iScience

CellPress
OPEN ACCESS

Review

Iterative multiscale and multi-physics computations for operando catalyst nanostructure elucidation and kinetic modeling

Ajin Rajan,[1,2] Anoop P. Pushkar,[1,2] Balaji C. Dharmalingam,[1] and Jithin John Varghese[1,*]

SUMMARY

Modern heterogeneous catalysis has benefitted immensely from computational predictions of catalyst structure and its evolution under reaction conditions, first-principles mechanistic investigations, and detailed kinetic modeling, which are rungs on a multiscale workflow. Establishing connections across these rungs and integration with experiments have been challenging. Here, operando catalyst structure prediction techniques using density functional theory simulations and *ab initio* thermodynamics calculations, molecular dynamics, and machine learning techniques are presented. Surface structure characterization by computational spectroscopic and machine learning techniques is then discussed. Hierarchical approaches in kinetic parameter estimation involving semi-empirical, data-driven, and first-principles calculations and detailed kinetic modeling via mean-field microkinetic modeling and kinetic Monte Carlo simulations are discussed along with methods and the need for uncertainty quantification. With these as the background, this article proposes a bottom-up hierarchical and closed loop modeling framework incorporating consistency checks and iterative refinements at each level and across levels.

INTRODUCTION

As heterogeneous catalysis involves spatiotemporal scales spanning the active site (typically an atom to a few atoms) to the catalytic reactor which is typically a packed bed of the catalyst pellets (up to a few meters), modeling and analysis of heterogeneously catalyzed phenomena requires a truly multiscale approach.[1,2] With typical measurables such as reactant conversion, product selectivity, and reaction rates intricately linked to the catalyst nanostructure, understanding, optimizing, and designing highly efficient catalytic reaction systems require fundamental insights on the interconnected phenomena across the scales.[3] Well-established techniques are available for modeling and analysis of catalytic phenomena at specific spatiotemporal scales, while integration across the scales continues to remain a challenge.[4] Hence, the state-of-the-art approaches in modeling heterogeneously catalyzed reactions are hierarchical both in spatiotemporal scales of the methods adopted and in the order of accuracy and computational effort of the methods.[4] With most catalysts being far from homogeneous, with the distribution of the active sites, and many of them undergoing substantial surface structural changes at the different stages of the catalytic cycle (pre-treatment, reaction, and so on), insights on actual operando surface structure of the catalysts remain obscure.[5] This further complicates and limits our understanding of the reaction chemistry on these catalyst surfaces.

Chemically homogeneous (made of similar atomic species) systems such as metal nanoparticles undergo morphology change due to substrate/gas species adsorption. Gases like CO, O_2, and NO and species like H_2O induce morphology changes in metal nanoparticles such as those of Pt and Au and nanoalloys such as that of Pt-Pd during reactions.[6–8] These gas species are common in reduction reactions such as methanol synthesis from syngas, NO_x selective catalytic reductions, and partial oxidation reactions with molecular oxygen, to mention a few. On the other hand, other catalyst systems such as metal oxides and metal carbides and so on are chemically heterogeneous and undergo morphological changes due to defect formation, substrate adsorption, and active participation in surface reactions.[9–12] Hence, *in situ* or operando characterization of heterogeneous catalysts is essential in understanding their complex structure and their structural dynamics due to their interactions with reactive gases under typical reaction

[1]Department of Chemical Engineering, Indian Institute of Technology Madras, Chennai, Tamil Nadu 600036, India

[2]These authors contributed equally

*Correspondence: jithinjv@iitm.ac.in

https://doi.org/10.1016/j.isci.2023.107029

conditions.[13] Operando description of catalyst morphology, surface structure, and active sites are essential to compute accurate reaction energetics and kinetics and to establish structure-activity relations. These structure-activity relations will eventually enable us to design catalysts for achieving better catalytic activity. As operando structure determination is still very challenging, computational operando catalyst modeling and design[14–16] is a new paradigm, emerging as an alternative.

Detailed kinetic modeling via a multiscale modeling framework to link the catalyst surface structure to its intrinsic surface activity and thereby experimentally observed catalytic performance is the key for design of better catalysts and engineering of catalytic reactions.[17] This is truly a challenging task and requires interdisciplinary efforts in integrating modeling techniques across the atomic to reactor scales, establishing feedback loops, and workflows between experimental and computational techniques and more recently in exploiting the potential of data-oriented approaches for both catalyst design and reaction engineering. For around two decades, density functional theory simulations have served as the basis for analysis of reaction mechanisms, estimation of kinetic parameters, and formulation of microkinetic models,[18] forming the backbone of the first-principles kinetic modeling framework that is widely used in gas-phase heterogeneous catalysis. While these kinetic models are excellent tools for qualitative comparison and screening of catalysts,[19] the quantitative reliability and parity of computed and experimental quantities have been far from the desirable limits.[20]

Catalytic processes are typically carried out at high temperatures, the dynamics of the active site ensembles, and the interacting substrates are most appropriately modeled using *ab initio* molecular dynamics simulations which enable accurate description of the underlying free energy landscape.[21] These are, however, computationally very expensive and find limited application in computational heterogeneous catalysis. First-principles-based kinetic modeling of liquid phase reaction have added complexities of the solvent interaction with the substrate and the catalyst, local transport effects over the catalyst surface, uncertainties due to dynamics in intermolecular interactions, and difficulties in energy calculations due to the need for sampling immensely large configurational space.[22,23]

Data-driven and machine-learning-based approaches are increasingly finding applications in all spheres of catalysis and catalytic reaction engineering. Machine-learning-based catalytic performance prediction is dealt with in two ways[24]: 1) a top-down approach where the training set is obtained from experiments (characterization) that describe catalyst structure such as particle size, surface area, electronic states, surface coordination, and so on and 2) a bottom-up approach where the training set is obtained from quantum chemical calculations involving typical descriptors such as energetics, band gaps, d-band centers, Bader charges, and so on.[24] The top-down approach often involves tedious data mining (literature search) exercises which are time-consuming and may lack quantitative accuracy in the predictions as it is often challenging to find high-quality training dataset for the desired reaction and catalyst compositions. On the other hand, the bottom-up approach demands considerable time and effort for building high-fidelity dataset for training the machine learning models (such as neural networks) for predicting accurate energetics, potential energy surface, and eventually the activity. A detailed review by Cova et al.[25] highlights the applications of deep-learning-based approaches for predicting reaction rates, reaction routes, molecular spectra, and so on.

A few recent review articles discuss various theoretical aspects, state-of-the-art techniques, and methodologies developed in the field of computational heterogeneous catalysis.[4,26–28] A detailed review by Chen et al.[4] discussed the theory of *first-principles*-based electronic calculations, microkinetic modeling, kinetic Monte Carlo, and *in silico* catalyst design methodologies. The review by Bruix et al.[28] addressed various aspects of multiscale modeling, including a brief overview and applications of data-driven methodologies in computational heterogeneous catalysis. Martinez et al.[26] discussed the applications of machine-learning-based techniques in catalysis and allied fields. Puliyanda et al.[27] discussed the recent applications of automated and data-driven techniques for reaction monitoring and mechanism elucidation.

This article is aimed at providing an overview of the complexities and approaches in modeling and analysis of heterogeneous surface phenomena and quantitative kinetic modeling, starting with operando catalyst structure modeling to kinetic parameter estimation and finally multiscale kinetic modeling. An iterative and hierarchical approach is proposed at each stage such as catalyst structure and active site prediction, reaction mechanism elucidation, kinetic parameter estimation, and multiscale kinetic modeling, with cross

validations with appropriate experimental data and using the current state-of-the-art approaches such as using data-driven models to build robust and accurate kinetic models capable of making accurate predictions. A detailed discussion on operando computational catalyst structure and morphology prediction, validation, and the overall iterative workflows are the highlights of the work.

The article is organized as follows. First-principles-based determination of operando catalyst morphology, surface structure, active site distribution on the relevant catalyst surfaces, and data-driven methods for catalyst screening and discovery are discussed in the section titled operando catalyst morphology and surface structure prediction. Heterogeneity in active sites on catalysts and identifying them are discussed in the section titled Active site nature and its impact on catalysis. Computational and data-driven approaches for the operando catalyst characterization and validation techniques that are conventionally used for validating the catalyst surface and active site details are described in the section titled computational and data-driven methods for operando catalyst characterization and validation. The section titled summary and iterative workflow for catalyst structure modeling summarizes the approaches in operando catalyst structure determination and validation and proposes the iterative refinement based on model and experimental inputs. The section titled reaction mechanism and network generation and validation discusses general practices of elucidation of reaction mechanisms using chemical intuition, reaction rules, data-driven techniques, and their validation for mapping of reaction networks comprising all possible elementary reaction steps using validated computational model catalysts. The section titled semi-empirical techniques describes semi-empirical techniques such as the UBI-QEP method and kinetic linear scaling relations used for estimation of the activation energy barriers for the elementary reaction steps. This is followed by a discussion on alternatives to semi-empirical techniques such as density functional theory (DFT) and data-driven methods which enable calculation of both the pre-exponential factor and the activation energy barrier. This is followed by a discussion on the uncertainties in the parameter estimation both in the DFT-based calculated energetics and parameters estimated through semi-empirical methods. We then summarize the reaction mechanism generation and kinetic parameter estimation highlighting the hierarchical approaches and decisions trees for desired accuracy across systems. The section titled mean field microkinetic modeling discusses the mean-field microkinetic modeling (MKM) technique as one of the most popular detailed kinetic modeling approaches, with the sub-sections describing the applications of the technique in different scenarios and caveats. This is followed by a discussion of the stochastic kinetic Monte Carlo (KMC) technique with typical applications, comparison with the MKM and caveats. We finally summarize, a unidirectional multiscale modeling framework and propose an iterative closed loop approach linking catalyst structure modeling with kinetic modeling, with multiple cross validations with experimental data across the scales.

OPERANDO CATALYST STRUCTURE MODELING AND VALIDATION

Operando characterization of heterogeneous catalysts requires sophisticated facilities which are still beyond the reach of most of the heterogeneous catalysis community. Hence, "Operando Catalyst Structure Determination" is a holy grail in heterogeneously catalyzed reactions. With advances in *first-principles* theoretical frameworks and high-performance parallel computing systems, the computational community has stepped up in developing tools and workflows that enable "operando computational catalyst structure predictions" with minimal inputs from experimental observables. A suite of first-principles calculations, primarily relying on DFT simulations, are now being employed in modeling and predicting the catalyst surface structure.[4,29] Modeling catalyst fluxionality and catalyst reconstruction at reaction condition may not be straightforward with DFT techniques alone and techniques like molecular dynamics (MD) and combinations with grand canonical Monte Carlo (GCMC) may be useful. Recently, DFT-aided machine learning techniques have also been employed to predict the catalyst structure under reaction environments.[29,30] The section titled operando catalyst morphology and surface structure prediction describes computational approaches widely used to study surface morphological changes and estimate the surface structure at equilibrium as a function of reactive gas environment and reaction temperatures, dynamical simulations addressing catalyst reconstruction and fluxionality changes, data-driven methods for catalyst screening, and discovery with selected examples from the literature demonstrating their application. The discussion here is primarily focused on metal, metal oxide, and metal carbide kind of catalysts and the catalyst materials like zeolites and metal-organic frameworks are not discussed in detail. Active sites of a type, where specific reaction steps happen on the catalyst surface, may be uniformly distributed on the catalyst surface, making them homogeneous surfaces. Sometimes there might be a distribution of active sites which differ in their catalytic potential, leading to active site heterogeneity, which complicates analyses of reaction

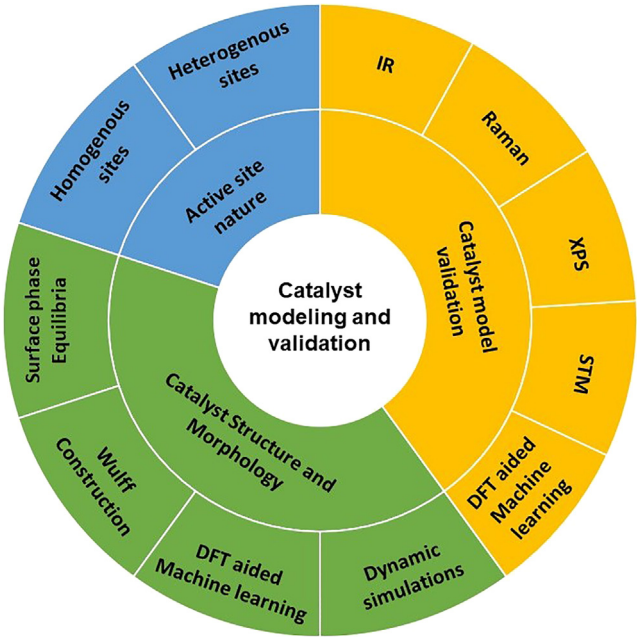

Figure 1. Conventional methods for computational operando catalyst structure modeling and validation

mechanisms, energetics, and kinetics. The section titled active site nature and its impact on catalysis briefly describes identification of active site heterogeneity and its implications. Upon predictions of the operando structural information, it is essential to validate these against experimentally reported observations. Full-fledged first-principles-based computational infrared (IR), Raman, X-ray photoelectron spectroscopy (XPS), and scanning tunneling microscopic (STM) analyses are suitable techniques for the surface structure and active site characterization and thereby validation of the active sites. These methods, along with a brief discussion on the recent advent of data-driven techniques for assisting computational and experimental operando characterization have also been discussed with selected examples from literature in the section titled Computational and Data-driven methods for operando catalyst characterization and validation. A schematic representation of all the contents discussed is given in Figure 1.

Operando catalyst morphology and surface structure prediction

The reaction conditions and chemical environments can significantly influence the catalyst morphology and the associated surface structure which in turn dictates the reaction mechanisms and energetics. A schematic representation of the implication of the catalyst morphology evolution on the surface structural changes and the reaction energetics is shown in Figure 2. The overall catalyst morphology determines the types of exposed crystal facets of the catalyst (nano)particle. The exposed crystal facets determine the specific atomic arrangement on the catalyst surface. This arrangement together with specific atomic ensembles which serve as the active sites for the various reaction steps determine the reaction mechanisms and the energetics along the possible reaction pathways. These microscopic features eventually determine the overall macroscopic observables such as species conversion, product selectivity/distribution, and reaction rates.

With selected examples from the literature, use of methods such as Wulff construction, *ab initio* thermodynamics, adsorption isotherm models, data-driven methods, MD simulations for estimation of the overall morphology of the catalyst (nano)particles, and its evolution with changes in the reaction environment/conditions are discussed in this section.

Catalyst morphology prediction

Catalyst morphological studies provide necessary information about the predominant crystal facet distributions of a catalyst during the reaction. When studied as a function of temperature and reactant pressure, the catalyst morphological changes at equilibrium are captured, and the corresponding stable facets can

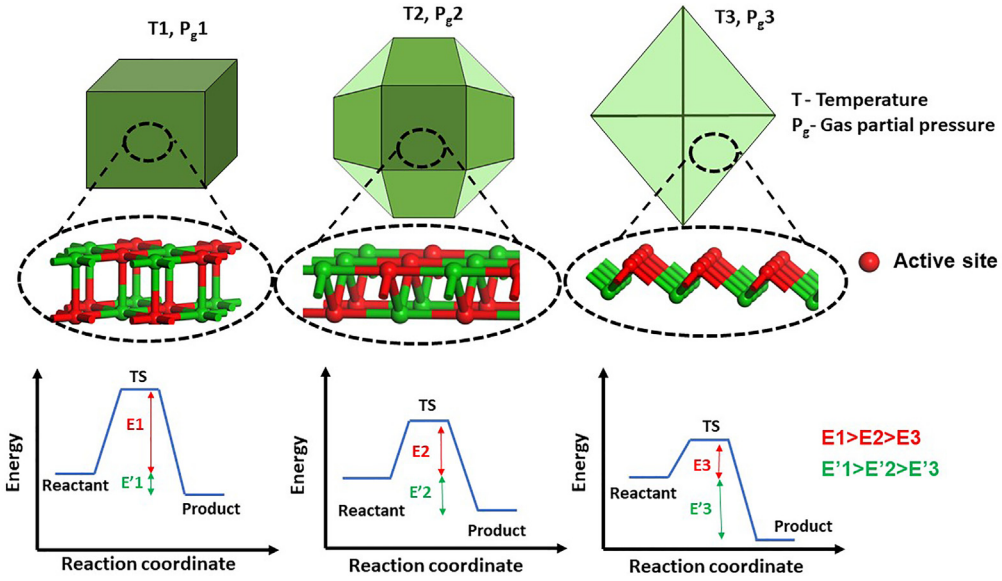

Figure 2. Schematic representation of the impact of the reaction conditions such as temperature and partial pressure of gases on the catalyst morphology, structure, and reaction energetics

An arbitrary catalyst is represented with different morphologies in the top panel. Arbitrary atomic arrangements representative of different crystal exposure is presented in the middle panel with balls of different color representing atoms of different elements. Energy profiles for an arbitrary reaction step (going from a reactant state to a product state via a transition state TS) happening on different crystal exposures are presented. Here, T_i and $P_{g,}i$ denotes reaction temperature and reactive gas partial pressure, respectively. E_i and $E_i^{'}$ refers to activation energy barriers and reaction energies. The energy comparisons are made with inclusion of effect of + and - signs.

be identified for subsequent reaction modeling. Computationally, catalyst nanoparticle shapes are predicted based on the Wulff construction,[31–34] which is a method of identifying a shape that minimizes the surface energies of different crystal facets for a fixed volume. This in turn relies on the Gibbs free energy for the formation of the different crystal surfaces per unit area (surface energy) which are calculated using *ab initio* thermodynamics techniques based on DFT energetics. For more details on the *ab initio* thermodynamics techniques, and catalyst shape and surface analyses using these techniques, refer to the ref.[14,35,36]. Many of the recent investigations employ Wulff construction to study the effects of reactant adsorption on the catalyst surface on the particle morphology evolution during reactions.[37–39] These analyses require the stable surface coverage of the reaction gas adsorbed on all the prominent low-index crystal facets of the catalyst material at the reaction conditions and this coverage can be estimated using equilibrium-based methodologies such as 1) analyses of the surface free energies of different surface coverages and generation of surface phase diagrams,[37] 2) using adsorption isotherm models,[13] and 3) from non-equilibrium-based methodologies such as from surface species concentrations from kinetic modeling such as MKM or KMC analysis.[40,41] Examples of each of these are discussed in the next few paragraphs and summarized in Table 1.

The implications of CO and H_2 environments on the equilibrium morphology evolution of cobalt nanoparticles were studied by Yu et al.[37] A stepwise CO and H (dissociated H_2) adsorption was studied over the (100), (111), (110), and (311) surfaces of Co nanoparticle. Saturated H coverage of 1 monolayer (ML) over (111), (311), and (100), and 0.78 ML for (110) surfaces and CO coverages of 0.78, 0.61, 0.78, and 0.5 ML over (111), (311), (100), and (110) surfaces were obtained from surface free energy analyses. The H_2 and CO environments increased the surface exposure of Co(311) surface from 0% (P_{H_2} = 10^{-10}atm) to 10% (P_{H_2} = 5 atm) and that of Co(110) from 0% (P_{CO} = 10^{-10}atm) to 6% (P_{CO} = 5 atm) at 675 K. These surfaces exposed substantial quantities of under-coordinated metal sites, and hence were expected to offer higher catalytic activity in terms of H_2 and CO dissociation during Fischer-Tropsch synthesis. Inoğlu et al.,[38] in their computational work, investigated the role of adsorbates on the morphological changes of Cu catalysts in a water gas shift (WGS) reaction environment. They reported that the presence of H_2S in the reaction system induced poisoning of the Cu surfaces due to strong S (from H_2S dissociation) adsorption and changed the

Table 1. Different catalyst morphology prediction techniques and corresponding references from literature

Method	Technique	Reference
Equilibrium-based methodologies	Wulff construction	[13] and [44]
	Surface phase diagrams	[37,42] and [47]
	Adsorption isotherm	[13]
Non-equilibrium-based methodologies	Microkinetic modeling (MKM)	[41] and [42]
	Kinetic Monte Carlo (KMC) analysis	[40] and [41]

crystal morphology significantly from one dominated by (111) and (100) facets to the one dominated by (110) facets, at high S chemical potentials.

The surface coverage of species at specific reaction temperature and partial pressure of the gas can be estimated using an appropriate adsorption isotherm model. Zhu B et al.[13] estimated the surface free energy of low-index Cu surfaces using a combination of the Langmuir adsorption isotherm model and DFT calculations of adsorption energy. This free energy was used in the Wulff construction scheme to study morphological evolution of copper nanoparticles in a humid atmosphere in the temperature range of 250 K–500 K. At high temperatures greater than 500 K where there is limited coverage of water on the surface, the nanoparticle was a cuboctahedron, dominated by Cu(111) and Cu(100) surfaces, while at intermediate temperatures of 380 K–420 K, the nanoparticle was a rhombic dodecahedron, dominated by Cu(110) surface.

Based on a microkinetic model built with a structureless catalyst, the most abundant reaction intermediates on the catalyst surface at the early zone of a packed catalyst bed during partial oxidation of methane on Rh/Al_2O_3 were identified as oxygen and at the later zone was identified as carbon monoxide with hydrogen, respectively, by Cheula et al.[42] The bulk nature of the catalyst (metal vs. metal oxide) at the partial oxidation condition was then identified using surface phase diagram analyses which showed Rh_2O_3 to be the dominant catalyst in the initial reactor zone and metallic Rh to be the dominant catalyst in the later part of the reactor. With Rh_2O_3 as the catalyst in the early zone of the reactor, the equilibrium morphology of the catalyst under conditions relevant for this zone was identified with predominant exposure of $Rh_2O_3(0001)$, Rh_2O_3 $(11\bar{0}2)$, and Rh_2O_3 $(11\bar{2}3)$ surfaces with the small contribution of the other surfaces diminishing along the axial direction in the reactor. In the later zone of the reactor with metallic Rh as the dominant catalyst and CO and H as the dominant adsorbed species, Rh nanoparticle with predominant exposure of Rh(111) surface followed by Rh(311) and Rh(100) surfaces was identified. This approach of utilizing surface coverage from kinetic model for the *ab initio* thermodynamic analysis is superior to the equilibrium surface energy calculations as it accounts for kinetic factors on the surface phenomena which are overlooked by the purely thermodynamical approach.

A few studies aim at developing morphology control strategies and hence employ Wulff construction as a tool in the computational screening of additives, promoters, or moderators,[43–45] as the nature of ligands plays a major role in surface morphology evolution. Domingo et al.[43] investigated the role of N-containing ligands in changing gold nanoparticle morphology. It was found that the morphology change was prominent in the case of specific ligands such as aniline, ethylene diamine, pyrrole, and trimethylamine, as the dominantly exposed facet on gold nanoparticles changed from Au(111) to Au(110). Among the studied ligands, ethylene diamine brought about the highest change in morphology, as the percentage exposure of Au(110) facets increased to 94% from 10% (over a bare gold nanoparticle).

Identification of descriptors for specific morphology control is crucial in screening and identification of potential modifiers to generate such morphologies. García-Mota et al.[44] studied the morphology evolution of supported Ag nanoparticles on α-$Al_2O_3(0001)$ using Wulff-Kaichev methodology and predicted surface energy descriptors for obtaining cubic nanoparticles that enable higher selectivity for ethylene epoxidation. It was reported that a surface energy ratio between the (100) and (111) facets of Ag should be less than $\frac{1}{\sqrt{3}}$, to obtain better ethylene oxide selectivity. This descriptor can be used to screen supports and modifiers that enable selective formation of cubic Ag nanoparticles for selective epoxidation of ethylene. Similar application of Wulff construction is seen in designing multiferroic materials. Ribeiro et al.[45] presented a methodology for designing $ATiO_3$ (A = Mn, Fe, Ni) materials, wherein the surface energy was used as the descriptor to selectively obtain (001) and (111) facets for enhanced superficial magnetic properties.

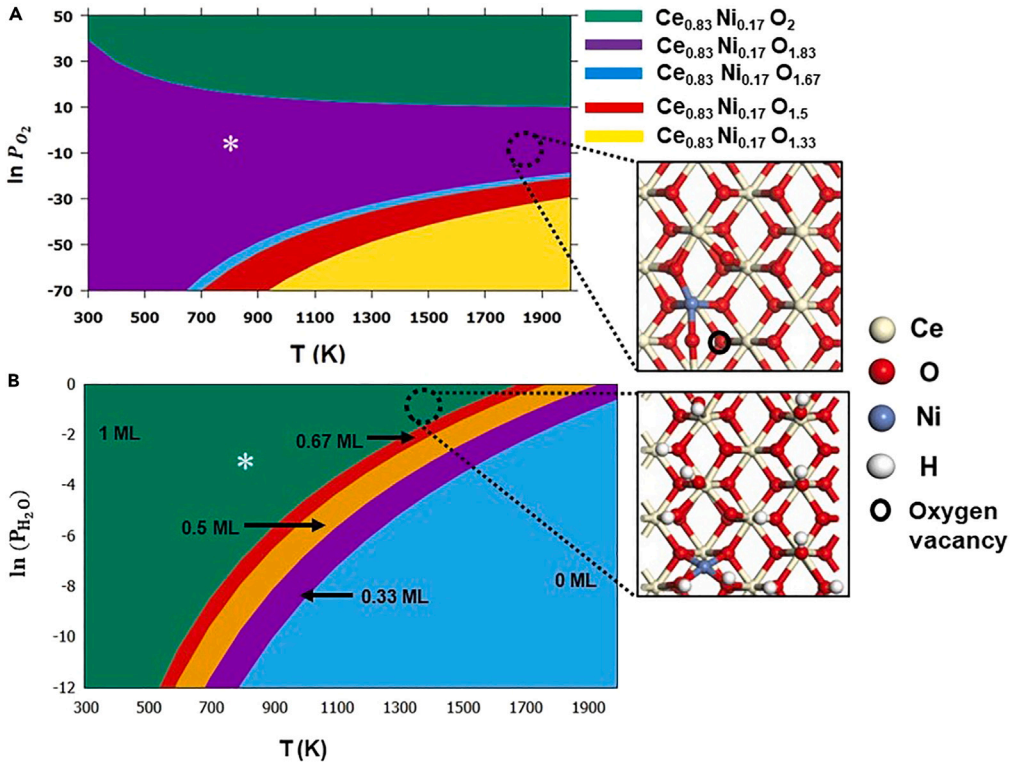

Figure 3. Surface phase equilibria studies depicting the stability of (a) various phases representing different extents of surface reduction (number of surface oxygen vacancies) on Ni-doped ceria nanorod surface and (b) hydroxyls on Ni-doped ceria nanorod surface with increasing coverage to complete monolayer as reported by Pushkar et al.[47] under typical propane oxidative dehydrogenation conditions

The surface atomic arrangement (top view) of the dominant surface phases predicted by the two-phase diagrams is shown to the right of each phase diagram. *Here asterisk in each phase diagram denotes the typical propane oxidative dehydrogenation reaction conditions.

Surface structure determination

Identifying the surface atomic arrangement and species composition, sometimes referred to as "surface phase", is an essential part of catalyst modeling. Surface phase equilibria studies are one of the commonly adopted thermodynamics-based methods to predict the stability of different phases of a catalyst surface under different reactive gas environments. They give valuable information on structural changes and surface stoichiometry, adsorbate coverages, and stability of clusters in multi-species catalysts. Prediction of defect concentrations over reducible metal oxide surfaces is crucial to model redox reactions. Typically, DFT calculations, together with *ab initio* thermodynamics calculations, are performed to obtain surface phase transitions at various reaction conditions. The Gibbs free energies for forming various surface phases as a function of the chemical potential of the reactive gas-phase species of interest are plotted. The intersection of these phase lines denotes phase transition, and the chemical potential corresponding to various phase transitions is used to obtain a surface phase diagram.

For ceria nanorod catalyst, a stoichiometric oxygen-dense phase was predicted to be stable under typical redox conditions,[46] while low-valent Ni-doped ceria nanorods were predicted to expose a partially reduced surface with 8.3% surface oxygen vacancy concentration and surface stoichiometry of $Ce_{0.83}Ni_{0.17}O_{1.83}(110)$[47] for a wide range of temperatures and oxygen pressures as shown in Figure 3A. Suthirakun et al.,[48] in their electrochemical H_2 oxidation studies over $Sr_2Fe_{1.5}Mo_{0.5}O_{6-\delta}$ (SFMO perovskites), predicted the optimal surface Fe/Mo ratios at typical solid oxide fuel cell operating conditions using *ab initio* thermodynamic studies. Their studies predicted lower Mo concentrations on the surface at typical anodic conditions, an essential piece of information required for modeling complex systems such as perovskites.[48]

Table 2. Different catalyst structure dynamics prediction techniques and corresponding references from literature

Technique	References
Grand canonical Monte Carlo (GCMC) simulations	[53]
Ab initio molecular dynamics (AIMD)	[54]
Grand canonical Monte Carlo (GCMC) and reactive force-field (ReaxFF) molecular dynamics simulations	[55]
Classical molecular dynamics simulations	[57]
Density functional theory (DFT) method	[58] and [59]

Substrate coverages are very common in surface catalysis. Surface phase equilibria studies also provide reasonable estimates of adsorbate coverages by analyzing the stability of each coverage at typical reaction conditions. Predicting substrate-covered surfaces is essential to capture the impact of adsorbed substrates on the reaction mechanisms and overall morphological changes during the reaction. Penschuke et al.[49] compared the stability of vanadia monomers (VO and VO_2) over ceria-supported vanadia nanoparticles and predicted the most stable monomer stoichiometry to be VO_2. Furthermore, among the various vanadia oligomers probable at higher vanadium loading, the trimeric V_3O_6 was estimated to be the most stable. The H and CO coverage studies reported for molybdenum carbide systems[50,51] revealed possible surface coverages relevant for various industrial reactions, such as WGS, CO hydrogenation to alcohols, and Fischer-Tropsch synthesis. Wang et al.[50] reported monolayer CO coverages of 9.64, 10.85, and 11.74 CO/nm^2 over the (101), (001), and (201) surface facets of β – Mo_2C catalysts. Similar studies for hydroxyl coverages over Ni-doped ceria nanorods reported the stability of 1 ML hydroxylated surface at typical propane oxidative dehydrogenation (ODH) conditions (Figure 3B).[47] Geng et al.[39] studied the H surface coverages on six different surfaces on HCP Co nanoparticles and reported fractional coverage ranging from 0 ML at high temperature/low hydrogen partial pressure condition to complete monolayer at low temperature and high hydrogen partial pressure.

Although equilibrium surface structure predictions are thermodynamically consistent, the adsorption-desorption and surface reaction kinetics are not usually accounted for in such analyses. Stochastic kinetic modeling techniques such as KMC simulations have been used to account for the substrate adsorption-desorption kinetics and accurately capture the coverages of metastable adsorbate species, which are otherwise overlooked in a typical thermodynamic analysis. Pilania et al.[52] used KMC simulations to predict a metastable O adatom-covered phase over $LaMnO_3$ catalyst surfaces which was not traced by typical *ab initio* thermodynamic calculations. Such studies give a more reliable operando catalyst structure as they do not assume equilibrium. However, these are computationally intensive compared to the equilibrium-based calculations.

Modeling and understanding catalyst structure dynamics

Typically, *ab initio* thermodynamics-based catalyst morphology prediction is relevant to and applicable under equilibrium conditions. However, there might be insufficient time to reach equilibrium at typical reaction conditions. The success of *ab initio* thermodynamics-based simulations would rely partially on the numbers of potential surface configurations sampled. However, manually exploring the vast possibilities for many materials is difficult and may not be exhaustive. Moreover, many reactions experience a dynamic chemical environment, changing surface coverage of species and enabling catalyst fluxionality. Hence, there is a need to investigate catalyst structure dynamics and structure under dynamic reaction conditions. Additionally, understanding the reconstruction process may be useful in identifying metastable catalyst structures and developing appropriate computational models in such cases. Some of the catalyst structure dynamics prediction techniques are summarized in Table 2.

Vaidish et al.[53] conducted *in situ* time-dependent STM study on Pt catalyst at $1.5*10^{-6}$ Torr CO pressure and higher at room temperature and noticed progressive detachment of sub-nanometer Pt clusters. At a CO pressure of 0.1 Torr, they observed more Pt clusters corresponding to a size range of 0.5–0.8 nm. The origin of the cluster reconstruction process was analyzed by Monte Carlo (MC) simulations at 0.0007, 0.5, and 450 Torr CO pressure, using machine learning inter-atomic potentials trained by DFT

calculations. Simulations revealed that higher CO coverage on the step sites of Pt caused rearrangement and extraction of step site atoms and, thereby, island formation. This study showed that appropriately designed MC simulations can qualitatively predict the catalyst features in operando conditions.

MD is also a powerful and widely used technique for operando modeling of the catalyst. Yongpeng et al.[54] studied reconstruction dynamics of Pd@Au core-shell nanoparticles at different CO coverages using *ab initio* molecular dynamics simulations. Very high CO coverage was found to increase the stability of the surface by limiting the surface atom diffusion into the second layer. Thomas et al.[55] used a combination of GCMC and reactive force-field (ReaxFF) MD simulations to obtain a reasonable catalyst model representing Pd clusters supported on a CeO_2 catalyst. GCMC simulations were performed at 500 K and 1 bar O_2 pressure to estimate the Pd-CeO_2(111) surface oxygen uptake. The resulting model was further simulated with ReaxFF MD simulations to identify the catalyst structure, with possibility of Pd diffusion into CeO_2. For more insights on the use of MD simulations in reaction kinetics analysis, readers are directed to the review article by Liu et al.[56]

Classical MD simulations are computationally less expensive than *ab initio* simulations. Mengjiao et al.[57] investigated the temperature effect on the morphology of iron and iron carbides (Fe_3C, Fe_5C_2, and Fe_7C_3) that are commonly used in high temperature applications such as carbon nanotube growth and Fischer-Tropsch synthesis using classical MD simulations. The MD simulations were performed with the modified embedded-atom method (MEAM) potential for Fe and Fe_3C and analytical bond-order potential (ABOP) for Fe_3C_2 and Fe_7C_3. A Wulff morphology was built using the surface energies obtained using MEAM and ABOP potential-based classical MD simulations of iron and different iron carbide nanoparticles. These nanoparticles exhibited the dominance of low Miller index surfaces for temperatures up to 1000 K and high index surfaces for temperatures above 1000 K. A significant change in exposure ratio of different surfaces across the considered temperature range of 100 K–1300 K was reported, suggesting the possibility of temperature-based morphology change.

Typically, supported and unsupported sub-nanometer catalytic clusters can exist in many low-energy isomers and the most stable isomer may not be the most catalytically active.[58] Additionally, isomers may interconvert at elevated temperatures due to the low barriers to be overcome and therefore exhibit fluxionality during catalysis. Rahul et al.[58] created initial geometries of free Pt_{13} clusters using the bond length distribution algorithm and optimized them using DFT simulations to identify the global minimum and the metastable isomers (within 0.4 eV energy difference from global minimum). A phase diagram was developed at different O_2 pressures and temperatures with these global minimum and metastable clusters to identify the most relevant structure corresponding to the gas-phase oxygen reduction reaction condition. Huanchen et al.[59] studied the fluxional behavior of Pt_7 cluster on Al_2O_3 support using the nudged elastic band-optimized minimum energy path of Pt_7 isomerization by considering different isomers corresponding to local/global minima as the endpoint. Most of the isomers corresponding to local minima were within an energy range of 0.6 eV of the global minimum. By considering the typical CO oxidation reaction temperature of 700 K, cluster reconstruction is possible for such small clusters. Hence metastable configurations of the clusters may need to be considered for the catalyst model for comprehensive evaluation. For more details on operando modeling, we refer to the reader to ref.[60,61].

Machine-learning-based predictive methods in structure determination

Predictive methods based on machine learning together with DFT can be employed in the identification of catalyst structures. The use of first-principles calculations in identifying the detailed site and coverage effects would require an enormous amount of computational power. Hence, with the advances in the machine learning concepts, models can be trained with experimental and computational data.[29] Specifically, the *in situ* surface characterization like IR spectroscopy using appropriate probe molecules has been very successful in unraveling the complex surfaces of simple catalysts by coupling them with computational IR calculations. Lansford et al.[29] demonstrated the use of IR spectroscopy as a method with enormous data generated from their first-principles calculations and combining it with the machine learning concepts for deducing the structure of Pt nanoparticle using CO as a probe molecule. Here, DFT-generated frequencies and intensities at low CO coverage for each binding type were taken as primary dataset on different Pt clusters and nanoparticles of multiple coordination sites representing facets, defects, edges, and corners. Outliers identified from the primary dataset were screened to generate secondary spectral data for each binding type and generalized coordination number (GCN). As the frequencies were very

sensitive to CO coverage, the dependence of IR data on coverage was a function of the binding type and independent of GCN. Finally, two neural network models were trained for the binding type and GCN using the secondary data as input for the inverse mapping matching of the experimental spectra and thus predicting the Pt surface microstructure with (111) and (110) as dominant facets.[29]

Another interesting study by Li et al.[62] used the stochastic surface walking-based neural network (SSW-NN) to self-learn a DFT global dataset of potential energy surfaces and predicted the desired catalytic surface phases of Pd-based nanoparticles required for achieving better yield in acetylene semi-hydrogenation to ethene and ethane. The authors were able to predict the hydride phase, i.e., Pd_4H_3(100) formed under a hydrogenation environment, which was involved in the deep hydrogenation of acetylene to ethane using their SSW-NN model. They confirmed the prediction by synthesizing a nanoparticle which predominantly exposed the predicted phase.[62]

Data-driven catalyst screening and discovery

Machine learning tools can be very handy for efficient, rapid (descriptor-based) catalyst screening and design,[62–68] especially in discovering new catalytic compositions when the chemistry of the individual constituents is relatively simple such as metals. The compositions so obtained can then be modeled and subjected to detailed first-principles-based structure, active site elucidation, and characterization/validation, as explained in previous sections. In their review of the available machine learning methods for efficient catalyst search, Musa et al.[68] compared the benefits of new-age ML-based techniques such as reinforcement learning over the conventional genetic algorithms, laid out protocols for choosing the right descriptors, and developing catalyst search tools. Data-driven predictions in heterogeneous catalysis are broadly classified into two types: forward and inverse predictions. The forward prediction objectives involve the prediction of catalytic performance with the known structural information. Machine-learning-based catalyst search/screening comes under inverse prediction objective, wherein for a required performance, the desired catalyst structure/composition is deduced.[67] Suzuki et al.[66] used the available data on the catalyst structure, performance descriptors, yield, and selectivity for oxidative coupling of methane (OCM) reaction and did an inverse prediction of the optimal compositions of various catalytic components for the design of novel catalysts. The authors predicted 20 promising candidate catalysts for OCM reactions by considering various alkali and alkaline earth metal combinations with MgO and La_2O_3 as the base metal oxides using various ML techniques such as non-linear regression models and cross-validation techniques.

DFT-aided machine-learning-based techniques have been used to discover multi-metallic high-entropy alloys (HEAs).[69,70] Lu et al.[70] developed a neural network (NN)-based model to predict the adsorption energy of hydroxyl (OH*)—a key intermediate in the oxygen reduction reactions. They used the same to decouple the ligand and coordination effects (determining factors in active site elucidation) on the performance of HEA catalyst (IrPdPtRhRu) for fuel cell applications. The authors developed a high-dimensional NN model that predicted the adsorption energies with high accuracy (MAE = 0.09 eV; root-mean-squared error (RMSE) = 0.12 eV). They further reduced the same to a linear scaling type model which can be used as an efficient tool for rapid catalyst screening and discovery of HEAs with a minimal loss of accuracy.[70] Similar studies exist on the discovery of atomically dispersed/single-atom catalysts for hydrogen evolution reaction applications.[63,64] Sun et al.[63] proposed a data-clustering algorithm called the fuzzy C-means model for data separation in ML and developed guidelines for the rapid catalyst screening and discovery of graphdiyne-based atomic catalysts. The authors considered all transition and lanthanide metals as heteroatoms and used reaction energies of H_2 adsorption, desorption, and dissociation as the descriptor for training their dataset, unlike the conventional choice of simple H adsorption energy. The developed ML model was used to predict H adsorption energies over the screened catalysts.

Although data mining-derived ML models can circumvent the computational costs incurred from pure first-principles-based structure prediction, they are limited applicability to simple systems such as metals and have a long way to go before they can be reliably applied for screening complex catalytic systems such as metal oxide, supported metal/metal oxides, and so on. Moreover, the predicted catalytic compositions are sometimes unphysical and might not be synthesized due to the lack of experimental synthesis procedures. Hence, first-principles-derived ML models need to be developed for more realistic structure predictions.

Active site nature and its impact on catalysis

Identifying the nature of active sites is an essential step in catalyst modeling and an important prerequisite for reaction modeling. The presence of energetically/electronically heterogeneous active sites impacts the overall reaction energetics via site-dependent mechanisms/energetics and ultimately determines the product distribution.[47,71,72] Except for simple systems such as metals and pristine metal oxide surfaces, most realistic catalyst systems exhibit active site heterogeneity in varying degrees. It is commonly identified using "activity descriptors," such as binding energies of specific species that describe the reactivity of various active sites on the catalyst surface. For example, H adsorption//binding energy over surface oxygen on metal oxides or adsorbed oxygen containing metal catalysts is commonly considered as the reactivity descriptor for reactions involving H abstraction.[71,73] Typical systems that exhibit site heterogeneity are single-atom alloys, metal oxide, metal carbides, doped, defect containing, and supported systems.[47,72]

The site heterogeneity in alumina-based catalysts is known to cause changes in the reaction mechanisms. Dixit et al.[71] investigated the site heterogeneity in the Lewis acid-base (Al-O) pairs of γ-Al_2O_3 and their significance on the propane non-oxidative dehydrogenation energetics. The site heterogeneity was captured for nine unique Lewis acid-base pairs each over Al_2O_3(100) and Al_2O_3(110) facets, using dissociative H_2 binding energy. The propane dehydrogenation mechanisms were site dependent, i.e., concerted mechanism over most active site pairs and sequential mechanism over the remaining pairs. The energetics varied significantly over each active site pair and was described with the help of a volcano plot relating the propane dehydrogenation turnover frequencies and H_2 binding energies.[71]

Sometimes, site heterogeneity might impact reaction energetics and not necessarily the nature of mechanisms. Pushkar et al.,[47] in their investigation of propane ODH over Ni-doped ceria nanorods, reported the significance of active site heterogeneity of surface oxygen caused due to local electronic rearrangement by Ni-doping-dictated surface reduction. On analyzing the hydrogen adsorption energies, four different types of surface oxygen were identified as significant contributors to the catalyst activity and were found to impact propane ODH energetics without influencing the nature of C-H activation steps.[47]

Another important implication of active site heterogeneity is the breaking of kinetic linear scaling relationships (LSRs), a commonly used tool for the kinetic parameter (activation energy barrier) prediction using a linear relationship developed for thermodynamic parameters like binding or reaction energies.[72]

Computational and data-driven methods for operando catalyst characterization and validation

Validating the computationally predicted equilibrium catalyst surface structure and active sites is critical for ensuring near-operando mechanistic insights and development of accurate first-principles detailed kinetic models from reaction network and energetics analyses. STM is an excellent technique for the direct validation of the catalyst surface structure as a one-to-one mapping can be obtained from experimental and computational analyses. *In situ* or operando vibrational spectroscopy, both IR and Raman are good but indirect choices for validation of catalyst surface structure and active sites, as computational IR and Raman analyses are fairly well-established protocols now.[74,75] A suite of non-vibrational spectroscopy techniques involve X-ray irradiation and subsequent analysis. Machine learning tools have been successfully employed for interpreting the spectra obtained from X-ray absorption-based spectroscopic techniques (XAS).[76–81] These include X-ray absorption near-edge structure (XANES), X-ray absorption fine structure, extended X-ray absorption fine structure (EXAFS), and electron energy loss-based spectroscopic techniques such as energy loss near-edge structure (ELNES) or extended electron energy loss fine structure (EXELFS).[82] A few applications of these techniques in establishing the catalyst/material surface structure are discussed in this section.

Infrared spectroscopy

IR spectroscopy is a powerful technique used to characterize the surface of catalyst structure using the vibrational frequencies of the adsorbates and these are obtained rapidly *in situ* for both the gas/solid and liquid/solid interfaces.[29] Vibrational spectra have been used to characterize adsorbate/surface vibrational frequencies for determining the surface of the catalysts including metals,[83] metal oxides,[84] and metal-organic frameworks.[85] Computational IR technique is used to validate the catalyst models by capturing the catalytic signatures on the surface of the catalyst observed in the experiments. IR intensities and frequencies are obtained using density functional perturbation theory (DFPT) and finite

difference-based methodologies, respectively, wherein the change in atomic polarizabilities with reference to an external electric field is estimated using DFT calculations. Most often, the computationally obtained frequencies are compared against experimentally observed frequencies from diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) for surface and gas-phase species and Fourier transform infrared (FTIR) analyses for surface species. With appropriate probe molecules, the nature of the surface sites on the catalyst is well reflected by the shifts in the vibrational frequencies of the characteristic modes of the probe molecules interacting with the surface.

Lansford et al.[86] employed a first-principles-based machine learning framework for detecting the microstructure of Pt catalyst surface using CO, NO, and C_2H_4 as the probe molecules. Here, they used C-O stretch frequencies as a descriptor for different Pt atop sites of similar geometries for a wide range of GCN values. They used the concept of interaction energy between the metal and adsorbates and scaled it across GCN to identify C_2H_4 as an ideal probe molecule for detecting different Pt coordination environments.[86] Penschke et al.[49] proposed different $(VO)_k$ and $(VO_2)_k$ (k = 1,2,3) species supported on CeO_2(111) catalyst and employed computational IR analyses for the structure validation. They captured the characteristic blue shift in the vanadyl (V=O) oxygen frequencies to validate the catalyst model. The computed IR frequencies exhibited a blue shift of 25 cm^{-1} for V_3O_6 trimer species in agreement with experimentally observed blue shift of 27 cm^{-1} for increase in the vanadia loadings from 0.3 V atoms/nm^2 to 0.7 V atoms/nm^2 in the CeO_2-based catalyst.[49]

While this is a powerful and easy-to-apply tool for a wide range of catalytic systems, the ideal probe molecules for the system at hand should be identified. Additionally, benchmarking of the computational methods for the vibrational spectra calculations should also be done for every probe-catalyst system separately.

Raman

Raman spectroscopy is a powerful technique to characterize catalyst structure (bulk and surface features), defects, adsorbates, and depositions (clusters and oligomers) and validate reaction mechanisms. It is governed by the Raman effect, caused due to the excited molecular vibrations in a sample when irradiated by monochromatic light. Hence, the incident light gets scattered inelastically and shifts its frequency. Computationally, the frequency shift is captured by vibrational frequency-based calculations, and the magnitude of the inelastic scattering (Raman intensities) is obtained by the derivative of the macroscopic dielectric tensor with respect to normal modes.[87] Within the DFT framework, the normal mode vibrations are calculated via a finite difference approach or DFPT, while the macroscopic dielectric tensors are obtained using DFPT.

In their detailed computational Raman analyses over ceria-based materials, Xu et al.[88] assigned the identities for the two Raman peaks observed for M-*O*-Ce bond vibrations in doped ceria systems, where M denotes the dopant metal, using bulk equivalent models of various ceria-based catalysts. A peak observed around 560 cm^{-1} was assigned to the M^{3+}-*O*-Ce^{4+} near oxygen defects, while the peak around 600 cm^{-1} was assigned to the M^{n+}-*O*-Ce^{4+} without oxygen defects. Such distinctions and detailed descriptions were difficult to capture in previous experimental Raman studies.[89,90]

Although, most computational Raman studies employ bulk catalyst models to explain the surface phenomena observed during reactions (due to easier prediction of bulk phonons/vibrations),[47,88] a few recent studies have also shown surface phonon prediction using surface models. Schilling et al.,[91] in their combined experimental and theoretical investigation, studied CeO_2(111) surface phonons and assigned the Raman band at 250 cm^{-1} to the longitudinal Ce-O stretch on the ceria surface. This is an important assignment as it is an indicator of the oxidation state of the Ce entities of CeO_2(111) nanoparticles. They also studied the peroxide (O_2^{2-}) vibrations, which are one of the important active oxygen species that influence oxidation reactions over the ceria surface. Various configurations of O_2^{2-} (based on their neighboring Ce states) were identified, and the corresponding Raman modes were predicted, which are otherwise untraceable in experiments. Such assignments are essential to predict the type of active sites and describe reaction mechanisms for oxidation reactions over the ceria surface.

XPS

XPS is a powerful technique for studying the electronic structure and active sites of catalyst surfaces. The surface atoms of the catalyst undergo electronic changes in their valence regions due to reduction,

oxidation, and interactions with adsorbates during chemical reactions. This can be captured by analyzing the response of core electrons via the calculation of core-level binding energies (CLBE). Experimentally, it is captured by the changes in the X-ray photoemissions,[92] while computationally, it is calculated using DFT-based electronic calculations. Due to the complexity of photoemission processes, theoretically calculated CLBEs aid in deconvolution and accurate interpretation of experimental spectra. Under the DFT framework, the core-level binding energies are calculated using initial (without accounting for core-hole screening) or final state (with core-hole screening) approximations. The difference between the core-level binding energies of a surface atom and its bulk equivalent, referred to as surface core-level shifts (SCLS), is used as a descriptor to explain various surface chemistries related to adsorption, site coverages, and defect formation.[92–96]

Posada-Borbón et al.[95] used SCLS of surface O atoms to explain the impact of oxygen vacancy formation on the surface reduction in ceria catalysts. A negative SCLS calculated for oxygen in the vicinity of the vacancy was ascribed to the destabilization of the surface due to the formation of Ce^{+3} species around the oxygen vacancy.[95] SCLS can also be used to predict the adsorption site, and hence the adsorbate-induced growth of catalysts or adsorbate islands.[97,98] Lousada et al.[97] predicted the extent of O coverage over Al(110) surfaces by comparing the trends in SCLS of surface and subsurface Al atoms of Al(110) with that of the surface Al of Al_2O_3. This provided a rational means of modeling Al(110)-exposed catalysts for oxidation studies with a right description of oxygen coverages/islands. Lizzit et al.[93] explained the nature of O chemisorption on the Ru(0001) and Rh(111) surfaces using the SCLS in Ru atoms. A similar trend in SCLS with an increase in O coverage indicated the presence of similar on-surface O chemisorption behavior in both Rh(0001) and Rh(111) surfaces, although O adsorption sites were different (hcp in the case of Ru(0001) and fcc in case of Rh(111)).[93] However, sole experimental XPS analysis would have failed to observe the distinctions in the active sites, as the spectra would have been similar for the two surfaces. SCLS is also used to resolve ambiguities in active site determination from experiments. Köhler et al.[94] confirmed the adsorption of CO on atop Rh site over Rh(111) surface as the Rh SCLS value corresponding to atop adsorption mode agreed well with that of the experimental core-level binding energy shift, while it did not for other sites (bridge and hollow sites).[94]

Although an indirect means, computational XPS analysis is becoming a useful tool for ascertaining active sites for adsorption and reactions. Sufficient caution needs to be exercised during the application of this technique as an accurate prediction of CLBE is contingent on the right choice of the catalyst models (lateral dimensions and number of atomic layers),[99] and computational parameters such as the application and value of on-site Coulombic interactions (Hubbard U), and the type of DFT exchange correlation functionals.[100]

STM or scanning tunneling electron microscopy (STEM)

High-resolution STM experiments are used to elucidate atomic-scale catalyst structure both qualitatively and quantitatively. It is used in studying surface defects,[101] growth,[102–104] adsorption and coverages,[105–107] cluster agglomeration,[108] and so on. However, one of the fundamental challenges for an experimentalist is the assignment of atomic identities for the observed STM images. This is easily resolved by theoretical STM simulations with the help of DFT-obtained charge densities of catalyst surface species, which can be used to explain the underlying surface structure. Hence, computational STM is a direct method of catalyst model validation, unlike the other techniques discussed in this article which are indirect in nature.

Quertite et al.,[101] in their work on the growth of NaCl over Ag(110), used a combination of experimental and simulated STM studies to explain the formation of Cl vacancies at one 1 ML coverage of NaCl film. The observed "depression" in the experimentally determined STM image was ascertained to be the Cl vacancies, as the simulated STM image of a Cl vacancy-containing model also showed a similar-sized depression.[101] The STM simulations also resolved ambiguities in predicting the right atomic compositions of the growing films. Chen et al.[103] observed two distinct types of step edges for the ceria growth (islands) over Pt(111) in their experimental STM images. However, the right description of the atomic species involved in the formation of these step edges was impossible solely by experiments. To explain this, four surfaces with different Ce and oxygen compositions were modeled, out of which two surfaces with step edge composition of O_i-O_s-Ce and O_s-O_i-Ce/O_i-Ce-O_s (where O_i and O_s denote interfacial and surface oxygens over Pt(111)) reproduced the experimentally observed step edges using DFT calculations and computational

STM analyses. Such studies not only explain the experimentally observed features but also gave realistic catalyst models for detailed mechanistic studies.

Data-driven methods for assisting computational and experimental operando characterization

The XANES, EXAFS, ELNES (analogous to XANES), and EXELFS (analogous to EXAFS) techniques are atom specific and are highly sensitive to the local chemical environments and capture the geometric and electronic effects around active sites, which ultimately impact the catalyst structures and reactivities. Interpretation and analyses of such spectra can substantially benefit from data-driven and computational approaches. For that, a reasonable dataset (hundreds to thousands) comprising the spectra of known structural features (descriptors such as bond lengths, coordination numbers, etc.) is required for a desired atom/element for training the supervised machine learning models. However, for systems with unknown descriptors, a larger dataset (thousands to millions) obtained from experiments and simulations can be subjected to unsupervised machine learning methods such as principal component analyses (PCA), multivariate analyses, and so on, for classification, dimensionality reduction, noise removal, and clustering to make meaningful inferences.

Liu et al.[79] used a supervised machine learning method involving a NN to understand the XANES spectra of copper oxide clusters using structural descriptors such as coordination numbers. The authors created 25 CuO and 30 Cu_2O cluster models by truncating the original bulk structures of Cu_2O and CuO along different low-index planes, such as (100) and (111) and to standard morphologies such as tetrahedral, octahedral, and cubic shapes. They also created surface slab models of (100) and (111), for developing the training set for the neural network. Their developed NN could extract important information such as average particle size and oxidation state using Cu-Cu coordination number as the descriptor.[79] Although metal-to-metal CN is a sufficient descriptor for interpreting XANES for metallic clusters,[81] consideration of metal-oxygen coordination numbers is also essential to predict the sizes of mixed phases (simultaneous presence of Cu, CuO, and Cu_2O) in metal oxide clusters. The unsupervised learning methods have been used to preprocess the XAS data and categorize them into XANES and EXAFS data. Using techniques such as PCA, *Z* score normalization, and k-means clustering, Kartashov et al.[77] developed software tools to pre-process the XAS data and offer a means of data mining to develop new insights about Pd nanoparticles. For more information on the machine-learning-based interpretation of X-ray absorption spectra, available databases for obtaining large-scale spectral datasets, and simulation codes for generating large datasets for XANES, the readers are suggested to refer to the review article by Timonshenko et al.[82]

Causes for incorrect predictions of catalyst structure

Poor catalyst modeling causes a mismatch with those obtained from the experiments. Reliable catalyst modeling requires rigorous testing and optimization of computational and catalyst model (supercell) parameters for DFT calculations. These include cutoff energy for the plane wave basis set, electronic and force convergence criteria for the self-consistent field calculations and the geometry optimization, respectively, k points density for the Brillouin zone sampling, lateral dimensions of the surface, number of atomic layers in the supercell, vacuum thickness above the surface, and so on. These are usually optimized against a suitable descriptor, such as formation (surface/phase) energies, adsorption energies of a probe molecule (describing a desired reaction), vibrational frequencies (absolute and shifts), binding energy shifts (CLBE for XPS analyses), and so on. Improper optimization of these parameters is one of the reasons for the failure of computational characterization having parity with experiments. One of the studies by Trinh et al.[99] on computational XPS analyses presented the need for optimization of some of the above-mentioned parameters, as the binding energy shifts (CLBEs) were found to be extremely sensitive to these parameters.[99] The other possible reasons include improper choice of the catalyst surface (crystal facet), catalyst phase, the surface coverage of species, and model types (oligomeric types, cluster, and periodic supercells). Finally, poor choice of experimental references also impacts catalyst modeling and validation as the quality and rigor of the experimental data that are chosen for developing catalyst models also determines the quality of the computational models.

Summary and iterative workflow for catalyst structure modeling

An ideal iterative workflow for rational catalyst modeling and validation is proposed in Figure 4. Among the first steps is the prediction of equilibrium catalyst morphology at reaction conditions using the commonly observed low-index crystal facets. Based on the most dominantly exposed crystal facets obtained from the

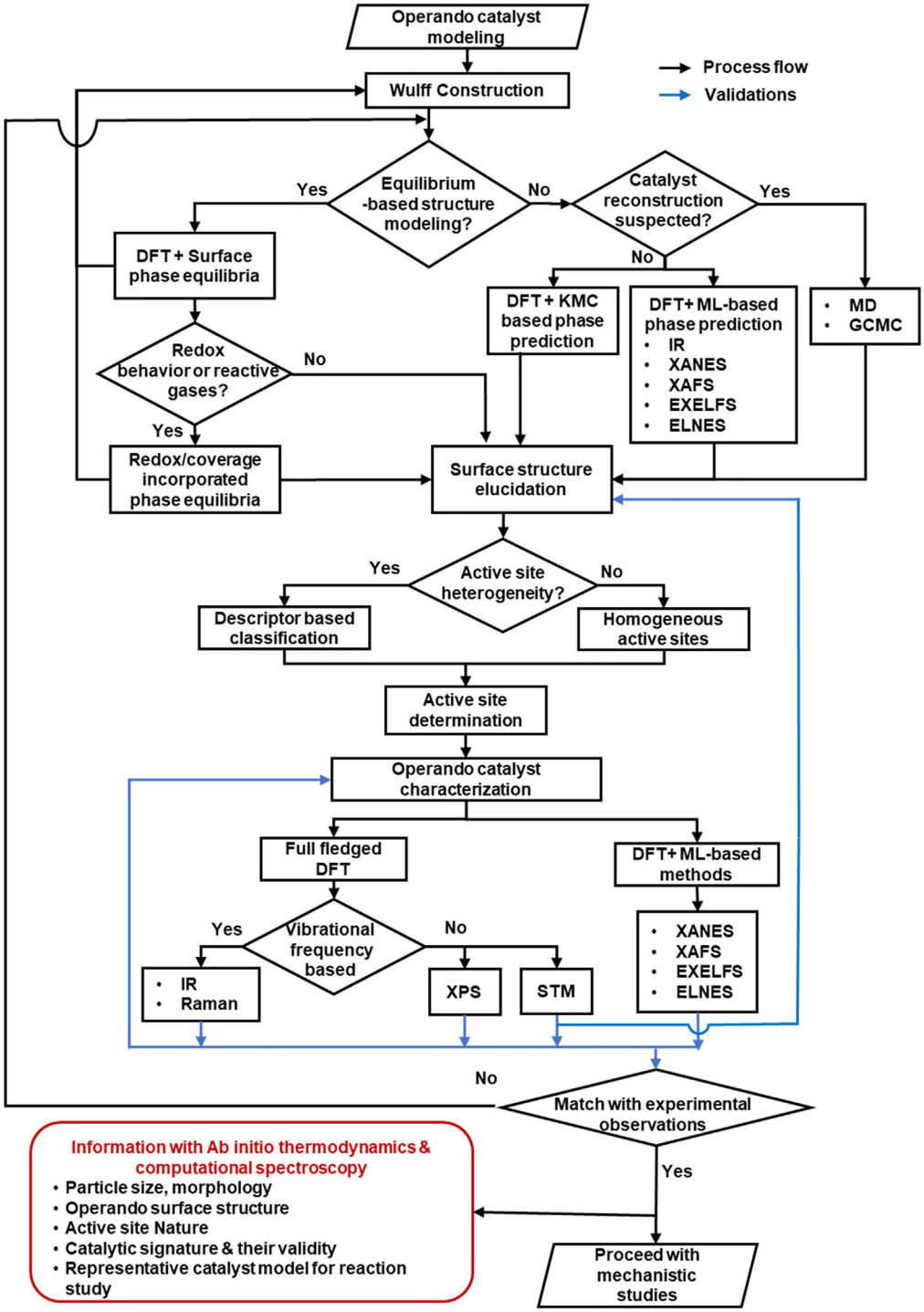

Figure 4. Proposed iterative workflow for rational operando catalyst structure prediction and validation

Wulff construction of the nanoparticle morphology, the surface models for the catalyst can be obtained. This is followed by an accurate description of the surface phase and surface atomic arrangement, either using equilibrium-based (surface phase equilibria) or kinetics-based simulation techniques. If the catalyst structure dynamics is expected to be relevant at reaction conditions, dynamic simulations may be performed to capture those effects. Otherwise, kinetic Monte Carlo or predictive (machine learning) methodologies may be useful. In carrying out surface phase equilibria studies, one must be cognizant of the

Table 3. Different reaction mechanism generators relevant for heterogeneous catalysis and their typical application in different reaction types

Mechanism generator	Commonly used reaction type
NetGen[120]	Deoxygenation[122] Oxidation[123] Oligomerization[122] Cracking[122] Alkylation[122] Aromatization[124]
Reaction Mechanism Generator (RMG)[117,118]	Reforming[125] Oxidation[126]
Rule Input Network Generator (RING)[127]	Dehydrogenation[128] Hydrodeoxygenation[128] Cracking[129]

presence of surface defects, redox behavior of the catalyst surfaces under the reactive gases in the reaction environment, and analyze surface phase equilibria accordingly. These determine the nature and stability of various active sites (inherent and induced) on the surface. Among these, the kinetics-based and predictive methods employing machine learning can be used for more rigorous prediction of surface phases which may normally get overlooked by phase equilibrium studies at the reaction conditions. The nature of active sites is determined using standard descriptor-based evaluations for capturing the heterogeneity present if any.

Finally, the developed catalyst models are to be validated using computational techniques such as IR, Raman, XPS, or STM (or combinations as relevant) against the available experimental evidence. Choosing the right type of technique for the validation is crucial and will depend on the type of catalyst system under consideration and the experimental data available for validation. STM offers a direct means of characterizing the structural features (such as sites) of the catalyst surface. Spectroscopic techniques can be used to explain and validate active sites, surface redox processes, surface reaction intermediates, reaction mechanisms, and so on. Failure to reasonably match the key catalytic signatures observed in the experiments suggests a poor choice or description of a surface model, and the process needs to be revisited (Refer to Figure 6 for more details on the iterative refinement of the catalyst model). Catalyst models that reproduce the experimentally reported catalytic signatures can be used for further reaction studies.

REACTION MECHANISM AND NETWORK GENERATION AND VALIDATION

Detailed reaction mechanism and complete reaction network, together with the corresponding energy profile, is a minimum requirement for the hierarchical detailed kinetic modeling of the reaction. Typically, reactions of small molecules with simple reaction networks are explored based on chemical intuition, breaking down the network into all possible elementary reaction steps. This approach may however be fraught with the risk of personal biases, resulting in incomplete or inappropriate networks. With recent developments in high-performance computing systems, application of DFT simulations to identify reaction mechanisms and mapping energy landscapes of whole reaction networks by calculating energy changes for each elementary step has become commonplace.[4,16,109] This is typically done by analysis of the minimum energy pathway connecting the reactant and the product states via the transition state (first-order saddle point in the potential energy profile) for each of the elementary reaction steps using DFT simulations. When all elementary steps are connected, it forms a continuous energy profile on an arbitrary reaction coordinate, representing the complete catalytic cycle. Such DFT potential energy profiles of reaction networks are now the backbone of reaction mechanism analysis at the atomistic scale.[4] An alternative is the analyses of the underlying free energy profile for the same reaction network, and this is routinely done by adding appropriate corrections using statistical mechanics to the DFT-derived potential energy to convert it to a free energy, ref.[110].

Rule-based reaction mechanism and network generation

As the size of the reactant or the number of species involved in the reaction increases, the number of elementary reactions in the network rapidly increases.[111] Although spectroscopic information on the

observed reaction intermediates is often helpful to identify pathways and eliminate some other possibilities, generation and analyses of large reaction networks are challenging. Here, the automatic reaction network generation algorithms which take information such as reactants, reaction rules, and reaction families as input for generation of the reaction network become extremely useful.[112] Although there are a wide range of reaction mechanism generators for gas-phase chemistries such as Genesys,[113] MechGen,[114] Kinetic Network Generator,[115] and Reaction,[116] a few that have applications for heterogeneous catalysis are listed in Table 3. The Reaction Mechanism Generator[117,118] was developed from the previous codes Exxon-Mobil Mechanism Generator,[119] NetGen,[120] and Exgas.[121] An automatically generated reaction network considers all possibilities, making it large and complex and analyses of such large networks would be computationally intractable.

Data-driven reaction mechanism and network generation

Chemometric techniques and data-driven models have been developed for monitoring reaction dynamics in order to interpret the multivariate (dependent on temperature, residence time, etc.) signals from spectroscopic data.[27,130] These have been shown to assist experimental elucidation of reaction mechanisms as they are carefully trained based on the existing spectral information and can also be used to develop hypotheses for carrying out new experiments. Puliyanda et al.[131] developed the hidden semi-Markov models for both offline and online monitoring of FTIR spectroscopic data, extracted the "latent" features (species identities) in the spectra using Bayesian networks, and interpreted the reaction mechanism and timescales for the identified modes, for low-temperature thermal cracking of Cold Lake bitumen. In another study for the same reaction system, a data injection/fusion model based on the joint non-negative matrix factorization and Bayesian networks was developed to extract the pseudo-component spectra and hypothesize the reaction mechanisms, respectively.[132] Sivaramakrishnan et al.,[130] in their study on the thermal upgrading of Athabasca bitumen, reported the use of the self-modeling multivariate curve resolution method and other statistical tools to develop a reaction mechanism with minimal prior chemical knowledge. For a more detailed understanding of the data-driven approaches in reaction mechanism generation, the readers are suggested to refer to the review by Puliyanda et al.[27]

Recently, many studies have come up with machine-learning-based reaction mechanism generation techniques in order to circumvent the computationally expensive DFT-based mechanistic studies with minimal DFT data for initial training or using data in available experimental databases.[133] Puliyanda et al.,[27] in their review, collated the recent studies that have reported automated and data-driven generation of reaction pathways. A detailed comparison of various rule-based (manually encoded and algorithmically extracted), quantum mechanical simulation-based, and machine learning-based methodologies to generate reaction networks for developing the mechanistic insights was made. Burés et al.[133] recently developed a large-scale artificial intelligence-guided reaction mechanism classification tool. The authors used neural network and deep learning tools to extract mechanistic information from the available experimental kinetic data for which they considered a wide range of reactions such as olefin metathesis, cycloaddition, alkene isomerization, C-H amination, photocatalyzed hydroalkoxylation, and so on. The developed AI model was reported to accurately (>99% categorical accuracy) predict the mechanisms for the aforementioned reactions, at various reaction conditions, and including those of catalyst activation and deactivation steps, which are normally inaccessible in traditional kinetic analyses. Unlike the case of conventional machine-learning-based reaction mechanism generators which require careful choice and assignment of reactivity descriptors[27] (obtained mostly with the aid of DFT), the proposed model by Burés et al.[133] relies only on the experimental dataset for developing the required training set and hence bypasses the inevitable human errors during the computational reaction and kinetic analyses.

Reaction mechanism validation

Validation of reaction mechanism is a prerequisite for conducting kinetic modeling. Conventionally, vibrational frequency-based techniques like IR and Raman spectroscopic analyses have been used in *in situ*/operando mode to assign molecular identities based on the identification of the functional groups from the characteristic spectroscopic signatures observed in the experiments and to obtain mechanistic insights.[49,134–137] The review by Zaera et al.[134] discusses the use of various IR-based techniques including transmission IR (TIR), FTIR, DRIFTS, reflection-absorption, and attenuated total reflection in detail. Modifications to the conventional DRIFTS analyses, such as synchrotron-based DRIFTS, modulation excitation (ME) with phase-sensitive detection (PSD), have yielded accurate descriptions of surface species during reactions.[134,136] Various Raman spectroscopy-based analyses such as time-dependent operando Raman,

ME-PSD Raman, surface-enhanced Raman, and shell-isolated nanoparticle-enhanced Raman have also been adopted for mechanistic studies.[137]

The computational analogs of vibrational spectroscopy-based techniques become very useful to directly assign the molecular identity of the surface species which would in turn help in identification of the reaction mechanism. This approach is to be adopted with a caveat that typically there is a deviation of the computed frequencies of the normal mode vibrations of the intermediates from the experimental frequencies by 10–20 cm^{-1}.[47,49] Pushkar et al.[47] employed computational IR analyses to assign the molecular identities of the surface intermediates observed during propane oxidation over Ni-doped ceria nanorods where formation of over-oxidized oxygenates such as acetone, ketene, and formaldehyde were confirmed. Another study reported the use of computational Raman spectroscopy in mechanistic analyses of WGS reaction over ceria-supported gold catalysts.[138] Characteristic vibrations reproduced for surface hydroxyl, peroxide, formate species, and Ce^{+3}-O vibrations explained the active role of the ceria support in WGS reactions and gave insights into the reaction mechanism.[138]

Hess et al.[139] studied the CO oxidation over Au/CeO_2 catalysts and confirmed the role of oxygen vacancies over the ceria surface in facilitating the adsorption of molecular O_2 and oxidation of CO to CO_2 using operando Raman analyses, which was previously predicted by Henkelman and co-workers[140] in their DFT studies. Another example would be that of the propane ODH studies by Schumacher et al.,[135] where they confirmed the active role of ceria surfaces in the first C-H activation of propane over VO_x/CeO_2 catalysts using operando DRIFTS with isotopic (C_3D_8) experiments and validated a similar claim made by Penschke et al.[49] in their DFT studies using H adsorption studies.

KINETIC PARAMETER ESTIMATION

With reaction mechanisms and reaction network identified, predicting the reaction kinetics is critical to quantify catalytic activity and to compare the performance of different catalysts. This requires estimation of the rate constants for all the elementary reaction steps in the network. Rate constant within the Arrhenius framework or within the transition state theory (TST) framework comprises the pre-exponential and exponential terms. Within TST, the pre-exponential and exponential terms are functions of the activation entropy change and activation enthalpy/energy barrier, respectively. In this review article, the pre-exponential term and activation barrier are referred to as the kinetic parameters.

Semi-empirical techniques

Semi-empirical techniques such as the unity bond index–quadratic exponential potential (UBI–QEP)[141] and transition state/kinetic scaling relations[142] can be used to estimate the activation barriers of elementary reactions. However, activation entropy change for elementary reactions is not easily acquired without the use of quantum mechanical calculations and often the pre-exponentials are obtained by fitting against the available experimental data. This section briefly introduces a couple of semi-empirical methods used in estimation of the activation barriers for kinetic modeling and discusses limitations of these techniques with examples from the literature.

UBI-QEP method

The UBI–QEP is a phenomenological approach modified from the earlier bond order conservation–Morse potential[143] method. The UBI–QEP method analytically describes the Morse potential with a polynomial description and bond index as an exponential function in the quadratic exponential potential.[144] Based on information such as binding strength, type of coordination, atom count, and so on, an appropriate equation comprising the bond dissociation energy and the heat of chemisorption is applied to estimate the activation barrier.[144] Evgeny et al.[141] described the UBI-QEP methods and its applications for the cases with different binding coordination, binding strength, and atomic size of molecules and were discussed in detail for different transition metal catalysts.

Maestri et al.[145] reported a 0.3 eV underestimation of the activation energy by the UBI-QEP method, compared to a DFT reference for surface hydroxyl (OH) dissociation (OH dissociation into O and H) on Rh(111) surface with a 0.5 ML of O coverage. The lateness of the transition state(TS) caused by the O coverage was identified as the reason for the error. A scaling factor was proposed to compensate for the underestimation which brought the predicted O-H activation barriers using UBI-QEP method within

a 10% uncertainty of the DFT reference. For the elementary steps in CO_2 reduction to methanol on Cu-based catalyst, Park et al.[146] improved the prediction of the activation energy barriers by the UBI-QEP method by using binding energy and enthalpy calculations from DFT simulations.

Evgeny et al.[141] compared the UBI-QEP-predicted CO binding energy on different transition metals such as Ni, Cu, Pt, Pd, Ag, and Co with the experimental data reported by Whitten et al.[147] and reported that the UBI-QEP estimations on 3d transition metal catalysts are in good agreement with experimental data while those on 4d and especially 5d metals exhibit deviations. Moreover, the accuracy of the UBI-QEP energetics is sensitive to the factors like 1) the accuracy of the estimated binding energy (BE), 2) the size of molecule, 3) the presence of repulsive interaction due to coverage of reactive/non-reactive intermediates such as O, OH, and H, 4) the nature of the catalyst material and the active sites, and 5) the extent of surface rearrangement.[141] Hence, appropriate corrections and considerations for these factors need to be incorporated while using the method.

LSR

Thermodynamic correlations of adsorption energies of certain species, with a structurally related and simpler species, on similar adsorption site types across different materials are typically called adsorption scaling relations.[148] More useful are the linear correlations such as the Brønsted–Evans–Polanyi (BEP) relation that links the reaction energy for elementary steps (thermodynamics) with activation energy barriers (kinetics)[149,150] as accurate determination of transition state of reaction steps and thereby the activation energy barrier calculation by DFT is computationally more intensive than calculations of the intermediates. In this context, scaling relations are an attractive means for estimation of activation barriers for kinetic modeling of complex reaction networks. An example of the use of LSRs to reduce the compute expense in large reaction network is discussed in the section titled Hierarchical iterative microkinetic modeling of large reaction networks. Like the UBI-QEP method, the LSRs, can at best only give estimates of the activation barriers for elementary steps and not pre-exponential factors which are critical components of kinetic expressions. LSRs are typically valid for simple catalytic systems such as metals, where the dissociation of the reactants and adsorption of the products occur on the similar site types, and hence the activation energy barriers and binding energies are coupled.[148–151]

Complex catalyst systems such as single-atom alloys (SAAs),[72,152,153] supported nanoclusters,[153] and ligand-modified systems typically exhibit active site heterogeneity.[72,152–155] These offer dissimilar active sites for reactant dissociation and product adsorption, resulting in the decoupling of the activation energy barriers and binding/reaction energies. This results in what is now commonly referred to as the breaking of the linear scaling relationship, which leads to deviations in the activation or binding energies predicted using a linear scaling relationship.[72,156] The extent of these deviations depends on the extent of geometric and electronic contributions to/from the catalyst system under consideration.[156] Some of the typical instances leading to the breaking of linear scaling relations include substrate spillover on the surface, secondary interactions, and confinement effects (in zeolites) during surface reactions.[72,153–155]

Darby et al.[72] attributed the deviations in LSR for the dissociation of H_2, CH_4, CO_2, and CH_3OH on SAAs of various metal systems such as Ni, Pt, Rh, Ir, and Pd to the large variation in the reaction energies for the dissociation of these molecules of 0.1–1.3 eV due to adsorbate spillover across the surface. The large deviations were caused due to the dissociation of these molecules on the dopant and adsorption of the product on the dissimilar metal host site. The deviations were also attributed to the valency of the bonding metal atom and the nature of the adsorption configuration. A similar observation was made by Sun et al.[152] over γ-Alumina-supported Pt/Cu SAAs during propane oxidative dehydrogenation. The Pt/Cu SAA offered weaker (−0.57 eV) propene adsorption due to π bonding over the single Pt atom, compared to the stronger di σ bonds in the case of adsorption on Pt (111) (−1.1 eV) and Pt_3Cu (−0.88 eV) alloy. As a result, the propene desorption barrier was lowered and the further dehydrogenation barrier increased over Pt/Cu SAA, hence showing a deviation from the linear scaling relationship developed between the propene dehydrogenation/desorption energy barrier and propane dehydrogenation energy barrier. Lustemberg et al.[153] reported the breaking of linear scaling relations on ceria-supported metal (Pt, Ni, or Co) systems (from a single atom to sub-nanometer clusters), wherein the strong metal-support interactions brought about significant deviations in methane dissociation barriers predicted from conventional linear scaling relationships.

Szécsényi et al.[154] reported the breaking of linear scaling relationships between the activation energy barriers and reaction energies for methane dissociation over different Fe active sites in Fe/ZSM-5 catalysts. The observed deviations were mainly attributed to the presence of confinement effects and adsorbed hydroxyl ligands in the zeolite micropores. The confinement in the zeolite micropores imposed geometric constraints on the transition states, destabilized them, and altered the activation energy barriers. The adsorbed hydroxyl ligands either stabilized the transition states by forming H bonds and reduced the activation energy barriers, or they destabilized the final states by losing the pre-existing H bonds (between the neighboring hydroxyls) during methane dissociation. The former reduced the activation energy barriers, while the latter increased the reaction energies.

Alternatives to semi-empirical techniques for kinetic parameter and reactivity descriptor estimation

Applications of the semi-empirical techniques are limited to some specific systems and are subject to the availability of parameters for the UBI-QEP method and existing LSRs for prediction of activations barriers. Moreover, these techniques are typically capable of predicting the activation barriers for the elementary reactions and not the pre-exponential factor. In addition to calculation of activation barriers, DFT simulations are now routinely used to calculate the activation entropy change for elementary reaction steps and thereby the pre-exponential factor in the transition state theory-based kinetic expressions. Harmonic normal mode vibrational frequency analysis within the DFT framework, together with *ab initio* thermodynamics calculations enable estimation of the activation entropy change from which the pre-exponential factor in the kinetic expression is estimated. For activation barrier estimation on complex catalyst systems containing dopants,[102] vacancy,[102] interface,[103] and involving coverage effects,[104] the typical semi-empirical techniques may not be appropriate choices and there is a need for DFT-based parameter estimation. Based on the ability to estimate activation entropy change and its applicability to a wide range of systems with reasonable accuracy, DFT calculations are now preferred over semi-empirical techniques. Saurabh et al.[157] built a kinetic model obtained from DFT-based kinetic parameter for formic acid decomposition reaction over Pt/C and Au/SiC catalysts and reported a good match with the experimental conversion and TOF by appropriately accounting for the quantity of different sites corresponding to different facets in the kinetic model. Although the DFT results are typically assumed to be reliable enough, the accuracy depends on the choice of spin state of the catalyst model,[158] the DFT exchange correlation functional used,[159] the valence state of the catalyst active site,[160] the impact of surface species coverage on kinetic and thermodynamics parameters,[161] presence of defects,[161] size of the catalyst model,[162] surface structure,[163–166] and approximations/limitations with catalyst model.[157,167]

As an alternative to traditional scaling relations, Baochuan et al.[168] proposed the use of graph-based machine learning models to estimate the binding energies and activation energies of elementary steps in ethanol synthesis from syngas on Rh catalyst. The parity plot of graph neural network and BEP relation-predicted activation energy barriers against the DFT-calculated activation energies were presented. It was shown that the BEP-predicted activation barriers showed significant deviations with RMSE of 0.35 eV while the graph neural network-predicted activation barriers showed a lower deviation with RMSE of 0.23 eV for barriers in the range of 0.5–1.5 eV. They suggested this method to be much more reliable in predicting the activation energies over a broader range of activation barriers compared to scaling relations.[168] The kinetic linear scaling relationships such as BEP relationships are usually developed to predict the activation energies for specific reaction steps on catalyst surfaces as they are mostly single-parameter-dependent (Reaction energies) correlations. The universal/linear scaling relationships developed for simple metal systems[151] rely mostly on the reaction energies where the changes in the chemical identities of the reactants and the surface restructuring of the catalyst during the course of reactions are neglected. Göltl et al.[169] proposed an adaptation of the well-known BEP relationships by adding certain correction terms. These terms described the changes in bonding between the catalyst surface and the adsorbed atom (C, N, O, and H), intermolecular interactions between the surface species, and interaction strength between the surface and the adsorbates.[169] Using techniques such as random forest regression, support vector regression, and gradient boosted regression, the authors developed generalized BEP relationships for metallic surfaces which predicted activation energies with an MAE of 0.11 eV.[169] A detailed review by Lewis-Atwell et al.[170] highlights various machine learning tools for developing data-driven models to predict activation energies and discusses some of the studies that have adopted these for reaction kinetic modeling.

Recent studies have employed machine learning tools to account for adsorbate interactions which prove to be computationally less expensive than KMC simulations, provided the necessary training dataset is generated.[171–173] Wang et al.[172] developed an NN model for describing the CO and NO adsorption over Au(111) surface using electric dipole moments, dipole-dipole interaction potential energy, angle between surface and molecule dipole, work function, and d-band center as the descriptors. The developed NN model predicted the adsorption energies (RMSE = 0.035 and 0.030 eV for CO and NO, respectively, over Au (111)) and transferred charges (RMSE = 0.004 and 0.001 eV for CO and NO, respectively, over Au (111)) accurately. This approach was also extended to CO and NO adsorption over Au(001) and Ag(111) surfaces without any significant loss of accuracy.[172] Another interesting study by Fung et al.[171] reported the use of density of states as the descriptor for developing the NN model, which accurately (MAE = 0.1 eV) described the adsorption energies, accounting for the surface-adsorbate interactions. The model was developed to account for the adsorption of a diverse set of species which included monoatomic (H, C, N, O, and S) and polyatomic (CH, CH_2, CH_3, NH, OH, and SH) species on 2000 different bimetallic alloys of 37 different transition and non-transition metals.[171] A recent review by Mou et al.[173] discusses more such studies that have developed machine-learning-based models to capture lateral adsorbate interactions for a better description of surface kinetics.

Li et al.[174] developed a machine-learning-based chemisorption model using artificial neural networks (ANNs) to accurately describe the CO adsorption energies over the (100) surface of Cu-based multimetallic (belonging to Group VIII and IB) catalysts for electrochemical CO_2 reduction. Using the descriptors such as local electronegativity, effective coordination number, ionic potential, electron affinity, and Pauling electronegativity, the authors trained the NN model to predict the CO adsorption energies which were in many cases comparable (RMSE ~ 0.12 eV) with those obtained using DFT calculations. Wang et al.[168] developed graph-based machine learning techniques to predict adsorbate energies, as this method accurately captured the variations in the adsorbate coverages. The models were developed using techniques such as graph convolutions, Weave, graph neural network, ensemble-median, and ensemble-mean, and were trained using the available DFT data (consisting of 315 species containing C, O, and H) on the Rh (111) system. Another study proposed a unique Adsorbate Chemical Environment-based Graph Convolution Neural Network model for screening atomistic configurations with varying adsorbates and their coverages, coordination environments, and varying catalytic surfaces.[175] This was used to model NO adsorption on a $Pt_3Sn(111)$ successfully which is a key reaction step in the electroreduction of nitrates. The reaction is typically characterized by high adsorbate coverages over low symmetry alloy surfaces and OH adsorption over Pt(221), which is a key step in oxygen reduction reaction, where significant direction-sensitive adsorbate-adsorbate interactions exist. For more such studies on machine-learning-based multi-metal alloy design and choice of suitable descriptors, the readers are suggested to refer to the review by Yang et al.[176]

Addressing the uncertainties in the estimated parameters

First-principles reaction kinetic modeling can be severely impacted by the introduction of errors due to uncertainties at various stages of kinetics and model parameter estimation which rely on basic DFT-based electronic energy calculation or semi-empirical energy calculations. Some approaches used in the literature to address the uncertainties associated with the estimated parameters are discussed in this section.

Estimating and addressing uncertainties in the DFT-estimated parameters

The uncertainty in DFT-predicted energetics can be estimated manually by carrying out DFT calculations for different exchange-correlation functionals and analyzing the variations in energies, but is computationally intensive.[4] An alternative is the use of BEEF-vdW functional—a framework consisting of an ensemble of functionals that are designed for systematic error prediction using Bayesian statistics.[4,28,177] The errors in the energies predicted using typical DFT exchange-correlation functionals and from the approximate descriptor-based linear scaling relationships get propagated to the kinetic models and impact the predicted reaction rates. Medford et al.[178] studied systematic error progression in the prediction of rates of ammonia synthesis over Fe and Ru catalysts and reported errors in turnover frequencies (TOFs) in the range of an order of magnitude over Fe and almost two orders of magnitude over Ru. Also, large non-trivial uncertainties in the TOFs were observed at lower temperatures in their Arrhenius plot, as the rates are more temperature sensitive to low temperatures. Furthermore, they showed that the extent of uncertainties in the predicted TOFs was more for inactive metals such as Pd and Ni compared to that on active Fe and Ru catalysts.[178]

Wang et al.[179] calculated the energies of reaction intermediates with an ensemble of 2000 functionals around the BEEF-vdW functional and used them for detailed analyses of uncertainties in reaction pathways via quantifying reaction rates/TOFs, surface coverages, and degree of rate control (DRC) analyses for dry reforming of methane (DRM) over Ni(111) and Pt(111) catalysts. Out of the 98 and 83 types of reaction pathways (developed without explicitly considering any reaction rules) over Ni(111) and Pt(111), respectively, three main reaction pathways (denoted as A, B, and C) were chosen for kinetic analyses (microkinetic modeling), based on their high frequencies of occurrence. The frequency of occurrence was considered as a pruning parameter that was defined based on the confidence of a DRM pathway being dominant within the temperature range of 823 K–1023 K. This BEEF-vdW-assisted reaction pathway pruning (reduction) based on reaction rates revealed that pathway A was more favorable over Ni(111) while B and C were more favorable over Pt(111).[179] Furthermore, the transition states for CHO* formation and CO_2 dissociation over Ni(111), and those of CH-OH* formation and OOH* dissociation were found to be more frequent (DRC analyses). Finally, the uncertainty quantification of TOFs revealed that the reduced DRM pathways were distributed between 10^{-8} and 10^{3} site^{-1}s^{-1} over Ni(111) and between 10^{-13} and $10^{2.5}$ site^{-1}s^{-1} over Pt(111).[179]

Estimating and addressing uncertainties in the parameters estimated through semi-empirical methods

Often choice of descriptor and its reliability causes uncertainty to acquire a reasonable estimation of energetics using LSR. Typically, the RMSE and the regression coefficients were considered to check the reliability of the descriptor.[168] In some cases, many descriptors may appear to be relevant and identifying the most reliable one itself may be a challenge. Using a data-driven approach, Wenbin et al.[180] identified potential descriptors for the adsorption enthalpies of oxygen during oxygen evolution reaction on doped RuO_2 and IrO_2 catalysts. They considered 31 potential descriptors and classified them as geometric and electronic properties of the surface among which a few exhibited more than 90% correlation with others and were eliminated from further analysis. Screened features were further analyzed with the sure independence screening and sparsifying operator method[181] to identify the most relevant descriptors.

Ulissi et al.[182] examined the uncertainty in predictions of catalytic activity during NH_3 decomposition over metal catalysts by manual perturbation of kinetic model parameters such as pre-exponential factor and the adsorbate bond indices (used in UBI-QEP equations). The motivation was to study the impact of an explicit assumption of a pre-exponential factor of 10^{13} for all elementary steps on the catalytic activity descriptors, such as the binding energies of atomic H (Q_H) and N (Q_N). An order of magnitude perturbation of pre-exponential factors brought about $\pm$ 0.9 kcal/mol change in Q_H and $\pm$ 2.9 kcal/mol change in Q_N; while a 25% change in the bond indices of a reaction pair brought about $\pm$ 0.7 and 3.1 kcal/mol changes in Q_H and Q_N, respectively. Furthermore, they also examined the impact of pre-exponential perturbation on the surface coverage-dependent binding energies as adsorbate-adsorbate lateral interactions at surface coverages of zero and 1/9 ML on Ni, Pt, and two monolayer bimetallics (Pt-Ni-Pt and Ni-Pt-Pt) as this is one of the important kinetic model parameters. At higher coverage, a significant shift of $\sim \pm$ 25 kCal/mol was observed in Q_H and Q_N, and Ni-Pt-Pt bimetallic was reported as the optimal catalyst as it offered the most favorable Q_H and Q_N. This observation was in good agreement with experiments, unlike the case at zero coverage where pure Ni and Pt were regarded as the optimal catalysts, as they offered better Q_H and Q_N compared to that of the bimetallics.

Summary of reaction mechanism generation and current approaches in kinetic parameter estimation

Typical methodologies adopted in the literature for estimating kinetic parameters are summarized in Figure 5. For relatively small reactant species undergoing simple reaction mechanisms on simple catalysts, the full reaction network may be manually built. In the absence of inputs from existing spectroscopic information and prior mechanistic studies, this task is heavily reliant on chemical intuition and experience of the investigator. However, for complex reaction networks of larger reactant species, an appropriate reaction mechanism generator may be used. The active site heterogeneity on the catalyst surface may necessitate multi-mechanistic investigations and the impact of coverage effects on reaction mechanism and energetics may require iterative studies. Semi-empirical methods such as the UBI-QEP which is computationally inexpensive have given reasonable estimates of activation barriers in conditions such as reactions of small molecules on 3d transition metals at relatively low pressure (low coverage). This can be used as a tool to obtain quick preliminary estimates of reaction energetics in a complex reaction network. When UBI-QEP estimates

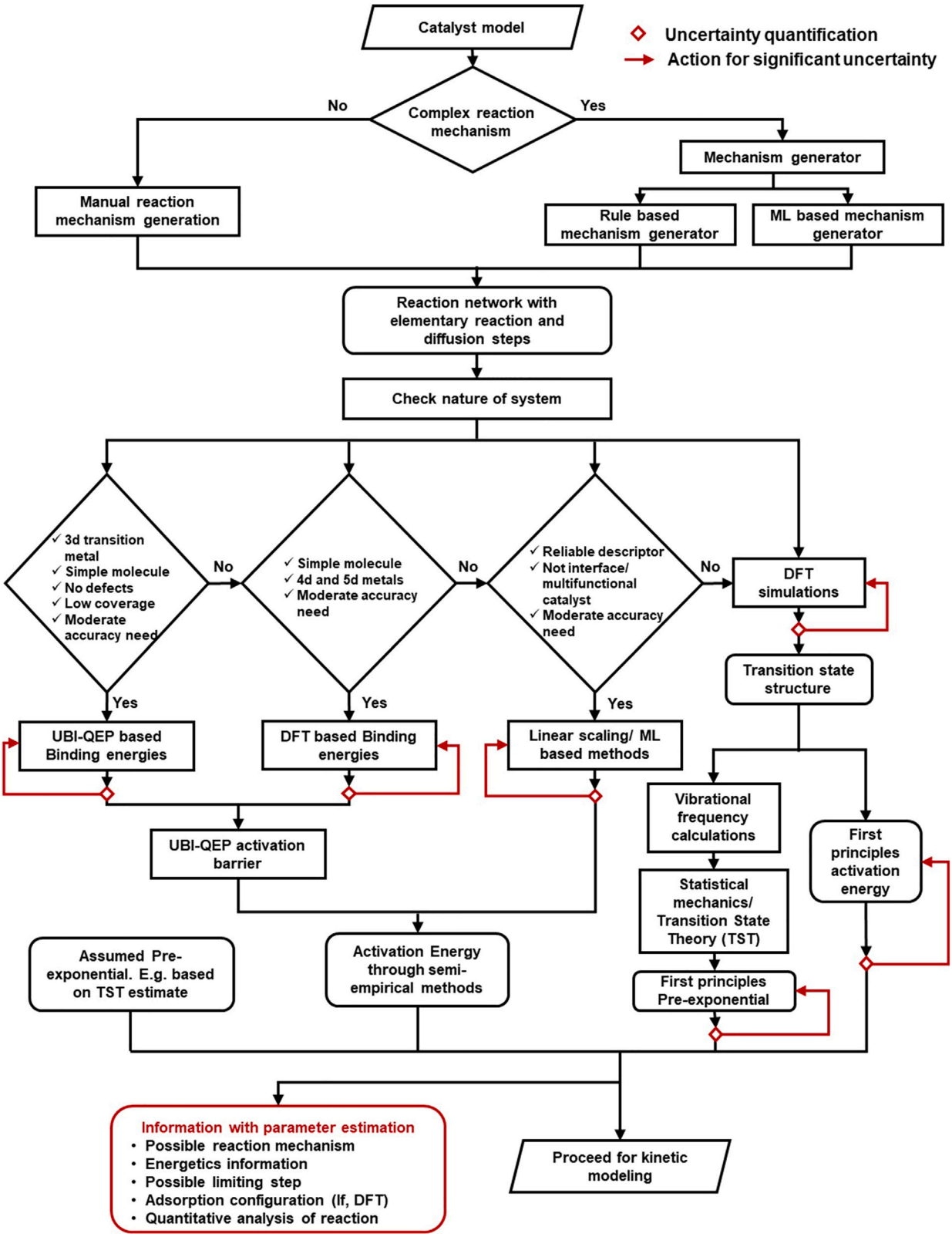

Figure 5. Proposed hierarchical workflow and decision trees for kinetic parameter estimation

are not sufficiently accurate, DFT calculated BEs may be combined with the UBI-QEP approach for improving the accuracy, especially for 4d and 5d transition metals. Kinetic linear scaling relations on appropriate descriptors have been established for a wide range of species/elementary steps for different catalyst classes. These can also be used for quick preliminary estimates of kinetic parameters, especially for complex reaction networks. However, these need to be used with caution, as the deviations in predictions can be significant for cases where the geometric and electronic features of the catalytic sites or the nature of species interactions on the surface are even marginally different from those for which the LSR was originally derived. For complex multi-component catalysts with reaction steps at the interfacial sites or for catalysts containing dopants, vacancies, and defects, the UBI-QEP/LSRs typically fail and direct DFT-based analyses are recommended. DFT analyses are also recommended when the surface species coverage effects are prominent or for the cases where high accuracy is needed. Ascertaining the fidelity of the computed kinetic parameters and quantifying the parametric uncertainties by using error estimating functionals, using data-centered approaches, or using experimental data injection are critical in ensuring reliable predictions by the kinetic models. Hierarchical approaches utilizing a judicious choice of lower rung semi-empirical methods with higher accuracy DFT techniques and iterative refinement of the parameters is a promising approach drawing compromises on accuracy and computational cost. For more insights and details on the methods and techniques for accurate estimation of reaction kinetics parameters, the readers are directed to the review article by Sebastian et al.[183]

KINETIC MODELING

A kinetic model represents the complete catalytic cycle and describes the temporal evolution of reactants, products, and various reaction intermediates during a chemical reaction. These models can predict macroscopic observables/measurables such as reactant conversions, turnover frequencies, product selectivity, and reaction orders at a wide range of conditions which can be validated with the existing experimental reports. A kinetic model built by coupling all the interconnected elementary surface reaction steps and capable of giving the overall reaction rate is typically described as a microkinetic model. Analysis of the microkinetic model requires solving a coupled set of differential equations for an appropriate reactor type. Depending on the complexity of the reaction system under consideration, the kinetic model can be a simple computationally efficient mean-field-based model that describes the reaction kinetics without accounting for the spatial correlations between the reaction intermediates or relying on a computationally expensive stochastic model that accounts for such interactions. In this section, we briefly discuss the applications of these approaches to describe heterogeneous catalytic reaction systems.

Mean-field microkinetic modeling

The mean-field approximation inherently neglects the attractive or repulsive interactions of a surface species with its neighboring surface intermediates. Hence, a mean-field microkinetic model is a system of coupled rate expressions of elementary surface reaction steps representing a complete catalytic cycle, where the species under consideration in each expression is treated as if they are unperturbed by their neighbors on the surface. Here, the fractional surface coverage of the reactants/intermediates/products is analogous to concentration/pressure used in liquid/gas-phase modeling. Microkinetic modeling becomes a crucial link of catalyst microstructure to readily measurable quantities such as TOF, conversion, and selectivity[184–187] from experimental studies at specific conditions which enable development of catalyst structure-activity relations. Hence, this also serves as an indirect tool for validating the computationally estimated kinetic parameters of the elementary steps which is not straightforward and often cumbersome. For a more detailed description of the methods and applications, the readers are directed to the review article by Motagamwala et al.[17] and for more details on the first-principles microkinetic modeling techniques, readers are requested to refer to the review article by Albert et al.[28]

Since the rates of reactions are described by kinetic and thermodynamic parameters and the coupled reactions of the developed mechanism are under thermodynamic constraints, one of the critical factors for successful microkinetic model development is to ensure thermodynamic consistency in parameter estimation.[188–190] Consistency needs to be ensured both in the enthalpic and entropic contributions in the kinetic parameters. Inconsistencies in calculating the kinetic parameters of elementary steps of the mechanism will lead to inaccuracies in the predictions of the microkinetic model and lead to incorrect representation/prediction of the reaction equilibrium. For more details on thermodynamic consistency in microkinetic

Table 4. Widely used microkinetic analyses for heterogeneous catalysis applications with representative references where these have been used

Analysis	Reference
Sensitivity analysis	186
Degree of rate control (DRC) analysis	193
Partial equilibrium analysis	185
Most abundant reaction intermediate (MARI) analysis	210

modeling and strategies to ensure it, please refer to articles by Mhadeshwar et al.,[189] Gossler et al.,[190] and Stotz et al.[190]

A few typical analyses of mean-field microkinetic models are discussed in section 5.1.1. Selected examples of microkinetic analyses of large and complex reaction networks where hierarchical techniques are adopted with iterative refinements are described in section 5.1.2. Cases where the mean-field approximation breaks down and an explicit coverage need to be incorporated in the estimation of energetics and kinetic parameters or cases where the catalyst structures/active sites are different at operando conditions due to species coverage on the surface, necessitating a catalyst model refinement are discussed in section 5.1.3. Caveats of MKM analyses are discussed in section 5.1.4.

Mean-field microkinetic models: Typical analyses and outcomes

Among the outcomes of the microkinetic analysis is the fractional surface coverage of reaction intermediates and products which enable determination of the most abundant reaction intermediate (MARI). Such species may in some cases block a significant number of active sites, and hence this information is helpful in exploring strategies to maximize the active site availability for seamless progress of the reaction. While the mean-field approximation works well for quite a few chemistries and catalytic systems, in many others, the lateral repulsive/attractive interactions from the abundant surface species might have to be considered for more accurate kinetic parameter estimation and thereby kinetic analyses. Moreover, the fractional surface coverage by intermediates may also change the stability and morphology of the catalyst surfaces. Hence, microkinetic analyses can be utilized for iterative refinement of the catalyst structure and kinetic parameters of the elementary reaction steps. Examples of this kind from literature are discussed in section 5.1.3. Some common analyses with MKM are summarized in Table 4.

The sensitivity analysis performed on a microkinetic model sheds light on the kinetically relevant elementary steps in the model and the extent of impact of reaction conditions or the nature of the catalyst have on these steps and thereby the overall impact on reaction rate, conversion, product selectivity, and so on. This is crucial in identifying critical reaction tuning parameters, and more importantly in the search of alternative more efficient catalysts that impact selectively the identified reaction steps. Typically, through brute force local sensitivity analyses,[191] normalized sensitivity is calculated. Srinivas et al.[186] studied the effect of various reaction parameters on sulfur poisoning on Ni catalyst for steam reforming reaction of biogas containing H_2S and found the H_2S sticking on the catalyst surface to be the sensitive step for sulfur poisoning. Hence, they proposed and demonstrated using the model that high-temperature operation minimizes sulfur poisoning. Additionally, sensitivity analysis is often helpful in refining the accuracy of the large kinetic models developed by lower order computational techniques by selectively updating only the sensitive steps. An example of this type is discussed in section 5.1.2.

In a large reaction network, chnage in the energies of the intermediates and transition states of a few elementary steps could independently change the net reaction towards the product of interest. The DRC is a helpful reaction engineering tool used to identify those elementary steps. A detailed discussion on the DRC analysis, applications, and examples is presented in the viewpoint by Campbell and coworkers.[192] Verena et al.[193] analyzed propane combustion reaction on Pd(211) surface using DFT-based microkinetic model. Through DRC analysis, the penultimate dehydrogenation step was found to limit the reaction.[11] On Ni-based oxygen carrier, Yuan et al.[194] identified H_2 decomposition as the controlling step by comparing the barriers of elementary steps in the syngas combustion reactions.

The partial equilibrium ratio defined as the ratio of the forward reaction rate to that of the sum of the forward and backward reaction rates of the elementary reaction helps to identify reaction steps with net forward or reverse flux. Information on reactions in partial equilibrium can be used to simplify the reaction network. Aswathy et al.[185] performed the partial equilibrium analysis for CO_2 methanation on Ru catalyst at 410°C, 330°C, and 250°C and reported only 6 out of 21 elementary steps in the network to have a forward reaction flux at the lower temperatures. At higher temperature of 410°C, only 2 of them showed a net forward flux, implying lower formation of CH_4 and thereby lower CH_4 selectivity. The partial equilibrium analysis in this case gave insights on the desirable operation window for the methanation reaction.

Several examples of microkinetic analyses are reported in the literature where the focus was to understand the surface reaction kinetics of various chemical reactions like methane oxidation and reforming, CO and CO_2 methanation, methanol synthesis from syngas over various catalysts, and so on. Delgado et al.[195] investigated the reaction mechanism of methane oxidation and reforming, comprising 52 elementary reactions over Ni catalyst. They developed the kinetic model after numerical adjustments to maintain thermodynamic consistency. The model predicted the product distribution for both partial oxidation and reforming of methane reactions in a temperature range of 373 K–1123 K. The modified kinetic model also predicted the effect of co-feed of products such as H_2 and H_2O which predicted a decrease in the surface carbon concentrations under reaction conditions which was in agreement with the experimental results that showed H_2 and H_2O to be coke deposition inhibitors. Zhu et al.[196] investigated the methane steam reforming reaction over Pt and Rh catalysts and their microkinetic results showed Rh to be more active than Pt catalyst due to the lower activation energy barrier for the rate controlling dissociative CH_4 adsorption step on Rh than Pt. They found that low-coordinated Rh atoms are the active sites for the dissociative methane adsorption. Schmider et al.[197] developed the microkinetic model for both CO and CO_2 methanation reactions over Ni-based catalysts.[197] The developed kinetic model was able to describe the methanation reaction for a wide range of temperature conditions, catalyst loadings, support materials, and reactant ratios. Campos et al.[198] developed a three-site microkinetic model for the methanol synthesis from syngas ($CO/CO_2/H_2$) over Cu(211) and Cu/Zn(211) catalysts. They considered the three sites as (i) Cu(211), (ii) fully Zn-covered Cu(211), and (iii) a separate Cu(211) site for H_2 and H_2O adsorption. They validated the microkinetic model with a total of 359 experimental data points from their own experiments and using those in the literature. However, their model showed discrepancies in predictions at certain conditions such as a combination of low temperature, high CO_2/CO concentration in the feed, and high values of gas hourly space velocities (GHSV). De Carvalho et al.[199] developed a microkinetic model for the WGS reaction on Ni-based catalysts which was able to predict the experimentally measured CO_2 conversion and apparent activation energies. Furthermore, they reduced the microkinetic model to a global one-step expression for the WGS reaction on Ni which was able to predict CO conversion, apparent activation energy for different conditions such as inlet compositions, volumetric flow rates, and temperatures. Sterk et al.[200] investigated the Sabatier reaction (CO_2 to CH_4) with DFT-based MKM on Ni(111), Ni(110), Ni(100), and Ni(211) facets. Among the different facets, Ni(110) surface was shown to be the most effective for the reaction. By analyzing the fractional contribution of the four nickel facets on Wulff-constructed metal nanoparticles of 1–9.5 nm, they predicted that the Ni nanoparticles sized around 1.7 nm to be catalytically most active. First-principles-based transient microkinetic study of CO hydrogenation on Co catalysts by Bart et al.[201] revealed that initially on the empty surface, the CO scission is the fastest step, which at a steady-state surface coverage is in competition with the CH_x formation (CO hydrogenation) steps. Minimizing the CH_x coverage was found to favor CO desorption and hence it was proposed that CO reabsorbs or diffuses to the edge sites to get activated for hydrogenation. Dharmalingam et al.[202] investigated the multifunctional $Cu/ZnO/ZrO_2/Al_2O_3$ catalyst for CO_2 reduction to methanol and showed that the thermodynamically consistent DFT-based reactor scale microkinetic model predicted the CO_2 conversion and product flow rates at the desired reaction temperature, GHSV, and inlet feed compositions. They also showed that the ZnO/Cu interface is active center for the reactions even though DFT calculations showed ZrO_2/Cu interface as favorable CO_2 adsorption sites which highlighted the importance of the coupling of DFT simulations with thermodynamically consistent reactor scale microkinetic models. The applicability of the MKM technique and analyses is not limited to heterogeneous thermocatalytic applications, such MKM techniques have been used in electrolysis cell applications[203,204] as well.

Hierarchical iterative microkinetic modeling of large reaction networks

The computational expense of the DFT-based kinetic modeling makes its application to large reaction networks prohibitive and hence alternative approaches, drawing a balance between computational expense

and accuracy, are desirable. One of the possibilities in this direction is to use a combination of semi-empirical and DFT-derived kinetic parameters, where the semi-empirical methods are applied to estimate parameters for the non-sensitive reaction steps while DFT data are used for the sensitive steps. This can typically be accomplished in a hierarchical approach where semi-empirical methods are used for the preliminary model development which is refined using DFT data for the sensitive reaction steps.

Yalan et al.[205] developed a microkinetic model for methane steam reforming over Co(0001), Ni(111), Ru(0001), Fe(110), Cu(111), Pd(111), Rh(111), and Pt(111) catalysts using UBI-QEP parameters and identified the sensitive reaction step. They then updated the parameters of the sensitive steps with DFT-calculated parameters. This approach led to an improvement in the accuracy of the prediction of the TOF of H_2 production which now was comparable to a pure DFT-based model, but with lower computational cost. Similarly, Jonathan et al.[206] adopted a hierarchical approach and updated the sensitive steps of the kinetic model for ethanol steam reforming on Pt catalyst which was developed using BEP relation with first-principles parameters. Analyses of the outcomes of the "pathways analysis" using the updated and the pure DFT models showed comparable accuracy of the updated semi-empirical model at minimal computational expense. This approach is believed to be a compromise in accuracy versus computational expense for larger reaction networks.

Surface species coverage and catalyst structure-dependent microkinetic modeling

UBI-QEP, LSRs, and typical DFT-based kinetic parameters are obtained at conditions representing vacuum or very low gas pressures which are reflected by the sparse population of the reactive species on the catalyst surface. The kinetic models based on these conventional approaches for reactions that are carried out at high pressures or with species that have very strong interactions with the catalyst are highly likely to be making erroneous predictions due to the "pressure gap" or "coverage gap." Hence, the effect of species coverage on the surface can be explicitly accounted for by appropriate corrections and species coverage containing simulation models. But it also needs to be kept in mind that the species coverage contributes to changing the catalyst surface structure and hence might need iterative model development and parameter estimation for accurate kinetic modeling.

Bhandari et al.[207] investigated formic acid decomposition over Pt catalyst and found that the DFT-calculated energetics over clean Pt(100) and Pt(111) surfaces required parameter adjustments to predict the experimentally determined apparent activation energies. Pt(100) was found to be poisoned by CO* under typical reaction conditions; hence, the CO coverage in the model was updated to 4/9 ML, the energetics were recalculated, and the MKM was updated. The updated MKM with the spectator CO*-assisted energetics were able to reproduce the experimentally measured apparent activation energies.[207] Ding et al.[208] compared the TOF predicted by the kinetic model for NO oxidation on Pt catalyst developed using BEP relations and the experimentally determined TOF for the same reaction and reported two orders of magnitude error in the TOF predicted by the semi-empirical kinetic model. The impact of lateral interactions of surface species on the adsorption energy and the energy of the TS was identified as the reason for the error. Their rigorous coverage-dependent model that accounted for the impact of lateral interactions on the TS and the adsorbates gave a TOF which was in excellent agreement with the experimental result.

Ground state or equilibrium catalyst states and thereby the active sites typically used for reaction modeling may not always accurately predict kinetics of reactions due to the existence of metastable catalyst structures with a different kind of active site distribution. Using the Gibbs free energy of formation of different Rh nanoparticle shapes under typical CO dissociation conditions (including explicit CO coverage) and by applying a Boltzmann probabilistic distribution at such conditions, Raffaele et al.[209] generated an ensemble of equilibrium and metastable nanoparticles of different shapes and sizes. The ensemble of nanoparticles had a distribution of active sites which had a higher percentage of the more active sites for CO dissociation than the equilibrium shape. This resulted in a two order of magnitude higher reaction rate predicted by the kinetic model with the ensemble of nanoparticles than the kinetic model with the ground state catalyst structure.

Through the MARI analysis of WGS reaction on Rh surface, Raffaele et al.[210] identified CO as the dominant surface species. With CO adsorbed on the surface, they performed Wulff construction by considering the reaction condition and CO coverage, to identify the equilibrium nanoparticle shape under reaction conditions. Rh(100), Rh(110), Rh(111), Rh(311), and Rh(331) facets with corresponding reaction energetics were

included in MKM; the Rh(111) facet was found to be the dominant active surface during WGS, whereas, the Rh(100) facet dominated the reverse WGS reaction rate. This information of the operando active surface and sites is typically missed while modeling only the stable surfaces.

Uncertainty quantification in MKM and machine learning tools to address them

Limitations in the catalyst model,[209] missing contributions of less stable facets,[211] choices of sticking coefficient,[212] break down of the mean-field approximation,[213] and neglect of multi-site reaction mechanisms are some of the factors which will lead to inaccurate predictions by the microkinetic model. Moreover, thermodynamic inconsistency can result in substantial deviations in the predictions of equilibrium-limited reactions and lead to unreliable kinetic models.[213] Although the surface reaction networks analyzed with DFT simulations are typically expected to satisfy the thermodynamic constraints, it must also ensure consistency in the overall catalytic cycle[188] and this needs to be thoroughly verified. To ensure thermodynamic consistency,[198] and many a time to obtain a better fit of model prediction against experimental data, kinetic parameter optimization is unavoidable. The reliability of such models needs to be ascertained by validating against experimental data obtained for a wide range of operating conditions.[184,185]

Machine learning tools can be handy in dealing with the typical inaccuracies present in computed enthalpies and entropies, uncertainty quantification, and other challenges such as multisite mechanisms and surface diffusion in MKM.[214,215] Rangarajan et al.,[214] in their review, highlighted some of the recent studies that have used NN or Gaussian process surrogate models with DFT-aided training to address the fast diffusion approximations made in mean-field-based microkinetic modeling. They also discussed techniques such as polynomial chaos expansion with DFT-aided training to address corrections to enthalpies and entropies. The authors suggested the use of multi-fidelity data such as those generated from combining the data obtained from simulations using the generalized gradient approximation, random phase approximation, and experiments for training the NN models.[214] In order to account for lateral interactions between the adsorbates and the presence of surface heterogeneities, Tian et al.[215] proposed an ANN-based model using a training set of macroscopic reaction rates obtained from the lattice Monte Carlo simulations. The model was applied for the CO oxidation reaction and the predicted reaction rates (ensemble averaged) showed RMSEs of $3.4x10^{-2}$, $7.0x10^{-3}$, and $3.4x10^{-3}$ for the oxidation, adsorption, and desorption processes, respectively, compared to those obtained via full-fledged KMC simulations. To render the microkinetic models globally predictive, one study[216] proposed a model refinement strategy using data injection from a set of experiments.

Wittreich et al.[217] proposed a framework based on group additivity (GA) with the incorporation of variances and correlations in the groups (of GA), for estimating uncertainties in reaction enthalpies and entropies in DFT-parameterized MKM models. This framework did not rely on analyses of energies with multiple exchange-correlation functionals or the use of BEEF-vdW functional. The authors used the framework to explain the impact of these uncertainties in the predicted conversion, reaction rates, selectivity, thermochemistry, reaction orders, and reaction barriers during propane complete oxidation and ethane oxidative dehydrogenation. Reaction orders showed a lower extent of variance (1 σ of $\pm$ 0.06–0.03) compared to the other model parameters such as activation energy barriers (1σ of 4–6 kcal/mol) and TOFs (1 σ of $\pm$ 1 order of magnitude).[217] Various other studies aimed at improving the predictive capabilities of first-principles-based microkinetic models have proposed different methodologies, such as use of data-driven models involving the Gaussian process (based on Bayesian statistics),[159] adaptive sparse grid approach,[218] and correlative global sensitivity analyses[219] based on Monte Carlo methods for uncertainty quantification.

Stochastic kinetic modeling by the KMC approach

The mean-field approximation-based approaches such as MKM predict overall observable rates, conversion, and selectivity trends without emphasizing on the spatial ordering of the adsorbates, site heterogeneity, and slow diffusion of surface species, which are quite common in heterogeneous catalysis.[220,221] Although some of these can be accounted as corrections in MKM, significant deviations can be expected due to violation of the law of mass action[222,223] unless explicitly incorporated. Under the KMC framework and Markovian master formalism, surface kinetics is described in a comprehensive manner by incorporating short- and long-range, many-body interactions between surface species and slow diffusion across surface sites.[224–229] Unlike the mean-field-approximated models, KMC uses a stochastic approach to study the adsorption/desorption, surface reaction, and diffusion events as state transitions, wherein each state is independent of its history. The probability of a transition is an exponential distribution based on its rate

constant, and the temporal evolution of these probabilities is captured by the Markovian master equation. For more details of the formulations, we recommend the readers to refer to other reviews[4,40,224] on this method and its applications. Many studies have used KMC in conjugation with DFT simulations to obtain steady-state surface coverages of dominant reactant/product species, and thereby accurate prediction of reaction orders and product selectivities.[52,227,229–232] However, one of the major drawbacks of KMC simulations is the high computational cost associated with it.

Typical applications and analyses of KMC simulations and uncertainty quantification and alternatives to KMC

Mei et al.,[229] in their kinetic studies of selective hydrogenation of acetylene to ethylene over the Pd(111) surface, highlighted the importance of the lateral interactions on the reaction kinetics. They established the same by studying the acetylene hydrogenation kinetics with and without the inclusion of lateral interactions (long-range interactions and from the surface interactions). Neglection of lateral interactions resulted in the prediction of stronger binding energies and overestimation of surface coverages of acetylene and hydrogen compared to those obtained in the presence of lateral interactions, as the nature of interactions was repulsive. This resulted in an underestimation of ethylene selectivity (7%–19%) as the kinetic model predicted the facile hydrogenation of acetylene and ethylene. However, the simulations that included lateral repulsion predicted much higher ethylene selectivity (79%–86%) due to low H coverage. Furthermore, the reaction order with respect to acetylene was also underestimated significantly (−0.95) in the absence of lateral interactions compared to the simulations that incorporated them (−0.52). Hence, the inclusion of lateral interactions played a major role in predicting the right overall selectivity of ethylene and reaction orders (−0.52 for C_2H_2 and 1.16 for H_2) that agreed well with the experimentally reported orders (−0.66 for C_2H_2 and 1.04 for H_2[233]). Yue et al.[234] analyzed the effect of CO and H_2 coverage for syngas combustion reaction on NiO catalyst using DFT-based KMC simulation. The results suggested that CO coverage limited the CO oxidation by weakening the CO adsorption and H_2 coverage promoted the CO oxidation. Using DFT-based KMC simulations, Jingde et al.[235] studied carbon nanotubes (CNT) growth mechanism on Ni(111) surface by considering CH_4 as a raw material. The sensitivity analysis revealed that the CNT growth was sensitive to the CH_4 dissociative adsorption step.

Cluster expansion Hamiltonian methods have also been employed to accurately capture the many body lateral interactions and hence precise adlayer description.[228,230] Papanikolaou et al.[230] considered different CE models with varying complexities from first nearest neighbor (1NN) interactions to 2NN and 3NN interactions. They incorporated them in modeling NO oxidation over Pt(111) surface and highlighted the importance of long-range interactions (between third nearest neighbors) in accurately describing the surface coverages of various adsorbates and kinetics. Piccinin et al.,[231] in their kinetic studies of CO oxidation over Pd (111), reported good agreement between the KMC predicted and experimentally reported catalyst performance. The one, two, and three body lateral interactions were incorporated to capture accurate surface coverages of CO and O adsorbates, which predicted accurately the reaction orders (0.51 and 0.8 with respect to O and CO, respectively) that matched well with the experimentally reported observations.[231]

Since KMC simulations are high-dimensional and stochastic, they lack closed-form expressions to relate the final desired output parameters (rate, coverage, etc.) with the model parameters, which pose a difficulty in the optimization and control of the multiscale model.[236,237] To address this, Rasoulian et al.[236] proposed a methodology to examine the uncertainty in the KMC-predicted microscopic rates for the adsorption, desorption, and migration steps during a general epitaxial thin-film growth process, using high-order power series (PSEs) to obtain probabilistic bounds in the various microscopic rates. Their methodology involved four steps: specification of the probability distribution function for the identified uncertainty parameter(s), evaluation of sensitivities in the microscopic rates with respect to the uncertain parameters, estimation of the probability distribution function of the rates using the PSEs (truncated to required orders by comparing the results against that obtained from brute-force KMC simulations), and calculation of upper and lower bounds of rates for a specific confidence level.[236] The authors described the uncertainty in the bulk gas-phase mole fraction using a first-order PSE, while for describing the uncertainty in some of the energetics (describing the bond strength and ease of desorption), they used second-order PSEs. They also showed that the PSE-based method of uncertainty quantification was at least two orders of magnitude less computationally intensive than the brute-force KMC methods.[236]

Chaffart et al.[237] proposed a computationally less intense, hybrid multiscale model consisting of continuum partial differential equations, stochastic partial differential equations (SPDEs), and ANNs as an alternative to the KMC method to describe a general thin-layer growth process. The film thickness was modeled using fourth-order SPDE, where its coefficients were predicted using the ANN, and the uncertainties in the model parameters, such as bulk gas (precursor) mole fraction and activation energies for surface reaction, migration, and desorption steps, were quantified to study their impact on the film thickness and hence the growth process.[237]

MKM vs. KMC and caveats of KMC modeling and analyses

Li et al.[41] compared the merits and demerits of MKM and KMC simulations with the inclusion of lateral interactions for a simple case study of CO oxidation reaction over Pt(111). It was reported that in the absence of slow diffusion processes, the computational costs incurred by KMC simulation was three orders of magnitude higher than that incurred by MKM simulations upon inclusion of lateral interaction effects via an iteratively solved coverage-dependent modeling (described by Grabow et al.[238]) and explicitly defined pairwise interaction energies in KMC. Although a similar homogeneous rate-determining step was predicted by both methods at high and low coverages, the TOF information obtained in each method differed by one order of magnitude.[41] Hence, the use of KMC and the associated high computational costs are justified for describing the kinetics and spatial coverages over catalyst surfaces with significant site heterogeneity, while for computational screening purposes and kinetic description over fairly homogeneous surfaces, MKM is sufficient.

KMC-based kinetic modeling offers an accurate means of capturing the temporal evolution of catalyst surfaces and adsorbate coverages. Compared to mean-field MKM models, these are very rigorous and computationally expensive as the input is completely first-principles based via DFT and the description of short- and long-range lateral interactions between surface adsorbates is quite tedious. Moreover, KMC simulations are very sensitive to the energy input and lattice sizes and require carefully simulated DFT results. Hence, KMC models need to be employed mindfully depending on the nature of the catalyst and reaction system in consideration.

SUMMARY AND PERSPECTIVES ON HIERARCHICAL AND ITERATIVE CATALYST STRUCTURE MODELING, REACTION MECHANISM ANALYSIS, AND DETAILED KINETIC MODELING

The accuracy of a hierarchical kinetic model and its predictive capabilities are contingent on the appropriateness of the chosen catalyst surface and the active sites, unambiguous establishment of the reaction mechanism, inclusion of all kinetically relevant elementary steps in the kinetic model, accuracy of the computational methods used to estimate the kinetic parameters, the choice of the kinetic modeling framework, and the validity of the assumptions made in the modeling.[157] Each of these are rungs on the multiscale ladder linking atomistic to reactor scale phenomena. Errors and uncertainties in any one of these rungs can result in predictions which can in some cases result in only minor inconsistencies at certain specific input reaction conditions[42,231] while in other cases can be completely unreliable across all reaction conditions.[179,239] Parity of the model predicted and experimental data pertaining to activity and selectivity and other observables are ultimate goals and there have been substantial efforts in this direction.[20] Unfortunately, the trade-off between accuracy and computational costs associated with generating validated output at each rung remains substantially high. In the last few years, developments in DFT methods have demonstrated the potential to reach chemical accuracy in specific cases[4] and thus have become the workhorse for catalyst structure and active site prediction, reaction mechanism and pathway analyses, and kinetic parameters estimation for detailed kinetic modeling. Such applications have become routine for simple metal catalysts and small molecules but remains a challenge for complex multicomponent nanostructured catalysts and for cases with large or complex reaction networks. These situations typically necessitate lower-order hierarchical treatment of the multiscale problem together with iterative refinements at different stages. Machine learning tools and data-driven approaches are now routinely included in the overall workflow for reaction mechanism identification/refinements, kinetic parameter estimation/refinement, and reaction kinetics analyses/refinement. These form the crux of this review article and some of these hierarchical multi-technique approaches and the potential refinement avenues have been discussed. The next paragraph is a summary of the unidirectional multiscale workflow from catalyst structure

predictions to kinetic modeling. In many cases, each of these elements are dealt with independently without integrating across the entire spatiotemporal scales.

For simple metal and metal oxide catalysts, equilibrium catalyst nanoparticle morphologies under typical reaction conditions and the surface state/phase under the influence of the reaction environment can be predicted using *ab initio* thermodynamics-based calculations. These enable identification of the catalytically relevant surface exposures of catalyst nanoparticles with specific atomic arrangement on such surfaces, including adsorbates and reactive gas species, with a caveat that these are equilibrium predictions with no accounting for kinetically controlled catalyst growth conditions. Understanding the catalyst nanoparticle surface dynamics and fluxionality of isomeric nanoparticles are also critical aspects in identifying operando catalyst surface sites for reaction investigations. Based on the homogeneity/heterogeneity of active sites, decision on mechanistic investigations is made. Also useful for this decision-making at this stage is information regarding the potential for significant impact of surface coverage of species and thereby lateral interactions on the calculated energetics/kinetic parameters or the possibility of slow surface diffusion steps involving prominent surface species. In the absence of significant heterogeneity of active sites, significant lateral interactions among surface species, and slow surface diffusion steps, the mean-field MKM is a powerful technique for kinetic modeling. When these factors are critical, then accurate modeling requires transitioning to the more expensive KMC framework. Both these detailed kinetic modeling frameworks require the complete reaction network of all possible elementary reaction steps in the catalytic cycle and for each of these steps. Building the complete reaction network from chemical intuition and *in situ* spectroscopic inputs appear relatively straightforward for reactions of small molecules on simple catalysts. Personnel and personal biases can impact the accuracy and completeness of such networks. But for larger molecules and for complex reaction networks, automatic reaction mechanism generators that work with simple input such as reactants, products, reaction rules, lumping rules, and so on become extremely handy. The kinetic parameters that allow the determination of a reaction rate constant for each of these elementary steps include the activation energy barriers and the pre-exponential factors. While activation barriers can be estimated using semi-empirical techniques such as the UBI-QEP method or use of kinetic scaling relations or by explicit first-principles (typically DFT) calculation, the pre-exponentials are best obtained from first-principles coupled statistical mechanics calculations. As kinetic parameter estimating is at the heart of the kinetic modeling activity, the choice of methods and the associated accuracy are critical for reliability of the model predictions. Depending on the nature of the system and availability of the pre-built scaling relations, preliminary estimates of activation barriers can be purely semi-empirical or a hybrid of DFT-semi-empirical. Complex catalysts and complex interactions of the species with the catalyst may necessitate the exclusive use of first-principles calculations. The KMC framework additionally requires the pairwise lateral interaction energies of the species, the kinetic parameters for the diffusion steps, and kinetic parameters for multi-site reaction mechanisms. The system of coupled species balance equations in the mean-field MKM can be solved for a specific reactor type to get information on spatiotemporal species evolution, overall reaction rates, product selectivity, and so on at specific input reaction conditions. These then serve as the crucial link between catalyst microstructure and catalyst performance enabling the formulation of structure activity relations.

The above unidirectional workflow from catalyst model and active site prediction to reaction mechanism analysis to kinetic parameter estimation and finally kinetic modeling may lead to unsatisfactory predictions and unreliable models due to the suite of approximations made at each of these levels for computational tractability. Reliability and applicability can be built into this workflow by cross validations with available experimental data at each stage and by adopting an iterative approach with iterations at each level or for the entire workflow as shown in Figure 6.

Establishing and validating the specific active sites on the surface structures identified from thermodynamic calculations can be done indirectly using computational spectroscopic techniques such IR, XPS, Raman, and directly using imaging techniques such as computational STEM analyses. If the computational predictions of catalyst structure do not match the experimental data, either the kinetic factors in catalyst synthesis overshadow the thermodynamic formation of the catalyst nanoparticles or the information on the reaction environment is incorrectly captured in the thermodynamic surface modeling and needs refinement. Such an iterative workflow for the operando catalyst structure prediction and validation was presented in Figure 4.

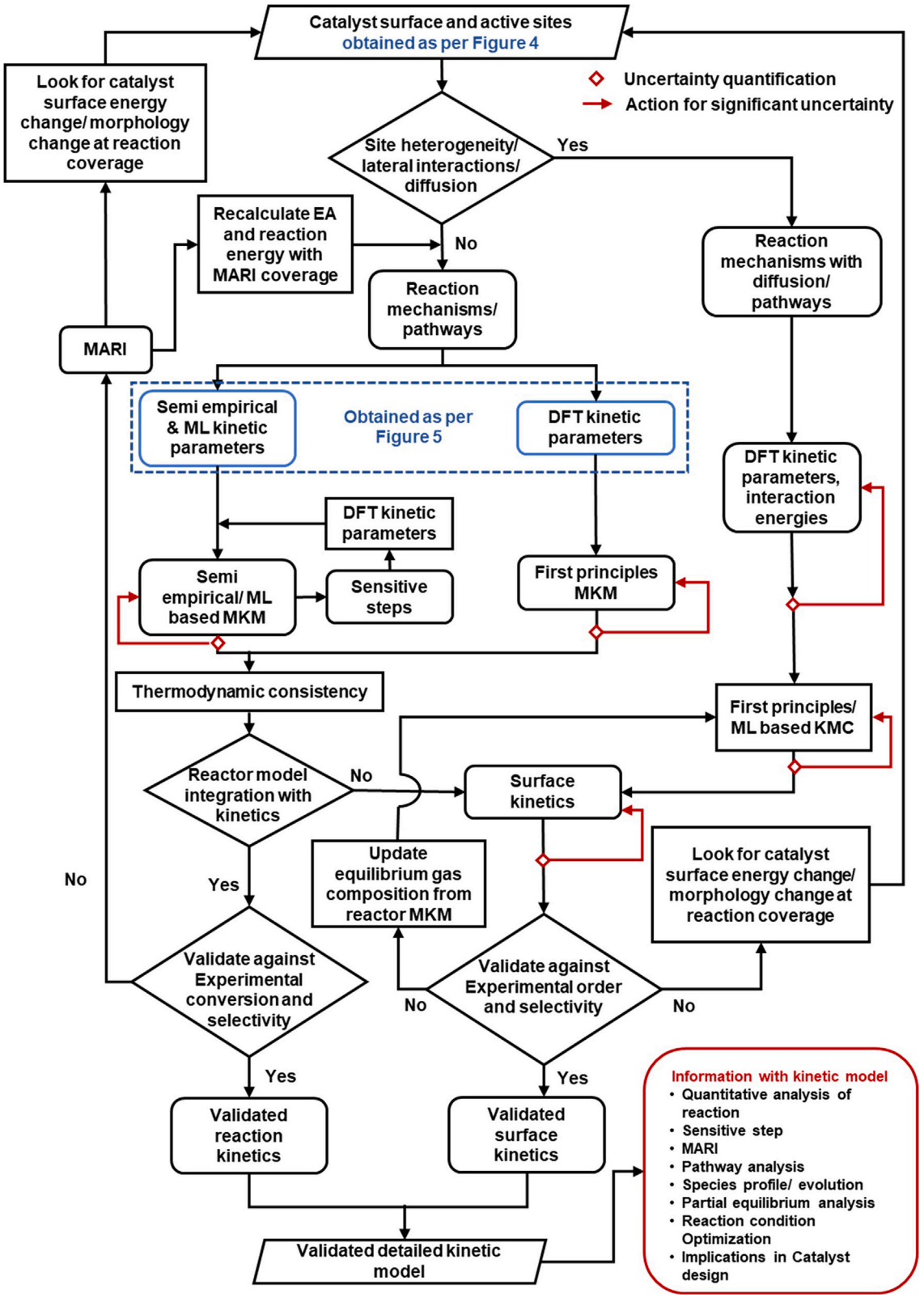

Figure 6. Proposed bottom-up iterative workflow for first-principles kinetic modeling and validation

For complex reaction mechanisms analyzed on validated computational catalyst models, a screening kinetic model can be built with preliminary estimates of kinetic parameters using semi-empirical or a hybrid approach of semi-empirical and DFT simulations. Refining of the model can be based on identification of the sensitive reaction steps that result in an overall impact on the reaction outcomes or using machine learning and data-driven approaches. Kinetic parameters for such sensitive steps can be re-estimated based on purely first-principles techniques or even higher-order first-principles techniques (more accurate DFT functionals or methods beyond DFT where possible) and by application of machine learning approaches. Uncertainty analyses and quantification are essential checkpoints to be included in the workflow for kinetic parameter estimation and modeling. A workflow for kinetic parameter estimation with choices and decision trees was presented in Figure 5.

Special care to ensure thermodynamic consistency is required during the kinetic modeling process. If the kinetic model predicts substantial coverage of reaction intermediates on the surface, a shift from the mean-field MKM to KMC might be necessary. Alternatively, the energetics and kinetic parameters for the elementary steps can be re-estimated by explicit or implicit coverage effects incorporated and "coverage-dependent" MKM can be performed. If refining the MKM with the coverage dependence or switching to KMC still results in poor prediction, experimental observables/measures, the catalyst model, and the corresponding active sites themselves may have to be revisited as shown in Figure 6. Species coverage on the surface leads to substantially different equilibrium shapes of catalysts and thereby exposed surfaces and active sites. Coverages of species from the screening kinetic models can be utilized to refine the catalyst structure under typical reaction conditions using thermodynamic calculations. Semi-empirical methods would in this case be inappropriate for estimation of the kinetic parameters, unless modified exclusively for incorporating the coverage effects. Such a complete cycle iteration as shown in Figure 6 is desirable for a truly first-principles bottom-up workflow and is likely to be accurate for most of the systems under consideration. However, this is too computationally intensive and is typically not adopted. Since typical first-principles-based kinetic models are bound to give poor predictions of catalytic performance due to varying degrees of uncertainties introduced in different stages such as catalyst structure and active site prediction, DFT-prediction of energetics, parameter estimation via semi-empirical correlations, and kinetic modeling,[182,216,236,237] uncertainty estimation and rectification is recommended at multiple stages.

Machine learning is an emerging tool and is finding widespread application in computational catalysis and molecular scale phenomena.[240,241] The applications of these techniques have been across the various themes discussed in this article and a few examples include determination of catalytically active surfaces and active sites,[68,242] optimization of geometries of intermediates and transition state on the catalyst surface,[243] surface reaction network analyses and generation,[244,245] predicting reaction barriers,[246] assessing the surface lateral interactions between adsorbed species,[173] correcting the mean-field assumption[215] and improving predictive capabilities of MKM,[214] and so on. Demonstration of accuracy and applicability of the techniques to a variety of catalytic systems and chemistries, and availability of curated data amenable for use with these techniques remain areas of concern and hence are areas with immense future potential. Hence, a combination of first-principles computational multiscale simulations and data-driven approaches are the likely drivers for computation-aided catalyst discovery, reaction optimization, and reaction engineering in the immediate future.

ACKNOWLEDGMENTS

The authors acknowledge financial support from the Indian Institute of Technology Madras, Chennai, India. J.J.V. also acknowledges financial support in the form of Scientific and Useful Profound Research Advancement (SUPRA) grant (SPR/2020/000461) from the Science and Engineering Research Board India.

AUTHOR CONTRIBUTION

J.J.V. conceived the idea for the article and received input from A.R., A.P.P., and C.D.B. A.R., A.P.P., and C.D.B. wrote the manuscript in consultation with J.J.V. J.J.V. revised the manuscript.

DECLARATION OF INTERESTS

There are no competing interests to declare.

REFERENCES

1. Maestri, M. (2017). Escaping the trap of complication and complexity in multiscale microkinetic modeling of heterogeneous catalytic processes. Chem comm *53*, 10244–10254. https://doi.org/10.1039/c7cc05740g.

2. Wehinger, G.D., Ambrosetti, M., Cheula, R., Ding, Z.B., Isoz, M., Kreitz, B., Kuhlmann, K., Kutscherauer, M., Niyogi, K., Poissonnier, J., et al. (2022). Quo vadis multiscale modeling in reaction engineering? – a perspective. Chem. Eng. Res. Des. *184*, 39–58. https://doi.org/10.1016/j.cherd.2022.05.030.

3. Keil, F.J. (2018). Molecular modeling for reactor design. Annu. Rev. Chem. Biomol. Eng. *9*, 201–227. https://doi.org/10.1146/annurev-chembioeng-060817-084141.

4. Chen, B.W.J., Xu, L., and Mavrikakis, M. (2021). Computational methods in heterogeneous catalysis. Chem. Rev. *121*, 1007–1048. https://doi.org/10.1021/acs.chemrev.0c01060.

5. Zaera, F. (2021). In-situ and operando spectroscopies for the characterization of catalysts and of mechanisms of catalytic reactions. J. Catal. *404*, 900–910. https://doi.org/10.1016/j.jcat.2021.08.013.

6. Tao, F.F., and Salmeron, M. (2011). In situ studies of chemistry and structure of materials in reactive environments. Science *331*, 171–174. https://doi.org/10.1126/science.1197461.

7. Zhang, S., Nguyen, L., Zhu, Y., Zhan, S., Tsung, C.K.F., and Tao, F.F. (2013). In-situ studies of nanocatalysis. Acc. Chem. Res. *46*, 1731–1739. https://doi.org/10.1021/ar300245g.

8. (2015). Frontiers of in situ electron microscopy, *40*, H. Zheng, Y.S. Meng, and Y. Zhu, eds., pp. 12–18. https://doi.org/10.1557/mrs.2014.305.

9. Michely, T., and Comsa, G. (1991). Temperature dependence of the sputtering morphology of Pt(111). Surf. Sci. *256*, 217–226. https://doi.org/10.1016/0039-6028(91)90865-P.

10. Pham, T.H., Duan, X., Qian, G., Zhou, X., and Chen, D. (2014). CO activation pathways of fischer-tropsch synthesis on χ-Fe 5C2 (510): direct versus hydrogen-assisted CO dissociation. J. Phys. Chem. C *118*, 10170–10176. https://doi.org/10.1021/jp502225r.

11. Radin, M.D., Rodriguez, J.F., Tian, F., and Siegel, D.J. (2012). Lithium peroxide surfaces are metallic, while lithium oxide surfaces are not. J. Am. Chem. Soc. *134*, 1093–1103. https://doi.org/10.1021/ja208944x.

12. Wilson, H.F., Tang, C., and Barnard, A.S. (2016). Morphology of zinc oxide nanoparticles and nanowires: role of surface and edge energies. J. Phys. Chem. C *120*, 9498–9505. https://doi.org/10.1021/acs.jpcc.6b01479.

13. Zhu, B., Xu, Z., Wang, C., and Gao, Y. (2016). Shape evolution of metal nanoparticles in water vapor environment. Nano Lett. *16*, 2628–2632. https://doi.org/10.1021/acs.nanolett.6b00254.

14. De Vrieze, J.E., Gunasooriya, G.K.K., Thybaut, J.W., and Saeys, M. (2019). Operando computational catalysis: shape, structure, and coverage under reaction conditions. Curr. Opin. Chem. Eng. *23*, 85–91. https://doi.org/10.1016/j.coche.2019.03.003.

15. Varghese, J.J. (2019). Computational design of catalysts for bio-waste upgrading. Curr. Opin. Chem. Eng. *26*, 20–27. https://doi.org/10.1016/j.coche.2019.08.002.

16. Shi, X., Lin, X., Luo, R., Wu, S., Li, L., Zhao, Z.J., and Gong, J. (2021). Dynamics of heterogeneous catalytic processes at operando conditions. JACS Au *1*, 2100–2120. https://doi.org/10.1021/jacsau.1c00355.

17. Motagamwala, A.H., and Dumesic, J.A. (2021). Microkinetic modeling: a tool for rational catalyst design. Chem. Rev. *121*, 1049–1076. https://doi.org/10.1021/acs.chemrev.0c00394.

18. Gokhale, A.A., Kandoi, S., Greeley, J.P., Mavrikakis, M., and Dumesic, J.A. (2004). Molecular-level descriptions of surface chemistry in kinetic models using density functional theory. Chem. Eng. Sci. *59*, 4679–4691. https://doi.org/10.1016/j.ces.2004.09.038.

19. Medford, A.J., Vojvodic, A., Hummelshøj, J.S., Voss, J., Abild-Pedersen, F., Studt, F., Bligaard, T., Nilsson, A., and Nørskov, J.K. (2015). From the Sabatier principle to a predictive theory of transition-metal heterogeneous catalysis. J. Catal. *328*, 36–42. https://doi.org/10.1016/j.jcat.2014.12.033.

20. Xie, W., Xu, J., Chen, J., Wang, H., and Hu, P. (2022). Achieving theory-experiment parity for activity and selectivity in heterogeneous catalysis using microkinetic modeling. Acc. Chem. Res. *55*, 1237–1248. https://doi.org/10.1021/acs.accounts.2c00058.

21. Piccini, G., Lee, M.S., Yuk, S.F., Zhang, D., Collinge, G., Kollias, L., Nguyen, M.T., Glezakou, V.A., and Rousseau, R. (2022). Ab initio molecular dynamics with enhanced sampling in heterogeneous catalysis. Catal. Sci. Technol. *12*, 12–37. https://doi.org/10.1039/d1cy01329g.

22. Saleheen, M., and Heyden, A. (2018). Liquid-phase modeling in heterogeneous catalysis. ACS Catal. *8*, 2188–2194. https://doi.org/10.1021/acscatal.7b04367.

23. Varghese, J.J., and Mushrif, S.H. (2019). Origins of complex solvent effects on chemical reactivity and computational tools to investigate them: a review. React. Chem. Eng. *4*, 165–206. https://doi.org/10.1039/c8re00226f.

24. Ma, S., and Liu, Z.P. (2020). Machine learning for atomic simulation and activity prediction in heterogeneous catalysis: current status and future. ACS Catal. *10*, 13213–13226. https://doi.org/10.1021/acscatal.0c03472.

25. Cova, T.F.G., and Pais, A.C.C. (2019). Deep learning for deep chemistry: optimizing the prediction of chemical patterns. Front. Chem. *7*, 809. https://doi.org/10.3389/fchem.2019.00809.

26. Rangel-Martinez, D., Nigam, K.D.P., and Ricardez-Sandoval, L.A. (2021). Machine learning on sustainable energy: a review and outlook on renewable energy systems, catalysis, smart grid and energy storage. Chem. Eng. Res. Des. *174*, 414–441. https://doi.org/10.1016/j.cherd.2021.08.013.

27. Puliyanda, A., Srinivasan, K., Sivaramakrishnan, K., and Prasad, V. (2022). A review of automated and data-driven approaches for pathway determination and reaction monitoring in complex chemical systems. Digit. Chem. Eng. *2*, 100009. https://doi.org/10.1016/j.dche.2021.100009.

28. Bruix, A., Margraf, J.T., Andersen, M., and Reuter, K. (2019). First-principles-based multiscale modeling of heterogeneous catalysis. Nat. Catal. *2*, 659–670. https://doi.org/10.1038/s41929-019-0298-3.

29. Lansford, J.L., and Vlachos, D.G. (2020). Infrared spectroscopy data- and physics-driven machine learning for characterizing surface microstructure of complex materials. Nat. Commun. *11*, 1513–1612. https://doi.org/10.1038/s41467-020-15340-7.

30. Zong, X., and Vlachos, D.G. (2022). Exploring structure-sensitive relations for small species adsorption using machine learning. J. Chem. Inf. Model. *62*, 4361–4368. https://doi.org/10.1021/acs.jcim.2c00872.

31. Wulff, G. (1901). XXV. Zur Frage der Geschwindigkeit des Wachsthums und der Auflösung der Krystallflächen. Z. für Kristallogr. - Cryst. Mater. *34*, 449–530. https://doi.org/10.1524/zkri.1901.34.1.449.

32. Marks, L.D. (1983). Modified Wulff constructions for twinned particles. J. Cryst. Growth *61*, 556–566. https://doi.org/10.1016/0022-0248(83)90184-7.

33. Winterbottom, W.L. (1967). Equilibrium shape of a small particle in contact with a foreign substrate. Acta Metall. *15*, 303–310. https://doi.org/10.1016/0001-6160(67)90206-4.

34. Roosen, A.R., McCormack, R.P., and Carter, W. (1998). Wulffman: a tool for the calculation and display of crystal shapes. Comput. Mater. Sci. *11*, 16–26. https://doi.org/10.1016/S0927-0256(97)00167-5.

35. Marks, L.D., and Peng, L. (2016). Nanoparticle shape, thermodynamics and kinetics. J. Phys. Condens. Matter *28*, 053001. https://doi.org/10.1088/0953-8984/28/5/053001.

36. Reuter, K. (2016). Ab initio thermodynamics and first-principles microkinetics for surface catalysis. Catal. Lett. *146*, 541–563. https://doi.org/10.1007/s10562-015-1684-3.

37. Yu, M., Liu, L., Jia, L., Li, D., Wang, Q., and Hou, B. (2020). Equilibrium morphology evolution of FCC cobalt nanoparticle under CO and hydrogen environments. Appl. Surf. Sci. *504*, 144469. https://doi.org/10.1016/j.apsusc.2019.144469.

38. Inoğlu, N., and Kitchin, J. (2009). Atomistic thermodynamics study of the adsorption and the effects of water-gas shift reactants on Cu catalysts under reaction conditions. J. Catal. *261*, 188–194. https://doi.org/10.1016/j.jcat.2008.11.020.

39. Geng, X., Liu, J., Yang, H., Guo, W., Bai, J., and Wen, X.D. (2022). Surface morphology evolution of cobalt nanoparticles induced by hydrogen adsorption: a theoretical study. New J. Chem. *46*, 9272–9279. https://doi.org/10.1039/d2nj00356b.

40. Pineda, M., and Stamatakis, M. (2022). Kinetic Monte Carlo simulations for heterogeneous catalysis: fundamentals, current status, and challenges. J. Chem. Phys. *156*, 120902. https://doi.org/10.1063/5.0083251.

41. Li, X., and Grabow, L.C. (2022). Evaluating the benefits of kinetic Monte Carlo and microkinetic modeling for catalyst design studies in the presence of lateral interactions. Catal. Today *387*, 150–158. https://doi.org/10.1016/j.cattod.2021.03.010.

42. Cheula, R., Soon, A., and Maestri, M. (2018). Prediction of morphological changes of catalyst materials under reaction conditions by combined: ab initio thermodynamics and microkinetic modeling. Catal. Sci. Technol. *8*, 3493–3503. https://doi.org/10.1039/c8cy00583d.

43. Domingo, M., Shahrokhi, M., Remediakis, I.N., and Lopez, N. (2018). Shape control in gold nanoparticles by N-containing ligands: insights from density functional theory and Wulff constructions. Top. Catal. *61*, 412–418. https://doi.org/10.1007/s11244-017-0880-3.

44. García-Mota, M., Rieger, M., and Reuter, K. (2015). Ab initio prediction of the equilibrium shape of supported Ag nanoparticles on α-Al2O3(0 0 0 1). J. Catal. *321*, 1–6. https://doi.org/10.1016/j.jcat.2014.10.009.

45. Ribeiro, R.A.P., Lacerda, L.H.S., Longo, E., Andrés, J., and de Lazaro, S.R. (2019). Towards enhancing the magnetic properties by morphology control of ATiO3 (A = Mn, Fe, Ni) multiferroic materials. J. Magn. Magn Mater. *475*, 544–549. https://doi.org/10.1016/j.jmmm.2018.12.002.

46. Botu, V., Ramprasad, R., and Mhadeshwar, A.B. (2014). Ceria in an oxygen environment: surface phase equilibria and its descriptors. Surf. Sci. *619*, 49–58. https://doi.org/10.1016/j.susc.2013.09.019.

47. Pushkar, A.P., and Varghese, J.J. (2022). Impact of surface-active site heterogeneity and surface hydroxylation in Ni doped ceria catalysts on oxidative dehydrogenation of propane. J. Catal. *413*, 681–691. https://doi.org/10.1016/j.jcat.2022.07.019.

48. Suthirakun, S., Ammal, S.C., Muñoz-García, A.B., Xiao, G., Chen, F., Zur Loye, H.C., Carter, E.A., and Heyden, A. (2014). Theoretical investigation of H2 oxidation on the Sr 2Fe1.5Mo0.5O6 (001) perovskite surface under anodic solid oxide fuel cell conditions. J. Am. Chem. Soc. *136*, 8374–8386. https://doi.org/10.1021/ja502629j.

49. Penschke, C., Paier, J., and Sauer, J. (2013). Oligomeric vanadium oxide species supported on the CeO_2 (111) surface: structure and reactivity studied by density functional theory. J. Phys. Chem. C *117*, 5274–5285. https://doi.org/10.1021/jp400520j.

50. Wang, T., Wang, S., Li, Y.W., Wang, J., and Jiao, H. (2012). Adsorption equilibria of CO coverage on β-Mo 2C surfaces. J. Phys. Chem. C *116*, 6340–6348. https://doi.org/10.1021/jp300422g.

51. Wang, T., Li, Y.W., Wang, J., Beller, M., and Jiao, H. (2014). Dissociative hydrogen adsorption on the hexagonal Mo2C phase at high coverage. J. Phys. Chem. C *118*, 8079–8089. https://doi.org/10.1021/jp501471u.

52. Pilania, G., Gao, P.X., and Ramprasad, R. (2012). Establishing the LaMnO3 surface phase diagram in an oxygen environment: an ab initio kinetic Monte Carlo simulation study. J. Phys. Chem. C *116*, 26349–26357. https://doi.org/10.1021/jp3083985.

53. Sumaria, V., Nguyen, L., Tao, F.F., and Sautet, P. (2023). Atomic-scale mechanism of platinum catalyst restructuring under a pressure of reactant gas. J. Am. Chem. Soc. *145*, 392–401. https://doi.org/10.1021/jacs.2c10179.

54. Yang, Y., Shen, X., and Han, Y.F. (2019). Diffusion mechanisms of metal atoms in Pd–Au bimetallic catalyst under CO atmosphere based on ab initio molecular dynamics. Appl. Surf. Sci. *483*, 991–1005. https://doi.org/10.1016/j.apsusc.2019.04.036.

55. Senftle, T.P., Van Duin, A.C.T., and Janik, M.J. (2017). Methane activation at the Pd/CeO2 interface. ACS Catal. *7*, 327–332. https://doi.org/10.1021/acscatal.6b02447.

56. Liu, W., Zhu, Y., Wu, Y., Chen, C., Hong, Y., Yue, Y., Zhang, J., and Hou, B. (2021). Molecular dynamics and machine learning in catalysts. Catalysts *11*, 1129. https://doi.org/10.3390/catal11091129.

57. Xing, M., Pathak, A.D., Sanyal, S., Peng, Q., Liu, X., and Wen, X. (2020). Temperature-dependent surface free energy and the Wulff shape of iron and iron carbide nanoparticles: a molecular dynamics study. Appl. Surf. Sci. *509*, 144859. https://doi.org/10.1016/j.apsusc.2019.144859.

58. Sharma, R.K., Nair, A.S., Bharadwaj, N., Roy, D., and Pathak, B. (2023). Role of fluxionality and metastable isomers in the ORR activity: a case study. J. Phys. Chem. C *127*, 217–228. https://doi.org/10.1021/acs.jpcc.2c06265.

59. Zhai, H., and Alexandrova, A.N. (2018). Local fluxionality of surface-deposited cluster catalysts: the case of Pt7 on Al2O3. J. Phys. Chem. Lett. *9*, 1696–1702. https://doi.org/10.1021/acs.jpclett.8b00379.

60. Zhai, H., and Alexandrova, A.N. (2017). Fluxionality of catalytic clusters: when it matters and how to address it. ACS Catal. *7*, 1905–1911. https://doi.org/10.1021/acscatal.6b03243.

61. Grajciar, L., Heard, C.J., Bondarenko, A.A., Polynski, M.V., Meeprasert, J., Pidko, E.A., and Nachtigall, P. (2018). Towards operando computational modeling in heterogeneous catalysis. Chem. Soc. Rev. *47*, 8307–8348. https://doi.org/10.1039/c8cs00398j.

62. Li, X.T., Chen, L., Wei, G.F., Shang, C., and Liu, Z.P. (2020). Sharp increase in catalytic selectivity in acetylene semihydrogenation on Pd achieved by a machine learning simulation-guided experiment. ACS Catal. *10*, 9694–9705. https://doi.org/10.1021/acscatal.0c02158.

63. Sun, M., Dougherty, A.W., Huang, B., Li, Y., and Yan, C.H. (2020). Accelerating atomic catalyst discovery by theoretical calculations-machine learning strategy. Adv. Energy Mater. *10*, 1903949. https://doi.org/10.1002/aenm.201903949.

64. Chen, A., Cai, J., Wang, Z., Han, Y., Ye, S., and Li, J. (2023). An ensemble learning classifier to discover arsenene catalysts with implanted heteroatoms for hydrogen evolution reaction. J. Energy Chem. *78*, 268–276. https://doi.org/10.1016/j.jechem.2022.11.035.

65. Gu, G.H., Noh, J., Kim, S., Back, S., Ulissi, Z., and Jung, Y. (2020). Practical deep-learning representation for fast heterogeneous catalyst screening. J. Phys. Chem. Lett. *11*, 3185–3191. https://doi.org/10.1021/acs.jpclett.0c00634.

66. Suzuki, K., Toyao, T., Maeno, Z., Takakusagi, S., Shimizu, K., and Takigawa, I. (2019). Statistical analysis and discovery of heterogeneous catalysts based on machine learning from diverse published data. ChemCatChem *11*, 4537–4547. https://doi.org/10.1002/cctc.201900971.

67. Li, S., Ting, J.Y., and Barnard, A.S. (2022). The impact of domain-driven and data-driven feature selection on the inverse design of nanoparticle catalysts. J. Comput. Sci. *65*, 101896. https://doi.org/10.1016/j.jocs.2022.101896.

68. Musa, E., Doherty, F., and Goldsmith, B.R. (2022). Accelerating the structure search of catalysts with machine learning. Curr. Opin. Chem. Eng. *35*, 100771. https://doi.org/10.1016/j.coche.2021.100771.

69. Yao, Y., Dong, Q., Brozena, A., Luo, J., Miao, J., Chi, M., Wang, C., Kevrekidis, I.G., Ren, Z.J., Greeley, J., et al. (2022). High-entropy nanoparticles: synthesis-structureproperty relationships and data-driven discovery.

Science *376*, 3103. https://doi.org/10.1126/science.abn3103.

70. Lu, Z., Chen, Z.W., and Singh, C.V. (2020). Neural network-assisted development of high-entropy alloy catalysts: decoupling ligand and coordination effects. Matter *3*, 1318–1333. https://doi.org/10.1016/j.matt.2020.07.029.

71. Dixit, M., Kostetskyy, P., and Mpourmpakis, G. (2018). Structure-activity relationships in alkane dehydrogenation on γ-Al2O3: site-dependent reactions. ACS Catal. *8*, 11570–11578. https://doi.org/10.1021/acscatal.8b03484.

72. Darby, M.T., Stamatakis, M., Michaelides, A., and Sykes, E.C.H. (2018). Lonely atoms with special gifts: breaking linear scaling relationships in heterogeneous catalysis with single-atom alloys. J. Phys. Chem. Lett. *9*, 5636–5646. https://doi.org/10.1021/acs.jpclett.8b01888.

73. Wang, W., Zhang, R., Liu, Z., Wang, W., Zhang, Q., and Wang, Q. (2022). Periodic DFT calculation for the formation of EPFRs from phenol on γ-Al2O3(110): site-dependent mechanism and the role of ambient water. J. Environ. Chem. Eng. *10*, 108386. https://doi.org/10.1016/j.jece.2022.108386.

74. Giannozzi, P., and Baroni, S. (1994). Vibrational and dielectric properties of C60 from density-functional perturbation theory. J. Chem. Phys. *100*, 8537–8539. https://doi.org/10.1063/1.466753.

75. Baroni, S., de Gironcoli, S., Dal Corso, A., and Giannozzi, P. (2001). Phonons and related crystal properties from density-functional perturbation theory. Rev. Mod. Phys. *73*, 515–562. https://doi.org/10.1103/RevModPhys.73.515.

76. Carbone, M.R., Topsakal, M., Lu, D., and Yoo, S. (2020). Machine-learning X-ray absorption spectra to quantitative accuracy. Phys. Rev. Lett. *124*, 156401. https://doi.org/10.1103/PhysRevLett.124.156401.

77. Kartashov, O.O., Chernov, A.V., Polyanichenko, D.S., and Butakova, M.A. (2021). XAS Data Preprocessing of Nanocatalysts for Machine Learning Applications. Materials *14*. https://doi.org/10.3390/ma14247884.

78. Rankine, C.D., Penfold, T.J., Madkhali, M.M.M., and Penfold, T.J. (2020). A deep neural network for the rapid prediction of X-ray absorption spectra. J. Phys. Chem. A *124*, 4263–4270. https://doi.org/10.1021/acs.jpca.0c03723.

79. Liu, Y., Marcella, N., Timoshenko, J., Halder, A., Yang, B., Kolipaka, L., Pellin, M.J., Seifert, S., Vajda, S., Liu, P., and Frenkel, A.I. (2019). Mapping XANES spectra on structural descriptors of copper oxide clusters using supervised machine learning. J. Chem. Phys. *151*, 164201. https://doi.org/10.1063/1.5126597.

80. Sarma, B.B., Maurer, F., Doronkin, D.E., and Grunwaldt, J.D. (2023). Design of single-atom catalysts and tracking their fate using operando and advanced X-ray spectroscopic tools. Chem. Rev. *123*, 379–444. https://doi.org/10.1021/acs.chemrev.2c00495.

81. Timoshenko, J., Lu, D., Lin, Y., and Frenkel, A.I. (2017). Supervised machine-learning-based determination of three-dimensional structure of metallic nanoparticles. J. Phys. Chem. Lett. *8*, 5091–5098. https://doi.org/10.1021/acs.jpclett.7b02364.

82. Timoshenko, J., and Frenkel, A.I. (2019). Inverting" X-ray absorption spectra of catalysts by machine learning in search for activity descriptors. ACS Catal. *9*, 10192–10211. https://doi.org/10.1021/acscatal.9b03599.

83. Belskaya, O.B., Danilova, I.G., Kazakov, M.O., Mironenko, R.M., Lavrenov, A.V., and Likholobov, V.A. (2012). FTIR spectroscopy of adsorbed probe molecules for analyzing the surface properties of supported Pt (Pd) catalysts. In Infrared Spectroscopy in Materials Science, Engineering and Technology, PT Theophile (InTech).

84. Hill, I.M., Hanspal, S., Young, Z.D., and Davis, R.J. (2015). DRIFTS of probe molecules adsorbed on magnesia, zirconia, and hydroxyapatite catalysts. J. Phys. Chem. C *119*, 9186–9197. https://doi.org/10.1021/jp509889j.

85. Chakarova, K., Strauss, I., Mihaylov, M., Drenchev, N., and Hadjiivanov, K. (2019). Evolution of acid and basic sites in UiO-66 and UiO-66-NH2 metal-organic frameworks: FTIR study by probe molecules. Microporous Mesoporous Mater. *281*, 110–122. https://doi.org/10.1016/j.micromeso.2019.03.006.

86. Lansford, J.L., and Vlachos, D.G. (2020). Spectroscopic probe molecule selection using quantum theory, first-principles calculations, and machine learning. ACS Nano *14*, 17295–17307. https://doi.org/10.1021/acsnano.0c07408.

87. Porezag, D., and Pederson, M.R. (1996). Infrared intensities and Raman-scattering activities within density-functional theory. Phys. Rev. B Condens. Matter *54*, 7830–7836. https://doi.org/10.1103/PhysRevB.54.7830.

88. Xu, Y., Wang, F., Liu, X., Liu, Y., Luo, M., Teng, B., Fan, M., and Liu, X. (2019). Resolving a decade-long question of oxygen defects in Raman spectra of ceria-based catalysts at atomic level. J. Phys. Chem. C *123*, 18889–18894. https://doi.org/10.1021/acs.jpcc.9b00633.

89. McBride, J.R., Hass, K.C., Poindexter, B.D., and Weber, W.H. (1994). Raman and x-ray studies of Ce 1− x RE x O 2− y , where RE=La, Pr, Nd, Eu, Gd, and Tb. J. Appl. Phys. *76*, 2435–2441. https://doi.org/10.1063/1.357593.

90. Nakajima, A., Yoshihara, A., and Ishigame, M. (1994). Defect-induced Raman spectra in doped CeO2. Phys. Rev. B *50*, 13297–13307. https://doi.org/10.1103/PhysRevB.50.13297.

91. Schilling, C., Hofmann, A., Hess, C., Verónica, M., and Ganduglia-Pirovano, V. (2017). Raman spectra of polycrystalline CeO 2 : a density functional theory study. Chem. Rev. *121*, 20834–20849. https://doi.org/10.1021/acs.jpcc.7b06643.

92. Nilsson, V., Van Den Bossche, M., Hellman, A., and Grönbeck, H. (2015). Trends in adsorbate induced core level shifts. Surf. Sci. *640*, 59–64. https://doi.org/10.1016/j.susc.2015.03.019.

93. Lizzit, S., Baraldi, A., Groso, A., Reuter, K., Ganduglia-Pirovano, M.V., Stampfl, C., Scheffler, M., Stichler, M., Keller, C., Wurth, W., and Menzel, D. (2001). Surface core-level shifts of clean and oxygen-covered Ru(0001). Phys. Rev. B *63*, 205419–205514. https://doi.org/10.1103/PhysRevB.63.205419.

94. Köhler, L., and Kresse, G. (2004). Density functional study of CO on Rh(111) [78]. Phys. Rev. B Condens. Matter *70*, 1–9. https://doi.org/10.1103/PhysRevB.70.165405.

95. Posada-Borbón, A., Bosio, N., and Grönbeck, H. (2021). On the signatures of oxygen vacancies in O1s core level shifts. Surf. Sci. *705*, 1–5. https://doi.org/10.1016/j.susc.2020.121761.

96. Büchner, C., Lichtenstein, L., Stuckenholz, S., Heyde, M., Ringleb, F., Sterrer, M., Kaden, W.E., Giordano, L., Pacchioni, G., and Freund, H.J. (2014). Adsorption of Au and Pd on ruthenium-supported bilayer silica. J. Phys. Chem. C *118*, 20959–20969. https://doi.org/10.1021/jp5055342.

97. Lousada, C.M., and Korzhavyi, P.A. (2018). First stages of oxide growth on Al(1 1 0) and core-level shifts from density functional theory calculations. Appl. Surf. Sci. *441*, 174–186. https://doi.org/10.1016/j.apsusc.2018.01.246.

98. Posada-Borbón, A., Hagman, B., Schaefer, A., Zhang, C., Shipilin, M., Hellman, A., Gustafson, J., and Grönbeck, H. (2018). Initial oxidation of Cu(100) studied by X-ray photo-electron spectroscopy and density functional theory calculations. Surf. Sci. *675*, 64–69. https://doi.org/10.1016/j.susc.2018.04.015.

99. Trinh, Q.T., Bhola, K., Amaniampong, P.N., Jérôme, F., and Mushrif, S.H. (2018). Synergistic application of XPS and DFT to investigate metal oxide surface catalysis. J. Phys. Chem. C *122*, 22397–22406. https://doi.org/10.1021/acs.jpcc.8b05499.

100. Pueyo Bellafont, N., Bagus, P.S., Sousa, C., and Illas, F. (2017). Assessing the ability of DFT methods to describe static electron correlation effects: CO core level binding energies as a representative case. J. Chem. Phys. *147*, 024106. https://doi.org/10.1063/1.4991833.

101. Quertite, K., Enriquez, H., Trcera, N., Bendounan, A., Mayne, A.J., Dujardin, G., El Kenz, A., Benyoussef, A., Dappe, Y.J., Kara, A., and Oughaddou, H. (2021). Electron beam analysis induces Cl vacancy defects in a NaCl thin film. Nanotechnology *33*,

095706. https://doi.org/10.1088/1361-6528/ac3c79.

102. Atrei, A., Cortigiani, B., and Ferrari, A.M. (2012). Epitaxial growth of TiO 2 films with the rutile (110) structure on Ag(100). J. Phys. Condens. Matter *24*, 445005. https://doi.org/10.1088/0953-8984/24/44/445005.

103. Chen, H., Huang, Z., Rong, W., Zhou, J., Xu, Z., Li, W., Lin, Y., Zhou, X., and Wu, K. (2020). Atomically resolved structure of monolayer ceria island on Pt(111). J. Phys. Chem. C *124*, 28531–28538. https://doi.org/10.1021/acs.jpcc.0c08303.

104. Nilius, N., Kozlov, S.M., Jerratsch, J.F., Baron, M., Shao, X., Viñes, F., Shaikhutdinov, S., Neyman, K.M., and Freund, H.J. (2012). Formation of One-dimensional electronic states along the step edges of CeO 2(111). ACS Nano *6*, 1126–1133. https://doi.org/10.1021/nn2036472.

105. Sung, D., Kim, D.H., and Hong, S. (2018). Functionalization of Ge(1 0 0) surface by adsorption of phenylthiol. Appl. Surf. Sci. *456*, 908–914. https://doi.org/10.1016/j.apsusc.2018.06.212.

106. Li, S.C., Jacobson, P., Zhao, S.L., Gong, X.Q., and Diebold, U. (2012). Trapping nitric oxide by surface hydroxyls on rutile TiO 2(110). J. Phys. Chem. C *116*, 1887–1891. https://doi.org/10.1021/jp209290a.

107. Chen, H., Fu, X., Jiang, J., Han, S., Gong, Z., Cui, Y., Li, Y., Liu, Y., Zhao, X., Huang, W., et al. (2019). CO and H2 Activation over g-ZnO Layers and w-ZnO(0001). ACS Catal. *10*, 1373–1382. https://doi.org/10.1021/acscatal.8b03687.

108. Baron, M., Abbott, H., Bondarchuk, O., Stacchiola, D., Uhl, A., Shaikhutdinov, S., Freund, H.-J., Popa, C., Ganduglia-Pirovano, M., and Sauer, J. (2009). Resolving the atomic structure of vanadia monolayer catalysts: monomers, trimers, and oligomers on ceria. Angew. Chem. *121*, 8150–8153. https://doi.org/10.1002/ange.200903085.

109. Chizallet, C. (2022). Achievements and expectations in the field of computational heterogeneous catalysis in an innovation context. Top. Catal. *65*, 69–81. https://doi.org/10.1007/s11244-021-01489-y.

110. Larmier, K., Liao, W.C., Tada, S., Lam, E., Verel, R., Bansode, A., Urakawa, A., Comas-Vives, A., and Copéret, C. (2017). CO2-to-Methanol hydrogenation on zirconia-supported copper nanoparticles: reaction intermediates and the role of the metal–support interface. Angew. Chem., Int. Ed. Engl. *56*, 2318–2323. https://doi.org/10.1002/anie.201610166.

111. Van De Vijver, R., Vandewiele, N.M., Bhoorasingh, P.L., Slakman, B.L., Seyedzadeh Khanshan, F., Carstensen, H.H., Reyniers, M.F., Marin, G.B., West, R.H., and Van Geem, K.M. (2015). Automatic mechanism and kinetic model generation for gas- and solution-phase processes: a perspective on best practices, recent advances, and future challenges. Int. J. Chem. Kinet. *47*, 199–231. https://doi.org/10.1002/kin.20902.

112. Vernuccio, S., and Broadbelt, L.J. (2019). Discerning complex reaction networks using automated generators. AIChE J. *65*, e16663. https://doi.org/10.1002/aic.16663.

113. Vandewiele, N.M., Van Geem, K.M., Reyniers, M.F., and Marin, G.B. (2012). Genesys: kinetic model construction using chemo-informatics. Chem Eng J. 207– *207–208*, 526–538. https://doi.org/10.1016/j.cej.2012.07.014.

114. Németh, A., Vidóczy, T., Héberger, K., Kúti, Z., and Wágner, J. (2002). MECHGEN: computer aided generation and reduction of reaction mechanisms. J. Chem. Inf. Comput. Sci. *42*, 208–214. https://doi.org/10.1021/ci010365i.

115. Di Maio, F.P., and Lignola, P.G. (1992). KING, a KInetic network generator. Chem. Eng. Sci. *47*, 2713–2718. https://doi.org/10.1016/0009-2509(92)87118-A.

116. Blurock, E.S. (1995). Reaction: system for modeling chemical reactions. J. Chem. Inf. Comput. Sci. *35*, 607–616. https://doi.org/10.1021/ci00025a032.

117. Goldsmith, C.F., and West, R.H. (2017). Automatic generation of microkinetic mechanisms for heterogeneous catalysis. J. Phys. Chem. C *121*, 9970–9981. https://doi.org/10.1021/acs.jpcc.7b02133.

118. Liu, M., Grinberg Dana, A., Johnson, M.S., Goldman, M.J., Jocher, A., Payne, A.M., Grambow, C.A., Han, K., Yee, N.W., Mazeau, E.J., et al. (2021). Reaction mechanism generator v3.0: advances in automatic mechanism generation. J. Chem. Inf. Model. *61*, 2686–2696. https://doi.org/10.1021/acs.jcim.0c01480.

119. Grenda, J.M., Androulakis, I.P., Dean, A.M., and Green, W.H. (2003). Application of computational kinetic mechanism generation to model the autocatalytic pyrolysis of methane. Ind. Eng. Chem. Res. *42*, 1000–1010. https://doi.org/10.1021/ie020581w.

120. Broadbelt, L.J., Stark, S.M., and Klein, M.T. (1994). Computer generated pyrolysis modeling: on-the-fly generation of species, reactions, and rates. Ind. Eng. Chem. Res. *33*, 790–799. https://doi.org/10.1021/ie00028a003.

121. Warth, V., Battin-Leclerc, F., Fournet, R., Glaude, P.A., Côme, G.M., and Scacchi, G. (2000). Computer based generation of reaction mechanisms for gas-phase oxidation. Comput. Chem. *24*, 541–560. https://doi.org/10.1016/S0097-8485(99)00092-3.

122. Dellon, L.D., Sung, C.Y., Robichaud, D.J., and Broadbelt, L.J. (2019). 110th anniversary: microkinetic modeling of the vapor phase upgrading of biomass-derived oxygenates. Ind. Eng. Chem. Res. *58*, 15173–15189. https://doi.org/10.1021/acs.iecr.9b03242.

123. Brydon, R.R.O., Peng, A., Qian, L., Kung, H.H., and Broadbelt, L.J. (2018). Microkinetic modeling of homogeneous and gold nanoparticle-catalyzed oxidation of cyclooctene. Ind. Eng. Chem. Res. *57*, 4832–4840. https://doi.org/10.1021/acs.iecr.8b00315.

124. Koninckx, E., Colin, J.G., Broadbelt, L.J., and Vernuccio, S. (2022). Catalytic conversion of alkenes on acidic zeolites: automated generation of reaction mechanisms and lumping technique. ACS Eng. Au *2*, 257–271. https://doi.org/10.1021/acsengineeringau.2c00004.

125. Lim, J.Y., Loy, A.C.M., Alhazmi, H., Fui, B.C.L., Cheah, K.W., Taylor, M.J., Kyriakou, G., and Yoo, C.K. (2022). Machine learning–assisted CO2 utilization in the catalytic dry reforming of hydrocarbons: reaction pathways and multicriteria optimization analyses. Int. J. Energy Res. *46*, 6277–6291. https://doi.org/10.1002/er.7565.

126. Kreitz, B., Lott, P., Bae, J., Blöndal, K., Angeli, S., Ulissi, Z.W., Studt, F., Goldsmith, C.F., and Deutschmann, O. (2022). Detailed microkinetics for the oxidation of exhaust gas emissions through automated mechanism generation. ACS Catal. *12*, 11137–11151. https://doi.org/10.1021/acscatal.2c03378.

127. Rangarajan, S., Bhan, A., and Daoutidis, P. (2012). Language-oriented rule-based reaction network generation and analysis: description of RING. Comput. Chem. Eng. *45*, 114–123. https://doi.org/10.1016/j.compchemeng.2012.06.008.

128. Wang, C., Wittreich, G.R., Lin, C., Huang, R., Vlachos, D.G., and Gorte, R.J. (2020). Hydrodeoxygenation of m-cresol over Pt-WOx/C using H2 generated in situ by n-hexane dehydrogenation. Catal. Lett. *14*, 913–921. https://doi.org/10.1007/s10562-019-03027-8.

129. Chen, C.J., Rangarajan, S., Hill, I.M., and Bhan, A. (2014). Kinetics and thermochemistry of C4-C6 olefin cracking on H-ZSM-5. ACS Catal. *4*, 2319–2327. https://doi.org/10.1021/cs500119n.

130. Sivaramakrishnan, K., Puliyanda, A., De Klerk, A., and Prasad, V. (2021). A data-driven approach to generate pseudo-reaction sequences for the thermal conversion of Athabasca bitumen. React. Chem. Eng. *6*, 505–537. https://doi.org/10.1039/d0re00321b.

131. Puliyanda, A., Li, Z., and Prasad, V. (2022). Real-time monitoring of reaction mechanisms from spectroscopic data using hidden semi-Markov models for mode identification. J. Process Control *117*, 188–205. https://doi.org/10.1016/j.jprocont.2022.07.011.

132. Puliyanda, A., Sivaramakrishnan, K., Li, Z., De Klerk, A., and Prasad, V. (2020). Data fusion by joint non-negative matrix factorization for hypothesizing pseudo-chemistry using Bayesian networks. React. Chem. Eng. *5*, 1719–1737. https://doi.org/10.1039/d0re00147c.

133. Burés, J., and Larrosa, I. (2023). Organic reaction mechanism classification using machine learning. Nature *613*, 689–695.

https://doi.org/10.1038/s41586-022-05639-4.

134. Zaera, F. (2014). New advances in the use of infrared absorption spectroscopy for the characterization of heterogeneous catalytic reactions. Chem. Soc. Rev. *43*, 7624–7663. https://doi.org/10.1039/c3cs60374a.

135. Schumacher, L., and Hess, C. (2021). The active role of the support in propane ODH over VOx/CeO2 catalysts studied using multiple operando spectroscopies. J. Catal. *398*, 29–43. https://doi.org/10.1016/J.JCAT.2021.04.006.

136. Schumacher, L., Weyel, J., and Hess, C. (2022). Unraveling the active vanadium sites and adsorbate dynamics in VOx/CeO2Oxidation catalysts using transient IR spectroscopy. J. Am. Chem. Soc. *144*, 14874–14887. https://doi.org/10.1021/JACS.2C06303.

137. Hess, C. (2021). New advances in using Raman spectroscopy for the characterization of catalysts and catalytic reactions. Chem. Soc. Rev. *50*, 3519–3564. https://doi.org/10.1039/D0CS01059F.

138. Schilling, C., and Hess, C. (2018). Elucidating the role of support oxygen in the Water−Gas shift reaction over ceria-supported gold catalysts using operando spectroscopy. ACS Catal. *9*, 1159–1171. https://doi.org/10.1021/acscatal.8b04536.

139. Lohrenscheit, M., and Hess, C. (2016). Direct evidence for the participation of oxygen vacancies in the oxidation of carbon monoxide over ceria-supported gold catalysts by using operando Raman spectroscopy. ChemCatChem *8*, 523–526. https://doi.org/10.1002/cctc.201501129.

140. Kim, H.Y., Lee, H.M., and Henkelman, G. (2012). CO oxidation mechanism on CeO 2-supported Au nanoparticles. J. Am. Chem. Soc. *134*, 1560–1570. https://doi.org/10.1021/ja207510v.

141. Shustorovich, E., and Sellers, H. (1998). The UBI-QEP method: a practical theoretical approach to understanding chemistry on transition metal surfaces. Surf. Sci. Rep. *31*, 1–119. https://doi.org/10.1016/s0167-5729(97)00016-2.

142. Motagamwala, A.H., Ball, M.R., and Dumesic, J.A. (2018). Microkinetic analysis and scaling relations for catalyst design. Annu. Rev. Chem. Biomol. Eng. *9*, 413–450. https://doi.org/10.1146/annurev-chembioeng-060817-084103.

143. Shustorovich, E. (1990). The bond-order conservation approach to chemisorption and heterogeneous catalysis: applications and implications. In Advances in catalysis (Elsevier), pp. 101–163.

144. Shustorovich, E.M., and Zeigarnik, A.V. (2006). The UBI-QEP method: basic formalism and applications to chemisorption phenomena on transition metal surfaces. Chemisorption energetics. Russ. J. Phys. Chem. *80*, 4–30. https://doi.org/10.1134/S003602440601002X.

145. Maestri, M., and Reuter, K. (2011). Semiempirical rate constants for complex chemical kinetics: first-principles assessment and rational refinement. Angew. Chem., Int. Ed. Engl. *50*, 1194–1197. https://doi.org/10.1002/anie.201006488.

146. Park, J., Cho, J., Lee, Y., Park, M.J., and Lee, W.B. (2019). Practical microkinetic modeling approach for methanol synthesis from syngas over a Cu-based catalyst. Ind. Eng. Chem. Res. *58*, 8663–8673. https://doi.org/10.1021/acs.iecr.9b01254.

147. Whitten, J., and Yang, H. (1996). Theory of chemisorption and reactions on metal surfaces. Surf. Sci. Rep. *24*, 55–124. https://doi.org/10.1016/0167-5729(96)80004-5.

148. Wang, Y., Montoya, J.H., Tsai, C., Ahlquist, M.S.G., Nørskov, J.K., and Studt, F. (2016). Scaling relationships for binding energies of transition metal complexes. Catal. Lett. *146*, 304–308. https://doi.org/10.1007/s10562-015-1667-4.

149. Wang, S., Vorotnikov, V., Sutton, J.E., and Vlachos, D.G. (2014). Brønsted–evans–polanyi and transition state scaling relations of furan derivatives on Pd(111) and their relation to those of small molecules. ACS Catal. *4*, 604–612. https://doi.org/10.1021/cs400942u.

150. Ichino, T., Takagi, M., and Maeda, S. (2019). A systematic study on bond activation energies of NO, N 2 , and O 2 on hexamers of eight transition metals. ChemCatChem *11*, 1346–1353. https://doi.org/10.1002/cctc.201801595.

151. Wang, S., Petzold, V., Tripkovic, V., Kleis, J., Howalt, J.G., Skúlason, E., Fernández, E.M., Hvolbæk, B., Jones, G., Toftelund, A., et al. (2011). Universal transition state scaling relations for (de)hydrogenation over transition metals. Phys. Chem. Chem. Phys. *13*, 20760–20765. https://doi.org/10.1039/C1CP20547A.

152. Sun, G., Zhao, Z.J., Mu, R., Zha, S., Li, L., Chen, S., Zang, K., Luo, J., Li, Z., Purdy, S.C., et al. (2018). Breaking the scaling relationship via thermally stable Pt/Cu single atom alloys for catalytic dehydrogenation. Nat. Commun. *9*, 4454. https://doi.org/10.1038/s41467-018-06967-8.

153. Lustemberg, P.G., Zhang, F., Gutiérrez, R.A., Ramírez, P.J., Senanayake, S.D., Rodriguez, J.A., Ganduglia-Pirovano, M.V., Senanayake, S.D., and Rodriguez, J.A. (2020). Breaking simple scaling relations through metal-oxide interactions: understanding room-temperature activation of methane on M/CeO2(M = Pt, Ni, or Co) interfaces. J. Phys. Chem. Lett. *11*, 9131–9137. https://doi.org/10.1021/acs.jpclett.0c02109.

154. Szécsényi, Á., Khramenkova, E., Chernyshov, I.Y., Li, G., Gascon, J., and Pidko, E.A. (2019). Breaking linear scaling relationships with secondary interactions in confined space: a case study of methane oxidation by Fe/ZSM-5 zeolite. ACS Catal. *9*, 9276–9284. https://doi.org/10.1021/acscatal.9b01914.

155. Gani, T.Z.H., and Kulik, H.J. (2018). Understanding and breaking scaling relations in single-site catalysis: methane to methanol conversion by FeIV=O. ACS Catal. *8*, 975–986. https://doi.org/10.1021/acscatal.7b03597.

156. Pérez-Ramírez, J., and López, N. (2019). Strategies to break linear scaling relationships. Nat. Catal. *2*, 971–976. https://doi.org/10.1038/s41929-019-0376-6.

157. Bhandari, S., Rangarajan, S., and Mavrikakis, M. (2020). Combining computational modeling with reaction kinetics experiments for elucidating the in situ nature of the active site in catalysis. Acc. Chem. Res. *53*, 1893–1904. https://doi.org/10.1021/acs.accounts.0c00340.

158. Shabbir, H., Pellizzeri, S., Ferrandon, M., Kim, I.S., Vermeulen, N.A., Farha, O.K., Delferro, M., Martinson, A.B.F., and Getman, R.B. (2020). Influence of spin state and electron configuration on the active site and mechanism for catalytic hydrogenation on metal cation catalysts supported on NU-1000: insights from experiments and microkinetic modeling. Catal. Sci. Technol. *10*, 3594–3602. https://doi.org/10.1039/d0cy00394h.

159. Tian, H., and Rangarajan, S. (2020). Leveraging thermochemistry data to build accurate microkinetic models. J. Phys. Chem. C *124*, 5740–5748. https://doi.org/10.1021/acs.jpcc.0c00491.

160. Xiao, T.-T., and Wang, G.-C. (2020). A DFT and microkinetic study of propylene oxide selectivity over copper-based catalysts: effects of copper valence states. Catal. Sci. Technol. *10*, 7640–7651. https://doi.org/10.1039/D0CY01611J.

161. van Helden, P., van den Berg, J.-A., and Weststrate, C.J. (2012). Hydrogen adsorption on Co surfaces: a density functional theory and temperature programmed desorption study. ACS Catal. *2*, 1097–1107. https://doi.org/10.1021/cs2006586.

162. Pascucci, B.M., Otero, G.S., Belelli, P.G., and Branda, M.M. (2019). Understanding the effects of metal particle size on the NO2 reduction from a DFT study. Appl. Surf. Sci. *489*, 1019–1029. https://doi.org/10.1016/j.apsusc.2019.05.318.

163. Zhou, R.J., Yan, W.Q., Cao, Y.Q., Zhou, J.H., Sui, Z.J., Li, W., Chen, D., Zhou, X.G., and Zhu, Y.A. (2022). Probing the structure sensitivity of dimethyl oxalate partial hydrogenation over Ag nanoparticles: a combined experimental and microkinetic study. Chem. Eng. Sci. *259*, 117830. https://doi.org/10.1016/j.ces.2022.117830.

164. An, J.W., and Wang, G.C. (2022). Titania crystal-plane-determined activity of copper cluster in water-gas shift reaction. Appl. Surf. Sci. *591*, 153145. https://doi.org/10.1016/j.apsusc.2022.153145.

165. Song, X., Sun, K., Hao, X., Su, H.Y., Ma, X., and Xu, Y. (2019). Facet-dependent of catalytic selectivity: the case of H2O2 direct synthesis on Pd surfaces. J. Phys. Chem. C

123, 26324–26337. https://doi.org/10.1021/acs.jpcc.9b07097.

166. Núñez, M., Lansford, J.L., and Vlachos, D.G. (2019). Optimization of the facet structure of transition-metal catalysts applied to the oxygen reduction reaction. Nat. Chem. *11*, 449–456. https://doi.org/10.1038/s41557-019-0247-4.

167. Deimel, M., Reuter, K., and Andersen, M. (2020). Active site representation in first-principles microkinetic models: data-enhanced computational screening for improved methanation catalysts. ACS Catal. *10*, 13729–13736. https://doi.org/10.1021/acscatal.0c04045.

168. Wang, B., Gu, T., Lu, Y., and Yang, B. (2020). Prediction of energies for reaction intermediates and transition states on catalyst surfaces using graph-based machine learning models. Mol. Catal. *498*, 111266. https://doi.org/10.1016/j.mcat.2020.111266.

169. Göltl, F., and Mavrikakis, M. (2022). Generalized brønsted-evans-polanyi relationships for reactions on metal surfaces from machine learning. ChemCatChem *14*, e202201108. https://doi.org/10.1002/cctc.202201108.

170. Lewis-Atwell, T., Townsend, P.A., and Grayson, M.N. (2022). Machine learning activation energies of chemical reactions. Wiley Interdiscip. Rev. Comput. Mol. Sci. *12*, e1593. https://doi.org/10.1002/wcms.1593.

171. Fung, V., Hu, G., Ganesh, P., and Sumpter, B.G. (2021). Machine learned features from density of states for accurate adsorption energy prediction. Nat. Commun. *12*, 88. https://doi.org/10.1038/s41467-020-20342-6.

172. Wang, X., Ye, S., Hu, W., Sharman, E., Liu, R., Liu, Y., Luo, Y., and Jiang, J. (2020). Electric dipole descriptor for machine learning prediction of catalyst surface-molecular adsorbate interactions. J. Am. Chem. Soc. *142*, 7737–7743. https://doi.org/10.1021/jacs.0c01825.

173. Mou, T., Han, X., Zhu, H., and Xin, H. (2022). Machine learning of lateral adsorbate interactions in surface reaction kinetics. Curr. Opin. Chem. Eng. *36*, 100825. https://doi.org/10.1016/j.coche.2022.100825.

174. Li, Z., Zhang, Z., Chen, X., Zhou, J., and Xiao, X.M. (2017). Feature engineering of machine-learning chemisorption models for catalyst design. Catal. Today *8*, 232–238. https://doi.org/10.1016/j.cattod.2016.04.013.

175. Ghanekar, P.G., Deshpande, S., and Greeley, J. (2022). Adsorbate chemical environment-based machine learning framework for heterogeneous catalysis. Nat. Commun. *13*, 5788. https://doi.org/10.1038/s41467-022-33256-2.

176. Yang, Z., and Gao, W. (2022). Applications of machine learning in alloy catalysts: rational selection and future development of descriptors. Adv. Sci. *9*, 2106043. https://doi.org/10.1002/advs.202106043.

177. Wellendorff, J., Lundgaard, K.T., Møgelhøj, A., Petzold, V., Landis, D.D., Nørskov, J.K., Bligaard, T., and Jacobsen, K.W. (2012). Density functionals for surface science: exchange-correlation model development with Bayesian error estimation. Phys. Rev. B *85*, 235149. https://doi.org/10.1103/PhysRevB.85.235149.

178. Medford, A.J., Wellendorff, J., Vojvodic, A., Studt, F., Abild-Pedersen, F., Jacobsen, K.W., Bligaard, T., and Nørskov, J.K. (2014). Assessing the reliability of calculated catalytic ammonia synthesis rates. Science *345*, 197–200. https://doi.org/10.1126/science.1253486.

179. Wang, B., Chen, S., Zhang, J., Li, S., and Yang, B. (2019). Propagating DFT uncertainty to mechanism determination, degree of rate control, and coverage analysis: the kinetics of dry reforming of methane. J. Phys. Chem. C *123*, 30389–30397. https://doi.org/10.1021/acs.jpcc.9b08755.

180. Xu, W., Andersen, M., and Reuter, K. (2021). Data-driven descriptor engineering and refined scaling relations for predicting transition metal oxide reactivity. ACS Catal. *11*, 734–742. https://doi.org/10.1021/acscatal.0c04170.

181. Ouyang, R., Curtarolo, S., Ahmetcik, E., Scheffler, M., and Ghiringhelli, L.M. (2018). SISSO: a compressed-sensing method for identifying the best low-dimensional descriptor in an immensity of offered candidates. Phys. Rev. Mater. *2*, 083802. https://doi.org/10.1103/PhysRevMaterials.2.083802.

182. Ulissi, Z., Prasad, V., and Vlachos, D.G. (2011). Effect of multiscale model uncertainty on identification of optimal catalyst properties. J. Catal. *281*, 339–344. https://doi.org/10.1016/j.jcat.2011.05.019.

183. Matera, S., Schneider, W.F., Heyden, A., and Savara, A. (2019). Progress in accurate chemical kinetic modeling, simulations, and parameter estimation for heterogeneous catalysis. ACS Catal. *9*, 6624–6647. https://doi.org/10.1021/acscatal.9b01234.

184. Afolabi, A.T.F., Li, C.Z., and Kechagiopoulos, P.N. (2019). Microkinetic modeling and reaction pathway analysis of the steam reforming of ethanol over Ni/SiO2. Int. J. Hydrogen Energy *44*, 22816–22830. https://doi.org/10.1016/j.ijhydene.2019.07.040.

185. Raghu, A.K., and Kaisare, N.S. (2020). Microkinetic modeling and analysis of CO2Methanation on ruthenium. Ind. Eng. Chem. Res. *59*, 16161–16169. https://doi.org/10.1021/acs.iecr.0c02685.

186. Appari, S., Janardhanan, V.M., Bauri, R., Jayanti, S., and Deutschmann, O. (2014). A detailed kinetic model for biogas steam reforming on Ni and catalyst deactivation due to sulfur poisoning. Appl. Catal. Gen. *471*, 118–125. https://doi.org/10.1016/j.apcata.2013.12.002.

187. Prašnikar, A., Jurković, D.L., and Likozar, B. (2021). Reaction path analysis of CO2 reduction to methanol through multisite microkinetic modeling over Cu/ZnO/Al2O3 catalysts. Appl. Catal., B *292*, 2–11. https://doi.org/10.1016/j.apcatb.2021.120190.

188. Mhadeshwar, A.B., Wang, H., and Vlachos, D.G. (2003). Thermodynamic consistency in microkinetic development of surface reaction mechanisms. J. Phys. Chem. B *107*, 12721–12733. https://doi.org/10.1021/jp034954y.

189. Gossler, H., Maier, L., Angeli, S., Tischer, S., and Deutschmann, O. (2018). CaRMeN: a tool for analysing and deriving kinetics in the real world. Phys. Chem. Chem. Phys. *20*, 10857–10876. https://doi.org/10.1039/c7cp07777g.

190. Stotz, H., Maier, L., Boubnov, A., Gremminger, A.T., Grunwaldt, J.D., and Deutschmann, O. (2019). Surface reaction kinetics of methane oxidation over PdO. J. Catal. *370*, 152–175. https://doi.org/10.1016/j.jcat.2018.12.007.

191. Sutton, J.E., Panagiotopoulou, P., Verykios, X.E., and Vlachos, D.G. (2013). Combined DFT, microkinetic, and experimental study of ethanol steam reforming on Pt. J. Phys. Chem. C *117*, 4691–4706. https://doi.org/10.1021/jp312593u.

192. Campbell, C.T. (2017). The degree of rate control: a powerful tool for catalysis Research. ACS Catal. *7*, 2770–2779. https://doi.org/10.1021/acscatal.7b00115.

193. Streibel, V., Aljama, H.A., Yang, A.C., Choksi, T.S., Sánchez-Carrera, R.S., Schäfer, A., Li, Y., Cargnello, M., and Abild-Pedersen, F. (2022). Microkinetic modeling of propene combustion on a stepped, metallic palladium surface and the importance of oxygen coverage. ACS Catal. *12*, 1742–1757. https://doi.org/10.1021/acscatal.1c03699.

194. Yuan, Y., Dong, X., and Ricardez-Sandoval, L. (2020). A multi-scale simulation of syngas combustion reactions by Ni-based oxygen carriers for chemical looping combustion. Appl. Surf. Sci. *531*, 147277. https://doi.org/10.1016/j.apsusc.2020.147277.

195. Delgado, K., Maier, L., Tischer, S., Zellner, A., Stotz, H., and Deutschmann, O. (2015). Surface reaction kinetics of steam- and CO2-reforming as well as oxidation of methane over nickel-based catalysts. Catalysts *5*, 871–904. https://doi.org/10.3390/catal5020871.

196. Zhu, T., van Grootel, P.W., Filot, I.A., Sun, S.G., van Santen, R.A., and Hensen, E.J. (2013). Microkinetics of steam methane reforming on platinum and rhodium metal surfaces. J. Catal. *297*, 227–235. https://doi.org/10.1016/j.jcat.2012.10.010.

197. Schmider, D., Maier, L., and Deutschmann, O. (2021). Reaction kinetics of CO and CO2Methanation over nickel. Ind. Eng. Chem. Res. *60*, 5792–5805. https://doi.org/10.1021/acs.iecr.1c00389.

198. Lacerda de Oliveira Campos, B., Herrera Delgado, K., Wild, S., Studt, F., Pitter, S., and Sauer, J. (2021). Surface reaction kinetics of the methanol synthesis and the water gas

shift reaction on Cu/ZnO/Al2O3. React. Chem. Eng. *6*, 868–887. https://doi.org/10.1039/d1re00040c.

199. De Carvalho, T.P., Catapan, R.C., Oliveira, A.A.M., and Vlachos, D.G. (2018). Microkinetic modeling and reduced rate expression of the water-gas shift reaction on nickel. Ind. Eng. Chem. Res. *57*, 10269–10280. https://doi.org/10.1021/acs.iecr.8b01957.

200. Sterk, E.B., Nieuwelink, A.E., Monai, M., Louwen, J.N., Vogt, E.T.C., Filot, I.A.W., and Weckhuysen, B.M. (2022). Structure sensitivity of CO2Conversion over nickel metal nanoparticles explained by micro-kinetics simulations. JACS Au *2*, 2714–2730. https://doi.org/10.1021/jacsau.2c00430.

201. Zijlstra, B., Broos, R.J., Chen, W., Filot, I.A., and Hensen, E.J. (2020). First-principles based microkinetic modeling of transient kinetics of CO hydrogenation on cobalt catalysts. Catal. Today *342*, 131–141. https://doi.org/10.1016/j.cattod.2019.03.002.

202. Dharmalingam, B.C., Koushik V, A., Mureddu, M., Atzori, L., Lai, S., Pettinau, A., Kaisare, N.S., Aghalayam, P., and Varghese, J.J. (2023). Unravelling the role of metal-metal oxide interfaces of Cu/ZnO/ZrO2/Al2O3 catalyst for methanol synthesis from CO2: insights from experiments and DFT-based microkinetic modeling. Appl. Catal., B *332*, 122743. https://doi.org/10.1016/j.apcatb.2023.122743.

203. Ren, B., Li, J., Wen, G., Ricardez−Sandoval, L., and Croiset, E. (2018). First-principles based microkinetic modeling of CO2 reduction at the Ni/SDC cathode of a solid oxide electrolysis cell. J. Phys. Chem. C *122*, 21151–21161. https://doi.org/10.1021/acs.jpcc.8b05312.

204. Ren, B., Wen, G., Ricardez–Sandoval, L., and Croiset, E. (2021). New mechanistic insights into CO2 reduction in solid oxide electrolysis cell through a multi-scale modeling approach. J. Power Sources *490*, 229488. https://doi.org/10.1016/j.jpowsour.2021.229488.

205. Wang, Y., Xiao, L., Qi, Y., Mahmoodinia, M., Feng, X., Yang, J., Zhu, Y.A., and Chen, D. (2019). Towards rational catalyst design: boosting the rapid prediction of transition-metal activity by improved scaling relations. Phys. Chem. Chem. Phys. *21*, 19269–19280. https://doi.org/10.1039/c9cp04286e.

206. Sutton, J.E., and Vlachos, D.G. (2015). Building large microkinetic models with first-principles[U+05F3] accuracy at reduced computational cost. Chem. Eng. Sci. *121*, 190–199. https://doi.org/10.1016/j.ces.2014.09.011.

207. Bhandari, S., Rangarajan, S., Maravelias, C.T., Dumesic, J.A., and Mavrikakis, M. (2020). Reaction mechanism of vapor-phase formic acid decomposition over platinum catalysts: DFT, reaction kinetics experiments, and microkinetic modeling. ACS Catal. *10*, 4112–4126. https://doi.org/10.1021/acscatal.9b05424.

208. Ding, Y., Xu, Y., Song, Y., Guo, C., and Hu, P. (2019). Quantitative studies of the coverage effects on microkinetic simulations for NO oxidation on Pt(111). J. Phys. Chem. C *123*, 27594–27602. https://doi.org/10.1021/acs.jpcc.9b08208.

209. Cheula, R., Maestri, M., and Mpourmpakis, G. (2020). Modeling morphology and catalytic activity of nanoparticle ensembles under reaction conditions. ACS Catal. *10*, 6149–6158. https://doi.org/10.1021/acscatal.0c01005.

210. Cheula, R., and Maestri, M. (2022). Nature and identity of the active site via structure-dependent microkinetic modeling: an application to WGS and reverse WGS reactions on Rh. Catal. Today *387*, 159–171. https://doi.org/10.1016/j.cattod.2021.05.016.

211. Genest, A., Silvestre-Albero, J., Li, W.Q., Rösch, N., and Rupprechter, G. (2021). The origin of the particle-size-dependent selectivity in 1-butene isomerization and hydrogenation on Pd/Al2O3 catalysts. Nat. Commun. *12*, 6098. https://doi.org/10.1038/s41467-021-26411-8.

212. Bosio, N., and Grönbeck, H. (2020). Sensitivity of Monte Carlo simulations to linear scaling relations. J. Phys. Chem. C *124*, 11952–11959. https://doi.org/10.1021/acs.jpcc.0c02706.

213. Salciccioli, M., Stamatakis, M., Caratzoulas, S., and Vlachos, D.G. (2011). A review of multiscale modeling of metal-catalyzed reactions: mechanism development for complexity and emergent behavior. Chem. Eng. Sci. *66*, 4319–4355. https://doi.org/10.1016/j.ces.2011.05.050.

214. Rangarajan, S., and Tian, H. (2022). Improving the predictive power of microkinetic models via machine learning. Curr. Opin. Chem. Eng. *38*, 100858. https://doi.org/10.1016/j.coche.2022.100858.

215. Tian, H., and Rangarajan, S. (2021). Machine-Learned corrections to mean-field microkinetic models at the fast diffusion limit. J. Phys. Chem. C *125*, 20275–20285. https://doi.org/10.1021/acs.jpcc.1c04495.

216. Prasad, V., Karim, A.M., Arya, A., and Vlachos, D.G. (2009). Assessment of overall rate expressions and multiscale, microkinetic model uniqueness via experimental data injection: ammonia decomposition on Ru/γ-Al 2O 3 for hydrogen production. Ind. Eng. Chem. Res. *48*, 5255–5265. https://doi.org/10.1021/ie900144x.

217. Wittreich, G.R., Gu, G.H., Robinson, D.J., Katsoulakis, M.A., and Vlachos, D.G. (2021). Uncertainty quantification and error propagation in the enthalpy and entropy of surface reactions arising from a single DFT functional. J. Phys. Chem. C *125*, 18187–18196. https://doi.org/10.1021/acs.jpcc.1c04754.

218. Döpking, S., Plaisance, C.P., Strobusch, D., Reuter, K., Scheurer, C., and Matera, S. (2018). Addressing global uncertainty and sensitivity in first-principles based microkinetic models by an adaptive sparse grid approach. J. Chem. Phys. *148*, 034102. https://doi.org/10.1063/1.5004770.

219. Sutton, J.E., Guo, W., Katsoulakis, M.A., and Vlachos, D.G. (2016). Effects of correlated parameters and uncertainty in electronic-structure-based chemical kinetic modeling. Nat. Chem. *8*, 331–337. https://doi.org/10.1038/nchem.2454.

220. Stamatakis, M., and Piccinin, S. (2016). Rationalizing the relation between adlayer structure and observed kinetics in catalysis. ACS Catal. *6*, 2105–2111. https://doi.org/10.1021/acscatal.5b02876.

221. Corte's, J., Valencia, E., and Puschmann, H. (1999). Monte Carlo and mean field theory studies of the effect of the next nearest neighbour sites of a square lattice on the monomer-dimer surface reaction. Phys. Chem. Chem. Phys. *1*, 1577–1581. https://doi.org/10.1039/a808011i.

222. Lu, J., Einhorn, S., Venkatarangan, L., Miller, M., Mann, D.A., Watkins, P.B., and LeCluyse, E. (2015). Theoretical investigation of the reaction mechanism of the hydrodeoxygenation of guaiacol over a Ru(0 0 0 1) model surface. J. Catal. *147*, 39–54. https://doi.org/10.1016/j.jcat.2014.11.003.

223. De Vrieze, J.E., Thybaut, J.W., and Saeys, M. (2018). Role of surface hydroxyl species in copper-catalyzed hydrogenation of ketones. ACS Catal. *8*, 7539–7548. https://doi.org/10.1021/acscatal.8b01652.

224. Lukkien, J.J., Segers, J.P.L., Hilbers, P.A.J., Gelten, R.J., and Jansen, A.P.J. (1998). Efficient Monte Carlo methods for the simulation of catalytic surface reactions. Phys. Rev. E *58*, 2598–2610. https://doi.org/10.1103/PhysRevE.58.2598.

225. Hansen, E.W., and Neurock, M. (2000). First-principles-based Monte Carlo methodology applied to O/Rh(100). Surf. Sci. *464*, 91–107. https://doi.org/10.1016/S0039-6028(00)00598-7.

226. Hansen, E.W., and Neurock, M. (2000). First-principles-based Monte Carlo simulation of ethylene hydrogenation kinetics on Pd. J. Catal. *196*, 241–252. https://doi.org/10.1006/jcat.2000.3018.

227. Reuter, K., and Scheffler, M. (2006). First-principles kinetic Monte Carlo simulations for heterogeneous catalysis: application to the CO oxidation at Ru O2 (110). Phys. Rev. B *73*, 045433–45517. https://doi.org/10.1103/PhysRevB.73.045433.

228. Stamatakis, M., and Vlachos, D.G. (2011). A graph-theoretical kinetic Monte Carlo framework for on-lattice chemical kinetics. J. Chem. Phys. *134*, 214115. https://doi.org/10.1063/1.3596751.

229. Mei, D., Sheth, P., Neurock, M., and Smith, C. (2006). First-principles-based kinetic Monte Carlo simulation of the selective hydrogenation of acetylene over Pd(111). J. Catal. *242*, 1–15. https://doi.org/10.1016/j.jcat.2006.05.009.

230. Papanikolaou, K.G., Darby, M.T., and Stamatakis, M. (2018). Adlayer structure and lattice size effects on catalytic rates predicted from KMC simulations: NO oxidation on Pt(111). J. Chem. Phys. *149*, 184701. https://doi.org/10.1063/1.5048787.

231. Piccinin, S., and Stamatakis, M. (2014). CO oxidation on Pd(111): a first-principles-based kinetic Monte Carlo study. ACS Catal. *4*, 2143–2152. https://doi.org/10.1021/cs500377j.

232. Huš, M., Kopač, D., and Likozar, B. (2020). Kinetics of non-oxidative propane dehydrogenation on Cr2O3 and the nature of catalyst deactivation from first-principles simulations. J. Catal. *386*, 126–138. https://doi.org/10.1016/j.jcat.2020.03.037.

233. Molero, H., Bartlett, B.F., and Tysoe, W.T. (1999). The hydrogenation of acetylene catalyzed by palladium: hydrogen pressure dependence. J. Catal. *181*, 49–56. https://doi.org/10.1006/jcat.1998.2294.

234. Yuan, Y., Dong, X., and Ricardez-Sandoval, L. (2022). A multi-scale model for syngas combustion on NiO oxygen carrier for chemical looping combustion: the role of nearest neighbors. Fuel Process. Technol. *229*, 107172. https://doi.org/10.1016/j.fuproc.2022.107172.

235. Li, J., Croiset, E., and Ricardez-Sandoval, L. (2015). Carbon nanotube growth: first-principles-based kinetic Monte Carlo model. J. Catal. *326*, 15–25. https://doi.org/10.1016/j.jcat.2015.03.010.

236. Rasoulian, S., and Ricardez-Sandoval, L.A. (2014). Uncertainty analysis and robust optimization of multiscale process systems with application to epitaxial thin film growth. Chem. Eng. Sci. *116*, 590–600. https://doi.org/10.1016/j.ces.2014.05.027.

237. Chaffart, D., and Ricardez-Sandoval, L.A. (2018). Optimization and control of a thin film growth process: a hybrid first principles/artificial neural network based multiscale modeling approach. Comput. Chem. Eng. *119*, 465–479. https://doi.org/10.1016/j.compchemeng.2018.08.029.

238. Grabow, L.C., Hvolbæk, B., and Nørskov, J.K. (2010). Understanding trends in catalytic activity: the effect of adsorbate-adsorbate interactions for Co oxidation over transition metals. In Topics in Catalysis, pp. 298–310. https://doi.org/10.1007/s11244-010-9455-2.

239. Lu, Y., Wang, B., Chen, S., and Yang, B. (2022). Quantifying the error propagation in microkinetic modeling of catalytic reactions with model-predicted binding energies. Mol. Catal. *530*, 112575. https://doi.org/10.1016/j.mcat.2022.112575.

240. Schlexer Lamoureux, P., Winther, K.T., Garrido Torres, J.A., Streibel, V., Zhao, M., Bajdich, M., Abild-Pedersen, F., and Bligaard, T. (2019). Machine learning for computational heterogeneous catalysis. ChemCatChem *11*, 3581–3601. https://doi.org/10.1002/cctc.201900595.

241. Guan, Y., Chaffart, D., Liu, G., Tan, Z., Zhang, D., Wang, Y., Li, J., and Ricardez-Sandoval, L. (2022). Machine learning in solid heterogeneous catalysis: recent developments, challenges and perspectives. Chem. Eng. Sci. *248*, 117224. https://doi.org/10.1016/j.ces.2021.117224.

242. Ulissi, Z.W., Tang, M.T., Xiao, J., Liu, X., Torelli, D.A., Karamad, M., Cummins, K., Hahn, C., Lewis, N.S., Jaramillo, T.F., et al. (2017). Machine-learning methods enable exhaustive searches for active Bimetallic facets and reveal active site motifs for CO2 reduction. ACS Catal. *7*, 6600–6608. https://doi.org/10.1021/acscatal.7b01648.

243. Yang, Y., Jiménez-Negrón, O.A., and Kitchin, J.R. (2021). Machine-learning accelerated geometry optimization in molecular simulation. J. Chem. Phys. *154*, 234704. https://doi.org/10.1063/5.0049665.

244. Ulissi, Z.W., Medford, A.J., Bligaard, T., and Nørskov, J.K. (2017). To address surface reaction network complexity using scaling relations machine learning and DFT calculations. Nat. Commun. *8*, 14621. https://doi.org/10.1038/ncomms14621.

245. Lan, T., and An, Q. (2021). Discovering catalytic reaction networks using deep reinforcement learning from first-principles. J. Am. Chem. Soc. *143*, 16804–16812. https://doi.org/10.1021/jacs.1c08794.

246. Singh, A.R., Rohr, B.A., Gauthier, J.A., and Nørskov, J.K. (2019). Predicting chemical reaction barriers with a machine learning model. Catal. Lett. *149*, 2347–2354. https://doi.org/10.1007/s10562-019-02705-x.