Title
Applications of Multiscale Atomistic Modeling for Materials Discovery

Author
Sours, Tyler

Publication Date
2023

Peer reviewed|Thesis/dissertation

Applications of Multiscale Atomistic Modeling for Materials Discovery

By

Tyler Sours

Dissertation

Submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Chemical Engineering

in the

Office of Graduate Studies

of the

University of California

Davis

Approved:

Ambarish Kulkarni, Chair

Roland Faller

Coleman Kronawitter

Committee in Charge

2023

Dedicated to my lovely wife, Marina.

Contents

List of Figures

List of Tables

Abstract

Applications of Multiscale Atomistic Modeling for Materials Discovery

With advances in computing technology and data science approaches, computational material design is becoming an increasingly reliable and powerful tool for guiding experimental investigations. By constructing databases of known or hypothetical structures of interest, atomistic simulations can be performed in parallel to screen all structures and identify the most promising candidate materials for a given application. This dissertation will highlight some of the diverse applications where computational screening can be implemented to gain new insight about different classes of materials. To begin, a study leveraging density functional theory (DFT) calculations to screen a class of bimetallic porphyrin-based metal-organic frameworks (MOFs) for electrocatalytic reduction of oxygen in fuel cell devices is presented. The highly tunable 3-dimensional pore spaces of these MOFs are shown to provide ideal catalytic environments that surpass the performance of commonly used 2-dimensional surface-based electrocatalysts (e.g., platinum). Next, a systematic approach that combines theory and experiment to characterize active sites in supported atomically dispersed catalysts is discussed. By creating a comprehensive DFT library of possible catalytic sites and comparing simulations with several complementary experimental characterization techniques, an atomic-level understanding of atomically dispersed platinum on magnesium oxide is elucidated. Finally, acknowledging the limitations of quantum mechanics-based simulations due to computational expense, I highlight how machine learning interatomic potentials (MLPs) are revolutionizing atomistic simulations and how easy-to-use open-source software packages are increasing the throughput of the scientific community. Specifically, I discuss the workflows behind generating a large DFT data set of pure silica zeolite configurations, training an accurate and transferable MLP for these systems, and calculating various material properties. Taken together, the work herein will showcase the versatility of molecular modeling approaches, while emphasizing the central roles of experimental collaborations, machine learning, and software development.

Acknowledgments

First, I would like to thank my family for their constant love and support and for allowing me the freedom to pursue my own interests. Thank you to my wife for keeping me somewhat sane during my graduate studies and during COVID. Thank you to the faculty in the ECH/MSE departments for creating an environment focused on interdisciplinary collaboration. Thank you to the staff (Ryan and Grace, in particular) for always quickly addressing any issue I encountered during my PhD. Thank you to all the members of the Kulkarni group for helpful discussions and general support: Rachita, Sam, Jiawei, Chenghan, Bobby, Dexter, Austin, Trevor, Kun-Lin, Sudheesh, Saurabh, Rajat, Siddharth, Daniel, Makena, and Artem. Thank you to Prof. Faller and Prof. Kronawitter for the feedback on my dissertation and Qualifying Exam. Finally, a special thank you to my mentor and friend, Prof. Ambarish Kulkarni. He has fully supported me in all my academic, professional, and personal endeavors, and I genuinely could not have made it to this point without him. As a first-generation student who wasn't entirely convinced that I belonged in a PhD program when I joined, Ambar's constant encouragement and belief in my abilities has been truly transformative in my self-confidence as a researcher. His heart is as large as his breadth of knowledge, and I will be forever grateful for everything he has done for me.

Chapter 1

Introduction

Emerging technologies in areas such as renewable energy require the discovery of new materials that exhibit ideal application-specific properties. With the existence of millions of potential materials from diverse classes like metal-organic frameworks and high entropy alloys, a purely experimental approach to screening even a small fraction of the vast material space is impossible. High-throughput computational screening approaches provide a way to narrow down the search space to a select few candidate materials that warrant additional investigation.[1–3] This investigation can come with more rigorous theoretical modeling or with experimental validation, and ideally a combination of both where experimental feedback is used to refine the theoretical models. These computational searches are becoming increasingly accurate, efficient, and exhaustive with advances in computing technology and wide-spread access to computing resources. Moreover, open-access release of the data produced in these searches[4,5] coupled with the rising popularity of data science and data-driven techniques is accelerating the rate at which we discover new materials and gain new insights about existing materials.

From electrocatalysis to materials characterization, the potential applications of computational screening studies are vast. While the science behind these applications can vary substantially, the similarities lie in the underlying workflows. These workflows typically consist of generating or collecting potential materials to create a search database, calculating properties of interest by performing atomistic simulations in parallel on all materials, and analyzing the resulting data to make observations about the materials

studied. This dissertation will highlight a couple of these diverse applications and investigate new machine learning methodologies for enhancing computational throughput. This section aims to outline the key foundational ideas behind the work presented in this dissertation, and it should be taken together with the content in Chapters 2–4.

1.1 Density Functional Theory

The majority of the results presented in this dissertation involves atomistic simulations using Density Functional Theory (DFT) calculations. DFT seeks an approximate solution to the time-independent Schrodinger equation and rests on the revolutionary theorems of Kohn and Hohenberg. These theorems prove that the ground state energy of a molecular system can be evaluated as a functional of electron density, reducing the 3N dimensionality of the Schrodinger equation to just three dimensions. Additionally, by minimizing this functional, the true electron density corresponding to the full Schrodinger equation solution is obtained. This is accomplished in practice by solving a set of single electron wave functions ($\psi_i(r)$), known as the Kohn-Sham equations,[6]

$$\left[\frac{-\hbar^2}{2m}\nabla^2 + V(r) + V_H(r) + V_{XC}(r)\right]\psi_i(r) = \epsilon_i\psi_i(r) \tag{1.1}$$

There are three potentials present on the left hand side of Equation 1.1: $V(r)$, $V_H(r)$, and $V_{XC}(r)$. The first describes the interaction between each electron and the fixed atomic nuclei. The second, $V_H(r)$, known as the Hartree potential, describes the electron-electron interactions using a mean-field approach that considers each electron interacting with the surrounding electron density at position r. The final potential, $V_{XC}(r)$ (the exchange-correlation potential), considers all additional quantum-mechanical effects and corrects for the self-interaction error in $V_H(r)$ induced by the mean-field approximation (i.e., each electron is interacting with the surrounding electron density, which includes the electron itself). The exchange-correlation potential is of particular practical importance to DFT, as many functionals exist of varying complexity that are more or less appropriate to use for any given atomic system. The most common functionals belong to a class that only use the first derivative of the electron density in $V_{XC}(r)$. This is known as the generalized

gradient approximation (GGA), and these functionals are thus broadly labeled as GGAs. More complicated forms of $V_{XC}(r)$ that more rigorously consider electron exchange effects are hybrid functionals, and Chapter 2 will discuss instances where these more complex (and thus more computationally expensive) functionals are necessary.

1.2 Computational Electrochemistry

Transitioning to an energy infrastructure featuring hydrogen as a fuel source is a promising solution to curb our reliance on fossil fuels given hydrogen's carbon-free emissions and ability to be sustainably produced from a renewable energy source.[7,8] A hydrogen fuel economy requires efficient electrochemical technology capable of producing hydrogen from water, then storing and later consuming hydrogen in a fuel cell device to deliver power on-demand. Fuel cells convert the energy stored in the chemical bond of hydrogen into an electrical current driven by the electrocatalytic oxygen reduction reaction (ORR) at the device's cathode, and further advancements in current technology are needed to make this process more feasible.[9,10] The high cost and thermodynamic inefficiency of the most commonly used ORR electrocatalysts (e.g., Pt/C) are seen as the primary bottlenecks in fuel cell development.[11] There is great demand for new materials, and computational screening methods coupled with electrochemical models employing DFT are enabling rapid discovery of promising new electrocatalysts to investigate experimentally.

1.2.1 Oxygen Reduction Reaction

The oxygen reduction reaction can follow the full four-electron reduction pathway to water, or the two-electron partial reduction pathway to hydrogen peroxide depending on the catalyst,

$$O_2 + 4(H^+ + e^-) \rightarrow 2H_2O; E^\circ = 1.23\ V \tag{1.2}$$

$$O_2 + 2(H^+ + e^-) \rightarrow H_2O_2; E^\circ = 0.7\ V \tag{1.3}$$

The four-electron reduction involves four proton-electron pair transfers and proceeds through three reaction intermediates: $*$OOH, $*$O, $*$OH,

$$O_2 + * + 4(H^+ + e^-) \rightarrow *OOH + 3(H^+ + e^-) \tag{1.4}$$

$$*OOH + 3(H^+ + e^-) \rightarrow *O + H_2O + 2(H^+ + e^-) \tag{1.5}$$

$$*O + 2(H^+ + e^-) \rightarrow *OH + H^+ + e^- \tag{1.6}$$

$$*OH + H^+ + e^- \rightarrow H_2O + * \tag{1.7}$$

where $*$ refers to the bare active site. The two-electron pathway involves only a single adsorbed $*$OOH intermediate and two successive proton-electron additions,

$$O_2 + 2(H^+ + e^-) \rightarrow *OOH + H^+ + e^- \tag{1.8}$$

$$*OOH + H^+ + e^- \rightarrow H_2O_2 \tag{1.9}$$

1.2.2 Computational Hydrogen Electrode

The works presented in Chapter 2 uses the computational hydrogen electrode (CHE) model developed by Nørskov[12] to incorporate the effect of the electrical potential into the theoretical predictions. Under the CHE model, the reference potential is set to that of the standard hydrogen electrode, thus at standard conditions (298 K, 1 bar, pH = 0), the chemical potential of a proton-electron pair is equal to that of hydrogen in the gas phase (i.e., the reaction $1/2H_2 \rightarrow H^+ + e^-$ is in equilibrium). This simple observation implies that the binding energies of the OXR intermediates can be then calculated using gaseous H_2 and H_2O as references,

$$\Delta E_{OOH} = E_{*OOH} - E_* - (2E_{H2O(g)} - 3/2E_{H2(g)}) \tag{1.10}$$

$$\Delta E_O = E_{*O} - E_* - (E_{H2O(g)} - E_{H2(g)}) \tag{1.11}$$

$$\Delta E_{OH} = E_{*OH} - E_* - (E_{H2O(g)} - 1/2E_{H2(g)}) \tag{1.12}$$

The electronic binding energies for the OXR intermediates are converted to free energies by including corrections for the change in zero-point energy (ΔZPE), entropic effects ($-T\Delta S$) and solvation (Δw).

$$\Delta G^0_{OXR} = \Delta E_{OXR} + \Delta ZPE - T\Delta S + \Delta w \tag{1.13}$$

The effect of the electrochemical cell potential is considered by shifting the free energies by neU where n is the number of proton-electron pairs and U is the applied potential vs RHE,

$$\Delta G_{OXR} = \Delta G^0_{OXR} - neU \tag{1.14}$$

The computational electrocatalytic performance is measured by the theoretical limiting potential (U_L), the potential at which one of the reaction steps stops being thermodynamically downhill. The individual reaction energies for each step are written as,

$$\Delta G_1 = \Delta G_{O2} - \Delta G_{OOH} = 4.92 - \Delta G_{OOH} \tag{1.15}$$

$$\Delta G_2 = \Delta G_{OOH} - \Delta G_O \tag{1.16}$$

$$\Delta G_3 = \Delta G_O - \Delta G_{OH} \tag{1.17}$$

$$\Delta G_4 = \Delta G_{OH} - \Delta G_{H2O} = \Delta G_{OH} \tag{1.18}$$

These equations are shown in the free energy diagram depicted in Figure 1.1. As each step consists of a single proton-electron pair transfer, the limiting potential for the overall reaction is equal to the least exergonic step in Equations 1.15-1.18,

$$U_L = min(\Delta G_1, \Delta G_2, \Delta G_3, \Delta G_4) \tag{1.19}$$

This limiting potential has been shown to correlate with experimental onset potentials,[13] and it can be related to the theoretical overpotential for 4e ORR as,

$$\eta = 1.23\ V - U_L \tag{1.20}$$

The overpotential represents the deviation from the equilibrium cell potential (1.23 V) and corresponds to thermodynamic inefficiencies (heat losses) in these devices. The limiting potential and overpotential serve as measures of catalyst performance.

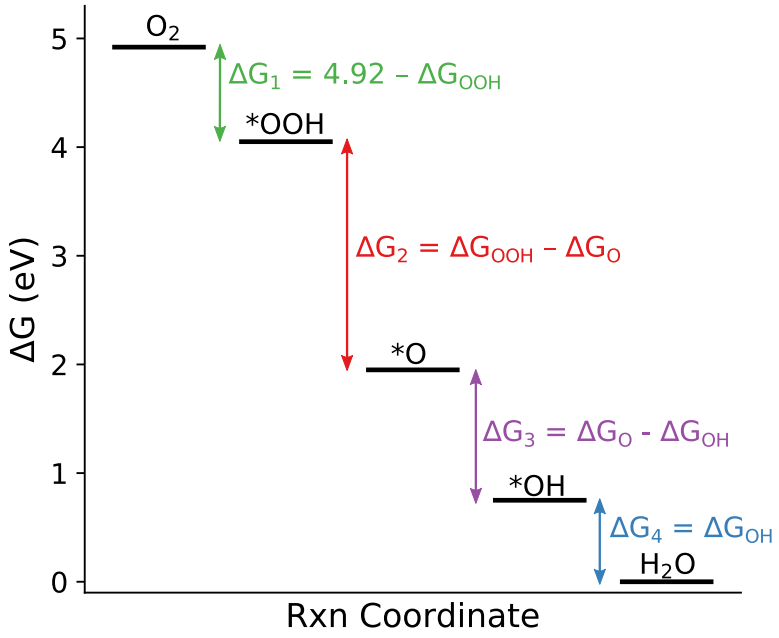

Figure 1.1. Example free energy diagram for the 4e ORR reaction. The differences in free energy between each reaction intermediate species correspond to each step's limiting potential, with the minimum of these corresponding to the overall limiting potential for the reaction.

1.2.3 Scaling Relationships and Volcano Plots

By inspection of Equation 1.19 and Figure 1.1, it can be seen that an ideal catalyst would bind each intermediate with a constant difference of 1.23 eV relative to the previous species (e.g., 3.69 eV for ∗OOH and 2.46 eV for ∗O). However, due to the existence of linear scaling between the reaction intermediate binding energies for a family of catalysts, individual tuning of the binding energies is impossible. For the commonly used 2-D surface catalysts (Pt, Pd, etc.), the binding energies have been shown to follow the relationships,[11]

$$\Delta G_{OOH} = \Delta G_{OH} + 3.2 \tag{1.21}$$

$$\Delta G_{O} = 2\Delta G_{OH} \tag{1.22}$$

These scaling relations exist due to the similar catalyst-oxygen bond that exists for all intermediates, with the slope of 2 for ΔG_O compared to ΔG_{OH} due to a double bond forming between the catalyst and oxygen adsorbate. As these energies are all correlated, a convenient consequence of these scaling relationships is the ability to predict the overall activity for a catalyst family as a function of only a single intermediate binding energy (typically chosen to be $*$OH). For 2-D surface catalysts, rearranging Equations 1.21 and 1.22 and substituting into Equations 1.15-1.17 shows each limiting potential can be written as functions of only ΔG_{OH} as,

$$\Delta G_1 = -\Delta G_{OH} + 1.72 \tag{1.23}$$

$$\Delta G_2 = -\Delta G_{OH} + 3.2 \tag{1.24}$$

$$\Delta G_3 = \Delta G_{OH} \tag{1.25}$$

and then the overall limiting potential corresponds to,

$$U_L = min(-\Delta G_{OH} + 1.72, -\Delta G_{OH} + 3.2, \Delta G_{OH}, \Delta G_{OH}) \tag{1.26}$$

This is typically represented graphically as a volcano-shaped plot with the legs of the volcano defined by the linear limiting potential equations above. Figure 1.2 shows the resulting volcano plot for common 2-D surface catalysts, and it can be seen that the peak of the volcano is well below the ideal value of 1.23 V - a direct consequence of the scaling relationships. For strongly binding metals on the left leg (blue in Figure 1.2, the overall activity is limited due to $*$ OH overbinding, while the metals on the right leg are limited due to $*$ OOH underbinding. This implies that 2-D surface catalysts will always

be subject to thermodynamic inefficiencies and an entirely new family of catalysts that circumvent these scaling relationships is needed to achieve more active electrocatalysts.

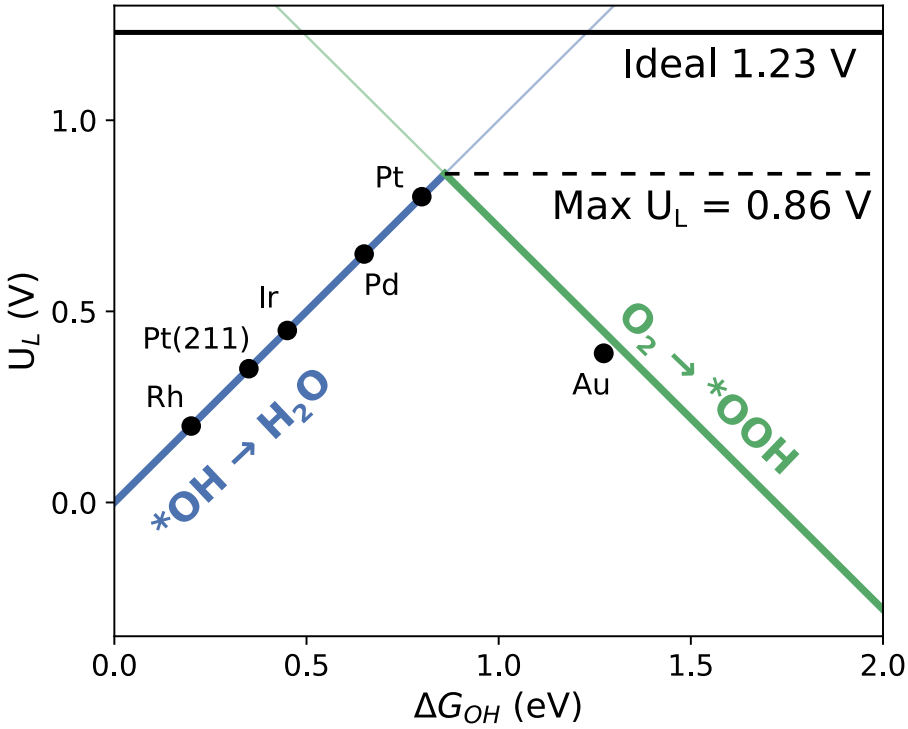

Figure 1.2. Volcano plot for common 2-D surface electrocatalysts showing the reduced maximum achievable efficiency due to linear scaling relationships. Reproduced from Kulkarni et al. 2018.[11]

As discussed in Chapter 2, a promising strategy to favorably alter these scaling relationships is through confinement of the reaction intermediates in a 3-D catalytic environment. To this end, metal-organic frameworks (MOFs), nanoporous materials consisting of metal nodes joined together by organic linking ligands, are promising due to the vast synthetic space enabling highly-customizable materials. By confining ORR to the 3-D MOF pore environments, selective stabilization of ∗OOH can be achieved. This selective stabilization favorably modifies the scaling relationships and shifts the right leg of the volcano up, resulting in a class of electrocatalysts with a higher achievable maximum limiting potential, as seen qualitatively in Figure 1.3.

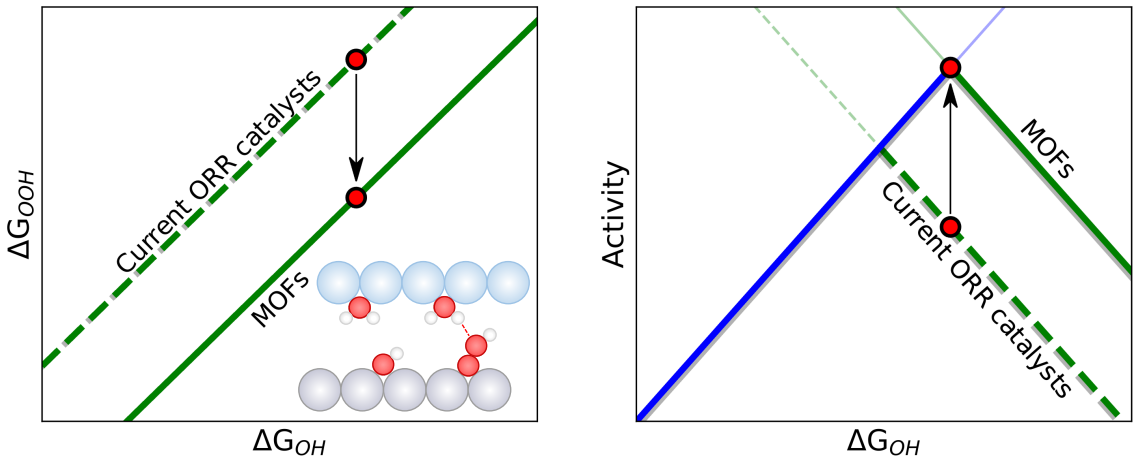

Figure 1.3. Effect of confining ORR inside 3-D MOF pore spaces allows for stabilization of ∗OOH relative to ∗OH, resulting in modified scaling behavior (left) and higher achievable limiting potentials (right).

1.3 Machine Learning Interatomic Potentials

While computational methods based on quantum mechanics, such as Density Functional Theory (DFT) used in this work, are often necessary to accurately describe many materials, these approaches suffer from high computational cost. With DFT, atomistic simulations are limited to small length (up to a few hundred atoms) and time scales (up to a few picoseconds), which significantly limits the feasible applications. High throughput screening of thousands or millions of materials or simulating long dynamic processes (e.g., diffusion with molecular dynamics (MD)) are rendered computationally intractable with DFT. Machine learning interatomic potentials (MLPs), also known as machine learning force fields (ML FFs), are revolutionizing atomistic simulations by significantly reducing computational cost compared to DFT without significant losses in accuracy. By training an MLP model on a suitable set of atomic configurations from first-principles calculations, the model learns the complex potential energy surface (PES) of the systems included in training.[14,15] Simulations can then be performed with the MLP performing the energy and force evaluations, circumventing the need to do any expensive quantum mechanical calculations. Numerous forms of MLPs exist, typically classified as either kernel or neural network methods, and new formulations and open-source software packages for working with MLPs are constantly appearing. The work in this dissertation utilizes the Deep Po-

tential (DP) MLP method[16] as implemented in the open-source DeePMD-kit[17] package, and the following sections will discuss the essential concepts behind this approach.

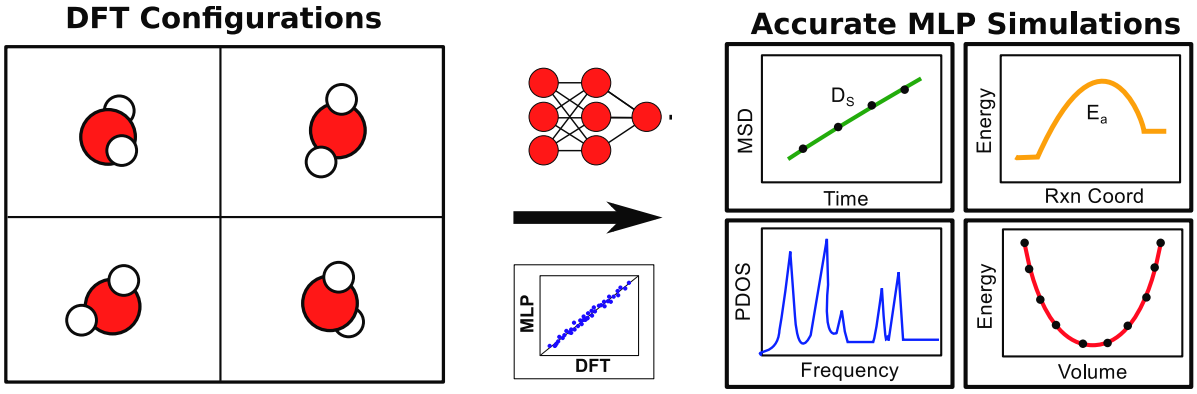

Figure 1.4. Effect of confining ORR inside 3-D MOF pore spaces allows for stabilization of ∗OOH relative to ∗OH, resulting in modified scaling behavior (left) and higher achievable limiting potentials (right).

1.3.1 Neural Networks

Neural Networks (NNs) are among the most popular machine learning algorithms and were originally developed to mimic biological neurons in the human brain.[18] They are powerful regression tools used to represent complex and highly nonlinear data by training on a large set of reference data. NNs function as connected layers of neurons that transform inputs to produce an output that is passed between layers if a threshold value is achieved. Deep NNs feature multiple layers and consist of an input layer, one or more hidden layers, and an output layer of interconnected neurons. At each neuron layer i, the vector of inputs ($\vec{x}_{i-1}$) from the previous layer's neurons are mapped to the output ($\vec{x}_i$) as,

$$\vec{x}_i = S_a(\mathbf{W}\vec{x}_{i-1} + \vec{b}_i) \tag{1.27}$$

where $\mathbf{W}$ is a matrix of weights for each input, $\vec{b}_i$ is a constant bias or threshold added to each weighted input, and S_a is an activation function to modulate the "firing" of the neuron. The weights and biases are assigned randomly and then optimized during training by minimizing a loss function (e.g., mean squared error) that measures the deviation of output values from a set of reference (training) data points.[19]

1.3.2 Deep Potential Method

The DP method has been applied to represent the PES of many diverse atomistic systems.[20–24] The generality of this approach can be attributed to six key considerations:

1. Model inputs consist of only atomic coordinates and chemical species (i.e., element of each atom).
2. The model is size extensive.
3. The model preserves natural symmetries: invariant for translational, rotational, and permutational symmetries.
4. No human intervention or parameterization is needed.
5. The model is smooth and continuously differentiable (necessary for force predictions).
6. The model is accurate for both finite and periodic systems.

The DP method evaluates the energy of a given atomic configuration by predicting single-atom energy contributions from each atom's local environment, which are then summed to obtain the overall system's energy prediction. This consists of first, for each atom i, transforming the Cartesian coordinates of each neighboring atom j within a specified cutoff distance (r_c) to localized coordinates relative to the position of atom i,

$$E = \sum E_i \tag{1.28}$$

$$\{x_{ij}, y_{ij}, z_{ij}\} \rightarrow \{s(r_{ij}), X_{ij}, Y_{ij}, Z_{ij}\} \tag{1.29}$$

$$X_{ij} = \frac{s(r_{ij})x_{ij}}{r_{ij}} \tag{1.30}$$

$$Y_{ij} = \frac{s(r_{ij})y_{ij}}{r_{ij}} \tag{1.31}$$

$$Z_{ij} = \frac{s(r_{ij})z_{ij}}{r_{ij}} \tag{1.32}$$

where r_{ij} is the distance between atom i and atom j and s is a function that gives more weight to closer atoms and avoids discontinuities at r_c by smoothing values to zero for distances between r_c and a smoothing cutoff r_s,

$$\begin{cases} \frac{1}{r_{ij}} & r_{ij} < r_s \\ \frac{1}{r_{ij}} \left\{ \frac{1}{2} cos \left[\pi \frac{(r_{ij} - r_s)}{(r_c - r_s)} \right] + \frac{1}{2} \right\} & r_s < r_{ij} < r_s \\ 0 & r_{ij} > r_c \end{cases}$$

These transformed inputs are then passed through an embedding neural network to obtain a set of descriptors with rotational and permutational symmetry invariance. Finally, these transformed descriptors are passed through a final fitting neural net that produces a single atom energy contribution, which is then summed to the total system energy as shown in Equation 1.28. The atomic forces for each atom i are also predicted as the negative gradient of the energy with respect to its position. The NN parameters are tuned by minimizing the loss function (L),

$$L = \frac{1}{B} \sum p_e |E - E^{DFT}|^2 + p_f |F - F^{DFT}|^2 \tag{1.33}$$

where B is the size of the training batch and p_e and p_f are weighting prefactors that control the contributions of the energy and force mean squared errors to the total loss function. Note that the virial stress tensor data can also be included in training which would add an additional term to the above loss function; however, stresses were not included in the training for the work presented in this dissertation.

1.4 References

[1] Jürgen Hafner, Christopher Wolverton, and Gerbrand Ceder. Toward Computational Materials Design: The Impact of Density Functional Theory on Materials Research. *MRS Bulletin*, 31(9):659–668, sep 2006. doi: 10.1557/mrs2006.174.

[2] Stefano Curtarolo, Gus L.W. Hart, Marco Buongiorno Nardelli, Natalio Mingo, Stefano Sanvito, and Ohad Levy. The high-throughput highway to computational materials design. *Nature Materials*, 12(3):191–201, 2013. doi: 10.1038/nmat3568.

[3] Keith T. Butler, Jarvist M. Frost, Jonathan M. Skelton, Katrine L. Svane, and Aron Walsh. Computational materials design of crystalline solids. *Chemical Society Reviews*, 45(22):6138–6146, 2016. doi: 10.1039/c5cs00841g.

[4] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin A. Persson. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, jul 2013. doi: 10.1063/1.4812323.

[5] Lowik Chanussot, Abhishek Das, Siddharth Goyal, Thibaut Lavril, Muhammed Shuaibi, Morgane Riviere, Kevin Tran, Javier Heras-Domingo, Caleb Ho, Weihua Hu, Aini Palizhati, Anuroop Sriram, Brandon Wood, Junwoong Yoon, Devi Parikh, C. Lawrence Zitnick, and Zachary Ulissi. Open Catalyst 2020 (OC20) Dataset and Community Challenges. *ACS Catalysis*, 11(10):6059–6072, 2021. doi: 10.1021/acscatal.0c04525.

[6] David S. Sholl and Janice A. Steckel. *Density Functional Theory: A Practical Introduction*. John Wiley & Sons, 2009.

[7] J. A. Turner. Sustainable Hydrogen Production. *Science*, 305(5686):972–974, August 2004. doi: 10.1126/science.1103197.

[8] Mark K. Debe. Electrocatalyst approaches and challenges for automotive fuel cells. *Nature*, 486(7401):43–51, 2012. doi: 10.1038/nature11115.

[9] Zhi Wei Seh, Jakob Kibsgaard, Colin F. Dickens, Ib Chorkendorff, Jens K. Nørskov, and Thomas F. Jaramillo. Combining theory and experiment in electrocatalysis: Insights into materials design. *Science*, 355(6321), 2017. doi: 10.1126/science.aad4998.

[10] Joseph H. Montoya, Linsey C. Seitz, Pongkarn Chakthranont, Aleksandra Vojvodic, Thomas F. Jaramillo, and Jens K. Nørskov. Materials for solar fuels and chemicals. *Nature Materials*, 16(1):70–81, 2016. doi: 10.1038/nmat4778.

[11] Ambarish Kulkarni, Samira Siahrostami, Anjli Patel, and Jens K Nørskov. Understanding Catalytic Activity Trends in the Oxygen Reduction Reaction. *Chemical Reviews*, 118(5):2302–2312, mar 2018. doi: 10.1021/acs.chemrev.7b00488.

[12] J K Nørskov, J Rossmeisl, A Logadottir, L Lindqvist, J. R. Kitchin, T. Bligaard, and H. Jónsson. Origin of the Overpotential for Oxygen Reduction at a Fuel-Cell Cathode. *The Journal of Physical Chemistry B*, 108(46):17886–17892, nov 2004. doi: 10.1021/jp047349j.

[13] Venkatasubramanian Viswanathan, Heine Anton Hansen, Jan Rossmeisl, and Jens K. Nørskov. Universality in oxygen reduction electrocatalysis on metal surfaces. *ACS Catalysis*, 2(8):1654–1660, 2012. doi: 10.1021/cs300227s.

[14] Jörg Behler. Perspective: Machine learning potentials for atomistic simulations. *Journal of Chemical Physics*, 145(17), 2016. doi: 10.1063/1.4966192.

[15] Jörg Behler. First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems. *Angewandte Chemie International Edition*, 56(42):12828–12840, oct 2017. doi: 10.1002/anie.201703114.

[16] Jiequn Han, Linfeng Zhang, Roberto Car, and Weinan E. Deep Potential: A General Representation of a Many-Body Potential Energy Surface. *Communications in Computational Physics*, 23(3):1–13, 2018. doi: 10.4208/cicp.oa-2017-0213.

[17] Han Wang, Linfeng Zhang, Jiequn Han, and Weinan E. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Computer Physics Communications*, 228:178–184, jul 2018. doi: 10.1016/j.cpc.2018.03.016.

[18] Warren S. McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4):115–133, dec 1943. doi: 10.1007/BF02478259.

[19] Grégoire Montavon, Geneviève Orr, and Klaus-Robert Müller. *Neural networks: tricks of the trade*, volume 7700. springer, 2012.

[20] Siddarth K. Achar, Linfeng Zhang, and J. Karl Johnson. Efficiently Trained Deep Learning Potential for Graphane. *Journal of Physical Chemistry C*, 125(27):14874–14882, 2021. doi: 10.1021/acs.jpcc.1c01411.

[21] Hsin Yu Ko, Linfeng Zhang, Biswajit Santra, Han Wang, E. Weinan, Robert A. DiStasio, and Roberto Car. Isotope effects in liquid water via deep potential molecular dynamics. *Molecular Physics*, 117(22):3269–3281, 2019. doi: 10.1080/00268976.2019.1652366.

[22] Marcos F. Calegari Andrade, Hsin-Yu Ko, Linfeng Zhang, Roberto Car, and Annabella Selloni. Free energy of proton transfer at the water–TiO 2 interface from ab initio deep potential molecular dynamics. *Chemical Science*, 11(9):2335–2341, 2020. doi: 10.1039/C9SC05116C.

[23] Jinzhe Zeng, Linfeng Zhang, Han Wang, and Tong Zhu. Exploring the Chemical Space of Linear Alkane Pyrolysis via Deep Potential GENerator. *Energy & Fuels*, 35 (1):762–769, jan 2021. doi: 10.1021/acs.energyfuels.0c03211.

[24] Tongqi Wen, Rui Wang, Lingyu Zhu, Linfeng Zhang, Han Wang, David J. Srolovitz, and Zhaoxuan Wu. Specialising neural network potentials for accurate properties and application to the mechanical response of titanium. *npj Computational Materials*, 7 (1):206, dec 2021. doi: 10.1038/s41524-021-00661-y.

Chapter 2

Circumventing Scaling Relations in Oxygen Electrochemistry Using Metal–Organic Frameworks

Tyler Sours, Anjli Patel, Jens Nørskov, Samira Siahrostami, Ambarish Kulkarni

Adapted with permission from J. Phys. Chem. Lett., 2020, 11, 10029–10036.

2.1 Abstract

It has been well-established that unfavorable scaling relationships between ∗OOH, ∗OH, and ∗O are responsible for the high overpotentials associated with oxygen electrochemistry. A number of strategies have been proposed for breaking these linear constraints for traditional electrocatalysts (e.g. metals, alloys, metal-doped carbons); such approaches have not yet been validated experimentally for heterogenous catalysts. Development of a new class of catalysts capable of circumventing such scaling relations remains an ongoing challenge in the field. In this work, we use density functional theory (DFT) calculations to demonstrate that bimetallic porphyrin-based MOFs (PMOFs) are an ideal materials platform for rationally designing the 3D active site environments for oxygen reduction reaction (ORR). Specifically, we show that the ∗OOH binding energy and the theoretical limiting potential can be optimized by appropriately tuning the transition metal active site, the oxophilic spectator, and the MOF topology. Our calculations predict theoretical limiting potentials as high as 1.07 V for Fe/Cr-PMOF-Al, which exceeds the Pt/C benchmark for 4e ORR. More broadly, by highlighting their unique characteristics, this

works aims to establish bimetallic porphyrin-based MOFs as a viable materials platform for future experimental and theoretical ORR studies.

2.2 Introduction

Oxygen electrochemistry in the form of oxygen reduction reaction (ORR, for fuel cells) and oxygen evolution reaction (OER, for water electrolyzers) lies at the heart of an environmentally sustainable hydrogen-based economy.[1] However, widespread adoption of these technologies has been limited due to the high cost and thermodynamic losses associated with ORR and OER electrocatalysts (referred together as OXR). Despite large research investments, only modest improvements have been achieved in the catalyst performance over the past decade; the best catalysts show onset potentials that deviate $\sim$ 0.3–0.4 V away from the ideal value (i.e. 1.23 V_{RHE}).[2–4] These losses have been attributed to the unfavorable scaling between the binding energies of OXR intermediates, in particular $*$OOH vs $*$OH.[2,5,6] Although recent discoveries of low-cost alternatives are encouraging,[7,8] it is likely that their performance is also constrained by similar intrinsic limitations. Despite slight material-specific variations, the robustness of the $*$OOH vs $*$OH linear correlation across different materials and computational methods is well-established.[9,10] Specifically, a slope of 0.9 - 1.0 is observed (due to the single bond with the surface), with an intercept of 2.8 – 3.2 eV (due to the peroxyl bond in $*$OOH).[5,11–14] Designing materials that circumvent this linear scaling relationship is crucial for the development of the next generation of OXR electrocatalysts.[11,15] A promising strategy towards this goal is the confinement of $*$OOH within a 3-D active site environment, which has been demonstrated for a few model systems and homogenous molecular catalysts. For instance, Rossmeisl et al.[16,17] have shown that diporphyrin motifs are capable of facilitating $*$OOH dissociation via two closely-spaced binding sites. This approach bypasses the limitations imposed by $*$OOH scaling in favor of the less restrictive dissociated $*$O + $*$OH pathway. Following some of the thermodynamic optimization guide-lines discussed by Calle-Vallejo and colleagues,[18] we aim to preferentially stabilize $*$OOH (relative to $*$OH) by tuning the surrounding active site environment.

2.3 Results and Discussion

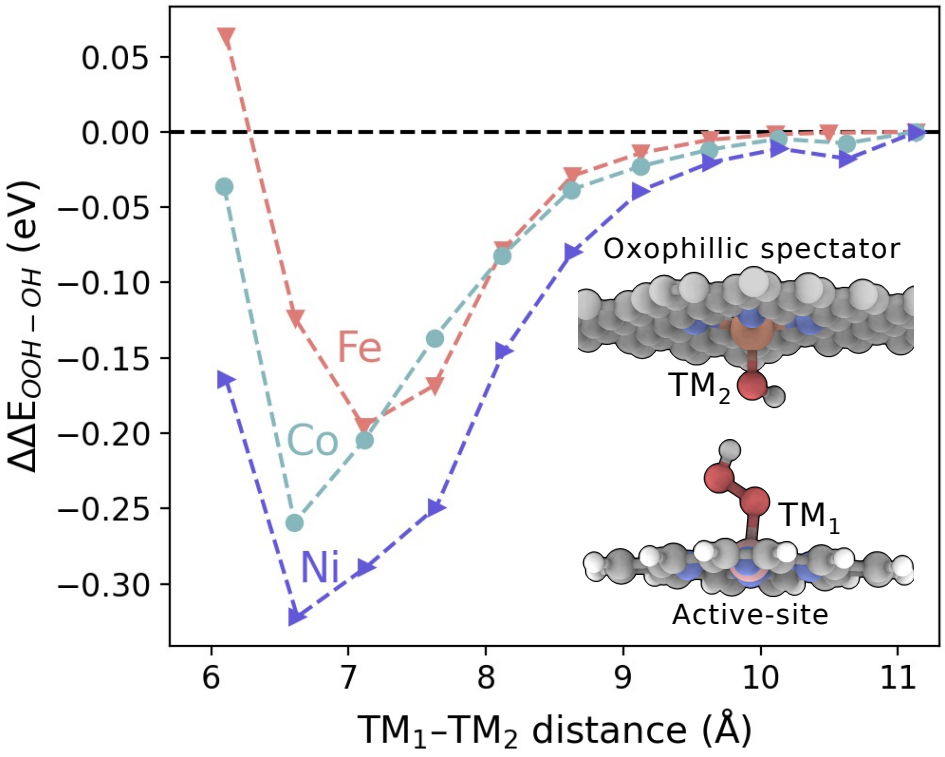

Figure 2.1. Preferential stabilization of ∗ OOH relative to ∗OH at different distances compared to the non-interacting system for various active sites (TM1 = Fe, red; Co, light blue; Ni, dark blue) with an oxophilic Fe-OH spectator (TM2). $\Delta\Delta E_{OOH-OH} = \Delta E_{OOH-OH} - \Delta E^{\infty}_{OOH-OH}$ where ΔE_{OOH-OH} represents the difference between the ∗OOH and ∗OH binding energies at various TM1-TM2 distances and $\Delta E^{\infty}_{OOH-OH}$ represents difference between the ∗OOH and ∗OH binding energies for the non-interacting systems (i.e., at large TM1-TM2 distances). Inset shows active TM1-OOH stabilized by hydrogen-bonding interactions with TM2-OH spectator embedded in graphene.

Similar to the co-facial Pacman molecules,[19–21] Figure 2.1 shows a related model system consisting of two transition metals (TMs). Specifically, Fe/N_4-doped (TM2 = Fe) graphene sheet is placed in close proximity with a TM porphyrin molecule (TM1 = Fe, Co and Ni). The Fe center embedded in the graphene acts as an inactive spectating species that preferentially stabilizes TM1-∗OOH via hydrogen bonding interactions. An additional stabilization of 0.2 – 0.35 eV is observed for ∗OOH at distances of ∼ 7.0 Å (using the BEEF-vdW[22] functional) compared to the non-interacting system (i.e. at large separations). Validation with other functionals (e.g. RPBE,[23] BLYP,[24,25] B3LYP[26] and

HSE06,[27,28] Grimme's D3BJ[29–31]) show that the stabilization (1) is not an artifact of using a specific functional, (2) is not dominated by van der Waals interactions (Figure 2.2) and (3) changes the TM1-O bond order affecting both the "slope" and the "intercept" of the scaling lines (see Figure 2.3). The preferential stabilization of ∗OOH relative to ∗OH arises due to the favorable confined geometry, which provides a design principle to guide the development of active OXR catalysts. This finding leads to an intriguing question: can we design other materials that leverage favorable spatial orientations of binding sites to circumvent OXR scaling relations? Although related concepts have been proposed,[32,33] to the best of our knowledge, no heterogenous catalysts have been experimentally proven to take advantage of these effects for ORR. Within the field of electrochemistry, it is critical to translate these strategies to an experimentally-synthesizable and industrially-scalable class of materials.[12]

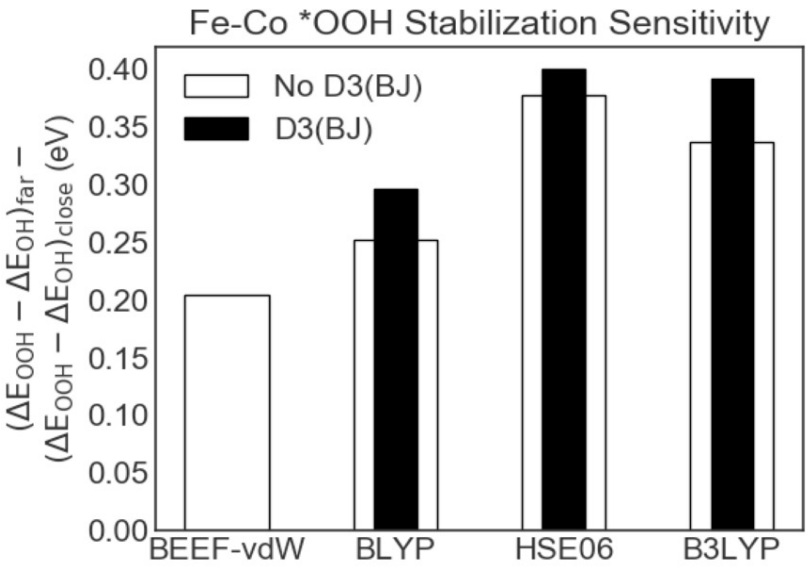

Figure 2.2. Functional dependency and Van der Waals contributions for the additional stability of ∗OOH relative to ∗OH for Co-porphine interacting with Fe-OH spectator embedded in graphene.

Metal-Organic Frameworks (MOFs), a class of nanoporous materials consisting of metal nodes interconnected by organic linkers, are uniquely attractive due to their experimentally realizable diversity of active sites, 3-D porous chemical environments, and topologies.[34–36] Although MOF electrochemistry is a relatively nascent direction, a va-

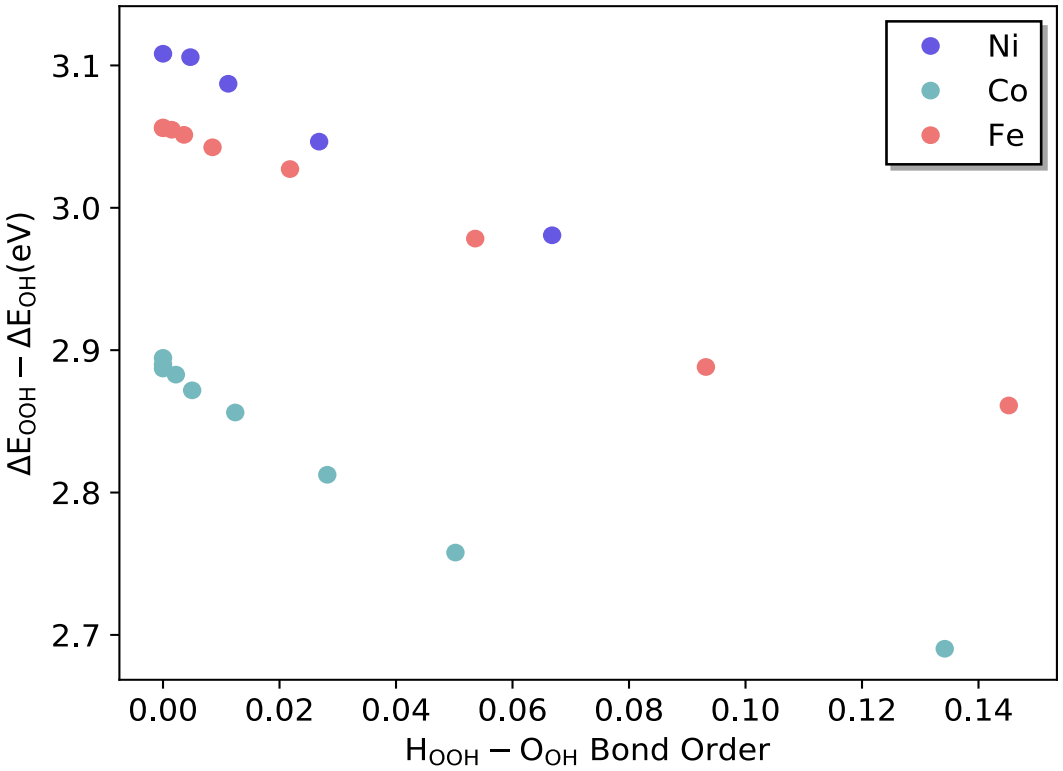

Figure 2.3. Difference in ∗OOH and ∗OH binding energies as a function of hydrogen bond strength formed between ∗OOH and the spectating Fe-OH for different metals.

riety of common MOFs, such as Co-PMOF-Al,[37] PCN-223(Fe),[38] PCN-226(Co),[39] and Ni-HAB[40] show intrinsic ORR activity. The experimental onset potentials reported for these MOF systems range from 0.7 – 0.83 V, which is comparable with other promising catalysts. Progress in MOF electrocatalysis, including possible challenges associated with MOF stability and charge transfer, has been summarized in a number of recent reviews.[41–44] In the remainder of this work, we use periodic density functional theory (DFT) calculations to show that bimetallic porphyrin-based, mixed-linker MOFs can be designed to preferentially stabilize ∗OOH. Specifically, by using a combination of active and oxophilic transition metal cations, we leverage the 3-D pore structure of MOFs to spatially orient the porphyrin binding sites and circumvent the scaling relations. We consider the full four-electron associative reduction pathway to water and the partial two-electron pathway to hydrogen peroxide (further details are presented in Chapter 1). It should be

noted that alternative mechanisms have been reported for some M-N_4 catalysts under certain conditions.[45–49] However, an exhaustive investigation into all possible mechanistic pathways is beyond the scope of this work. The theoretical limiting potential (U_L), defined as the highest potential at which all reaction steps remain thermodynamically downhill, is used to evaluate electrocatalytic performance. Although U_L is determined entirely from thermodynamics, it has been shown to correlate with ORR activity.[6] Figure 2.6a shows the structure of PMOF-Al,[37,50] which consists of TCPP ligands (TCPP = tetrakis(4-carboxyphenyl)porphyrin) stacked on top of each other to form a 3-D structure connected via a 1-D Al-oxide chain. PMOF-Al is water stable and has been shown to be experimentally active for ORR.[37] The spacing between the two linkers is $\sim$ 6.7 Å, which compares well with the favorable region in Figure 2.1. As different linkers can be incorporated in MOFs during synthesis or using post-synthetic methods,[51–53] we used DFT calculations to explore mixed-linker PMOF-Al configurations consisting of alternating TM1/TCPP and TM2/TCPP linkers (TM1 = Cr, Mn, Fe, Co, Ni; TM2 = Cr, Mn, Fe). Pourbaix diagrams are used to determine the coverage and oxidation states of the TMs while allowing simultaneous binding on both sides of the linker (Fig. 2.4). All possible spin combinations are explored to determine the lowest energy electronic structure for each TM (Table 2.4). Entropic corrections are calculated using the harmonic approximation (Table 2.5), and solvation contributions are estimated using VASPsol.[54] The solvent isosurface (Figure 2.5) confirms that the solvent penetrates the MOF cavity. Other solvation approaches are possible,[55] but are beyond the scope of this work. The calculated absolute solvation effects are summarized in Table 2.1.

Table 2.1. Absolute solvation effects (eV) for each porphyrin system.

Adsorbate	TCPP	PMOF-Al No spectator	PMOF-Al Fe-OH spectator
Bare	-1.18	-0.88	-0.91
$*$OOH	-1.39	-1.12	-1.05
$*$O	-1.43	-1.15	-1.14
$*$OH	-1.37	-1.10	-1.12

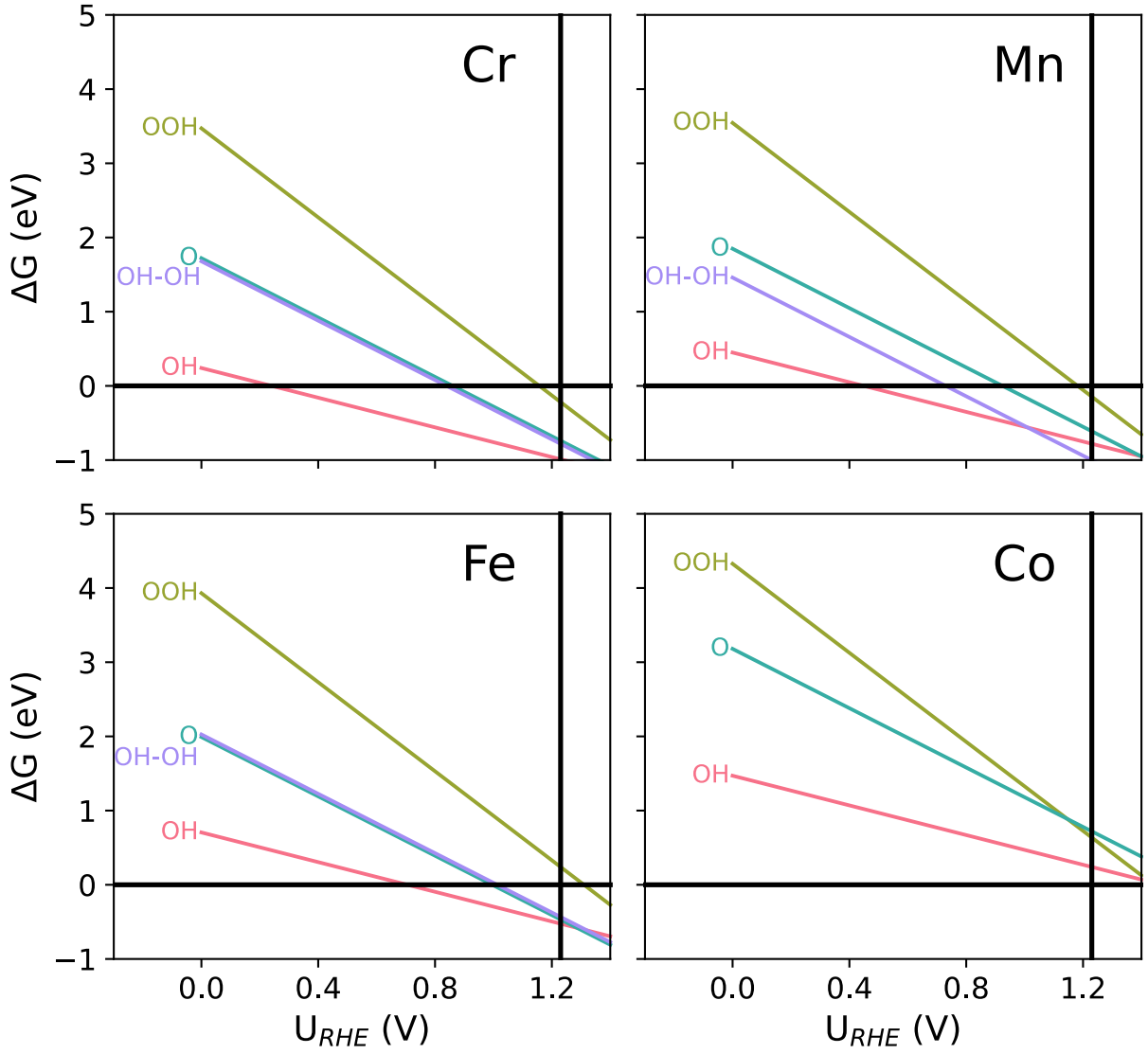

Figure 2.4. TPSSh-D3BJ predicted Pourbaix diagrams showing the coverage of each metal active site at increasing potential.

Figure 2.6b shows the DFT-calculated free energy diagram for the Co active site in monometallic and bimetallic PMOF-Al using TPSSh-D3BJ[56] functional. The theoretical limiting potential for the molecular Co/TCPP analog ($U_L = 0.52$ V) is consistent with the monometallic Co/PMOF system ($U_L = 0.59$ V, blue). Our calculated limiting potentials agree reasonably with the experimentally measured onset potentials (0.75 V),[37] further confirming the suitability of the computational protocol.

Similar calculations with RPBE-D3BJ predicts incorrect spin states and TM geometry, while the HSE06-D3BJ functional under-predicts binding energies for Co/TCPP and

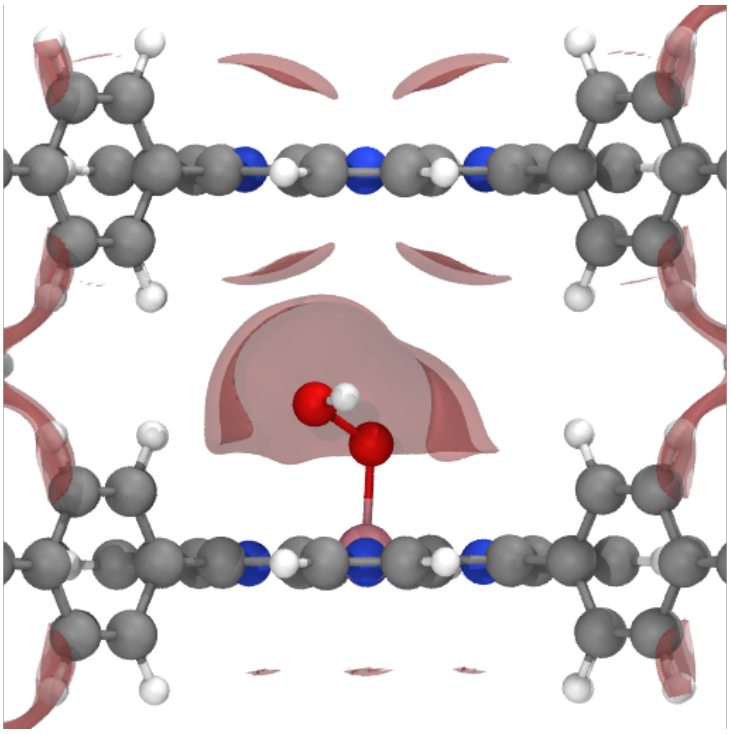

Figure 2.5. Solvent isosurface predicted by VASPsol for Co-∗OOH in the no spectator PMOF-Al system.

Fe/TCPP. These findings are inconsistent with experimental results and are not discussed further (see section 2.5.2 for details). Detailed benchmarking of different functionals (e.g. ωB97,[57] MN15[58] and others[59]) with coupled cluster theory and/or multireference methods for various transition metals is beyond the scope of this work.[55,60–62]

Compared to Co-PMOF-Al with no spectator, Figure 2.6b shows that the presence of an oxophilic spectator (Fe-OH, green) improves the predicted activity. For instance, the theoretical limiting potential for the bimetallic Co/Fe-OH/PMOF-Al catalyst (0.81 V) is comparable to the Pt/C benchmark (U_L = 0.8 V). The improved activity originates from the additional 0.22 eV stabilization of ∗OOH due to the presence of the Fe-OH spectator. Simulations of other oxophilic TMs (see Mn-OH and Cr-OH in Figure 2.6) confirms that the presence of the spectator ligand significantly affects ∗OOH binding; the other OXR adsorbates (i.e. ∗OH and ∗O) are largely unaffected. While the above discussion is limited to the Co active site, we perform additional DFT calculations with Ni, Fe, Cr, and Mn active sites. For each TM, the most stable coverage predicted by the Pourbaix diagram

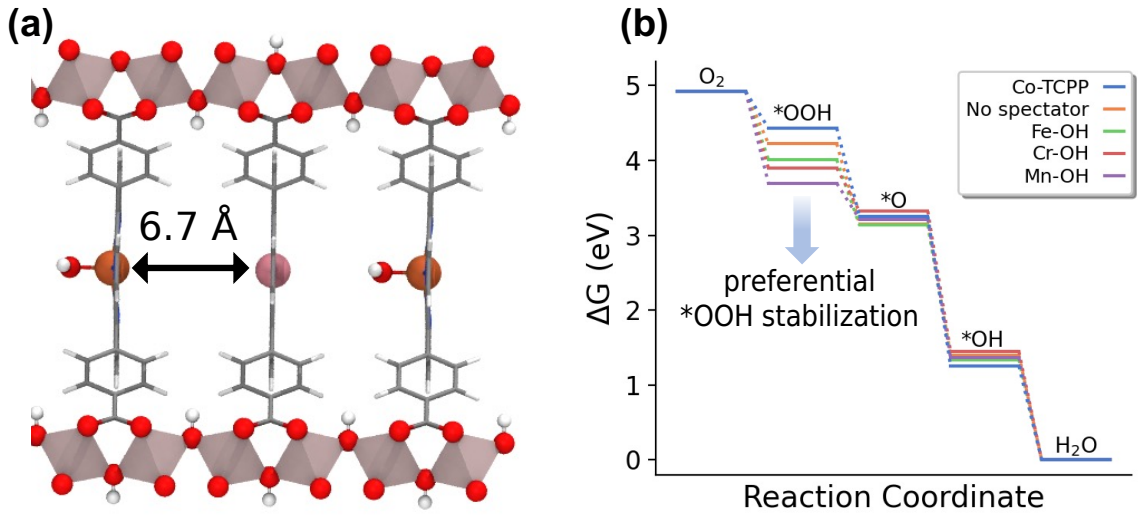

Figure 2.6. (a) Structure of bimetallic PMOF-Al with Co active site and Fe-OH spectator, (b) Free energy diagram (TPSSh-D3BJ) showing the preferential stabilization of ∗OOH. Color scheme: Co (pink), Fe (brown), C (grey), O (red), and H (white).

is used (e.g. Cr-OH, Mn-OH). As shown in Fig. 3a, we predict similar scaling slopes for TCPP (0.66, orange) and monometallic PMOF-Al (0.60, blue). The deviation from unity likely arises due to the changes in the spin states for ∗OOH and ∗OH intermediates and differences in the active site coordination geometry (e.g. square planar Co vs. distorted square pyramidal Cr-OH). Interestingly, previous work by Busch et al. has shown that the slope of the scaling line depends on the choice of the DFT functional.[46]

More interestingly, the presence of the Fe-OH spectator results in a favorable deviation from the above scaling behavior. Specifically, we observe a preferential ∗OOH stabilization of up to 0.4 eV for various TM active sites (Fig. 2.8). A smaller deviation is observed for strongly binding TMs (e.g. Fe-∗OOH, 0.07 eV). In contrast, weakly binding TMs (e.g. Ni-∗OOH) are stabilized to a larger extent (0.39 eV). For a given TM active site, the extent of ∗OOH stabilization also depends on the identity of the spectator. For instance, Mn-OH spectator (purple, Figure 2.7a) results in systematically higher stabilization than Fe-OH spectator. In all cases, the ∗O and ∗OH binding energies do not change significantly ($<$ 0.15 eV).

The trends in ∗OOH stabilization can be explained by the increasing electronegativity of the spectating metal (Mn $<$ Cr $<$ Fe). Spectating metals with lower electronegativity

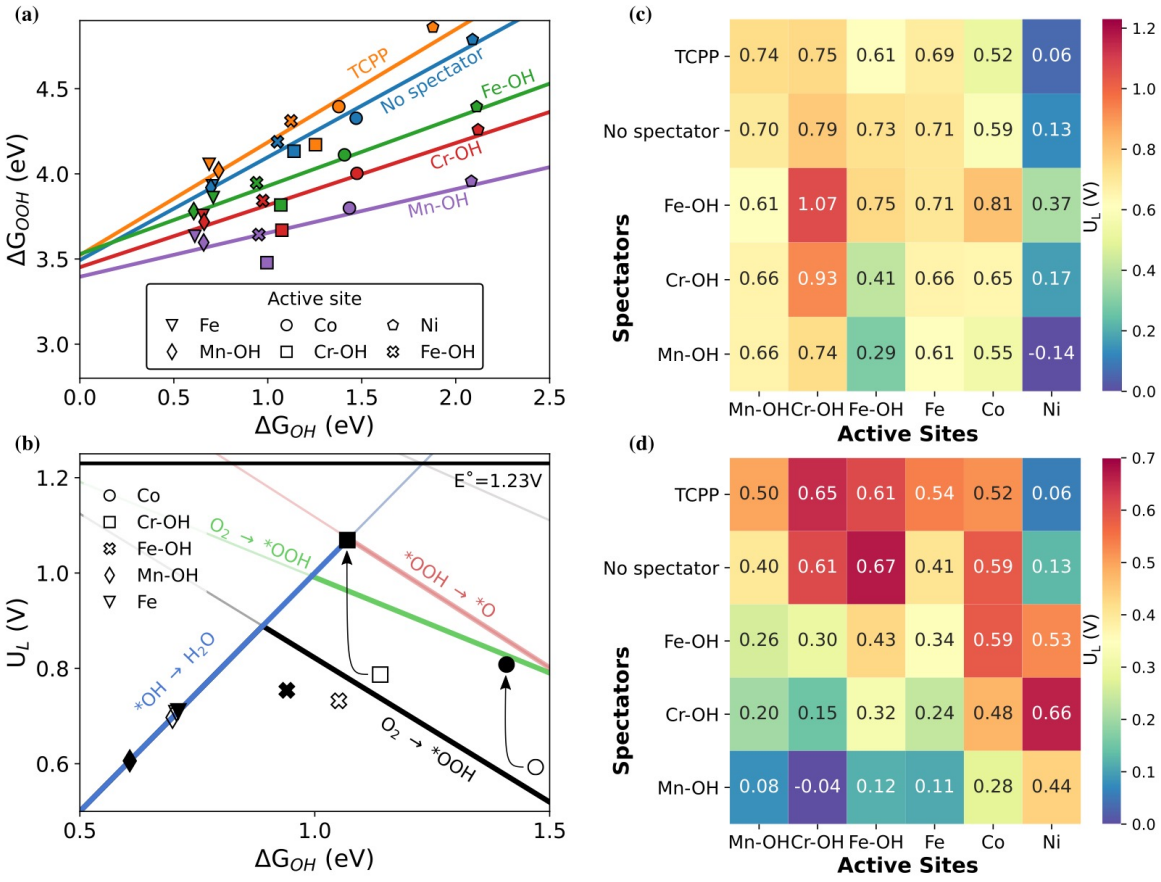

Figure 2.7. (a) Comparison of *OOH vs. *OH scaling lines for TCPP, monometallic and bimetallic PMOF-Al with different active sites and spectators, (b) 1-D volcano plot showing the improvement in activity (black arrows) due to the presence of Fe-OH spectator (filled symbols, red and green volcano legs) compared to the no spectator (empty symbols, black volcano leg) scenario. TPSSh-D3BJ calculated limiting potential for (c) 4e and (d) 2e ORR for various bimetallic MOFs.

allow for more charge to be localized on the spectator ligand (OH), resulting in a stronger hydrogen bond and increased stability of the ∗OOH intermediate. Table 2.2 shows the differences in the calculated bond orders for Co active site with different spectators. In particular, the H_{OOH}–O_{OH} bond order increases from 0.23 (Fe-OH spectator) to 0.33 (Mn-OH spectator) confirming a stronger hydrogen bond. Bond order analysis also reveals that the presence of the spectator localizes more charge on the ∗OOH adsorbate compared to the no spectator system. The Co-O_{OOH} bond order decreases from 0.57 to 0.48 with the Mn-OH spectator. Although the weakened TM-oxygen bond may favor the 2e pathway, detailed investigation of kinetic effects and product selectivity (i.e. H_2O_2 vs. H_2O) is

beyond the scope of this work.

As the spectator stabilizes ∗OOH, the improvement in the ORR limiting potential is well-explained by the 1-D volcano plot in Figure 2.7b. Specifically, for TMs that lie on the right leg of the volcano (i.e. activity is limited by $O_2 \rightarrow$ ∗OOH), the spectator stabilizes ∗OOH, improves the ∗OOH vs. ∗OH scaling, and favorably shifts the right leg of the volcano. This is depicted by black arrows in Figure 2.7b corresponding to an improvement in the activity of Co and Cr-OH active sites due to the Fe-OH spectator (green line, filled symbols) compared to the no spectator case (black line, empty symbols). The predicted limiting potential for the bimetallic Cr/Fe-PMOF-Al catalyst is 1.07 V, which exceeds the performance of the benchmark Pt/C catalyst (U_L = 0.8 eV). TPSSh-D3BJ suggests Cr-OH active site possesses high activity for all spectator combinations, and the limiting potentials for Cr-OH TCPP (0.75 V) and no spectator (0.79 V) agree very well with the experimental half-wave potential (0.77 V) of a recently reported Cr-N_4 catalyst.[63] The strongly-binding metals on the left leg of the volcano (e.g. Mn-OH) do not benefit from the addition of the spectator, as ∗OOH formation is not limiting for these catalysts.

Table 2.2. Calculated DDEC bond orders for Co–∗OOH with different spectators.

Spectator	Pauling electronegativity	Bond 1*	Bond 2*	Bond 3*	Bond 4*
No spectator	–	0.57	1.25	0.79	–
Fe-OH	1.83 (Fe)	0.52	1.30	0.66	0.23
Cr-OH	1.66 (Cr)	0.49	1.32	0.63	0.24
Mn-OH	1.55 (Mn)	0.48	1.36	0.56	0.33

*Refer to Figure 2.9 for identifying numbered bonds.

We emphasize that circumventing scaling relations via ∗OOH stabilization does not guarantee enhanced activity, and only catalysts in which the potential-determining step (i.e. the least thermodynamically favorable reaction step) involves ∗OOH are influenced by this approach.[64,65] Moreover, although the potential-dependent step has been shown to correlate with kinetics,[6,66,67] it may not provide a quantitative description of the reaction rates.[68–71] These results are further summarized in the heat-map in Figure 2.7c, which

includes Mn-OH and Cr-OH spectators. While Fe-OH spectator tends to improve the 4e limiting potentials, Mn-OH and Cr-OH do not have the same effect. Specifically, Mn-OH and Cr-OH overstabilize $*$OOH, and the activity becomes limited by the $*$OOH $\rightarrow$ $*$O step (see 1-D volcanoes in Figure 2.10). As the $*$OOH $\rightarrow$ $*$O step is often not limiting for metals and alloys, we emphasize that molecular and MOF-based ORR catalysts require a more careful analysis beyond the traditional $*$OOH vs $*$OH scaling for metals and alloys. Single atom M-N_4 catalysts only allow binding of the $*$O intermediate at the on-top site (as opposed to the more favorable hollow sites for metallic surfaces), which has already been shown to destabilize binding of $*$O.[72] This destabilization of $*$O combined with overstabilization of $*$OOH presents a new bottleneck for 4e ORR, where formation of $*$O from $*$OOH limits the achievable potential. While these effects are observed in all bimetallic combinations, the moderate stabilization of $*$OOH due to Fe-OH spectator allows for an improvement in theoretical overpotential, notably for the Co and Cr-OH active center. However, stronger stabilization seen with Mn-OH and Cr-OH spectators yields a reduction in catalytic performance as $*$OOH is stabilized to the point where $*$O formation is unfavorable.

Recognizing that this $*$O destabilization relative to metal surfaces may compromise the selectivity of these systems towards the 4e reduction to water, we extend our analysis to the 2e oxygen reduction reaction to form hydrogen peroxide. Figure 2.7d summarizes the calculated 2e limiting potentials for all active site and spectator combinations. As expected, for stronger binding active sites that lie on the left leg of the volcano, incorporation of an oxophilic spectator overstabilizes $*$OOH and reduces the 2e ORR activity. However, for weakly binding metals on the far right of the volcano, a noticeable improvement in theoretical limiting potential is obtained. For the Ni active site, TPSSh-D3BJ predicts highly active limiting potentials of 0.53 (0.66) V when combined with Fe-OH (Cr-OH) spectators, which represents a 0.4 eV improvement over the no spectator scenario. The above results indicate that porphyrin-based MOFs are a promising platform for tuning the binding energies of OXR adsorbates. Unlike metals and alloys where only the composition[73] or strain[74] can be varied, the 3-D active sites in MOFs offer unique

degree of control that is unavailable in other materials. Specifically, in addition to appropriately choosing the bimetallic system (i.e. TM1 and TM2), we can now potentially design MOF topologies to further tune the adsorbate binding energies for OXR. As an illustrative example, we consider a model system consisting of two interacting porphine molecules (Co active site, Fe-OH spectator) at varying distances. As shown in Figure 2.11a, parallel configurations can be considered as analogs of PMOF-Al with different 1-D metal oxide chains (e.g. substituting Al with Ga or In[75]), or other similar rod-based MOFs, such as MIL-173.[76] Additionally, a pair of angled (37°) porphines is chosen to mimic the topology of the porphyrinic Zr_6-oxo cluster MOF, PCN-225.[77] The distance between the TMs is varied to illustrate the reticular approach in MOF synthesis;[78] we demonstrate the effects of varying the linker size while retaining the same topology. Fig. 4a shows a high degree of control on reaction thermodynamics can be achieved by tuning the separation distance between adjacent porphyrin motifs, and the distances of several known MOFs are highlighted with vertical dotted lines. For both the parallel and angled configurations, shorter distances ($\sim$ 7 Å) enable stronger $*$OOH stabilization that suggest a preference towards 4e ORR. Interestingly, TPSSh-D3BJ predicts a range of $*$OOH binding energies that encompass the 2e ORR optimum (4.2 eV, dashed black line in Figure 2.11a). For the model systems considered here, we find that intrinsic catalytic activity is insensitive to small topological changes in porphyrin orientation (i.e. parallel versus angled). These results show that weakly $*$OOH binding active sites can be optimized for 2e peroxide synthesis by appropriately choosing the spectator transition metal, optimizing the TM1-TM2 separation, and changing the MOF topology. We acknowledge that these results alone are not conclusive of reaction selectivity; detailed kinetic studies and experimental measurements are necessary to validate our predictions.

2.4 Conclusion

By using state-of-the-art DFT calculations, we have demonstrated that porphyrin-based MOFs are an ideal materials platform for rationally designing 3-D active site environments for ORR. To the best of our knowledge, these computational predictions represent the first

experimentally-synthesizable heterogenous catalysts (i.e. MOFs) where the 3-D structure of the active site can be intentionally designed to circumvent the ∗OOH vs. ∗OH scaling relations. In addition to possible electrochemical stability concerns, other factors that impact the efficacy of MOF electrocatalysts, such as charge transport,[79–82] counterion and substrate diffusion,[81,82] explicit solvation,[55] etc. represent on-going research directions in the group. We anticipate that this work will motivate further computational studies and experimental validation of MOF-based electrocatalysts for ORR.

2.5 Computational Methods

2.5.1 DFT Calculations

Periodic density functional theory (DFT) calculations are performed using the projector augmented wave method as implemented in the Vienna ab initio simulation package (VASP). All energies are calculated using a 400 eV plane-wave cutoff. Only the Γ-point is sampled owing to the large MOF unit cells. A range of generalized gradient approximation (BEEF-vdW,[22] BLYP,[24,25], RPBE[23]) and hybrid (B3LYP,[26], HSE06,[27,28], and TPSSh[56]) functionals are used to examine the sensitivity of our results. Dispersion corrections are considered using the DFT-D3 method with Becke-Johnson damping.[29–31]. Electronic energies are converged to 10^{-6} eV. All structures are relaxed until the forces are less than 0.03 eV/Å for RPBE and 0.1 eV/Å for HSE06 and TPSSh. A lower threshold is used for the hybrid functionals owing to the high computational costs; differences in binding energies are less than 0.01 eV. Full relaxations were performed for RPBE and HSE06. TPSSh relaxations were performed with HSE06 geometry and allowing the N_4-M-adsorbate moieties to relax due to computational cost. All possible spin states are considered. Implicit solvation corrections are implemented using VASPSol.[54] The finite displacement method (0.015 Å) is used to calculate the entropic corrections and zero-point energies. Bond orders are calculated using the density derived electrostatic and chemical (DDEC) charge method.[83,84]

2.5.2 Functional Comparison and Justification

A detailed investigation of various functionals including RBE, TPSSh and HSE06 (all with Grimme's D3BJ correction) is performed to identify a suitable computational protocol for these systems. Specifically, we choose TPSSh-D3BJ as it correctly predicts the spin-states while also agreeing with the available experimental observations for N_4-Fe and N_4-Co catalysts. Our results show that RPBE tends to overstabilize the low-spin states, and overpredicts OXR binding energies for Co. In particular, Co-N_4 catalysts are known to be active for 2e ORR, but RPBE predicts Co to lie on the left leg of the volcano (strong $*$OH binding), indicating low 2e ORR activity. On the other hand, HSE06 generally predicts the correct spin-states, but strongly underbinds the ORR intermediates. Fe-N_4 catalysts are known to be excellent for 4e ORR, and several recent studies have suggested that the favorable reactions energetics are due to the formation of an additional axial OH ligand on the metal center. HSE06 shows that N_4-Fe-OH will be completely inactive for ORR due to weaker $*$OOH binding, and N_4-Fe is only weakly active. The calculated theoretical limiting potentials obtained with Co are also underpredicted for HSE06. These findings are summarized in Table 2.3, and all results obtained with RPBE and HSE06 are presented in Figures 2.13-2.16. Similar results (incorrect spin states for GGA functionals and under-binding due to higher exact exchange) have been observed for various other transition metal porphyrin complexes. While we acknowledge uncertainty in our free energy and implicit solvation corrections, we believe TPSSh provides the most reliable comparisons with experiments. Although detailed benchmarking of electronic structure methods and solvation is currently underway, we do not anticipate differences in the qualitative trends predicted using TPSSh-D3BJ functional.

RPBE predictions were found to overestimate the binding energies of the ORR intermediates. The tightly bound adsorbates were thus less influenced by the adjacent spectator and less stabilization was seen with this functional, as seen in Figure 2.12. The performance of the bimetallic PMOF systems as predicted by RPBE is shown in Figure 2.13, and the scaling behavior is shown in Figure 2.14.

Table 2.3. Comparison of experimental and theoretical activity of Fe and Co catalytic sites in PMOF-Al.

	Experimental potential (V)		Theoretical U_L (V)		
Active site	Onset	Half-wave	RPBE	HSE06	TPSSh
N_4-Co	0.75	0.55	-0.11	0.35	0.59
N_4-Fe	0.86	0.74	0.34	0.29	0.71
$N_4-Fe-OH$	0.86	0.74	0.75	-0.26	0.73

Table 2.4. TPSSh-D3BJ predicted relative spin-state energies of ORR intermediates with Co active site.

Species	Unpaired electrons	Relative Energy (eV)
N_4-Co	1	0.0
	3	0.4
	5	2.4
$N_4-Co-OH$	0	0.0
	2	0.2
	4	0.3
N_4-Co-O	1	0.0
	3	0.0
	5	0.5
$N_4-Co-OOH$	0	0.1
	2	0.0
	4	0.4

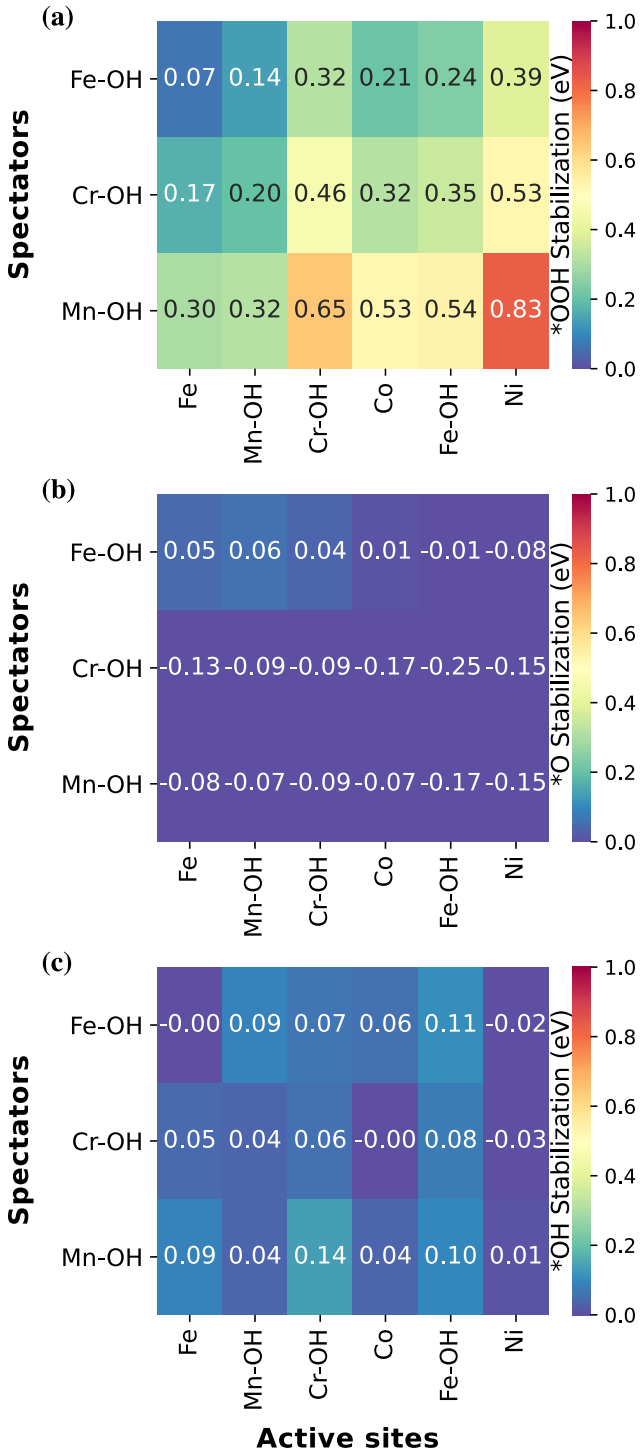

Figure 2.8. Stabilization of (a) *OOH, (b) *O, and (c) *OH for each active site and spectator relative to the no spectator system.

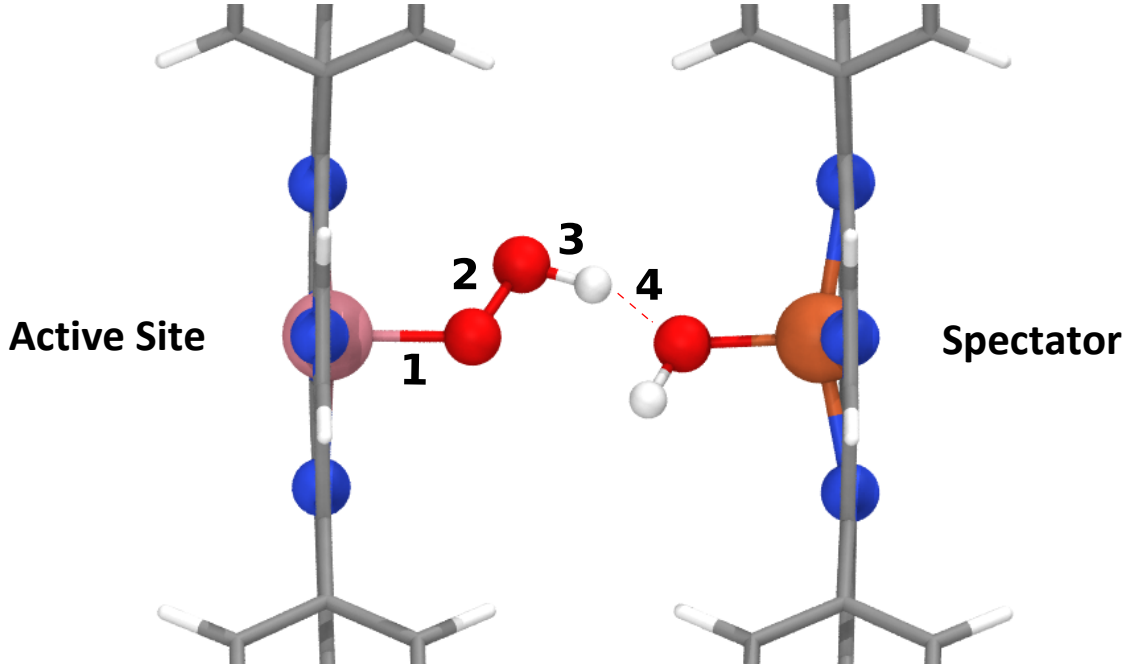

Figure 2.9. Hydrogen-bond interaction between ∗OOH and TM-OH spectator. Calculated bond orders for 1-4 are shown in Table 2.2.

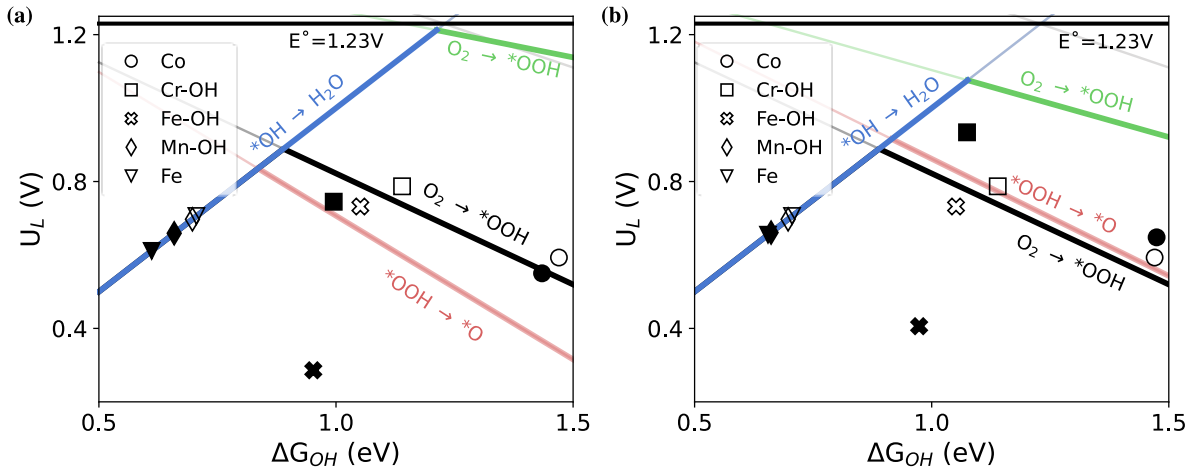

Figure 2.10. Volcano plots for (a) Mn-OH and (b) Cr-OH spectators (filled symbols) showing a decrease in performance relative to no spectator (empty symbols) due to overstabilization of the ∗OOH intermediate

Table 2.5. Free energy corrections applied to ORR binding energies.

Adsorbate	$\Delta G - \Delta E$ (eV)
∗OOH	0.11
∗O	-0.08
∗OH	0.18

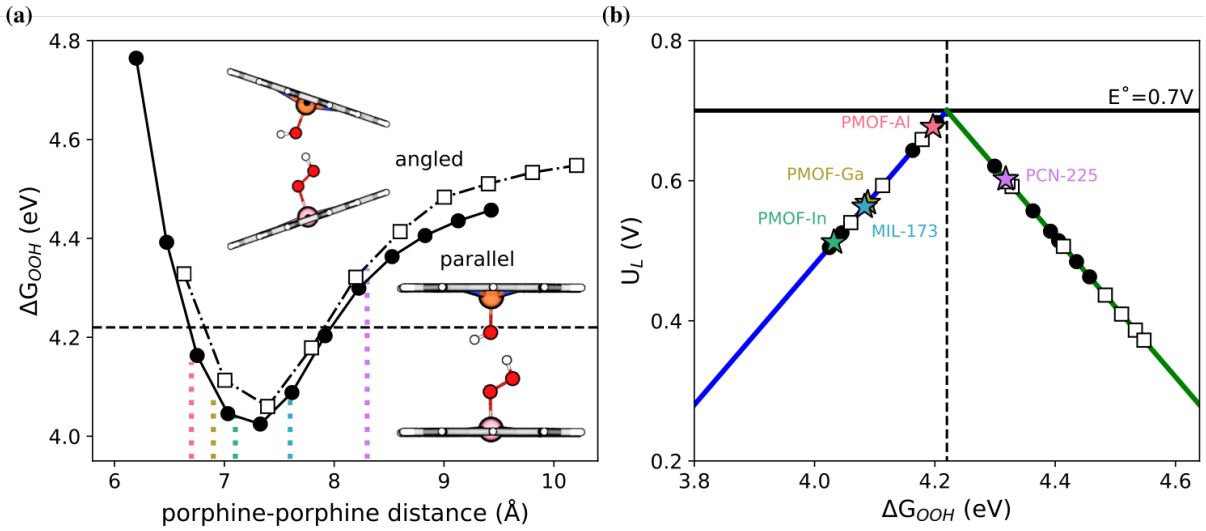

Figure 2.11. (a) The ∗OOH binding energy for the two parallel (filled circles) and angled (empty squares) interacting porphine molecules at various distances, (b) TPSSh-D3BJ calculated 2e limiting potentials for the model systems. Limiting potentials for PMOF-Al (red), PMOF-Ga (yellow), PMOF-In (green), MIL-173 (blue), and PCN-225 (purple) are predicted from their measured porphyrin separation distance.

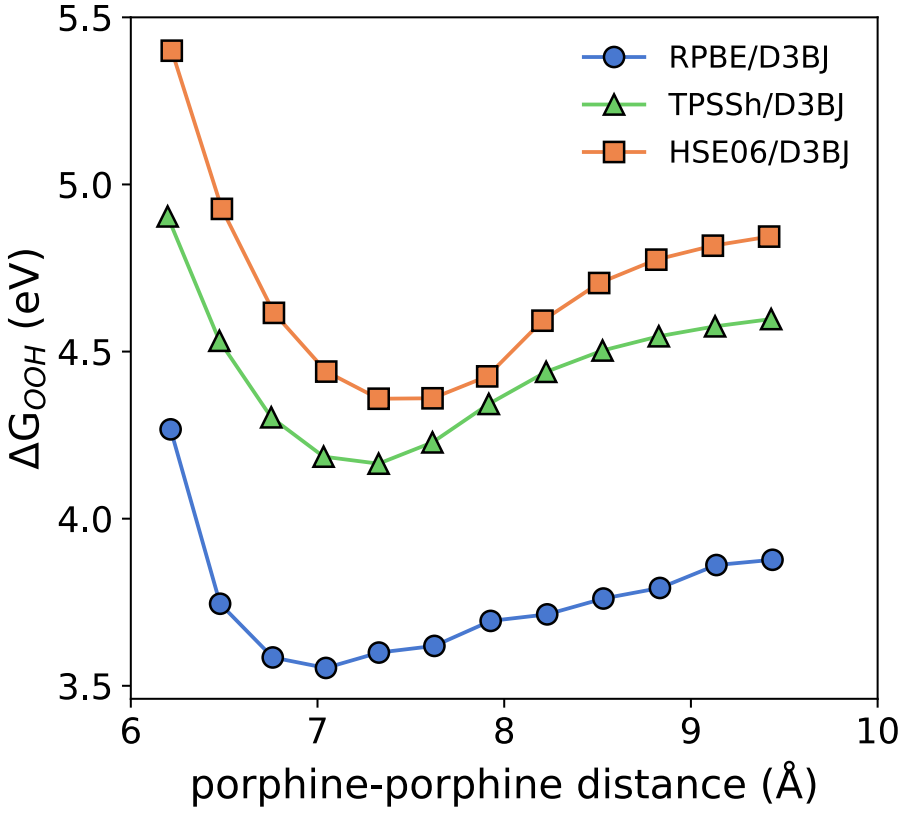

Figure 2.12. Functional comparison for *OOH binding for parallel porphine molecules with Co active site and Fe-OH spectator at various distances.

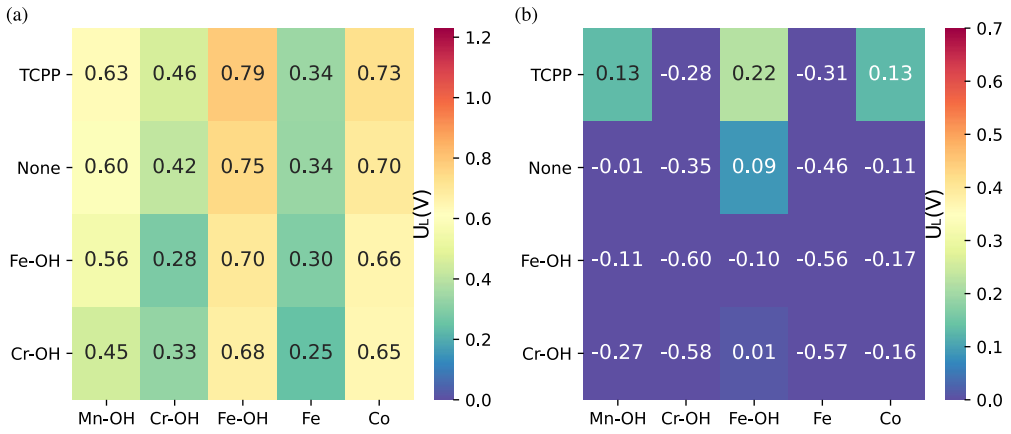

Figure 2.13. RPBE-D3BJ predicted theoretical limiting potentials for (a) 4e and (b) 2e ORR for different spectator and active site combinations.

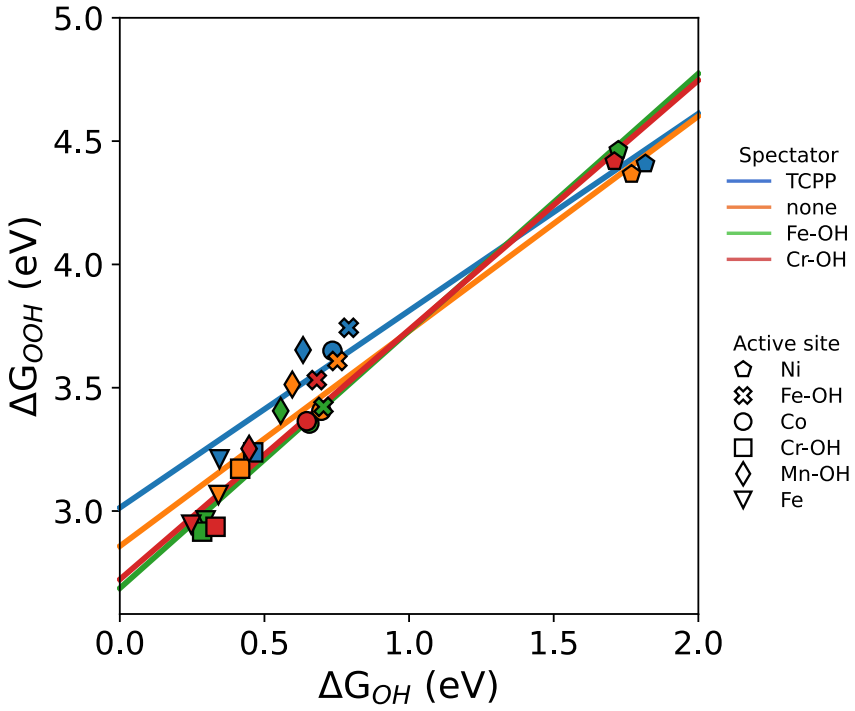

Figure 2.14. RPBE-D3BJ predicted *OOH vs. *OH scaling for different spectators.

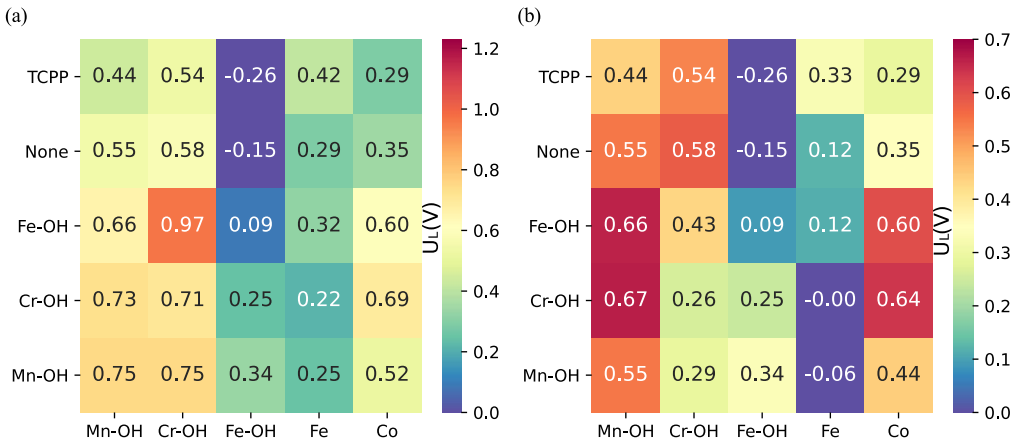

Figure 2.15. HSE06-D3BJ predicted theoretical limiting potentials for (a) 4e and (b) 2e ORR for different spectator and active site combinations.

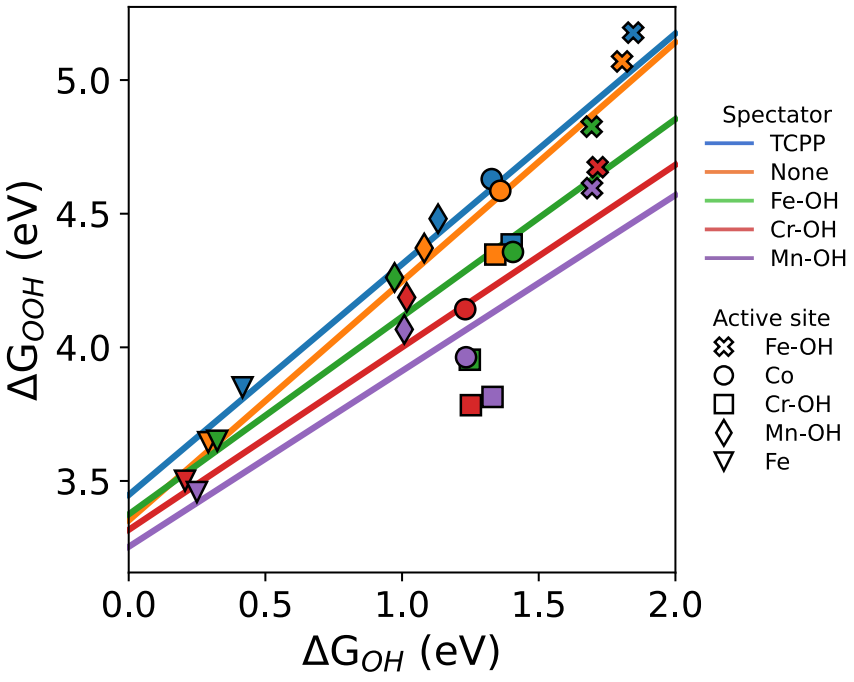

Figure 2.16. HSE06-D3BJ predicted *OOH vs. *OH scaling for different spectators.

2.6 References

[1] Mark K. Debe. Electrocatalyst approaches and challenges for automotive fuel cells. *Nature*, 486(7401):43–51, 2012. doi: 10.1038/nature11115.

[2] Zhi Wei Seh, Jakob Kibsgaard, Colin F. Dickens, Ib Chorkendorff, Jens K. Nørskov, and Thomas F. Jaramillo. Combining theory and experiment in electrocatalysis: Insights into materials design. *Science*, 355(6321), 2017. doi: 10.1126/science.aad4998.

[3] Michel Lefèvre, Eric Proietti, Frédéric Jaouen, and Jean Pol Dodelet. Iron-based catalysts with improved oxygen reduction activity in polymer electrolyte fuel cells. *Science*, 324(5923):71–74, 2009. doi: 10.1126/science.1170051.

[4] Jiazhan Li, Mengjie Chen, David A. Cullen, Sooyeon Hwang, Maoyu Wang, Boyang Li, Kexi Liu, Stavros Karakalos, Marcos Lucero, Hanguang Zhang, Chao Lei, Hui Xu, George E. Sterbinsky, Zhenxing Feng, Dong Su, Karren L. More, Guofeng Wang, Zhenbo Wang, and Gang Wu. Atomically dispersed manganese catalysts for oxygen reduction in proton-exchange membrane fuel cells. *Nature Catalysis*, 1(12):935–945, 2018. doi: 10.1038/s41929-018-0164-8.

[5] Isabela C. Man, Hai Yan Su, Federico Calle-Vallejo, Heine A. Hansen, José I. Martínez, Nilay G. Inoglu, John Kitchin, Thomas F. Jaramillo, Jens K. Nørskov, and Jan Rossmeisl. Universality in oxygen evolution electrocatalysis on oxide surfaces. *ChemCatChem*, 3(7):1159–1165, 2011. doi: 10.1002/cctc.201000397.

[6] Venkatasubramanian Viswanathan, Heine Anton Hansen, Jan Rossmeisl, and Jens K. Nørskov. Universality in oxygen reduction electrocatalysis on metal surfaces. *ACS Catalysis*, 2(8):1654–1660, 2012. doi: 10.1021/cs300227s.

[7] Yuanjun Chen, Shufang Ji, Chen Chen, Qing Peng, Dingsheng Wang, and Yadong Li. Single-atom catalysts: Synthetic strategies and electrochemical applications. *Joule*, 2(7):1242–1264, 2018. doi: 10.1016/j.joule.2018.06.019.

[8] Chengzhang Wan, Xiangfeng Duan, and Yu Huang. Molecular design of single-atom catalysts for oxygen reduction reaction. *Advanced Energy Materials*, 10(14):1–19, 2020. doi: 10.1002/aenm.201903815.

[9] Rune Christensen, Heine A. Hansen, Colin F. Dickens, Jens K. Nørskov, and Tejs Vegge. Functional independent scaling relation for orr/oer catalysts. *Journal of Physical Chemistry C*, 120(43):24910–24916, 2016. doi: 10.1021/acs.jpcc.6b09141.

[10] Fang Song, Lichen Bai, Aliki Moysiadou, Seunghwa Lee, Chao Hu, Laurent Liardet, and Xile Hu. Transition metal oxides as electrocatalysts for the oxygen evolution reaction in alkaline solutions: An application-inspired renaissance. *Journal of the American Chemical Society*, 140(25):7748–7759, 2018. doi: 10.1021/jacs.8b04546.

[11] Federico Calle-Vallejo, José Ignacio Martínez, and Jan Rossmeisl. Density functional studies of functionalized graphitic materials with late transition metals for oxygen reduction reactions. *Physical Chemistry Chemical Physics*, 13(34):15639–15643, 2011. doi: 10.1039/c1cp21228a.

[12] Ambarish Kulkarni, Samira Siahrostami, Anjli Patel, and Jens K Nørskov. Understanding catalytic activity trends in the oxygen reduction reaction. *Chemical Reviews*, 118(5):2302–2312, 3 2018. doi: 10.1021/acs.chemrev.7b00488.

[13] Hua Bing Tao, Junming Zhang, Jiazang Chen, Liping Zhang, Yinghua Xu, Jingguang G. Chen, and Bin Liu. Revealing energetics of surface oxygen redox from kinetic fingerprint in oxygen electrocatalysis. *Journal of the American Chemical Society*, 141(35):13803–13811, 2019. doi: 10.1021/jacs.9b01834.

[14] Marc T.M. Koper. Thermodynamic theory of multi-electron transfer reactions: Implications for electrocatalysis. *Journal of Electroanalytical Chemistry*, 660(2):254–260, 2011. doi: 10.1016/j.jelechem.2010.10.004.

[15] Aleksandra Vojvodic and Jens K. Nørskov. New design paradigm for heterogeneous catalysts. *National Science Review*, 2(2):140–143, 2015. doi: 10.1093/nsr/nwv023.

[16] Hao Wan, Thomas Mandal Østergaard, Logi Arnarson, and Jan Rossmeisl. Climbing the 3d volcano for the oxygen reduction reaction using porphyrin motifs. *ACS Sustainable Chemistry and Engineering*, 7(1):611–617, 2019. doi: 10.1021/acssuschemeng.8b04173.

[17] Hao Wan, Anders W. Jensen, María Escudero-Escribano, and Jan Rossmeisl. Insights in the oxygen reduction reaction: From metallic electrocatalysts to diporphyrins. *ACS Catalysis*, 10(11):5979–5989, 6 2020. doi: 10.1021/acscatal.0c01085.

[18] Nitish Govindarajan, Marc T. M. Koper, Evert Jan Meijer, and Federico Calle-Vallejo. Outlining the scaling-based and scaling-free optimization of electrocatalysts. *ACS Catalysis*, 9(5):4218–4225, 5 2019. doi: 10.1021/acscatal.9b00532.

[19] C. K. Chang, H. Y. Liu, I. Abdalmuhdi, and H. Y. Liu. Electroreduction of oxygen by pillared cobalt cofacial diporphyrin catalysts. *Journal of the American Chemical Society*, 106(9):2725–2726, 1984. doi: 10.1021/ja00321a055.

[20] James P Collman, Paul S Wagenknecht, and James E Hutchison. Molecular catalysts for multielectron redox reactions of small molecules: The "cofacial metallodiporphyrin" approach. *Angewandte Chemie International Edition in English*, 33(15-16): 1537–1554, 9 1994. doi: 10.1002/anie.199415371.

[21] Joel Rosenthal and Daniel G Nocera. Role of proton-coupled electron transfer in o–o bond activation. *Accounts of Chemical Research*, 40(7):543–553, 7 2007. doi: 10.1021/ar7000638.

[22] Jess Wellendorff, Trent L Silbaugh, Delfina Garcia-Pintos, Jens K Nørskov, Thomas Bligaard, Felix Studt, and Charles T Campbell. A benchmark database for adsorption bond energies to transition metal surfaces and comparison to selected {dft} functionals. *Surface Science*, 640:36–44, 10 2015. doi: 10.1016/j.susc.2015.03.023.

[23] B. Hammer, L. B. Hansen, and J. K. Nørskov. Improved adsorption energetics within density-functional theory using revised perdew-burke-ernzerhof functionals. *Physical*

Review B - Condensed Matter and Materials Physics, 59(11):7413–7421, 1999. doi: 10.1103/PhysRevB.59.7413.

[24] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Physical Review A*, 38(6):3098–3100, 9 1988. doi: 10.1103/PhysRevA.38.3098.

[25] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37(2):785–789, 1 1988. doi: 10.1103/PhysRevB.37.785.

[26] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *Journal of Physical Chemistry®*, 98(45):11623–11627, 1994. doi: 10.1021/j100096a001.

[27] Jochen Heyd, Gustavo E. Scuseria, and Matthias Ernzerhof. Hybrid functionals based on a screened coulomb potential. *Journal of Chemical Physics*, 118(18):8207–8215, 2003. doi: 10.1063/1.1564060.

[28] V. Aliaksandr Krukau, Oleg A. Vydrov, Artur F. Izmaylov, and Gustavo E. Scuseria. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *Journal of Chemical Physics*, 125(22), 2006. doi: 10.1063/1.2404663.

[29] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. *Journal of Chemical Physics*, 132(15), 2010. doi: 10.1063/1.3382344.

[30] Stefan Grimme, Stephan Ehrlich, and Lars Goerigk. Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry*, 32(7):1456–1465, 5 2011. doi: 10.1002/jcc.21759.

[31] Daniel G.A. Smith, Lori A. Burns, Konrad Patkowski, and C. David Sherrill. Revised damping parameters for the d3 dispersion correction to density functional theory. *Journal of Physical Chemistry Letters*, 7(12):2197–2203, 2016. doi: 10.1021/acs.jpclett.6b00780.

[32] Andrew S. Rosen, Justin M. Notestein, and Randall Q. Snurr. Structure-activity relationships that identify metal-organic framework catalysts for methane activation. *ACS Catalysis*, 9(4):3576–3587, 2019. doi: 10.1021/acscatal.8b05178.

[33] Ágnes Szécsényi, Elena Khramenkova, Ivan Yu Chernyshov, Guanna Li, Jorge Gascon, and Evgeny A. Pidko. Breaking linear scaling relationships with secondary interactions in confined space: A case study of methane oxidation by fe/zsm-5 zeolite. *ACS Catalysis*, 9(10):9276–9284, 2019. doi: 10.1021/acscatal.9b01914.

[34] David J. Tranchemontagne, José L. Tranchemontagne, Michael O'keeffe, and Omar M. Yaghi. Secondary building units, nets and bonding in the chemistry of metal–organic frameworks. *Chemical Society Reviews*, 38(5):1257–1283, 2009. doi: 10.1039/b817735j.

[35] Alexander Schoedel, Mian Li, Dan Li, Michael O'Keeffe, and Omar M. Yaghi. Structures of metal-organic frameworks with rod secondary building units. *Chemical Reviews*, 116(19):12466–12535, 2016. doi: 10.1021/acs.chemrev.6b00346.

[36] Weigang Lu, Zhangwen Wei, Zhi Yuan Gu, Tian Fu Liu, Jinhee Park, Jihye Park, Jian Tian, Muwei Zhang, Qiang Zhang, Thomas Gentle, Mathieu Bosch, and Hong Cai Zhou. Tuning the structure and function of metal-organic frameworks via linker design. *Chemical Society Reviews*, 43(16):5561–5593, 2014. doi: 10.1039/c4cs00003j.

[37] Mathieu Lions, Jean-Bernard Tomasino, Raphael Chattot, Brain Abeykoon, Nathalie Guillou, Thomas Devic, Aude Demessence, Luis Cardenas, Frederic Maillard, Alexandra Fateeva, J.-B. Tommasino, Raphael Chattot, Brain Abeykoon, Nathalie Guillou, Thomas Devic, Aude Demessence, Luis Cardenas, Frederic Maillard, and Alexandra

Fateeva. Insights into the mechanism of electrocatalysis of oxygen reduction reaction by a porphyrinic metal organic framework. *Chem. Commun.*, 53(48):6496–6499, 2017. doi: 10.1039/C7CC02113E.

[38] Pavel M Usov, Brittany Huffman, Charity C Epley, Matthew C Kessinger, Jie Zhu, William A Maza, and Amanda J Morris. Study of electrocatalytic properties of metal–organic framework pcn-223 for the oxygen reduction reaction. *ACS Applied Materials & Interfaces*, 9(39):33539–33543, 10 2017. doi: 10.1021/acsami.7b01547.

[39] Magdalena Ola Cichocka, Zuozhong Liang, Dawei Feng, Seoin Back, Samira Siahrostami, Xia Wang, Laura Samperisi, Yujia Sun, Hongyi Xu, Niklas Hedin, Haoquan Zheng, Xiaodong Zou, Hong-Cai Zhou, and Zhehao Huang. A porphyrinic zirconium metal–organic framework for oxygen reduction reaction: Tailoring the spacing between active-sites through chain-based inorganic building units. *Journal of the American Chemical Society*, 142(36):15386–15395, 9 2020. doi: 10.1021/jacs.0c06329.

[40] Elise M Miner, Tomohiro Fukushima, Dennis Sheberla, Lei Sun, Yogesh Surendranath, and Mircea Dinca. Electrochemical oxygen reduction catalysed by ni3(hexaiminotriphenylene)2. *Nature Communications*, pages 1–7, 2016. doi: 10.1038/ncomms10942.

[41] Yuxia Xu, Qing Li, Huaiguo Xue, and Huan Pang. Metal-organic frameworks for direct electrochemical applications. *Coordination Chemistry Reviews*, 376:292–318, 2018. doi: https://doi.org/10.1016/j.ccr.2018.08.010.

[42] D. M. D'Alessandro. Exploiting redox activity in metal-organic frameworks: Concepts, trends and perspectives. *Chemical Communications*, 52(58):8957–8971, 2016. doi: 10.1039/c6cc00805d.

[43] Marcello B. Solomon, Tamara L. Church, and Deanna M. D'Alessandro. Perspectives on metal–organic frameworks with intrinsic electrocatalytic activity. *CrystEngComm*, 19(29):4049–4065, 2017. doi: 10.1039/C7CE00215G.

[44] Wei Wang, Xiaomin Xu, Wei Zhou, and Zongping Shao. Recent progress in metal-organic frameworks for applications in electrocatalytic and photocatalytic water splitting. *Advanced Science*, 4(4):1600371, 4 2017. doi: 10.1002/advs.201600371.

[45] Rongrong Chen, Haixia Li, Deryn Chu, and Guofeng Wang. Unraveling oxygen reduction reaction mechanisms on carbon-supported fe-phthalocyanine and co-phthalocyanine catalysts in alkaline solutions. *The Journal of Physical Chemistry C*, 113(48):20689–20697, 12 2009. doi: 10.1021/jp906408y.

[46] Michael Busch. Water oxidation: From mechanisms to limitations. *Current Opinion in Electrochemistry*, 9:278–284, 6 2018. doi: 10.1016/j.coelec.2018.06.007.

[47] Shyam Kattel, Plamen Atanassov, and Boris Kiefer. A density functional theory study of oxygen reduction reaction on non-pgm fe-nx-c electrocatalysts. *Physical Chemistry Chemical Physics*, 16(27):13800–13806, 2014. doi: 10.1039/c4cp01634c.

[48] Hao Cao, Guang Jie Xia, Jie Wei Chen, Hui Min Yan, Zhen Huang, and Yang Gang Wang. Mechanistic insight into the oxygen reduction reaction on the mn-n4/c single-atom catalyst: The role of the solvent environment. *Journal of Physical Chemistry C*, 124(13):7287–7294, 2020. doi: 10.1021/acs.jpcc.0c00352.

[49] Lixiang Zhong and Shuzhou Li. Unconventional oxygen reduction reaction mechanism and scaling relation on single-atom catalysts. *ACS Catalysis*, 10(7):4313–4318, 4 2020. doi: 10.1021/acscatal.0c00815.

[50] Alexandra Fateeva, Philip A Chater, Christopher P Ireland, Asif A Tahir, Yaroslav Z Khimyak, V Paul Wiper, James R Darwent, and Matthew J Rosseinsky. A water-stable porphyrin-based metal–organic framework active for visible-light photocatalysis. *Angewandte Chemie International Edition*, 51(30):7440–7444, 7 2012. doi: 10.1002/anie.201202471.

[51] Zhenqiang Wang and Seth M. Cohen. Postsynthetic modification of metal–organic frameworks. *Chemical Society Reviews*, 38(5):1315–1329, 2009. doi: 10.1039/b802258p.

[52] Kristine K. Tanabe and Seth M. Cohen. Postsynthetic modification of metal–organic frameworks—a progress report. *Chemical Society Reviews*, 40(2):498–519, 2011. doi: 10.1039/c0cs00031k.

[53] Min Kim, John F. Cahill, Yongxuan Su, Kimberly A. Prather, and Seth M. Cohen. Postsynthetic ligand exchange as a route to functionalization of 'inert' metal-organic frameworks. *Chemical Science*, 3(1):126–130, 2012. doi: 10.1039/c1sc00394a.

[54] Kiran Mathew and Richard G Hennig. Implicit self-consistent description of electrolyte in plane-wave density-functional theory. *arXiv*, pages 1–6, 2016.

[55] Anjli M. Patel, Stefan Ringe, Samira Siahrostami, Michal Bajdich, Jens K. Nørskov, and Ambarish R. Kulkarni. Theoretical approaches to describing the oxygen reduction reaction activity of single-atom catalysts. *Journal of Physical Chemistry C*, 122 (51):29307–29318, 2018. doi: 10.1021/acs.jpcc.8b09430.

[56] Viktor N. Staroverov, Gustavo E. Scuseria, Jianmin Tao, and John P. Perdew. Comparative assessment of a new nonempirical density functional: Molecules and hydrogen-bonded complexes. *Journal of Chemical Physics*, 119(23):12129–12137, 2003. doi: 10.1063/1.1626543.

[57] Da Jeng Chai and Martin Head-Gordon. Systematic optimization of long-range corrected hybrid density functionals. *Journal of Chemical Physics*, 128(8), 2008. doi: 10.1063/1.2834918.

[58] Haoyu S. Yu, Xiao He, Shaohong L. Li, and Donald G. Truhlar. Mn15: A kohn-sham global-hybrid exchange-correlation density functional with broad accuracy for multi-reference and single-reference systems and noncovalent interactions. *Chemical Science*, 7(8):5032–5051, 2016. doi: 10.1039/c6sc00705h.

[59] Michael Busch, Alberto Fabrizio, Sandra Luber, Jürg Hutter, and Clemence Corminboeuf. Exploring the limitation of molecular water oxidation catalysts. *Journal of Physical Chemistry C*, 122(23):12404–12412, 2018. doi: 10.1021/acs.jpcc.8b03935.

[60] Carlo Alberto Gaggioli, Samuel J. Stoneburner, Christopher J. Cramer, and Laura Gagliardi. Beyond density functional theory: The multiconfigurational approach to model heterogeneous catalysis. *ACS Catalysis*, 9(9):8481–8502, 2019. doi: 10.1021/acscatal.9b01775.

[61] Mariusz Radoń and Kristine Pierloot. Binding of co, no, and o2 to heme by density functional and multireference ab initio calculations. *Journal of Physical Chemistry A*, 112(46):11824–11832, 2008. doi: 10.1021/jp806075b.

[62] Mariusz Radoń. Spin-state energetics of heme-related models from dft and coupled cluster calculations. *Journal of Chemical Theory and Computation*, 10(6):2306–2321, 2014. doi: 10.1021/ct500103h.

[63] Ergui Luo, Hao Zhang, Xian Wang, Liqin Gao, Liyuan Gong, Tuo Zhao, Zhao Jin, Junjie Ge, Zheng Jiang, Changpeng Liu, and Wei Xing. Single-atom cr-n4 sites designed for durable oxygen reduction catalysis in acid media. *Angewandte Chemie - International Edition*, 58(36):12469–12475, 2019. doi: 10.1002/anie.201906289.

[64] Nitish Govindarajan, Juan M. García-Lastra, Evert Jan Meijer, and Federico Calle-Vallejo. Does the breaking of adsorption-energy scaling relations guarantee enhanced electrocatalysis? *Current Opinion in Electrochemistry*, 8:110–117, 2018. doi: 10.1016/j.coelec.2018.03.025.

[65] Oriol Piqué, Francesc Illas, and Federico Calle-Vallejo. Designing water splitting catalysts using rules of thumb: Advantages, dangers and alternatives. *Physical Chemistry Chemical Physics*, 22(13):6797–6803, 2020. doi: 10.1039/d0cp00896f.

[66] Heine A. Hansen, Venkatasubramanian Viswanathan, and Jens K. NØrskov. Unifying kinetic and thermodynamic analysis of 2 e- and 4 e - reduction of oxygen on metal surfaces. *Journal of Physical Chemistry C*, 118(13):6706–6718, 2014. doi: 10.1021/jp4100608.

[67] Marc T.M. Koper. Analysis of electrocatalytic reaction schemes: Distinction between

rate-determining and potential-determining steps. *Journal of Solid State Electrochemistry*, 17(2):339–344, 2013. doi: 10.1007/s10008-012-1918-x.

[68] Kai S. Exner and Herbert Over. Kinetics of electrocatalytic reactions from first-principles: A critical comparison with the ab initio thermodynamics approach. *Accounts of Chemical Research*, 50(5):1240–1247, 5 2017. doi: 10.1021/acs.accounts.7b00077.

[69] Kai S. Exner. Beyond the traditional volcano concept: Overpotential-dependent volcano plots exemplified by the chlorine evolution reaction over transition-metal oxides. *Journal of Physical Chemistry C*, 123(27):16921–16928, 2019. doi: 10.1021/acs.jpcc.9b05364.

[70] Kai S Exner. A universal descriptor for the screening of electrode materials for multiple-electron processes: Beyond the thermodynamic overpotential. *ACS Catalysis*, pages 12607–12617, 10 2020. doi: 10.1021/acscatal.0c03865.

[71] Kai S. Exner, Iman Sohrabnejad-Eskan, and Herbert Over. A universal approach to determine the free energy diagram of an electrocatalytic reaction. *ACS Catalysis*, 8 (3):1864–1879, 2018. doi: 10.1021/acscatal.7b03142.

[72] Sungeun Yang, Arnau Verdaguer-Casadevall, Logi Arnarson, Luca Silvioli, Viktor Čolić, Rasmus Frydendal, Jan Rossmeisl, Ib Chorkendorff, and Ifan E.L. Stephens. Toward the decentralized electrochemical production of h2o2: A focus on the catalysis. *ACS Catalysis*, 8(5):4064–4081, 2018. doi: 10.1021/acscatal.8b00217.

[73] Sanjeev Mukerjee, Supramaniam Srinivasan, Manuel P. Soriaga, and James McBreen. Role of structural and electronic properties of pt and pt alloys on electrocatalysis of oxygen reduction: An in situ xanes and exafs investigation. *Journal of The Electrochemical Society*, 142(5):1409–1422, 1995. doi: 10.1149/1.2048590.

[74] María Escudero-Escribano, Paolo Malacrida, M. H. Hansen, U. G. Vej-Hansen, Amado Velazquez-Palenzuela, Vladamir Tripkovic, J. Schiotz, Jan Rossmeisl, Ifan

E L Stephens, and Ib Chorkendorff. Tuning the activity of pt alloy electrocatalysts by means of the lanthanide contraction. *Science*, 352(6281):73–76, 4 2016. doi: 10.1126/science.aad8892.

[75] Timo Rhauderwiek, Steve Waitschat, Stefan Wuttke, Helge Reinsch, Thomas Bein, and Norbert Stock. Nanoscale synthesis of two porphyrin-based mofs with gallium and indium. *Inorganic Chemistry*, 55(11):5312–5319, 2016. doi: 10.1021/acs.inorgchem.6b00221.

[76] Georges Mouchaham, Brian Abeykoon, Mónica Giménez-Marqués, Sergio Navalon, Andrea Santiago-Portillo, Maame Affram, Nathalie Guillou, Charlotte Martineau, Hermenegildo Garcia, Alexandra Fateeva, and Thomas Devic. Adaptability of the metal(iii,iv) 1,2,3-trioxobenzene rod secondary building unit for the production of chemically stable and catalytically active mofs. *Chemical Communications*, 53(54): 7661–7664, 2017. doi: 10.1039/c7cc04215a.

[77] Hai Long Jiang, Dawei Feng, Kecheng Wang, Zhi Yuan Gu, Zhangwen Wei, Ying Pin Chen, and Hong Cai Zhou. An exceptionally stable, porphyrinic zr metal-organic framework exhibiting ph-dependent fluorescence. *Journal of the American Chemical Society*, 135(37):13934–13938, 2013. doi: 10.1021/ja406844r.

[78] Omar M. Yaghi, Michael O'Keeffe, Nathan W. Ockwig, Hee K. Chae, Mohamed Eddaoudi, and Jaheon Kim. Reticular synthesis and the design of new materials. *Nature*, 423(6941):705–714, 2003. doi: 10.1038/nature01650.

[79] Saied Md Pratik, Laura Gagliardi, and Christopher J. Cramer. Engineering electrical conductivity in stable zirconium-based pcn-222 mofs with permanent mesoporosity. *Chemistry of Materials*, 32(14):6137–6149, 2020. doi: 10.1021/acs.chemmater.0c01847.

[80] Sameer Patwardhan and George C. Schatz. Theoretical investigation of charge transfer in metal organic frameworks for electrochemical device applications. *Journal of Physical Chemistry C*, 119(43):24238–24247, 2015. doi: 10.1021/acs.jpcc.5b06065.

[81] Meng Cai, Quentin Loague, and Amanda J. Morris. Design rules for efficient charge transfer in metal-organic framework films: The pore size effect. *Journal of Physical Chemistry Letters*, 11(3):702–709, 2020. doi: 10.1021/acs.jpclett.9b03285.

[82] Itamar Liberman, Ran Shimoni, Raya Ifraemov, Illya Rozenberg, Chanderpratap Singh, and Idan Hod. Active-site modulation in an fe-porphyrin-based metal-organic framework through ligand axial coordination: Accelerating electrocatalysis and charge-transport kinetics. *Journal of the American Chemical Society*, 142(4): 1933–1940, 2020. doi: 10.1021/jacs.9b11355.

[83] Thomas A. Manz and David S. Sholl. Chemically meaningful atomic charges that reproduce the electrostatic potential in periodic and nonperiodic materials. *Journal of Chemical Theory and Computation*, 6(8):2455–2468, 2010. doi: 10.1021/ct100125x.

[84] Thomas A. Manz. Introducing ddec6 atomic population analysis: Part 3. comprehensive method to compute bond orders. *RSC Advances*, 7(72):45552–45581, 2017. doi: 10.1039/c7ra07400j.

Chapter 3

A Theory-Guided X-ray Absorption Spectroscopy Approach for Identifying Active Sites in Atomically Dispersed Transition-Metal Catalysts

Y. Chen[#], R. Rana[#], T. Sours[#], F. D. Vila, S. Cao, T. Blum, J. Hong, A. S. Hoffman, C.-Y. Fang, Z. Huang, C. Shang, C. Wang, J. Zeng, M. Chi, C. X. Kronawitter, S. R. Bare, B. C. Gates and A. R. Kulkarni

[#]Authors contributed equally. My contributions to this work consist of all theoretical predictions excluding the EXAFS and XANES spectra simulations. All experimental work was performed by collaborators.

Adapted with permission from J. Am. Chem. Soc., 2021, 143, 20144–20156.

3.1 Abstract

Atomically dispersed supported metal catalysts offer new properties and the benefits of maximized metal accessibility and utilization. The characterization of these materials, however, remains challenging. Using atomically-dispersed Pt supported on crystalline MgO (chosen for its well-defined bonding sites for Pt) as a prototypical example, in this work, we show how high-throughput density functional theory calculations (for assessing all the potentially stable Pt sites) combined with automated EXAFS analysis can lead to unbiased identification of isolated, surface-enveloped platinum cations as the catalytic species for CO oxidation. The catalyst has been characterized by atomic-resolution imaging, EXAFS, and HERFD-XANES spectroscopies; the proposed Pt site are in full

agreement with experiment. This theory-guided workflow leads to rigorously determined structural models and provides a more detailed picture of the structure of the catalytically active sites than what is currently possible with conventional EXAFS analysis. As this approach is efficient and agnostic to the metal, support, and catalytic reaction, we posit that it will be of broad interest to the materials characterization and catalysis communities.

3.2 Introduction

Atomically dispersed metals on reducible (CeO_2, TiO_2, Fe_3O_4) and nonreducible (Al_2O_3, MgO, ZrO_2) metal oxide supports have drawn wide attention recently as catalysts for a wide range of reactions including CO oxidation, water-gas shift, and other oxidations and reductions.[1–8] These catalysts offer the advantages of maximum accessibility of the metals, which are often expensive (e.g., Pt, Pd, Ir, Rh), and catalytic properties distinct from those of the conventional supported metal clusters and nanoparticles. Like most single transition metal (TM) atoms in organometallic complexes, those on metal oxide supports are positively charged, having properties that are strongly influenced by their surroundings.[9–12] These surroundings (bonding environments) are challenging to identify, because the support surfaces are inherently heterogeneous. Even when the most incisive experimental methods are used (e.g., aberration corrected scanning transmission electron microscopy (STEM), X-ray absorption spectroscopy (XAS), and probe molecule experiments with spectroscopies), the results often provide only average structural models of the dominant metal bonding environment; the possible contributions of minority catalytic species are often overlooked.[13] Common research approaches involve using experimentally determined structure parameters (e.g., bond lengths and coordination numbers from extended x-ray absorption fine structure (EXAFS) spectroscopy) with complementary information characterizing local geometry and electronic structure from x-ray absorption near-edge structure (XANES) spectra. These techniques are often combined to create plausible atomistic models that are subsequently investigated using density functional theory (DFT) calculations to derive quantitative structural and mechanistic insights.[7,14,15] More recent approaches have utilized machine learning techniques to gain structural in-

formation.[16] We emphasize that, as conventionally practiced, the initial development of atomistic models relies predominantly on thorough EXAFS modeling, which is often quite demanding and time intensive for the user, especially when multiple scattering paths are considered. Today's "hands-on" EXAFS modeling approaches fall short of determining the existence of multiple metal bonding environments—a central limitation in the field. Moreover, structures inferred for the sites may not be consistent with all the complementary experimental results. As atomically dispersed catalysts continue to be reported for new reactions, there is a clear potential scientific benefit to develop computationally guided approaches for characterization of these materials, specifically for the quantitative identification of the metal site environment(s).

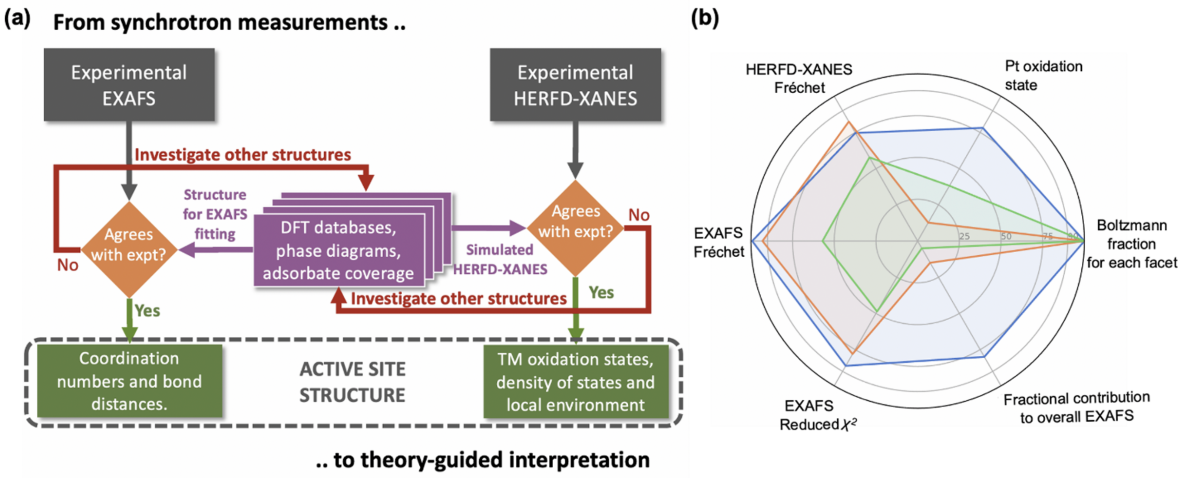

Figure 3.1. (a) Schematic representation of theory-guided workflow for identifying the Pt active sites stabilized within or on a MgO lattice. (b) Radar plot for the most stable [100], [100]Mg-vac, and [310] Pt configurations showing that the [100]Mg-vac/sub1 (blue) is more consistent with our experimental observations than [100]/sub0/$*O_2$ (orange) or [310]/pos1/$*O_2$ site (green). Structure models are presented in Figure 3.3.

As a point of departure from reported investigations, we now describe a theory-led workflow for characterizing atomically dispersed supported metal catalysts. Our approach combines state-of-the-art characterization techniques (high-angle annular dark field (HAADF) STEM, EXAFS spectroscopy, and high-energy resolution fluorescence detection (HERFD) XANES spectroscopy) with DFT-level theory for critical identification of the local environments for a prototypical catalyst—atomically dispersed Pt cations on MgO (Pt/MgO). We chose MgO as a prototypical support because it (a)

is nonreducible, (b) consists of low-atomic-number elements for excellent contrast with Pt in STEM, and (c) is available as robust, high-area crystalline materials used in industrial catalysts.[5,10] These crystalline supports are expected to present limited numbers of sites for stable anchoring of isolated metal atoms, making Pt/MgO an excellent test case to exemplify our methodology. Figure 3.1a is a summary of our strategy that combines microscopy and conventional EXAFS analyses (Sections 3.3.1 and 3.3.2), large-scale DFT calculations (Section 3.3.3), automated DFT-based EXAFS analyses (Section 3.3.4), and FEFF-XANES spectroscopy (Section 3.3.5) to characterize Pt/MgO and identify the structure most consistent with data obtained through all the complementary techniques. As shown throughout this report, the outcome of this approach, represented as a radar plot (Figure 3.1b, details in Table 3.S1), identifies sub-surface Pt sites within MgO (denoted Pt^{+4}/[100]Mg-vac) to be the most populated sites for Pt atoms for this system. Although the approach is applied for atomically dispersed Pt/MgO (Section 3.3.6), we emphasize that the strategy is agnostic to the metal, the oxide support, and the reaction. To our best knowledge, this approach is unique in the field and represents a significant step forward towards the goal of comprehensive, validated integration of theoretical and experimental methods to describe the structure and catalytic function of isolated noble metal cations stabilized by a metal oxide support.

3.3 Results

3.3.1 Synthesis and structural characterization.

Atomically dispersed Pt on MgO, prepared from ethanol-water solutions of K_2PtCl_4 and slurried MgO powder, was calcined at 700 °C. Low Pt loadings (0.05 wt%) were targeted to minimize the number of different support surface sites that the Pt atoms occupy. TEM images and x-ray diffraction (XRD) patterns show that MgO was present as the cubic phase consisting of approximately 200-nm diameter crystallites; there was no XRD evidence of metallic Pt (Figure 3.S1). The calcined MgO was highly dehydroxylated, with only a weak OH band in the IR spectrum, at 3741 cm^{-1} (Figure 3.S2). HAADF-STEM images show atomically dispersed Pt in the absence of clusters or nanoparticles (Figure

3.2, Figure 3.S3). The bright spots in yellow circles (Figure 3.2a) show isolated Pt atoms along Mg columns in the (110) projection.[17] Intensity profiles along the X–Y line in Figure 3.2b (Figure 3.2c) confirm the isolation of the Pt atoms.

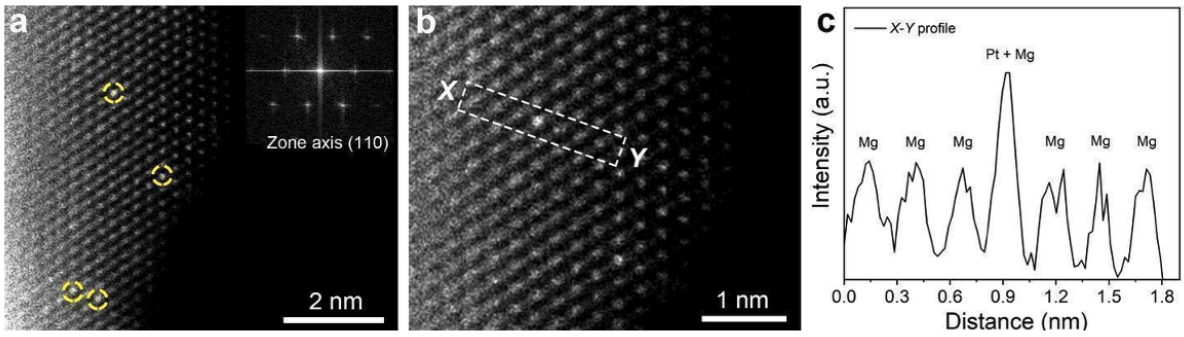

Figure 3.2. (a) HAADF-STEM image (along the (110) zone axis) showing that the MgO-supported Pt atoms (yellow circles) were atomically dispersed. (b) Higher magnification image showing a single isolated Pt atom, and (c) corresponding intensity profile from the X–Y line scan in panel b

3.3.2 Conventional EXAFS analysis.

Pt L3-edge EXAFS spectra were modeled in Artemis[18] using traditional methods. The best-fit model comprises three scattering paths: Pt–O (2.05 ± 0.01 Å), with a coordination number (CN) of 6.7 ± 0.8, and two Pt–Mg (3.01 ± 0.04 and 3.15 ± 0.04 Å), with a total CN of 11 (Figures 3.S4-3.S6, Tables 3.S2-3.S4). Addition of a Pt–Pt scattering path (expected at ∼ 2.7 Å for Pt^0 clusters) resulted in non-meaningful results,[19] confirming the atomic dispersion of Pt demonstrated by HAADF-STEM. The continuous Cauchy wavelet transform (CCWT) heatmap of the sample (Figure 3.S7a) shows two main scattering features arising from short and long distances.[20] This analysis provides evidence that the longer features at 5 $Å^{-1}$ arise from a lower-Z scattering atom than Pt (i.e., Mg) as it does not match the features of the CCWT for Pt metal (Figure 3.S7b). The scattering feature with a short-distance, low-Z scattering component was best-modeled as a Pt–O path. The higher-R (distance) component was modeled as two Pt–Mg paths at longer distances. We emphasize that this conventional analysis of the modeled EXAFS data is insufficient to determine more than a single average, best-fit structure. Thus, it is difficult to define the actual bonding site. To address the possible presence of more than

one significant supported Pt species and to exhaustively examine the possible Pt bonding environments, we developed an automated, high-throughput DFT-based workflow for EXAFS data fitting that reduces user intervention, and thereby improves the objectivity of the data interpretation.

3.3.3 A comprehensive library of DFT-optimized structures.

We used a theory-guided characterization approach to elucidate the local Pt bonding environment, that is, using DFT calculations (PBESol functional, implemented in VASP)[21] to create a comprehensive library of all plausible Pt/MgO structures. As summarized in Figure 3.3a-c, we considered three representative MgO facets (terrace sites: [100], Mg vacancy terrace sites: [100]Mg-vac, step sites: [310]), various adsorbates ($*O$, $*O_2$), vacancy types (O_{vac}, Mg_{vac}), and sub-surface Pt locations.[22] These structures spanning a wide range of local Pt environments were used to construct the temperature-dependent phase diagram (using pMuTT, harmonic approximation for vibrational entropy) for each site.[23] We chose to analyze each facet separately as the mechanism of Mg-vacancy formation (and therefore, Mg chemical potential (μ_{Mg})) during the high-temperature calcination is not known.

The phase diagrams in Figure 3.3d-f show the relative stabilities of various Pt sites (i.e., locations and adsorbate coverages) for the [100], [100]Mg-vac, and [310] facets. Our calculations for the stoichiometric [100] and [310] surfaces show that surface Pt ([100]/sub0/$*O_2$) and leading step edge ([310]/pos1/$*O_2$) sites are favored under our experimental conditions (T = 300 K, P_{O2} = 1 bar), respectively. Although the results demonstrate weaker O_2 binding for [100]/sub1 (O_2 desorption is predicted at T > 600 K), the unsaturated [310]/pos1 (orange gold triangles) and [310]/pos2 (orange gold squares, 0.1 eV less stable) sites are characterized by strong $*O$ adsorption associated with the undercoordinated Pt atom. In contrast, no oxygen adsorption is predicted on the most stable [100]Mg-vac/sub1 structure at room temperature. Next, we compared the relative stabilities of Pt atoms at the surface and in sub-surface layers for the various MgO facets. Significantly, for the stoichiometric [100] facet, Pt sites in the first (second) sub-surface layer are 1.3 (2.1) eV less favorable energetically than those at the MgO surface (Figure

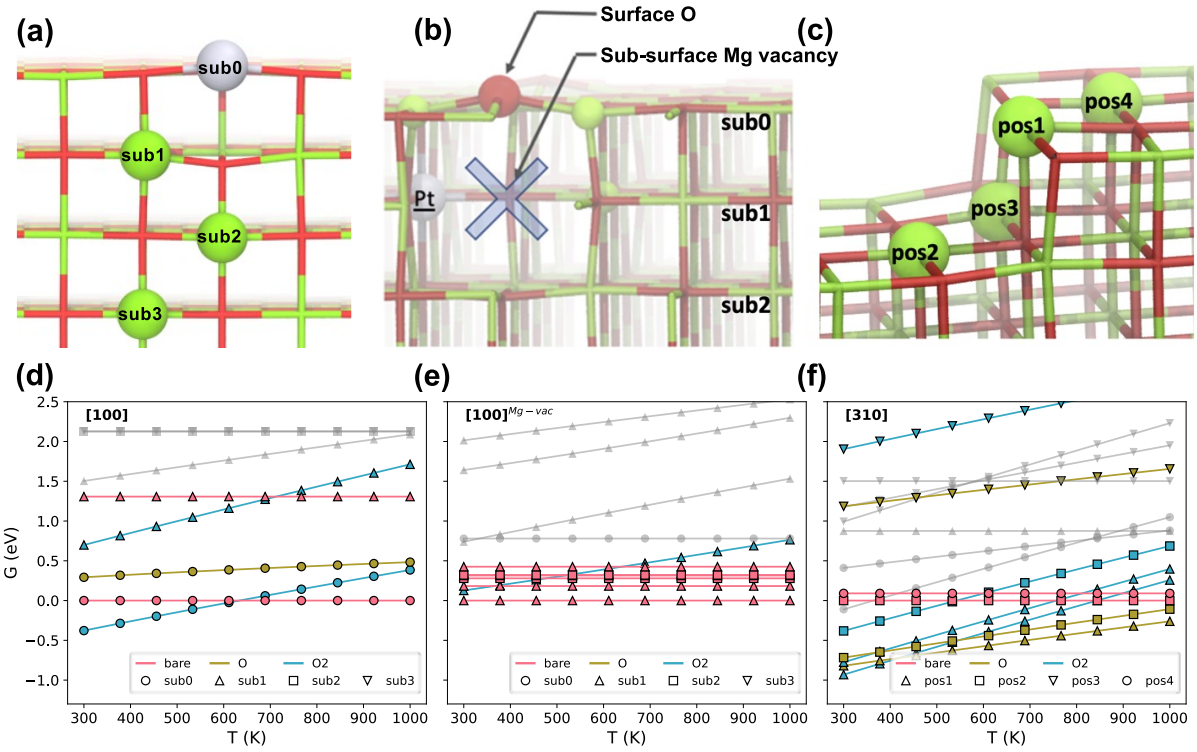

Figure 3.3. Atomistic models for the (a) [100], (b) [100]Mg-vac, and (c) [310] MgO facets considered in this study. Color scheme: Mg (green), O (red), Pt (grey). The corresponding temperature-dependent phase diagrams are summarized in (d), (e), and (f). Notation scheme for [100] and [100]Mg-vac MgO terraces: top surface (sub0), first (sub1), second (sub2), and third (sub3) sub-surface layers. For the [310] facet, we considered various step site positions including leading step edge (pos1), lower step edge (pos2), underneath step edge (pos3), and behind step edge (pos4). The phase diagrams were calculated relative to the most-stable "bare" Pt-configuration (i.e., without adsorbates). All DFT-optimized structures are available as an Atomic Simulation Environment (ASE) database file[24] in the Supporting Information.

3.3d-f). Although surface Pt sites are preferred for the stoichiometric [100] surface (i.e., [100]/sub0, circles in Figure 3.3d), the calculations show that creation of a Mg-vacancy (denoted as [100]Mg-vac) changes the relative stabilities. Specifically, in contrast to that in the [100] facet, Pt substitution in the first sub-surface layer ([100]Mg-vac/sub1) is energetically more favorable than that in the surface layer (0.8 eV less stable, [100]Mg-vac/sub0) or the second sub-surface layer (0.3 eV less stable, [100]Mg-vac/sub2) (Figure 3.3d-f). These computational results are based on the PBESol functional; similar calculations with other functionals (RPBE/D3(BJ) and PBE/D3(BJ)) did not qualitatively change the trends (Figure 3.S8).

3.3.4 Automated DFT-based workflows for fitting EXAFS data with improved thoroughness and objectivity.

Although conventional EXAFS modeling indicates high Pt–O and Pt–Mg CNs and the absence of a Pt–Pt contribution, we emphasize that this "hands-on" approach is limited to analyzing a few plausible, average scattering paths and does not identify a specific bonding configuration. This approach is therefore not sufficient to characterize the possible anisotropy of the individual Pt–O and Pt–Mg contributions. To overcome this limitation, we used the open-source X-ray analysis package Larch to perform EXAFS analysis of all 47 DFT-optimized structures mentioned above.[25] The DFT-optimized structures were used "as-is" in the fitting, with fixed CNs and spatial orientations of Mg and O atoms in the fitting. In contrast to the typical approach whereby the average nearest-neighbors fits (described by the average bond lengths and CNs) are used to create structural models, we considered all relevant scattering paths (> 250 in some cases) for each DFT-optimized structure. Notwithstanding the increased sophistication of this approach, it provides a typical EXAFS fit for each candidate structure in a few minutes, without user intervention.

As the contribution of each scattering path to the total EXAFS depends on various parameters, it is important to limit the total number of variables in the fitting. This efficiency is achieved by (1) hierarchically classifying the individual paths into distinct categories based on their effective scattering distances (R_{eff}) and the identity of the scattering atom (if $R_{eff} < 3.2$ Å) and (2) limiting the number of EXAFS variables optimized across each category. Specifically, for each DFT-optimized structure, our algorithm uses the following EXAFS fitting parameters: one energy alignment parameter (denoted as $\Delta E0$); a single value is used for all paths three distance scaling parameters, α_1, α_2, and α_3, one each for three different R_{eff} ranges, where $R_{eff,optimized} = \alpha_i \times R_{eff,DFT}$ and five mean square variations in path lengths that depend on the scatterer identity (e.g., $\sigma^2_{1,Pt-O}$ and $\sigma^2_{1,Pt-Mg}$ if $R_{eff} < 3.2$ Å) and the effective scattering distance (e.g., σ^2_3, σ^2_4, σ^2_5). The hierarchical classification is summarized in Figure 3.5 and Table 3.1. The details of this approach and strategy of this categorization are discussed in the Supporting Information.

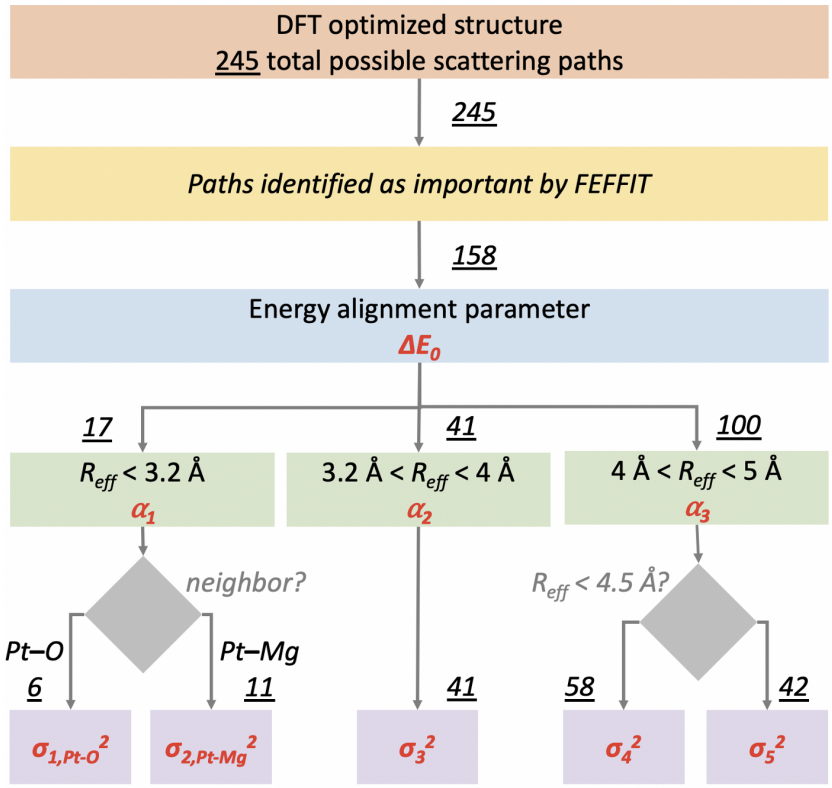

Figure 3.4. Flow-chart for hierarchically classifying all the possible paths into categories. The underlined numbers show the number of paths in the hierarchical classification approach, and the text in red identifies fitted parameters.

The scattering path classification scheme described above ensures that (1) the constraints on the number of fitting parameters (calculated on the basis of the Nyquist criterion) are not violated and (2) physically relevant characteristics of scattering paths (e.g., different mean square variations in path length, that is, $\sigma^2_{1,Pt-O}$ and $\sigma^2_{1,Pt-Mg}$ for Pt–O and Pt–Mg scatterers) are faithfully captured. The use of DFT structures enables modeling of the longer effective scattering distances ($\sim$ 5 Å), including contributions from multiple scattering paths using distance-dependent fitting parameters (e.g., σ_i^2, and α_i) without overfitting the data. These points are well-illustrated by our analyses for the [100]Mg-vac/sub1 site (Figure 3.4), whereby only 9 parameters (Table 3.1 and 3.2) are used to model all 158 unique scattering paths.

Table 3.1. Summary of hierarchical approach used to classify all possible scattering paths into five different categories. Only 9 parameters (shown in red) are optimized in the automated EXAFS analyses.

Category	**Classification criterion**	**Number of paths**	**Variables**		
1	$R_{eff} < 3.2$ Å and Pt–O scatter	6	ΔE_0	α_1	$\sigma^2_{1,Pt-O}$
2	$R_{eff} < 3.2$ Å and Pt–Mg scatter	11			$\sigma^2_{1,Pt-Mg}$
3	3.2 Å $< R_{eff} <$ 4 Å	41		α_2	σ^2_3
4	4 Å $< R_{eff} <$ 4.5 Å	58		α_3	σ^2_4
5	4.5 Å $< R_{eff} <$ 5 Å	42			σ^2_5

Table 3.2. Optimized values of the 9 fitting parameters for the [100]Mg-vac/sub1 site.

Category	ΔE_0 **(eV)**	α_i	10^3 x σ_i^2 (Å^2)	$R_{eff,optimized}$ (Å^2)
1	4.8 ± 1.9	0.994 ± 0.005	3.2 ± 1.0	2.01, 2.01, 2.02, 2.03, 2.04, 2.05
2			6.7 ± 1.6	2.96, 2.97, 2.98, 2.98, 2.99, 2.99, 3.00, 3.00, 3.01, 3.01, 3.02
3		0.987 ± 0.011	6.3 ± 7.0[a]	3.40 – 3.79[b]
4		0.996 ± 0.001	6.2 ± 5.4[a]	4.02 – 4.40[c]
5			5.3 ± 4.3[a]	4.59 – 4.98[d]

The above fitting workflow was repeated for each unique DFT structure. In addition to the reduced-χ^2 metric (commonly used by the XAS community), we quantified the agreement between the experimental and simulated EXAFS spectra for each DFT structure by using the Fréchet distance in R-space for the magnitude portions of the data and the model (denoted as δF, which measures the similarity between two curves). The histogram in Figure 3.5a shows the reduced-χ^2 of the EXAFS fit corresponding to each DFT structure, along with the Fréchet distance (a lower value is better), corresponding to each structure. In general, we observed that the [100]Mg-vac structures (blue bars in Figure 3.5a) are in better agreement with experiment (reduced-$\chi^2 < 20$, $\delta F < 2$). Although some of the [100] and [310] Pt configurations (e.g., [100]/sub1 and [310]/pos3/$*O_2$) show satisfactory agreement with experimental EXAFS (reduced-$\chi^2 < 35$, $\delta F < 4.5$), our DFT calculations show these configurations to be highly unstable (1.5 eV and 0.99 eV, respectively).

This interplay between thermodynamic stability of a structure and its consistency with the EXAFS data is quantitatively illustrated by comparing the quality of the EXAFS fits with the DFT-calculated Boltzmann fractions (at 300 K) (Figure 3.5b). Unsurprisingly, only the most stable structures ([100]/sub0/$*O_2$ (orange circle), [100]Mg-vac/sub1 (blue triangle), and [310]/pos1/$*O_2$ (green triangle)) were observed in any significant number (based on Boltzmann distributions). The EXAFS fits corresponding to the three stable Pt configurations are presented in Figure 3.5c-e, showing that the [100]Mg-vac/sub1 (CN_{Pt-O} = 6, CN_{Pt-Mg} = 11) model agrees most closely with experiment. Specifically, although the [100]/sub0/$*O_2$ structure agrees satisfactorily with the Pt–O scattering path (CNPt-O = 7), the Pt–Mg path is underestimated, corresponding to the much lower Pt–Mg contribution ($CN_{Pt-Mg} = 8$). We emphasize that the [310]/pos3 site (as suggested by Sarma et al.)[26] has coordination numbers similar to those characterizing [100]Mg-vac/sub1, but it does not agree well with the EXAFS data (Figure 3.S10). These results demonstrate that both the effects of coordination number (i.e., [100]/sub0/$*O_2$ vs. [100]Mg-vac/sub1) and the local spatial orientation of the scattering atoms (i.e., [310]/pos3 vs. [100]Mg-vac/sub1) are captured in our EXAFS fitting approach. At this stage, it is useful to highlight an

additional advantage of our EXAFS fitting procedure, with [100]Mg-vac/sub1 used as an illustrative example. Instead of using the average Pt–O (and/or Pt–Mg) distances and coordination numbers as the fitting parameters (as done conventionally), in our approach each individual Pt–O (total 6 for [100]Mg-vac/sub1 as an example) and Pt–Mg (total 11) scattering path is considered separately. This procedure allows us to capture the anisotropies of the local bonding environment, as indicated by the DFT-calculated bond lengths. Overall, we considered 158 total paths, which included single, double, and multiple scattering paths (involving up to 5 scattering atoms), using only use 9 fitting variables (Tables 3.1 and 3.2) to achieve unprecedented agreement with the experimental EXAFS across the entire R-range (Figure 3.5g). As the entire workflow relies on physically motivated models, it is possible to bolster the self-consistency between DFT structures and EXAFS fits. Beyond providing good models for the EXAFS fits, the DFT calculations also provide force constants that can be used to generate ab-initio vibrational mean-square relative displacements for the EXAFS (i.e., σ^2).[27–29] Specifically, the values for σ^2 for Pt–Mg (0.005 Å^2) and Pt–O (0.003 Å^2) obtained from DFT-calculated harmonic frequencies are consistent with the results of the EXAFS fitting (Pt–Mg: 0.0067 ± 0.001 Å^2, Pt–O: 0.0032 ± 0.001 Å^2). This point highlights the uniqueness of our approach: (1) the DFT-optimized structures are used to perform an objective, thorough EXAFS fitting, and (2) the EXAFS fitting parameters (themselves determined using the automated Python workflows) are consistent with independently calculated thermal disorders (from DFT vibrations). This level of detail has not been reported for any atomically dispersed catalysts, and we posit that it represents a significant step forward in automated analysis, self-consistent interpretation of EXAFS data.

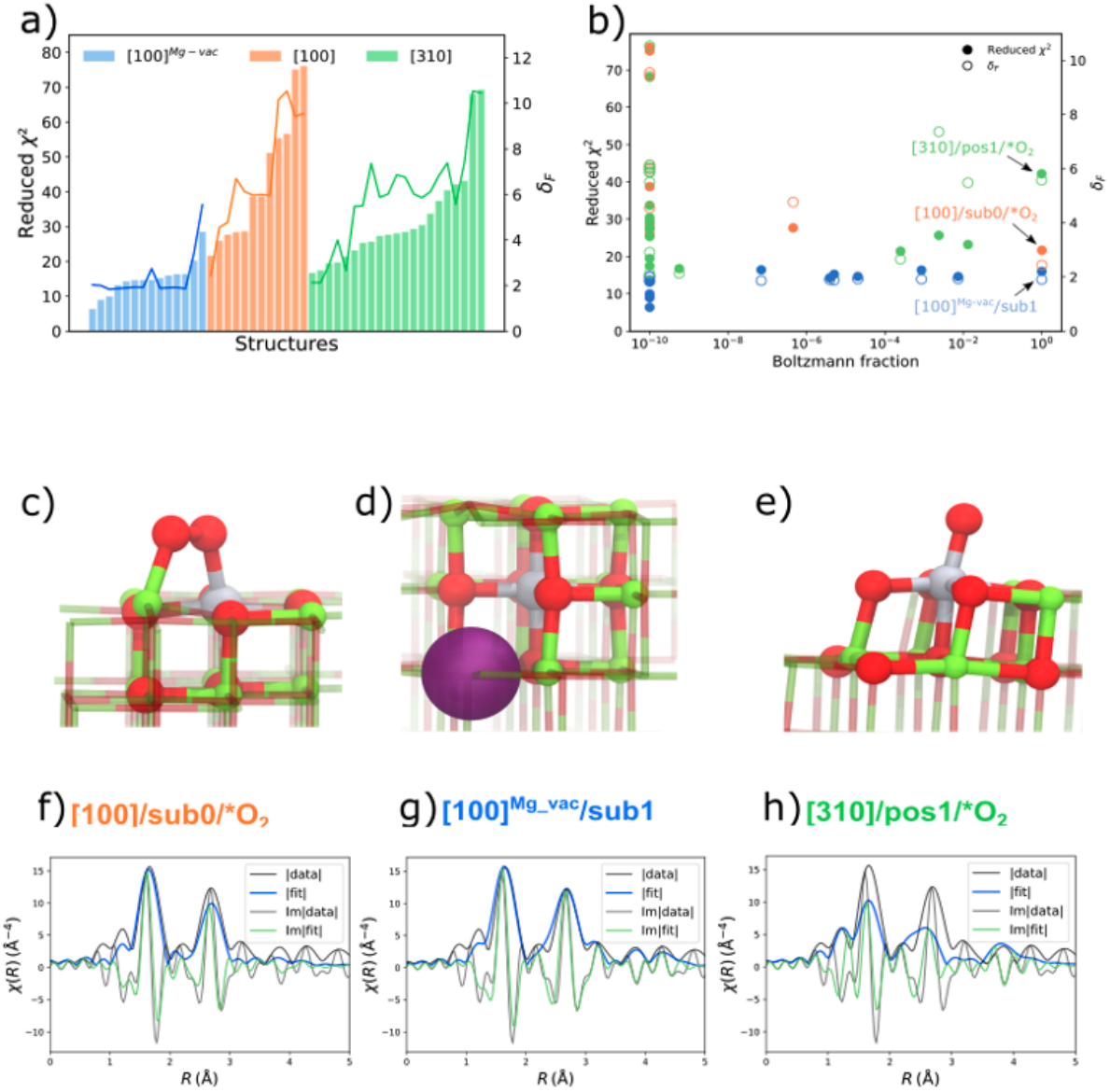

Figure 3.5. (a) Reduced-χ^2 and Fréchet distance (δ_F) of EXAFS analysis for all the DFT-optimized structures, (b) Boltzmann fraction calculations for all facets. The DFT optimized geometries of the most stable (c) [100]/sub0/$*O_2$ (d) [100]Mg-vac/sub1, and (e) [310]/pos1/$*O_2$ structures are shown with the corresponding EXAFS fits in (f), (g), and (h) showing the magnitude (fit: blue, experiment: black) and imaginary portions (fit: green, experiment: black) of Fourier transforms. The k-range of 2.2–12.5 $Å^{-1}$ and the R-range of 1.0–5.0 Å were used for the fits. Colors: Mg (green), O (red), Pt (grey). The purple sphere in (d) represents the sub-surface Mg-vacancy.

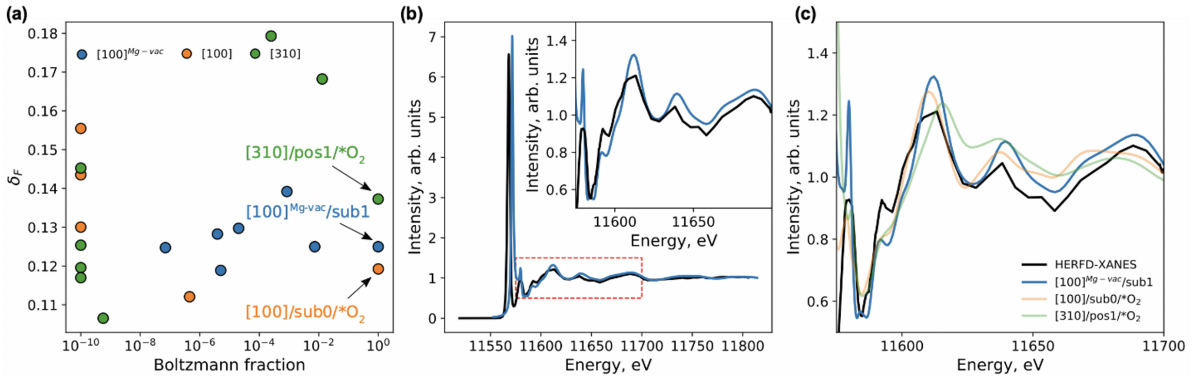

Figure 3.6. (a) Fréchet distances from comparison of HERFD-XANES data with FEFF-simulated XANES and DFT-calculated Boltzmann fractions. (b) Simulated and experimental XANES spectra for [100]Mg-vac/sub1. Inset (dashed red box) shows the region of the spectra from 11575–11700 eV. (c) Comparison of FEFF-simulated XANES spectra of the most stable structures from each facet ([100]Mg-vac/sub1, [100]/sub0/$*O_2$, and [310]/pos1/$*O_2$) with HERFD-XANES data from 11575–11700 eV.

3.3.5 XANES evidence of structure.

In addition to the EXAFS, we also consider the XANES data (Figure 3.S11; data characterizing reference compounds are shown for comparison). XANES spectra provide evidence of the Pt electronic structure, as the intensity of the white line at the Pt L3 edge is a measure of the unoccupied d-states and occupied p-states (2p3/2 → 5d3/2 or 5d5/2).[30] The oxidation state of the supported Pt is close to that of Pt^{4+} in $H_2Pt(OH)_6$ and markedly different from that of the Pt^{2+} in $Pt(acac)_2$ (acac = acetylacetonato, $[C_5H_7O_2]^-$) and Pt^0 in Pt foil (Figure 3.S11). These data imply that isolated Pt^{4+} was present in six-coordinate octahedral geometry in cation vacancy sites, consistent with the STEM and EXAFS data. To account for charge balance, one Pt^{4+} in an MgO sub-surface site requires one Mg^{2+} vacancy, which implies a Pt-Mg coordination number of 11 and not 12, consistent with the EXAFS analyses. These results further agree with DFT-calculated Bader charges that show consistent oxidation states for the sub-surface Pt^{4+} with one Mg vacancy and the Pt^{4+} in $H_2Pt(OH)_6$ (q_{Bader} = +1.35e and +1.41e, respectively). These values are much higher than q_{Bader} = +0.86e for the Pt^{2+} in $Pt(acac)_2$.

We also collected HERFD-XANES spectra; HERFD-XANES minimizes the limitations of energy resolution affected by core-hole lifetime broadening and beamline optics in conventional XANES.[31] A comparison of the HERFD and conventional XANES data (Figure 3.S11) shows the advantages of HERFD in resolving near-edge features including (a) the stable near-zero signal in the pre-edge; (b) strong intensities and sharpness without a long tail characterizing the white line; and (c) the clear resonance feature with little broadening beyond the white line. The HERFD-XANES data provide insight into the catalyst structure beyond what is determined by the EXAFS data and DFT-guided analysis.[32] The HERFD-XANES data were compared with FEFF-simulated XANES spectra based on the aforementioned DFT-determined structures (Figure 3.6).[33,34] The agreement between experimental and simulated XANES spectra (quantified using the Fréchet distance (δ F)) is compared with the thermodynamic stability (Figure 3.6a) of various Pt sites. Figure 3.6b, c show a comparison of the experimental HERFD and the FEFF simulations for each of the most stable structures predicted from DFT-calculated Boltzmann fractions:

[100]/sub0/$*O_2$ (99.2%), [100]Mg-vac/sub1 (99.9 %), and [310]/pos1/$*O_2$ (98.4%). The FEFF results were shifted 4.5 eV and scaled to match the normalized experimental results for values > 11650 eV. Of all the stable structures, both [100]/sub0/$*O_2$ and [100]Mg-vac/sub1 show good agreement with the experimental XANES; however, by comparing the corresponding EXAFS spectra in Figure 3.6f and g, we see that the Mg-vacancy provides a notable improvement over [100]/sub0/$*O_2$. Although some [310] structures are more consistent with experiment than these ($\delta F < 0.125$), they are significantly less stable (Boltzmann fraction $< 10^{-9}$), and, we infer, unlikely to exist. The most stable [310] structures show poor agreement with experiment (Figure 3.6c (green)). Taking together the unique combination of EXAFS and HERFD-XANES fits of [100]Mg-vac/sub1 and the DFT-predicted stability, our results show that the Pt/MgO catalyst is well-characterized as an atomically dispersed Pt^{4+} ion embedded in the first subsurface layer (sub1) located adjacent to a Mg-vacancy site. As shown in Figure 3.S12, our HERFD-XANES analyses and DFT calculations do not identify the exact location of the Mg-vacancy site; we used the most stable structure, Mg-vacancy in the surface layer (sub0) for further analysis. To analyze the origin of the various features in the HERFD, Figure 3.S13 shows the locally projected, l-dependent density of states (l-DOS) of the [100]Mg-vac/sub1 site. Overall, the Pt L3 HERFD spectra resemble the Pt d-DOS, as indicated by the vertical dashed lines. The white line is composed largely of Pt d-DOS with a small contribution from the O p-DOS, with its position being largely insensitive to the local disorder. The second peak ($\sim$ 9 eV above the Fermi level) shows contributions from both the O and Mg atoms as well as the local Pt s-DOS, indicating that this feature corresponds to non-local states and arises at least in part from bonding of Pt to the near-neighboring atoms (Figure 3.S13).

3.3.6 CO oxidation catalyzed by Pt/MgO.

We characterized the Pt/MgO as a catalyst for CO oxidation as a probe reaction. The light-off curves (Figure 3.7a) show that the reaction became readily measurable at approximately 180 °C, with the CO conversion reaching nearly 100% at 280 °C under our conditions. Separate experiments were carried out to determine reaction rates (turnover

frequencies, TOF, rates per Pt atom) from data obtained at low conversions (<5%), which were shown to be differential by the linear dependence of conversion on inverse space velocity at temperatures of 180–200 °C (Figure 3.7b,c). An Arrhenius plot (Figure 3.7d) indicates an apparent activation energy of 79 ± 2 kJ mol^{-1}. Data showing conversion as a function of time onstream in the flow reactor under steady-state conditions (Figure 3.S14) indicate that the catalyst was stable, retaining its activity for as long as it was onstream (48 h) at 210 °C. The catalyst remained white in color after use. The HAADF-STEM images of used Pt/MgO samples (after three light-off experiments) show that Pt was still atomically dispersed (Figure 3.S15), even at a high temperature (300 °C). As further checks of the catalyst stability, we recorded in-operando HERFD-XANES data during CO oxidation at 210 °C (Figure 3.S16), demonstrating unchanged spectra.

This catalytic probe reaction provided yet another opportunity to verify the identity of the support sites for Pt. Figure 3.7e shows trends in CO adsorption for the stable [100]Mg-vac/sub1 Pt terrace site at various temperatures. Unsurprisingly, stronger CO adsorption was found for the undercoordinated Pt site (i.e., [100]/sub0); we predicted CO adsorption at temperatures as high as about 83 °C (Figure 3.S17). In contrast, the sub-surface Pt site ([100]Mg-vac/sub1) is characterized by weaker CO binding (0.25 eV at 83 °C); no CO adsorption was predicted at room temperature. The trends for the [100]Mg-vac/sub1 site are consistent with (1) our IR experiments that do not show significant CO uptakes at room temperature (Figure 3.S18) and (2) Sarma's report[26] of low-temperature (-163 °C) IR spectra of CO on a sample similar to ours; a 2166 cm^{-1} ν_{CO} band was observed, agreeing within error with our calculated frequency (2141 cm^{-1}).

It is significant, and a challenge to our interpretation, that our Pt/MgO catalyst, which consists of sub-surface, highly coordinated Pt sites, does not bind CO at room temperature but nonetheless is catalytically active for CO oxidation. Thus, we turned to DFT calculations (RPBE/D3-BJ functional, nudged elastic band and dimer method for barriers) to investigate various possible mechanisms of CO oxidation on the [100]Mg-vac/sub1 site, and the two most favorable mechanisms are shown in Figure 3.8. Indeed, our calculations show that the insertion of CO into the surface-bound O_2 is the rate limiting

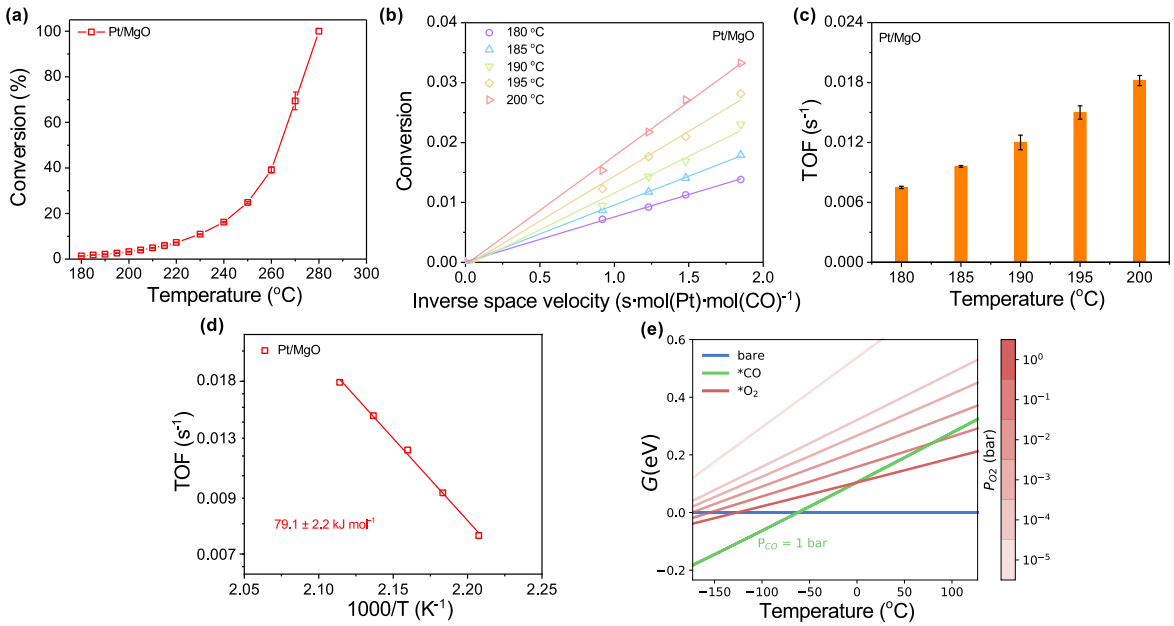

Figure 3.7. (a) Light-off curve characterizing CO oxidation catalyzed by Pt/MgO. Error bars represent standard deviation determined from three independent measurements. The once-through plug-flow reactor was heated from room temperature at a rate of 2 °C/min to 300 °C. The feed was 5.0% CO in helium flowing at 4.0 mL (NTP)/min + 5.0% O_2 in helium flowing at 16.0 mL (NTP)/min; the catalyst mass was 100 mg. (b) Demonstration of differential conversion at various temperatures. Low-conversion (<5%) data were obtained under the conditions stated in a, except that the feed gas was 5.0% CO in helium flowing at 5.0, 6.0, or 8.0 mL (NTP)/min + 5.0% O_2 in helium flowing at 20.0, 24.0, or 32.0 mL (NTP)/min; the catalyst mass was 100 mg. (c) TOF values determined from slopes of lines in b, with errors. (d) Arrhenius plot. (e) DFT-calculated CO and O_2 adsorption at various temperatures and O_2 partial pressures.

step—the calculated enthalpic barrier (105.3107.4 kJ/mol) is in satisfactory agreement with experimentally determined apparent activation energies (Figure 3.8a); similar carbonate species were proposed by Sarma et al.[26] for Pt at the step site on MgO. The addition of the subsurface Pt and Mg vacancy activates the bound O_2 and allow for CO insertion to form the intermediate carbonate. Charge density difference plots showing the change in the electronic structure of the $*O_2$ compared to pristine MgO are shown in Figure 3.S19. Although the role of the Pt atom in this mechanism differs greatly to what has been reported by Sarma, the calculated free energy barrier at 300 °C in this work is slightly lower (169 kJ/mol compared to 173 kJ/mol[26]). Additionally, a similar Eley–Rideal mechanism has been seen in a MgO/Ag system.[35] A second mechanism con-

sisting of the abstraction of a surface oxygen by CO to form an oxygen vacancy as the rate limiting step is shown in Figure 3.8b. The similarity of the free energy barrier of these two mechanisms (170.2 kJ/mol versus 168.8 kJ/mol) make it challenging to discern which mechanism is more favorable and both appear reasonable according to our calculations.

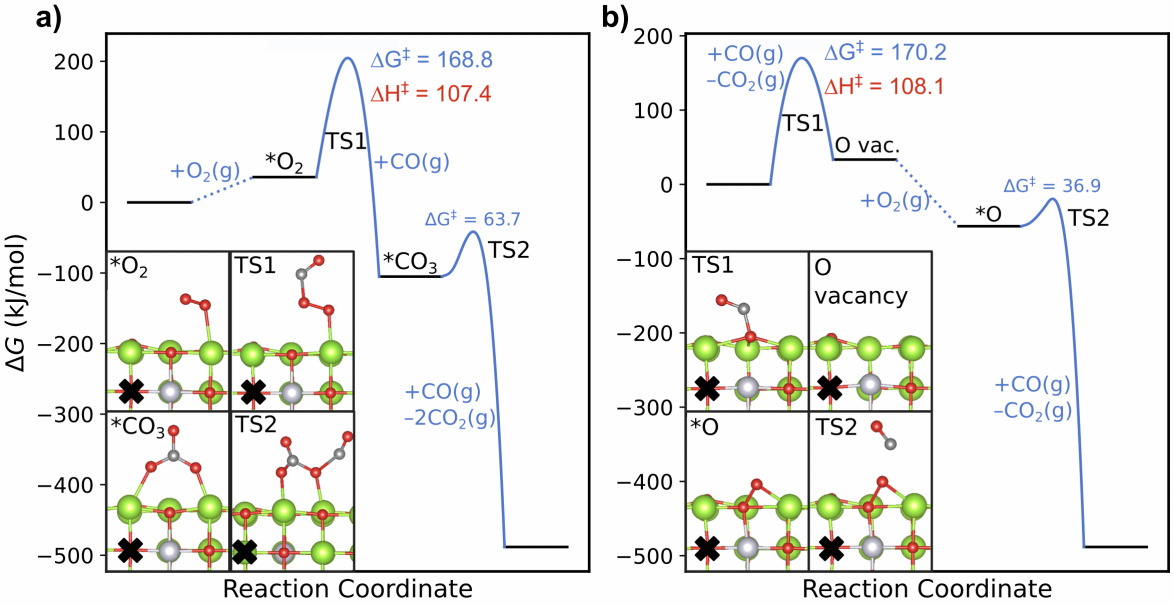

Figure 3.8. Reaction free energy diagrams for CO oxidation on [100]Mg-vac/sub1 surface at 300 °C. Mechanism a) consists of an Eley–Rideal type mechanism with the formation of a carbonate intermediate, while mechanism b) proceeds through the formation of an oxygen vacancy above the subsurface Pt. Calculated free energy (blue) and enthalpic (red) barriers are annotated for each transition state. Insets show the optimized structures of the key intermediate species with the location of the Mg vacancy marked with a black cross (colors: Mg, green; O, red; Pt, light grey; C, dark grey).

A pertinent question is whether the catalyst underwent structural changes under the conditions of CO oxidation catalysis. The EXAFS data recorded after catalysis were not substantially different from those observed before, with [100]Mg-vac/sub1 again emerging as the best-fit model of the EXAFS spectra (Figure 3.S20). This comparison in in line with the catalyst performance data, indicating that the MgO- supported platinum sites are stable under CO oxidation conditions.

3.4 Discussion

The intense recent attention paid to atomically dispersed supported metal catalysts reflects the promise of catalysts with valuable new properties and the prospect that some will be so simple in structure as to propel significant advances in the understanding of surface catalysis broadly at the atomic level.[9,36] Some literature reflects the notion that these catalysts are straightforward to understand because the metals are atomically dispersed[36], but this view overlooks the intrinsic complexity of the supports and the need to incorporate the metal–support combinations in realistic models of the catalysts. Thus, there is a motivation to work with metal–support-site combinations that are nearly unique and can be understood in depth. Consequently, researchers have been motivated to use crystalline metal oxides as supports and to use low metal loadings to create catalysts that can be approximated as isolated metals in unique, stable surroundings. The characterization data presented here validate that approach. They have a high degree of internal consistency and show that the structure of the catalyst consisting of Pt in a loading of only 0.05 wt% on crystalline MgO powder is well represented with a single, stably encapsulated—yet still catalytically active—structure that meets the criterion of near structural uniqueness. A full set of complementary experimental characterizations combined with theoretical verification was needed to test the hypothesis that such a catalyst could be characterized structurally with some confidence. The DFT-guided EXAFS fitting approach described here represents, we posit, a substantial advance in the characterization of such catalysts. Although previous approaches represent attempts to develop DFT models that are motivated by results of conventional EXAFS analyses, we instead used a library of DFT-optimized Pt structures to identify the single best representation of the all the experimental results (e.g., EXAFS, HERFD-XANES, IR spectra of adsorbed CO, and catalytic kinetics) collected in this investigation. Going forward, we anticipate that further development of this combination of characterization techniques will help guide the choice of metal–support combinations to advance the field of atomically dispersed supported metal catalysts; help advance the understanding of the broad class of supported catalysts; and help guide the development of new and improved methods for

understanding of even more complex surface structures.

3.5 Experimental and Computational Methods

3.5.1 Synthesis of Pt/MgO samples.

MgO powder (1000 mg, US Research Nanomaterials) with a manufacturer-specified specific surface area of 25 m2/g was dispersed in 250 mL of ethanol in a 500-mL beaker with magnetic stirring. Aqueous K_2PtCl_4, 20 mL of 0.25 mM solution, and 40 mL of ethanol were transferred into the stirred beaker through a syringe pump at a rate of 5 mL/h. The resultant solid was collected by centrifugation, washed with distilled water, and then dried overnight in air in an oven at 80 °C. The resultant powder was calcined in flowing O_2 (10 mL(NTP)/min) and N_2 (40 mL(NTP)/min) as the temperature was ramped from room temperature to 120 °C and held for 2 h to remove any residual water and organics. The sample was then further calcined in a mixture of flowing O_2 (10 mL(NTP)/min) and N2 (40 mL(NTP)/min) as the temperature was ramped at a rate of 5 °C/min until a final temperature of 700 °C was reached, which was held for 4 h.

3.5.2 X-ray absorption spectroscopy.

XAS was carried out at beamlines 4-1 and 6-2 at the Stanford Synchrotron Radiation Lightsource (SSRL). At beamline 4-1, which is a side station on a 20-pole wiggler beamline, the ring SPEAR3 was operated in top-off mode with a storage ring energy of 3 GeV and 500 mA stored current. A Si double-crystal (220) monochromator was detuned by 20–25% of maximum intensity to minimize harmonics at the Pt L3 edge. XAS data were collected in fluorescence yield mode with a 30-element germanium solid-state detector array with samples pressed into a pellet at 25 °C. For energy calibration, a Pt foil reference was placed between the ion chambers upbeam and downbeam of the sample so that its spectrum was measured simultaneously with that of the sample.

At beamline 6-2, which is fed by a 56-pole, 0.9 Tesla wiggler, high-energy resolution fluorescence detection HERFD XANES experiments were conducted with a liquid-nitrogen-cooled double-crystal Si (311) monochromator to select the energy of the incident beam. A Rowland circle spectrometer (radius 1 m) equipped with three spherically bent Si (800)

analyzers and a silicon drift detector were used to select the Pt-Lα emission line. A Pt foil was scanned in the transmission mode for initial energy calibration. For ex-situ HERFD-XANES experiments, the sample was pressed into a pellet at 25 °C. In each in-operando HERFD-XANES experiment, approximately 50 mg of catalyst sample was loaded into a flow-through cell, a Kapton tube (i.d. = 2.8 mm), connected to a treatment gas line.[37] The compositions of effluent gases flowing from the cell were measured with an online mass spectrometer (Hiden HPR20). The Pt/MgO was first heated in helium flowing at 20 mL(NTP)/min as the temperature was ramped from room temperature to 210 °C at a rate of 5 °C/min. Then the feed gas was switched to a mixture of CO flowing at 0.2 mL(NTP)/min + O_2 flowing at 0.8 mL(NTP)/min + helium flowing at 19 mL(NTP)/min with the sample in the cell held at 210 °C for 1 h as CO oxidation catalysis took place. Thereafter, the reactor was cooled to room temperature and the gas feed switched to helium flowing at 20 mL(NTP)/min. HERFD-XANES spectra were collected periodically during all these steps.

3.5.3 IR spectroscopy.

Transmission IR spectra of powder samples in the ν_{O-H} region were determined with a Bruker IFS 66v/S spectrometer with a resolution of 4 cm^{-1}. Approximately 10 mg of sample was loaded between two KBr windows, and spectra were recorded at room temperature with the sample under vacuum, with an average of 256 scans per spectrum. IR spectra of samples with adsorbed CO were determined with approximately 40 mg of sample pressed into a wafer loaded into a cell (In-situ Research Instruments, South Bend, IN) that served as a flow reactor fed with various gases, and transmission spectra of catalysts in the presence of these gases were recorded with a liquid-nitrogen-cooled MCT detector.

3.5.4 CO oxidation catalysis in a conventional laboratory plug-flow reactor.

The catalyst samples were evaluated for CO oxidation in a once-through plug-flow reactor, with products analyzed with an online mass spectrometer (Hiden Analytical HPR20)

equipped with a secondary electron multiplier detector used in multiple ion detection mode. Samples of catalyst powder (100 mg) that had passed through a 40 to 60 mesh sieve were loaded into a quartz tube reactor (i.d. = 4 mm), with the upstream and downstream sections packed with quartz wool. In experiments to determine light-off curves for CO oxidation, the feed was a mixture of 5.0% CO in helium flowing at 4.0 mL(NTP)/min + 5.0% O_2 in helium flowing at 16.0 mL(NTP)/min; the pressure was atmospheric. The reactor was heated from room temperature at a rate of 2 °C/min with the gases flowing and then held at 300 °C for 20 min before the reactor was cooled down. In a separate flow reactor system, near-steady-state conversion data were obtained under the conditions stated in the preceding paragraph, except that the feed gas was a mixture of 5.0% CO in helium flowing at 5.0, 6.0, or 8.0 mL(NTP)/min + 5.0% O_2 in helium flowing at 20.0, 24.0, or 32.0 mL(NTP)/min. Catalytic reaction rates were calculated from low (differential) conversions (<5%) determined at temperatures in the range of 180–200 °C. In experiments determining the dependence of conversion on time on stream, the samples were kept on stream for up to 48 h to demonstrate stability at 210 °C; the feed was 5.0% CO in helium flowing at 4.0 mL(NTP)/min and 5.0% O_2 in helium flowing at 16.0 mL(NTP)/min.

3.5.5 Scanning transmission electron microscopy.

STEM imaging experiments were performed on a JEOL 200CF (NEOARM) transmission electron microscope with an acceleration voltage of 200 kV. The specimens were prepared by a direct dispersion of powder samples on lacey carbon grids. The HAADF images were acquired with a convergence angle of 28.5 mrad and an inner collection angle of 55 mrad.

3.5.6 Transmission electron microscopy.

TEM images of as-prepared sample powders loaded onto copper grids were recorded with a Hitachi H-7700 microscope at an acceleration voltage of 100 kV.

3.5.7 X-ray diffraction crystallography

XRD patterns of the samples were collected on a Philips X'Pert Pro Super diffractometer with a monochromatized Cu Kα radiation source and a wavelength of 0.1542 nm.

3.5.8 Inductively coupled plasma mass spectrometry.

The Pt loadings in the catalysts were determined by inductively coupled plasma mass spectrometry with a Thermo Scientific XSERIES 2 instrument.

3.5.9 DFT calculations.

Periodic DFT calculations were performed using the projector augmented wave method as implemented in the Vienna ab initio simulation package (VASP).[21] Energies were calculated using a 500 eV plane-wave cutoff with a $2 \times 1 \times 1$ Monkhorst-Pack k-point grid. A range of generalized gradient approximation functionals (PBE, RPBE, and PBESol) were used. Electronic energies were converged to 10^{-6} eV, and all structures were relaxed until the forces were less than 0.05 eV/Å. The finite displacement method (0.02 Å) was used to calculate the entropic and zero-point energy corrections. For CO oxidation mechanism calculations, a 400 eV plane-wave cutoff was used with the RPBE functional, and dispersion corrections were considered with the DFT-D3 method with Becke-Johnson damping. Barriers were calculated using the climbing image nudged elastic band and dimer methods.

3.5.10 Conventional XAS data analysis.

The EXAFS and XANES data were analyzed with the Demeter package18. Pre-processing of data included alignment, edge calibration, deglitching, normalization, background subtraction, and conversion of data into a chi file for data fitting, performed with Athena. The energy at the Pt L3 edge, determined by the first inflection point of the absorption edge data characterizing the reference Pt foil, was calibrated to the reported energy, 11564.0 eV. EXAFS data were modeled with the Artemis package in Demeter. The continuous Cauchy wavelet transform (CCWT)[20] was performed on the extracted EXAFS to visualize the data and assess how many scattering paths would be needed to fit the data. Larch[25] was used to perform the CCWT on EXAFS spectra at the Pt L3 edge characterizing Pt/MgO samples and Pt foil.

3.5.11 Automated DFT-based fitting of EXAFS data.

The EXAFS data were analyzed with the open-source X-ray Larch package.[25] Data processing, such as alignment, edge calibration, deglitching, normalization, and background

subtraction, was performed using the Python interface to Larch. The DFT-optimized structures were used to generate a feff.inp file using an in-house code. These feff.inp files were run in Larch, with FEFF used for generating scattering paths. The code uses Matplotlib for plotting the fitting results.

3.5.12 FEFF and DOS modeling.

All HERFD simulations used DFT-optimized (PBE-D3) structures and were performed with FEFF 9 using SCF and FMS cutoff radii of 6 and 9 Å, respectively, which ensured convergence of the spectra.[33,34] The representative atomic potentials were chosen to reproduce chemical distinguishability for each atom type. On the basis of our experience with Pt-containing samples, we chose not to use a core-hole in the calculations.[38,39] To properly simulate HERFD rather than the default XANES in FEFF, we also removed 1.8 eV from the default core-hole lifetime broadening of 5.2 eV. The self-energy was modeled using the density-dependent MPSE (many-pole self-energy) dielectric function approach based on a weighted average of the atomic loss functions.[40] Moreover, to provide converged results up to about 200 eV above the edge, the maximum angular momenta for the site basis set were raised to 5, 4, and 4 for Pt, O, and Mg, respectively. Further, vibrational disorder was added using single-scattering Debye-Waller factors based on a correlated Debye model. The Debye temperature for this model was estimated for each system using average force constants for the first-shell Pt–O bonds obtained from the DFT structural simulations.[27,28]

3.6 References

[1] Joaquin Resasco, Leo DeRita, Sheng Dai, Joseph P Chada, Mingjie Xu, Xingxu Yan, Jordan Finzel, Sergei Hanukovich, Adam S Hoffman, George W Graham, Simon R Bare, Xiaoqing Pan, and Phillip Christopher. Uniformity Is Key in Defining Structure–Function Relationships for Atomically Dispersed Metal Catalysts: The Case of Pt/CeO2. *Journal of the American Chemical Society*, 142(1):169–184, jan 2020. doi: 10.1021/jacs.9b09156.

[2] Leo DeRita, Joaquin Resasco, Sheng Dai, Alexey Boubnov, Ho Viet Thang, Adam S Hoffman, Insoo Ro, George W Graham, Simon R Bare, Gianfranco Pacchioni, Xiaoqing Pan, and Phillip Christopher. Structural evolution of atomically dispersed Pt catalysts dictates reactivity. *Nature Materials*, 18(7):746–751, 2019. doi: 10.1038/s41563-019-0349-9.

[3] Yifeng Zhu, Simuck F Yuk, Jian Zheng, Manh-Thuong Nguyen, Mal-Soon Lee, Janos Szanyi, Libor Kovarik, Zihua Zhu, Mahalingam Balasubramanian, Vassiliki-Alexandra Glezakou, John L Fulton, Johannes A Lercher, Roger Rousseau, and Oliver Y Gutiérrez. Environment of Metal–O–Fe Bonds Enabling High Activity in CO2 Reduction on Single Metal Atoms and on Supported Nanoparticles. *Journal of the American Chemical Society*, 143(14):5540–5549, apr 2021. doi: 10.1021/jacs.1c02276.

[4] Zailei Zhang, Yihan Zhu, Hiroyuki Asakura, Bin Zhang, Jiaguang Zhang, Maoxiang Zhou, Yu Han, Tsunehiro Tanaka, Aiqin Wang, Tao Zhang, and Ning Yan. Thermally stable single atom Pt/m-Al2O3 for selective hydrogenation and CO oxidation. *Nature Communications*, 8(1):16100, 2017. doi: 10.1038/ncomms16100.

[5] Adam S Hoffman, Louise M Debefve, Shengjie Zhang, Jorge E Perez-Aguilar, Edward T Conley, Kimberly R Justl, Ilke Arslan, David A Dixon, and Bruce C Gates.

Beating Heterogeneity of Single-Site Catalysts: MgO-Supported Iridium Complexes. *ACS Catalysis*, 8(4):3489–3498, apr 2018. doi: 10.1021/acscatal.8b00143.

[6] Yongwoo Kwon, Tae Yong Kim, Gihun Kwon, Jongheop Yi, and Hyunjoo Lee. Selective Activation of Methane on Single-Atom Catalyst of Rhodium Dispersed on Zirconia for Direct Conversion. *Journal of the American Chemical Society*, 139(48): 17694–17699, dec 2017. doi: 10.1021/jacs.7b11010.

[7] Lulu Li, Xin Chang, Xiaoyun Lin, Zhi-Jian Zhao, and Jinlong Gong. Theoretical insights into single-atom catalysts. *Chem. Soc. Rev.*, 49(22):8156–8178, 2020. doi: 10.1039/D0CS00795A.

[8] Sharon Mitchell and Javier Pérez-Ramírez. Single atom catalysis: a decade of stunning progress and the promise for a bright future. *Nature Communications*, 11(1): 4302, 2020. doi: 10.1038/s41467-020-18182-5.

[9] Yizhen Chen, Hanlei Sun, and Bruce C Gates. Prototype Atomically Dispersed Supported Metal Catalysts: Iridium and Platinum. *Small*, 17(16):2004665, apr 2021. doi: https://doi.org/10.1002/smll.202004665.

[10] Maria Flytzani-Stephanopoulos and Bruce C Gates. Atomically Dispersed Supported Metal Catalysts. *Annual Review of Chemical and Biomolecular Engineering*, 3(1): 545–574, jun 2012. doi: 10.1146/annurev-chembioeng-062011-080939.

[11] Manoja K Samantaray, Valerio D'Elia, Eva Pump, Laura Falivene, Moussab Harb, Samy Ould Chikh, Luigi Cavallo, and Jean-Marie Basset. The Comparison between Single Atom Catalysis and Surface Organometallic Catalysis. *Chemical Reviews*, 120 (2):734–813, jan 2020. doi: 10.1021/acs.chemrev.9b00238.

[12] Ruixuan Qin, Kunlong Liu, Qingyuan Wu, and Nanfeng Zheng. Surface Coordination Chemistry of Atomically Dispersed Metal Catalysts. *Chemical Reviews*, 120(21): 11810–11899, nov 2020. doi: 10.1021/acs.chemrev.0c00094.

[13] Bruce C Gates. Atomically Dispersed Supported Metal Catalysts: Seeing Is Believing. *Trends in Chemistry*, 1(1):99–110, 2019. doi: https://doi.org/10.1016/j.trechm.2019.01.004.

[14] Selina K Kaiser, Zupeng Chen, Dario Faust Akl, Sharon Mitchell, and Javier Pérez-Ramírez. Single-Atom Catalysts across the Periodic Table. *Chemical Reviews*, 120 (21):11703–11809, nov 2020. doi: 10.1021/acs.chemrev.0c00576.

[15] Yubing Lu, Jiamin Wang, Liang Yu, Libor Kovarik, Xiwen Zhang, Adam S Hoffman, Alessandro Gallo, Simon R Bare, Dimosthenis Sokaras, Thomas Kroll, Vanessa Dagle, Hongliang Xin, and Ayman M Karim. Identification of the active complex for CO oxidation over single-atom Ir-on-MgAl2O4 catalysts. *Nature Catalysis*, 2(2):149–156, 2019. doi: 10.1038/s41929-018-0192-4.

[16] Janis Timoshenko and Anatoly I Frenkel. "Inverting" X-ray Absorption Spectra of Catalysts by Machine Learning in Search for Activity Descriptors. *ACS Catalysis*, 9 (11):10192–10211, nov 2019. doi: 10.1021/acscatal.9b03599.

[17] S J Pennycook and D E Jesson. High-resolution Z-contrast imaging of crystals. *Ultramicroscopy*, 37(1):14–38, 1991. doi: https://doi.org/10.1016/0304-3991(91)90004-P.

[18] B Ravel and M Newville. ATHENA, ARTEMIS, HEPHAESTUS: data analysis for X-ray absorption spectroscopy using IFEFFIT. *Journal of Synchrotron Radiation*, 12(4):537–541, jul 2005. doi: 10.1107/S0909049505012719.

[19] Lichen Liu, Miguel Lopez-Haro, Christian W Lopes, Chengeng Li, Patricia Concepcion, Laura Simonelli, Jose J Calvino, and Avelino Corma. Regioselective generation and reactivity control of subnanometric platinum clusters in zeolites for high-temperature catalysis. *Nature Materials*, 18(8):866–873, 2019. doi: 10.1038/s41563-019-0412-6.

[20] Manuel Munoz, Pierre Argoul, and Francois Farges. Continuous Cauchy wavelet transform analyses of EXAFS spectra: A qualitative approach. *American Mineralogist*, 88(4):694–700, apr 2003. doi: 10.2138/am-2003-0423.

[21] G Kresse and J Hafner. Ab initio molecular dynamics for liquid metals. *Physical Review B*, 47(1):558–561, jan 1993. doi: 10.1103/PhysRevB.47.558.

[22] Stefano Prada, Livia Giordano, and Gianfranco Pacchioni. Charging of Gold Atoms on Doped MgO and CaO: Identifying the Key Parameters by DFT Calculations. *The Journal of Physical Chemistry C*, 117(19):9943–9951, may 2013. doi: 10.1021/jp401983m.

[23] Jonathan Lym, Gerhard R Wittreich, and Dionisios G Vlachos. A Python Multiscale Thermochemistry Toolbox (pMuTT) for thermochemical and kinetic parameter estimation. *Computer Physics Communications*, 247:106864, 2020. doi: https://doi.org/10.1016/j.cpc.2019.106864.

[24] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N Groves, Bjørk Hammer, Cory Hargus, Eric D Hermes, Paul C Jennings, Peter Bjerre Jensen, James Kermode, John R Kitchin, Esben Leonhard Kolsbjerg, Joseph Kubal, Kristen Kaasbjerg, Steen Lysgaard, Jón Bergmann Maronsson, Tristan Maxson, Thomas Olsen, Lars Pastewka, Andrew Peterson, Carsten Rostgaard, Jakob Schiøtz, Ole Schütt, Mikkel Strange, Kristian S Thygesen, Tejs Vegge, Lasse Vilhelmsen, Michael Walter, Zhenhua Zeng, and Karsten W Jacobsen. The atomic simulation environment—a Python library for working with atoms. *Journal of Physics: Condensed Matter*, 29 (27):273002, 2017. doi: 10.1088/1361-648X/aa680e.

[25] Matthew Newville. Larch: An Analysis Package for XAFS and Related Spectroscopies. *Journal of Physics: Conference Series*, 430(1):12007, 2013. doi: 10.1088/1742-6596/430/1/012007.

[26] Bidyut B Sarma, Philipp N Plessow, Giovanni Agostini, Patricia Concepción, Norbert Pfänder, Liqun Kang, Feng R Wang, Felix Studt, and Gonzalo Prieto. Metal-Specific Reactivity in Single-Atom Catalysts: CO Oxidation on 4d and 5d Transition Metals

Atomically Dispersed on MgO. *Journal of the American Chemical Society*, 142(35): 14890–14902, sep 2020. doi: 10.1021/jacs.0c03627.

[27] Fernando D Vila, J J Rehr, H H Rossner, and H J Krappe. Theoretical x-ray absorption Debye-Waller factors. *Physical Review B*, 76(1):14301, jul 2007. doi: 10.1103/PhysRevB.76.014301.

[28] Fernando D Vila, Scott T Hayashi, and John J Rehr. Efficient Calculation of the Negative Thermal Expansion in ZrW2O8 , 2018.

[29] James E Penner-Hahn. X-ray Absorption Spectroscopy. In *eLS*. may 2005. ISBN 9780470015902. doi: https://doi.org/10.1038/npg.els.0002984.

[30] Hemma Mistry, Farzad Behafarid, Simon R Bare, and B Roldan Cuenya. Pressure-Dependent Effect of Hydrogen Adsorption on Structural and Electronic Properties of Pt/γ-Al2O3 Nanoparticles. *ChemCatChem*, 6(1):348–352, jan 2014. doi: https://doi.org/10.1002/cctc.201300783.

[31] Adam S Hoffman, Dimosthenis Sokaras, Shengjie Zhang, Louise M Debefve, Chia-Yu Fang, Alessandro Gallo, Thomas Kroll, David A Dixon, Simon R Bare, and Bruce C Gates. High-Energy-Resolution X-ray Absorption Spectroscopy for Identification of Reactive Surface Species on Supported Single-Site Iridium Catalysts. *Chemistry – A European Journal*, 23(59):14760–14768, oct 2017. doi: https://doi.org/10.1002/chem.201701459.

[32] Florian Maurer, Jelena Jelic, Junjun Wang, Andreas Gänzler, Paolo Dolcet, Christof Wöll, Yuemin Wang, Felix Studt, Maria Casapu, and Jan-Dierk Grunwaldt. Tracking the formation, fate and consequence for catalytic activity of Pt single sites on CeO2. *Nature Catalysis*, 3(10):824–833, 2020. doi: 10.1038/s41929-020-00508-7.

[33] John J Rehr, Joshua J Kas, Micah P Prange, Adam P Sorini, Yoshinari Takimoto, and Fernando Vila. Ab initio theory and calculations of X-ray spectra. *Comptes Rendus Physique*, 10(6):548–559, 2009. doi: https://doi.org/10.1016/j.crhy.2008.08.004.

[34] John J Rehr, Joshua J Kas, Fernando D Vila, Micah P Prange, and Kevin Jorissen. Parameter-free calculations of X-ray spectra with FEFF9. *Physical Chemistry Chemical Physics*, 12(21):5503–5513, 2010. doi: 10.1039/B926434E.

[35] Anders Hellman, Simon Klacar, and Henrik Grönbeck. Low temperature co oxidation over supported ultrathin mgo films. *Journal of the American Chemical Society*, 131 (46):16636–16637, 2009. doi: 10.1021/ja906865f.

[36] Melike Babucci, Adisak Guntida, and Bruce C Gates. Atomically Dispersed Metals on Well-Defined Supports including Zeolites and Metal–Organic Frameworks: Structure, Bonding, Reactivity, and Catalysis. *Chemical Reviews*, 120(21):11956–11985, nov 2020. doi: 10.1021/acs.chemrev.0c00864.

[37] Adam S Hoffman, Joseph A Singh, Stacey F Bent, and Simon R Bare. In situ observation of phase changes of a silica-supported cobalt catalyst for the Fischer–Tropsch process by the development of a synchrotron-compatible in situ/operando powder X-ray diffraction cell. *Journal of Synchrotron Radiation*, 25(6):1673–1682, nov 2018. doi: 10.1107/S1600577518013942.

[38] Fernando D Vila, John J Rehr, Shelly D Kelly, and Simon R Bare. Operando Effects on the Structure and Dynamics of PtnSnm/γ-Al2O3 from Ab Initio Molecular Dynamics and X-ray Absorption Spectra. *The Journal of Physical Chemistry C*, 117 (24):12446–12457, jun 2013. doi: 10.1021/jp403931z.

[39] F Vila, J J Rehr, J Kas, R G Nuzzo, and A I Frenkel. Dynamic structure in supported Pt nanoclusters: Real-time density functional theory and x-ray spectroscopy simulations. *Physical Review B*, 78(12):121404, sep 2008. doi: 10.1103/PhysRevB.78.121404.

[40] J J Kas, J Vinson, N Trcera, D Cabaret, E L Shirley, and J J Rehr. Many-pole model of inelastic losses applied to calculations of XANES. *Journal of Physics: Conference Series*, 190(1):12009, 2009. doi: 10.1088/1742-6596/190/1/012009.

3.7 Supporting Information

Table 3.S1. Radar Plot Scoring Metrics

Structure	HERFD-XANES Fréchet	Pt oxidation state	Boltzmann fraction for each facet (%)	Fractional contribution to overall EXAFS (%)	EXAFS Reduced-χ^2	EXAFS Fréchet
$[100]^{Mg\text{-}vac}$/sub1	75	78	100	80	86	99
[100]/sub0/*O_2	83	13	100	15	78	93
[310]/pos1/*O_2	58	39	100	5	72	36

Each metric in the radar plot is determined computationally and is given a score from 0–100 to show how well the metric compares with the corresponding experimental result. The metrics are scored as: **HERFD-XANES Fréchet:** Percentile score of Fréchet distance for all FEFFsimulated XANES spectra compared with HERFD-XANES data (shown in Figure 3.S12); 0 represents the highest calculated Fréchet distance (worst match with experiment), and 100 represents the lowest Fréchet distance (best match with experiment). **Pt oxidation state:** Deviation of Pt oxidation state from +4 as determined from Bader charge analysis compared to $H_2Pt(OH)_6$ reference. **Boltzmann fraction for each facet:** DFT-calculated Boltzmann fraction relative to each facet ([100], [100]Mg-vac, and [310]) at 300 K as presented in the phase diagrams of Figure 3.3e-f. **Fractional contribution to overall EXAFS:** Contribution of the most stable structure from the 3 facets to the EXAFS data (see Figure 3.S20 for further discussion). **EXAFS Fréchet and Reduced-χ^2:** Percentile scores of Fréchet distances and Reduced-χ^2 values for EXAFS simulations.

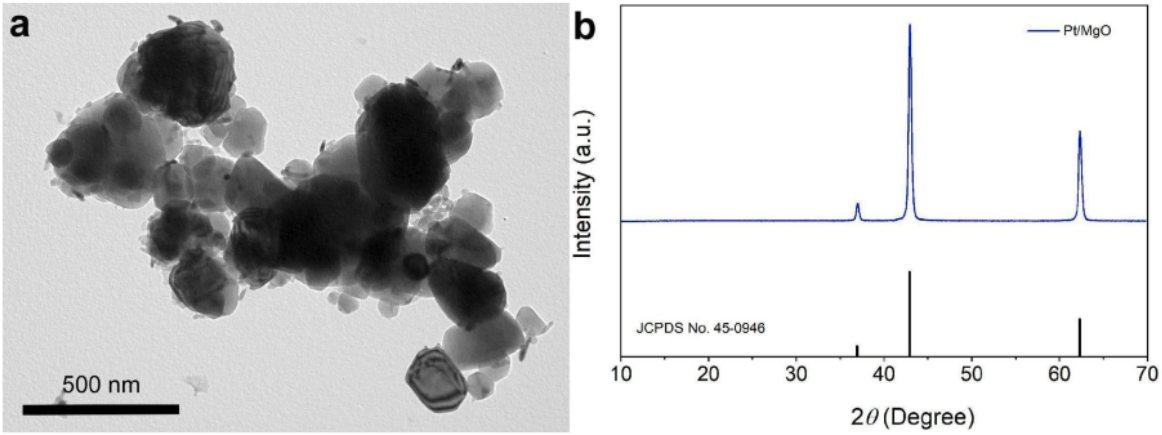

Figure 3.S1. TEM image and XRD pattern of Pt/MgO.

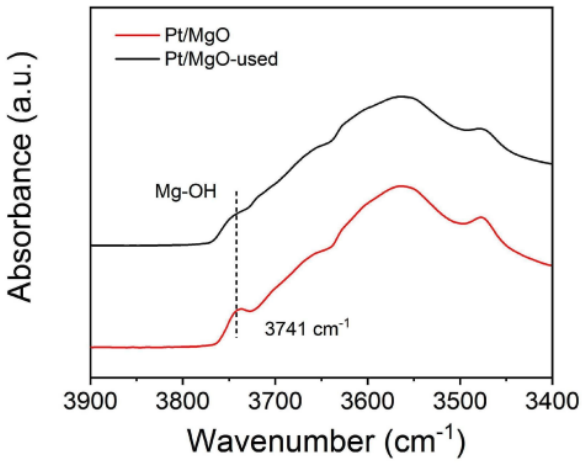

Figure 3.S2. IR spectra of ν_{O-H} region of Pt/MgO and the sample after use as a CO oxidation catalyst in three independent light-off experiments.

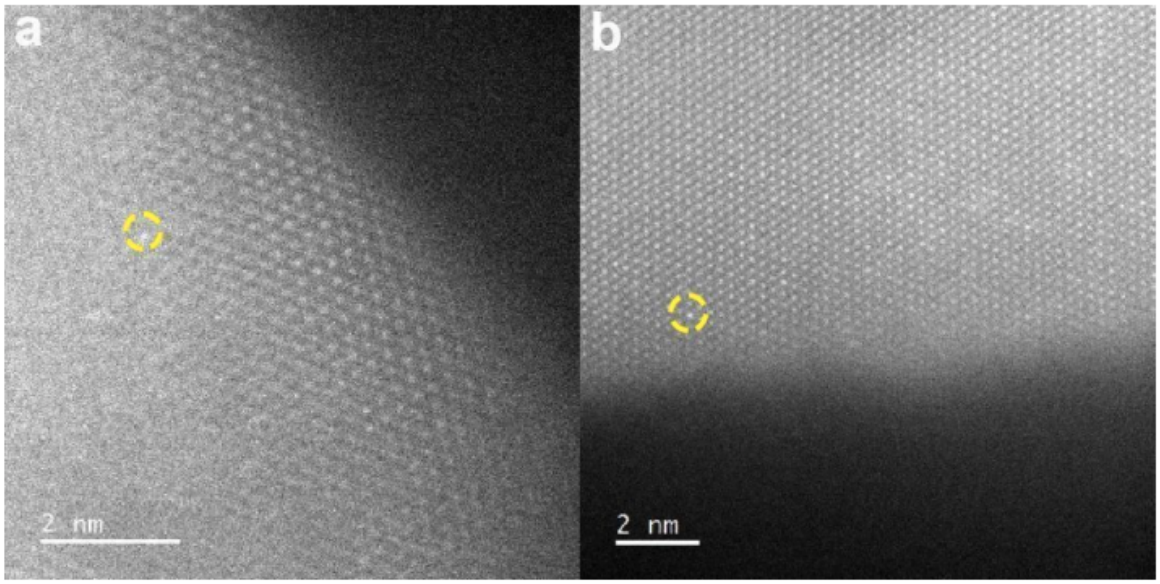

Figure 3.S3. Multiple HAADF-STEM images of Pt/MgO samples showing that the MgO supported platinum (yellow circles) was atomically dispersed.

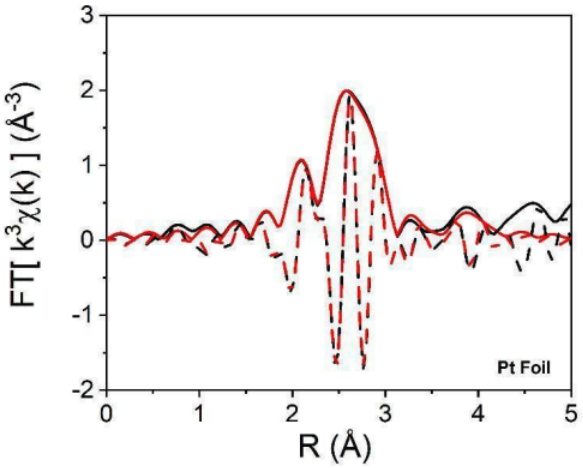

Figure 3.S4. EXAFS data at Pt L3 edge and EXAFS modelling of platinum foil. The data and the fits are shown in black and red, respectively. The magnitude and the imaginary portions of the Fourier transforms are shown as solid and dashed lines, respectively (k2-weighted).

Table 3.S2. Summary of EXAFS fit parameters for Pt foil reference.

Sample	Shell	CN[a]	R^a (Å)	$\Delta\sigma^2$ (x10^3)[a] (Å^2)	$\Delta E_0{}^a$ (eV)	$S_0{}^2$	R-factor
Pt foil	Pt–Pt$_1$	12.0	2.76 (0.00)	4.2 (0.3)	7.8 (0.6)	0.79 (0.06)	1.0 %
	Pt–Pt$_2$	6.0	3.91 (0.01)	6.4 (1.1)			

[a]*CN*, coordination number; *R*, distance between absorber and scatterer atoms; $\Delta\sigma^2$, disorder term; ΔE_0, inner potential correction. Details of data analysis for Pt foil: *k* range: 3.65–13.76 Å^{-1}; *R* range: 1–4.2 Å. Error is reported inside the parentheses (accuracies).

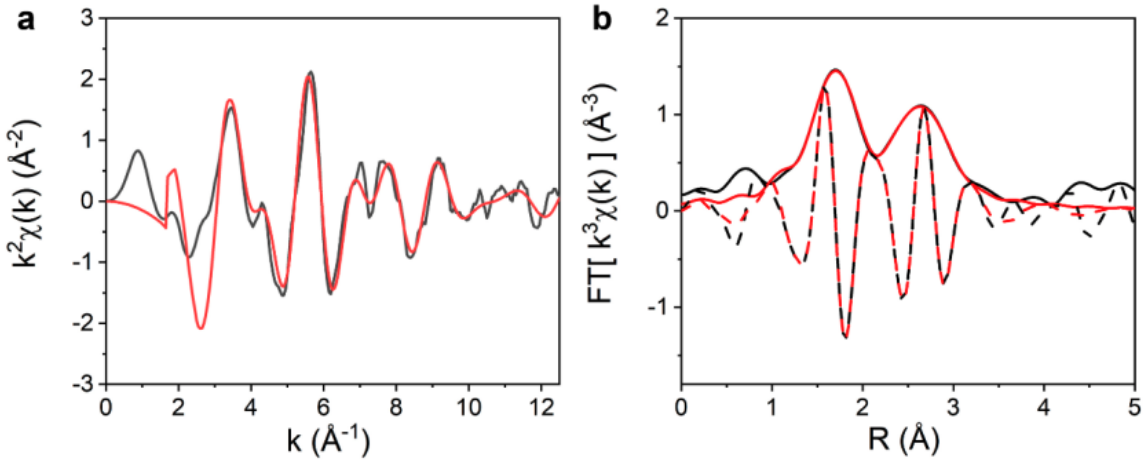

Figure 3.S5. EXAFS collected at the Pt L3 edge characterizing the Pt/MgO sample [black] and the best-fit EXAFS model (3-scattering path) representing the data [red] presented in the k2-weighted (a) EXAFS and (b) magnitude and imaginary components of the Fourier transform.

Table 3.S3. Best fit mdoel characterizing Pt L3 edge EXAFS of Pt/MgO sample.

Sample	Shell	CN[a]	R^a (Å)	$\Delta\sigma^2$ (x10^3)[a] ($Å^2$)	$\Delta E_0{}^a$ (eV)	R-factor	reduced χ^2
Pt/MgO	Pt–O	6.7 (0.8)	2.05 (0.01)	4.9 (1.5)	14.2 (1.7)	0.15%	246
	Pt–Mg	4.8 (1.0)	3.15 (0.04)	1.0 (4.0)	10.6 (2.0)		
	Pt-Mg	6.2 (2.4)	3.01 (0.04)	1.0 (4.0)	10.6 (2.0)		

[a]*CN*, coordination number; *R*, distance between absorber and scatterer atoms; $\Delta\sigma^2$, disorder term; ΔE_0, inner potential correction. Details of data analysis for Pt/MgO: *k* range: 3.80–11.68 $Å^{-1}$; *R* range: 1.0–3.2 Å. Error is reported inside the parentheses (accuracies). The $S_0{}^2$ term was determined to be 0.79 by fitting the first two single-scattering paths of fcc-Pt to the Pt foil.

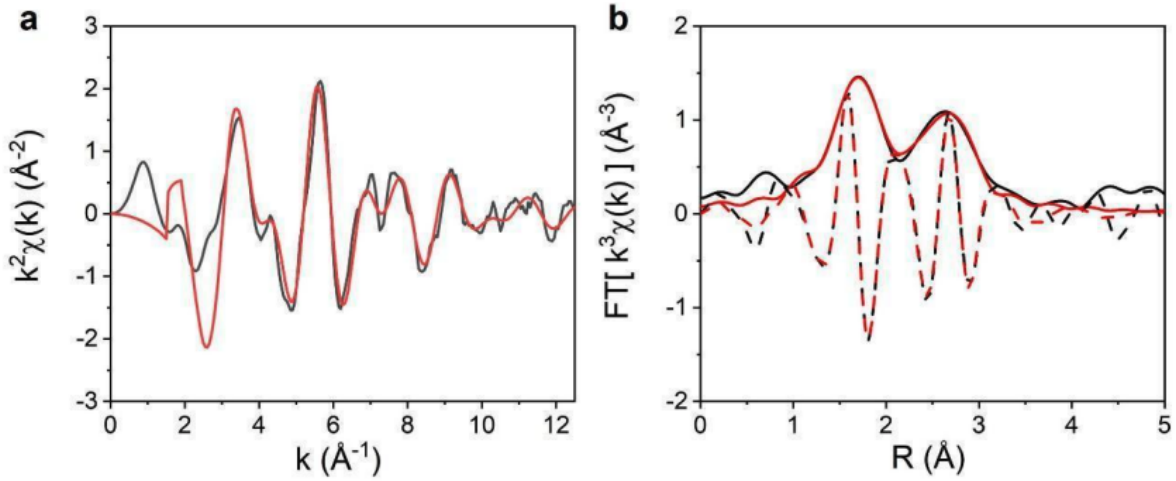

Figure 3.S6. EXAFS collected at the Pt L3 edge characterizing the Pt/MgO sample [black] and the 2-scattering path EXAFS model representing the data [red] presented in the k2-weighted (a) EXAFS and (b) magnitude and imaginary components of the Fourier transform.

Table 3.S4. 2-scattering path EXAFS model characterizing the Pt L3 edge EXAFS of the Pt/MgO sample.

Sample	Shell	CN[a]	R^a (Å)	$\Delta\sigma^2$ (x10^3)[a] (Å^2)	ΔE_0[a] (eV)	R-factor	reduced χ^2
Pt/MgO	Pt–O	7.0 (0.7)	2.05 (0.01)	5.4 (1.3)	13.8 (1.4)	0.47%	194
	Pt–Mg	12.3 (1.5)	3.05 (0.01)	8.0 (1.6)	8.9 (1.2)		

[a]*CN*, coordination number; *R*, distance between absorber and scatterer atoms; $\Delta\sigma^2$, disorder term; ΔE_0, inner potential correction. Details of data analysis for Pt/MgO: *k* range: 3.80–11.65 Å^{-1}; *R* range: 1.0–3.2 Å. Error is reported inside the parentheses (accuracies). The S_0^2 term was determined to be 0.79 by fitting the first two single-scattering paths of fcc-Pt to the Pt foil.

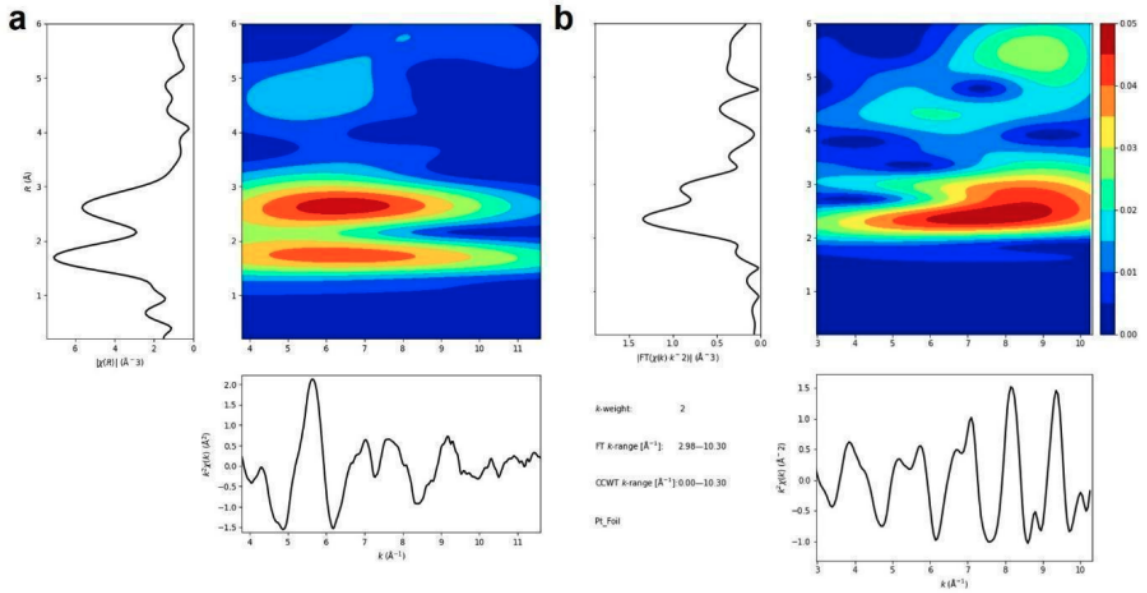

Figure 3.S7. k2-weighted CCWT of (a) Pt/MgO sample (k2-weighted EXAFS of the sample presented in the bottom plot. A k-range of 3.80-11.65 Å^{-1} was used to generate the magnitude of the Fourier transform, left plot.); (b) Pt foil (k2-weighted EXAFS of the sample presented in the bottom plot. A k-range of 2.98-10.30 Å^{-1} was used to generate the magnitude of the Fourier transform, left plot).

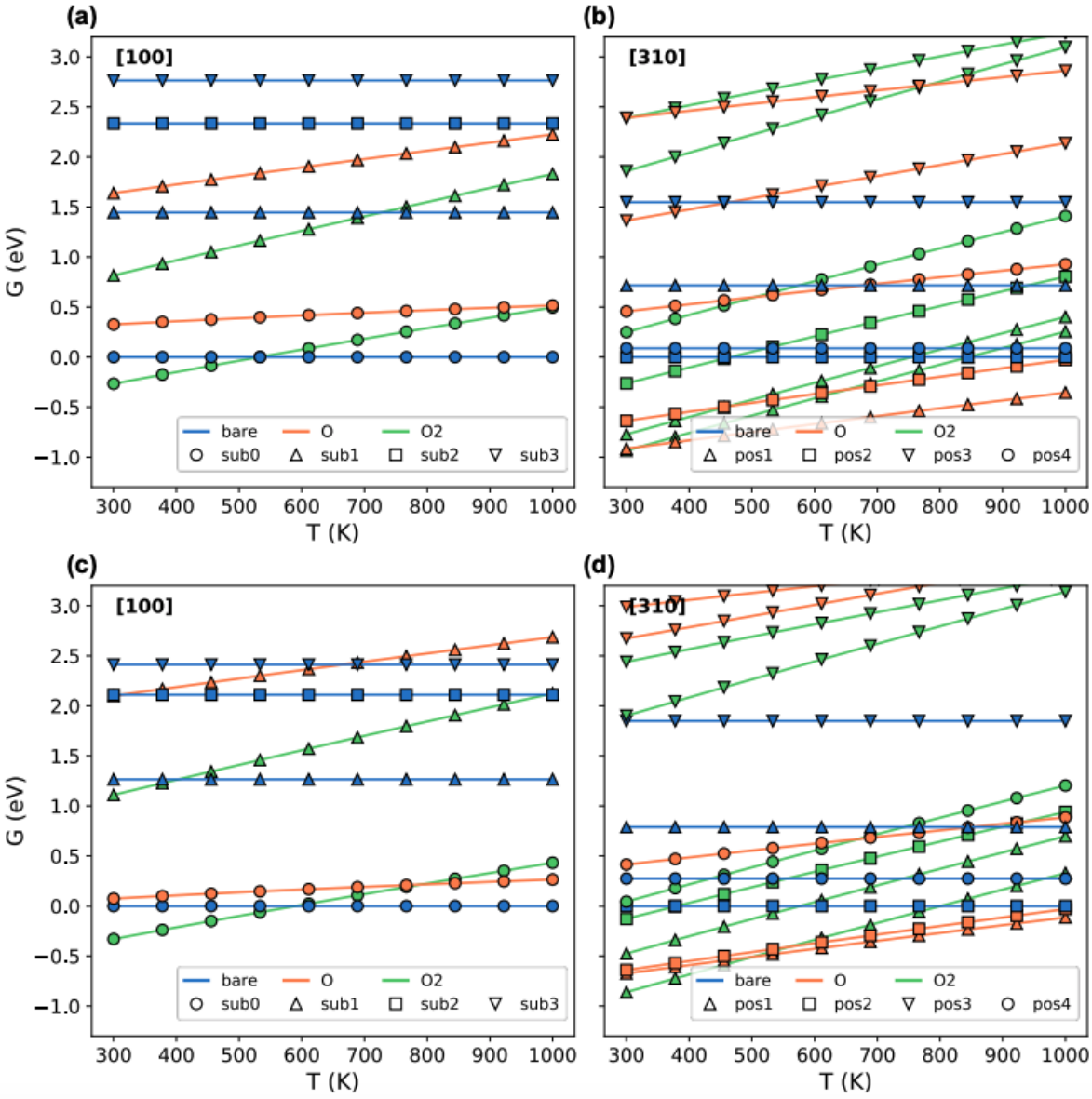

Figure 3.S8. PBE-D3BJ calculated temperature-dependent phase diagrams for (a) [100] and (b) [310] structures. RPBE-D3BJ calculated temperature-dependent phase diagrams for (c) [100] and (d) [310] structures.

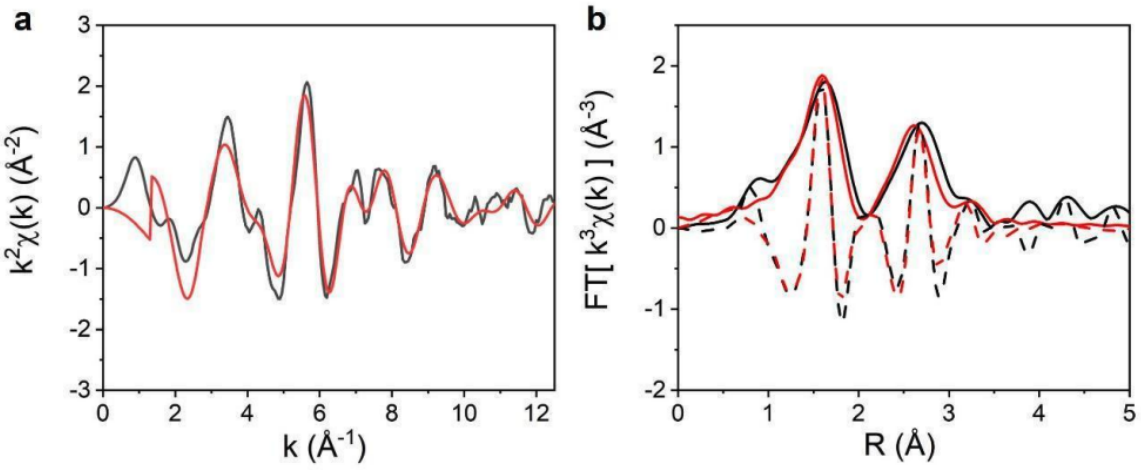

Figure 3.S9. EXAFS collected at the Pt L3 edge characterizing the Pt/MgO sample [black] and the fits performed using [100]Mg-vac DFT structure representing the data [red] presented in the k2-weighted (a) EXAFS and (b) magnitude and imaginary components of the Fourier transform.

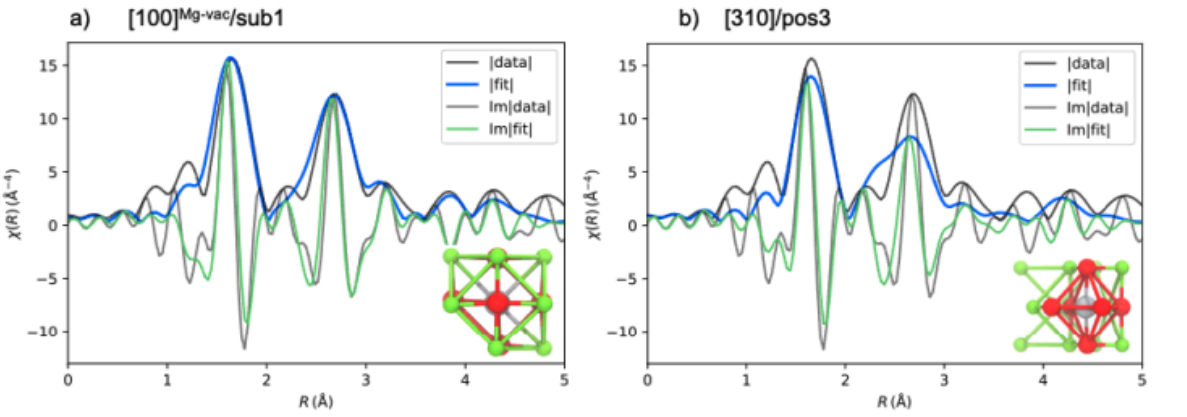

Figure 3.S10. Illustration of the capability of our fitting approach to capture the anisotropic behavior of structures with the same coordination number (Pt-O = 6 and Pt-Mg = 11). R-space EXAFS fits of (a) [100]Mg-vac/sub1 and (b) [310]/pos3 showing the magnitude and imaginary portions of experimental results in black and the magnitude and imaginary portions of the model in blue and green, respectively. The k-range of 2.2–12.5 $Å^{-1}$ and the R-range of 1.0–5.0 Å were used for the fits. Colors: Mg (green), O (red), Pt (grey).

Table 3.S5. The EXAFS fitting parameters using [100]Mg-vac/sub1 as the DFT model structure.

Sample	Shell	CN[a]	R^a (Å)	$\Delta\sigma^2$ (x10^3)[a] (Å^2)	ΔE_0[a] (eV)	R-factor	reduced χ^2
	Pt-O_1	3	2.01 (0.01)	4.3 (1.5)			
	Pt-O_2	3	2.04 (0.01)	4.3 (1.5)			
Pt/MgO	Pt-Mg_1	4	2.99 (0.02)	7.3 (2.1)	6.5 (1.3)	0.06	1824
	Pt-Mg_2	7	3.02 (0.02)	7.3 (2.1)			
	Pt-O_3	5	3.60 (0.01)	4.3 (1.5)			

[a]*CN*, coordination number; *R*, distance between absorber and scatterer atoms; $\Delta\sigma^2$, disorder term; ΔE_0, inner potential correction. Details of data analysis for Pt/MgO: *k* range: 2.20–12.50 Å^{-1}; *R* range: 1.0–3.5 Å. Error is reported inside the parentheses (accuracies). The DFT structures were used as-is with fixed coordination numbers. The S_0^2 term was determined to be 0.79 by fitting the first two single-scattering paths of fcc-Pt to the Pt foil.

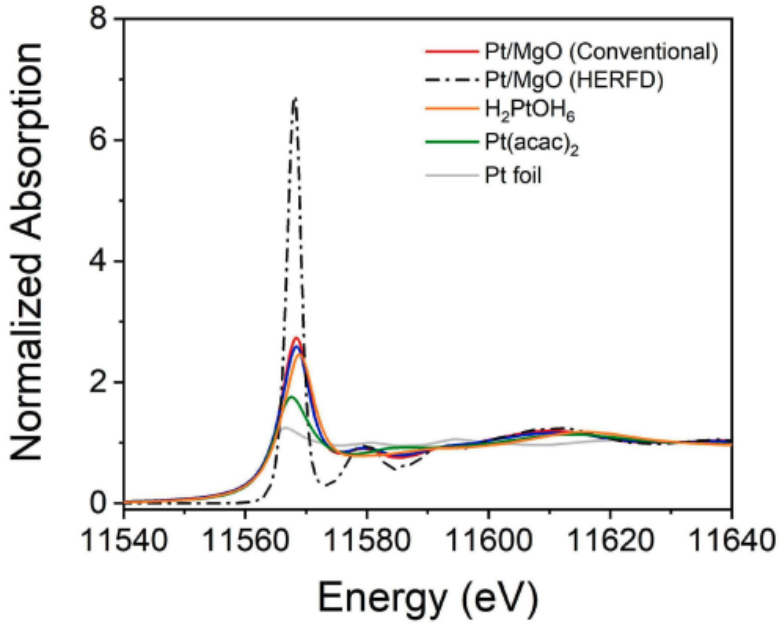

Figure 3.S11. HERFD-XANES spectrum at the Pt L3 edge of Pt/MgO (black dash dot line) and conventional XANES spectra at the Pt L3 edge of Pt/MgO (red line); $H_2Pt(OH)_6$ (orange line); $Pt(acac)_2$ (green line); and Pt foil (grey line). Data were obtained at room temperature under static conditions.

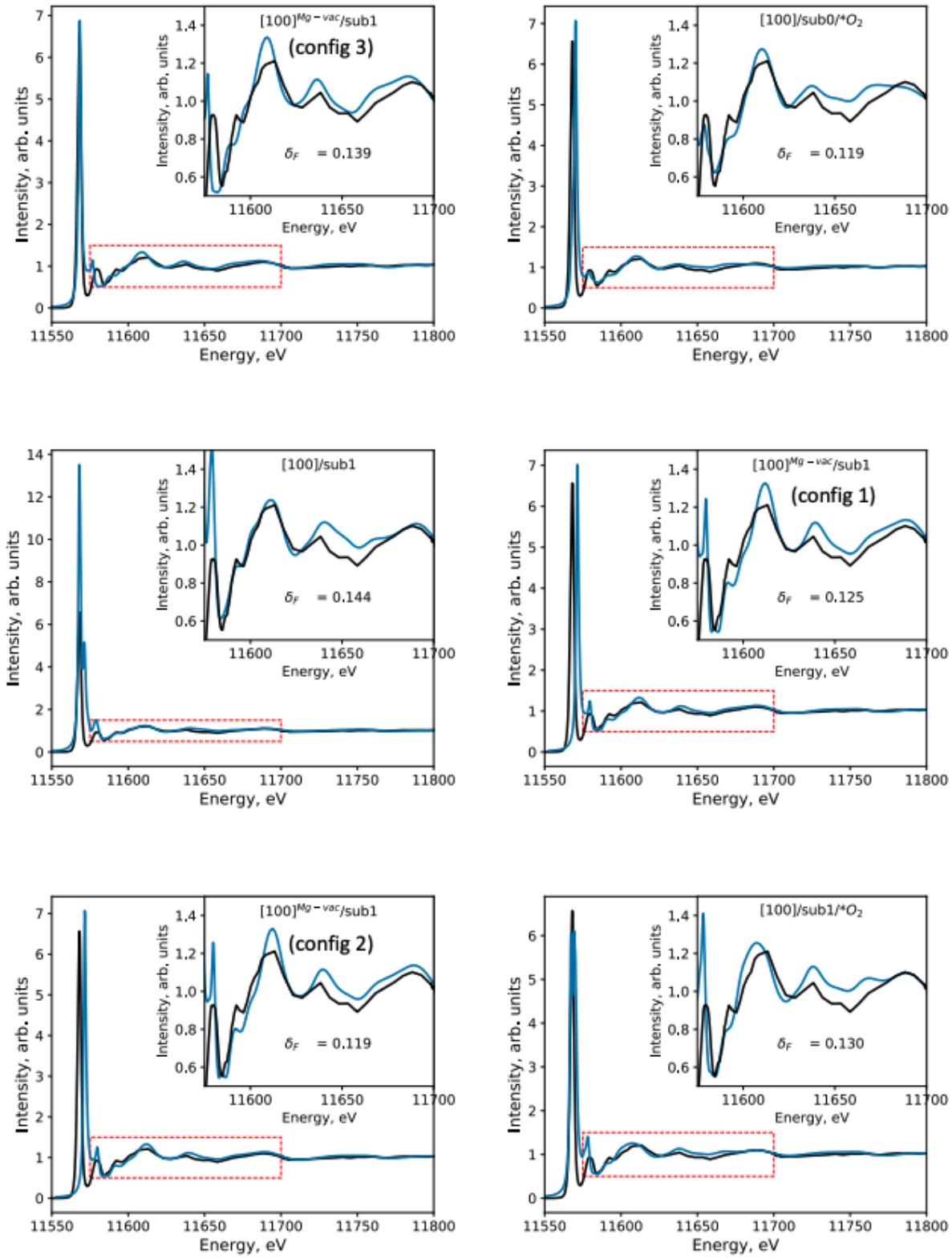

Figure 3.S12. FEFF-simulated XANES spectra (blue) and experimental HERFD-XANES results (black) for feasible DFT-calculated structures. Insets denote the region from 11575–11700 eV (dashed red box).

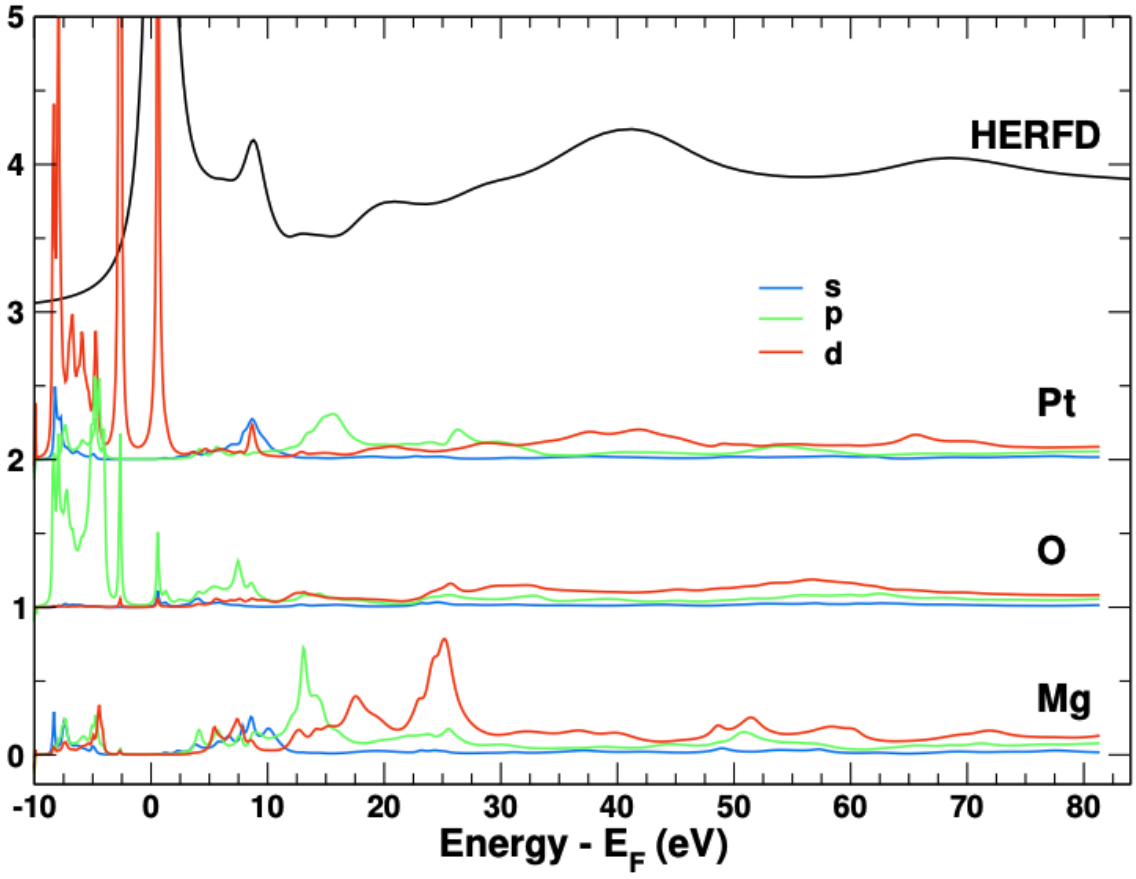

Figure 3.S13. Locally projected, l-dependent density of states for [100]Mg-vac/sub1 site with HERFD-XANES spectrum relative to the Fermi level (E_F).

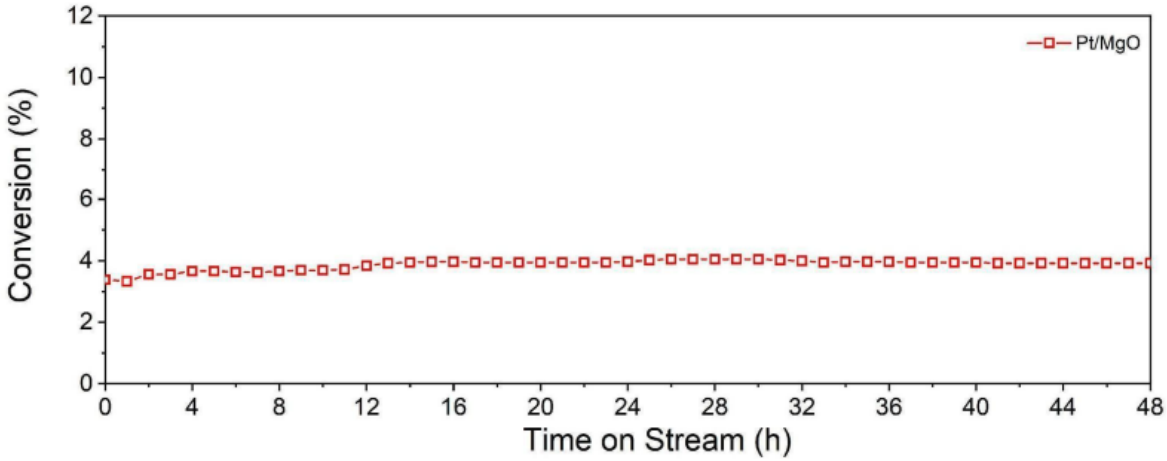

Figure 3.S14. Stability of Pt/MgO catalyst for CO oxidation in a once-through flow reactor operated at 210 °C and atmospheric pressure with a steady flow of reactant consisting of 5.0% CO in helium flowing at a rate of 4.0 mL(NTP)/min and 5.0% O_2 in helium flowing at a rate of 16.0 mL(NTP)/min.

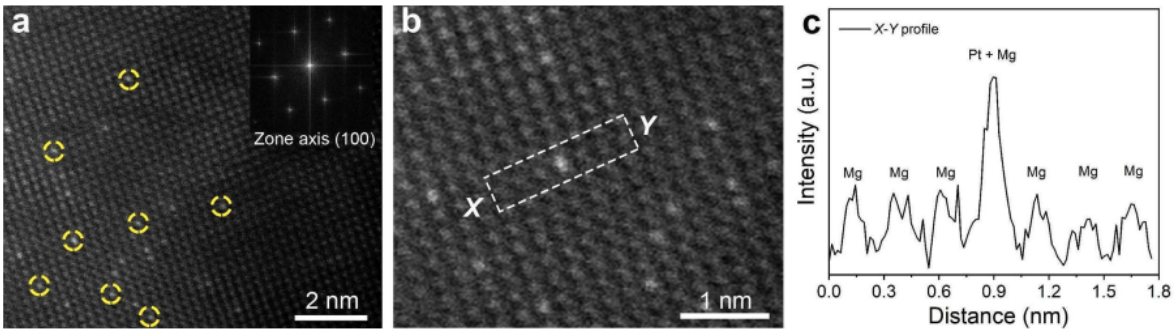

Figure 3.S15. (a) HAADF-STEM image of Pt/MgO-used sample (Pt/MgO after it had been used as a catalyst in three independent light-off CO oxidation measurements) along the (100) zone axis. Pt atoms are marked in yellow circles; the inset shows the corresponding FFT pattern. (b) Magnified HAADF-STEM image of a. (c) Corresponding intensity profile from the X–Y line scan in b.

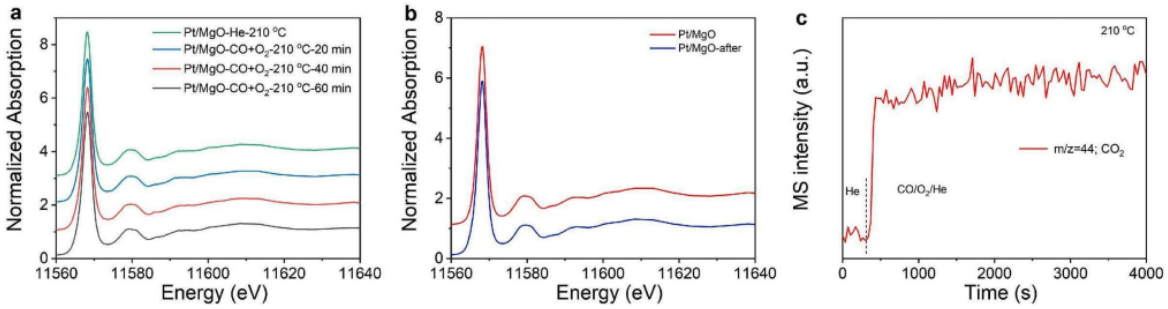

Figure 3.S16. In-operando HERFD-XANES spectra at the Pt L3 edge of Pt/MgO and mass spectrometry results. (a) HERFD-XANES data collected at 210 °C with the sample (approximately 50 mg) in helium flowing at 20 mL(NTP)/min and then after replacement of the helium stream with a mixture of CO flowing at 0.2 mL(NTP)/min + O_2 flowing at 0.8 mL(NTP)/min + He flowing at 19 mL(NTP)/min for 1 h. (b) HERFD-XANES spectra collected with sample in helium flowing at 20 mL(NTP)/min at room temperature; the data characterize Pt/MgO before and after it had been used as a catalyst for CO oxidation at 210 °C for 1 h. (c) Mass spectrometry results characterizing effluent gas recorded during in-operando HERFDXANES experiments (shown in a) characterizing Pt/MgO (mass 44 represents CO_2).

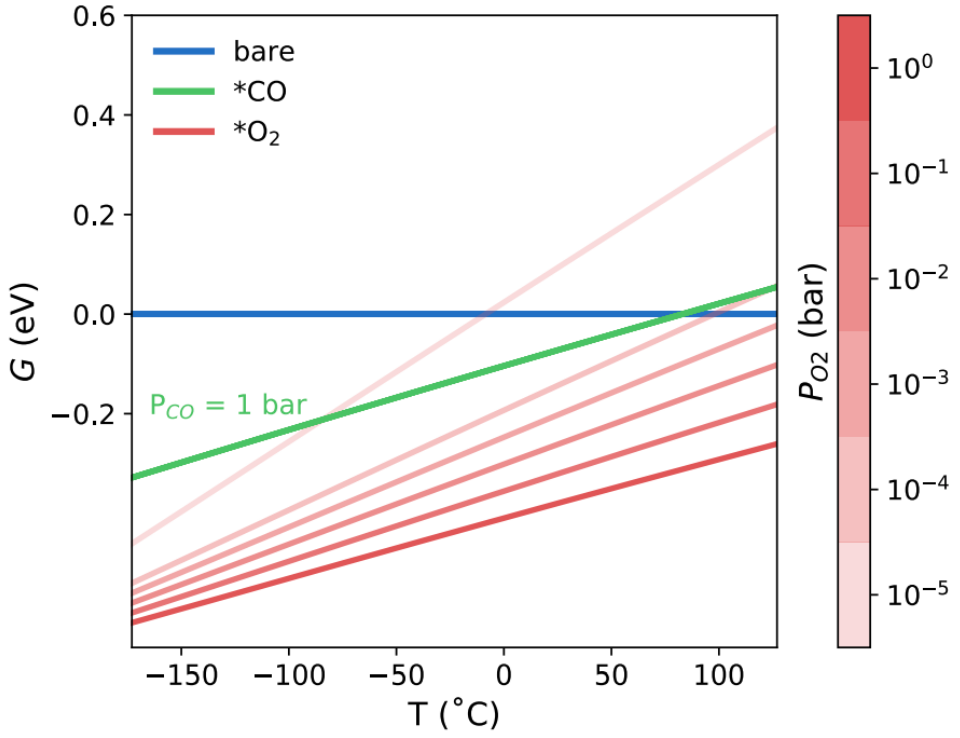

Figure 3.S17. PBESol-calculated temperature-dependent CO and O_2 adsorption on [100]/sub0 for O_2 partial pressures from 10^{-5} to 100 bar. CO adsorption is predicted until 83 °C is reached for low O_2 partial pressures.

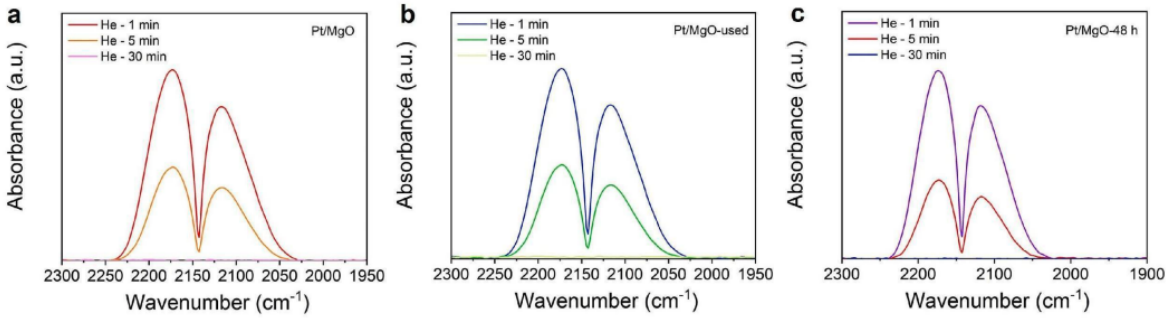

Figure 3.S18. IR spectra of (a) Pt/MgO; (b) Pt/MgO-used (Pt/MgO after it had been used as a catalyst in three independent light-off CO oxidation measurements); and (c) Pt/MgO-used (Pt/MgO after it had been used as a catalyst in three independent light-off CO oxidation measurements) in flowing helium recorded at 1, 5, and 30 min after the beginning of flow of 10% CO in helium at room temperature and 1 atm. The bands at 2174 and 2118 cm−1 represent the gas-phase CO.

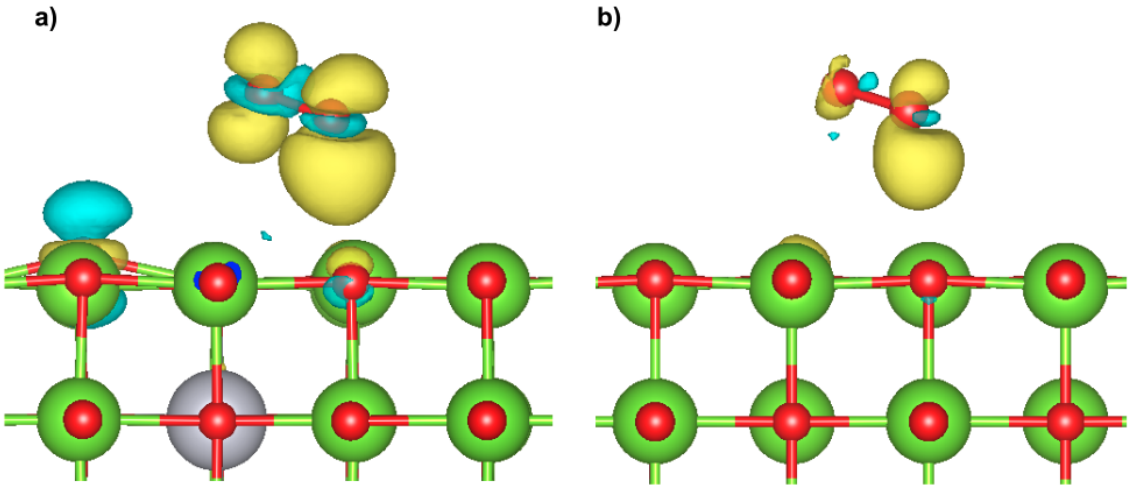

Figure 3.S19. Differences in charge density ($\pm$0.001 electron/Å^3) upon binding of O_2 on a) [100]Mg-vac/sub1 and b) pure ideally crystalline MgO. Charge gain and depletion are represented with yellow and blue, respectively. The depletion of charge between the O—O bond in Pt/ MgO enables CO insertion to form bound $*CO_3$ during CO oxidation catalysis.

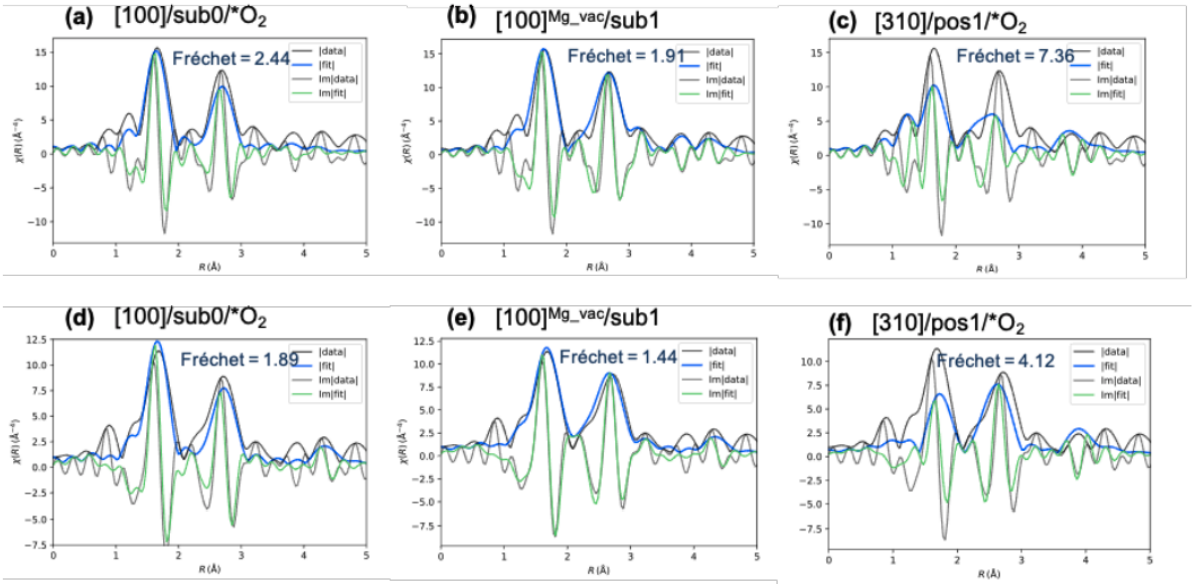

Figure 3.S20. EXAFS fits of Pt/MgO as is (a–c) (i.e., as synthesized before CO oxidation catalysis) and used (d–e) catalyst (i.e., after it had been used as a catalyst in three independent light-off CO oxidation measurements) for the three most stable structures [100]/sub0/$*O_2$, [100]Mg-vac/sub1, and [310]/pos1/$*O_2$ showing the magnitude and imaginary portions of Fourier transforms. The magnitude is shown with experimental results in black and the model in blue, and the imaginary portions are shown in black and green, respectively. The k-range of 2.2–12.5 Å^{-1} and the R-range of 1.0–5.0 Å were used for the fits.

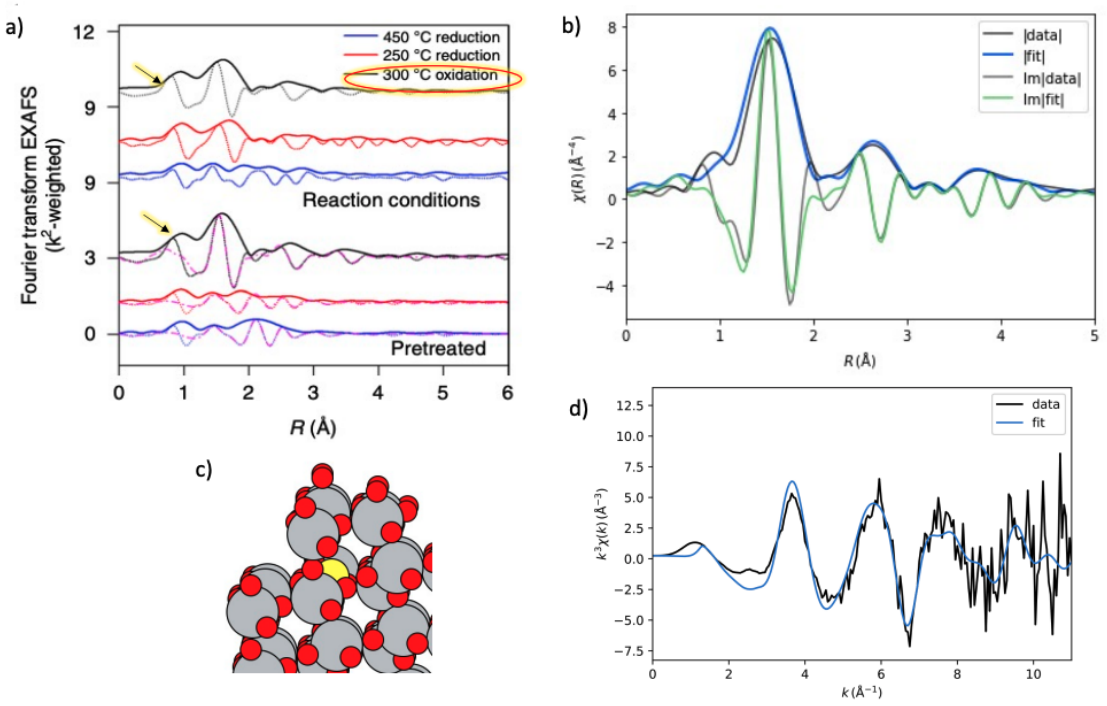

Figure 3.S21. Testing the applicability of QuantEXAFS on well-defined system (Pt/TiO_2) - a) EXAFS fits reported in the paper derived from conventional approach; EXAFS fits using QuantEXAFS on the same data (300 °C oxidation); fit from QuantEXAFS in b) R-space and d) k-space; c) DFT optimized structure used for fitting the data in QuantEXAFS ('hypothesized' to be representative of the Pt site in the original manuscript). Red: O, gray: Ti, and yellow: Pt. The k-range of 2.0–10.5 Å^{-1} and the R-range of 1.0–5.0 Å were used for the fits. Adapted with permission from DeRita et al., Nature Materials volume 18, pages 746–751. Copyright 2019 Springer Nature.

Chapter 4

Predicting Structural Properties of Pure Silica Zeolites Using Deep Neural Network Potentials

Tyler G. Sours and Ambarish R. Kulkarni

Adapted with permission from J. Phys. Chem. C, 2023, 127, 1455–1463.

4.1 Abstract

Machine learning potentials (MLPs) capable of accurately describing complex *ab initio* potential energy surfaces (PES) have revolutionized the field of multiscale atomistic modeling. In this work, using an extensive density functional theory (DFT) dataset (denoted as Si-ZEO22) consisting of 219 unique zeolite topologies (350,000 unique DFT calculations) found in the International Zeolite Association (IZA) database, we have trained a DeePMD-kit MLP to model the dynamics of silica frameworks. The performance of our model is evaluated by calculating various properties that probe the accuracy of the energy and force predictions. This MLP demonstrates impressive agreement with DFT for predicting zeolite structural properties, energy-volume trends, and phonon density of states. Furthermore, our model achieves reasonable predictions for stress-strain relationships without including DFT stress data during training. These results highlight the ability of MLPs to capture the flexibility of zeolite frameworks and motivates further MLP development for nanoporous materials with near-*ab initio* accuracy.

4.2 Introduction

Accurate and efficient calculation of the energies and forces of atomistic systems remains one of the leading challenges in computational chemistry. While *ab initio* approaches rooted in quantum mechanics, e.g., Density Functional Theory (DFT), often yield reliable results, large scale simulation of system dynamics with DFT remains impractical. For instance, predicting self diffusivity coefficients, phase transitions, and phonon spectra using molecular dynamics (MD) often requires millions of force and energy evaluations. Traditionally, generic or DFT-parameterized force fields (FFs) are used for such computationally demanding simulations. While the simplicity of the FF methods enables longer simulation timescales for larger systems, these approaches are often less accurate than *ab initio* simulations. Even for FFs derived from DFT calculations, the rigid analytical form of bonded (e.g., harmonic, Morse etc.) and non-bonded (e.g., 12-6 Lennard-Jones, Buckingham etc.) potentials often results in systematic deviations.[1]

In contrast to the simple analytical form of classical FFs, machine learning potentials (MLPs) have emerged as a flexible alternative to describe complex potential energy surfaces. Specifically, by training the model on a suitable set of first principles data that spans the relevant configuration space of a system, an MLP is able to evaluate the PES at accuracy close to the *ab initio* method at significantly lower computational cost. Several different MLP forms have been proposed, which are broadly classified as either kernel methods or neural network methods. Kernel methods, such as GAP[2] and sGDML[3], employ kernel functions (e.g., SOAP[4]) to assess the similarity of atomic configurations and interpolate the energy from known data points. Neural network methods calculate single atomic energy contributions by using a set of symmetry invariant descriptors that capture the local environment of each atom as inputs to various neural network architectures. Popular neural network potentials include ANI[5] and DeePMD[6,7], and newer message-passing networks like PhysNet[8], SchNet[9], and SpookyNet[10]. New MLPs continue to appear in the literature, and several reviews exist describing and comparing the current state of the art models.[11–15]

Open-source releases of MLP software have enabled researchers to develop their own

force fields for various systems including small molecules, nanoparticles, and metal surfaces. However, to the best of our knowledge, similar approach have not been used for zeolites. Siliceous zeolites are polymorphs composed of the SiO_2 formula unit with significant industrial use.[16,17] Given the chemical simplicity and the existence of over 200 unique topologies, and hundreds of thousands of theoretical structures,[18] zeolites are ideally suited for demonstrating the capabilities of MLPs.

Many industrially relevant applications of zeolites involve small molecules diffusing through the porous framework over relatively long time-scales. As including framework flexibility is necessary to accurately model diffusion and adsorption phenomena in zeolites,[19,20] it is important to develop MLPs that accurately model dynamics of the framework. Thus, the central goals of this work are to develop a DFT dataset that rigourously samples the atomic configuration space of pure silica zeolites, and train and validate an MLP using the Deep Potential (DP) method implemented in DeePMD-kit.[7]

The DP method represents the system energy as the sum of single atomic energies that are determined from descriptors that capture the localized interactions between each atom and its neighbors within a specified cutoff distance. For a given atom, the relative coordinates of the local environment (i.e., the neighboring atoms) are passed through an encoding network to obtain symmetry invariant descriptors. These descriptors are then mapped to single atomic energies via an additional fitting neural network. This approach has shown promising results for describing the dynamics of both small molecules[21–25] and periodic bulk materials.[26–31] Additionally, DeePMD-kit provides seamless integration with the Atomic Simulation Environment (ASE)[32], the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS)[33], and several other popular molecular simulation platforms.

In this work, a large DFT dataset is generated using 219 of the 248 siliceous zeolite topologies found in the IZA database; all topologies with fully connected frameworks and less than 400 atoms were included. A single DP model, trained on 187 of these topologies to obtain a generalized silica MLP, is shown to accurately predict energies and forces of DFT configurations not included in the training set. This analysis is extended by using our

DP model to calculate properties not explicitly included in the training, and the results are compared with the DFT predictions. Our results show excellent agreement between DP and DFT for structural properties, equations of state (EOS), and phonon density of states (PDOS). We also demonstrate the ability of DP to model stress-strain behavior and give reasonable predictions of mechanical properties even when *ab initio* stress data are not used during training. While other ML models have been developed to predict some of these properties purely from zeolite geometric descriptors,[34–36] we test how well a DFT-trained MLP can directly calculate these properties by evaluating the PES. Our results are also compared with those from the BKS force field[37] (used as a prototypical example of a classical force field), and we find that the DP model provides significantly more accurate results. We end our analysis by calculating the above properties for an additional set of 32 topologies (not included in the training) to demonstrate the transferability of the model. Taken together, by highlighting the efficacy of the DeePMD-kit formulation for silica zeolites, this study lays a foundation for future exploration of more complex materials such as those containing extra-framework cations and adsorbates.

4.3 Computational Methods

4.3.1 Training Set Generation

DFT NVT-MD was used to generate the initial training set for the DP model. The Vienna *ab initio* simulation package (VASP) was used with the PBE[38] functional for DFT calculations. Dispersion corrections were considered with the DFT-D3 method with Becke-Johnson damping (D3BJ).[39–41] Only the Γ-point was used for k-space sampling. A plane-wave cutoff of 400 eV was was used, and electronic energies were converged to 10^{-5} eV. Configurations were obtained from MD trajectories ($\geq$ 1.5 ps simulation time, 0.5 fs timestep) at three different temperatures: 300, 600, and 900 K. Snapshots from these trajectories were taken every 10 timesteps and were used to train an initial DP model. This model was then used with LAMMPS to generate 100 ps NPT-MD trajectories at 0.1, 1.0, and 10.0 bar. After equilibrating the system (298 K, 10 ps), the temperature was ramped from 298 K to 1000 K over the course of the simulation. This approach provides

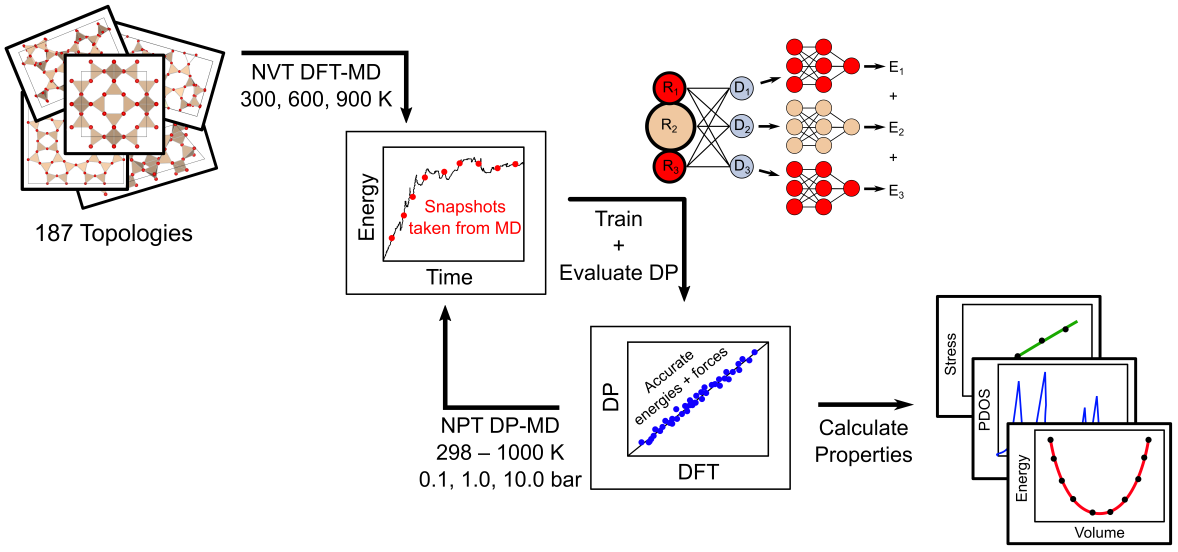

Figure 4.1. Schematic overview of procedure used to train DP model. Initial model was trained on configurations from 3 ps NVT DFT-MD runs at 300, 600, and 900 K. The initial DP was then used to generate 100 ps NPT DP-MD trajectories at pressures of 0.1, 1.0, and 10.0 bar with the temperature linearly ramped from 298-1000 K. Snapshots from every 1000 time steps were selected to obtain new uncorrelated configurations for training final DP model that is used to predict various structural properties of silica zeolites.

a diverse set of configurations at various temperatures and pressures. Snapshots of each system at every 1000 timesteps were extracted for a total of 600 configurations (200 × 3 pressures). The energies and forces of the new configurations were evaluated with DFT, and the model was retrained including these results. This procedure is illustrated in Figure 4.1. Due to computational cost, fewer configurations were collected for large unit cell topologies; a full list of the topologies and corresponding dataset sizes are included in the SI. While not used in this work, we note that the DPGEN[42,43] training protocol can be used to select snapshots for training.

4.3.2 Model Training

The DP training parameters were adjusted to achieve an architecture that balances the accuracy of the energy and force predictions as well as the time to train and evaluate new configurations. The primary parameters we considered for tuning were the size of the embedding neural net, size of the fitting neural net, atomic cutoff radius, initial learning rate (LR), and the number of training steps (Tables 4.1-4.5 and Figures 4.2-

4.6). Other "smaller" parameters (e.g., batch size and decay rate) were found to have a negligible effect on prediction accuracy and thus have been omitted. It was discovered that DeePMD-kit provides a very robust platform that produces reasonable results for most architectures chosen; however, it is still advantageous to tune the parameters to find the least complex architecture necessary to achieve the desired accuracy in the interest of reducing computational expense.

Table 4.1. Hyperparameter configurations for tuning the embedding neural net size.

Model	**Embedding Net**	Fitting Net	Cutoff	Initial LR	Training Steps
A	**(8, 16, 32)**	(64, 64, 64)	6.0 Å	0.0005	1×10^6
B	**(16, 32, 64)**	(64, 64, 64)	6.0 Å	0.0005	1×10^6
C	**(32, 64, 128)**	(64, 64, 64)	6.0 Å	0.0005	1×10^6
D	**(32, 64, 128)**	**(128, 128, 128)**	6.0 Å	0.0005	1×10^6

Table 4.2. Hyperparameter configurations for tuning the fitting neural net size.

Model	Embedding Net	**Fitting Net**	Cutoff	Initial LR	Training Steps
A	(16, 32, 64)	**(32, 32, 32)**	6.0 Å	0.0005	1×10^6
B	(16, 32, 64)	**(64, 64, 64)**	6.0 Å	0.0005	1×10^6
C	(16, 32, 64)	**(128, 128, 128)**	6.0 Å	0.0005	1×10^6

Table 4.3. Hyperparameter configurations for tuning the atomic cutoff.

Model	Embedding Net	Fitting Net	**Cutoff**	Initial LR	Training Steps
A	(16, 32, 64)	(64, 64, 64)	**5.0 Å**	0.0005	1×10^6
B	(16, 32, 64)	(64, 64, 64)	**6.0 Å**	0.0005	1×10^6
C	(16, 32, 64)	(64, 64, 64)	**7.0 Å**	0.0005	1×10^6

The configurations collected from each run were first shuffled and then split into 80% training data, 10% validation data used by DeePMD-kit during the training process, and an additional 10% testing data. The cutoff radius is 6.0 Å with smoothing beginning at 5.5 Å. The embedding net was set to 3 layers with (16, 32, 64) neurons, respectively. The fitting net was also set to 3 layers with (64, 64, 64) neurons. The model was trained for

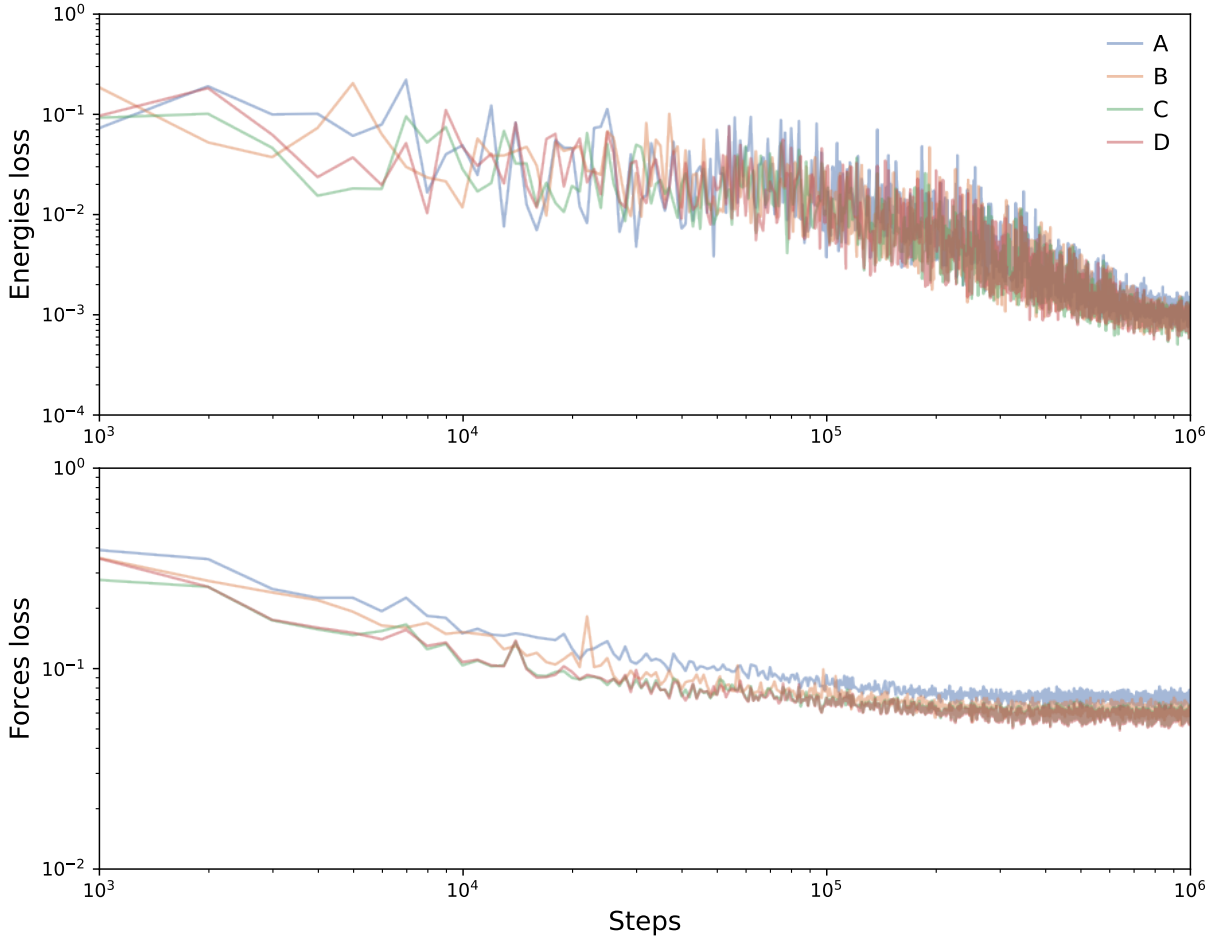

Figure 4.2. Validation set learning curves for the model hyperparameter combinations shown in Table 4.1 for the (top) energy and (bottom) force loss function contributions.

Table 4.4. Hyperparameter configurations for tuning the initial learning rate.

Model	Embedding Net	Fitting Net	Cutoff	**Initial LR**	Training Steps
A	(16, 32, 64)	(64, 64, 64)	6.0 Å	**0.001**	1×10^6
B	(16, 32, 64)	(64, 64, 64)	6.0 Å	**0.0005**	1×10^6
C	(16, 32, 64)	(64, 64, 64)	6.0 Å	**0.0001**	1×10^6

2×10^7 steps with the learning rate starting at 5×10^{-4} and exponentially decaying to 5×10^{-8}. The prefactors for the energy and force contributions to the loss function were set to $p_e^{start} = 0.02$, $p_e^{limit} = 1$, $p_f^{start} = 1000$, $p_f^{limit} = 1$. Validation set learning curves for several model architectures are shown in Figures 4.2-4.6. The hyperparameters selected were found to provide a reasonable balance between accuracy and training/evaluation

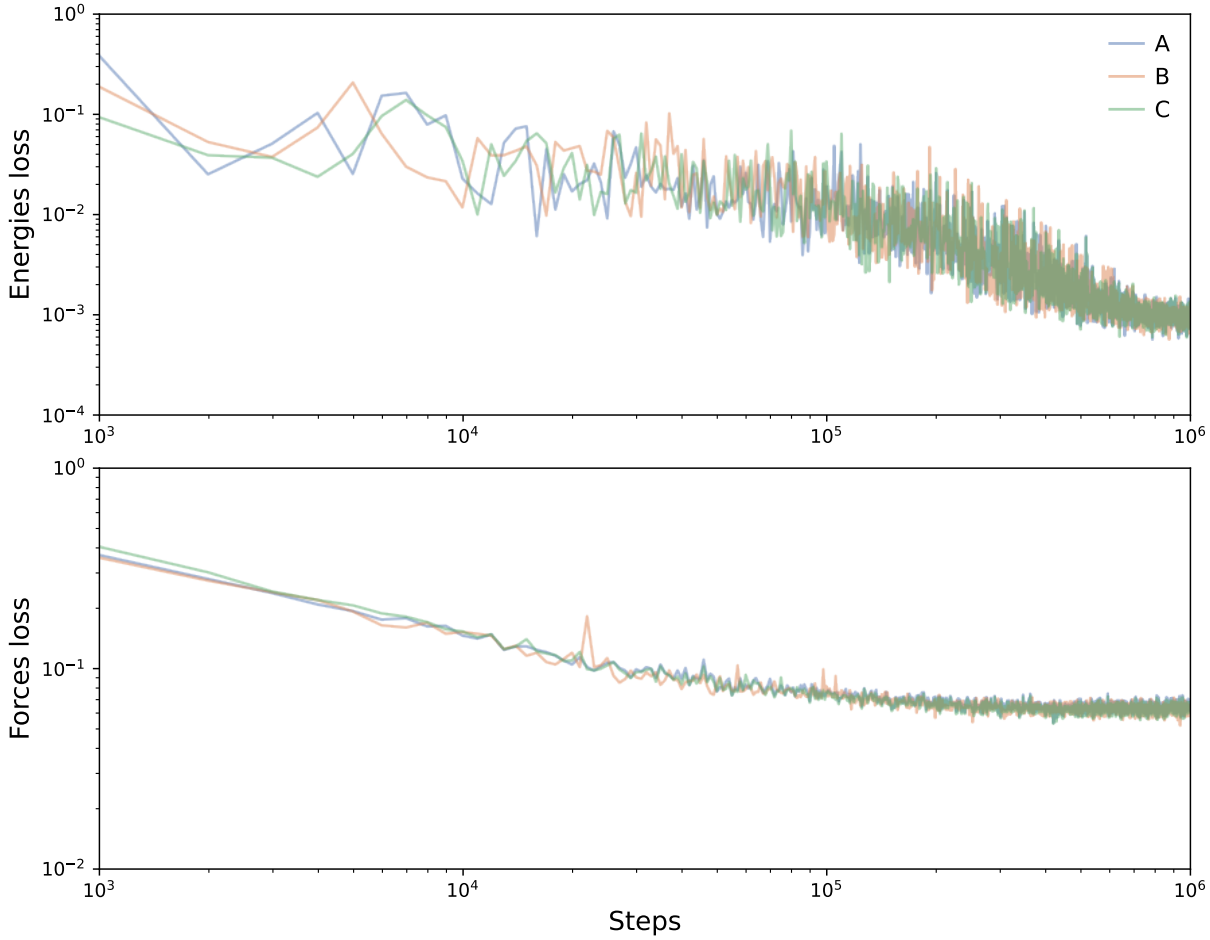

Figure 4.3. Validation set learning curves for the model hyperparameter combinations shown in Table 4.2 for the (top) energy and (bottom) force loss function contributions.

Table 4.5. Model hyperparameter configurations for tuning the number of training steps

Model	Embedding Net	Fitting Net	Cutoff	Initial LR	**Training Steps**
A	(16, 32, 64)	(64, 64, 64)	6.0 Å	0.0005	5×10^5
B	(16, 32, 64)	(64, 64, 64)	6.0 Å	0.0005	1×10^6
C	(16, 32, 64)	(64, 64, 64)	6.0 Å	0.0005	5×10^6
D	(16, 32, 64)	(64, 64, 64)	6.0 Å	0.0005	2×10^7

time (Figure 4.7). The complete input file of all parameters used for training the final model is hosted at https://github.com/tysours/Si-ZEO22/blob/main/supporting_info/SiZeo_dpmd_params.json.

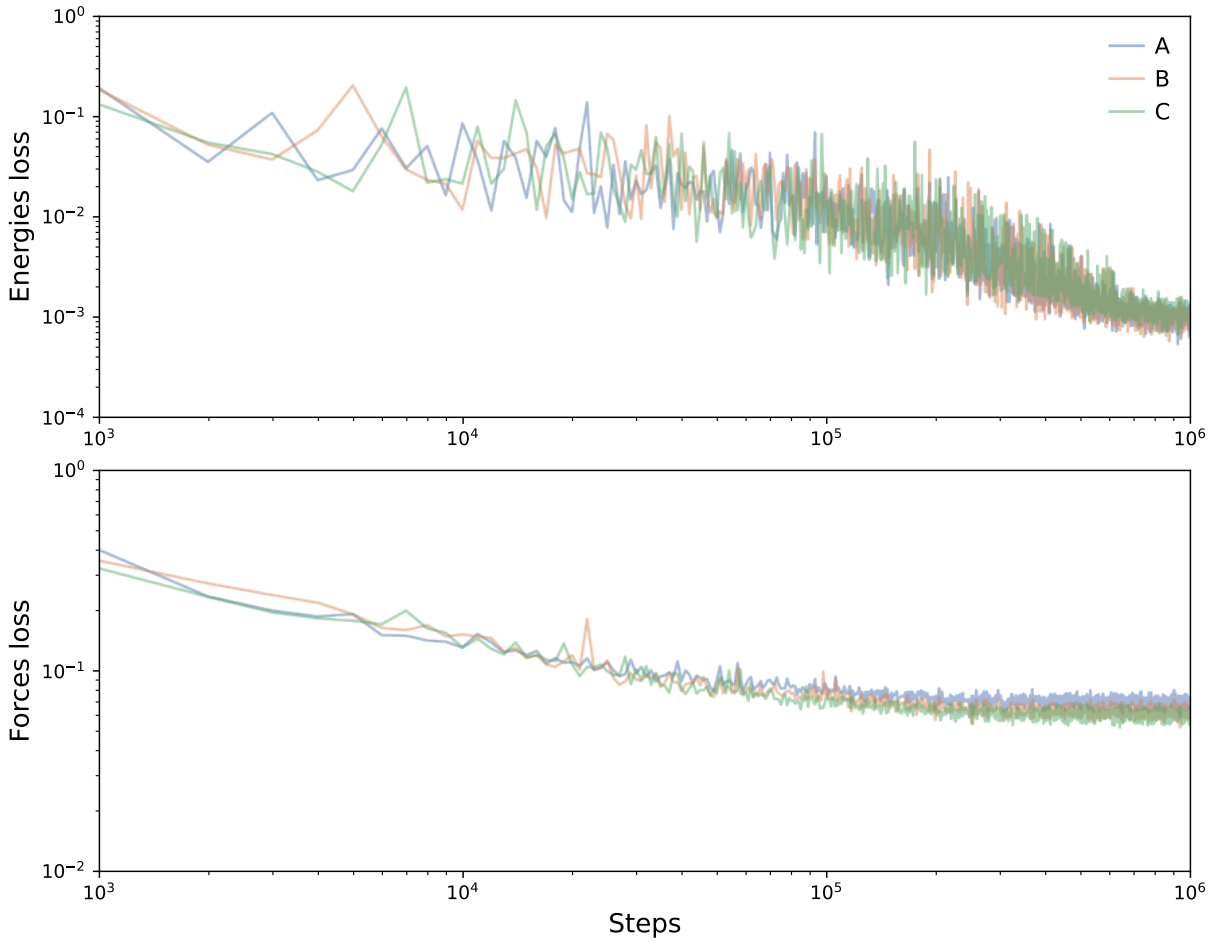

Figure 4.4. Validation set learning curves for the model hyperparameter combinations shown in Table 4.3 for the (top) energy and (bottom) force loss function contributions.

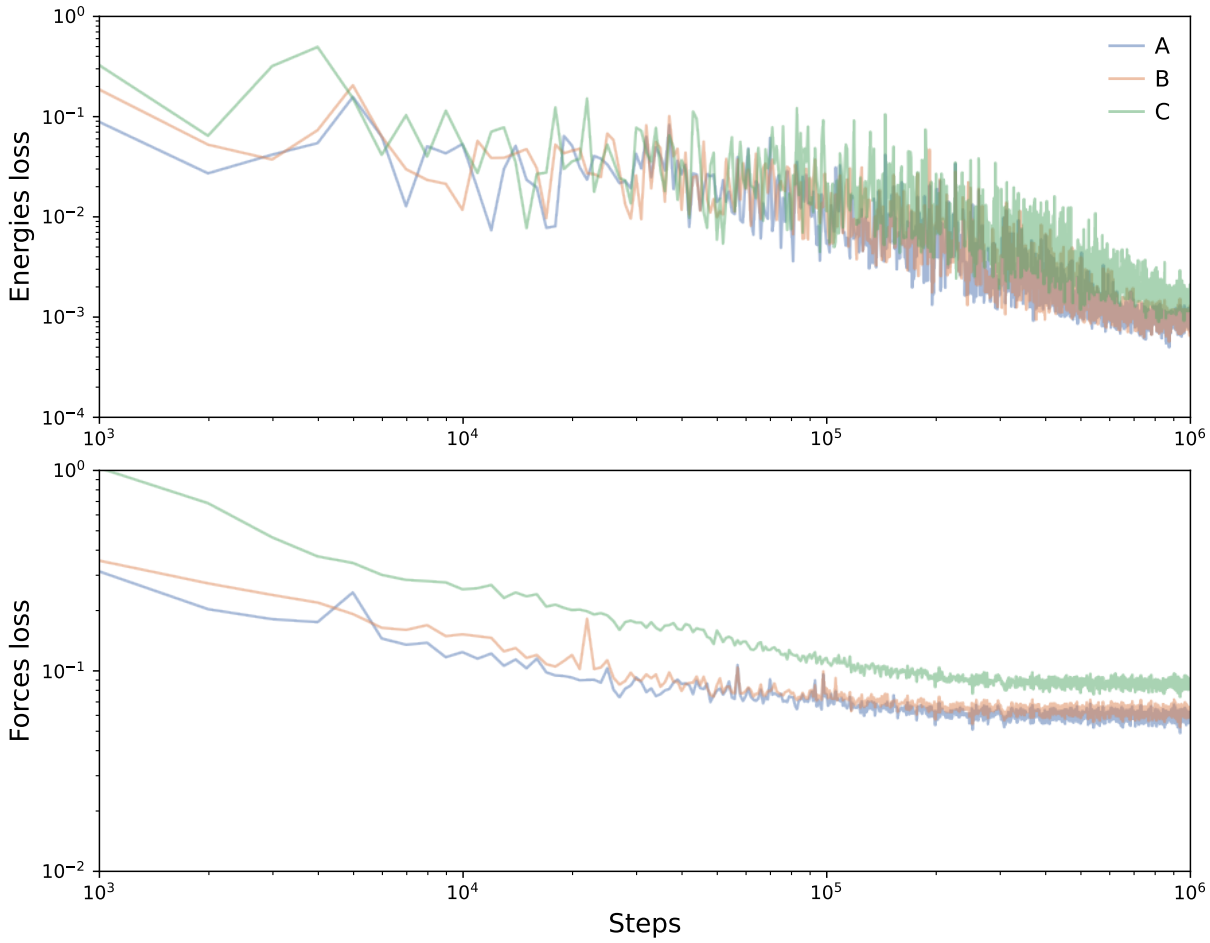

Figure 4.5. Validation set learning curves for the model hyperparameter combinations shown in Table 4.4 for the (top) energy and (bottom) force loss function contributions.

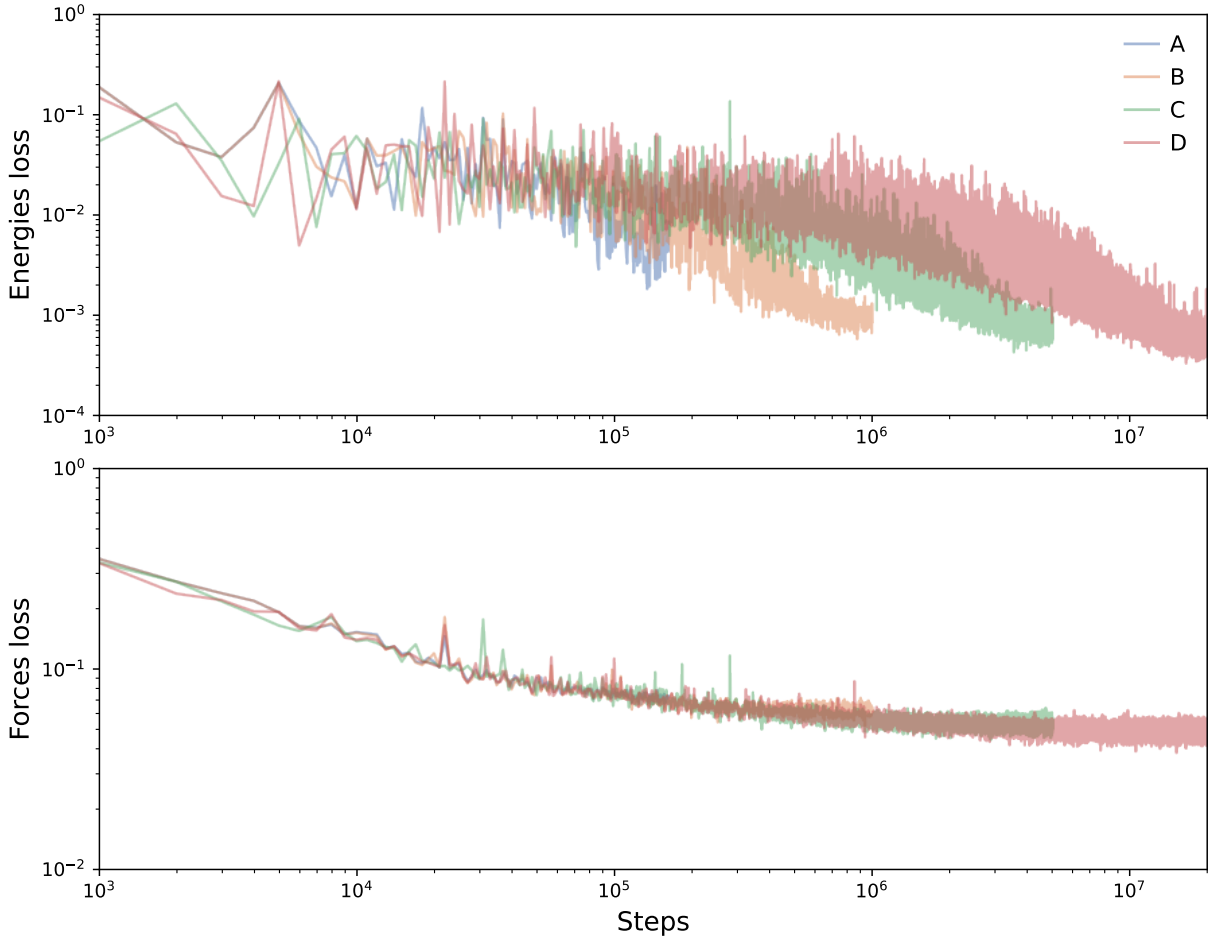

Figure 4.6. Validation set learning curves for the model hyperparameter combinations shown in Table 4.5 for the (top) energy and (bottom) force loss function contributions.

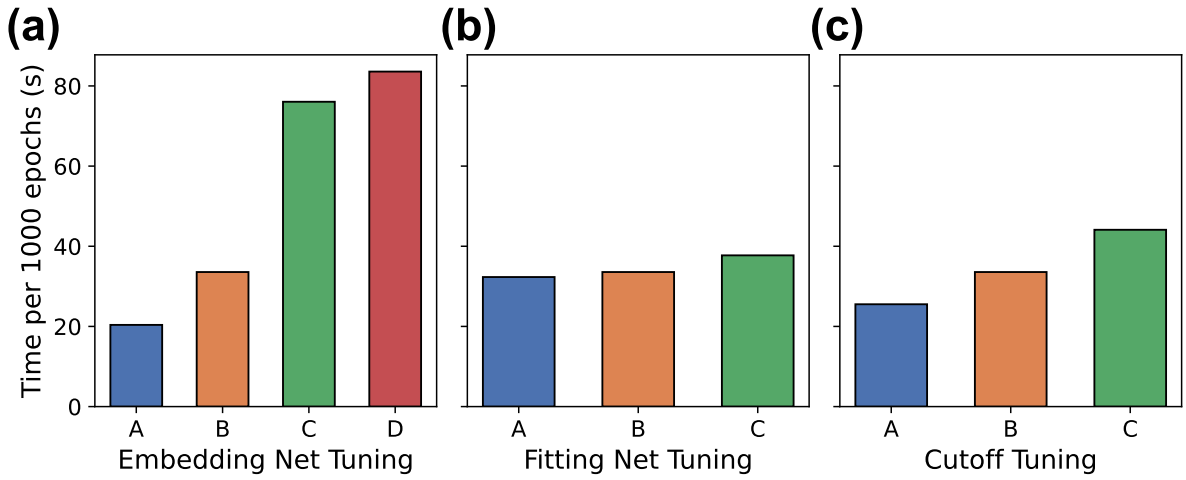

Figure 4.7. Relative training times per 1,000 epochs for different hyperparameter combinations for tuning the (a) embedding neural net (Table 4.1), (b) fitting neural net (Table 4.2), and (c) cutoff (Table 4.3).

4.4 Results and Discussion

4.4.1 Model Performance

The accuracy of the energy and force predictions for the trained DP model were evaluated using testing data that was unseen during model training (10% of the original dataset for each topology was set aside for post-training testing). The parity plots comparing DP predictions to DFT values for the energies (per SiO_2 unit) and forces are shown in Figure 4.8a and 4.8b, respectively. DP was found to be able to predict DFT values with excellent accuracy, as seen by the MAE of 2.6×10^{-3} eV/SiO_2 for energies and 3.9×10^{-2} eV/Å for forces. Note that the data shown in Figure 4.8a-b corresponds to the combined test sets of all 187 training topologies considered. Predictions for some topologies were found to be more or less accurate than others, and the complete list of MAE values for all individual topology test sets is included in the SI.

To further demonstrate our DP's ability to predict energies and forces on configurations outside of the training set, and to probe for any potential sampling biases arising from only including short DFT-MD trajectories in the initial training set, an additional 20,000 step DFT-MD run at 298 K was completed for CHA topology. The energies and forces of all configurations of the trajectory were evaluated with DP, and the MAE for the entire trajectory was found to be 0.95×10^{-4} eV/SiO_2 for the energy and 2.0×10^{-2} eV/Å for the force predictions. Figure 4.8c shows the DFT energies (black line) of the final 500 fs snippet from the trajectory with DP evaluations overlaid (dashed orange line).

4.4.2 Structural Properties

The structures of all 187 topologies were relaxed using DP and compared to DFT optimizations. The normalized distribution of all Si-O bond lengths and O-Si-O angles for all relaxed structures are shown in Figure 4.9a and 4.9b, respectively. The distributions for both angles and bond lengths with DP match almost perfectly with the DFT distributions, highlighting the remarkable ability of DP to replicate relaxed *ab initio* geometries.

The percent error distribution in calculated lattice constants relative to DFT values for DP (orange) and BKS (green) are shown in Figure 4.9c, where positive and negative errors

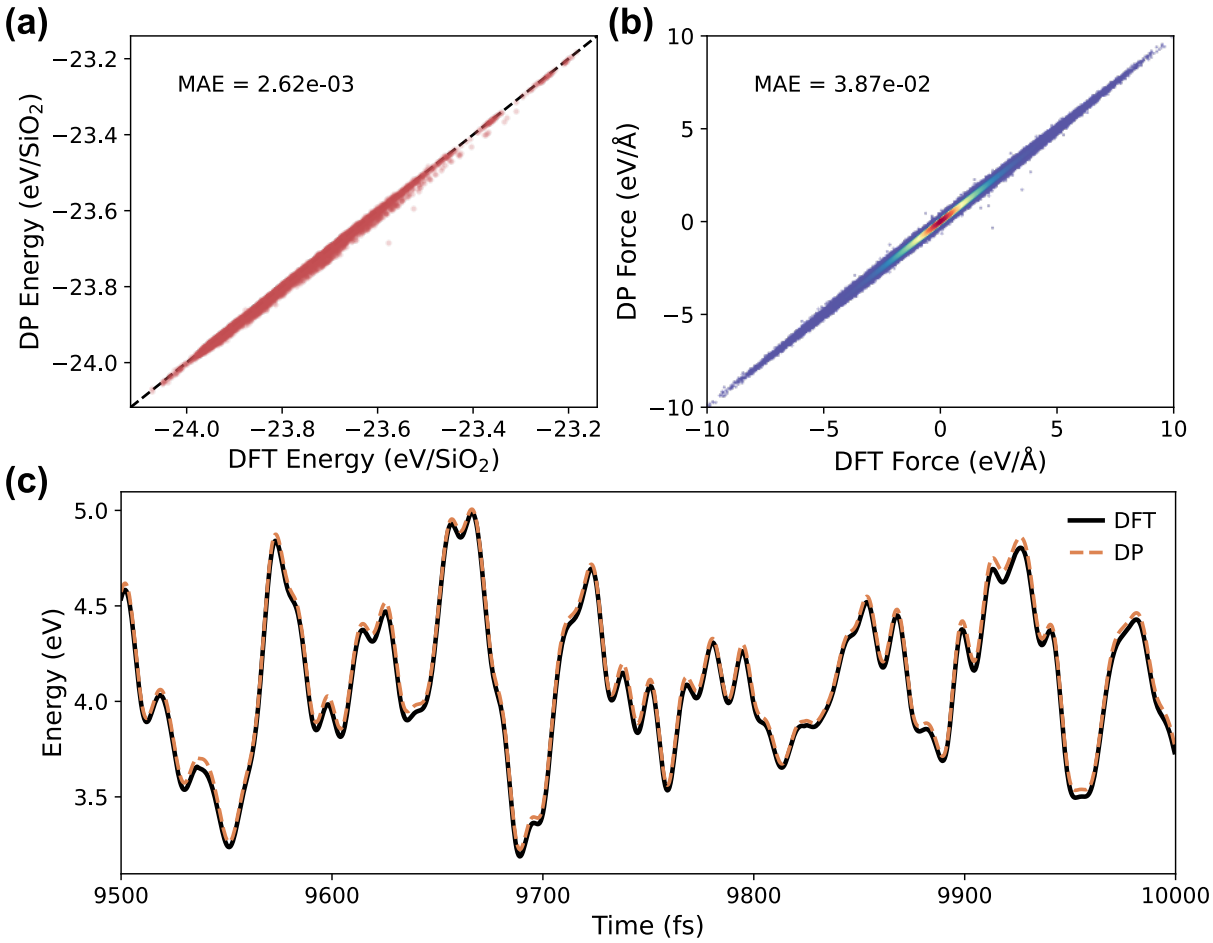

Figure 4.8. Parity plots comparing DP-predicted (a) energies and (b) forces with corresponding DFT values for test dataset not seen during training. (c) Energy relative to relaxed structure from the final 500 fs of a 10,000 fs DFT-MD (solid black line) run with DP predictions overlaid (dashed orange line) for CHA topology at 298 K.

correspond to overestimation and underestimation of lattice constants, respectively. The narrow distribution centered at 0% error for DP implies excellent agreement with DFT. BKS shows a wider distribution centered at positive error, indicating a slight tendency to overestimate the lattice constants compared to DFT. These results show that a DP trained on higher energy MD snapshots can still produce very similar global minima to the DFT PES.

We acknowledge that recently reported classical zeolite force fields[44,45] may show better performance than the BKS model used in Figure 4.9. As the central goal of this study is to develop a MLP model that shows similar accuracy to the DFT data, the comparison

with other classical force fields (beyond the BKS model) is beyond the scope of this work. Interested readers are referred to the seminal work of Sastre for more in-depth comparison across different zeolite force fields.[20,46,47]

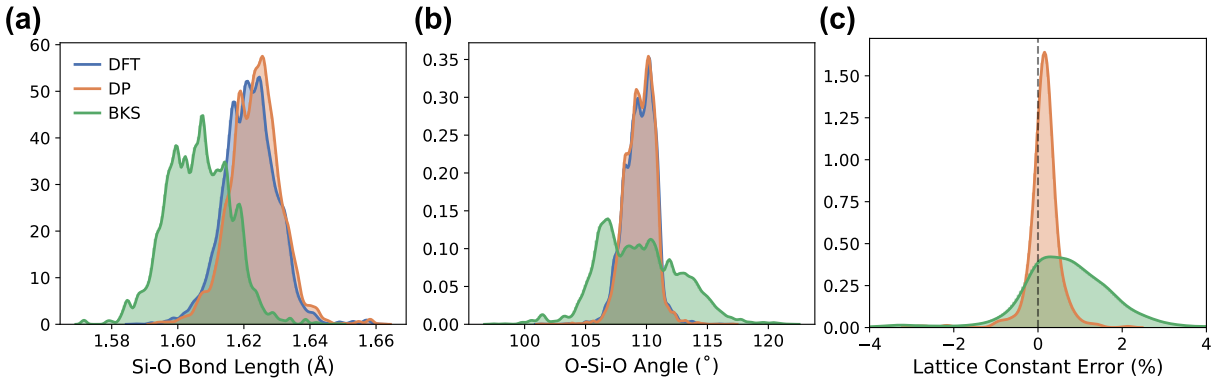

Figure 4.9. Normalized distributions of (a) Si-O bond lengths and (b) O-Si-O angles for relaxed geometries of the 187 topologies included in the training set for DFT, DP, and BKS. (c) Normalized distribution of percent errors relative to DFT of optimized lattice constants for DP and BKS. Vertical dashed black line denotes zero error (perfect agreement with DFT lattice constant).

4.4.3 Equation of State

Energy versus volume curves at 0 K were generated with DFT, DP, and BKS to assess how well DP can predict energies of systematically varied cell volumes. The resulting data were fit to the third-order Birch-Murnaghan EOS,

$$E(V) = E_0 + \frac{9V_0B_0}{16}\left\{\left[\left(\frac{V_0}{V}\right)^{2/3} - 1\right]^3 B_0' + \left[\left(\frac{V_0}{V}\right)^{2/3} - 1\right]^2 \left[6 - 4\left(\frac{V_0}{V}\right)^{2/3}\right]\right\}$$

where E_0 and V_0 are the energy and volume of the relaxed structure, respectively, and B_0 and B_0' are the bulk modulus (a property that describes the resistance to uniform compression/expansion) and its derivative. Thus, the bulk modulus can be determined from fitting energy-volume data to an EOS and serves as an additional metric for evaluating the performance of DP.

Taking the topologies of CHA, FER, and RHO as examples, the energy-volume curves and EOS fits are shown in Figure 4.10 for 15 volumes across ±5% volumetric strain. The DP data aligns very well with DFT, while BKS noticeably deviates. The similar curvature

of the EOS fits for DP and DFT suggests DP can accurately calculate bulk moduli. Additionally, the similar location of V_0 (the volume corresponding to the minimum energy) is further evidence that DP can accurately predict lattice constants. The higher curvature of the BKS energy-volume data implies that BKS overestimates the bulk moduli, and, similarly, the values of V_0 imply BKS overestimates the lattice constants for these topologies.

We extended this analysis to all other topologies included in the dataset, and the resulting parity plots comparing predicted bulk moduli values with DFT values for DP and BKS are shown in Figure 4.10d and Figure 4.10e, respectively. For computational efficiency, only 5 volumes were used with ±2% volumetric strain. The RMSE of bulk moduli calculated with DP was found to be 8.6 GPa, while BKS values had an RMSE of 31.4 GPa. Again, we see that BKS has a tendency to overestimate the bulk moduli in comparison to DFT.

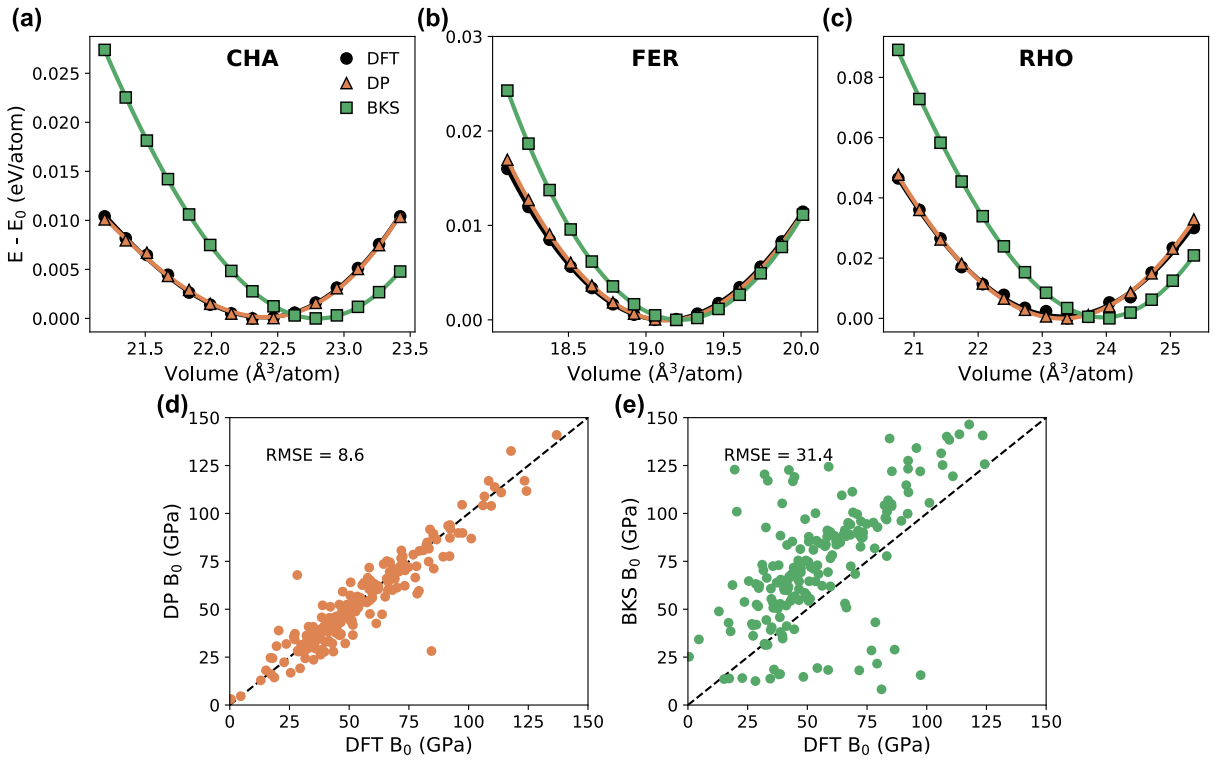

Figure 4.10. Energy-volume curves with third-order Birch-Murnaghan EOS fit for (a) CHA, (b) FER, and (c) RHO topologies for 15 volumes across ±5% volumetric strain. Parity plots comparing DFT with (d) DP and (e) BKS for bulk moduli calculated from EOS fits for all 187 topologies using 5 volumes across ±2% volumetric strain.

4.4.4 Mechanical Properties

Second-order elastic constants were calculated with Elastool[48] using the optimized high-efficiency strain-matrix set (OHESS) using 5 strains (±2% amplitude) for each deformation.[49] The elastic constants were used to compute Voigt-Reuss-Hill (VRH) averages of the bulk (K_{VRH}) and shear (G_{VRH}) moduli for 172 topologies. Figure 4.11 shows the agreement of DP and BKS with DFT for K_{VRH} and G_{VRH}. DP is able to predict K_{VRH} quite accurately for topologies with values less than around 60 GPa; however, there is a noticeable drop in accuracy for stiffer materials with high K_{VRH} values, with DP consistently underestimating bulk moduli relative to DFT. This suggests that that our DP model struggles to reproduce the expected stress-strain behavior for stiff topologies with high stress tensor values. Additionally, while DP tends to underestimate G_{VRH}, the overall predictions are comparable to the BKS predictions.

Accurate calculation of elastic constants using stress-strain relations requires accurate stresses, so the DFT calculated bulk and shear moduli were calculated using a 700 eV plane-wave cutoff to ensure convergence of the stress tensor components. We note that it is possible to train a DP model including virial stress error in the loss function, and doing so would likely improve the accuracy of the mechanical property calculations. However, the DFT training set configurations were calculated using a plane-wave energy cutoff of 400 eV, and higher cutoffs are needed to converge the stress tensor components. Therefore, it would not be appropriate to use the stress values for training. Notwithstanding these limitations, it is quite impressive that DP can produce reasonable predictions of mechanical properties that were calculated with a 700 eV cutoff even though the training data consists entirely of configurations calculated at 400 eV. We note that it is necessary to include stress data in the training and ensure the appropriate basis set is used to give reliable stress tensors to train on, as shown by the accurate calculations obtained in other work.[30] We also note that better agreement may be obtained by using methods that calculate elastic constants from energy-strain relationships as opposed to stress-strain. However, a detailed investigation into mechanical properties is beyond the scope of this work, and we elected to use stress-strain approaches to examine the accuracy of DP-calculated stresses

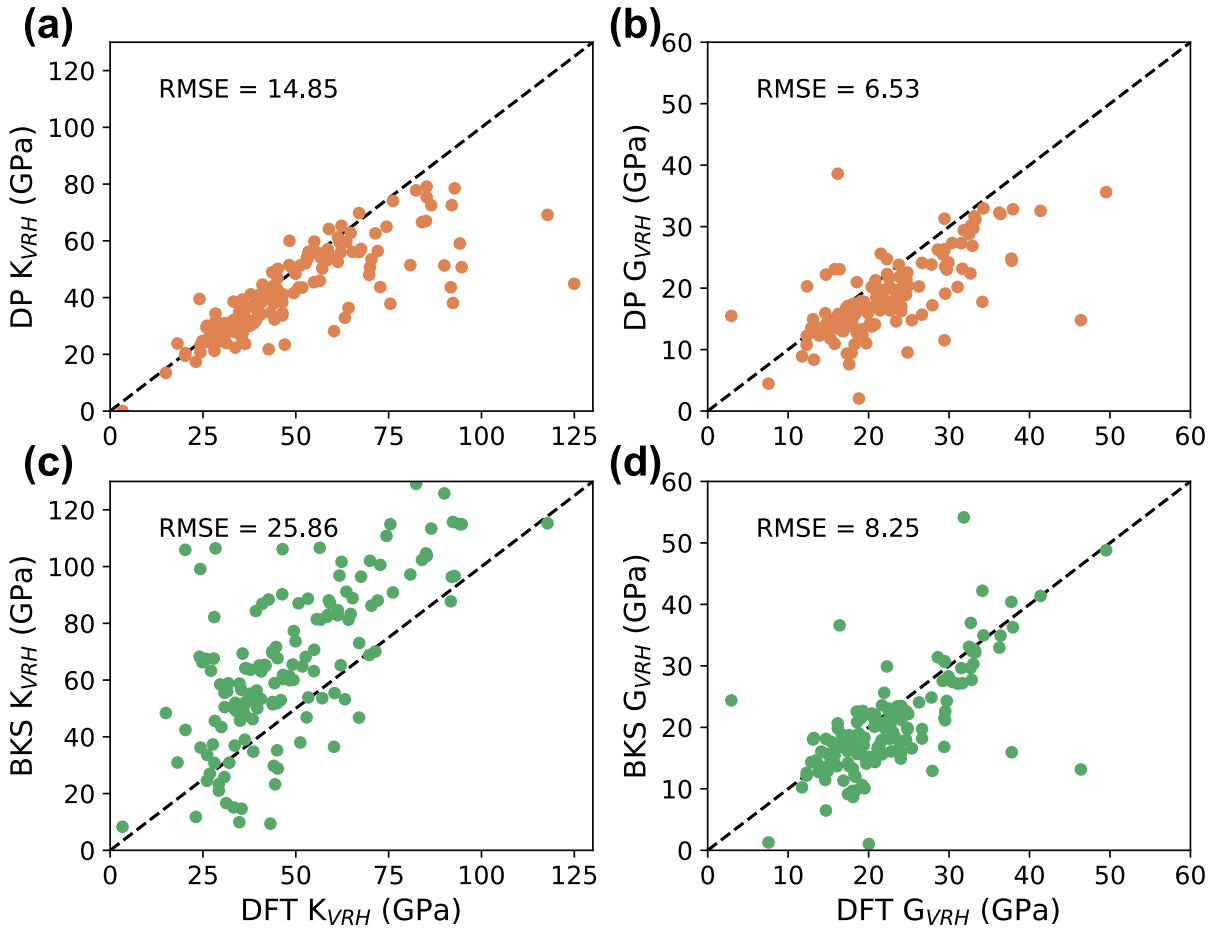

Figure 4.11. Parity plots comparing DFT VRH averages with DP-calculated (a) bulk moduli and (b) shear moduli and BKS-calculated (c) bulk moduli and (d) shear moduli.

when not included in training.

4.4.5 Phonon Density of States

The PDOS of CHA (chosen due to lower DFT computational cost) was calculated at 900 K to assess DP's ability to calculate vibrational modes. Atomic velocities from MD trajectories were used to calculate PDOS from the Fast Fourier transform of the velocity autocorrelation function. An MD trajectory of 10 ps was used for the DFT PDOS (black in Figure 4.12), and 50 ps were used for DP and BKS (orange and green, respectively) PDOS calculations. As the PDOS is calculated from the changes in atomic positions, which are determined by the atomic force calculations at each MD step, it provides a good metric to probe the accuracy of the DP forces. We see good agreement in the frequencies of

the vibrational modes between DFT and DP, while the intensity of the peaks is generally consistent but with some disagreement at a few frequencies. BKS shows a tendency to overpredict vibrational mode frequencies with broader and less intense peaks. These data demonstrate the suitability of the DP model for predicting phonon modes of silica zeolites at close to DFT accuracy.

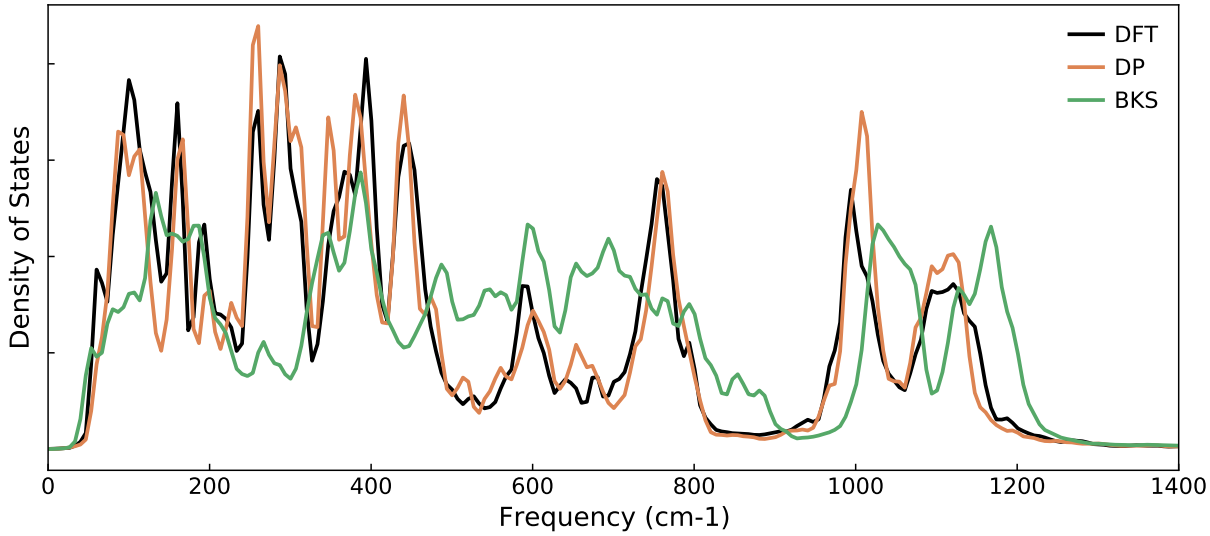

Figure 4.12. PDOS of CHA at 900 K calculated from the velocity autocorrelation function from an NVT-MD trajectory for DFT, DP, and BKS.

4.4.6 Model Transferability

The previous results assessed DP's ability to predict properties of the 187 topologies included in the model's training. We now examine a testing set of 32 topologies from our dataset (not used during training) to see how DP performs for topologies completely unseen by our model. The optimized geometries of these 32 zeolites were obtained using both DFT and our DP model to assess DP's ability to predict PES minima for new topologies. As seen in Figure 4.13a-c, the DP model continues to show impressive agreement with DFT for optimized geometries of new topologies. Both the optimized Si-O bond length and O-Si-O angle distributions align almost perfectly with DFT, and the calculated lattice constants agree with DFT typically within 1% error for the majority of topologies considered.

The calculations of bulk moduli (B_0) from fitting energy-volume data and bulk and

shear moduli (K_{VRH} and G_{VRH}, respectively) from elastic constants using stress-strain data were repeated for the test set of topologies. As shown by the excellent agreement between DP and DFT calculated EOS bulk moduli in Figure 4.13d, we find our DP is transferable and capable of mapping out PES of unseen topologies by learning the PES of many similar structures. High transferability suggests that MLPs may be ideally suited for applications involving high-throughput screening of large zeolite databases by calculating a property of interest at near-DFT accuracy. The DP model also yields reasonable K_{VRH} (Figure 4.13e) and G_{VRH} (Figure 4.13f) predictions for the testing topologies with accuracy on-par with that obtained for the training topologies. We reiterate that DFT stresses were not included in training, so it should be expected that K_{VRH} and G_{VRH} (calculated using stresses) be less accurate than B_0 (calculated using energies) for both the testing and training topologies.

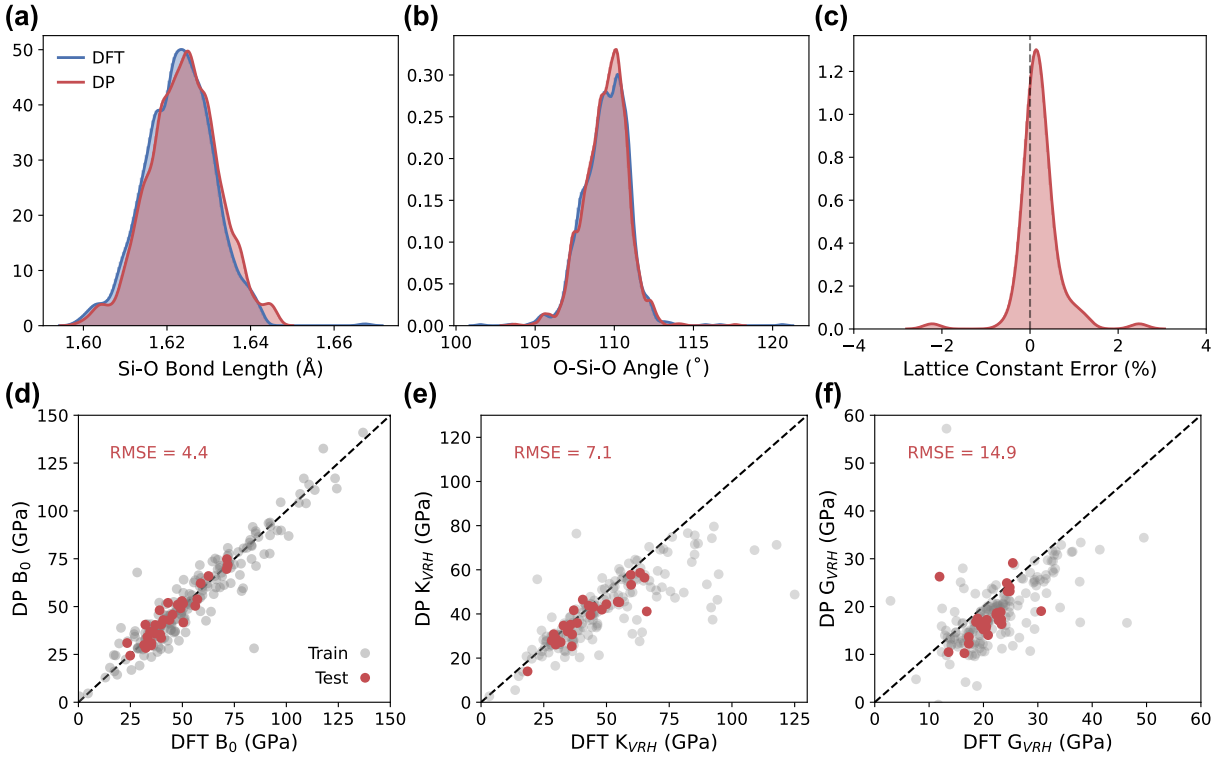

Figure 4.13. Comparison of normalized distributions of (a) Si-O bond lengths and (b) O-Si-O angles for optimized DFT and DP geometries of 32 topologies foreign to the trained model. (c) Normalized distribution of percent errors relative to DFT of optimized lattice constants for DP. (d) Bulk moduli calculated from EOS fits, and (e) bulk moduli and (f) shear moduli calculated from stress-strain relationships for the 32 testing topologies (red) compared to the 187 topologies included in training (gray).

4.4.7 Computational Cost

We end our discussion with a brief analysis of the computational efficiency of DFT and DP. The average time per MD step was found for DP and DFT for eight randomly selected topologies of varying system size. The DFT and DP calculations were both performed using 32 cores (2.3 GHz Intel® Xeon™ Processor E5-2698 v3) for a direct comparison of performance. Although using a fixed number of cores neglects potential scaling differences between DP and DFT with increasing CPU cores, an exhaustive cost analysis across different parallelization schemes is beyond the scope of this work. Figure 4.14 shows the speedup (ratio of DFT and DP time per MD step) for increasing system size. For our pure silica zeolite systems, we found DP to be > 1000 times faster than the corresponding DFT calculation, with the more favorable scaling of DP with increasing number of atoms leading to improved speedup for larger systems. Coupled with the accuracy of the results discussed previously, we conclude that the DP approach significantly improves the accuracy-cost trade off for these materials. Note that the above results were obtained with the CPU version of DeePMD-kit; using GPUs could lead to better performance and improved parallelization.

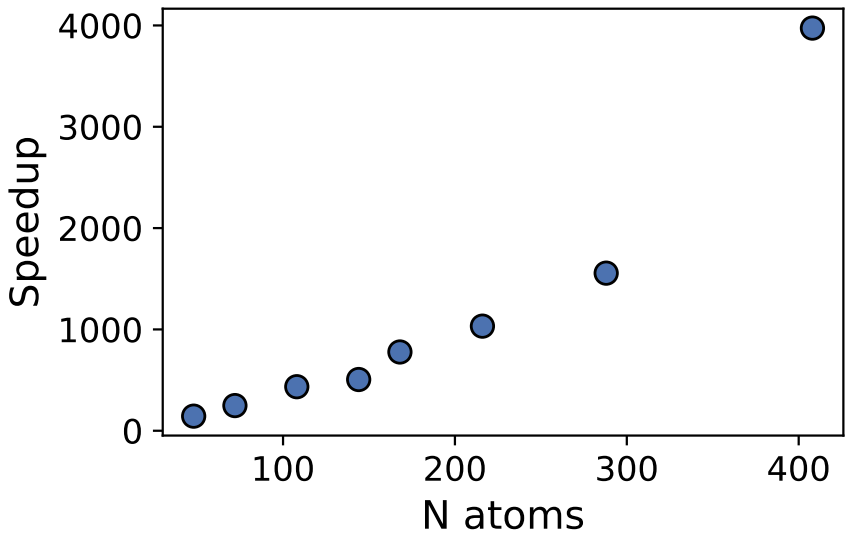

Figure 4.14. Computational speedup and scaling for DP compared to DFT for various silica topologies. The eight topologies chosen correspond to ACO, GME, CHA, MOR, SAO, STI, MFI, and IWS in order of increasing number of atoms.

4.5 Conclusion

In this work, a diverse DFT dataset was generated consisting of 219 pure silica zeolite topologies for training MLPs. DeePMD-kit was used to train a single DP on 187 of the 219 topologies (32 were set aside as a test set for model transferability) that accurately reproduces the *ab initio* PES of silica. We assessed the ability of the DP to calculate properties that were not explicitly trained for through energy and force evaluations. We have shown excellent agreement with DFT structural properties, as seen by nearly identical tetrahedral SiO_4 geometry and lattice constants in structures relaxed by DFT and DP. The accuracy of the energies and forces was also highlighted by good agreement with DP and DFT for energy-volume curves (EOS) and finite temperature PDOS calculated from MD velocities. Mechanical properties from elastic constants calculated from stress-strain relationships were found to show reasonable agreement, with large improvement likely to be gained from including DFT stresses during training. We also tested how our model performs at calculating these same properties for the 32 testing topologies not included during training, and we found comparable accuracy to the training set topologies, suggesting a generalized DP applicable for any pure silica zeolite structure. Our results indicate an MLP trained on *ab initio* data can successfully model zeolite framework dynamics. We are currently extending the DP approach to model the diffusion of small molecules and metal nanoclusters in zeolites and metal-organic frameworks (MOFs). Our findings provide a promising avenue to develop DP-based MLPs for zeolites, and are broadly relevant to the nanoporous modeling community. Additionally, we anticipate that the silica zeolite dataset developed in this work (denoted as Si-ZEO22) will motivate the development of other MLPs for this important class of industrially relevant materials.

4.6 References

[1] Ambarish R. Kulkarni and David S. Sholl. Screening of Copper Open Metal Site MOFs for Olefin/Paraffin Separations Using DFT-Derived Force Fields. *J. Phys.*

Chem. C, 120(40):23044–23054, 2016. doi: 10.1021/acs.jpcc.6b07493.

[2] Albert P. Bartók, Mike C. Payne, Risi Kondor, and Gábor Csányi. Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons. *Phys. Rev. Lett.*, 104(13):1–4, 2010. doi: 10.1103/PhysRevLett.104.136403.

[3] Stefan Chmiela, Huziel E. Sauceda, Klaus-Robert Müller, and Alexandre Tkatchenko. Towards exact molecular dynamics simulations with machine-learned force fields. *Nat. Commun.*, 9(1):3887, 12 2018. doi: 10.1038/s41467-018-06169-2.

[4] Albert P. Bartók, Risi Kondor, and Gábor Csányi. On representing chemical environments. *Phys. Rev. B*, 87(18):184115, 5 2013. doi: 10.1103/PhysRevB.87.184115.

[5] J. S. Smith, O. Isayev, and A. E. Roitberg. ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost. *Chem. Sci.*, 8(4): 3192–3203, 2017. doi: 10.1039/C6SC05720A.

[6] Linfeng Zhang, Jiequn Han, Han Wang, Wissam A. Saidi, Roberto Car, and Weinan E. End-to-end Symmetry Preserving Inter-atomic Potential Energy Model for Finite and Extended Systems. *Adv. Neural Inf. Process. Syst.*, 2018-Decem(NeurIPS 2018): 4436–4446, 5 2018.

[7] Han Wang, Linfeng Zhang, Jiequn Han, and Weinan E. DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics. *Comput. Phys. Commun.*, 228:178–184, 7 2018. doi: 10.1016/j.cpc.2018.03.016.

[8] Oliver T. Unke and Markus Meuwly. PhysNet: A Neural Network for Predicting Energies, Forces, Dipole Moments, and Partial Charges. *J. Chem. Theory Comput.*, 15(6):3678–3693, 6 2019. doi: 10.1021/acs.jctc.9b00181.

[9] K. T. Schütt, H. E. Sauceda, P.-J. Kindermans, A. Tkatchenko, and K.-R. Müller. SchNet – A deep learning architecture for molecules and materials. *J. Chem. Phys.*, 148(24):241722, 6 2018. doi: 10.1063/1.5019779.

[10] Oliver T. Unke, Stefan Chmiela, Michael Gastegger, Kristof T. Schütt, Huziel E. Sauceda, and Klaus Robert Müller. SpookyNet: Learning force fields with electronic degrees of freedom and nonlocal effects. *Nat. Commun.*, 12(1):7273, 2021. doi: 10.1038/s41467-021-27504-0.

[11] Max Pinheiro, Fuchun Ge, Nicolas Ferré, Pavlo O. Dral, and Mario Barbatti. Choosing the right molecular machine learning potential. *Chem. Sci.*, 12(43):14396–14413, 2021. doi: 10.1039/d1sc03564a.

[12] Jörg Behler and Gábor Csányi. Machine learning potentials for extended systems: a perspective. *Eur. Phys. J. B*, 94(7):142, 7 2021. doi: 10.1140/epjb/s10051-021-00156-1.

[13] Yunxing Zuo, Chi Chen, Xiangguo Li, Zhi Deng, Yiming Chen, Jörg Behler, Gábor Csányi, Alexander V. Shapeev, Aidan P. Thompson, Mitchell A. Wood, and et al. Performance and Cost Assessment of Machine Learning Interatomic Potentials. *J. Phys. Chem. A*, 124(4):731–745, 2020. doi: 10.1021/acs.jpca.9b08723.

[14] Pascal Friederich, Florian Häse, Jonny Proppe, and Alán Aspuru-Guzik. Machine-learned potentials for next-generation matter simulations. *Nat. Mater.*, 20(6):750–761, 2021. doi: 10.1038/s41563-020-0777-6.

[15] Oliver T. Unke, Stefan Chmiela, Huziel E. Sauceda, Michael Gastegger, Igor Poltavsky, Kristof T. Schütt, Alexandre Tkatchenko, and Klaus-Robert Müller. Machine Learning Force Fields. *Chem. Rev.*, 121(16):10142–10186, 8 2021. doi: 10.1021/acs.chemrev.0c01111.

[16] Scott M. Auerbach, Kathleen A. Carrado, and Prabir K. Dutta. *Handbook of Layered Materials*. CRC Press, Boca Raton, 3 2004. doi: 10.1201/9780203021354.

[17] Cristina Martínez and Avelino Corma. Inorganic molecular sieves: Preparation, modification and industrial application in catalytic processes. *Coord. Chem. Rev.*, 255(13-14):1558–1580, 7 2011. doi: 10.1016/j.ccr.2011.03.014.

[18] Ramdas Pophale, Phillip A. Cheeseman, and Michael W. Deem. A database of new zeolite-like materials. *Phys. Chem. Chem. Phys.*, 13(27):12407–12412, 2011. doi: 10.1039/c0cp02255a.

[19] N. E. R. Zimmermann, S. Jakobtorweihen, E. Beerdsen, B. Smit, and F. J. Keil. In-Depth Study of the Influence of Host-Framework Flexibility on the Diffusion of Small Gas Molecules in One-Dimensional Zeolitic Pore Systems. *J. Phys. Chem. C*, 111(46):17370–17381, 11 2007. doi: 10.1021/jp0746446.

[20] Aldo F. Combariza, Diego A. Gomez, and German Sastre. Simulating the properties of small pore silica zeolites using interatomic potentials. *Chem. Soc. Rev.*, 42(1): 114–127, 2013. doi: 10.1039/c2cs35243e.

[21] Hsin Yu Ko, Linfeng Zhang, Biswajit Santra, Han Wang, E. Weinan, Robert A. DiStasio, and Roberto Car. Isotope effects in liquid water via deep potential molecular dynamics. *Mol. Phys.*, 117(22):3269–3281, 2019. doi: 10.1080/00268976.2019.1652366.

[22] Linfeng Zhang, Han Wang, Roberto Car, and Weinan E. Phase Diagram of a Deep Potential Water Model. *Phys. Rev. Lett.*, 126(23):236001, 6 2021. doi: 10.1103/PhysRevLett.126.236001.

[23] Jinzhe Zeng, Liqun Cao, Mingyuan Xu, Tong Zhu, and John Z. H. Zhang. Complex reaction processes in combustion unraveled by neural network-based molecular dynamics simulation. *Nat. Commun.*, 11(1):5713, 12 2020. doi: 10.1038/s41467-020-19497-z.

[24] Jinzhe Zeng, Linfeng Zhang, Han Wang, and Tong Zhu. Exploring the Chemical Space of Linear Alkane Pyrolysis via Deep Potential GENerator. *Energy & Fuels*, 35 (1):762–769, 1 2021. doi: 10.1021/acs.energyfuels.0c03211.

[25] Siddarth K Achar, Jacob J Wardzala, Leonardo Bernasconi, Linfeng Zhang, and J Karl Johnson. Combined Deep Learning and Classical Potential Approach for

Modeling Diffusion in UiO-66. *J. Chem. Theory Comput.*, 18(6):3593–3606, 6 2022. doi: 10.1021/acs.jctc.2c00010.

[26] Nan Xu, Yao Shi, Yi He, and Qing Shao. A Deep-Learning Potential for Crystalline and Amorphous Li-Si Alloys. *J. Phys. Chem. C*, 124(30):16278–16288, 2020. doi: 10.1021/acs.jpcc.0c03333.

[27] R. Li, E. Lee, and T. Luo. A unified deep neural network potential capable of predicting thermal conductivity of silicon in different phases. *Mater. Today Phys.*, 12:100181, 3 2020. doi: 10.1016/j.mtphys.2020.100181.

[28] Marcos F. Calegari Andrade, Hsin-Yu Ko, Linfeng Zhang, Roberto Car, and Annabella Selloni. Free energy of proton transfer at the water–TiO 2 interface from ab initio deep potential molecular dynamics. *Chem. Sci.*, 11(9):2335–2341, 2020. doi: 10.1039/C9SC05116C.

[29] Alejandro Rodriguez, Stephen Lam, and Ming Hu. Thermodynamic and Transport Properties of LiF and FLiBe Molten Salts with Deep Learning Potentials. *ACS Appl. Mater. Interfaces*, 13(46):55367–55379, 2021. doi: 10.1021/acsami.1c17942.

[30] Tongqi Wen, Rui Wang, Lingyu Zhu, Linfeng Zhang, Han Wang, David J. Srolovitz, and Zhaoxuan Wu. Specialising neural network potentials for accurate properties and application to the mechanical response of titanium. *npj Comput. Mater.*, 7(1): 206, 12 2021. doi: 10.1038/s41524-021-00661-y.

[31] Siddarth K. Achar, Linfeng Zhang, and J. Karl Johnson. Efficiently Trained Deep Learning Potential for Graphane. *J. Phys. Chem. C*, 125(27):14874–14882, 7 2021. doi: 10.1021/acs.jpcc.1c01411.

[32] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E. Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N. Groves, BjØrk Hammer, Cory Hargus, and et al. The atomic simulation environment—a Python library for working with atoms. *J. Phys. Condens. Matter*, 29(27):273002, 7 2017. doi: 10.1088/1361-648X/aa680e.

[33] Aidan P. Thompson, H. Metin Aktulga, Richard Berger, Dan S. Bolintineanu, W. Michael Brown, Paul S. Crozier, Pieter J. in 't Veld, Axel Kohlmeyer, Stan G. Moore, Trung Dac Nguyen, and et al. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput. Phys. Commun.*, 271:108171, 2 2022. doi: 10.1016/j.cpc.2021.108171.

[34] Jack D. Evans and François Xavier Coudert. Predicting the Mechanical Properties of Zeolite Frameworks by Machine Learning. *Chem. Mater.*, 29(18):7833–7839, 2017. doi: 10.1021/acs.chemmater.7b02532.

[35] Namjung Kim and Kyoungmin Min. Accelerated Discovery of Zeolite Structures with Superior Mechanical Properties via Active Learning. *J. Phys. Chem. Lett.*, 12(9): 2334–2339, 2021. doi: 10.1021/acs.jpclett.1c00339.

[36] Maxime Ducamp and François Xavier Coudert. Prediction of Thermal Properties of Zeolites through Machine Learning. *J. Phys. Chem. C*, 126(3):1651–1660, 2022. doi: 10.1021/acs.jpcc.1c09737.

[37] B.W.H. van Beest and G.J. Kramer. Force Fields for Silicas and Aluminophosphates Based on Ab Initio Calculations. *Phys. Rev. Lett.*, 64(16):1955–1958, 1990.

[38] John P Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized Gradient Approximation Made Simple. *J. Chem. Phys.*, 77(18):3865–3868, 1996.

[39] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.*, 132(15):154104, 4 2010. doi: 10.1063/1.3382344.

[40] Stefan Grimme, Stephan Ehrlich, and Lars Goerigk. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.*, 32(7):1456–1465, 5 2011. doi: 10.1002/jcc.21759.

[41] Daniel G. A. Smith, Lori A. Burns, Konrad Patkowski, and C. David Sherrill. Revised Damping Parameters for the D3 Dispersion Correction to Density Functional Theory. *J. Phys. Chem. Lett.*, 7(12):2197–2203, 6 2016. doi: 10.1021/acs.jpclett.6b00780.

[42] Linfeng Zhang, De Ye Lin, Han Wang, Roberto Car, and E. Weinan. Active learning of uniformly accurate interatomic potentials for materials simulation. *Phys. Rev. Mater.*, 3(2):1–9, 2019. doi: 10.1103/PhysRevMaterials.3.023804.

[43] Yuzhi Zhang, Haidi Wang, Weijie Chen, Jinzhe Zeng, Linfeng Zhang, Han Wang, and Weinan E. DP-GEN: A concurrent learning platform for the generation of reliable deep learning based potential energy models. *Comput. Phys. Commun.*, 253:107206, 8 2020. doi: 10.1016/j.cpc.2020.107206.

[44] German Sastre. Computational study of diffusion of propane in small pore acidic zeotypes AFX and AEI. *Catal. Today*, 226:25–36, 5 2014. doi: 10.1016/j.cattod.2013.07.021.

[45] An Ghysels, Samuel L.C. Moors, Karen Hemelsoet, Kristof De Wispelaere, Michel Waroquier, German Sastre, and Veronique Van Speybroeck. Shape-Selective Diffusion of Olefins in 8-Ring Solid Acid Microporous Zeolites. *J. Phys. Chem. C*, 119 (41):23721–23734, 10 2015. doi: 10.1021/acs.jpcc.5b06010.

[46] Jordi Toda, Avelino Corma, Raed H. Abudawoud, Mohamed S. Elanany, Ibrahim M. Al-Zahrani, and German Sastre. Influence of force fields on the selective diffusion of para -xylene over ortho -xylene in 10-ring zeolites. *Mol. Simul.*, 41(16-17):1438–1448, 11 2015. doi: 10.1080/08927022.2015.1047370.

[47] Diego Bermúdez and German Sastre. Calculation of pore diameters in zeolites. *Theor. Chem. Acc.*, 136(10):116, 10 2017. doi: 10.1007/s00214-017-2143-6.

[48] Zhong-Li Liu, C.E. Ekuma, Wei-Qi Li, Jian-Qun Yang, and Xing-Ji Li. ElasTool: An automated toolkit for elastic constants calculation. *Comput. Phys. Commun.*, 270: 108180, 1 2022. doi: 10.1016/j.cpc.2021.108180.

[49] Zhong-Li Liu. High-efficiency calculation of elastic constants enhanced by the optimized strain-matrix sets. *arXiv*, pages 20–22, 1 2020. doi: 10.48550/ARXIV.2002.00005.

Chapter 5

Summary and Future Directions

The work presented here represents several of the diverse applications of atomistic modeling to gain insight about a class of materials. The first section detailed how DFT predictions were used with theoretical electrochemical models to evaluate bimetallic porphyrin-based MOFs as catalysts for oxygen reduction in fuel cell devices. Certain synergetic combinations of transition metal porphyrin complexes in the MOF linkers were shown to provide ideal catalytic environments with selective stabilization of the $*OOH$ intermediate, resulting in favorable modifications of the infamous scaling relationships that bottleneck the performance of today's commonly used 2-dimensional surface catalysts. Specifically, bimetallic Cr and Fe Al-PMOF was identified as a highly active catalyst with a theoretical limiting potential of 1.07 V (compared to 0.8 V with Pt, the most common ORR catalyst).

As electrochemical oxygen reduction is inherently a multiscale process with oxygen diffusion, charge transfer, and the intrinsic catalytic activity all influencing the overall reaction rate,[1] there are many under-explored avenues of research that need to be addressed to refine the theoretical models used for this newer class of nanoporous electrocatalysts. In particular, the solvent environment around the reaction centers is poorly understood in these MOF materials, as the topological features of the MOF heavily influence solvent transport.[2] Accurate description of the solvent is essential for ORR since the oxygen-based reaction intermediates are stabilized by hydrogen bonding interactions with the surrounding water molecules.[3] Solvation models that explicitly include water molecules

in the simulations are necessary but computationally costly due to the large configurational space of solvated systems.[4] To this end, the methodologies for training accurate machine learning potentials discussed in Chapter 4 can be directly applied to water in MOFs. A DFT training data set of water interacting with each adsorbate in the MOF can be easily developed. In short, molecular dynamics (MD) simulations at various temperatures and pressures with a range of water loadings surrounding each adsorbate can be used to generate DFT configurations. Several MD runs with different initial starts from randomly shuffling the water molecule locations coupled with temperature and pressure ramping would efficiently sample the PES of the solvated ORR system. Iteratively training an MLP with the active learning procedure previously discussed would result in a flexible ML-powered force field to combine with existing methodologies for calculating solvation effects in ORR.

Chapter 3 introduced an approach for characterizing atomically dispersed transition metal catalysts using a suite of experimental and theoretical techniques. This work leveraged high throughput computational screening as a means to validate experimental measurements for difficult to characterize active sites. State-of-the-art experimental x-ray absorption spectra were compared to theoretical equivalents for a library of DFT calculated potential active site structures to elucidate a sub-surface Pt^{4+} embedded in MgO. Experimental CO oxidation apparent activation energies were compared with DFT-calculated energy barriers to validate the activity of the identified catalytic site. This work highlights the powerful synergy of designing experiments that incorporate theoretical feedback (and vice-versa) and the importance of diverse, interdisciplinary collaborations.

The work presented in Chapter 4 developed a large DFT data set (SiZEO-22) and trained a generalized MLP that was shown to be accurate for representing the PES of pure silica zeolites. This was demonstrated by calculating various unrelated properties that involve performing atomistic simulations and evaluating energies and forces of atomic configurations not included in training. Specifically, the MLP accurately predicted relaxed geometries (PES minima), energy-volume relationships, phonon/vibrational properties, and mechanical properties (shear and bulk moduli). The open-access release of the Si-

ZEO22 data set provides researchers with an exhaustive set of DFT data for an industrially significant class of materials. As the field of MLP research is fairly nascent with new methodologies constantly being developed, Si-ZEO22 will serve as an invaluable resource for validating new MLP formulations.

This project played a pivotal role in the Kulkarni group's transition to machine-learning based molecular modeling. It is our first publication involving the usage of MLP approaches, and the data, workflows, and software developed during the course of the project were used to establish a computational infrastructure in a new frontier of research for our group. Additionally, zeolites are a central class of material to our group's research, and by demonstrating the effectiveness of MLPs for describing the flexible dynamics of the frameworks, this work can be extended to model more complex phenomena in nanoporous materials (such as reactions and gas adsorption). To this end, several ongoing works such as gold nanoparticle diffusion in zeolites, formation of subsurface Pt sites in MgO, small molecule diffusion in metal-organic frameworks, and methane activation in Cu-exchanged zeolites are all currently underway in our group. The lessons learned and technical difficulties encountered in this work have allowed me to establish a role as a mentor for guiding these studies.

We have already made significant progress with applying a DeePMD-kit MLP to calculate methane activation reaction barriers for thousands of Cu-exchanged zeolite sites. The training set was created with rigorous sampling of the configuration space around the reaction coordinate for 30 initial CuOCu sites. MD runs with the initial state (CH_4), final state (CH_3^+), and constrained images of the reaction coordinate from DFT nudged-elastic band (NEB) calculations for these 30 sites were used to train an initial model. This initial MLP was then used to perform high throughput NEBs to generate many configurations along the CH_4 to CH_3^+ reaction coordinate to add to the training set. Several training iterations were performed with these 30 sites and then extended to predict the energy barriers for 5,400 unique CuOCu sites, increasing the feasible search space relative to DFT by a factor of around 100. Additional training iterations were performed, each time adding more DFT data from reaction coordinate images of these sites generated by

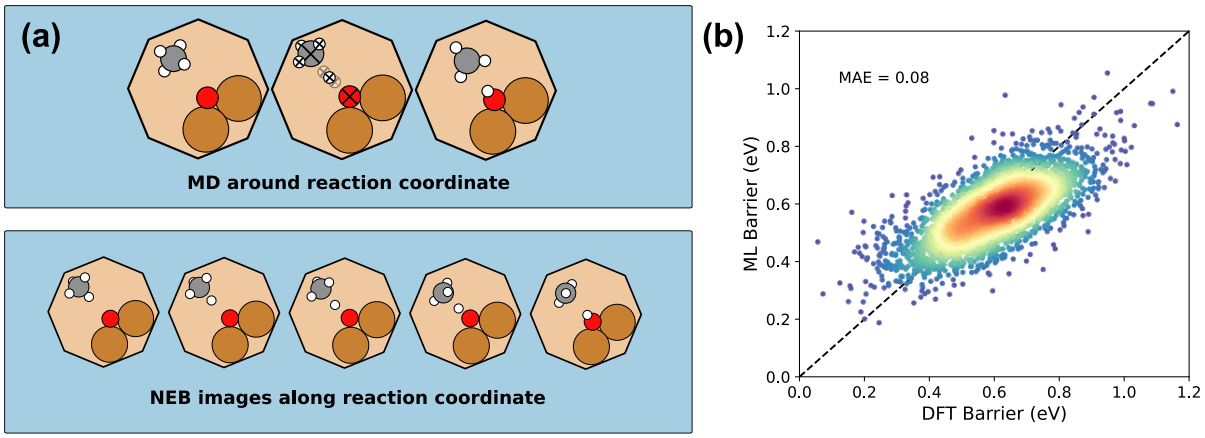

Figure 5.1. (a) Sources of training data for training DeePMD-kit MLP to model CH_4 activation in Cu-exchanged zeolites. (b) Preliminary results of MLP-calculated energy barriers for 5,400 CuOCu sites compared to DFT values.

MLP NEBs. Figure 5.1 summarizes the DFT training set constituents and preliminary predictions of the 5,400 site reaction barrier energies compared to DFT. We expect to submit this work to be published in Spring 2023.

Additionally, these MLP studies have motivated investigation into GPU-accelerated MD simulations. Figure 5.2 shows benchmark calculations for a 143 atom unit cell and 4x4 supercell of an initial MLP model trained on the data from the Pt MgO system detailed in Chapter 3. As the computational cost of DFT scales approximately $O(N^3)$ with increasing number of atoms, simulations involving more than a few hundred atoms are highly unfeasible with current computing power. The linear scaling of MLPs enable simulations with larger molecular systems,[5] with GPU-powered simulations scaling extremely favorably.

The significantly lowered cost of simulating large systems with near-DFT accuracy with GPU-powered MLP models presents many exciting opportunities. For instance, characterization of defects in MOFs is a new area of active research that would benefit from these approaches. Defective MOF materials are challenging to characterize due to the formation of nanoregions of assorted defect configurations that require using large supercells to simulate.[6–8] By training an MLP on various unit cells and small supercells of various pristine and defective MOFs (different variations of missing node and linker defects), the MLP can then be applied to explicitly model larger nanoregions of these

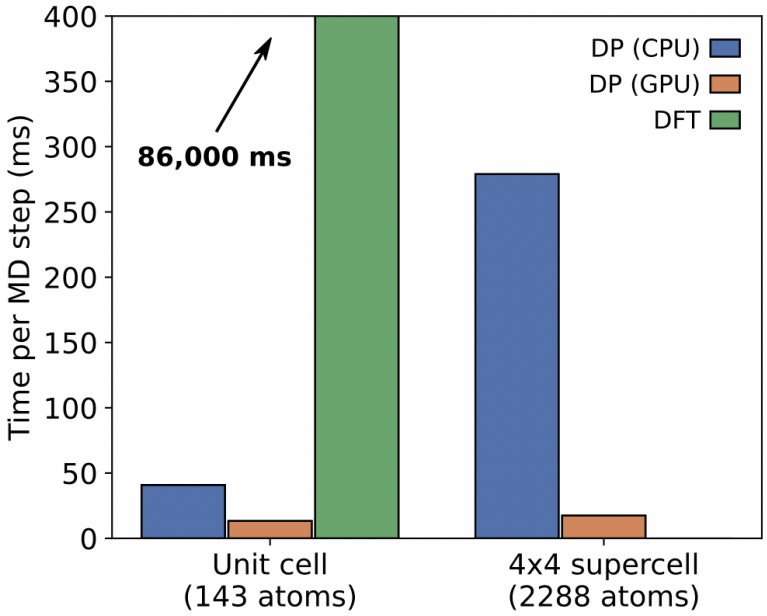

Figure 5.2. Computational cost comparison for DFT and MLP (DP) using CPU and GPU computing resources.

materials and probe specific defect configurations.

Multiscale atomistic modeling continues to be applied in novel ways to discover new materials and provide atomic-level insight to experimental observations. The work presented here is a small contribution to a global effort directed at leveraging computational workflows to automate scientific discovery. With impending advancements in computing and machine learning coupled with the growing collection of publicly available material and chemical databases, computational approaches will serve a crucial role in driving technological progress in renewable energy, pharmaceuticals, and more.

References

[1] Itamar Liberman, Ran Shimoni, Raya Ifraemov, Illya Rozenberg, Chanderpratap Singh, and Idan Hod. Active-Site Modulation in an Fe-Porphyrin-Based Metal-Organic Framework through Ligand Axial Coordination: Accelerating Electrocatalysis and Charge-Transport Kinetics. *Journal of the American Chemical Society*, 142(4):1933–1940, 2020. doi: 10.1021/jacs.9b11355.

[2] Meng Cai, Quentin Loague, and Amanda J. Morris. Design Rules for Efficient Charge Transfer in Metal-Organic Framework Films: The Pore Size Effect. *Journal of Physical Chemistry Letters*, 11(3):702–709, 2020. doi: 10.1021/acs.jpclett.9b03285.

[3] Anjli M. Patel, Stefan Ringe, Samira Siahrostami, Michal Bajdich, Jens K. Nørskov, and Ambarish R. Kulkarni. Theoretical Approaches to Describing the Oxygen Reduction Reaction Activity of Single-Atom Catalysts. *Journal of Physical Chemistry C*, 122(51):29307–29318, 2018. doi: 10.1021/acs.jpcc.8b09430.

[4] Nawras Abidi, Kang Rui Garrick Lim, Zhi Wei Seh, and Stephan N. Steinmann. Atomistic modeling of electrocatalysis: Are we there yet? *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 11(3):1–27, 2021. doi: 10.1002/wcms.1499.

[5] Alejandro Rodriguez, Stephen Lam, and Ming Hu. Thermodynamic and Transport Properties of LiF and FLiBe Molten Salts with Deep Learning Potentials. *ACS Applied Materials and Interfaces*, 13(46):55367–55379, 2021. doi: 10.1021/acsami.1c17942.

[6] Lingmei Liu, Zhijie Chen, Jianjian Wang, Daliang Zhang, Yihan Zhu, Sanliang Ling, Kuo-Wei Huang, Youssef Belmabkhout, Karim Adil, Yuxin Zhang, Ben Slater, Mohamed Eddaoudi, and Yu Han. Imaging defects and their evolution in a metal–organic framework at sub-unit-cell resolution. *Nature Chemistry*, 11(7):622–628, jul 2019. doi: 10.1038/s41557-019-0263-4.

[7] Ying Wu, Haipeng Duan, and Hongxia Xi. Machine Learning-Driven Insights into Defects of Zirconium Metal–Organic Frameworks for Enhanced Ethane–Ethylene Separation. *Chemistry of Materials*, 32(7):2986–2997, 2020. doi: 10.1021/acs.chemmater.9b05322.

[8] Lucas S R Cavalcante, Makena A Dettmann, Tyler Sours, Dong Yang, Luke L Daemen, Bruce C Gates, Ambarish R Kulkarni, and Adam J Moulé. Elucidating correlated defects in metal organic frameworks using theory-guided inelastic neutron scattering spectroscopy. *Materials Horizons*, 10(1):187–196, 2023. doi: 10.1039/D2MH00914E.