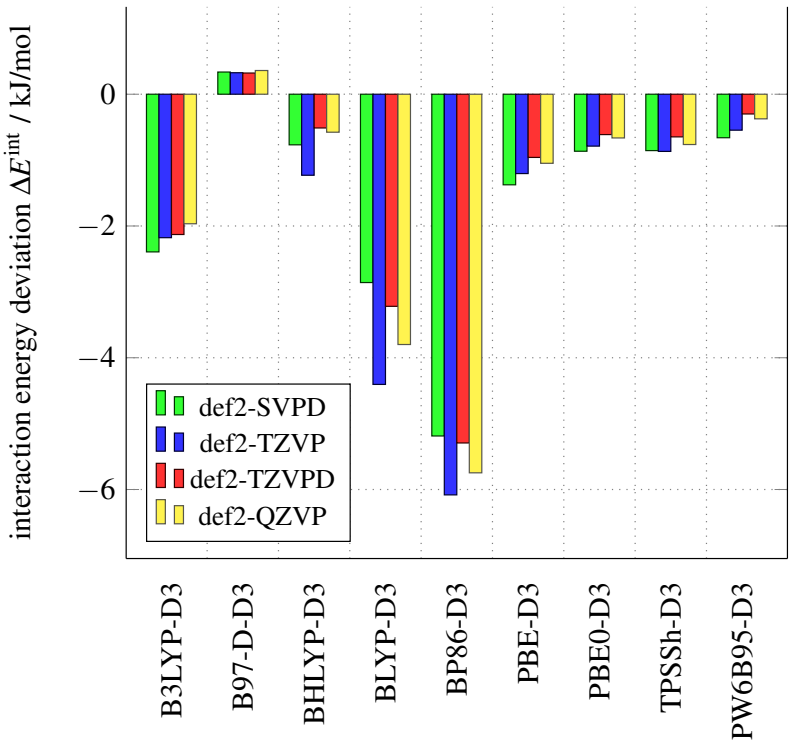


Figure 5 The H_2O-H benchmark results for different basis sets and functionals including dispersion correction. Energies are plotted as differences to the reference energy from CCSD(T)-F12a/VTZ-F12 calculations, see (3.1).

of 3.34 Å.

The opposite problem happens for B97-D and B97-D-D3 optimizations. Here, the attraction is considered close to being repulsive, with interaction energy values around $E^{\text{int}}_{\text{B97-D}} \approx -0.03$ kJ/mol. While the resulting deviation of around 0.37 kJ/mol does not look too dramatic in Figures 4 and especially Figure 5, it is a qualitative error. The separation between radical H and water O is around 5.80 Å, which is close to describing two isolated systems.

The inclusion of the D3 dispersion correction is not beneficial to the interaction energy values of H_2O-H, which can be seen by comparing Figure 4 to Figure 5. Indeed, while most functionals without dispersion corrections have errors of between -0.2 and $+0.2$ kJ/mol for at least one basis set, the D3 corrected versions deviate stronger than -5 kJ/mol. The contribution of dispersion in H_2O-H is therefore strongly overestimated by the D3 correction. The best results for dispersion corrected functionals are obtained by PW6B95-D3/def2-TZVPD with a deviation of -0.30 kJ/mol.

Optimum geometry. Comparing RMSD data of different functionals yields similar results as the energy comparison. Functionals without the D3 correction predict very accurate geometries with RMSD < 0.01 Å for all but TPSS and B97-D. Despite worse energy results, many D3 corrected versions still find accurate optimum geometries. Exceptions are B3LYP-D3, BLYP-D3, BP86-D3, which find minima similar to the already mentioned TPSS minimum, with the H radical too close to the water O.

Concerning the change of optimum geometry with respect to the choice of basis set, most basis sets lead to the same optimum geometry for nearly all functionals. Again, the def2-SVP basis set and the def2-TZVPP basis set often predict wrong, i.e. too attractive, geometries.

3.1.2 H_2O-H_2O Interaction

Reference energy. The reference energy for the H_2O-H_2O dimer is at $E^{\text{int}}_{\text{CC}} = -20.80$ kJ/mol, where the CP correction contributes $+0.18$ kJ/mol.

Basis sets. We use three promising functionals to investigate the difference between basis sets in Figure 6. Again, the def2-SVP basis set is not considered, since it yields errors between -12 kJ/mol and -15 kJ/mol for all three functionals. The figure shows that while increasing the cardinal number of the basis set improves the results, inclusion of additional diffuse functions seems to be even more beneficial. The basis set def2-SVPD is more accurate than def2-TZVP and def2-TZVPD is more accurate than def2-QZVP. This may well be due to reduction of the BSSE, which depends on the ability of basis sets to describe electrons far from the atomic nucleus. The inclusion of polarization functions as in def2-TZVPP also affects the results positively, however not to the extent of additional diffuse functions. The def2-TZVPPD basis set is again an improve-

ment to the def2-TZVPP basis set, it is also slightly more accurate than the def2-TZVPD basis set. Indeed, for all three functionals the def2-TZVPPD and def2-QZVPD results are strikingly close to one another, which could be an indication that the basis set truncation error is low and that the intrinsic quality of the methods is reached. In that case, the three functionals analyzed in Figure 6 are all capable of yielding excellent approximations to our coupled-cluster data, and the results at the def2-TZVPD basis set are already close to the methods' intrinstic error.

Concerning speedup, here def2-TZVPD takes on average 1.16 times longer than def2-TZVP for a single energy and gradient calculation, and def2-QZVP takes 2.18 times longer than def2-TZVPD. The results so far speak in favor of the def2-TZVPD basis set. Due to the size of the quantum mechanical part of the QM/MM system we will later consider, it is computationally too expensive to use def2-TZVPD for the whole QM part. We will therefore use a hybrid basis set with the more important atoms described by the def2-TZVPD basis set and the less important ones by the def2-TZVP basis set. We will describe this further subdivision of the QM domain in more detail in Section 3.3.1.

We should therefore determine functionals that are both good for def2-TZVP and def2-TZVPD, while their accuracy for def2-TZVPD is more important since the adsorption site will be closer to the def2-TZVPD atoms.

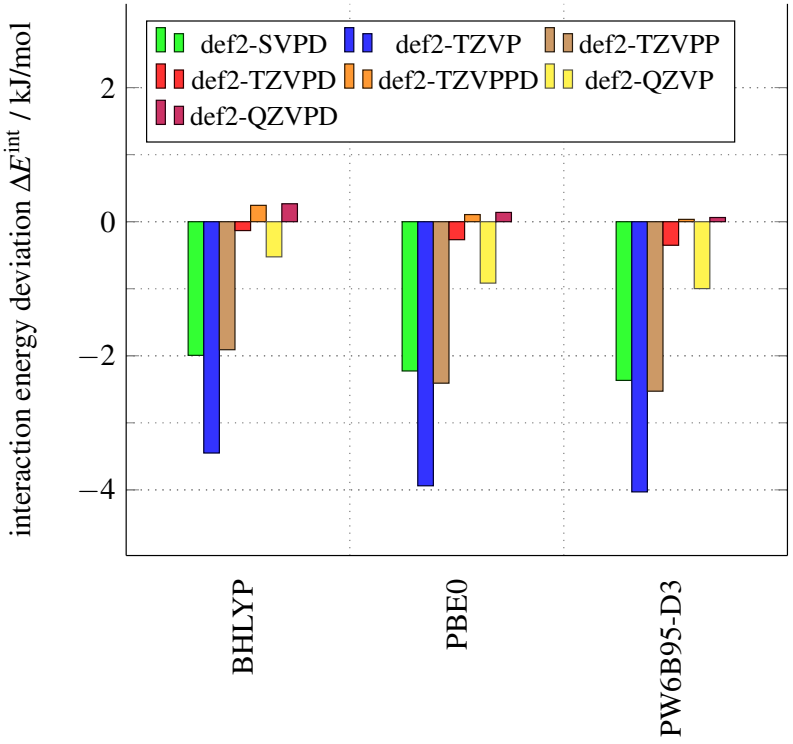


Figure 6 The H_2O-H_2O benchmark results for BHLYP, PBE0 and PW6B95-D3 and all basis sets used in the study (excluding def2-SVP). Energies are plotted as differences to the reference energy from CCSD(T)-F12a/VTZ-F12 calculations, see (3.1).

Functionals and dispersion correction. In Figure 7, we take a look at the differences between dispersion corrected and not dispersion corrected results in the def2-TZVP and def2-TZVPD basis sets. The stronger attraction between the two H_2O molecules prevents errors as for the B97-D functional in the previous section, so all functionals expect attractive interaction. Inclusion of the D3 dispersion correction again predicts stronger attraction, while the additional diffuse functions of the def2-TZVPD basis set weaken the attraction. These effects can combine to yield good results, for example for B97-D-D3 or PW6B95-D3.

The inaccuracy of some functionals may be due to the bad treatment of dispersion in DFT. But inclusion of the D3 correction does not resolve the issue for cases where the interaction is already too attractive. This is most striking for the def2-TZVP basis set, which predicts too attractive interaction for all functionals but B97-D. Generally, every functional has its most accurate results – either in the standard or the dispersion corrected version – at the def2-TZVPD basis set. Excellent results at the def2-TZVPD level are obtained by BHLYP, PBE0 and PW6B95-D3, but they are all not too accurate in the def2-TZVP basis set.

Good compromises between the def2-TZVP and def2-TZVPD basis sets are therefore given by B3LYP, TPSS, TPSSh and PW6B95, which all give mediocre results in both basis sets. BHLYP, PBE0 and PW6B95-D3 are highly accurate for def2-TZVPD but not very good with def2-TZVP. B97-D-D3 is somewhat between those functionals, being less accurate for def2-TZVPD but still

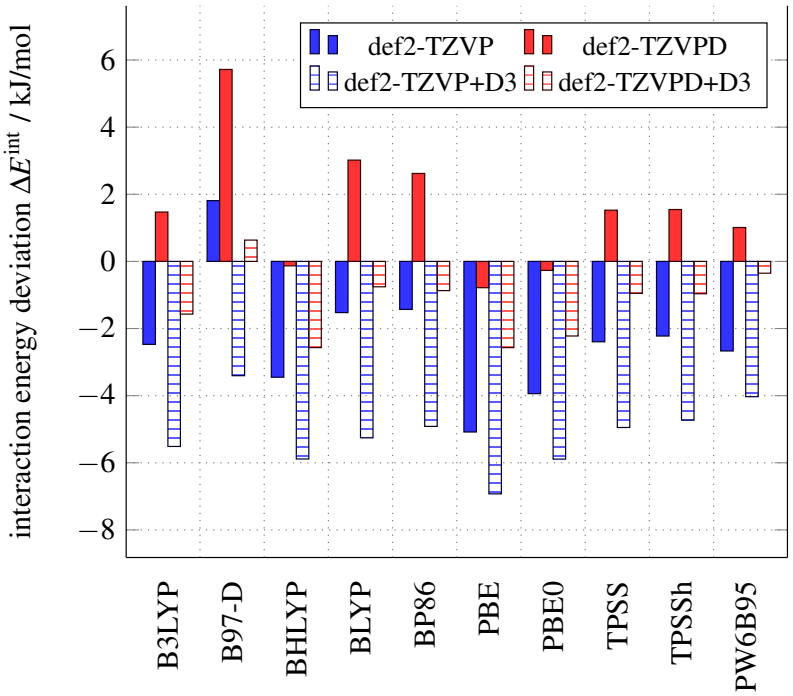


Figure 7 The $H_2O - H_2O$ benchmark results for the def2-TZVP and def2-TZVPD basis sets. The results with dashed bars include dispersion corrections. Energies are plotted as differences to the reference energy from CCSD(T)-F12a/VTZ-F12 calculations, see (3.1).

better for def2-TZVP than e.g. PBE0. All other functionals are not accurate enough for def2-TZVP or def2-TZVPD.

Optimum geometry. For H_2O-H_2O, the RMSD values are unproblematic for any basis set beyond def2-SVP (where an entirely wrong geometry is predicted). def2-TZVPP is not too exact, typically it has double the RMSD value of all other basis sets but def2-SVP. The worst RMSD is reached by PBE-D3/def2-TZVPP, it is of 0.063 Å, which is still just a slight displacement of H atoms.

Concerning functionals, especially B3LYP and PW6B95 both with and without dispersion corrections and BHLYP without the dispersion correction predict RMSDs of less than 0.008 Å to the reference geometry. The rest is between 0.02 and 0.05 Å. Therefore, H_2O-H_2O interaction geometries are well described by any functional in a basis set bigger than def2-SVP.

3.1.3 $H_2O-{}^3O$ Interaction

Reference energy. It already became clear at the beginning of this section that the $H_2O-{}^3O$ interaction is more tricky for post-HF methods because some geometries in the optimization process are not well described by a single Slater determinant. At the optimum geometry, a single-reference approach is fortunately possible, so our results for this benchmark have the full credibility of the

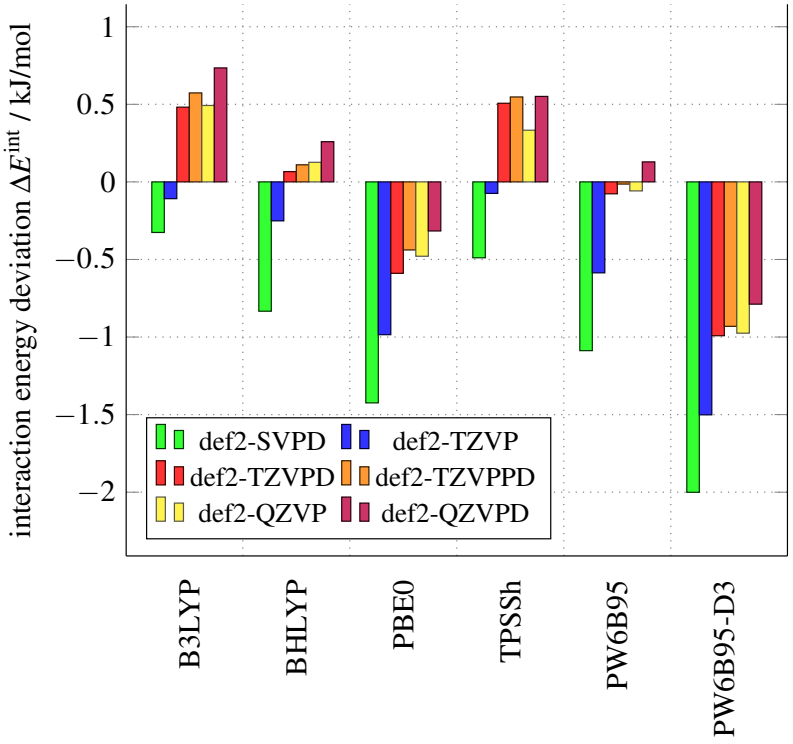


Figure 8 The $H_2O-{}^3O$ benchmark results for the def2-TZVP and def2-TZVPD basis sets. Energies are plotted as differences to the reference energy from CCSD(T)-F12a/VTZ-F12 calculations, see (3.1).

CCSD(T)-F12a method. Only the CP correction is taken from an MRCISD-F12 calculation, but it is as small as 0.10 kJ/mol and therefore the error here is unlikely to deteriorate our results. The CP corrected interaction energy is $E_{\mathrm{CC}}^{\mathrm{int}} = -6.72$ kJ/mol.

Basis sets. Consider Figure 8. In contrast to the previous two systems, the def2-SVPD basis set tends to yield the worst results. It always gives the most attractive energy. More reliable results seem to require triple-ζ basis sets. And again, bigger basis sets than def2-TZVPD are not necessarily more accurate. For B3LYP, BHLYP and TPSSh, the reference value is somewhere between def2-TZVP and def2-TZVPD. In the last two test systems, six functionals have shown a good overall performance.

Functionals and dispersion correction. While this was not so clear in the previous benchmarks, here the PW6B95 functional is superior to the PW6B95-D3 alternative. For basis sets bigger than def2-TZVP, PW6B95 is a strikingly accurate functional in this benchmark. However, the most accurate seems to be BHLYP, with good results even for the def2-TZVP basis set.

Optimum geometry. Of the functionals used in Figure 8, the optimum geometries are very close to the reference geometry. Only PBE0 has an RMSD of around 0.05 Å for def2-TZVPD and def2-TZVPP, which is still small. Apart from that, none has an RMSD greater than 0.01 Å for any basis set. This does not apply to all functionals in this benchmark, many of those excluded in Figure 8 have RMSD values greater than 0.5 Å.

3.1.4 Summary

We have decided on the def2-TZVP/def2-TZVPD hybrid basis. Both basis sets have proven to be able to yield good results, especially the def2-TZVPD basis is in many cases a good compromise between accuracy and computational cost. Since the adsorbate molecules can be placed close to water molecules described by def2-TZVPD, a functional's results for the def2-TZVPD basis set are the most important ones, while results in the def2-TZVP basis set should still be reasonable but not necessarily too accurate.

Under these conditions, we find that PBE0, BHLYP and PW6B95-D3 should be very good functionals for describing our system. Of these, especially BHLYP seems to have a great accuracy for both H_2O-H and $H_2O-{}^3O$. PBE0 is also quite accurate in both cases. For PW6B95-D3, the interaction with oxygen may become problematic.

The functionals B3LYP, TPSSh and PW6B95 are also good options for functionals. While the results for H_2O-H_2O interaction are not the most accurate ones in the def2-TZVPD basis set, their higher accuracy with def2-TZVP makes them good candidates. Additionally, PW6B95 was better in describing both H_2O-H and $H_2O-{}^3O$ interaction than its dispersion corrected counterpart.

For everything that follows, we will therefore only use results obtained with BHLYP, PBE0,

PW6B95-D3, B3LYP, TPSSh or PW6B95. Especially on adsorption under interstellar condition, there is not much experimental data available, therefore our only indication of good results may be the agreement between the different functionals that seemed to yield credible results by the standards of this benchmark.

3.2 The Gas Phase

Before we progress to the ice surface, we can use this section to get to know some key properties of the molecules we want to study as adsorbates on the ice surface. The gas-phase systems we wish to study are small enough to be admissible for CCSD(T)-F12a calculations. Therefore, the results presented in this section can be used to further evaluate the functionals we considered most reliable in the previous benchmark.

3.2.1 Energies of Formation

Absolute energy values for different functionals can not be compared properly. Instead, we need to compare energy differences to a reference value. We study only molecules with O and H atoms in them, so for a molecular species X with M_H H atoms and M_O O atoms, we define a standard energy of formation by

$$E_\mathrm{X}^0 = E_\mathrm{X} - M_\mathrm{H}\frac{E_{\mathrm{H}_2}}{2} - M_\mathrm{O}\frac{E_{^3\mathrm{O}_2}}{2}. \tag{3.3}$$

If two electronic structure methods agree well, they should obviously agree well with respect to E^0. We include ZPE corrections into E^0. For each functional, the correction is calculated at the corresponding optimum geometry for the functional without dispersion correction.

An energy defined by (3.3) has two problems. For once, we can not say how accurately the energies for H_2 and $^3O_2$ themselves are predicted. This leads to the second problem that inaccurate data for these will possibly make the entire set of results for a functional seem less accurate instead of only a few data points. We have to bear this in mind for the comparison.

Table 1 includes data for all functionals we decided upon in the benchmark section. Note that the formerly considered PW6B95-D3 functional is not included, since the dispersion correction only affects intermolecular interaction. Our calculations however treated isolated molecules, for which PW6B95 covers the results of both functionals.

To get estimates of the overall performance of the functionals, we give the *mean absolute deviation* (MAD) and minimum (MIN) and maximum (MAX) deviations as well as the overall mean deviation from the CCSD(T)-F12a calculation for each functional. The MAD agreement is best for PW6B95 and B3LYP. We can also see immediately that B3LYP and TPSSh always yield greater formation energies than the CCSD(T)-F12a calculations. Indeed, only BHLYP yields more

Table 1 Energies of formation for DFT functionals, according to (3.3). All values are ZPE corrected with the same method (including CCSD(T)-F12a). All energies in kJ/mol.

	B3LYP	BHLYP	PBE0	TPSSh	PW6B95	CCSD(T)-F12a
H	217.04	213.17	204.96	222.13	212.72	216.45
H_2O	−215.65	−224.54	−222.74	−197.51	−220.11	−238.28
H_2O_2	−98.26	−97.16	−104.82	−85.65	−101.84	−129.08
OH	42.89	16.85	43.01	51.30	46.06	36.49
HO_2	21.75	19.85	19.16	25.28	21.96	14.75
^1O	519.48	482.52	543.69	538.46	521.61	451.63
^3O	252.00	204.88	254.89	246.85	253.28	245.13
$^1O_2$	162.16	177.97	171.07	163.27	160.38	120.91
MAD	22.92	25.24	26.77	30.77	22.94	
MAX	67.86	57.05	92.06	86.84	69.98	
MIN	0.60	−40.26	−11.49	1.72	−3.73	
MEAN	22.92	9.44	23.90	30.77	22.01	

exothermic formation energies than the reference calculations.

As we can see in Table 1, the predictive power of the DFT functionals varies strongly among the molecular species. Especially the case of ^1O and $^1O_2$ seems to be problematic. We had to use spin-unrestricted Kohn–Sham DFT (UKS) to force the desired spin multiplicity. But the agreement is also not very good for H_2O_2, where spin-unrestricted Kohn–Sham DFT (RKS) could be used. The PBE0 functional is off by 92.06 kJ/mol for the case of ^1O, followed by a disagreement of 86.84 kJ/mol for TPSSh and disagreements of 69.96 and 67.86 kJ/mol for PW6B95 and B3LYP, respectively. The best agreement is achieved by BHLYP with 30.90 kJ/mol of difference. However, the BHLYP functional predicts E^0 very badly for $^1O_2$, with an error of 40.26 kJ/mol, while the other functionals differ by less than 10 kJ/mol, and it is also the only functional to differ by more than 50% from the reference energy of OH. Generally, all functionals predict too high values for E^0 and show strong deviations from the reference calculation for at least one molecular species. Since the values are not too far off for H_2 while showing considerable problems with molecules composed solely of oxygen, this could be an indication that the energetic description of $^3O_2$ is challenging to DFT functionals in general. Then again, the rather good agreement to the reference data for all functionals but BHLYP in the case of ^3O does not support this claim if we do not expect a cancellation of error.

Further comparison shows that for most molecules, the functionals closest to the CCSD(T)-F12a energies are B3LYP and PW6B95, which also yield similar results. This mainly means that they provide energy values that are not as high as those of other functionals.

Table 2 Reaction energies for DFT functionals. E^{ZPE} and E are reaction energies with and without ZPE correction, respectively. We also give the value of the ZPE correction ΔE^{ZPE}. CC-F12 is shorthand for CCSD(T)-F12a. The closest energy to CCSD(T)-F12a is highlighted in boldface. All energies in kJ/mol.

		B3LYP	BHLYP	PBE0	TPSSh	PW6B95	CC-F12
$OH + H \longrightarrow H_2O$	E^{ZPE}	−475.58	−454.56	−470.70	−470.94	**−478.89**	−491.23
	E	−509.17	−489.42	−504.76	−504.53	−512.99	−525.29
	ΔE^{ZPE}	33.58	34.87	34.05	33.59	34.09	34.06
$H_2O + {}^1O \longrightarrow H_2O_2$	E^{ZPE}	−402.09	**−355.15**	−425.77	−426.60	−403.33	−342.42
	E	−415.75	−370.59	−440.19	−439.92	−417.56	−355.80
	ΔE^{ZPE}	13.66	15.44	14.41	13.31	14.23	13.38
$OH + OH \longrightarrow H_2O_2$	E^{ZPE}	−184.03	−130.86	−190.84	−188.24	**−193.96**	−202.06
	E	−209.22	−158.07	−216.83	−213.19	−219.86	−227.13
	ΔE^{ZPE}	25.19	27.21	25.99	24.94	25.90	25.07
$HO_2 + H \longrightarrow H_2O_2$	E^{ZPE}	**−337.05**	−330.18	−328.94	−333.06	−336.51	−360.28
	E	−369.45	−364.20	−362.04	−365.28	−369.52	−392.71
	ΔE^{ZPE}	32.40	34.02	33.11	32.22	33.01	32.44
$OH + {}^3O \longrightarrow HO_2$	E^{ZPE}	−273.14	−201.88	−278.73	**−272.87**	−277.38	−266.88
	E	−287.98	−218.17	−294.10	−287.56	−292.69	−281.88
	ΔE^{ZPE}	14.84	16.29	15.36	14.69	15.31	15.00
$^3O_2 + H \longrightarrow HO_2$	E^{ZPE}	−195.30	−193.32	−185.80	**−196.85**	−190.76	−201.70
	E	−222.37	−221.89	−213.34	−223.80	−218.26	−229.48
	ΔE^{ZPE}	27.07	28.57	27.54	26.95	27.51	27.78
$E^{\mathrm{ZPE}}_{\mathrm{DFT}} - E^{\mathrm{ZPE}}_{\mathrm{CC\text{-}F12}}$	MAD	19.36	41.30	26.58	23.19	19.58	
	MIN	−59.67	−12.72	−83.35	−84.18	−60.91	
	MAX	23.22	71.20	31.34	27.22	23.76	
	MEAN	−1.27	37.66	−4.01	−4.28	−3.83	

3.2.2 Reactions

We already described gas-phase reactions in Section 2.1 by (2.4). For simplicity, we drop the (g) subscript.

We calculated DFT energies in the def2-TZVPD basis set with the B3LYP, BHLYP, PBE0, TPSSh and PW6B95 functionals. We also calculated CCSD(T)-F12a/VTZ-F12 energy data. The energies are given at the optimum geometry for each method. All results were supplemented by a ZPE correction.

The studied reactions are

$$OH + H \longrightarrow H_2O, \tag{3.4a}$$

$$H_2O + {}^1O \longrightarrow H_2O_2, \tag{3.4b}$$

$$OH + OH \longrightarrow H_2O_2, \tag{3.4c}$$

$$HO_2 + H \longrightarrow H_2O_2, \tag{3.4d}$$

$$OH + {}^3O \longrightarrow HO_2, \tag{3.4e}$$

$${}^3O_2 + H \longrightarrow HO_2. \tag{3.4f}$$

The ^{1}O in (3.4b) is in an excited state compared to the ground state ^{3}O, with an excitation energy difference of 206.50 kJ/mol according to the CCSD(T)-F12a results in Table 1. The reaction with H_2O to the more stable H_2O_2 will therefore be barrierless. All the other reactions are cases of radical recombination, which means that they can be predicted to be barrierless as well. For the same reasons, all reactions are also strongly exothermic.

Table 2 contains all the relevant information on reaction energies. We give ZPE corrected data and the energy values without ZPE correction as well as the value of the correction itself. Alongside the DFT results, we also give the CCSD(T)-F12a/VTZ-F12 results.

We give again the MAD, MIN, MAX and MEAN deviations. The most accurate functionals by MAD are again PW6B95 and B3LYP, with values of 20.70 and 21.29 kJ/mol, respectively, but even these have deviations of up to -60.91 and -59.67 kJ/mol, respectively.

The BHLYP functional shows an intersting behaviour. It usually predicts the reactions to be less exothermic than the other functionals do. For that reason, it usually agrees less with the reference energy. Only for the case of reaction (3.4b), this is beneficial. This reaction does involve the ^{1}O atom, for which BHLYP described the formation energy most accurately. The opposite is the case for reaction (3.4c), where the agreement between BHLYP and the reference data is much worse than the one of the other functionals. And again, for the formation energy BHLYP was not as accurate as the other functionals for OH.

The agreement between CCSD(T)-F12a and DFT functionals is never better than a difference of 5 kJ/mol. Still, for most cases we can find at least one functional within 10 kJ/mol of errror. When we later analyze these reactions on a surface, we can hope that a functional that provided good results in the gas phase will have a tendency to provide good results on the surface. However, there is no certainty, therefore this “most accurate functional by tendency” is only the best guess we can make. More thorough research could be done with an extended benchmark.

While the results here do not clearly identify specific functionals for which the agreement be-

tween CCSD(T)-F12a reaction energies and the functional's reaction energies should be particularly good, the data shows that it is in principle possible to describe the reactions (3.4) by DFT. Still, the results for CCSD(T)-F12a must be considered the most accurate and the DFT results are mere approximations. The approximation is of clearly lesser quality than for interaction energies.

3.3 Adsorption and Reactions on the Ice Surface

This section is all concerned with molecular adsorption on ice surfaces and the reaction processes of adsorbed species. We will first give an account of what kind of surface model we use and how well the ideal crystalline description is maintained by the QM/MM description of the surface. Within QM/MM, the computation of the ZPE correction will have to be discussed. Then, we will give ZPE corrected adsorption energies for different functionals and molecular species. We only consider neutral molecules. After that, based on the adsorption data, reaction energies are given and compared among functionals as well as to gas-phase reaction energies and experimental data.

3.3.1 The Surface Model

3.3.1.1 The Fletcher Surface

Barnes established a first description of water at temperatures as low as 90 K in 1929.[76] He could not locate the hydrogen atoms due to their low X-ray scattering power,[77] but he was able to experimentally verify that the oxygen atoms are aligned hexagonally on layers. The hydrogen atoms were then mostly described by statistical distributions. While at high temperatures the exact location of the hydrogen atoms will always be unordered due to thermal fluctuations, there are more regular phases of crystalline water at low temperatures. We follow an approach presented by Fletcher that minimizes surface free energy of the (0001) surface of hexagonal ice I_h at low temperatures.[78]

It is best understood by consindering Figure 9. Starting from a layer of hexagonally arranged O atoms (which are not arranged on a plane but alteratingly above and below the plane), one can connect each O atom to its four nearest neighbors. On these designated hydrogen bonds, H atoms can be placed close to one of the two O atoms taking part in the bond. This leads to a multitude of possible arrangements. Fletcher's approach is to draw parallel lines into the lattice to group the molecules in rows. Molecules in the same row have the same orientation. In Figure 9, these rows can be defined by vertical lines.

This ordering comes from first considering only the surface molecules. These have a broken hydrogen bond outward, which may or may not have a hydrogen atom on it. When grouping those broken bonds with hydrogen atoms and those without hydrogen atoms in alternating rows, the

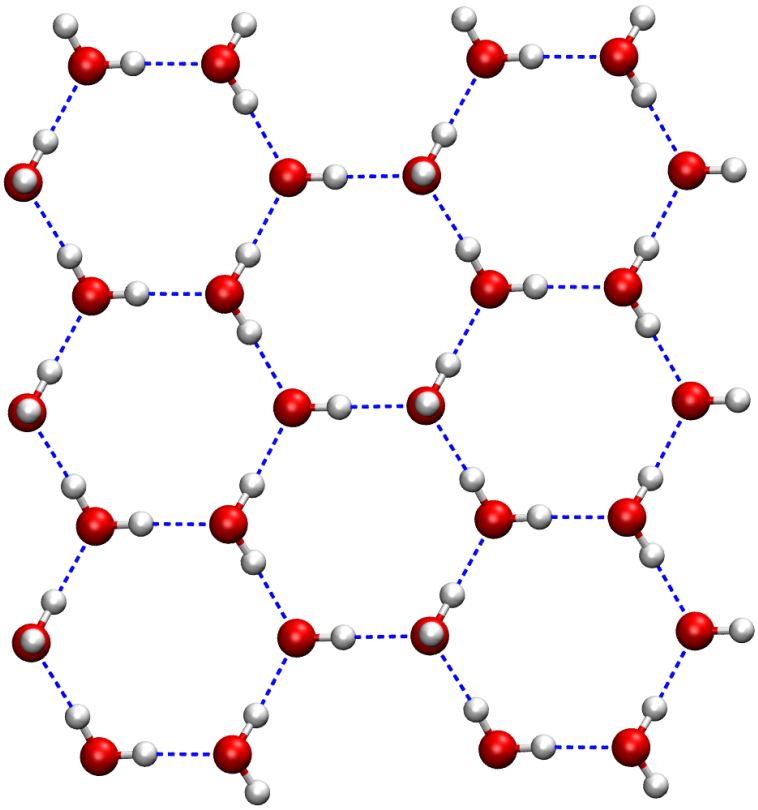


Figure 9 The (0001) basal Fletcher surface. Hexagonal O (red) structure with vertical lines of identical H (white) orientation. H-bonds are indicated by dashed blue lines. The O atoms are not all in the same plane.

structure described above is extended to the full system.

Such a crystalline surface must be considered as an idealization. First of all, since we are interested in an ice surface on a grain, the interaction between the water molecules and the grain at the grain surface will surely lead to changes in the water crystal, probably both in O position as well as in H orientation. Research in that respect was presented by Cabrera Sanfelix *et al.*[20] on a graphite surface, where a water dimer adsorbed on the surface has different bond angles than the gaseous water dimer. Therefore, an undisturbed crystalline surface will only be reasonable if the water ice is several monolayers thick, which is unlikely since one only encounters "dirty" ices in the interstellar medium.[10] A second idealization is that the crystalline structure of water ice can only be true to the perfect hexagonal structure if a layer like in Figure 9 is within the crystal, i.e. there are more such layers above and below. Then, there is a balance of force. If the layers above are removed to describe a surface layer, the forces from inside the crystal will surely deform the ideal surface layer of Figure 9. We will see this effect in Section 3.3.2.

Instead of choosing a cuboid subdomain of the Fletcher surface – which would be more adequate

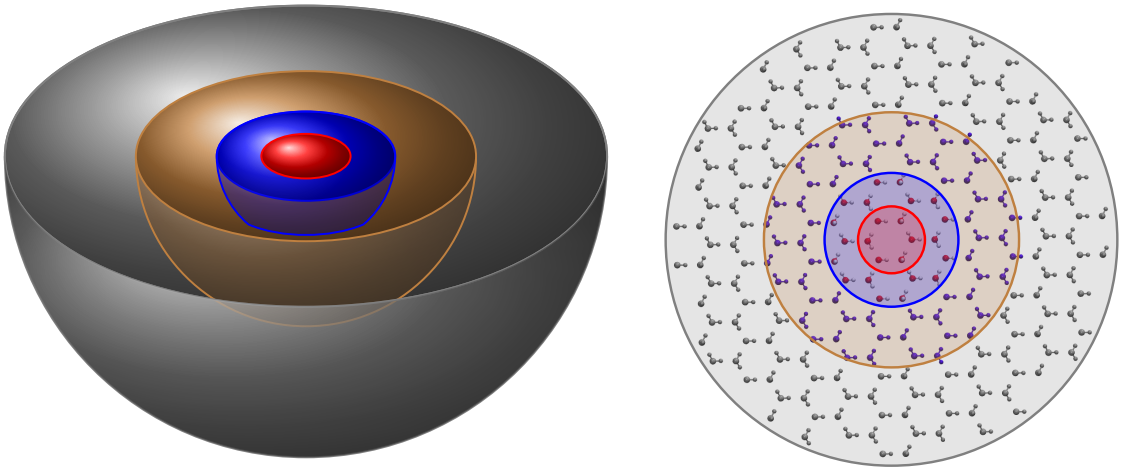


Figure 10 The QM/MM decomposition of the Fletcher surface. The picture to the right shows the surface along its normal, the picture to the left is a schematic representation. The gray and the brown area make up the MM region, where water molecules are described by the TIP3P potential. In the gray region, they are frozen. The blue and red region make up the QM region. Atoms in the blue region are described by the def2-TZVP basis set. The central ring constituting the red region uses the def2-TZVPD basis set.

to the surface's translational symmetry – we will consider a hemisphere as depicted in Figure 10. We do so because the surface model shall be used in future studies to investigate reactions. These depend on energy dissipation processes throughout the surface, which will propagate spherically from a reaction site. The symmetry of the model is chosen to be adequate to that, with the designated reaction site at the central ring.

3.3.1.2 The QM/MM Region

We start from an infinite fully-ordered water I_h crystal with a Fletcher surface. We choose a normal basal (0001) surface as in Figure 9, not a prismatic surface. We define a central ring, that is one of the hexagonal water rings of the surface. We take a hemispherical cut of radius 25 Å centered at the center of mass of the central ring. The hemispherical cut does not separate H_2O molecules internally, instead an H_2O molecule is fully incorporated in the system if its O atom is within the cutoff radius. The full number of QM/MM water molecules in the system is 1151, with 1151 O atoms and 2302 H atoms.

The decomposition into QM/MM domains is described in Figure 10. At a sphere of radius 8 Å we separate the system into two domains. The domain within this sphere is the QM domain and the domain outside of it is the MM domain. The QM region then contains three layers of molecules. On the surface layer, there is the central hexagonal ring of water molecules and the six rings adjacent to it. The second layer has a water ring and the six molecules forming hydrogen bonds with the O atoms of the ring. On the third layer, there is only one ring left. This would total to

24+12+6 = 42 QM molecules. However, the QM description is only necessary for regions where chemistry takes place. In our study, this will all happen close to the surface layer. Therefore, going three layers deep with the QM domain is not necessary and the the six atoms of the lowest ring are also treated by MM. This makes the system slightly less symmetric, but for the problems discussed here, the decrease in computation time is enough to make up for that: calculations with 42 QM molecules take on average 1.4 times as long as calculations with 36 QM molecules. Therefore, we decide on a total of 36 QM molecules. Of these, only the central ring at the surface is described by the def2-TZVPD basis set, the rest uses the def2-TZVP basis set.

As for the MM domain, it is again separated at 15 Å of the system center. Those molecules within that radius can change their position (active molecules) in the optimizations we later carry out. The rest of the MM molecules is fixed in position (frozen molecules) to yield boundary conditions. The 10 Å thick envelope of frozen molecules may not be necessary as a boundary, but on the other hand it has virtually no effect on computation time to include these additional MM atoms, and if necessary they can be turned into further active atoms. There is a total of 1115 MM molecules, of which 225 are active MM molecules and 890 frozen MM molecules.

The geometry of the original Fletcher surface has to be changed to fit the requirements of the TIP3P potential. In the original formulation the hydrogen atoms are placed directly on the hydrogen bonds, that is connecting lines between each oxygen atom and its four nearest neighbors. This process yields bond angles of 109.50° with an in principle variable bond length. The latter can be easily adjusted to the TIP3P requirements by simple stretching. This does not change the bond angles which are still in disagreement with the bond angles of the TIP3P potential fixed at 104.52°. This is mended by simply taking a perfect Fletcher surface with adjusted bond lenghts $|\mathbf{r}_{\mathrm{O-H1}}| = |\mathbf{r}_{\mathrm{O-H2}}| = 0.9572$ Å and changing the angles to the TIP3P value while maintaining the orientation of the water dipole moment, that is the unit vector corresponding to $\mathbf{r}_{\mathrm{O-H1}} + \mathbf{r}_{\mathrm{O-H2}}$ stays the same for each molecule. This means symmetrical bending of all bond angles with respect to the dipole moment.

We must also remark on the MM force field used for adsorbates. Since our QM region is rather big, we assume that the MM parameters (only the LJ parameters matter to the coupling) for the adsorbates are of minor importance if they stay above the central ring, but they should still be chosen sensibly. We studied only adsorbates composed of H and O. We decided on an isolated description of the atoms in the adsorbate, with the TIP3P LJ potential according to atomic species. Since the interaction of atoms within the adsorbate molecule is described by QM alone and is independent of the coupling, this approach seems sensible. The water LJ parameters should still be reasonable for the isolated atoms.

3.3.1.3 Setting Up the Model

Now we talked about how to define the surface geometry and how to decompose the system into QM and MM domains. To give guidance for a general approach to set up a surface model, we want to repeat the steps we took.

1. **Geometry.** Set up the lattice for oxygen atoms with hexagonal rings. Include the hydrogen atoms to arrive at the perfectly ordered Fletcher surface. If desired, this ordered surface can lose some order by changing the hydrogen orientation. It may be even further disordered by a short molecular dynamics calculation applying a classical force field and a finite temperature, which should lead to some form of ASW. When the surface satisfies the requirements, take a hemispherical cut. You can then add adsorbates or perform other manipulations to adjust the system to your requirements.

2. **Subdomains.** Select QM atoms and MM molecules. Divide the MM molecules into frozen and active molecules.

3. **Force field.** Decide on a force field for the MM region. We used TIP3P and believe that the details of the force field have only a minor impact on the result of the calculation due to the large QM region, but for smaller QM regions, its effect may be more considerable. Note that force field parameters are required also for the adsorbates.

4. **DFT.** Decide on the details of your DFT calculations. Our benchmark favors def2-TZVPD, therefore we recommend a def2-TZVP/def2-TZVPD basis set and the functionals discussed at the end of Section 3.1.

5. **QM/MM.** Set up the communication between QM and MM region. Note that you may require MM parameters for your QM atoms and that on the other hand your QM calculations should support point charges.

Following this approach was very easy for us, since we used the ChemShell interface. It supports a QM/MM interface where the QM and MM theories can be defined independently of each other. They are then coupled to yield a hybrid theory, which is provided to DL-FIND alongside the initial geometry.

3.3.2 First Geometry Optimizations

To verify the agreement between the theoretically assumed Fletcher surface and the methods we want to use, we calculated RMSD values (3.2) between the ideal Fletcher surface and the surface

Table 3 RMSD values for the pure force field calculation with TIP3P and QM/MM calculations with different functionals. Deviations to perfect Fletcher surface, TIP3P optimized surface and B3LYP QM/MM optimized surface, respectively.

	to Fletcher	to TIP3P	to B3LYP
TIP3P	0.179 Å	0.000 Å	0.101 Å
B3LYP	0.183 Å	0.101 Å	0.000 Å
BHLYP	0.182 Å	0.096 Å	0.010 Å
PBE0	0.193 Å	0.103 Å	0.031 Å
TPSSh	0.191 Å	0.103 Å	0.025 Å
PW6B95	0.183 Å	0.102 Å	0.013 Å
PW6B95-D3	0.187 Å	0.101 Å	0.015 Å

after QM/MM optimization. Note that the factor $\frac{1}{M}$ in (3.2) is not defined by the full number of QM/MM atoms in the system but by the number of active QM/MM atoms.

Data for the RMSD values are presented in Table 3. We include the pure MM optimization with TIP3P. Generally, the results can be considered to be good. For all systems, the same effect is visible: the active H_2O molecules sink slightly into the surface. We did already have reasons to expect that, as explained in Section 3.3.1. This initial sinking in process leads to an RMSD of less than 0.2 Å, and the displacement is stronger the further a molecule is away from the frozen molecule region. Below the surface layer, the displacements due to this effect are not very strong. The RMSD among the DFT optima is very small, they deviate from the B3LYP optimum geometry by an RMSD of 0.031 Å or less. This is small compared to the $O-H$ separation of 0.98 Å and the $O-O$ separation of around 2.70 Å. Even those functionals we excluded in the benchmark yield optimum geometries within boundaries of an RMSD of 0.063 Å to B3LYP. Between the functionals of Table 3 and the TIP3P potential, there is an RMSD of around 0.1 Å. The source of this deviation is that the DFT water does not have the $\theta(\mathrm{HOH})$ bond angle restriction that TIP3P water has to satisfy.

We can therefore say that the QM/MM descriptions we decided on all find an optimum geometry that is close enough to the ideal Fletcher geometry.

We can also compare the behaviour of the QM and MM part during the optimization process starting at the initial crystalline geometry. Figure 11 gives a typical example with the BHLYP functional. The QM, MM and QM/MM energies are given as differences to their values at the energy minimum. At the perfectly crystalline system, the energy difference is too high for the plot axis, it goes up to 226 kJ/mol for the QM part and 43 kJ/mol for the MM part. After the first few iterations, one can see that the energy changes in both domains are of the same order of magnitude. The overall energy seems to converge more or less smoothly. There is a codependence visible where a more attractive interaction in the MM region is accompanied by a less attractive interaction in the QM region and vice versa. We take this as an implication that the interaction

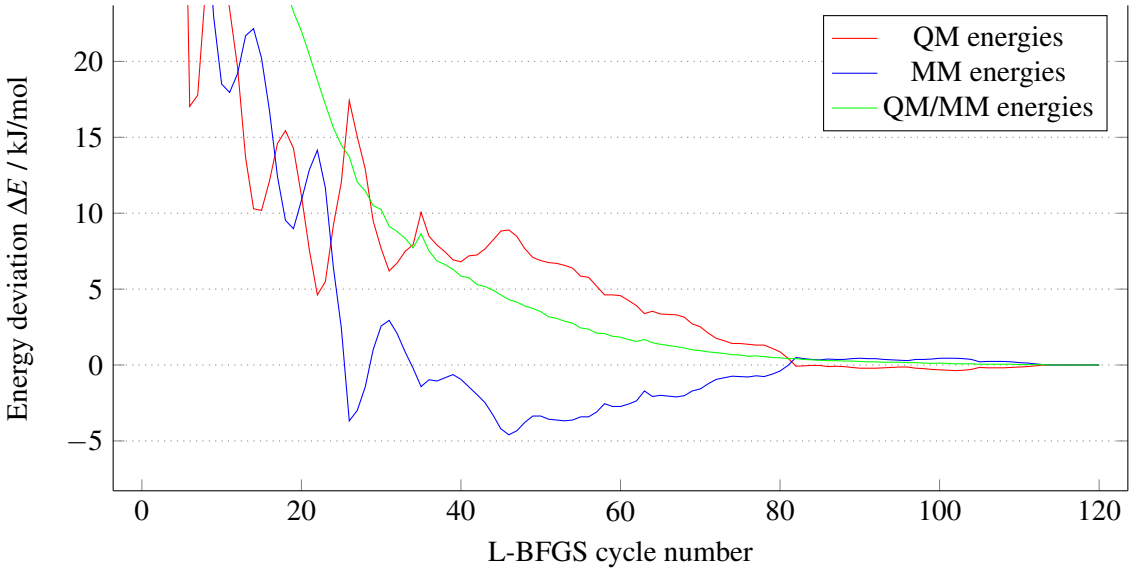


Figure 11 The QM/MM energy changes during a typical optimization. The energies are deviations from their value at the optimum geometry. Functional: BHLYP.

between the two region works well and that neither one of them is outbalanced by too strong energy gradients in the other one. We also see this as an indication that active MM atoms should be included in a geometry optimization to allow proper deformations of the QM region.

3.3.3 ZPE Corrections

Now we have established a model that should be able to describe ^{3}O, H and H_2O adsorption to some satisfaction. We can compute optimum geometries and their energies for different adsorbates and obtain the adsorption energy according to (2.3). Before we do so, it is necessary to give a few words on the ZPE correction in this case.

With the QM/MM coupling, we can still compute analytical gradients,but analytical Hessians are not available. There is no way around computing finite differences, so two atomic diplacements per coordinate per atom, leading to six displacements per atom. Computations for more complicated systems, especially s-HO_2 calculations with their doublet character, require a sensible choice of atoms to be displaced. On the other hand of course, enough atoms should be displaced such that the finite difference maintains good results.

The MM atoms do not need to be displaced. As for the QM region, we come back to the argument that the atoms relevant to adsorption are the central ring of six H_2O molecules plus adsorbate, that is those atoms that we describe with the def2-TZVPD basis set. The idea here is that when an adsorbate comes into the system and is adsorbed within the central ring, then its impact on the vibrational ground state of the system will mostly affect the atoms within the ring. The effect on atoms further away should be considerably smaller. However, the question if the cutoff for this

Table 4 Different choices for the ZPE correction. For a reference, the corresponding adsorption energies are given for the B3LYP functional. All energies in kJ/mol.

	E_{B3LYP}	$\Delta E^{\mathrm{all}}_{\mathrm{B3LYP}}$	$\Delta E^{\mathrm{ring}}_{\mathrm{B3LYP}}$	$\Delta E^{\mathrm{ring}}_{\mathrm{PBE0}}$
s-H	−1.71	5.03	3.28	2.84
s-H_2	−2.65	8.42	6.80	6.81
s-H_2O	−48.01	15.98	14.84	14.45
s-H_2O_2	−41.76	9.28	10.09	9.51
s-OH	−44.99	13.60	12.11	12.05
s-HO_2	−65.48	13.17	11.66	11.59
s-^3O	−16.17	2.69	3.35	3.04
s-$^1O_2$	−8.41	2.46	3.86	4.96
s-$^3O_2$	−2.62	3.06	3.90	3.86

effect should only include the central ring must be further investigated. The upshot is obvious: With M being the number of atoms in the adsorbate, these $18+M$ atoms need $108+6M$ displacements (plus one initial geometry). This is much less than the original $648+6M$ displacements for $108+M$ QM atoms.

We will call this selective correction the ZPE/ring correction in contrast to the ZPE/all correction for all QM atoms. Corresponding energies are then called E^{ring} and E^{all}, respectively, and the corrections are ΔE^{all} and ΔE^{ring}.

We can compare the two corrections in the next section.

3.3.4 Adsorption

Of the molecules relevant in Section 3.2, we examine adsorption energies for all species but ^1O, since it is chemisorbed by the surface to form some conformation of H_2O_2. This is a result of the barrierless and very exothermic reaction $H_2O + {}^1O \longrightarrow H_2O_2$, as described in (3.4b).

Before discussing adsorption data, we want to discuss the relevance of the ZPE correction and the differences between ZPE/all and ZPE/ring as well as the differences between the ZPE/ring correction for B3LYP and PBE0, which we hope to represent the variability of the ZPE correction for different functionals.

In general, the value of the correction term varies from adsorbate to adsorbate. All three corrections lie roughly between 2 and 16 kJ/mol. That makes them – if there is no cancellation of error of some other sort – necessary for the correct description of adsorption processes, since the uncorrected adsorption energies themselves lie between −1.71 and −65.48 kJ/mol. Usually, the ZPE correction is between 15% and 30% of the uncorrected energy. The exceptions are s-H, s-H_2 and s-^3O, which we will discuss later. The value of the ZPE correction is also typically greater than the variation of the uncorrected adsorption energy of different functionals, as we can see when

comparing to Table 5.

The ZPE/all and ZPE/ring energies do not deviate strongly in absolute value, all of them less than 2 kJ/mol and all but s-H and s-HO by less than 1.5 kJ/mol. All corrections are positive. It seems safe to say that the ZPE/ring correction may not be extremely close to the ZPE/all correction, but it shows the same qualitative behaviour. It seems alright to use the correction of the smaller system with an error to the correction for the full system of ± 2 kJ/mol.

Comparing the correction terms for B3LYP and PBE0, we generally find good agreement. With the exception of s-$^1O_2$, the energies differ by less than 0.6 kJ/mol. Given the uncertainty between ZPE/ring and ZPE/all, the differences between the PBE0 correction and the B3LYP correction are negligible. It seems reasonable to generalize what we already discovered for the gas phase result of Section 3.2: the ZPE correction does not change greatly between different functionals. We will therefore use the ZPE/all correction for the B3LYP functional for all following data.

Talking about the differences between ZPE corrected adsorption energies and those without correction, there are the interesting cases of s-H, s-H_2 and s-$^3O_2$. In all three cases, the uncorrected E predicts weakly attractive interaction, while the corrected E^{all} and E^{ring} both predict repulsive interaction. That means that the binding site found by the uncorrected DFT method is probably not the desired optimum. In such a case, it would of course be interesting whether there is an optimum geometry for the ZPE corrected functional that is still attractive. One should assume so, since it would seem unphysical if the surface and one of the molecules would repel one another. The problem here may be more the B3LYP functional than the ZPE correction alone.

Still, the ZPE correction in such a weakly adsorbed case is probably erroneous since the minimum is probably very shallow. This means that approximating the PES by a harmonic potential does not reproduce the PES correctly.

This result implies that the adsorbates s-H, s-H_2 and s-$^3O_2$ are only very weakly bound to the surface in this 0 K limit.

Let us now turn to Table 5, where ZPE corrected adsorption energies are given. Where there is experimental data available, it is given in the last column. So far, not many studies exist on measuring adsorption under interstellar conditions. We therefore also include some data for OH, ^3O and $^1O_2$ adsorption on bare amorphous silicates, which were gathered in order to gain a better understanding of water ice formation on bare silicates.[80,82] Other data include H_2O adsorption on an Au surface covered in crystalline water ice[79] and $^3O_2$ adsorption on porous amorphous water ice.[81] We have an estimation of the difference between adsorption on water ice and on the bare grain by comparing the experimental data for s-^3O, where the difference is below 3 kJ/mol including the error bars. While this does not have to apply to all adsorbates, a difference of similar magnitude is at least plausible, such that the experimental values for adsorption energies on bare

Table 5 Comparison of adsorption energies for different density functionals. All energies are given with ZPE/all correction for B3LYP. PW is shorthand for PW6B95. Energies can be compared to the average $E_{\mathrm{avg}}^{\mathrm{ads}}$ and experimental data. MAD, MIN, MAX and MEAN are give with respect to $E_{\mathrm{avg}}^{\mathrm{ads}}$. Energies in kJ/mol.

	B3LYP	BHLYP	PBE0	TPSSh	PW	PW-D3	$E_{\mathrm{avg}}^{\mathrm{ads}}$	experimental, E_{ads}
s-H	3.32	3.77	3.37	−5.05	3.29	1.27	1.66	
s-H_2	5.77	5.11	3.57	4.55	4.65	2.47	4.35	
s-H_2O	−32.03	−35.07	−36.90	−30.92	−33.51	−41.17	−34.93	-48.00 ± 0.50^{a}
s-H_2O_2	−32.47	−35.99	−37.64	−31.04	−34.85	−44.33	−36.05	
s-OH	−31.39	−32.88	−35.69	−31.10	−32.78	−39.66	−33.92	-13.77 to -39.58^{b}
s-HO_2	−52.30	−49.10	−56.97	−54.10	−51.08	−57.80	−53.56	
s-^3O	−13.48	−5.16	−11.25	−13.22	−12.85	−18.13	−12.35	-13.80 ± 0.50^{c}
								-15.38 ± 0.75^{c}
s-$^1O_2$	−5.95	−7.51	−11.44	−9.44	−7.44	−12.53	−9.05	
s-$^3O_2$	0.45	−0.65	−1.52	−0.52	26.68	7.04	$-0.56^{\dagger}$	-7.52^{d}
MAD	2.06	1.93	1.74	2.29	4.17	4.85		
MIN	−1.13	−0.14	−3.42	−6.72	−0.50	−8.28		
MAX	3.58	7.19	1.71	5.02	27.23	7.60		
MEAN	1.81	1.88	−1.12	0.40	4.06	-3.16		

[a] Adsorption on Au grain covered in crystalline water. For other grain materials between 42.15 and 49.79 kJ/mol. Fraser *et al.* 2001.[79]
[b] Adsorption on an amorphous silicate. He, Vidali 2014.[80]
[c] Adsorption on porous amorphous water ice and amorphous silicate, respectively. He *et al.* 2015.[81]
[d] Adsorption on an amorphous silicate. He, Jing, Vidali 2014.[82]
[†] Average excluding both PW and PW-D3.

grains should be a rough estimate to the adsorption energies on water surfaces on grains.

Table 5 also contains a column with average values for each adsorbate. For s-$^3O_2$, PW6B95 and PW6B95-D3 are excluded from the average because of the two very unphysical values. We will use this average energy when discussing reactions.

When there is no experimental data available, we can only compare the functionals to one another. For both s-H and s-H_2, the problem of repulsive interaction due to ZPE correction seems to affect most functionals. TPSSh disagrees with the others by still maintaining a negative sign for s-H, but for s-H_2, all functionals agree. This is at least an indication that the physisorption of H and H_2 is very fragile and that rather oxygen and hydrogenated oxygen species have longer residence times than hydrogen. For $^3O_2$, the case is a little puzzling. PW6B95 predicts a strong repulsion between surface and $^3O_2$, which is supported by PW6B95-D3. This looks like an error in some computation script, however we were unable to find such an error. We must therefore consider PW6B95 to be unreliable for $^3O_2$ adsorption. The other functionals predict very weak attraction, only B3LYP predicts a very weak repulsion. We thus infer that $^3O_2$ is probably only weakly bound to the surface, too, and must have a short residence time.

While each functional does have its "slips" for some adsorbates, the first four functionals of Table

5 seem to be in good accord. For many cases, one of the two PW6B95 functionals is furthest from an average value of all six functionals. This is especially true for O and O_2. Since it is either PW6B95 or PW6B95-D3, this could be an indication that neither the PW6B95 nor the D3 dispersion correction are universally reliable for the study at hand.

Comparing the results to experimental findings, the agreement is more qualitative than quantitative. As already mentioned, the systems described by experiment differ from the idealized system we study. Most noteworthy is that the data for s-H_2O does not agree very well with experiment although it was among our benchmark tests. A reason for that (and other disagreements) may very well be that in our system, the H_2O molecule is placed above the surface, while the experimental source of Fraser used thermal-programmed desorption (TPD).[79] In TPD, the H_2O atoms are removed from the crystal water itself. Water in the crystal must be tighter bound than add-on water as we describe, which explains the energy difference.

It is often the case that the experimental values mostly lie somewhere between the ZPE corrected energy value and the uncorrected one for many functionals. There are a number of potential sources of error. One could be the already discussed problem of the ZPE correction to describe shallow minima. Still, the ZPE correction represents a physical necessity and its inclusion can not be questioned generally. Another source of error is the system's strongly idealized geometry, which is not only due to the QM description but also to the water geometry of the MM description, which would make it difficult to allow for systems with distorted bond angles as described by Cabrera Sanfelix *et al.* for the case of adsorption of a water dimer on graphite.[20] We will come back to possible improvements to the model in Section 4.

On the other hand, the broad agreement between B3LYP, BHLYP, PBE0 and TPSSh allows for the assumption that the energies calculated for the model at hand may be accurate to within a few kJ/mol. Since we chose the functionals because they were best able to describe interaction energies in Section 3.1, this hope becomes more plausible. However, due to their rather bad performance for O species, the PW6B95 functionals have lost some credibility.

3.3.5 Reactions

With the adsorption energy data and corresponding gas-phase energies at the def2-TZVPD level, we can calculate reaction energies. We will only consider reactions of the Eley–Rideal mechanism in which a molecule from the gas phase approaches an adsorbed molecule to form a new molecular species, as described in the introduction. The product will for our study always be a single adsorbed molecule. With our data, one could also investigate cases where two product molecules remain of which none, one or both remain adsorbed on the surface.

Table 6 Surface reaction energies with Eley–Rideal type reactions. All energies are ZPE corrected with $\Delta E^{\mathrm{ZPE/all}}_{\mathrm{B3LYP}}$. PW is short for PW6B95. The functionals most accurate for the gas-phase reactions are highlighted by boldface again (cf. Table 2). The last column contains the hybrid reaction energy defined in (3.6). The deviations of these are listed in the last four rows for each functional.

		B3LYP	BHLYP	PBE0	TPSSh	PW	PW-D3	$E^{\mathrm{react}}_{\mathrm{hybrid}}$
s-OH $+H_{(g)}$	$\longrightarrow$ s-H_2O	−476.22	−458.03	−472.39	−470.76	**−480.13**	**−480.92**	−492.71
s-H_2O $+{}^1O_{(g)}$	$\longrightarrow$ s-H_2O_2	−402.54	**−357.85**	−427.27	−426.38	−405.24	−407.50	−344.17
s-OH $+OH_{(g)}$	$\longrightarrow$ s-H_2O_2	−185.12	−137.04	−194.02	−187.84	**−197.12**	**−200.16**	−205.33
s-HO_2 $+H_{(g)}$	$\longrightarrow$ s-H_2O_2	**−317.22**	−318.69	−310.31	−309.82	−320.89	−323.99	−343.39
s-OH $+{}^3O_{(g)}$	$\longrightarrow$ s-HO_2	−294.05	−219.54	−300.54	**−295.71**	−296.15	−296.10	−287.00
s-^3O $+OH_{(g)}$	$\longrightarrow$ s-HO_2	−311.96	−248.31	−325.41	**−313.50**	−316.45	−318.00	−308.95
s-$^3O_2$ $+H_{(g)}$	$\longrightarrow$ s-HO_2	−248.05	−243.27	−241.72	**−250.31**	−268.95	−256.14	−254.74
$E_{\mathrm{DFT}} - E^{\mathrm{react}}_{\mathrm{hybrid}}$	MAD	19.71	40.13	27.26	24.70	19.32	17.04	
	MIN	−58.37	−13.68	−83.10	−82.21	−61.07	−63.33	
	MAX	26.17	68.29	33.07	33.57	22.49	19.39	
	MEAN	0.16	36.22	−5.06	−2.58	−6.95	−6.65	

We use ER versions of the gas-phase reactions (3.4), namely

$$\text{s-OH} + \mathrm{H_{(g)}} \longrightarrow \text{s-}\mathrm{H_2O}, \tag{3.5a}$$

$$\text{s-}\mathrm{H_2O} + {}^1\mathrm{O_{(g)}} \longrightarrow \text{s-}\mathrm{H_2O_2}, \tag{3.5b}$$

$$\text{s-OH} + \mathrm{OH_{(g)}} \longrightarrow \text{s-}\mathrm{H_2O_2}, \tag{3.5c}$$

$$\text{s-}\mathrm{HO_2} + \mathrm{H_{(g)}} \longrightarrow \text{s-}\mathrm{H_2O_2}, \tag{3.5d}$$

$$\text{s-OH} + {}^3\mathrm{O_{(g)}} \longrightarrow \text{s-}\mathrm{HO_2}, \tag{3.5e}$$

$$\text{s-}{}^3\mathrm{O} + \mathrm{OH_{(g)}} \longrightarrow \text{s-}\mathrm{HO_2}, \tag{3.5f}$$

$$\text{s-}{}^3\mathrm{O_2} + \mathrm{H_{(g)}} \longrightarrow \text{s-}\mathrm{HO_2}. \tag{3.5g}$$

Note that both (3.5e) and (3.5f) are ER versions of (3.4e), but with different reactant adsorbates. The reaction energies for (3.5) are given in Table 6. All reaction energies are computed according to (2.7).

For each molecule, we use the energy values at the optimized geometries of each functional plus the ZPE/all correction with the B3LYP functional. Note that we have an ambiguity in the choice of ZPE corrections for the reactions with $\mathrm{OH_{(g)}}$. This is the only case where the gaseous reactant does have a non-vanishing ZPE correction. This means that we can either use the correction term demanded by the individual functional or always the one demanded by B3LYP. We chose the B3LYP correction for the $\mathrm{OH_{(g)}}$ molecules, too, which makes the whole ZPE correction more consistent. This choice may be challenged, however it changes the energy of the two reactions

at hand by never more than 1.05 kJ/mol, and this biggest change appears for BHLYP, which is generally not in good agreement with the rest of the reaction energies anyway.

Recalling the strong deviations between DFT and CCSD(T)-F12a gas-phase reaction energies, we should not assume any DFT calculation to yield ER reaction energies too close to the physically correct value. But we know that we can decompose an ER reaction energy into a difference between adsorption energies and the gas-phase reaction energy. We can expect that the gas-phase reaction is described more accurately by CCSD(T)-F12a than by any of the DFT functionals. The benchmark in Section 3.1, on the other hand, gives reason to assume that the adsorption energies, which are basically interaction energies between the surface and the adsorbate, are described rather accurately by DFT functionals. We do not know which functional describes the adsorption best, so we take the average values of Table 5. Our recommendation of reaction energy is then the hybrid

$$E^{\text{react}}_{\text{hybrid}} = \Delta E^{\text{ads}}_{\text{avg}} + E^{\text{react}}_{\text{CCSD(T)-F12a}}. \tag{3.6}$$

For a reaction of ER type (2.6), $\Delta E^{\text{ads}}_{\text{avg}}$ is

$$\Delta E^{\text{ads}}_{\text{avg}} = E^{\text{ads}}_{\text{avg,Z}} - E^{\text{ads}}_{\text{avg,X}}, \tag{3.7}$$

with the values of $E^{\text{ads}}_{\text{avg}}$ given in Table 5.

We highlight the functionals that were closest to the CCSD(T)-F12a results for the gas phase by boldface. In the cases where PW6B95 was closest, we highlight both PW6B95 and PW6B95-D3. Since the adsorption energy does not vary too strongly between functionals, the functionals usually stay as close to the coupled-cluster results as they did in the gas phase. This is clearly visible when comparing the MAD, MIN, MAX and MEAN deviations. There are small shifts in approximative power among the functionals, the boldface functionals are now not always the best approximations to $E^{\text{react}}_{\text{hybrid}}$. But this is not really a matter of accuracy, since the change in difference to the reference value depends only on the adsorption energy average $E^{\text{ads}}_{\text{avg}}$, that is a functional can get closer to $E^{\text{react}}_{\text{hybrid}}$ by over- or underestimating its respective ΔE^{ads} as in (3.7). This is merely a cancellation of error. Therefore, the seemingly increased accuracy particularly in PW6B95-D3 results are consequences of the strong deviations from the average documented in the previous subsection.

We can only note that it would of course also be possible to use a more consistent correction than the simple average value, i.e. to decide on a single functional to yield ΔE^{ads}. From the results so far we would recommend B3LYP to be this functional, since the B3LYP results for adsorption energy are closest to the average value when comparing the overall MAD in Table 5.

Since we assume that $E^{\text{react}}_{\text{hybrid}}$ is reasonably accurate, we can now compare different reactions. The adsorption energies of s-OH, s-H_2O and s-H_2O_2 differ by less than 2 kJ/mol among one another.

Therefore, the first three surface reactions must have similar reaction energies to their gas-phase counterparts. They become more exothermic by between -1.48 and -3.58 kJ/mol. The only reaction to become less exothermic is (3.5d), the difference being $+16.89$ kJ/mol. This is because HO_2 is very strongly bound to the surface. With that, it is also no surprise that those reactions that have HO_2 as a product are much more exothermic, especially reaction (3.5g), where the reactant is the only very weakly bound s-$^3O_2$. This reaction is more exothermic by -53.02 kJ/mol. This is also a special case where the bad description of s-$^3O_2$ brings PW6B95-D3 close to the CCSD(T)-F12a results.
We can see a big energy difference between the related reactions (3.5e) and (3.5f). (3.5e) is $+21.95$ kJ/mol less exothermic than (3.5f). This, too, is the difference in adsorption energy between the strongly bound OH and the weaker bound ^{3}O.

3.4 Possible Application

The data we offered so far are mainly energy differences. Of course, the surface model can be used to derive other data relevant to astrochemistry, such as reaction rates, diffusion rates or tunneling rates. These could be further used as simulation input. We will not go into detail here but give a small outlook meant to exemplify other investigations possible on the surface.

3.4.1 Binding Sites

In the previous calculations, we did not investigate how strongly the optimum geometry of adsorption dependends on an initial geometry and if there are other local energy minima that molecules adsorbing to the surface can occupy. It is well possible, especially for those molecules that have additional internal and rotational degrees of freedom, that our search algorithm did not even find the most attractive geometry possible.
In this binding site study, we want to look a little deeper into these problems. For simplicity, we choose the s-^{3}O system. ^{3}O has spherical symmetry, therefore the binding energy depends only on the position of ^{3}O in space (and the corresponding surface deformation, which is assumed to be uniquely defined by adsorbate position). The average binding energy for all functionals was $E_{\text{avg}}^{\text{ads,ZPE}} = -12.35$ kJ/mol. The ZPE correction was $\Delta E_{\text{B3LYP}}^{\text{all}} = 2.69$ kJ/mol, which means an uncorrected binding energy of $E_{\text{avg}}^{\text{ads}} = -15.04$ kJ/mol.
We used the BHLYP functional with a ZPE correction free adsorption energy of -7.85 kJ/mol, which is the weakest bound energy. For this energy, there is reason to assume that there could be a more attractive binding site somewhere in the central ring. To arrive at different binding sites, we used a function to create initial data on the ring. Let the center of mass of the oxygen atoms of the central ring be located at the origin $(0,0,0)^\intercal$ and let y be normal to the Fletcher surface. We can

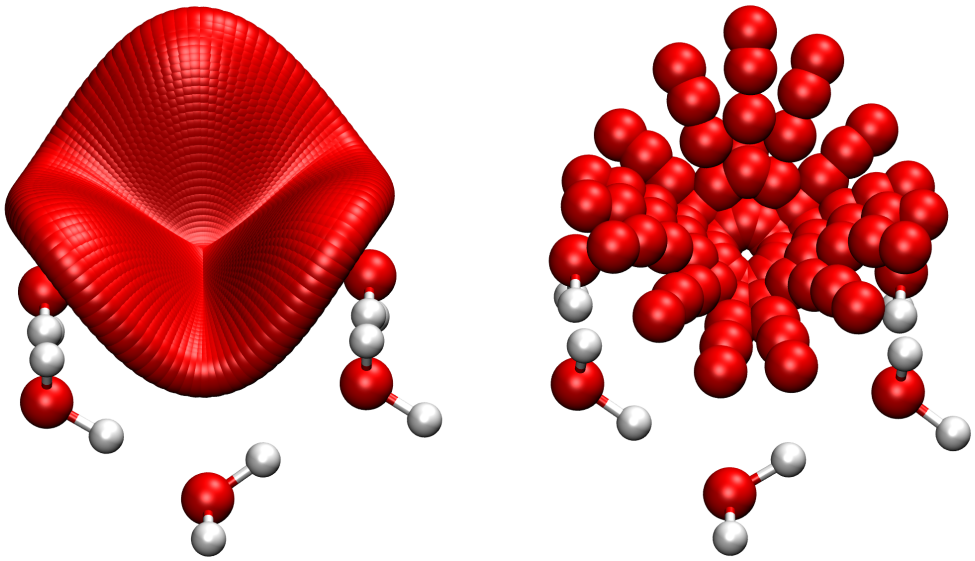


Figure 12 The initial guesses for binding sites. The left is a visualization of the function used, the right is the actual sites used as initial geometry. The molecules below are the central ring.

transform (x,z) in polar coordinates with $x = r\cos\phi$, $z = r\sin\phi$, $0 \leq r \leq r_{\max}$, $0 \leq \phi < 2\pi$. The direction of the unit vector $\mathbf{e}_x$ (that is $\mathbf{e}_r$ for $\phi = 0$) and the value of $r_{\max}$ are chosen to ensure that there is an oxygen atom at the xyz position vector $(r_{\max}, y < 0, 0)^\intercal$. In Figure 12, the vector could be pointing towards the oxygen atom in the front. We found $r_{\max} = 2.4$ Å.

In this coordinate system, we used the initial position

$$y(r,\phi) = y_0 + \left(A - B\cos(3\phi)\right)\sin^2\left(\frac{r}{r_{\max}}\frac{\pi}{2}\right). \tag{3.8}$$

This initial guess is just to avoid unphysical starting positions with too far-off or too close ^{3}O atoms. To that end, the parameters y_0, A and B are included, which we chose to be $y_0 = 1.5$ Å, $A = 1.5$ Å, $B = 0.8$ Å.

We covered the area $(r,\phi) \in [0, r_{\max}] \times [0, 2\pi)$ with equidistantly distributed points according to $r_k = r_{\min} + k(r_{\max} - r_{\min})/(N-1)$ and $\phi_l = 2\pi l/M$, $0 \leq k < N$ and $0 \leq l < M$, with $r_{\min} = 0.4$ Å and point numbers $N = 5$, $M = 15$, leading to a total of 75 initial geometries. The equidistant distribution is chosen to cover as many local minima on the x-z plane as possible.

The oxygen atoms are added to the optimum geometry of the bare surface, where the central ring is nearly undistorted.

We then used the L-BFGS algorithm to find new energy minima. Of the 75 initial guesses, four did not converge in the six day computation time limit of the JUSTUS cluster. The remaining initial positions were optimized and yielded the energies plotted in Figure 13. One can see that most prominently, two adsorption energy values are found. One around -8.4 kJ/mol and one around

-7.85 kJ/mol. In our first study, we only found the less attractive geometry, which is directly above the center of mass at $(0,2,0)^\intercal$, so close to the initial guess there. Other minima are usually above lines connecting oxygen atoms. The most attractive geometries profit of hydrogen bonds. The worst case, with an energy value of only -5.01 kJ/mol, has the adsorbate positioned on a broken H bond. This minimum is probably not very stable, it is only assumed by a single computation, all others have energies of at least -6.70 kJ/mol. The strongest adsorption of -8.64 kJ/mol is also reached by only a single calculation. Therefore, we conclude that there are two attractive energy minima to which more than half of the initial geometries are drawn.

We recommend a central binding site at $(0,1.97,0)^\intercal$. This is not the strongest binding site, but all optimum geometry searches starting close to it converged there, so we can infer that there is no other energy minimum close by. At greater distance from the center, we recommend further binding sites above hydrogen atoms pointing out of the surface at a distance of 2.17 Å to the hydrogen atom.

This becomes interesting when we consider hopping of adsorbates. In Figure 12, the optima with $E < -8$ kJ/mol lie to the left and right. But then, due to quantum tunneling or temperature-dependent vibrations, an O atom located above the center of mass can hop into another optimum at the boundary of the ring, we may call it a transmission optimum, and continue hopping into other rings from there. This mechanism should be favorable for hopping ratios, but then there should be a preferred direction of hopping on a Fletcher surface.

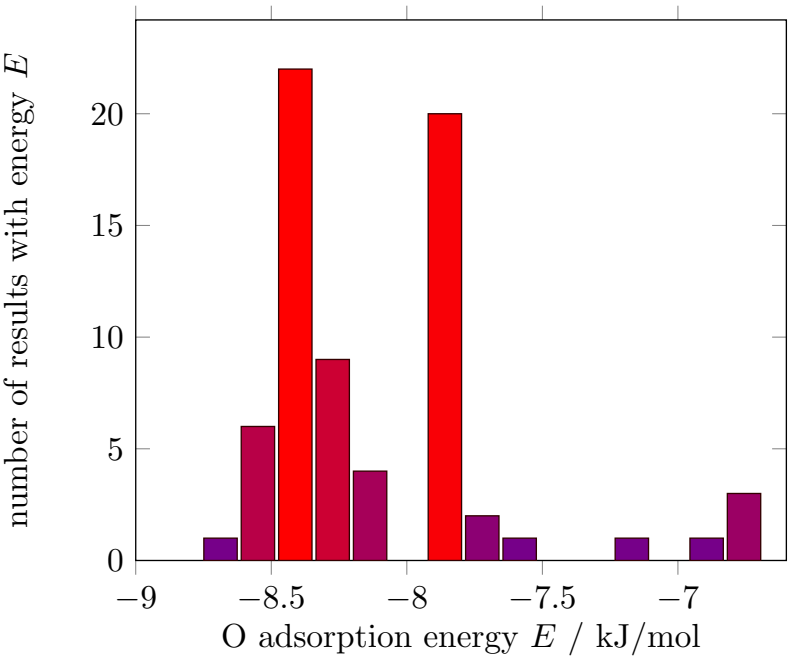


Figure 13 Population of optimum geometry energies. Of the 75 initial guesses according to (3.8), 71 converged to a minimum. One geometry at $E = -5.01$ kJ/mol is not included. Energies in kJ/mol.

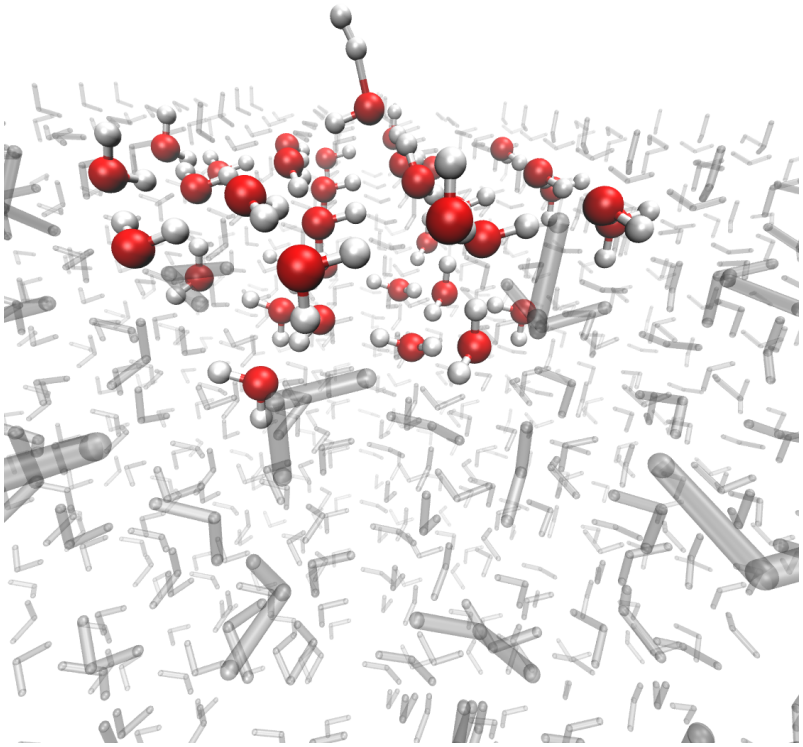


Figure 14 $OH + H_2$ transition state on Fletcher surface. Calculated with B3LYP. Transparent atoms are in the MM region, red (O) an white (H) atoms in the QM region.

This may be interesting to study. However, one has to bear in mind that the Fletcher surface is only an idealization in a zero temperature limit. Even at temperatures in the cold ISM, the hydrogen atoms must give up their order. To what extent this happens could be crucial to further analysis.

3.4.2 Transition State

As mentioned before, transition states are important to reaction kinetics. The energy difference between a system at the transition state and a system at the optimum geometry yields the activation energy, a key ingredient to reaction kinetics.

We studied the system $OH + H_2$, which is interesting for the reaction $OH + H_2 \longrightarrow H_2O + H$. This reaction is part of the water formation scheme[1] and was studied experimentally by Oba *et al.*[83] We did not use any coupled-cluster references in this test but started dircetly with DFT calculations. We used the B3LYP functional because it was closest to the average value for adsorption energies in the previous section. We first calculated a gas-phase transition state using the dimer method (cf. Section 2.5) and the def2-TZVPD basis set. The transition state geometry was placed on the relaxed Fletcher surface obtained with B3LYP and the hybrid def2-TZVP/def2-TZVPD basis set. Unfortunately, the orientation of the $OH + H_2$ subsystem on the surface is not clear.

We assumed that the OH radical should be at least as close to the surface as the H_2 molecule, since its adsorption energy is predicted to be far more attractive. We therefore chose an initial geometry where the OH is placed above the center of the central ring, the H_2 is slightly above it. To get a better guess for the initial geometry, we then performed dimer method iteration with the smaller def2-SVP basis set, despite it being not too good at describing geometries in the benchmarks. Still, the dimer method typically requires more than one energy evaluation per iteration cycle, which makes it computationally very expensive for the hybrid basis set if the initial geometry is too far from the transition state geometry. And indeed, it took def2-SVP 99 iteration cycles with a total of 405 energy and gradient evaluations to reach the transition state. The subsequent optimization with the hybrid basis did, however, also require 87 iteration cycles with 477 energy and gradient evaluations. The RMSD between the initial geometry and the def2-SVP transition state is 0.11 Å (0.91 Å for $OH + H_2$ alone) and the RMSD between the def2-SVP transition state and the hybrid basis transition state is 0.08 Å (0.22 Å). As a comparison, the RMSD between the optimum geometry (that is the energy minimum) and the transition state at the hybrid basis set level is 0.06 Å (0.90 Å).

We can make some notes on the differences between our initial geometry and the two transition states for different basis sets. The first one is that in this case, def2-SVP seems to yield geometries that are reasonably close to the hybrid basis geometry. That means that the geometries are not good enough to consider the def2-SVP result a good approximation, but they are close enough to serve as initial geometries. On the other hand, we can not be sure whether the initial guess provided by def2-SVP is necessary at all, since the majority of the convergence cycles occur at geometries that are already very close to the transition state (when compared by RMSD value). This explains the only slight difference in iteration cycles. The approach may still be sensible for larger surface reactions. As a third note, the effect of orientation of the adsorbate in the initial geometry may be an interesting topic for future research.

Relevant data for the bond distances and angles can be found in Figure 15. Figure 14 shows the transition state on the surface, where we have a better look at how the surface adjusts to the adsorbate. Most notably, the H atoms pointing out of the surface are tilted towards the O atom of OH, while the H atom in OH is oriented towards an H_2O molecule with a clipped hydrogen bond pointing out. The H_2 molecule is too far from the surface to relevantly deform it.

Comparing gas-phase and adsorbate geometries for the transition state, we find very good agreement. Both transition states are predicted nearly planar, which is not true for the pre-reactive complex. Bond lengths do not change greatly between transition state and minimum geometry, but the separation between the two molecules decreases in the transition state. If we compare the gas-phase geometry to the adsorbate geometry for the minimum, the disagreement is stronger.

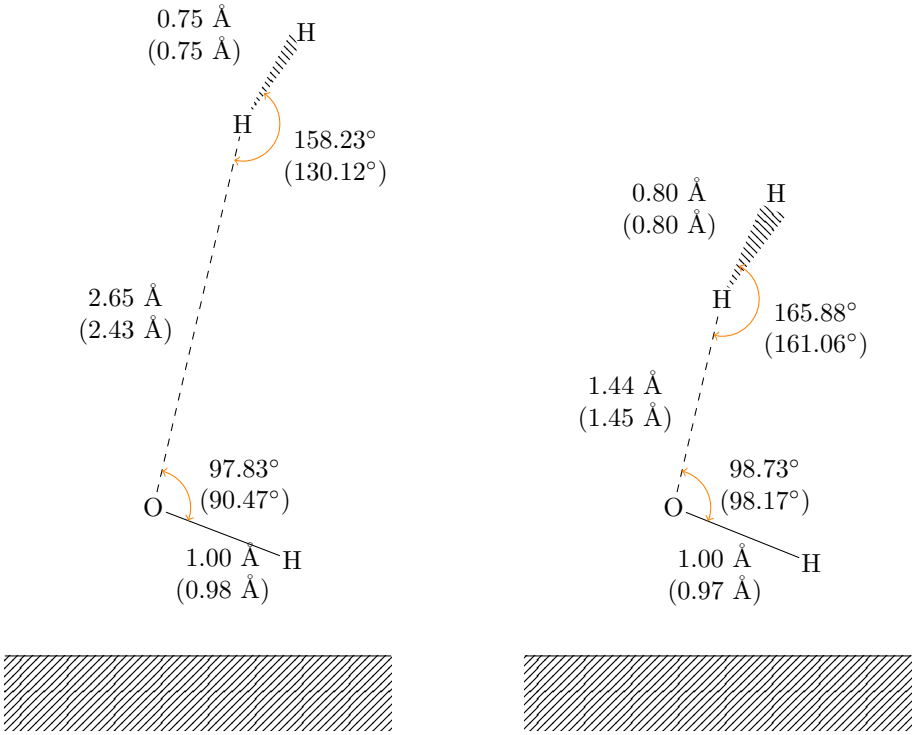


Figure 15 Left: Optimum geometry for $OH + H_2$. Right: Transition state for $OH + H_2$. Values are given for the adsorbate, values in parentheses are for the gas phase. The dihedral angles for the optimum geometry are 50.21° (33.58°) and for the transition state 2.58° (1.32°).

The H_2 molecule has an increased separation from the OH radical. The "upper" H atom in H_2 tilts towards the surface. The stronger deformation may be attributed to the fact that the interaction between OH and H_2 is weaker in the optimum geometry and that therefore the surface affects the optimum geometry stronger.

A reaction energy diagram is presented in Figure 16.

We find the activation energy for the reaction $OH + H_2 \longrightarrow H_2O + H$ to be 7.16 kJ/mol and the reaction energy to be -44.03 kJ/mol. This does not compare well to experimental data by Romanzin *et al.*,[84] who found an activation energy of 22.19 kJ/mol and a reaction energy of -96.49 kJ/mol. A source of error could be the neglect of a finite temperature in the optimization. Part of the problems may also come from problems of DFT describing the breaking and rearrangement of bonds of the given reactions, which we already encountered for gas-phase reactions in Section 3.2.2.

The difference between gas-phase reaction energies and surface reaction energies is not very strong, it affects the activation energy by only $+0.06$kJ/mol, the reaction energy is changed by $+1.48$kJ/mol. This small change could come from the fact that we put the H_2 molecule above the OH molecule, so it does not interact strongly with surface the in both the pre-reactive complex and the transition state.

On the surface, the influence of the ZPE/ring correction is small for the transition state compared

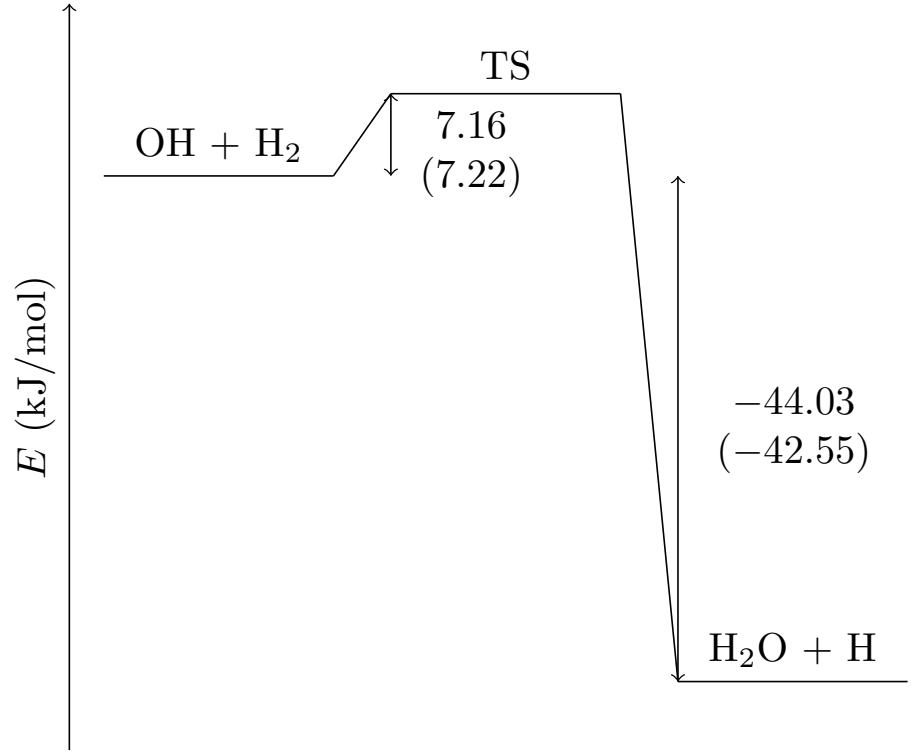


Figure 16 Energy diagram for the reaction $OH + H_2 \longrightarrow H_2O + H$ on the surface. Values in parentheses are the corresponding gas-phase energies.

to the initial $OH + H_2$ energy, whereas the correction has a stronger effect on the product state, where it makes up 18% of the energy. In the gas-phase, the ZPE correction makes up 11.5% and 13.9% of the energy for the transition state and the product, respectively.

4 Summary and Conclusion

We have designed a crystalline water I_h Fletcher surface model for QM/MM calculation of adsorption processes and chemical reactivity on the surface. With this surface, we can give recommendations on the DFT functionals that can be used in the QM part, which we base on an interaction energy benchmark. The ZPE correction for the QM part was discussed and we compared two different approaches, where we came to the conclusion that the computationally less expensive ZPE/ring correction still provides results with acceptable accuracy.

The adsorption energies calculated with the different functionals show not too high disagreement among one another, however the ZPE corrected values yield unphysical results for weakly adsorbed molecular species. We calculated average values, which are the best recommendation we can give. The B3LYP functional is the functional closest to the average by MAD.

With the adsorption energies, we can calculate surface reaction eneriges of the ER type based on gas-phase reaction energies at the CCSD(T)-F12a/VTZ-F12 level of theory. With the good results

of the interaction energy benchmark and the accuracy of CCSD(T)-F12a for reaction energies, we can hope that the accuracy of the surface reaction energies is good.

In two small investigations, we considered the ability of the model to analyze different binding sites for a specific adsorbate and the ability to describe transition states. We found two main classes of binding sites for the ^{3}O radical, and we found that the calculation of transition states is computationally feasible.

Topics for further research could be the evaluation of orientation dependence in both the adsorption of molecules and the calculation of transition states. One could also look for a more systematic approach to analyze binding sites, which may obtain a PES for some adsorbate.

Also, research in kinetic properties of surface reactions could be a next big step. It can be hoped that insight into Langmuir–Hinshelwood type reactions can be gained from the surface model, including diffusion rates for adsorbates as well as surface reaction rates. One could also evaluate the influence of tunneling on surface reactions and surface diffusion.

Since the crystalline surface is a rather strong idealization, the model could be further adjusted to include a grain surface. Such a model can describe adsorption processes on ices only a few monolayers thick, thereby describing early stages of water ice formation.

Additionally, less ordered surfaces are of interest to contemporary research. An investigation of which of the results for the crystalline water I_h surface also apply to amorphous solid water and how strongly adsorption energies can vary dependent on the surface composition should be interesting. The same goes for incorporating impurities into the surface, most importantly CO_2. Esepcially the accuracy of the ZPE/ring correction may greatly vary from surface type to surface type, while additionally being not so well-defined for less symmetrical surfaces.

In conclusion, the model established here is only a starting point for a possibly wide range of theoretical research around the topic of interstellar grain mantle composition and reactions. It seems appropriate to further advance the application of *ab initio* calculations in this field, which would help both experimentalists and theoreticians to a better understanding of what exactly is happening in interstellar clouds and possibly a more detailed explanation of water formation in the interstellar medium within the next few decades.

Acknowledgments

For this work, I received a lot of support from my group and many steps were planned together. Therefore, I decided to use “we” in this thesis.

I want to thank Prof. Johannes Kästner for his support, Dr. Thanja Lamberts for proofreading and Mr. Jan Meisner for his introduction to the topic, his helpful suggestions and his company along

the way. I also thank Dr. Jens Smiatek for assisting the exam.
This research was supported in part by the bwHPC initiative and the bwHPC-C5 project[70] provided through associated compute services of the JUSTUS HPC facility at the University of Ulm.

References

[1] E. F. van Dishoeck, E. Herbst, and D. A. Neufeld, "Interstellar Water Chemistry: From Laboratory to Observations," *Chemical Reviews*, 2013.

[2] D. A. Williams and E. Herbst, "It's a dusty universe: surface science in space," *Surface Science*, 2002.

[3] R. J. Trumpler, "ABSORPTION OF LIGHT IN THE GALACTIC SYSTEM," *Publications of the Astronomical Society of the Pacific*, 1930.

[4] H. A. Zook, "Spacecraft Measurements of the Cosmic Dust Flux," in *Accretion of Extraterrestrial Matter Throughout Earth's History*, pp. 75–92, 2001.

[5] A. J. Westphal, R. M. Stroud, H. A. Bechtel, F. E. Brenker, A. L. Butterworth, G. J. Flynn, D. R. Frank, Z. Gainsforth, J. K. Hillier, F. Postberg, A. S. Simionovici, V. J. Sterken, L. R. Nittler, C. Allen, D. Anderson, A. Ansari, S. Bajt, R. K. Bastien, N. Bassim, J. Bridges, D. E. Brownlee, M. Burchell, M. Burghammer, H. Changela, P. Cloetens, A. M. Davis, R. Doll, C. Floss, E. Grun, P. R. Heck, P. Hoppe, B. Hudson, J. Huth, A. Kearsley, A. J. King, B. Lai, J. Leitner, L. Lemelle, A. Leonard, H. Leroux, R. Lettieri, W. Marchant, R. Ogliore, W. J. Ong, M. C. Price, S. A. Sandford, J.-A. S. Tresseras, S. Schmitz, T. Schoonjans, K. Schreiber, G. Silversmit, V. A. Sole, R. Srama, F. Stadermann, T. Stephan, J. Stodolna, S. Sutton, M. Trieloff, P. Tsou, T. Tyliszczak, B. Vekemans, L. Vincze, J. V. Korff, N. Wordsworth, D. Zevin, and M. E. Zolensky, "Evidence for interstellar origin of seven dust particles collected by the Stardust spacecraft," *Science*, 2014.

[6] B. Draine, "Interstellar Dust Grains," *Annual Review of Astronomy and Astrophysics*, 2003.

[7] E. E. Becklin and G. Neugebauer, "Observations of an Infrared Star in the Orion Nebula," *The Astrophysical Journal*, 1967.

[8] F. C. Gillett and W. J. Forrest, "Spectra of the Becklin–Neugebauer point source and the Kleinmann–Low nebula from 2.8 to 13.5 microns.," *The Astrophysical Journal*, 1973.

[9] K. Hiraoka, T. Miyagoshi, T. Takayama, K. Yamamoto, and Y. Kihara, "Gas-Grain Processes for the Formation of CH_4 and H_2O: Reactions of H Atoms with C, O, and CO in the Solid Phase at 12 K," *The Astrophysical Journal*, 1998.

[10] A. A. Boogert, P. A. Gerakines, and D. C. Whittet, "Observations of the Icy Universe," *Annual Review of Astronomy and Astrophysics*, 2015.

[11] K. M. Ferrière, "The interstellar environment of our galaxy," *Reviews of Modern Physics*, 2001.

[12] A. Léger, J. Klein, S. de Cheveigne, C. Guinet, D. Defourneau, and M. Belin, "The 3.1 micron Absorption in Molecular Clouds Is Probably Due to Amorphous H_2O Ice," *Astronomy and Astrophysics*, 1979.

[13] P. Jenniskens, D. F. Blake, M. A. Wilson, and A. Pohorille, "High-Density Amorphous Ice, the Frost on Interstellar Grains," *The Astrophysical Journal*, 1995.

[14] A. G. G. M. Tielens and W. Hagen, "Model Calculations of the Molecular Composition of Interstellar Grain Mantles," *Astronomy and Astrophysics*, 1982.

[15] H. Cuppen and E. Herbst, "Simulation of the formation and morphology of ice mantles on interstellar grains," *The Astrophysical Journal*, 2007.

[16] "Langmuir–Hinshelwood mechanism," in *Gold Book*, International Union of Pure and Applied Chemistry (IUPAC).

[17] K. J. Laidler, "A glossary of terms used in chemical kinetics, including reaction dynamics (IUPAC Recommendations 1996)," *Pure and Applied Chemistry*, 1996.

[18] R. Papoular, "On water ice formation in interstellar clouds," *Monthly Notices of the Royal Astronomical Society*, 2005.

[19] H. Bergeron, N. Rougeau, V. Sidis, M. Sizun, D. Teillet-Billy, and F. Aguillon, "OH Formation from O and H Atoms Physisorbed on a Graphitic Surface through the Langmuir–Hinshelwood Mechanism: A Quasi-Classical Approach," *The Journal of Physical Chemistry A*, 2008.

[20] P. Cabrera Sanfelix, S. Holloway, K. Kolasinski, and G. Darling, "The structure of water on the (0001) surface of graphite," *Surface Science*, 2003.

[21] J. P. Perdew and Y. Wang, "Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation," *Physical Review B*, 1986.

[22] J. P. Perdew and Y. Wang, "Accurate and simple analytic representation of the electron-gas correlation energy," *Physical Review B*, 1992.

[23] C. S. Lin, R. Q. Zhang, S. T. Lee, M. Elstner, T. Frauenheim, and L. J. Wan, "Simulation of Water Cluster Assembly on a Graphite Surface," *The Journal of Physical Chemistry B*, 2005.

[24] T. P. M. Goumans, C. R. A. Catlow, and W. A. Brown, "Hydrogenation of CO on a silica surface: An embedded cluster approach," *The Journal of Chemical Physics*, 2008.

[25] T. P. M. Goumans, C. R. A. Catlow, W. A. Brown, J. Kästner, and P. Sherwood, "An embedded cluster study of the formation of water on interstellar dust grains," *Physical Chemistry Chemical Physics*, 2009.

[26] G. Henkelman, B. P. Uberuaga, and H. Jónsson, "A climbing image nudged elastic band method for finding saddle points and minimum energy paths," *The Journal of Chemical Physics*, 2000.

[27] D. E. Woon, "Modeling Gas-Grain Chemistry with Quantum Chemical Cluster Calculations. I. Heterogeneous Hydrogenation of CO and H_2CO on Icy Grain Mantles," *The Astrophysical Journal*, 2002.

[28] H. bin Xie, Y. hong Ding, and C. chung Sun, "Reaction mechanism of oxygen atoms with cyanoacetylene in the gas phase and on water ice," *The Astrophysical Journal*, 2006.

[29] C. Møller and M. S. Plesset, "Note on an Approximation Treatment for Many-Electron Systems," *Physical Review*, 1934.

[30] J. Gauss and D. Cremer, "Analytical evaluation of energy gradients in quadratic configuration interaction theory," *Chemical Physics Letters*, 1988.

[31] L. J. Karssemeijer, A. Pedersen, H. Jónsson, and H. M. Cuppen, "Long-timescale simulations of diffusion in molecular solids," *Physical Chemistry Chemical Physics*, 2012.

[32] P. Hohenberg and W. Kohn, "Inhomogeneous Electron Gas," *Physical Review*, 1964.

[33] W. Kohn and L. J. Sham, "Self-Consistent Equations Including Exchange and Correlation Effects," *Physical Review*, 1965.

[34] A. D. Becke, "Density-functional thermochemistry. III. The role of exact exchange," *The Journal of Chemical Physics*, 1993.

[35] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Physical Review A*, 1988.

[36] J. P. Perdew, "Density-functional approximation for the correlation energy of the inhomogeneous electron gas," *Physical Review B*, 1986.

[37] C. Lee, W. Yang, and R. G. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density," *Physical Review B*, 1988.

[38] J. P. Perdew, M. Ernzerhof, and K. Burke, "Rationale for mixing exact exchange with density functional approximations," *The Journal of Chemical Physics*, 1996.

[39] S. Grimme, "Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction," *Journal of Computational Chemistry*, 2006.

[40] J. Tao, J. P. Perdew, V. N. Staroverov, and G. E. Scuseria, "Climbing the density functional ladder: Nonempirical meta–generalized gradient approximation designed for molecules and solids," *Physical Review Letters*, 2003.

[41] Y. Zhao and D. G. Truhlar, "Design of density functionals that are broadly accurate for thermochemistry, thermochemical kinetics, and nonbonded interactions," *The Journal of Physical Chemistry A*, 2005.

[42] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, "Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields," *The Journal of Physical Chemistry*, 1994.

[43] A. D. Becke, "A new mixing of Hartree–Fock and local density-functional theories," *The Journal of Chemical Physics*, 1993.

[44] V. N. Staroverov, G. E. Scuseria, J. Tao, and J. P. Perdew, "Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes," *The Journal of Chemical Physics*, 2003.

[45] C. Adamo and V. Barone, "Toward reliable density functional methods without adjustable parameters: The PBE0 model," *The Journal of Chemical Physics*, 1999.

[46] Y. Zhao and D. G. Truhlar, "The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals," *Theoretical Chemistry Accounts*, 2007.

[47] TURBOMOLE V6.6 2014, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.

[48] M. Valiev, E. Bylaska, N. Govind, K. Kowalski, T. Straatsma, H. V. Dam, D. Wang, J. Nieplocha, E. Apra, T. Windus, and W. de Jong, "NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations," *Computer Physics Communications*, 2010.

[49] A. Schäfer, H. Horn, and R. Ahlrichs, "Fully optimized contracted Gaussian basis sets for atoms Li to Kr," *The Journal of Chemical Physics*, 1992.

[50] F. Weigend, M. Häser, H. Patzelt, and R. Ahlrichs, "RI-MP2: optimized auxiliary basis sets and demonstration of efficiency," *Chemical Physics Letters*, 1998.

[51] F. Weigend, F. Furche, and R. Ahlrichs, "Gaussian basis sets of quadruple zeta valence quality for atoms H–Kr," *The Journal of Chemical Physics*, 2003.

[52] D. Rappoport and F. Furche, "Property-optimized Gaussian basis sets for molecular response calculations," *The Journal of Chemical Physics*, 2010.

[53] T. Anacker and J. Friedrich, "New Accurate Benchmark Energies for Large Water Clusters: DFT Is Better Than Expected," *Journal of Computational Chemistry*, 2014.

[54] E. S. Kryachko and E. V. Ludeña, *Energy density functional theory of many-electron systems*, vol. 4. 2012.

[55] A. D. Becke and E. R. Johnson, "Exchange-hole dipole moment and the dispersion interaction revisited," *The Journal of Chemical Physics*, 2007.

[56] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu," *The Journal of Chemical Physics*, 2010.

[57] S. Grimme, S. Ehrlich, and L. Goerigk, "Effect of the damping function in dispersion corrected density functional theory," *Journal of Computational Chemistry*, 2011.

[58] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *The Journal of Chemical Physics*, 1983.

[59] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, and et al., "CHARMM: The Biomolecular Simulation Program," *Journal of Computational Chemistry*, 2009.

[60] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, and et al., "All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins," *The Journal of Physical Chemistry B*, 1998.

[61] A. Warshel and M. Levitt, "Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme," *Journal of Molecular Biology*, 1976.

[62] J. Kästner, J. M. Carr, T. W. Keal, W. Thiel, A. Wander, and P. Sherwood, "DL-FIND: An Open-Source Geometry Optimizer for Atomistic Simulations," *The Journal of Physical Chemistry A*, 2009.

[63] D. C. Liu and J. Nocedal, "On the limited memory BFGS method for large scale optimization," *Mathematical Programming*, 1989.

[64] G. Henkelman and H. Jónsson, "A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives," *The Journal of Chemical Physics*, 1999.

[65] I. T. Todorov, W. Smith, K. Trachenko, and M. T. Dove, "DL_POLY_3: new dimensions in molecular dynamics simulations via massive parallelism," *Journal of Materials Chemistry*, 2006.

[66] ChemShell, a Computational Chemistry Shell, see www.chemshell.org.

[67] S. Metz, J. Kästner, A. A. Sokol, T. W. Keal, and P. Sherwood, "ChemShell—a modular software package for QM/MM simulations," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2013.

[68] 2012. MOLPRO, version 2012.1, a package of ab initio programs, see http://www.molpro.com/.

[69] W. Humphrey, A. Dalke, and K. Schulten, "VMD – Visual Molecular Dynamics," *Journal of Molecular Graphics*, 1996.

[70] 1996. bwHPC and bwHPC-C5 (http://www.bwhpc-c5.de) funded by the Ministry of Science, Research and the Arts Baden-Württemberg (MWK) and the Germany Research Foundation (DFG).

[71] G. Knizia, T. B. Adler, and H.-J. Werner, "Simplified CCSD(T)-F12 methods: Theory and benchmarks," *The Journal of Chemical Physics*, 2009.

[72] K. E. Yousaf and K. A. Peterson, "Optimized auxiliary basis sets for explicitly correlated methods," *The Journal of Chemical Physics*, 2008.

[73] T. Shiozaki, G. Knizia, and H.-J. Werner, "Explicitly correlated multireference configuration interaction: MRCI-f12," *The Journal of Chemical Physics*, 2011.

[74] S. R. Langhoff and E. R. Davidson, "Configuration interaction calculations on the nitrogen molecule," *International Journal of Quantum Chemistry*, 1974.

[75] W. Kabsch, "A discussion of the solution for the best rotation to relate two sets of vectors," *Acta Crystallographica Section A*, 1978.

[76] W. H. Barnes, "The Crystal Structure of Ice between 0 ° C. and −183 ° C.," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 1929.

[77] N. Fletcher, "The freezing of water," *Science Progress (Oxford)*, 1966.

[78] N. H. Fletcher, "Reconstruction of ice crystal surfaces at low temperatures," *Philosophical Magazine Part B*, 1992.

[79] H. J. Fraser, M. P. Collings, M. R. McCoustra, and D. A. Williams, "Thermal desorption of water ice in the interstellar medium," *Monthly Notices of the Royal Astronomical Society*, 2001.

[80] J. He and G. Vidali, "Experiments of Water Formation on Warm Silicates," *The Astrophysical Journal*, 2014.

[81] J. He, J. Shi, T. Hopkins, G. Vidali, and M. J. Kaufman, "A New Determination of the Binding Energy of Atomic Oxygen On Dust Grain Surfaces: Experimental Results and Simulations," *The Astrophysical Journal*, 2015.

[82] J. He, D. Jing, and G. Vidali, "Atomic oxygen diffusion on and desorption from amorphous silicate surfaces," *Physical Chemistry Chemical Physics*, 2014.

[83] Y. Oba, N. Watanabe, T. Hama, K. Kuwahata, H. Hidaka, and A. Kouchi, " WATER FORMATION THROUGH A QUANTUM TUNNELING SURFACE REACTION, OH + H_2 , AT 10 K ," *The Astrophysical Journal*, 2012.

[84] C. Romanzin, S. Ioppolo, H. M. Cuppen, E. F. van Dishoeck, and H. Linnartz, "Water formation by surface O_3 hydrogenation," *The Journal of Chemical Physics*, 2011.