







This research has been conducted at the **Center for Molecular Modeling**, in collaboration with:

- Centre for Surface Chemistry and Catalysis, University of Leuven (Prof. R. Ameloot, João Marreiros)
- Instituto de Tecnología Química, Universitat Politècnica de València (Dr. F. X. Llabrés i Xamena, Dr. Francisco García Cirujano)
- Van 't Hoff Institute for Molecular Sciences, Universiteit Van Amsterdam (Prof. E. J. Meijer, Prof. B. Ensing)



# Contents

<b>List of Abbreviations</b>	<b>iii</b>
<b>List of Symbols</b>	<b>v</b>
<b>Samenvatting</b>	<b>vii</b>
<b>Summary</b>	<b>ix</b>
<b>I Towards modeling catalytic processes in UiO-66 metal-organic framework at operating conditions</b>	<b>1</b>
<b>1 Introduction</b>	<b>3</b>
1.1 Catalysis . . . . .	3
1.2 Metal Organic Frameworks . . . . .	4
<b>2 Modeling metal organic frameworks</b>	<b>7</b>
2.1 Framework topology . . . . .	7
2.2 Theoretical methods . . . . .	10
2.3 Free energy . . . . .	16
2.4 Exploring the free energy surface . . . . .	20
<b>II Published papers</b>	<b>23</b>
<b>Bibliography</b>	<b>25</b>



# List of Abbreviations

---

<b>AIM</b>	Atomic layer deposition in MOF
<b>BEA</b>	Beta polymorph A
<b>BDC</b>	Benzene-1,4-dicarboxylate
<b>BPDC</b>	Biphenyl-4,4'-dicarboxylate
<b>BTC</b>	Benzene-1,3,5-tricarboxylic acid
<b>CMM</b>	Center for Molecular Modeling
<b>CN</b>	Coordination number
<b>CV</b>	Collective variable
<b>DFT</b>	Density functional theory
<b>DMMP</b>	Dimethyl methylphosphonate
<b>DRIFT</b>	Diffuse reflectance infrared Fourier transform
<b>FAU</b>	Faujasite
<b>FER</b>	Ferrierite
<b>FES</b>	Free energy surface
<b>FHVA</b>	Full Hessian Vibrational Analysis
<b>FTIR</b>	Fourier-transform infrared
<b>GCMC</b>	Grand canonical Monte Carlo
<b>HCl</b>	Hydrochloric acid
<b>IR</b>	Infrared
<b>IUPAC</b>	International Union of Pure and Applied Chemistry
<b>IZA-SC</b>	Structure Commission of the International Zeolite Association
<b>LFEP</b>	Lowest free energy path
<b>M-L</b>	Metal-linker coordination
<b>MD</b>	Molecular dynamics
<b>MFI</b>	Mordenite Framework Inverted
<b>MOF</b>	Metal-organic framework
<b>MOR</b>	Mordenite
<b>MPV</b>	Meerwein-Ponndorf-Verley
<b>MTD</b>	Metadynamics
<b>MTK</b>	Martyna-Tobias-Klein
<b>MTO</b>	Methanol-to-olefins
<b>NEB</b>	Nudged elastic band
<b>PHVA</b>	Partial Hessian Vibrational Analysis
<b>PES</b>	Potential energy surface
<b>PXRD</b>	Powder X-Ray diffraction
<b>RE</b>	Replica exchange
<b>SBU</b>	Secondary building unit
<b>SCF</b>	Self-consistent field

<b>SXRD</b>	Single-crystal X-ray diffraction
<b>TPDC</b>	<i>p</i> -terphenyl-4,4"-dicarboxylate
<b>TFA</b>	Trifluoroacetic acid
<b>TG-DSC</b>	Gravimetric characterization techniques
<b>TPS</b>	Transition path sampling
<b>TS</b>	Transition state
<b>TST</b>	Transition state theory
<b>UFF</b>	Universal force field
<b>UiO</b>	Universitetet i Oslo (University of Oslo)
<b>US</b>	Umbrella sampling
<b>WHAM</b>	Weighted histogram analysis method
<b>ZPE</b>	Zero point energy
<b>ZSM-5</b>	Zeolite Socony Mobil-5

---

# List of Symbols

---

$\ddagger$	Transition state
$\Delta E^\ddagger$	Electronic energy barrier
$\Delta H^\ddagger$	Enthalpy barrier
$\Delta G^\ddagger$	Free energy barrier
$B$	Bulk modulus
$E$	Electronic energy
$E_{ZPE}$	Zero-point vibrational energy
$F$	Helmholtz free energy
$G$	Gibbs free energy
$H$	Enthalpy
$h$	Planck constant
$k$	Rate coefficient unimolecular reaction
$k_B$	Boltzmann constant
$N$	Number of particles
$N_A$	Avogadro constant
$N_{dof}$	Number of degrees of freedom
$p$	Pressure
$P$	Product
$q_X$	Partition function of species $X$
$q_{vib}$	Vibrational partition function
$R$	Reactant
$S$	Entropy
$T$	Temperature
$U$	Internal energy
$Q$	Partition function
$\nu_i$	Vibrational frequency

---



# Samenvatting

Summary in Dutch [here](#)



# Summary

Summary in English here



# List of Figures

2.1	Extended cluster model cut from the periodic structure of UiO-66. The cluster contains the active site, the brick and the linkers in the closest proximity to the active site. . . . .	9
2.2	Representation of the unit cells containing the defect. In blue, the conventional 4-brick unit cell, in orange, the 2-brick unit cell used for the calculations. The two different bricks are highlighted in orange. The 10-fold coordinated brick has two terephthalate linkers missing, one at site A and one at the opposite site B. . . . .	10
2.3	The two lowest PES in the BO approximation for a diatomic molecule. In blue, the electronic PES for the ground state and first electronic excited state (UV-Vis transition). In orange, the vibrational energy levels (IR transitions), in grey the rotational levels (microwave transitions). . . . .	11
2.4	Representation of some of the molecular modes taken into account in a generic force field model. . . . .	12
2.5	Schematic representation of the potential energy surface and the stationary points. . . . .	14
2.6	representation of the atoms taken into account in the PHVA approach. Top: a schematic representation of a reactive process in nanoporous material, bottom: a snapshot from the static calculations where the atoms of the active site and the adsorbates are highlighted . . . . .	17



## **Part I**

**Towards modeling catalytic  
processes in UiO-66  
metal-organic framework at  
operating conditions**



# 1

## Introduction

### 1.1 Catalysis

Catalysis, from the greek for “down” and “loosen”, is at the heart of almost every industrially relevant chemical process. A catalyst intervenes in the reaction mechanism, allowing chemical species to meet each other and react with a specific mechanism, without being consumed. It can not only increase the rate of a given reaction, allowing processes which would not occur spontaneously, but can also steer the reaction towards specific end products. Catalysts can be divided in homogeneous and heterogeneous, according to the phase where they are located with respect to the reactants. Homogeneous catalysts share the same phase with the reactants, and can be for instance molecules dissolved in a solvent, such as organometallic compounds. These type of catalysts are used in many industrial processes, but have the drawback of being difficult to separate from the products. From a modeling and experimental point of view, they are the easiest to study, and it is easier to hypothesize reaction mechanisms from experimental data. Heterogeneous catalysts, on the other hand, are located in a different phase with respect to the reactants, and for this reason they have the advantage of being easily separable from the reaction products. Even if the term does not refer to a specific phase, often heterogeneous catalysts is synonymous for solid catalysts. These materials must possess specific active sites that easy to reach for the reactants, where these can adsorb, react, desorb, and ultimately diffuse back in the bulk to leave space for a new cycle. For these processes to occur, the

number of active sites and the area of contact between the two phases, or surface area must be sufficiently high. In this sense, nanoporous materials, which possess pores with diameter of  $\geq 100$  nm, have drawn a lot of attention for their enormous catalytic potential. Their pore structure provides them with an exceptionally high surface area, facilitating diffusion of reactants inside the material, and allowing shape selectivity to give specific products. The most industrially relevant class of nanoporous materials is certainly zeolites, which are used extensively in the petrochemical industry. In general, the study of solid heterogeneous catalysts is more difficult than their homogeneous counterparts due to their higher complexity. It is not always clear, in fact, where exactly the active sites may be located in the material, and how they interact with molecules.

In this work of thesis, we focus on a new class of nanoporous materials which combine the crystalline nanoporous structure of zeolites with the versatility in synthesis of metal complexes and for this reason has shown great potential for catalysis.

## 1.2 Metal Organic Frameworks

Metal organic frameworks (MOFs) are one of the most intriguing class of materials of current science. These promising materials, first called 'coordination polymers' were discovered in the late 50's, but only at the end of the century with the works of Robson, Kitagawa, Yaghi and Ferey, the scientific community started to understand their full potential. If at first these materials were seen as a laboratory curiosity, with the focus on the discovery of new structures, in the last few decades, the field has seen an incredible explosion in scientific and industrial interest, with new applications being explored. MOFs are hybrid nanoporous materials that are composed by metal or metal-oxo clusters connected by multtopic organic linkers, to form multidimensional pore structures. Their particular building block design allows the creation of an almost infinite number of crystalline structures with different topology and chemical composition. In principle, by modulating linkers and bricks, many properties can be tuned, such as pore shape and size, functionalization, surface area, or response to chemical and physical stimuli. Moreover, multiple physical or chemical functionalities can be integrated in the crystals at the same time. This tunability, along with their high cristallinity, metal content and porosity, allows their application in different industrially relevant fields, such as catalysis, gas storage and separation, drug delivery or sensing. More specific applications are being further explored, such as warfare agents decomposition, magnetic applications, or membrane separation. After the discovery of MOFs, with their tunability and ease in functionalization, the study of their application in catalysis followed naturally. Their high porosity, in particular, allows chemical species to diffuse in the pores and access the active sites, and the high metal content offers the possibility to have many guest interactions sites. Exploiting this property, a plethora of MOFs has been synthesized possessing unsaturated metal sites within the pores which

offer different Lewis acidity. In this sense, provided the structures are stable in the solvent, MOFs can be considered true heterogeneous catalysts, where the active sites are inherent part of the framework.

## Soft Porous Crystals

### Post-synthetic modification

When it is not possible to introduce functionalities with direct synthesis, post synthetic modification, which makes use of the building block design, has become a well established procedure that allows the preparation of MOF materials with specific functionality. Via this technique, it is possible to modify the crystal after the synthesis, allowing to finely tune the properties of the material. Post synthetic modifications include encapsulation of guest molecules or nanoparticles in the pores, modifications of the linkers, or post synthetic exchange (PSE), where building blocks are dynamically exchanged.

## Zr-based MOFs

The metal-ligand coordination bonds that make MOFs so tunable is also regarded as one of their main drawbacks, as it is responsible for a less stable structure when compared to already established nanoporous catalysts such as zeolites. The first generation of MOFs suffered from low stability that limited their range of applications, together with their high production cost.

Recently, a class of robust MOFs have been synthetized having the necessary thermal, mechanical and chemical stability to function at operating conditions. Zr-based MOFs

Furukawa jacs 2014, Bai RSC 2016

which exploit the robustness of the Zr-O bond, show an outstanding stability and are at present time one of the most studied classes of MOFs. The majority of these materials is characterized by Zr(IV) atoms in a high coordination state, which interact strongly with the oxygens of carboxylate linkers of various topology. The SBU from which most of the Zr-based MOFs stem is a  $\text{Zr}_6\text{O}_4(\text{OH})_4$  cluster in which each of the 6 zirconium is connected to 4 oxygen atoms in the brick, and can form 4 other bonds with ligands, accommodating up to 24 metal-ligand bonds per cluster. Each zirconium atom forms a total of 8 coordination bonds giving a rich range of possible structures that can be formed. Zirconium atoms that remain undercoordinated in these materials are Lewis acid sites where reactants can adsorb. Different structures can be constructed with this SBU and different linker topology and connectivity ranging from 12 to 6. The higher the connectivity and the number of metal-ligand bonds, the higher the stability of the material. However, active sites in these materials are present only when the brick is somehow undercoordinated.

Zirconium is an ubiquitous metal that is present in biological systems and has low toxicity, as well as limited cost. For this reason, this class of MOFs is promising for many applications.

## **UiO-66**

The most widely studied MOF up to date belongs to the class of Zr-based MOFs. UiO-66, composed by  $\text{Zr}_6\text{O}_4(\text{OH})_4$  bricks and terephthalate (BDC) linkers

## **Defects on UiO-66**

### **Active sites for catalysis on UiO-66**

### **Outline and goal of the thesis**

In this work, the properties of catalytically active sites on UiO-66 were studied by means of molecular modeling.

In Chapter 2, a theoretical overview of the state of the art modeling techniques is given. Particular attention is drawn on how these techniques can be applied in the case of MOF materials to obtain insight into structural and catalytic properties at operating conditions.

In Chapter 3, the main results of this PhD thesis are summarized. All results are the result of fruitful collaborations and have been published in international peer-reviewed journals.

In Chapter 4, some conclusions and perspectives are given.

# 2

## Modeling metal organic frameworks

The understanding of catalytic processes in MOFs is a very challenging task. MOFs are materials of complex nature, and reactive processes in these materials are elusive and difficult to track on a purely experimental basis. Molecular simulations offer an alternative approach that starts from the construction of models that can explain, complement, and predict experiments. With growing computational power, computational models can aim at giving a more and more accurate description of materials at operating conditions, narrowing the gap between theory and experiment. Often, the structural properties and chemical transformations that take place on the active sites need to be investigated using a combination of multiple computational techniques, that allow to tackle the problem from different points of view. In this chapter, an overview of the different computational methods that can be used to study reactive processes in MOFs will be given.

check this part later

### 2.1 Framework topology

A crucial decision when performing simulations lies in the choice of the model system, and what should be included in it. When choosing a model to represent the system under study, there is always a fine balance between accuracy and computational cost. On the one hand, it is crucial to use a computational model that captures all the relevant properties of the material and mimics the experimental

structure as close as possible. On the other, it is often convenient to approximate and neglect certain properties in favor of a larger scale description of the processes. The focus in this work are the active sites that can be used for catalysis, therefore an accurate electronic description of this region of the material is imperative. Nevertheless, the activity of these sites for chemical reactions can also be influenced by other factors, such as the pore size or functionalization. Therefore, to describe active sites in MOFs and other nanoporous materials, which can have rather large unit cells and non-periodic structural defects, the first question that needs to be asked is how to account for periodicity. Two conceptually different approaches, which are described below, can be used.

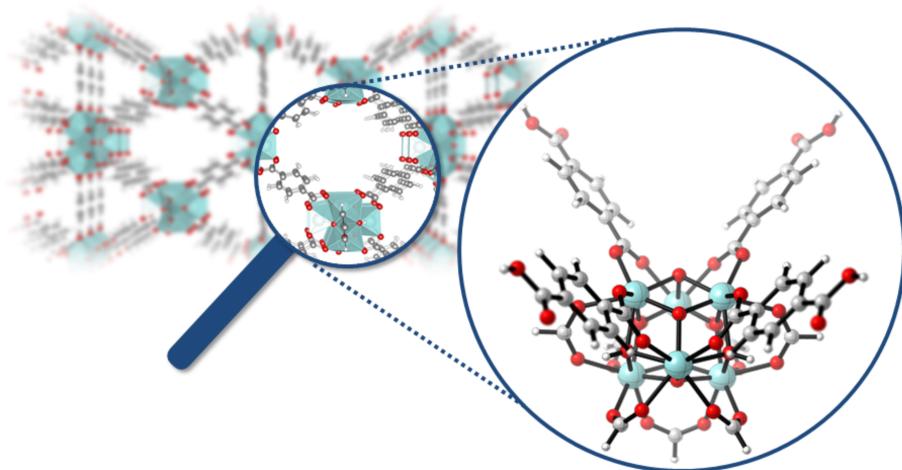
## Extended cluster model

A very computationally efficient approach consists in neglecting periodicity and extracting a finite cluster of atoms from the periodic structure. This cluster model, displayed in Figure 2.1, contains the active sites and their surroundings but consists in a limited number of atoms, which decrease the computational cost. This allows both a more accurate treatment of the electronic structure, and a screening of different possible geometrical configurations of adsorbates, which is useful in the search for transition states. Moreover, very efficient transitions state searching algorithms have been developed for such systems in Gaussian, the most widely used code for cluster calculations. When cutting a cluster, particular attention has to be drawn to the termination of bonds and the charge compensation, that have to be done in the most realistic way. The rest of the crystal structure does not surround the external cluster atoms. Some of these atoms need to be fixed in order to mimic the periodic environment and prevent nonphysical deformations that would affect an estimation of the entropy [1].

Cluster calculations are an excellent way to benchmark and do a first qualitative screening of reactions and possible configurations and have been for long the standard computational tool when studying reaction in nanoporous materials. However, they are not adequate to correctly describe complex reactive processes, where confinement effect of the pores and structural rearrangements can play an important role. The role of solvent in the pores can also be crucial for the outcome of a reaction and cannot be explicitly studied by cluster models. Periodic calculations resolve this shortcoming, and as computational power grows, the heterogeneous catalysis community is shifting towards these more expensive, however more accurate models.

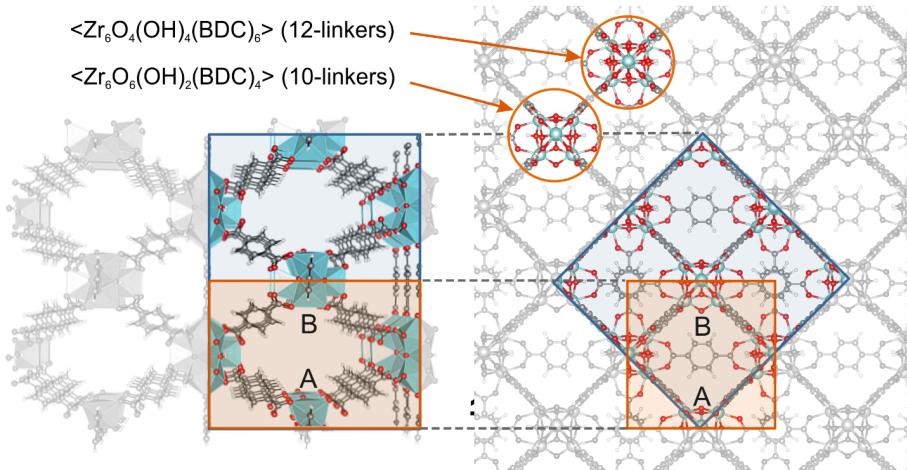
## Periodic model

Periodic models enable to describe the whole topology of the framework. These calculations make use of periodic boundary conditions (PBC), which allow to simulate bulk phases with a limited number of atoms. In this model, the unit cell is replicated infinite times in each direction. When one atom disappears from one side



**Figure 2.1:** Extended cluster model cut from the periodic structure of UiO-66. The cluster contains the active site, the brick and the linkers in the closest proximity to the active site.

of the unit cell it will reappear on the opposite side and each atom interacts with its neighbors in the same unit cell but also in the adjacent ones. Spurious interactions between the atoms can be avoided by applying a *minimum image convention*, for which each atom interacts with its nearest neighbor or periodic image. In the case of long range interactions such as the electrostatic other techniques need to be used, such as Ewald summation [2], where the potential is divided into a short range contribution, calculated in real space, and a long range contribution, calculated in reciprocal space using a Fourier transform. In the case of UiO-66, the conventional unit cell [3] contains 4 Zr bricks (Figure 2.2, blue). In the calculations of this thesis, BDC linkers have been removed in the unit cell to introduce defects which are active sites in catalysis. Different amounts of missing linkers with different topologies have been considered. An interesting topology is the one denoted as type 6 in the work of Rogge et al. [4] that is characterized by a channel which offers good perspectives for the diffusion of guest molecules. This unit cell (displayed in blue in Figure 2.2) can be reduced by symmetry to a 2-brick unit cell (in orange, Figure 2.2) which offers the best compromise between accuracy and computational cost. This reduced unit cell is used in most of the calculations performed in this thesis.



**Figure 2.2:** Representation of the unit cells containing the defect. In blue, the conventional 4-brick unit cell, in orange, the 2-brick unit cell used for the calculations. The two different bricks are highlighted in orange. The 10-fold coordinated brick has two terephthalate linkers missing, one at site A and one at the opposite site B.

## 2.2 Theoretical methods

### Electronic energy methods

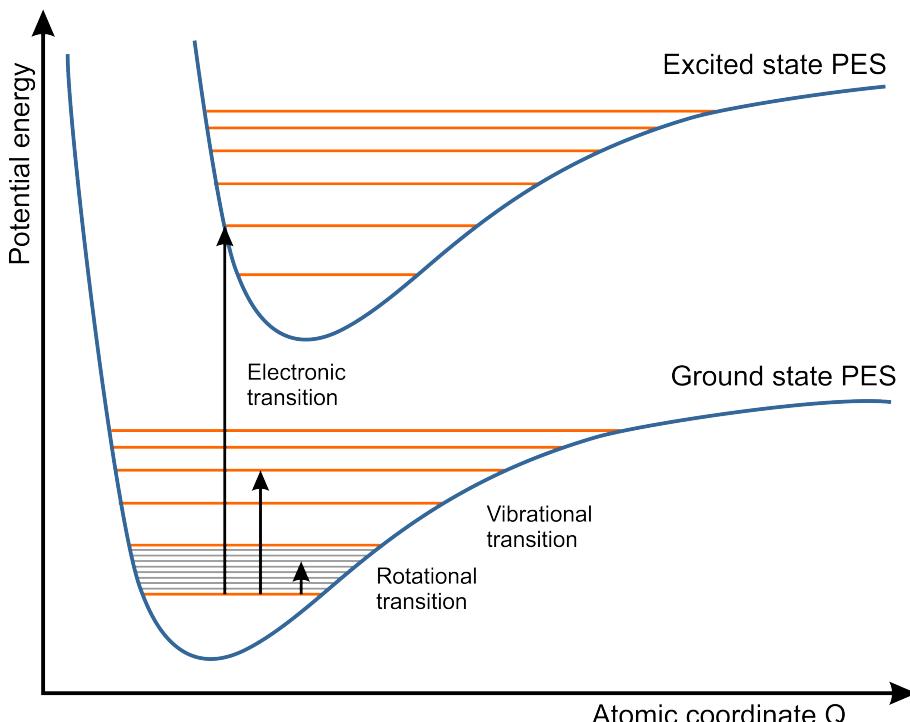
A basic quantity to study any chemical or physical transformation is the potential energy surface (PES), which is a function of the coordinates of all the atoms of the system. The PES is always the reference quantity in our simulations and every atomic configuration can be represented as a point in this hypersurface, with a given value of potential energy (Figure 2.3). Ideally, by calculating the value of the PES for each atomic configuration we can obtain all information on the system and on the transformations that can occur. However, the complexity of this surface escalates quickly with the number of atoms, and the sampling of its relevant regions represents the main challenge of molecular simulations. The information gained by exploring the PES is tightly connected to the experimental observables. Statistical physics acts like a bridge between the microscopic insight that is gained through molecular simulations and the macroscopic properties which are measured experimentally. In principle, all macroscopic properties of a system can be derived from its wavefunction. To calculate it, the stationary Schrödinger equation is solved:

$$\hat{H} |\psi\rangle = E |\psi\rangle$$

where  $\psi$  is the wavefunction,  $\hat{H}$  is the Hamiltonian of the system, and  $E$  is the total energy. The resolution of this equation is at the heart of computational

chemistry and will in principle provide the exact description of matter, but it is nevertheless extremely difficult to solve for most of the electron systems. The presence of electron-electron interactions makes it a highly coordinated problem, and for this reason, different approximations need to be applied, to remove the interactions that have a minimal contribution to the energy.

### Born–Oppenheimer approximation

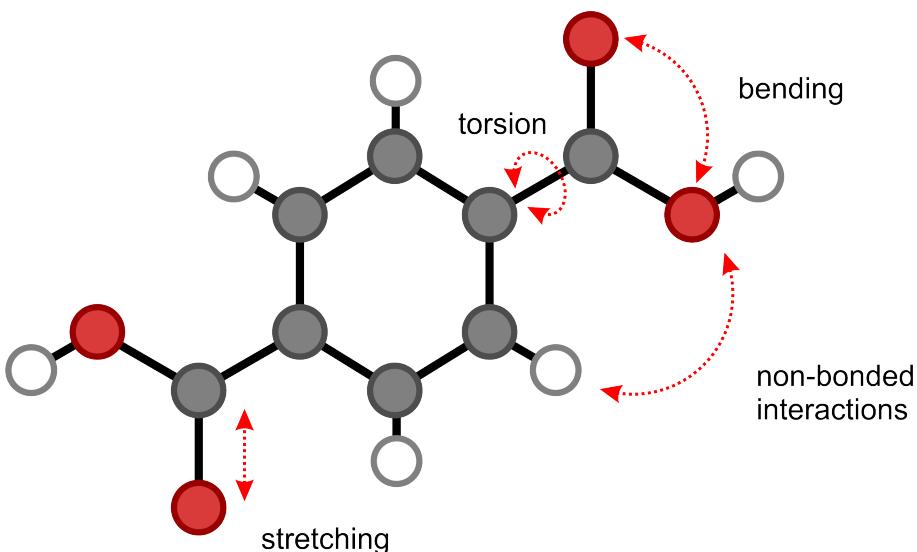


**Figure 2.3:** The two lowest PES in the BO approximation for a diatomic molecule. In blue, the electronic PES for the ground state and first electronic excited state (UV-Vis transition). In orange, the vibrational energy levels (IR transitions), in grey the rotational levels (microwave transitions).

For all calculations performed in this work, we rely on the so-called Born–Oppenheimer (BO) approximation [5]. In this treatment, nuclei are considered as classical points which move in the potential energy surface generated by the electrons (Figure 2.3). This way, to each nuclear configuration a corresponding electronic energy can be assigned, and nuclear coordinates enter in the Schrödinger equation only as parameters, allowing to construct a BO surface, or PES. This approximation holds since nuclei are much slower than electrons, therefore the motion of electron is instantaneous from the nuclei point of view. This approximation is not always possible, especially when dealing with light nuclei such as hydrogen.

In these cases, nuclear quantum effects can have an impact on the measured properties [6]. In most cases, the electronic ground state is also not interacting with the higher electronic states because of the high energy difference. In the BO approximation, the electronic energy levels are also considered fully separated and do not interact with each other. For this reason, the approximation is also called adiabatic approximation. Additional interactions have to be considered when two surfaces lie close to each other, for instance in the neighborhood of conical intersections.

## Force Fields



**Figure 2.4:** Representation of some of the molecular modes taken into account in a generic force field model.

The simplest way to describe interactions between atoms which determine the PES is the so-called “balls and springs” model. In this treatment, all interactions are represented by interatomic classical potentials which are parametrized to reproduce the results of more accurate quantum mechanical calculations. In this work, generic force field calculations have been used in some cases to give preliminary input structures for more costly *ab initio* calculations, through which the description of chemical transformations is possible. Force fields are often constituted by harmonic potentials which do not allow bonds being broken and formed (Figure 2.4). Reactive force fields, such as ReaxFF [7] are currently being developed, but their application in complex heterogeneous reactions is still an ongoing chemical challenge and is out of the scope of this work. For this reason, the description of reactive processes needs a more advanced treatment, where electronic distributions are explicitly taken into account.

## Density Functional Theory

Density functional theory (DFT) has become the method of choice for the study of chemical systems, due to its good trade-off between computational cost and accuracy of the obtained results. DFT began in the 1920's with the work of Thomas and Fermi [8, 9], but it was only in the '60s that it became a complete and accurate theory, shown in the work of Kohn, Hohenberg and Sham [10]. The fundamental property that DFT describes is electron density as opposed to many body electron wavefunctions, which allows to reduce enormously the number of variables in the case of complex systems. Two fundamental theorems by Hohenberg and Kohn state that there is a unique relation between electronic density and total wavefunction, therefore the ground state density allows us to determine all properties of the system. Moreover, the ground state density can be obtained from a minimization of the total energy functional with a variational method by solving the so-called Kohn-Sham equations. The global minimum value of the functional determines the exact ground state of the system. This way it is possible to obtain the total energy of the system and the forces which act on the atoms, two quantities which are needed in all the simulations performed in this work.

In principle, DFT is an exact method, but the minimization of energy is far from trivial. Kohn and Sham [11] introduced a method which replaces the many-body problem with an auxiliary system of non-interacting particles, allowing a fast solution of the eigenvalue problem. What needs to be added in this treatment is an additional functional which describes exchange and correlation. Nowadays one of the greatest challenges in DFT consists in the search for an accurate expression for the exchange-correlation functional. The simplest is known as Local Density Approximation (LDA) initially proposed by Kohn and Sham [11] and can also be adapted to include spin in the Local Spin Density Approximation (LSDA) [12]. A more refined method is the Generalized Gradient Approximation (GGA) which involves the calculation of the gradient of electron density and includes functionals such as B88 [13], LYP [14] and PBE [15, 16], used in this thesis. More recent functionals are the so-called hybrid functionals, which include the Hartree-Fock (HF) exchange, such as B3LYP [13, 14, 17], which is a combination of B88, LYP and LDA with HF, and PBE0 [18], which mixes PBE with HF. These functionals can give a more accurate electronic description of the system but are computationally very expensive. As compromise between accuracy and computational cost, what is often performed in the simulation of this thesis is a geometry optimization with PBE, and a single point calculation to refine the energies with B3LYP.

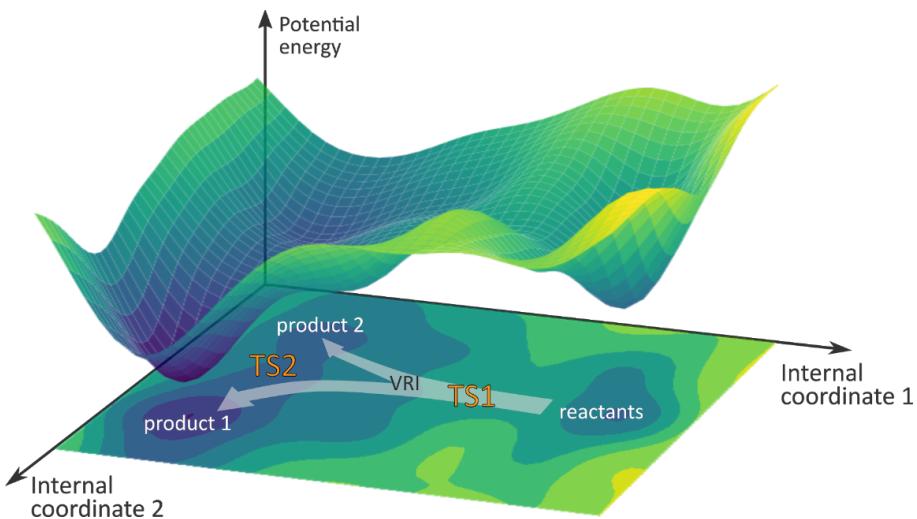
## Dispersion interactions

In this thesis we often encounter non-covalent interactions which need to be treated with high accuracy, such as the adsorption of guest molecules on the Zr Lewis acid sites or interaction between solvent molecules. One of the challenges of DFT methods is the description of long range dispersive interactions such as

London forces, which are commonly referred to as van der Waals interactions. These interactions are due to many particle electron correlation effects which are present also in absence of charges and can have a significant impact on the non-covalent interaction energy. To tackle this problem, various dispersion schemes have been proposed. One of the most used is currently the Grimme–D3 method [19], where a damped  $-C_6R^{-6}$  function is added to the DFT functional. Recently, more advanced dispersion schemes have been developed, such as the many body dispersion scheme [20], or the one of Tkatchenko and collaborators [21], although for the systems we are studying not many benchmarks of these new methods have been performed so far [22].

## Geometry optimization

In order to obtain molecular structures that have physical significance and their relative energies, the arrangement of the atoms needs to be optimized. There are generally two types of molecular structures that we need to find in our simulations, the equilibrium geometries, which correspond to minima of the PES, and the transition state geometries, which correspond to first order saddle points, as displayed in Figure 2.5. These points are characterized by null first derivatives of the energy (the total forces acting on each atom are sufficiently close to zero), all positive second derivatives for local minima and one negative second derivative for first order saddle points, which correspond to transition states.



**Figure 2.5:** Schematic representation of the potential energy surface and the stationary points.

The geometrical optimization of reactants and products consists in a minimization of the energy along the nuclear coordinates. Often the starting point is the experimental structure which can be obtained from diffraction data. In the most

used codes several minimization methods are implemented, each characterized by a different computational cost and robustness, such as steepest descent, conjugated gradient or simulated annealing. The algorithms will find local minima, and do not guarantee that the system will be in a global minimum, therefore the minimizations must start from a sufficiently good guess. The search for transition states is far from trivial and often requires an iterative procedure involving different methods and requiring a good knowledge about the system and chemical process under study. In the calculations performed in this thesis, we often start from an equilibrium structure and as a first guess, we adapt the bond lengths and angles to be close to the transition state with a molecular editor such as Zeobuilder [23]. These bond lengths are then fixed, and the rest of the structure is reoptimized. The Hessian of this partly optimized structure needs to be then computed, and the vibrational modes analyzed, to check which (if any) negative frequency corresponds to the transition state. The system can be optimized again without the constraint using an improved dimer method along a selected eigenvector [24] which is followed by the optimization with a quasi-Newton method [25]. In some difficult cases, the TS search can be initiated on simpler cluster models and optimized in a code in which methods are usually implemented to directly find the TS structure. This way, a first guess of the TS geometry might be obtained and further transferred to the periodic model. In both transition states and minima, if there are superfluous negative frequencies, these need to be removed. Often, it is sufficient to minimize the energy along that vibrational mode. Single point energy calculations can be performed for different values of the displacement along that mode, and the lowest point in energy can be used as starting point for a subsequent minimization of all the coordinates.

### Cell optimization

In the case of periodic systems, not only the structure, but also the unit cell needs to be optimized. This is not trivial, as when using a finite plane wave basis set the number of plane waves depends on the volume of the unit cell. If the volume changes during the optimization, artificial forces which go under the name of Pulay stress can arise. This would require many iterations to optimize the volume. In this thesis, another approach was used [26] which relies on an equation of state fit. For a given volume, for instance taken from experimental data, the unit cell is optimized. Then a set of equally spaced different volumes is defined and for each of these points the geometry and unit cell parameters are optimized. This way it is possible to construct an energy-volume curve, which for a rigid system can be fitted with a Birch-Murnaghan equation of state, allowing to extract the volume  $V_0$  which corresponds to the minimum electronic energy  $E_0$ .

$$E(V) = E_0 + \frac{9V_0B_0}{16} \left\{ \left[ \left( \frac{V_0}{V} \right)^{\frac{2}{3}} - 1 \right]^3 B'_0 + \left[ \left( \frac{V_0}{V} \right)^{\frac{2}{3}} - 1 \right]^2 \left[ 6 - 4 \left( \frac{V_0}{V} \right)^{\frac{2}{3}} \right] \right\}$$

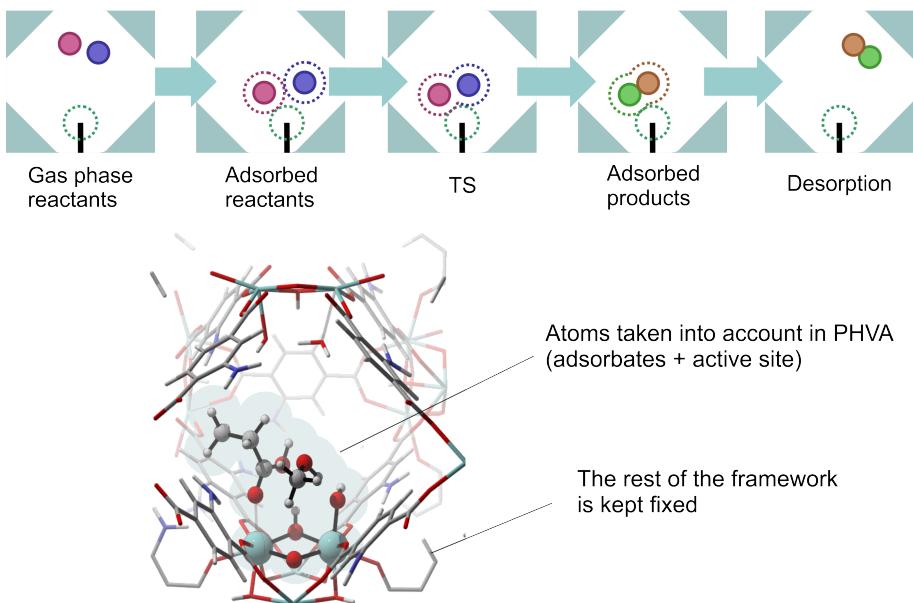
Where  $B_0$  and  $B'_0$  are the bulk modulus and its derivative. A new structure is then generated at this given volume and coordinates and unit cell parameters are optimized again.

## Molecular vibrations

As seen in the previous paragraph, for many purposes in this thesis we need to calculate the second order derivatives (Hessian matrix) of the PES, which are associated to molecular vibrations. First of all, the Hessian gives us information about the curvature of the surface and the nature of the stationary points encountered during the minimization. The second order derivatives are obtained by displacing the atoms in the three directions and calculating the energies, then the Hessian is diagonalized to determine the eigenvectors that correspond to the vibrational motions. From the Hessian we can calculate the vibrational frequencies, which open the door to a lot more information on the system than a single point calculation. Single point calculations are performed at 0 K, but even at this temperature nuclei vibrate around their equilibrium positions, and this movements are responsible for the vibrational entropy. We can approximate these motions with those of harmonic oscillators, by using the vibrational frequencies constructed from the Hessian. These frequencies can then be used to estimate the value of the vibrational entropy at finite temperatures, as will be explained later. In the calculations performed in this thesis, due to computational limits, a partial Hessian approach was used when dealing with reactions, as implemented in the TAMkin toolkit [27]. The quantity that needs to be derived from these calculations is the change in free energy, which mainly depends on the parts of the system that change during the reaction, in the case of a heterogeneous catalyst the active site and the adsorbed reactants. Therefore, restricting the entropy calculations only to this part of the system is a good approximation that allows to decrease enormously the computational cost [28]. This approach has been used in the calculation of the free energy barriers for the Fischer esterification on UiO-66 (**PAPER I**), where the atoms taken into account were the adsorbed reactants and four atoms of the active sites in their immediate proximity, as displayed in Figure 2.6.

## 2.3 Free energy

The central thermodynamic quantity that determines the outcome of a reaction is the free energy change associated to the process. In general, a chemical system will undergo changes in a direction that minimizes its free energy, until an equilibrium is reached. Knowing the difference in free energy between reactants and products allows us to know the equilibrium constant for a given reaction. The Gibbs Free energy can be decomposed in an enthalpic and an entropic contribution, that can be evaluated from the simulations knowing the molecular partition functions. Initially, the total internal energy  $U$  has to be obtained from the electronic energy



**Figure 2.6:** representation of the atoms taken into account in the PHVA approach. Top: a schematic representation of a reactive process in nanoporous material, bottom: a snapshot from the static calculations where the atoms of the active site and the adsorbates are highlighted

$\varepsilon_0$ , the zero-point vibrational energy  $E_{ZPE}$  and the molecular partition function  $Q$  at constant number of particles  $n$  and volume  $V$ :

$$U = U_0 + RT^2 \left( \frac{\partial \ln Q}{\partial T} \right)_{n,V}$$

$$U_0 = \varepsilon_0 + E_{ZPE}$$

Where  $R$  is the gas constant, equal to the product  $N_A \cdot k_B$  between Avogadro's number and Boltzmann constant. The molecular partition function  $Q$  can be split in its translational, rotational, and vibrational components:

$$Q = Q_{trans} Q_{rot,ext} Q_{vib}$$

The enthalpy  $H$  corresponds to the total energy plus the work associated to the change in volume.

$$H = U + p_0 V$$

The entropy  $S$  can be directly obtained from the partition function:

$$S = R \ln Q + RT \left( \frac{\partial \ln Q}{\partial T} \right)_{n,V}$$

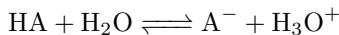
Finally, the Gibbs free energy  $G$  will be:

$$G = H - TS = U_0 + p_0 V - RT \ln Q$$

Where in the case of non-interacting particles  $p_0 V = RT$  following the ideal gas law.

## Equilibrium

As explained above, there is a tight connection between free energy and equilibrium concentrations in chemical reactions. As an example, an equilibrium which is of outmost importance in chemistry is the acidic dissociation of species in aqueous solution:



The acidic dissociation constant ( $K_a$ ) is an important equilibrium constant in chemistry and is equal to the ratio between the concentration of products and reactants when the reaction reaches the equilibrium. It is often reported with its negative decimal logarithm as  $pK_a$ .

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

The equilibrium constant is equal to the Gibbs free energy change from reactants to products.

$$\Delta G = RT \ln K_a$$

$$pK_a = \frac{\ln 10}{RT} \Delta G$$

## Transition state theory

A chemical reaction is a process that through rearrangement of the atoms transforms one stable state into another. Every elementary reaction can be represented as a minimum energy path connecting two minima along the potential energy surface. Furthermore, along this reaction path the existence of a saddle can be postulated, which is the highest point in energy that needs to be crossed to go to the product state. The saddle point is typically called transition state or activated complex and it is the basis for the transition state theory developed by Eyring in the 1930's, one of the most successful chemical theories which allows to explain reaction rates of elementary chemical reactions. The assumption of the theory is that there is a quasi-equilibrium between reactants and activated complex, and the rate constant can be obtained by the size of the energy barrier and by the frequency at which the system can cross the barrier. This is possible because the barrier acts as a bottleneck in the reaction, its crossing is a rare event and all the kinetics depends only on it. For a unimolecular reaction, the rate constant can be derived from the partition functions of reactant, transition state and their energy difference:

$$k(T) = \frac{k_B T}{h} \frac{q_{TS,\ddagger}}{q_R} e^{-\frac{\Delta E^\ddagger}{k_B T}}$$

Where  $k_B$  is the Boltzmann constant,  $h$  is the Planck constant,  $q_R$  and  $q_{TS,\ddagger}$  are the molecular partition functions of reactants and activated complex for all coordinates except the reaction coordinate, evaluated from the zero-point vibrational level.

$$q_{vib,i} = \prod_{i=1}^{N_{dof}} \frac{1}{1 - e^{-\frac{h\nu_i}{k_B T}}}$$

Where  $N_{dof}$  are the number of vibrational degrees of freedom of the system. The energy difference  $\Delta E^\ddagger$  includes electronic energy and zero-point vibrational energy difference at 0 K:

$$\Delta E^\ddagger = E_0^{TS} - E_0^R + \Delta E_{0,vib}$$

$$\Delta E_{0,vib} = \sum_{i=0}^{N_{dof}-1} \frac{h\nu_i^{TS,\ddagger}}{2} - \sum_{i=0}^{N_{dof}} \frac{h\nu_i^R}{2}$$

This theory has some limitations, and may fail in the case of labile intermediates, when nuclei deviate from a classical behavior, or at high temperatures. For a given reaction, in fact, there will be many paths characterized by different barriers, and at low temperature only the lowest one will be likely to be crossed. When the kinetic energy is high enough, many other paths will be activated. The transition state will occupy a larger region of the PES, and it will not be possible to derive entropy from the vibrational partition functions.

## 2.4 Exploring the free energy surface

Static calculations, where molecular vibrations are approximated using harmonic oscillators, can fail to give an accurate representation of the entropy when there is a high configurational freedom. When the PES is flat with respect to  $k_B T$ , the system at equilibrium can evolve in a larger region of the PES and move along more than one minimum. In this case, vibrational frequencies are anharmonic and it is not possible to represent the system by approximating around one single minimum. Therefore, static calculations are not always sufficient in describing the system at operating conditions. In this view, molecular dynamics (MD) techniques, which follow the time evolution of the system, can resolve this shortcoming.

### Ab initio Molecular Dynamics

From MD simulations, thermodynamic properties such as free energy can be obtained taking into account a whole region of the PES instead of a single point. This is based on the ergodic theorem, that states that the time average of equilibrium properties is equal to the ensemble average, in the limit of a sufficient long simulation. MD simulations are based on solving Newton's equations of motion:

$$M_i \ddot{\mathbf{R}}_i = \mathbf{F}_i = -\nabla_i V$$

where  $M_i$  and  $\mathbf{R}_i$  are the mass of a given nucleus and its coordinates,  $\mathbf{F}_i$  the forces that act on it, which correspond to the gradient  $\nabla_i V$  of the PES. There are many ways to calculate these quantities and to integrate the equations of motion, and at present time, chemists and physicists can choose between a plethora of MD techniques which span a whole range of complexity, accuracy and computational cost. In the calculations performed in this thesis, potential energy and forces on the PES are calculated from first principles by means of DFT to account for the full dynamic behavior of the material by ab initio molecular dynamics (AIMD). The calculation of electronic properties which define the PES is decoupled from the propagation of nuclear motions, in a method called Born-Oppenheimer Molecular Dynamics (BOMD). Other famous AIMD methods, which differ by how the calculations of electronic potential and the equation of motion are combined, are the Car-Parrinello MD (CPMD) [29], where a fictitious electronic kinetic energy is added to the lagrangian, or the Ehrenfest MD, based on the namesake theorem [30, 31]. The first MD calculations were performed in the microcanonical (NVE) ensemble, where total energy, number of particles and volume are fixed. However, in experiments it is often the temperature that is fixed, not the energy. In general, the choice of the ensemble depends on the thermodynamic quantities that need to be determined. Nowadays there are many thermodynamic ensembles in which the simulation can be performed. The most convenient for a comparison with experiments are the canonical (NVT), with fixed number of molecules, volume and temperature, or the isothermal-isobaric (NpT), with fixed number of molecules and temperature, but where the volume can fluctuate. In order to have a fixed average

temperature, some control of the kinetic energy of the atoms is needed. Various thermostats, which differ by speed and robustness, are implemented in every MD code. In this thesis, Nose'-Hoover thermostat was used, where the system is connected to a heat bath. The pressure is also controlled in simulations by means of a barostat. The most commonly used is the one developed by Martyna, Tobias, and Klein (MTK) [32].

### Towards modeling at operating conditions

With the growth in computational time, the new challenge is constituted by modeling the system at operating conditions. Many chemical reactions, especially when performed at mild conditions, involve the presence of a solvent. In order to move closer to modeling the system at operating conditions, the solvent in the pores can also be taken into account. This adds a lot of degrees of freedom to the system, and for this reason often an implicit description of the solvent is done, such as in the Periodic Continuum Model (PCM). In the case of the work in this thesis, however, it is necessary to fully model the solvent molecules, as they are actively involved in proton transfers. To do so, the number of solvent molecules that can fit in the unit cell needs to be estimated. Monte Carlo method (MC) is an alternative approach to MD to explore the PES for complex systems. It was initially developed for the calculation of multidimensional integrals and is nowadays largely used in chemistry, especially when dealing with adsorption. In the framework of this thesis, it has been applied in the Gran Canonical ensemble (fixed chemical potential, volume and temperature) to determine the number of solvent molecules that could fill the pores of the material at standard conditions.

### Enhanced MD methods

MD simulations can offer valuable insights into the behaviour of a chemical system at equilibrium conditions. From these simulations, many properties can be extracted, such as equilibrium geometries, vibrational spectra, diffusion coefficients, structural parameters etc. Configurations associated to higher (or lower) values of potential energies will be sampled for shorter (or longer) times, and in principle, if a certain process is sufficiently sampled, based on the ergodic theorem we can know its equilibrium constant, and in turn the free energy barrier associated to it. However, chemical reactions, where bonds are broken and formed, are generally rare events that will not be sampled with a regular exploration of the PES. If the barrier is high compared to the  $k_B T$ , the probability that such event would spontaneously occur during the simulation time is practically none. For this reason, different enhanced sampling techniques have been developed to enhance the sampling of certain regions of the PES.

### Choice of collective variable

The PES is a highly dimensional surface, defined by the positions of all atoms in the system. However, often a reactive process can be described by few important coordinates called “collective variables” (CV), which are projections of the high dimensional space. In fact, what can be considered a chemical configuration is an ensemble of microstates that are different in terms of absolute coordinates of each atom, but all contribute to the same macrostate. In the simplest case, a single CV can represent the reaction coordinate for the process.

The choice of the right collective variable is crucial to describe the correct process.

what are the main choices, distances, CN, angles

### Metadynamics

### Umbrella sampling

## **Part II**

# **Published papers**



# Bibliography

- [1] K. D. Wispelaere, L. Vanduyfhuys, and V. V. Speybroeck, "Chapter 6 - entropy contributions to transition state modeling," in *Modelling and Simulation in the Science of Micro- and Meso-Porous Materials* (C. R. A. Catlow, V. V. Speybroeck, and R. A. van Santen, eds.), pp. 189 – 228, Elsevier, 2018.
- [2] P. P. Ewald, "Die berechnung optischer und elektrostatischer gitterpotentiale," *Annalen der Physik*, vol. 369, no. 3, pp. 253–287, 1921.
- [3] J. H. Cavka, S. Jakobsen, U. Olsbye, N. Guillou, C. Lamberti, S. Bordiga, and K. P. Lillerud, "A new zirconium inorganic building brick forming metal organic frameworks with exceptional stability," *Journal of the American Chemical Society*, vol. 130, no. 42, pp. 13850–13851, 2008. PMID: 18817383.
- [4] S. M. J. Rogge, A. Bavykina, J. Hajek, H. Garcia, A. I. Olivos-Suarez, A. Sepúlveda-Escribano, A. Vimont, G. Clet, P. Bazin, F. Kapteijn, M. Daturi, E. V. Ramos-Fernandez, F. X. Llabrés i Xamena, V. Van Speybroeck, and J. Gascon, "Metal–organic and covalent organic frameworks as single-site catalysts," *Chem. Soc. Rev.*, vol. 46, pp. 3134–3184, 2017.
- [5] M. Born and R. Oppenheimer, "Zur quantentheorie der moleküle," *Annalen der Physik*, vol. 389, no. 20, pp. 457–484, 1927.
- [6] M. Ceriotti, W. Fang, P. G. Kusalik, R. H. McKenzie, A. Michaelides, M. A. Morales, and T. E. Markland, "Nuclear quantum effects in water and aqueous systems: Experiment, theory, and current challenges," *Chemical Reviews*, vol. 116, no. 13, pp. 7529–7550, 2016. PMID: 27049513.
- [7] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard, "Reaxff: A reactive force field for hydrocarbons," *The Journal of Physical Chemistry A*, vol. 105, no. 41, pp. 9396–9409, 2001.
- [8] L. H. Thomas, "The calculation of atomic fields," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 23, no. 5, p. 542–548, 1927.
- [9] E. Fermi, "Eine statistische methode zur bestimmung einiger eigenschaften des atoms und ihre anwendung auf die theorie des periodischen systems der elemente," *Zeitschrift für Physik*, vol. 48, pp. 73–79, Jan 1928.

- [10] P. Hohenberg and W. Kohn, "Inhomogeneous electron gas," *Phys. Rev.*, vol. 136, pp. B864–B871, Nov 1964.
- [11] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," *Phys. Rev.*, vol. 140, pp. A1133–A1138, Nov 1965.
- [12] S. H. Vosko, L. Wilk, and M. Nusair, "Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis," *Canadian Journal of Physics*, vol. 58, no. 8, pp. 1200–1211, 1980.
- [13] A. D. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," *Phys. Rev. A*, vol. 38, pp. 3098–3100, Sep 1988.
- [14] C. Lee, W. Yang, and R. G. Parr, "Development of the colle-salvetti correlation-energy formula into a functional of the electron density," *Phys. Rev. B*, vol. 37, pp. 785–789, Jan 1988.
- [15] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.*, vol. 77, pp. 3865–3868, Oct 1996.
- [16] J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple [phys. rev. lett. 77, 3865 (1996)]," *Phys. Rev. Lett.*, vol. 78, pp. 1396–1396, Feb 1997.
- [17] A. D. Becke, "Density-functional thermochemistry. iii. the role of exact exchange," *The Journal of Chemical Physics*, vol. 98, no. 7, pp. 5648–5652, 1993.
- [18] C. Adamo and V. Barone, "Toward reliable density functional methods without adjustable parameters: The pbe0 model," *The Journal of Chemical Physics*, vol. 110, no. 13, pp. 6158–6170, 1999.
- [19] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu," *The Journal of Chemical Physics*, vol. 132, no. 15, p. 154104, 2010.
- [20] T. Bučko, S. Lebègue, T. Gould, and J. G. Ángyán, "Many-body dispersion corrections for periodic systems: an efficient reciprocal space implementation," *Journal of Physics: Condensed Matter*, vol. 28, p. 045201, jan 2016.
- [21] A. Ambrosetti, A. M. Reilly, R. A. DiStasio, and A. Tkatchenko, "Long-range correlation energy calculated from coupled atomic response functions," *The Journal of Chemical Physics*, vol. 140, no. 18, p. 18A508, 2014.
- [22] J. Wieme, K. Lejaeghere, G. Kresse, and V. Van Speybroeck, "Tuning the balance between dispersion and entropy to design temperature-responsive flexible metal-organic frameworks," *Nature communications*, vol. 9, no. 1, p. 4899, 2018.

- [23] T. Verstraelen, V. Van Speybroeck, and M. Waroquier, "Zeobuilder: A gui toolkit for the construction of complex molecular structures on the nanoscale with building blocks," *Journal of Chemical Information and Modeling*, vol. 48, no. 7, pp. 1530–1541, 2008. PMID: 18543904.
- [24] A. Heyden, A. T. Bell, and F. J. Keil, "Efficient methods for finding transition states in chemical reactions: Comparison of improved dimer method and partitioned rational function optimization method," *The Journal of Chemical Physics*, vol. 123, no. 22, p. 224101, 2005.
- [25] W. H. Press, W. H. Press, B. P. Flannery, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, and W. T. Vetterling, *Numerical recipes in Pascal: the art of scientific computing*, vol. 1. Cambridge University Press, 1989.
- [26] D. E. Vanpoucke, K. Lejaeghere, V. Van Speybroeck, M. Waroquier, and A. Ghysels, "Mechanical properties from periodic plane wave quantum mechanical codes: The challenge of the flexible nanoporous mil-47 (v) framework," *The Journal of Physical Chemistry C*, vol. 119, no. 41, pp. 23752–23766, 2015.
- [27] A. Ghysels, T. Verstraelen, K. Hemelsoet, M. Waroquier, and V. Van Speybroeck, "Tamkin: a versatile package for vibrational analysis and chemical kinetics," 2010.
- [28] A. Ghysels, D. Van Neck, and M. Waroquier, "Cartesian formulation of the mobile block hessian approach to vibrational analysis in partially optimized systems," *The Journal of chemical physics*, vol. 127, no. 16, p. 164108, 2007.
- [29] R. Car and M. Parrinello, "Unified approach for molecular dynamics and density-functional theory," *Physical review letters*, vol. 55, no. 22, p. 2471, 1985.
- [30] P. Ehrenfest, "Bemerkung über die angenäherte Gültigkeit der klassischen Mechanik innerhalb der Quantenmechanik," *Zeitschrift für Physik A Hadrons and Nuclei*, vol. 45, no. 7, pp. 455–457, 1927.
- [31] D. Marx and J. Hutter, *Ab initio molecular dynamics: basic theory and advanced methods*. Cambridge University Press, 2009.
- [32] G. J. Martyna, D. J. Tobias, and M. L. Klein, "Constant pressure molecular dynamics algorithms," *The Journal of Chemical Physics*, vol. 101, no. 5, pp. 4177–4189, 1994.

