Catalysis Science & Technology

Ab initio molecular dynamics with enhanced sampling in heterogeneous catalysis

Journal:	*Catalysis Science & Technology*
Manuscript ID	CY-MRV-07-2021-001329.R1
Article Type:	Minireview
Date Submitted by the Author:	17-Sep-2021

Complete List of Authors:	Piccini, GiovanniMaria; Eidgenossische Technische Hochschule Zurich Informationszentrum Chemie Biologie Pharmazie, D-CHAB Lee, Mal-Soon; PNNL, FCSD Yuk, Simuck; United States Military Academy at West Point, Department of Chemistry and Life Science Zhang, Difan; Pacific Northwest National Laboratory Collinge, Gregory; Pacific Northwest National Laboratory, Kollias, Loukas; Pacific Northwest National Laboratory Nguyen, Manh-Thuong; Pacific Northwest National Laboratory, Catalysis Science Glezakou, Vanda; Pacific Northwest National Laboratory, PCSD Rousseau, Roger; Pacific Northwest National Laboratory, pyshical sciences diviosn

Ab initio molecular dynamics with enhanced sampling in heterogeneous catalysis

GiovanniMaria Piccini[1,2,*], Mal-Soon Lee[1], Simuck F. Yuk[1,3], Difan Zhang[1], Greg Collinge[1], Loukas Kollias[1], Manh-Thuong Nguyen[1], Vassiliki-Alexandra Glezakou[1,*], Roger Rousseau[1*]

[1] *Basic & Applied Molecular Foundations, Physical and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA 99352, USA.*

[2] *Istituto Eulero, Università della Svizzera italiana, Via Giuseppe Buffi 13, Lugano, Ticino, Switzerland*

[3] *Department of Chemistry and Life Science, United States Military Academy, West Point, NY 10996, USA.*

Abstract: Ab initio molecular dynamics simulations combined with enhanced sampling techniques are becoming widespread methods to investigate chemical phenomena in catalytic systems. These techniques automatically include finite temperature effects, anharmonicity, and collective dynamics in their robust description of enthalpic and entropic contributions, which can have significant impact on reaction free energy landscapes. This contrasts with standard ab initio static approaches that are based on assessing reaction free energies from various coarse-grained descriptions of the reaction potential energy surface. Enhanced sampling ab initio molecular dynamics opens the way to first principles simulations of systems of increasing complexity like solid/liquid catalytic interfaces. In this work, we aim at guiding the reader through the basis of these techniques, summarizing their fundamental theoretical and practical aspects, and reviewing the relevant literature in the field. After a brief introduction to the problem, we will illustrate the

advantage of using molecular simulations to include finite temperature effects, examine the most common ab initio techniques currently in use, describe their application to solid state heterogeneous catalysts, and finally critically review the most popular enhanced sampling techniques used in computational catalysis.

*Corresponding author:

GiovanniMaria Piccini: giovannimaria.piccini@gmail.com

Vassiliki-Alexandra Glezakou: vanda.glezakou@pnnl.gov

Roger Rousseau: roger.rousseau@pnnl.gov

Introduction

Modeling the reactivity of heterogeneous catalytic systems is nontrivial and poses significant challenges[1-7]. At low temperature and coverage and for systems with well-defined catalytic sites ab initio static calculations deliver many important results allowing interpretation of experimental evidence[8-19]. However, such conditions are often largely ideal, as catalytic processes at realistic experimental conditions imply dealing with a much higher levels of complexity. Common catalytic production processes work at high temperatures and pressures where conditions are far from ideal. Moreover, modern heterogeneous catalytic systems are often characterized by complex interfaces between the solid catalyst and a substrate from the liquid phase (pure or in solution) at room temperature, confined environments, and soft mater scaffolds such as enzymes or liquid catalysts at elevated temperatures. In these scenarios, dynamic and entropic effects play a fundamental role in determining the reactivity of the systems and static ab initio methods face severe limitations in describing their activity.[2] In this context, molecular dynamics (MD) simulations become a fundamental tool for modelling and describing catalytic chemical reactions at realistic experimental conditions.[2, 20-30]

By integrating the classical equations of motions of the nuclei, MD simulations provide one the most accurate approaches to account for finite temperature effects, i.e., entropy. In MD simulations, the time evolution of a chemical system is propagated by integrating numerically Newton's second principle of dynamics $\mathrm{F} = -\nabla \mathrm{V}$, where V is the potential energy function. For sufficiently long MD trajectories one can estimate the Boltzmann distribution of a system at a given temperature and thus its free energy. In most of its successful applications, MD simulations rely on the definition of empirical classic potentials that parametrize forces arising from bond stretching, angles bends, torsions, electrostatic, and van der Waals interactions. For a broad variety

of systems ranging from biological molecules[31] to materials[32], these potentials have shown great capabilities and allow the exploration of the dynamics of systems exhibiting incredible complexity with unprecedented atomistic detail. However, by construction, these potentials do not allow bond breaking or formation, a feature that lies at the core of chemical reactivity, arising from the quantum mechanical electronic properties of the molecular system.

In order to model the type of problems that deal with the electronic properties of molecular systems, one must solve the electronic Schrödinger equation. This produces energy and forces (potential) that allow one to describe the dynamics of a reaction according to its explicit electronic structure. This approach is termed ab initio molecular dynamics (AIMD).[33-36] Unfortunately, this comes at a great expense in terms of computational cost, as the solution of the Schrödinger equation is usually performed iteratively and scales exponentially with the total number of electrons in the system. Density functional theory (DFT) has provided the computational chemistry community with an immensely versatile and efficient tool to reduce the computational cost of electronic structure calculations since its computational cost scales cubically with the total number of electrons in the systems (N^3). This allows calculating up to several hundreds of atoms in the simulation unit cell with reasonable accuracy. Despite its lower computational cost compared to other more accurate quantum chemical methods, DFT still requires much more computational power than classical potentials, hindering significantly its use in MD simulations. The first example of a combination of DFT and MD simulations has been provided by the Car-Parrinello method[35], where the solution of the electronic problem is handled by integrating the nuclei and electronic orbitals coupled equations of motions simultaneously, thus avoiding expensive iterative matrix diagonalizations. Subsequently, the increase of computational power and the enhanced performances of ab initio algorithms allowed the direct use of Born-Oppenheimer potentials to

propagate the dynamics (BOMD)[36, 37]. If the system's size is overwhelmingly large, one can use hybrid methods to calculate, under specific conditions, the total energy and potential of a system. In such cases, the reactive region is treated at the quantum mechanical level (QM) and the remainder is described by a classical molecular mechanics potential (MM). These approaches are commonly referred to as QM/MM methods[38-40]. These last two methods have become nowadays the most used approaches to ab initio molecular dynamics and are playing an increasingly significant role in computational catalysis. More recently, it has been shown how reactive force fields can be obtained ad hoc for specific systems by training deep neural networks or other machine learning schemes on ab initio data. Given the availability of efficient GPU[41, 42] based architectures and algorithms, the applications of these methods are increasing in the simulation community, providing energy and gradients with ab initio accuracy at a cost comparable to classical force fields.

Ab initio methods alone, however, are not sufficient to simulate reactive events in catalysis, particularly those with high free energy barriers□i.e., those that are rate limiting or rate controlling. It is well known that chemical reactions are activated processes. This means that relevant metastable states of a reaction (reactants, products, and intermediates) are separated by high free energy barriers. For this reason, the probability of transition between states is very low and these events are rare compared to low-energy or unactivated events. In the field of molecular simulations this is a well-known issue, commonly referred to as the "time-scale problem".[43-45] Typically, to ensure energy conservation in a MD simulation one needs to integrate the equations of motion with a time step on the order of fractions of femtoseconds. A chemical reaction usually takes place in a time scale that is orders of magnitude larger than 10^{-15} s. For example, for a reaction at 300 K with a free energy barrier of 75 kJ mol^{-1} the average escape time to reach the products state is 2

seconds. Covering this time span with a femtosecond time step several times to collect sufficient statistics on the event is simply intractable.

To mitigate or circumvent this problem, several methods have been proposed to accelerate sampling of rare events. These are denoted enhanced sampling methods. Starting from the seminal work of Torrie and Valleau on non-Boltzmann sampling known as Umbrella sampling[46], a plethora of methods appeared in the literature like parallel tempering,[47, 48] hyperdynamics,[49] transition path sampling,[50] metadynamics,[51] and many others[52]. These methods allow sampling of rare events, enabling control and investigation of reaction mechanisms and estimates of the free energy profiles of the reaction. What differentiates these approaches from more standard methods in quantum chemistry, e.g., the nudged elastic band (NEB) method[53], is that they automatically include with great accuracy the entropic effects that are fundamental in processes like chemical catalysis. As such, there is no need to approximate the free energy of stationary and transition states using models like the rigid roto-translator/harmonic oscillator. In addition, the free energy, ΔG (ΔA), can be decomposed into an enthalpic component, $\Delta H = \Delta U + P\Delta V$ (energetic component, ΔU), and an entropic component, ΔS, thereby enabling one to understand temperature effects upon reaction barriers and rates.

Furthermore, as data science has evolved rapidly in recent years, data-driven methods such as machine learning (ML) and artificial neural networks (NN) have drawn more attention in a variety of fields such as materials science,[54, 55] catalysis,[56-58] and separations.[59] Since conventional AIMD simulations have been challenged due to the multiscale complexity of catalytic reactions, their data-driven methods have helped scientists to improve computational modeling and better understanding fundamental insights of catalysis in AIMD simulations. Therefore, we also discuss the application of these methods in the aspects of improving AIMD for studying catalytic reactions.

Other applications of ML and NNs to static quantum calculations, computational screening, or pathway exploration in catalysis are beyond our scope and readers can refer to other resources.[56, 58, 60-64]

The overall purpose of this review is to discuss AIMD combined with enhanced sampling for the study of catalysis. There is currently a rise in the number of studies in catalysis using these tools, and it is the feeling of the authors that it is highly appropriate to provide: (i) an overview of the most commonly encountered methods in the field, (ii) a discussion of the strengths, weaknesses and pitfalls of the various approaches, and (iii) a series of examples that demonstrate what one can learn from these types of studies. This article is a complement to the recent perspective article published by the authors in ACS Catalysis[2] which emphasizes areas in catalysis where entropy and correlated dynamical motion necessitate a statical mechanics based approach which goes beyond the simulations based on static models (ex. elevated temperatures, confined spaces and solid-liquid interfaces). In contrast, the current article focuses specifically on AIMD/enhanced sampling method and how they have been employed in catalysis. In this article we will review the fundamental theoretical and practical aspects of heterogeneous catalysis modeling using AIMD simulations for problems in catalysis as well as highlight some of the most recent developments and applications of enhanced sampling techniques to relevant problems in catalytic reactivity.

Entropy from unconstrained MD

Molecular dynamics offers the most accurate way to include finite temperature effects in modeling realistic systems at operando conditions. In recent years, ab initio molecular dynamics (AIMD) have also become one of commonly used tools to calculate free-energy changes and their deconvoluted enthalpic and entropic components in not only heterogeneously catalyzed systems,[2, 6, 26] but also in homogeneously catalyzed systems as well[65-70]. The changes in free energy can

quickly be estimated based on the distribution of states extracted from the AIMD trajectory through $\Delta G = -k_B T \ln(p/p_0)$, where p and p_0 represent the probabilities of being in the state of interest and the reference state, respectively. One of the more well-received approaches to accounting for anharmonicity is to employ the quasi-harmonic approximation (QHA) using AIMD data. Alternatively, enhanced sampling methodologies with constrained MD, such as metadynamics[71-73] or the Blue-Moon approach,[74-76] can be applied to calculate entropies with the relationships, $\Delta S = -d\Delta G/dT$ or $\Delta S = (\Delta U - \Delta G)/T$, using the free energy (ΔG) and internal energy (ΔU) obtained from the simulation as discussed in detail later. It is noteworthy that unconstrained AIMD only enables us to calculate ΔS of the end point states, i.e., initial and final states of elementary reactions, while the enhanced sampling methods furthermore allow us to determine ΔS of transition states. In the QHA approach,[25, 26, 77] the vibrational density of states (VDOS) is obtained from the Fourier transform of the velocity autocorrelation function using extracted velocities from the equilibrated AIMD data, as follows:

$$D(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} \langle v(\mathrm{t}_0) \cdot v(\mathrm{t}_0 + t) \rangle dt \qquad (1)$$

where v represents the velocity and the angular brackets is defined as the statistical average over time. This equation allows us to properly describe the anharmonicity of systems at a given temperature.[78, 79] The resulting VDOS, from eqn. (1), can further be employed to properly weight the harmonic partition function via:

$$S_{vib} = k_B(3N - n) \int_0^{\infty} \left[\frac{\hbar\omega}{2k_B T} coth\left(\frac{\hbar\omega}{2k_B T}\right) - ln\left(2sinh\left(\frac{\hbar\omega}{2k_B T}\right)\right) \right] D(\omega) d\omega \qquad (2)$$

where N represents the number of atoms and $(3N - n)$ is defined as the number of degrees of freedom in the system[75]. For gas-phase molecules, the translational and rotational modes are

projected out and are processed separately. Translational and rotational entropy from AIMD trajectories can then be calculated as follows:

$$S_{trans} = k_B ln\left[\left(\frac{24\pi e k_B T}{h^2}\right)^{3/2} \sigma_x \sigma_y \sigma_z\right] \quad (3)$$

$$S_{rot} = k_B ln\left(\frac{8\pi^2 \sqrt{I_A I_B I_C}}{\sigma_s}\left[\frac{2\pi e k_B T}{h^2}\right]^{3/2}\right) \quad (4)$$

where e = exp(1) is the Euler number, k_B is the Boltzmann constant, h represents the Planck constant, σ_x, σ_y, and σ_z are defined as the principal root-mean-square fluctuations of the center of molecular mass as obtained from an AIMD trajectory. I_A, I_B, and I_C are defined as the average principal moments of inertia, and σ_s is the symmetry number. The overall scheme of QHA methodology is summarized in Figure 1. By employing QHA, we can not only define a separate enthalpy and entropy term (needed to extrapolate over a range of temperatures not just the one you computed at) but also decompose the entropy via projection methods to understand its origin.

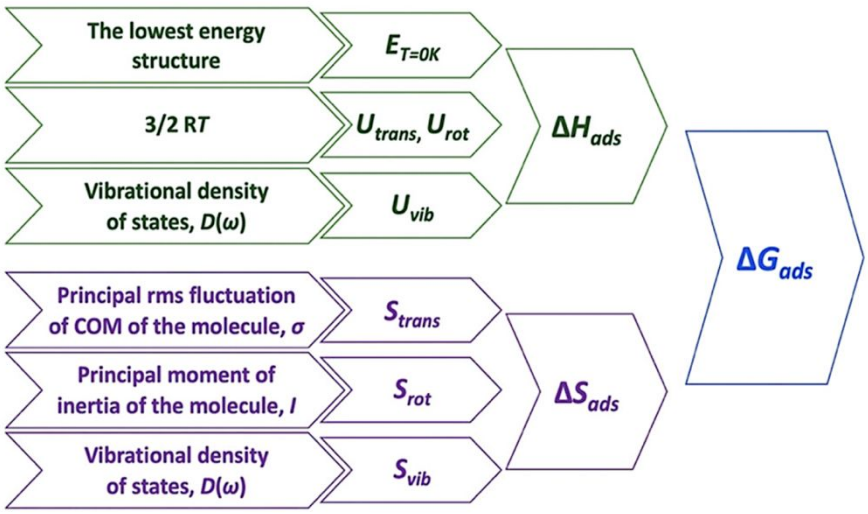

Figure 1. Overall scheme of adsorption free energies derivation based on the QHA. Reprinted with permission from Ref. [2]. Copyright 2020 American Chemical Society.

The computational determination of enthalpic and entropic contributions is a crucial step for understanding the detailed free energy landscape of complex catalyzed systems. Unlike enthalpic contributions to the free energy change which can be qualitatively predicted from the "0 K" DFT energy of a few well-defined minima on the potential energy surface, the entropic component cannot be directly deduced from static DFT calculations. Instead, the harmonic approximation (HA) is frequently applied for studying strongly bound adsorbates at low temperatures where the surface-adsorbate bonding are minimally perturbed. HA depends on the second-order derivative of the Born-Oppenheimer energy surface around the equilibrium. The resulting vibrational modes are associated with high vibrational frequencies, which are adequately described within the limits of HA. In contrast, the influence of anharmonicity as well as collective dynamics in catalytic systems leads to inaccuracy in calculating the entropic contribution using the standard HA approach alone.[80] The atomic fluctuations, shifting away from equilibrated positions, induce atoms and molecules to access the anharmonic regions of the potential energy surface. Such effects become prominent for many adsorbates that are weakly bound, solvated, confined, or at high temperatures. Thus, the determination of entropy via anharmonic consideration is important to gain precise thermodynamic and kinetic perspectives of complex catalytic systems at finite temperatures.[2] Table 1 and Figure 2 show calculated entropies of gas-phase ethanol at high temperatures using HA and QHA which are compared with absolute entropies obtained from NASA polynomials, evidencing the importance of anharmonic effect on entropy even for simple molecules[25]. Table 1 clearly shows strong overestimation of translational entropy but underestimation of vibrational entropy with HA leading to larger difference of entropy from NASA polynomials than QHA.

Table 1. Computed entropy terms (in J/mol/K) from standard approach and harmonic approximations to the vibrations (HA) and the quasi-harmonic approximation (QHA) compared with absolute entropies as obtained from NASA polynomials.[25]

T(K)	QHA				HA				NASA polynomials	% diff for QHA(HA)
	S_{trans}	S_{rot}	S_{vib}	S_{tot}	S_{trans}	S_{rot}	S_{vib}	S_{tot}		
300	114	96	67	277	157	94	22	273	282	-1.8 (-3.2)
700	133	107	125	365	174	104	73	351	359	+1.6 (-2.2)

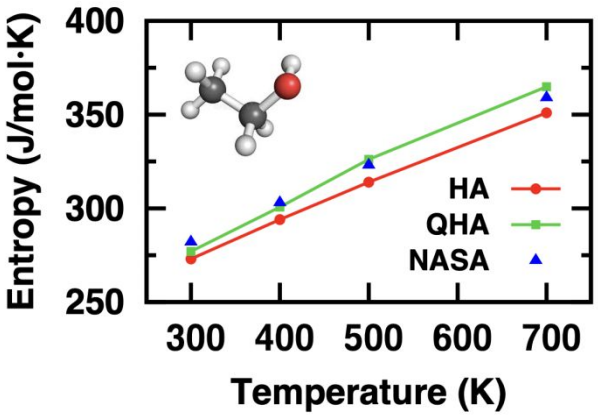

Figure 2. Computed entropies for gas-phase ethanol as a function of temperature using HA and QHA compared with those from NASA polynomials. Adapted with permission from Ref. [25]. Copyright 2016 American Chemical Society.

It is not surprising that multiple approaches have been developed to properly account for anharmonic effects and their influences on entropy.[79, 81-83] For instance Campbell and Seller found an interesting relationship to derive the entropy of adsorbed molecules, including methanol and alkanes, based on their gas-phase entropy.[84] The entropies of such molecules were calculated to be higher than those estimated by HA and can be better described by their gas-phase entropy, as shown below:

$$S^o_{ads}(T) = 0.70 S^o_{gas}(T) - 3.3R \quad (5)$$

where $S^o_{gas}(T)$ represents the standard gas-phase entropy of the molecule, $S^o_{ads}(T)$ is defined as the standard entropy of adsorbates, and R is the molar gas constant. Campbell and Seller's study indicates that even such relatively simple systems are affected by anharmonic motion that makes the standard HA approaches inadequate for estimating entropy and hence equilibrium constants and rate constants for reaction. These approaches are excellent but still rely on expansion about well-defined minmax on an energy surface and thus are appropriate for relatively simple systems (low coverage, anharmonic but no large deformations due to collective oscillations). It is fundamental at this stage of the discussion to point out that the ultimate goal of these calculations is to obtain free energy differences concerning adsorption and reactive activated process, whereas the absolute quantities can be used to merely observe the effect of different approximations on the estimation of such.

In our previous studies of 2D and 3D zeolitic systems, we were able to properly account for the anharmonic effect through the usage of QHA. Figure 3 shows a comparison of Gibbs free energies of adsorbed ethanol over 3D H-ZSM-5 computed with QHA and compared against HA and calorimetric measurements.[25, 26] Here anharmonic effects are large and can lead to errors of several 10s of kJ/mol as compared with experiment as shown in Figure 3. The authors showed a good quantitative agreement between the experimental and theoretical adsorption energies of ethanol. We also showed that the change of entropy from first to second ethanol adsorption on H-ZSM-5 cannot be reproduced with HA.[25]

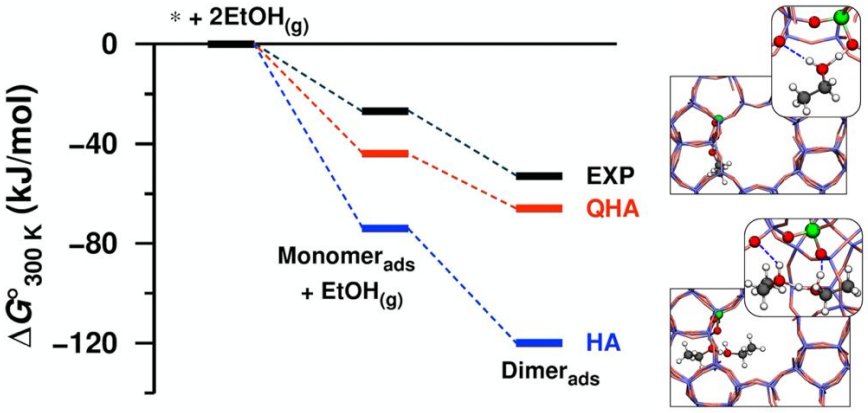

Figure 3. Standard Gibbs free energy of ethanol adsorption obtained from HA and QHA compared with experiments and the most stable structures (at 0 K) of the adsorbed ethanol monomer and dimer on the Brønsted acid site of H-ZSM-5. Reprinted with permission from Ref. [25]. Copyright 2016 American Chemical Society.

For low-coverage systems, QHA gives very compatible results with the more rigorous approach of Piccini and Sauer which more readily incorporates more accurate electronic structure calculations and explicitly accounts for quantum chemical anharmonicity and mode counting from

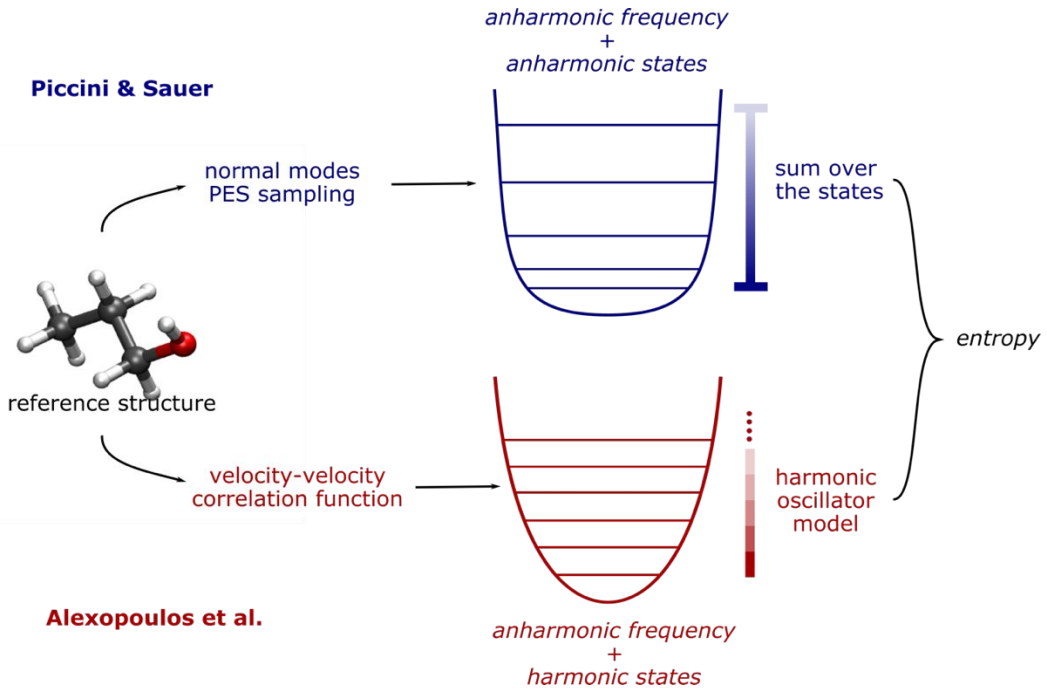

Figure 4. Schematic comparison of the Piccini & Sauer method for thermal anharmonic correction[85] and the QHA by Alexopoulos et al.[25]

hybrid QM/QM methods.[85] Compared to QHA, this approach accounts correctly for the unequal spacing of the vibrational quantum states. To be certain, QHA accurately estimates the anharmonic fundamental frequencies by integrating the velocity-velocity correlation functions in the time domain. However, when it comes to estimates of the thermodynamic functions from a statistical mechanics point of view, the thermal contribution of each normal mode is calculated using a harmonic model, where the spacing between vibrational states is equally spaced. Conversely, in terms of computing the thermodynamic functions, the approach proposed by Piccini and Sauer uses a numerical sum-over-the-states partition function for each anharmonic normal mode, which accounts for the uneven spacing (see Fig. 4 for this comparison). For each single normal mode, this represents a moderate contribution in terms of entropy, and for small systems, it is usually irrelevant. However, for condensed phase systems containing hundreds of atoms, each single anharmonic normal mode contribution can quickly sum up to a non-negligible quantity that may affect the estimation of the thermodynamic functions. Nonetheless, the method proposed by Piccini and Sauer relies on a static picture of the problem under study, where referenced molecular geometries (reactants, products, intermediates, and transition structures) represent well the fundamental thermodynamic states. For these reasons, such a method could be difficult to apply when collective molecular and thermal phenomena become relevant, like in the presence of a solid-liquid interface or very high temperatures.

In extended or confined systems where adsorption phenomena are relevant, the effects of anharmonicity in the estimation of the entropic term are very large. QHA is able to capture these effects, especially in highly fluxional systems for which it is not possible to identify single reference structures. Such systems include those at high temperatures, solid-liquid interfaces, and porous materials such as zeolite or metal-organic framework (MOF). As previously mentioned,

the main limitations of the QHA method reside in the statistical mechanical treatment of the vibrations. Namely, these limitations are (i) all the modes are considered as harmonic fundamental frequencies and no corrections for overtones or combination bands are possible (in most cases, these contributions are often not very large when estimating entropy but may be more substantial for zero-point energies (ZPE), (ii) depending on the size of the system, rather long trajectories are needed to correctly extract the anharmonic contributions from the integration of the velocity-velocity correlation function, (iii) nuclear quantum effects are accounted for at a semiclassical level and are included only in the ZPE, and (iv) QHA can only sample minima that are separated by small energy barriers.

Dynamics on ab initio potentials

In catalytic processes where chemical bonds between atoms are broken or formed, we must account for explicit electronic contributions to the interaction between atoms to correctly predict their thermodynamics/kinetics. On this basis, methods based on electronic structure calculations are commonly used. Keep in mind that while reactive force fields (like ReaxFF[86]) have been used to simulate bond forming and bond breaking processes,[87] they do have problems that can badly affect the accuracy of simulations such as their lack of transferability[88, 89] and unreasonable charge distributions of structures that are far from equilibrium[90]. Semi-empirical methods like tight-binding based approaches are also used by the catalysis community,[91] but they also have significant limitations arising from the parametrization of their Hamiltonians[92]. First principles-based methods, largely based on gradient corrected DFT, are most widely used in computational catalysis. In this section, we will summarize in detail the two most common AIMD techniques for direct simulations of (classical) trajectories, namely, Born-Oppenheimer (BO) approximation[36, 37] and (extended Lagrangian) Car-Parrinello (CP) molecular dynamics[35].

BO molecular dynamics

At the heart of electronic structure theory of a system consisting of N electrons and M nuclei is the Schrodinger equation

$$H\psi_{tot}(r,R;t) = i\hbar\frac{\partial}{\partial t}\psi_{tot}(r,R;t) \quad (6)$$

where $\hbar$ is the reduced Planck constant, {r, R} are the electronic and nuclear coordinates respectively, and H is the Hamiltonian operator given by

$$\begin{aligned} H &= -\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 - \sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 + \frac{1}{4\pi\varepsilon_0}\left(\sum_{i<j}\frac{e^2}{|r_i - r_j|} - \sum_{i,I}\frac{e^2 Z_I}{|r_i - R_I|} + \sum_{I<J}\frac{e^2 Z_I Z_J}{|R_I - R_J|}\right) = -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 + \frac{1}{4\pi\varepsilon_0}\sum_{I<J}\frac{e^2 Z_I Z_J}{|R_I - R_J|} + H_{el} \\ &= -\sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 + V_{NN} + H_{el} = T_N + V_{NN} + H_{el} \end{aligned} \quad (7)$$

where m_e is the electron mass; ε_0 is the vacuum permittivity; and M_I and Z_I are the mass and charge of nucleus I, respectively.

Because of the significant mass difference between the electron and nuclei (the latter are almost 2000 times heavier than the former), we can assume that the electronic wavefunction responds instantaneously to the nuclear variation. The nuclei can be treated as "fixed" while the electrons move in their electrostatic field. The electronic wavefunction is then obtained by solving the following equation.

$$H_{el}\psi_{el}(r,R) = E_{el}(R)\psi_{el}(r,R) \quad (8)$$

The time-independent electronic Schrodinger equation (8) is the major part of the BO approximation. It can be shown that by applying differentiation rules one can separate the dynamics of the nuclei and treat them at the classical level. In such a way, one obtains the classical equation of motion of nuclei:

$$M_I\ddot{R}_I(t) = -\nabla_I[V_{NN} + E_{el}(R)] \quad (9)$$

Eqn (9) shows that nuclei move on an energy surface determined by the repulsion between nuclei and the BO potential energy surface determined by the quantum mechanical treatment of the electrons. In practice, different integration algorithms for the equations of motion, such as Verlet[93] or leapfrog[94], are used to propagate the dynamics of the nuclei. It is important to note, however, that the classical level treatment of nuclear motions fails to address the contribution of quantum effects in proton reactions as in some cases tunneling is significant.[95-97]

To this end, we need to determine the BO potential energy surface. To do that, one can apply several electronic structure methods. BO AIMD, based on DFT, is probably the most widely used AIMD approach. The availability of DFT codes and the easy combination of a DFT code and a MD routine are amongst the reasons for this. Serval codes provide BOMD. The ONETEP[98], FHI-aims[99], CP2K[100, 101] CASTEP[102] and VASP[103] are amongst the most popular BOMD codes.

CP molecular dynamics

In BO MD, the electronic wavefunctions are obtained from the solution of the time-independent Schrodinger equation, the time evolution (dynamics) of the electrons is therefore not considered[104]. In CPMD, a "fictitious" dynamics for orbitals is introduced and incorporated in the form of the following extended Lagrangian:

$$L^{CP} = \sum_i \mu\langle\dot{\phi}_i|\dot{\phi}_i\rangle + \sum_I \frac{\dot{R}_I^2}{2M_I} - [\langle\psi|H_{el}|\psi\rangle + V_{NN}] + \sum_{i,j} \wedge_{ij}(\langle\psi_i|\psi_j\rangle - \delta_{ij}) \quad (10)$$

here μ is the fictitious mass of the electrons, the notation ψ indicates the electronic wave function expanded in the $\{\phi_i\}$ orbital set. The orbitals are also subject to an orthogonality constraint, which is ensured by a set of Lagrange multipliers $\wedge_{ij}$. From the Euler-Lagrange equations, one can obtain the equations of motion for the nuclei and orbitals

$$M_I\ddot{R}_I = -\frac{\partial}{\partial R_I}(\langle\psi|H_{el}|\psi\rangle + V_{NN}) + \sum_{i,j}\Lambda_{ij}\frac{\partial}{\partial R_I}\langle\psi_i|\psi_j\rangle \quad (11)$$

$$\mu\ddot{\phi}_i(t) = -\frac{\delta}{\delta\phi_i^*}(\langle\psi|H_{el}|\psi\rangle + V_{NN}) + \sum_j \Lambda_{ij}\phi_j \quad (12)$$

The last equation indicates that no explicit wavefunction optimization is required in CPMD. However, initial ground state wavefunctions are needed. This can be archived with DFT. As time elapses, the fictitious dynamics of electrons keeps the electron system close to its ground state. As localized basis sets significantly add complexity to the implementation of the CP algorithm, plane-wave basis functions are usually used.[105] CPMD[106] and Quantum ESPRESSO[107] are popular CPMD codes.

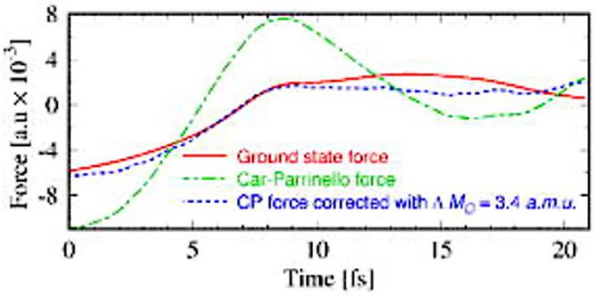

Figure 5. Samples of force components on D_2O molecules based on a 21 fs segment of CPMD trajectory. The Car-Parrinello force (green) and the Car-Parrinello force corrected with constant mass corrections at three points along the trajectory (blue) is compared against the ground-state force at the same ionic positions. Adapted from Ref. [108], with the permission of AIP Publishing

Comparatively, the CPMD and BOMD approaches should give comparative results when both are conducted carefully. The relative choice between the two thus depends on the system under investigation. In general, the CPMD approach requires smaller time steps than BOMD but does not require any iterations of the SCF wavefunction and are fully time revisable unlike BOMD[109]. However, care must be chosen in determining the fictious electron mass as to not destroy the underlaying adiabatic decoupling of nucleic and fictious election degrees of freedom required for the CP approximation to work[108, 110]. Accounting for the errors in CPMD forces is not

a trivial task since they cannot be refined simply based on a constant correction to the masses of the ions, as shown in Figure 5.

Enhancing the sampling of catalytic reactions

As discussed above, MD is a very powerful tool for simulating several properties of many chemical systems. A catalytic reaction is activated processes that bring reactants to products (i.e., chemical reactions) that are of interest to chemistry—and heterogenous catalysis in particular. From a thermodynamic perspective, reactants, products, and intermediates of a reaction are local free energy minima separated by energy barriers. AIMD studies directly observe catalytic elementary steps occurred during simulation time when a study is not associated with a rate limiting step, or is not limited by surface kinetics, or strongly exothermic reactions with low energy barriers. However, energy barriers are high that are much larger than a Boltzmann factor, k_BT. For this reason, the probability of observing a transition event is very low, and the process is associated with extremely long-time scales that are much larger than those reachable in MD. For a long time, this limitation has hindered the application of MD to problems of interest in chemistry and catalysis.

One strategy for circumventing this problem is the application of an external fictitious bias potential to the system in order to discourage the simulation from sampling already visited regions of phase space. The first example of this technique can be found in the landmark paper of Torrie and Valleau[46] in which umbrella sampling (US) was introduced. There, the authors proposed to add an external bias along a designated collective variable (CV)—i.e., a function $s(R)$ of the Cartesian atomic coordinates, R, of the system—that describes the thermodynamic process of interest. A bond distance in a chemical bond breaking or a dihedral angle in a conformational isomerization are examples of CVs. In such a way, US achieves a non-Boltzmann sampling of the problem at a given temperature, as the bias potential that prevents the system from exploring

already visited regions of the free energy landscape modifies the probability distribution function, reducing the effective height of the barriers in the biased ensemble.

Since the US method was introduced, a large variety of methods have been proposed to overcome limitations related to high free energy barriers in MD simulation and thereby permit the study of rare events (see, e.g., Ref. 52). Here, we will focus only on methods based on biasing along a CV to obtain a free energy profile since these methods are the most appropriate to high precision studies of specific chemical reactions. Namely, we will focus on two families of CV-based methods: gradient-based ones, like the Blue-Moon ensemble method, and non-Boltzmann sampling, like US metadynamics (MetaD). The latter will be described in a bit more detail as it has become one of the most popular and widely applied methods in chemical simulations and also allows for the calculation of reaction rates. Other methods involving many replicas of the systems at different conditions like parallel tempering,[47, 48] or free energy perturbation methods[111, 112] are powerful tools as well but not the most commonly used in AIMD for studying chemical reactivity due to prohibitively large number of configurations required to converge the statistics.

Collective variables for chemical reactions

Collective variables (CVs), $s = s(\mathbf{R})$ are functions of the atomic coordinates $\mathbf{R}$ and are meant to describe the progress of a reaction from reactants to products. Ideally, the best collective variable is the reaction coordinate, i.e., the line that follows the minimum free energy path from reactants to products. In practice, this function is not known a priori, and its identification is completely non-trivial. Thus, approximations are made in terms of simpler chemical parameters such as bond lengths, angles, torsions, coordination numbers, taken alone or combined in lower dimensional forms, e.g., in linear or non-linear combinations. The fundamental requirement for a CV is to discriminate between the relevant states of a reaction and *it must include all the slow degrees of*

freedom that are relevant to the process. These are fundamental requirements to ensure proper sampling and, thus, convergence of the free energy estimation.

Typical examples can be found in organic S_N2 reactions, where the distances between the central carbon atom and the leaving group, and the one between the carbon atom and the nucleophile discriminate well between reactants and products.[113, 114] Often, complex chemical reactions involve several degrees of freedom and the use of these degrees of freedom as independent collective variables implies dealing with large multidimensional free energy surfaces. This comes with several disadvantages. First, the physical and chemical interpretation of the results is compromised since high dimensional surfaces are not easily readable and projection onto lower dimensions is needed. Second, from a technical point of view, the convergence rate and efficiency of most CV-based free energy methods drops dramatically when many degrees of freedom are treated independently (see, e.g., Ref. [71]). A practical solution is to combine the most relevant CVs into low dimensional, possibly 1D, using, e.g., linear or non-linear combinations[115]. For simple reactions, chemical intuition can be used straightforwardly to determine the optimal coefficients of the combination. However, for systems of increasing complexity this may be a daunting task. Recently, it has been shown that low dimensional CVs in the form of simple linear combinations can be obtained through automating the concept of chemical intuition [116-118]. Starting from the statistical information provided by the fluctuations of the system in the reference free energy basins, i.e., reactants and products, one can use simple classification rules and algorithms to obtain linear combinations of several chemical descriptors, like distances, angles, coordination numbers etc., to be used as optimal CVs. Later, this approach has been further developed beyond its linear formulation using neural networks to introduce non-linearity, providing an even more flexible tool to obtain low dimensional CVs.[119]

If one has access not only to the information provided by the local basins of attraction but also to the dynamics of the rare event it is possible, in principle, to obtain even more accurate CVs. If this information is available from unbiased simulations (usually only for conformational changes in biomolecular systems[120] or from biased ones one can extract optimal collective variables from the transition dynamics. It is not our purpose to dig deep into the details of these specialized methods first developed by Noè and Pande and later applied to the enhanced sampling problems[117, 121, 122] but they represent a very promising and efficient solution to approximate the reaction coordinates.

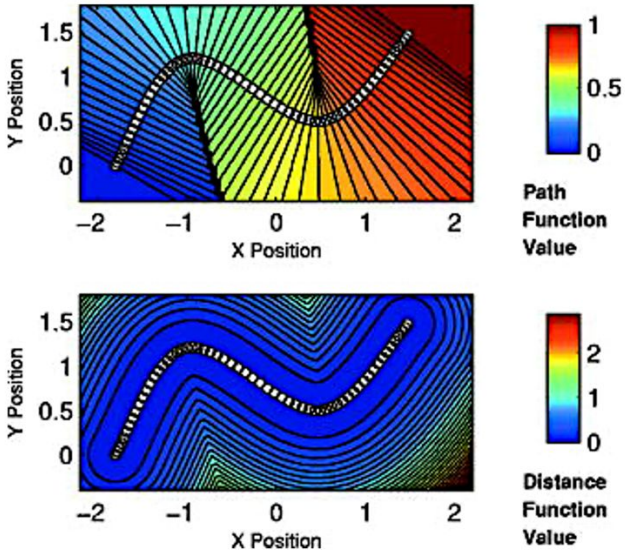

Figure 6. Top: Surface plot of variable s for sixty points in two dimensions (white dots). Note that isolines are perpendicular to the path in its neighborhood. Bottom: Contour plot of variable z in two-dimensional space shows that its definition can be approximately considered as a measure of the distance from the path itself. Reprinted from Ref. [123], with the permission of AIP Publishing.

Alternatively, if one knows a priori the mechanism of a reaction, i.e. can approximate a reaction path through a series of plausible molecular structures, special CVs, commonly referred to as path collective variables (PathCVs) can be used to sample the chemical reaction.[123] The development of these CVs does not differ substantially to the grounding ideas of the nudged elastic band (NEB) method [124] and PathCVs can be seen as the finite temperature version of it. In simple terms, having

a series of images (molecular structures) representing the change from reactants to products, one can build two PathCVs, s (Path Function Value) and z (Distance Function Value), that describe the progress of a reaction and the distance from the minimum free energy path, respectively (see Figure 6).

PathCVs are a very powerful method but require the knowledge of the reaction mechanism beforehand, which is often not available. A very promising variant of this method, based on coordination numbers metrics, is way more robust for sampling chemical reactions and does not require the knowledge of intermediate structures.[125] Topology-based methods such as Social PeRmutation INvarianT (SPRINT) coordinates represent a third alternative to the problem and have been used for chemical reaction problems, as well.[126] However, despite their great flexibility, SPRINT coordinates may suffer from a limited resolving power in the case of homogeneous bulk systems with uniform coordination patterns. In this context, Pietrucci et al. used these Path and SPRINT CVs to investigate the formation of fullerene from graphene flakes using biased AIMD simulations (see Figure 7).

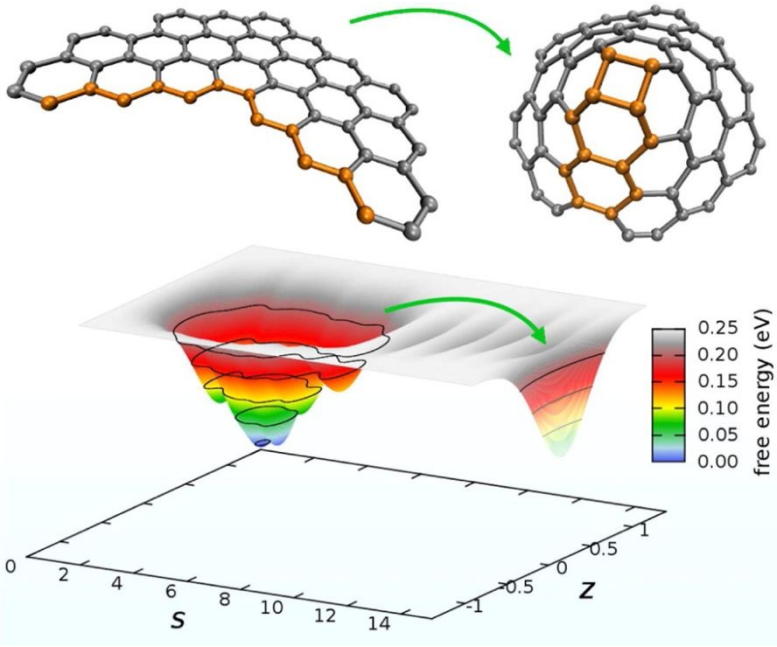

Figure 7. Graphene nanoflake: zipping into a nanocone and free-energy landscape as a function of the path collective variables. Reprinted with permission from Ref. [127]. Copyright 2014 American Chemical Society.

Methods to enhance sampling along collective variables

As previously mentioned, enhanced sampling techniques can be divided into two major families, one based on sampling along collective variables and one based on tempering techniques. Here, we focus only on CV-based methods as they are the most commonly used in chemical reaction sampling and have led to major results in computational catalysis. The reason is that catalysis is often associated with large changes in the entropy of the system and CV-based methods increase the probability of crossing entropic barriers in CV space; hence exploring configurations of largely different entropy. Such a family can be further subdivided into two complementary approaches: gradient methods like thermodynamic integration (TI),[128, 129] Blue-Moon ensemble,[74] or adaptive biasing force algorithm,[130] and non-Boltzmann sampling like umbrella sampling[46] or metadynamics[51]. Figure 8 summarizes the main features of the two method families.

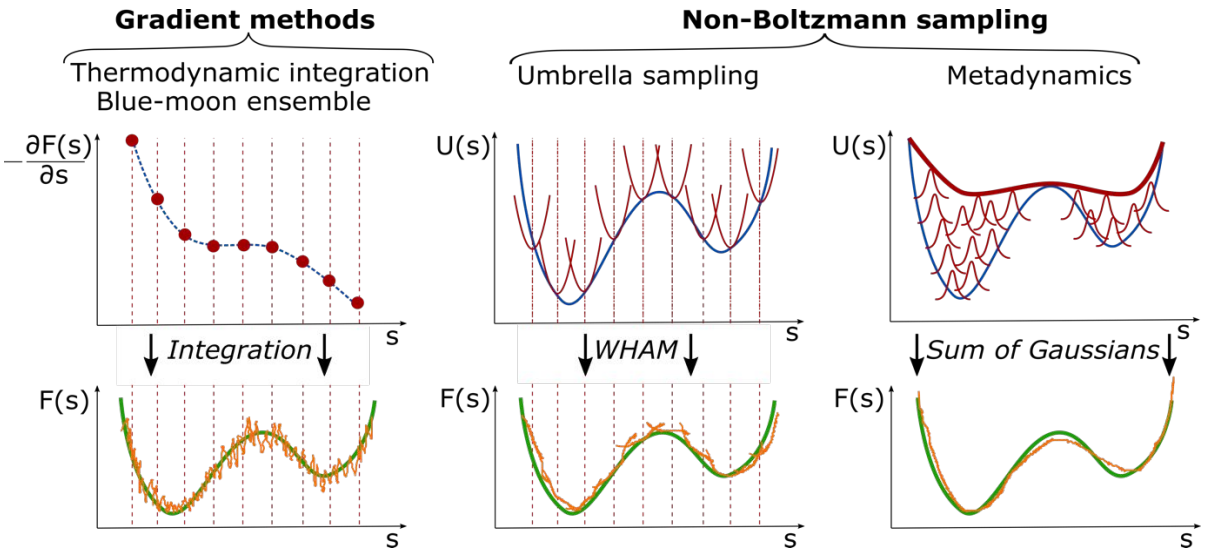

Figure 8. Schematic representation of gradient and non-Boltzmann sampling methods. The upper panel reports the potential (blue) that is sampled during the simulation and the integration points or applied bias (red). The lower panel is related to the top one by the method used for reconstructing the free energy profile (orange) that approximates the real unknown free energy (green).

Gradient methods

These methods aim at reconstructing the free energy along a designated CV by integrating its gradient ($\partial \mathrm{F}/\partial \mathbf{s}$). The first example of this technique is represented by thermodynamic integration, where one calculates the mean force (average gradient) of the free energy at specific points of the CVs and subsequently reconstruct the free energy by numerical integration as $\mathrm{F}(\mathbf{s}) - \mathrm{F}(\mathbf{s_0}) = \int_{s_0}^{s} (\partial \mathrm{F}/\partial \mathbf{s})\, \mathrm{d}\mathbf{s}$. Often, in order to estimate the mean force at a point, s_i, of the CV space, one needs to constrain the value of the CV to sample a proper ensemble of configurations of the system. This is usually done by using holonomic constraints but introduces fictitious forces that must be "weighed out" to reconstruct the equilibrium free energy. This method, proposed originally by Ciccotti et al.,[74] is referred to as the Blue-Moon ensemble, as it aims at observing events so rare they occur "once every blue moon". Choosing ahead of time the quadrature points, $\{s_i\}$, for integrating the free energy may be rather cumbersome for complex systems and more practical solutions that do not depend dramatically on this choice are desirable. This is the essence of

adaptive methods like the adaptive-biasing force[74, 131] which allow for a more flexible choice of the initial configurations and, thus, make life easier when it comes to converging the free energy estimate along the CV.

Non-Boltzmann sampling

These methods aim at reconstructing the equilibrium free energy profile of a specific activated process from a non-Boltzmann probability distribution. To achieve this goal, sampling is performed adding a bias potential to the underlying potential energy surface along a designated CV. By doing so, one prevents the system from sampling already visited regions of the configurational space, thus, eventually favoring transition between metastable states, i.e., the local minima of the free energy along the CV. By changing the potential energy landscape along a CV, $V_b(\mathbf{s}) = V(\mathbf{s}) + \Delta V(\mathbf{s})$, one modifies the ensemble in which the sampling occurs, and equilibrium properties can be reconstructed by statistical reweighting procedures. In the seminal work introducing US, Torrie and Valleau proposed that non-Boltzmann sampling can be achieved by adding a static bias potential in the form $\Delta V(\mathbf{s}) = -k_B T \log w$ where w is a trial function that aims at flattening the probability distribution over the CV space, thus removing high free energy barriers separating metastable states. The function w is thought of as an "umbrella" covering both easy and hard to sample regions of the configurational space and must be found by trial-and-error procedures.[52] This technique represents a big step forward in terms of free energy sampling. However, its application to complex systems is largely hindered by the trial-and-error procedure to determine the bias potential. To overcome these technical limitations, a more practical solution has been proposed, where a series of local restrain potentials (windows) are placed along fixed points over the CV space, similarly to TI methods, and different replicas of the simulations are performed for each window. The final Boltzmann equilibrium distribution is then reconstructed

by the weighted histogram analysis method (WHAM)[132] that combines properly all the information collected from the replicas in one single statistical ensemble.

As mentioned, the main difficulty of US lies in the choice and design of a proper bias potential or, in the case of multi-replica simulations, in the large number of calculations required to ensure convergence. A great alternative to this approach is represented by metadynamics.[51, 133] The fundamental idea behind MetaD can be found in previous formulations like "local elevation",[134] where an on-the-fly and time-dependent bias in the form of sum of Gaussian functions is deposited along some order parameters, e.g. dihedral angles of a protein, to push the system away from its free energy basins of attraction. The great advancement of MetaD is the possibility of reconstructing the unbiased free energy landscape of the process under study, and thus allowing for quantitative measurement of thermodynamic and kinetic properties. Given its popularity and wide application range in the field of computational catalysis, here will follow a more detailed section on this method in its well-tempered variant (WT-MetaD)[133], clarifying the fundamental theoretical and practical aspects.

Well-tempered Metadynamics in a nutshell

In WT-MetaD, a history dependent biased potential is deposited along one or more CVs {**s**} at fixed time intervals. Such a bias takes the form of a sum of Gaussian kernels whose height decreases as long as the simulation proceeds

$$\Delta V_n(\mathbf{s},t) = \sum_{t' = \tau_G, 2\tau_G,\ldots}^{n} w(t) \exp\left(-\frac{|\mathbf{s}(t) - \mathbf{s}(t')|}{2\delta \mathbf{s}}\right) \quad (13)$$

where $w(t)$ is the Gaussian height at time t, τ_G is the time interval at which Gaussians are deposited, and $2\delta s$ is the width of the Gaussian kernel. In standard MetaD $w(t)$ remains constants throughout the whole simulations, whereas in WT-MetaD its value is set according to $w(t) = \omega_0 \tau_G \exp$

$(-\Delta V(\mathbf{s},t)/k_B\Delta T)$, where ω_0 is the initial "deposition rate", with units in Gaussian height per time, and ΔT is a parameter controlling the excursion velocity of the system from free energy minima. It can be demonstrated[133] that WT-MetaD converges smoothly to the true free energy times a multiplicative constant and has the advantage that only specific barriers are crossed by tuning the ΔT parameter.

During a MetaD simulation the system is stochastically pushed away from already visited regions of the CV space. In this way, the fluctuations within the basin of attraction are enhanced and transition to other metastable states are accelerated. This gives access not only to a full exploration of the desired chemical rare event but allows one to determine the shape of the free energy landscape and calculate free energy differences including all the relevant entropic contributions without the need of approximations, e.g., harmonic or anharmonic frequency calculations, to estimate the thermodynamics of the process. As a general rule, the quality of the results must be supported by a convergence test of the free energy landscape obtained. This can be done by simply integrating the free energy within the basins of attraction corresponding to reactants and products and calculate the difference ΔF every N time steps during the simulation. In such a way, one can monitor the oscillating behavior of the free energy difference between the two refence basins which should approach a specific value asymptotically. The oscillations are due to the system being pushed forth and back among the free energy basins. Therefore, to converge a free energy surface, one needs several recrossing of the barriers separating them. Only when such oscillations are small within a specific range (optimally $\pm\ 1\ k_BT$) the simulation can be considered converged.

From the free energy profile, one may also derive an approximate estimation of the free energy barrier although this quantity is strongly related to the quality of the CV used to sample

and project the free energy process. A more quantitative approach that gives access to reaction rates is provided by a variant of MetaD referred to as infrequent metadynamics[72], a method inspired by Voter's hyperdynamics approach[49]. In this method one starts with N independent replicas of the system from a reference metastable state configuration, say the reactants state, initialized with different velocities and properly equilibrated. Then, a MetaD bias is applied on each replica independently using very long deposition intervals (long τ_G) and the simulations are stopped when the system has escaped the initial metastable state. The slow deposition rate ensures a rather low probability of depositing a bias on the transition state. In such a way, using transition state theory one can calculate for each replica the physical (unbiased) escape time t^* simply by reweighting the MetaD simulation time t_{MetaD} by the ensemble average of the bias as $t^* = t_{MetaD} < \exp(\beta\Delta V(\mathbf{s})) >_V$, where $\beta = 1/k_BT$ is the inverse temperature. From the collected escape time of each replica, one can obtain a cumulative distribution probability function, fit it to a Poisson-like distribution, and calculate reaction rates as $k_{R\to P} = 1/ < t^* >$ where the term in angular brackets is the mean escape time for all the N replicas obtained from the Poisson fit[135]. In this context, Xie and Barton used infrequent metadynamics simulations (InMetaD) to calculate the activation energy of glucose 6-phosphate hopping between association states in electrolytes (see Figure 9).

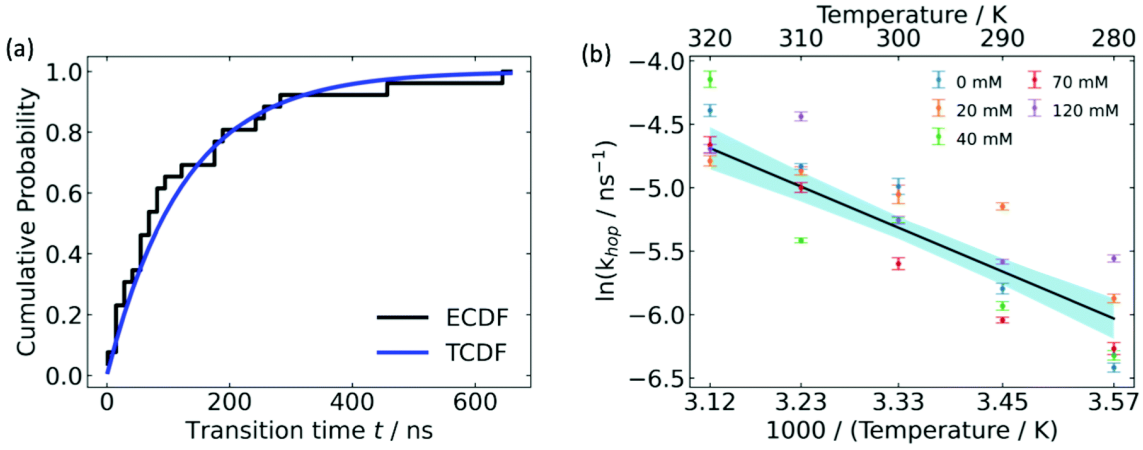

Figure 9. Hopping activation energy by InMetaD. (a) Cumulative probability distribution of transition time from independent InMetaD simulations at IS = 0 mM, 310 K. Theoretical cumulative probability distribution (TCDF) is in blue, and the stepwise empirical cumulative probability distribution (ECDF) in black was fitted from the unbiased transition times, *t*. (b) Arrhenius plot of the hopping rate at varying ionic strength. The black line represents an Arrhenius fit for all ionic strengths. The shaded area in blue is the 95% confidence interval. Reproduced from Ref. [136] (under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence) with permission from the Royal Society of Chemistry.

Data-driven methods for improving AIMD simulations

As mentioned above, data-driven methods such as machine learning and neural networks have been applied to the study of molecular simulations in catalytic reactions.[137-141] However, directly employing these methods to AIMD simulations is still under-developed, and their benefits could generally be considered in three categories. First is that these approaches have directly accelerated AIMD simulations where the expensive and repetitive energy and force computations required in AIMD simulations have resulted in significant bottlenecks.[142, 143] These ML-based algorithms have been proposed to learn past configurations of AIMD on-the-fly and predict energy and force of new configurations faster than conventional AIMD simulations.[144] The second categorical advantage is that these methods enable the generation of force field parameters with quantum accuracy based on established AIMD simulations for catalytic systems.[56, 145, 146] The reactive dynamics on catalysts can be explored under conditions comparable to experiments.[147] They can also push the AIMD simulations to longer time scale and larger simulated system sizes within acceptable simulation times.[148] Similarly, conventional AIMD simulations have difficulty exploring solvent and particle size effects in catalysis due to the need for enough statistical sampling, but the neural network force field derived from AIMD data can provide a powerful tool, e.g., for complicated structure and composition identification under real reaction conditions.[149, 150] Lastly but not least, these methods can provide some insights by the in-depth analysis of the large

amount of AIMD-produced data in catalysis simulations. For instance, ML is applied to interpret AIMD simulations to extract relevant information in chemical reactions with minimal prior input of knowledge on these reactions.[151, 152]

Applications of enhanced sampling AIMD in computational catalysis

Here, we review the most relevant applications of different CV based enhanced sampling techniques in catalysis to provide a better understanding of when it is appropriate to use this method and what can be learned from this approach. We focus on nanoporous materials and soli-liquid interfaces where these methods are the most ubiquitously encountered due the nature of confinement effects which challenge simple harmonic approximations. Here, we review the application of enhanced sampling techniques to the study of the reactivity in three important classes of catalytic systems, namely zeolites, metal organic frameworks (MOFs), and solid-liquid interfaces.

Zeolites

Acidic zeolites are among the most widely used materials in catalysis and play an important role in the petrochemical industry[153] and more recently in the conversion of biomasses[154-158]. The catalytic activity in these materials is locally concentrated in the proximity of the Brønsted acid site (BAS), where a bridging OH group is shared by an aluminum and a silicon atom belonging to the material framework (see Figure 10).

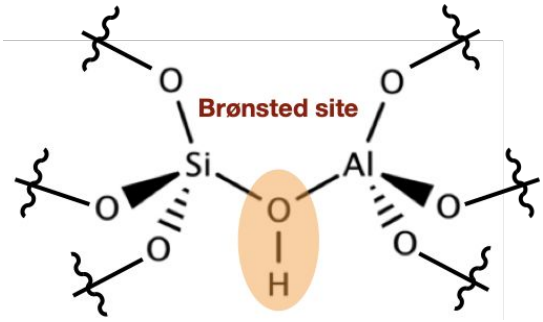

Figure 10. Schematic representation of a BAS showing the bridging hydroxyl group embedded in the zeolite framework.

This specific structural feature of acidic zeolites is a great advantage in terms of computational modelling as not many heterogeneous systems are as well characterized. This gave rise to a large research effort to describe from first principles the reactivity inside these materials, shedding light on the mechanistic aspects as well as their thermodynamics and kinetics. One of the most relevant industrial reactions that is catalyzed efficiently by zeolites is the methanol to hydrocarbons (MTH) process, where a new carbon-carbon bond is formed between a light alkene molecule and a zeolite-activated methanol (see Figure 11). This reaction is very challenging from a computational point of view as its reactivity is driven by the adsorption of methanol at the BAS as well as the supramolecular complex formed with the physiosorbed alkene.

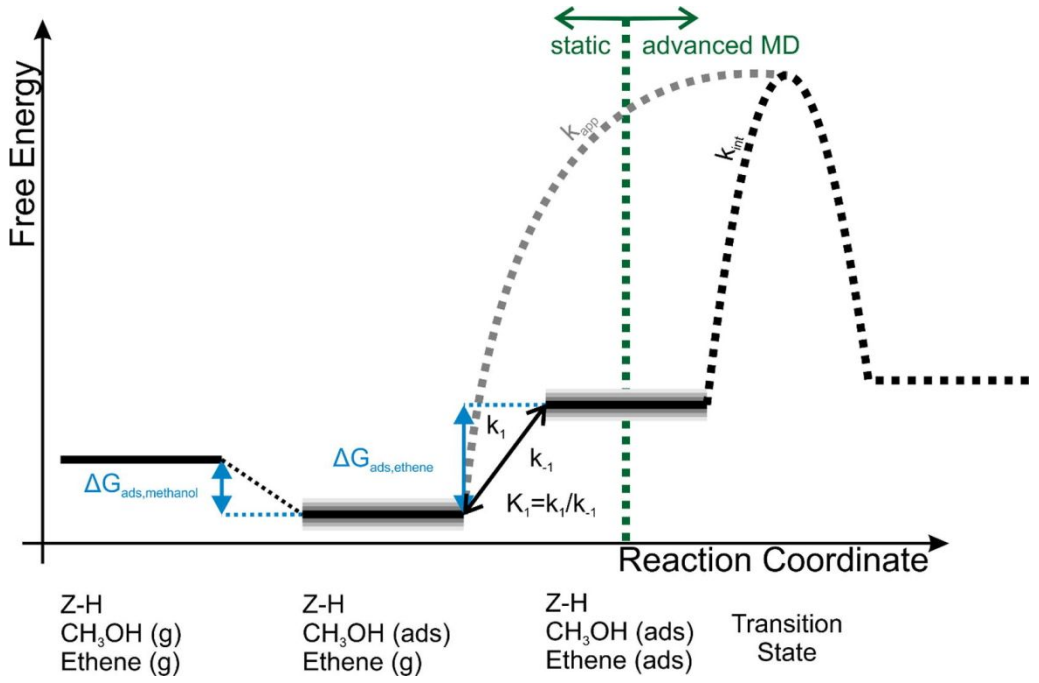

Figure 11. Schematic representation of the relation between apparent and intrinsic kinetics. For the calculation of $\Delta G_{ads,ethene}$ various schemes are used, shown by the grey uncertainty in the co-adsorption step of ethene. The value of k_{int} is obtained dynamically in this work.[159] Reprinted from Journal of Catalysis, Vol. 388, Bailleul, S., Dedecker, K., Cnudde, P., Vanduyfhuys, L., Waroquier, M., & Van Speybroeck, V., Ab initio enhanced sampling kinetic study on MTO ethene methylation reaction, 38-5, Copyright 2020, with permission from Elsevier.

From the point of view of enhanced sampling simulations, the group of van Speybroek has provided important methodological and application advancements in modeling the MTO reaction in zeolites. In a series of articles,[159-163] they use different CV-based enhanced sampling techniques to study the methylation of different olefinic and aromatic substrates in acidic zeolites. Using TI, US, MetaD and variationally enhanced sampling (VES)[164] – a variational approach to metadynamics – the authors provided a dynamical picture of the reaction mechanisms[162, 163] and qualitative estimates and trends over olefinic chemical series of the reaction barriers[159-161] along the CV. The use of enhanced sampling techniques combined with MD simulations allows one to describe the complex reactivity of these processes at real experimental conditions, where in static approaches one may face limitations.

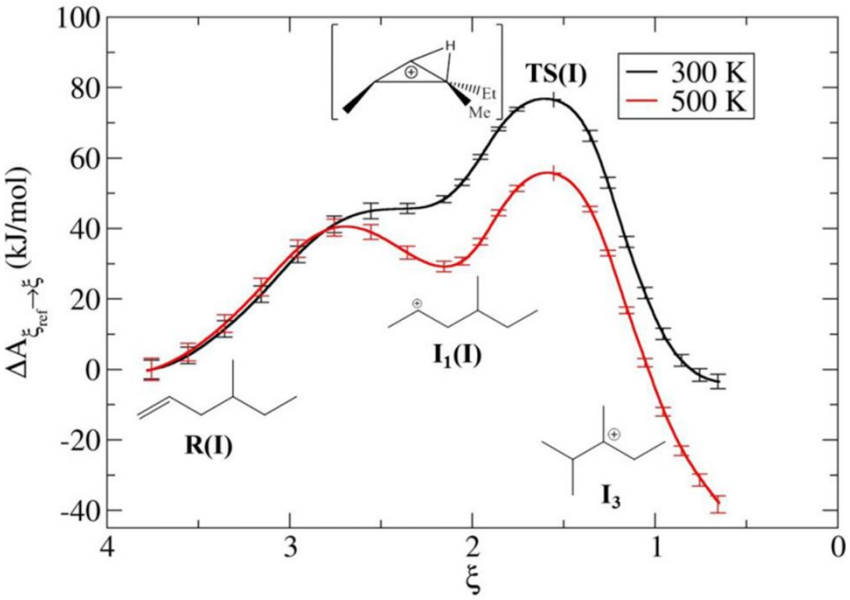

Figure 12. Dynamic features of transition states for isomerization of alkenes over chabazite at two different temperatures based on enhanced sampling AIMD simulations—in this case the Blue Moon ensemble method. Reprinted with permission from Ref. 163. Copyright 2019 American Chemical Society.

Zeolites are also used very frequently at the industrial level for alkene cracking to optimize the size of hydrocarbons. It is generally assumed that the mechanism of the reaction proceeds via protonation of the olefinic C=C bond by means of the BAS of the zeolite generating a carbeniun ion intermediate that subsequently undergo to β-scission leading to shorter carbon chains. Recently, Bučko, Chizallet and coworkers investigated the zeolites catalyzed acidic cracking of alkenes focusing on the skeletal isomerization[165, 166] and β-scission mechanisms[167] using AIMD in combination with Blue Moon ensemble calculations. Some of these computational demonstrations of unique chemistry over zeolites can be seen in Figure 12. In such a way, they have sampled the possible reaction pathways leading to different products involving carbenium ion intermediates, shedding light on the mechanistic aspects driving the catalytic reaction acceleration and the selectivity induced by the material confinement. Van Speybroek et al. used ab initio MetaD to

additionally investigate the stability of cracking intermediates in zeolite H-ZSM-5, which confirmed that carbenium ions are cracking intermediates rather than alkoxides.[168]

More recently, the application of zeolites in the conversion of alcohols have been studied at large. Alcohols are very abundant compounds in lignocellulosic biomass streams.[169, 170] In order to be useful for other chemical processes involving synthesis steps like carbon-carbon bond formation, they must be converted into olefins. Acidic dehydration followed by Friedel-Crafts alkylation with aromatic compounds leads to useful supplements for non-fossil fuels. It has been shown experimentally that zeolites/water interfaces are able to catalyze efficiently alcohol dehydration[171, 172] and Friedel-Crafts alkylation[173-175], leading to high-performance industrial production of heavier substrates. The group of Lercher at the Technical University of Munich - TUM (Germany) and Pacific Northwest National Laboratory - PNNL (USA) has pioneered the field shedding light on the fundamental experimental details of these reactions.

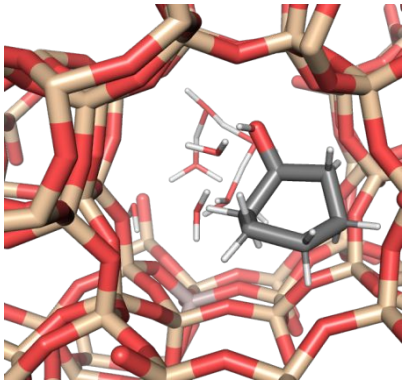

Figure 13. A cyclohexanol/water solution confined in zeolite H-ZSM-5.

From a computational point of view, these alcohol conversion systems represent a further challenge as the reaction takes place at the solid/liquid interface between the zeolite and the alcohol aqueous solution.[176-178] In this case, the catalytic site is not a well-defined Brønsted site but a dynamic molecular interface where the acidic proton is delocalized in the solution confined in the

pores of the zeolite (see Figure 13). Despite the intense experimental characterization of these processes, a detailed and comprehensive mechanistic understanding at the molecular level is still lacking.

First, one needs to characterize the mobility of the Brønsted proton in the water cluster confined in the tight hydrophobic zeolite walls. In this regard, Grifoni et al. have recently published a comprehensive study correlating two fundamental aspects, namely zeolite framework structure and water loading.[179] They investigated the acidic deprotonation and diffusion process of four different zeolites of increasing pore size in the presence of 1 up to 8 adsorbed water molecules around the catalytic site using full ab initio periodic MetaD simulations. The results of the calculations provide evidence that at low water loadings, the standard free energy of the formed complexes is dominated by enthalpy and is associated with the acid strength of the Brønsted acidic site and the space around it. Conversely, the entropy increases linearly with the concentration of water in the pores and favors proton solvation which is independent of the pore size/shape.

One of the first insights into the dehydration mechanism of ethanol/water mixtures in zeolite has been proposed by Gounder et al. where they used MetaD to explore the mechanism of the reaction.[180] At lower water pressure, the inhibition effects of water leads the formation of transition states around the periphery of H_2O clusters. At higher water pressure, the extended hydrogen-bonded networks are presented within the water-ethanol cluster, which is accompanied by strengthening water inhibition effects. Such interesting dynamic phenomena were captured due to the combined works of AIMD and MetaD simulations. Fois et al. also used Blue-Moon sampling to study Ti-based zeolites. They identified kinetically vs. thermodynamically favorable structures resulting from the interactions of water and hydrogen peroxide with the metal center in Ti-based zeolites and explained their relevance to the catalytic properties of zeolites [181]. Recently, Hack et

al. found that the excess charge is transferred from the zeolite lattice to highly coordinated water molecules within the water cluster using the combined approach of 2D IR spectroscopy and AIMD simulation[182]. This study provided valuable understandings on the dynamic natures of H-bond switching and reorientation within water clusters under the tight confinement environment of zeolites.

Metal-Organic Frameworks (MOFs)

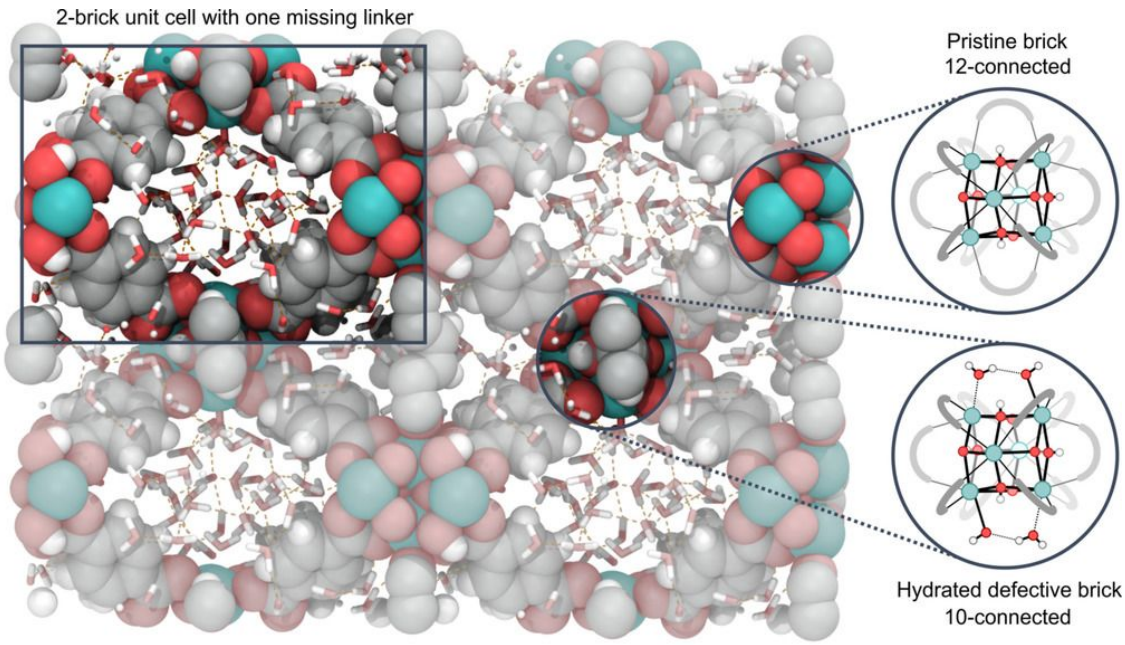

Figure 14. Representation of the defective 2-brick unit cell of UiO-66 used in the simulations, with solvent in the pores of the material. The unit cell is composed by one 12-connected pristine brick and one 10-connected hydrated defective brick, which are displayed on the right. Reprinted from Ref. [183] under a creative commons license: https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode).

MOFs are novel materials with remarkable porosity and surface area. This makes them excellent candidates for separations and catalysis [184-188]. Nevertheless, their complex hybrid structure of metal clusters connected by organic ligands makes them difficult to model due to the rich

chemistry of inorganic-organic interactions. In this respect, recent studies have used AIMD with enhanced sampling methods to investigate catalytic processes involving MOFs.

More precisely, Caratelli et al. studied UiO-66 (see Figure 14) and postulated that Lewis acid sites can be formed due to ligands temporarily detaching after interacting with protic solvents [183]. This dynamic behavior is of particular interest as these sites can be the active sites in catalytic processes [183]. They used coordination number CVs in order to study this process using AIMD with metadynamics. Heshmat investigated CO_2 hydrogenation catalyzed by a frustrated Lewis pair bonded to UiO-66. In this work, the use of metadynamics provides new insights into the reaction mechanism. More precisely, a previously undiscovered stepwise mechanism that is kinetically similar to the concerted mechanism suggested by static DFT calculations is identified using biased AIMD simulations [189].

Haigis et al. used AIMD simulation with metadynamics bias acting on a bond distance CV in order to evaluate the free energy landscape of interactions between MIL-53(Ga) and water [190]. The organic linker detaches from the metal center in the presence of water following the cleavage of a gallium – carboxylic oxygen bond. Also, water inside MOF channels lowers the free energy barrier substantially, thus facilitating the process [190]. Consequently, this work explains how framework stability is deteriorated in presence of water and suggests ways to overcome this issue[191, 192].

Defects are also known to affect their mechanical stability; hence mechanistic studies are needed to understand their formation and regulate their emergence during MOF synthesis. Large system sizes and long-time scales are needed to study the thermodynamics and kinetics of growth and understand defect formation. Recently, we investigated the role of solvent and spectator ions in the formation of defects based on microsecond-long metadynamics simulations relying on

classical potentials.[192] In this context, we used AIMD to simulate the formation of building units in order to derive the classical potential (see Figure 15).[191, 192]

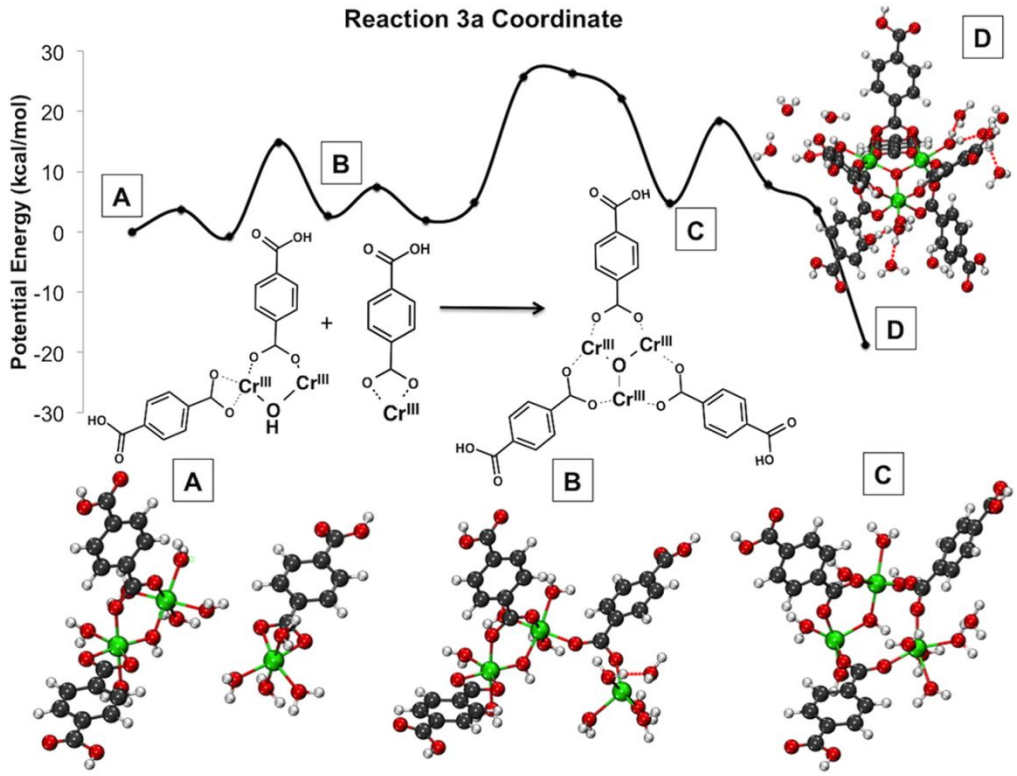

Figure 15. Reaction 3 coordinates and energy profile. In (D) the central bridging oxygen joining the three chromiums is formed, whereas in (C) an OH group joins two chromium atoms. Not all water molecules shown explicitly for clarity. Green is Cr(III), red is O, gray is C, and white is H. All Cr atoms are Cr(III). Reprinted with permission from Ref. [191]. Copyright 2014 American Chemical Society.

Although metadynamics is a powerful tool that can improve sampling during AIMD simulations in the field of catalysis, other enhanced sampling methods have been successfully employed in this context. For instance, Hajek et al. used AIMD simulations with umbrella sampling in order to investigate structural changes during activation processes. They focused on UiO-66 and biased the coordination number between the Zr atoms of the metal center and the bridging hydroxyl oxygen. Through these findings, they explained how linkers can be much more mobile than metal centers as they detach and reattach to hydroxyl groups inside the MOF. This

results in a dynamic acidity, observed inside the MOF, that can have an important role in catalysis[193]. Ming et al. used thermodynamic integration to estimate the free energy barrier of water insertion in MOF-5 [194]. Above a threshold of adsorbed water molecules in the MOF, these facilitate the insertion of more water molecules as the free energy barrier for this process is rather low due to the cleavage of Zn – O bonds. Demuynck et al. used both classical and ab initio MD as well as various enhanced sampling methods to study the MIL-53(Al) breathing behavior [195]. They highlighted the necessity for AIMD simulations due to dispersion corrections and anharmonic effects contributing to the free energy profile relevant to the breathing process. In a later study, the same group used time-structure based independent component analysis (tICA) on information gained from replica exchange (RE) simulations, namely the tICA-RE protocol [196]. This helps the identification of slow modes in the replica exchange trajectories that can be biased using enhanced sampling methods. This way they could formulate CVs which characterize breathing of MOFs. This analysis showed that internal dihedral angles are better CVs than volume as they reveal a stepwise breathing mechanism and identify metastable states in this process.

Enhanced sampling is ordinarily used to study catalytic processes, but there are also unbiased AIMD simulations of MOFs in catalysis. For example, Vandichel et al. studied the stability of UiO-66 in presence of water that is important for its use in catalytic processes [197]. In this work, they explain how the presence of linker defects facilitates dehydroxylation and water removal (when defects occur in neighboring linkers for the latter) and how defects affect mechanical stability. Zhang et al. assessed the catalytic properties of TMOF-1. They studied the energetics of CO_2 diffusion using AIMD simulations. They postulate that CO_2 can easily diffuse through the channels of this renewable heterogeneous catalyst as the free energy landscape of CO_2 diffusion is rather flat [198]. Bellarosa et al. investigated structural changes within IRMOF-1 using Born-

Oppenheimer MD [199]. Zn-IRMOF-1 degrades more easily than Mg-IRMOF-1 as the latter maintains its structure due to the strong coordination between Mg and the oxygen of water. More elaborately, this strong interaction hinders additional water molecules approaching the same metal center [199].

Chen et al. used AIMD simulations in conjunction with experiments to investigate the effect of temperature and CO_2 adsorption on MIL-53(Sc) breathing, hence characterizing the change in pore geometry caused by external stimuli [200]. At last, Xue et al. used AIMD simulations to investigate interactions between HKUST-1 and water [201]. This MOF has attracted interest in heterogeneous catalysis, but its low stability in humid environments limits its use in applications; hence interactions between the MOF and water need further investigation. Simulations revealed that strong vibrations are induced by water molecules in the pores of HKUST-1. Also, framework stability is affected by temperature and water concentration [201].

Heterogeneous catalysis at solid/liquid interfaces

Sitting at an opposite side of nanoporous catalytic materials like zeolites or MOFs we find catalytic surfaces. Here, a much larger variety of materials in terms of composition, structure, and physico-chemical properties is available, ranging from insulating or semiconducting metal oxides to metallic systems, from clean surfaces in vacuum to multifaceted nanoparticles in solution. This great experimental versatility, however, often comes with a great disadvantage in terms of computational modelling. First of all, in most cases such surfaces are not characterized by well-defined catalytic sites, whereas the active part of the material is delocalized all over the surface or situated ambiguously on structural defects of the surface like vacancies or color centers up to kinks or terraces. In such a situation modeling the reaction may be rather cumbersome and could take a lot of work for screening the possible reactive sites. Furthermore, nanoparticles or large catalytic

surfaces including catalytic defect sites often constitute several thousands of atoms which hampers the application of ab initio methods for sampling long trajectories. In such cases, one often needs to reduce the complexity of the problem using a slab surface model, where the first surface layer and a number of sublayers sufficient to simulate the bulk properties of the material are used to mimic the real material. However, even under such reduced complexity conditions, the model may require up to hundreds of atoms, often including transition or post-transition metals of the *d* or *f* block.

Despite modelling challenges, AIMD simulations have been employed to study structural and dynamic properties as well as adsorption/desorption of organic molecules at the solid-liquid interface. Gaigeot et al. employed DFT-based MD simulations to study properties of water orientation and hydrogen bonding at solid oxide/liquid water interfaces.[202] They showed that different chemical properties such as acidity of oxide surfaces differentiate water organization at the interfaces. Cheng and Sprik investigated a compact electric double layer at the rutile TiO_2-water interface by adding protons to bridging oxygens or removing oxygens from H_2O molecules adsorbed on terminal Ti cation sites from DFT-based MD simulations[203]. Yuk et al. investigated the hydrogenation of benzaldehyde on Pt-group metals in aqueous phase using AIMD simulations, where they show that the site competition between H and benzaldehyde is dependent on the metal surface.[204] They also observe the strong effect of water on the hydrogenation.

Several researchers successfully used enhanced sampling techniques to simulate reactivity at catalytic surfaces. Among many examples, oxide surfaces have been largely studied for their application in sustainable energy conversion catalysis. Different crystalline phases of titanium oxide have attracted a lot of attention as they can play a relevant role in the photocatalytic water splitting[205] and CO_2 reduction[206]. Recently, Andrade et al. (See Figure 16) used US in combination

with ab initio derived neural network potentials to investigate the proton transfer process of bulk liquid water and surface hydroxylation of TiO_2 anatase (101).[207]

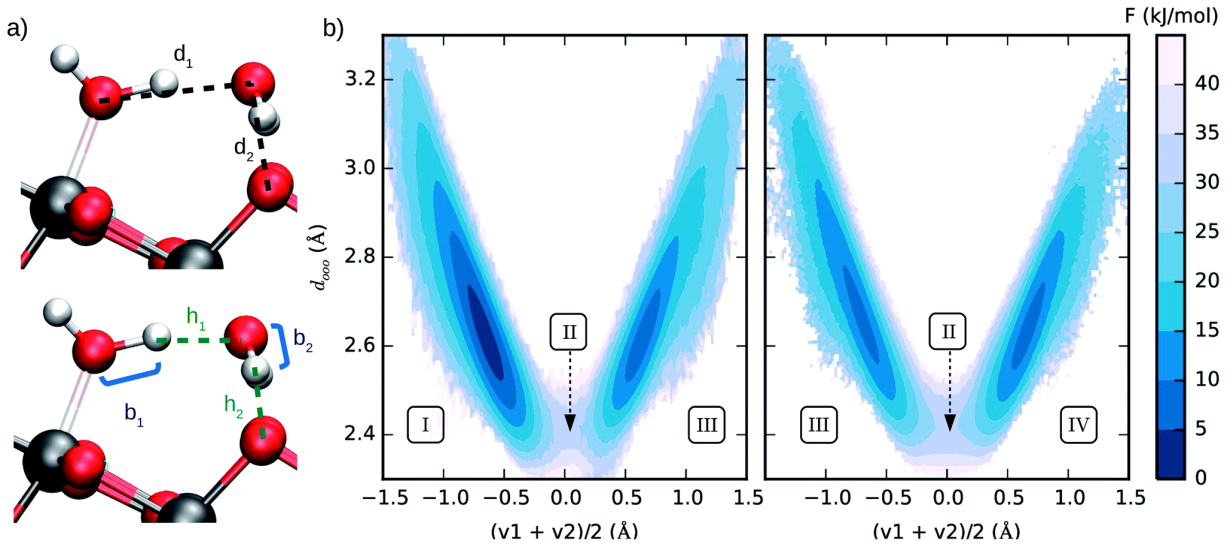

Figure 16. (a) Interatomic distances used to define the two collective variables (CVs) describing proton transfer reactions. The first CV, $d_{ooo} = (d_1 + d_2)/2$, represents the average distance between hydrogen-bonded oxygen atoms connecting proton-donor and proton-acceptor sites. The second CV, $(v_1 + v_2)/2$ ($v_i = b_i - h_i$), describes the progress of the proton transfer reaction, with positive and negative values corresponding to the dissociated and the molecular states, respectively. (b) Free energy surface of water dissociation (left) and proton transport (right) at the anatase (101)–water interface. Roman numerals indicate the adsorption state of water at specific regions in the free energy plot. Molecular water is more stable than the dissociated state by 8.0 ± 0.9 kJ mol−1, with a free energy barrier of 32 ± 4 kJ mol−1 separating these states. Reproduced from Ref. [207] (under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence) with permission from the Royal Society of Chemistry.

Often the complexity of the molecular motions at solid/liquid interfaces may lead to very slow converge of the unbiased AIMD sampling due the large number of degrees of freedom relevant to the solvent rearrangements. This is often practically impossible from a simulation point of view given the prohibitive computational demand. In such cases, it is mandatory to enhance the sampling including solvent degrees of freedom. In another work, Li et al. studied the dissociation of water on defective rutile TiO_2 (110) surfaces using ab initio MetaD simulations.[208] Zinc oxide can also be used to reduce and convert CO into more profitable products like methanol. Marx et al. used ab initio MetaD to investigate the mechanism and reaction pathways of this process over

the ZnO (000$\underline{1}$) surface.[209] Michel et al. used a combination of experimental observations and ab initio metadynamics to understand the stability of reactive alumina surfaces in water.[210] Fabris et al. used a combination of ab initio MetaD and US to investigate the hydroxylation and reduction of ceria surfaces.[211] Lee et al. studied water film formation, cation mobilization and carbonation at sc-CO_2/anorthite interfaces by combining AIMD simulations and Blue-Moon ensemble simulations. They showed that a water monolayer can be formed even at the low water concentrations at the interface. Furthermore, the presence of water monolayer accelerates the formation of cation vacancies followed by carbonation reaction at electron-rich terminal oxygen sites[212].

Catalysis at metallic surfaces is often more demanding from a computational point of view due to the fact that the small band gap restricts the use of linear scaling DFT methods and necessitates the use of either small cells or short simulation times □ both factors that limit the reliability of the results. Nonetheless, examples of the application of enhanced sampling techniques can also be found in this field. Yang et al. studied the adsorption of hydrogen at water/Pt(111) interface through a combination of AIMD/Blue-Moon ensemble simulations and kinetics.[213] They discovered that H_2 adsorption is strongly suppressed in the presence of the liquid as a result of both enthalpic (change in work function of Pt(111)) and entropic (restructuring of the interface) drivers. Sun and Jiang reported an AIMD study of the $CH_2 \rightleftharpoons CH + H$ reaction on Ni(111)[214] using integrated tempering[215]. Okumura et al. studied the reactive adsorption of hydrogen on a Pt surface using MetaD on a neural network derived potential. Very recently, de Pablo et al. reported the investigation of the free energy Landscape for nitrogen dissociation on Ru(0001)[216] using a neural-network-based enhanced sampling method called combined-force frequency (CFF)[217], a ML gradient-based method.

5. Conclusions

The combination of ab initio molecular dynamics with enhanced sampling methods provides the computational catalysis community with a powerful and flexible tool to study reactivity in complex and structured reaction environments from first principles. Through these molecular simulations, one can examine catalysts in a holistic manner accounting for both the active site and its environment. This contribution to the understanding of the atomistic and molecular aspects of catalysis enables us to move beyond simple models (which are designed for rapid screening of materials) to models that better reflect the innate complexity of catalysts under operating conditions, including elevated temperatures and a mixture of products, reactants, and inhibiters/co-catalysts. In this respect, AIMD and other statistical methods are slowly becoming vital to complementing experimental evidence, refining our understanding of reactivity in complex environments, and ultimately using this knowledge to design better catalytic materials/processes. In parallel, the entrance of machine learning and data science techniques in the mainstream of molecular simulations is boosting the applicability of modeling techniques to a wide variety of problems and improving in silico materials design. It is the hope of the authors that this review will help researchers in the computational catalysis community in orienting themselves in this vast field as well as guiding experimental scientists in understanding the immense potential that these techniques possess to improve process understanding with atomistic detail.

Acknowledgments

GMP, MSL, MTN, LK, and VAG acknowledge support from the US Department of Energy (DOE), Office of Science, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, and Biosciences. SFY, DZ, GC, and RR were supported by funding from US-DOE, Office of

Energy Efficiency and Renewable Energy, Biotechnology Office's Consortium for Computationally Physics and Chemistry. Pacific Northwest National Laboratory is operated by Battelle for DOE under Contract DE-AC05-76RL01830.

References

1. Grajciar, L.; Heard, C. J.; Bondarenko, A. A.; Polynski, M. V.; Meeprasert, J.; Pidko, E. A.; Nachtigall, P., Towards operando computational modeling in heterogeneous catalysis. *Chemical Society Reviews* **2018,** *47* (22), 8307-8348.
2. Collinge, G.; Yuk, S. F.; Nguyen, M.-T.; Lee, M.-S.; Glezakou, V.-A.; Rousseau, R., Effect of Collective Dynamics and Anharmonicity on Entropy in Heterogenous Catalysis: Building the Case for Advanced Molecular Simulations. *ACS Catalysis* **2020,** *10* (16), 9236-9260.
3. Stamatakis, M.; Vlachos, D. G., Unraveling the Complexity of Catalytic Reactions via Kinetic Monte Carlo Simulation: Current Status and Frontiers. *ACS Catalysis* **2012,** *2* (12), 2648-2663.
4. Jørgensen, M.; Grönbeck, H., Perspectives on Computational Catalysis for Metal Nanoparticles. *ACS Catalysis* **2019,** *9* (10), 8872-8881.
5. Nørskov, J. K.; Bligaard, T.; Rossmeisl, J.; Christensen, C. H., Towards the computational design of solid catalysts. *Nature Chemistry* **2009,** *1* (1), 37-46.
6. Chen, B. W. J.; Xu, L.; Mavrikakis, M., Computational Methods in Heterogeneous Catalysis. *Chemical Reviews* **2021,** *121* (2), 1007-1048.
7. *Computational Catalysis*. Royal Society of Chemistry: Cambridge, 2013.
8. Alavi, A.; Hu, P.; Deutsch, T.; Silvestrelli, P. L.; Hutter, J., CO Oxidation on Pt(111): An Ab Initio Density Functional Theory Study. *Physical Review Letters* **1998,** *80* (16), 3650-3653.
9. Saeys, M.; Reyniers, M.-F.; Marin, G. B.; Neurock, M., Density Functional Study of Benzene Adsorption on Pt(111). *The Journal of Physical Chemistry B* **2002,** *106* (30), 7489-7498.
10. Liu, Z.-P.; Hu, P.; Alavi, A., Catalytic Role of Gold in Gold-Based Catalysts: A Density Functional Theory Study on the CO Oxidation on Gold. *Journal of the American Chemical Society* **2002,** *124* (49), 14770-14779.
11. Digne, M.; Sautet, P.; Raybaud, P.; Euzen, P.; Toulhoat, H., Use of DFT to Achieve a Rational Understanding of Acid–basic Properties of γ-Alumina Surfaces. *Journal of Catalysis* **2004,** *226* (1), 54-68.
12. Hansgen, D. A.; Vlachos, D. G.; Chen, J. G., Using First Principles to Predict Bimetallic Catalysts for the Ammonia Decomposition Reaction. *Nature Chemistry* **2010,** *2* (6), 484-489.
13. Tao, F.; Dag, S.; Wang, L.-W.; Liu, Z.; Butcher, D. R.; Bluhm, H.; Salmeron, M.; Somorjai, G. A., Break-Up of Stepped Platinum Catalyst Surfaces by High CO Coverage. *Science* **2010,** *327* (5967), 850.
14. Grabow, L. C.; Mavrikakis, M., Mechanism of Methanol Synthesis on Cu through CO2 and CO Hydrogenation. *Acs Catalysis* **2011,** *1* (4), 365-384.
15. Behrens, M.; Studt, F.; Kasatkin, I.; Kühl, S.; Hävecker, M.; Abild-Pedersen, F.; Zander, S.; Girgsdies, F.; Kurr, P.; Kniep, B.-L.; Tovar, M.; Fischer, R. W.; Nørskov, J. K.; Schlögl, R., The Active Site of Methanol Synthesis over Cu/ZnO/Al₂O₃ Industrial Catalysts. *Science* **2012,** *336* (6083), 893.
16. McFarland, E. W.; Metiu, H., Catalysis by Doped Oxides. *Chemical Reviews* **2013,** *113* (6), 4391-4427.
17. Thiel, W., Computational Catalysis—Past, Present, and Future. *Angewandte Chemie International Edition* **2014,** *53* (33), 8605-8613.
18. Saavedra, J.; Doan, H. A.; Pursell, C. J.; Grabow, L. C.; Chandler, B. D., The Critical Role of Water at the Gold-titania Interface in Catalytic CO Oxidation. *Science* **2014,** *345* (6204), 1599.
19. Kulkarni, A.; Siahrostami, S.; Patel, A.; Nørskov, J. K., Understanding Catalytic Activity Trends in the Oxygen Reduction Reaction. *Chemical Reviews* **2018,** *118* (5), 2302-2312.

20. Frenkel, D.; Smit, B., *Understanding Molecular Simulation: From Algorithms to Applications*. Elsevier Science: 2001.
21. Liu, X.; Wen, X.; Hoffmann, R., Surface Activation of Transition Metal Nanoparticles for Heterogeneous Catalysis: What We Can Learn from Molecular Dynamics. *ACS Catalysis* **2018,** *8* (4), 3365-3375.
22. Boero, M.; Parrinello, M.; Terakura, K., First Principles Molecular Dynamics Study of Ziegler−Natta Heterogeneous Catalysis. *Journal of the American Chemical Society* **1998,** *120* (12), 2746-2752.
23. Termath, V.; Haase, F.; Sauer, J.; Hutter, J.; Parrinello, M., Understanding the Nature of Water Bound to Solid Acid Surfaces. Ab Initio Simulation on HSAPO-34. *Journal of the American Chemical Society* **1998,** *120* (33), 8512-8516.
24. van Speybroeck, V.; Meier, R. J., A recent development in computational chemistry: chemical reactions from first principles molecular dynamics simulations. *Chemical Society Reviews* **2003,** *32* (3), 151-157.
25. Alexopoulos, K.; Lee, M.-S.; Liu, Y.; Zhi, Y.; Liu, Y.; Reyniers, M.-F. o.; Marin, G. B.; Glezakou, V.-A.; Rousseau, R.; Lercher, J. A., Anharmonicity and confinement in zeolites: structure, spectroscopy, and adsorption free energy of ethanol in H-ZSM-5. *Journal of Physical Chemistry C* **2016,** *120* (13), 7172-7182.
26. Yuk, S. F.; Lee, M.-S.; Collinge, G.; Zhang, J.; Padmaperuma, A. B.; Li, Z.; Polo-Garzon, F.; Wu, Z.; Glezakou, V.-A.; Rousseau, R., Mechanistic Understanding of Catalytic Conversion of Ethanol to 1-Butene over 2D-Pillared MFI Zeolite. *The Journal of Physical Chemistry C* **2020,** *124* (52), 28437-28447.
27. Wang, Y.-G.; Mei, D.; Glezakou, V.-A.; Li, J.; Rousseau, R., Dynamic formation of single-atom catalytic active sites on ceria-supported gold nanoparticles. *Nature Communications* **2015,** *6* (1), 6511.
28. Wang, Y.-G.; Yoon, Y.; Glezakou, V.-A.; Li, J.; Rousseau, R., The Role of Reducible Oxide–Metal Cluster Charge Transfer in Catalytic Processes: New Insights on the Catalytic Mechanism of CO Oxidation on Au/TiO_2 from ab Initio Molecular Dynamics. *Journal of the American Chemical Society* **2013,** *135* (29), 10673-10683.
29. Xu, C.-Q.; Lee, M.-S.; Wang, Y.-G.; Cantu, D. C.; Li, J.; Glezakou, V.-A.; Rousseau, R., Structural Rearrangement of Au–Pd Nanoparticles under Reaction Conditions: An ab Initio Molecular Dynamics Study. *ACS Nano* **2017,** *11* (2), 1649-1658.
30. Wang, Y.-G.; Cantu, D. C.; Lee, M.-S.; Li, J.; Glezakou, V.-A.; Rousseau, R., CO Oxidation on Au/TiO2: Condition-Dependent Active Sites and Mechanistic Pathways. *Journal of the American Chemical Society* **2016,** *138* (33), 10467-10476.
31. Hollingsworth, S. A.; Dror, R. O., Molecular Dynamics Simulation for All. *Neuron* **2018,** *99* (6), 1129-1143.
32. Harrison, J. A.; Schall, J. D.; Maskey, S.; Mikulski, P. T.; Knippenberg, M. T.; Morrow, B. H., Review of force fields and intermolecular potentials used in atomistic computational materials research. *Applied Physics Reviews* **2018,** *5* (3), 031104.
33. Marx, D. In *Ab initio molecular dynamics: Theory and Implementation*, 2000.
34. Tuckerman, M. E., Ab initio molecular dynamics: basic concepts, current trends and novel applications. *Journal of Physics: Condensed Matter* **2002,** *14* (50), R1297-R1355.
35. Car, R.; Parrinello, M., Unified Approach for Molecular Dynamics and Density-Functional Theory. *Physical Review Letters* **1985,** *55* (22), 2471-2474.
36. Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulos, J. D., Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. *Reviews of Modern Physics* **1992,** *64* (4), 1045-1097.
37. Born, M.; Oppenheimer, R., Zur Quantentheorie der Molekeln. *Annalen der Physik* **1927,** *389* (20), 457-484.

38. Laino, T.; Mohamed, F.; Laio, A.; Parrinello, M., An Efficient Real Space Multigrid QM/MM Electrostatic Coupling. *Journal of Chemical Theory and Computation* **2005,** *1* (6), 1176-1184.
39. Laio, A.; VandeVondele, J.; Rothlisberger, U., A Hamiltonian electrostatic coupling scheme for hybrid Car–Parrinello molecular dynamics simulations. *The Journal of Chemical Physics* **2002,** *116* (16), 6941-6947.
40. Woo, T. K.; Margl, P. M.; Blöchl, P. E.; Ziegler, T., A Combined Car−Parrinello QM/MM Implementation for ab Initio Molecular Dynamics Simulations of Extended Systems: Application to Transition Metal Catalysis. *The Journal of Physical Chemistry B* **1997,** *101* (40), 7877-7880.
41. Lu, D.; Wang, H.; Chen, M.; Lin, L.; Car, R.; E, W.; Jia, W.; Zhang, L., 86 PFLOPS Deep Potential Molecular Dynamics simulation of 100 million atoms with ab initio accuracy. *Computer Physics Communications* **2021,** *259*, 107624.
42. Ufimtsev, I. S.; Martinez, T. J., Quantum Chemistry on Graphical Processing Units. 3. Analytical Energy Gradients, Geometry Optimization, and First Principles Molecular Dynamics. *Journal of Chemical Theory and Computation* **2009,** *5* (10), 2619-2628.
43. Voter, A. F.; Montalenti, F.; Germann, T. C., Extending the Time Scale in Atomistic Simulation of Materials. *Annual Review of Materials Research* **2002,** *32* (1), 321-346.
44. Valsson, O.; Tiwary, P.; Parrinello, M., Enhancing Important Fluctuations: Rare Events and Metadynamics from a Conceptual Viewpoint. *Annual Review of Physical Chemistry* **2016,** *67* (1), 159-184.
45. Peters, B., Chapter 1 - Introduction. In *Reaction Rate Theory and Rare Events Simulations*, Peters, B., Ed. Elsevier: Amsterdam, 2017; pp 1-17.
46. Torrie, G. M.; Valleau, J. P., Nonphysical Sampling Distributions in Monte Carlo Free-energy Estimation: Umbrella Sampling. *Journal of Computational Physics* **1977,** *23* (2), 187-199.
47. Earl, D. J.; Deem, M. W., Parallel tempering: Theory, applications, and new perspectives. *Physical Chemistry Chemical Physics* **2005,** *7* (23), 3910-3916.
48. Swendsen, R. H.; Wang, J.-S., Replica Monte Carlo Simulation of Spin-Glasses. *Physical Review Letters* **1986,** *57* (21), 2607-2609.
49. Voter, A. F., Hyperdynamics: Accelerated Molecular Dynamics of Infrequent Events. *Physical Review Letters* **1997,** *78* (20), 3908-3911.
50. Dellago, C.; Bolhuis, P. G.; Geissler, P. L., Transition Path Sampling. In *Advances in Chemical Physics*, 2002; pp 1-78.
51. Laio, A.; Parrinello, M., Escaping free-energy minima. *Proceedings of the National Academy of Sciences* **2002,** *99* (20), 12562.
52. Abrams, C.; Bussi, G., Enhanced Sampling in Molecular Dynamics Using Metadynamics, Replica-Exchange, and Temperature-Acceleration. *Entropy* **2014,** *16* (1).
53. Jonsson, H.; Mills, G.; Jacobsen, K. W., Nudged elastic band method for finding minimum energy paths of transitions. In *Classical and Quantum Dynamics in Condensed Phase Simulations*, WORLD SCIENTIFIC: 1998; pp 385-404.
54. Chen, C.; Zuo, Y.; Ye, W.; Li, X.; Deng, Z.; Ong, S. P., A Critical Review of Machine Learning of Energy Materials. *Advanced Energy Materials* **2020,** *10* (8), 1903242.
55. Schleder, G. R.; Padilha, A. C. M.; Acosta, C. M.; Costa, M.; Fazzio, A., From DFT to machine learning: recent approaches to materials science–a review. *Journal of Physics: Materials* **2019,** *2* (3), 032001.
56. Li, Y.; Li, H.; Pickard, F. C.; Narayanan, B.; Sen, F. G.; Chan, M. K. Y.; Sankaranarayanan, S. K. R. S.; Brooks, B. R.; Roux, B., Machine Learning Force Field Parameters from Ab Initio Data. *Journal of Chemical Theory and Computation* **2017,** *13* (9), 4492-4503.
57. Serra, J. M.; Corma, A.; Chica, A.; Argente, E.; Botti, V., Can artificial neural networks help the experimentation in catalysis? *Catalysis Today* **2003,** *81* (3), 393-403.

58. Yang, W.; Fidelis, T. T.; Sun, W.-H., Machine Learning in Catalysis, From Proposal to Practicing. *ACS Omega* **2020,** *5* (1), 83-88.
59. Jawad, J.; Hawari, A. H.; Javaid Zaidi, S., Artificial neural network modeling of wastewater treatment and desalination using membrane processes: A review. *Chemical Engineering Journal* **2021,** *419*, 129540.
60. Kitchin, J. R., Machine learning in catalysis. *Nature Catalysis* **2018,** *1* (4), 230-232.
61. Li, Z.; Wang, S.; Xin, H., Toward artificial intelligence in catalysis. *Nature Catalysis* **2018,** *1* (9), 641-642.
62. Schlexer Lamoureux, P.; Winther, K. T.; Garrido Torres, J. A.; Streibel, V.; Zhao, M.; Bajdich, M.; Abild-Pedersen, F.; Bligaard, T., Machine Learning for Computational Heterogeneous Catalysis. *ChemCatChem* **2019,** *11* (16), 3581-3601.
63. Ma, S.; Liu, Z.-P., Machine Learning for Atomic Simulation and Activity Prediction in Heterogeneous Catalysis: Current Status and Future. *ACS Catalysis* **2020,** *10* (22), 13213-13226.
64. Toyao, T.; Maeno, Z.; Takakusagi, S.; Kamachi, T.; Takigawa, I.; Shimizu, K.-i., Machine Learning for Catalysis Informatics: Recent Applications and Prospects. *ACS Catalysis* **2020,** *10* (3), 2260-2297.
65. Lee, M.-S.; McGrail, B. P.; Glezakou, V.-A., Microstructural Response of Variably Hydrated Ca-rich Montmorillonite to Supercritical CO2. *Environmental Science & Technology* **2014,** *48* (15), 8612-8619.
66. Praveen, C. S.; Comas-Vives, A.; Copéret, C.; VandeVondele, J., Role of Water, CO_2, and Noninnocent Ligands in the CO_2 Hydrogenation to Formate by an Ir(III) PNP Pincer Catalyst Evaluated by Static-DFT and ab Initio Molecular Dynamics under Reaction Conditions. *Organometallics* **2017,** *36* (24), 4908-4919.
67. Rousseau, R.; Schenter, G. K.; Fulton, J. L.; Linehan, J. C.; Engelhard, M. H.; Autrey, T., Defining Active Catalyst Structure and Reaction Pathways from ab Initio Molecular Dynamics and Operando XAFS: Dehydrogenation of Dimethylaminoborane by Rhodium Clusters. *Journal of the American Chemical Society* **2009,** *131* (30), 10516-10524.
68. Vidossich, P.; Lledós, A.; Ujaque, G., First-Principles Molecular Dynamics Studies of Organometallic Complexes and Homogeneous Catalytic Processes. *Accounts of Chemical Research* **2016,** *49* (6), 1271-1278.
69. O'Hagan, M.; Shaw, W. J.; Raugei, S.; Chen, S.; Yang, J. Y.; Kilgore, U. J.; DuBois, D. L.; Bullock, R. M., Moving Protons with Pendant Amines: Proton Mobility in a Nickel Catalyst for Oxidation of Hydrogen. *Journal of the American Chemical Society* **2011,** *133* (36), 14301-14312.
70. Woo, T. K.; Margl, P. M.; Ziegler, T.; Blöchl, P. E., Static and Ab Initio Molecular Dynamics Study of the Titanium(IV)-Constrained Geometry Catalyst ($CpSiH_2NH$)Ti-R^+. 2. Chain Termination and Long Chain Branching. *Organometallics* **1997,** *16* (15), 3454-3468.
71. Barducci, A.; Bonomi, M.; Parrinello, M., Metadynamics. *WIREs Computational Molecular Science* **2011,** *1* (5), 826-843.
72. Tiwary, P.; Parrinello, M., From Metadynamics to Dynamics. *Physical Review Letters* **2013,** *111* (23), 230602.
73. Tiwary, P.; Parrinello, M., A Time-Independent Free Energy Estimator for Metadynamics. *The Journal of Physical Chemistry B* **2015,** *119* (3), 736-742.
74. Ciccotti, G.; Ferrario, M., Blue Moon Approach to Rare Events. *Molecular Simulation* **2004,** *30* (11-12), 787-793.
75. Carter, E. A.; Ciccotti, G.; Hynes, J. T.; Kapral, R., Constrained reaction coordinate dynamics for the simulation of rare events. *Chemical Physics Letters* **1989,** *156* (5), 472-477.
76. Sprik, M.; Ciccotti, G., Free energy from constrained molecular dynamics. *The Journal of Chemical Physics* **1998,** *109* (18), 7737-7744.

77. De Sousa, R. L.; Leite Alves, H. W., Ab Initio Calculation of the Dynamical Properties of PPP and PPV. *Brazilian Journal of Physics* **2005,** *36*, 501-504.
78. Sediki, A.; Snoek, L. C.; Gaigeot, M.-P., N–H+ vibrational anharmonicities directly revealed from DFT-based molecular dynamics simulations on the Ala7H+ protonated peptide. *International Journal of Mass Spectrometry* **2011,** *308* (2), 281-288.
79. Piccini, G.; Sauer, J., Effect of Anharmonicity on Adsorption Thermodynamics. *Journal of Chemical Theory and Computation* **2014,** *10* (6), 2479-2487.
80. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*. John Wiley & Sons: 2005.
81. Campbell, C. T.; Sprowl, L. H.; Arnadottir, L., Equilibrium Constants and Rate Constants for Adsorbates: Two-Dimensional (2D) Ideal Gas, 2D Ideal Lattice Gas, and Ideal Hindered Translator Models. *Journal of Physical Chemistry C* **2016,** *120* (19), 10283-10297.
82. Jorgensen, M.; Gronbeck, H., Adsorbate Entropies with Complete Potential Energy Sampling in Microkinetic Modeling. *Journal of Physical Chemistry C* **2017,** *121* (13), 7199-7207.
83. Njegic, B.; Gordon, M. S., Predicting accurate vibrational frequencies for highly anharmonic systems. *The Journal of Chemical Physics* **2008,** *129* (16), 164107.
84. Campbell, C. T.; Sellers, J. R. V., The Entropies of Adsorbed Molecules. *Journal of the American Chemical Society* **2012,** *134* (43), 18109-18115.
85. Piccini, G.; Alessio, M.; Sauer, J., Ab initio study of methanol and ethanol adsorption on Brønsted sites in zeolite H-MFI. *Physical Chemistry Chemical Physics* **2018,** *20* (30), 19964-19970.
86. van Duin, A. C. T.; Strachan, A.; Stewman, S.; Zhang, Q.; Xu, X.; Goddard, W. A., ReaxFFSiO Reactive Force Field for Silicon and Silicon Oxide Systems. *The Journal of Physical Chemistry A* **2003,** *107* (19), 3803-3811.
87. Chenoweth, K.; van Duin, A. C. T.; Goddard Iii, W. A., The ReaxFF Monte Carlo Reactive Dynamics Method for Predicting Atomistic Structures of Disordered Ceramics: Application to the Mo3VOx Catalyst. *Angewandte Chemie International Edition* **2009,** *48* (41), 7630-7634.
88. Furman, D.; Wales, D. J., Transforming the Accuracy and Numerical Stability of ReaxFF Reactive Force Fields. *The Journal of Physical Chemistry Letters* **2019,** *10* (22), 7215-7223.
89. Senftle, T. P.; Hong, S.; Islam, M. M.; Kylasa, S. B.; Zheng, Y.; Shin, Y. K.; Junkermeier, C.; Engel-Herbert, R.; Janik, M. J.; Aktulga, H. M.; Verstraelen, T.; Grama, A.; van Duin, A. C. T., The ReaxFF Reactive Force-Field: Development, Applications and Future Directions. *npj Computational Materials* **2016,** *2* (1), 15011.
90. Han, Y.; Jiang, D.; Zhang, J.; Li, W.; Gan, Z.; Gu, J., Development, applications and challenges of ReaxFF reactive force field in molecular simulations. *Frontiers of Chemical Science and Engineering* **2016,** *10* (1), 16-38.
91. Woodcock, H. L.; Hodošček, M.; Brooks, B. R., Exploring SCC-DFTB Paths for Mapping QM/MM Reaction Mechanisms. *The Journal of Physical Chemistry A* **2007,** *111* (26), 5720-5728.
92. Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G., Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Physical Review B* **1998,** *58* (11), 7260-7268.
93. Verlet, L., Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Physical Review* **1967,** *159* (1), 98-103.
94. Hockney, R. W., The potential calculation and some applications. In *Methods in Computational Physics* Academic Press: New York, 1970; pp 135-211.
95. Costentin, C.; Robert, M.; Savéant, J.-M., Electrochemical concerted proton and electron transfers. Potential-dependent rate constant, reorganization factors, proton tunneling and isotope effects. *Journal of Electroanalytical Chemistry* **2006,** *588* (2), 197-206.

96. Fermann, J. T.; Auerbach, S., Modeling proton mobility in acidic zeolite clusters: II. Room temperature tunneling effects from semiclassical rate theory. *The Journal of Chemical Physics* **2000,** *112* (15), 6787-6794.
97. Hammes-Schiffer, S., Hydrogen Tunneling and Protein Motion in Enzyme Reactions. *Accounts of Chemical Research* **2006,** *39* (2), 93-100.
98. Skylaris, C.-K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C., Introducing ONETEP: Linear-scaling density functional simulations on parallel computers. *The Journal of Chemical Physics* **2005,** *122* (8), 084119.
99. Blum, V.; Gehrke, R.; Hanke, F.; Havu, P.; Havu, V.; Ren, X.; Reuter, K.; Scheffler, M., Ab initio molecular simulations with numeric atom-centered orbitals. *Computer Physics Communications* **2009,** *180* (11), 2175-2196.
100. VandeVondele, J.; Borštnik, U.; Hutter, J., Linear Scaling Self-Consistent Field Calculations with Millions of Atoms in the Condensed Phase. *Journal of Chemical Theory and Computation* **2012,** *8* (10), 3565-3573.
101. VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J., Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Computer Physics Communications* **2005,** *167* (2), 103-128.
102. Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. I. J.; Refson, K.; Payne, M. C., First principles methods using CASTEP. *Zeitschrift für Kristallographie - Crystalline Materials* **2005,** *220* (5-6), 567-570.
103. Kresse, G.; Hafner, J., Ab initio molecular dynamics for liquid metals. *Physical Review B* **1993,** *47* (1), 558-561.
104. Remler, D. K.; Madden, P. A., Molecular dynamics without effective potentials via the Car-Parrinello approach. *Molecular Physics* **1990,** *70* (6), 921-966.
105. Hutter, J., Car–Parrinello molecular dynamics. *WIREs Computational Molecular Science* **2012,** *2* (4), 604-612.
106. CPMD *http://www.cpmd.org/*.
107. Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; Dal Corso, A.; de Gironcoli, S.; Fabris, S.; Fratesi, G.; Gebauer, R.; Gerstmann, U.; Gougoussis, C.; Kokalj, A.; Lazzeri, M.; Martin-Samos, L.; Marzari, N.; Mauri, F.; Mazzarello, R.; Paolini, S.; Pasquarello, A.; Paulatto, L.; Sbraccia, C.; Scandolo, S.; Sclauzero, G.; Seitsonen, A. P.; Smogunov, A.; Umari, P.; Wentzcovitch, R. M., QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *Journal of Physics: Condensed Matter* **2009,** *21* (39), 395502.
108. Tangney, P., On the theory underlying the Car-Parrinello method and the role of the fictitious mass parameter. *The Journal of Chemical Physics* **2006,** *124* (4), 044111.
109. Kühne, T. D.; Krack, M.; Mohamed, F. R.; Parrinello, M., Efficient and Accurate Car-Parrinello-like Approach to Born-Oppenheimer Molecular Dynamics. *Physical Review Letters* **2007,** *98* (6), 066401.
110. Galli, G.; Parrinello, M., Ab-Initio Molecular Dynamics: Principles and Practical Implementation. In *Computer Simulation in Materials Science: Interatomic Potentials, Simulation Techniques and Applications*, Meyer, M.; Pontikis, V., Eds. Springer Netherlands: Dordrecht, 1991; pp 283-304.
111. Jorgensen, W. L.; Thomas, L. L., Perspective on Free-Energy Perturbation Calculations for Chemical Equilibria. *Journal of Chemical Theory and Computation* **2008,** *4* (6), 869-876.
112. Zwanzig, R. W., High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *The Journal of Chemical Physics* **1954,** *22* (8), 1420-1426.
113. Hwang, J. K.; King, G.; Creighton, S.; Warshel, A., Simulation of free energy relationships and dynamics of SN2 reactions in aqueous solution. *Journal of the American Chemical Society* **1988,** *110* (16), 5297-5311.

114. Fu, C. D.; Oliveira, L. F. L.; Pfaendtner, J., Assessing Generic Collective Variables for Determining Reaction Rates in Metadynamics Simulations. *Journal of Chemical Theory and Computation* **2017,** *13* (3), 968-973.
115. M. Sultan, M.; Pande, V. S., tICA-Metadynamics: Accelerating Metadynamics by Using Kinetically Selected Collective Variables. *Journal of Chemical Theory and Computation* **2017,** *13* (6), 2440-2447.
116. Mendels, D.; Piccini, G.; Parrinello, M., Collective Variables from Local Fluctuations. *The Journal of Physical Chemistry Letters* **2018,** *9* (11), 2776-2781.
117. Sultan, M. M.; Pande, V. S., Automated design of collective variables using supervised machine learning. *The Journal of Chemical Physics* **2018,** *149* (9), 094106.
118. Piccini, G.; Mendels, D.; Parrinello, M., Metadynamics with Discriminants: A Tool for Understanding Chemistry. *Journal of Chemical Theory and Computation* **2018,** *14* (10), 5040-5044.
119. Bonati, L.; Rizzi, V.; Parrinello, M., Data-Driven Collective Variables for Enhanced Sampling. *The Journal of Physical Chemistry Letters* **2020,** *11* (8), 2998-3004.
120. Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E., How Fast-Folding Proteins Fold. *Science* **2011,** *334* (6055), 517.
121. McCarty, J.; Parrinello, M., A variational conformational dynamics approach to the selection of collective variables in metadynamics. *The Journal of Chemical Physics* **2017,** *147* (20), 204109.
122. Piccini, G.; Polino, D.; Parrinello, M., Identifying Slow Molecular Motions in Complex Chemical Reactions. *The Journal of Physical Chemistry Letters* **2017,** *8* (17), 4197-4200.
123. Branduardi, D.; Gervasio, F. L.; Parrinello, M., From A to B in free energy space. *The Journal of Chemical Physics* **2007,** *126* (5), 054103.
124. Henkelman, G.; Uberuaga, B. P.; Jónsson, H., A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *The Journal of Chemical Physics* **2000,** *113* (22), 9901-9904.
125. Pietrucci, F.; Saitta, A. M., Formamide Reaction Network in Gas Phase and Solution via a Unified Theoretical Approach: Toward a Reconciliation of Different Prebiotic Scenarios. *Proceedings of the National Academy of Sciences* **2015,** *112* (49), 15030.
126. Pietrucci, F., Novel Enhanced Sampling Strategies for Transitions Between Ordered and Disordered Structures. In *Handbook of Materials Modeling: Methods: Theory and Modeling*, Andreoni, W.; Yip, S., Eds. Springer International Publishing: Cham, 2020; pp 597-619.
127. Pietrucci, F.; Andreoni, W., Fate of a Graphene Flake: A New Route toward Fullerenes Disclosed with Ab Initio Simulations. *Journal of Chemical Theory and Computation* **2014,** *10* (3), 913-917.
128. Frenkel, D.; Smit, B., Chapter 1 - Introduction. In *Understanding Molecular Simulation (Second Edition)*, Frenkel, D.; Smit, B., Eds. Academic Press: San Diego, 2002; pp 1-6.
129. Kirkwood, J. G., Statistical Mechanics of Fluid Mixtures. *The Journal of Chemical Physics* **1935,** *3* (5), 300-313.
130. Darve, E.; Pohorille, A., Calculating free energies using average force. *The Journal of Chemical Physics* **2001,** *115* (20), 9169-9183.
131. Comer, J.; Gumbart, J. C.; Hénin, J.; Lelièvre, T.; Pohorille, A.; Chipot, C., The Adaptive Biasing Force Method: Everything You Always Wanted To Know but Were Afraid To Ask. *The Journal of Physical Chemistry B* **2015,** *119* (3), 1129-1151.
132. Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A., THE weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry* **1992,** *13* (8), 1011-1021.
133. Barducci, A.; Bussi, G.; Parrinello, M., Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method. *Physical Review Letters* **2008,** *100* (2), 020603.

134. Huber, T.; Torda, A. E.; van Gunsteren, W. F., Local elevation: A method for improving the searching properties of molecular dynamics simulation. *Journal of Computer-Aided Molecular Design* **1994,** *8* (6), 695-708.
135. Fleming, K. L.; Tiwary, P.; Pfaendtner, J., New Approach for Investigating Reaction Dynamics and Rates with Ab Initio Calculations. *The Journal of Physical Chemistry A* **2016,** *120* (2), 299-305.
136. Xie, Y.; Calabrese Barton, S., Infrequent metadynamics study of rare-event electrostatic channeling. *Physical Chemistry Chemical Physics* **2021,** *23* (23), 13381-13388.
137. Noé, F.; Tkatchenko, A.; Müller, K.-R.; Clementi, C., Machine Learning for Molecular Simulation. *Annual Review of Physical Chemistry* **2020,** *71* (1), 361-390.
138. Lorenz, S.; Groß, A.; Scheffler, M., Representing high-dimensional potential-energy surfaces for reactions at surfaces by neural networks. *Chemical Physics Letters* **2004,** *395* (4), 210-215.
139. Lorenz, S. Reactions on Surfaces with Neural Networks. Doctoral Thesis, Technische Universität Berlin, Berlin, 2001.
140. Lorenz, S.; Scheffler, M.; Gross, A., Descriptions of surface chemical reactions using a neural network representation of the potential-energy surface. *Physical Review B* **2006,** *73* (11), 115431.
141. Lorenz, S.; Scheffler, M.; Gross, A., Surface physics, nanoscale physics, low-dimensional systems-Descriptions of surface chemical reactions using a neural network representation of the potential-energy surface. *Physical Review-Section B-Condensed Matter* **2006,** *73* (11), 115431-115431.
142. Ang, S. J.; Wang, W.; Schwalbe-Koda, D.; Axelrod, S.; Gómez-Bombarelli, R., Active learning accelerates ab initio molecular dynamics on reactive energy surfaces. *Chem* **2021,** *7* (3), 738-751.
143. Jinnouchi, R.; Karsai, F.; Kresse, G., Making free-energy calculations routine: Combining first principles with machine learning. *Physical Review B* **2020,** *101* (6), 060201.
144. Botu, V.; Ramprasad, R., Adaptive machine learning framework to accelerate ab initio molecular dynamics. *International Journal of Quantum Chemistry* **2015,** *115* (16), 1074-1083.
145. Chmiela, S.; Tkatchenko, A.; Sauceda, H. E.; Poltavsky, I.; Schütt, K. T.; Müller, K.-R., Machine learning of accurate energy-conserving molecular force fields. *Science Advances* **2017,** *3* (5), e1603015.
146. Gao, T.; Kitchin, J. R., Modeling Palladium Surfaces with Density Functional Theory, Neural Networks and Molecular Dynamics. *Catalysis Today* **2018,** *312*, 132-140.
147. Shakouri, K.; Behler, J.; Meyer, J.; Kroes, G.-J., Accurate Neural Network Description of Surface Phonons in Reactive Gas–Surface Dynamics: N2 + Ru(0001). *The Journal of Physical Chemistry Letters* **2017,** *8* (10), 2131-2136.
148. Jia, W.; Wang, H.; Chen, M.; Lu, D.; Lin, L.; Car, R.; Weinan, E.; Zhang, L. In *Pushing the Limit of Molecular Dynamics with Ab Initio Accuracy to 100 Million Atoms with Machine Learning*, SC20: International Conference for High Performance Computing, Networking, Storage and Analysis, 9-19 Nov. 2020; 2020; pp 1-14.
149. Artrith, N.; Kolpak, A. M., Grand canonical molecular dynamics simulations of Cu–Au nanoalloys in thermal equilibrium using reactive ANN potentials. *Computational Materials Science* **2015,** *110*, 20-28.
150. Artrith, N.; Kolpak, A. M., Understanding the Composition and Activity of Electrocatalytic Nanoalloys in Aqueous Solvents: A Combination of DFT and Accurate Neural Network Potentials. *Nano Letters* **2014,** *14* (5), 2670-2676.
151. Häse, F.; Galván, I. F.; Aspuru-Guzik, A.; Lindh, R.; Vacher, M., Machine learning for analysing ab initio molecular dynamics simulations. *Journal of Physics: Conference Series* **2020,** *1412*, 042003.

152. Häse, F.; Fdez. Galván, I.; Aspuru-Guzik, A.; Lindh, R.; Vacher, M., How machine learning can assist the interpretation of ab initio molecular dynamics simulations and conceptual understanding of chemistry. *Chemical Science* **2019,** *10* (8), 2298-2307.
153. Vermeiren, W.; Gilson, J. P., Impact of Zeolites on the Petroleum and Petrochemical Industry. *Topics in Catalysis* **2009,** *52* (9), 1131-1161.
154. Li, Y.; Li, L.; Yu, J., Applications of Zeolites in Sustainable Chemistry. *Chem* **2017,** *3* (6), 928-949.
155. Lima, C. G. S.; Moreira, N. M.; Paixão, M. W.; Corrêa, A. G., Heterogenous green catalysis: Application of zeolites on multicomponent reactions. *Current Opinion in Green and Sustainable Chemistry* **2019,** *15*, 7-12.
156. Ennaert, T.; Van Aelst, J.; Dijkmans, J.; De Clercq, R.; Schutyser, W.; Dusselier, M.; Verboekend, D.; Sels, B. F., Potential and challenges of zeolite chemistry in the catalytic conversion of biomass. *Chemical Society Reviews* **2016,** *45* (3), 584-611.
157. Serrano, D. P.; Melero, J. A.; Morales, G.; Iglesias, J.; Pizarro, P., Progress in the design of zeolite catalysts for biomass conversion into biofuels and bio-based chemicals. *Catalysis Reviews* **2018,** *60* (1), 1-70.
158. Perego, C.; Bosetti, A.; Ricci, M.; Millini, R., Zeolite Materials for Biomass Conversion to Biofuel. *Energy & Fuels* **2017,** *31* (8), 7721-7733.
159. Bailleul, S.; Dedecker, K.; Cnudde, P.; Vanduyfhuys, L.; Waroquier, M.; Van Speybroeck, V., Ab initio enhanced sampling kinetic study on MTO ethene methylation reaction. *Journal of Catalysis* **2020,** *388*, 38-51.
160. Moors, S. L. C.; De Wispelaere, K.; Van der Mynsbrugge, J.; Waroquier, M.; Van Speybroeck, V., Molecular Dynamics Kinetic Study on the Zeolite-Catalyzed Benzene Methylation in ZSM-5. *ACS Catalysis* **2013,** *3* (11), 2556-2567.
161. Van der Mynsbrugge, J.; Visur, M.; Olsbye, U.; Beato, P.; Bjørgen, M.; Van Speybroeck, V.; Svelle, S., Methylation of benzene by methanol: Single-site kinetics over H-ZSM-5 and H-beta zeolite catalysts. *Journal of Catalysis* **2012,** *292*, 201-212.
162. De Wispelaere, K.; Ensing, B.; Ghysels, A.; Meijer, E. J.; Van Speybroeck, V., Complex Reaction Environments and Competing Reaction Mechanisms in Zeolite Catalysis: Insights from Advanced Molecular Dynamics. *Chemistry – A European Journal* **2015,** *21* (26), 9385-9396.
163. De Wispelaere, K.; Bailleul, S.; Van Speybroeck, V., Towards molecular control of elementary reactions in zeolite catalysis by advanced molecular simulations mimicking operating conditions. *Catalysis Science & Technology* **2016,** *6* (8), 2686-2705.
164. Valsson, O.; Parrinello, M., Variational Approach to Enhanced Sampling and Free Energy Calculations. *Physical Review Letters* **2014,** *113* (9), 090601.
165. Rey, J.; Raybaud, P.; Chizallet, C.; Bučko, T., Competition of Secondary versus Tertiary Carbenium Routes for the Type B Isomerization of Alkenes over Acid Zeolites Quantified by Ab Initio Molecular Dynamics Simulations. *ACS Catalysis* **2019,** *9* (11), 9813-9828.
166. Rey, J.; Gomez, A.; Raybaud, P.; Chizallet, C.; Bučko, T., On the origin of the difference between type A and type B skeletal isomerization of alkenes catalyzed by zeolites: The crucial input of ab initio molecular dynamics. *Journal of Catalysis* **2019,** *373*, 361-373.
167. Rey, J.; Bignaud, C.; Raybaud, P.; Bučko, T.; Chizallet, C., Dynamic Features of Transition States for β-Scission Reactions of Alkenes over Acid Zeolites Revealed by AIMD Simulations. *Angewandte Chemie International Edition* **2020,** *59* (43), 18938-18942.
168. Cnudde, P.; De Wispelaere, K.; Van der Mynsbrugge, J.; Waroquier, M.; Van Speybroeck, V., Effect of temperature and branching on the nature and stability of alkene cracking intermediates in H-ZSM-5. *Journal of Catalysis* **2017,** *345*, 53-69.
169. Isikgor, F. H.; Becer, C. R., Lignocellulosic biomass: a sustainable platform for the production of bio-based chemicals and polymers. *Polymer Chemistry* **2015,** *6* (25), 4497-4559.

170. Fatma, S.; Hameed, A.; Noman, M.; Ahmed, T.; Shahid, M.; Tariq, M.; Sohail, I.; Tabassum, R., Lignocellulosic Biomass: A Sustainable Bioenergy Source for the Future. *Protein Pept Lett* **2018,** *25* (2), 148-163.
171. Liu, Y.; Vjunov, A.; Shi, H.; Eckstein, S.; Camaioni, D. M.; Mei, D.; Baráth, E.; Lercher, J. A., Enhancing the catalytic activity of hydronium ions through constrained environments. *Nature Communications* **2017,** *8* (1), 14113.
172. Hintermeier, P. H.; Eckstein, S.; Mei, D.; Olarte, M. V.; Camaioni, D. M.; Baráth, E.; Lercher, J. A., Hydronium-Ion-Catalyzed Elimination Pathways of Substituted Cyclohexanols in Zeolite H-ZSM5. *ACS Catalysis* **2017,** *7* (11), 7822-7829.
173. Liu, Y.; Baráth, E.; Shi, H.; Hu, J.; Camaioni, D. M.; Lercher, J. A., Solvent-determined mechanistic pathways in zeolite-H-BEA-catalysed phenol alkylation. *Nature Catalysis* **2018,** *1* (2), 141-147.
174. Zhao, C.; Camaioni, D. M.; Lercher, J. A., Selective catalytic hydroalkylation and deoxygenation of substituted phenols to bicycloalkanes. *Journal of Catalysis* **2012,** *288*, 92-103.
175. Zhao, C.; Lercher, J. A., Selective Hydrodeoxygenation of Lignin-Derived Phenolic Monomers and Dimers to Cycloalkanes on Pd/C and HZSM-5 Catalysts. *ChemCatChem* **2012,** *4* (1), 64-68.
176. Pfriem, N.; Hintermeier, P. H.; Eckstein, S.; Kim, S.; Liu, Q.; Shi, H.; Milakovic, L.; Liu, Y.; Haller, G. L.; Baráth, E.; Liu, Y.; Lercher, J. A., Role of the ionic environment in enhancing the activity of reacting molecules in zeolite pores. *Science* **2021,** *372* (6545), 952.
177. Wang, M.; Jaegers, N. R.; Lee, M.-S.; Wan, C.; Hu, J. Z.; Shi, H.; Mei, D.; Burton, S. D.; Camaioni, D. M.; Gutiérrez, O. Y.; Glezakou, V.-A.; Rousseau, R.; Wang, Y.; Lercher, J. A., Genesis and Stability of Hydronium Ions in Zeolite Channels. *Journal of the American Chemical Society* **2019,** *141* (8), 3444-3455.
178. Zhi, Y.; Shi, H.; Mu, L.; Liu, Y.; Mei, D.; Camaioni, D. M.; Lercher, J. A., Dehydration Pathways of 1-Propanol on HZSM-5 in the Presence and Absence of Water. *Journal of the American Chemical Society* **2015,** *137* (50), 15781-15794.
179. Grifoni, E.; Piccini, G.; Lercher, J. A.; Glezakou, V.-A.; Rousseau, R.; Parrinello, M., Confinement effects and acid strength in zeolites. *Nature Communications* **2021,** *12* (1), 2630.
180. Bates, J. S.; Bukowski, B. C.; Greeley, J.; Gounder, R., Structure and solvation of confined water and water–ethanol clusters within microporous Brønsted acids and their effects on ethanol dehydration catalysis. *Chemical Science* **2020,** *11* (27), 7102-7122.
181. Fois, E.; Gamba, A.; Spanó, E., Competition between Water and Hydrogen Peroxide at Ti Center in Titanium Zeolites. An ab Initio Study. *The Journal of Physical Chemistry B* **2004,** *108* (28), 9557-9560.
182. Hack, J. H.; Dombrowski, J. P.; Ma, X.; Chen, Y.; Lewis, N. H. C.; Carpenter, W. B.; Li, C.; Voth, G. A.; Kung, H. H.; Tokmakoff, A., Structural Characterization of Protonated Water Clusters Confined in HZSM-5 Zeolites. *Journal of the American Chemical Society* **2021**.
183. Caratelli, C.; Hajek, J.; Meijer, E. J.; Waroquier, M.; Van Speybroeck, V., Dynamic Interplay between Defective UiO-66 and Protic Solvents in Activated Processes. *Chemistry – A European Journal* **2019,** *25* (67), 15315-15325.
184. Pascanu, V.; González Miera, G.; Inge, A. K.; Martín-Matute, B., Metal–Organic Frameworks as Catalysts for Organic Synthesis: A Critical Perspective. *Journal of the American Chemical Society* **2019,** *141* (18), 7223-7234.
185. Alhumaimess, M. S., Metal–organic frameworks and their catalytic applications. *Journal of Saudi Chemical Society* **2020,** *24* (6), 461-473.
186. Wang, Q.; Astruc, D., State of the Art and Prospects in Metal–Organic Framework (MOF)-Based and MOF-Derived Nanocatalysis. *Chemical Reviews* **2019,** *120* (2), 1438-1511.
187. Huang, Y.-B.; Liang, J.; Wang, X.-S.; Cao, R., Multifunctional metal–organic framework catalysts: synergistic catalysis and tandem reactions. *Chemical Society Reviews* **2017,** *46* (1), 126-157.

188. Ali, M.; Pervaiz, E.; Noor, T.; Rabi, O.; Zahra, R.; Yang, M., Recent advancements in MOF-based catalysts for applications in electrochemical and photoelectrochemical water splitting: A review. *International Journal of Energy Research* **2021,** *45* (2), 1190-1226.
189. Heshmat, M., Alternative Pathway of CO2 Hydrogenation by Lewis-Pair-Functionalized UiO-66 MOF Revealed by Metadynamics Simulations. *The Journal of Physical Chemistry C* **2020,** *124* (20), 10951-10960.
190. Haigis, V.; Coudert, F.-X.; Vuilleumier, R.; Boutin, A.; Fuchs, A. H., Hydrothermal Breakdown of Flexible Metal–Organic Frameworks: A Study by First-Principles Molecular Dynamics. *The Journal of Physical Chemistry Letters* **2015,** *6* (21), 4365-4370.
191. Cantu, D. C.; McGrail, B. P.; Glezakou, V.-A., Formation Mechanism of the Secondary Building Unit in a Chromium Terephthalate Metal–Organic Framework. *Chemistry of Materials* **2014,** *26* (22), 6401-6409.
192. Kollias, L.; Cantu, D. C.; Tubbs, M. A.; Rousseau, R.; Glezakou, V.-A.; Salvalaglio, M., Molecular Level Understanding of the Free Energy Landscape in Early Stages of Metal–Organic Framework Nucleation. *Journal of the American Chemical Society* **2019,** *141* (14), 6073-6081.
193. Hajek, J.; Caratelli, C.; Demuynck, R.; De Wispelaere, K.; Vanduyfhuys, L.; Waroquier, M.; Van Speybroeck, V., On the intrinsic dynamic nature of the rigid UiO-66 metal–organic framework. *Chemical Science* **2018,** *9* (10), 2723-2732.
194. Ming, Y.; Kumar, N.; Siegel, D. J., Water Adsorption and Insertion in MOF-5. *ACS Omega* **2017,** *2* (8), 4921-4928.
195. Demuynck, R.; Rogge, S. M. J.; Vanduyfhuys, L.; Wieme, J.; Waroquier, M.; Van Speybroeck, V., Efficient Construction of Free Energy Profiles of Breathing Metal–Organic Frameworks Using Advanced Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation* **2017,** *13* (12), 5861-5873.
196. Demuynck, R.; Wieme, J.; Rogge, S. M. J.; Dedecker, K. D.; Vanduyfhuys, L.; Waroquier, M.; Van Speybroeck, V., Protocol for Identifying Accurate Collective Variables in Enhanced Molecular Dynamics Simulations for the Description of Structural Transformations in Flexible Metal–Organic Frameworks. *Journal of Chemical Theory and Computation* **2018,** *14* (11), 5511-5526.
197. Vandichel, M.; Hajek, J.; Ghysels, A.; De Vos, A.; Waroquier, M.; Van Speybroeck, V., Water coordination and dehydration processes in defective UiO-66 type metal organic frameworks. *CrystEngComm* **2016,** *18* (37), 7056-7069.
198. Zhang, G.; Wei, G.; Liu, Z.; Oliver, S. R. J.; Fei, H., A Robust Sulfonate-Based Metal–Organic Framework with Permanent Porosity for Efficient CO2 Capture and Conversion. *Chemistry of Materials* **2016,** *28* (17), 6276-6281.
199. Bellarosa, L.; Calero, S.; López, N., Early stages in the degradation of metal–organic frameworks in liquid water from first-principles molecular dynamics. *Physical Chemistry Chemical Physics* **2012,** *14* (20).
200. Chen, L.; Mowat, J. P. S.; Fairen-Jimenez, D.; Morrison, C. A.; Thompson, S. P.; Wright, P. A.; Düren, T., Elucidating the Breathing of the Metal–Organic Framework MIL-53(Sc) with ab Initio Molecular Dynamics Simulations and in Situ X-ray Powder Diffraction Experiments. *Journal of the American Chemical Society* **2013,** *135* (42), 15763-15773.
201. Xue, W.; Zhang, Z.; Huang, H.; Zhong, C.; Mei, D., Theoretical Insights into the Initial Hydrolytic Breakdown of HKUST-1. *The Journal of Physical Chemistry C* **2019,** *124* (3), 1991-2001.
202. Gaigeot, M.-P.; Sprik, M.; Sulpizi, M., Oxide/water interfaces: how the surface chemistry modifies interfacial water properties. *Journal of Physics: Condensed Matter* **2012,** *24* (12), 124106.
203. Cheng, J.; Sprik, M., The electric double layer at a rutile TiO2water interface modelled using density functional theory based molecular dynamics simulation. *Journal of Physics: Condensed Matter* **2014,** *26* (24), 244108.

204. Yuk, S. F.; Lee, M.-S.; Akhade, S. A.; Nguyen, M.-T.; Glezakou, V.-A.; Rousseau, R., First-principle investigation on catalytic hydrogenation of benzaldehyde over Pt-group metals. *Catalysis Today* **2020**.
205. Fujishima, A.; Honda, K., Electrochemical Photolysis of Water at a Semiconductor Electrode. *Nature* **1972,** *238* (5358), 37-38.
206. White, J. L.; Baruch, M. F.; Pander, J. E.; Hu, Y.; Fortmeyer, I. C.; Park, J. E.; Zhang, T.; Liao, K.; Gu, J.; Yan, Y.; Shaw, T. W.; Abelev, E.; Bocarsly, A. B., Light-Driven Heterogeneous Reduction of Carbon Dioxide: Photocatalysts and Photoelectrodes. *Chemical Reviews* **2015,** *115* (23), 12888-12935.
207. Calegari Andrade, M. F.; Ko, H.-Y.; Zhang, L.; Car, R.; Selloni, A., Free energy of proton transfer at the water–TiO2 interface from ab initio deep potential molecular dynamics. *Chemical Science* **2020,** *11* (9), 2335-2341.
208. Wang, H.-L.; Hu, Z.-P.; Li, H., Dissociation of liquid water on defective rutile TiO2 (110) surfaces using ab initio molecular dynamics simulations. *Frontiers of Physics* **2018,** *13* (3), 138107.
209. Kiss, J.; Frenzel, J.; Nair, N. N.; Meyer, B.; Marx, D., Methanol synthesis on ZnO(0001¯). III. Free energy landscapes, reaction pathways, and mechanistic insights. *The Journal of Chemical Physics* **2011,** *134* (6), 064710.
210. Réocreux, R.; Girel, É.; Clabaut, P.; Tuel, A.; Besson, M.; Chaumonnot, A.; Cabiac, A.; Sautet, P.; Michel, C., Reactivity of shape-controlled crystals and metadynamics simulations locate the weak spots of alumina in water. *Nature Communications* **2019,** *10* (1), 3139.
211. Negreiros, F. R.; Camellone, M. F.; Fabris, S., Effects of Thermal Fluctuations on the Hydroxylation and Reduction of Ceria Surfaces by Molecular H2. *The Journal of Physical Chemistry C* **2015,** *119* (37), 21567-21573.
212. Lee, M.-S.; Peter McGrail, B.; Rousseau, R.; Glezakou, V.-A., Structure, dynamics and stability of water/scCO2/mineral interfaces from ab initio molecular dynamics simulations. *Scientific Reports* **2015,** *5* (1), 14857.
213. Yang, G.; Akhade, S. A.; Chen, X.; Liu, Y.; Lee, M.-S.; Glezakou, V.-A.; Rousseau, R.; Lercher, J. A., The Nature of Hydrogen Adsorption on Platinum in the Aqueous Phase. *Angewandte Chemie International Edition* **2019,** *58* (11), 3527-3532.
214. Sun, G.; Jiang, H., Ab initio molecular dynamics with enhanced sampling for surface reaction kinetics at finite temperatures: CH2⇌ CH + H on Ni(111) as a case study. *The Journal of Chemical Physics* **2015,** *143* (23), 234706.
215. Gao, Y. Q., An integrate-over-temperature approach for enhanced sampling. *The Journal of Chemical Physics* **2008,** *128* (6), 064105.
216. Lee, E. M. Y.; Ludwig, T.; Yu, B.; Singh, A. R.; Gygi, F.; Nørskov, J. K.; de Pablo, J. J., Neural Network Sampling of the Free Energy Landscape for Nitrogen Dissociation on Ruthenium. *The Journal of Physical Chemistry Letters* **2021,** *12* (11), 2954-2962.
217. Sevgen, E.; Guo, A. Z.; Sidky, H.; Whitmer, J. K.; de Pablo, J. J., Combined Force-Frequency Sampling for Simulation of Systems Having Rugged Free Energy Landscapes. *Journal of Chemical Theory and Computation* **2020,** *16* (3), 1448-1455.