Surface Adsorption on Interstellar Ice I_h

Thomas Bissinger

[width=7cm]./Figures/unistuttgart.jpg

Master Thesis supervised by

Prof. Dr. rer. nat. J. Kästner

M. Sc. Jan Meisner

Institute for Theoretical Chemistry, Universität Stuttgart, December 2015

Surface Adsorption on Interstellar Ice I_h

T. Bissinger

January 22, 2016

Abstract

Molecule formation in the interstellar medium is thought to be assisted by surface reactions on ice mantles of interstellar dust grains. We present a model for a water I_h Fletcher surface within a QM/MM description. For MM, we use the TIP3P force field, for the QM region we propose a def2-TZVP/def2-TZVPD hybrid basis set and a selection of functionals which are compared to CCSD(T)-F12A/VTZ-F12 gas phase interaction energies in a benchmark. We give formation and reaction energies for gas phase molecules and adsorption and reaction energies for molecules on the surface. Combining adsorption energies and gas phase reaction energies, it is possible to gain very accurate results for surface reaction energies. We propose the special ZPE/ring correction for the surface which yields similar results to the most accurate ZPE/all correction at a fraction of its computational costs. An outlook on binding site analysis and transition state search shows that the model is fit for further application.

1 Introduction

Interstellar chemistry is the key ingredient to understanding the molecular abundancies 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 happens in interstellar clouds. With modern telescopes it is possible to measure the molecular abundancies in the interstellar medium (ISM) to a high degree of accuracy. It became evident that the reaction rates governing the formation of molecules can not properly be explained by gas phase chemistry alone.

A prominent theory to mend this discrepancy is to consider the contribution of surface reaction on interstellar dust grains. Interstellar dust was first observed by Trumpler[1], who attributed interstellar extinction 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[2].

Apart from dust, a wide variety of molecules is observed in the ISM. In 1976, Merrill *et al.*[3] discovered and absorption feature which they identified with the H_2O stretching vibration. The presence of water was controversial at first, but definite proof of the existence of water in the ISM was proven at latest by the Interstellar Space Observatory in 2004 by Gibb *et al.*[4].

While at high temperatures the water will not freeze on the grains, at temperatures as low as in molecular clouds (between 10 and 20 K[5]), some of the grains are believed to have icy mantles, the mantle material is mostly composed of H_2O and CO. In the case of water, one faces *amorphous solid water* (ASW), but at very low temperatures one may find a state close to the crystalline I_h state of frozen water.

In 1982 Tielens and Hagen[6] published an influential paper proposing a network of water formation reactions by hydrogenation of O, O_2 and O_3. The network was refined by later studies[7]. The modern version gives an account of the most relevant reactions that lead to water formation. Depending on temperatures, densities and radiation, different reactions of the network are assigned varying importance dependent on the interstellar region.

While experimental research in the field of surface reaction 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 analysed by Cabrera Sanfelix *et al.* [8] in 2003. They analysed the early accretion of water on a bare graphite grain using density functional theory (DFT) with PW91 functional, in which they found that the water is physisorbed on the surface. 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.

Other computational approaches to ice adsorption did so far not employ an electronic structure theory directly 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 is legit and does indeed not only yield good approximations to experimental data but also gives rise to new conclusions, it still needs further verification.
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 physisorption on 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 while we give a selection of DFT functionals that should yield accurate results for adsorption energies and a basis set recommendation.
We will describe the theory underlying the model in the next section. Section 3 then describes the benchmarking we performed on smaller test systems to determine the best functionals and basis sets to use to describe intermolecular interaction and therefore adsorption. After that, we give our results for adsorption energies and surface reaction energies in Section 5. As a suggestion of further applications, we give an idea of how one can analyse binding sites and transition states for adsorabtes. Finally, there will be concluding remarks and an outlook on possible further application for our findings in Section 7.

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 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 of both molecules $\mathrm{X}-\mathrm{Y}$. We also consider the *adsorption energy* of a molecule X on the ice surface S. We call this system s-X. If not declared elsewise, all appearing energies are electronic ground state energies. We wil 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 $\mathrm{X}-\mathrm{Y}$, which will in general depend on the distance and the orientation of the two molecules, to be $E_{\mathrm{X}-\mathrm{Y}}$. Then, the interaction energy between the two molecules is the energy given by

$$E^{\mathrm{int}}_{\mathrm{X}-\mathrm{Y}} := E_{\mathrm{X}-\mathrm{Y}} - E_\mathrm{X} - E_\mathrm{Y}. \tag{1}$$

We did not include spatial dependence of $E_{\mathrm{X}-\mathrm{Y}}$ into the above definition. A map $(R_{\mathrm{X}-\mathrm{Y}}, \Omega_\mathrm{X}, \Omega_\mathrm{Y}) \mapsto E^{\mathrm{int}}_{\mathrm{X}-\mathrm{Y}}$ with the center of mass separation $R_{\mathrm{X}-\mathrm{Y}}$ 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.
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 $\mathrm{X}-\mathrm{Y}$, that is the global minimum of the PES $E_{\mathrm{X}-\mathrm{Y}}$ and therefore $E^{\mathrm{int}}_{\mathrm{X}-\mathrm{Y}}$. If the interaction between X and Y were purely repulsive, that is $E^{\mathrm{int}}_{\mathrm{X}-\mathrm{Y}} > 0$ 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 around the initial geometry we specify, by which we will then classify the strength of the repulsion. Even for attractive potentials, that is potentials where $E^{\mathrm{int}}_{\mathrm{X}-\mathrm{Y}} < 0$ for some geometries, 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.
The adsorption energy is mostly similar. 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

$$E^{\mathrm{ads}}_\mathrm{X} := E_{\text{s-X}} - E_\mathrm{X} - E_\mathrm{S}. \tag{2}$$

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 is quite sensible 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 $(\mathbf{r}_i)_i \mapsto E^{\mathrm{ads}}_{\text{s-X}}$, 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 $E^{\mathrm{ads}}_\mathrm{X}$ are called *binding geometries*, and the position of the center of mass of the molecule X in a binding geometry is called a *binding site*. Exploring binding sites and the strength of the binding $E^{\mathrm{ads}}_\mathrm{X}$ can be useful to simulation.
We will also give reaction energies. In this work, we distinguish two kinds of reactions: Gas phase reactions, where isolated molecules come in contact, and Eley-Rideal (ER)[?] or surface

reactions as described in the introduction. A gas phase reaction is described by a scheme

$$\mathrm{X_{(g)}} + \mathrm{Y_{(g)}} \longrightarrow \mathrm{Z_{(g)}}, \tag{3}$$

where the subscript (g) indicates that gas phase. The initially isolated reactants $\mathrm{X_{(g)}}$ and $\mathrm{Y_{(g)}}$ combine to form a product $\mathrm{Z_{(g)}}$. Of course, there are also reactions possible with more than one product, we will however not consider these.
The reaction energy consumed or released in this process is

$$E_{\mathrm{gas}}^{\mathrm{react}} := E_{\mathrm{Z_{(g)}}} - E_{\mathrm{X_{(g)}}} - E_{\mathrm{Y_{(g)}}}. \tag{4}$$

Positive $E_{\mathrm{gas}}^{\mathrm{react}}$ means an *endothermic*, negative $E_{\mathrm{gas}}^{\mathrm{react}}$ an *exothermic* reaction. When neglecting quantum tunneling, only barrierless exothermic reactions occur in the gas phase since there is no source to provide the necessary energy for an endothermic 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

$$\text{s-X} + \mathrm{Y_{(g)}} \longrightarrow \text{s-Z}, \tag{5}$$

much similar to (3).
The $E_{\mathrm{ER}}^{\mathrm{react}}$ reaction energy is

$$E_{\mathrm{ER}}^{\mathrm{react}} := E_{\text{s-Z}} - E_{\text{s-X}} - E_{\mathrm{Y_{(g)}}}. \tag{6}$$

Especially exothermic reactions are interesting, since in that case the surface has to absorb the energy difference between reactants and product. If this is possible, the product s-Z stays intact, otherwise it will react further.
ER reactions, adsorption energies and gas phase reactions can be related. We can see that when starting from (6) and inserting (2), which yields

$$\begin{aligned} E_{\mathrm{ER}}^{\mathrm{react}} :=& E_{\text{s-Z}} - E_{\text{s-X}} - E_{\mathrm{Y_{(g)}}} \\ =& E_{\mathrm{Z}}^{\mathrm{ads}} + E_{\mathrm{S}} + E_{\mathrm{Z_{(g)}}} - \left(E_{\mathrm{X}}^{\mathrm{ads}} + E_{\mathrm{S}} + E_{\mathrm{X_{(g)}}}\right) - E_{\mathrm{Y_{(g)}}} \\ =& E_{\mathrm{Z}}^{\mathrm{ads}} - E_{\mathrm{X}}^{\mathrm{ads}} + E_{\mathrm{Z_{(g)}}} - E_{\mathrm{X_{(g)}}} - E_{\mathrm{Y_{(g)}}} \\ =& E_{\mathrm{Z}}^{\mathrm{ads}} - E_{\mathrm{X}}^{\mathrm{ads}} + E_{\mathrm{gas}}^{\mathrm{react}}. \end{aligned} \tag{7}$$

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 (7) 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.

The superscripts and subscripts on E^{int}, E^{ads} and E^{react} may be ignored if it is sufficiently clear which energy is meant.
We also want to introduce a third type of energy, the *zero-point (vibrational) energy*. We describe it for some general system A that may contain any arrangement of atoms. For all calculations we do, 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. This is usually done by including the zero-point vibrational energy of the atoms. It is computed from the *Hessian* matrix $\mathbf{H}$ of the system,

$$\mathbf{H}(\mathbf{r}_1, \ldots, \mathbf{r}_N) := \left(\frac{\partial^2 E}{\partial \mathbf{r}_i \partial \mathbf{r}_j}(\mathbf{r}_1, ..., \mathbf{r}_N)\right)_{1 \leq i,j \leq N} \tag{8}$$

by

$$E^{\mathrm{ZPE}} = E + \frac{1}{2}\mathrm{tr}\left[\mathbf{H}\right] = E + \Delta E^{\mathrm{ZPE}}, \tag{9}$$

where $\mathrm{tr}\left[\mathbf{H}\right]$ is the *trace* of the matrix $\mathbf{H}$, that is the sum of the eigenvalues of $\mathbf{H}$. In the harmonic approximation, which should be accurate for the vibrational ground state, the eigenvalues of $\mathbf{H}$ are proportional to the eigenfrequencies of the system. We call the term ΔE^{ZPE} the *zero-point (vibratinal energy) correction*.
We will always use the superscript in E^{ZPE} if we want to denote energies that are corrected with ΔE^{ZPE} in (9).

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 usually looks for solution of the *time-independent Schrödinger equation*

$$\hat{H}\left|\Psi\right\rangle = E\left|\Psi\right\rangle. \tag{10}$$

This equation holds for all non-relativistic quantum mechanical particles. Within the *Born-Oppenheimer* approximation, one can separate the dynamics of the atomic cores from the dynamics of the electrons. We will treat the cores in a more or less classical way, therefore we will for now focus on solving the Schrödinger equation for $N < \infty$ electrons, that is the Hamiltonian of our system is

$$\hat{H} = -\sum_{i=1}^{N} \frac{\hbar^2}{2m_e}\nabla_i^2 + \frac{e^2}{4\pi\varepsilon_0} \sum_{1 \leq i < j \leq N} \frac{1}{\left|\mathbf{r}_i - \mathbf{r}_j\right|} + V(\mathbf{r}^N). \tag{11}$$

$\mathbf{r}_i$ is the space 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 core geometry and movement, where $\mathbf{r}^N$ is the vector containing all electron coordinates. We will have for $K < \infty$ atomic cores

$$\begin{aligned} V(\mathbf{r}^N) = &- \sum_{A=1}^{K} \frac{\hbar^2}{2m_A} \nabla_A^2 + \frac{e^2}{4\pi\varepsilon_0} \sum_{A=1}^{K} \sum_{B=1}^{K} \frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \\ &- \frac{e^2}{4\pi\varepsilon_0} \sum_{A=1}^{K} \sum_{i=1}^{N} \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|}. \end{aligned} \tag{12}$$

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

$$V(\mathbf{r}^N) = V_{\mathrm{cc}} + V_{\mathrm{ce}}(\mathbf{r}^N) \tag{13}$$

with

$$V_{\mathrm{ce}}(\mathbf{r}^N) = \sum_{i=1}^{N} \tilde{V}(\mathbf{r}_i) = -\frac{e^2}{4\pi\varepsilon_0} \sum_{A=1}^{K} \sum_{i=1}^{N} \frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|}. \tag{14}$$

There will be more than one solution to (10), so one can construct the set of all solutions $\{|\Psi_i\rangle \,|i \in \mathbb{N}_0\}$ with corresponding energy eigenvalues E_i. $\{|\Psi_i\rangle\}$ always is the complete basis of some $\mathbb{C}$-vector space $\mathcal{H}_a$, where the subscript a denotes antisymmetry according to the *Pauli principle*

$$\begin{aligned} &\langle \mathbf{x}_1, \ldots, \mathbf{x}_l, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_N | \Psi_i \rangle \\ &= \quad \Psi_i(\mathbf{x}_1, \ldots, \mathbf{x}_l, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_N) \\ &= -\Psi_i(\mathbf{x}_1, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_l, \ldots, \mathbf{x}_N), \end{aligned} \tag{15}$$

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

If $\mathcal{H}_a \subseteq \mathcal{L}^2$, which is usually the case, the set of solutions can be chosen 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 $\hat{H}$ is not necessarily bounded from above. 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

$$\begin{aligned} E_0 &= \min_{|\Psi\rangle} \left\{ \langle \Psi | \hat{H} | \Psi \rangle \right\}, \\ |\Psi_0\rangle &= \arg\min_{|\Psi\rangle} \left\{ \langle \Psi | \hat{H} | \Psi \rangle \right\}. \end{aligned} \tag{16}$$

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 (10) has a variety of equivalent counterparts. One of them is the key to the approach of DFT. When multiplying (10) by the bra $\langle \Psi |$, one can interpret the resulting energy as a (non-linear) functional of the wavefunction by

$$E : \mathcal{H}_a \to \mathbb{R}, \qquad E[|\Psi\rangle] = \langle \Psi | \hat{H} | \Psi \rangle, \tag{17}$$

which would mean that the ground state energy can be found by minimizing the functional $E[|\Psi\rangle]$ according to (16). But so far, the minimisation 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

$$\rho(\mathbf{x}_1, ..., \mathbf{x}_N) = |\Psi_0(\mathbf{x}_1, ..., \mathbf{x}_N)|^2, \tag{18}$$

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

$$\rho(\mathbf{r}_1, s_1) = \int \mathrm{d}\mathbf{x}_2 \cdots \int \mathrm{d}\mathbf{x}_N \, |\Psi_0(\mathbf{x}_1, ..., \mathbf{x}_N)|^2, \tag{19}$$

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. Therefore, the task of solving the Schrödinger equation (10) 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. The popular ansatz by Kohn and Sham[?] is called Kohn-Sham (KS-)DFT, and the energy expression is given by

$$\begin{aligned} E[\rho] = &V_{\mathrm{cc}} + \int \mathrm{d}\mathbf{r} \, \tilde{V}(\mathbf{r}) \rho(\mathbf{r}) + \frac{e^2}{2} \int \mathrm{d}\mathbf{r} \int \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r}) \rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \\ &+ T_s[\rho] + E_{\mathrm{xc}}[\rho] \end{aligned} \tag{20}$$

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

$$\left(-\frac{\hbar^2}{2m} \nabla^2 + V_{\mathrm{eff}}(\mathbf{r}) \right) \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}) \tag{21}$$

with the *effective potential*

$$V_{\mathrm{eff}}(\mathbf{r}) = V_{\mathrm{cc}} + V(\mathbf{r}) + \frac{e^2}{2} \int \mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta \rho(\mathbf{r})}. \tag{22}$$

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

$$V_{\mathrm{xc}} = \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta \rho(\mathbf{r})}. \tag{23}$$

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

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2 . \tag{24}$$

With that, the kinetic energy functional is given by

$$T_s[\rho] = \sum_{i=1}^{N} \int \mathrm{d}\mathbf{r}\, \phi_i^*(\mathbf{r}) \left(-\frac{\hbar^2}{2m} \nabla^2 \right) \phi_i(\mathbf{r}). \tag{25}$$

Now, the remaining problem before one can find self-consistent solution to (21) with (22) and use them to minimize (20) is to find an expression for $E_{\mathrm{xc}}[\rho]$ (see below).
After one decides on some functional for E_{xc}, one needs to find *self-consistent* solution to (21). Since the KS orbitals ϕ_i define the potential V_{eff} in (22), these solutions are called the *self-consistent field* (SCF). (21) is usually solved iteratively, where one chooses some initial guess for the ϕ_i, then solves (21) 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 $V_{\mathrm{xc}}(\mathbf{r})$ does only depend on the the value of $\rho(\mathbf{r})$, which may be sensible for slowly varying densities. The more general *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[?]. These functionals are called *hybrid* functionals. Rather recently, a variety of *meta-hybrid GGA* functionals was introduced that promise to be potentially even more accurate.
We chose functionals from the GGA, Hybrid and Meta-GGA class.
GGA: BP86, BLYP, PBE and B97-D[9].
Hybrid: B3LYP, BHLYP and PBE0.
Meta-GGA: TPSS, TPSSH and PW6B95.
The functionals B97-D and PW6B95 (with additional dispersion correction, see below) and the PBE0 functional were among the most accurate functionals in a study by Anacker and Friedrich[10] on water-water interaction. They also included the meta-GGA functional M06. While M06 is indeed an accurate and fast functional, in our tests it had the strange deficiency of convergence failures when unbonded 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 equilibrium (energy minimum) geometry, but for even slightly longer separations of H and H_2O lead to oscillations in 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[11] and NWCHEM[?] calculations), and it could not be helped 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 (24) 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 $\{\chi_\mu \mid 1 \le \mu \le M\} = X \subset \mathcal{L}^2(\mathbb{R}^3)$. 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 (21) 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 may 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 does however compensate 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, def2-TZVP and def2-QZVP basis sets[?]. We also used the def2-SVPD, def2-TZVPD and def2-QZVPD basis sets with additional diffuse functions, the def2-TZVPP basis set with additional polarisation functions and finally the def2-TZVPPD basis set with both additional polarisation 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 (1) can not be calculated by the difference between the system X − 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 − Y with the mass and charge of all atoms in Y set to zero (dummy atom), and vice versa. For more than two interacting molecules, this *counterpoise correction* (CP) gets more complicated, see for example Anacker and Friedrich[10].

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 ignore the BSSE. 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. Since water-water interaction is strongly affected by non-covalent bonds, a correct account of dispersion is desirable for our approach to the system. Some of the functionals we use here are already (at least slighty) optimised to include dispersion, namely PBE0, TPSS, TPSSH and B97-D. Others were not originally designed to consider systems dependent on dispersion. But there were a variety of correction terms proposed[12,13], of which we will use Grimme's D3 correction[14]. Its most basic idea is to add a dispersion correction to the energy E_{KS} obtained from minimizing (20). This yields

$$E_{\text{DFT-D3}} = E_{\mathrm{KS}} + E_{\mathrm{disp}}, \tag{26}$$

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

$$E^{(2)} = -\sum_{AB}\sum_{n=6,8} s_n \frac{C_n^{AB}}{r_{AB}^n} f_{d,n}(r_{AB}) \tag{27}$$

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

$$f_{d,n}(r_{AB}) = \frac{1}{1+6(r_{AB}/s_{r,n}R_0^{AB})^{-\alpha_n}}. \tag{28}$$

R_0^{AB} is a cutoff-radius and α_n is called a "steepness"-parameter. They are both chosen independent 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.

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 overestimation of dispersion or still underestimate it at some circumstances. We will later present results of a small benchmark test including multiple basis sets and functionals from which we will determine the functionals with the most potential to describe water-water interaction and therefore the system we are about to study.

2.3 Molecular Mechanics

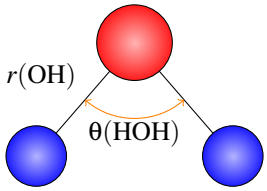

Figure 1 The geometry of TIP3P water. O is red, H is blue. Each atom coincides with a site with charge q and LJ-paramteres ε and σ.

When there is no bond formation or breaking to be expected, quantum mechanical calculations can be replaced by classical calculation, that is the position of the atomic cores is again a sharply determinable quantity and motion is governed by forces. That is, all particles move in a classical 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.*[15]. The geometry can be seen in Figure 1. The three sites of the TIP3P potential coincide with the atoms forming H_2O. Each site has a charge q_α with *Coulomb interaction*

$$V_{\alpha\beta}^{\mathrm{C}}(r_{\alpha\beta}) = \frac{1}{4\pi\varepsilon_0}\frac{q_\alpha q_\beta}{r_{\alpha\beta}}, \tag{29}$$

where α,β are the site types and $r_{\alpha\beta}$ is the distance between the sites α and β. Sites also interact by Van-der-Waals forces

which are described by a *Lennard-Jones* (LJ) potential

$$V_{\alpha\beta}^{\mathrm{LJ}}(r) = 4\varepsilon_{\alpha\beta}\left[\left(\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}}\right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}}\right)^{6}\right]. \tag{30}$$

The original TIP3P model only had LJ interaction between the oxygen atoms. In the CHARMM[?] implementation, the H atoms are LJ sites as well. Therefore, a TIP3P water molecule i moves in the potential of all other TIP3P waters by

$$V_i(\mathbf{r}_i) = \sum_{j\neq i}\sum_{\alpha,\beta}\left[V_{\alpha\beta}^{\mathrm{C}}(r_{\alpha\beta}(\mathbf{r}_i,\mathbf{r}_j)) + V_{\alpha\beta}^{\mathrm{LJ}}(r_{\alpha\beta}(\mathbf{r}_i,\mathbf{r}_j))\right]. \tag{31}$$

The first sum runs over all other water molecules and the second over all sites in both molecules. (31) only contains 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 [16].

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 [17] is nowadays only one among several schemes. We chose an approximation in which the QM part was subject to the external potential defined by the MM charges and the MM part interacted with the QM atoms by the LJ potential (30). That way, the electrostatic interaction is covered by the QM part while the MM part accounts for VdW interaction.

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)

$$\mathbf{F}(\mathbf{R}^M) = -\nabla V(\mathbf{R}^M) \tag{32}$$

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, (32) must be integrated in a discretised 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 algorithm [18] with the default tolerances to follow the minimum energy path.
Evaluating the gradient in (32) 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 (31). For the QM region, there are schemes to obtain analytical results for the gradient for a given wave function [19].

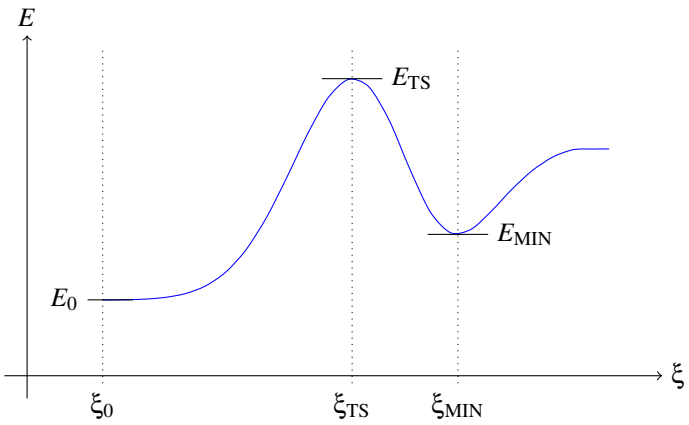

Figure 2 A cut of a PES along reaction coordinate ξ. The reaction coordinate starts at some initial ξ_0 with energy E_0 and follows a path of minimum energy. The minimum with E_{MIN} is located at ξ_{MIN}. The transition state with E_{TS} is a maximum at ξ_{TS}.

To evaluate our model, 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 geometry with initial *reaction coordinate* ξ_0 and then follows the negative gradient to a maximum (the transition state) and after that reaches a new minimum by continuing on its route along the gradient. In Figure 2, an example of such a path between reaction coordinates ξ_0 and ξ_{MIN} is plotted.
In this work, we obtain transition state with the *dimer method*[?] . It uses two images of the system with a slight offset, then rotates them to find the minimum eigenvalue of the Hessian matrix (8). It follows these "gradients" by an iterative scheme a saddle point is reached. This is the transition state.
Transition states are important for calculating tunneling rates. These are interesting for the analysis of molecule formation on ice surfaces since at temperatures as low as in the interstellar medium, 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 men-

tioned, the optimisations are carried out with DL-FIND[18], as are the dimer method and the calculation of Hessian matrices. DFT calculations are performed with TURBOMOLE[11]. Force field calculations are performed by CHARMM[20]. These programs were interfaced via ChemShell[21], which also provided the QM/MM coupling.
For the post-HF calculations in the benchmark, we used the MOLPRO[22] package.
All molecular visualisations are created with VMD[23].
The calculations were performed on the bwForCluster JUSTUS of the German federal country of Baden-Württemberg. Shared memory parallel calculations with up to 16 CPUs were performed. The main limit was the maximal computation time of six days.

3 Benchmarking

In the previous section, we did introduce the basic notions of the theoretical structure underlying our computations. We did already mention 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 the later 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[10]. We did also already name the basis sets we want to compare alongside the functionals in Section 2.2.3.
We used three different benchmarks, namely the interaction energies of an H_2O-H, an H_2O-H_2O and an $H_2O-{}^3O$ system. The reference calculations were in all cases performed with the MOLPRO[22] package at the CCSD(T)-F12A? level of theory with the VTZ-F12? 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 MRCI-F12/VTZ-F12 level of theory 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 MRCI-F12/VTZ-F12 calculation with Davidson correction? 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 require only small CP corrections of $0.05-0.18$ kJ/mol.
We performed energy minimisations 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 of optimum geometry $\mathbf{R}^M_{\mathrm{DFT}} = (\mathbf{r}^1_{\mathrm{DFT}}, \dots, \mathbf{r}^M_{\mathrm{DFT}})$ from the reference optimum geometry $\mathbf{R}^M_{\mathrm{CC}} = (\mathbf{r}^1_{\mathrm{CC}}, \dots, \mathbf{r}^M_{\mathrm{CC}})$ can be quantified by the *root mean square deviation* (RMSD). It is given by

$$\Delta_{\mathrm{RMSD}}(\mathbf{R}^M_{\mathrm{DFT}}; \mathbf{R}^M_{\mathrm{CC}}) = \sqrt{\frac{1}{M}\sum_{k=1}^{M} \left|\mathbf{r}^k_{\mathrm{DFT}} - \mathbf{r}^k_{\mathrm{CC}}\right|^2}. \tag{33}$$

We will give Δ_{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 benchmark is the accuracy with which numerical integration was carried out. In the TURBOMOLE code, we used an m4 integration grid with keyword "$scfconv 8". The results were compared to the more accurate m5 grid and "$scfconv 9", which meant no change in energy. Therefore, we recommend the slightly less accurate integration technique due to it being computationally faster.

3.1 H_2O-H Interaction

First information about H_2O-H interaction can be found in Figure 3 and Figure 4, where we compared functionals (with and without dispersion correction, respectively) in different basis sets. We compared energy values of the form

$$\Delta E^{\mathrm{int}} = E^{\mathrm{int}}_{\mathrm{DFT}} - E^{\mathrm{int}}_{\mathrm{CC}}. \tag{34}$$

$E^{\mathrm{int}}_{\mathrm{DFT}}$ stands for the interaction energy (1) of the system computed with a certain DFT functional and basis set, and $E^{\mathrm{int}}_{\mathrm{CC}}$ is the CP corrected interaction energy with CCSD(T)-F12A/VTZ-F12. For H_2O-H, we found $E^{\mathrm{int}}_{\mathrm{CC}} = -0.40$ kJ/mol, so we have an only weakly bonded system. The CP correction contributes $+0.05$ kJ/mol to the energy. Before discussing different functionals and the dispersion corrections, we can have a look at the differences between the four

*In the $H_2O-{}^3O$ case, a first minimum was calculated at the MRCI-F12/VTZ-F12? level of theory, where it was found that the wavefunction can be well described by a single-reference method. From there, the CCSD(T)-F12A minimisation was carried out.

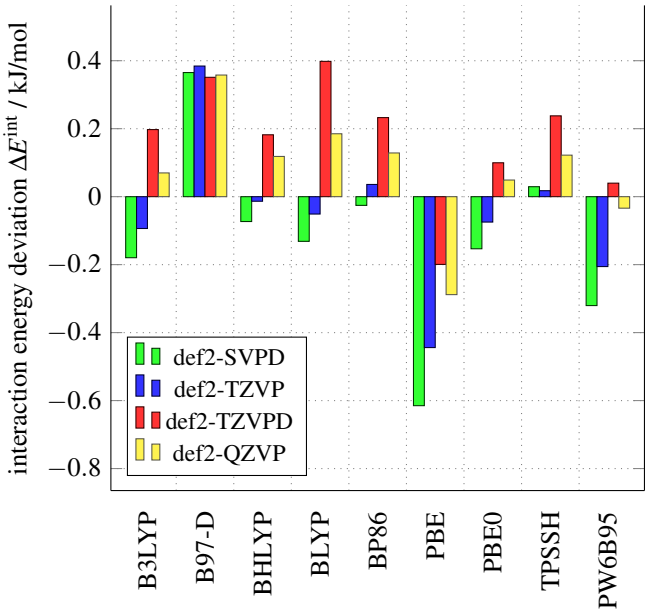

Figure 3 The H_2O-H benchmark results for different basis sets without dispersion correction. Energies are plotted as differences to the reference energy from CCSD(T)-F12A/VTZ-F12 calculations, see (34).

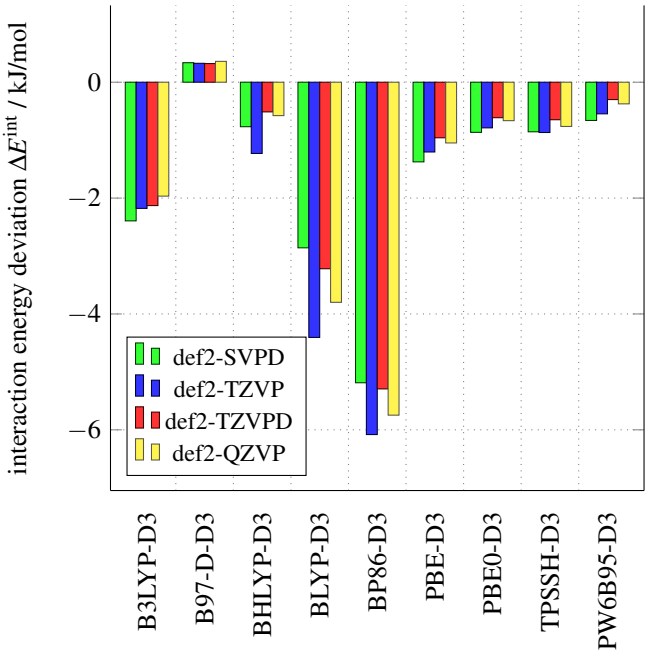

Figure 4 The H_2O-H benchmark results for different basis sets including dispersion correction. Energies are plotted as differences to the reference energy from CCSD(T)-F12A/VTZ-F12 calculations, see (34).

basis sets used.Figure 3 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. Only the BHLYP, PBE0 and PW6B95 functional had errors between -2 and -5 kJ/mol with def2-SVP. 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 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 both figures. With additional dispersion, it tends to yield similar accuracy as def2-TZVP, while usually being inferior to def2-TZVPD for most functionals. Still, without dispersion correction, its accuracy is between def2-TZVP and def2-TZVPD and only rarely better than both. def2-QZVP takes approximately 1.5 times as long to complete a single SCF iteration as def2-TZVPD. From a first glance at the "better" functionals in this case, one could therefore consider the def2-TZVP and def2-TZVPD basis sets as good compromises between accuracy and computational cost.

The first general result on functionals is that 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 3 to Figure 4. 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.

We will therefore mainly focus on the data in Figure 3. The results here tend to be very accurate. In fact, the deviations are within the error bonds of a CCSD(T)-F12A calculation, so it is hard to favor a functional of the given choice. Still, at least two 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 of 3.34 Å.

The opposite problem happens for B97-D and B97-D-D3 optimisations. 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 3 and 4, it is a qualitative mistake. The separation between radical H and water O is around 5.80 Å, which is close to describing two isolated systems.

Comparing RMSD data of different functionals yields similar results as the energy comparison. Functionals without the D3 correction predict very accurate geometries with $\Delta_{\text{RMSD}} <$

0.01 Å for all but TPSS and B97-D. Despite worse energy results, many D3 corrected versions still find accurate optimum geometries. Exclusions are B3LYP-D3, BLYP-D3, BP86-D3, which do find optima similar to the already mentioned TPSS optimum, with the H radical too close to the water O. The BHLYP functional generally yields good geometries, only for the def2-TZVP basis set it also predicts 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. Only the def2-SVP basis set and the def2-TZVPP basis set often lead to wrong, i.e. too attractive, geometries.

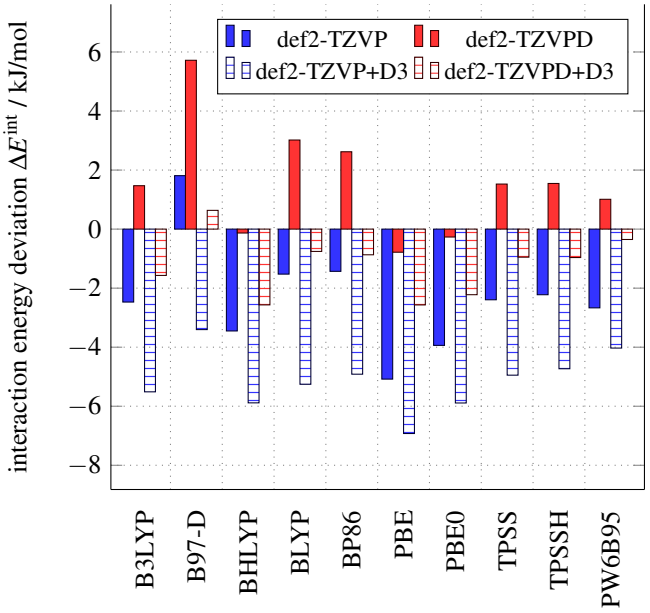

Figure 5 The H_2O-H_2O benchmark results for the def2-TZVP and def2-TZVPD basis sets. The results with dashed bars include dispersion corrections. Energies are plotted as differences to the reference energy from CCSD(T)-F12A/VTZ-F12 calculations, see (34).

3.2 H_2O-H_2O Interaction

Again, we compare energies to the reference energy as in (34). The reference energy for the H_2O-H_2O dimer is at $E_{CC}^{int} = -20.80$ kJ/mol, where the CP correction contributes $+0.18$ kJ/mol. In Figure 5, 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 does prevent errors as for the B97-D functional in the previous section, so all functionals expect attractive interaction. Inclusion of the D3 dispersion correction does again predict 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. This was also very accurate for H_2O-H interaction, so this benchmark confirms it as a good choice of 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.

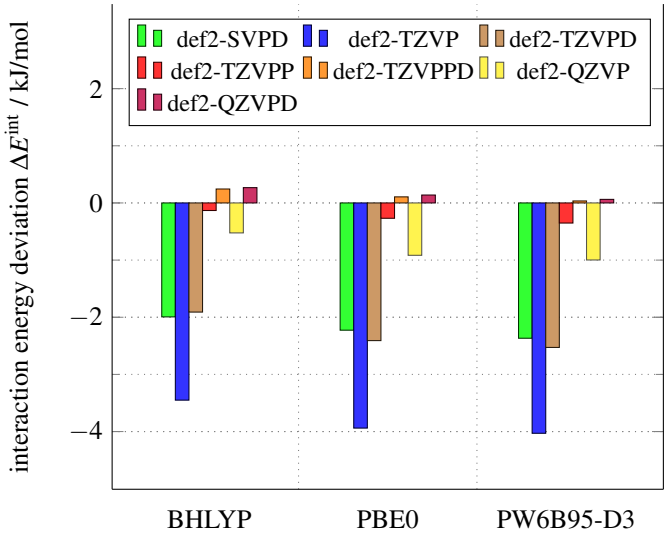

Figure 6 The H_2O-H_2O benchmark results for BHLYP, PBE0 and PW6B95-D3 and all basis sets used in the study (excluding def2-SVP). Energies are plotted as differences to the reference energy from CCSD(T)-F12A/VTZ-F12 calculations, see (34).

We use these three most promising functionals to investigate the difference between basis sets in Figure 6. Only the def2-SVP basis set is not considered, since it yields errors of between -12 kJ/mol and -15 kJ/mol for all three functionals. The figure shows that while increasing the ζ-order of the basis set does indeed improve 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 core. The inclusion of polarisation functions as in def2-TZVPP does also affect the results positively, however not to the extent of additional diffuse functions. The def2-TZVPPD basis set is again an improvement to the def2-TZVPP basis set, it is also stronger

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 methods of 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.

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 for the whole QM part. We will therefore use a hybrid system 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 5.1. 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 5 shows that good compromises between the two basis sets are 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 better for def2-TZVP. All other functionals are not accurate enough for def2-TZVP or def2-TZVPD.

For H_2O-H_2O, the RMSD values are entirely unproblematic for any basis set beyond def2-SVP. The worst RMSD is reached by B97-D-D3/def2-TZVPD, it is of 0.076 Å, which is just a slight displacement of H atoms. Especially B3LYP and PW6B95 both with and without dispersion corrections predict RMSDs of less than 0.008 Å to the reference geometry, BHLYP is rougly at 0.010 Å and 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.

3.3 $H_2O-{}^3O$ Interaction

It already became clear at the beginning of this section that $H_2O-{}^3O$ interaction is more tricky for post-HF methods because some geometries in the optimisation 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 CCSD(T)-F12A method. Only the CP correction is taken from an MRCI-F12 calculation, but it is as small as 0.10 kJ/mol and therefore unlikely to deteriorate our results. The CP corrected interaction energy is $E_{CC}^{int} = -6.72$ kJ/mol.

In the last two test systems, six functionals show a good overall performance. Figure 7 depicts the energy data of the $H_2O-{}^3O$ benchmark for these functionals. In contrast to the previous two

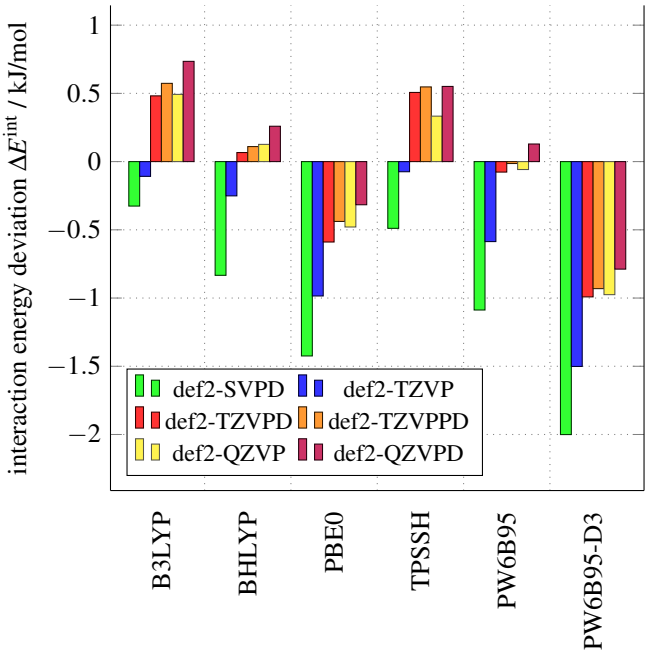

Figure 7 The $H_2O-{}^3O$ benchmark results for the def2-TZVP and def2-TZVPD basis sets. The results with dashed bars include dispersion corrections. Energies are plotted as differences to the reference energy from CCSD(T)-F12A/VTZ-F12 calculations, see (34).

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.

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.

Of the functionals used in Figure 7, none has an RMSD greater than 0.01 Å. They do therefore all yield good approximations to the optimum geometry. This does not apply to all functionals in this benchmark, many of those excluded here have RMSD values greater than 0.5 Å.

3.4 Conclusions

We have already 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 among the most accurate. Since the adsorption process can easily be restricted to by close to only atoms described by def2-TZVPD, a functional's results for the def2-TZVPD ba-

sis set have to be the most important ones, while results in the def2-TZVP basis should still be reasonable but not necessarily too accurate.

Under these conditions, we find that PBE0, BHLYP and PW6B95-D3 are excellent 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. PW6B95 is even better for both H_2O-H and $H_2O-{}^3O$, but it is considerably worse for water-water interaction, therefore the PW6B95-D3 functional seems preferable in a highly water-dominated system, especially since the (reference) water-water interaction energy is of -20.80 kJ/mol.

The functionals B3LYP and TPSSH are also good options for functionals, despite not being recommended by Anacker and Friedrich[10]. They are however slightly inferior to BHLYP and PBE0.

For everything that follows, we will therefore mostly use results obtained with BHLYP, PBE0, PW6B95-D3, B3LYP or TPSSH. Especially on adsorption under interstellar condition, there is not much experimental data available, therefore our only indication of good results may be agreement between the different functionals that seemed to yield credible results by the standards of this benchmark.

4 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.

4.1 Pure Molecule Energy Data

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

$$E_X^0 = E_X - \frac{M_H}{2}E_{H_2} - \frac{M_O}{2}E_{^3O_2}. \tag{35}$$

$^3O_2$ is the triplet oxygen. The spin multiplicity of a molecule will only be given if the choice is not obvious. This is the case for the pure oxygen species ^1O, ^3O, $^1O_2$ and $^3O_2$. All other species have singlet or dublet spin.

The value is of only minor physical importance, but if two electron strutcture 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. Note that this means that we only have the ZPE corrected value of the energy minimum of the not ZPE corrected PES. But evaluating the ZPE corrected PES is computationally expensive and will get exceedingly expensive when the system size enlarges to the water surface later on.

An energy adjusted by (35) 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 inaccurate 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 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 exothermic formation energies than the reference calculations.

As we can see in Table 1, the approximative power of the DFT functionals varies strongly among the molecular species. Especially the case of ^1O and $^1O_2$ seems to be problematic, but the agreement is also not very good for H_2O_2. 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 func-

Table 1 Energies for DFT functionals, according to (35). All values are ZPE corrected. 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	

tionals 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.

4.2 Reactions

We already described gas phase reaction in Section 2.1 by (3). For simplicity, we drop the (g) subscript.

We calculated DFT energies in the def2-TZVPD basis set with the B3LYP, BHLYP, PBE0, TPSSH, PW6B95 and PW6B95-D3 functionals. We did also calculate CCSD(T)-F12A/VTZ-F12 energy data. All results were supplemented by a ZPE correction at the optimum geometry obtained by DL-FIND.

The sudied reactions are

$$OH + H \longrightarrow H_2O, \tag{36a}$$
$$H_2O + {}^1O \longrightarrow H_2O_2, \tag{36b}$$
$$OH + OH \longrightarrow H_2O_2, \tag{36c}$$
$$HO_2 + H \longrightarrow H_2O_2, \tag{36d}$$
$$OH + {}^3O \longrightarrow HO_2, \tag{36e}$$
$${}^3O_2 + H \longrightarrow HO_2. \tag{36f}$$

The ^1O in (36b) is in an excited state compared to the ground state ^3O, with a excitation energy difference of 206.50 kJ/mol according to the most precise result in Table 1. The reaction with the radical OH to the more stable H_2O 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 more endothermic reactions than the other functionals. This leads to a decrease in agreement with the reference energy. Only for the case of reaction (36b), this is beneficial. This reaction does involve the ^1O atom, for which BHLYP described the formation energy most accurately. The opposite is the case for reaction (36c), 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. It seems to be the case that if Table 1 implies that a functional describes a molecular species well or badly, a reaction energy containing this molecular species will also be accurate or inaccurate. This can be seen as an indication that the descprition of the accuracy of the functionals for the different molecular species in Table 1 is not too strongly contaminated by the functional's respective errors in the description of H_2 and $^3O_2$.

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 analyse 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

Table 2 Reaction energies for DFT functionals. E^{ZPE} and E are data with and without ZPE correction, respectively. We also give the value of the ZPE correction ΔE^{ZPE}. The closest energy to CCSD(T)-F12A is highlighted in boldface. All energies in kJ/mol.

		B3LYP	BHLYP	PBE0	TPSSH	PW6B95	CCSD(T)-F12A
$OH + H \longrightarrow H_2O$	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
$E^{\mathrm{ZPE}}_{\mathrm{DFT}} - E^{\mathrm{ZPE}}_{\mathrm{CCSD(T)\text{-}F12A}}$	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	

make.

While the results here do not clearly identify functionals for which the agreement between CCSD(T)-F12A reaction energies and the functional's reaction energies should be particularly good, they do show that it is in principle possible to describe the reactions (36) 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.

5 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.

5.1 The Surface Model

5.1.1 The Fletcher Surface

Barnes[24] established a first model of water at temperatures as low as 90 K in 1929. He could not locate the hydrogen atoms due to their low X-ray scattering power[25], 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, crystalline water at very low temperatures may very well have a more regular structure. We follow an approach presented by Fletcher[26] that minimizes surface free energy at low temperatures.

It is given in Figure 8. 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 hydro-

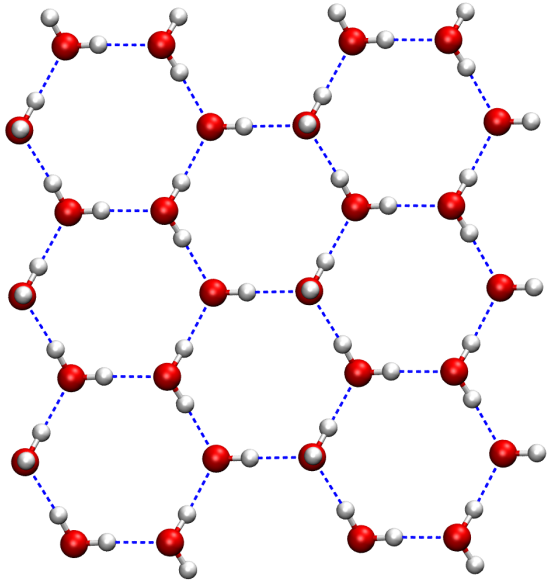

Figure 8 The Fletcher surface. Hexagonal O (red) structure with vertical lines of identical H (white) orientation. H-bonds are indicated by dashed blue lines. The O atoms are not all in the same plane.

gen bonds, H atoms can be placed close to one of the two O atoms taking part in the bond. This can lead 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 8, 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 structure described above is extended to the full system.

Such a crystalline surface must be considered as an idealisation. For once, 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 surface, probably both in O position as well as in H orientation. Research in that respect was presented by Cabrera Sanfelix *et al.*[8] 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 only the case in the cold interstellar medium. 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 8 is wihtin 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 an ideal surface layer. We will see this effect in Section 5.2.

Instead of choosing a cuboid subdomain of the Fletcher surface – which would be more adequate to the surface's translational symmetry – we will consider a hemisphere as depicted in Figure 9. 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.

5.1.2 The QM/MM Region

Of a fully-ordered water I_h crystal with a Fletcher surface, we take a hemispherical cut of radius 25 Å centered on the center of a hexagonal ring of water molecules. We choose a normal basal surface as in Figure 8, not a prismatic surface. 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 9. 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 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. 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 QMmolecules. 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 optimisations we later carry out. The rest of the MM molecules is fixed in position (frozen molecules) to yield boundary conditions. The 10 Å 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

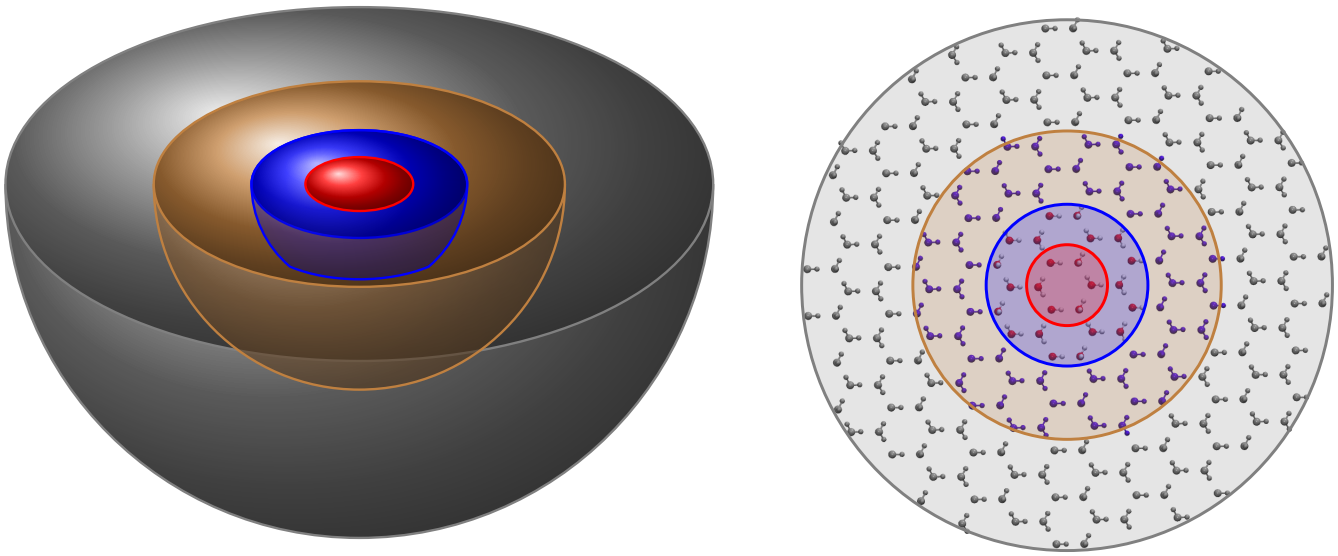

Figure 9 The QM/MM decomposition of the Fletcher surface. The picture to the right shows the surface from above, the picture to the left is a schematic representation. The gray and the brown area make up the MM region, where water molecules are described by the TIP3P potential. In the gray region, they 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.

include these additional MM atoms, and if necessary they can be turned into further active atoms. There is a total of 1121 MM molecules, of which 231 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 a in princple 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 appropriate value while maintaining the orientation of the water dipole moment, that is the vector $\mathbf{r}_{\mathrm{O-H1}} + \mathbf{r}_{\mathrm{O-H2}}$ stays the same for each molecule. This means symmetrical bending of all bond angles with respect to the dipole moment.

5.2 First Geometry Optimisations

To verify the agreement between the theoretically assumed Fletcher surface and the methods we want to use on it, we calculated RMSD values (33) between the ideal Fletcher surface and the surface it changes into after QM/MM optimisation.

Table 3 RMSD values for the all classical calculation with TIP3P and QM/MM calculations with different functionals. The results are ordered with increasing RMSD.

TIP3P	0.179 Å		
BHLYP	0.182 Å	PW6B95-D3	0.187 Å
B3LYP	0.183 Å	PBE0	0.193 Å
PW6B95	0.183 Å	TPSSH	0.191 Å

Note that the factor $\frac{1}{M}$ in (33) 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 optimisation with TIP3P. Generally, the results can be considered 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 **??**. This effect is not sinking process leads to an RMSD of order < 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, 0.035 Å and less. Even those functionals we excluded in the benchmark yield optimum geometries within boundaries of an RMSD of 0.063 Å. Between the functionals of Table 3 and the TIP3P

potential, there is an RMSD of roughly 0.1 Å. The source of this deviation is that the DFT water does not have the θ(HOH) bond angle that TIP3P water has to satisfy.

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

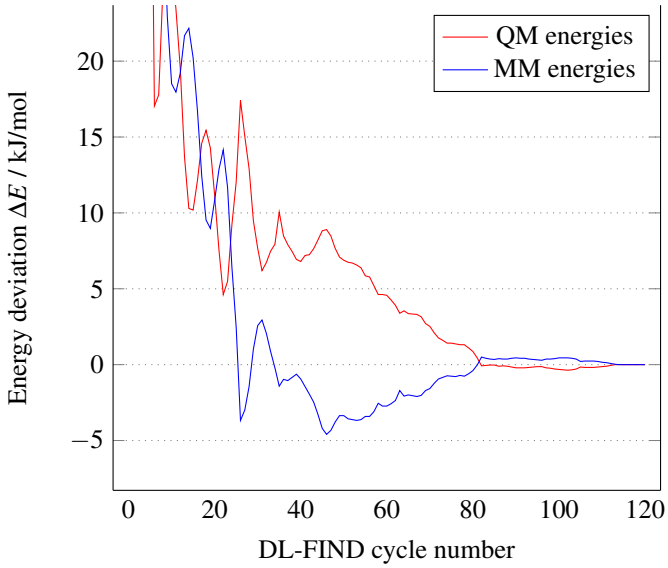

Figure 10 The QM/MM energy changes during a typical optimisation. The energies are deviations from their value at the optimum geometry. Functional: BHLYP.

We can also compare the behaviour of the QM and MM part during the optimisation process starting at the initial crystalline geometry. Figure 10 gives a typical example with the BHLYP functional. Both QM and MM energies are differences to their values at the energy minimum. At the fully 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. This is again an implication that the actual optimum geometry has to be tighter bound and that the effect is strongest for the QM region, although the high energies in the QM region must also be due to the wrong geometry. After the first few iterations, one can see that the energy changes in both domains are of the same order of magnitude. 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 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.

5.3 ZPE Corrections

Now we have established a model that should be able to describe at least ^{3}O, H and H_2O adsorption to some satisfaction. We can compute optimum geometries and their energies for different adsorbates and gain the adsorption energy according to (2). Before we do so, it is necessary to give a few words on the ZPE correction in this case.

Due to the QM/MM coupling, no analytical Hessian matrices can be computed. There is no way around computing finite differences, so two atomic diplacements per coordinate per atom, leading to six displacements for each atom. Computations for more complicated systems, especially s-HO_2 calculations with their dublet 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.

Obviously, the MM atoms do not need to be displaced. If their electrons are not described by quantum mechanics then neither should their atomic cores get a quantum mechanical correction. 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 do 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 less. However, the question if the cutoff for this 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 $54+3M$ displacements (plus one initial geometry). This is much less than the original $324+3M$ 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.

5.4 Adsorption

Of the molecules relevant in Section 4, we examine adsorption energies for all species but ^{1}O, 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 (36b).

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 general variability of the ZPE correction for dif-

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

ferent functionals.

In general, the value of the corrective term does vary 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. 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. And since there is no knowing whether or not ZPE/all is accurate either, it seems alright to use the correction of the smaller system while accepting error bars of ±2 kJ/mol.

Comparing the corrective 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 4: 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 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. Due to the computational cost of a single ZPE correction, an optimization on a more sophisticated PES will be very difficult to achieve.

In any case, it is safe to say that the adsorbates s-H, s-H_2 and s-$^3O_2$ are only very weakly bound to the surface.

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 included 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[28,30]. Other data include H_2O adsorption on an Au surface covered in crystalline water ice[27] and $^3O_2$ adsorption on porous amorphous water ice[29]. 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 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. Due to its usual deviation from the other functionals, PW6B95-D3 is not included into the averaging. For s-$^3O_2$, PW6B95 is aditionally excluded because of the very unphysical value of 26.68 kJ/mol. We will use this average energy when discussing reactions.

Where 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 are 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 input values, however we were unable to find such an error. We must therefore consider PW6B95 to be 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 $^3O_2$, too, is probably only weakly bound to the surface 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

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_{des}
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, -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[†]	−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 [27]

[b] Adsorption on an amorphous silicate. He, Vidali 2014 [28]

[c] Adsorption on porous amorphous water ice and amorphous silicate, respectively. He *et al.* 2015 [29]

[d] Adsorption on an amorphous silicate. He, Jing, Vidali 2014 [30]

[†] Average excluding both PW and PW-D3.

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. Of course, some of these possibly erroneous results could stem from DL-FIND getting stuck in the wrong minima.

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 is far off while the experiment yielding the results should be the one closest to our model. The agreement would be better if we neglected $\Delta E^{\mathrm{all}} = 15.98$, with extremely good results for B3LYP (cf. Table 4). It is even the case that the experimental values mostly lie somewhere between the ZPE corrected energy value and the uncorrected one for many functionals. But this should not question the reliability of the ZPE correction, since it represents a physical necessity and is here carried out close to the best precision available. We should rather question sources of error like 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.* [8] for the case of adsorption of a water dimer on graphite. We will come back to possible improvements to the model in the final section.

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, this hope becomes more plausible. However, due to their rather bad performance for O species, the PW6B95 functionals have lost some credibility.

5.5 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 surrounding gas approaches an adsorbed molecule to form a new molecular species, as described in Section 2.1. 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.

We use Eley-Rideal versions of the gas phase reaction (36), namely

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 average values. 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 $+\mathrm{H_{(g)}}$	$\longrightarrow$ s-$\mathrm{H_2O}$	−476.22	−458.03	−472.39	−470.76	**−480.13**	**−480.92**	−492.71
s-$\mathrm{H_2O}$ $+{}^1\mathrm{O_{(g)}}$	$\longrightarrow$ s-$\mathrm{H_2O_2}$	−402.54	**−357.85**	−427.27	−426.38	−405.24	−407.50	−344.17
s-OH $+\mathrm{OH_{(g)}}$	$\longrightarrow$ s-$\mathrm{H_2O_2}$	−185.12	−137.04	−194.02	−187.84	**−197.12**	**−200.16**	−205.33
s-$\mathrm{HO_2}$ $+\mathrm{H_{(g)}}$	$\longrightarrow$ s-$\mathrm{H_2O_2}$	**−317.22**	−318.69	−310.31	−309.82	−320.89	−323.99	−343.39
s-OH $+{}^3\mathrm{O_{(g)}}$	$\longrightarrow$ s-$\mathrm{HO_2}$	−294.05	−219.54	−300.54	**−295.71**	−296.15	−296.10	−287.00
s-^{3}O $+\mathrm{OH_{(g)}}$	$\longrightarrow$ s-$\mathrm{HO_2}$	−311.96	−248.31	−325.41	**−313.50**	−316.45	−318.00	−308.95
s-$^3\mathrm{O_2}$ $+\mathrm{H_{(g)}}$	$\longrightarrow$ s-$\mathrm{HO_2}$	−248.05	−243.27	−241.72	**−250.31**	−268.95	−256.14	−254.74
$E_{\mathrm{DFT}} - E^{\mathrm{ads}}_{\mathrm{avg}}$	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	

$$\text{s-OH} + \mathrm{H_{(g)}} \longrightarrow \text{s-}\mathrm{H_2O}, \tag{37a}$$
$$\text{s-}\mathrm{H_2O} + {}^1\mathrm{O_{(g)}} \longrightarrow \text{s-}\mathrm{H_2O_2}, \tag{37b}$$
$$\text{s-OH} + \mathrm{OH_{(g)}} \longrightarrow \text{s-}\mathrm{H_2O_2}, \tag{37c}$$
$$\text{s-}\mathrm{HO_2} + \mathrm{H_{(g)}} \longrightarrow \text{s-}\mathrm{H_2O_2}, \tag{37d}$$
$$\text{s-OH} + {}^3\mathrm{O_{(g)}} \longrightarrow \text{s-}\mathrm{HO_2}, \tag{37e}$$
$$\text{s-}^3\mathrm{O} + \mathrm{OH_{(g)}} \longrightarrow \text{s-}\mathrm{HO_2}, \tag{37f}$$
$$\text{s-}^3\mathrm{O_2} + \mathrm{H_{(g)}} \longrightarrow \text{s-}\mathrm{HO_2}. \tag{37g}$$

Note that both (37e) and (37f) are Eley-Rideal versions of (36e), but with different reactant adsorbates.

The reaction energies for (37) are given in Table 6. All reaction energies are computed according to (6).

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 corrective term demanded by the individual functional or always the one demanded by B3LYP. We chose to use 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 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, 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

$$E^{\mathrm{react}}_{\mathrm{hybrid}} = \Delta E^{\mathrm{ads}}_{\mathrm{avg}} + E^{\mathrm{react}}_{\mathrm{CCSD(T)\text{-}F12A}}. \tag{38}$$

For a reaction s-X $+\,\mathrm{Y_{(g)}} \longrightarrow$ s-Z, $\Delta E^{\mathrm{ads}}_{\mathrm{avg}}$ is

$$\Delta E^{\mathrm{ads}}_{\mathrm{avg}} = E^{\mathrm{ads}}_{\mathrm{avg,Z}} - E^{\mathrm{ads}}_{\mathrm{avg,X}}, \tag{39}$$

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

We do again 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 $E^{\mathrm{react}}_{\mathrm{hybrid}}$. But this is not really a matter of accuracy, since the change in difference to the reference value does only depend on the adsorption energy average $E^{\mathrm{ads}}_{\mathrm{avg}}$, that is a functional can get closer to $E^{\mathrm{react}}_{\mathrm{hybrid}}$ by over- or underestimating its respective ΔE^{ads} as in (39). This is merely a cancellation of error. Therefore, the seemingly in-

creased accuracy particularly in PW6B95-D3 results are consequences of the strong deviations from the average documented in the previous subsection.
There is therefore nothing new to say about the approximative power of DFT functionals to the recommended value. 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}. We will not give the values here, but from the results so far we would recommend B3LYP to be this functional, since the ZPE correction was already determined by B3LYP and 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 more endothermic is (37d), 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 (37g), 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 also see a big energy difference between the related reactions (37e) and (37f). (37e) is $+21.95$ kJ/mol more endothermic than (37f). This, too, is the difference in adsorption energy between the strongly bound OH and the weaker bound ^{3}O.

6 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, diffusion or tunneling rates. These could be further used as simulation input. We will not go into detail here but give a small outlook to what other investigations are possible on the surface. The research in this section is not meant to be complete in any way.

6.1 Binding Sites

In the previous calculations, we did not investigate how strongly the optimum geometry for an adsorption is dependent 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 orientational 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 does only depend on the position of ^{3}O in space (and the corresponding surface deformation, which we neglect here). The average binding energy for all functionals was $E^{\text{ads,ZPE}}_{\text{avg}} = -12.35$ kJ/mol. The ZPE correction was $\Delta E^{\text{all}}_{\text{B3LYP}} = 2.69$ kJ/mol, which means an uncorrected binding energy of $E^{\text{ads}}_{\text{avg}} = -15.04$ kJ/mol.

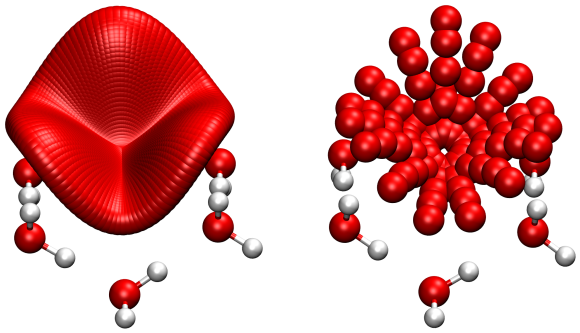

Figure 11 The initial guesses for binding sites. The left is a visualization of the function used, the right is the actual sites used as initial geometry. The molecules below are the central ring.

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 transform (x,z) in polar coordinates with $x = r\cos\phi$, $z = r\sin\phi$, $0 \le r \le r_{\max}$, $0 \le \phi < 2\pi$. 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 position $(r_{\max}, y < 0, 0)^\intercal$. In Figure 11, 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

$$y(r,\phi) = y_0 + \Big(A - B\cos(3\phi)\Big)\sin^2\left(\frac{r}{r_{\max}}\frac{\pi}{2}\right). \qquad (40)$$

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 aequidistantly distributed points according to $r_k = r_{\min} + k(r_{\max} - r_{\min})/(N-1)$ and $\phi_l = 2\pi l/M$, $0 \le k < N$ and $0 \le l < M$, with $r_{\min} = 0.4$ Å and point numbers $N = 5$, $M = 15$, leading to a

total of 75 initial geometries. The aequidistant distrubtion 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 slightly distorted.

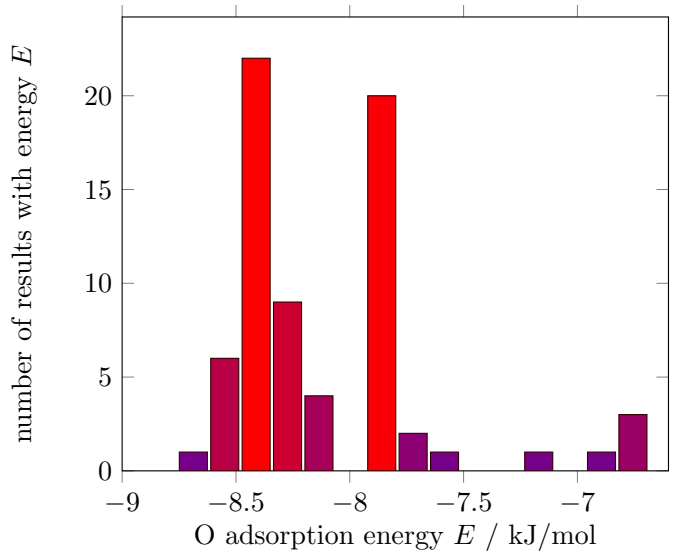

Figure 12 Population of optimum geometry energies. Of the 75 initial guesses according to (40), 71 did converge to a minimum. One geometry at $E = -5.01$ kJ/mol is not included. Energies in kJ/mol.

We then used the DL-FIND algorithm to find new energy minima. The resulting energies are plotted in Figure 12. 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 strongly attractive (in a numerical sense) energy minima to which more than half of the initial geometries are drawn.
We recommend the binding site $(0,2,0)^\intercal$ despite it being not the most attractive. The reason for that choice is that all other binding sites are located more or less above the central ring. This means that for simulations, they are not clearly within one ring or the other.
However, this becomes interesting when we consider hopping of adsorbates. In Figure 11, 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 central ring can hop into another optimum, we may call it a transmission optimum, and continue hopping into other rings from there. This mechanism should be favourable for hopping ratios, but then there should be a preferred direction of hopping on a Fletcher surface.
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.

6.2 Transition State

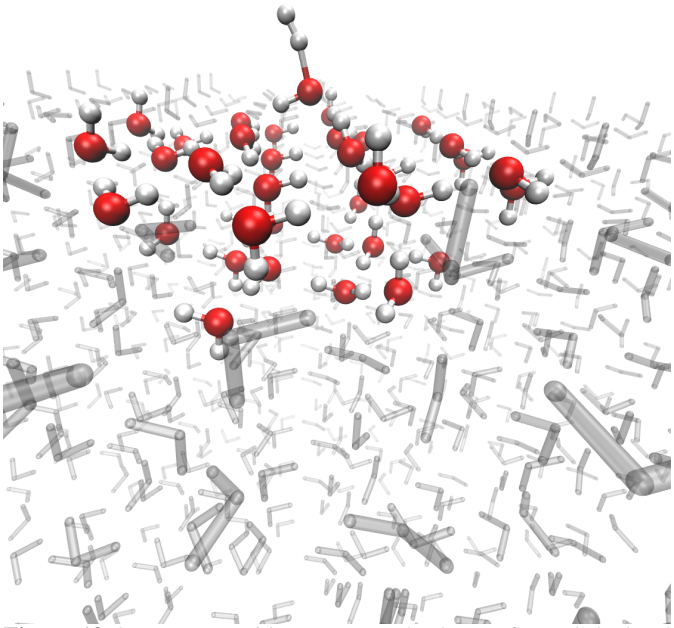

Figure 13 $OH + H_2$ transition state on Fletcher surface. Calculated with B3LYP. Transparent atoms are in the MM region, red (O) an white (H) atoms in the QM region.

We did already note that transition states are important to reaction kinetics. The energy difference between a system at the transition state and a system at the optimum geometry yield the activation energy, a key ingedrient to the activation energy.
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 proposed by Thielens and Hagen[?].
We did not use any coupled-cluster references in this test but started dircetly with DFT calculations. Since it was closest to the average value for adsorption energies in the previous section, we used the B3LYP functional. 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 an optimization 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. Reasonably close does mean 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 14, which does however not give a good impression of surface deformation. Figure 13 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 a 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, we find very

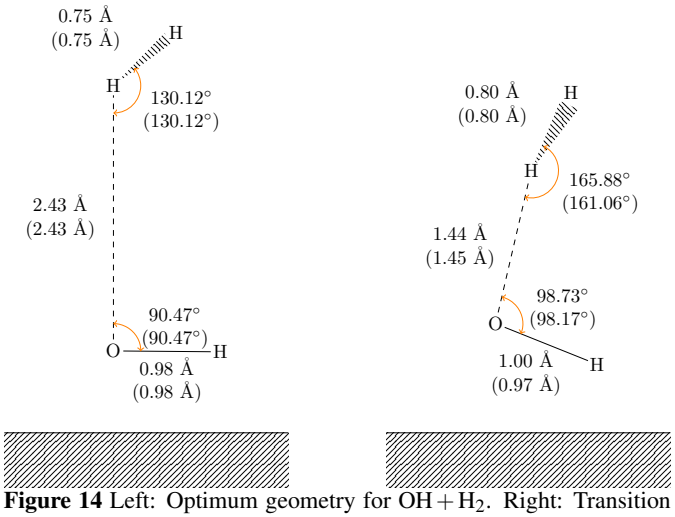

Figure 14 Left: Optimum geometry for $OH + H_2$. Right: Transition state for $OH + H_2$. Values are given for the adsorbate, values in brackets are for the gas phase. The dihedral angles for the optimum geometry are 33.58° (33.58°) and for the transition state 2.58° (1.32°).

good agreement in both optimum as well as transition state geometry. Both transition states are predicted nearly planar, which is not true for the optimum geometry. Bond lengths do not change greatly between transition state and optimum geometry, but the separation between the two molecules decreases.

7 Summary and Conclusion

We have designed a crystalline water I_h Fletcher surface model for QM/MM treatment. 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 EL type based on gas phase reaction eneriges 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 without too much scientific rigor, we considered the ability of the model to analyse different bind-

ing 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 analyse 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 idealisation, 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. 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.

Acknowledgment

I want to thank Professor J. Kaestner for his support and Mr. Jan Meisner for his introduction to the topic, his helpful suggestions and his company along the way.

References

[1] R. J. Trumpler, "ABSORPTION OF LIGHT IN THE GALACTIC SYSTEM," *Publications of the Astronomical Society of the Pacific*, vol. 42, no. 248, pp. 214–227, 1930.

[2] B. Draine, "Interstellar Dust Grains," *Annual Review of Astronomy and Astrophysics*, vol. 41, pp. 241–289, sep 2003.

[3] K. Merrill, R. Russell, and B. Soifer, "Infrared observations of ices and silicates in molecular clouds," *The Astrophysical Journal*, vol. 207, pp. 763–769, 1976.

[4] E. L. Gibb, D. C. B. Whittet, A. C. A. Boogert, and A. G. G. M. Tielens, "Interstellar Ice: The Infrared Space Observatory Legacy," *The Astrophysical Journal Supplement Series*, vol. 151, pp. 35–73, mar 2004.

[5] K. M. Ferrière, "The interstellar environment of our galaxy," *Reviews of Modern Physics*, vol. 73, pp. 1031–1066, dec 2001.

[6] A. G. G. M. Tielens and W. Hagen, "Model Calculations of the Molecular Composition of Interstellar Grain Mantles," *Astronomy and Astrophysics*, vol. 114, pp. 245–260, 1982.

[7] C. Romanzin, S. Ioppolo, H. M. Cuppen, E. F. van Dishoeck, and H. Linnartz, "Water formation by surface O3 hydrogenation," *The Journal of Chemical Physics*, vol. 134, no. 8, p. 084504, 2011.

[8] P. Cabrera Sanfelix, S. Holloway, K. Kolasinski, and G. Darling, "The structure of water on the (0001) surface of graphite," *Surface Science*, vol. 532-535, pp. 166–172, Jun 2003.

[9] S. Grimme, "Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction," *Journal of Computational Chemistry*, vol. 27, no. 15, pp. 1787–1799, 2006.

[10] T. Anacker and J. Friedrich, "New Accurate Benchmark Energies for Large Water Clusters: DFT Is Better Than Expected," *Journal of Computational Chemistry*, vol. 35, pp. 634–643, Jan 2014.

[11] "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`.."

[12] A. D. Becke and E. R. Johnson, "Exchange-hole dipole moment and the dispersion interaction revisited," *J. Chem. Phys.*, vol. 127, no. 15, p. 154108, 2007.

[13] 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," *J. Chem. Phys.*, vol. 132, no. 15, p. 154104, 2010.

[14] S. Grimme, S. Ehrlich, and L. Goerigk, "Effect of the damping function in dispersion corrected density functional theory," *Journal of Computational Chemistry*, vol. 32, pp. 1456–1465, Mar 2011.

[15] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," *J. Chem. Phys.*, vol. 79, no. 2, p. 926, 1983.

[16] 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," *J. Phys. Chem. B*, vol. 102, pp. 3586–3616, Apr 1998.

[17] 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*, vol. 103, pp. 227–249, May 1976.

[18] 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*, vol. 113, pp. 11856–11865, Oct 2009.

[19] K. R. Brorsen, F. Zahariev, H. Nakata, D. G. Fedorov, and M. S. Gordon, "Analytic Gradient for Density Functional Theory Based on the Fragment Molecular Orbital Method," *Journal of Chemical Theory and Computation*, vol. 10, pp. 5297–5307, Dec 2014.

[20] 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*, vol. 30, pp. 1545–1614, Jul 2009.

[21] "Chemshell, a computational chemistry shell, see www.chemshell.org."

[22] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, *et al.*, "Molpro, version 2012.1, a package of ab initio programs," 2012. see.

[23] W. Humphrey, A. Dalke, and K. Schulten, "VMD – Visual Molecular Dynamics," *Journal of Molecular Graphics*, vol. 14, pp. 33–38, 1996.

[24] W. H. Barnes, "The Crystal Structure of Ice between 0 degrees C. and - 183 degrees C," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 125, no. 799, pp. 670–693, 1929.

[25] N. Fletcher, "The freezing of water," *Science Progress (Oxford)*, vol. 54, pp. 227–241, 1966.

[26] N. H. Fletcher, "Reconstruction of ice crystal surfaces at low temperatures," *Philosophical Magazine Part B*, vol. 66, pp. 109–115, Jul 1992.

[27] 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*, vol. 327, no. 4, pp. 1165–1172, 2001.

[28] J. He and G. Vidali, "Experiments of Water Formation on Warm Silicates," *The Astrophysical Journal*, vol. 788, p. 50, May 2014.

[29] 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*, vol. 801, no. 2, p. 120, 2015.

[30] J. He, D. Jing, and G. Vidali, "Atomic oxygen diffusion on and desorption from amorphous silicate surfaces," *Physical Chemistry Chemical Physics*, vol. 16, no. 8, p. 3493, 2014.