



Edited by Vincent Blay, Luis F. Bobadilla, Alejandro Cabrera-García

Zeolites and Metal-Organic Frameworks

From Lab to Industry

Amsterdam
University
Press

Amsterdam University Press



ATLANTIS
PRESS

Cover illustration: © 2010 James Coleman

Cover design: Coördesign, Leiden

Lay-out: Konvertus/Newgen

Amsterdam University Press English-language titles are distributed in the US and Canada by the University of Chicago Press.

ISBN 978 94 6298 556 8

e-ISBN 978 90 4853 671 9

DOI 10.5117/9789462985568

NUR 971

Creative Commons License CC BY NC ND

(<http://creativecommons.org/licenses/by-nc-nd/3.0>)



© The authors / Atlantis Press B.V. / Amsterdam University Press B.V., Amsterdam 2018

All rights reserved. Without limiting the rights under copyright reserved above, no part of this book may be reproduced, stored in or introduced into a retrieval system, or transmitted, in any form or by any means (electronic, mechanical, photocopying, recording or otherwise) without the written permission of both the copyright owner and the author of the book.

Every effort has been made to obtain permission to use all copyrighted illustrations reproduced in this book. Nonetheless, whosoever believes to have rights to this material is advised to contact the publisher.

Amsterdam University Press

9. Computational Chemistry Experiment Possibilities

Bartłomiej M. Szyja¹ and Danny Vanpoucke^{2,3}

¹Wrocław University of Science and Technology, Faculty of Chemistry, Division of Fuels Chemistry and Technology, ul. Gdańska 7/9, 50–344 Wrocław, Poland.

²UHasselt - Hasselt University, Institute for Materials Research (IMO), 3590 Diepenbeek, Belgium.

³IMOMECE, IMEC vzw, 3590 Diepenbeek, Belgium.

Keywords: Force fields, density functional theory, potential energy surface, molecular dynamics

1. Introduction

Thanks to a rapid increase in the computational power of modern CPUs, computational methods have become a standard tool for the investigation of physico-chemical phenomena in many areas of chemistry and technology. The area of porous frameworks, such as zeolites, metal-organic frameworks (MOFs) and covalent-organic frameworks (COFs), is not different. Computer simulations make it possible, not only to verify the results of the experiments, but even to predict previously inexistent materials that will present the desired experimental properties. Furthermore, computational research of materials provides the tools necessary to obtain fundamental insight into details that are often not accessible to physical experiments.

The methodology used in these simulations is quite specific because of the special character of the materials themselves. However, within the field of porous frameworks, *density functional theory* (DFT) and *force fields* (FF) are the main actors. These methods form the basis of most computational studies, since they allow the evaluation of the *potential energy surface* (PES) of the system.

2. Energy calculations

Force fields

The simplest, and consequently least computationally demanding, way of calculating the energy of a system is through the force field method, sometimes referred to as *Molecular Mechanics* (MM). This method evaluates the energy of the system by making use of classical mechanics concepts and formulations. The specific energy expression is given by a sum of energy terms. These can be single particle terms (the kinetic energy terms) or multi-particle terms (the potential energy terms), the latter are often parameterized. For example, bond vibrations can be described using the harmonic oscillator. In this case, the energy expression would take the form:

$$E = \frac{1}{2} kx^2 \quad (1)$$

where E is the potential energy of the system, k is the force constant and x is the displacement from the optimal bond length. This approach uses only one parameter: k . The harmonic oscillator approximates the energy of the system reasonably well for displacements close to the equilibrium position; however, far from the equilibrium deviations become large. As a result, it fails to describe the dissociation of a bond, which is not captured at all by the parabolic curve of the harmonic oscillator.

In contrast, the Morse potential is a much better approximation of the energy profile of a bond (Figure 1), but it has a slightly more complex energy expression with three parameters. The Morse potential has the following form:

$$E = D(1 - e^{-a(r_0 - r)})^2 \quad (2)$$

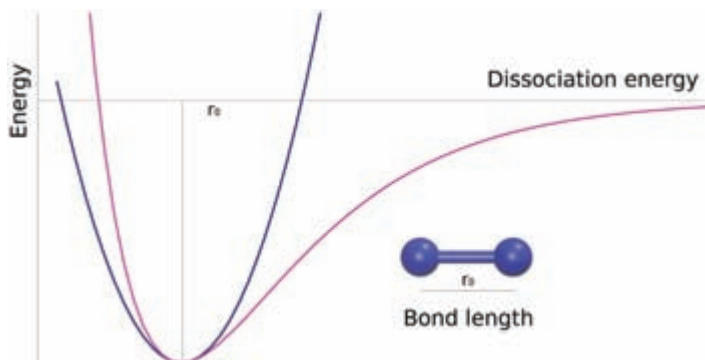


Figure 1. Harmonic oscillator (blue) and Morse (pink) potential energy curves. r_0 represents the optimal bond length of a diatomic molecule, the minimum of the potential energy.

where D is the dissociation energy, r_0 gives the equilibrium position, and a controls the width of the potential well.

Although the Morse potential only describes diatomic molecules, it has the right shape to describe any bond between two atoms. As such, similar terms can be added to describe bonds in case of more complex geometries. Because any given atom of a system undergoes several interactions at the same time, a typical force field has an energy expression consisting of several terms representing the different types of interactions:

$$E = E_{\text{bond}} + E_{\text{angle}} + E_{\text{dihedral}} + E_{\text{vdW}} + E_{\text{coulomb}} \quad (3)$$

The first three terms describe the bonding or intramolecular interactions, namely the energies related to the stretching of the covalent bonds, to the bending of the angles formed by three atoms, and to the twisting of the dihedral angles along the bonds. These interactions are usually described with the energy expression of the harmonic oscillator (equation 1). The last two terms describe the non-bonding or intermolecular interactions. The interactions of the van der Waals type are usually described by the Lennard-Jones potential:

$$E_{\text{vdW}} = A / r^{12} - B / r^6 \quad (4)$$

where A and B are the specific parameters for the interaction.

The Coulomb –or electrostatic– interactions depend on the charges q_1 and q_2 in the system:

$$E_{\text{Coulomb}} = (q_1 \cdot q_2) / (\epsilon \cdot r) \quad (5)$$

The parametrization of a force field is a very complex and difficult procedure. In almost all cases, the parameters need to match one another, and even changing the atomic charge calculation procedure is not something that should be taken lightly. Instead, one usually wants to use the force field assigned charges.

Several force fields have been developed for the study of porous frameworks. One example of a zeolite force field is proposed by van Beest, Kramer and van Santen (BFS force field) [1]. It is used to model the properties of aluminosilicates, such as pore sizes [2]. While this force field can only be applied to model the intrinsic properties of the solid, other force fields are specialized in the host-guest interactions, such as the YAFF force field used to study the interaction between CO_2 molecules and differently functionalized MIL-47(V) MOFs [3].

Besides the obvious advantage of simplicity, force field methods have important limitations as well. These result from neglecting any type of electronic interactions. For example, this method cannot be used to investigate optical effects. More importantly, bond formation and cleavage, which are the result of electronic interactions, render force fields useless for the investigation of catalytic phenomena or chemisorption.

Density Functional Theory

Electronic phenomena can, however, be investigated using Quantum Mechanical methods. These methods are often referred to as *ab-initio* (meaning ‘from the beginning’ or ‘first principles’)¹ because, unlike force fields, they only require physical constants and involve no fitting to other data. In the case of porous periodic frameworks, the vast majority of quantum mechanical calculations are performed using Density Functional Theory (DFT). Such calculations are far more complex, time- and resources-demanding than force field calculations. This puts a limit on the available size of the computational model, and leads to so-called “model” vs. “method” accuracy, which will be discussed later on.

DFT is the current gold standard in the description of chemical reactivity and, along with *ab-initio* methods such as Hartree-Fock or post-Hartree-Fock, it is capable of describing the electronic state of the system. Contrary to *ab-initio* methods, DFT does not involve solving the Schrödinger equation to find the many-electron wave function of the system, but it solves instead a different set of equations (the Hohenberg-Kohn-Sham equations) which relate the total energy of the system with the electron density.

One of the attractive features of the DFT method is the fact that the electron density can be directly and experimentally measured (e.g. through X-ray diffraction) and is therefore much easier to understand than the concept of wave function, which is not a quantum mechanical observable. Moreover, the electron density is only a function of position, thus it only depends on three parameters: the spatial coordinates. This makes it much easier to treat mathematically.

The foundations of the DFT method were provided by the work of Hohenberg, Kohn and Sham [4,5], who demonstrated that the properties of the system are a functional of the electron density. Furthermore, the total energy can be decomposed into the kinetic energy of the electrons and the potential energy of nuclear-electron and electron-electron interactions. The motion of the nuclei can be neglected – this is called

the *Born-Oppenheimer approximation* and is based on the fact that nuclei are much heavier than electrons. As a result, the nuclear motion is orders of magnitude slower than the electronic motion, such that the former can be considered static on the timescales relevant for electron interactions.

As a result, the expression for the total energy of the system can be written as a sum of four terms:

$$E = T_{\text{e}} + \int v(r)\rho(r)dr + J[\rho] + E_{\text{xc}}[\rho] \quad (6)$$

where T_{e} is the kinetic energy of the electrons, the second term describes the interactions between the nuclei and the electrons, and $J[\rho]$ is the term for the static Coulombic self-interaction of the electron cloud. These terms can be calculated directly from fundamental physical expressions. This is not the case for the exchange-correlation term, E_{xc} . This term is needed because the J term does not take into account the repulsion of the electrons caused by the correlation of their motion or the repulsion of the electrons with the same spin due to the Pauli exclusion principle. The XC subscript in this term represents these two factors: eXchange and Correlation. Although DFT is in theory an exact method, its practical implementation is approximate because no general analytic formulation of the E_{xc} functional is known. As a result of the never ending quest for this functional, many differently defined exchange-correlation functionals have seen the light of day.

The most simple and least accurate approaches rely only on the value of the electron density in the given spatial position, and are known as LDA (Local Density Approximation). A better solution is the GGA (Generalized Gradient Approximation), which, in addition to the electron density at the given point, also takes into account its gradient (the first derivative of the density). This results in an improvement of the accuracy of the method but, obviously, at the expense of a (slight) increase in the computational effort. With this in mind, it was proposed that by adding another term, the second derivative of the density, accuracy could be improved even further. This led to so called meta-GGA functionals, which often offered an improvement although more limited than that of GGA functionals over LDA ones.

Another approach to improve the E_{xc} functional is through the inclusion of the exact Hartree-Fock exchange energy in the E_{xc} term in addition to the approximated DFT exchange-correlation energy. This gives a hybrid functional, in which the amount of the HF exchange is defined

by a mixing coefficient, which is fitted to a set of test systems. An important difference between LDA and GGA, on the one hand, and hybrid functionals, on the other, is that the former are local functionals only requiring information on the electron density at the position where the functional is calculated, while the latter are non-local functionals, where information on the electron density in the entire system is needed. This increases the computational cost of hybrid functionals by several orders of magnitude.

Despite the fact that high level or hybrid functionals often predict electronic properties that are in very good agreement with the experiment, the DFT method still suffers from two important shortcomings. One is the poor description of dispersion interactions. As such, when van der Waals interactions are of interest, one needs to extend DFT to account for the dispersion interactions. A simple and computationally inexpensive extension dubbed DFT-D, was devised by Grimme and coworkers, with currently two parametrization sets available in many DFT codes: D2 and D3 [6,7]. The second issue is more fundamental: the electron density is a ground state property and the excited states are usually beyond a reasonable description. This issue can be addressed by so called time-dependent DFT but even then the accuracy is not perfect.

'Partial' solutions

There are two approaches that attempt to combine the advantages of force fields and first principles methods in a single solution. The first one is QM/MM (Quantum Mechanics/Molecular Mechanics) [8]. This method allows a separation of the computational model into two (or more) zones – one zone is described by the force field (MM) and the other by quantum-mechanics (QM). The advantage of this method is that the direct vicinity of the reactive site is described by the method that is fully aware of the electronic effects, and as such allows for investigations of chemical reactions. The rest of the model is described by the force field, and provides a simplified description of the framework (Figure 2).

Although QM/MM methods are generally very efficient compared to purely QM described systems, they are still rather tricky to handle. The regions simulated by the QM need to be carefully selected, because moving the region's boundaries can influence the results as well as the computing time. The choice of the size of the QM region also depends strongly on the range of interactions present and can require sizeable QM regions to get atomic properties converged in the presence of active sites [9].

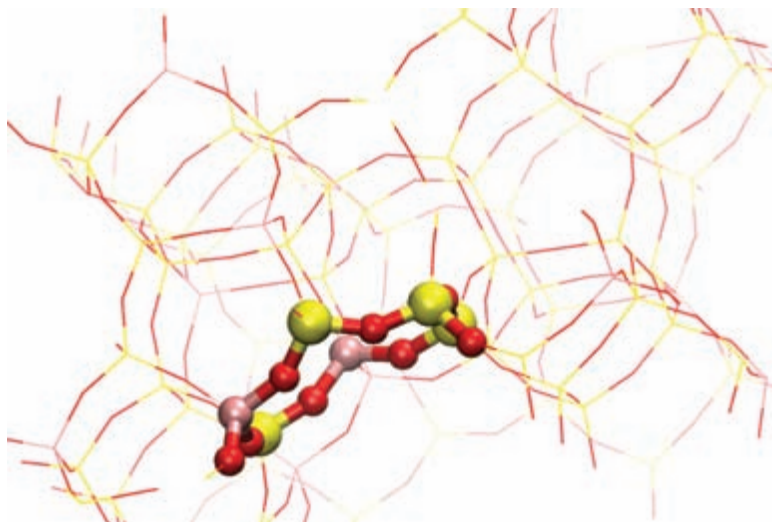


Figure 2. Example of QM/MM regions in TS-1 system, which is isostructural with the MFI framework. The titanium, silicon and oxygen atoms are shown in pink, yellow and red, respectively. The energy of the framework is described by MM, and represented by thin lines, the QM region (the 6-membered ring) is shown as balls and sticks.

Another approach is the *reactive force field*, ReaxFF, developed by Adri van Duin [10]. This method does not split the system into parts, but treats all the atoms on the same level of theory. The calculations are less resource-intensive, because they do not rely on a calculation of the electronic state of the system. Instead, the extensive set of parameters combined with a quite complex energy expression makes it possible to describe different kinds of interactions between the atoms. The most important feature is obviously the formation and cleavage of the chemical bonds. With a proper set of parameters, the reactivity of the system can be investigated with good accuracy. In addition, depending on the direct vicinity of the particular atom, the bonding between it and its neighbour might change – as in singly, doubly or triply bonded C atoms with different number of hydrogens bound to them. The description of the energy of a system using the reactive force field is much more complex than that of classical force fields, although these simulations are still much more efficient than any QM method, which allows investigation of much larger systems.

3. The Potential Energy Surface (PES)

Having the model constructed and the method selected, one can proceed to investigate the reactivity of the system. The reaction pathways run via specific points on the potential energy surface (PES), called *stationary*

points. These points in the potential energy surface are *local minima* (all the gradients are zero) and they reflect stable geometries of the system. Small displacement of the atoms in any direction from the optimal positions increases the potential energy of the system or, in other words, the gradients next to the minimum become positive, which leads to a force pulling them back to their optimal position. Finding the geometries of the systems in those specific points is called *geometry optimization* or *energy minimization*. This is the most basic technique in most areas of computational chemistry and solid-state physics.

Finding a minimum of a function of many variables (coordinates of the atoms) is a non-trivial task and for problems related to real-life materials it cannot be done analytically. Instead, numerical methods are used. The optimization starts with an evaluation of the initial energy of the system and gradients with respect to the position of the atoms. Next, the atoms are displaced according to the gradients, which allow a decrease in the potential energy of the system. This is followed by the evaluation of the potential energy and a calculation of the gradients in the new coordinates. The whole procedure is repeated until the convergence is achieved.

There are different methods for finding the minimum and they generally differ in the way they move towards the minimum (Figure 3). The steepest descent method finds the largest negative gradient in each iteration and follows that direction. This usually leads to oscillations in the followed direction, resulting in more steps being needed. The conjugate gradients (CG) method and the methods derived thereof make a correction in the current gradient based on the previous step(s). This usually allows a reduction of the total number of steps taken. The Newton-Raphson method finds the direction to the minimum based on the calculated curvature of the PES close to the initial point.

As it is the case in almost all areas of computational research, a compromise is often necessary. The CG and Broyden-Fletcher-Goldfarb-Shanno (BFGS) methods have become the current gold standards. Both of them keep track of the previously taken steps and adaptively change the direction. At the same time they keep the computational demand for each iteration relatively small, especially in the limited-memory version of BFGS.

Finding the local minima corresponding to the desired geometries allows one to estimate the thermodynamics of the process. Transition from one minimum to another on the PES corresponds to the change in geometry of the system, where some bonds might be broken or formed, i.e. the

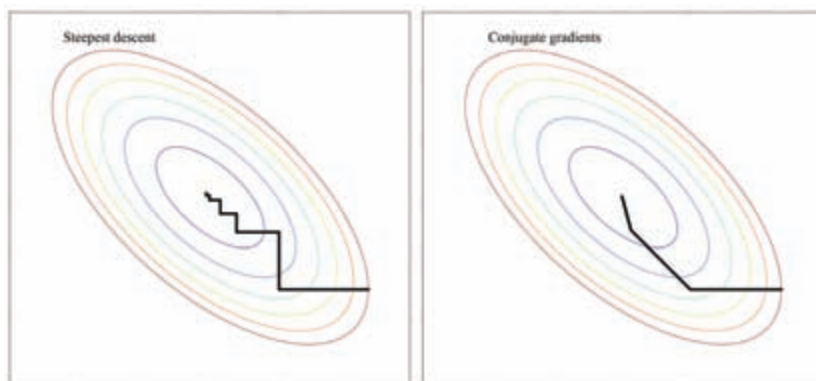


Figure 3. Example of Steepest Descent (left) and Conjugate Gradient (right) method of finding the minimum of a second order polynomial.

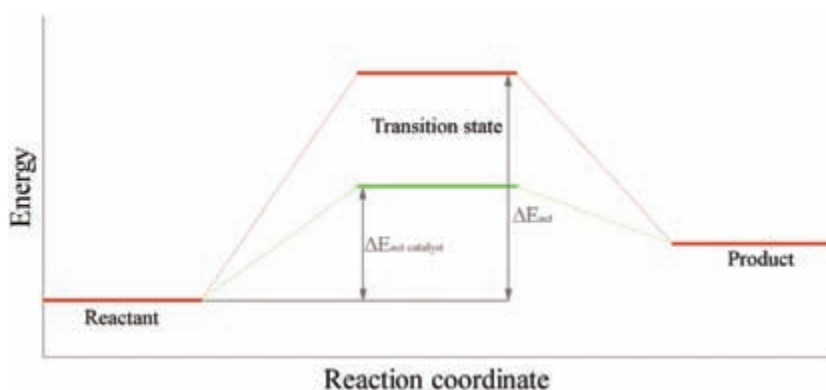


Figure 4. Schematic representation of the transition state stabilization. The non-catalysed path is shown in red, while the catalysed path is shown in green. The presence of the catalyst does not influence the reactant or product energy but it lowers the activation energy.

chemical reaction has occurred. The change in the energy of the system determines if the reaction is exo- or endothermic.

This, however, is not a full or sufficient description of the reactivity, because it is not only thermodynamics that determines the possible reaction pathways. Kinetics is represented on the PES as the height of the barrier in the transition from one minimum to another. These points are extremely important in the studies of catalytic effects. A catalyst increases the rate of the reaction without influencing its thermodynamics because it lowers the activation energy (i.e. the barrier height). This effect is based on the formation of an active complex of the reactant and the catalyst, which is more stable than the transition state in the absence of a catalyst (Figure 4).

The activation barriers correspond to other types of points on the PES, often referred to as *saddle points* or *transition states* (TS). Simple geometry optimizations are insufficient to find these points, and other algorithms are needed. Depending on the model used, different approaches are available.

For cluster models (section 6) the Berny algorithm of ‘eigenvector following’ is usually the method of choice [11]. It is based on the optimization of a candidate for the TS, which might be either guessed at or found from a series of constrained optimizations (e.g. a particular structural parameter related to the reaction coordinate is varied to locate the structure of maximum energy within this trajectory). In the Berny algorithm, the calculation starts with the evaluation of the Hessian matrix (the matrix of second derivatives of the energy, which describes the local curvature of the PES). Then, a structure optimization is carried out in which the algorithm follows the eigenvector corresponding to most negative eigenvalue of the hessian matrix. As a result, the structure will evolve on a path on the PES, which goes through the TS.

Because the calculation of the Hessian in a periodic model is highly time-consuming, another approach is needed. A method used frequently is called the Nudged Elastic Band (NEB) method [12]. This method does not rely on the Hessian matrix. Instead, the pathway leading from the reactant to the product is divided into several steps. The geometries at these steps, called images, are generated by means of interpolation of the atomic positions. The energy of each image is calculated, which results in an energy profile along the path. Subsequently the structure of the images is optimized in parallel. An artificial spring force between the images, similar to springs between beads, is used to keep the images equally spaced. This makes the NEB method especially well-suited for tracking the transition states along the reaction pathway [13]. Another advantage besides avoiding the calculation of the Hessian is that the NEB method does not need a good initial guess. Unfortunately, there are also two main disadvantages: the first is that most of the computation time is spent on calculations of energies and gradients of images other than the TS candidate. The second is that the exact saddle point might be missed, because the energy of the images is minimized and images tend to be fit along the pathway initially interpolated.

An improvement on the NEB method is the Climbing Image variant. This method neglects the forces from the virtual springs on the image with the highest energy. The image is then optimized to maximize its energy along the path defined by the other images (i.e. the *reaction coordinate*) and to minimize its energy in the other directions. The principle of the NEB method is shown in Figure 5.

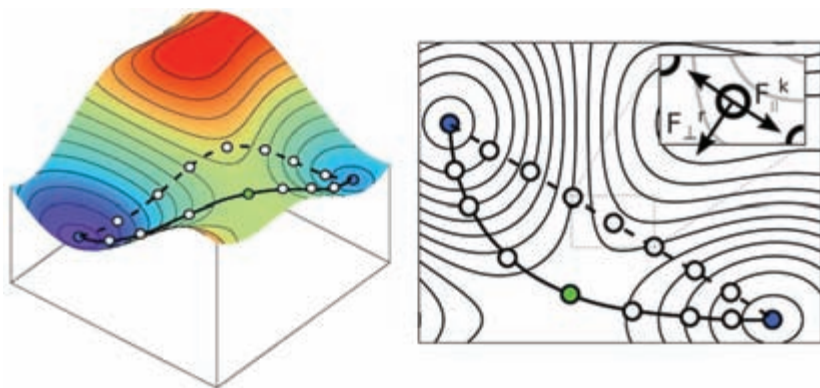


Figure 5. The principle of the NEB method of finding the saddle point between two minima. The climbing image is shown as a green point. Figure by Antoine Kraych. Reprinted with permission from [14].

Regardless of the objective, the optimization procedure is iterative, and as such it never ends exactly on a stationary point but rather approaches it. Hence, a condition is needed upon which the procedure will be terminated. Generally, one needs to calculate the energy of the system and in this case the procedure should be terminated when the energy changes are below a certain threshold. The accuracy of the DFT method is in the order of magnitude of 10 kJ mol^{-1} , thus an energy difference of 1 kJ mol^{-1} is more than sufficient. However, when vibrational frequencies are desired, this criterion might not be sufficient. In this case, it is useful to use additional criteria, such as the maximum force acting on the atom not exceeding $0.01 \text{ eV/\AA}$.

4. Molecular Dynamics

In contrast to the static calculations of the stationary points on the PES, molecular dynamics (MD) allows a study of the motion of particles.

Diffusion of guest molecules

In the case of favourable reaction kinetics, diffusion limitations might play a role in the efficiency of the overall process. This is especially so in small pore size zeolites, in which much effort is devoted to the synthesis of mesoporous or hierarchical materials to alleviate this problem, as discussed in previous chapters of this book. Diffusion of guest molecules is

also important when porous membranes are used, for instance, in the separation of gas mixtures.

Diffusion is a mass transport process and it cannot be investigated by static calculations as is possible for chemical reactions. The motion of the atoms and molecules needs to be considered explicitly. This can be accomplished by a method called Molecular Dynamics. This method calculates the evolution of the system over time by solving Newton's equation of motion:

$$F = ma \quad (7)$$

where F is the force acting on a particular atom, m represents its mass, and a its acceleration. There are many different methods to integrate Newton's equation of motion over time in molecular dynamics and all of them rely on the iterative calculation of the properties of the system in consecutive time steps. Assuming that the time interval between the steps is short, the acceleration can be safely assumed to be constant in this interval, and the equation of motion can be expressed as a Taylor's expansion:

$$r(t + \Delta t) = r(t) + v(t)\Delta t + \frac{1}{2}a(t)\Delta t^2 \quad (8)$$

$$v(t + \Delta t) = v(t) + \frac{1}{2}[a(t) + a(t + \Delta t)]\Delta t \quad (9)$$

where $r(t)$ and $r(t + \Delta t)$ represent the spatial coordinates of the given atom at a given time (t) and after a timestep ($t + \Delta t$), respectively. Analogously $v(t)$ and $v(t + \Delta t)$, and $a(t)$ and $a(t + \Delta t)$ stand for the velocity and the acceleration at these times. These equations are known as the Velocity Verlet algorithm [15].

Repeating these steps iteratively to cover the desired timespan generates a sequence of geometries vs. time, called a trajectory. These points represent different conformations of the same system but they all belong to the same *ensemble*.² The course of the MD simulation is schemed in Figure 6.

Two ensembles are mostly used for simulations in solid frameworks: NVE and NVT. The former represents a micro-canonical ensemble, corresponding to an isolated system with a constant number of particles (N), volume (V), and total energy (E). In this ensemble, neither mass nor energy can be exchanged with the environment. The latter is the canonical ensemble, with temperature (T) held constant by means of a thermostat. This ensemble is better suited, for example, to simulations where adsorption takes place or energy change is expected for another reason. Upon adsorption, energy related to the interactions between the components of the system is released and can be dissipated by the external thermostat.

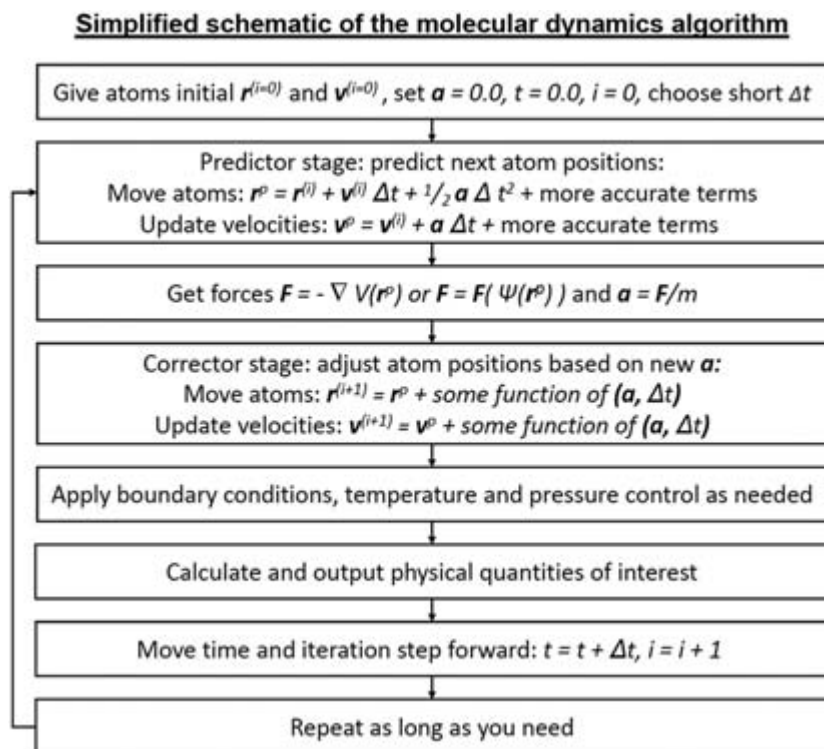


Figure 6. The algorithm for Molecular Dynamics simulation. By Knordlun, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=41419953>.

The NPH and NPT ensembles with constant number of particles (N), pressure (P), enthalpy (H) and temperature (T) represent systems that are in contact with a barostat. These are rarely used in simulation of zeolites because the crystalline structure of these materials is generally rigid.

The proper initiation of an MD simulation is important, because at the beginning the velocities of the atoms are unknown. Typically, initial velocities are assigned randomly but special conditions have to be met. Firstly, the magnitude of the velocity vectors needs to match the desired temperature, especially when the NVE ensemble is used. Secondly, the directions of the velocity vectors must not be aligned otherwise, the centre of mass of the system would translate.

The random numbers used in the initialization of each MD run, obviously, do not depend on physical interactions between the atoms. However, the further the simulation proceeds, the less important they are, because each consecutive iteration of the MD run relies directly on the previous one, and new positions and velocities are obtained by integration of

Newton's equation of motion. Thus, the first part of the trajectory is only used to equilibrate the system in the desired state, and is usually characterized by large fluctuations of temperature or energy. The *equilibration run* is followed by the *data collection* (or *production*) run, for which calculated physical quantities can be trusted.

The length of the run should be sufficiently long to be able to observe the phenomenon of interest. In the case of calculating diffusion coefficients, the length of the simulations should be in the order of nanoseconds. On the other hand, the timestep taken in the run should be small enough to properly sample the vibrations of the atoms. That means that if a timestep is 0.5 fs, 2 million iterations are needed to observe the evolution of the system in 1 ns. Depending on the system size and computational power available, a compromise often has to be made.

Assuming that the MD run was sufficiently long, the diffusion coefficient can be calculated from the mean square displacement (MSD) of the atoms by using Einstein's formula. MSD is defined as follows:

$$MSD = \frac{1}{N} \sum_{n=1}^N (x_n(t) - x_n(o))^2 \quad (10)$$

where N is the total number of particles, $x_n(t)$ and $x_n(o)$ are the positions of the n -th particle at time t and o , respectively. The plot of the MSD as a function of time should be a straight line, the slope being proportional to the diffusion coefficient. If the MSD has a concave curvature, then the motion of the particles is impeded: they cannot diffuse away from its starting point. It has to be stressed that in micro- or mesoporous materials the diffusion of guest molecules is always impeded due to more (zeolites) or less (MOFs) rigid surfaces of the pores. Zeolites of the MOR topology have a large, 12-membered ring pore in z -direction, which allows for relatively easy diffusion. These channels intersect with smaller channels, formed by 8-membered rings, where the diffusion is limited, especially for bulkier molecules. This would be visible as the lines of different slopes when MSD is calculated only along the particular direction.

On the contrary, if the MSD increases faster than at linear rate, then another mechanism of transport takes place. This effect is called superdiffusion, and is usually a result of specific interactions of the guest molecules with the environment, being either the pore walls or other guest molecules. Superdiffusion in zeolites and MOFs is frequently associated with the confined space effects, where the geometry (i.e. size and shape) of the guest molecule closely matches the geometry of the channel.

IR spectra

Another useful application of the MD is predicting infrared spectra for a given model. In the most basic approach, the calculation of the IR spectra can be performed for a harmonic oscillator model with the assumption that the PES behaves like the harmonic potential in the direct neighbourhood of the minimum. This involves the calculation of vibrational frequencies directly from the derivatives of the potential energy. This method, however, neglects anharmonicity and, in order to improve the results, empirical scaling factors need to be used. In addition, the presence of the solvent can also greatly influence the spectra, especially when hydrogen bonds between the solvent and the analysed species are formed [16].

The MD allows for the proper simulation of the bulk phase together with the solvent, where different configurations, such as the arrangement of the hydrogen bonds and the orientation of the solvent molecules, can easily be sampled. Carrying out the MD simulation on the DFT level ensures the proper electronic structure to be maintained during the run.

The procedure is quite straightforward: to calculate the IR spectra, it is necessary to obtain the dipole information for the investigated system and to localize the molecular orbitals to obtain the Wannier centres. The length of the simulation needed to obtain reasonable results varies with the system type. A system with explicit solvent molecules generally needs more time to sample the phase space, while a solid system can be studied in a much shorter time. If the Wannier centres are written to the trajectory every 5th step, approximately 50000 trajectory steps should be sufficient.

The IR spectra can then be obtained from the Fourier transformation of the dipole moment autocorrelation function:

$$I(\omega) = \int \langle M(t) \cdot M(0) \rangle \cos(\omega t) dt \quad (11)$$

where I is the intensity $M(t)$ is the total dipole moment in time t and ω is the vibrational frequency.

5. Monte Carlo methods

While MD simulations are of *deterministic* nature (i.e. given exactly the same initial positions and velocities of the atoms, identical results are obtained),³ Monte Carlo (MC) methods represent a *stochastic* approach.

These methods rely on a random number generator to evaluate the probability of a particular state of the system.

Perhaps the most important variant of the MC method is the Grand Canonical Monte Carlo (GCMC). The simulations are carried out in the μVT ensemble, in which, in addition to volume (V) and temperature (T), the chemical potential (μ) is also kept constant. This represents the system coupled to an infinite reservoir with which it can exchange particles. It is a system where not only heat but also mass is exchanged.

Having an initial state of the system, with atoms in the desired positions, a new state of the system is then generated in one of the following ways: (1) a particle is added to a randomly generated location, (2) an existing particle is removed, (3) an existing particle is translated or rotated. Such a newly generated state is either accepted as one of the probable states in the phase space, or rejected as an unlikely one. The probability of this acceptance is based on the change of the total energy of the system.

This procedure needs to be repeated a statistically meaningful number of times after which the system can be considered in thermodynamic equilibrium. This can be exploited to obtain, for example, the adsorption isotherms if the simulations are carried out at varying pressures but at the same temperature, see section 10.7.

6. Models

Zeolites and MOFs are crystalline solids with a well-defined porosity, which is the result of the interconnection of distinct building units (Figure 7). In the case of zeolites, a single class of building units, generally silica or alumina tetrahedra, referred to as T-sites, is used. MOFs, on the other hand, consist of two classes of building units. In this case the framework is formed by metal(-oxide) nodes which are connected through organic linkers (in the case of COFs the metal(-oxide) nodes are replaced by organic nodes). The periodic crystalline nature of these porous frameworks has implications in virtually all aspects of modelling: from the diffusion of small molecules through the boundaries of the periodic cell, to the Ewald method for charge summation or the Plane-Wave basis set in the description of molecular orbitals.

For both zeolites and MOFs two types of models are generally available: *cluster* models and *periodic* models, each with its own strengths and limitations. As a result, one should choose the type of model depending on the particular phenomenon to be investigated.

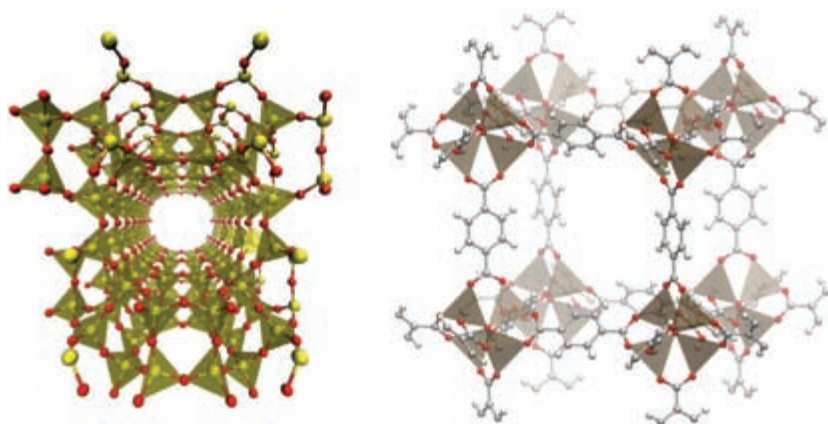


Figure 7. Left: structure of MOR framework. Silicon atoms (shown as yellow spheres) are located in the centres of the tetrahedra (T-sites). Oxygen atoms are shown as red spheres. Right: structure of MOF-5. ZnO nodes represented as tetrahedra, which are interconnected by the terephthalate linkers.

Cluster models

A cluster model is obtained by cutting a finite piece out of the crystalline periodic structure. This piece is centred on the specific atoms or sites of interest. Their finite nature makes cluster models simpler and easier to handle within a computational setup. However, they do not represent the full crystalline structure and may suffer from finite size and boundary effects. In order to deal with the cleaved bonds at the edges of the cluster, the uncoordinated sites need to be passivated with hydrogens or hydroxyls. Cluster models are typically used to study the reactivity of the active sites of zeolites (for instance, Brønsted acid sites, BAS) or MOFs (interactions with a node or linker of the framework).

The main advantage of cluster models is their reduced size. Depending on the particular reaction of interest, the model can contain from a dozen to a few hundreds of atoms. This allows the application of accurate methods to describe the electronic structure of the system and, consequently, the study of effects resulting from this electronic structure. In addition, finding TS is much easier on clusters than it is in periodic systems.

Figure 8 shows a model used to study the effect of substituting Si by Al in the zeolite framework on the propylene protonation reaction [17]. The cluster model consists of three T-sites, out of which two are Si and one is Al. A proton, coordinated with one of the bridging oxygen atoms compensates the charge imbalance and the dangling bonds are saturated by hydrogen atoms.

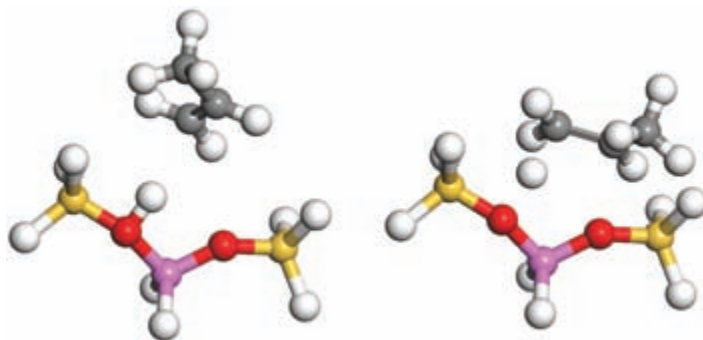


Figure 8. The cluster model of the small section of the ZSM-5 framework used in the simulation of the propylene protonation reaction [17].

This model can be considered extreme on the scale of accuracy. While its relatively small size allows one to use a high level method (such as MP2 or CCSD) to obtain accurate results, the model itself has too many limitations preventing the achievement of a qualitatively accurate picture. In the given example, the small cluster model does not take into account interactions with the other parts of the zeolite framework, especially the long-range electrostatic interactions. A larger cluster model would improve the qualitative picture but would limit the level of methodology that could be used given the same computational resources available (time, memory, and CPU-power). This effect is called the *method vs. model* accuracy trade-off.

Periodic models

In contrast to a cluster model, a periodic model represents a perfect infinitely repeating crystal. Because of its periodicity, it is possible to represent such a system considering only a single finite unit cell, which is infinitely duplicated in every direction. Periodic boundary conditions prevent atoms from exiting the system and rule out the existence of boundary effects. The use of a periodic model also has an important consequence when calculating the total energy of a system. For a simple system, like a molecule or a cluster, the energy contribution to the total energy coming from the electrons is simply the sum of the energies of these electrons. The latter are obtained by solving the Schrödinger equation, which provides a set of possible energy levels the electrons may occupy. Although the same is still true for a periodic system, the calculation of the energies is more complicated as the energies are also a function of the wavevector k of the

electron. This wavevector k gives the periodicity in three dimensions of the specific energy state (cf. section on electronic structure). Consequently, the energies or states, which are just single valued numbers in the case of molecules, become three-dimensional surfaces in a four-dimensional space (1 dimension for energy and 3 dimensions for the wavevector), in what is called reciprocal space, for solids and other periodic systems. The energy of a single state is now obtained through integration of this surface. Looking back at a molecule, one could say that for a molecule only a single wavevector k exists: $k = (0,0,0)$ or no periodicity in any direction. This special point is called the Γ -point. This means that for a molecule (or a cluster) the reciprocal space is reduced to a single point. In contrast, the finite size of a periodic unit cell gives rise to a reciprocal space with a finite size. As a result, the energy of an electron needs to be calculated for each point in the reciprocal space. Luckily, the periodicity of our system allows one to limit these calculations to only a part of the reciprocal space, called the first Brillouin zone. Assume that the energy surface (imagine it as a function $f(x,y,z)$) in the first Brillouin zone can be accurately described by a grid of 10 points in each direction, then one would have to solve the ‘molecular’ Schrödinger equation, we started with, 1000 times. This makes periodic calculations computationally much more demanding, since each point in the reciprocal space needs to be considered. Fortunately, the large size of zeolite and MOF unit cells means that their corresponding reciprocal space is relatively small, making it possible to reasonably sample the reciprocal space by only considering the Γ -point, as is done by many authors. However, if very accurate results are required, such a sampling turns out to be insufficient, as was shown in an accuracy study of the flexible MOF: MIL-47(V) [18].

Because periodic models give access to the full reciprocal space of a crystalline material, they are ideally suited to a detailed investigation of the electronic structure of porous frameworks. Since long-range effects are included in a qualitatively accurate manner, it is also easier to investigate interactions of the spin of unpaired electrons with neighbouring metal nodes. This, however, all comes at a significant computational cost. Fortunately, the vast increase in computational resources in recent decades and the steady improvement of methodologies and algorithms in computational materials science have made it possible to handle larger and larger models in routine simulations, including zeolite reactivity or the breathing behaviour of flexible MOFs. Nowadays, it is possible to perform high quality studies of porous frameworks using fully periodic models. Two examples which will be discussed further in this section are: (1) the study of the CO_2 hydrogenation by the FAU zeolite and Ir_4 cluster supported on it, and

(2) the electronic structure of the breathing of MIL-47(V) MOF and how this can be linked back to the experiments.

7. Electronic structure of porous frameworks

A central topic in the first principles investigation of periodic crystals is the study of their electronic structure. This can be done by solving either the Schrödinger equation or the HKS equations of the system. In both cases, this means solving an eigenvalue problem, where the eigenvalues are the possible energies of the particles in the system, i.e. the electrons.⁴ In contrast to a molecule or cluster, electrons in a solid experience a periodic potential, and their movement can be described using a ‘Bloch wave’. Such a Bloch wave consists of two parts: a periodic function $u(\mathbf{r})$ with the same periodicity as the crystal, and a plane wave:

$$\varphi_{\mathbf{n}\mathbf{k}}(\mathbf{r}) = u_{\mathbf{n}\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}} \quad (12)$$

where $u_{\mathbf{n}\mathbf{k}}$ is a periodic function in real space, n is the band index (similar to the one found when solving the eigenvalue problem for a molecule), and $\mathbf{k}$ is the wave vector related to the momentum of the electron in the crystal. The wave vector $\mathbf{k}$ gives the periodicity of the specific state and can take any value within the reciprocal space of the simulation cell. As a result, varying the crystal momentum ($p=\hbar\mathbf{k}$) will lead to a change in the energy of the electron. While for electrons in a molecule the energy states form a discrete 1D spectrum, each energy state of an electron in a periodic solid gives rise to a 4D energy surface. To visualize the latter, one can plot the energy states/bands as a function of their position in reciprocal space, following a path along the high symmetry lines of the system. Since each electron gives rise to a single energy band, such a representation quickly starts to look like spaghetti. Figure 9 shows the example of diamond, which contains 2 atoms in its unit cell.

In the case of a MOF or a zeolite, with a hundred or more atoms in their primitive unit cell, the band structure picture contains too much information. A simplified picture can be obtained in the density of states (DOS). The density of states shows the density of bands in the reciprocal space as a function of the energy (i.e. it integrates the number of bands over the first Brillouin zone) presenting a picture comparable to the energy spectrum obtained for a molecule. In addition, by projecting the energy bands of a system onto the atomic orbitals, it is possible to obtain a so-called

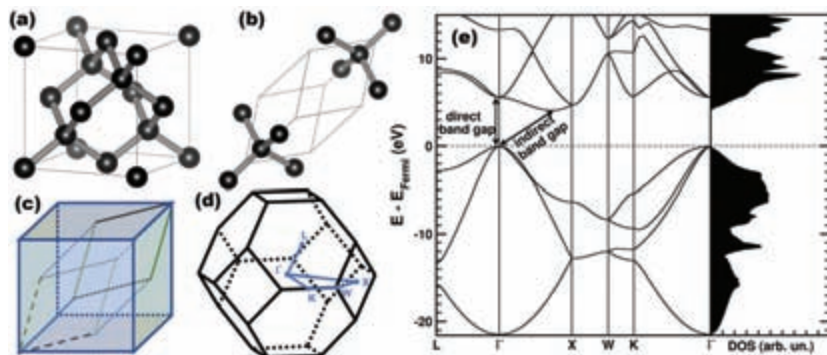


Figure 9. a) Ball-and-stick representation of a conventional unit cell of diamond. b) Ball-and-stick representation of the primitive unit cell of diamond. Note that only the centre atoms of the two clusters are inside the cell, the other atoms belong to neighbour cells (they were added to clarify the local environment of the C atoms in the primitive unit cell). c) Relationship between the conventional and primitive unit cells, in case of diamond, or any other material with face-centred cubic symmetry, such as for example the UiO-66 MOF. d) First Brillouin zone of a system with face-centred cubic symmetry. Some high symmetry lines are indicated. e) The electronic band structure and density of states of diamond, using a combined representation typical in solid-state literature.

local-DOS or LDOS. This teaches us which atoms or even atomic orbitals give rise to the specific peaks seen in the DOS. As such it can be used to discover which states make up the edges of the band gap, and are thus relevant for excitations and luminescence. Figure 10 shows an example of a DOS and LDOS for the luminescent and breathing MOF COK-69(Ti) [19]. The two different spin states ('up and down' or 'majority and minority' in the context of solid state physics) are indicated in black and red, respectively, with one spin state flipped down. The specific MOF mentioned here presented very clear luminescent behaviour, turning bright blue upon irradiation. This can be understood in terms of the protonation of the oxo-group of the Ti clusters. To retain charge balance upon this protonation, one of the Ti atoms needs to undergo a valence change from Ti^{IV} to Ti^{III} . This is clearly visible in the electronic structure through the appearance of a gap state. Using the LDOS of the different Ti centres this gap state could be attributed to the Ti^{III} atoms.

The above approach presents a way to investigate the electron distribution as a function of their energy. However, sometimes it is also interesting to investigate the spatial distribution of the electrons. This can be done through the direct visualization of the electron density or through density difference plots. It is also possible to partition the electron density between the atoms in an atoms-in-molecules (AIM) type of fashion providing atomic charges (useful for constructing a force field). This partitioning can

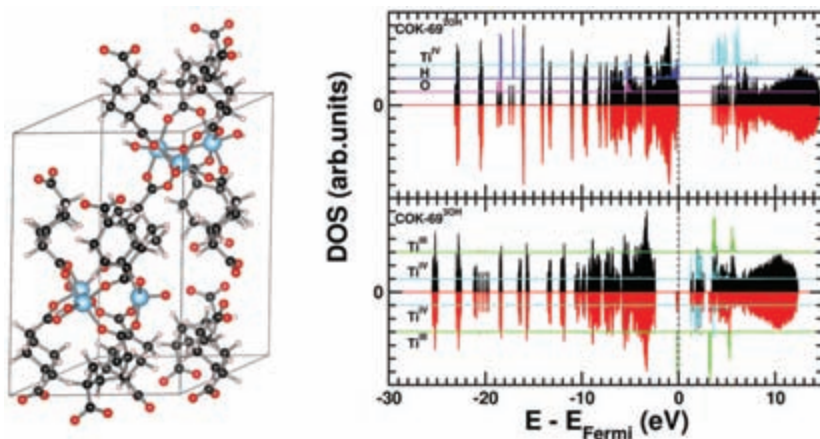


Figure 10. Left: ball-and-stick representation of the primitive unit cell of the COK-69(Ti) MOF. Atomic positions of Ti are indicated in blue. Right: spin polarized DOS of the COK-69(Ti) with different titanium clusters. In the case of COK-692OH, each Ti-cluster has two hydroxy groups, leading to three Ti^{IV} atoms, while the Ti-clusters of COK-693OH have three hydroxy groups, giving rise to a single Ti^{III} in each cluster. This Ti^{III} gives rise to a gap state and luminescent behaviour.

be done in several different ways: Bader's QTAIM, Mulliken charges, Hirshfeld charges, etc. Each of these methods comes with its own strengths and weaknesses, and one has to be careful when comparing their results [20]. The main problem observed by Hamad and coworkers was the fact that, upon geometry deformation, such as in case of breathing, some schemes may give rise to significant changes in the calculated charges, making them less suitable. In this regard, the Hirshfeld-I partitioning scheme performs remarkably well [21]. Tests performed by one of the present authors showed no significant variations in the obtained charges over the full volume range covered by the breathing MIL-47 MOF [18].

As noted earlier, DFT is a ground state theory. One important consequence of that is the well-known band gap problem of local functionals, which boils down to a general underestimation of the band gap width for all semiconductors and insulators. Luckily, it mainly entails a downward shift of the conduction bands, meaning that the shape of those bands is qualitatively accurate. Several solutions are available in the solid-state community, varying from scissor-operators, DFT+U, the use of hybrid functionals or Green's function approaches. Despite this limitation, DFT generally provides very good atomic structures for periodic systems (in the case of porous frameworks it is often useful to include van der Waals contributions via a corrective scheme such as DFT-D). This means that one can use local DFT calculations, to obtain optimized structures, in tandem with hybrid

functional calculations, to obtain high quality electronic structures. Using such an approach the computational cost is kept in check while results are obtained in good agreement with experimental measurements [22].

8. Practical investigations of porous frameworks using DFT

Example 1: CO₂ hydrogenation on Ir₄/FAU system

This example shows the importance of the fully periodic model of the zeolite lattice [23]. The model used is a combination of the cluster model of the Ir₄ [24] and a realistic model of FAU with a reduced cell size [25]. The cluster is supported in one of the 6-membered rings in the faujasite cage while the surrounding lattice provides the reaction environment. The Ir₄ cluster is able to generate hydrides, while the BAS of the FAU framework provides the protons, both hydrogen species are needed in the reaction.

The reaction starts with the CO₂ activation on the Ir₄ nanoparticle, which is partially hydrogenated while the top site remains available to bind a CO₂ molecule. The first step in the process is the hydrogen transfer to the oxygen atom of CO₂. This can be accomplished in two ways: either by an attack of the hydride from the Ir₄ cluster, or by a proton from the Brønsted acid site (BAS). The former is accompanied by a high barrier of 143 kJ mol⁻¹ because the hydrogen atom, formally a hydride, is used to protonate the oxygen. The latter, however, requires a water assist because the BAS is located farther from the CO₂, but in such case the barrier is only 98 kJ mol⁻¹. Such an observation could not have been made if the cluster model had been used in the simulation and the presence of distant BAS had been neglected.

The result of the first step is the carboxyl bound to the Ir₄ cluster. The second step can lead, either to formic acid (FA) by hydrogenation of the C atom, or to carbon monoxide by protonation of the hydroxyl O. These pathways are shown in Figure 11. The pathway to FA is kinetically preferred because it occurs via the [8–9] transition state, which is associated with the barrier of 114 kJ mol⁻¹, compared to the [8–11] transition state, associated with a barrier of 146 kJ mol⁻¹. On the other hand, the latter is thermodynamically preferred because it leads to a strongly bound carbonyl complex, more stable than FA by 86 kJ mol⁻¹. This complex should be considered undesired, because the release of the CO from the Ir₄ cluster requires as much as 246 kJ mol⁻¹, which is a prohibitive amount. The CO thus poisons the active site by blocking it for adsorption of the reactants.

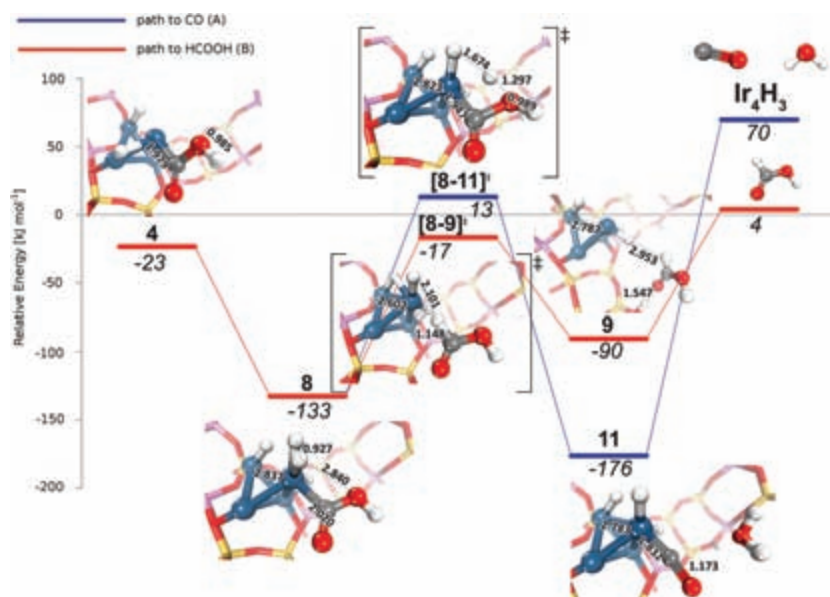


Figure 11. Reaction pathways for the CO (blue) or FA (red) from the Ir-bound carboxyl.

Example 2: Electronic structure of breathing MOFs

Structure optimization and breathing

Porous frameworks exist in all possible shapes and forms, and their modular nature gives rise to an almost infinite number of possible systems. In the case of MOFs, one class of MOFs attracted attention due to its breathing behaviour. Breathing is observed when a material can undergo reversible volume changes under the influence of an external stimulus. This can be a temperature or pressure change, or the presence of a specific gas. In breathing MOFs, these volume changes are large (easily 30 % or more). Many MOFs with MIL-47/53 topology are known to present breathing behaviour. Of these, MIL-47(V) is considered a rigid example as it only shows breathing under application of an external pressure of about 100 MPa (or 1 kbar). This flexibility makes structure optimization a difficult exercise, since a poor computational setting may push it over the barrier between large and narrow pore geometry. The Pulay stress, an effect, which has no relevant size in modern day calculations on standard solids, can be large enough in MOFs to do the trick [18]. To circumvent this problem, structures can be optimized under the constraint of a fixed volume. The equilibrium volume is then obtained through an equation of state fit performed on a set of calculated structures with (slightly) different volumes. Fitting the parameters

Table 1. Effect of the magnetic configuration of the MIL-47(V) MOF on its mechanical properties.

Magnetic configuration	P_t (MPa)	B_o (GPa)	E_f (meV/V atom)
AF	124	8.12	0
AF + FM	102	7.17	36
FM	83	5.95	70

of the state equation also has the advantage of allowing a calculation of the transition pressure P_t in a breathing MOF. An important feature of the MIL-47(V) MOF not mentioned before is the unpaired electron present on each of the V centres. The four unpaired electrons in the unit cell of MIL-47(V) can have different orientations. This gives rise to ferromagnetic (FM) or antiferromagnetic (AF) configurations on the vanadyl chains. Calculations show the AF configuration to be the most stable one, as can be seen from the relative energies, E_p , shown in the Table. Furthermore, calculations also show that these spin configurations have a significant influence on the mechanical properties (bulk modulus, B_o , and transition pressure, P_t) of the MOF, as shown in Table 1 [26].

From single crystal experiments, the ground state configuration is expected to be AF, which is corroborated by the calculations [27]. However, linking the calculated transition pressures to the results obtained from mercury intrusion experiments that indicated an FM configuration was generally present [28]. These apparently contradictory results can also be elucidated making use of high accuracy computational data. Starting from optimized geometries for the two spin configurations, Bogaerts and coworkers simulated and fitted X-ray diffraction patterns for the MIL-47(V) MOF. The simulated results were compared to experimental spectra from single crystal and powders, leading to the conclusion that single crystal samples were best represented by the AF configuration, while powder samples were best represented by the FM configuration [29].

Electronic structure: from spin coupling to splitting orbitals

In the previous section, it was noted that spin configuration plays a crucial role in the mechanical properties of MIL-47(V). As the vanadyl chains are linked via relatively long benzenedicarboxylate (BDC) linkers, one may wonder how strongly the spins of different chains are coupled. Making use

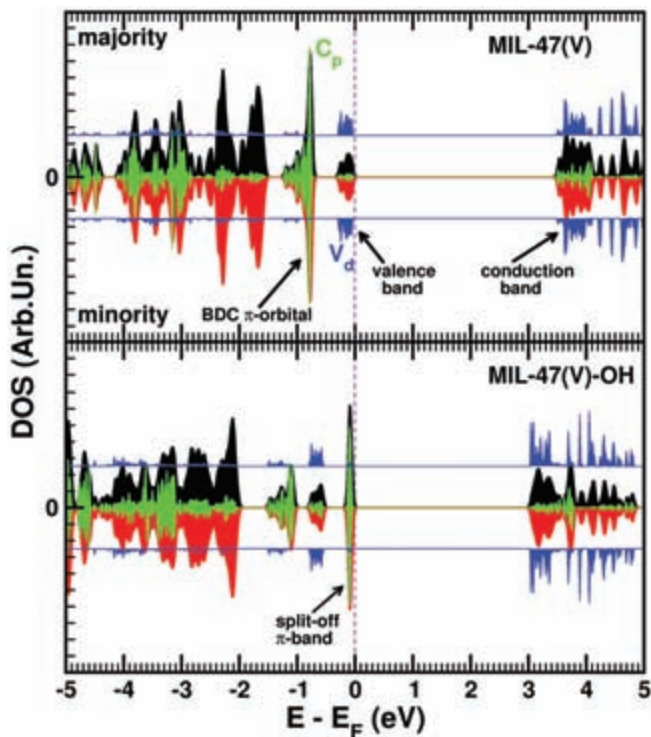


Figure 12. DOS and LDOS for the non-functionalized MIL-47(V) (top) and the hydroxy-functionalized MIL-47(V) (bottom).

of a simple spin Hamiltonian, it is possible to calculate the intra- and inter-chain coupling. Although the obtained values are quite sensitive to the function used, the qualitative picture is constant. For the MIL-47(V) the inter-chain coupling is about 2 orders of magnitude weaker than the intra-chain coupling, making the vanadyl chains quasi-1D systems.

The different spin configurations also have an influence on the electronic band structure. However, things get more interesting upon functionalization of the BDC linker. Looking at the band structure of the MIL-47(V) host material (Figure 12), one finds that the band gap edges are dominated by states originating from the V centres. The π -orbital associated with the BDC linker is located about 1 eV below the Fermi level. However, upon functionalization, this orbital splits and the split-off π orbital is shifted upward toward the Fermi level. The upward shift depends on the functional group and the number of functional groups on the linker. In the hydroxy group, the upward shift becomes so big that the split-off π orbital moves above the V valence band, effectively reducing the band gap [30]. This

behaviour is not unique for the MIL-47(V) and has also been observed for the linker-functionalized UiO-66(Zr).

Example 3: Separation of CO₂ and CH₄ in UiO-66

Microporous solids such as zeolites and MOFs are often referred to as *molecular sieves*, because of their well-defined porosity that matches the dimensions of individual molecules. The frameworks interact with molecules of particular species differently, which enables the separation of gas mixtures. In this example, the separation of CO₂ and CH₄ molecules by UiO-66 will be discussed.

GCMC simulations were carried out on a model of UiO-66 consisting of only one cubic unit cell of dimensions of $a=b=c=20.75$ Å. The interactions between the host and guest molecules was described by the Force Field based on the parameters by Prakash et al. [31] Only van der Waals and Coulomb type interactions were taken into account because the structure of both MOF and guest molecules remained rigid during simulations, meaning that the bonding interactions were always exactly the same. Figure 13 shows the adsorption isotherms for both types of guest molecules in the pressure range up to 1000 kPa. It can be seen that the amount of adsorbed CO₂ is much higher than that of CH₄. In addition, the temperature does not seem to have an influence on the adsorption of CH₄, while it

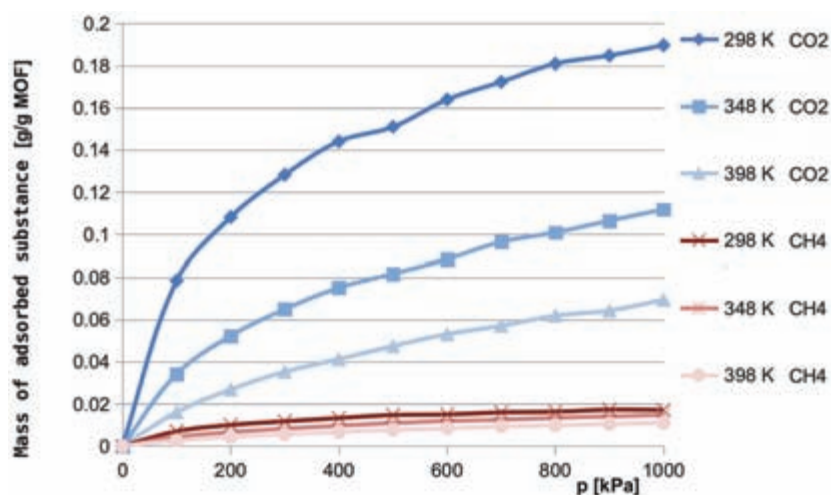


Figure 13. Adsorption isotherms for the mixture of CO₂ and CH₄ in UiO-66 in 298–398 K temperature range.

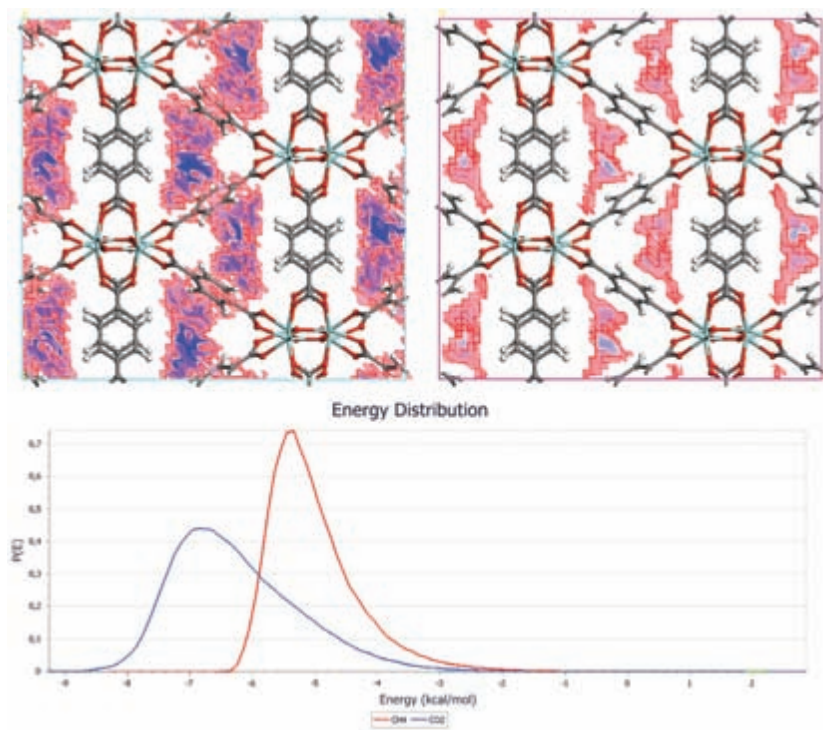


Figure 14. a) Distribution of CH_4 (left) and CO_2 (right) molecules inside the UiO-66 MOF resulting from GCMC simulation. The colour represents the relative interaction energy – blue, white and red represent the strongest, intermediate and weakest interactions, being related to the probability of the guest molecules occupying that location. b) Energy distribution profile for CH_4 and CO_2 in UiO-66.

significantly affects the isotherms for CO_2 . The best separation properties are thus observed at the highest pressure and the lowest temperature.

Figure 14 a) shows the distribution of the guest molecules inside the host pores. It shows that the CO_2 molecules form a more diffuse ‘cloud’, which suggests that they fit into the thin pores slightly better. On the contrary, methane, which is a bulkier molecule, is located only in specific areas. This observation is also consistent with the slight influence of the temperature or pressure on the adsorption of CH_4 . The properties of the methane molecules allow them to occupy only some ‘spots’ and neither a temperature decrease nor a pressure increase allows the accommodation of more guests. The situation seems different for CO_2 molecules, which have a linear shape, allowing them to adjust better to the pore sizes and shapes of the MOF and making other locations available for adsorption. Interestingly, sites available for CH_4 adsorption are mostly coincident with those for CO_2 , which suggests

the competition for these locations might take place. Finally, the energy histogram shows the probability of finding a molecule with a particular interaction energy inside the pores of UIO-66, Figure 14 b). It shows one peak for each of the species, which corresponds to the preferred adsorption site. The peak for CH₄ is sharper, suggesting a better localization of this adsorption site. In contrast, the peak in the CO₂ curve is smaller and broader. This means that the molecules are not limited to one particular location and can 'spill' over a larger area in the pores. The separation abilities of the UIO-66 in the given mixture illustrated by the isotherms above is the result of a stronger interaction of the framework with the CO₂ molecules. The peaks of the curves are separated by approximately 2 kcal mol⁻¹ in the figure, meaning that the interaction with CO₂ is stronger and the presence of this species leads to a stronger stabilization of the system than the presence of CH₄.

References

- [1] van Beest, B.W., Kramer, G.J., van Santen, R.A. *Phys. Rev. Lett.* 1990, 64, 1955.
- [2] Shi, H., Migués, A.N., Auerbach, S.M. *Green Chem.* 2014, 16, 875.
- [3] Llewellyn, P.L., Bourrelly, S., Vaguer, C., Heymans, N., Leclerc, H., Ghoufi, A., Bazin, P., Vimont, A., Daturi, M., Devic, T., Serre, C., De Weireld, G., Maurin, G. *J. Phys. Chem. C*, 2013, 117, 962.
- [4] Hohenberg, P., Kohn, W. *Phys. Rev.* 1964, 136, B864.
- [5] Kohn, W., Sham, L.J. *Phys. Rev.* 1965, 140, A1133.
- [6] Grimme, S. *J. Comput. Chem.* 2006, 27, 1787.
- [7] Grimme, S., Antony, J., Ehrlich, S., Krieg, H. *J. Chem. Phys.* 2010, 132, 154104.
- [8] Warshel, A., Levitt, M. *J. Mol. Biol.* 1976, 103, 227.
- [9] Vanpoucke, D.E.P., Oláh, J., De Proft, F., Van Speybroeck, V., Roos, G.J. *Chem. Inf. Model.* 2015, 55, 564.
- [10] van Duin, A.C.T., Dasgupta, S., Lorant, F., Goddard, W.A. *J. Phys. Chem. A* 2001, 105, 9396.
- [11] Li, X., Frisch, M.J. *J. Chem. Theory Comput.* 2006, 2, 835.
- [12] Mills, G., Jónsson, H. *Phys. Rev. Lett.* 1994, 72, 1124.
- [13] Henkelman, G., Jónsson, H. *J. Chem. Phys.* 2000, 113, 9978.
- [14] <http://www.rheoman.eu/>
- [15] Allen, M.P., Tildesley, D.J. *Computer Simulation of Liquids*, Oxford University Press, 1989.
- [16] Thomas, M., Brehm, M., Fligg, R., Vöhringer, P., Kirchner, B. *Phys. Chem. Chem. Phys.* 2013, 15, 6608.
- [17] Dziekoński, P., Sokalski, W.A., Szyja, B., Leszczynski, J. *Chem. Phys. Lett.* 2002, 364, 133.
- [18] Vanpoucke, D.E.P., Lejaeghere, K., Van Speybroeck, V., Waroquier, M., Ghysels, A. *J. Phys. Chem. C* 2015, 119, 23752.
- [19] Bueken, B., Vermoortele, F., Vanpoucke, D.E.P., Reinsch, H., Tsou, C.C., Valvekens, P., De Baerdemaeker, T., Ameloot, R., Kirschhock, C.E.A., Van Speybroeck, V., Mayer, J.M., De Vos, D. *Angew. Chem. Int. Ed.* 2015, 54, 13912.
- [20] Hamad, S., Balestra, S.R.G., Bueno-Perez R., Calero, S., Ruiz-Salvador, A.R. *J. Solid State Chem.* 2015, 223, 144.
- [21] Vanpoucke, D.E.P., Bultinck, P., Van Driessche, I. *J. Comput. Chem.* 2013, 34, 405.

- [22] Hendrickx, K., Vanpoucke, D.E.P., Leus, K., Lejaeghere, K., Van Yperen-De Deyne, A., Van Speybroeck, V., Van Der Voort, P., Hemelsoet, K. *Inorg. Chem.* 2015, 4, 10701.
- [23] Szyja, B.M., Smykowski, D., Szczygieł, J., Hensen, E.J.M., Pidko, E.A. *ChemCatChem*, 2016, 8, 2500.
- [24] Ferrari, A.M., Neyman, K.M., Mayer, M., Staufer, M., Gates, B.C., Rösch, N. *J. Phys. Chem. B* 1999, 103, 5311.
- [25] Pidko, E.A., Mignon, P., Geerlings, P., Schoonheydt, R.A., van Santen, R.A. *J. Phys. Chem. C* 2008, 112, 5510.
- [26] Vanpoucke, D.E.P., Jaeken, J.W., De Baerdemacker, S., Lejaeghere, K., Van Speybroeck, V. Beilstein, *J. Nanotechnol.* 2014, 5, 1738.
- [27] Barthelet, K., Marrot, J., Riou, D., Férey, G. *Angew. Chem. Int. Ed.* 2002, 41, 281.
- [28] Yot, P.G., Ma, Q., Haines, J., Yang, Q., Ghoufi, A., Devic, T., Serre, C., Dmitriev, V., Férey, G., Zhong, C., Maurin, G. *Chem. Sci.* 2012, 3, 1100.
- [29] Bogaerts, T., Vanduyfhuys, L., Vanpoucke, D.E.P., Wieme, J., Waroquier, M., Van Der Voort, P., Van Speybroeck, V. *Cryst. Eng. Comm.* 2015, 17, 8612.
- [30] Vanpoucke, D.E.P. *J. Phys. Chem. C* 2017, 121, 8014.
- [31] Prakash, M., Jobic, H., Ramsahye, N.A., Nouar, F., Damasceno-Borges, D., Serre, C., Maurin, G. *J. Phys. Chem. C* 2015, 119, 23978.

Notes

- 1 The use of the term “ab-initio” can be debated as hybrid functionals (such as B3LYP) and DFT+U do contain a fitting component. In addition, due to historical reasons the IUPAC definition of ‘ab-initio’, aimed at distinguishing QM and empirical methods, also excludes any non-wave function based method, as it was inconceivable in the early days of Quantum Chemistry that a non-wave-function based description could exist. To accommodate this difference between physics and chemistry-oriented views, DFT is often referred to as ‘first principles’. In this chapter, we will follow the IUPAC definition of ab-initio referring to wave function based implementations of QM, in contrast to the density-based implementation used in DFT.
- 2 The ensemble is defined as the collection of all possible systems that have different microscopic states but the same macroscopic or thermodynamic state.
- 3 Strictly speaking the results might still differ due to the accumulation of the round-off errors or a different precision of numbers stored in a memory/disk. This is known as the butterfly effect.
- 4 In the HKS-equations this is not exactly true, however, it is possible to relate the obtained states to the electron states. So, for practical purposes, we can continue as if they are.