BUNSEN-MAGAZIN · 21. JAHRGANG · 1/2019

Giovanni Maria Piccini

REMOVING BARRIERS TO UNDERSTAND CHEMISTRY

INTRODUCTION

Experimental, theoretical and computational investigations of reaction mechanisms represent one of the most challenging fields in chemistry. A deep understanding of activated processes allows chemists to comprehend the nature of chemical transformations and helps designing molecules in order to govern the reaction. Besides the fundamental academic interest, the study of chemical reactions has a great impact in everyday life.

Understanding a chemical reaction is not simple. To properly investigate these events one must understand how and why a reaction proceeds from reactants to products with a specific mechanism, and how fast this process is. Changes in the molecular structure, ambient conditions, different solvation environments, allow chemist to interfere with the chemical transformation process in order to achieve the desired products. Therefore, the atomistic details of the reaction mechanism are fundamental to reach this goal.

Quantum chemistry (QC) has been very successful in unravelling the nature of many chemical reactions at the molecular level.[1] Several computational techniques and new physical chemistry concepts have been developed bridging the computational results to the experimental findings.[2] Reaction paths[3] and coordinates,[4] reactive trajectories,[5] potential energy minima and saddle points[6] have become customary objects in computational chemistry.[7] However, their application is extremely limited if not meaningless to systems for which the structural complexity and fluxionality is very high, e.g. solvated systems.

A valid alternative is represented by *ab initio* Molecular dynamics (MD).[8] MD simulations represent a powerful tool for studying a vast range of chemical systems. Many equilibrium properties can be calculated and their related atomistic details investigated. Since decades these techniques have been applied to several phenomena in different kind of systems from materials to biological macromolecules. Unfortunately, many important processes occur on a time scale that is larger by orders of magnitude than the one accessible by standard MD simulations. These phenomena are commonly referred to as rare events.

Dr. Giovanni Maria Piccini
ETH Zürich
Professur für Computational Science, LUI
USI-Campus Via Giuseppe Buffi 13, 6900 Lugano, Switzerland
Tel.: +41 58 666 48 03
E-Mail: g.piccini@phys.chem.ethz.ch

To circumvent this problem several methods and computational techniques have been proposed in order to enhance the sampling of these events.[9] This is achieved by accelerating the transitions between metastable states, thus, broadening the probability distribution between them. Because of the direct relationship between probability distribution and free energy, these methods are often referred as free energy methods. A large family of these techniques is based on the concept of the collective variable (CV),[10] a parameter that characterizes and distinguishes the metastable states of interest. Among all, a very popular and widely used method to enhance the sampling between free energy minima is Metadynamics (MetaD).[11, 12] This approach makes use of an adaptive and history dependent bias potential that enhances the fluctuations within the free energy basins, thus, favouring transitions. MetaD allows the exploration of the configurational space, the reconstruction of the free energy surface (FES), and the determination of the kinetics of the process.[13, 14] In spite of its popularity and vast use in several application, MetaD has been much less frequently applied to chemical reactivity. Among all the 3000 and more citations of the landmark paper by Laio and Parrinello “Escaping free energy minima” introducing the MetaD algorithm,[11] only about 80 papers report its application to chemical reactions.[15] The research presented in the present report aims at developing new methods and approaches for studying chemical reactions within the framework of Metadynamics. Based on the previous works and on recent promising developments in the field, together with the development of faster computers and efficient algorithms, MetaD could potentially become a routine method in the study of chemical reactions in complex systems competing with the aforementioned well established methods. Applications to relevant problems in chemistry will be tackled. They range from homogeneous and heterogeneous catalysis up to the role of water in chemistry such as acid/base reactions and anticancer drugs activity. The overall and most ambitious goal of this project is to move forward the frontiers in quantum chemical simulations.

METHODS

Metadynamics

The central concept in quantum chemistry is the potential energy surface (PES).[16] Within the Born-Oppenheimer approximation the PES embodies most of the information regarding the chemical properties of a molecular system. Formally, the PES is a 3N-6 dimensional function V(**R**) that maps the potential

among the nuclei for a given configuration **R** = (R_1, R_2, . . . , R_{3N-6}), where N is the number of atoms in the system. To characterize the chemically relevant configurations, namely reactants, products, and transition structures, one must locate on the PES minima and first order saddle points respectively.[6, 17] Traditionally, different geometry optimization algorithms have been used to achieve this goal.[18–22] Unfortunately, due to the large number of degrees of freedom this task may be rather cumbersome. Moreover, no guarantee that the located stationary points correspond to the correct structure is given. Fig 1 shows a schematic representation of a PES projection on 2 of the 3N-6 dimensions of the system. Several local minima are present as well as many saddle points between them. Understanding if an optimized reference structure corresponds correctly to the state configuration of interest still requires a large amount of chemical intuition. The "roughness" of the PES is mainly a consequences of the fact that no temperature effects are considered when evaluating the system potential energy. This means that besides quantum zero-point vibrations at 0 K the nuclei are frozen in local configurations. On the contrary, at finite temperature many local minima corresponding to the structures characterizing the metastable thermodynamics states may be explored.

Thus, the inclusion of finite temperature effects, i.e. entropy, results in a reduction of the complexity of the problem as several narrow minima are smoothed into shallower valleys (see Fig. 1). Including the entropy means that the function we are considering is no longer the potential energy but the free energy. This comes with several advantages. The free energy can be obtained selecting a few parameters that characterize and distinguish well the metastable states of interest without loss of generality. We call these parameters collective variables (CVs)[10] and they are defined as functions of the microscopic atomic Cartesian coordinates.

Therefore, it seems straightforward to make use of the free energy rather than the potential to calculate most of the transition properties for an activated process. In this way one would explore and characterize the different local minima associated to reactants and products, e.g fluxional systems, and at the same time get information about the dynamics involved in a transition. Unfortunately, in most cases the estimation of the free energy is extraordinarily slow, if not prohibitive, in a standard MD trajectory due to the presence of high energy barriers separating the metastable states. In other words, for such atomistic trajectories the system will remain stuck in one of the free energy basins, just fluctuating around the local minimum. To circumvent this problem several methods have been proposed to achieve ergodic sampling and are named after this as *enhanced sampling* or *free energy methods*.[9]

Here, we will focus on a particular class of enhanced sampling methods that rely on collective variable biasing, namely MetaD.[11–13] Other techniques that do not make a direct use of the concept of CV to explore the free energy landscape are as well important and powerful but less adequate for the goals we want to achieve.

The idea behind CV based enhanced sampling methods is to add a bias potential along a set of CVs in order to force the system visiting those region of the free energy space in which the equilibrium probability distribution is small. The first example of this strategy can be traced down to the 1977 landmark paper of Torrie and Valleau introducing umbrella sampling. In their work the authors proposed to add a bias to the underlying PES derived by trial-and-error in order to sample all the CV space of interest. Of course, finding a "good" bias potential for each problem may be rather cumbersome most of the times. MetaD overcomes these problems by introducing a progressively growing bias potential along the CVs. This history dependent bias is commonly built by a sum of repulsive Gaussian kernels preventing the system visiting regions that have been already explored.

Fig. 2 shows the effect of MetaD biasing on a one dimensional double well potential. The free energy basins are filled by the adaptive bias during the simulation, thus, enhancing the fluctuations around the local minima and favouring transitions across the barrier. From a statistical point of view the effect of the bias is a progressive broadening of the probability distribution within the free energy basins, i.e. filling the gap of non zero probability in the barrier region.

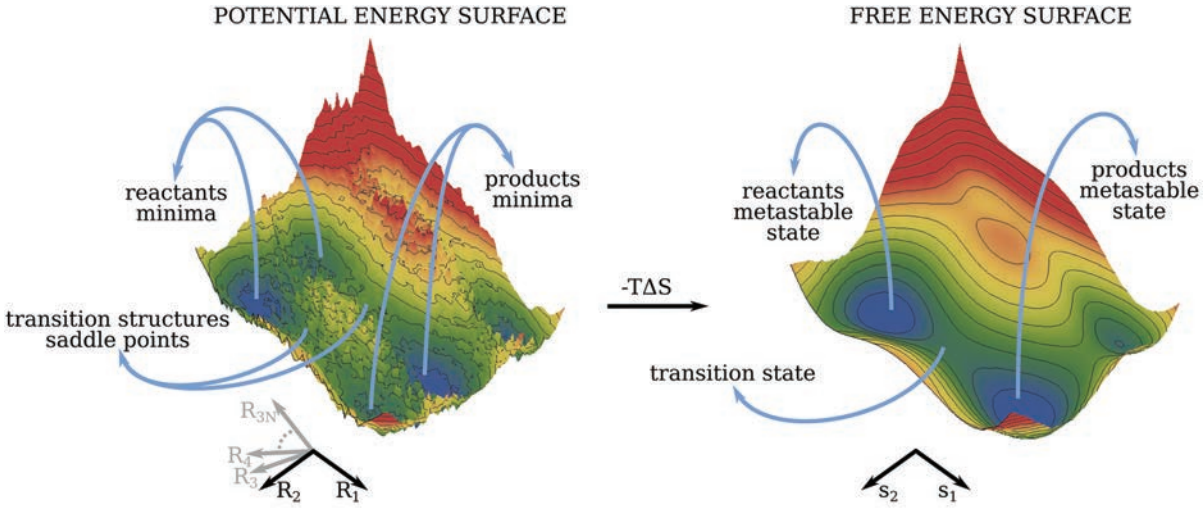

Fig. 1: The effect of entropy on the potential energy landscape along a 2-dimensional representations.

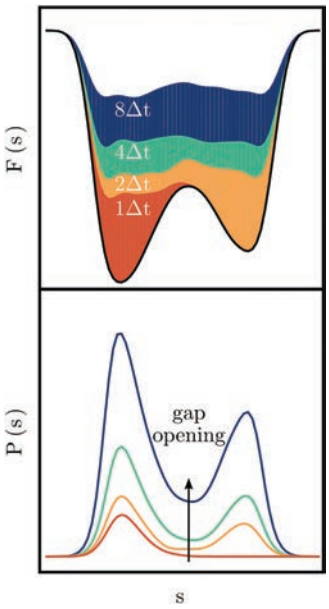

Fig. 2: The effect of MetaD on a double well potential energy surface (upper panel) associated to the distortion of the probability distribution (lower panel) along a generic CV s. As more bias is deposited inside the free energy basins a broadening of the total probability distribution is observed including a gap filling in the regions where P(s) is close to zero.

The main advantages of MetaD is the fact that apart from some few simple parameters (height and width of the Gaussians kernels) no a priori knowledge on the the shape of the free energy is needed to design the bias. Moreover, the statistical features of the system like the probability distribution and, therefore, the free energy can be accurately reconstructed by means of reweighting techniques.[23, 24]

Development of collective variables for chemical reactions

The characterization of metastable thermodynamic states by means of specific order parameters lies at the hearth of free energy methods. MetaD makes a direct use of these descriptors to explore and represent the free energy landscape of an activated process.

Often, the complexity of a chemical reaction makes the intuitive selection of a proper set of free energy descriptors extremely hard.[15, 25, 26] It is therefore fundamental to develop and implement new CVs and methods allowing MetaD to become a routine approach in the computational study of chemical reactions.

In most cases, a chemical reaction involves a rearrangement of the molecular bonding topology. Reactants and products differ from each other by the presence or absence of different types of chemical bonds. Crossing the reaction barrier means spending energy to simultaneously form and break such bonds. A typical example is the S_N2 reaction $CH_3F + Cl^- \rightarrow CH_3Cl + F^-$. In order to transit from reactants to products (see Fig. 3) distance d_1 must elongate while d_2 shrinks and vice versa for the backward reaction.

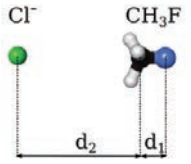

Fig. 3: Carbon-halide distances used in the description of an S_N2 reaction.

Using d_1 and d_2 as independent CVs one obtains the FES reported on the left of Fig. 4. In a simple case like this the use of both distances is perfectly reasonable. However, in more complex scenarios more descriptors may be necessary. In such cases a MetaD simulation may converge poorly and, most importantly, the physical interpretation may lose in clarity and simplicity. By means of chemical intuition it is possible to reduce the dimensionality of the problem. In the case of the S_N2 reaction, a valid solution may be combining the whole information in an antisymmetric linear combination such as $s = d_1 - d_2$. Using this CV, instead of d_1 and d_2 independently, one gets a FES that describes very efficiently and much more clearly the process.
Recently, we have introduced a new method able to pack a large set of state descriptors into low dimensional CVs. The method uses the information extracted from the local fluctuations (means and multivariate variance) of these descriptors within the reactants and products free energy basins. We modified Fisher's discriminants[27] approach to operate a dimensionality reduction on the CVs. We called this method Harmonic Linear Discriminant Analysis (HLDA).[28]
In this way, given a arbitrary large set of state descriptors one can operate a dimensionality reduction to one single variable that embodies all the information needed to enhance the sampling along an optimal direction.
A practical example is provided by a typical organic chemistry reaction, the Diels-Alder cycloaddition (see Fig. 5). We studied the classical [4+2] Cycloaddition of 1,3-butadiene and ethene. The essence of a chemical reaction can be summarized as a combination of simultaneous bonds breaking and formation. This is accompanied by a time-crucial conformational reorganization[29] and a local change of distances. Such a bond reorganization can be a dramatic event involving breaking and/or formation of π-type bonds but can also be accompanied by local electronic rearrangements resulting in a non-local strengthening or weakening of other bonds like the one to conversion.

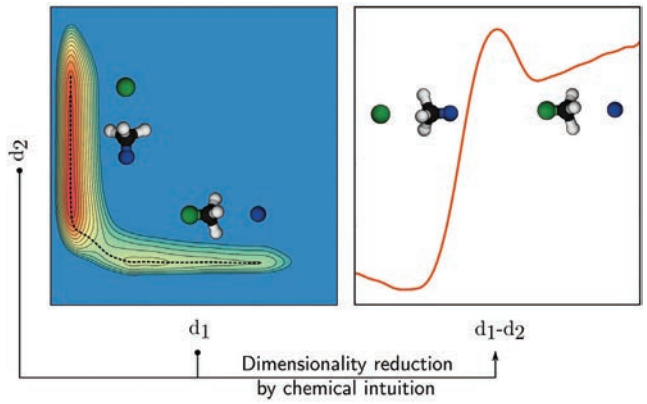

Fig. 4: FES obtained applying MetaD along d_1 and d_2 independently (left) and along the reduced CV $s = d_1 - d_2$ (right).

The chemical scheme for a classical Diels-Alder reaction is reported in Fig. 5. The peculiarity of a Diels-Alder reaction lays in the fact that the formation of the two bonds in the product state implies that three π-bonds of the reactant state become two σ-bonds while a σ-bond becomes a π-bond.

Therefore, although it is clear from Fig. 5 that distances d_4 and d_6 will play a major role in the reaction progression as they

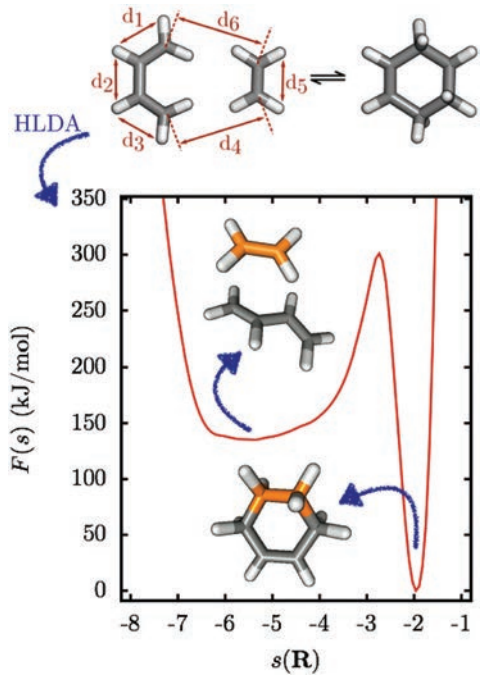

Fig. 5: Free energy profile for the Diels-Alder reaction of 1,3-butadiene with ethene as a function of the HLDA CV obtained using the set of six bond distances as basis descriptor.

govern the approach of the two molecules to each other, the concerted elongation-contraction of distances d_1, d_2,d_3, and d_5 will play a significant role. Having the two distributions for the six distances considered we applied HLDA to them.

To demonstrate the power of such a method we report in Fig. 5 the Free-energy profile along s_{HLDA}. The HLDA CV besides being rather efficient clearly brings out the chemistry of the problem with a wide entropic basin and a narrow and deep enthalpic state. Furthermore, the configurations extracted from those that are in the apparent transition state do correspond to those that are described in standard text books (see Fig. 6). In the final state two possible configurations are possible, namely *endo* and *exo*. In our case they are symmetry related by reflection. However, if some specific asymmetry of the reactants would be present this may lead to the well known endo-exo chirality. Such a procedure can be applied to more complex cases where the bonding transformation cannot be clearly attributed by means of simple chemical intuition. Moreover, as the discrimination of the states is so well defined the method can be applied to understand more deeply the nature of the transition state that is key in the understanding of many chemical properties.
HLDA can also be extended to treat multiple metastable states (MC-HLDA[30]). In order to illustrate the power of MC-HLDA in the study of chemical reactions we start with a simple S_N2 nucleophilic substitution, namely the substitution of a chlorine atom in the compound CH_2Cl_2 by a Cl^- anion. For this system reactants and products are formally identical and the statistics of the state descriptors in the free-energy basins is permutationally equivalent. To enhance sampling between the three basins we employed the distances between the chlorine atoms and the central carbon atom similarly to what has been done recently by Pfaender et al.[31] and by our group[32] (see Fig. 7). These three descriptors represent a good choice in describing

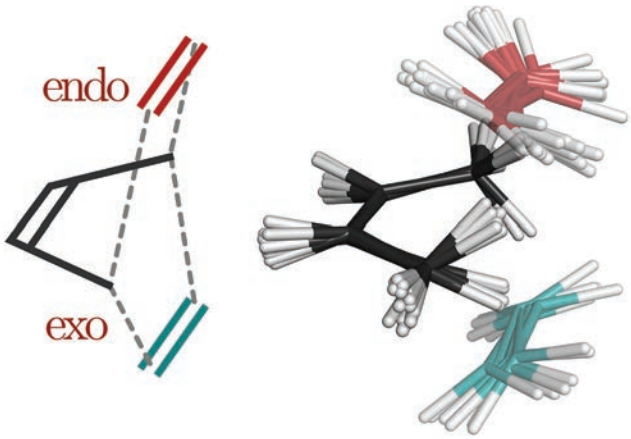

Fig. 6: Overlap of an ensemble of configurations extracted from the barrier top region (s(R) -2.8) resembling the typical text book example of a Diels-Alder transition structure. Here the potentially chiral stereospecific configurations endo and exo are depicted in magenta and cyan.

the slow motions of the atoms in going from reactants to products. More precisely, they are related to the slowest collective oscillations, i.e. low vibrational frequencies, as they involve the dynamics of the heaviest atoms of the system.[33, 34]

Fig. 8 reports the free energy surface (FES) obtained by enhancing the sampling along the two CVs by means of metadynamics. The variables provide an excellent discrimination of the states neatly reflecting the 3-fold symmetry of the problem.

To extract further information we calculated the free energy profile along the minimum free energy path (MFEP).[35, 36] The latter is obtained by taking a series of points lying along a guess path and minimized according to the nudged elastic band algorithm (NEB). In principle such a MFEP is not required to go necessarily through the transition states. However, in the present case it does and the conformations extracted from the apparent transition state do correspond to the classical back-side nucleophilic attack as depicted in blue in Fig. 8. This is a clear indication of the quality of the CVs. The estimated barrier along this path is about 50 kJ/mol. The backside attack is the most probable route in a S_N2 reaction mechanism and it would be responsible for the chiral center inversion if any chirality were present. However, a closer inspection of the FES revealed the existence of another reaction pathway much higher in terms of free energy barriers. The MFEP analysis shows that along this possible, although much less probable path, the system must overcome a barrier in the order of 200 kJ/mol. Rather surprisingly, the analysis revealed the presence of a high energy intermediate in which the nucleophile and the leaving group coordinate the Cl^- atom bonded to the central carbon atom. In physical organic chemistry this mechanism is usually referred to as front-side attack and it is known experimentally[37] and theoretically[38–40] to be high in terms of energy barriers, thus, unlikely to happen. In fact, this mechanism is expected to retain the chirality of the reactants a feature that is almost never observed in pure S_N2 reactions. Therefore, it is clear that the nature of the S_N2 chirality inversion is a conse-

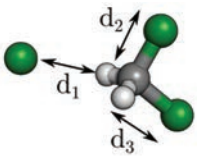

Fig. 7: Carbon-chlorine distances used as descriptors for the S_N2 reaction.

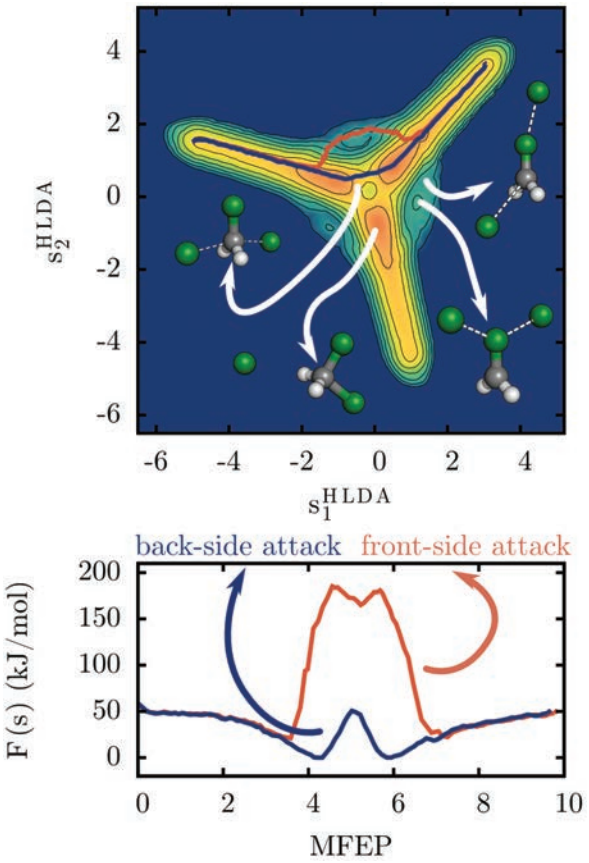

Fig. 8: Upper panel: free energy surface for the nucloephilic substitution of a chlorine atom in dichloromethane showing the two possible reaction paths and associated relevant reference structures. Lower panel: minimum free energy paths along for the two possible mechanisms of the reaction, namely the low-barrier back-side attack (blue line), and the high-barrier front-side attack (orange line).

quence of the chemical kinetics as the barrier for the back-side attack mechanism is much lower than the front-side one.
It is worth underlining that no information on the route the system can take to go from reactants to products has been given as an input. All this wealth of information on the system was hidden in the simple statistics that one collects from a short monitoring of the local fluctuations in the free energy basins.

Another fundamental type of organic reaction is the electrophilic addition. One such process is the hydrobromination of propene. Here, a hydrogen bromide molecule is used to break the propene double bond by adding the hydrogen atom and the bromine to the carbon atoms involved in the double bond. Such a reaction can give rise to two different isomers depending on which carbon atoms the halide group will bind to (see Fig. 9).
It is known that the halide atom prefers to bond to the most substituted carbon. This because the mechanisms starts with the abstraction of the acidic hydrogen by a carbon atom of the double bond implying the formation of a carbocation. Carbocations are much more stable if the carbon center is surrounded by other carbon groups rather than by hydrogen atoms. In the case of propene the central carbon atom has a hydrogen atom formally substituted by a methyl group. Therefore, this stabilizes the transient carbocation with respect to the less substituted carbon in the double bond. The carbocation is formed in the transition region and is unstable, thus, the preference of

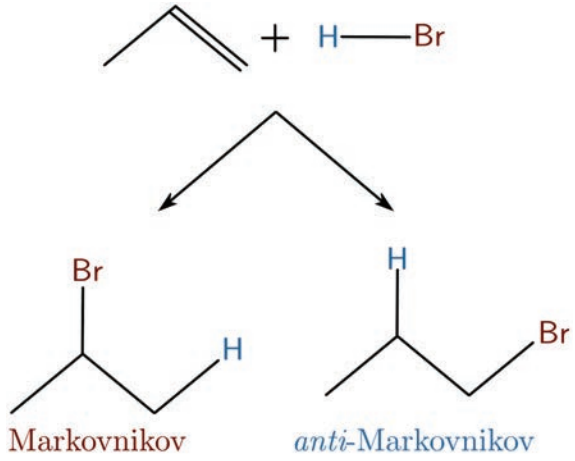

Fig. 9: Reaction scheme for the hydrobromination of propene showing the two possible isomers classified as Markovnikov and *anti* -Markovnikov products.

the bromine atom to sit in the most substituted carbon has to be interpreted as a kinetic rather that a thermodynamic effect. These are the so called Markovnikov rules[41–43] and the associated reaction outcomes are referred as Markovnikov and anti-Markovnikov products (see Fig. 9).

To study this particular reaction we applied MC-HLDA to the three states associated to reactants, Markovnikov, and anti-Markovnikov products states. We used as descriptors the five distances illustrated in Fig. 10. These CVs are both able to break and form the desired bonds but at the same time they embody the necessary information to discriminate between Markovnikov and anti-Markovnikov products.

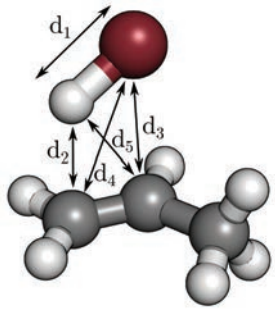

Fig. 10: Fundamental distances used as descriptors for the hydrobromination of propene.

Fig. 11 reports in the upper panel the FES of the hydrobromination reaction. The lower elongated minimum is associated to the reactants state. This state can evolve into the Markovnikov and anti-Markovnikov product states by crossing two different barriers. The lower panel of Fig. 11 shows the free energy profile along the MFEP. Two things are worth noticing in this plot. Firstly, the difference in thermodynamic stability of the Markovnikov and anti-Markovnikov products is almost negligible. Secondly, the barriers separating the states differ from each other by about 80 kJ/mol. Recalling that the transition probability for a system to transit from one metastable state to another is proportional to the exponential of the free energy barrier it is clear that the nature of the Markovnikov's regioselectivity is almost purely kinetic as deduced from empirical observations.

To further support our conclusion we compare the barriers obtained with metadynamics with the ones calculated by means of a NEB optimization of 60 images for both the Markovnikov and anti-Markovnikov route at 0 K on the potential energy surface. The gray dashed dotted line in the lower panel of Fig. 11 represents the energy of the images optimized by means of the NEB algorithm. It is clear that metadynamics is able to repro-

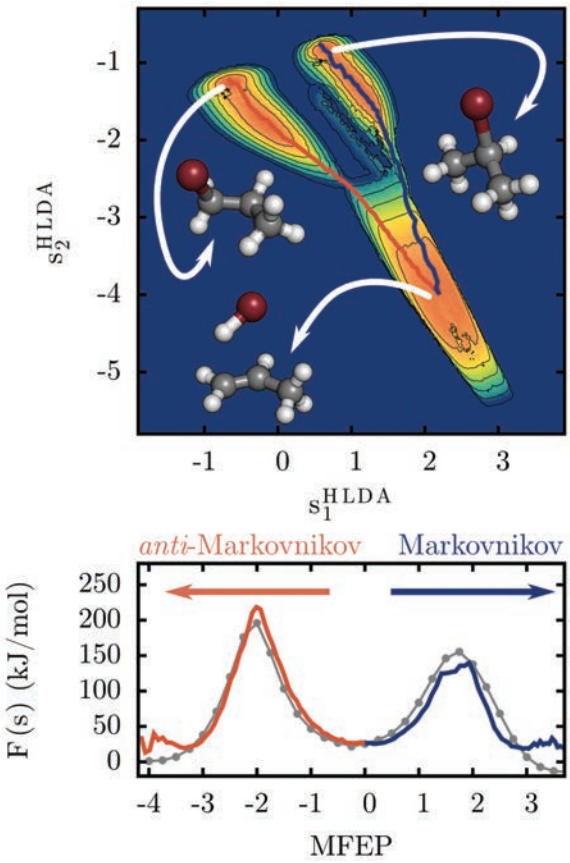

Fig. 11: Upper panel: free energy surface for the hydrobromination of propene showing the two possible reaction paths and associated reactants and products structures. Lower panel: minimum free energy paths along for the two possible mechanisms of the reaction, namely the high-barrier *anti* -Markovnikov's route (orange line), and the low-barrier Markovnikov's route (blue line). The grey dashed line represent the results of the NEB optimization at 0 K.

duce pretty well the energy landscape (see also Ref.[44]). One must keep in mind that metadynamics works on the free energy surface at finite temperature whereas the NEB optimization is applied on the potential energy surface without any account for temperature effects. This fact is the origin of the slight differences in terms of barrier heights between the two methods. The large advantage of metadynamics is that all the entropic effects, even the anharmonic ones, are automatically included since the simulation is performed at finite temperature. Moreover, NEB optimization may be a rather complicated method if the complexity of the reaction rises with an increasing number of important degrees of freedom. This happens when the interpolated images between reactants and products lie far from the ideal minimum energy path resulting in a collection of unphysical configurations. Again, our method does not imply the knowledge of what lies in between reactants and products, as this will be explored automatically by metadynamics directly on the free energy surface, but it rather suggests the proper direction to follow in order to connect them.

METADYNAMICS SIMULATIONS IN COMPUTATIONAL CATALYSIS.

Catalysis is the industrial workhorse of chemistry. A large interest from both academia and chemical industry is focused on the understanding of the fundamental reaction steps that accelerate a chemical reaction.

Computational methods allow unravelling the atomistic details of the catalytic processes.[45] Unfortunately, this is not a trivial task. Metadynamics can be an extremely powerful tool in determining the free energy profile of a catalysed reaction and the relative rates.[14] Essential to this purpose is the choice of the CVs used to characterize the reaction dynamics. HLDA is a formidable technology to reduce the complexity of these problems.

Encouraged by the success of this method in solving many chemical problems we started focusing on a much more complex process. We considered zirconocene catalyzed polymerization of propene.[46–49] For these types of reactions the homogeneous catalysts allows high stereospecificity control of the growing polymer chain. By definition, a catalyst lowers the activation barriers in a reaction. A stereospecific catalysts lowers only the reaction barrier dividing the reactants from the products with a well defined stereochemistry. Understanding the nature of these mechanisms at the atomistic level is fundamental in designing better materials for more efficient processes.

We have performed a preliminary study on these type of systems and the results are very encouraging. We designed the most simple model for a zirconocene complex (see Fig. 12). At this stage of the study no explicit solvent was considered. Moreover, the two cyclopentadienyl groups being symmetric and free to rotate, this model complex is not stereospecific. Nonetheless, it allows us to understand the basis of the activated process in view of more complex situations. Considering the local fluctuations of the four descriptors reported in Fig. 12 for both reactants and products states HLDA provides a simple CV that embodies all the complex rearrangements of the bonding topology for this reaction. The free energy profile (see Fig. 12) allows us to shed light on the fundamental reaction steps. These include the complexation of the zirconium with the pro-

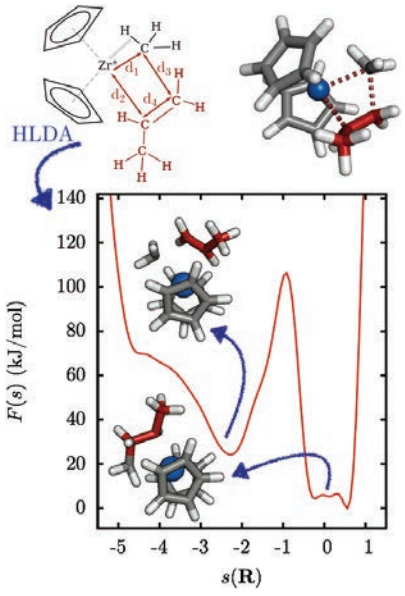

Fig. 12: Free energy profile for the zirconocene catalysed polymerization of propene using the HLDA CV obtained combining the four distances reported in the upper panel figure.

pene -bond represented by the barrierless process in the left basin of the FES and the thermal rearrangement of the growing polymer chain represented by the ripples on the right basis. These first results are comforting and provide us with the first clear information regarding the catalytic process.

ANTI-CANCER DRUGS REACTIVITY

Recently we have studied the effect of water in the reduction mechanism of an important anticancer drug: asplatin.[61] In the presence of a reducing agent like ascorbate, asplatin[62, 63] is reduced to cisplatin and releases the two axial ligands, namely an aspirin ion having an anti-inflammatory effect, and a hydroxyl group. The reduction process implies the formation of a water molecule by abstraction of an hydrogen atom from the ascorbate by the hydroxyl leaving group. However, an alternative route is possible by forming a geminal diol on the ascorbate instead of releasing a water molecule.

Preliminary studies have been conducted for a model system in the gas phase. Using MC-HLDA we obtained two CVs starting from a rather general set of descriptors accounting for the coordination of the complex. Fig. 13 reports the FES associated to this reaction. Higher level calculations are under preparation that involve the inclusion of several water molecules in a periodic simulation box while full *ab initio* density functional theory calculations at the PBE[59] level will be used. Also in this case, a preliminary study on a simplified model system

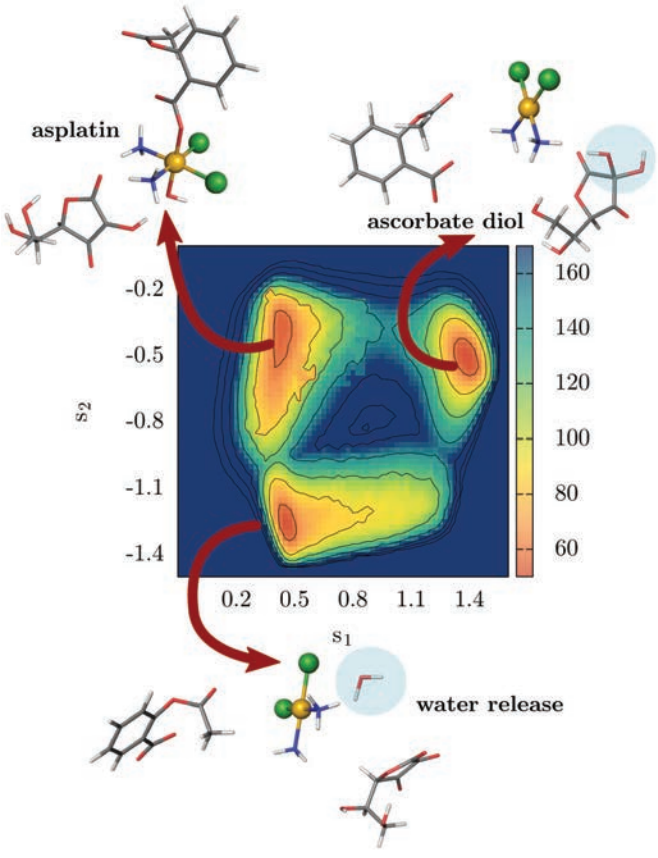

Fig. 13: FES for the reduction of asplatin to cisplatin and water or cisplatin and ascorbate diol. Energy in kJ/mol.

together with a relatively cheap electronic structure method ensures a proper choice of the CVs by means of HLDA. The future research strategy consists in adding increasing layers of complexity, such as the explicit solvent and periodic boundary conditions, while refining the accuracy of the *ab initio* method used to calculate energies and gradients.

ENHANCING THE SAMPLING OF ACID/BASE EQUILIBRIUM

A large variety of chemical reactions occur in a solvent. Among all, water is ubiquitous and plays a fundamental role in inorganic and organic chemistry as well as in biological systems. A very important feature that makes water a special solvent is that in most cases it is not just a spectator but plays a key role in the reactivity.
Acid/base reaction dynamics in water represents one of the computational chemistry holy grails.[50] Traditionally, the calculation of acid-base dissociation constant has been tackled using the static approach dividing the problem of solvation-dissociation by means of Born-Haber cycles. However, although very accurate quantum chemical calculations can be used to calculate the energies of the single states of the Born-Haber cycle, in many cases this approach is not able to grasp the complexity of the problem due to the high fluxionality of solvated systems at finite temperatures. Therefore, several MD free energy methods have been used to cope with the problem of solvation and diffusion of the acid-base groups.[51–53] Unfortunately, due to the elusive nature of the species that form and diffuse through the water network it is hard to generalize the description of the free energy landscape for these processes.

Recently, we suggested a change of paradigm in the way we look at this problem. We consider the fast moving hydrogen atoms as a sea of particles that can migrate from the acid/base species to the water molecules and *vice versa* with relative ease. The solvent is treated as a large macromolecule in which a defect, i.e. the hydronium or hydroxyl ion, is formed after the dissociation. However, due to the hydrogen bond network in water the identification of the charged species may be rather cumbersome. Typically, oxygen-hydrogen coordination numbers have been used to identify such species. Unfortunately, applying a fixed spherical cutoff centered on the oxygen

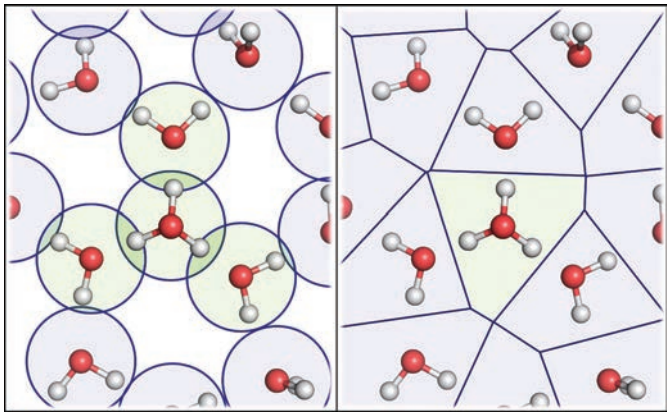

Fig. 14: Coordination number calculated with a fixed spherical cutoff (left) and using a Voronoi tessellation (right).

atoms may result in large ambiguities (see Fig. 14, left panel). To circumvent this problem, we introduced a dynamical and univocal definition of the coordination number using a Voronoi tessellation. This means that no empty space is left when calculating the atoms coordination (see Fig. 14, right panel) but more importantly no ambiguity due to the sphere overlap is possible. This ensures that the definition and identification of the hydronium or hydroxide group is univocal.

Within this perspective, we use this adaptive coordination number to build a pair of CVs, n and d, that by means of MetaD favour acid/base dissociation events and drive the diffusion of the water charged species into the solvent respectively. We applied this method to the dissociations of a weak acid, a weak base, and an amphoteric compound, namely acetic acid, ammonia and the bicarbonate ion respectively. Fig. 15 reports the FES obtained via MetaD using *ab initio* energies and gradients at the PBE+D level for the three cases considered. Considering the most stable configuration for each of them as the reference "undissociated" point we observe a deep minimum for n and d around zero. This is the region in the CV space associated to the absence of defects in water. The bias acting on the CVs eventually makes a dissociation event happen. For the acid species this means the right positive branch is explored with d characterizing the displacement of the conjugate base with respect to the hydronium. On the contrary, for the ammonia solution the formation of the hydroxide and ammonium ions corresponds to the exploration of the left negative branch. For the bicarbonate ion both acid and base branches are explored resembling the amphoteric nature of the system.[54]

These calculations allowed us to study with atomistic detail the thermodynamics of the acid/base equilibria calculating the pK_a at the *ab initio* level. Currently, we are investigating more challenging systems. One is the zwitterionic equilibrium of glycine.[55, 56] These methods allows a very general exploration of the mechanisms and understanding the peculiar behaviour in water of the fundamental building blocks of biological molecules. The future plan is to extensively apply the method to more complex amino acids and more importantly to the study of the acid/base properties of small peptides.

Another challenging class of problems relates to heterogeneous catalysis. Most catalyst act by modifying the local pH conditions of the environment lowering the free energy barriers of a reaction. Nanoporous materials such as zeolites are characterized by the presence of crystallographically well defined Brønsted acidic sites.[57, 58] We aim at using liquid dense water within the zeolite channels to probe the acidity of the Brønsted site. In this way we can understand how the framework influences the dissociation constants of the different and shed light on the diffusion process of the acidic proton in a confined environment. Test calculations have been performed on the acidic zeolite chabazite. MetaD simulations at the PBE[59]+D[60] level using the dissociation and displacement CVs described above are under study, and the preliminary results look very promising. Once the results on the acidic chabazite will be definitive the plan is to move our attention to different types of zeolite frameworks and relate the topology of the crystal structure with its acid/base properties. This project is undoubtedly very ambitious and fraught with several possible diffculties but represents also a potential step forward in understanding the fundamental chemical properties of a very important class of industrial catalysts.

CONCLUSIONS

The computational study of chemical reactions is a challenging field. Metadynamics can become a routine tool in the study of complex chemical transformations where standard methods usually fail. To achieved this goal new methods and computational protocols must be developed with the framework of Metadaynamics. In this report two recent important developments in this field have been presented able to cope with the high complexity of different chemical problems, namely complex chemical rearrangements and acid/base dissociation in aqueous media. These methods and their future extensions may open the possibility to study highly complex situations, offering experimental chemists the opportunity to understand and visualize with atomistic details the ongoing phenomena underlying a chemical reaction. We are confident that these methods may find a large range of applications, from molecular biology/medicine to heterogeneous and homogeneous catalysis.

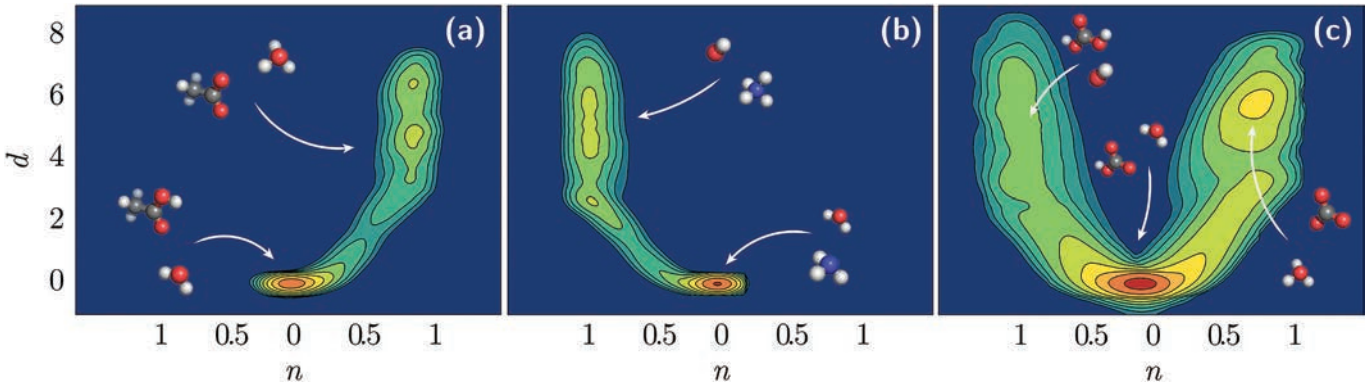

Fig. 15: FESs for the acid/base dissociation of acetic acid (a), ammonia (b), and bicarbonate (c).

REFERENCES

[1] Truhlar, D. G.; Steckler, R.; Gordon, M. S. *Chem. Rev.* 1987, **87**, 217–236.

[2] Friedrich, B.; Herman, Z.; Zahradnik, R.; Havlas, Z. *Adv. Quant. Chem.* 1988, **19**, 257–288.

[3] Gonzalez, C.; Schlegel, H. B. *J Chem. Phys* 1989, **90**, 2154–2161.

[4] Fukui, K. *J. Phys. Chem.* 1970, **74**, 4161–4163.

[5] POLANYI, J. C. *Acc. Chem. Res.* 1972, **5**, 161–&.

[6] Schlegel, H. B. *J. Comput. Chem.* 2003, **24**, 1514–1527.

[7] Friedrich, B.; Herman, Z.; Zahradnik, R.; Havlas, Z. *Adv. Quant. Chem.*; Elsevier, 1988; Vol. **19**; pp 257–288.

[8] Marx, D.; Urg, J. *Quantum* 2000, 1–149.

[9] Abrams, C.; Bussi, G. *Entropy* 2014, **16**, 163–199.

[10] Hartmann, C.; Latorre, J. C.; Ciccotti, G. *Eur. Phys. J. Spec. Top.* 2011, **200**, 73–89.

[11] Laio, A.; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.* 2002, **99**, 12562–12566.

[12] Barducci, A.; Bussi, G.; Parrinello, M. *Phys. Rev. Lett.* 2008, **100**.

[13] Barducci, A.; Bonomi, M.; Parrinello, M. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* 2011, **1**, 826–843.

[14] Tiwary, P.; Parrinello, M. *Phys. Rev. Lett.* 2013, **111**, 230602.

[15] Zheng, S.; Pfaendtner, J. *Mol. Simul.* 2015, **41**, 55–72.

[16] Cramer, C. J. *2nd John Wiley Sons* 2004, 550–555.

[17] Simons, J.; Joergensen, P.; Taylor, H.; Ozment, J. *J. Phys. Chem.* 1983, **87**, 2745–2753.

[18] Schlegel, H. B. *Mod. Electron. Struct. Theory Part I*; World Scientific, 1995; pp 459–500.

[19] Baker, J. *J. Comput. Chem.* 1986, **7**, 385.

[20] Peng, C.; Ayala, P. Y.; Schlegel, H. B.; Frisch, M. J. *J. Comp. Chem.* 1996, **17**, 49.

[21] Billeter, S. R.; Turner, A. J.; Thiel, W. *Phys. Chem. Chem. Phys.* 2000, **2**, 2177–2186.

[22] Govind, N.; Petersen, M.; Fitzgerald, G.; King-Smith, D.; Andzelm, J. *Comput. Mater. Sci.* 2003, **28**, 250–258.

[23] Bonomi, M.; Barducci, A.; Parrinello, M. *J. Comput. Chem.* 2009, **30**, 1615–1621.

[24] Tiwary, P.; Parrinello, M. *J. Phys. Chem. B* 2015, **119**, 736–742.

[25] Ensing, B.; De Vivo, M.; Liu, Z.; Moore, P.; Klein, M. L. *Acc. Chem. Res.* 2006, **39**, 73–81.

[26] Fu, C. D.; Oliveira, L. F. L.; Pfaendtner, J. *J. Chem. Theory Comput.* 2017, **13**, 968–973.

[27] Fisher, R. A. *Ann. Eugen.* 1936, **7**, 179–188.

[28] Mendels, D.; Piccini, G.; Parrinello, M. *J. Phys. Chem. Lett.* **0**, 2776–2781.

[29] Piccini, G.; Polino, D.; Parrinello, M. *J. Phys. Chem. Lett.* 2017, **8**, 4197–4200.

[30] Piccini, G.; Mendels, D.; Parrinello, M. *J. Chem. Theory Comput.* 2018, **14**, 5040–5044, PMID: 30222350.

[31] Fleming, K. L.; Tiwary, P.; Pfaendtner, J. *J. Phys. Chem. A* 2016, **120**, 299–305.

[32] Piccini, G.; McCarty, J. J.; Valsson, O.; Parrinello, M. *J. Phys. Chem. Lett.* 2017, **8**, 580–583.

[33] Vande Linde, S. R.; Hase, W. L. *J. Chem. Phys.* 1990, **93**, 7962–7980.

[34] Tucker, S. C.; Truhlar, D. G. *J. Phys. Chem.* 1989, **93**, 8138–8142.

[35] Branduardi, D.; Faraldo-Gomeź, J. D. *J. Chem. Theory Comput.* 2013, **9**, 4140–4154.

[36] Ensing, B.; Laio, A.; Gervasio, F. L.; Parrinello, M.; Klein, M. L. *J. Am. Chem. Soc.* 2004, **126**, 9492–9493.

[37] Dougherty, R. C.; Dalton, J.; David roberts, J. *Org. Mass Spectrom.* 1974, **8**, 77–79.

[38] Schlegel, H. B.; Mislow, K.; Bernardi, F.; Bottoni, A. *Theor. Chim. Acta* 1977, **44**, 245–256.

[39] Szabó, I.; Czakó, G. *Nat. Commun.* 2015, **6** .

[40] Xie, J.; Hase, W. L. *Science (80-.).* 2016, **352**, 32–33.

[41] Sæthre, L. J.; Thomas, T. D.; Svensson, S. *J. Chem. Soc. Perkin Trans.* **2** 1997, 749–756.

[42] Aizman, A.; Contreras, R.; Galván, M.; Cedillo, A.; Santos, J. C.; Chamorro, E. *J. Phys. Chem. A* 2002, **106**, 7844–7849.

[43] Yang, Z.; Ding, Y.; Zhao, D. *ChemPhysChem* 2008, **9**, 2379–2389.

[44] Franzen, S.; Cochran, K. H.; Weng, J.; Bartolotti, L.; Delley, B. *Chem. Phys.* 2016, **464**, 46–54.

[45] Thiel, W. *Angew. Chem. Int. Ed.* 2014, **53**, 8605–8613.

[46] Kaminsky, W.; Sinn, H. *Zeitschrift für Phys. Chemie*; Springer Science & Business Media, 1988; Vol. **115**; p 442.

[47] Spaleck, W.; Antberg, M.; Dolle, V.; Klein, R.; Rohrmann, J.; Winter, A. *New J. Chem.* 1990, **14**, 499–503.

[48] Brintzinger, H. H.; Fischer, D.; Mulhaupt,¨ R.; Rieger, B.; Waymouth, R. M. *Angew. Chemie Int. Ed. English* 1995, **34**, 1143.

[49] Motta, A.; Fragalà, I. L.; Marks, T. J. *J. Chem. Theory Comput.* 2013, **9**, 3491–3497.

[50] Alongi, K. S.; Shields, G. C. *Annu. Rep. Comput. Chem.*; Elsevier, 2010; Vol. **6**; pp 113–138.

[51] Tummanapelli, A. K.; Vasudevan, S. *J. Phys. Chem. B* 2014, **118**, 13651–13657.

[52] de Alba Ortiz, A.; Tiwari, A.; Puthenkalathil, R. C.; Ensing, B. *J. Chem. Phys.* 2018, **149**, 72320.

[53] Park, J. M.; Laio, A.; Iannuzzi, M.; Parrinello, M. *J. Am. Che. Soc.* 2006, **128**, 11318–11319.

[54] Grifoni, E.; Piccini, G.; Parrinello, M. *Artic. Prep.*

[55] Tuñoń, I.; Silla, E.; Ruiz-Lopeź, M. F. *Chem. Phys. Lett.* 2000, **321**, 433–437.

[56] Campo, M. G. *J. Chem. Phys.* 2006, **151**, 1–8.

[57] Lo, C.; Trout, B. *J. Catal.* 2004, **227**, 77–89.

[58] Derouane, E. G.; Vedrine, J. C.; Pinto, R. R.; Borges, P. M.; Costa, L.; Lemos, M.; Lemos, F.; Ribeiro, F. R. *Catal. Rev.* 2013, **55**, 454–515.

[59] Perdew., J. P.; Burken, K.; Ernzerhof, M. *Phys. Rev. Lett.* 1996, **77**, 3865–3868.

[60] Grimme, S. *J Comput Chem* 2006, **27**, 1787–1799.

[61] Cheng, Q.; Shi, H.; Wang, H.; Min, Y.; Wang, J.; Liu, Y. *Chem. Comm.* 2014, **50**, 7427–7430.

[62] Pathak, R. K.; Marrache, S.; Choi, J. H.; Berding, T. B.; Dhar, S. *Angew. Chem. Int. Ed.* 2014, **53**, 1963–1967.

[63] Ponte, F.; Russo, N.; Sicilia, E. *Chem. Eur. J.* 2018, **24**, 9572–9580.