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 11, 2016

Abstract

This abstract explains what happens. Benchmark, adsorption, maybe something about the interstellar playground.

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 composition of molecular compunds 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. The dust has been measured[?] , so the question is not so much if the reaction happens but rather to quantify its effect.

At temperatures as low as in cold interstellar clouds (between 10 and 100 K), the grains are covered in ices, most prominently 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.

Various works have already considered surface reactions.

(Here we will cite some experimentalists).

The research described above typically chose a model to describe the processes taking place in the experiment and adjusted simulation data – including parameters like the diffusion coefficient – to fit the experimental data. That would mean: if surface reaction processes are the key to filling the gap between observed and predicted reaction rates, the input parameters might be good approximations to the actual coefficients. This is a legit approach since one can not easily think of other processes than surface and gas phase reactions to lead to molecular formation.

However, there is not yet a proper *ab initio* theoretical calculation of surface diffusion. This means that evaluating the parameters found in simulation is a very difficult task – there is neither a recommended value nor are there any error bars to such a value. This work tries to make a first step toward the accurate simulation of surface adsorption, diffusion and reaction of small molecules on interstellar ices. Its aim is to establish a model of crystalline I_h water in which simulations can be carried out. The main mathematical tool for describing the chemistry of this surface is a subdomain treated by *density functional theory* (DFT) within a bigger domain where the interaction is modelled by *molecular mechanics* (MM). The two domains are coupled by a QM/MM coupling scheme.

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 for the actual system. After that, we give our results for adsorption energies in Section 5 and 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 $X-Y$. 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 $X-Y$, which will in general depend on the distance and the orientation of the two molecules, to be E_{X-Y}. Then, the interaction energy between the two molecules is the energy given by

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

We did not include spatial dependency of E_{X-Y} into the above definition. A map $(R_{X-Y}, \Omega_X, \Omega_Y) \mapsto E^{\text{int}}_{X-Y}$ with the center of mass separation R_{X-Y} and the molecular orientation Ω_X and Ω_Y is called the *potential energy surface* (PES) of the intermolecular interaction. It may also contain internal deformations of the molecule.
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 $X-Y$, that is the global minimum of the PES E_{X-Y} and therefore E^{int}_{X-Y}. If the interaction between X and Y were purely repulsive, that is $E^{\text{int}}_{X-Y} > 0$ for all geometries, the global minimum would not be well-defined since it requires inifinite separation of X and Y in an arbitrary direction. However, the algorithms we use will converge to a local minimum 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 $E^{\text{int}}_{X-Y} < 0$ 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

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

Again, we did not include the dependency 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 $(\mathbf{r}_i)_i \mapsto E^{\text{ads}}_{\text{s-}}X$, where the vector $\mathbf{r}_i$ is the coordinate of the i-th atom in s-X, $1 \le i \le N$ for some N. The different energy minima of this map to $E^{\text{ads}}_{\text{s-}}X$ 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 $E^{\text{ads}}_{\text{s-}}X$ can be useful to simulation.
It is also possible to calculate *reaction energies* E^{react} for adsorbed species. We will do this for a situation where species X is adsorbed on the surface to s-X. From the surrounding gas, a molecule of species Y approaches. The two react to form s-Z. The reaction energy consumed or released in this process is

$$E^{\text{react}} := E_{\text{s-Z}} - E_{\text{s-X}} - E_Y. \tag{3}$$

Positive E^{react} means an *endothermic*, negative E^{react} an *exothermic* reaction. 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. The reaction can be schematically described by

$$\text{s-X} + \text{Y}_{(\text{g})} \longrightarrow \text{s-Z}, \tag{4}$$

where the subscripted $\text{Y}_{(\text{g})}$ is the gas-phase Y.
The superscripts 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 $\mathbf{r}_i$, $1 \le i \le N$ 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,

$$\mathbf{H}(\mathbf{r}_1, \ldots, \mathbf{r}_N) := \left(\frac{\partial^2 E}{\partial \mathbf{r}_i \partial \mathbf{r}_j}(\mathbf{r}_1, ..., \mathbf{r}_N) \right)_{1 \le i,j \le N} \tag{5}$$

by

$$E^{\text{ZPE}} = E + \frac{1}{2}\text{tr}\left[\mathbf{H}\right] = E + \Delta E^{\text{ZPE}}, \tag{6}$$

where $\text{tr}\left[\mathbf{H}\right]$ 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 (6).

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*

$$\hat{H}\left|\Psi\right\rangle = E\left|\Psi\right\rangle. \tag{7}$$

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

$$\hat{H} = -\sum_{i=1}^{N} \frac{\hbar^2}{2m_e}\nabla_i^2 + \frac{e^2}{4\pi\varepsilon_0}\sum_{1\le i<j\le N}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + V(\mathbf{r}^N). \tag{8}$$

$\mathbf{r}_i$ 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

$$\begin{aligned} V(\mathbf{r}^N) = &-\sum_{A=1}^{K}\frac{\hbar^2}{2m_A}\nabla_A^2 + \frac{e^2}{4\pi\varepsilon_0}\sum_{A=1}^{K}\sum_{B=1}^{K}\frac{Z_A Z_B}{|\mathbf{R}_A - \mathbf{R}_B|} \\ &-\frac{e^2}{4\pi\varepsilon_0}\sum_{A=1}^{K}\sum_{i=1}^{N}\frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|}. \end{aligned} \tag{9}$$

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

$$V(\mathbf{r}^N) = V_{\text{cc}} + V_{\text{ce}}(\mathbf{r}^N) \tag{10}$$

with

$$V_{\text{ce}}(\mathbf{r}^N) = \sum_{i=1}^{N}\tilde{V}(\mathbf{r}_i) = -\frac{e^2}{4\pi\varepsilon_0}\sum_{A=1}^{K}\sum_{i=1}^{N}\frac{Z_A}{|\mathbf{R}_A - \mathbf{r}_i|}. \tag{11}$$

There will be more than one solution to (7), 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*

$$\begin{aligned} &\langle \mathbf{x}_1, \ldots, \mathbf{x}_l, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_N | \Psi_i \rangle \\ &= \quad \Psi_i(\mathbf{x}_1, \ldots, \mathbf{x}_l, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_N) \\ &= -\Psi_i(\mathbf{x}_1, \ldots, \mathbf{x}_k, \ldots, \mathbf{x}_l, \ldots, \mathbf{x}_N), \end{aligned} \tag{12}$$

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

$$\begin{aligned} E_0 &= \min_{|\Psi\rangle}\left\{\langle\Psi|\hat{H}|\Psi\rangle\right\}, \\ |\Psi_0\rangle &= \arg\min_{|\Psi\rangle}\left\{\langle\Psi|\hat{H}|\Psi\rangle\right\}. \end{aligned} \tag{13}$$

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 (7) has a variety of equivalent counterparts. One of them is the key to the approach of DFT. When multiplying (7) by the bra $\langle\Psi|$, one can interpret the resulting energy as a (non-linear) functional of the wavefunction by

$$E : \mathcal{H}_a \to \mathbb{R}, \qquad E[|\Psi\rangle] = \langle\Psi|\hat{H}|\Psi\rangle, \tag{14}$$

which would mean that the ground state energy can be found by minimizing the functional $E[|\Psi\rangle]$ according to (13). 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

$$\rho(\mathbf{x}_1, ..., \mathbf{x}_N) = |\Psi_0(\mathbf{x}_1, ..., \mathbf{x}_N)|^2, \tag{15}$$

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

$$\rho(\mathbf{r}_1, s_1) = \int \mathrm{d}\mathbf{x}_2 \cdots \int \mathrm{d}\mathbf{x}_N \, |\Psi_0(\mathbf{x}_1, ..., \mathbf{x}_N)|^2, \tag{16}$$

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 (7) 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

$$\begin{aligned} E[\rho] = &V_{\text{cc}} + \int \mathrm{d}\mathbf{r}\, \tilde{V}(\mathbf{r})\rho(\mathbf{r}) + \frac{e^2}{2}\int \mathrm{d}\mathbf{r}\int \mathrm{d}\mathbf{r}'\frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \\ &+ T_s[\rho] + E_{\text{xc}}[\rho] \end{aligned} \tag{17}$$

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

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V_{\mathrm{eff}}(\mathbf{r})\right)\phi_i(\mathbf{r}) = \varepsilon_i\phi_i(\mathbf{r}) \tag{18}$$

with the *effective potential*

$$V_{\mathrm{eff}}(\mathbf{r}) = V_{\mathrm{cc}} + V(\mathbf{r}) + \frac{e^2}{2}\int \mathrm{d}\mathbf{r}'\,\frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho(\mathbf{r})}. \tag{19}$$

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

$$V_{\mathrm{xc}} = \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho(\mathbf{r})}. \tag{20}$$

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

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2. \tag{21}$$

With that, the kinetic energy functional is given by

$$T_s[\rho] = \sum_{i=1}^{N}\int \mathrm{d}\mathbf{r}\;\phi_i^*(\mathbf{r})\left(-\frac{\hbar^2}{2m}\nabla^2\right)\phi_i(\mathbf{r}). \tag{22}$$

Now, the remaining problem before one can find self-consistent solution to (18) with (19) and use them to minimize (17) 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 (18). Since the KS orbitals ϕ_i define the potential V_{eff} in (19), these solutions are called the *self-consistent field* (SCF). (18) is usually solved iteratively, where one chooses some initial guess for the ϕ_i, then solves (18) 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 Methods

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 classes of 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[1].
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[2] 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[3] 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 (21) 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 (18) 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 dependency 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 com-

putational 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[2].

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[4,5], of which we will use Grimme's D3 correction[6]. Its most basic idea is to add a dispersion correction to the energy E_{KS} obtained from minimizing

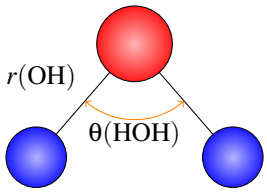

Figure 1 The geometry of TIP3P water. O is red, H is blue. Each atom coincides with a site with charge q and LJ-paramteres ε and σ.

(17). This yields

$$E_{\text{DFT-D3}} = E_{\mathrm{KS}} + E_{\mathrm{disp}}, \tag{23}$$

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

$$E^{(2)} = -\sum_{AB}\sum_{n=6,8} s_n \frac{C_n^{AB}}{r_{AB}^n} f_{d,n}(r_{AB}) \tag{24}$$

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

$$f_{d,n}(r_{AB}) = \frac{1}{1+6(r_{AB}/s_{r,n}R_0^{AB})^{-\alpha_n}}. \tag{25}$$

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

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.*[7]. 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*

$$V^{\mathrm{C}}_{\alpha\beta}(r_{\alpha\beta}) = \frac{1}{4\pi\varepsilon_0}\frac{q_\alpha q_\beta}{r_{\alpha\beta}}, \tag{26}$$

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

$$V^{\mathrm{LJ}}_{\alpha\beta}(r) = 4\varepsilon_{\alpha\beta}\left[\left(\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}}\right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r_{\alpha\beta}}\right)^{6}\right]. \tag{27}$$

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

$$V_i(\mathbf{r}_i) = \sum_{j\neq i}\sum_{\alpha,\beta}\left[V^{\mathrm{C}}_{\alpha\beta}(r_{\alpha\beta}(\mathbf{r}_i,\mathbf{r}_j)) + V^{\mathrm{LJ}}_{\alpha\beta}(r_{\alpha\beta}(\mathbf{r}_i,\mathbf{r}_j))\right]. \tag{28}$$

The first sum runs over all other water molecules and the second over all sites in both molecules. (28) 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 dependency 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[8].

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[9] 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 (27). That way, the electrostatic interaction is covered by the QM part while the MM part accounts for VdW interaction.

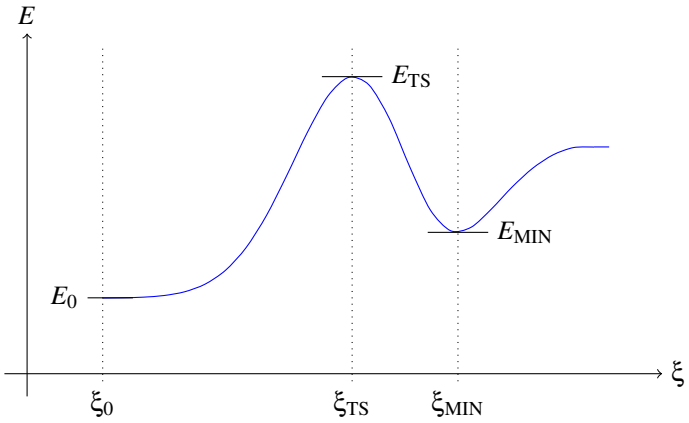

Figure 2 A cut of a PES along reaction coordinate ξ. The reaction coordinate starts at some initial ξ_0 with energy E_0 and follows a path of minimum energy. The minimum with E_{MIN} is located at ξ_{MIN}. The transition state with E_{TS} is a maximum at ξ_{TS}.

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)

$$\mathbf{F}(\mathbf{R}^M) = -\nabla V(\mathbf{R}^M) \tag{29}$$

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, (29) 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[10] with the default tolerances to follow the minimum energy path.
Evaluating the gradient in (29) 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 (28). For the QM region, there are schemes to obtain analytical results for the gradient for a given wave function[11].
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 (5). 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 mentioned, the optimisations are carried out with DL-FIND[10], as are the dimer method and the calculation of Hessian matrices. DFT calculations are performed with TURBOMOLE[3]. Force field calculations are performed by CHARMM[12]. These programs were interfaced via ChemShell[13], which also provided the QM/MM coupling.

For the post-HF calculations in the benchmark, we used the MOLPRO[14] package.

All molecular visualisations are created with VMD[15].

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[2]. 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[14] 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 $\mathbf{R}^M_{\mathrm{DFT}} = (\mathbf{r}^1_{\mathrm{DFT}}, \ldots, \mathbf{r}^M_{\mathrm{DFT}})$ from the reference optimum geometry $\mathbf{R}^M_{\mathrm{CC}} = (\mathbf{r}^1_{\mathrm{CC}}, \ldots, \mathbf{r}^M_{\mathrm{CC}})$ can be quantified by the *root mean square deviation* (RMSD). It is given by

$$\Delta_{\mathrm{RMSD}}(\mathbf{R}^M_{\mathrm{DFT}}; \mathbf{R}^M_{\mathrm{CC}}) = \sqrt{\frac{1}{M}\sum_{k=1}^{M} \left|\mathbf{r}^k_{\mathrm{DFT}} - \mathbf{r}^k_{\mathrm{CC}}\right|^2}. \tag{30}$$

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.

[*] 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.

3.1 H_2O-H Interaction

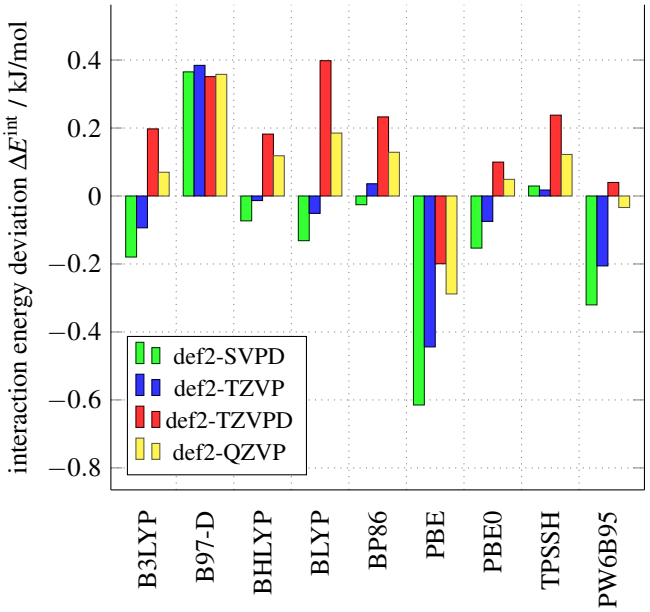

Figure 3 The H_2O-H benchmark results for different basis sets without dispersion correction. Energies are plotted as differences to the reference energy from CCSD(T)-F12A/VTZ-F12 calculations, see (31).

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

$$\Delta E^{\text{int}} = E^{\text{int}}_{\text{DFT}} - E^{\text{int}}_{\text{CC}}. \tag{31}$$

$E^{\text{int}}_{\text{DFT}}$ stands for the interaction energy (1) of the system computed with a certain DFT functional and basis set, and $E^{\text{int}}_{\text{CC}}$ is the CP corrected interaction energy with CCSD(T)-F12A/VTZ-F12. For H_2O-H, we found $E^{\text{int}}_{\text{CC}} = -0.40$ 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 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.

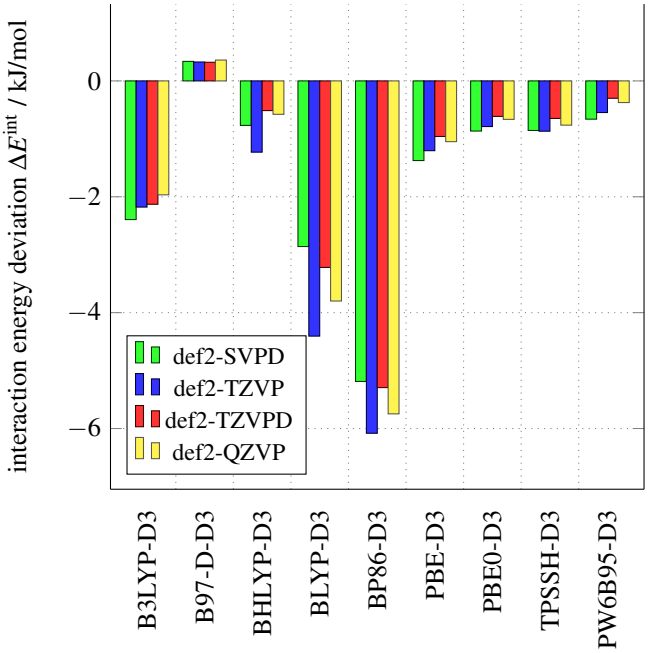

Figure 4 The H_2O-H benchmark results for different basis sets including dispersion correction. Energies are plotted as differences to the reference energy from CCSD(T)-F12A/VTZ-F12 calculations, see (31).

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.

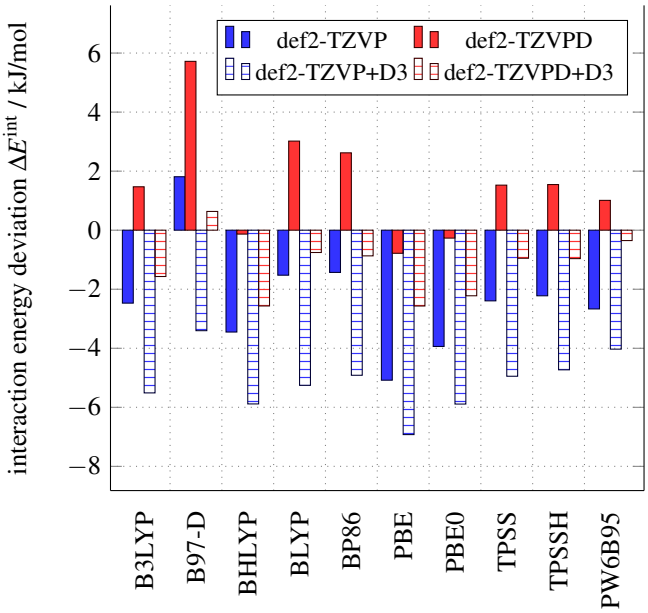

Figure 5 The H_2O-H_2O benchmark results for the def2-TZVP and def2-TZVPD basis sets. The results with dashed bars include dispersion corrections. Energies are plotted as differences to the reference energy from CCSD(T)-F12A/VTZ-F12 calculations, see (31).

3.2 H_2O-H_2O Interaction

Again, we compare energies to the reference energy as in (31). The reference energy for the H_2O-H_2O dimer is at $E^{\text{int}}_{\text{CC}} = -20.80$ kJ/mol, where the CP correction contributes $+0.18$ kJ/mol. 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.

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

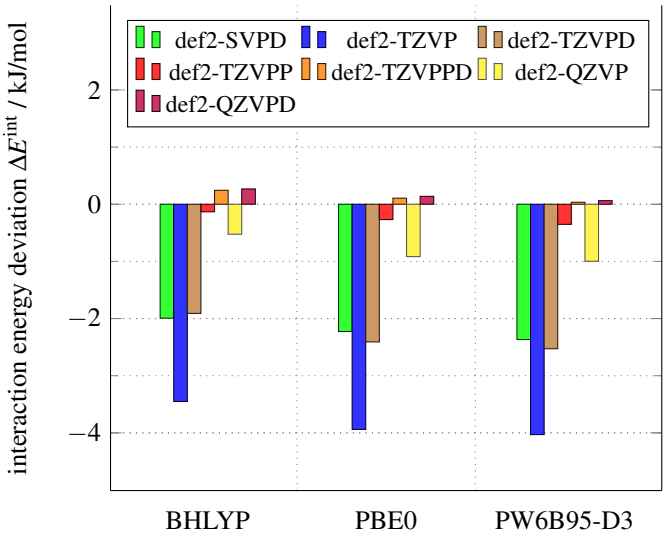

Figure 6 The H_2O-H_2O benchmark results for BHLYP, PBE0 and PW6B95-D3 and all basis sets used in the study (excluding def2-SVP). Energies are plotted as differences to the reference energy from CCSD(T)-F12A/VTZ-F12 calculations, see (31).

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 $E_{\text{CC}}^{\text{int}} = -6.72$ kJ/mol.

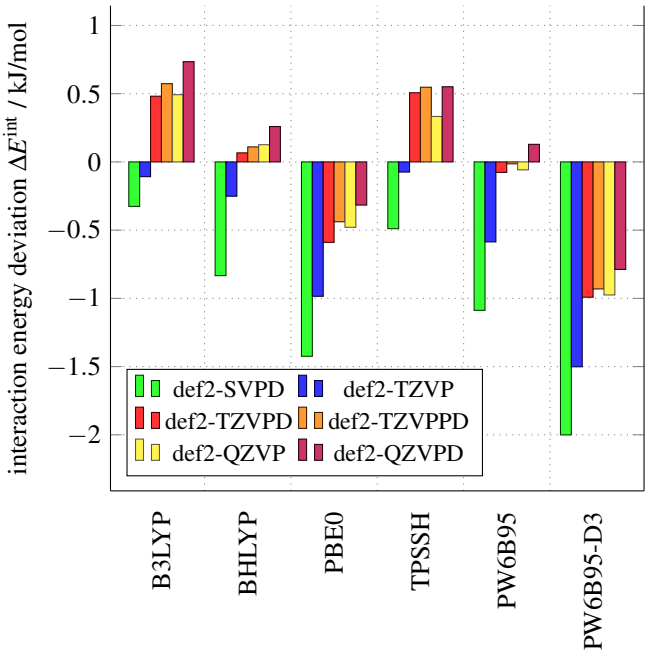

Figure 7 The $H_2O-{}^3O$ benchmark results for the def2-TZVP and def2-TZVPD basis sets. The results with dashed bars include dispersion corrections. Energies are plotted as differences to the reference energy from CCSD(T)-F12A/VTZ-F12 calculations, see (31).

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 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 basis 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[2]. 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 adsorbents 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 sort of non-literal atomisation energy by

$$E_X^0 = E_X - \frac{M_H}{2}E_{H_2} - \frac{M_O}{2}E_{{}^3O_2}. \tag{32}$$

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

The value is of no physical importance, but if two functionals 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.

An energy adjusted by (32) 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 make the results for the entire functional seem inaccurate. But to have a first comparison of the functionals, we have to accept these problems.

Note that the formerly considered PW6B95-D3 functional is not included in Table 1, since the dispersion correction only affects intermolecular interaction. Our calculations however treated isolated molecules.

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. Generally, all functionals predict too high values for E^0. Since the values are not too far off for H_2 while showing considerable problems with molecules composed solely of xygen, this could be an indication that the energetic description of ${}^3O_2$ is challenging to DFT functionals in general. Then again, the rather good agreement to the reference data for all functionals but BHLYP in the case of ${}^3O$ does not support this claim if we do not expect a cancellation of error.

Further comparison shows that for most molecules, the functionals closest to the CCSD(T)-F12A energies are B3LYP and

Table 1 Energies for DFT functionals, according to (32). 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
HO	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

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

A gas phase reaction is a simpler version of the surface reaction we introduced in the theoretical section, the scheme being

$$X_{(g)} + Y_{(g)} \longrightarrow Z_{(g)}, \tag{33}$$

much similar to (4). 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

$$HO + H \longrightarrow H_2O, \tag{34a}$$
$$H_2O + {}^1O \longrightarrow H_2O_2, \tag{34b}$$
$$HO + HO \longrightarrow H_2O_2, \tag{34c}$$
$$HO_2 + H \longrightarrow H_2O_2, \tag{34d}$$
$$HO + {}^3O \longrightarrow HO_2, \tag{34e}$$
$$^1O_2 + H \longrightarrow HO_2, \tag{34f}$$
$$^3O_2 + H \longrightarrow HO_2. \tag{34g}$$

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.
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 data for E^{ZPE}. From this data, it is already clear that no functional can be considered highly accurate for reaction energies. The most accurate ones by MAD are 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 (34b), this is beneficial
The results of Table 1 already
While the agreement between CCSD(T)-F12A and the DFT functionals is never better than a difference of 5 kJ/mol, there are cases where the agreement is particularly bad.

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. Then, we will give 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. Finally, to show an outlook on possible further application of this model, a transition state for $HO + H_2$ on the surface will be given alongside a corresponding IRC calculation.

5.1 The Surface Model

5.1.1 The Fletcher Surface

Barnes[16] 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[17], 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

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}. All energies in kJ/mol.

		B3LYP	BHLYP	PBE0	TPSSH	PW6B95	CCSD(T)-F12A
$HO_{(g)} + H_{(g)} \longrightarrow H_2O_{(g)}$	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
$H_2O_{(g)} + {}^1O_{(g)} \longrightarrow H_2O_{2(g)}$	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
$HO_{(g)} + HO_{(g)} \longrightarrow H_2O_{2(g)}$	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
$HO_{2(g)} + H_{(g)} \longrightarrow H_2O_{2(g)}$	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
$HO_{(g)} + {}^3O_{(g)} \longrightarrow HO_{2(g)}$	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
$^1O_{2(g)} + H_{(g)} \longrightarrow HO_{2(g)}$	E^{ZPE}	-357.45	-371.28	-356.87	-360.12	-351.14	-322.61
	E	-384.60	-399.86	-384.50	-387.18	-378.72	-350.98
	ΔE^{ZPE}	27.14	28.58	27.63	27.06	27.58	28.36
$^3O_{2(g)} + H_{(g)} \longrightarrow HO_{2(g)}$	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	21.29	42.22	27.54	24.98	20.70	
	MIN	-59.67	-48.67	-83.35	-84.18	-60.91	
	MAX	23.22	71.20	31.34	27.22	23.76	
	MEAN	-5.47	26.87	-7.79	-8.44	-6.91	

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[18] 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 on two alterating heights), one can connect each O atom to its four nearest neighbors. On these designated hydrogen bonds, H atoms can be placed close to one of the two O atoms taking part in the bond. This 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.

5.1.2 The QM/MM Region

Of a fully-ordered water I_h crystal with a Fletcher surface, we take a semispherical 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 semispherical 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.

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

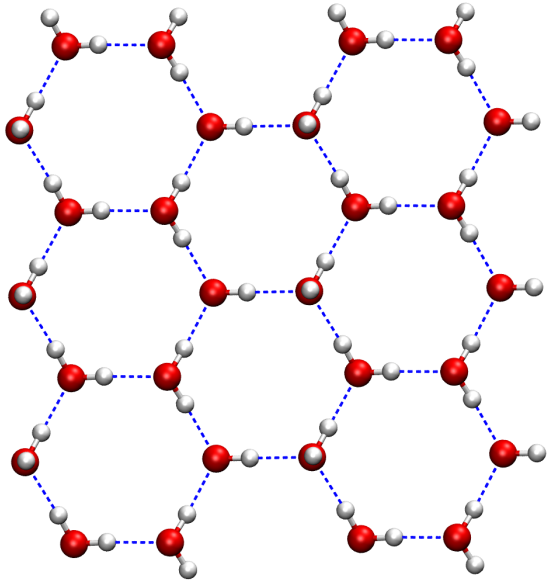

Figure 8 The Fletcher surface. Hexagonal O (red) structure with vertical lines of identical H (white) orientation. H-bonds are indicated by dashed blue lines. The O atoms are not all in the same plane.

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 may not be necessary as a boundary, but on the other hand it has virtually no effect on computation time to include these additional MM atoms, and if necessary they can be turned into further active atoms. There is a total of 1121 MM molecules, of which 231 are active MM molecules and 890 frozen MM molecules.

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

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, leading bond angles of 109.50°. These are in disagreement with the bond angles of the TIP3P potential fixed at 104.52°. To mend this problem, we simply take a perfect Fletcher surface and change the angles to the appropriate value while maintaining the orientation of the water dipole, that is the vector $\mathbf{r}_{O-H1} + \mathbf{r}_{O-H2}$ stays the same for each molecule.

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 (30) between the ideal Fletcher surface and the surface it changes into after QM/MM optimisation. Note that the factor $\frac{1}{M}$ in (30) 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. This can be expected, since the initial geometry is gained by cutting a perfect crystal at a plain, so the molecules that are now at the surface do not feel the attractive hydrogen bonds of the removed molecules anymore and move closer to the remaining molecules within the surface. This effect is not very strong, since the RMSD is only of order < 0.2 Å.

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.

We can also compare the behaviour of the QM and MM part during the optimisation process starting at the initial crystalline geometry. Figure 9 gives a typical example with the BHLYP functional. Both QM and MM energies are differences to their

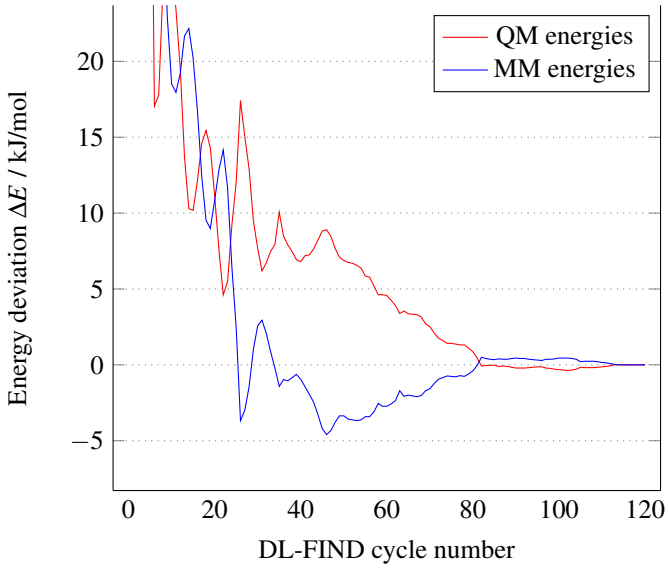

Figure 9 The QM/MM energy changes during a typical optimisation. The energies are deviations from their value at the optimum geometry. Functional: BHLYP.

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. Again an implication that the actual optimum geometry has to be tighter bound. From there on, one can see that the energy changes 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.

5.3 ZPE Corrections

Now we have established a model that should be able to describe at least 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, not fast analytical Hessian matrices can be computed. There is no way around computing finite differences, so two atomic diplacements for each coordinate per atom, leading to six displacements per atom. Computations for more complicated systems, especially s-HO_2 calculations, require a sensible choice of atoms to be displaced. On the other hand of course, enough atoms should be chosen for diplacement to maintain 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 those that we do describe with the def2-TZVPD basis set, that is the central ring of six H_2O molecules and adsorbent. The idea here is that when an adsorbent comes into the system and is adorbed 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, 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 adsorbent, these $18+M$ atoms need $54+3M$ displacements (plus one initial geometry). This is much less than the original $324+3M$ displacements for $108+M$ atoms.

We will call this selective correction the ZPE/ring correction in contrast to the ZPE/all correction for all QM atoms.

We can compare the two corrections in the next section.

5.4 Adsorption

If we assume the ZPE/all correction to be accurate, we can evaluate the ZPE/ring correction when comparing results for the B3LYP functional in Table 4. Of the molecules relevant in Section 4, we examine adsorption energies for all species but ^1O, since is chemisorbed by the surface to form some conformation of H_2O_2.

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-HO	-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

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 different functionals.

In general, the value of the corrective term does vary from adsorbent to adsorbent. 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 in 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 surface-molecule 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.

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 HO, ^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[20,21]. Other data include H_2O adsorption on an Au surface covered in crystalline water ice[19] and $^1O_2$ adsorption on porous amorphous water ice[22]. Although the experiments on bare silicates do not quite resemble our system, we consider the data they provide to give good guidance since they are at least influenced by adsorption on the freshly formed water ice surfaces on the probes.

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. $^3O_2$, which was also critical in our discussion of the ZPE correction, is not the unanimously considered repulsive. The PW6B95 method predicts strong repulsion, which is to some extent supported by PW6B95-D3, but the rest of the functionals seems to still predict a weak attraction or in the B3LYP case a very weak repulsion.

While each functional does have its "slips" for some adsorbents, the first four functionals of Table 5 seem to be in good accord. For many cases, one of the two PW6B95 functionals is furthest from an average value of all six functionals. This is especially true for O and O_2. Since it is either PW6B95 or PW6B95-D3, this could be an indication that neither the PW6B95 nor the D3 dispersion correction are universally reliable for the study at hand. 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 $\Delta E^{\text{all}} = 15.98$, 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 du 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.* [23] 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. Due to their rather bad performance for O species, the PW6B95 functionals have lost some credibility.

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 in kJ/mol.

	B3LYP	BHLYP	PBE0	TPSSH	PW	PW-D3	experimental, E_{des}
s-H	3.32	3.77	3.37	-5.05	3.29	1.27	
s-H_2	5.77	5.11	3.57	4.55	4.65	2.47	
s-H_2O	-32.03	-35.07	-36.90	-30.92	-33.51	-41.17	-48.00 ± 0.50[a]
s-H_2O_2	-32.47	-35.99	-37.64	-31.04	-34.85	-44.33	
s-HO	-31.39	-32.88	-35.69	-31.10	-32.78	-39.66	-13.77 to -39.58[b]
s-HO_2	-52.30	-49.10	-56.97	-54.10	-51.08	-57.80	
s-^3O	-13.48	-5.16	-11.25	-13.22	-12.85	-18.13	-14.67[c]
s-$^1O_2$	-5.95	-7.51	-11.44	-9.44	-7.44	-12.53	-7.52[c],-13.80 ± 0.50[d]
s-$^3O_2$	0.45	-0.65	-1.52	-0.52	26.68	7.04	

[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 [19]

[b] Adsorption on an amorphous silicate. He, Vidali 2014 [20]

[c] Adsorption on an amorphous silicate. He, Jing, Vidali 2014 [21]

[d] Adsorption on porous amorphous water ice. He *et al.* 2015 [22]

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 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_{avg}
s-HO $+H_{(g)}$	$\longrightarrow$ s-H_2O	-476.22	-458.03	-472.39	-470.76	-480.13	-480.92	-473.07
s-H_2O $+^1O_{(g)}$	$\longrightarrow$ s-H_2O_2	-402.54	-357.85	-427.27	-426.38	-405.24	-407.50	-404.46
s-HO $+HO_{(g)}$	$\longrightarrow$ s-H_2O_2	-185.12	-137.04	-194.02	-187.84	-197.12	-200.16	-183.55
s-HO_2 $+H_{(g)}$	$\longrightarrow$ s-H_2O_2	-317.22	-318.69	-310.31	-309.82	-320.89	-323.99	-316.82
s-HO $+^3O_{(g)}$	$\longrightarrow$ s-HO_2	-294.05	-219.54	-300.54	-295.71	-296.15	-296.10	-283.68
s-^3O $+HO_{(g)}$	$\longrightarrow$ s-HO_2	-311.96	-248.31	-325.41	-313.50	-316.45	-318.00	-305.61
s-$^1O_2$ $+H_{(g)}$	$\longrightarrow$ s-HO_2	-403.80	-414.31	-402.88	-404.70	-395.22	-396.95	-402.98
s-$^3O_2$ $+H_{(g)}$	$\longrightarrow$ s-HO_2	-248.05	-243.27	-241.72	-250.31	-268.95	-256.14	-251.41
$E_{\mathrm{DFT}} - E_{\mathrm{avg}}$	MAD	3.49	31.37	10.86	7.28	9.26	8.78	
	MIN	-10.37	-11.33	-22.81	-21.91	-17.54	-16.61	
	MAX	3.36	64.14	9.68	7.01	7.76	6.03	
	MEAN	-2.17	28.07	-6.62	-4.68	-7.32	-7.27	

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.

A selection of energies for Eley-Rideal reactions is given in Table 6. They are surface-mediated versions of the reactions discussed in Section 4. All reaction energies are computed according to (3) and the reactions on the left are given in a scheme like (4).

We use the minimum energy values of each method 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 $HO_{(g)}$. 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 $HO_{(g)}$ 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 big change appears for BHLYP, which is generally not in good agreement with the rest of the reaction energies.

Recalling the strong deviations between DFT and CCSD(T)-F12A gas phase reaction energies, we must admit that we can not know which of the reaction eneriges in table 6 is close to the truth. Therefore, we will give an average

value for each reaction. There is no reason to hope that the process of averaging should help to find an accurate value for the reaction energy, especially since the $_{\text{DFT}}$ gas phase reaction energies were often incongruent with the CCSD(T)-F12A data. The sample of reactions is also too small to claim that a functional that is close to the average in most cases should be close to the average value of DFT energies for all surface reactions composed of these molecular species.
But to offer recommendations for reaction energies which may be of further use to simulations, it seems reasonable to provide a value that is somewhat democratically chosen from all the functionals. And since no functional was particularly close to the reference data in Section 4, the choice of taking a simple mean value comes naturally.
We can calculate the MAD from the average values, which is minimal for the B3LYP functional. With a maximum absolute deviation of 10.37 kJ/mol, B3LYP can be seen as not too strongly "oscillating".

6 Possible Application

6.1 Binding Sites

6.2 Transition State and IRC

7 Conclusion

Acknowledgment

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

References

[1] 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.

[2] 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.

[3] "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`.."

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

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

[14] 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.

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

[16] 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.

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

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

[19] 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.

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

[21] 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.

[22] 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.

[23] P. C. 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.