NORTHWESTERN UNIVERSITY

Quantum-Chemical Screening of Redox-Active Metal–Organic Frameworks

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Chemical and Biological Engineering

By

Andrew S. Rosen

EVANSTON, ILLINOIS

September 2021

ABSTRACT

Quantum-Chemical Screening of Redox-Active Metal–Organic Frameworks

Andrew S. Rosen

Metal–organic frameworks (MOFs) are a class of crystalline materials composed of metal nodes connected by organic linkers. Due to their high degree of synthetic tunability, MOFs have been considered for a wide range of applications, including many that rely on a change in oxidation state. While most MOFs are generally considered to be redox-inactive, a growing number of MOF structures have been synthesized that can support redox processes, oftentimes via the presence of open-shell transition metal cations and/or redox non-innocent linkers. These so-called redox-active MOFs have been investigated for challenging catalytic oxidation reactions, the selective adsorption of reducible gas species, and next-generation electronic devices.

In this dissertation, a computational screening approach based on density functional theory calculations is used to gain insight into the reactive properties of redox-active MOFs for three main application areas. The first portion of this dissertation is focused on the catalytic oxidation of strong C–H bonds, such as those of light alkanes, via the formation of high-valent metal-oxo and metal-oxyl species at the inorganic nodes of MOFs. Switching focus to adsorption processes, the second portion of this dissertation is centered around the design of MOFs with redox-active metal centers that can selectively bind O_2 over N_2 via charge transfer interactions. In the final portion of this dissertation, a high-throughput periodic DFT workflow is used to create the first large-scale quantum-mechanical property database for MOFs, which is then used to train machine learning

models that can guide the discovery of MOFs with targeted band gaps. Throughout this work, several methodological studies are also carried out to better understand the qualitative and quantitative shortcomings of different density functional approximations with the goal of making more actionable predictions in future computational screening efforts. Collectively, this dissertation demonstrates the ability to use high-throughput quantum-mechanical simulations to discover new structure–property relationships, identify promising MOFs for challenging oxidation reactions, and more efficiently explore the vast expanse of MOF chemical space.

ACKNOWLEDGMENTS

First and foremost, I wish to thank my fantastic advisors Randy Snurr and Justin Notestein whose guidance and unwavering support helped make my graduate school experience such a positive one. I am incredibly fortunate to have had two advisors who gave me the academic freedom to pursue a wide range of topics, helped keep me on track when I would occasionally go astray, and always had faith in me throughout the course of my graduate studies. I am grateful for all they have taught me over the years.

I thank my committee members Omar Farha and Chris Wolverton as well as my collaborators whose stimulating discussions and ideas helped shape this thesis, including Alán Aspuru-Guzik, Haoyuan Chen, Laura Gagliardi, Timur Islamoglu, M. Rasel Mian, Debmalya Ray, and Zhenpeng Yao. It has also been an absolute privilege to work with and mentor Shaelyn Iyer who was paramount in bringing several machine learning ideas to life. Thank you to Scotty Bobbitt, Benjamin Bucior, Kenton Hicks, Scott Nauert, and Megan Wasson who led exciting research projects that I had the pleasure of working on throughout my graduate studies.

Many individuals in the Department of Chemical & Biological Engineering (ChBE) have my utmost thanks for helping me throughout my time at Northwestern. While there are far too many people to list here, I would like to take the opportunity to specifically thank the prior and current members of the Snurr and Notestein research groups, the friends and colleagues of my cohort, the ChBE staff, and the many ChBE faculty members I have been fortunate to interact with over the last several years.

On a slightly less personal note, I acknowledge financial support from the U.S. Department of Defense for a National Defense Science and Engineering Graduate Fellowship, the International Institute for Nanotechnology for a Ryan Fellowship, and The Graduate School at Northwestern University for a Presidential Fellowship that collectively allowed me to pursue whatever research topics my heart desired. Similarly, I thank the many supercomputing facilities and computing staff I have relied on throughout my graduate studies, without which none of this work would have been possible.

Finally, I thank my wife Alexandra Brumberg for the constant encouragement she has provided throughout my graduate school career. This thesis would not be what it is without her love, support, and willingness to hear me rant about DFT yet another time.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

Chapter 1: INTRODUCTION

1.1 Metal–Organic Frameworks

1.1.1 Overview

Metal–organic frameworks (MOFs) are a class of highly tunable, extended solids composed of metal ions or clusters connected via organic ligands.[1] The constituent inorganic nodes and organic linkers that make up a given MOF are known as secondary building units and dictate the resulting physical and chemical properties of the material. By considering different combinations of node and linker building blocks, as well as the network topology describing their connectivity,[2] it is possible to design a MOF that is specifically tailored for an application of interest.[3] In many ways, MOFs can be thought of as bringing the synthetic tunability of molecular systems to the solid state, bridging the gap between these two otherwise distinct classes of materials.[4]

MOFs have been studied for an enormous range of applications, including gas storage and separations,[5,6] heterogeneous catalysis,[7,8] energy storage and conversion,[9] sensing,[10] drug delivery,[11] water purification,[12] and much more.[1] Beyond these many potential uses, MOFs are an important class of materials for gaining a greater fundamental understanding of a variety of chemical processes. For instance, in the area of heterogeneous catalysis, MOFs can be used to discover structure–reactivity relationships by systematically exchanging the metals at the nodes and/or functionalizing the linkers. The relatively uniform and spatially isolated nature of the metal sites in MOFs can also provide insight into the properties of proposed active site motifs and reaction intermediates that may be difficult to isolate in more conventional materials.[13]

1.1.2 Redox Activity

Historically, most experimentally synthesized MOFs have metals that cannot readily undergo changes in formal oxidation state. In these scenarios, the MOF is said to be redox-inactive. With respect to the inorganic nodes, most MOFs are composed of metals that generally cannot be oxidized or reduced, such as Zn(II), Cd(II), Ag(I), Al(III), Mg(II), Zr(IV), Al(III), and Sc(III) (in order of decreasing frequency of occurrence[14]). MOFs composed of Cu(II) ions, which are the second most common metals in MOFs after Zn(II),[14] are also rarely suitable for oxidation processes. For the purposes of charge transport, the commonly used carboxylate linkers often inhibit electron transfer between the metal nodes, and most experimentally synthesized MOFs are best-classified as electronic insulators with large band gaps.[15,16]

Recently, there has been significant interest in the design of redox-active MOFs,[15–19] as there are countless applications that are reliant on the transfer of electrons, such as the selective binding of reducible guest molecules, catalytic oxidation reactions, electrocatalysis, and energy storage processes. For reactions that involve a change in oxidation state at the metal center, MOFs with metals that have more easily accessible redox couples have been synthesized, including (but not limited to) MOFs with redox-active Ti(III), V(III), Cr(II), Mn(II), Fe(II), Co(II), Ni(II), and Cu(I) cations.[20–27] In some cases, the linkers can exhibit redox non-innocent behavior and undergo a change in oxidation state, even if the reaction of interest takes place at the metal center.[28] For (opto)electronic applications, MOFs have also been synthesized with redox-active linkers as well as guest molecules that can facilitate intraframework charge transport.[16]

1.2 High-Throughput Computational Screening

With the development of robust computational screening workflows,[29–32] high-throughput quantum-chemical calculations have been used to investigate the electronic, energetic, and structural properties of hundreds of thousands of materials over the last decade.[33–39] Especially when coupled with recent advances in machine learning, high-throughput computational methods have made it possible to discover promising new materials *in silico*[40–43] as an alternative to discovery based solely on intuition, trial-and-error, or serendipity. The need for a high-throughput approach to materials discovery is particularly paramount when the composition space is combinatorial in nature, as the number of unique materials to consider will often far exceed the number that can be studied via the more conventional hypothesis-synthesize-test cycle.

At the time of writing, tens of thousands of MOFs have been experimentally characterized,[14,44] and virtually unlimited more can be proposed.[45] Given the vast chemical space of MOF structures, high-throughput computational screening approaches have proven to be extremely useful for accelerating the discovery of promising new MOFs.[46–52] Most studies in this area are based on classical simulations, most notably grand canonical Monte Carlo (GCMC) simulations that can be used to model gas adsorption in MOFs. High-throughput GCMC simulations have been successfully used to identify top-performing MOFs with respect to CH_4 storage capacity,[53] CO_2 capture,[54] H_2 storage,[55] and O_2 uptake,[56] among several other applications.[48] Extending high-throughput computational screening of MOFs to applications requiring quantum chemical calculations, however, remains a significant challenge and has been largely unexplored in comparison.

1.3 Density Functional Theory

1.3.1 Hohenberg-Kohn Theorems and the Kohn-Sham Equations

Currently, the most widely used computational method to predict the electronic and energetic properties of MOFs is density functional theory (DFT).[57] In short, DFT is based on two key fundamental principles derived by Hohenberg and Kohn in 1964.[58] The first Hohenberg-Kohn theorem states that the ground-state energy E obtained from the Schrödinger equation is a unique functional of the system's electron density $\rho(\mathbf{r})$, which itself encodes the probability that any of the N electrons in the system are at a given position $\mathbf{r}$. The second Hohenberg-Kohn theorem states that the ground-state electron density is the one that minimizes the aforementioned energy functional, $E[\rho(\mathbf{r})]$. Collectively, these two theorems imply that if one can compute the ground-state electron density, then the energy – in addition to the wavefunction and all derivable quantum-mechanical properties of the system – can also be readily obtained.

On their own, the Hohenberg-Kohn theorems do not provide an answer as to how the ground-state electron density can be computed. Within the framework of DFT, this can be directly addressed using the so-called Kohn-Sham equations,[59] which transform the original many-body problem of electrons into one based on an artificial system of non-interacting electrons with the same density as the true system. In the Kohn-Sham formalism, the ground-state electron density is expressed in terms of an eigenvalue equation involving single-electron wavefunctions of non-interacting electrons $\psi_i(\mathbf{r})$ and energies of the corresponding electron orbitals ε_i, given by

$$\left[-\frac{\hbar^2}{2m_\mathrm{e}}\nabla^2 + V(\mathbf{r}) + V_\mathrm{H}(\mathbf{r}) + V_\mathrm{XC}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}). \tag{1.1}$$

The first term on the left-hand side represents the kinetic energy of the electron, where $\hbar$ is the reduced Planck's constant, m_{e} is the mass of the electron, and ∇^2 is the Laplacian operator. The second term, $V(\mathbf{r})$, represents the potential energy of the electron-nuclei interactions. The third term, known as the Hartree potential $V_{\mathrm{H}}(\mathbf{r})$, is given by

$$V_{\mathrm{H}}(\mathbf{r}) = e^2 \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'. \tag{1.2}$$

where e is the elementary charge. The Hartree potential describes the repulsive Coulombic interaction between the electron and total electron density. The final term on the left-hand side of the Kohn-Sham equations is the exchange-correlation potential $V_{\mathrm{XC}}(\mathbf{r})$, which is given by

$$V_{\mathrm{XC}}(\mathbf{r}) \equiv \frac{\delta E_{\mathrm{XC}}[\rho(\mathbf{r})]}{\delta \rho(\mathbf{r})} \tag{1.3}$$

where $E_{\mathrm{XC}}[\rho(\mathbf{r})]$ is the exchange–correlation energy functional. The exchange-correlation potential accounts for the quantum-mechanical factors not defined elsewhere in the Kohn-Sham equations and is the term needed to make the system of N non-interacting electrons have the same electron density as N interacting electrons.

In practice, the Kohn-Sham equations are solved using an iterative approach, typically referred to as converging the self-consistent field (SCF). In this multi-step approach, an initial guess for the electron density is defined, denoted $\rho_{\mathrm{guess}}(\mathbf{r})$. With this trial electron density, the single-particle wavefunctions $\psi_i(\mathbf{r})$ are computed using Equation 1.1, from which the electron density $\rho(\mathbf{r})$ is calculated as

$$\rho(\mathbf{r}) = 2 \sum_i \psi_i^*(\mathbf{r}) \psi_i(\mathbf{r}). \tag{1.4}$$

If $\rho(\boldsymbol{r}) \neq \rho_{\text{guess}}(\mathbf{r})$, the procedure is repeated with a new trial electron density until the provided electron density and that computed from Equation 1.4 are equal (within some numerical tolerance). Once the SCF is considered converged, the electron density has been obtained. Returning to the many-body system of interest, the newly obtain electron density can be used to calculate the total energy and the resulting quantum-mechanical properties of interest. The total energy expression for the many-body system is given by

$$E[\rho(\mathbf{r})] = -\frac{\hbar^2}{2m_{\text{e}}}\sum_i \int \psi_i^*(\mathbf{r})\,\nabla^2\psi_i(\boldsymbol{r})\,d\mathbf{r} + \int V(\mathbf{r})\rho(\mathbf{r})\,d\mathbf{r} + \frac{e^2}{2}\iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}d\mathbf{r}' + E_{\text{XC}}[\rho(\mathbf{r})]. \quad (1.5)$$

1.3.2 Density Functional Approximations

The key ingredient missing from the aforementioned discussion is the precise definition of the exchange-correlation functional, $E_{\text{XC}}[\rho(\mathbf{r})]$. While the true exchange-correlation functional is unknown, various density functional approximations (DFAs) have been developed to solve the Kohn-Sham equations and make DFT a practical method for computing the electronic structure properties of a given chemical system.[60,61]

At the time of writing, there are approximately 600 exchange-correlation functionals incorporated in the popular Libxc library,[62] and many more are likely to be developed in the forthcoming years. The various density functionals are often classified by what factors they rely on in the calculation of $E_{\text{XC}}[\rho(\mathbf{r})]$. In order of increasing computational cost, there is the local density approximation (LDA)[58] that depends only on the local electron density $\rho(\mathbf{r})$; generalized gradient approximation (GGA) functionals[63] that depend on the local density and its gradient $\nabla\rho(\mathbf{r})$; meta-GGA functionals[64] that depend on the density, its gradient, and its Laplacian $\nabla^2\rho(\mathbf{r})$ (or the non-interacting kinetic energy density); and hybrid (meta-)GGA functionals,[65] which

incorporate some fraction of exact-exchange from Hartree-Fock theory[66,67] in addition to the factors described for GGA or meta-GGA functionals.[60] Higher "rungs" of this ladder of density functionals are often associated with increasingly accurate quantum-mechanical simulations.[61] In practice, however, there is rarely a clear consensus on the optimal density functional for a given application, and there are many factors to consider beyond a functional's performance on a particular benchmark dataset.[68] One the most important aspects to consider is that there is often a strong tradeoff between the computational cost and presumed accuracy of a given density functional. For instance, while hybrid meta-GGA functionals are routinely used in the study of organic molecules, they can often be computationally prohibitive for materials with many valence electrons, including some MOFs.

1.3.3 Periodic and Cluster Calculations

When using DFT to model the properties of MOFs, an important decision that must be made is whether to model the structure using its full crystalline structure or a molecular analogue.[4] As with most aspects of computational chemistry, there are benefits and tradeoffs associated with both approaches.

When modeling the full crystallographic unit cell with periodic boundary conditions, the simulated material often closely matches that being studied experimentally, excluding possible crystalline defects. Unfortunately, the relatively high computational cost associated with periodic DFT calculations of MOFs means that state-of-the-art density functionals, particularly those in the hybrid(-meta) GGA class, can be computationally intractable for frequent use. An alternative approach is to carve a representative molecular cluster from the crystalline structure that can be modeled with greater computational efficiency and permits the use of higher levels of theory.[69]

While not all properties can be appropriately modeled using a cluster model (e.g. the electronic band structure), there are many chemical phenomena that are local in nature and do not depend strongly on the behavior of atoms far away from the reactive site of interest.[70]

For the purposes of high-throughput computational screening, it should nonetheless be noted that cluster calculations involve many human-guided decisions that are difficult to automate. Given an arbitrary MOF structure, there are no robust algorithms that can systematically carve an appropriately sized cluster model that does not neglect important medium-to-long range electrostatic interactions or van der Waals interactions, such as pore-based confinement effects. Similarly, it can often be ambiguous as to where to best terminate the cluster, how to cap the resulting dangling bonds, and which atoms should be kept rigid to appropriately mimic the bulk constraints of the crystalline framework. Regardless of the approach taken, it is always important to understand the model's limitations to ensure that the reported trends and results can be used to make actionable suggestions for future experiments.

1.4 Outline of Research

1.4.1 Thesis Outline

Broadly, this thesis can be broken down into three main topical areas: 1) structure–reactivity relationships to help guide the discovery of MOFs capable of catalytically oxidizing strong C–H bonds; 2) design principles for MOFs that can selectively bind O_2 over N_2; 3) a database of MOF electronic structure properties that can be used to train and benchmark machine learning models to circumvent computationally expensive DFT calculations.

The original research portion of this thesis begins with Chapter 2, which lays out the computational framework for a fully automated, high-throughput periodic DFT screening

procedure tailored for MOF structures.[32] As a proof-of-concept, this workflow is used to screen a database of experimentally synthesized MOFs for the catalytic oxidation of methane. This theme is revisited and refined in Chapter 3, where a curated dataset of MOFs with coordinatively unsaturated metal centers is studied for the oxidation of methane to methanol. The main result of Chapter 3 is that there is often an indirect relationship between the thermodynamic stability of the terminal metal-oxo active site and its reactivity towards cleaving the C–H bond of methane.[71] The topic of C–H bond activation is continued in Chapter 4, which is focused on a highly tunable family of metal–triazolate frameworks (discussed again in Chapter 6 in the context of selective O_2 binding).[72] Chapter 4 introduces the concept of antiferromagnetically enhanced reactivity, where the alignment of the unpaired electrons at the metal center and terminal oxo ligand can greatly impact the predicted reactivity towards strong C–H bonds.

In Chapter 5, DFT-based computational screening is now used to study MOFs for the task of air separation via the selective binding of O_2 or N_2 at coordinatively unsaturated metal sites.[73] Through this systematic study, several design rules are uncovered. As one example, it is demonstrated both computationally and experimentally that exchanging the Cl^- bridging ligands in a family of metal–triazolate frameworks with more basic OH^- ligands can greatly increase the binding strength of O_2 at the metal centers without a substantial increase in the binding strength of N_2. Continuing on this topic, Chapter 6 takes a more methodological approach and is focused on comparing the behavior of common DFAs when studying redox-dependent adsorption of small molecules at the metal centers of MOFs.[74] By comparing the results of GGA, meta-GGA, and GGA+*U* simulations, the effects of electron self-interaction error are uncovered.

Up to this point, the aforementioned work has all involved small- to moderate-scale computational screening. In Chapter 7, a high-throughput DFT workflow is now used to develop a freely accessible, large-scale database of MOF electronic structure properties called the Quantum MOF (QMOF) Database.[75] The data underlying the QMOF Database is then used to train and benchmark various machine learning models for the rapid prediction of MOF band gaps without the need to carry out computationally expensive DFT calculations. This theme is continued in Chapter 8, which describes an augmented QMOF Database with computed electronic structure properties at multiple levels of theory. This new data is then used to better understand the fundamental limitations of commonly employed DFAs and can, in the future, enable the prediction of high-fidelity electronic structure properties at a small fraction of the computational cost otherwise needed to carry out hybrid-level DFT calculations.

In Chapter 9, the themes discussed throughout this thesis are put in the context of the recently published literature to provide a forward-looking perspective on the many challenges and opportunities within the area of data-driven MOF catalyst design and discovery. In Chapter 10, the thesis concludes with a brief summary of the key takeaways from the presented work.

A list of the publications and manuscripts in preparation that make up the chapters of this theses is included below:

- Chapter 2: **A.S. Rosen**, J.M. Notestein, R.Q. Snurr. "Identifying Promising Metal−Organic Frameworks for Heterogeneous Catalysis via High-Throughput Periodic Density Functional Theory." *J. Comput. Chem.*, 40, 1305–1318 (2019).
- Chapter 3: **A.S. Rosen**, J.M. Notestein, R.Q. Snurr. "Structure−Activity Relationships that Identify Metal−Organic Framework Catalysts for Methane Activation." *ACS Catalysis*, 9, 3576–3587 (2019).

- Chapter 4: **A.S. Rosen**, J.M. Notestein, R.Q. Snurr. "High-Valent Metal-Oxo Species at the Nodes of Metal–Triazolate Frameworks: The Effects of Ligand Exchange and Two-State Reactivity for C–H Bond Activation." *Angew. Chem. Int. Ed.*, 59, 19494–19502 (2020).
- Chapter 5: **A.S. Rosen**, M. Rasel Mian, T. Islamoglu, H. Chen, O.K. Farha, J.M. Notestein, R.Q. Snurr. "Tuning the Redox Activity of Metal−Organic Frameworks for Enhanced, Selective O_2 Binding: Design Rules and Ambient Temperature O_2 Chemisorption in a Cobalt−Triazolate Framework." *Journal of the American Chemical Society*, 142, 4317–4328 (2020).
- Chapter 6: **A.S. Rosen**, J.M. Notestein, R.Q. Snurr. "Comparing GGA, GGA+*U*, and Meta-GGA Functionals for Redox-Dependent Binding at Open Metal Sites in Metal–Organic Frameworks." *Journal of Chemical Physics,* 152, 24101 (2020)
- Chapter 7: **A.S. Rosen**, S.M. Iyer, D. Ray, Z. Yao, A. Aspuru-Guzik, L. Gagliardi, J.M. Notestein, R.Q. Snurr. "Machine Learning the Quantum-Chemical Properties of Metal–Organic Frameworks for Accelerated Materials Discovery." *Matter*, 4, 1578–1597 (2021)
- Chapter 8: **A.S. Rosen**, J.M. Notestein, R.Q. Snurr. "Predicting the Electronic Properties of Metal–Organic Frameworks with Big Data: Probing the Limits of Density Functional Theory." *In preparation.*
- Chapter 9: **A.S. Rosen**, J.M. Notestein, R.Q. Snurr. "Realizing the Data-Driven, Computational Discovery of Metal–Organic Framework Catalysts." *In preparation*.

1.4.2 Other Projects

In addition to the research outlined in Chapters 2–10, I have had the privilege to work on several other projects over the course of my graduate studies, which are briefly summarized below.

One of my first projects at Northwestern University was the use of *ab initio* free energy calculations to generate comprehensive phase diagrams of the catalytically relevant edge sites of MoS_2.[76] I have also been involved with several collaborations. The first of these collaborations involved a DFT-based investigation of the plausible active site species in a SiO_2-supported CuO_x catalyst.[77] In a separate collaboration, I used computational methods to help understand the packing structure of Ce-oxo clusters.[78] Currently, I am also mentoring an undergraduate student,

Shaelyn Iyer, on a project involving high-throughput computational screening of monocopper transition metal complexes for their ability to form terminal metal-oxo/metal-oxyl sites.

On the topic of MOFs, I helped develop the Python interface for the MOFid code, which can be used to provide a unique identifier for MOFs to aid search and cheminformatics analyses.[79] I have also used DFT calculations to better understand the relative energetic stabilities of various MOF polymorphs[80] and to probe the reactivity of proposed active site motifs in Zr-MOFs for the dissociation of H_2.[81] In an ongoing project not included in this thesis, I have explored the formation and reactivity of various bridging and terminal metal-oxo species in a family of metal–triazolate frameworks for the activation of strong C–H bonds.

While not discussed in this thesis, the aforementioned papers and manuscripts in preparation are listed below for reference:

- **A.S. Rosen**, J.M. Notestein, and R.Q. Snurr. "Comprehensive Phase Diagrams of MoS_2 Edge Sites Using Dispersion-Corrected DFT Free Energy Calculations." *Journal of Physical Chemistry C*, 122, 15318–15329 (2018).
- S.L. Nauert, **A.S. Rosen**, H. Kim, R.Q. Snurr, P.C. Stair, J.M. Notestein. "Evidence for Copper Dimers in Low-Loaded CuO_x/SiO_2 Catalysts for Cyclohexane Oxidative Dehydrogenation." *ACS Catalysis*, 87, 9775–9789 (2018).
- B.J. Bucior, **A.S. Rosen**, M. Haranczyk, Z. Yao, M.E. Ziebel, O.K. Farha, J.T. Hupp, J.I. Siepmann, A. Aspuru-Guzik, R.Q. Snurr. "Identification Schemes for Metal−Organic Frameworks to Enable Rapid Search and Cheminformatics Analysis." *Crystal Growth and Design*, 9, 6682–6697 (2019).
- N.S. Bobbitt, **A.S. Rosen**, R.Q. Snurr. "Topological Effects on Separation of Alkane Isomers in Metal–Organic Frameworks." *Fluid Phase Equilibria*, 519, 112642 (2020).
- M.C. Wasson, X. Zhang, K. Otake, **A.S. Rosen**, S. Alayoglu, M.D. Krzyaniak, Z. Chen, L.R. Redfern, L. Robison, F.A. Son, Y. Chen, T. Islamoglu, J.M. Notestein, R.Q. Snurr, M.R. Wasielewski, O.K. Farha. "Supramolecular Porous Assemblies of Atomically Precise Catalytically Active Cerium-Based Clusters." *Chemistry of Materials*, 32, 8522–8529 (2020).

- K.E. Hicks, **A.S. Rosen**, Z.H. Syed, R.Q. Snurr, O.K. Farha, J.M. Notestein. "Zr_6O_8 Node-Catalyzed Butene Hydrogenation and Isomerization in the Metal–Organic Framework NU-1000." *ACS Catalysis*, 10, 14959–14970 (2020).
- **A.S. Rosen**, J.M. Notestein, R.Q. Snurr. "Investigating Redox-Active Metal–Organic Frameworks with Terminal and Bridging Metal-Oxo Species for the Activation of Strong C–H Bonds." *In preparation.*

Chapter 2: A High-Throughput Periodic Density Functional Theory Workflow to Identify Promising Metal−Organic Frameworks

This chapter describes the development of a fully automated, high-throughput periodic density functional theory (DFT) workflow for screening promising MOF candidates, with a specific focus on applications in catalysis. As a proof-of-concept, the high-throughput workflow is used to screen MOFs containing open metal sites from the Computation-Ready, Experimental MOF database for the oxidative C–H bond activation of methane. The results from the screening process suggest that, despite the strong C–H bond strength of methane, the main challenge is identifying MOFs with open metal sites that can be readily oxidized at moderate reaction conditions.

This chapter is adapted from the following peer-reviewed article: A.S. Rosen, J.M. Notestein, R.Q. Snurr. "Identifying Promising Metal−Organic Frameworks for Heterogeneous Catalysis via High-Throughput Periodic Density Functional Theory." *J. Comput. Chem.*, **40**, 1305–1318 (2019).

2.1 Introduction

With the advent of materials informatics toolkits and software enabling the development of high-throughput (HT) screening workflows,[31,82–87] HT periodic density functional theory (DFT) has been used to construct databases of electronic, energetic, and structural properties for hundreds of thousands of inorganic materials from first-principles calculations.[29,30,33,35,38,40,88] Metal−organic frameworks (MOFs), a novel class of highly porous crystalline materials, are well-suited for computational screening studies due to the modular nature of their inorganic nodes and organic linkers.[53] To date, there are tens of thousands of experimentally synthesized MOFs,[44] and numerous crystal structure databases have been developed from both known[14,44,89] and hypothetical[53,90] MOF structures.

HT screening of these databases is particularly appealing, as it allows for a greater number of MOFs to be investigated than would be possible experimentally. One of the main goals of HT screening of MOF crystal structure databases is to reduce the time-to-discovery of MOFs with desired chemical and physical properties. In the area of gas storage and separations, HT grand canonical Monte Carlo simulations using classical force fields have been successfully used to identify top-performing MOFs with respect to CH_4 storage capacity,[53] CO_2 capture,[54] H_2 storage,[55] and O_2 uptake.[56] However, extending HT screening of MOFs to applications requiring quantum chemical calculations, such as catalysis, remains unexplored.

Due in part to the large unit cells of many MOFs, the most common approach when modeling MOFs for any catalytic reaction is to crop and terminate the periodic structure to create a finite-sized cluster model of the proposed active site, often consisting of no more than a few dozen atoms.[69] An appropriate choice of where to artificially terminate the MOF unit cell is often

not immediately obvious, and this approach is therefore not amenable to HT screening of MOFs with widely varying structures and topologies. Finite cluster models can also introduce artificial boundary effects that have the potential to influence charge delocalization[17] and pore-based confinement effects.[91,92] The use of periodic DFT to represent the full crystallographic unit cell naturally resolves these issues, and most implementations of periodic DFT are well-suited for highly parallel calculations that can be used to treat the larger number of atoms in each simulation. Nevertheless, to the best of our knowledge, there has never been a fully automated, HT periodic DFT screening study for any catalytic reaction using MOFs.

In this work, we have developed a HT workflow based on periodic DFT to screen large numbers of MOFs for promising catalytic candidates. As a proof-of-concept, we use this HT workflow to screen MOFs with coordinatively unsaturated metal sites, also known as open metal sites (OMSs), for oxidative C–H bond activation. Due to the large economic demand for a catalyst that can directly convert methane to methanol[93] and motivated by prior work involving MOFs for methane conversion,[94–99] we specifically consider the partial oxidation of methane as the reaction of interest. In the process, we demonstrate the feasibility of a HT-DFT screening workflow for MOF catalysis and make several recommendations for future work involving HT-DFT screening studies of MOFs.

2.2 Methods

2.2.1 General Scheme

Motivated by similar schemes for computational catalysis screening of bulk metals and alloys,[100,101] the general approach for HT-DFT of MOF catalysts can be outlined as follows:

(1) Determine a set of catalytic descriptors that can be used to correlate catalytic activity with readily computed quantities, such as those based on adsorption energies, reaction energies, or electronic structure properties.[102–107]

(2) Identify or construct a dataset of MOF crystal structures to study. Tools such as Pymatgen,[82,108] Zeo++,[109] and PoreBlazer[110] can be used to select MOFs with specific metals, coordination environments, and pore sizes relevant to the given reaction of interest.

(3) Using DFT, optimize the unit cell volume, unit cell shape, and atomic positions for each MOF.

(4) Starting from the optimized MOF structures, initialize the positions of atomic and molecular adsorbates required to predict catalytic activity via (1).

(5) Using DFT, relax the atomic positions of the structures generated via (4).

(6) Compute the catalytic descriptors of interest to rank MOF candidates.

(7) For promising MOF candidates, generate the potential energy landscape for the proposed mechanism and perform detailed electronic structure analyses to better understand the reaction profile.

In the following sections, we describe how to fully automate this process when studying MOFs with spatially isolated active sites, given previously determined catalytic descriptors. We then apply this procedure to screen MOFs for the ability to oxidatively activate the C–H bond of methane.

2.2.2 Catalytic Descriptors for Oxidation C–H Bond Activation

When studying materials such as MOFs with spatially isolated active sites, the expected mechanism for the conversion of methane to methanol is the radical-rebound mechanism, as shown

in Figure 2.1.[111] For HT screening purposes, we focus on the oxidation of the metal and H-abstraction steps of this mechanism, as they dictate the overall conversion of methane by influencing the number of metal oxide active sites and the activity of each site, respectively. Recently, Nørskov and coworkers have proposed a universal transition state (TS) scaling relationship for the C–H bond activation of methane.[105] This TS scaling relationship can be used to accurately predict the energy of the corresponding TS over a wide range of heterogeneous catalysts based on the strength that an H atom binds to the metal oxide active site, denoted E_{H}. This H-affinity linear scaling relationship has been benchmarked for a wide range of materials, including cation-exchanged zeolites, bulk metal oxides, transition metal surfaces, and MOFs.[105]

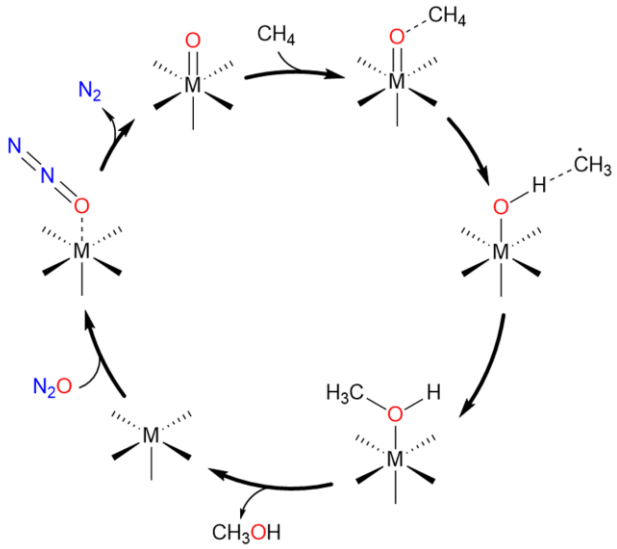


Figure 2.1. Radical-rebound mechanism for the oxidative C–H bond activation of methane at a coordinatively unsaturated metal site (M) using an N_2O oxidant. The precise coordination environment and M–O bond order are dependent on the given MOF. In this work, we mainly focus on oxidation of the metal center and subsequent H-abstraction.

Here, we define the H-affinity as

$$E_{\mathrm{H}} = E_{\mathrm{MOF-OH}} - E_{\mathrm{MOF-O}} - \frac{1}{2}E_{\mathrm{H_2}} \tag{2.1}$$

where $E_{\mathrm{MOF-OH}}$ and $E_{\mathrm{MOF-O}}$ are the electronic energies of the metal site with adsorbed OH and O species, respectively, and $E_{\mathrm{H_2}}$ is the electronic energy of H_2. Note that unlike the original description of E_{H},[105] Equation 2.1 uses H_2 as the H-reference, as justified in the Appendix. With this, we can use the universal scaling relationship[105] of

$$E^*_{\mathrm{TS,C-H}} = 0.75E_{\mathrm{H}} + 1.96\ \mathrm{eV} \tag{2.2}$$

to predict the energy of the TS, $E^*_{\mathrm{TS,C-H}}$, with respect to the initial oxidized state, as schematically illustrated in Figure A.1. With Equations 2.1 and 2.2, the C–H bond activation barrier can then be readily computed via

$$E^*_{\mathrm{a,C-H}} = E^*_{\mathrm{TS,C-H}} - \left[E_{\mathrm{MOF-O-CH_4}} - \left(E_{\mathrm{MOF-O}} + E_{\mathrm{CH_4}}\right)\right] \tag{2.3}$$

where $E_{\mathrm{MOF-O-CH_4}}$ is the electronic energy of methane adsorbed to the metal oxide active site, and $E_{\mathrm{CH_4}}$ is the electronic energy of gas-phase methane. The asterisks in Equations 2.2 and 2.3 are used to denote quantities obtained from a TS scaling relationship rather than directly computed using a TS finding algorithm; this difference is expected to be about ~10 kJ/mol, on average, based on the work of Nørskov and coworkers.[105]

In contrast with the C–H bond activation step, there is currently no reported universal TS scaling relationship for the step in which the metal site is oxidized. Instead, we consider the extrinsic oxidation reaction energy using N_2O as the proposed oxidant, defined as

$$\Delta E_{\mathrm{ox}} = \left(E_{\mathrm{MOF-O}} + E_{\mathrm{N_2}}\right) - \left(E_{\mathrm{MOF}} + E_{\mathrm{N_2O}}\right) \tag{2.4}$$

to determine the thermodynamic favorability of oxidation, as has been done in prior work.[105,106] Here, E_{MOF}, $E_{\mathrm{N_2}}$, and $E_{\mathrm{N_2O}}$ are the electronic energies of the guest-free MOF, gas-phase N_2, and

gas-phase N_2O, respectively. We refer to Equation 2.4 as an "extrinsic" reaction energy, as the energies for N_2 and N_2O are for the isolated gas-phase species, not in the adsorbed state.

2.2.3 Initial Dataset Construction

The starting dataset considered in this work is an 838 MOF subset[112] of the Computation-Ready, Experimental (CoRE) MOF database.[89] The MOFs in this database were originally obtained from the Cambridge Structural Database (CSD) with free and bound solvents removed. The 838 MOFs in this database have previously been optimized at the PBE-D3(BJ) level of theory via the CP2K code,[113] which uses mixed Gaussian and plane-wave basis sets.[114] From this 838 MOF dataset, 168 unique MOFs were selected for analysis based on a high likelihood of having OMSs following optimization and having pore-limiting diameters of at least 3.0 Å as determined by Zeo++[109] (refer to the Appendix for additional details).

2.2.4 Adsorbate Initialization

One of the necessary aspects of any HT screening workflow is that the entire process should be fully automated with minimal user-intervention. For heterogeneous catalysis and other surface science applications, adsorption energies are commonly computed but typically involve the user manually specifying an initial guess for the location of the adsorbate that is later optimized using DFT. This inherently limits the number of materials that can be screened, especially since computational catalysis screening often involves the calculation of numerous adsorption energies per catalytic candidate. In addition, the choice of a good initial location for a given adsorbate can significantly reduce the length of the corresponding geometry optimization by ensuring the structure is relatively close to a minimum in the potential energy surface. We summarize methods to automate this adsorbate initialization process below.

An easy-to-implement set of geometrical rules, which we refer to as the vector-sum method, can be used to systematically locate physically relevant adsorption sites on undercoordinated atoms for the purposes of HT-DFT calculations. As an example, consider the chemisorption of a single O atom to an unsaturated Cu site at the node of the widely studied material HKUST-1[115] (Figure 2.2a). For simplicity, we define the binding site (in this case, a Cu atom) as species $i = 0$, and all atoms in the first coordination sphere are given indices $i = 1$ through n, where n is the coordination number. The proposed metal binding site can be automatically determined using one of a variety of OMS detection algorithms.[14,109,116] A distance vector, $\boldsymbol{r}_{0,i}$, is then computed for each coordinating atom via

$$\boldsymbol{r}_{0,i} = \boldsymbol{p}_0 - \boldsymbol{p}_i \tag{2.5}$$

where $\boldsymbol{p}_i$ is the position vector of atom i. Note that $\boldsymbol{r}_{0,i}$ should be the minimum-image distance vector, taking into account the periodicity of the unit cell.

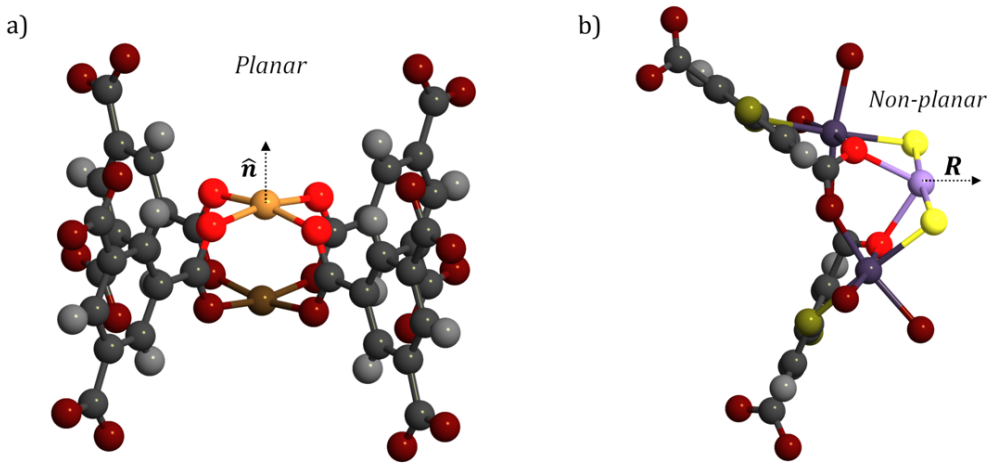


Figure 2.2. (a) Initializing the position of adsorbates at undercoordinated atoms with planar coordination environments is done in the direction of the unit normal vector to the plane with the fewest neighbors, $\hat{\boldsymbol{n}}$. (b) For non-planar coordination environments, the adsorbate is initialized in the direction resulting from the sum of the normalized distance vectors formed between the coordinating atoms and the central atom, $\boldsymbol{R}$.

We then determine the planarity of the coordinating atoms via two metrics. The first method involves total least-squares regression of the positions of the coordinating atoms to the equation of a plane, $ax + by + cz + d = 0$. As a second method, we calculate the sum of normalized distance vectors, $\boldsymbol{R}$, as

$$\boldsymbol{R} = \sum_{i=1}^{n} \frac{\boldsymbol{r}_{0,i}}{||\boldsymbol{r}_{0,i}||} \tag{2.6}$$

We suggest using both metrics as measures of planarity and, through iterative testing, have found that a coordination environment is typically well-described as planar if the root mean square error in the planar fit is less than 0.25 Å or $||\boldsymbol{R}|| < 0.25$ Å.

If the coordination environment is planar, the adsorbate's position $\boldsymbol{p}_{\text{ads}}$ is given by

$$\boldsymbol{p}_{\text{ads}} = \boldsymbol{p}_0 \pm \alpha \hat{\boldsymbol{n}} \tag{2.7}$$

where $\hat{n}$ is the unit normal vector to the best-fit plane and α is a bond distance scale-factor that is dependent on the proposed adsorbate. For instance, in this work we use $\alpha = 2$ Å for the chemisorption of an O atom to an OMS. To determine the sign in Equation 2.7, we calculate the number of neighbors within a cutoff distance r_{cut} and choose the direction with fewer neighboring atoms. In this work, we use a value of $r_{\text{cut}} = 2.5$ Å. The choice of sign is important for many MOFs, especially those with paddlewheel secondary building units such as HKUST-1.[115] If the coordination environment is not planar, the desired adsorption site p_{ads} is given by

$$\boldsymbol{p}_{\text{ads}} = \boldsymbol{p}_0 + \alpha \boldsymbol{R} \tag{2.8}$$

Collectively, this approach attempts to maximize the symmetry of the molecular geometry, such as the formation of a trigonal bipyramidal geometry from the original seesaw structure shown in Figure 2.2b. Naturally, a modified approach is needed for coordination numbers of three or less, which we describe in the Appendix along with the method for determining the atoms within the first coordination sphere. While the example shown in Figure 2.2 is for a monatomic adsorbate, this procedure can be readily extended for the adsorption of small molecules as well.[117]

To initialize the position of molecular adsorbates, a different method based on a molecular mechanics-based potential energy grid (PEG) generated for each MOF can be used to identify physically plausible adsorbate positions. In this work, we consider the adsorption of CH_4 near the metal oxide active site of each MOF following oxidation of a single metal center. The PEG for each MOF is calculated using a single-site TraPPE[118] CH_4 probe in RASPA[119] with a grid discretization of 0.1 Å in x, y, and z (Figure 2.3). The lowest energy position within a cutoff distance $r_{\mathrm{cut,PEG}}$ from the proposed active site is taken as the initial position for the adsorbate. Since we used a single-site CH_4 probe, the central C atom was placed at the lowest energy site within $r_{\mathrm{cut,PEG}} = 3.0$ Å, and one of the four H atoms in CH_4 was made colinear with the O atom at the active site and the central C atom of CH_4. The remaining three H atoms are arranged to satisfy the tetrahedral geometry of CH_4.

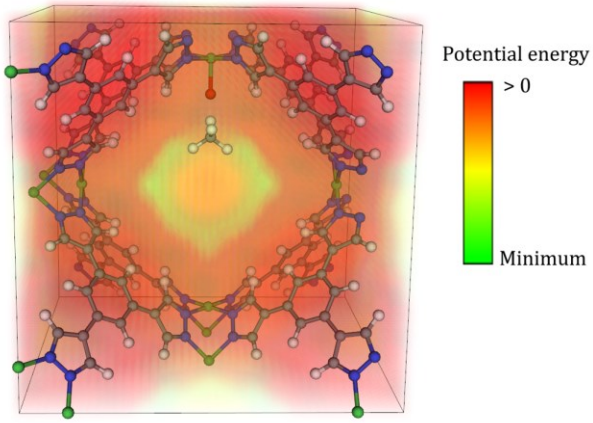


Figure 2.3. Potential energy grid obtained using a single-site CH_4 probe in an example MOF with a hypothetical $[NiO]^{2+}$ active site. The position of the CH_4 molecule has been initialized in the low-energy adsorption site. The repulsive regions are colored in dark red, and the most energetically favorable region is shown in light green. Color key: Ni (green), O (red), N (blue), C (gray), H (white).

2.2.5 Implementation of Methods

Additional details regarding the implementation of the adsorbate initialization algorithms used in this study can be found in the Appendix. An open-source Python code referred to as the MOF Adsorbate Initializer (MAI) is made publicly available to readily perform fully automated adsorbate initialization workflows such as those described in this work.[117] This code makes use of Pymatgen[82] and the Atomic Simulation Environment (ASE)[83] to carry out the adsorbate initialization process and has optional interfaces to Zeo++[109] or Open Metal Detector[14,120] for the automated detection of OMSs in MOFs. For PEG-based adsorbate initialization, the code is compatible with multiple volumetric data formats for the energy as a function of (x, y, z) coordinates, including PEGs computed from RASPA[119] or PorousMaterials.jl.[121].[a]

[a] A related code for larger adsorbates, deemed the MOF Big Adsorbate Initializer (MBAI),[625] was developed by Haoyuan Chen and can be used for larger adsorbates than those studied in the present chapter

2.2.6 High-Throughput Density Functional Theory

Given that the unit cells of MOFs can contain hundreds of atoms and that multiple reaction energies and kinetic barriers are needed to predict catalytic activity, it is essential that a robust and efficient HT-DFT workflow is established. In this section, we describe the multi-stage optimization scheme used to successfully perform the calculations in this work (~500 DFT calculations on materials with an average of ~200 atoms per Niggli-reduced unit cell).

2.2.7 Constant Parameters

All periodic DFT calculations are performed with the Vienna *ab initio* Simulation Package (VASP) v.5.4.1[122,123] and the projector-augmented wave (PAW) pseudopotentials[124] outlined in Table A.1 of the Appendix. The electron exchange-correlation is described by the Perdew-Burke-Ernzerhof (PBE) functional,[125] and Grimme's D3 dispersion correction[126] with Becke-Johnson (BJ) damping[127] is used to account for van der Waals (vdW) dispersion interactions. We use the PBE functional as a reasonable balance between computational cost and accuracy for the purposes of HT screening.[128,129] Due to the electronically insulating nature of most MOFs, Gaussian smearing of the band occupancies with a small smearing width of σ = 0.01 eV is used prior to extrapolation to the 0 K limit. Transition state calculations are done using an automated procedure involving the climbing image nudged elastic band method[130] and the dimer method[131] as described in the Appendix.

2.2.8 Electronic Optimization

The default electronic optimization algorithm used in this work is a preconditioned conjugate gradient (CG) algorithm (also referred to as the "all bands simultaneous update of orbitals" algorithm), which is suggested for both large and insulating materials.[132,133] An additional benefit of this algorithm is that it is not heavily reliant on the choice of Pulay density mixing parameters,[134]

which are used to achieve convergence of the self-consistent field (SCF) but can be highly material-specific.[29,33,122,135] This is in contrast with electronic optimization routines more commonly used for modeling metallic systems (e.g. blocked Davidson, RMM-DIIS),[122] which generally require frequent monitoring of the SCF convergence and *post hoc* tweaking of the mixing parameters to resolve problematic convergence issues in a HT-DFT workflow.[29,33]

2.2.9 Multi-Stage Geometry Optimization

As is common in computational catalysis studies, the first step of the HT-DFT workflow involves optimizing the unit cell of each bare MOF, including the cell shape, cell volume, and all internal degrees of freedom. As shown in Figure 2.4**,** we performed a multi-step relaxation scheme inspired in part by the one used in constructing the Open Quantum Materials Database (OQMD).[30]

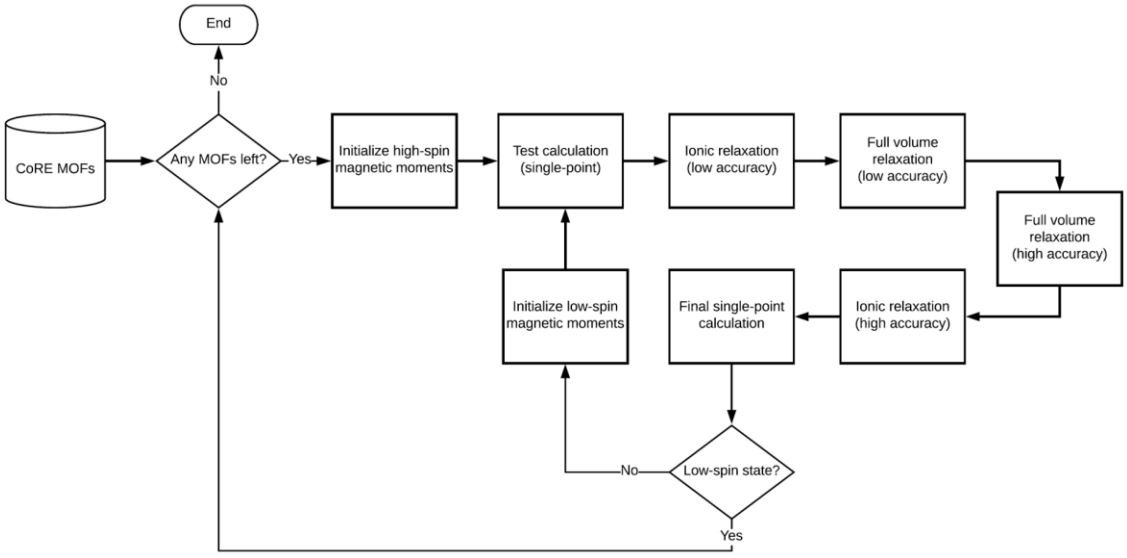


Figure 2.4. Schematic summarizing the high-throughput periodic density functional theory workflow for performing volume relaxations (including the cell parameters, cell shape, and internal degrees of freedom) of bare MOFs.

In the first step, a test single-point calculation is performed to confirm that there are no VASP-related errors. Any errors in this stage (e.g. too high a degree of parallelization requested

for a particularly small MOF) and throughout the workflow (e.g. convergence failures) are fixed on-the-fly using a procedure similar to the Custodian tool of the Materials Project.[29,33] Following the test calculation, a preliminary low-accuracy structure relaxation is performed at fixed cell shape and volume, which is necessary for the efficient geometry optimization of most MOFs due to the large number of atoms per unit cell. The default maximum plane-wave energy cutoff specified by the pseudopotentials (generally 400 eV) and a k-point density (KPPA) of 100 k-points/number of atoms was used, as arranged using Pymatgen.[82]

The initial relaxation of atomic positions is done in two main stages. In the first stage, we use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm with a line search mechanism that ensures the energy and absolute value of the force decrease monotonically. This algorithm is available in ASE[83] and is used until the maximum force on all individual atoms is less than 5 eV/Å. Compared to the default CG or quasi-Newton (QN) algorithms in VASP, we found that the BFGS line search algorithm was able to resolve high forces without atoms in the MOF unit cell moving to unphysical locations far away from their starting positions. Once the 5 eV/Å threshold is reached, the CG algorithm in VASP is used until the maximum force is less than 0.05 eV/Å.

Following the low-accuracy relaxation of atomic positions, a full volume relaxation (i.e. cell shape, cell volume, atomic positions) is performed at an initial low-accuracy setting. To prevent Pulay stresses,[136] the plane-wave kinetic energy cutoff is raised to 520 eV. All other parameters are unchanged. In the subsequent high-accuracy volume relaxation, the density of the k-point grid is increased from 100 KPPA to 1000 KPPA, and the force-tolerance is tightened to 0.03 eV/Å. A k-point grid density of 1000 KPPA is also currently used in the Materials Project[33] and has been previously used to model MOFs.[128]

A final high-accuracy relaxation of atomic positions is performed after the high-accuracy volume relaxation to confirm that the atomic positions are fully converged when the cell shape is kept fixed, and then a single-point calculation is performed to store the final wavefunction and charge density. Starting from the converged high-spin structure, a low-spin spin-initialization is considered. If performed, the lowest energy structure from the two initialized spin states is used for further stages of the HT screening workflow. The key DFT parameters are summarized in Table 2.1 (with all parameters listed in Table A.2).

Table 2.1. Periodic DFT parameters used in the optimizing guest-free MOFs. ENCUT is the plane-wave kinetic energy cutoff, KPPA is the k-point density (with the corresponding k-point grids generated using Pymatgen[82]), and $F_{\max}$ is the force-convergence criterion. The "Volume?" and "Positions?" columns specify if the volume and/or atomic positions were allowed to change.

Stage	Volume?	Positions?	ENCUT (eV)	KPPA	$F_{\max}$ (eV/Å)
1: Atomic positions (low-accuracy)	No	No	400	100	0.05
2: Volume (low-accuracy)	Yes	Yes	520	100	0.05
3: Volume (high-accuracy)	Yes	Yes	520	1000	0.03
4: Atomic positions (high-accuracy)	No	Yes	520	1000	0.03
5: Single-point (high-accuracy)	No	No	520	1000	N/A

The percent difference in computed cell volumes between the low-accuracy and high-accuracy volume relaxations is shown in Figure 2.5a. For the vast majority of tested MOFs, the cell volume is nearly identical at the low- and high-accuracy settings. As shown in Figure 2.5b, the deviation in cell volumes is most pronounced for MOFs with a small number of atoms per unit cell. While 100 KPPA and 1000 KPPA corresponds to the same k-point grid (consisting of just

the Γ-point) for 33% of the MOFs in this analysis, 90% of MOFs had a volume change of less than ±1% when going from the low- to high-accuracy settings.

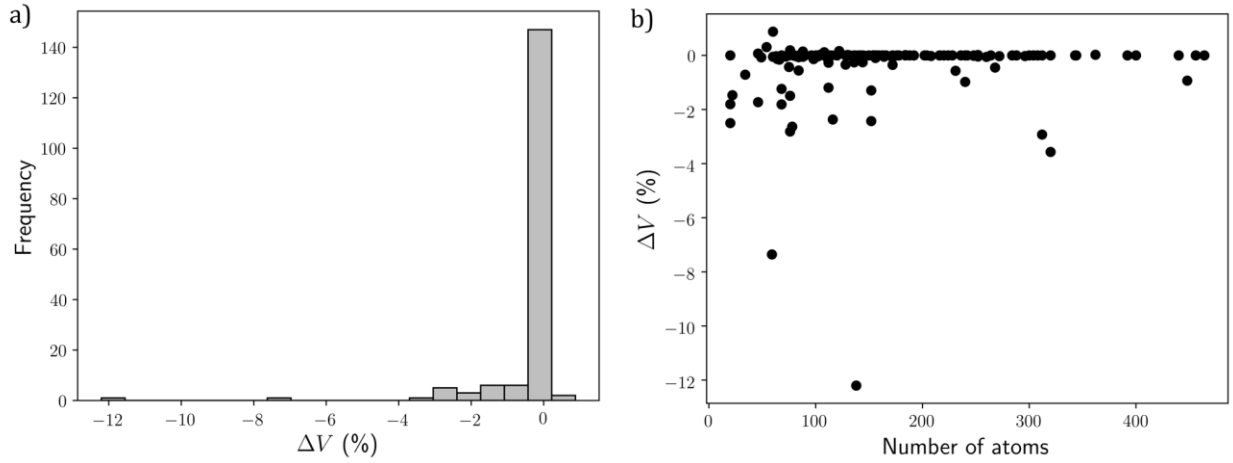


Figure 2.5. (a) Percent deviation in the cell volume, ΔV, computed with low-accuracy settings (cutoff = 520 eV, KPPA = 100, $F_{\max} < 0.05$ eV/Å) compared to high-accuracy settings (cutoff = 520 eV, KPPA = 1000, $F_{\max} < 0.03$ eV/Å), defined as $\Delta V = \left(V_{\text{high acc.}} - V_{\text{low acc.}}\right)/V_{\text{low acc.}}$. (b) Percent deviation in cell volumes as a function of the number of atoms per Niggli-reduced unit cell. Data shown in this figure is for the high-spin initialization cycle of the high-throughput workflow.

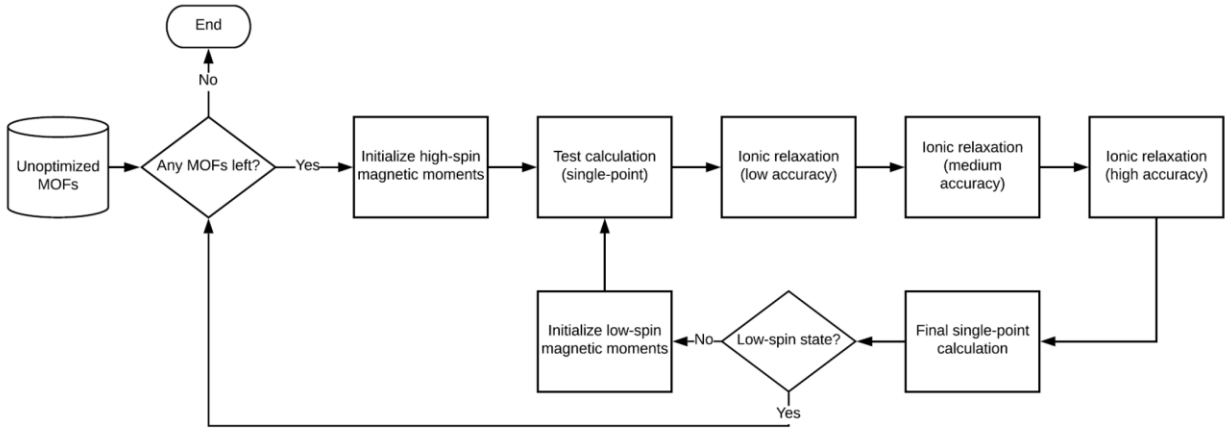


Figure 2.6. Schematic summarizing the high-throughput periodic density functional theory workflow for performing structure relaxations of MOFs with adsorbates.

The multi-stage optimization procedure for relaxing the atomic positions of MOFs with adsorbates (at fixed cell shape and volume) is summarized in Figure 2.6 and is similar to the

procedure for volume relaxations of the bare MOFs. At the medium-accuracy structure relaxation stage, the k-point density is increased from 100 KPPA to 1000 KPPA, but the default (400 eV) plane-wave kinetic energy cutoff is still used. For both the medium- and high-accuracy structure relaxations, the Fast Inertial Relaxation Engine (FIRE) algorithm[137] as implemented in VTST Tools[138] is used instead of the CG algorithm for relaxing the atomic positions. This was found to be necessary for optimizing the structures of MOFs with adsorbates, as the atoms far away from the adsorption site are already close to their optimized positions and the PES is very shallow, resulting in the CG algorithm failing to converge for nearly every MOF-adsorbate system studied in this work. The FIRE algorithm, which is a molecular dynamics method, was found to reach convergence in these otherwise problematic cases. This is consistent with the findings of Bitzek et al.[137] who demonstrated that the FIRE algorithm is both more efficient and robust than the typical CG and QN schemes when the forces are sufficiently low. For this reason, we also switch to the FIRE algorithm in any other part of the HT-DFT workflow if the default CG algorithm fails. The key DFT parameters are summarized in Table 2.2.

Table 2.2. Periodic DFT parameters used in the optimizing MOFs with adsorbates. ENCUT is the plane-wave kinetic energy cutoff, KPPA is k-point density (with the corresponding k-point grids generated using Pymatgen[82]), and $F_{\max}$ is the force-convergence criterion. The "Volume?" and "Positions?" columns specify if the volume and/or atomic positions were allowed to change.

Stage	Volume?	Positions?	ENCUT (eV)	KPPA	$F_{\max}$ (eV/Å)
1: Low-accuracy	No	Yes	400	100	0.05
2: Medium-accuracy	No	Yes	400	1000	0.05
3: High-accuracy	No	Yes	520	1000	0.03
4: Single-point	No	No	520	1000	N/A

Figure 2.7a and Figure 2.7b emphasize the value in using a multi-stage workflow, as the lower accuracy runs often result in nearly converged geometries and therefore greatly accelerate

the HT-DFT workflow. Since energy differences are typically desired for catalytic applications, we also computed the H-affinity previously shown in Equation 2.1 at the low-, medium-, and high-accuracy settings. Similar to what was shown in Figure 2.5, the results in Figure 2.7c emphasize the need to use 1000 KPPA over the much coarser 100 KPPA k-point density. However, as shown in Figure 2.7d, the increase in cutoff from 400 eV to 520 eV (and tighter force-convergence of 0.03 eV/Å compared to 0.05 eV/Å) does not significantly change the computed H-affinity beyond 1−2 kJ/mol on average, which is far below the expected accuracy from the choice of exchange-correlation functional and any TS scaling relationships. That being said, due to the nearly converged geometry at the end of the medium-accuracy run, the high-accuracy calculations exhibit rapid convergence (particularly when starting from the previously converged wavefunctions), so we decided to still include this step in the workflow.

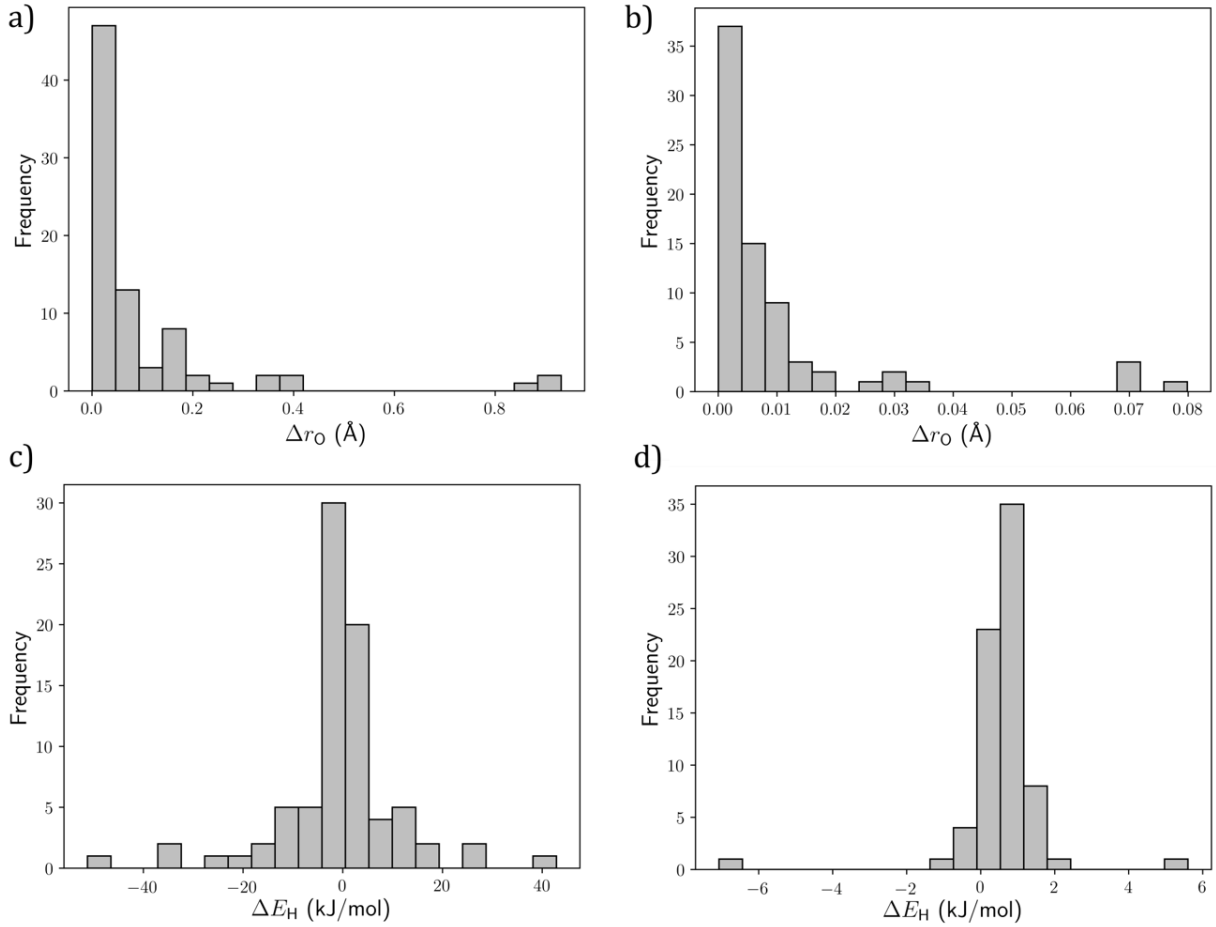


Figure 2.7. a) Deviation in the position of the O adsorbate, Δr_O, for the formation of the active site between the low-accuracy and medium-accuracy run. b) Deviation in Δr_O between the medium-accuracy and high-accuracy run. c) Deviation in H-affinity, ΔE_H, computed with low-accuracy settings (cutoff = 400 eV, KPPA = 100, $F_{max} < 0.05$ eV/Å) compared to medium-accuracy settings (cutoff = 400 eV, KPPA = 1000, $F_{max} < 0.05$ eV/Å), defined as $\Delta E_H = E_{H,med\ acc.} - E_{H,low\ acc.}$. d) ΔE_H computed with medium-accuracy settings and high-accuracy settings (cutoff = 520 eV, KPPA = 1000, $F_{max} < 0.03$ eV/Å), defined as $\Delta E_H = E_{H,high\ acc.} - E_{H,med\ acc.}$.

2.2.10 Spin-Polarization

To account for spin-polarization, both high-spin and low-spin initial magnetic moments are considered in a procedure motivated in part by the Materials Project[29,33] and OQMD.[30] In the high-spin case, d-block elements (i.e. Sc−Cu, Y−Ag, Hf−Au) are initialized in a high-spin state of 5 μ_B (Bohr magnetons), all f-block elements (i.e. La−Lu, Ac−Lr) are initialized with 7 μ_B, all metals

and semi-metals in groups 12−17 are initialized with 0.1 μ_B, and all s- and p-block metals are initialized with no spin. Once the high-spin run is completed, if the absolute values of the converged magnetic moment for each atom is less than 0.1 μ_B (or if there were no d- or f-block metals in the MOF), a low-spin configuration is not performed. Otherwise, a low-spin initial configuration is applied to the converged structure from the high-spin initialization, and the structure is reoptimized. For the low-spin case, the only difference is that the d- and f-block elements are initialized with 0.1 μ_B instead of 5 μ_B and 7 μ_B, respectively. If at the end of any step of the HT-DFT workflow the low-spin calculation converges to the magnetic moments that resulted from the high-spin calculation, the low-spin calculation is aborted. Two spin states are deemed to be equal if the converged magnetic moments are all within ±0.05 μ_B. While the high-spin initialization is sufficient for many of the MOFs studied in this work, the low-spin initialization can result in significantly more stable structures for select MOFs, as shown in Figure 2.8.

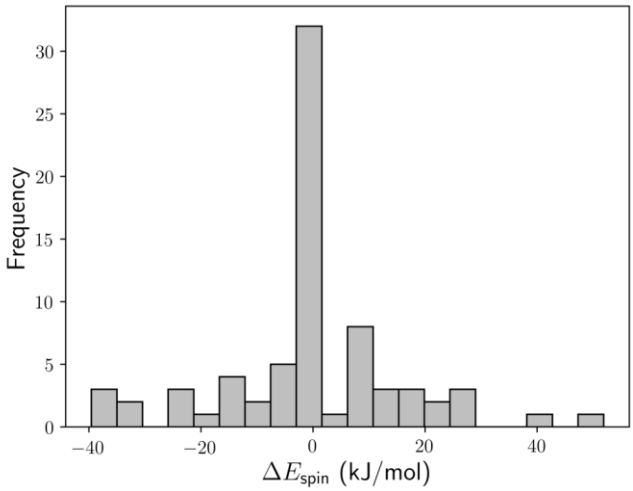


Figure 2.8. Difference in energy between the converged high-spin and low-spin electronic energies, $\Delta E_{spin} = E_{high\ spin} - E_{low\ spin}$, of MOFs where a full low-spin calculation was performed.

2.2.11 Implementation of Methods

Additional details regarding the implementation of the periodic DFT screening workflow used in this study can be found in the Appendix. An open-source Python code referred to as PyMOFScreen is made publicly available to readily perform fully automated periodic DFT calculations of MOFs using the related workflows to those described in this work.[139] As with the MOF Adsorbate Initializer, PyMOFScreen makes extensive use of Pymatgen[82] and ASE[83] to set up and carry out the DFT calculations. We note that analogous workflows could be readily constructed using the highly flexible Atomate package as well.[85]

2.3 Results and Discussion

With the developed HT screening workflow, we identified MOFs in an 838 MOF subset of the CoRE MOF database[112] that contained OMSs, performed volume relaxations on each MOF, initialized the positions of relevant adsorbates, and performed structure relaxations on each MOF with bound adsorbates to calculate adsorption energies. The volume relaxation (i.e. E_{MOF}) and three relaxations of just the atomic positions (i.e. $E_{\mathrm{MOF-O}}$, $E_{\mathrm{MOF-OH}}$, $E_{\mathrm{MOF-O-CH_4}}$) all achieved 99−100% convergence of the screened MOFs in this work, indicative of the robustness of the developed HT-DFT workflow. With the DFT-computed adsorption energies, we subsequently used Equations 2.1−2.4 to calculate the catalytic descriptors of interest. The results of the catalytic screening for oxidative C–H bond activation of methane are discussed below.

As shown in Figure 2.9a, the vast majority of screened MOFs have strongly endothermic oxidation reaction energies in the presence of an N_2O oxidant. Thermodynamically, this suggests that there will only be a small (potentially negligible) population of metal-oxo active sites present at moderate reaction conditions. Kinetically, since the lower-bound for the activation barrier is set

by the reaction energy for endothermic reactions, this also implies that the rate of N_2O activation is likely to be very low. As a result of this phenomenon, we expect that few MOFs in the 838 CoRE MOF subset have accessible metal sites that can be readily oxidized using N_2O.

Nonetheless, it is clear from Figure 2.9b that if a metal-oxo active site can be formed, the barrier for C–H activation would likely be sufficiently low for the catalytic conversion of methane for most MOFs considered in this work, as every MOF tested in this study has a methane C–H activation barrier below 100 kJ/mol, comparable to many cation-exchanged zeolites that can activate methane.[140] It is common in the computational catalysis literature to focus on the C–H activation step due to the large 440 kJ/mol bond dissociation energy of methane (when not in the presence of a catalyst).[141] However, the results shown in Figure 2.9 suggest that the main challenge in terms of high-throughput screening is not identifying MOFs with low C–H activation barriers, but rather identifying MOFs with OMSs that can be readily oxidized. Indeed, despite the many thousands of MOFs that have been synthesized to date, there are relatively few that have been experimentally shown to exhibit redox-active OMSs for catalytic applications.[17]

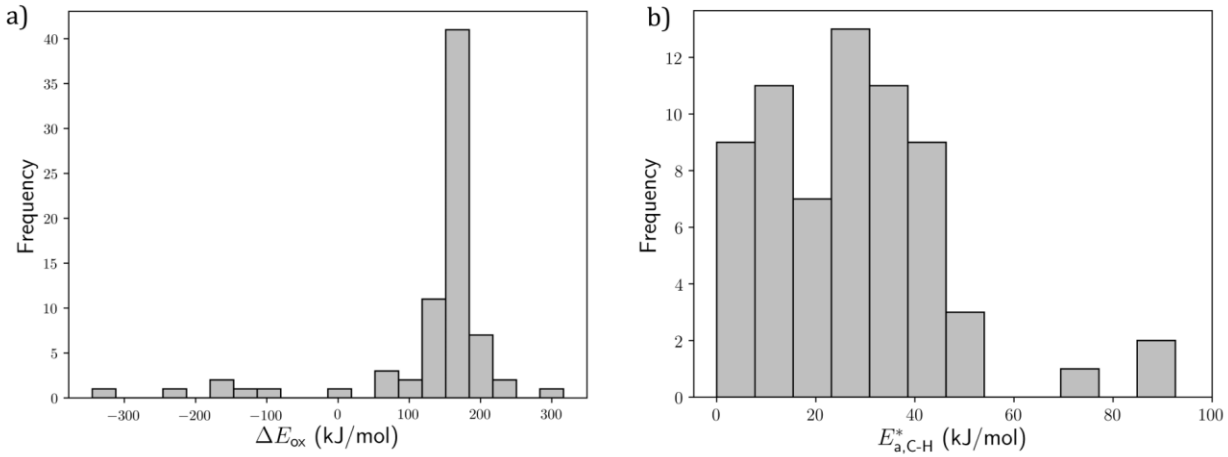


Figure 2.9. a) Reaction energy for N_2O activation, ΔE_{ox}, defined in Equation 2.4. b) Methane C–H activation barrier at metal oxide active site, $E^*_{a,C-H}$, defined in Equation 2.3.

As one example from the HT-DFT analysis, consider $Cu_3(dmtrz)_2(ox)_2$ (Hdmtrz = 3,5-dimethyl-1H-1,2,4-triazole, ox = oxalate, CSD refcode = LIFWEE)[142] shown in Figure 2.10. The structure of $Cu_3(dmtrz)_2(ox)_2$ has mono(μ-aquo)dicopper(II,II) sites that can be desolvated around 125 °C, leaving behind adjacent square-planar dicopper(II) sites.[142] In principle, the dicopper(II) species can be oxidized to form mono(μ-oxo)dicopper(III) species. This mono(μ-oxo)dicopper(III) species is extremely active toward the C–H activation of methane, with a predicted barrier of $E^*_{a,C-H} = 40$ kJ/mol. Despite this high reactivity toward H-abstraction, it can be expected that $Cu_3(dmtrz)_2(ox)_2$ would not be promising for oxidative C–H bond activation, as the computed reaction energy for formation of the mono(μ-oxo)dicopper(III) species is highly endothermic ($\Delta E_{ox} = 127$ kJ/mol).

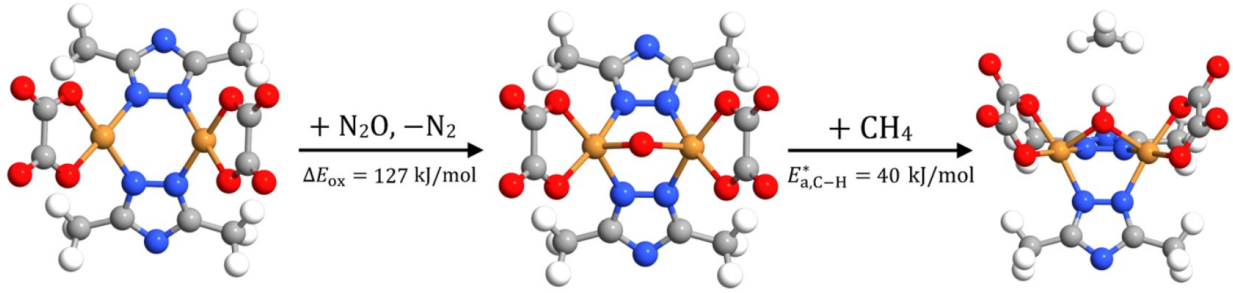


Figure 2.10. Formation of the metal oxide active site and subsequent H-abstraction of methane with $Cu_3(dmtrz)_2(ox)_2$. Color key: Cu (orange), C (gray), N (blue), O (red), H (white). Structures are DFT-optimized (only a representative portion of the unit cell is shown for clarity).

One approach to increase the likelihood of accessing low C–H bond activation barriers with MOFs containing OMSs is to increase the temperature during the oxidation step. This is a common and often necessary approach in the Cu-exchanged zeolite literature, where a multi-step conversion process is performed with O_2 or N_2O activation around 450 °C and subsequent methane

activation around 125–200 °C.[143–146] However, applying this approach to most MOFs is likely to be challenging, as the majority of MOFs are not capable of withstanding harsh thermal treatment without significant decomposition of the framework, especially in the presence of air or other oxidizing agents.[147] One of the more thermally and chemically stable MOFs synthesized to date is $Ni_3(btp)_2$ (H_3btp = 1,3,5-tris(1*H*-pyrazol-4-yl)benzene, CSD refcode = UTEWOG),[148] which is stable in air up to 430 °C and is included in the database of MOFs screened in this work (its structure was previously shown in Figure 2.3). From the present work, we predict that $E^*_{a,C-H} =$ 25 kJ/mol. However, we also predict that $\Delta E_{ox} = 119$ kJ/mol, which is likely to prevent sufficient oxidation of the metal site even at elevated temperatures near the stability limit of $Ni_3(btp)_2$.

Another general approach to increase the activity of MOFs for oxidative C–H bond activation when N_2O activation is the rate-limiting step is to consider the use of a more reactive oxidizing agent. As one example, H_2O_2 could be used and would act as a greener oxidant since it only releases water as the byproduct. In addition to being more kinetically reactive than N_2O, the standard enthalpy of reaction for MOF + H_2O_2 → MOF−O + H_2O compared to MOF + N_2O → MOF−O + N_2 is more thermodynamically favorable by 82 kJ/mol at the PB3-D3(BJ) level of theory. H_2O_2 has recently been used as the oxidant in the conversion of methane to methanol with Fe-containing Al-MIL-53[98,149] as well as graphene-confined single Fe atoms[150] and multiple works involving cation-exchanged zeolites.[92,151,152] The use of other strong oxidizing agents is also likely worth considering for the purposes of gaining experimental insight into oxidative C–H bond activation on MOF-supported metal-oxo species. For instance, the strong oxidant $K_2S_2O_8$ has been used for the conversion of methane to acetic acid with V-containing MIL-47 and MOF-48 catalysts.[95] Although the most industrially desirable oxidant is air or O_2, the use of strong oxidants

is expected to significantly aid experimental investigation of MOFs capable of activating the strong C–H bond of methane.

While seven MOF structures were identified as having exothermic ΔE_{ox} values, they are unlikely to be experimentally realizable, as they have unexpected structural defects or solvents that are unlikely to be removed. These instances can often be traced back to challenges in experimentally resolving the positions of ligands or charge-balancing ions during X-ray diffraction studies and/or their removal in construction of the CoRE MOF database. As an example, we highlight one MOF with a predicted exothermic oxidation reaction energy but an ambiguous crystal structure. This MOF, referred to as Rh-BMOF-1 (B = 4,4'-bipyridine, CSD refcode = TERFUT), is a post-synthetically modified, cyclometalated MOF obtained from the reaction of a Rh(I) precursor and a 2-phenylpyridine-5,4'-dicarboxylic acid (dcppy) ligand.[153] BMOF-1 and the related DMOF-1 (D = DABCO = 1,4-diazabicyclo[2.2.2]octane) can be synthesized with Rh(I) or Ir(I) species anchored between a carbanion and a nitrogen atom of the nearby pyridine in the dcppy ligand.[153] Rh-BMOF-1 is predicted to have $\Delta E_{\mathrm{ox}} = -153$ kJ/mol and $E^{*}_{\mathrm{a,C-H}} = 91$ kJ/mol when using the as-published crystal structure with two-coordinate Rh(I) cations.

It is likely that the highly exothermic ΔE_{ox} value can be attributed to atypically undercoordinated Rh(I) sites. Since Rh(I) and Ir(I) are d^8 metals, it is more likely that they form square-planar geometries via the coordination of two additional ligands, as has been found in a related Rh(I)-containing MOF by Sumby and coworkers.[154] As an example, the presence of CO or CH_3CN ligands greatly reduces the redox activity of the Rh(I) and Ir(I) sites, although the resulting C–H activation barrier is expected to be lower if the active site were to form (Figure 2.11). Since the exact nature of the bound ligands has not been experimentally determined due to low

occupancies of the cyclometalated species in the crystal structure,[153] we focus on other screened MOFs for the remainder of the analysis.

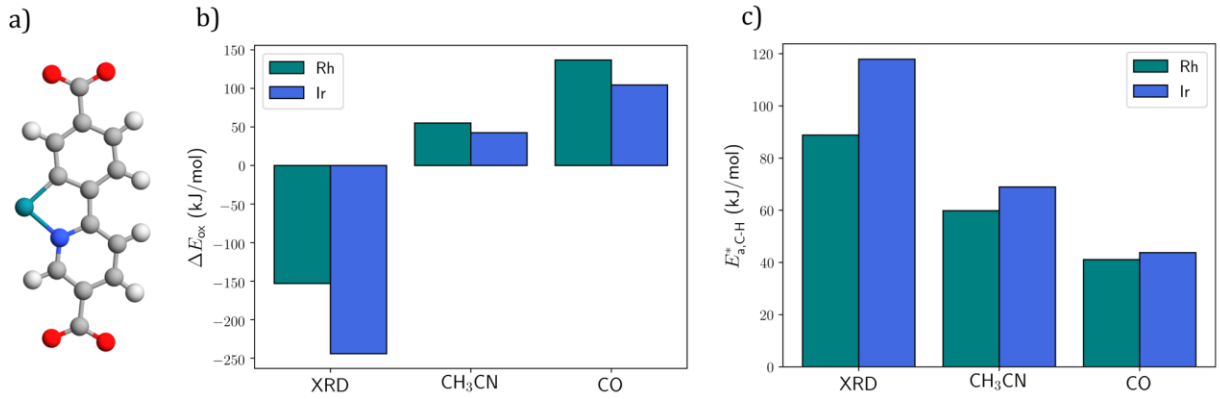


Figure 2.11. a) Rh(I) environment in cyclometalated DMOF-1 based on the structure from X-ray diffraction (XRD) (i.e. no additional ligands). The structure is DFT-optimized (only a representative portion of the unit cell is shown for clarity). Color key: Rh (teal), C (gray), N (blue), O (red), H (white). b) Extrinsic reaction energy for oxidation via N_2O, ΔE_{ox}, defined in Equation 2.4. c) Methane C–H activation barrier, $E^*_{a,C-H}$, at metal oxide active sites, defined in Equation 2.3. XRD, CH_3CN, and CO refer to the two-coordinate metal, square-planar metal environment with two CH_3CN ligands, and square-planar metal environment with two CO ligands. Energetics are identical for the interpenetrated BMOF-1.

From the HT-DFT procedure, the MOF with the lowest oxidation reaction energy and a fully resolved crystal structure is $Cu_8I_4(dmtrz)_4$ (CSD refcode = CUNFOH01), which contains Cu(I)–I cubane tetramers as well as two-coordinate Cu(I) cations,[155] of which only the latter are expected to be readily accessible for oxidation by N_2O. The HT-DFT analysis suggests that the two-coordinate Cu(I) sites in $Cu_8I_4(dmtrz)_4$ have $\Delta E_{ox} = 103$ kJ/mol and $E^*_{a,C-H} = 36$ kJ/mol. With the goal of identifying Cu(I) cations that can be more easily oxidized, we screened the 2019 CoRE MOF database[14] for MOFs with two-coordinate Cu(I) species, a PLD greater than 3.0 Å, and less than 400 atoms in the Niggli-reduced unit cell. From this procedure, we identified $Cu_2(tqpt)$ (H_2tqpt = 6,6,14,14-tetramethyl-6,14-dihydroquin-oxalino-[2,3-*b*]phenazinebistriazole, CSD

refcode = URUWEL), also known as CFA-8,[156] which has the topology shown in Figure 2.12a. CFA-8 contains two-coordinate (linear), three-coordinate (trigonal planar), and four-coordinate (tetrahedral) Cu(I) species in the framework. We will refer to these species as Cu_A, Cu_B, and Cu_C, respectively.

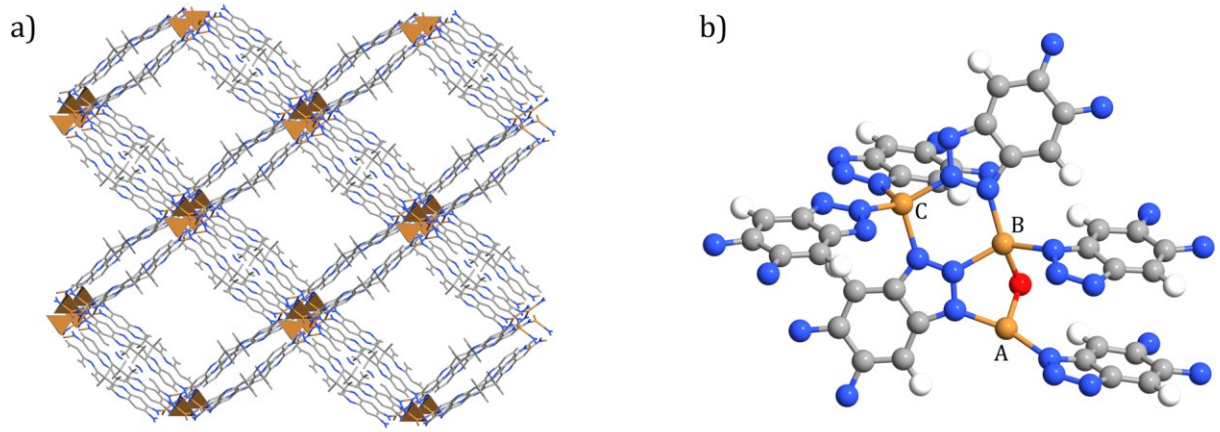


Figure 2.12. a) Structure of CFA-8 with Cu(I) ions lining the pore along the c-axis. H atoms omitted for clarity. b) Ball-and-stick model of distinct Cu sites in CFA-8 and a proposed mono(μ-oxo)dicopper(II) active site following activation of N_2O. Color key: Cu (orange), C (gray), N (blue), O (red), H (white). Structures are DFT-optimized.

Due to the close proximity of Cu(I) species in the framework, we hypothesized that the thermodynamics for oxidation via N_2O could be significantly more favorable than that of $Cu_8I_4(dmtrz)_4$ via the formation of a mono(μ-oxo) motif bridging two neighboring Cu(I) sites (Figure 2.12b). Based on the HT-DFT calculations, the extrinsic oxidation reaction energy is $\Delta E_{ox} = 47$ kJ/mol, which is significantly more thermodynamically achievable than the other MOFs screened in this study. The proposed $[Cu_2O]^{2+}$ core of oxidized CFA-8 contains a bent geometry with a Cu_A–O–Cu_B bond angle of 103° and equal Cu_A–O and Cu_B–O bond lengths of 1.88 Å. The Cu_A–Cu_B bond length decreases from 3.40 Å in the bare MOF to 2.96 Å in the oxidized state.

The potential energy diagram of the proposed mechanism for oxidative C–H bond activation of methane in this MOF is shown in Figure 2.13. For the proposed mono(μ-oxo) sites in CFA-8, we calculate that $E_{\mathrm{a,C-H}} = 57$ kJ/mol, which is essentially identical to the experimentally observed ~65 kJ/mol barrier attributed to similar mono(μ-oxo) dicopper(II) active sites in Cu-ZSM-5 that can catalytically convert methane to methanol.[143] It is also predicted that methanol, the desired product, should readily desorb at ambient conditions, which is commonly problematic for cation-exchanged zeolites.[157] We note that the barrier for N_2O activation to form the $[Cu_2O]^{2+}$ site is relatively high with a predicted value of 152 kJ/mol, such that it still expected to dictate the overall reaction kinetics for this system. Improved catalytic performance can potentially be achieved by considering more reactive oxidants. Provided a transient mono(μ-oxo) active site can be formed, CFA-8 would likely be a promising candidate for oxidative C–H bond activation reactions.

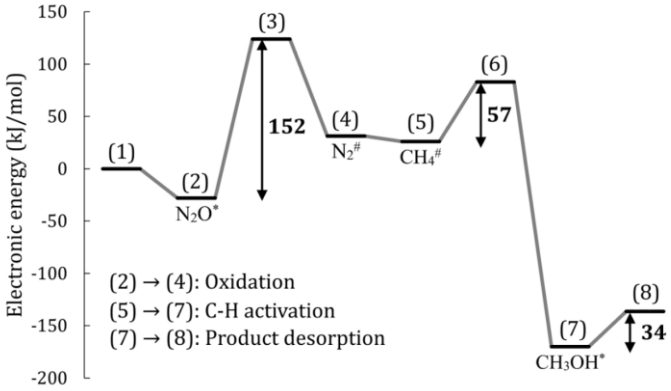


Figure 2.13. Potential energy diagram for oxidation of the metal centers via N_2O, subsequent H-abstraction of methane, and formation/desorption of methanol with CFA-8. (1) MOF + N_2O (g) + CH_4 (g); (2) N_2O*; (3) transition state for N_2O activation; (4) $N_2^{\#}$; (5) $CH_4^{\#}$; (6) transition state for H-abstraction; (7) CH_3OH*; (8) MOF + CH_3OH (g) + N_2 (g). Here, * denotes adsorption at the metal center and # denotes adsorption at the metal-oxo site formed via oxidation. Since the barrier for the radical rebound of $\bullet CH_3$ is expected to be small compared to the other barriers, this step is omitted.

To obtain a better understanding of the proposed $[Cu_2O]^{2+}$ active site of CFA-8, we calculated the density-derived electrostatic and chemical (DDEC) partial atomic charges[158] and Bader spin density[159] for each atom. As shown in Table 2.3, the DDEC charges on the Cu(I) sites increase from 0.26 to 0.53–0.57 following N_2O activation. The large increase and near-equal magnitude of the partial atomic charges on the Cu species is consistent with the oxidation of the Cu(I) sites to mono(μ-oxo)dicopper(II) species. At the TS for N_2O activation, the DDEC analysis suggests that the Cu_B site is more easily oxidized than the Cu_A site, with the Cu_B site being almost entirely oxidized from Cu(I) to Cu(II) at the TS.

Formally, the H-abstraction from methane should decrease the charge of the active site by one. Instead of equally reducing the charge on both Cu(II) sites, the DDEC analysis indicates that the Cu_A site is more significantly reduced, whereas the Cu_B retains most of its charge. The Bader spin densities provide additional insight into this phenomenon. Since it is expected that a Cu(I) species should have no net spin, the Bader spin densities suggest that the Cu_A and Cu_B sites are likely best described as being in the 1+ and 2+ oxidation state following H-abstraction, respectively. The Bader spin density analysis also highlights the radical-like character of the bridging oxo species prior to H-abstraction, which results in the low methane C–H activation barrier of 57 kJ/mol. Finally, the formation of methanol closes the catalytic cycle, with both the Cu_A and Cu_B sites returning to their original 1+ oxidation states.

Table 2.3. DDEC6 partial atomic charges[158] and Bader spin densitiy[159] for the proposed active site of CFA-8 throughout the radical-rebound mechanism. Cu_A and Cu_B refer to the Cu sites shown in Figure 2.12b. $S = 1$ and $S = 3$ refer to the closed-shell singlet and open-shell triplet spin multiplicities, respectively. $S = 1^*$ refers to an open-shell singlet state with an antiferromagnetically coupled methyl radical.

	Bare MOF	TS for formation of $[Cu_2O]^+$ site	Oxidized MOF	TS for C–H activation of CH_4	MOF with adsorbed $•CH_3$	MOF with adsorbed CH_3OH
DDEC partial atomic charge						
Cu_A	0.26	0.34	0.53	0.40	0.36	0.27
Cu_B	0.26	0.52	0.57	0.54	0.52	0.28
O	--	−0.50	−0.61	−0.65	−0.72	−0.45
Magnitude of Bader spin density						
Cu_A			0.38	0.04	0.11	
Cu_B			0.48	0.33	0.40	
O			0.76	0.11	0.15	
Spin multiplicity						
	$S = 1$	$S = 1$	$S = 3$	$S = 1^*$	$S = 1^*$	$S = 1$

2.4 Conclusion

High-throughput periodic DFT is a promising method for accelerating the discovery of MOFs for various applications, but there are numerous technical challenges that must be addressed before it can be routinely used to design and/or discover MOF candidates. In this work, we have developed a robust and automated workflow for the high-throughput screening of MOFs using periodic DFT, specifically focusing on applications in heterogeneous catalysis. We describe appropriate choices for electronic and structural optimization algorithms, treatment of spin states, and methods for automating the calculation of adsorption energies at open metal sites.

As a proof-of-concept, we applied this workflow to screen MOFs with open metal sites from an 838 MOF subset of the CoRE MOF database[89,112] for the oxidative C–H bond activation of methane using an N_2O oxidant. From this high-throughput procedure, it was found that oxidation of the metal – and not the C–H bond activation of methane – is the step with the largest barrier for the majority of screened MOFs in this work. Based on this finding, we expect that the development of new MOF datasets focused on low-valence, redox-active open metal sites will be

central to the discovery of MOF-based heterogeneous catalysts that can directly convert methane to methanol at ambient conditions as well as other oxidative C–H bond activation reactions. With the computational framework outlined in this work, we hope that high-throughput periodic density functional theory will become a more mainstream tool for designing and identifying MOFs with unique physicochemical properties.

Chapter 3: Structure−Activity Relationships that Identify Metal−Organic Framework Catalysts for Methane Activation

In this chapter, computational screening based on periodic density functional theory (DFT) is used to study a diverse set of experimentally derived metal−organic frameworks (MOFs) with accessible metal sites for the oxidative activation of methane. We find that the thermodynamic favorability of forming the metal-oxo active site has a strong, inverse correlation with the reactivity toward C−H bond activation for a wide range of MOFs. This scaling relationship is found to hold over MOFs with varying coordination environments and metal compositions, provided the bonds of the framework atoms are conserved. The need to conserve bonds is an important constraint on the correlations but also demonstrates a route to intentionally break the scaling relationship to generate novel catalytic reactivity. Periodic trends are also observed across the dataset of screened MOFs, with later transition metals forming less stable but more reactive metal-oxo active sites. Collectively, the results in this work provide robust rules-of-thumb for choosing MOFs to investigate for the activation of methane at moderate reaction conditions.

This chapter is adapted from the following peer-reviewed article: A.S. Rosen, J.M. Notestein, R.Q. Snurr. "Structure−Activity Relationships that Identify Metal−Organic Framework Catalysts for Methane Activation." *ACS Catalysis*, **9**, 3576–3587 (2019). This work was highlighted in *C&EN*.[160]

3.1 Introduction

Globally, over ten billion cubic meters of natural gas are produced per day to meet ever-increasing energy demands.[161] Depending on the source, natural gas can be composed of up to 96% methane by volume, with the remainder consisting mainly of other light alkanes.[161] The current method to convert the methane found in natural gas into methanol, a liquid fuel and feedstock for valuable chemical derivatives, is via a multi-step industrial process that involves the intermediate production of synthesis gas and requires harsh operating conditions.[162] Since this process is neither efficient nor economically viable for deployment at smaller scales, such as near natural gas extraction sites, a significant portion of the world's supply of natural gas is wastefully flared.[163] This results in a multibillion dollar loss of value as well as a significant increase in the amount of greenhouse gases released into the atmosphere.[163,164] The discovery of a catalyst that can directly convert methane (with its large C−H bond dissociation enthalpy of 4.55 eV at 298 K)[165] into methanol at moderate temperatures and pressures is therefore extremely desirable, although such a catalyst has yet to be fully realized.[157,166,167]

Recently, there have been numerous experimental studies on metal−organic frameworks (MOFs) for the catalytic upgrading of natural gas-derived light alkanes,[95–99,168,169] due in part to their ability to support well-defined, spatially isolated metal sites.[94,170] The high degree of chemical and topological tunability of MOFs, enabled by the vast library of inorganic nodes and organic linkers,[171,172] also makes them particularly well-suited for high-throughput computational screening studies.[46,53,173,174] Despite the promise of computationally guided catalyst design via high-throughput periodic density functional theory (DFT),[40,101,175] modeling of MOFs for the partial oxidation of light alkanes has traditionally relied on small-scale, system-specific studies

often involving finite cluster models.[107,176–179] Given the many thousands of MOFs that can be experimentally synthesized[44,89] and the computational cost associated with modeling the large unit cells of most MOFs,[69] there exists a significant need for robust and widely applicable design principles that would accelerate the process of discovering MOF catalysts that can activate the strong C−H bond of methane.

One of the most crucial aspects of predictive catalyst design is the identification of readily computed descriptors, such as adsorption or reaction energies, that can be used for the efficient prediction of other key kinetic and thermodynamic parameters.[101,104,180–184] For the partial oxidation of methane via the radical rebound mechanism (Figure 3.3a), Nørskov and coworkers have identified a universal relationship between the H-affinity, ΔE_{H} (Figure 3.1b, Equation 2.1), and the energy of the transition state (TS) for C−H bond activation, $E_{\mathrm{TS,C-H}}$ (Figure 3.1c, Equation 2.2).[105] The linear relationship between ΔE_{H} and $E_{\mathrm{TS,C-H}}$ has been benchmarked for a wide range of materials (i.e. cation-exchanged zeolites, bulk metal oxides, transition metal surfaces, metal nanoparticles, decorated graphene nanosheets, and MOFs) and has a low mean absolute error (MAE) of 0.11 eV such that the H-affinity can be interpreted as a robust measure of active site reactivity without the need to perform computationally prohibitive TS calculations.[105] Nonetheless, even with the H-affinity relationship or the related Brønsted-Evans-Polanyi (BEP) relationship,[104] calculating the kinetic barrier for C−H activation, $E_{\mathrm{a,C-H}}$ (Figure 3.1c, Equation 2.6), as well as the thermodynamic favorability of forming the metal-oxo active site, ΔE_{O} (Figure 3.1b, Equation 2.8), still requires many DFT calculations for each MOF, significantly limiting the number of MOFs that can be investigated.

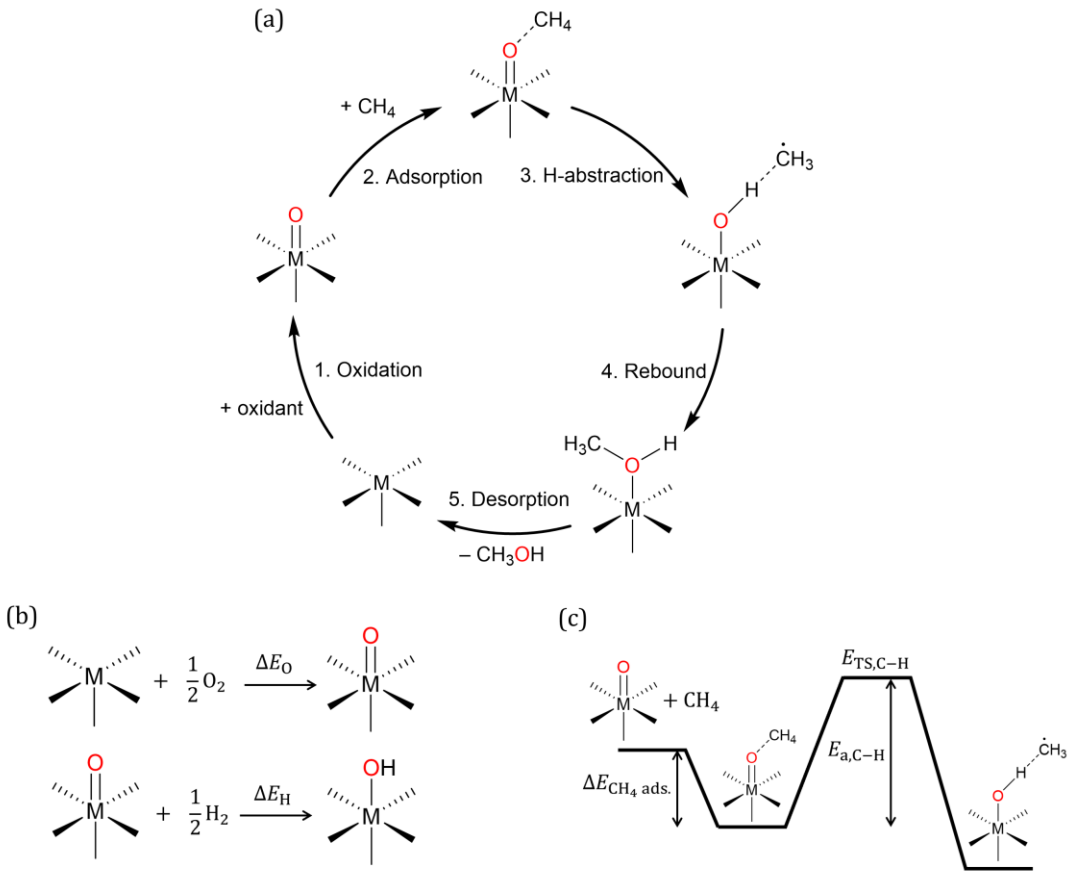


Figure 3.1. a) Radical-rebound mechanism for the catalytic partial oxidation of methane to methanol. The specific coordination environment and M−O bond order are dependent on the MOF. b) Depiction of ΔE_{O} (active site formation energy using O_2 as the reference) and ΔE_{H} (H-affinity using H_2 as the reference). c) Depiction of $\Delta E_{\mathrm{CH_4\ ads.}}$ (methane adsorption energy), $E_{\mathrm{TS,C-H}}$ (energy of the transition state for methane C−H bond activation with respect to the energy of the M−O site and gas-phase CH_4), and $E_{\mathrm{a,C-H}}$ (barrier for the C−H bond activation of methane).

In this work, we use high-throughput periodic DFT[29,32,40] to efficiently screen a diverse set of 60 MOFs with accessible metal sites (Figure 3.2 and Appendix) for the catalytic activation of methane. As a result of this screening study, we show that a single parameter – the active site formation energy – can be used to predict both the reactivity toward C−H bond activation and the thermodynamic stability of the metal-oxo active site. We find this relationship to hold across all the MOFs screened in this work, regardless of coordination environment, metal type, and degree

of charge delocalization, so long as the bonds of the remaining framework atoms are conserved during the reaction. Periodic trends with regard to active site stability and reactivity are observed, with later transition metals forming less thermodynamically favored but more reactive metal-oxo active sites for C−H activation. As a result of this screening study, the binding strength of the oxo species appears to be a powerful descriptor for assessing the ability of isolated metal sites in MOFs to oxidatively activate strong C−H bonds, especially for the purposes of rapidly identifying promising MOF candidates for further investigation.

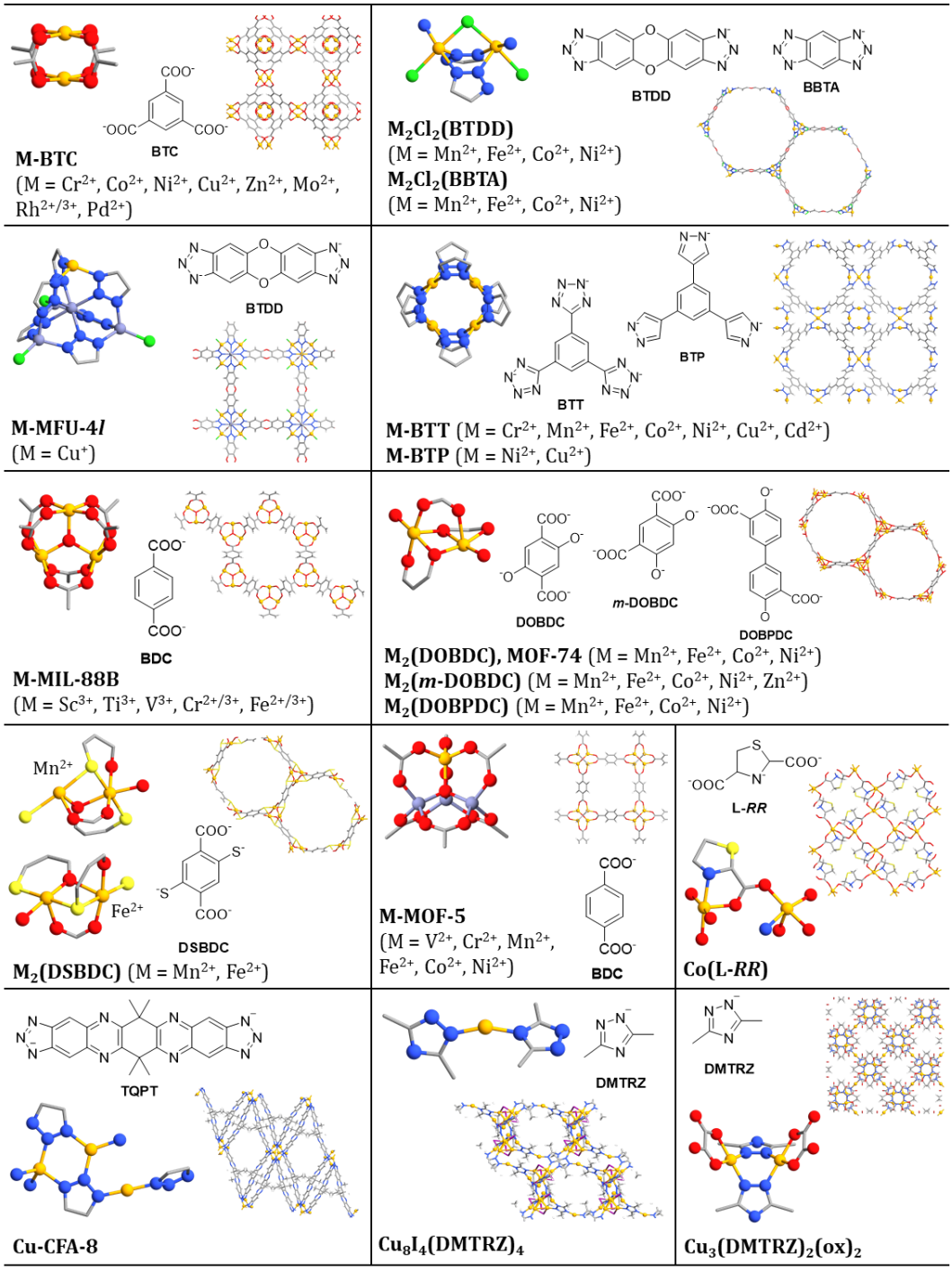


Figure 3.2. Overview of the 60 MOFs considered in this work. Additional details for each of the investigated MOFs can be found in the Appendix. Color key: transition metal cation (orange), N (blue), O (red), C (gray), S (yellow), Cl (green), I (purple), Zn (light blue).

3.2 Methods

3.2.1 MOF Selection

The MOF families M-BTC,[115,185–191] M_2Cl_2(btdd),[192,193] M_2Cl_2(bbta),[193,194] M-MFU-4*l*,[27] M-BTT,[22,195–198] M-BTP,[148] M-MIL-88B,[199,200] M-MOF-74,[201,202] M_2(*m*-dobdc),[203] M_2(dobpdc),[204] M_2(dsbdc),[205,206] M-MOF-5,[207–210] and Co(L-*RR*)[211] were chosen because at least one member of each family has been shown to support open metal sites that are redox-active[25,168,207,212–214] and/or can chemisorb O_2[20,22,24,27,185,215–217] (or is structurally similar to a MOF with such properties). Cu-CFA-8,[156] $Cu_8I_4(dmtrz)_4$,[155] and $Cu_3(dmtrz)_2(ox)_2$[142] were highlighted in prior work[32] and are included here due to their unique active site structures (as discussed in the Appendix). Additional details regarding the motivation and pertinent modeling decisions for each of the MOFs considered in this study are included in the Appendix.

3.2.2 Energetic Parameters

For the radical-rebound mechanism shown in Figure 3.1a, Nørskov and coworkers have identified that the energy of the TS for methane C−H bond activation is directly related to the strength that an H atom binds to the active site, referred to as the H-affinity.[105,111,140] Using H_2 as the reference state, the H-affinity, ΔE_{H}, can be calculated as

$$\Delta E_{\mathrm{H}} = E_{\mathrm{MOF-OH}} - E_{\mathrm{MOF-O}} - \frac{1}{2}E_{\mathrm{H_2}} \tag{3.1}$$

here $E_{\mathrm{MOF-OH}}$ and $E_{\mathrm{MOF-O}}$ are the electronic energies of the metal site with adsorbed OH and O species, respectively, and $E_{\mathrm{H_2}}$ is the electronic energy of H_2. With this definition, the H-affinity TS scaling relationship[105] of

$$E_{\mathrm{TS,C-H}} = 0.75\Delta E_{\mathrm{H}} + 1.96\ \mathrm{eV} \tag{3.2}$$

can be used to predict the energy of the TS for H-abstraction of methane, $E_{\mathrm{TS,C-H}}$. The value for $E_{\mathrm{TS,C-H}}$ is with respect to the infinitely separated M−O site and gas-phase CH_4. Note that unlike the original definition of ΔE_{H},[105] Equation 3.2 has been modified to use H_2 (rather than O_2 and H_2O) as the H-reference. It has been proposed that the relationship between ΔE_{H} and $E_{\mathrm{TS,C-H}}$ of different light hydrocarbons are linearly offset by their C−H bond strengths,[105,218] so the scaling relationships developed in this work can potentially be extended to species other than methane as well.

Using Equations 3.1 and 3.2, the C−H bond activation barrier, $E_{\mathrm{a,C-H}}$, can be computed via

$$E_{\mathrm{a,C-H}} = E_{\mathrm{TS,C-H}} - \Delta E_{\mathrm{CH_4\ ads.}} \tag{3.3}$$

where

$$\Delta E_{\mathrm{CH_4\ ads.}} = E_{\mathrm{MOF-O-CH_4}} - E_{\mathrm{MOF-O}} - E_{\mathrm{CH_4}} \tag{3.4}$$

and $\Delta E_{\mathrm{CH_4\ ads.}}$ is the adsorption energy of CH_4 at the metal-oxo active site. While a BEP relationship[181] may result in an even more accurate prediction of $E_{\mathrm{a,C-H}}$ when studying a closely related subset of catalytic candidates,[219] the H-affinity relationship is the only approach that has been shown to hold over a diverse set of catalysts.[105]

To account for the thermodynamic stability of the proposed metal-oxo active site, we compute the metal-oxo formation energy. As in prior literature,[105–107,111,140] we present the metal-oxo formation energy using O_2 as the reference state, defined a

$$\Delta E_{\mathrm{O}} = E_{\mathrm{MOF-O}} - E_{\mathrm{MOF}} - \frac{1}{2}E_{\mathrm{O_2}} \tag{3.5}$$

Here, E_{O_2} is the electronic energy of O_2 corrected to account for limitations in the description of triplet O_2 by common exchange-correlation functionals,[220–222] as described in the Appendix. A variety of gas-phase energies are reported in Table B.6 to allow for direct comparison with prior work, regardless of the choice of reference state. While all trends discussed in this work are independent of the choice of reference state, the consideration of a different oxidant will naturally result in values for ΔE_O that are linearly offset by an additive constant. This also implies a shift in the boundaries shown in Figure 3.4 such that the choice of oxidant is another important design decision to consider when selecting promising MOFs for experimental investigation.

To summarize, without structure−activity relationships beyond the H-affinity descriptor, computing ΔE_O and $E_{a,C-H}$ requires the energies associated with four DFT-optimized structures per MOF: 1) the bare MOF, E_{MOF}; 2) the MOF with metal-oxo active site, E_{MOF-O}; 3) the MOF with the partially reduced active site following H-abstraction, E_{MOF-OH}; 4) the MOF with methane adsorbed at the active site, $E_{MOF-O-CH_4}$. All of these computed energies are reported in Table B.1 for ease of reference.

In this work, we specifically focused on formation of the metal-oxo site, methane adsorption, and H-abstraction, as they dictate whether the catalyst can activate methane. Achieving high selectivity toward methanol is a separate but related problem[223] that cannot be fully addressed without a greater understanding of how to successfully activate methane in the first place. We note that the kinetic barrier for the rebound of the methyl radical is often small compared to either the barrier for C−H activation or formation of the metal-oxo active site. In addition, both the rebound step and the thermodynamic barrier for methanol desorption provide insight into the selectivity

toward methanol, rather than the conversion of methane, the latter of which is the focus of the present study.

3.2.3 Periodic Density Functional Theory

All DFT calculations were performed with the Vienna *ab initio* Simulation Package (VASP) v.5.4.1[122,123] and the VASP v.5.4 projector-augmented wave pseudopotentials.[124] The electron exchange-correlation was described by the generalized gradient approximation (GGA) functional developed by Perdew, Burke, and Ernzerhof (PBE)[125] with Grimme's D3 dispersion correction scheme[126] and Becke-Johnson (BJ) damping.[224] Since periodic DFT was used for the computational screening procedure, the entire crystallographic unit cell of each MOF was included in all calculations. The cell volume, cell shape, and internal degrees of freedom were optimized for each MOF starting from the experimental crystal structure (when available) with all solvents removed. After the volume relaxation procedure, adsorbates were systematically added,[32] and the atomic positions were optimized to a minimum in the potential energy surface (at a fixed cell volume).

A 520 eV plane-wave kinetic energy cutoff and k-point grid[225] consisting of 1000 k-points/number of atoms were used in the final stage of the high-throughput periodic DFT workflow for each MOF, as computed using Pymatgen[82] and benchmarked previously.[32] In addition, all forces were converged to within 0.03 eV/Å. A preconditioned conjugate gradient (CG) algorithm referred to as the "all bands simultaneous update of orbitals" algorithm[132,133] was used to achieve convergence of the self-consistent field, and we employed Gaussian smearing of the band occupancies with a small smearing width of σ = 0.01 eV, followed by extrapolation of the electronic energy to the $T = 0$ K limit. A combination of the CG and Fast Inertial Relaxation

Engine (FIRE)[137] algorithms were used to achieve convergence for geometry optimizations.[32] Spin-polarization was considered for all MOFs, as discussed in the Appendix.

In general, it is often found that GGA-level functionals disfavor high spin states, whereas many hybrid-level functionals artificially stabilize high spin states.[226–229] While ~80% of the bare MOFs in this work with d^3 to d^8 transition metals were predicted to have high spin ground states, we note this is one possible limitation of the PBE exchange-correlation functional. Nonetheless, the PBE-D3(BJ) level of theory was chosen due to its ability to reasonably describe adsorption at transition metal sites at a modest computational cost,[222] which is essential for carrying out the large number of structural relaxations in the present study. Including the full crystallographic unit cell via periodic DFT calculations avoids the possibility of artificial boundary effects associated with smaller unit cells and allows for a more rapid and automated screening process than could be achieved using finite cluster models.[32] The PBE functional with dispersion corrections is also known to accurately capture the structures of MOFs.[128,129]

Partial atomic charges, spin densities, and effective bond orders were computed using the DDEC6 method,[158,230–232] which has been shown to give physically reasonable results for porous solids, including MOFs.[231,233,234] The open-source, Python-based high-throughput periodic DFT screening workflow[139] makes use of the Atomic Simulation Environment (ASE)[83] and Pymatgen[82] for automating the VASP calculations.[32] As described in prior work,[32] the MOF Adsorbate Initializer (MAI)[117] was used to automatically initialize the positions of adsorbates at the open metal sites,[32] and RASPA[119] was used to compute the potential energy grids required to initialize the structures for methane adsorption. Additional computational details can be found in the Appendix, including Tables B.4−B.7.

3.2.4 Data Availability

The data that support the findings of this study (including DFT-optimized structures, energies, and computed physicochemical properties) are publicly available via Zenodo with the identifier doi:10.5281/zenodo.1734640.[235] Relevant energetic quantities are also tabulated in the Appendix for ease of reference. The codes used to carry out the DFT screening in VASP and to initialize the positions of adsorbates are publicly available.[117,139]

3.3 Results and Discussion

3.3.1 Active Site Formation Energy as a Unifying Descriptor

Given the previously determined robust linear relationship between the H-affinity, ΔE_{H}, and the TS energy for methane C−H activation, $E_{\mathrm{TS,C-H}}$,[105] we set out to calculate ΔE_{H} as a measure of active site reactivity and ΔE_{O} as a measure of the thermodynamic favorability of forming the active site for the MOFs in Figure 3.2. We emphasize that both the active site stability and reactivity toward C−H activation are necessary to consider when evaluating a series of MOF candidates, as methane conversion cannot be achieved without both criteria being satisfied.[32]

As shown in Figure 3.3, we identify a general trend that less stable metal-oxo active sites (i.e. more endothermic ΔE_{O} values) are associated with higher reactivities toward H-abstraction (i.e. more exothermic ΔE_{H} values and lower TS energies for C−H bond activation). Specifically, we observe that a single scaling line given by

$$\Delta E_{\mathrm{H}} = -0.43\Delta E_{\mathrm{O}} - 1.92 \text{ eV} \tag{3.6}$$

can be used to directly relate the active site formation energy to the H-affinity. The correlation between these two parameters is substantial, with $r^2 = 0.94$ and a MAE of only 0.12 eV, well within expected errors from DFT. Since the linear energy relationship in Figure 3.3 holds over the

wide variety of MOFs screened in this work, this implies that active site stability can be used as a reliable descriptor for predicting the reactivity toward C−H bond activation for MOFs containing accessible, spatially isolated metal sites.

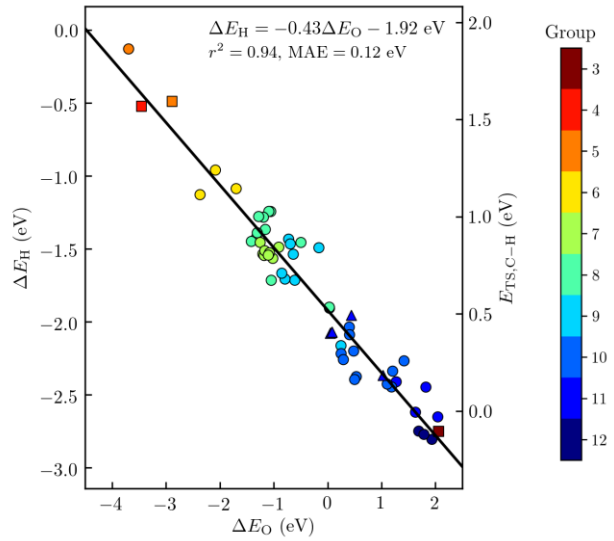


Figure 3.3. H-affinity, ΔE_H, as a function of the active site formation energy, ΔE_O, for the screened MOFs. The best-fit line, $\Delta E_H = -0.43\Delta E_O - 1.92$ eV, has $r^2 = 0.94$ and a mean absolute error (MAE) of 0.12 eV. The energy of the transition state (TS) for methane C−H activation is computed from ΔE_H via the H-affinity TS scaling relationship. Symbol color refers to the metal group number. Symbol shape indicates the formal oxidation state of the metal site prior to oxidation as 1+ (▲), 2+ (●), or 3+ (■).

While Equation 3.6 enables the prediction of ΔE_H, and by extension $E_{TS,C-H}$, the more useful kinetic parameter is $E_{a,C-H}$. Since the activation energy and TS energy differ by the energy of methane adsorption (Figure 3.1c, Equation 3.3), the methane binding strength for each catalytic candidate is needed. For adsorption at the metal-oxo active sites in MOFs, however, we find that the methane adsorption energy is effectively constant, as demonstrated in Figure B.1. The average methane adsorption energy for the screened MOFs is predicted to be −0.18 eV, as evident from the best-fit line given by

$$E_{\mathrm{a,C-H}} = 0.99E_{\mathrm{TS,C-H}} + 0.18\ \mathrm{eV} \tag{3.7}$$

which has $r^2 = 0.99$ and a MAE of 0.03 eV (Figure B.1a). The consistently small value for the methane adsorption energy (Figure B.1b) suggests that the process is mainly governed by dispersive interactions consistent with weak physisorption, similar to what has been predicted for various trinuclear motifs in the zeolite ZSM-5.[219] The MOF with the largest deviation from the average −0.18 eV adsorption energy is Co(L-*RR*) (L-*RR* = (*R*,*R*)-thiazolidine-2,4-dicarboxylate),[211] which is predicted to have a methane adsorption energy of −0.28 eV due to increased van der Waals interactions as a result of its particularly small pore size[211,215] (Figure B.2). This suggests that even MOFs with pores approaching the kinetic diameter of methane are reasonably well-described by Equation 3.7, especially for screening purposes.

The near-constant methane adsorption energy makes it possible to predict the barrier for C−H bond activation, instead of just the TS energy, from the active site formation energy. This is illustrated in Figure 3.4, with the modified scaling relationship of

$$E_{\mathrm{a,C-H}} = -0.32\Delta E_{\mathrm{O}} + 0.69\ \mathrm{eV} \tag{3.8}$$

which has $r^2 = 0.94$ and a small MAE of 0.09 eV. Note that since the near-constant methane adsorption energy was not known in advance, we did not assume it was a constant value in proposing the scaling relationship given by Equation 3.8. Nonetheless, it is only possible to express the stability−reactivity linear scaling relationship in terms of $E_{\mathrm{a,C-H}}$ if $E_{\mathrm{TS,C-H}}$ is directly proportional to $E_{\mathrm{a,C-H}}$, as shown in Figure B.1a. The main benefit of the widely applicable scaling relationship given by Equation 3.8 is that only the active site formation energy is required to predict both the stability of the active site and the C−H bond activation barrier. This implies that both the

active site stability and reactivity can be reliably determined from just two straightforward energy calculations (i.e. that of the bare and oxidized MOF), without having to directly calculate the H-affinity or methane adsorption energy.

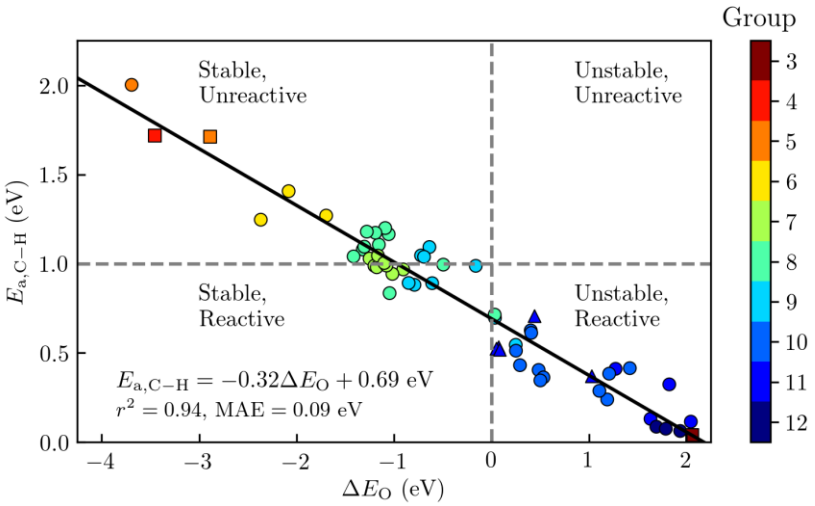


Figure 3.4. Predicted barrier for the C−H bond activation of methane, $E_{a,C-H}$, as a function of the metal-oxo formation energy, ΔE_O. The best-fit line, $E_{a,C-H} = -0.32\Delta E_O + 0.69$ eV, has $r^2 = 0.94$ and a mean absolute error (MAE) of 0.09 eV. As a visual aid, we classify MOFs with $E_{a,C-H} < 1$ eV as being reactive toward C−H bond activation and MOFs with $\Delta E_O < 0$ as having thermodynamically favored active sites when using O_2 as the reference state. Symbol color refers to the group number of the metal in the periodic table. Symbol shape indicates the formal oxidation state of the metal site prior to oxidation as 1+ (▲), 2+ (●), or 3+ (■).

Consistent with the Sabatier principle,[182] the ideal catalyst is likely one with an active site formation energy that is not so exothermic that the C−H activation barrier is too high but also not so endothermic that the population of active sites is too low to achieve sufficient methane conversion. As shown in Figure 3.4, the scaling line does indeed cross through the desirable "stable, reactive" region, providing support that it is feasible to synthesize MOFs that can form metal-oxo active sites for the activation of methane. For reference, our calculations indicate that Fe-MOF-74 is in the "stable, unreactive" region with $\Delta E_O = -1.16$ eV and $E_{a,C-H} = 1.11$ eV. This is consistent with the fact that Fe-MOF-74 can oxidize ethane to ethanol but has not been shown

to activate the stronger C−H bond of methane.[168] Reassuringly, the present work suggests that there is significant improvement that can be gained over Fe-MOF-74 in terms of catalytic performance. It is notable that many of the MOFs with Fe^{2+} or Cu^{+} cations, which are commonly investigated for the activation of strong C−H bonds, are along the "unreactive/reactive" and "stable/unstable" boundaries, respectively. Based on this finding, the design of MOFs with more reactive Fe(IV)-oxo species or more thermodynamically favored Cu(III)-oxo motifs are promising areas for future research.[b]

3.3.2 Periodic Stability and Reactivity Trends

The color gradient in Figure 3.4 reveals that the thermodynamic favorability for forming the metal-oxo active site generally decreases with increasing group number, with reaction energies ranging from approximately −4 eV to +2 eV depending on the identity of the metal species (and oxidation state). As shown in Figure 3.5a, MOFs with accessible V^{2+} (group 5) or Cr^{2+} (group 6) cations are predicted to form thermodynamically favored metal-oxo active sites, whereas MOFs containing late transition metals such as Ni^{2+} (group 10) are harder to oxidize. An analogous manifestation of this trend is apparent when considering the ligand field diagrams for terminal metal-oxo transition metal complexes, which tend to have an increased occupation of antibonding molecular orbitals at higher d electron counts (the so-called "oxo wall" separating stable, terminal metal-oxo complexes composed of early transition metals from those containing late transition metals),[236] thereby reducing the M−O bond order and active site stability.

[b] Following the publication of this work, it was shown that MIL-100(Fe) – a structural analogue of Fe-MIL-88B (which resides in the "stable, reactive" region of Figure 3.4) – can selectively oxidize methane to methanol in the presence of N_2O.[626,627]

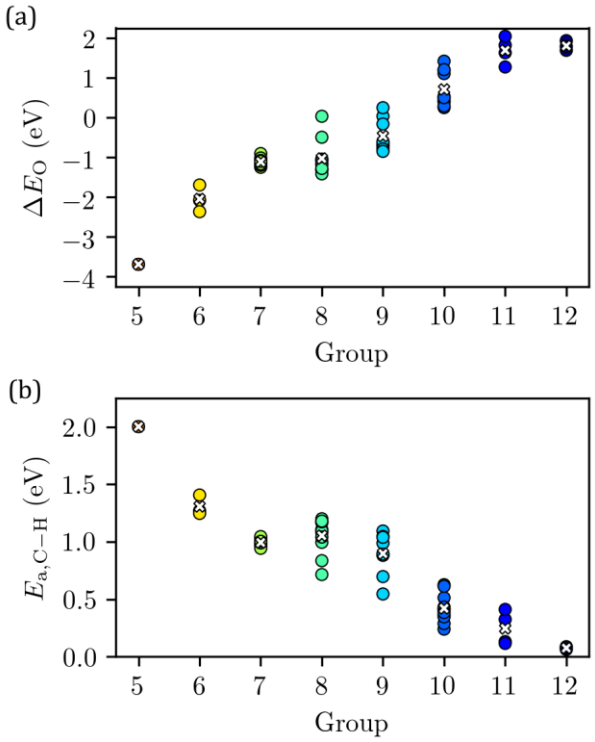


Figure 3.5. a) Active site formation energy, ΔE_O, and b) barrier for methane C−H activation, $E_{a,C-H}$, as a function of group number for the screened MOFs containing transition metals in the 2+ state prior to oxidation. The mean value for each group number is represented by a white "x".

While transition metals with higher d electron counts are less likely to form a stable metal-oxo species, they also tend to be more reactive toward activating the C−H bond of methane (Figure 3.5b). In the extreme limit, such as for purely hypothetical $[ZnO]^{2+}$ (group 12) active sites, the metal-oxyl radical is so reactive that the C−H bond activation of methane is likely to be nearly barrierless. However, such sites are not expected to be catalytically relevant at moderate reaction conditions because they are unlikely to form. Since a highly endothermic formation energy for the metal-oxo active site necessarily implies that there is also a high kinetic barrier to formation, both the kinetic and thermodynamic feasibility of active site formation are expected to be poor for cases where the C−H bond activation of methane is extremely low. Similarly, readily oxidized MOFs, such as those with exposed V^{2+} sites that can form $[VO]^{2+}$ species, are also unlikely to be ideal

candidates due to relatively low reactivities toward C−H bond activation. This design principle can be used as a robust rule-of-thumb for identifying plausible experimental targets. For instance, Ti-MIL-101 is known from experiments to strongly and irreversibly chemisorb O_2,[20] but for this reason it can be immediately assumed to be unreactive toward the C−H bond activation of methane. Indeed, for the closely related Ti-MIL-88B (which has analogous trimetallic nodes and terephthalate linkers) included in our screening study, we predict that $E_{\mathrm{a,C-H}}$ is 1.74 eV.

3.3.3 Additional Physicochemical Descriptors

While there is a strong correlation between ΔE_{O} and ΔE_{H}, it is important to identify which factors contribute to these binding energies. Spin density on the abstracting oxygen atom, ρ_{O}, is one property commonly thought to affect the reactivity toward C−H activation, although the generality of this effect has been a topic of significant debate.[237–239] As shown in Figure 3.6a, we find a qualitative correlation between the DDEC-computed[158] ρ_{O} and ΔE_{H}, and MOFs with low barriers for C−H activation tend to have some degree of radical-like character on the oxo species. Nonetheless, the linearity of this relationship is relatively weak ($r^2 = 0.73$) for the wide range of MOF structures in this work. As one example, Ni_2Cl_2(btdd) is predicted to have a lower barrier toward H-abstraction than Cu-MFU-4*l* ($E_{\mathrm{a,C-H}}$ = 0.35 eV and 0.53 eV, respectively), yet Ni_2Cl_2(btdd) is predicted to have nearly two times less spin density on the oxo ligand than Cu-MFU-4*l* (ρ_{O} = 0.46 and 0.87, respectively). Similar behavior to that shown in Figure 3.6a is observed when computing ρ_{O} via a Bader partitioning scheme[159] (Figure B.3a).

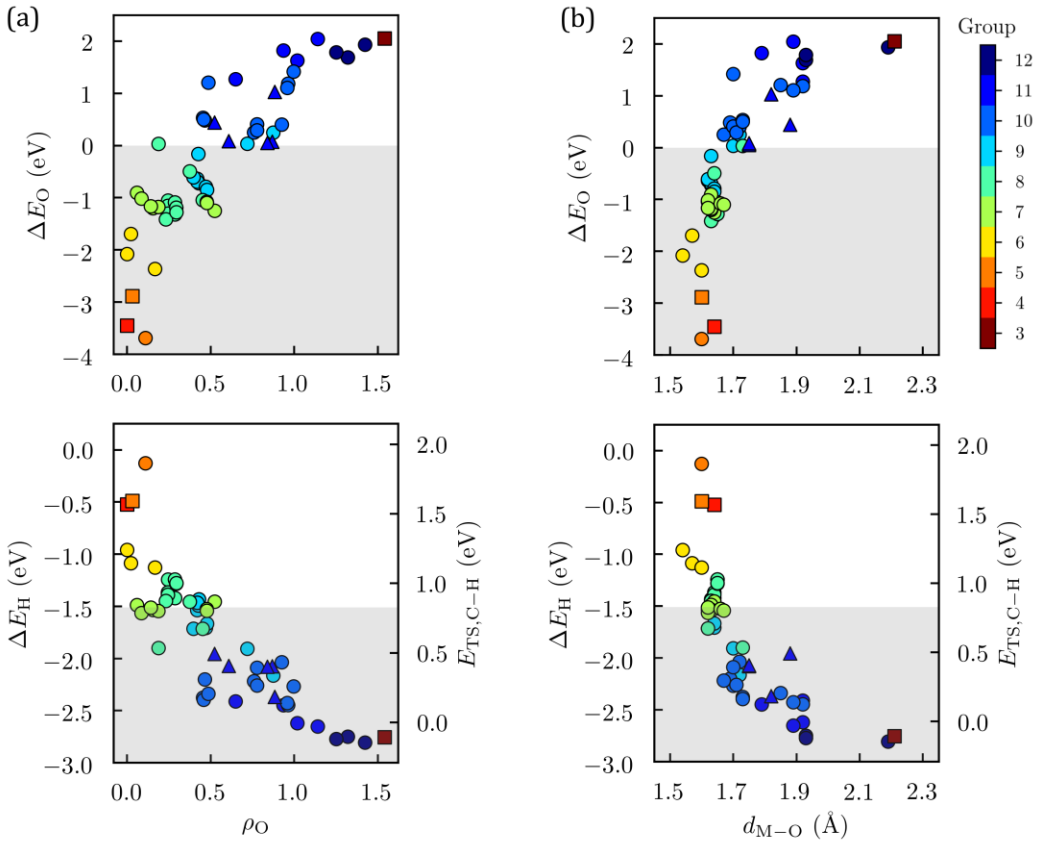


Figure 3.6. a) Active site formation energy, ΔE_O, and H-affinity, ΔE_H, as a function of the absolute spin density on the oxo ligand computed via the DDEC6[158] method, ρ_O. b) ΔE_O and ΔE_H as a function of the M−O distance. The gray shaded area refers to the "stable, reactive" region in Figure 3.4 (i.e. $\Delta E_O < 0$ eV and $\Delta E_H < -1.51$ eV, the latter of which corresponds to approximately $E_{a,C-H} < 1$ eV). Both ρ_O and d_{M-O} are computed for the oxidized state of the MOF. Symbol color refers to the group number of the metal in the periodic table. Symbol shape indicates the formal oxidation state of the metal site prior to oxidation as 1+ (▲), 2+ (●), or 3+ (■).

The results in Figure 3.6a illustrate that ρ_O plays an important role, but it is certainly not the only factor that should be considered and does not entirely dictate the reactivity of the abstracting oxo species. Even so, the value of ρ_O can potentially be used as a rough guideline for catalyst design purposes. If the goal is to design a MOF that can readily form a metal-oxo active site and activate methane, we predict that ρ_O should ideally be ~0.5 to increase the likelihood of having a C−H activation barrier less than 1 eV but an exothermic active site formation energy. The

ideal M−O bond distance appears to be ~1.65−1.7 Å based on an analogous argument (Figure 3.6b).

Within the framework of the metal-oxo wall theory,[236] periodic trends in the stability of terminal metal-oxo complexes can be directly related to the M−O bond order, which as a result may impact the reactivity toward C−H activation. To probe this behavior, we computed effective bond orders based on the DDEC method[231] and observed that lower DDEC-based M−O bond orders are qualitatively correlated with reduced active site stabilities and increased reactivities toward C−H activation (Figure B.3b). The spin density on the oxo ligand is inversely correlated with the M−O bond order (Figure B.4), which is consistent with the fact that DDEC bond orders are computed based on partitions of the electron and spin magnetization densities.[231] It has also previously been proposed that the spin multiplicity can be a major factor governing C−H activation processes, such as via the exchange-enhanced reactivity model;[240,241] however, we find no strong relationships between spin state (or spin density on the metal) and reactivity, likely due to the diverse geometries, coordination environments, ligands, metals, and resulting electronic structures present in the investigated MOF dataset. Additional correlations can be found in Figure B.5.

3.4 Breaking the Scaling Relationship

The presence of a single active site stability−reactivity scaling line and the effectively constant methane physisorption energy can be used to significantly accelerate the MOF screening process, but these results also imply that possible values for the active site stability and reactivity are limited to those along the scaling line. In general, attempts to identify cases where a material class "breaks" from a given scaling relationship are highly sought after for catalyst design purposes and to better understand otherwise anomalous behavior.[106,242–244] We hypothesize that one approach to break

from the proposed scaling relationship is by coupling the formation of the active site with a separate bond-breaking or bond-making event, as this process may change ΔE_O without significantly altering the reactivity of the metal-oxo species compared to that predicted from the scaling line.

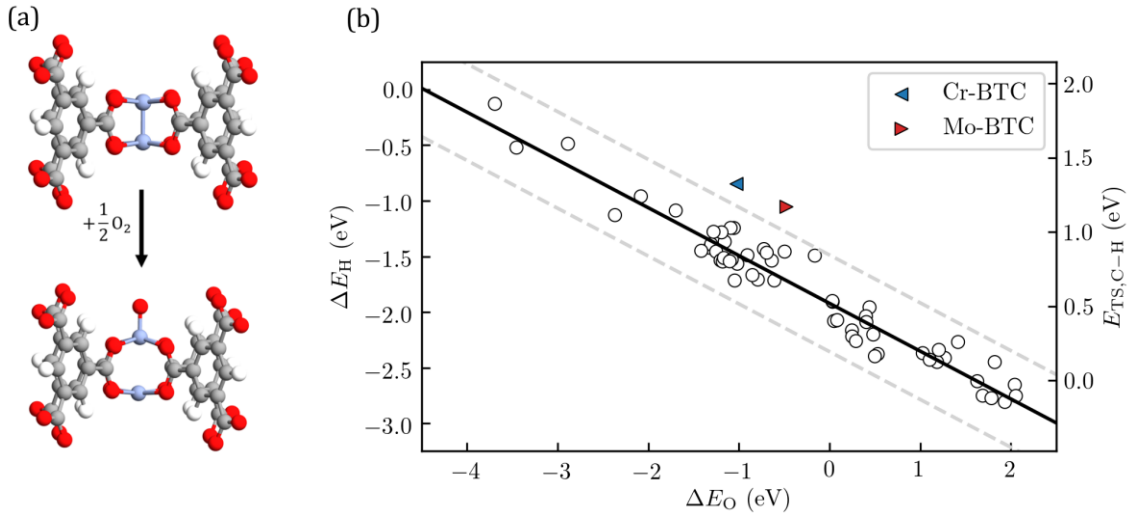


Figure 3.7. a) Increase in the metal-metal distance (and subsequent decrease in metal-metal bond order) in Cr-BTC and Mo-BTC following formation of a $[MO]^{2+}$ active site. A representative portion of the unit cell is shown for Cr-BTC. Color key: Cr (light blue), O (red), C (grey), H (white). b) H-affinity, ΔE_H, as a function of the active site formation energy, ΔE_O, for Cr-BTC (blue, ◀) and Mo-BTC (red, ▶), which significantly deviate from the scaling relationship of $\Delta E_H = -0.43\Delta E_O - 1.92$ eV (black solid line). For reference, the grey dashed lines represent three standard deviations from the scaling line.

For example, it has previously been suggested that Cr-BTC and Mo-BTC (BTC^{3-} = 1,3,5-benzenetricarboxylate) exhibit metal-metal bonding interactions.[185,189] As depicted in Figure 3.7a, we predict that the M−M distance in Cr-BTC will increase from 1.90 Å to 2.85 Å, while that in Mo-BTC will increase from 2.11 Å to 2.68 Å upon formation of a metal-oxo active site. These increases correspond to a significant reduction in the M−M bonding interaction, with a decrease in the DFT-computed M−M bond order of 0.80 and 0.92 for Cr-BTC and Mo-BTC, respectively,

based on the DDEC method.[231] These calculations are also consistent with prior experimental observations of bond lengths in these materials.[185,189]

For Cr-BTC and Mo-BTC, we find that the DFT-calculated H-affinities are 0.63 eV and 0.66 eV more endothermic than those predicted by the scaling relationship using the computed ΔE_{O} values (Figure 3.7b). We attribute this to the breaking of a metal-metal bond upon formation of the metal-oxo active site. While not to the same degree as with Cr-BTC and Mo-BTC, many of the other MOFs that deviate most from the scaling line exhibit changes in coordination geometry (e.g. cation-substituted MOF-5,[207] as shown in Figure B.6). Related to the concept of concerted bond-breaking or bond-making events, oxidation reactions with mechanisms that involve reversible linker displacement[245–247] may result in situations where the MOF breaks from the proposed linear scaling relationship. In addition to the energetics associated with breaking the metal-linker bond, this process can potentially lead to positive or negative deviations from the scaling line by introducing changes to the coordination number, coordination geometry, and/or oxidation state of the metal. One could also envision a scenario where noncovalent interactions (e.g. H-bonding) would stabilize the metal-oxo species without significantly altering the reactivity of the metal-oxo species toward C−H activation (Figure B.7).[c]

In prior work, it was proposed that materials that more effectively delocalize changes in charge upon formation of the metal-oxo active site exist on a less reactive scaling line (i.e. more endothermic ΔE_{H} for a given ΔE_{O}) than materials that are more electronically insulating.[105] To address this, we computed the degree of charge delocalization (Equation B.1) for each MOF and

[c] Following publication of this work, Sours and Patel et al.[424] demonstrated a similar idea for electrocatalytic oxygen reduction with MOFs.

found that the MOFs considered in this work have ligands ranging from nearly redox-innocent to highly redox non-innocent when the proposed metal-oxo active site is formed (Figure B.8, Table B.3). We conclude that the degree of charge delocalization does not result in distinct scaling relationships for the MOFs analyzed in this work. In fact, if charge delocalization can help stabilize otherwise reactive metal-oxo species,[248] then we anticipate that ΔE_{O} can be greatly improved without causing the MOF to shift to a significantly less reactive scaling line. As a result, we would recommend deliberate incorporation of redox non-innocent ligands in MOFs[28,249] to increase the likelihood of designing MOFs for oxidative C−H bond activation reactions in cases where formation of the active site is the rate-limiting step.

3.5 Conclusion

By using periodic density functional theory to screen 60 MOF structures with a wide range of metal sites and coordination environments for the oxidative C−H bond activation of methane, we observe a clear tradeoff between the thermodynamics of forming the metal-oxo active site, ΔE_{O}, and reactivity of the active site toward C−H activation. We propose a linear scaling relationship relating the active site formation energy and the H-affinity, ΔE_{H}, that is generally applicable to MOFs with spatially isolated open metal sites. Through the scaling relationship between ΔE_{O} and ΔE_{H}, the consistent methane physisorption energy among MOFs with accessible metal-oxo sites, and the universal H-affinity scaling relationship,[105] it is possible to predict both active site stability and the barrier for methane C−H activation from just ΔE_{O}, reducing the computational burden from at least four DFT calculations per MOF to only two. This has significant implications for the predictive design of MOF catalysts from first-principles calculations and provides a robust rule-of-thumb for choosing which MOFs to investigate experimentally. We also show that it is possible

to design materials that significantly break from the proposed scaling line by coupling oxidation of the metal center with changes in bonding behavior of the remaining framework atoms. As a whole, we anticipate that the structure−activity relationships developed in this work will enable greatly accelerated computational screening and provide a set of promising catalyst design criteria for MOFs that can directly activate methane at moderate reaction conditions.

Chapter 4: The Effects of Ligand-Exchange and Two-State Reactivity for C–H Bond Activation

In this chapter, quantum-chemical calculations are used to investigate a family of metal−organic frameworks (MOFs) containing triazolate linkers, M_2X_2(bbta) (M = metal, X = bridging anion, H_2bbta = 1*H*,5*H*-benzo(1,2-d:4,5-d′)bistriazole), for their ability to form terminal metal-oxo sites and subsequently activate the C−H bond of methane. By varying the metal and bridging anion in the framework, we show how to significantly tune the reactivity of this series of MOFs. The electronic structure of the metal-oxo active site is analyzed for each combination of metal and bridging ligand, and we find that significant spin density localized on the oxo ligand (at stationary points along the reaction coordinate) is not necessarily the primary driving force behind low predicted C−H activation barriers. Nonetheless, the important role that the spin density can have on the reactivity of individual spin channels is discussed. For the Mn- and Fe-containing frameworks in particular, the high reactivity can be attributed to the formation of antiferromagnetically coupled spin density on the metal binding site and the abstracting O atom at the transition state for C−H bond activation, which is not necessarily present in the initial $[MO]^{2+}$ state. In cases where there is a ferromagnetic to antiferromagnetic transition *en route* to the transition state, this is an example of two-state reactivity, albeit one with a fixed spin multiplicity.

This chapter is adapted from the following peer-reviewed article: A.S. Rosen, J.M. Notestein, R.Q. Snurr. “High-Valent Metal-Oxo Species at the Nodes of Metal–Triazolate Frameworks: The Effects of Ligand Exchange and Two-State Reactivity for C–H Bond Activation.” *Angew. Chem. Int. Ed.*, **59**, 19494–19502 (2020).

4.1 Introduction

High-valent transition metal-oxo species have been invoked as reactive intermediates for a wide range of oxidation reactions with both enzymatic and synthetic catalysts.[250–252] Many terminal oxo complexes are powerful oxidants and serve as promising active site motifs for the activation of strong C−H bonds, such as those of light alkanes, at moderate reaction conditions.[250–254] However, terminal oxo complexes are often difficult to isolate and probe experimentally as a direct consequence of their inherent reactivity, especially those containing mid-to-late transition metals.[236,255–257] In the synthesis of molecular metal-oxo complexes, sterically bulky ligands are typically employed to protect highly reactive metal-oxo units, as it can be difficult to prevent them from dimerizing or deactivating in solution.[258] A different strategy is to leverage the well-defined metal sites incorporated within the inorganic nodes of metal−organic frameworks (MOFs) to support reactive metal-oxo species in a heterogeneous catalyst.[94] This approach takes advantage of the site isolation, uniformity, and periodicity of metal sites when present as part of the secondary building units of MOFs.[259] Enabled by the vast library of inorganic nodes and organic linkers, the highly tunable nature of MOFs also makes it possible to design active sites with tailored coordination environments for a given reaction of interest.[171,172]

Despite the many thousands of MOFs that have been synthesized to date,[14,44] the structural diversity of MOFs has hardly been explored for this challenging class of reactions. As a result, relatively little is known about how to enhance the reactive properties of these materials despite the modular nature of the underlying secondary building units.[3] To date, nearly all MOFs studied for their potential ability to support terminal metal-oxo species contain metals connected to one

another with carboxylate linkers that yield an all-oxido coordination environment, as is the case for the metal sites of Fe-MOF-74 (and its structural analogues),[168,177,178,260,261] Mn-exchanged MOF-5,[262] $M_3(btc)_2$ (M = Cr, Fe, Co, Ni, Cu, Zn; H_3BTC = 1,3,5-benzenetricarboxylic acid),[263] FeM_2-PCN-250 (M = Fe, Mn, Co, Ni; PCN = Porous Coordination Network),[264] and Fe-MIL-100 (MIL = Materials Institut Lavoisier).[179,265] Fe-BTT (H_3BTT = 1,3,5-benzenetristetrazolate), which has tetrazolate linkers, is one exception; however, it is believed that framework defects (rather than the computationally investigated, crystallographic framework sites) are responsible for the material's reactivity toward ethane in the presence of N_2O.[116]

Metal−triazolate frameworks represent one family of MOFs that have not yet been widely explored for activating the strong C−H bonds of light alkanes. Like many metal−azolate frameworks (MAFs),[266] metal−triazolate frameworks with open metal sites often contain anionic ligands within the first coordination sphere of the metal, which can typically be post-synthetically exchanged.[267] This represents a promising route to alter the reactivity of a given framework, beyond the more common design handles of metal identity and linker functionalization. This is further motivated by prior studies of transition metal-oxo complexes, which have shown that the presence and identity of coordinating, anionic ligands can greatly enhance rates of hydrogen atom transfer reactions.[254,268,269] Of the previously synthesized MAFs containing open metal sites, M-BTT,[195] M-BTTri (H_3BTTri = 1,3,5-tri(1*H*-1,2,3-triazol-5-yl)benzene),[216,270] and M-BDTriP ($H_3BDTriP$ = 5,5′-(5-(1*H*-pyrazol-4-yl)-1,3-phenylene)bis(1*H*-1,2,3-triazole))[216] all have μ_4-Cl^- ligands, $M_2Cl_2(bbta)$ (H_2bbta = 1*H*,5*H*-benzo(1,2-d:4,5-d′)bistriazole)[271] and $M_2Cl_2(btdd)$ (H_2btdd = bis(1*H*-1,2,3-triazolo[4,5-b],[4′,5′-i])dibenzo[1,4]dioxin))[192] have μ-Cl^- ligands, and M-MFU-4*l*

(MFU = Metal−Organic Framework Ulm) has terminal Cl^- ligands[212,272] that can potentially be exchanged. Post-synthetic ligand-exchange in the aforementioned materials has led to the discovery of MOFs with enhanced water uptake,[273] O_2 binding strengths,[73] hydrogen storage properties,[274] and turnover frequencies for CO_2 reduction[275] among several other applications.

MOFs with triazolate linkers may also be promising for stabilizing reactive metal-oxo active sites given that triazolate groups are stronger donors than the more widely used carboxylate linkers.[276] In the context of one-electron oxidation reactions and experiments involving the chemisorption of O_2 at open metal sites, increasing the basicity of the linker – a proxy for its donor strength – has been shown to increase the redox activity of the framework.[73,216,276] Separate studies have also shown that this often serves to increase the framework's chemical stability.[276] The greater redox activity of the framework is particularly relevant, as mechanistic studies based on density functional theory (DFT) calculations frequently indicate that the reactivity of MOFs with open metal sites are limited by their ability to form high-valent metal-oxo active species.[32,264,265]

In the present study, we focus on the family of metal−triazolate frameworks with the general formula M_2X_2(bbta) (M = divalent metal, X = monovalent bridging anion). M_2X_2(bbta) is topologically analogous to the MOF-74 family, containing a high density of coordinatively unsaturated metal sites with square pyramidal coordination geometries and a honeycomb-like crystal structure (Figure 4.1). Many members of the M_2X_2(bbta) family are relatively stable, maintaining their crystallinity up to 400 °C in a N_2 atmosphere and withstanding exposure to H_2O_2, Cl_2, Br_2, NH_3, and other reactive gases.[25,192,194] Perhaps most notably, like several of the aforementioned MAFs, the M_2X_2(bbta) family is highly tunable, both with regards to the metal

identity as well as the bridging monovalent anions connecting each metal center. To date, M_2Cl_2(bbta) (M = Mn, Fe, Co, Ni),[d] $M_2(OH)_2$(bbta) (M = Co, Ni), Ni_2F_2(bbta), Ni_2Br_2(bbta), and the large-pore analogues Cu_2Cl_2(btdd) and $V_2Cl_{2.8}$(btdd) have all been experimentally synthesized.[193,194,271,273,275,277–279] For reference, the pores of these materials are quite large, with diameters of approximately 13 Å and 23 Å for M_2X_2(bbta) and M_2X_2(btdd), respectively.[280]

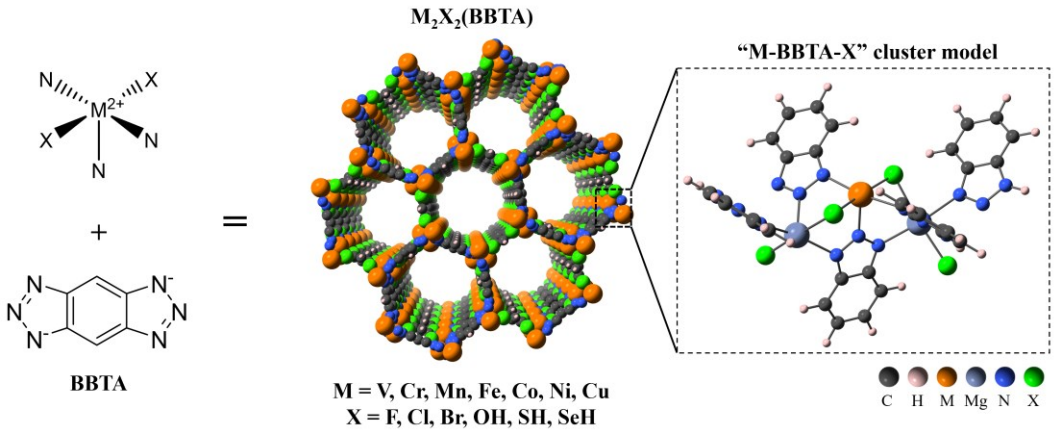


Figure 4.1. Structure of the honeycomb-like M_2X_2(bbta) framework and the corresponding Mg-diluted finite cluster model, denoted M-BBTA-X, used in this work. We consider M-BBTA-X frameworks composed of M = {V, Cr, Mn, Fe, Co, Ni, Cu} and X = {F, Cl, Br, OH, SH, SeH}.

Using DFT, we investigate the effects of exchanging the metal cation (M = V, Cr, Mn, Fe, Co, Ni, Cu) and bridging ligand (X = F, Cl, Br, OH, SH, SeH) on the framework's ability to form reactive metal-oxo motifs for the oxyfunctionalization of strong C−H bonds. While some of these combinations of metals and ligands have not yet been synthesized or may not yield frameworks capable of forming (meta-)stable metal-oxo motifs, we include a wide range of metals and ligands for the purposes of identifying overarching structure−reactivity relationships across this family of

[d] Since publication of this work, Cu_2Cl_2(bbta) has also been experimentally synthesized.[628]

tunable materials. As a result of this study, we show how the choice of linker and anions within the first coordination sphere can be used to increase the stability of high-valent metal-oxo sites that are reactive toward strong C−H bonds in the M_2X_2(bbta) family of metal−triazolate frameworks. We also use the M_2X_2(bbta) family to demonstrate the important role of electron spin for C−H bond activation via terminal metal-oxo species, while simultaneously clarifying several misconceptions about the radical-like character of the metal-oxo motif.

4.2 Methods

Charge-neutral finite cluster models containing 75 atoms (X = F, Cl, Br) or 79 atoms (X = OH, SH, SeH) were carved from periodic structures, as depicted in Figure 4.1. As done in several prior studies,[281] the clusters were diluted with Mg^{2+} cations to reduce the computational complexity, particularly with regards to converging to the ground state spin states. We refer to the Mg-diluted models as M-BBTA-X. We consider M-BBTA-X frameworks composed of M = {V, Cr, Mn, Fe, Co, Ni, Cu} and X = {F, Cl, Br, OH, SH, SeH}. For the purposes of identifying structure−property relations, all structures were assumed to have the square pyramidal coordination environment of the experimentally synthesized Mn, Fe, Co, and Ni bbta-based frameworks. Periodic DFT calculations in VASP[122,123] indicate that the cluster models are accurate representations of the larger periodic frameworks (Figure C.7).

Finite cluster DFT calculations were carried out via the Gaussian 16, Rev. A03 program[282] using the hybrid-level B3LYP exchange−correlation functional,[283–285] Grimme's D3 dispersion correction[126] with Becke-Johnson (BJ) damping,[127] and the def2-TZVP[286] basis set. Oxidation reactions are known to suffer from significant self-interaction error without Hartree-Fock

exchange, so the use of a hybrid functional is expected to be important for accurately describing the thermodynamics of metal-oxo formation. For comparison purposes, select calculations were also carried out using the M06-L functional,[287] as shown in Figure C.8. However, we note that using the M06-L and M06[288] functionals leads to significant difficulties in converging the self-consistent field (particularly for X = OH, SH, SeH) regardless of the electronic optimization algorithm, which prevents their widespread use throughout this screening study. As described in the "Model Assumptions" section in the Appendix, while the B3LYP functional has previously been shown to over-stabilize high-spin states in several cases,[228,229] the B3LYP-D3(BJ) and M06-L functionals predict the same ground state spin state for every M-BBTA-X (X = F, Cl, Br) cluster model studied in this work, both in the guest-free and oxidized states.

An "ultrafine" integration consisting of 99 radial shells and 590 angular points per shell was used for all calculations, and symmetry constraints were disabled. Partially constrained geometry optimizations were carried out to approximate the rigidity of the framework such that only the framework atoms within the first and second coordination spheres of the transition metal binding site were allowed to relax (in addition to any guest molecules). As is common for first-row transition metals, there are several physically plausible spin states for each structure reported in this work. Generally, results presented in the main text correspond to a predicted high-spin ground state, with a few exceptions (such as for the intermediate-spin $[CoO]^{2+}$ sites). Full details regarding the spin state analysis can be found in Figures C.1−C.6 and Tables C.1−C.15. Partial atomic charges were evaluated using the Charge Model 5 (CM5) method[289] and spin densities were calculated using a Hirshfeld population analysis.[290] Wiberg bond orders were computed in the

natural atomic orbital basis via the natural bonding orbital (NBO) method implemented in the NBO v.7.0.2 code.[291,292]

Thermochemical corrections were applied using the ideal gas, harmonic oscillator, rigid rotor, and particle-in-a-box approximations[293] with analytically computed vibrational frequencies. As implemented in GoodVibes v.3.0.1,[294] quasi-harmonic corrections to the enthalpy were applied with a vibrational frequency cutoff of 50 cm^{-1}.[295] A frequency scaling factor of 0.985 was applied for the thermochemical corrections.[296] All thermochemical quantities were computed at 298.15 K.

Automatic density fitting of the Coulomb integrals was used to increase the computational efficiency of the M06-L/def2-TZVP calculations. For the B3LYP-D3(BJ)/def2-TZVP calculations, the spin states in Table C.1 were confirmed to be local minima or transition states based on the lack or presence of one imaginary vibrational mode, respectively. Vibrational analyses were not carried out for the less stable spin states (except where thermochemical corrections were computed, such as in Figure 4.7) due to the high computational cost of these calculations and the large number of investigated spin states. Additional computational details can be found in the Appendix.

4.3 Results and Discussion

Throughout this work, we primarily focus on two key reactivity parameters: the thermodynamic favorability of forming a terminal metal-oxo site and the subsequent barrier for C−H activation. The enthalpy of metal-oxo formation, ΔH_O, can be described by

$$\Delta H_{\mathrm{O}} = H_{\mathrm{MOF-O}} - H_{\mathrm{MOF}} - 0.5H_{\mathrm{O_2}} \quad (1)$$

where $H_{\mathrm{MOF-O}}$, H_{MOF}, and $H_{\mathrm{O_2}}$ represent the enthalpies of the oxidized MOF, adsorbate-free MOF, and gas-phase O_2, respectively (Figure 4.2a). While O_2 is a natural choice for the gaseous reference state, this decision is in principle arbitrary, and all ΔH_{O} values would simply be linearly offset if another molecule (e.g. N_2O, H_2O_2) were used as the reference state. As a probe for each MOF's ability to activate strong C−H bonds, we computed the barrier for the H-abstraction of methane (Figure 4.2b). Methane was chosen due to the societal need for a catalyst that can directly convert methane to methanol at moderate reaction conditions[157,166,167] and its small size that permits the computational investigation of many catalyst structures. The trends in this work are expected to be generalizable to light hydrocarbons beyond methane, as C−H activation barriers tend to scale with the bond dissociation enthalpies.[297] The apparent activation enthalpy for breaking the C−H bond of methane is given by

$$\Delta H_{\mathrm{a,C-H}}^{\mathrm{app.}} = H_{\mathrm{C-H}}^{\ddagger} - H_{\mathrm{MOF-O}} - H_{\mathrm{CH_4}} \quad (2)$$

where $H_{\mathrm{C-H}}^{\ddagger}$ is the enthalpy of the transition state for H-abstraction, and $H_{\mathrm{CH_4}}$ is the enthalpy of gas-phase CH_4. The apparent and intrinsic activation enthalpy, $\Delta H_{\mathrm{a,C-H}}$, are related to one another via

$$\Delta H_{\mathrm{a,C-H}} = \Delta H_{\mathrm{a,C-H}}^{\mathrm{app}} + \left|\Delta H_{\mathrm{CH_4\ ads.}}\right| \quad (3)$$

where $\Delta E_{\mathrm{CH_4\ ads.}}$ is the enthalpy of methane adsorption at the metal-oxo site in the MOF. Prior work has shown that, assuming no significant variations in the degree of confinement, the adsorption energy of methane at isolated metal-oxo sites in MOFs (and cation-exchanged

zeolites[219]) is effectively constant with a value of approximately 17 kJ/mol, consistent with weak physisorption.[71] As described in detail in the Appendix (Tables C.1−C.15 and Figures C.1−C.6), we considered multiple plausible spin states for each DFT calculation and report results for the low-energy spin states in most cases.

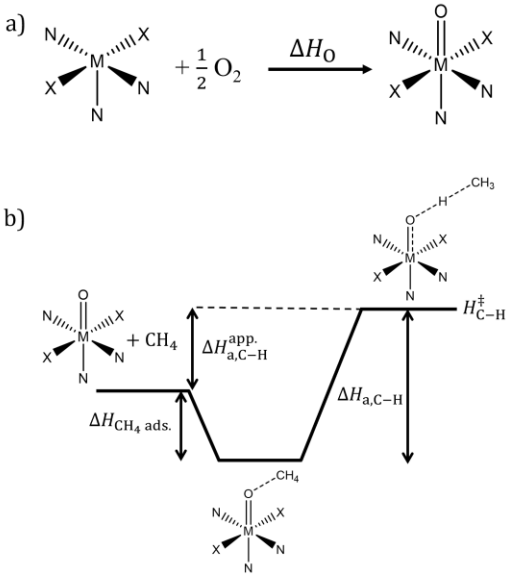


Figure 4.2. a) Depiction of the metal-oxo formation enthalpy, ΔH_{O}. b) Depiction of the apparent, $\Delta H^{\mathrm{app}}_{\mathrm{a,C-H}}$, and intrinsic, $\Delta H_{\mathrm{a,C-H}}$, activation enthalpies for cleaving the C−H bond of methane as they relate to the transition state energy, $H^{\ddagger}_{\mathrm{C-H}}$, and methane adsorption energy, $\Delta H_{\mathrm{CH_4\ ads.}}$.

The thermodynamic feasibility of forming the metal-oxo active site as a function of both metal and bridging ligand identity in the M-BBTA-X series of MOFs is shown in Figure 4.2a. As found in prior studies,[71,264] there are clear periodic trends such that later transition metals are more difficult to oxidize than earlier transition metals. This is to be expected, given that higher d electron counts will lead to a greater occupation of π-antibonding molecular orbitals that weaken the metal-oxo bond.[236] One notable exception is for the Mn-containing frameworks, which have comparable

(if not more endothermic) metal-oxo formation enthalpies than the Fe-containing frameworks for all bridging ligands except μ-OH^-. From the perspective of forming a terminal metal-oxo site, these results suggest that there is not always a thermodynamic benefit in using Mn over Fe in this family of materials. Many of the Mn- and Fe-containing frameworks are also close to thermoneutral with regards to the ΔH_O value, whereas the Cr- and Co-containing frameworks have highly exothermic and endothermic $[MO]^{2+}$ formation enthalpies, respectively, when using an O_2 reference state. We note that this does not necessarily imply $[MO]^{2+}$ active sites in Co-containing frameworks would be impossible to generate. Rather, if formed, the population of these sites would be expected to be low since recombination to yield gas-phase O_2 would be thermodynamically preferable.

The identity of the bridging ligand can also greatly influence the thermodynamic feasibility of forming a metal-oxo species. In general, the use of OH (and to a lesser extent, SH or SeH) bridging ligands rather than the halides is recommended for the purposes of forming metal-oxo active sites (Figure 4.2a). After clustering the results by X = {F, Cl, Br} or X = {OH, SH, SeH}, more basic bridging ligands tend to yield metal centers with more exothermic ΔH_O values (e.g. $\Delta H_{O,M-bbta-F} < \Delta H_{O,M-bbta-Cl} < \Delta H_{O,M-bbta-Br}$) for the early transition metals. This is analogous to the predicted trends for O_2 adsorption at the open metal sites of the M_2X_2(bbta) family,[73] wherein more basic bridging ligands resulted in a greater degree of electron-donation to the metal center, increasing the thermodynamic favorability of the O_2 binding process. Nonetheless, this behavior becomes less pronounced as one goes across the periodic table (Figure 4.2b and Figure 4.2c), with an eventual switch for the late transition metal cations (i.e. M = Co, Ni, Cu) for which more basic ligands tend to yield less thermodynamically stable metal-oxo sites

(e.g. $\Delta H_{\mathrm{O,M-bbta-Br}} < \Delta H_{\mathrm{O,M-bbta-Cl}} < \Delta H_{\mathrm{O,M-bbta-F}}$). We attribute this phenomenon to a tradeoff between the basicity of the bridging ligand (which increases as $Br^- < Cl^- < F^-$) and the partial charge on the metal (which increases as $F^- < Cl^- < Br^-$ based on electronegativity trends). For the early transition metals that can readily form the hypothesized $[MO]^{2+}$ site, a ligand with a greater basicity is expected to better stabilize the increased charge on the metal. However, for the later transition metals where little or no change in charge on the metal occurs (Figure C.9), the basicity of the ligand has less of an influence than its electronegativity.

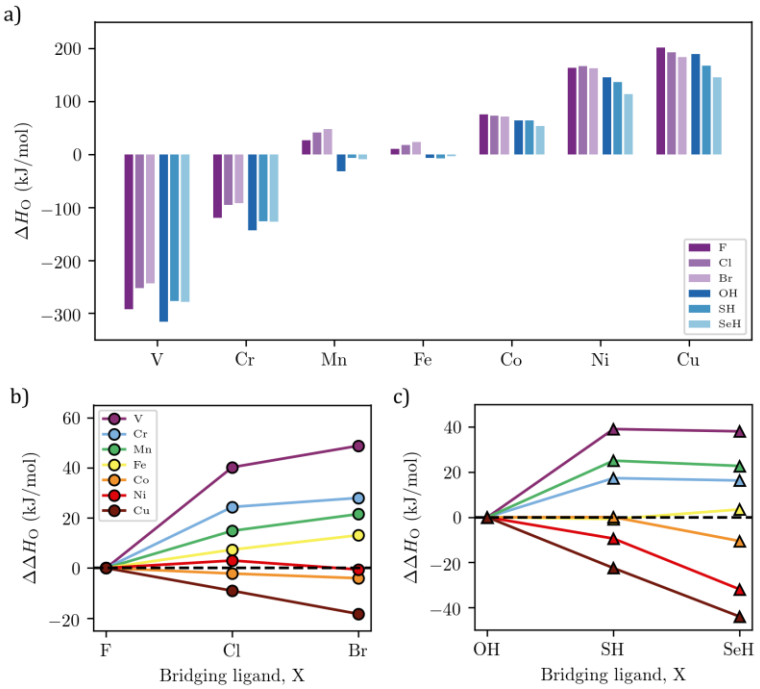


Figure 4.3. a) Enthalpy of metal-oxo formation, ΔH_O, in M-BBTA-X as a function of metal binding site and bridging ligand identity. (b, c) Relative change in ΔH_O, denoted $\Delta\Delta H_O$, as a function of bridging ligand with respect to X = F (b, circles) and X = OH (c, triangles) for a given transition metal. The sign convention is such that a positive value of $\Delta\Delta H_O$ indicates a more endothermic value of ΔH_O compared to the reference and vice versa. Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

Beyond the choice of anionic bridging ligands, the donor strength of the organic linker is also likely to influence the redox activity of the framework in cases where an $[MO]^{2+}$ site can form. To test this, we compared the M−O formation enthalpies of the Fe-BBTA-X cluster models with those of $Fe_{0.1}Mg_{1.9}$(dobdc) (H_4dobdc = 2,5-dihydroxybenzene-1,4-dicarboxylic acid) and $Fe_3(\mu_3\text{-}O)(HCOO)_6$ (Figure C.10), the latter of which is representative of the node in the Fe-MIL-100/101 (and Fe-PCN-250) series. These MOFs were chosen as points of comparison, as they have been shown to activate strong C−H bonds via a presumed Fe(IV)-oxo intermediate[168,264,265] and have carboxylate linkers, which are weaker donors than the triazolate linkers of the M_2X_2(bbta) series.[276]

At the B3LYP-D3(BJ)/def2-TZVP level of theory,[126,127,283–286] ΔH_O is predicted to be +24 kJ/mol for $Fe_{0.1}Mg_{1.9}$(dobdc) and +35 kJ/mol for $Fe_3(\mu_3\text{-}O)(HCOO)_6$. While Fe-BBTA-Br has a ΔH_O value nearly equivalent to that of $Fe_{0.1}Mg_{1.9}$(dobdc), all other members of the Fe-BBTA-X series have more favorable Fe(IV)-oxo formation enthalpies, with Fe-BBTA-X (X = OH, SH, SeH) having ΔH_O values that are ~34 kJ/mol more exothermic than $Fe_{0.1}Mg_{1.9}$(dobdc) and ~45 kJ/mol more exothermic than Fe-MIL-100/Fe-MIL-101/Fe-PCN-250 (Table C.16). This supports the hypothesis that the N donors of the more strongly donating triazolate linker is better stabilizing the oxidized state of the metal center than the more weakly donating – but commonly studied – carboxylate linkers. Although the donor strength of the linker is not the only factor that can influence the metal-oxo formation enthalpy (other properties, such as coordination number and spin state are expected to play a role), the use of triazolate linkers, in comparison to the widely used carboxylate linkers, is one route to increase the thermodynamic favorability of forming terminal metal-oxo sites in MOFs with open metal sites.

Up to this point, we have simply referred to the $[MO]^{2+}$ sites as "metal-oxo" species in line with formal oxidation state counting rules. However, the true electronic structure of these $[MO]^{2+}$ sites is expected to exist on a spectrum with a varying degree of oxo and oxyl character (Figure 4.4a).[252] Several complementary methods can be used to assign a "best" resonance structure. Here, we use the partial atomic charge on the O atom, q_O, the spin density on the O atom, ρ_O, and the computed M−O bond order to probe the properties of the $[MO]^{2+}$ site. The partial atomic charge is not the same as the expected oxidation state,[298] so we have computed the partial charge on the O atoms of H_2O and H_2O_2 as reference points for O^{2-} and O^{-} oxidation states, respectively.

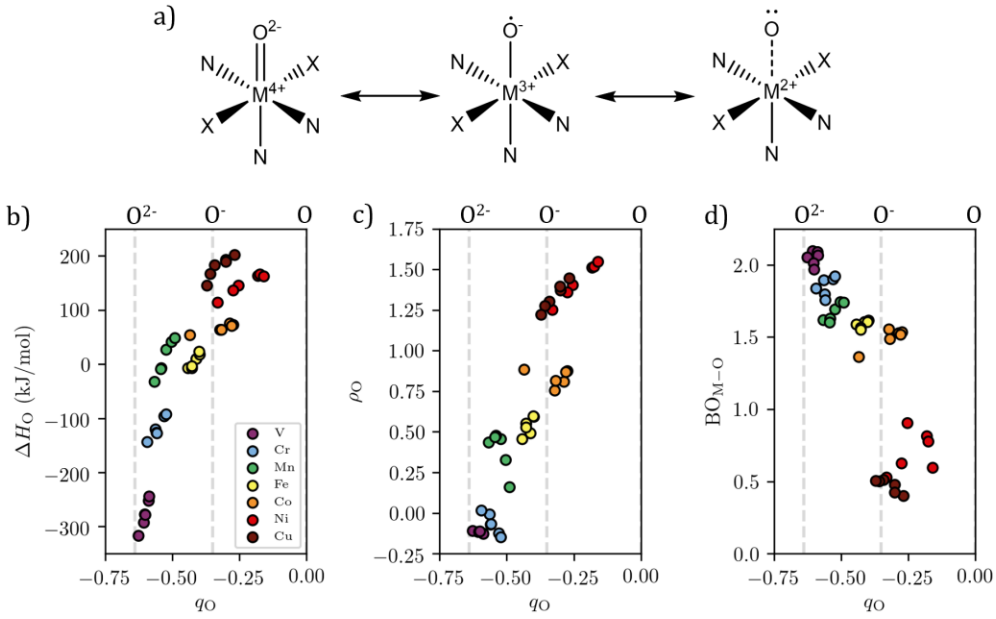


Figure 4.4. a) Schematic of several different resonance structures associated with the $[MO]^{2+}$ motif, showing the range of bond orders, oxidation states, and unpaired electrons. The rightmost structure is an extreme hypothetical case with no formal bond between the metal and O atom. b) Metal-oxo formation enthalpy, ΔH_O, as a function of the CM5 partial atomic charge on the O atom, q_O. c) Hirshfeld spin density on the O atom of the $[MO]^{2+}$ site, ρ_O, as a function of q_O. d) Wiberg M−O bond order in the natural atomic orbital basis, BO_{M-O}, as a function of q_O. For all subplots, a negative ρ_O value represents an antiferromagnetic alignment with the spin density on the metal center. Reference q_O values for H_2O and H_2O_2 are shown as dashed vertical lines, which

correspond to O formal oxidation states of −2 and −1, respectively. Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

As shown in Figure 4.4b, the oxo/oxyl ligand exhibits a wide range of partial charges. In addition, the thermodynamic feasibility of forming the $[MO]^{2+}$ site is correlated with the degree of charge transfer for a given metal identity. Based on the partial charge and spin density analysis (Figure 4.4c), the V-, Cr-, Mn-, and Fe-containing MAFs can be best described as having at least some degree of oxo-like character. The Co-, Ni-, and Cu-containing frameworks, however, are best described as having predominantly oxyl character. The computed M−O bond orders (Figure 4.4d) are in agreement with this electronic structure assignment, albeit with a clearer divide between the Co-containing MAFs and the Ni- or Cu-containing MAFs. The computed bond orders are qualitatively consistent with what would be expected on the basis of the "oxo wall" principle.[236]

With a more thorough understanding of the $[MO]^{2+}$ sites in the M-BBTA-X materials, we shift our focus to their reactivity, specifically with regards to activating the strong C−H bond of methane, which has a bond dissociation enthalpy of 439 kJ/mol at 298 K.[165] Excluding the V-BBTA-X series, the metal-oxo site in every MOF studied in this work has an apparent C−H activation enthalpy less than 100 kJ/mol, with many having an apparent activation enthalpy around 50 kJ/mol or less (Figure 4.4). In contrast with what is typically assumed,[299] we emphasize that activating the C−H bond of methane is predicted to be relatively easy in general, and efforts should instead be focused on forming the $[MO]^{2+}$ active site. Importantly, the low activation enthalpies for breaking the C−H bond of methane are not restricted to late transition metals, some of which (particularly the Ni- and Cu-containing frameworks) would be unlikely to form the pre-requisite

$[MO]^{2+}$ sites. Unlike the metal-oxo formation enthalpies in Figure 4.2, the differences in reactivity as a function of bridging ligands are less consistent, not showing any widespread trends except that the X = {OH, SH, SeH} bridging ligands in the Mn-BBTA-X series have significantly higher C−H activation barriers than the halide bridging ligands.

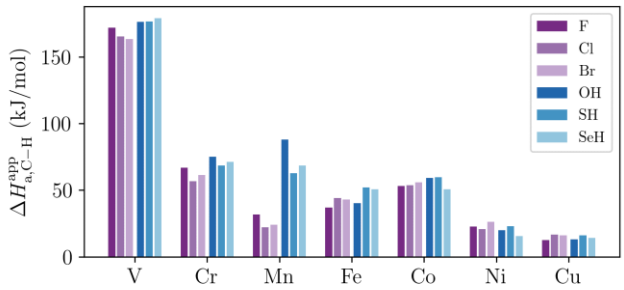


Figure 4.5. Apparent enthalpic barrier for C−H activation of methane, $\Delta H^{app.}_{a,C-H}$, in M-BBTA-X as a function of metal binding site and bridging ligand identity. Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

There has been significant debate in the literature regarding the effect of spin density on the reactivity of metal-oxo complexes.[237,238,300] Some have claimed that localized spin density on the oxo ligand is a necessary condition for them to be reactive,[237] whereas others claim that this is not necessarily the case.[238] On the basis of the apparent C−H activation barriers in Figure 4.5, we can conclude that significant spin density on the oxo ligand is not always required, as the Cr-BBTA-X series and several members of the Mn-BBTA-X series have a very low degree of spin density (Figure 4.4c) yet are highly reactive. Significant spin density on the oxo ligand at the transition state is also not always necessary, as exemplified with the Fe-BBTA-X series, which has negligible spin density (Figure 4.4a) but an apparent activation enthalpy of ~50 kJ/mol (Figure 4.5). These conclusions are in agreement with the perspective of Saouma and Mayer, where the

authors state that, while spin density on the oxo ligand can influence the reactivity, it is not an inherent requirement.[238,239] Indeed, prior work has shown that there is often,[71,176] but not always,[219] an inverse correlation between spin density on the abstracting atom and the barrier for C−H activation of the substrate. For the metal−triazolate frameworks studied here, there is not always strong correlation between the spin density on the oxo ligand and the C−H activation barrier (Figure C.13).

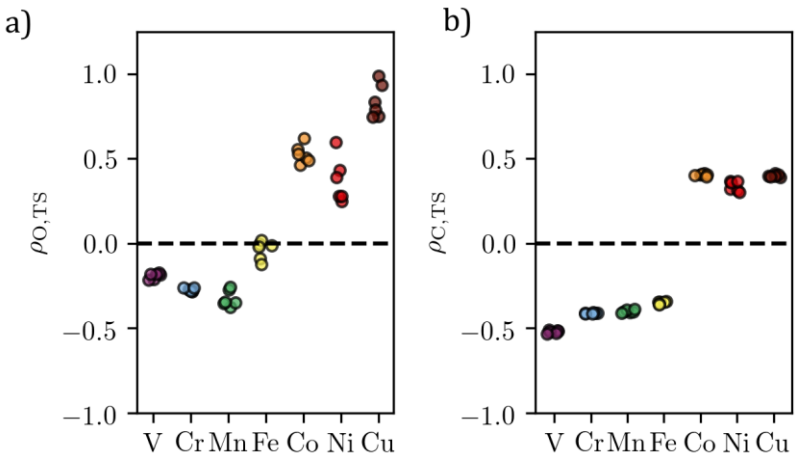


Figure 4.6. a) Hirshfeld spin density on the O atom of the metal-oxo site at the transition state (TS) for C−H activation, $\rho_{O,TS}$. b) Hirshfeld spin density on the C atom of the methane molecule at the TS, $\rho_{C,TS}$. For both subplots, a negative value for ρ represents an antiferromagnetic alignment with the spin density on the metal center. Horizontal jitter is added to distinguish overlapping points. Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

One important subtlety is the sign of the spin density on the oxo ligand with respect to the unpaired electron density on the metal. In this work, we have defined the metal as having a positive (i.e. spin up) spin density. Therefore, if $\rho_O < 0$, the oxo ligand is antiferromagnetically aligned with any spin density on the metal (i.e. it is a broken-symmetry state), whereas $\rho_O > 0$ indicates a ferromagnetic alignment. As shown in Figure 4.6a, the V-, Cr-, Mn-, and Fe-containing MAFs all

have an antiferromagnetically stabilized metal-oxo bond at the transition state for C–H activation, whereas the later transition metals have a ferromagnetically aligned metal-oxo (or, more likely, metal-oxyl) bond. Similar behavior is observed with regards to the spin density on the methyl radical (Figure 4.6b), which adopts the same spin orientation as the oxo ligand.

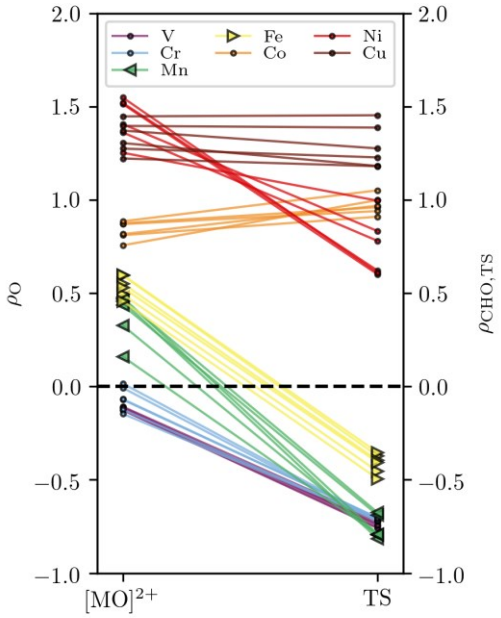


Figure 4.7. Comparison of the Hirshfeld spin density on the O atom of the metal–oxo site (before C−H activation), ρ_O, and the Hirshfeld spin density of the three-center C−H−O radical at the transition state (TS) for C−H activation, $\rho_{CHO,TS}$. A negative value for ρ represents an antiferromagnetic alignment with the spin density on the metal center. Mn and Fe are highlighted via larger triangle symbols. Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

By comparing the spin density analysis before and at the transition state for C−H activation, there are several catalytic candidates where the initial $[MO]^{2+}$ state has a ferromagnetically aligned ground state, but a spin flip must occur as the transition state is approached (Figure 4.7). This appears unique to the Mn- and Fe-containing frameworks (although the ferromagnetic and antiferromagnetic states for the $[MnO]^{2+}$ species are often close in energy, as shown in Table C.11).

The relative enthalpies of various spin states for two representative MOFs, Mn-BBTA-F and Fe-BBTA-F, are illustrated in Figure 4.7 and indicate that this change from a parallel alignment of spins (red) to antiparallel alignment (blue) occurs in both frameworks without a change in the overall spin multiplicity of the system. We note that an antiferromagnetically aligned iron-oxo complex could only be isolated computationally if a constrained geometry optimization was performed with the Fe−O bond artificially stretched beyond that of the initial state, mimicking the increase in bond length associated with the transition state. Therefore, this spin state appears to only form along the C−H activation reaction coordinate. This type of spin state behavior has previously been predicted in an $[FeO]^{2+}$ complex, where it was concluded that a spin-coupled electron pair must form as the Fe−O bond elongates to form a reactive oxyl radical *en route* to the transition state.[301,302]

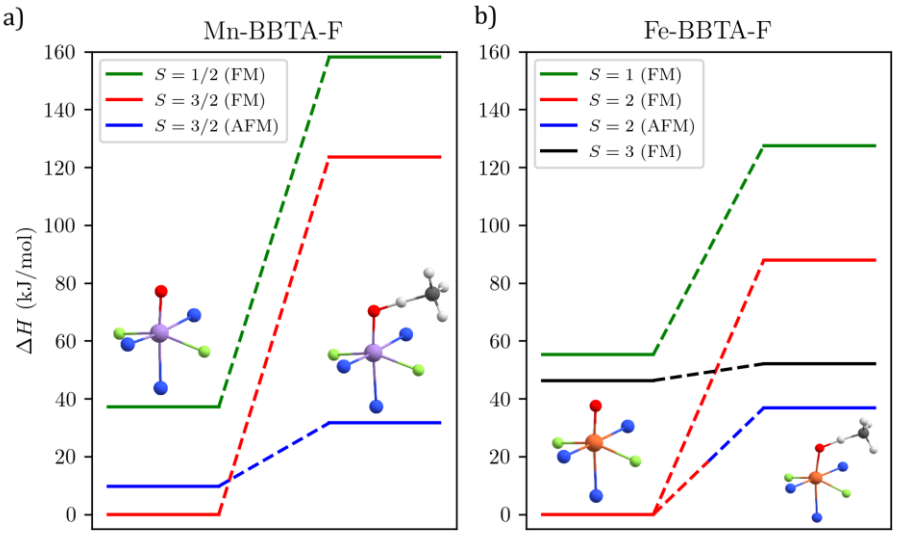


Figure 4.8. Apparent C−H activation enthalpies at the terminal metal-oxo sites of a) Mn-BBTA-F and b) Fe-BBTA-F for different spin states. A spin state labeled "FM" (ferromagnetic) or "AFM" (antiferromagnetic) refers to the alignment of the metal spin density and the spin density on the oxo ligand (or three-center C−H−O radical at the transition state). Representative portions of the

lowest energy structures are shown. Atom color key: Mn (purple), Fe (orange), O (red), Cl (green), N (blue), C (black), H (white). Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

The apparent change from ferromagnetic to antiferromagnetic coupling represents a pronounced example of two-state reactivity (TSR),[303,304] except unlike most TSR scenarios where there is a spin crossing between two different spin multiplicities, here the spin transition occurs between two arrangements of spins for a fixed spin multiplicity. Comparing the C−H activation barriers for Mn-BBTA-F along several relevant spin channels as one example, the pathway involving antiferromagnetic coupling of the unpaired electrons on the metal and the three-center C−H−O radical (Figure 4.8a, blue) reduces the activation enthalpy by 92 kJ/mol with respect to the ferromagnetically aligned transition state (Figure 4.8a, red). A similarly large activation enthalpy difference of 87 kJ/mol is observed for Mn-BBTA-Br (Table C.18). This phenomenon is also present for the Fe-containing frameworks, as has been noted previously by Neese and coworkers.[301,302,305] For Fe-BBTA-F, the antiferromagnetically aligned transition state (Figure 4.8b, blue) reduces the barrier for C−H activation by 51 kJ/mol with respect to the ferromagnetically coupled transition state (Figure 4.8b, red). A similar 39 kJ/mol enhancement is seen for Fe-BBTA-Br (Table C.19).

A population analysis at the transition state indicates that the magnitude of the spin density on the abstracting O atom is approximately equal in the ferromagnetic and antiferromagnetic high-spin states for Mn-BBTA-F (Table 4.1). Instead of the degree of radical-like character, this drastic difference in activation enthalpies is better explained within the context of exchange-enhanced reactivity.[240,241,306,307] Exchange-enhanced reactivity implies that an increase in the number of

identical-spin electrons on the transition metal center can stabilize the transition state via a greater number of exchange interactions, provided this exchange stabilization is greater than the energy required to place the electron in an otherwise vacant orbital. In the case of methane activation, this electron comes from the cleavage of the C−H bond and can either be added to an empty $3d$ orbital on the metal or paired up with an electron of the opposite spin in a singly occupied molecular orbital.[240] As shown in Table 4.1, for both Mn-BBTA-F and Fe-BBTA-F, the pathway that increases the spin density (i.e. increases the spin-identical electrons and thereby exchange interactions) on the metal center is the one with the greatly reduced C−H activation barrier. Most often, exchange-enhanced reactivity is used to rationalize an increase in the spin multiplicity at the transition state for C−H activation. Here, however, the spin multiplicity is constant because the electron is added to the metal in combination with a spin-flip on the abstracting O atom that forms the three-center C−H−O radical.

Table 4.1. Hirshfeld spin density, ρ, for the metal (M) and oxo ligand (O) in the $[MO]^{2+}$ initial state and at the transition state (TS) for the C−H activation of methane for various spin states of Mn-BBTA-F and Fe-BBTA-F. The spin density on the C atom of methane at the TS is also shown. The spin state containing dashed lines could not be isolated computationally. Bolded values correspond to the lowest energy structure. Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

MOF	S	$\rho_{M,[MO]^{2+}}$	$\rho_{O,[MO]^{2+}}$	$\rho_{M,TS}$	$\rho_{O,TS}$	$\rho_{C,TS}$
Mn-BBTA-F	3/2 (AFM)	3.53	−0.76	**3.65**	**−0.38**	**−0.41**
	3/2 (FM)	**2.44**	**0.46**	1.99	0.37	0.43
	1/2 (FM)	0.66	0.34	0.05	0.41	0.43
Fe-BBTA-F	3 (FM)	4.13	1.25	4.14	0.85	0.36
	2 (AFM)	---	---	**4.04**	**−0.09**	**−0.34**
	2 (FM)	**3.19**	**0.49**	2.88	0.36	0.42
	1 (FM)	1.27	0.74	0.96	0.50	0.43

Finally, returning to the concept of spin density on the oxo ligand, it is worth noting that if one focuses on a single spin channel in the examples shown in Figure 4.8, the lowest barrier corresponds to the active site with the most radical-like character on the abstracting O atom (i.e. S = 3/2 (AFM) for Mn-BBTA-F and S = 3 (FM) for Fe-BBTA-F, as shown in Table 4.1). This suggests that when attempting to rationalize the barrier heights of individual spin channels for a single catalytic candidate, spin density on the oxo ligand can be illustrative. However, the possibility of a spin transition during the C−H activation process is one factor that prevents a direct relationship between the radical-like character on the oxo ligand and the activation enthalpy of the process when calculated by comparing the enthalpies of the most stable initial state and transition state.

4.4 Conclusion

Significant attention in recent years has been focused on using MOFs with coordinatively unsaturated metal sites for the purposes of oxidation catalysis. In MOFs where a terminal metal-oxo species is the proposed active site, one of the known challenges is stabilizing the fleeting nature of the metal-oxo complex, if it can even be formed at all. In this work, we studied a highly tunable family of metal−triazolate frameworks with the general formula M_2X_2(bbta) (M = divalent metal, X = monovalent bridging anion, H_2bbta = 1*H*,5*H*-benzo(1,2-d:4,5-d′)bistriazole). The results of the quantum-chemical investigation highlight how anion-exchange of the bridging ligands can be used to tune the reactivity of MOFs for oxidation catalysis. In particular, MOFs with μ-OH^-, μ-SH^-, and μ-SeH^- ligands tend to be better at stabilizing the metal-oxo species than μ-F^-, μ-Cl^-, and μ-Br^- ligands. The basicity of the bridging ligand is also qualitatively correlated

with the thermodynamic favorability of forming the $[MO]^{2+}$ site, with trends that are a function of metal identity. With regards to the reactivity of the hypothesized metal-oxo sites, relatively low (i.e. < 75 kJ/mol) apparent C−H activation barriers are predicted for all the frameworks in this work except those containing V^{2+} cations. These results suggest that future efforts should primarily be focused on stabilizing the $[MO]^{2+}$ site rather than on improving their ability to activate strong C−H bonds.

For the Mn- and Fe-containing frameworks in particular, the high reactivity can be attributed to the formation of antiferromagnetically coupled spin density on the metal binding site and the abstracting O atom at the transition state for C−H bond activation, which is not necessarily present in the initial $[MO]^{2+}$ state. In cases where there is a ferromagnetic to antiferromagnetic transition *en route* to the transition state, this is an example of two-state reactivity, albeit one with a fixed spin multiplicity. The more reactive pathway is the one that increases the spin density on the metal center, consistent with the concept of exchange-enhanced reactivity. In addition, we note that the presence of significant spin density on the abstracting O atom at either the initial $[MO]^{2+}$ state or the C−H activation transition state is not always directly related to low C−H activation barriers. We expect that these trends are not restricted to the M_2X_2(bbta) family of materials and may be used to rationalize the reactivities of other inorganic, bioinorganic, and enzymatic catalysts wherein terminal metal-oxo species are believed to carry out the oxidation of strong C−H bonds.

Chapter 5: TUNING THE REDOX ACTIVITY OF METAL–ORGANIC FRAMEWORKS FOR ENHANCED, SELECTIVE O_2 BINDING

This chapter describes the use of computational screening methods based on periodic density functional theory to investigate O_2 and N_2 adsorption at the coordinatively unsaturated metal sites of several MOF families. A variety of design handles are identified that can be used to modify the redox activity of the metal centers, including changing the functionalization of the linkers, anion exchange of bridging ligands, and altering the formal oxidation state of the metal. As a result, we show that it is possible to tune the O_2 affinity at the open metal sites of MOFs for applications involving the strong and/or selective binding of O_2. In contrast with O_2 adsorption, N_2 adsorption at open metal sites is predicted to be relatively weak across the MOF dataset, with the exception of MOFs containing synthetically elusive V^{2+} open metal sites. As one example from the screening study, we predicted that exchanging the μ-Cl^- ligands of M_2Cl_2(bbta) (H_2bbta = 1*H*,5*H*-benzo(1,2-d:4,5-d′)bistriazole) with μ-OH^- groups would significantly enhance the strength of O_2 adsorption at the open metal sites without a corresponding increase in the N_2 affinity. Experimental investigation of Co_2Cl_2(bbta) and $Co_2(OH)_2$(bbta) confirms that the former exhibits weak physisorption of both N_2 and O_2, whereas the latter is capable of chemisorbing O_2 at room temperature in a highly selective manner. The O_2 chemisorption behavior is attributed to the greater electron-donating character of the μ-OH^- ligands and the presence of H-bonding interactions between the μ-OH^- bridging ligands and the reduced O_2 adsorbate.

All the density functional theory (DFT) calculations in this work were carried out by A.S. Rosen. M. Rasel Mian and T. Islamoglu synthesized, characterized, and tested the O_2 adsorption

behavior of Co_2X_2(bbta) (X = Cl, OH). H. Chen carried out the grand canonical Monte Carlo simulations and helped calculate the O_2/N_2 selectivity. All authors provided feedback on the manuscript.

This chapter is adapted from the following peer-reviewed article: A.S. Rosen, M. Rasel Mian, T. Islamoglu, H. Chen, O.K. Farha, J.M. Notestein, R.Q. Snurr. "Tuning the Redox Activity of Metal−Organic Frameworks for Enhanced, Selective O_2 Binding: Design Rules and Ambient Temperature O_2 Chemisorption in a Cobalt−Triazolate Framework." *Journal of the American Chemical Society*, **142**, 4317–4328 (2020).

5.1 Introduction

Since the discovery and characterization of hemoglobin in the 1800s, significant effort has been focused on designing synthetic transition metal complexes that can bind O_2 for a diverse range of applications, including oxidation catalysis, electrocatalytic oxygen reduction, and the separation of oxygen from air.[308–311] Over the last several decades, metal−dinitrogen complexes have also been actively investigated, with much attention focused on catalytic nitrogen activation, the separation of nitrogen from methane-rich natural gas, and the purification of N_2 from air.[310,312–314] For each of these purposes, the binding strength of O_2 and/or N_2 has a significant influence on the performance of the investigated material. This is especially true for the design of an adsorbent that can separate O_2 from N_2, as the material must be able to selectively bind one molecule over the other at ambient temperatures and to do so reversibly.[309] An adsorbent that could perform this separation at moderate conditions with high selectivity would be desirable for industrial and medical processes that require purified O_2,[315] as the conventional method to separate O_2 from air at a large scale involves cryogenic distillation, an energy-intensive and expensive process.[316] While cation-exchanged zeolites are currently used as adsorbents for air separation purposes,[315] they typically favor the adsorption of N_2 over O_2 and are inefficient compared to O_2-selective materials for the production of O_2 via air separation.

Metal−organic frameworks (MOFs) have been widely studied for applications involving the binding of small molecules due to their well-defined structures and high degree of tunability made possible by modifying the inorganic nodes and organic linkers that compose each material.[171,172,317–320] MOFs with coordinatively unsaturated metal sites, also known as open metal

sites, are of particular interest for the selective binding of O_2 or N_2 due to the possibility of charge transfer between the adsorbate and exposed metal cations.[321]

Despite active research in this area,[22,24,185,216,322–325] significant improvements are still needed to achieve a greater O_2/N_2 selectivity at near-ambient conditions before MOFs with open metal sites can be used for the purification of O_2 from air. Several MOFs with Cr(II) open metal sites, including $Cr_3(btc)_2$ (H_3btc = benzene-1,3,5-tricarboxylic acid) and Cr-BTT ($Cr_3[(Cr_4Cl)_3(BTT)_8]_2$, H_3BTT = benzene-1,3,5-tris(1*H*-tetrazole)),[22,185] show strong and selective binding of O_2 over N_2 but exhibit a loss in capacity with multiple cycles. Oftentimes, MOFs with open metal sites exhibit only weak physisorption of O_2 and N_2, especially at near-ambient temperatures.[324,326] On both ends of this spectrum, $Fe_2(dobdc)$ (H_4dobdc = 2,5-dihydroxybenzene-1,4-dicarboxylic acid) reversibly binds O_2 over N_2 at low temperatures, but temperatures above 226 K result in the formation of superoxo species that irreversibly oxidize the framework.[24] Among the more promising MOFs investigated for O_2/N_2 separation, Co-BTTri ($Co_3[(Co_4Cl)_3(BTTri)_8]_2$, H_3BTTri = 1,3,5-tri(1*H*-1,2,3-triazol-5-yl)benzene) and Co-BDTriP ($Co_3[(Co_4Cl)_3(BDTriP)_8]_2$, $H_3BDTriP$ = 5,5′-(5-(1*H*-pyrazol-4-yl)-1,3-phenylene)bis(1*H*-1,2,3-triazole)) can reversibly bind O_2 over N_2, although their selectivity towards O_2 is significantly reduced at temperatures above 195 K.[216]

Given the enormous library of inorganic nodes and organic linkers[44,89] and the significant time involved in synthesizing, characterizing, and testing new MOFs for O_2/N_2 separation, there is a need for design rules that can be used to predictively tune the O_2 and N_2 affinities at the open metal sites of MOFs. Computational screening methods based on first-principles calculations are appealing avenues to achieve this goal, as multiple distinct MOF families can be studied in a

systematic fashion and on a much larger number of materials than would be feasible experimentally. However, much of the prior computational work has been carried out on a small scale, often focusing on MOFs directly motivated by experiments[217,327,328] or on the prototypical (yet structurally unrelated) M_2(dobdc) and $M_3(btc)_2$ MOF families.[326–331] This has limited the identification of broadly generalizable principles that can be used to identify and predictively design new MOFs for applications involving the adsorption of O_2 or N_2 at redox-active open metal sites. Nonetheless, for these systems, theory-driven studies have proven quite useful for understanding the effect of metal identity,[330] preferred gas binding geometries,[328,330] electronic structure of the $[M{-}O_2]^{2+}$ species,[327] and dynamics of competitive gas adsorption.[329] Prior work has also shown that the O_2/N_2 selectivity can be correlated with the difference in computed binding energies,[330] although the ambient temperature O_2/N_2 selectivity is often significantly lower than that at cryogenic conditions.[216,324] The discovery of MOFs that selectively bind O_2 over N_2 at room temperature remains a significant challenge.

In the present study, we leverage a high-throughput periodic density functional theory (DFT) workflow[32] to investigate O_2 and N_2 adsorption at the open metal sites of several MOF families (Figure 5.1). We have specifically chosen structurally related pairs of MOFs to aid in identifying structure−property relationships that can be used to guide the design of MOFs with enhanced affinities for O_2 and/or N_2 at open metal sites, primarily (albeit not exclusively) for air separation applications. These MOF families include $M(II)_2$(dobdc),[201,202,332] $M(II)_2$(dsbdc) (H_4dsbdc = 2,5-disulfhydrylbenzene-1,4-dicarboxylic acid),[205,206] $M(II)_2Cl_2$(bbta) (H_2bbta = 1*H*,5*H*-benzo(1,2-d:4,5-d′)bistriazole),[193,194,271] $M(II)_2(OH)_2$(bbta),[273,275,277] $M(II)M(III)_2(\mu_3$-O)$(bdc)_3$ (H_2bdc = benzene-1,4-dicarboxylic acid),[199,200] and $M(III)_3(\mu_3$-O)$(bdc)_3$OH,[199,200] which

were selected due to their ability to support several coordinatively unsaturated, first row transition metal cations as well as their structural tunability.

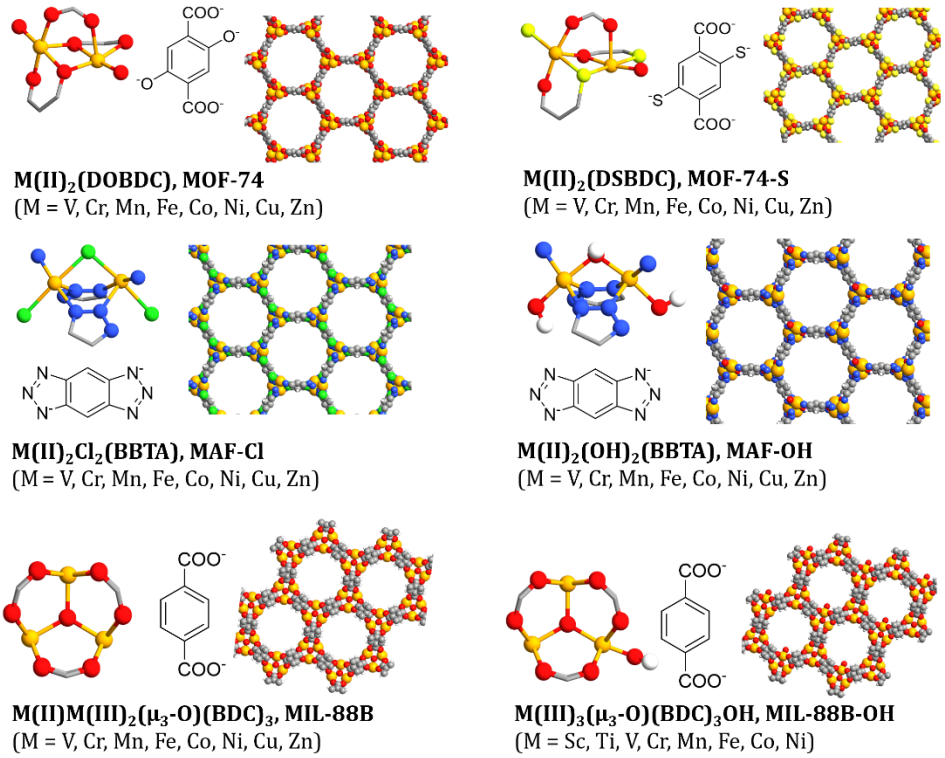


Figure 5.1. MOF families that are the primary focus of this work, with each row representing a pair of structurally related materials. The inorganic nodes, organic linkers, and crystal structures are highlighted for each MOF family. Color key: M (orange), O (red), S (yellow), N (blue), Cl (green), C (gray), H (white).

As a result of our computational screening workflow and the design rules established in this work, we predict that exchanging the μ-Cl^- ligands of Co_2Cl_2(bbta) with μ-OH^- ligands to form the isostructural framework $Co_2(OH)_2$(bbta) should significantly enhance the O_2 affinity at the open Co(II) sites without strengthening their interaction with N_2. This prediction is further validated experimentally, providing support that MOFs with open metal sites that are otherwise redox-inactive can be made to chemisorb O_2 by means of post-synthetic ligand-exchange,

particularly when using basic bridging ligands that also enable H-bonding interactions with the reduced O_2 adsorbate.

5.2 Methods

5.2.1 Constructing Initial Model Structures

To aid in the identification of structure−property relationships, the members of each MOF family were modeled as being in an isostructural series using the coordination geometries shown in Figure 5.1. For brevity, we will often refer to these MOFs as MOF-74 or MOF-74-S, MAF-X (MAF = Metal−Azolate Framework, X = bridging anion), and MIL-88B (MIL = Materials Institut Lavoisier) or MIL-88B-OH, as shown in Figure 5.1. We considered the presence of V^{2+}, Cr^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, and Zn^{2+} cations for MOFs with M^{2+} binding sites and Sc^{3+}, Ti^{3+}, V^{3+}, Cr^{3+}, Mn^{3+}, Fe^{3+}, Co^{3+}, and Ni^{3+} cations for MOFs with M^{3+} binding sites. The MOF-74, MOF-74-S, and MAF-X families support divalent metal cations, whereas the MIL-88B family can support formally trivalent (in the case of MIL-88B-OH) or mixed divalent/trivalent (in the case of MIL-88B) cations depending on the presence of terminal OH groups, as depicted in Figure D.1.

Some of the metal cations considered in this work have yet to be experimentally incorporated in the investigated MOF families; nonetheless, a wide range of $3d$ metals were considered for the purposes of identifying trends related to O_2 and N_2 adsorption at open metal sites in MOFs. As adopted in prior computational studies of MOFs,[281] only a single $3d$ metal cation was included per primitive simulation unit cell during the screening process to reduce the computational expense and prevent challenges associated with accurately identifying complex ground state spin orderings. The remaining metal species were exchanged with closed-shell cations, namely Mg^{2+} or Al^{3+} depending on the formal oxidation state of the metal in the as-

synthesized state (Figure D.1). Similar structures have been achieved experimentally for a variety of MOFs, such as $Mg_{1.9}Fe_{0.1}$(dobdc)[168] and Fe-doped Al-MIL-53.[98] Various mixed-metal motifs have also been incorporated in the trimetallic nodes of the PCN-250 (PCN = Porous Coordination Network) and MIL-88 families.[200,333]

Adsorbates were systematically added to the open metal sites in an automated fashion using the MOF Adsorbate Initializer (MAI) program.[32,117] O_2 conventionally adsorbs to mononuclear transition metal complexes in either a bent, end-on (η^1) mode or a side-on (η^2) mode, whereas stable M−N_2 complexes are generally only found with near-linear, end-on (η^1) adsorption modes.[310,334] The initialized structures generated using MAI were constructed accordingly. For O_2, an η^1 mode with an M−O−O bond angle of 120° and an η^2 mode with an M−O_{MP}−O bond angle of 90° were constructed, where M is the metal binding site and O_{MP} represents the midpoint between the two O atoms in molecular oxygen. For N_2, an η^1 mode with an M−N−N bond angle of 180° was constructed. In all cases, a single guest molecule was introduced to the Niggli-reduced unit cell, and the lowest energy conformations resulting from the DFT optimization of these initial structures were considered for further analysis. The CrystalNN bonding topology algorithm[335,336] implemented in Pymatgen[82] was used to categorize if the lowest energy adsorbate configuration exhibits an η^1 or η^2 adsorption mode (or approximately no binding at all, which we denote η^0) following structural optimization.

5.2.2 Quantum-Chemical Calculations

For the DFT screening procedure, adsorption energies were evaluated using the Vienna *ab initio* Simulation Package (VASP) v.5.4.1[122,123] with v.5.4 of the VASP-recommended[337] projector-augmented wave (PAW) pseudopotentials.[124] The unit cell shapes and volumes for the diluted

MOFs were optimized using the PBE-D3(BJ) level of theory,[125–127] as PBE with dispersion corrections has been shown to accurately capture the lattice constants of a variety of MOFs.[128,129] Following this, the lattice constants were kept fixed and the atomic positions refined using the M06-L meta-generalized gradient approximation (GGA) exchange-correlation functional.[287] All subsequent calculations involving the adsorption of any guest molecules were carried out at the M06-L level of theory with a fixed simulation unit cell. While hybrid-level functionals would be expected to reduce the tendency to overbind O_2,[326] their high computational cost precludes their use when screening a large number of MOF−adsorbate systems, each with several possible spin states and adsorption modes. O_2 and N_2 adsorption energies computed using GGA and GGA+U approaches will be reported in forthcoming work and show similar structure−property relationships to those reported here.[74]

As benchmarked and described in detail previously,[32] a multi-stage automated workflow built upon the Atomic Simulation Environment[83] was used to robustly and efficiently carry out the periodic DFT calculations in this work. The final high-accuracy settings used in the optimization workflow include a 520 eV plane-wave kinetic energy cutoff, k-point density of approximately 1000/number of atoms (as computed using Pymatgen's automatic k-point density generator[82]), and a force-convergence criterion of 0.03 eV/Å.[32] Spin-polarization was considered for each MOF, as summarized in Table D.1. The properties of the lowest energy spin state found from this procedure are reported in all analyses. The preconditioned conjugate gradient "all bands simultaneous update of orbitals" algorithm[132,133,338] was used to converge the self-consistent field (SCF), and Gaussian smearing of the band occupancies with a smearing width of 0.01 eV was employed (with electronic energies extrapolated back to the 0 K limit). The SCF was considered converged when the change

in electronic energy was less than 10^{-6} eV. The accurate-precision keyword was enabled in VASP, and all symmetry operations were disabled. Isolated gas-phase species were modeled in the center of a periodic box with 10 Å of vacuum space on each side of the molecule and k-point sampling at the Γ-point only.

O_2 and N_2 adsorption energies were computed as

$$\Delta E_{O_2} = E_{MOF-O_2} - E_{MOF} - E_{O_2} \quad (5.1)$$

$$\Delta E_{N_2} = E_{MOF-N_2} - E_{MOF} - E_{N_2} \quad (5.2)$$

where E_{MOF-O_2} and E_{MOF-N_2} are the electronic energies of the MOF with an O_2 or N_2 guest species at the open metal site, E_{MOF} is the electronic energy of the bare MOF, and E_{O_2} and E_{N_2} are the electronic energies of isolated O_2 and N_2, respectively. All computed physicochemical properties (e.g. optimized structures, spin states, binding modes, partial charges, spin densities), and additional methodological details can be found in the Appendix.

Periodic DFT calculations of the non-diluted $Co_2(OH)_2$(bbta) and Co_2Cl_2(bbta) frameworks were carried out at the PBE-D3(BJ) level of theory with a Hubbard U correction[339] of 3.3 eV on the *d* levels of the Co sites chosen because of its ability to reproduce oxidation energies in cobalt oxides.[340] This level of theory accurately reproduces both the lattice constants and O_2 adsorption energy in the structurally similar Co-BTTri framework (Table D.7). Additional computational details, such as the most stable spin states, can be found in the Appendix, with all computational results hosted at the Zenodo repository with DOI: 10.5281/zenodo.2652475.

5.2.3 Experimental Methods

Co_2Cl_2(bbta) and $Co_2(OH)_2$(bbta) were synthesized, characterized, and investigated for their O_2 and N_2 uptake. The MOFs were synthesized based on previously reported procedures,[194,277,280] with $Co_2(OH)_2$(bbta) prepared by soaking 50 mg of the parent Co_2Cl_2(bbta) framework overnight in 25 mL of 1 M KOH solution at room temperature. Powder X-ray diffraction (PXRD), N_2 physisorption surface area using the Brunauer–Emmett–Teller (BET) method, X-ray photoelectron spectroscopy (XPS), and diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) were carried out to characterize the materials and to confirm the full exchange of μ-Cl^- ligands with μ-OH^- in synthesizing $Co_2(OH)_2$(bbta). Additional details regarding the experimental methods can be found in the Appendix.

5.3 Results and Discussion

5.3.1 Overview of Screening Results and Periodic Trends

Using a high-throughput periodic DFT workflow,[32] we screened the MOFs shown in Figure 5.1 for the adsorption of O_2 and N_2. As summarized in Figure 5.2, a wide spread of O_2 and N_2 adsorption energies are observed across the MOF dataset. Nonetheless, the range in O_2 binding energies is approximately three times larger than that of the N_2 binding energies, suggesting that the chemisorption of O_2 is more common than that of N_2 at the MOF open metal sites. There are also broad periodic trends in the O_2 affinity at the open metal sites in MOFs, with early transition metals generally resulting in significantly more exothermic adsorption energies than later transition metals for a given MOF family and oxidation state (Figure 5.2a). This is consistent with the fact that d orbital energies tend to decrease (i.e. become more stable) with increasing d electron count, which reduces the degree of electron density transferred from the metal to the O_2 adsorbate.[310]

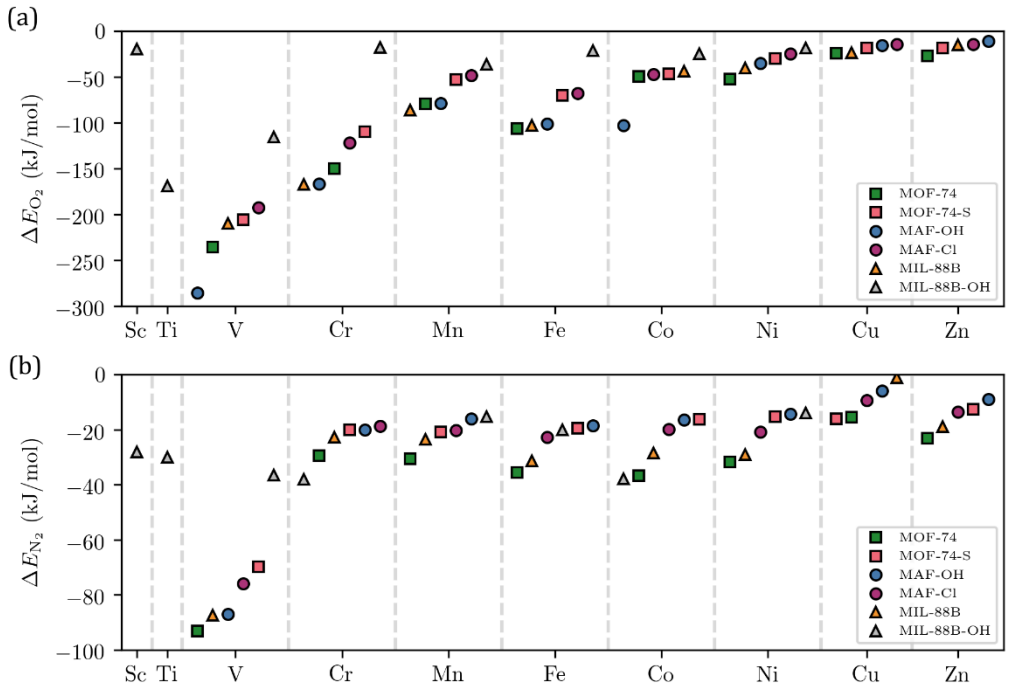


Figure 5.2. a) O_2 adsorption energy, ΔE_{O_2}, and b) N_2 adsorption energy, ΔE_{N_2}, at the M06-L/PAW level of theory for each MOF shown in Figure 5.1. For each metal, the MOFs are organized from strongest to weakest binding of the guest molecule.

Depending on the interaction between the metal center and O_2 adsorbate, the metal−dioxygen complex can take on one of several possible electronic states that exist on the spectrum of M(II)−O_2 ↔ M(III)−O_2^- ↔ M(IV)−O_2^{2-}. By comparing the sum of Bader partial atomic charges on the O_2 adsorbate with the sum of Bader partial charges on the O atoms of gaseous H_2O_2, HO_2, and O_2, it is possible to evaluate the degree of O_2 reduction for each MOF. As shown in Figure 5.3a, the exothermicity of O_2 adsorption is correlated with the degree of charge transfer to the O_2 adsorbate, implying that redox-active open metal sites are needed for strong O_2 chemisorption. In addition, the results in Figure 5.3a indicate that the reduction levels of O_2 are best thought of as continuous, in contrast with the quantized perspective often invoked in the literature.[310]

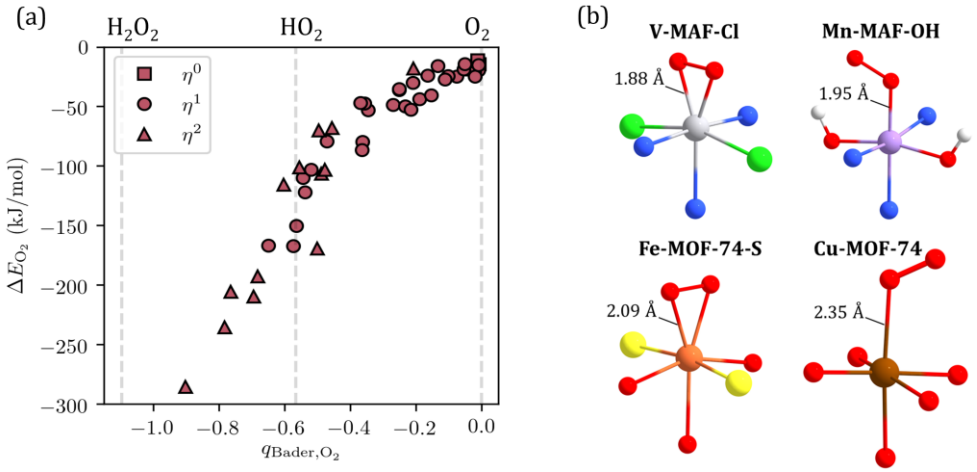


Figure 5.3. a) Relationship between the O_2 adsorption energy, ΔE_{O_2}, and sum of Bader partial atomic charges on the O_2 adsorbate, q_{Bader,O_2}, at the M06-L/PAW level of theory categorized by negligible (η^0, squares), end-on (η^1, circles) and side-on (η^2, triangles) binding modes. The dashed vertical lines indicate reference q_{Bader,O_2} values for free H_2O_2 (i.e. O_2^{2-}), HO_2 (i.e. O_2^-), and O_2 (neutral molecule). b) Example DFT-optimized geometries at the M06-L/PAW level of theory (only a representative portion of the binding site is shown). Color key: V (silver), Mn (purple), Fe (orange), Cu (brown), O (red), S (yellow), N (blue), Cl (green), H (white).

DFT-computed stretching frequencies, bond distances, spin densities, and partial atomic charges of the O_2 and N_2 guest molecules collectively indicate that the MOFs in this work support dioxygen-, superoxo-, and peroxo-like binding of O_2, whereas the triple bond of N_2 stays largely intact for most of the MOFs (Figure D.2). In terms of the most favorable O_2 binding modes for the MOFs in Figure 5.1, side-on adsorption was observed at all the Ti^{3+}, V^{2+}, V^{3+}, Cr^{3+}, and Fe^{2+} sites, whereas a bent, end-on mode occurred for the remaining transition metals (excluding those with negligible binding, such as the Sc^{3+}, Cu^{2+}, and Zn^{2+} sites), as shown in Table D.1. In most cases, the side-on binding modes are associated with a greater degree of superoxo character than the end-on binding mode (Figure 5.3a). Select MOFs are shown in Figure 5.3b to highlight the different O_2 adsorption behavior across the MOF dataset.

The binding of N_2 typically involves π-backbonding interactions from the metal d orbitals into the π^* orbitals of N_2, and the extent of π-backbonding is expected to be greater in low-valent, relatively electropositive transition metals.[310] As such, one would expect qualitatively similar periodic trends for both the O_2 and N_2 binding energies, with stronger binding for the divalent, early transition metals. While MOFs with V^{2+} sites tend to adsorb N_2 much more than strongly than all divalent cations with higher d electron counts, this periodic trend appears to be less pronounced than that for O_2 binding (Figure 5.2). The MOF with the largest N_2 binding energy at the M06-L/PAW level of theory is V-MOF-74, which was previously predicted to be a promising candidate for the separation of N_2 from CH_4 via strong π-backbonding interactions between the V^{2+} cations and N_2, although its synthesis has yet to be realized.[341] Based on the results in Figure 5.2, it appears that this behavior is not unique to V-MOF-74 and that several other MOFs with V^{2+} open metal sites, such as the yet-to-be-synthesized V-MAF-OH, would have similarly strong bonding interactions with N_2 (Figure 5.2b, Figure D.4).[e] However, the even stronger and highly exothermic adsorption of O_2 in these materials points to their likely instability in air.

For a given oxidation state, open metal sites composed of early transition metals are more likely to selectively bind O_2 over N_2 (Figure D.5), as has been noted previously.[330] Despite this, we caution against the design of MOFs with readily oxidized early transition metals for this application, as the highly exothermic O_2 adsorption energies are likely to result in irreversible chemisorption or, potentially, decomposition of the framework. This behavior has caused several previously investigated MOFs with early transition metals, such as M-MIL-101 (M = Ti, V),[20,217]

[e] Following the publication of this work, $V_2Cl_{2.8}$(btdd) was experimentally synthesized and shown to selectively adsorb N_2 at the V(II) sites.[279]

to be unsuitable for applications where reversible O_2 adsorption is necessary. Additionally, MOFs with exposed, low-valent early transition metals (particularly V^{2+} and Cr^{2+}) have not been widely synthesized to date.[22,185,207,342,343] At the other end of the first row of transition metals, MOFs with Ni, Cu, and Zn cations are unlikely to result in high O_2/N_2 selectivities, as they are not predicted to strongly adsorb O_2 at the metal centers. We instead recommend the identification and design of MOFs with redox-active Mn^{2+}, Fe^{2+}, or Co^{2+} sites that can be made to strongly favor the adsorption of O_2 over N_2 without forming metal−peroxo complexes that are often associated with irreversible O_2 binding in MOFs.[20,24,344]

5.3.2 Trends Between Related MOF Families

By examining the results for structurally related MOF families, it is possible to discern how subtle changes to the linkers and nodes of MOFs can be used to significantly modify the O_2 and N_2 binding strengths. Starting with the MOF-74 and MOF-74-S families, Figure 5.4 shows that the presence of thiophenoxide groups in MOF-74-S results in weaker adsorption of O_2 than the phenoxide groups of MOF-74, a trend that holds for all the metal cations considered in this work. The O_2 adsorption energy trend is the opposite of what might be predicted solely on the basis of electronegativity differences between O and S. Even though the metal centers of MOF-74 have higher partial atomic charges than their corresponding MOF-74-S analogues (Table D.2), the former still binds O_2 more strongly. Although likely not the sole factor, as noted in prior work for a series of ketoiminato/thioiminato Co(II) complexes,[345] these trends can be attributed in part to the greater donating character of the oxido (as opposed to sulfido) bridges. Figure 5.4 also shows that MOF-74 tends to bind N_2 more strongly than MOF-74-S. We note that the experimentally synthesized Mn-MOF-74-S structure is known to have seesaw-like open metal sites (Figure D.6),

in contrast with the M-MOF-74 family and Fe-MOF-74-S.[206] This seesaw structure for Mn-MOF-74-S is also predicted to have a lower affinity for O_2 and N_2 than its Mn-MOF-74 counterpart (Table D.3).

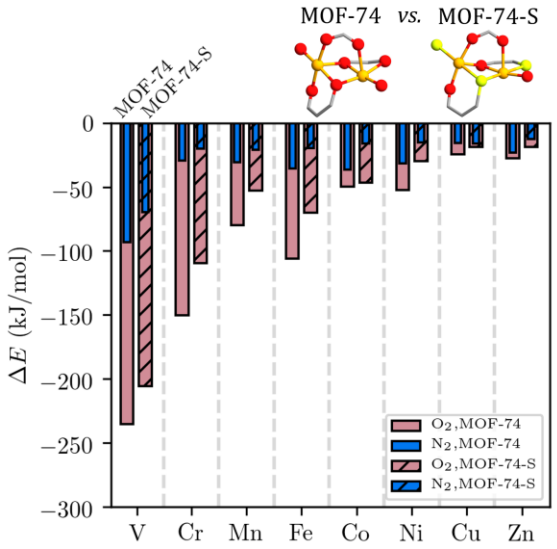


Figure 5.4. O_2 adsorption energy (red) and N_2 adsorption energy (blue) at the M06-L/PAW level of theory for MOF-74 (no hatch, left bars) and MOF-74-S (diagonal hatch, right bars) as a function of atomic number of the open metal site.

Significant differences in the O_2 binding strength are also found between the MAF-Cl and MAF-OH families, which have identical triazolate linkers but different bridging anions connecting the extended chain of metals. Based on the periodic DFT calculations, we predict that the triazolate MOFs with μ-OH^- species generally bind O_2 more strongly than their analogues with μ-Cl^- species (Figure 5.5). The only exception is for the Cu- and Zn-containing MOFs, where changing the bridging ligands does not greatly influence the O_2 binding affinity since neither Cu^{2+} nor Zn^{2+} are expected to be oxidized by O_2.

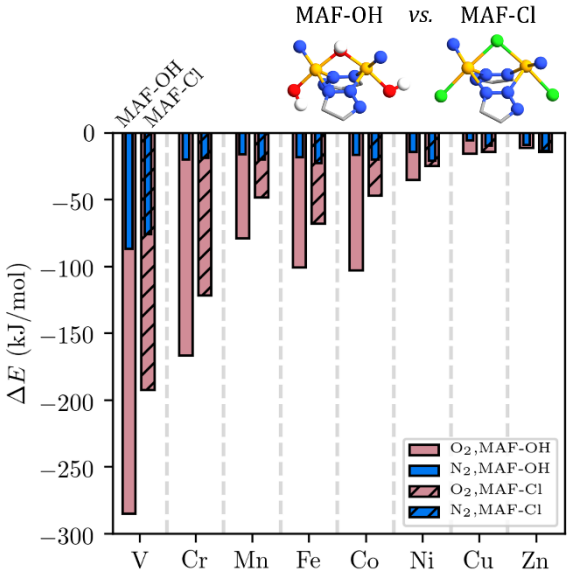


Figure 5.5. O_2 adsorption energy (red) and N_2 adsorption energy (blue) at the M06-L/PAW level of theory for MAF-OH (no hatch, left bars) and MAF-Cl (diagonal hatch, right bars) as a function of atomic number of the open metal site.

Based on the DFT-optimized structures, we predict that the MAF-OH family allows for H-bonding between the bridging μ-OH^- species and the O_2 species bound to the metal sites (Figure D.7), provided the O_2 molecule has been partially reduced. The presence of H-bonding is believed to stabilize metal−dioxygen species in enzymatic systems, such as hemoglobin,[346,347] and can be used to provide further control over the O_2 adsorption energy in MOFs where this functionality is possible. In addition, the μ-Cl^- ligands and μ-OH^- ligands have notably different basicities. Several correlations have been noted in prior work that relate greater basicities of the coordinating ligands to stronger O_2 affinities,[344,348] as is predicted here for MAF-OH. In these cases, the basicity of the ligand acts as a proxy for its donating ability, which serves to stabilize the oxidized state of the metal site. With the exception of proposed V^{2+} open metal sites as previously discussed, both the MAF-Cl and MAF-OH families are predicted to weakly physisorb N_2 (i.e. $|\Delta E_{N_2}| < 23$ kJ/mol). They also show weaker affinities for N_2 than the analogous members of the M-MOF-74 series. We

note that the substitutions invoked for both the MOF-74/MOF-74-S and MAF-OH/MAF-Cl families are with respect to the equatorial ligands. If the modified ligands were in the axial position, the adsorption process can also be influenced by structural *trans* effects, which may need to be considered.[349–352]

To further emphasize the degree of control over the O_2 adsorption energies that is possible via this anion exchange approach, the equatorial bridging anions in MAF-Cl were exchanged with a wider range of species. Varying both the metal and bridging anion, we considered M = V^{2+}–Ni^{2+} and bridging ligands of (in order of increasing pK_a of the parent acid)[353,354] μ-Br^-, μ-Cl^-, μ-F^-, μ-SH^-, and μ-OH^-. As shown in Figure 5.6, there is a qualitative trend in the O_2 adsorption energy as a function of the bridging ligand, with stronger O_2 binding found in MOFs that have more basic ligands. In addition, the ionic radius of the bridging ligand may play a complementary role in this family of MOFs, as smaller bridging anions allow for closer contact between the metals and their coordinating, equatorial donor ligands (Table D.4). Based on the results in Figure 5.6, it is clear that the reactivity of the metal site can be finely tuned via judicious selection of the metal and bridging ligand. Nonetheless, of the Mn-, Co-, and Ni-containing MOFs in Figure 5.6, only Co-MAF-F, Co-MAF-SH, and Co-MAF-OH are predicted to bind O_2 more strongly than their M-MOF-74 counterpart – an important comparison given that M-MOF-74 (M = Mn, Co, Ni) does not chemisorb O_2 at room temperature.[28,326] All members of the Co-MAF-X series investigated in this work are predicted to have weak interactions with N_2 (Table D.5).

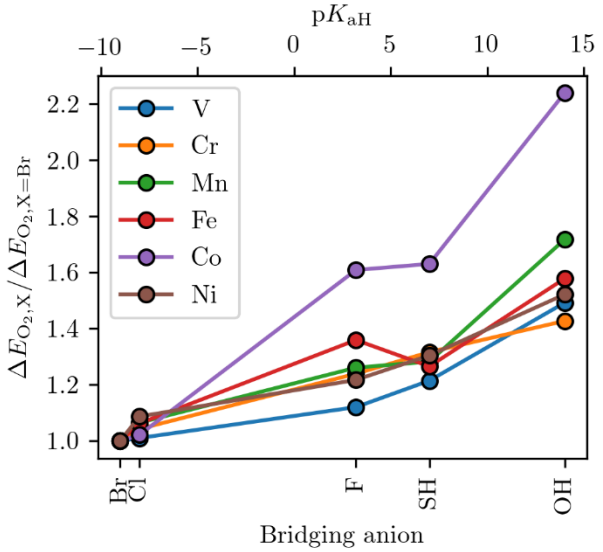


Figure 5.6. O_2 adsorption energy as a function of the bridging anion, X, for several M-MAF-X structures at the M06-L/PAW level of theory. For a given metal, the adsorption energies are relative to X = Br. The bridging anions are listed as a function of their basicity, as measured by the pK_a of the conjugate acid in water, pK_{aH}.

In addition to altering the atomic number of the metal or the basicity of the ligands, one can also consider modifying the oxidation state of the binding site. This is exemplified with the MIL-100 and MIL-101 families, where the presence or lack of charge-balancing OH^- anions at the trimetallic nodes can be used to tune the formal oxidation state of the metal (Figure D.1, Table D.6).[355,356] Since the MIL-100 and MIL-101 series have over one thousand atoms per unit cell, we chose to study the MIL-88B family, which has the same trimetallic node and terephthalic acid linkers as MIL-101 but is more computationally tractable due to its smaller unit cell size. For this family of MOFs, the formally trivalent transition metals exhibit significantly weaker adsorption of O_2 than the lower valent analogues (Figure 5.7), as higher oxidation states reduce the ability of the metal site to transfer electron density to the O_2 adsorbate. A similar trend is observed for N_2 adsorption, although both Cr^{3+} and Co^{3+} sites of MIL-88B-OH are predicted to adsorb N_2 at least as strong (if not stronger) than their MIL-88B counterparts. In agreement with prior experiments

that have shown trimetallic nodes of Cr-MIL-100 and Cr-MIL-101 selectively adsorb N_2 over O_2,[357] the closely related Cr-MIL-88B-OH with Cr^{3+} open metal sites is predicted to adsorb N_2 more strongly than O_2 by approximately 20 kJ/mol. Given these results, we would not suggest MOFs with trivalent cations for the selective binding of O_2 over N_2, although they may be of interest for selective N_2 binding.

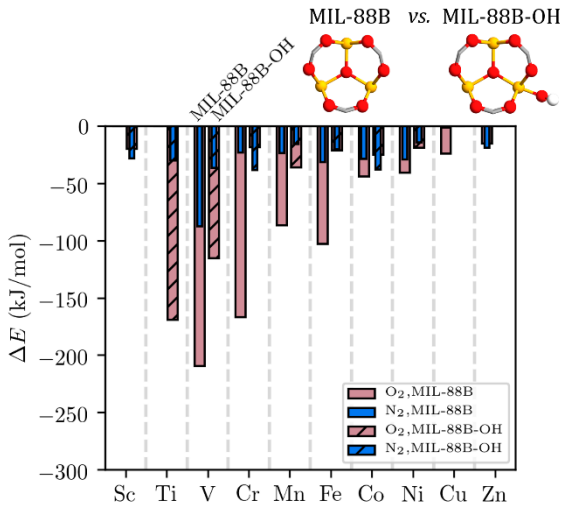


Figure 5.7. O_2 adsorption energy (red) and N_2 adsorption energy (blue) at the M06-L/PAW level of theory for MIL-88B (no hatch, left bars) and MIL-88B-OH (diagonal hatch, right bars) as a function of atomic number of the open metal site.

5.3.3 Enhanced, Selective Binding of O_2 in $Co_2(OH)_2(bbta)$

Of the MOFs screened in this work, we chose to investigate Co-MAF-OH , as its O_2 adsorption energy far exceeds that of Co-MAF-Cl and all other cobalt-containing MOFs considered in this study (Figure 5.2a and Figure 5.6). When combined with the fact that N_2 is predicted to weakly physisorb at the open metal sites (Figure 5.2b), Co-MAF-OH is predicted to have the largest difference in O_2 and N_2 binding energies of all the Mn-, Fe-, Co-, Ni-, Cu-, and Zn-containing MOFs studied in this work (Figure D.5). Perhaps most importantly, Co-MAF-OH serves as a clear example to highlight the effect of bridging ligand identity on the thermodynamic favorability of O_2 binding; of all the tested transition metals, it is predicted to have the greatest difference in O_2

binding energy when the μ-Cl^- ligands of the parent M-MAF-Cl framework are exchanged with μ-OH^- ligands (Figure 5.6). The corresponding non-diluted structure with all cobalt cations at the nodes is known to be stable when exposed to air as well as water, and it is thermally stable up to ~400 °C in a N_2 environment.[275,277] For clarity, we will refer to the non-diluted variants of Co-MAF-OH and Co-MAF-Cl as $Co_2(OH)_2$(bbta) and Co_2Cl_2(bbta), respectively.

We predict that the most energetically favored adsorption mode for O_2 bound at the coordinatively unsaturated cobalt sites of $Co_2(OH)_2$(bbta) is a bent, end-on mode with a Co−O−O angle of 116° and a relatively short Co−O_2 bond distance of 1.88 Å. The distance between the distal O atom of the O_2 adsorbate and the H atom of the nearby μ-OH^- group is 2.02 Å, such that there are stabilizing H-bonding interactions. Based on an analysis of the Charge Model 5 (CM5) partial atomic charges,[289] Bader spin densities,[159] O−O distance, and O−O stretching frequency, the thermodynamically favored $[CoO_2]^{2+}$ complex in $Co_2(OH)_2$(bbta) can be best characterized as a low-spin (S = 1/2) cobalt−superoxo species (Tables D.10 and D.15). The $[CoO_2]^{2+}$ site of Co_2Cl_2(bbta) is, instead, best described as an antiferromagnetically aligned, Co(II)−dioxygen complex (Table D.11). The charge density difference upon O_2 adsorption indicates that there is significant electron transfer from both the metal center and the adjacent ligands of $Co_2(OH)_2$(bbta) (Figure 5.8a), in contrast with Co_2Cl_2(bbta) where little reduction of the O_2 molecule occurs (Figure 5.8b). The H-bonding interaction between the μ-OH^- group of $Co_2(OH)_2$(bbta) and the distal atom of the O_2 adsorbate is also readily observed (Figure 5.8a).

At the PBE-D3(BJ)+U/PAW level of theory – chosen to enable direct comparison with the experimentally determined O_2 binding enthalpy in Co-BTTri (Table D.7) – the enthalpy of O_2 adsorption in $Co_2(OH)_2$(bbta) is predicted to be ΔH_{O_2} = −45 kJ/mol in the low-coverage limit

compared to ΔH_{O_2} = −19 kJ/mol for Co_2Cl_2(bbta) at the same level of theory. In contrast, N_2 adsorption at the open Co(II) sites of $Co_2(OH)_2$(bbta) is predicted to be very weak, with ΔH_{N_2} = −14 kJ/mol. The predicted O_2 adsorption enthalpy with $Co_2(OH)_2$(bbta) is stronger than that of the triazolate-bound Co(II) sites of Co-BTTri (Q_{st} = −34 kJ/mol) and comparable to that of the pyrazolate-bound Co(II) sites in the mixed triazolate/pyrazolate framework Co-BDTriP (Q_{st} = −47 kJ/mol at low-loadings), with similarly weak N_2 adsorption.[216]

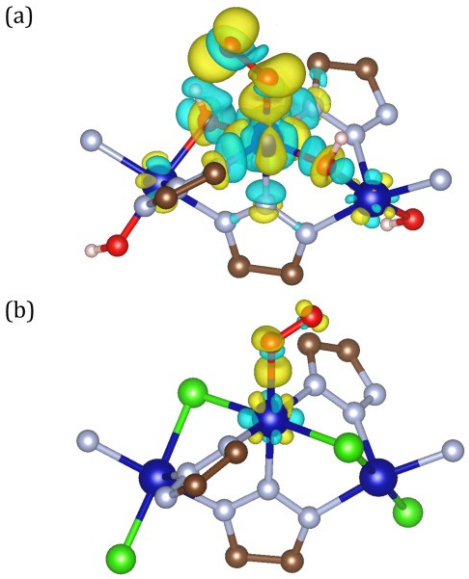


Figure 5.8. Charge density difference upon O_2 adsorption in a) $Co_2(OH)_2$(bbta) and b) Co_2Cl_2(bbta) at the PBE-D3(BJ)+*U*/PAW level of theory. The yellow and cyan surfaces represent a gain or loss of electron density, respectively (isovalue = 0.005 e^-/bohr^3). Only a representative portion of the periodic structure is shown. Color key: Co (dark blue), O (red), N (light blue), C (brown), H (white).

Assuming the presence of high-spin Co(II) species, as predicted at the PBE-D3(BJ)+*U*/PAW level of theory (Table D.9), the Co−ligand bond distances decrease by upwards of ~0.1 Å upon formation of a low-spin $[CoO_2]^{2+}$ site (Table D.16). This atypical strengthening of the bonding interactions between the coordinating ligands and the cobalt center upon O_2 adsorption is due to the smaller ionic radius of low-spin Co(III) compared to high-spin Co(II)[358] and further stabilizes

the $[CoO_2]^{2+}$ complex. As might be expected based on spin-crossover behavior in several previously synthesized metal−azolate frameworks,[193,216,267,270] the spin states of the Co(II) sites in $Co_2(OH)_2$(bbta) are relatively close in energy (Table D.9). The adsorption energies and charge density difference plot for hypothetical low-spin Co(II) sites are reported in the Appendix (Figure D.9, Tables D.19−D.23) and indicate that O_2 chemisorption should occur regardless of the spin state of the binding sites.

Based on our predictions, we chose to experimentally investigate Co_2Cl_2(bbta) and $Co_2(OH)_2$(bbta) to support the hypothesis that exchanging the μ-Cl^- ligands of Co_2Cl_2(bbta) with μ-OH^- ligands would increase the O_2 affinity of the Co(II) open metal sites. The PXRD patterns are consistent with those that would be expected for Co_2Cl_2(bbta) and $Co_2(OH)_2$(bbta) based on prior reports of their crystal structures (Figure D.11).[277] The IR spectra, via the presence of an O−H stretch at 3650 cm^{-1} for $Co_2(OH)_2$(bbta) but not for Co_2Cl_2(bbta), and the XPS spectra, via the presence of a broad 197.8 eV peak in the Cl 2p core level spectrum of Co_2Cl_2(bbta) but not $Co_2(OH)_2$(bbta), both indicate that the chloride anions have been exchanged with hydroxide anions in the $Co_2(OH)_2$(bbta) material (Figures D.13−D.16).

The O_2 and N_2 isotherms at 298 K for Co_2Cl_2(bbta) and $Co_2(OH)_2$(bbta) are shown in Figure 5.9. There is a sharp rise in the O_2 adsorption isotherm at low pressures for $Co_2(OH)_2$(bbta), indicative of chemisorption. This is accompanied by a large hysteresis in the O_2 adsorption and desorption branches, which suggests O_2 is chemically bound to the open metal sites. In stark contrast, the 298 K O_2 isotherm for Co_2Cl_2(bbta) shows a much shallower rise in the O_2 uptake at low pressures without any signs of chemisorption. Collectively, these results are consistent with an increased O_2 affinity enabled by post-synthetically exchanging the μ-Cl^- ligands with μ-OH^-

ligands, as predicted by DFT. Grand canonical Monte Carlo (GCMC) simulations further corroborate the finding that Co_2Cl_2(bbta) solely exhibits O_2 physisorption, whereas O_2 chemisorption is present for $Co_2(OH)_2$(bbta) (Figure D.21).

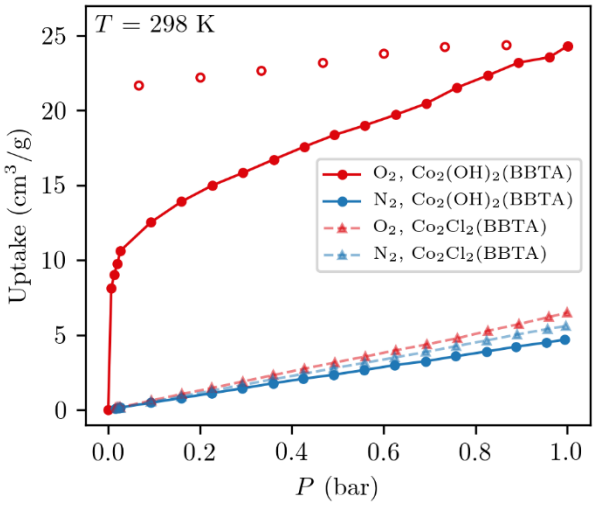


Figure 5.9. Experimentally measured O_2 (red) and N_2 (blue) isotherms for $Co_2(OH)_2$(bbta) (circles) and Co_2Cl_2(bbta) (triangles) at 298 K. The O_2 adsorption and desorption branches for $Co_2(OH)_2$(bbta) are shown as closed and open symbols, respectively.

As expected on the basis of the DFT calculations, the N_2 isotherms at 298 K for Co_2Cl_2(bbta) and $Co_2(OH)_2$(bbta) are both consistent with weak physisorption and are similar for both materials (Figure 5.9). As a result of the large difference in O_2 and N_2 binding in $Co_2(OH)_2$(bbta), ideal adsorbed solution theory[359,360] predicts an O_2/N_2 selectivity of 49 at 298 K, 0.21 bar O_2, and 0.79 bar N_2 (Figure D.17), which is particularly high compared to that of most other MOFs reported in the literature for air separation applications.[24,185,216,323,324,361–363] For reference, Co-BTTri has a selectivity of 41 at 195 K, but this greatly decreases to only 13 at 243 K; Co-BDTriP has a selectivity of 40 at 243 K, with a significantly lower selectivity expected at 298 K.[216]

The binding energies resulting from the DFT screening procedure describe the thermodynamic favorability of O_2 and N_2 adsorption, making it possible to gain fundamental insights and identify promising MOF candidates to investigate in greater detail. While the binding strengths of O_2 and N_2 are perhaps the most important properties to compute, they are not the only factors to consider once MOF promising candidates are identified. In the case of $Co_2(OH)_2$(bbta), a significantly longer time is needed for the O_2 uptake to reach equilibrium than for Co_2Cl_2(bbta) (Figure D.18). $Co_2(OH)_2$(bbta) is predicted to undergo an S = 5/2 to S = 1/2 transition (i.e. ^{4}Co(II) + $^3O_2 \rightarrow {}^2[CoO_2]^{2+}$), and large differences in the net spin multiplicity often result in slow kinetics.[364] In this scenario, the rate of room temperature O_2 adsorption could likely be increased by stabilizing low-spin Co(II) sites ($\Delta S = -1$) or by considering metals that remain high-spin upon O_2 adsorption, as is predicted for $Fe_2(OH)_2$(bbta) ($\Delta S = 0$) and $Mn_2(OH)_2$(bbta) ($\Delta S = -2$) (Table D.1), both of which have not yet been synthesized.

Overall, the ability of $Co_2(OH)_2$(bbta) to strongly bind O_2 suggests that this and related MOFs may be of interest for oxidation reactions and other processes involving the presence of redox-active open metal sites, especially when compared to the Cl-containing analogue. Based on the pronounced difference in O_2 affinities for Co_2Cl_2(bbta) and $Co_2(OH)_2$(bbta), post-synthetic ligand-exchange is likely to serve as a promising and accessible route to control the O_2 binding affinity in a wide range of MOFs.

5.4 Conclusion

Through the use of periodic density functional theory, we modeled the adsorption of O_2 and N_2 at the open metal sites of several pairs of structurally related MOF families. A comparison of the $M(II)_2$(dobdc) and $M(II)_2$(dsbdc) frameworks suggests that oxido bridges are preferable to sulfido

bridges with regards to the adsorption of both O_2 and N_2. Investigation of a topologically similar framework with triazolate (rather than carboxylate) linkers, $M(II)_2X_2(bbta)$ (X = F, Cl, Br, OH, SH), indicates that the choice of equatorial, bridging anions in this material can be used to greatly alter the O_2 binding affinity without significantly changing the N_2 binding affinity. We show that there is a qualitative correlation between the pK_a of the bridging anion and the strength of O_2 adsorption at the open metal sites, with more basic bridging ligands resulting in more redox-active metal centers. In general, we find that the strength of O_2 binding is a strong function of the degree of charge transfer. Consistent with this, early transition metals in low-valence states strongly bind O_2 due to the presence of readily oxidized metal centers. With regards to N_2 binding, most of the MOFs in this work are not capable of significantly reducing the N_2 molecule, with the primary exception being MOFs that contain synthetically elusive V^{2+} sites, regardless of the coordination environment of the binding site.

As a result of the quantum-chemical screening process and the trends described in this work, we predicted that exchanging the μ-Cl^- ligands of $Co_2Cl_2(bbta)$ with more basic μ-OH^- groups to form $Co_2(OH)_2(bbta)$ would significantly increase the O_2 affinity of the Co(II) open metal sites without strengthening their interaction with N_2. The strong affinity for O_2 at the Co(II) centers is associated with a significant degree of charge transfer that enables the formation of thermodynamically favored cobalt−superoxo species with the possibility of H-bonding interactions between the O_2 adsorbate and μ-OH^- groups. Experimentally determined adsorption isotherms at room temperature support the DFT predictions that $Co_2(OH)_2(bbta)$ chemisorbs O_2, whereas O_2 weakly physisorbs at the open metal sites of $Co_2Cl_2(bbta)$. Weak interactions with N_2 occur in both materials. More generally, we have demonstrated that anion exchange can serve as

a simple route to tune the redox activity of coordinatively unsaturated metal sites incorporated within the inorganic nodes of MOFs, enabling the strong and selective room temperature binding of O_2 in an otherwise unreactive framework.

Chapter 6: Comparing Density Functional Approximations for Redox-Dependent Binding at Open Metal Sites in Metal–Organic Frameworks

In this chapter, we compare commonly used generalized gradient approximation (GGA), GGA+*U*, and meta-GGA exchange-correlation functionals in modeling redox-dependent binding at open metal sites in MOFs, using O_2 and N_2 as representative small molecules. We find that the self-interaction error inherent to the widely used PBE GGA predicts metal sites that are artificially redox-active, as evidenced by their strong binding affinities, short metal−adsorbate bond distances, and large degree of charge transfer. The incorporation of metal-specific, empirical Hubbard *U* corrections based on the transition metal oxide literature systematically reduces the redox activity of the open metal sites, often improving agreement with experiment. Additionally, the binding behavior shifts from strong chemisorption to weaker physisorption as a function of *U*. The M06-L meta-GGA typically predicts binding energies between those of PBE-D3(BJ) and PBE-D3(BJ)+*U* when using empirically derived *U* values from the transition metal oxide literature. Despite the strong sensitivity of the binding affinities toward a given DFA, the GGA, GGA+*U*, and meta-GGA approaches often yield similar qualitative trends and structure−property relationships.

This chapter is adapted from the following peer-reviewed article: A.S. Rosen, J.M. Notestein, R.Q. Snurr. “Comparing GGA, GGA+*U*, and Meta-GGA Functionals for Redox-Dependent Binding at Open Metal Sites in Metal–Organic Frameworks.” *Journal of Chemical Physics,* **152**, 24101 (2020).

6.1 Introduction

Metal−organic frameworks (MOFs) are a class of highly porous, crystalline materials that have been widely considered for applications in heterogeneous catalysis, gas storage, separations, and several other areas of research.[171] MOFs with coordinatively unsaturated metal sites, particularly those that are redox-active,[17] are of significant interest due to their ability to strongly and/or selectively bind small molecules via a judicious choice of inorganic nodes and organic linkers.[365] With the thousands of MOFs that have been synthesized to date[44] and the many more that can potentially be synthesized,[53,90] quantum-chemical screening approaches are essential for the efficient exploration of the vast chemical space of possible MOF structures for redox-dependent adsorption applications.

In plane-wave, periodic density functional theory (DFT) studies, the most commonly used exchange-correlation functional is that of Perdew, Burke, and Ernzerhof (PBE),[125] due in part to its relatively modest computational cost.[366] This is especially the case when modeling the large unit cells of most MOFs, where hybrid functionals are often computationally prohibitive. However, generalized gradient approximation (GGA) functionals such as PBE are known to exhibit many-electron self-interaction error (SIE), in which each electron interacts with the total electron density, including its own.[367] The many-electron SIE in density functional approximations (DFAs) tends to overly delocalize electrons where they should otherwise be well-localized.[368] As has been emphasized for transition metal oxides, this delocalization error can greatly influence reaction energies when a change in redox state of the metal occurs.[340] By analogy, one might expect that redox-dependent adsorption of small molecules at the open metal sites of MOFs would be

highly influenced by the SIE as well. The investigation of this phenomenon is the focus of the present study.

Despite the wide variety of MOFs with open metal sites that have been synthesized to date, prior studies comparing GGA and GGA+*U* methods for small molecule binding at open metal sites in MOFs have almost entirely focused on the MOF family M_2(dobdc) (H_4dobdc = 2,5-dihydroxybenzene-1,4-dicarboxylic acid) and adsorbates, such as CO_2, that bind mainly via electrostatic and dispersive interactions rather than charge transfer with the metal site.[369–371] To address this gap in the literature and test the generality of structure−property trends, we have studied several MOF families with open metal sites using first-row transition metals from V to Cu. These MOFs include M_2(dobdc),[201,202,372] M_2(dsbdc) (H_4dsbdc = 2,5-disulfhydrylbenzene-1,4-dicarboxylic acid),[205,206] M_2Cl_2(bbta) (H_2bbta = 1*H*,5*H*-benzo(1,2-d:4,5-d′)bistriazole),[193,194,271] $M_2(OH)_2$(bbta),[275,277] and M-MIL-88B (MIL = Materials Institut Lavoisier),[199,200] as shown in Figure 6.1. The first four MOF families have formally divalent open metal sites, whereas M-MIL-88B can have divalent or trivalent open metal sites depending on the presence of charge-balancing anions, which we will distinguish from one another as MIL-88B (with the formula M(II)$M(III)_2(\mu_3\text{-O})(bdc)_3$ (H_2BDC = benzene-1,4-dicarboxylic acid)) and MIL-88B-OH (with the formula $M(III)_3(\mu_3\text{-O})(bdc)_3(OH)$, respectively.

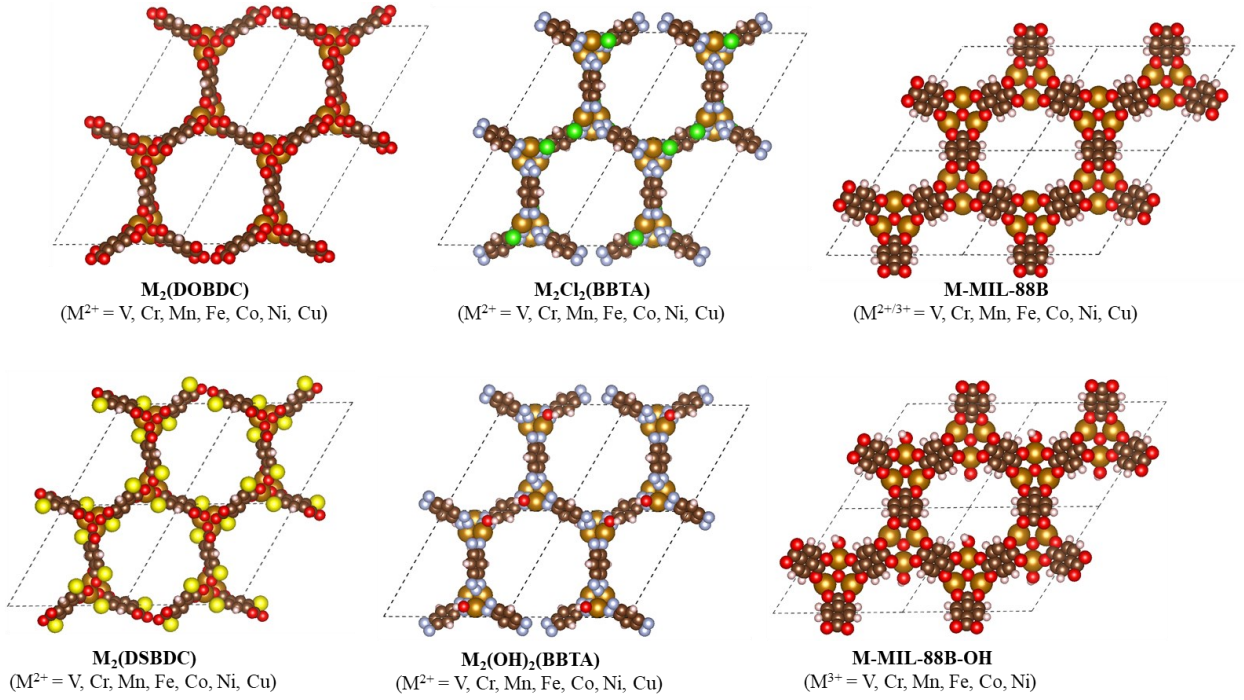


Figure 6.1. Summary of the 41 MOFs that are the primary focus of this work. The simulation unit cells are shown as dashed black lines. Color key: M (gold), O (red), N (blue), S (yellow), Cl (green), C (brown), H (white).

To investigate the effect of the DFA on redox-dependent adsorption in the MOFs shown in Figure 6.1, we focus on the adsorption of O_2 and N_2, as these molecules can access a variety of possible redox states depending on the metal−guest bond strength[310] and are relevant for applications in both catalysis (e.g. oxidation reactions) and gas separations (e.g. separation of O_2 from air). To evaluate the influence of the many-electron SIE, we directly compare the PBE GGA functional with and without a DFT+U treatment, as the incorporation of Hubbard U corrections can address the problem of electron over-delocalization.[373] The physicochemical properties computed at the GGA and GGA+U levels of theory are then compared with those from the meta-GGA M06-L functional,[287,288] which has been widely adopted by computational chemists studying $3d$ transition metal complexes.[281] In carrying out this study, we demonstrate how the SIE present with the commonly used PBE GGA functional can greatly influence the redox-dependent

adsorption of small molecules and how PBE+*U* (and to a lesser extent, M06-L) addresses these limitations. Notably, despite the large differences between the three DFAs, all three tend to yield similar structure−property relationships. The implications for future computational screening studies are discussed.

6.2 Methods

DFT calculations were carried out using the Vienna *ab initio* Simulation Package (VASP) v.5.4.1[122,123] with the entire crystallographic unit cell considered in all calculations. As previously mentioned, calculations were performed at three different levels of theory representing GGA, GGA+*U*, and meta-GGA DFAs. The PBE exchange-correlation functional[125] was used in the GGA (denoted PBE-D) and GGA+*U* (denoted PBE-D+*U*) treatments, with the inclusion of dispersion corrections via Grimme's D3 scheme[126] and Becke−Johnson (BJ) damping.[127] Since M06-L implicitly accounts for medium-range dispersion forces, the M06-L calculations did not include additive dispersion corrections.

For the PBE-D+*U* calculations, the approach by Dudarev and coworkers[339] was employed such that an effective *U* value is the only required input parameter. Unless otherwise stated, the Hubbard *U* values were taken as V: 3.1 eV, Cr: 3.5 eV, Mn: 3.8 eV, Fe: 4.0 eV, Co: 3.3 eV, Ni: 6.4 eV, and Cu: 4.0 eV applied to the metal 3*d* sites, motivated by prior work of Ceder and coworkers who fit them to reproduce experimental oxidation energies of transition metal oxides.[340] While Hubbard *U* values can be determined *ab initio* via linear response theory,[369] this procedure is less practical when studying a diverse set of materials and does not necessarily result in the empirically best *U* values to reproduce experimentally determined adsorption energies.

We have adopted the same computational workflow as in our prior work,[71,73] which we summarize as follows. A multi-stage automated periodic DFT screening workflow[32] based on the Atomic Simulation Environment (ASE) 3.18.0[83] was used to efficiently carry out the structural optimizations and energy evaluations in this work. The final high-accuracy settings used in the optimization procedure include a 520 eV plane-wave kinetic energy cutoff, k-point density of approximately 1000/number of atoms (as computed and arranged using the automatic k-point density generator in Pymatgen 2019.9.16[82]), and a force-convergence criterion of 0.03 eV/Å. The VASP-recommended v.5.4 projector-augmented wave (PAW) pseudopotentials[124] were used for all calculations. Spin-polarization was considered for each MOF using both a high- and low-spin initialization (refer to Table E.1 for more details). The preconditioned conjugate gradient "all band simultaneous update of orbitals" algorithm[132,133,338] was used to converge the self-consistent field (SCF), and Gaussian smearing of the band occupancies with a smearing width of 0.01 eV was employed (with electronic energies extrapolated back to the 0 K limit). The SCF was considered converged when the change in electronic energy was less than 10^{-6} eV. The accurate-precision keyword was enabled in VASP, and all symmetry operations were disabled. Aspherical contributions to the gradient corrections inside the PAW spheres were included for meta-GGA and GGA+U calculations.

The primitive unit cell was taken for each MOF and was diluted with closed-shell Mg^{2+} or Al^{3+} species such that only a single $3d$ transition metal cation remained. This was done to reduce the computational cost and make it easier to identify the ground state spin state, as done in prior computational studies of MOFs.[281] The DFT-computed properties at the M06-L level of theory were obtained from our prior work.[73] Since PBE with dispersion corrections has been shown to

accurately capture the lattice parameters of MOFs,[128,129] the cell shapes and volumes of the diluted MOFs were relaxed with the PBE-D functional. Subsequent calculations were carried out at the specified level of theory with a fixed simulation unit cell but allowing the atomic positions to change. Adsorbates were systematically added to the $3d$ transition metal sites using the MOF Adsorbate Initializer (MAI) program v.1.1,[32,117] considering η^1–O_2, η^2–O_2, and η^1–N_2 adsorption modes (Figure 6.2), as described previously.[73] The lowest energy structures at the PBE-D level of theory were adopted as initial guesses for the PBE-D+U calculations. All DFT-computed properties presented in this work are for the lowest energy spin state (Table E.1) and adsorption mode (Table E.2) unless otherwise noted.

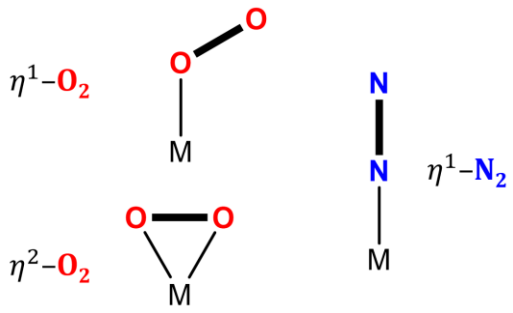


Figure 6.2. Depiction of the η^1–O_2, η^2–O_2, and η^1–N_2 binding modes considered in this work. Bold lines are drawn for the O–O and N–N bonds, as the bond order will vary depending on the metal binding site.

Adsorption energies were calculated as

$$\Delta E_{O_2} = E_{MOF-O_2} - E_{MOF} - E_{O_2} \quad (1)$$

and

$$\Delta E_{N_2} = E_{MOF-N_2} - E_{MOF} - E_{N_2} \quad (2)$$

where $E_{\mathrm{MOF-O_2}}$ and $E_{\mathrm{MOF-N_2}}$ are the electronic energies of the MOF with a single adsorbed O_2 or N_2 guest molecule per unit cell, E_{MOF} is the electronic energy of the bare MOF, and $E_{\mathrm{O_2}}$ and $E_{\mathrm{N_2}}$ are the electronic energies of gas-phase O_2 and N_2. Partial atomic charges and spin densities reported in the main text were computed via a Bader analysis[159] using v.1.03 of the Bader code by the Henkelman group.[138] DDEC6[158,230] partial atomic charges and spin densities (reported in the supplementary dataset) were computed using Chargemol v.09-26-2017.[374] All data used to produce the figures in this work, including DFT-optimized structures, energies, and several other properties can be found at the Zenodo repository with DOI: 10.5281/zenodo.3817991.[375]

6.3 Results and Discussion

6.3.1 Charge and Spin Density of the Bare MOFs

We start by analyzing the MOFs at the PBE-D, M06-L, and PBE-D+*U* levels of theory prior to the adsorption of any guest molecules. As shown in Table 6.1, the partial atomic charges of the coordinatively unsaturated transition metals vary slightly depending on the density functional. In most cases, the computed partial atomic charges are the lowest at the PBE-D level of theory. When comparing the PBE-D and PBE-D+*U* simulations directly, all but eight of the 41 MOFs studied in this work exhibited an increase in charge on the transition metal when Hubbard *U* corrections were applied. The exceptions are the five V^{2+}-containing MOFs, Cr-MIL-88B, Fe-MIL-88B, and $Fe_2(OH)_2$(bbta). When compared to the PBE-D simulations, all MOFs optimized with the M06-L functional had an increase in charge on the transition metal binding site with the exception of Cr-MIL-88B. Collectively, these results suggest that one shortcoming of the PBE exchange-correlation functional is that there is too much electron density on the open metal site, which is shifted to the surrounding framework atoms with the PBE-D+*U* and M06-L functionals. The observation that the partial charge on the metal often, but not always, increases when Hubbard *U*

corrections are applied has also been noted by Zhao and Kulik for various open-framework transition metal solids.[376]

Table 6.1. Median values of the Bader charge, $q_{\mathrm{M,Bader}}$, and (absolute) Bader spin density, $|\rho_{\mathrm{M,Bader}}|$, on the coordinatively unsaturated metal site for each transition metal and level of theory for the MOFs shown in Figure 6.1.

	median$(q_{\mathrm{M,Bader}})$			median$(\lvert\rho_{\mathrm{M,Bader}}\rvert)$		
Metal	PBE-D	M06-L	PBE-D+*U*	PBE-D	M06-L	PBE-D+*U*
V	1.63	1.65	1.55	2.31	2.53	2.65
Cr	1.40	1.45	1.42	3.57	3.70	3.77
Mn	1.44	1.50	1.48	4.52	4.62	4.69
Fe	1.36	1.43	1.37	3.62	3.74	3.73
Co	1.16	1.33	1.27	0.99	2.66	2.68
Ni	1.08	1.26	1.23	1.10	1.67	1.76
Cu	1.04	1.16	1.07	0.53	0.62	0.61

The (absolute) spin density on the transition metal binding site was also computed at each of the three levels of theory. In general, the spin density of the binding site tends to increase as PBE-D < M06-L < PBE-D+*U* (Table 6.1). For every MOF considered in this work, the spin density predicted at the PBE-D level of theory was the lowest of the three DFAs. Notably, the Co- and Ni-containing MOFs appear to have the largest change in spin density when going from the PBE-D functional to M06-L or PBE-D+*U*. This is a direct consequence of the PBE exchange-correlation functional artificially disfavoring high-spin states.[228,229] Out of the 41 MOFs investigated in this work, six had a different spin state depending on the employed functional, and all had either Co or Ni sites. Five of these MOFs contain formally divalent metal cations. In each case, these five MOFs were predicted to have low-spin ground-states at the PBE-D level of theory but high-spin ground-states at the M06-L and PBE-D+*U* levels of theory (Table 6.2). For reference, prior experimental and theoretical work has shown that both Ni_2(dobdc) and Co_2Cl_2(btdd) (H_2btdd = bis(1*H*-1,2,3-triazolo[4,5-b],[4',5'-i])dibenzo[1,4]dioxin)), the large-pore analogue of Co_2Cl_2(bbta), exhibit

high-spin ground-states,[25,377] in contrast with the low-spin results from PBE-D. This suggests that both the M06-L and PBE-D+U approaches correct for the erroneous favoring of low-spin states associated with the PBE-D level of theory. One MOF with formally trivalent Co sites (i.e. Co-MIL-88B-OH) also had a different ground state spin state depending on the employed functional (Table E.1).

Table 6.2. The five MOFs in this study with divalent open metal sites that had different ground-state spin-states depending on the employed density functional approximation. The difference in energy between the high- and low-spin state, $\Delta E_{\mathrm{HS-LS}}$, is shown. Negatives values correspond to a high-spin ground-state and *vice versa*.

	$\Delta E_{\mathrm{HS-LS}}$ (kJ/mol)		
MOF	PBE-D	M06-L	PBE-D+U
Co_2(dsbdc)	+18	−4	−28
Co_2Cl_2(bbta)	+17	−11	−32
$Co_2(OH)_2$(bbta)	+18	−3	−25
Ni_2(dobdc)	+12	−33	−47
Ni_2(dsbdc)	+24	−26	−43

6.3.2 O_2 and N_2 Adsorption Energies

With an understanding of how the three DFAs influence the properties of the bare metal sites, we now consider the adsorption of O_2 for the various MOFs in this work. The distribution of O_2 binding energies for each level of theory is shown in Figure 6.3a. By identifying the extrema, medians, and peaks in the binding energy distributions, it becomes apparent that the PBE-D functional results in the most exothermic adsorption energies among the three levels of theory, whereas the PBE-D+U functional results in the least exothermic adsorption energies. In fact, the inclusion of the Hubbard U correction weakens the degree of O_2 binding with the metal site for every investigated MOF (Figure 6.3b). This is consistent with prior GGA+U studies of transition metal oxides, where the SIE associated with GGA functionals results in an overestimation of the oxidation energies by favoring the oxidized state over the reduced state.[340] Similar behavior has also been observed when modeling O_2 binding to iron heme complexes.[378]

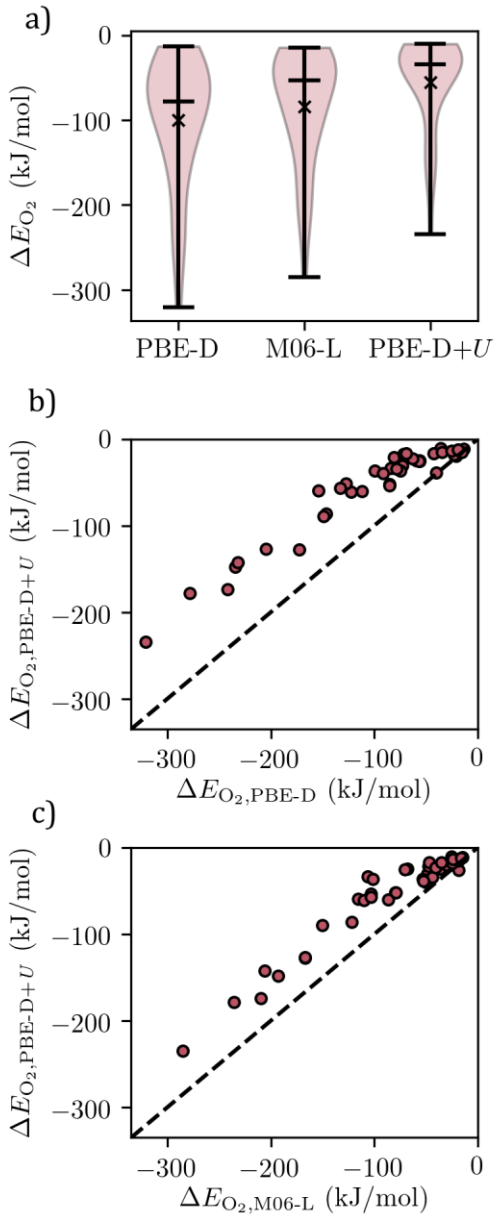


Figure 6.3. a) Violin plot representing the distribution of O_2 binding energies, ΔE_{O_2}. The extrema and median are marked by horizontal lines, and the mean is marked by an X. b) Comparison of O_2 binding energies computed using the PBE-D+U and PBE-D functionals. c) Comparison of O_2 binding energies computed using the PBE-D+U and M06-L functionals. The dashed lines represent perfect agreement between the two levels of theory.

Despite the range of metal elements and empirically-derived U values, the difference in O_2 adsorption energies at the PBE-D and PBE-D+U levels of theory are systematically offset from

one another, particularly for $\Delta E_{O_2,\mathrm{PBE-D}} < -50$ kJ/mol. Analogous trends are found with the M06-L functional, which predicts O_2 binding strengths between PBE-D and PBE-D+*U* (Figure 6.3c). The approximately linear relationships between the various DFAs tends to level off at small binding energies (e.g. $\Delta E_{O_2,\mathrm{PBE-D}} > -50$ kJ/mol), which can be attributed to a switch between chemisorption and physisorption of the O_2 guest molecule. For instance, MOFs with open metal sites that cannot be readily oxidized (e.g. Cu^{2+}) are not expected to significantly reduce the O_2 molecule, and inclusion of a Hubbard *U* correction therefore has only a minor impact on ΔE_{O_2} for these systems.

Both GGA+*U* and hybrid functionals are designed to reduce electron−electron SIEs and so numerous parallels inevitably emerge between the two approaches.[376,379,380] When modeling O_2 adsorption at coordinatively unsaturated metal sites of inorganometallic complexes, it has been shown that the incorporation of Hartree-Fock exchange leads to more endothermic ΔE_{O_2} values compared to pure GGA functionals.[326,364,381,382] Since the use of hybrid functionals in plane-wave, periodic DFT studies of MOFs is rarely computationally tractable, it is promising that the inclusion of a Hubbard *U* correction mimics the more endothermic binding strengths often found with the more expensive (and typically more accurate) hybrid functionals. It is also reassuring that the M06-L functional shifts the binding energies to be more endothermic than those at the PBE-D level of theory despite not including Hartree-Fock exchange to correct for the electron SIE.

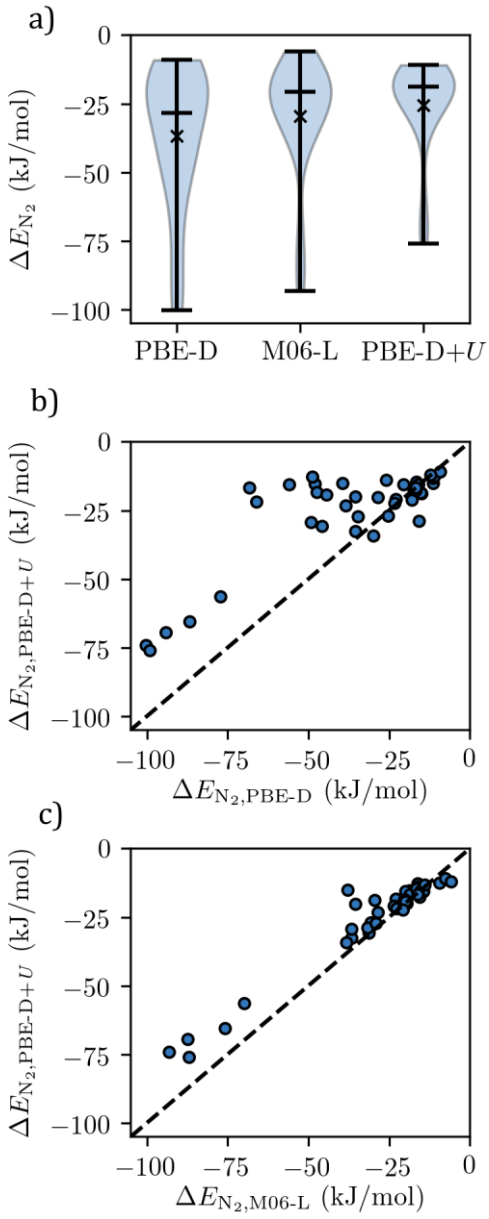


Figure 6.4. a) Violin plot representing the distribution of N_2 binding energies, ΔE_{N_2}. The extrema and median are marked by horizontal lines, and the mean is marked by an X. b) Comparison of N_2 binding energies computed using the PBE-D+U and PBE-D functionals. c) Comparison of N_2 binding energies computed using the PBE-D+U and M06-L functionals. The dashed lines represent perfect agreement between the two levels of theory.

With regards to N_2 binding, on average, the U correction tends to weaken the M−N_2 interaction compared to U = 0 eV (Figure 6.4a) but does so less consistently than for the M−O_2

interactions (Figure 6.4b). Similar to what was discussed for O_2 adsorption, if $\Delta E_{N_2,\mathrm{PBE-D}}$ is fairly weak (e.g. $\Delta E_{N_2,\mathrm{PBE-D}} > -30$ kJ/mol), the PBE-D and PBE-D+*U* calculations give roughly the same N_2 binding strengths due to weak interactions between the N_2 guest and the metal sites. For $\Delta E_{N_2,\mathrm{PBE-D}}$ values between −70 and −30 kJ/mol, the corresponding PBE-D+*U* binding strengths are all −30 to −15 kJ/mol and are mostly not correlated with the PBE-D binding strengths. Compared with the PBE-D functional, this creates a clearer division between MOFs that strongly bind N_2 (i.e. the V^{2+}-containing MOFs) and the remainder that do not. Similar behavior was observed in the range of weak O_2 adsorption (note the flat region at the top-right of Figure 6.3b) but is less pronounced due to the wider range of O_2 adsorption energies. The N_2 binding strengths at the M06-L level of theory are, overall, quite similar to those computed at the PBE-D+*U* level of theory, particularly for the weaker adsorption region of $\Delta E_{N_2} > -50$ kJ/mol.

The difference in binding energies between the PBE-D and PBE-D+*U* levels of theory can be quite large for many of the MOFs investigated in this work. For certain MOFs, the differences in ΔE_{O_2} can be as large as ~100 kJ/mol and the differences in ΔE_{N_2} can be as large as ~50 kJ/mol. The significant weakening of the metal−guest bond for chemisorbed O_2 and N_2 species suggests that the delocalization error present with the PBE-D functional can have a profound impact on the predicted energetics. This is in contrast with many prior GGA+*U* adsorption studies of MOFs that have mainly focused on the binding of small molecules that are highly stabilized by electrostatic interactions rather than charge transfer from the metal to the guest species. For instance, prior studies on M_2(dobdc) have shown that the inclusion of a Hubbard *U* correction can change CO_2 binding energies by up to 8.5 kJ/mol depending on the transition metal and *U* value.[369,370] Since the chemisorption of O_2 and N_2 are redox processes that alter the oxidation states of the metal and

guest molecule,[310] it is expected that there is significantly less cancellation of the SIE between the bare MOFs and MOF−adsorbate complexes, resulting in the large differences in binding energies observed in the present study.

While the results in Figure 6.3 and Figure 6.4 highlight the sensitivity of the O_2 and N_2 adsorption energies, it is worth investigating if important qualitative trends hold regardless of the DFA. To test this, the O_2 and N_2 binding strengths for each MOF and at each level of theory are shown in Figure 6.5. For determining structure−property relationships between related MOF families, it appears that PBE-D, M06-L, and PBE-D+*U* generally yield the same overall conclusions. For instance, M_2(dobdc) is found to bind O_2 stronger than its S-containing analogue M_2(dsbdc), $M_2(OH)_2$(bbta) is found to bind O_2 stronger than the closely related M_2Cl_2(bbta), the metal cations of MIL-88B bind O_2 stronger than their higher valent MIL-88B-OH analogues, and early transition metals tend to bind O_2 stronger than late transition metals for a given isostructural series of MOFs (as discussed in greater detail in prior work[73]). Since the majority of the MOFs studied in this work do not strongly bind N_2, there are fewer trends to compare with regards to N_2 binding, but for all three levels of theory, the V^{2+}-containing MOFs exhibit the most exothermic N_2 adsorption energies.

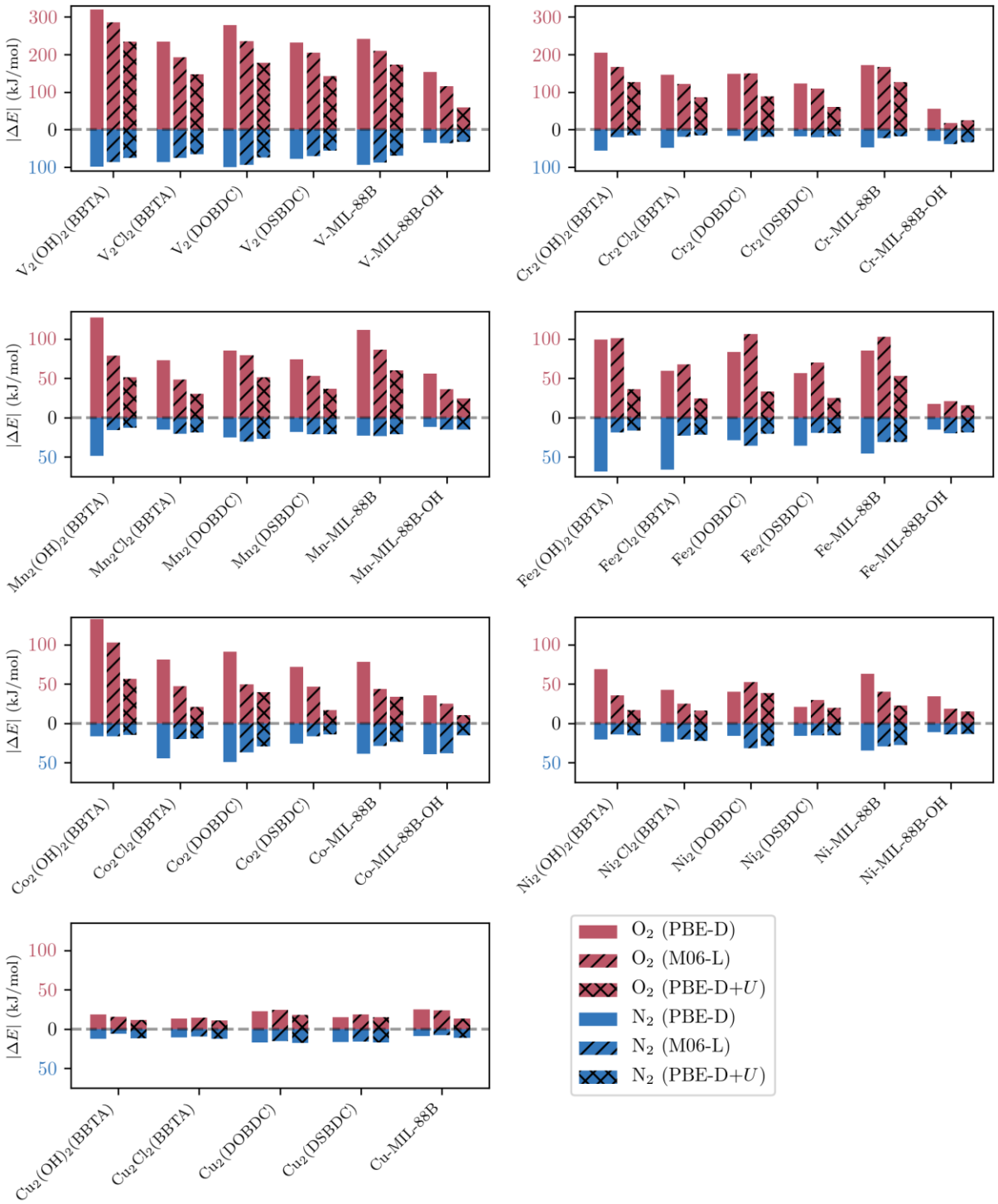


Figure 6.5. Absolute binding energy of O_2 (top bars, red) and N_2 (bottom bars, blue) for each MOF considered in this work at the PBE-D (left bars, no hatch), M06-L (middle bars, diagonal hatch), and PBE-D+U (right bars, cross hatch) levels of theory. For visual clarity, the binding strengths of

O_2 are displayed above the horizontal dashed line, whereas the binding strengths of N_2 are displayed below the horizontal dashed line; however, adsorption is exothermic in all cases.

Qualitative trends related to relative binding energies, however, are more sensitive to the level of theory, particularly when identifying MOFs that can selectively adsorb N_2 over O_2. As one example, Cr-MIL-88B-OH is predicted to preferentially adsorb N_2 over O_2 with PBE-D+*U* and M06-L but not PBE-D (for reference, the closely related MOF Cr-MIL-100 is known[357] to preferentially adsorb N_2 over O_2). Regardless of the choice of DFA, the large sensitivity of the O_2 (and occasionally N_2) binding energies suggests that numerous levels of theory should be compared when suggesting a particularly promising MOFs for applications involving the selective binding of O_2 or N_2.

The results in Figure 6.5 also reinforce the observation that, for chemisorbed O_2 species, the magnitude of the binding strengths often decreases as PBE-D > M06-L > PBE-D+*U*. However, this trend appears to break down in select cases, mainly when there are errors present in both the redox activity and the spin states. For example, the $\Delta E_{\mathrm{O_2,PBE-D}}$ and $\Delta E_{\mathrm{O_2,PBE-D+}U}$ values are similar for both Ni_2(dobdc) and Ni_2(dsbdc), but this is merely due to anomalous error cancellation in the PBE-D level of theory due to an artificial stabilization of low-spin Ni(II) centers, as was shown in Table 2. For the Fe-containing MOFs, $\Delta E_{\mathrm{O_2,PBE-D}}$ tends to be less exothermic than $\Delta E_{\mathrm{O_2,M06-L}}$, in contrast with the trends observed for the other metals considered in this work. This can also be attributed to an over-stabilization of the low-spin states at the PBE-D level of theory, which reduces the thermodynamic favorability of O_2 adsorption in an unphysical manner (Table E.3).

6.3.3 Degree of Charge Transfer and Geometrical Properties

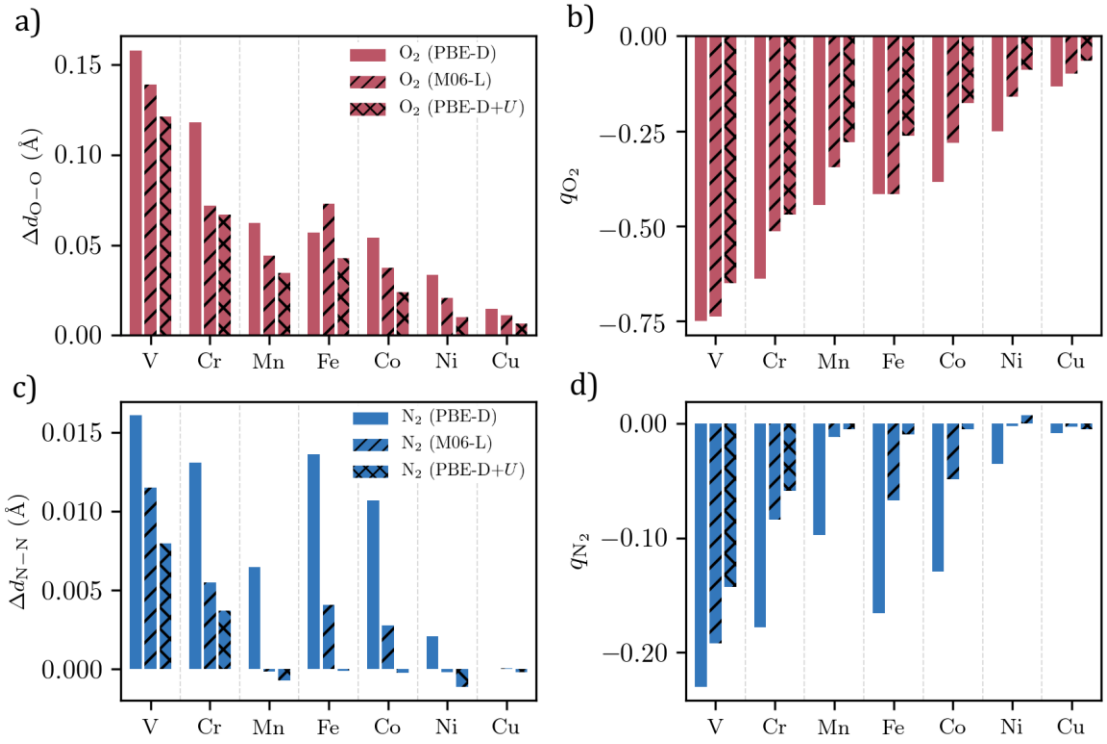


Figure 6.6. a) Mean signed difference (MSD) in the O−O distance of the O_2 guest molecule with respect to free O_2, Δd_{O-O}, for a given metal. b) Average sum of Bader charges on the O_2 guest molecule, q_{O_2}, for a given metal. c) MSD in the N−N distance of the N_2 guest molecule with respect to free N_2, Δd_{N-N}, for a given metal. d) Average sum of Bader charges on the N_2 guest molecule, q_{N_2}, for a given metal. Positive values for Δd_{O-O} and Δd_{N-N} indicate larger bond distances than the free adsorbate and *vice versa*. The left (no hatch), middle (diagonal hatch), and right (cross hatch) set of bars represent the PBE-D, M06-L, and PBE-D+U levels of theory, respectively.

To further rationalize the changes in adsorption energies with the three DFAs considered in this work, we also investigated the redox states of the bound adsorbates, using the PBE-D calculations as a baseline. We determined the redox state of the adsorbate via two readily computed metrics. The first is a measure of how stretched the O−O and N−N bonds are compared to the gas-phase values for a given level of theory. The second is the sum of partial atomic charges on the adsorbate. As shown in Figure 6.6a, while the O−O distance generally increases upon adsorption, the adsorbed O_2 species at the M06-L and PBE-D+U levels of theory have O−O bond distances

closer to that of free O_2, with the PBE-D+U level of theory resulting in the least stretched O−O bonds. Furthermore, for a given transition metal, the magnitude of the partial charge of the O_2 adsorbate generally decreases as PBE-D > M06-L > PBE-D+U (Figure 6.6b). Both of these observations are fully consistent with the adsorption energy trends previously shown in Figure 6.3a and indicate that the O_2 molecule is most significantly reduced at the PBE-D level of theory, followed by M06-L and PBE-D+U. Analogous trends are observed for N_2 adsorption at the open metal sites of MOFs (Figure 6.6c and Figure 6.6d).

In prior work by Neaton and coworkers, it was found that increasing U from 0 to 5.3 eV caused the CO_2 binding energy at the open metal sites of Co_2(dobdc) to change from approximately −32 kJ/mol to −36 kJ/mol.[369] This strengthened interaction between the metal site and the guest molecule was attributed to more spatially localized d states with the inclusion of a Hubbard U correction, which allowed the CO_2 molecule to adsorb closer to the Co^{2+} open metal site, increasing the electrostatic contribution to the binding energy.[369] Motivated by this observation, we computed the M−O and M−N distances for each MOF and with each DFA (Figure 6.7). In contrast with the aforementioned CO_2 example, the inclusion of the Hubbard U correction increases the M−O and M−N distance for the majority of MOFs in this work. This finding is a consequence of weaker binding strengths when U > 0 eV and suggests that, even if the d states are more localized, the reduced degree of charge transfer is significantly more important than any changes to the electrostatic and dispersive interactions for these systems. Consistent with the trends in adsorption energies (Figure 6.3a and Figure 6.4a), the M−O and M−N distances tend to increase as PBE-D < M06-L < PBE-D+U.

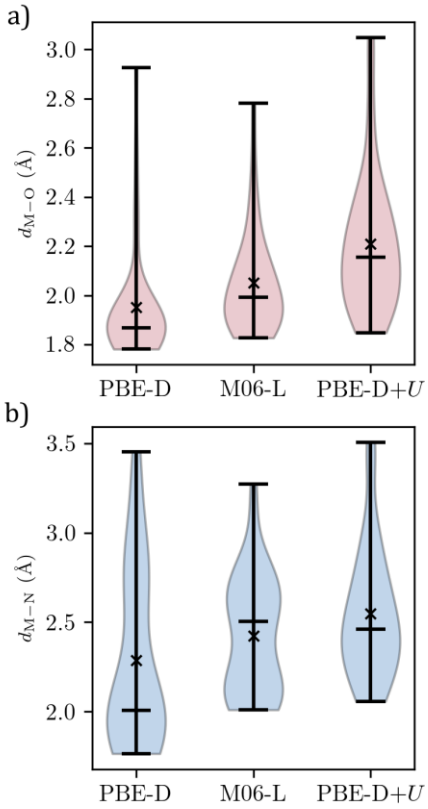


Figure 6.7. Violin plots of the distance between a) the metal binding site and the closest O atom of the O_2 guest, d_{M-O}, and b) the metal binding site and the closest N atom of the N_2 guest, d_{M-N}, at the PBE-D, M06-L, and PBE-D+U levels of theory. The extrema and median are marked by horizontal lines, and the mean is marked by an X.

6.3.4 Determining Empirically Ideal *U* Values

The previous sections have highlighted how the redox-dependent binding of O_2 and N_2 are influenced by common GGA, meta-GGA, and GGA+U approaches. Here, we solely focus on the GGA+U approach and consider the impact that variations in U have on the predicted O_2 adsorption energies (i.e. considering a range of U values beyond the empirically selected values based on the work of Ceder and coworkers[340]). We chose to investigate five MOFs with a wide range of experimentally determined isosteric heats of O_2 adsorption: M_2(dobdc) (M = Fe, Co, Ni),[24,326] Co-BTTri (H_3BTTri = 1,3,5-tri(1H-1,2,3-triazol-5-yl)benzene),[216] and Cr-BTT (H_3BTT = benzene-1,3,5-tris(1*H*-tetrazole)).[22] The structures of Co-BTTri and Cr-BTT are shown in Figure 6.8. Of

these MOFs, Cr-BTT and Fe_2(dobdc) are known to chemisorb O_2 via partial oxidation of the metal sites,[22,24] Co-BTTri can bind O_2 but does so with a relatively small degree of charge transfer,[216] and M_2(dobdc) (M = Co, Ni) exhibit only physisorption of O_2 at room temperature.[326] The MOFs were not diluted with Mg^{2+} cations for this analysis in order to aid direct comparison with experiments. Additional methodological details can be found in the Appendix.

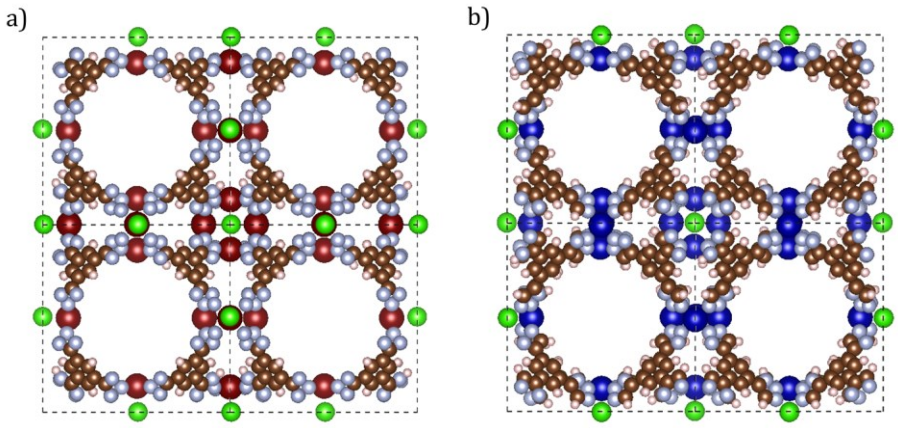


Figure 6.8. Structures of a) Cr-BTT and b) Co-BTTri. Aside from the different metals, the MOFs differ in the number of N atoms per linker. The simulation unit cells are shown as dashed black lines. Color key: Cr (maroon), Co (dark blue), Cl (green), N (light blue), C (brown), H (white).

As shown in Figure 6.9a, the O_2 adsorption energies change monotonically with U, which is consistent with adsorption behavior modeled in several prior GGA+U studies.[369,378,383] In particular, we observe that larger U values result in weaker binding of O_2 to the metal centers of each of the investigated MOFs. For highly exothermic O_2 binding, increasing U results in an approximately linear decrease in the strength of O_2 adsorption. However, the change in ΔE_{O_2} as a function of U begins to plateau as the O_2 adsorption behavior switches from chemisorption (with oxidation of the metal center) to physisorption (with the adsorbate largely stabilized by electrostatic and dispersive interactions). This transition point is different depending on the MOF but is generally in the range of −40 to −20 kJ/mol.

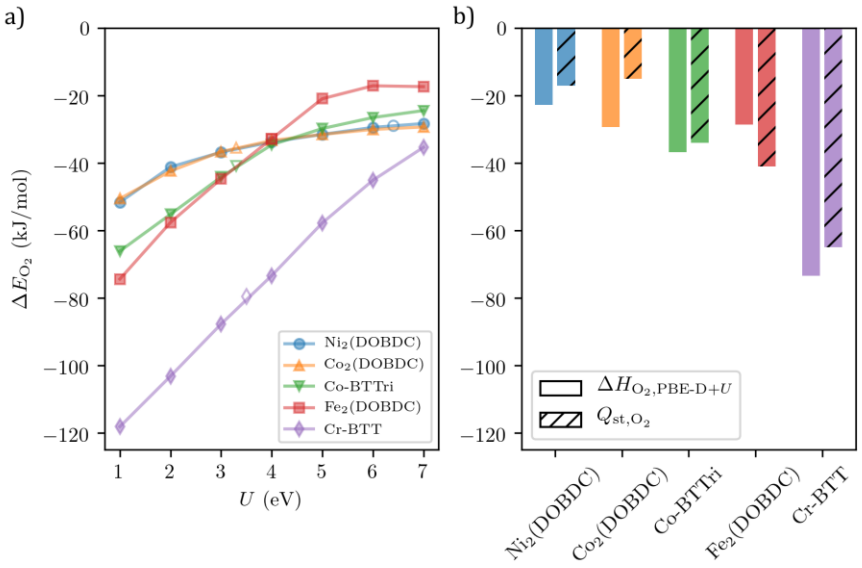


Figure 6.9. a) O_2 adsorption energy, ΔE_{O_2}, with the PBE-D functional and U = 1 – 7 eV for Ni_2(dobdc), Co_2(dobdc), Co-BTTri, Fe_2(dobdc), and Cr-BTT. The ΔE_{O_2} values corresponding to the U values from Ceder and coworkers[340] are marked with open symbols for ease-of-reference. b) O_2 adsorption enthalpies at the PBE-D level of theory using U = {Ni: 6.4 eV, Co: 3.3 eV, Fe: 4.0 eV, Cr: 3.5 eV}, $\Delta H_{O_2,\text{PBE-D+}U}$, alongside the corresponding experimentally determined (low-loading) isosteric heats of O_2 adsorption, Q_{st,O_2}. Unit cell shapes and volumes of the guest-free structures were relaxed for each U value.

It is also worth comparing the predicted O_2 adsorption enthalpies with the experimentally measured (low loading) isosteric heats. For the MOFs where O_2 is known to chemisorb to the metal sites via charge-transfer interactions (i.e. Co-BTTri, Fe_2(dobdc), and Cr-BTT), the metal oxide-based U values yield adsorption enthalpies that are within 20 kJ/mol (and more often 10 kJ/mol) of the experimentally determined isosteric heats of O_2 adsorption as shown in Figure 6.9b. Although by no means perfect, these U values represent a vast improvement over the exclusion of Hubbard U corrections altogether. This is reassuring given that the U values of Ceder and coworkers are based solely on reproducing the oxidation energies of transition metal oxides,[340] which have very different electronic structure environments than those of isolated metal sites

incorporated within the nodes of MOFs. For the MOFs where it is known that O_2 is physisorbed at the metal centers (i.e. Co_2(dobdc), Ni_2(dobdc)), the inclusion of a Hubbard *U* correction does not necessarily reduce the O_2 binding affinity to the experimental isosteric heats even at large values of *U*, indicating that this is a fundamental limitation of the PBE exchange-correlation functional. The *U* values for M_2(dobdc) (M = Fe, Co, Ni) calculated via linear response theory (i.e. U = 6.5 eV, 5.3 eV, and 6.7 eV for Fe_2(dobdc), Co_2(dobdc), and Ni_2(dobdc), respectively),[369] while derived from first-principles, also do not result in strong agreement with experiment. In the case of Fe_2(dobdc), for instance, the transition metal oxide *U* value of 4.0 eV already significantly underpredicts the magnitude of the O_2 adsorption enthalpy. A value of U = 6.5 eV would lead to essentially no binding at all based on the results in Figure 6.9a.

6.4 Conclusion

Using periodic DFT, we investigated the adsorption of O_2 and N_2 at the open metal sites of 41 MOFs at three different levels of theory, including PBE with dispersion corrections (PBE-D), PBE-D with empirical Hubbard *U* corrections that reproduce oxidation energies in transition metal oxides[340] (PBE-D+*U*), and the M06-L functional. For the bare MOFs (i.e. prior to the adsorption of any guest molecules), the PBE-D functional erroneously disfavors high-spin states, which is corrected by the M06-L and PBE-D+*U* functionals. This is also reflected in the absolute spin density on the metal site, which tends to increase as PBE-D < M06-L < PBE-D+*U*. With regards to the partial atomic charge on the metal site, PBE-D tends to result in the lowest (i.e. least positive) charge for most MOFs. The inclusion of Hubbard *U* corrections often increases the partial charge on the metal, except for the V^{2+} open metal sites, in which it decreases.

By directly comparing the PBE-D and PBE-D+U levels of theory, it is apparent that the many-electron SIE associated with the PBE functional often leads to overly exothermic O_2 adsorption energies, which are always made less exothermic with the inclusion of a Hubbard U correction. The M06-L functional, despite not explicitly correcting for the many-electron SIE, generally improves on the PBE-D functional by decreasing the magnitude of the adsorption energies so that they are between PBE-D and PBE-D+U. For N_2 adsorption, the U correction also reduces the binding strength on average. However, since few MOFs in this study were found to strongly bind N_2 and there is less charge transfer associated with weakly physisorbed species, this weakening effect is less consistent for N_2 adsorption. The M06-L and PBE-D+U functionals predict similar N_2 adsorption energies, with the M06-L functional resulting in slightly stronger binding than PBE-D+U on average.

Clear trends are observed with regards to the redox state and bond distances associated with the O_2 and N_2 guest molecules. The negative charge on the adsorbates increased in magnitude in the order PBE-D+U < M06-L < PBE-D, indicating that the weaker adsorption energies are associated with a reduced degree of electron transfer to the bound guest molecules. Analogous behavior was observed for the O−O and N−N bond distances, with PBE-D resulting in the longest (i.e. weakest) bonds. Despite the GGA+U method resulting in more highly localized $3d$ orbitals, the inclusion of a +U correction increased the M−O_2 and M−N_2 distances, once again consistent with the significantly weaker adsorption energies when correcting for the GGA-related delocalization error.

Although the qualitative trends between related MOF families are predicted to be the same in most cases, we recommend against using the PBE-D functional for redox-dependent adsorption

processes given its artificial disfavoring of high-spin states and tendency to yield overly exothermic binding strengths. The inclusion of empirical Hubbard U corrections is a computationally efficient way of addressing many of the errors associated with the PBE-D functional and is arguably the better option when studying a large number of MOFs for applications involving redox-dependent binding at open metal sites. The M06-L functional serves as a middle-ground between PBE-D and PBE-D+U in many cases (except for O_2 adsorption in MOFs with Fe^{2+} sites) and has the advantage that it does not rely on a user-specified and (in principle) material-specific U value, although the O_2 binding strengths are likely to still be overestimated compared to experiments and the computational cost is greater than that of PBE-D+U.

Finally, we emphasize that the quantitative values for the O_2 and N_2 adsorption energies are highly sensitive to the delocalization error, especially when the adsorption is relatively strong (i.e. when charge transfer is present). Given the nearly unlimited number of MOFs that can potentially be realized, there is a significant need for the development of computationally inexpensive methods to reliably predict redox-dependent adsorption energies at open metal sites in MOFs, such as further validation of U values that are tailored for $3d$ transition metal sites incorporated at the nodes of MOFs. The development of such methods would greatly accelerate the design of promising MOFs for catalysis, gas separations, and related applications where the redox-dependent binding of small molecules governs the success of a given material candidate.

Chapter 7: Machine Learning the Quantum-Chemical Properties of Metal–Organic Frameworks for Accelerated Materials Discovery

In this chapter, we introduce the QMOF Database – a publicly available database of computed quantum-chemical properties for 14,000+ experimentally synthesized MOFs. Throughout this study, we demonstrate how machine learning models trained on the QMOF Database can be used to rapidly discover MOFs with targeted electronic structure properties, using the prediction of DFT-computed band gaps as a representative example. We conclude by highlighting several MOFs predicted to have low band gaps, a challenging task given the electronically insulating nature of most MOFs.

This chapter is adapted from the following peer-reviewed article: A.S. Rosen, S.M. Iyer, D. Ray, Z. Yao, A. Aspuru-Guzik, L. Gagliardi, J.M. Notestein, R.Q. Snurr. "Machine Learning the Quantum-Chemical Properties of Metal–Organic Frameworks for Accelerated Materials Discovery." *Matter*, **4**, 1578–1597 (2021). This work was featured on the cover of *Matter*, in Northwestern Engineering News, and previewed in the journal *Patterns*.[384] A.S. Rosen designed the project and carried out the bulk of the work presented in this chapter. S.M. Iyer helped benchmark and test the various machine learning models. D. Ray assisted in interpreting the density of states plots. Z. Yao provided feedback on the motivation and machine learning. All authors provided feedback on the manuscript.

7.1 Introduction

Over the last several years, significant attention has been focused on the design of novel metal−organic frameworks (MOFs), a class of materials composed of discrete inorganic nodes connected to one another via organic linkers. One of the main advantages of MOFs is that they often have predictable and atomically defined structures with properties that are directly related to the choice of underlying metal and organic building blocks.[1] In this way, it becomes possible to impart physical and chemical functionality specifically tailored for a given application.[3] To date, tens of thousands of MOFs have been synthesized,[14,44] and a nearly unlimited number can be proposed[53,90,385] by considering different combinations of constituent building blocks. Due to the enormous set of possible framework compositions, structures, and resulting properties,[386] it remains difficult to discover truly top-performing MOFs for a particular application based solely on chemical intuition, conventional trial-and-error experimental testing, or serendipity alone.

High-throughput computational screening approaches based on classical simulations have proven extremely useful for more efficiently exploring the vast combinatorial space of MOF structures.[46,48] Recently, the large quantities of data generated during these computational screening studies have led to the development of machine learning (ML) models[387] that can accelerate the MOF design and discovery process even further. ML-assisted screening studies have been successfully applied to the discovery of MOFs suitable for H_2 storage,[388–390] CO_2 separation/capture,[391–393] and numerous other applications predominantly (although not exclusively[394,395]) in the area of gas storage and separations.[48,50,396] Nonetheless, similar efforts remain almost entirely unexplored for the many applications where the properties of interest are best described by quantum mechanical models,[4] such as those based on the electronic, optical,

magnetic, and/or catalytic properties of MOFs. Beyond the sheer number of possible MOFs that can be realized, the large number of atoms in MOF crystal structures often makes it computationally demanding to carry out even moderate-scale quantum-chemical screening studies, further magnifying the need for ML approaches in this area.

To date, the most relevant studies focused on training ML models to predict the quantum-chemical properties of MOFs are those of Raza et al.,[397] Korolev et al.,[398] and Kancharlapalli et al.[399] who developed ML models that can predict the partial atomic charges of MOFs in the Computation-Ready, Experimental (CoRE) MOF database.[89,234] Beyond these fundamental studies on partial charge prediction, however, there remains a significant gap in the literature, particularly for the discovery of MOFs with desired electronic structure properties. To the best of our knowledge, the only prior work in this area is that of He et al.[400] who trained binary classification models to predict whether inorganic solids in the Open Quantum Materials Database (OQMD)[30,34] are metallic or nonmetallic. Without retraining on MOF data, a multi-model voting procedure was then used to predict the metallic or nonmetallic behavior of 2932 MOFs in the CoRE MOF database,[89] which do not have computed band gaps. Of the six identified materials with near-zero band gap at the PBE level of theory,[125] all are best-described as metal–cyanide/thiocyanate cluster complexes and none have H atoms in the structure. This is likely due in large part to the extreme differences between the OQMD, which consists almost entirely of inorganic compounds, and the CoRE MOF database. Furthermore, the fidelity of the metallic materials was not considered, leading to highlighted structures such as $[CdC_4]_n$ that should actually be $[Cd(CN)_2]_n$.[401]

In the present study, we leverage a recently developed high-throughput periodic DFT workflow tailored for MOF structures[32] to construct a large-scale database of MOF quantum

mechanical properties. This publicly available dataset – the Quantum MOF (QMOF) database[402] – contains computed properties for 15,713 experimentally characterized MOFs after structure relaxation via DFT, including but not limited to optimized geometries, energies, band gaps, charge densities, density of states, partial charges, spin densities, and bond orders. We anticipate that the QMOF Database will serve two primary purposes: 1) materials discovery using the as-deposited data; 2) the evaluation and development of novel ML algorithms to reduce, or circumvent altogether, the need for otherwise expensive DFT calculations.

To demonstrate the utility of the data generated via the high-throughput DFT workflow, we use the QMOF Database to develop several ML models for the prediction of MOF band gaps from nothing more than an encoding of the experimental (i.e. unrelaxed) crystal structures, drastically decreasing the number of computationally demanding quantum mechanical simulations that would need to be carried out in future screening studies. Beyond serving as a proof-of-concept, an ML model that can predict MOF band gaps is particularly desirable, as most MOFs are known to be electronically insulating,[16] which limits their potential use in electrocatalysis, sensing, energy storage, and other applications where some degree of electrical conductivity is necessary.[16,17,403–405] We identify a top-performing band gap regression model based on a crystal graph convolutional neural network[406] and show how dimensionality reduction techniques can be used to discover overarching structure–property relationships for the identification of MOFs with targeted electronic structure properties. We conclude by highlighting several iron MOFs with low band gaps identified for the first time in this work.

7.2 Results and Discussion

7.2.1 Generation and Overview of the QMOF Database

Prior to carrying out any periodic DFT calculations, a dataset of starting structures must be assembled. There are several databases of MOF structures that have been published to date.[14,44,53,89,90,112] However, it is imperative to note that existing databases of synthesized MOFs cannot be used as-is for quantum-chemical screening purposes. If even a single atom is missing or duplicated in a MOF crystal structure, the resulting DFT calculations are unlikely to be physically meaningful. Put another way, the simulation unit cell is expected to be charge-neutral unless otherwise specified; any additional or missing electron in the system ruins the integrity of the resulting charge density and, therefore, all the quantum-chemical properties derived from it. These situations can arise as a result of deficiencies in the deposited experimental crystal structure and/or in the dataset curation process when generating a database of MOF crystal structures. Therefore, in this work we aim to start with a comparatively "clean" dataset of crystal structures for high-throughput computational investigation, one we will refer to as a suitably "DFT-ready" dataset of MOFs.

We considered the list of materials identified as MOFs from both the Cambridge Structural Database (CSD) MOF subset[44] and the 2019 Computation-Ready, Experimental (CoRE) MOF database,[14] the latter of which contains a relatively small number of MOFs not present in the former. All starting structures were taken directly from the CSD by querying the corresponding CSD reference code ("refcode"), and free (i.e. unbound) solvents were automatically removed from the frameworks. We chose to take the initial structures directly from the CSD as a matter of consistency and so that we could make use of valuable CSD meta-data[407] (e.g. unresolved atoms, charged structures) associated with each crystal structure. From this set of experimental crystal

structures, we constructed a smaller DFT-ready subset of 42,349 non-disordered MOF structures ("QMOF-42349") after an extensive suite of automated fidelity checks, as summarized in Figure F.1. This process serves to filter out many problematic MOFs with omitted H atoms, fractional occupancies, deleted framework atoms, lone (i.e. unbonded) atoms, overlapping atoms, an improper number of charge-balancing ions, and other structural issues that have been discussed in several recent studies.[79,407–412] Of these 42,349 experimental crystal structures, a subset of materials with 300 atoms or fewer per primitive cell was considered such that high-throughput DFT calculations could be carried out in an efficient manner. Full structure relaxations (including cell volume and atomic positions) were carried out via a multi-stage workflow[32] (Table F.2) at the PBE-D3(BJ)[125–127] level of theory with the Vienna *ab initio* Simulation Package (VASP).[122,123] Additional methodological details regarding the dataset construction, DFT calculations, and ML methods can be found in the Supplemental Information.

The high-throughput periodic DFT workflow was successfully completed for 15,713 MOFs, and several DFT-computed properties were tabulated following the structure relaxation process, a selection of which are listed in Figure 7.1.[f] Of these, band gaps are likely to be of interest for electronic and optical properties, especially in the search for (semi)conducting MOFs[16–18,413] or screening for photocatalytic materials.[414] Electronic energies, particularly if converted to formation energies, may provide insight into the relative stability of MOFs.[415] Machine learning the charge density[416] is a potential way to bypass a large portion of the calculations performed with

[f] Following publication of this work, ~2600 unique MOFs were added to the QMOF Database, including ~1200 MOFs taken from the CSD MOF Subset, ~1200 hypothetical MOFs taken from the Boyd & Woo dataset,[629] ~150 MOF-74 and MOF-5 analogues,[630,631] and ~50 hypothetical Zr MOFs made using ToBaCCo.[90,621,632] Additional MOF structures are planned for forthcoming releases, including more hypothetical MOFs and a curated set[399] of experimental MOFs from the 2019 CoRE MOF Database.[14] The maximum number of atoms per cell was raised to 500.

Kohn-Sham DFT.[417–419] Both the charge density and density of states can provide insight into the electronic structure in addition to serving as promising features to predict a variety of other quantum-chemical properties.[420,421] Partial atomic charges,[158,230,232,289] bond orders,[231] and spin densities[158,230] have a wide range of potential use-cases, from describing electrostatic interactions in classical simulations of MOFs[234,422] to serving as descriptors to better understand trends in catalytic reactions[71,423,424] and small molecule binding.[73] Furthermore, the DFT-optimized structures can be used as starting points for further quantum-chemical calculations and for analyzing geometric properties of MOFs. In addition to the curated data mentioned in Figure 7.1, all output data from the DFT calculations are made publicly available so that other properties of interest can be readily investigated.

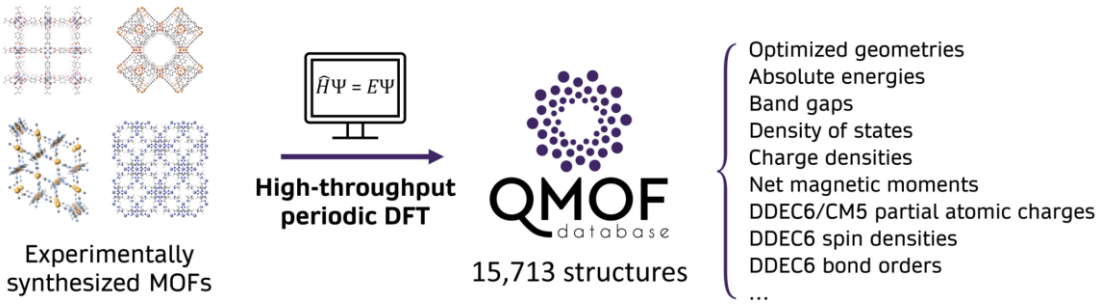


Figure 7.1. Selected DFT-computed properties for the structurally relaxed MOFs made available in the QMOF Database.

Prior to highlighting how this data can be used in practice, we first investigated several properties of the QMOF Database. As shown in Figure 7.2A, the QMOF Database contains MOFs with chemical elements that span nearly the entire periodic table, which is beneficial for the development of transferable ML models. As anticipated, there is also a large number of MOFs in the QMOF Database containing Cu, Zn, and Cd, which compose the three most common types of inorganic nodes in the MOF literature.[79] Nonetheless, we note that some types of MOF families are likely underrepresented in the QMOF Database due in part to the dataset curation process,

which filters out any MOFs with missing atomic coordinates or partial occupancies. These situations are likely to arise in MOFs with complex proton topologies that cannot be resolved from X-ray diffraction alone (e.g. Zr- and Hf-based MOFs[425,426]),[g] MOFs with defects in the crystal structure,[427] and MOFs that have undergone post-synthetic modification.[428,429]

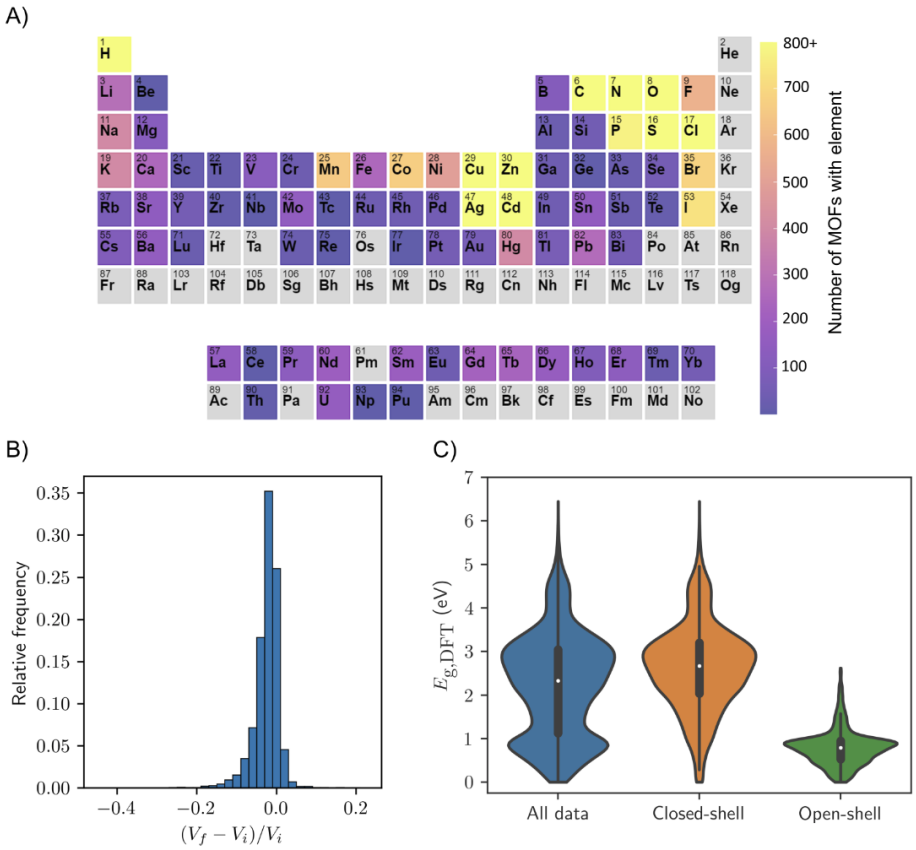


Figure 7.2. A) Number of MOFs in the QMOF Database containing a given element. All elements that occur in greater than 800 structures are capped at 800 for ease of visualization. These include: C (15,713), H (15,713), N (12,892), O (12,821), Cu (2,882), S (2,684), Zn (2,665), Cd (2,538), Cl (1,687), and Ag (1,213). Elements in gray are not present in any structure. B) Histogram of the fractional change in cell volume before (V_i) and after (V_f) structure relaxation at the PBE-D3(BJ) level of theory for the MOFs in the QMOF Database. C) Violin plots of the DFT-computed band gaps, $E_{g,DFT}$, at the PBE-D3(BJ) level of theory for the MOFs in the QMOF Database. Separate distributions are shown for the entire dataset (15,713 entries), the closed-shell MOFs (12,169 entries), and the open-shell MOFs (3,544 entries). Open-shell character is defined here as having

[g] Since the original publication of this work, a dataset of Zr- and Hf-containing MOFs has been added to the QMOF Database.

a DDEC6 atomic spin density with a magnitude greater than 0.1. A box plot, showing the extrema and interquartile range, is included in each violin with the median marked by a white dot.

When looking at the geometries before and after structure relaxation, we find that 96.6% of the DFT-optimized MOFs had a change in cell volume less than 10% (Figure 7.2B), suggesting that the removal of free solvent does not drastically alter the structural properties for most of the MOFs in this work. In the case of flexible MOFs, multiple conformations are often included in the QMOF Database, which is important since they may exhibit different electronic structure properties.[430] As depicted in Figure F.3, three distinct conformations of the flexible MOF Fe(bdp) (H_2bdp = 1,4-benzenedipyrazole)[431] are included in the QMOF Database (refcodes: QUPZIM, QUPZIM01, QUPZIM02), one of which is on the extreme end of the distribution shown in Figure 7.2A (Table F.5). This is not surprising given that high pressures of CH_4 are needed to stabilize the given open-pore configuration of Fe(bdp).[431]

The distribution of DFT-computed band gaps for the fully optimized structures at the PBE-D3(BJ) level of theory is shown in Figure 7.2C and indicates that there is a wide spread of values from nearly 0 eV to 6.45 eV. The band gaps are not normally distributed and instead are bimodal, with peaks centered around 0.9 eV and 2.9 eV. This can be attributed to different distributions associated with closed- and open-shell materials in the QMOF Database (Figure 7.2C), the latter of which have significantly lower band gaps at the PBE-D3(BJ) level of theory on average. With regards to partial atomic charges, a wide spread of values is also obtained (Figure F.5A). In comparing the partial atomic charges before and after structure relaxation, we find that 92.4% of the ~1.2 million data points have an absolute difference less than 0.05 q_e, and 98.9% of the points have an absolute difference less than 0.1 q_e (Figure F.5B). As has been observed on a smaller scale

in prior work,[112,234] it can be safely assumed that the partial charges remain essentially unchanged upon structure relaxation in most cases.

As a brief demonstration for how the data generated via the high-throughput DFT workflow could be used directly, we identified any porous framework materials with high-spin Fe species following the high-throughput DFT workflow. High-spin Fe complexes are known to be promising for oxidation catalysis, in particular for the activation of strong C–H bonds, and recent work has focused on stabilizing such motifs in MOFs for this purpose.[98,116,168] This query of the QMOF Database resulted in six unique MOFs, as shown in Figure F.6. Providing validation of this screening approach, two of the six MOFs – Fe_2(dobdc) (H_4dobdc = 2,5-dihydroxybenzene-1,4-dicarboxylic acid) (refcode: COKNOH)[432] and Fe_2(dobpdc) (H_4dobpdc = 4,4′-dihydroxy-(1,1′-biphenyl)-3,3′-dicarboxylic acid) (refcode: MALSIE)[204] – have already been shown to oxidize strong C–H bonds.[168,177,433] Another two of the six MOFs – Fe_2Cl_2(bbta) (H_2bbta = 1*H*,5*H*-benzo(1,2-d:4,5-d′)bistriazole) (refcode: HAYYUE)[193] and Fe_2Cl_2(btdd) (H_2btdd = bis(1*H*-1,2,3-triazolo[4,5-*b*],[4′,5′-*i*])dibenzo[1,4]dioxin) (refcode: HAYZAL)[193] – have been computationally investigated for their use in oxidation reactions.[71,72,74] Prior experimental studies suggest that the aforementioned MOFs exhibit high-spin Fe sites.[168,193,433]

7.2.2 Machine Learning Models for Band Gap Prediction

Beyond analyzing the DFT-computed properties directly, the QMOF Database now makes it possible to train a wide range of ML models specifically tailored for MOFs, which are likely to have their own distinct feature space compared to isolated molecules and inorganic solids. This serves two primary purposes. The first is more theoretical: featurization methods (i.e. how each MOF structure is numerically encoded) and ML algorithms that are well-suited for other materials

may not be equally suitable for MOFs, so this database of quantum-chemical properties can serve as a testing ground to benchmark new ML methods.[h] The Materials Project[33] and OQMD[30,34] in particular have accelerated this research direction for inorganic solids, and the QM9 dataset[434,435] (as one example) has done the same for small molecule chemistry. The second purpose of this new database is to apply these rapid yet accurate ML models to accelerate the materials discovery process, now with the ability to train these models directly on properties computed for MOFs.

In this work, we have chosen to develop an ML regression model that can rapidly predict the DFT-computed band gaps of MOFs. Specifically, we aim to predict the computed band gaps of the DFT-optimized structures from the un-optimized, experimentally resolved MOF crystal structures such that no quantum-chemical calculations need to be carried out. To achieve this, all ML models are trained on the band gaps of the DFT-optimized structures but take representations of the corresponding unrelaxed experimental structures as the input. Since the development of an ML regression model that can predict the band gaps of MOF crystal structures has not been achieved before, we trained several ML models using a variety of common featurization methods to benchmark each approach. These featurization methods are graphically summarized in Figure 7.3 for a representative material IRMOF-1 (IRMOF = isoreticular MOF),[210] also known as MOF-5 (Figure 7.3A). For the purposes of training ML models throughout this work, we specifically focus on a de-duplicated subset of 14,482 materials in the QMOF Database ("QMOF-14482") that have gone through the full periodic DFT volume relaxation process.

[h] Following publication of this work, a paper was published using the QMOF Database to benchmark various graph neural network models.[499]

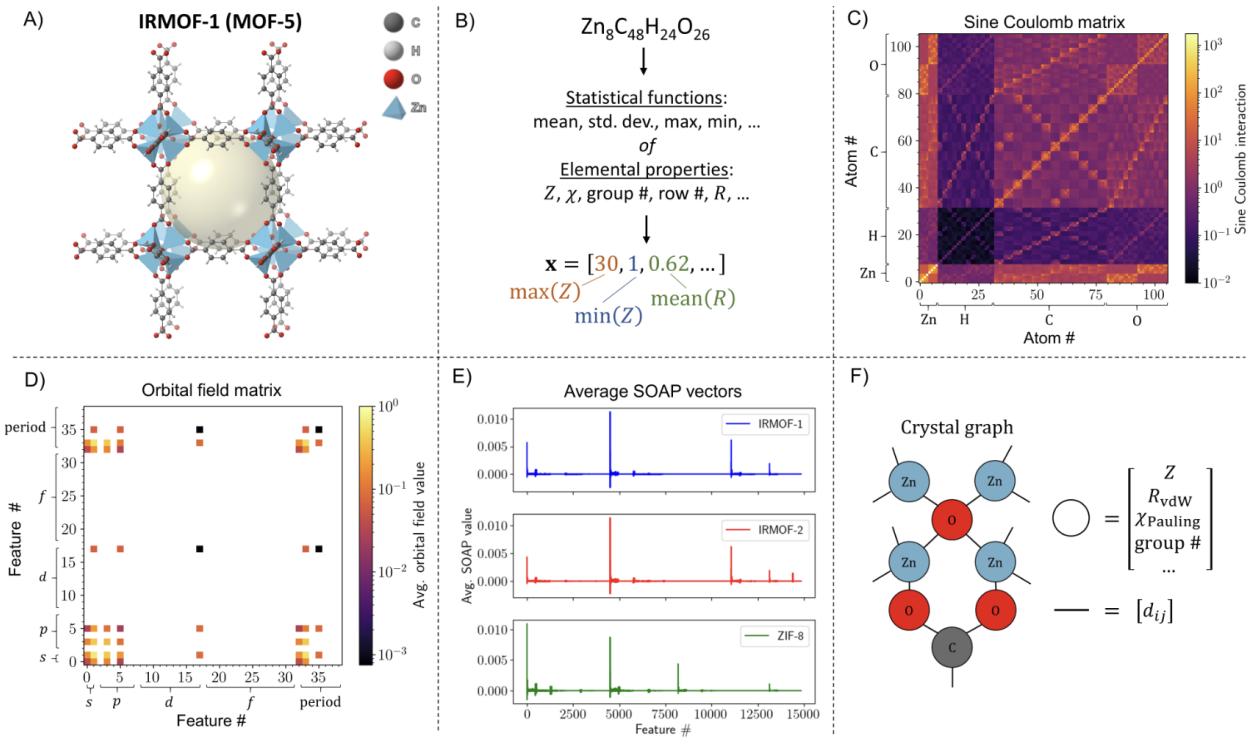


Figure 7.3. A) IRMOF-1 structure. B) Examples of composition-based features. C) Sine Coulomb matrix showing the interaction values between each pair of atoms. D) Orbital field matrix showing the average interaction value between each pair of orbital- or period-based features. Only non-zero values are shown. E) Averaged SOAP fingerprint of IRMOF-1 compared to IRMOF-2 and ZIF-8. Taking the dot product of any two vectors yields an unnormalized similarity score. F) Schematic of a crystal graph with example node (circle) and edge (line) embeddings (only a representative portion is shown for clarity).

The simplest featurization methods considered in this work are the feature sets of He et al.[400] (with 45 statistical attributes of elemental properties, denoted "Stoichiometric-45") and Meredig and Agrawal et al.[436] (with 103 attributes describing the elemental fractions from H–Lr and 17 statistical attributes of elemental properties, denoted "Stoichiometric-120"), which rely solely on the chemical composition of each material (Figure 7.3B). In addition, we consider several structure-sensitive featurization approaches, including the sine Coulomb matrix[437] that encodes pairwise electrostatic interactions between nuclei in a material (Figure 7.3C, Equation F.4) and the orbital field matrix[438] that encodes the distribution of valence electrons in each coordination

environment of a material (Figure 7.3D). The smooth overlap of atomic positions (SOAP)[439,440] is another structure-sensitive descriptor considered in this work, which can be used to compute the similarity between a pair of local atomic environments – and, by extension, a pair of structures – by representing the atoms as Gaussians (i.e. "smoothed positions") and comparing the spatial overlap in the resulting atomic density fields (Figure F.2 and Equations F.5–F.9). In all of the aforementioned examples, these features are used to develop a kernel ridge regression[441] (KRR) model (Equations F.1–F.3). Motivated by prior work on inorganic solids, we also investigated the use of a crystal graph convolutional neural network (CGCNN),[406] wherein an approximate crystal graph is generated for each MOF, with each node in the graph representing an atom and each edge representing the bonds that connect the atoms (Figure 7.3F). More detailed descriptions and full methodological details for each featurization method and ML model architecture can be found in the Supplemental Information.

As shown in Table 7.1, the KRR models trained on composition-based features (i.e. Stoichiometric-45 and Stoichiometric-120) are able to capture some of the band gap trends with mean absolute errors (MAEs) of 0.43 – 0.44 eV (with respect to the DFT-computed values) on the out-of-sample testing set. Nonetheless, these methods are still quite limited for regression purposes given that they do not encode any information about the structural properties of the MOF. In terms of structure-sensitive methods, taking an eigenvalue spectrum of the sine Coulomb matrix fares worse than the stoichiometry-based features, yielding a testing set MAE of 0.53 eV (Table 7.1). This can likely be traced back to the required use of zero-padding in the sine Coulomb matrix to ensure constant-length feature vectors between MOFs with different numbers of atoms per unit cell. The KRR model using a flattened orbital field matrix as the feature set is more accurate than

the model based on the sine Coulomb matrix but shows only a minor improvement over the stoichiometry-based features. Overall, SOAP performs the best of all tested KRR descriptor sets, with an MAE of 0.36 eV and $R^2 = 0.82$ on the testing set. The marked improvement in performance with SOAP is especially clear when comparing the parity plots of the different KRR models (Figure F.7).

Table 7.1. Summary of the testing set mean absolute error (MAE), coefficient of determination (R^2), and Spearman rank-order correlation coefficient (ρ) for several machine learning methods to predict the computed band gaps of MOFs from their deposited crystal structures with free solvent removed. Kernel ridge regression was used for all featurization methods except for the crystal graphs of CGCNN, for which a convolutional neural network was constructed. The testing set statistics are shown, averaged over five runs (using different random seeds for data splitting) with ±1 standard deviation shown. For all models, 80% of the QMOF-14482 dataset was used for training. The MAE for a dummy model that predicts the mean band gap (2.220 eV) for all the MOFs is shown for reference.

ML Method	MAE (eV)	R^2	ρ
Constant mean model	0.973	—	—
Sine Coulomb matrix	0.529 ± 0.008	0.643 ± 0.012	0.787 ± 0.008
Stoichiometric-45	0.437 ± 0.004	0.743 ± 0.006	0.842 ± 0.004
Stoichiometric-120	0.433 ± 0.010	0.750 ± 0.009	0.847 ± 0.005
Orbital field matrix	0.417 ± 0.008	0.763 ± 0.010	0.863 ± 0.003
SOAP	0.357 ± 0.008	0.822 ± 0.010	0.910 ± 0.003
CGCNN	0.274 ± 0.008	0.876 ± 0.011	0.932 ± 0.005

Notably, CGCNN significantly outperforms all the aforementioned KRR models, achieving an MAE of 0.27 eV and $R^2 = 0.88$ (Table 7.1). As a point of reference, a trivial model that simply predicts the mean band gap for every MOF would have an MAE of 0.97 eV, suggesting that CGCNN captures much of the underlying chemistry. The performance of the CGCNN model for MOF band gaps is comparable, if not slightly better, than state-of-the-art ML band gap models trained on inorganic solids from the OQMD and Materials Project as well as the organic crystals from the Organic Materials Database (OMDB).[37,406,442,443] It is also worth noting that the

experimentally measured band gaps of MOFs can vary by several tenths of an eV depending on the synthesis and post-treatment conditions.[444] As such, an MAE less than 0.3 eV is promising for the identification of structure–property trends and for sorting material candidates by band gap, the latter of which is further justified by the CGCNN's high Spearman rank-order correlation coefficient of $\rho = 0.93$. For context, it took ~8 minutes (7 minutes for a one-time encoding of the crystal graphs and 1 minute to evaluate the neural network) on a modern laptop computer to predict the band gaps of all 14,482 MOFs in the QMOF-14482 set using the CGCNN model. In stark contrast, it took over 1.5 million hours (~170 years) of computing time on the Stampede2 supercomputer[445,446] to carry out the structure relaxations and compute the band gaps via DFT.

The learning curves for each of the six models are shown in Figure 7.4A, highlighting the testing set MAE as a function of the training set size. Of all the individual models, CGCNN has the lowest MAE regardless of training set size. While SOAP has a worse testing set MAE compared to the simpler stoichiometric models when trained on fewer than 1000 MOFs, SOAP has a significantly higher learning rate such that it performs much better for larger training set sizes (although still underperforms compared to CGCNN). Reassuringly, the MAEs of the top-performing CGCNN and SOAP methods have not plateaued with respect to the training set size over the range of values considered in this work (i.e. up to ~10^4 training points). This indicates that both CGCNN and SOAP are capable of encoding the MOF crystal structures with sufficient uniqueness between structures and that the performance of the ML algorithms could be further improved if a greater number of training examples were provided. The testing set parity plot for the CGCNN trained on 80% of the QMOF-14482 MOF dataset is shown in Figure 7.4B. As one

would expect based on the relatively low MAE and high R^2, the agreement with the DFT predictions is quite strong, and this is true across a wide range of band gap values.

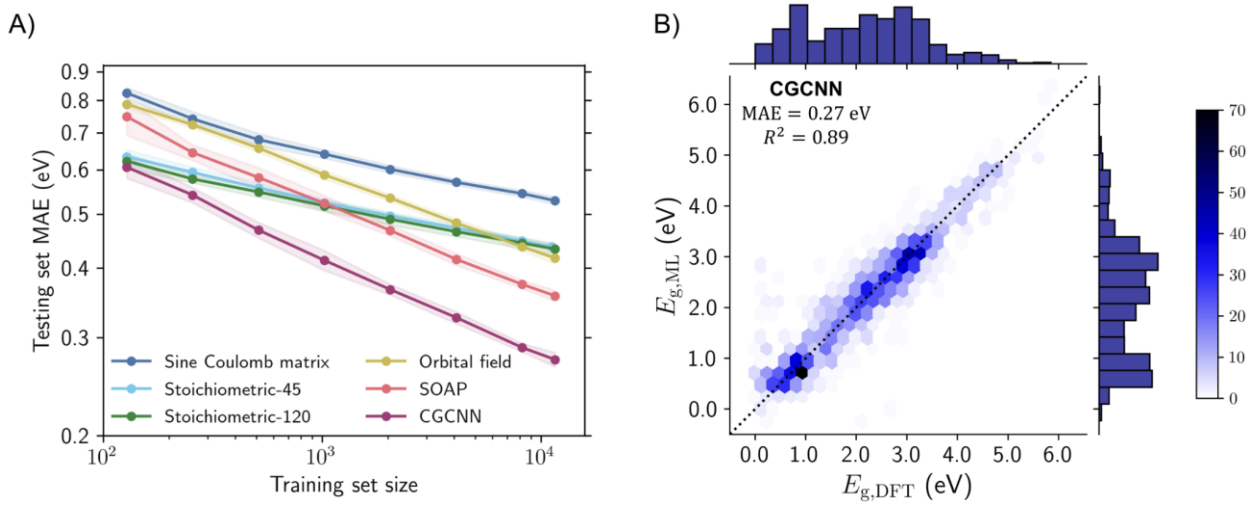


Figure 7.4. A) Mean absolute error (MAE) for band gap predictions on the testing set as a function of training set size for various machine learning methods. Each point represents the average MAE over five runs with different random seeds for data splitting, and the shaded region represents ±1 standard deviation. The data are shown on log–log axes. B) Testing set parity plot for the CGCNN model with hexagonal binning, comparing the machine learning band gaps, $E_{g,ML}$, to the PBE-D3(BJ) band gaps of the DFT-optimized structures, $E_{g,DFT}$. The color bar indicates the number of MOFs in each bin, and the line of parity is shown as a dashed line. Histograms summarizing the distribution of $E_{g,DFT}$ and $E_{g,ML}$ data are displayed parallel to the x- and y-axes, respectively.

7.2.3 Dimensionality Reduction for Structure–Property Analysis

While the kernel-based methods have a higher MAE than CGCNN when predicting MOF band gaps, the underlying descriptors can be readily used for dimensionality reduction – an unsupervised learning task that can cluster structurally similar MOFs in feature space for the purposes of identifying interpretable structure–property relationships. Using the uniform manifold approximation and projection (UMAP) algorithm to carry out the dimensionality reduction,[447,448] the distance between each MOF in the reduced space can be related to the distance in feature space, such that clusters of points tend to have similar structures (Equation F.10). By overlaying the DFT-

computed band gaps over the UMAP, regions of low and high band gap can emerge, making it possible to identify otherwise subtle structure–property trends.

As an example, selecting several MOFs in region A of the SOAP-based UMAP (Figure 7.5) yields materials with long, linear alkane-based linkers (e.g. refcode: NEZMEM[449]), which consistently have high band gaps regardless of the coordinating metal. The low band gap MOFs are more scattered throughout the reduced feature space, but as one example, region B of Figure 7.5 contains framework materials with linkers consisting of various TCNQ (TCNQ = 7,7,8,8-tetracyano-quinodimethane) derivatives, with several of these materials previously shown to have high electrical conductivities (e.g. BISVUW[450], FAFJAZ[451]). The projection in Figure 7.5 can be used to find MOFs that are structurally similar to a given material of interest as well. For instance, $Cu[Ni(pdt)_2]\cdot C_2H_2$ (pdt^{2-} = 2,3-pyrazinedithiolate) (refcode: HIVPOU[452]) is in the QMOF-14482 dataset, and it is known to be one of the rare examples of a three-dimensional, porous framework that exhibits room temperature electrical conductivity.[452] Perhaps unsurprisingly, one of the closest points to $Cu[Ni(pdt)_2]\cdot C_2H_2$ is the isostructural framework $Cu[Cu(pdt)_2]\cdot C_2H_2$ (refcode: WIHQEM[453]) (region C), which has also been studied for its conductive properties.[454,455] In general, we find that the SOAP-based UMAP places greater emphasis on the similarity of the organic linkers rather than the metal identity, likely due to the averaging scheme used in the generation of the similarity kernel (Table F.6). Modifications to the SOAP encoding that better account for the discrete building block nature of MOFs, such as variations on the recently developed coarse-grained SOAP (cg-SOAP) method,[456] may yield improvements in the future.

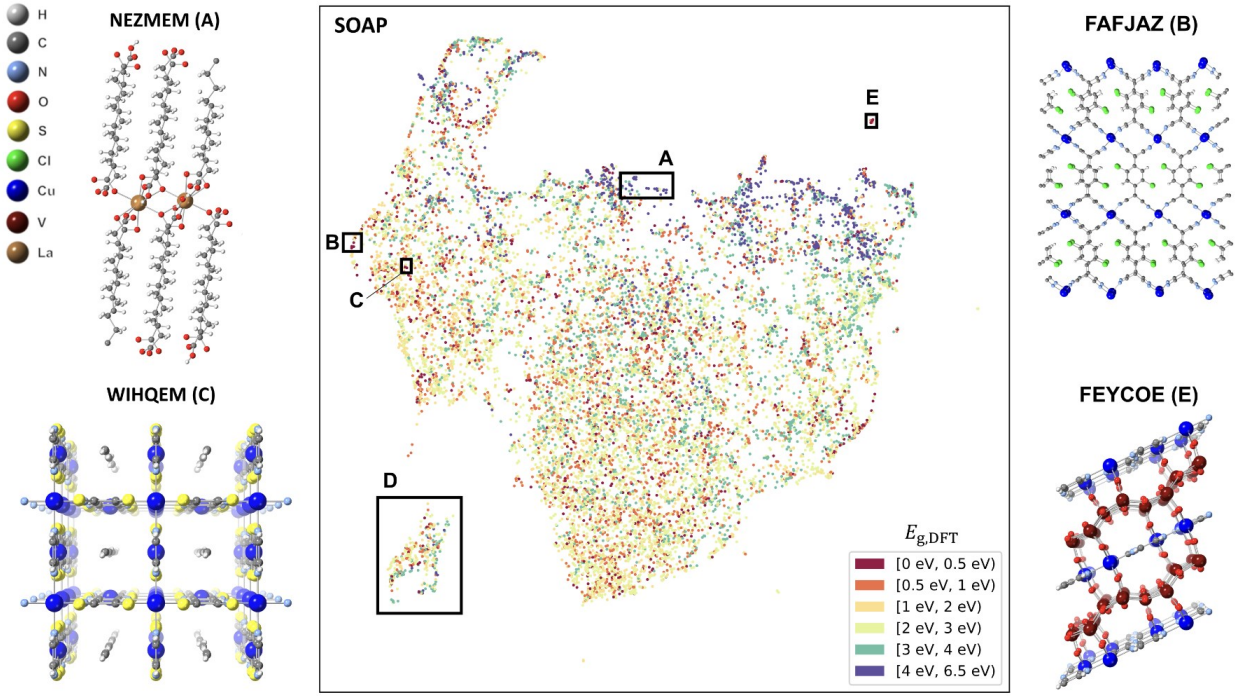


Figure 7.5. Unsupervised structural dimensionality reduction performed using UMAP, with a distance matrix obtained from the SOAP average similarity kernel of the unrelaxed structures in the QMOF-14482 dataset. The PBE-D3(BJ) band gaps of the DFT-optimized structures, $E_{g,DFT}$, are overlaid on the UMAP. Selected MOFs in the projection are highlighted.

Similar to what has been done in prior work with revised autocorrelation functions,[457] we can use the SOAP similarity kernel to understand the diversity of structures in the QMOF-14482 dataset and identify structural outliers. The most apparent example is the isolated cluster of points in region D of Figure 7.5. Investigation of these crystal structures indicates that they are predominantly frameworks with high fluorine content, such as MOFs with fluorinated linkers (e.g. MUQCEH[458], HADMOR[459]) or metal–fluoride species (e.g. EMEJAJ[460]), which leads to a large difference in the average SOAP fingerprint compared to most other MOFs in the dataset. The isolated region E of Figure 7.5 where there is a low band gap cluster contains polyoxovanadate-based MOFs, some of which have already been investigated for their conductive and electrocatalytic properties (e.g. FEYCOE[461], XEHYEP[462]).

While the SOAP-based UMAP is useful for identifying local trends in feature space, significantly greater clustering is observed when using the Stoichiometric-120 encoding. As is evident in Figure 7.6A, the UMAP based on the Stoichiometric-120 encoding largely partitions the MOF chemical space by the maximum atomic number in each chemical formula. The variations within a given cluster are due to more subtle differences in the elemental fractions and compositional features that compose the Stoichiometric-120 descriptor. Notably, the band gaps are well-separated between and within each cluster in the reduced space (Figure 7.6B). For these reasons, the Stoichiometric-120 UMAP is one useful way to obtain a global view of the QMOF Database. For instance, we find that the QMOF-14482 dataset closely overlaps with both the larger QMOF-42349 dataset it was drawn from and the separate CoRE MOF 2019 database[14] based on the reduced space of Stoichiometric-120 features (Figures F.8 and F.9). The band gap data in Figure 7.6B also emphasizes how the Zn-containing MOFs (east cluster, $\max(Z) = 30$) tend to have lower band gaps than MOFs with first-row transition metals (south-east cluster, $\max(Z) = 23 – 29$) at the PBE-D3(BJ) level of theory. To enable additional data exploration, interactive versions of the UMAPs are available in the supporting dataset.[402]

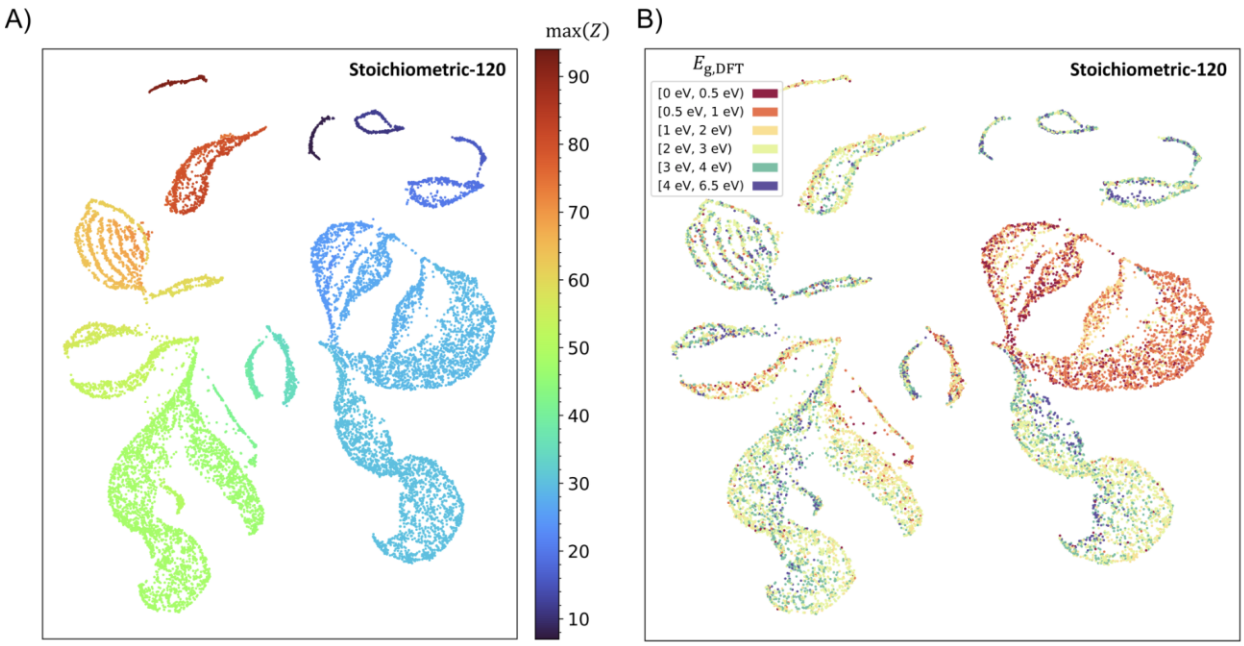


Figure 7.6. Unsupervised dimensionality reduction performed using UMAP, with a distance matrix obtained using a Euclidean distance metric of the Stoichiometric-120 encodings for the structures in the QMOF-14482 dataset. The a) maximum atomic number in each structure, $max(Z)$, and b) PBE-D3(BJ) band gaps for the corresponding DFT-optimized structures, $E_{g,DFT}$, are overlaid on the UMAPs.

7.2.4 Highlighting Notable Low Band Gap MOFs

We conclude by highlighting several framework materials identified in this work that have low band gaps, motivated in part by the search for a greater number of (semi)conducting MOFs. It should be noted that while the PBE-D3(BJ) level of theory makes it possible to generate a sufficiently large database for the purposes of ML model development and to identify structure–property relationships, it is known to underestimate band gaps like essentially all generalized gradient approximation functionals.[463,464] As such, we carried out full structure relaxations and corresponding band gap calculations using the hybrid-level HSE06-D3(BJ) functional[465–467] on select materials to generate more accurate band gap predictions. As a point of reference, materials with band gaps in excess of ~4 eV are often classified as electronic insulators, including many of the most commonly studied MOFs (e.g. MOF-5,[444] UiO-66 (UiO = Universitetet i Oslo),[468] ZIF-8

(ZIF = zeolitic imidazolate framework)[469]).[16,463] Generally, lower band gaps are necessary to support electrical conductivity (although it is not the sole factor required for achieving high electrical conductivities[16]).

When the CGCNN model is used to predict the band gaps of all 42,349 structures that compose the QMOF-42349 dataset, one of the lowest band gap material is predicted to be $Ag(DCl)_2$ (DCl = 2,5-Cl,Cl-*N*,*N*′-dicyanoquinone diamine) (refcode: OTARUX[470]), which is known from experiments to exhibit metallic character via organic radicals that connect the Ag(I) cations.[470] The introduction of radical or redox-active linking units is a well-established strategy to increase the electrical conductivity of framework materials.[16] Although $Ag(DCl)_2$ is arguably best described as a coordination polymer, one notable MOF in the QMOF-42349 dataset with a low predicted band gap and a radical-containing linker is (TTF)[$\{Rh_2(CH_3CO_2)_4\}_2$TCNQ] (TTF = tetrathiafulvalene) (refcode: WAQMEJ[471]) – a pillared layer framework material built from Rh(II) paddlewheels and a TTF–TCNQ charge-transfer salt (Figure 7.7A). The HSE06-D3(BJ) band gap for this material is found to be particularly small with a value of 0.71 eV, which can be directly attributed to a reduced conduction band minimum (CBM) from the TTF and TCNQ components (Figure 7.7D). Furthermore, the valence band maximum (VBM) also exhibits hybridization between the 4*d* orbitals of Rh and 2*p* orbitals of C and N atoms belonging to the radical TCNQ linker, which is important for applications involving electron transport. In contrast, one of the most electronically insulating structures in the QMOF-42349 dataset based on CGCNN-predicted band gap is the non-porous coordination polymer $Sr[C_2H_4(SO_3)_2]$ (refcode: GUTYAW[472]), which has an HSE06-D3(BJ) band gap of 8.36 eV (Figure F.12).

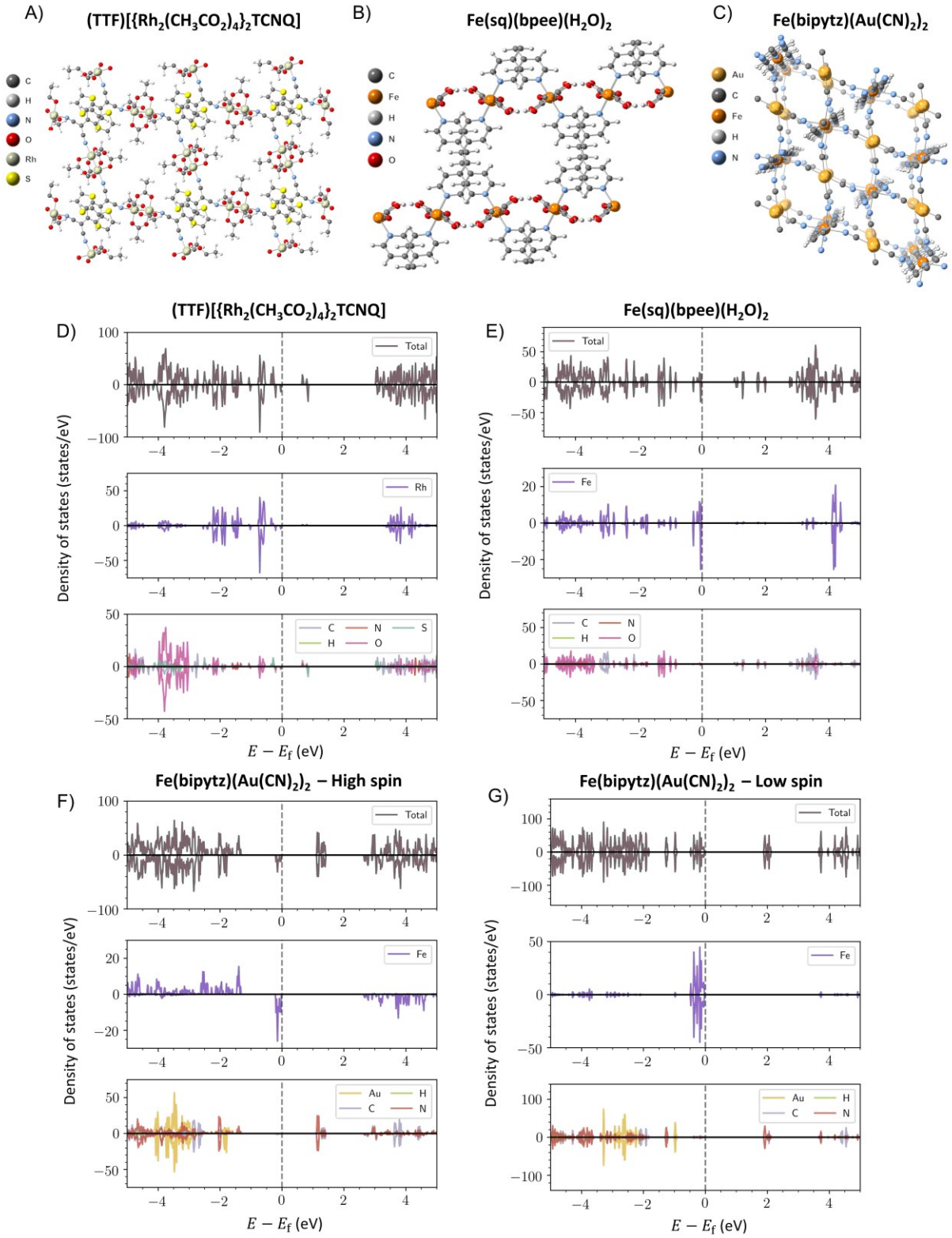


Figure 7.7. Structures of A) (TTF)[{$Rh_2(CH_3CO_2)_4$}$_2$TCNQ], B) Fe(sq)(bpee)(H_2O)$_2$, and C) Fe(bipytz)($Au(CN)_2$)$_2$. Total and projected density of states (DOS) at the HSE06-D3(BJ) level of

theory for D) (TTF)[{$Rh_2(CH_3CO_2)_4$}$_2$TCNQ], E) Fe(sq)(bpee)$(H_2O)_2$, F) Fe(bipytz)$(Au(CN)_2)_2$ (high spin), and G) Fe(bipytz)$(Au(CN)_2)_2$ (low spin). The energy, E, in eV is shown with respect to the Fermi level, E_f. For MOFs with magnetic character, DOS values above and below the horizontal zero line refer to the spin-up and spin-down channels, respectively.

Consistent with prior experimental work,[473] we also find several Fe-containing materials in the QMOF-42349 dataset with low band gaps, many of which have not yet been studied for their electronic properties. One representative example is Fe(sq)(bpee)$(H_2O)_2$ (bpee = 1,2-bis(4-pyridyl)ethylene; sq = squarate) (refcode: RAXNEK[474]), shown in Figure 7.7B, which has a band gap of 1.06 eV at the HSE06-D3(BJ) level of theory. The high-spin Fe(II) species in an octahedral crystal field with $t_{2g}^4e_g^2$ electron configuration dominate the VBM in this material, whereas the bpee linker (as opposed to the bridging sq species or inorganic node) make up the conduction band edge (Figure 7.7E).

Another noteworthy example is the three-dimensional porous framework material Fe(bipytz)$(Au(CN)_2)_2$ (bipytz = 3,6-bis(4-pyridyl)-1,2,4,5-tetrazine) (refcode: LOJLAZ[475]), shown in Figure 7.7C. At the HSE06-D3(BJ) level of theory, we find that the high spin state exhibits a band gap of 1.17 eV (Figure 7.7F) – similar to that of Fe(sq)(bpee)$(H_2O)_2$. The projected density of states indicates that the Au(I) species are unrelated to the relatively low band gap; instead, the low band gap can be attributed to the combination of Fe(II) and bipytz linker. Fe(bipytz)$(Au(CN)_2)_2$ is known to be a spin-crossover framework (with a sharp spin transition around 290 K),[475] and we find the low spin HSE06-D3(BJ) band gap to be 1.95 eV (Figure 7.7G), suggesting that the material may have tunable electronic properties as a function of temperature. For the low spin case, the VBM is composed of Fe 3*d* orbitals and the CBM is composed of N 2*p* orbitals. The reduction in band gap from low spin to high spin state can be rationalized on the basis

of crystal field theory. In the high spin state, the Fe(II) centers have a $t_{2g}^4e_g^2$ electronic configuration, whereas in the low spin state they have a $t_{2g}^6e_g^0$ electron configuration. This occupation of the e_g orbitals in the high spin state is directly related to the predicted ~0.8 eV reduction in the band gap compared to the low spin state. For both highlighted Fe-containing frameworks, the band gaps are lower – or comparable in the low spin state for Fe(bipytz)$(Au(CN)_2)_2$ – to those of several iron-containing MOFs that have been studied for their conductive properties, such as Fe_2(dobdc), Fe_2(dsbdc) (H_4dsbdc = 2,5-disulfhydrylbenzene-1,4-dicarboxylic acid), and Fe(bpz).[473,476] Collectively, these findings demonstrate the practical utility of the QMOF Database for identifying MOFs with targeted quantum-chemical properties.

7.3 Conclusion

In this work, we have developed a database of quantum-chemical properties for over 14,000 MOF structures (the "QMOF Database")[402] via a high-throughput periodic DFT workflow.[32] DFT-computed geometries, energies, band gaps, density of states, partial charges, spin densities, bond orders, and related electronic structure properties are made publicly available.[402] We highlight how this database can be used to identify MOFs with targeted electronic structure properties and then develop several ML models to predict the DFT-computed band gaps using descriptors derived from the un-optimized MOF crystal structures. A crystal graph convolutional neural network (CGCNN)[406] is found to achieve high predictive performance for this task, making it possible to circumvent large numbers of computationally expensive DFT calculations in future studies. While not as accurate as CGCNN for regression purposes, we show that both the smooth overlap of atomic positions (SOAP)[439,440] and composition-based features[436] can be used to discover otherwise subtle structure–property relationships in the QMOF Database via unsupervised

dimensionality reduction techniques. Finally, we show how top-performing ML models generated from the database of DFT-computed properties can be used to aid in the discovery of MOFs with desired quantum-chemical properties – in this case, discovering MOFs with low band gaps that could be suitable candidates to consider further for applications where electrical conductivity is necessary.

Importantly, the QMOF Database now makes it possible to pursue several important research directions that are reliant on a large database of quantum-chemical properties for MOFs beyond those directly discussed in this work. For instance, with the success of transfer learning,[442,477] multi-task learning,[478] and Δ-ML[479] methods in materials research, the QMOF Database can serve as a valuable resource to increase the accuracy – and reduce the required training set size – for ML models tasked with the prediction of new MOF properties not present in the QMOF Database. Since the output of any ML models will depend on the chosen density functional approximation, related transfer learning approaches may also prove useful in generalizing ML model predictions to other levels of theory using the PBE-D3(BJ) data as a starting point. Instead of relying on representation approaches that were originally designed for inorganic solids or small molecules, the QMOF Database can also be used to develop better methods for the encoding of MOF structures in ML models. Even outside the areas of high-throughput DFT screening, data mining, and ML, there are countless possible use-cases for the QMOF Database. As just one example, the DFT-generated properties in the QMOF Database could be used to develop and/or benchmark (semi-)empirical methods (e.g. tight binding approaches[480] or molecular mechanics force fields[481]) with the hopes of achieving high accuracies for MOF structures.

We conclude by noting that the QMOF Database should be considered a living resource; several updates to the QMOF Database are planned for the future, and we welcome the development of subsets, modifications, and supplements to the database that suit the diverse needs of the MOF community. With all this in mind, we anticipate that the QMOF Database will accelerate the material design and discovery process while being specifically tailored for the chemical space of experimentally realized MOF structures.

Chapter 8: Probing the Limits of Density Functional Theory in Predicting the Electronic Structure Properties of Metal–Organic Frameworks: A Big Data Approach

In this chapter, we expand upon the QMOF Database to include computed electronic structure properties at levels of theory beyond the widely used generalized gradient approximation (GGA) functional PBE. With this new data, we obtain a better understanding of the limitations of commonly used GGA functionals and how different density functional approximations influence the prediction of computed MOF electronic structure properties, with a specific focus on band gaps and partial atomic charges. By changing the fraction of Hartree-Fock exchange, we find that there are both quantitative and qualitative differences in the computed band gaps across the different levels of theory. Furthermore, the presence of magnetic character results in unusually large differences in band gaps between the (meta-)GGA functionals and screened hybrid functionals. With regards to partial charges, we find that both the tested meta-GGA and screened hybrid functionals shift electron density away from the metal center and onto the ligand atoms with respect to the PBE reference. Nonetheless, the magnitude of these changes in partial charge is found to be much smaller than the differences observed when using an alternate charge partitioning scheme.

8.1 Introduction

Metal–organic frameworks (MOFs) have been widely studied over the last two decades due to their high degree of synthetic tunability, which makes it possible to tailor their physical and chemical properties for a given application of interest.[1] While much attention has been focused on the use of MOFs for industrial gas storage and separations, the design of MOFs with targeted electronic properties has become a topic of recent interest as well.[16,482] Through a judicious selection of inorganic node and organic linker, MOFs have been proposed for novel (opto)electronic devices, electrocatalysts, photocatalysts, sensors, and energy storage devices, among many other applications.[16] However, with tens of thousands of MOFs that have been experimentally synthesized[44] and virtually unlimited more that can be proposed,[53] it is often difficult to identify promising MOF candidates with the ideal set of electronic properties.

The advent of machine learning (ML) and related big data approaches have made it possible to more efficiently search through MOF chemical space, and high-throughput computational screening in general can often provide insight into previously unknown structure–function relationships.[46–51,396,483] In recent work, a high-throughput density functional theory (DFT) workflow[32] was used to construct a database of quantum-chemical properties for thousands of MOFs (and coordination polymers), known as the QMOF Database.[402] Like many databases of material properties generated from high-throughput periodic DFT calculations,[30,33,35,37,484] the electronic structure properties within the QMOF Database were computed with the relatively inexpensive PBE[125] exchange-correlation functional, including dispersion corrections. While PBE is particularly useful for generating the large quantities of material property data at a relatively modest computational cost, the electron self-interaction error of generalized gradient

approximation (GGA) functionals like PBE is known to greatly influence the predicted electronic properties.[367] Most infamously, PBE is known to severely underpredict band gaps,[464] and the degree to which there may be qualitative (as opposed to merely quantitative) errors is not well-established. This inherently limits the practical utility of high-throughput computational screening approaches for the accelerated design and discovery of MOFs with desired electronic structure properties.

For relatively simple inorganic compounds, several approaches have been taken in prior studies to increase the accuracy of ML-predicted band gaps trained on high-throughput DFT calculations in a computationally tractable manner. Arguably, the most intuitive route is to train ML models on experimental band gap data[485] or an ensemble of both theoretical and experimental band gap data.[486] Unfortunately, this a challenging task for MOFs, as there remain relatively few reports of experimentally measured MOF band gaps, and reported values can vary by several tenths of an eV depending on the synthesis conditions and crystallinity of the material.[16] A separate approach is to carry out higher accuracy DFT calculations on a subset of materials with the hope of training a ML model that can make more actionable predictions. Recently, databases of computed band gaps at higher levels of theory have been published for inorganic compounds with this goal in mind,[487,488] although no such resource currently exists for MOFs.

In the present work, we complement the existing dataset of PBE-quality electronic structure properties in the QMOF Database with analogous data computed using the HLE17[489] (meta-GGA) and HSE06[465,466] (screened hybrid) functionals. By analyzing the electronic structure properties calculated at these levels of theory, we uncover severe theoretical limitations associated with the more computationally efficient (meta-)GGA density functionals that prevent them from achieving

quantitatively – and sometimes qualitatively – accurate band gap predictions for MOFs and related structures with respect to hybrid functionals. Since it is well-established that different density functional approximations (DFAs) can alter the underlying charge density, we also investigated trends related to the computed partial atomic charges. In general, we find that the different levels of theory predict similar partial charge values; however, the meta-GGA and screened hybrids tend to shift electron density away from the metal centers and onto the ligand environments compared to the PBE point of reference. We anticipate that the computational data and trends presented in this work will make it possible to achieve both rapid and highly accurate predictions of MOF band gaps that can greatly accelerate the materials design and discovery process.

8.2 Methods

All plane-wave, periodic DFT calculations were carried out using the Vienna *ab initio* Simulation Package (VASP)[122,123] 5.4.4 and the Atomic Simulation Environment (ASE)[83] 3.20.0b1. DFT-optimized structures at the PBE-D3(BJ)[125–127] level of theory were obtained from the QMOF Database.[75] With these structures, static (i.e. single-point) HLE17[489] and HSE06[465,466] calculations were performed. HSE06 calculations were carried out with 10% Hartree-Fock (HF) exchange in addition to the default 25% HF exchange. For brevity, we will refer to the PBE-D3(BJ), HLE17//PBE-D3(BJ), HSE06 (10% HF exchange)//PBE-D3(BJ), and HSE06 (25% HF exchange)//PBE-D3(BJ) calculations as PBE, HLE17, HSE06*, and HSE06, respectively.

HLE17 was selected because prior benchmarking studies[463,489] suggest that it can greatly improve the prediction of semiconductor band gaps without the need for computationally expensive hybrid functionals or system-specific parameters, as is the case with the commonly employed GGA+*U* approach.[373] HSE06 was selected because it is currently considered the gold-

standard for predicting the electronic properties of MOFs with DFT,[4,463] although it can be computationally prohibitive for many larger MOF structures. For HSE06, we consider the standard 25% HF exchange as well as 10% HF exchange since, for some materials, 25% HF exchange can overcorrect the band gap underprediction problem of PBE.[490,491]

For materials without magnetic character, the band gap is defined as the energy difference between the conduction band minimum (CBM) and valence band maximum (VBM). For materials with magnetic character, there are potentially multiple possible band gaps that can be considered.[492] Unless otherwise stated, we define the band gap for spin-polarized systems as $\min(\mathrm{CBM}_{\uparrow}, \mathrm{CBM}_{\downarrow}) - \max(\mathrm{VBM}_{\uparrow}, \mathrm{VBM}_{\downarrow})$, where $\uparrow$ and $\downarrow$ refer to the spin-up and spin-down channels, respectively. Nonetheless, we note that the aforementioned definition can occasionally result in a band gap that is associated with a formally spin-forbidden transition, as depicted in Figure G.4. Using the band gap defined by $\min(\mathrm{CBM}_{\uparrow} - \mathrm{VBM}_{\uparrow}, \mathrm{CBM}_{\downarrow} - \mathrm{VBM}_{\downarrow})$ avoids the possibility of a spin-forbidden process. Regardless of which band gap definition is employed, the trends and conclusions reported throughout this work remain unchanged (Figure G.5).

Additional methodological details regarding the DFT calculations and dataset curation can be found in the Appendix.

8.3 Results and Discussion

8.3.1 Band Gap Comparison

In order to develop ML-guided models that can directly guide future experimental efforts, it is essential to understand the behavior and potential limitations of various levels of theory when predicting MOF electronic structure properties. As such, we begin by comparing the DFT-predicted band gaps for 11,122 structures in the QMOF Database with the PBE (GGA: 0% HF

exchange), HLE17 (meta-GGA: 0% HF exchange), HSE06* (screened hybrid: 10% HF exchange), and HSE06 (screened hybrid: 25% HF exchange) functionals.

As shown in Figure 8.1, we observe pronounced differences amongst the various DFAs. Starting with the box plots, we find that of the four levels of theory tested in this work, PBE predicts the lowest band gaps in general. Including HF exchange — as with HSE06* and HSE06 — tends to increase the predicted band gap values, with the relative increase depending on the fraction of HF exchange in the selected functional. Qualitatively, the HSE06* and HSE06 results are more reflective of prior experimental studies,[16] which suggest that the vast majority of MOFs are electronically insulating and that relatively few exhibit semi-conducting or metallic character. Switching focus to the HLE17 meta-GGA, we find that the median band gap value is within 0.10 eV of the HSE06* calculations, suggesting that the parameterization of this functional can partially improve upon the band gap underprediction problem of PBE despite not incorporating HF exchange.

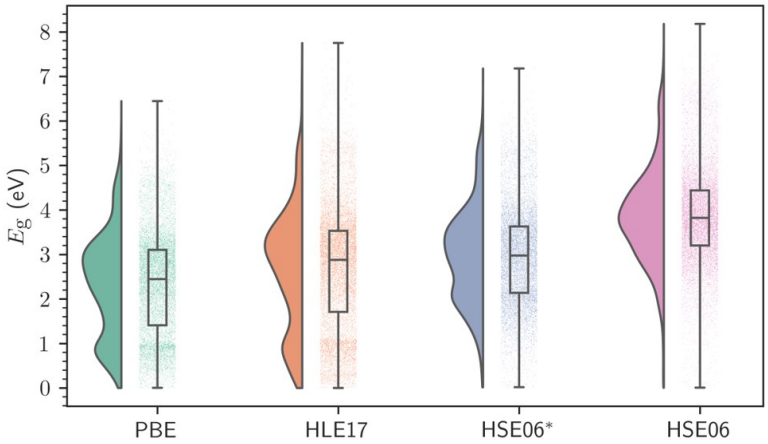


Figure 8.1. Raincloud plots (i.e. combined violin plot, box plot, and strip plot) for the DFT-computed band gaps, E_g, of 11,122 structures in the QMOF Database at the PBE, HLE17, HSE06*, and HSE06 levels of theory. The strip plots show all the data at that level of theory (jittered horizontally for ease-of-visualization). The box plots show the extrema (whisker tails),

interquartile range (box boundaries), and median (horizontal line). The violin plots show the probability density of the data.

The aforementioned summary statistics, however, only tell part of the story. When comparing the violin plots in Figure 8.1, it is immediately clear that the shape of the band gap distribution can vary significantly depending on the DFA. As noted in prior work on the QMOF Database,[75] the PBE-computed band gap data exhibits two separate distributions with peaks around 0.90 eV and 2.93 eV (Figure 8.1). A qualitatively similar distribution of band gaps is obtained with the HLE17 functional but not for the HSE06* and HSE06 functionals. In fact, at the HSE06 level of theory, the data is reasonably well-described by a single Gaussian distribution with a peak around 3.81 eV (Figure 8.1).

The two distinct distributions in the PBE band gap data can be attributed to closed-shell and open-shell materials, with the latter exhibiting lower band gaps on average (Figure 8.2a). When including 10% HF exchange with HSE06*, the degree of overlap between the closed-shell and open-shell band gap distributions is partway between that of PBE and HSE06 (Figure 8.2a). Taking the hybrid-quality calculations as the more accurate reference point,[463] these findings suggest that the PBE functional exhibits severe quantitative and qualitative shortcomings when applied to a wide range of MOF structures. Although HLE17 increases the median band gap of the dataset, it does not drastically change the bimodal nature of the band gap distribution found with PBE. Instead, HLE17 primarily increases the band gaps of the closed-shell frameworks (Figure 8.2a). When directly comparing the median predicted band gaps for the PBE, HSE06*, and HSE06 calculations, we find that there is a linear correlation with the fraction of HF exchange (Figure 8.2b) within the range of 0 – 25% HF exchange considered in this work. Assuming linear behavior in this region, it can be concluded that the median band gap across the dataset changes by ~0.05

eV per percent of HF exchange for the closed-shell frameworks and ~0.10 eV per percent of HF exchange for the open-shell frameworks.

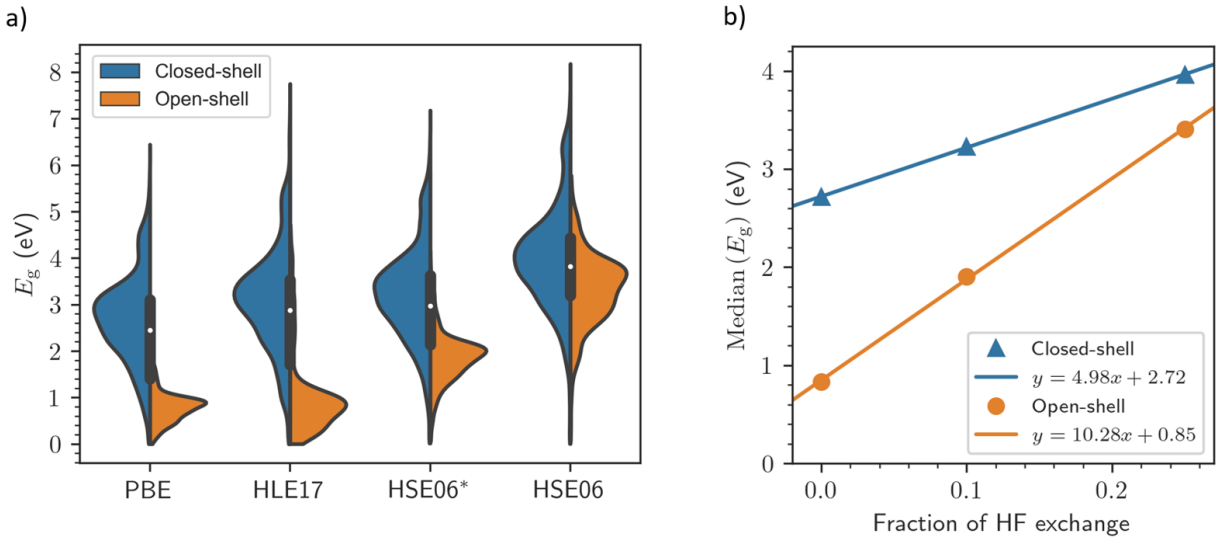


Figure 8.2. a) Violin plots of the predicted band gaps, E_g, for 11,122 structures in the QMOF Database at the PBE, HLE17, HSE06*, and HSE06 levels of theory. The left and right sides of each violin plot include structures with closed-shell (8970 structures) and open-shell (2152 structures) character, respectively. A box plot is included inside each violin, highlighting the extrema (whisker edges), interquartile range (box boundaries), and median (white dot) of the band gap data at the specified level of theory. b) Median band gap as a function of the fraction of Hartree-Fock (HF) exchange where 0% = PBE, 10% = HSE06*, and 25% = HSE06. The blue triangles and orange circles are the median band gaps for the closed-shell and open-shell structures, respectively. The solid lines display the linear best-fit equations.

While Figure 8.1 and Figure 8.2 show how the entire dataset changes with different density functionals, it is also important to investigate the degree of correlation between the various levels of theory. As shown in the parity plots in Figure 8.3, nearly every MOF has a larger predicted band gap at the HSE06* (Figure 8.3b) and HSE06 (Figure 8.3c) levels of theory than with PBE. This is also the case for most of the closed-shell MOFs at the HLE17 level of theory, particularly when $E_{g,PBE}$ is above ~1 eV (Figure 8.3a). For the closed-shell frameworks (Figure G.6), there is a linear correlation between the computationally inexpensive PBE-quality band gaps and those calculated with the more accurate HSE06* and HSE06 functionals (as well as the HLE17 functional). As

shown in Figure G.6c, a simple linear equation of the form $1.09E_{\mathrm{g,PBE}} + 1.04$ eV can predict HSE06 band gaps with an R^2 value of 0.92, provided the frameworks do not exhibit magnetic character and are not classified as metallic. Similar linear equations can be obtained for HLE17 and HSE06* for the closed-shell structures (Figure G.6a and Figure G.6b). The correlation between PBE and the hybrid functionals is weaker for MOFs with magnetic character, hence the larger degree of scatter in the low $E_{\mathrm{g,PBE}}$ range of Figure 8.3b and Figure 8.3c.

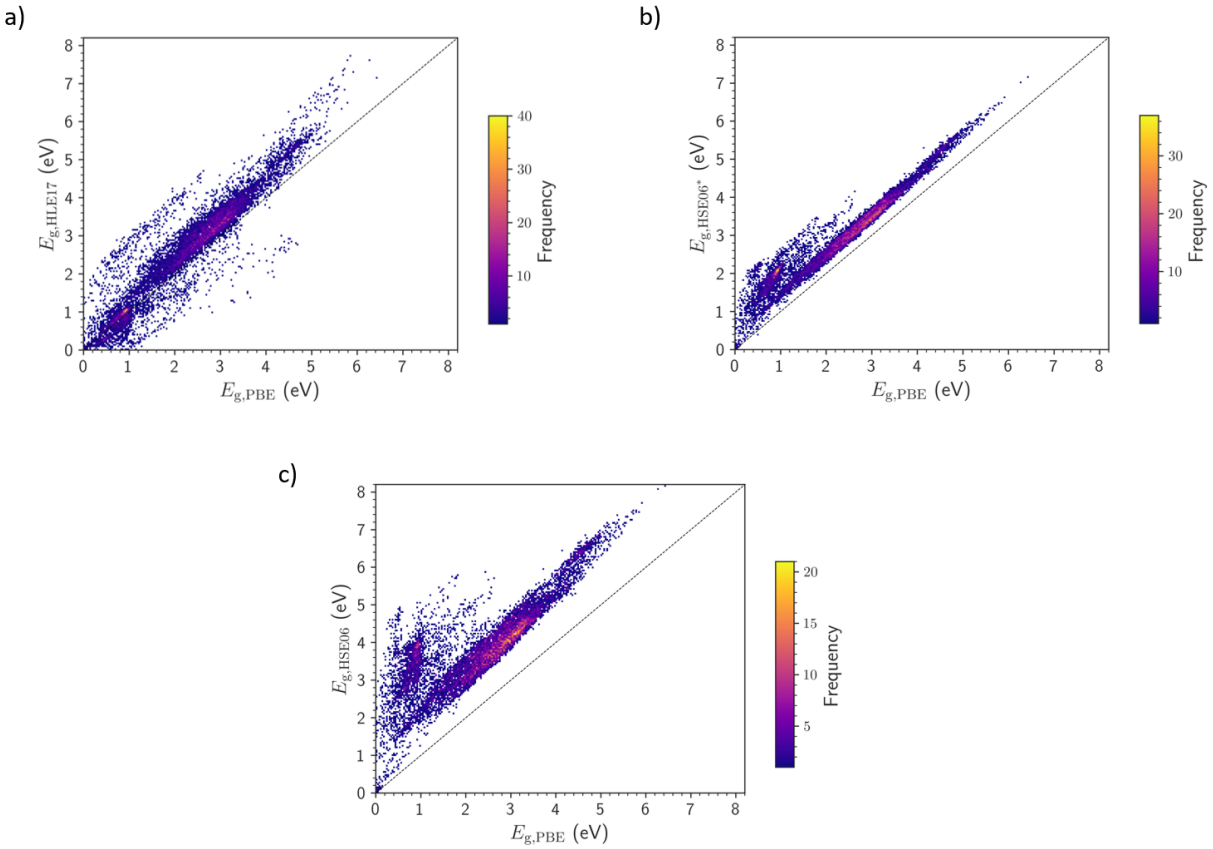


Figure 8.3. Parity plots of the computed band gaps, E_{g}, for 11,122 structures in the QMOF Database at various levels of theory. a) HLE17 vs. PBE; b) HSE06* vs. PBE; c) HSE06 vs. PBE. Given the large dataset size, the parity plots are shown as 2D histograms with the color bar reflecting the frequency of points in each bin. The line of parity is shown as a diagonal line.

8.3.2 Partial Charge Comparison

Beyond band gaps, it is well-established that different DFAs can change how the charge density is distributed in a given material.[493] This is particularly important because all the derivable quantum-mechanical properties are obtained from the underlying electron density. Furthermore, partial atomic charges (which are often computed directly from the underlying charge density) are widely used in molecular simulations of MOFs and can be used to interpret trends when modeling redox processes and chemical reactions.[234,494] With this in mind, we compared ~960,000 partial atomic charges calculated using the sixth-generation Density Derived Electrostatic and Chemical (DDEC6) partitioning scheme[158,230,231] at the PBE, HLE17, HSE06*, and HSE06 levels of theory.

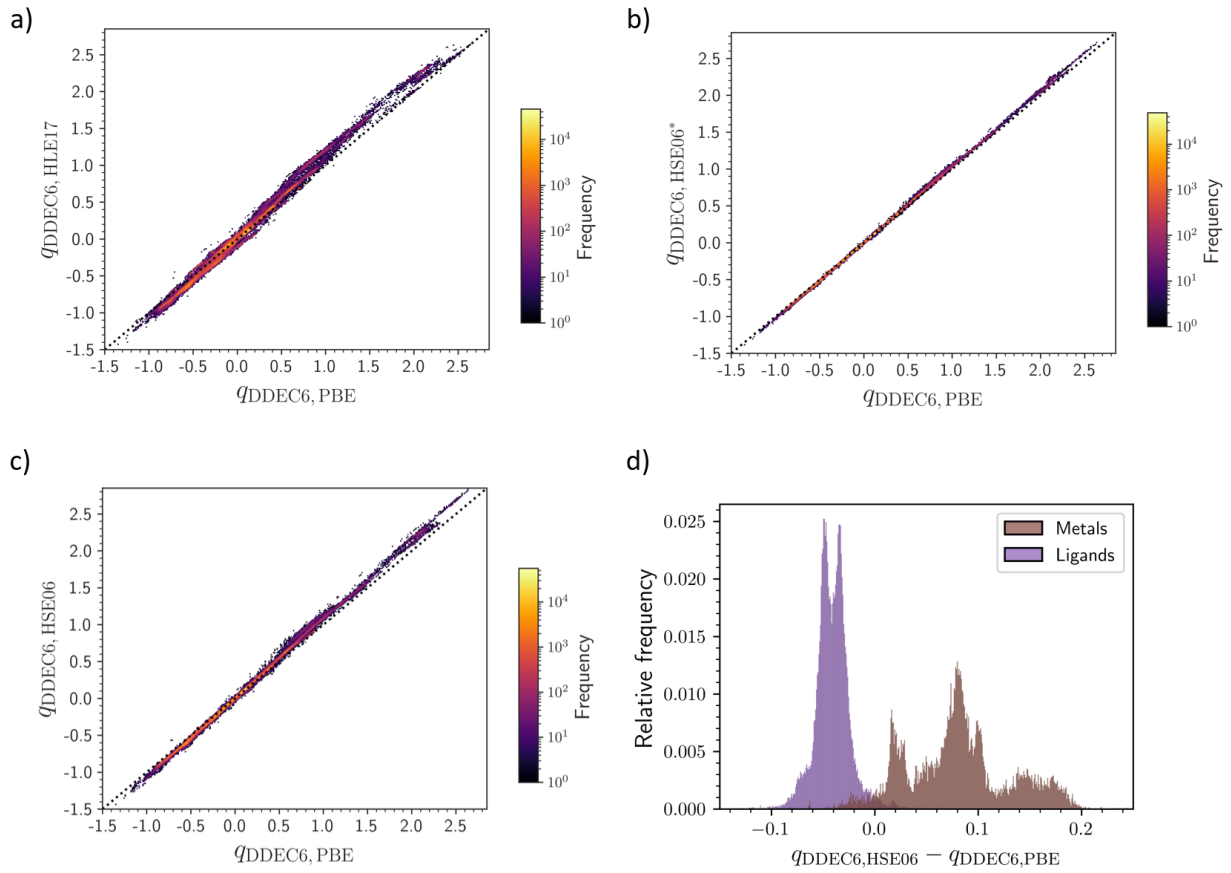


Figure 8.4. Parity plots of the DDEC6 partial atomic charges, q_{DDEC6}, for 958,465 atoms based on charge densities at various levels of theory: a) HLE17 vs. PBE; b) HSE06* vs. PBE; c) HSE06 vs.

PBE. Given the large dataset size, each parity plot is a 2D histogram with the logarithmic color bar reflecting the frequency of points in each bin. The line of parity is shown as a dashed diagonal line. d) A histogram of the change in DDEC6 charges between the PBE and HSE06 levels of theory for the (semi-)metals and ligand atoms within the first coordination sphere.

As shown in Figure 8.4a, the HLE17- and PBE-quality DDEC6 partial atomic charges are highly correlated across the entire dataset, with most points falling within 0.04 charge units from the line of parity. When investigating the computed partial charges at the HSE06* level of theory, we find that the values are closer to the PBE reference than for HLE17 (Figure 8.4b), suggesting that 10% HF exchange does not have a large role on the underlying charge density. However, when increasing the HF exchange to 25% with HSE06, a slightly larger difference can be observed (Figure 8.4c). By focusing solely on the (semi-)metal elements and the atoms within their first coordination spheres, we find that – compared to the PBE reference – there is often a loss of electron density at the metal and corresponding gain of electron on the surrounding ligands when using the HSE06 functional (Figure 8.4d). These trends are consistent with previous partial charge analyses carried out on transition metal complexes and open-framework solids.[376,380,495] Given the large partial charge dataset in the present work, we can conclude that this shifting of electron density occurs for an enormously diverse range of metal–ligand environments and can be taken as a rule-of-thumb in most cases. While there are differences in the partial atomic charges between the different levels of theory, they are relatively minor deviations in general. The overall strong agreement suggests that the less expensive PBE-quality charges are likely suitable when carrying out high-throughput computational screening studies of MOFs.

On a related note, it is worth comparing the effect of different charge partitioning schemes for a given level of theory. As shown in Figure 8.5, the differences between Bader,[159,496] DDEC6,[158,230,232] and Charge Model 5 (CM5)[289] partial atomic charges (at the PBE level of theory)

tend to be far larger than any differences observed when changing the DFA, similar to what was observed for several inorganic solids by Choudhuri and Truhlar.[497] For instance, large deviations are often observed for the S and P atoms of SO_4^{2-} and PO_4^{2-} groups, which have partial atomic charges upwards of ~2.4 charge units higher with the Bader method than the DDEC6 method (Figure G.9). For applications where a given charge partitioning scheme has not been thoroughly benchmarked, we therefore recommend comparing several types of partial charges and further investigating any substantial differences.

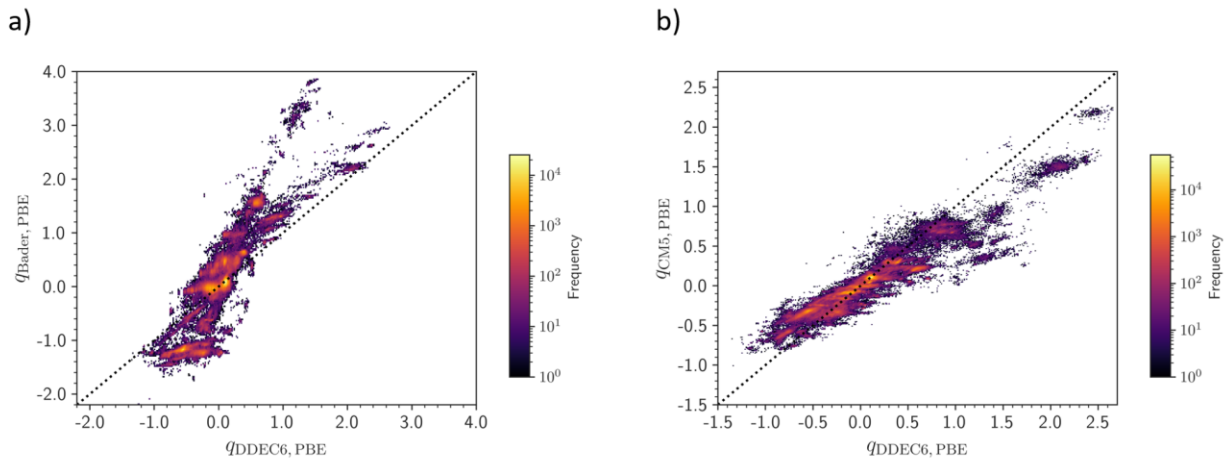


Figure 8.5. a) Parity plot of the partial atomic charges, q, for 882,682 atoms computed using the Bader and DDEC6 charge partitioning schemes at the PBE level of theory. b) Parity plot of the partial atomic charges, q, for 1,789,278 atoms computed using the CM5 and DDEC6 charge partitioning schemes at the PBE level of theory. Given the large dataset size, the parity plots are shown as 2D histograms with the logarithmic color bar reflecting the frequency of points in each bin. The line of parity is shown as a dashed diagonal line.

8.4 Future Work

This project is currently ongoing, and several additions are planned prior to its eventual publication. Currently, I am collaborating with members of the Persson group at the University of California, Berkeley and Lawrence Berkeley National Laboratory to host the QMOF Database – including the results at these new levels of theory – as a custom application on the Materials

Project.[33] This will help ensure that the QMOF Database is aligned with the findable, accessible, interoperable, and reusable (FAIR)[498] guiding principles of data-driven research. I am also currently collaborating with Dr. Victor Fung at Oak Ridge National Laboratory – who has developed the graph convolutional network (GCN) benchmarking program MatDeepLearn[499] – to construct ML band gap models tailored for this new HLE17, HSE06*, and HSE06 data (for ML-based partial charge prediction, we refer the reader to several ML models[397–399] that have been shown to accurately predict DDEC6 and CM5 charges for MOFs, an example of which is shown in Figure G.10). The current plan is to train multi-task[500] and/or multi-fidelity[501,502] GCN models on the new dataset of MOF band gaps. With the multi-task learning approach, the GCN can simultaneously predict PBE, HLE17, HSE06*, and HSE06 band gaps given a graph-based encoding of a MOF structure. With the multi-fidelity learning approach, the larger amounts of low-fidelity PBE data can potentially be used to increase the learning rate of a GCN trained on higher fidelity HLE17/HSE06*/HSE06 data, which has fewer data points. With these new ML models, I also plan to carry out a dimensionality reduction on the trained GCN embeddings to visualize what the GCN is learning and to aid in the identification of otherwise difficult to discern structure–property trends.

8.5 Conclusion

With a newly generated dataset of electronic structure properties for ~11,000 MOFs (and coordination polymer) in the QMOF Database,[75] we compare the performance of different density functional approximations for the prediction of band gaps and partial atomic charges. When comparing DFT-computed band gaps with the commonly used PBE functional against those that incorporate some fraction of Hartree-Fock exchange, we observe that PBE almost universally

results in a lower band gap prediction. This difference is largely systematic for MOFs without magnetic character and can be empirically corrected through a simple linear relationship for structures that are semi-conductors or insulators. For MOFs with some degree of magnetic character, an even larger – and slightly less predictable – disparity between band gap predictions is observed as a function of the fraction of Hartree-Fock exchange. With respect to the PBE results, the meta-GGA HLE17 is found to increase the band gaps of the MOFs without magnetic character such that they are roughly similar to values predicted using the HSE06 screened hybrid functional with 10% Hartree-Fock exchange. However, HLE17 does not significantly alter the behavior of the MOFs with magnetic character.

When investigating the partial atomic charges, which are reflective of the underlying charge density for a given density functional approximation, we find that there are slight systematic differences amongst the different levels of theory. For both the HLE17 meta-GGA and screened hybrid functionals, electron density localized on the (semi-)metals is lower than with PBE, and the opposite is true for the coordinating ligand atoms. Nonetheless, these changes in the partial atomic charges are relatively minor compared to the differences arising from using different charge partitioning schemes.

Chapter 9: Realizing the Data-Driven, Computational Discovery of Metal–Organic Framework Catalysts

Metal–organic frameworks (MOFs) have been widely investigated for challenging catalytic transformations due to their well-defined structures and high degree of synthetic tunability. These features, at least in principle, make MOFs ideally suited for a computational approach towards catalyst design and discovery. Nonetheless, the widespread use of data science and machine learning to accelerate the discovery of MOF catalysts has yet to be substantially realized. In this chapter, we provide an overview of recent work that sets the stage for future high-throughput computational screening and machine learning studies involving MOF catalysts. This is followed by a discussion of several challenges currently facing the broad adoption of data-centric approaches in MOF computational catalysis, and we share possible solutions that can help propel the field forward.

9.1 Introduction

Metal–organic frameworks (MOFs) are a class of porous solids composed of metal ions or clusters connected by organic ligands. Due to their high degree of crystallinity and synthetic tunability, MOFs have been widely studied as novel catalytic materials for a variety of chemical transformations.[8] The atomically precise structures of most MOFs is particularly advantageous for heterogeneous catalysis, as it makes it possible — at least in principle — to achieve fine-tuned control of the active site environment in ways that are typically reserved for molecular transition metal complexes.

Through different combinations of inorganic and organic building blocks, thousands of MOFs have been experimentally synthesized, and virtually unlimited more can be proposed. As such, the traditional hypothesis-driven approach to materials discovery is inherently limited in its ability to identify the optimal MOF for a given catalytic reaction. Recent advances in data science, high-throughput computational screening, and machine learning (ML) represent a complimentary route and can be used to rapidly identify promising MOFs from the vast combinatorial space of inorganic nodes, organic linkers, and topologies. Data-driven materials discovery, particularly based on high-throughput computational screening, has been used to discover top-performing MOFs for numerous gas storage and separation processes.[48] Despite these successes — and the increasingly widespread use of data science techniques in the field of heterogeneous catalysis[503] — there remain extraordinarily few large-scale, data-driven studies of MOF-based catalysts.

In this review, we first provide a brief overview of some recent studies that lay the groundwork for future big data and ML studies on MOF catalysts. We then discuss several challenges that must be overcome in order to fully enable a computational materials discovery

pipeline for MOF catalysts. Drawing inspiration from successful uses of ML and materials informatics in adjacent application areas, we also highlight multiple opportunities to address the challenges currently facing data-driven MOF catalysis. Several of these themes are displayed in Figure 9.1.

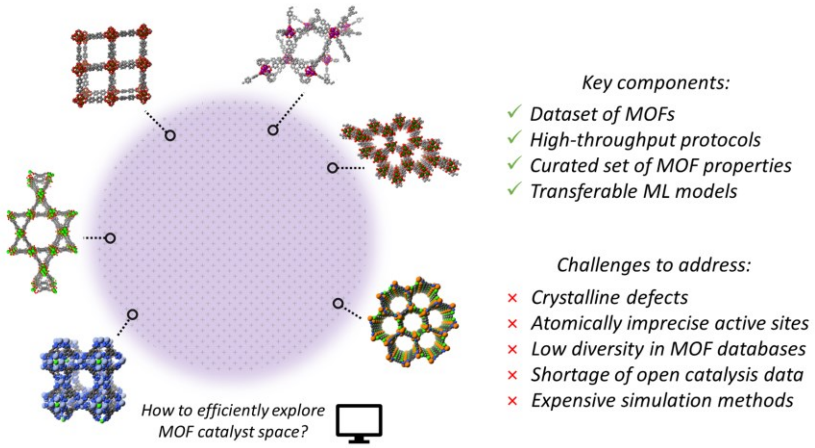


Figure 9.1. The enormous combinatorial space of inorganic nodes, organic linkers, and topologies results in vast array of plausible MOF structures to consider for a given catalytic transformation. Several of the key components for a successful data-driven catalyst discovery approach are listed alongside some of the challenges currently facing the MOF computational catalysis community.

9.2 Laying the Groundwork

In this section, we provide an overview of recent work involving high-throughput screening methods, MOF databases, and ML to predictively identify promising MOF catalysts and uncover important structure–property relationships. While high-throughput methods can refer to experiments or simulations, we will focus predominantly on the latter given the state of the published catalysis literature. High-throughput catalytic experiments with MOFs are uncommon at the time of writing, with an important exception of several prior studies by Cohen and coworkers[504,505] who used high-throughput assays to identify promising MOF catalysts for nerve agent degradation. For a detailed review on high-throughput synthesis, characterization, and experimental testing of MOFs, we refer the reader to an article by Clayson and coworkers.[506]

9.2.1 Small-to-Moderate-Scale Computational Screening

Truly large-scale, high-throughput screening studies remain rare for MOF catalysts; nonetheless, several small-to-moderate-scale computational screening studies have been used to identify promising MOFs for ethylene dimerization,[507] light alkane oxidation,[71,508,509] nerve agent hydrolysis,[510] CO_2 fixation to epoxides,[511] electrocatalytic CO_2 reduction,[512] oxygen electrochemistry,[424] and alcohol dehydrogenation,[513] among other reactions.[514] A common theme amongst these studies is the use of density functional theory (DFT) calculations to predict catalytically relevant reaction energies and activation energies across a curated set of MOF families. Brønsted–Evans–Polanyi relationships, linear scaling relationships, and various structure–property relationships are then developed to gain chemical insight into the predicted reactivity trends.[514] As one example, Sours and Patel et al.[424] used DFT calculations to study ~30 porphyrin-based MOFs with different transition metal (TM) cations for the oxygen reduction reaction. The authors found that the presence of an oxophilic spectator (in this case, a nearby TM–OH group) could stabilize adsorbed *OOH species via H-bonding interactions without substantially altering the corresponding *OH and *O binding energies, potentially improving the overall electrocatalytic performance.[514] Although the relatively small datasets in the aforementioned work preclude the use of modern data science techniques, these studies demonstrate the ability of computational methods to guide future experimental efforts.

9.2.2 Approaching High-Throughput Computational Screening

One of the first database-driven computational screening studies of MOF catalytic candidates was that of Vogiatzis et al.[116] who focused on the identification of MOFs that can activate N_2O for the purpose of selectively oxidizing light alkanes. Starting from a dataset of 5000 experimentally determined MOF crystal structures, the authors algorithmically identified MOFs with

coordinatively unsaturated Fe sites, which are often studied for oxidation catalysis. Quantum-mechanical simulations based on DFT were then carried out in a conventional low-throughput mode to investigate the catalytic properties for seven candidates of interest. Scaling up this approach further, Rosen et al.[32] developed a fully automated computational workflow that initializes the positions of small molecule adsorbates and then carries out the necessary structure relaxations with periodic DFT to systematically identify promising MOF catalysts. The authors then used this platform to screen a dataset of 60 MOF structures with a variety of inorganic nodes, organic linkers, and topologies to identify structure–reactivity relationships that can guide the discovery of MOFs for the catalytic activation of methane.[71]

Aside from thermal catalysis and electrocatalysis, MOFs have been widely investigated for their potential photocatalytic properties.[515] Arguably the most important properties of a MOF photocatalyst are its band gap and the energies of the corresponding band edges, which dictate the energy of light that can be absorbed and the redox properties, respectively. The energies of the valence band maximum and conduction band minimum must be aligned to a common reference, which can be done in an automated fashion by using the point in the MOF pore with the smallest variation in the electrostatic potential (i.e. the one closest to vacuum level).[516] With this protocol, Fumanal et al.[516] developed energy-based descriptors that allow for a high-throughput determination of the UV-vis absorption capabilities and photo-redox properties of a dataset of MOFs. In a separate study by Fumanal et al.,[517] the authors developed a computational workflow to predict the charge separation and charge mobility properties of MOFs in a high-throughput manner. These methodological advances pave the way for future computational screening studies of MOF photocatalysts.

9.2.3 Machine Learning

With regards to the use of ML for MOF catalyst discovery, there are currently few examples in the published literature. Arguably the closest study is that of Schweitzer and Archuleta et al.,[518] who constructed an ML model to determine the most important features dictating the DFT-computed binding energy of various atomic and small molecule adsorbates on simplified model systems of MOF-encapsulated nanoparticles. Li et al.[519] used experimentally reported turnover frequencies (TOFs) from 106 published experiments to train binary classification models that can predict whether a MOF might be active for carbon dioxide fixation; however, the low frequency of many of the one-hot encoded features (i.e. metals and linkers that appear only once in the dataset) and differences in how the TOFs are reported likely limits the predictive capabilities of the developed models. Although the number of ML studies specifically focused on MOF catalysis is limited, it should be noted that data-driven approaches in other application areas can likely be of significant value for catalyst discovery. For instance, ML models have been developed to predict the thermal, mechanical, and hydrolytic stability of MOFs from structural features,[394,520,521] the crystal morphology of MOFs from synthesis parameters,[522] and the geometry of bound adsorbates directly from X-ray absorption near-edge structure spectra,[523] all of which can potentially aid the design of promising MOF catalysts.

9.3 Challenges and Opportunities

Perhaps unsurprisingly, the primary factor currently holding back data-driven MOF catalysis is a lack of structured data from which to construct robust and generalizable models. While automated synthesis, characterization, and catalytic testing efforts have occasionally been undertaken,[506] the data (especially those involving “failed” experiments) are rarely made publicly available in a machine-readable format, and the total number of data points is often too small to build transferable

ML models. From a computational perspective, high-throughput DFT calculations represent a natural way to generate large amounts of catalytically relevant data. In this section, we highlight some of the most important challenges facing data-driven computational catalysis with MOFs and suggest several ways that they can be addressed in future work.

9.3.1 High-Throughput DFT Property Databases for MOF Catalysis

For catalytic reactions taking place on inorganic surfaces, there are multiple databases of DFT-computed adsorption energies (and occasionally, barrier heights) that can be used to develop data-driven models.[36,524] However, no such resource currently exists for MOFs. The closest analogue in the MOF field is that of the Quantum MOF (QMOF) Database,[75] which contains DFT-computed geometric and electronic structure properties for 18,000+ MOFs and coordination polymers in their as-synthesized states. With algorithms to automatically position adsorbates at the metal sites of MOFs,[32] one can envision constructing a database of DFT-computed adsorption energies, particularly if the guest species are small and unlikely to induce complicated changes in spin state. Until such a resource becomes available, ML models based on transfer learning[525] may be able to take advantage of the large amounts of data available for other material classes (e.g. inorganic surfaces, transition metal complexes) to accurately predict MOF catalytic properties with a comparatively small amount of MOF-specific training data.

9.3.2 Computational Screening of Atomically Imprecise MOFs

High-throughput computational screening of MOFs is also inherently tied to the quality of the underlying crystal structures being modeled. Unfortunately, due to the presence of crystallographic disorder as well as the difficulty in resolving hydrogen atoms and charge-balancing ions, databases of experimentally synthesized MOF structures can rarely be screened without correcting or

filtering out potentially problematic structures — a task that is far from trivial to automate. Fortunately, several approaches can be taken to increase the structural fidelity of a given MOF dataset. In addition to simple geometric heuristics that check for mis-bonded atoms and unlikely coordination environments, it may be possible to use one of several ML models that can rapidly predict partial atomic charges in MOFs[397–399] to identify potential anomalies in a MOF crystal structure database. Jablonka et al.[526] demonstrated a similar concept with their ML-guided oxidation state prediction model, which was trained on user-reported data in the Cambridge Structural Database and was accurate enough to identify MOFs with erroneous oxidation state assignments. If developed, an ML model that could predict bond orders would also be able to efficiently flag structures with erroneously over- or under-coordinated framework atoms, similar to what has been achieved with DFT-computed effective bond orders.[412]

Similarly, while MOFs are often touted as having atomically precise structures, many of the most promising MOF catalysts in the published literature have undergone post-synthetic modifications (e.g. metal- or linker-exchange, linker functionalization, atomic layer deposition) that introduce uncertainty in the active site structure.[429] Many MOFs are also known to exhibit defects in the crystal structure (e.g. missing nodes or linkers), which may greatly influence the catalytic properties of a given MOF but can be difficult to identify for the purposes of computational modeling. As one example, in the aforementioned screening study by Vogiatzis et al.,[116] Fe-BTT (BTT^{3-} = 1,3,5-benzenetristetrazolate) was computationally identified as a promising MOF for oxidation catalysis when using N_2O as an oxidant. Although the material was able to convert ethane to ethanol in experiments, spectroscopic measurements and catalyst cycling experiments led the authors to conclude that unidentified framework defects were likely

responsible for the catalytic activity.[116] For large-scale computational screening studies that are likely to emerge in the future, it may prove beneficial to deliberately introduce carefully selected defects in databases of otherwise "pristine" MOF structures to help ensure that the resulting data-driven models can accurately capture the complexities of experimentally synthesized MOF catalysts.

9.3.3 Ensuring a Diverse MOF Dataset

An alternate route to maximize the structural fidelity of a given MOF dataset is to use one of several hypothetical MOF construction tools[45,90,385,527] to computationally construct MOFs from a curated set of inorganic and organic building blocks. One cautionary aspect of this approach, however, is that databases of hypothetical MOFs can cover different regions of chemical space and/or lack the chemical diversity of experimentally synthesized MOF structures. Using an unsupervised dimensionality reduction technique, Moosavi et al.[457] demonstrated that hypothetical MOF databases often lack significant diversity with regards to the inorganic nodes, which are of crucial importance for heterogeneous catalysis. Many of the hypothetical MOF databases that have been published to date are also tailored to gas storage and separation processes, where variations in pore geometry and linker length are likely to play a larger role than in many catalytic reactions, particularly those involving small molecules. To address potential issues related to structural diversity, a judicious selection of building blocks must be chosen, preferably before any computational catalyst screening studies are even carried out. This may be made easier with dimensionality reduction approaches like those of Moosavi et al.[457] as well as larger databases of MOF building blocks, like that of Lee et al.[45] who amassed a dataset of over several hundred

inorganic and organic building blocks, most of which are derived from experimentally reported MOF crystal structures.

9.3.4 Automated Construction of Representative Cluster Models

When carrying out any theoretical study of MOF catalytic candidates, an important methodological decision is whether the system should be modeled as a crystalline lattice or as a truncated cluster model.[4] The former approach more closely resembles the true MOF structure, but the larger model system often limits the use of more accurate levels of theory. Carving a representative cluster model from the crystalline lattice takes advantage of the local nature of many catalytic reactions and results in a smaller model system that can be studied with more expensive computational methods. However, there are no automated approaches to determine the ideal size of the cluster model, where to terminate the cluster, how to cap the resulting dangling bonds, and which atoms should be kept rigid to appropriately mimic the bulk constraints of the crystalline framework. For high-throughput computational screening studies that involve a diverse range of MOF catalysts, the fully periodic approach is often the only tractable option. Recently developed MOF deconstruction codes that can split up a MOF into its constituent building blocks[45,79,457] represent a first step towards automating the construction of MOF cluster models for high-throughput computational catalysis.

9.3.5 Accelerated Identification of Ground State Magnetic Orderings

The automated identification of ground state spin states for MOFs where open-shell character is possible, as is the case for most MOFs with 3*d* transition metal cations,[4] represents yet another challenge for high-throughput computational MOF catalysis. The choice of spin state can drastically change the predicted reaction energies and barrier heights, and a separate DFT

calculation must be carried out for each plausible spin state to determine the lowest energy electron configuration. Even enumerating the different spin states can be a challenge in a high-throughput setting, particularly for MOFs that exhibit antiferromagnetism[4] or in cases where the metal oxidation state is not known *a priori*. Reassuringly, ML models may be able to ease the computational burden of screening MOFs with magnetic character. As one example, the aforementioned ML model by Jablonka et al.[526] can be used to accurately assign oxidation states for each metal center in a MOF, which can then help narrow down the list of possible spin multiplicities. There are also many recent studies to draw inspiration from the homogeneous catalysis literature, such as the ML models developed by Taylor et al.[528] to predict the ground state spin states of transition metal complexes. High-throughput workflows have been developed to accurately and efficiently identify the ground state magnetic ordering in inorganic solids, which may be extensible to MOFs as well.[529]

9.4 Outlook

Looking into the future when automated theoretical calculations and experiments are likely to be more widespread in the field of MOF catalysis, an integrated materials discovery platform augmented by machine learning and data science can be envisioned. This concept has been referred to as "digital reticular chemistry"[530] and has the potential to revolutionize the MOF design and discovery process, including for the accelerated design of MOF catalysts tailored for challenging chemical transformations. In addition to the continued development of high-throughput methods,[506] we stress that it is essential for openly accessible data repositories to be developed in the area of MOF catalysis for this digital reticular chemistry future to be realized. Ideally, these MOF property databases should have standardized data formats and rich metadata,[531] in addition

to abiding by the so-called FAIR guiding principles[498] (i.e. that the data is findable, accessible, interoperable, and reusable).

Collectively, the last several years have resulted in a gradual shift towards data-driven MOF catalysis, with several examples of small-to-moderate-scale screening studies published in the literature. Although a relative lack of carefully curated, catalytically relevant data currently limits the widespread adoption of ML in MOF catalysis research, we anticipate that this will rapidly change as recent advances in high-throughput screening approaches — both computational and experimental — become more widely adopted.

Chapter 10: CONCLUDING THOUGHTS

10.1 Summary

In this dissertation, density functional theory (DFT) calculations were used to explore the properties and reactivity of redox-active MOFs. In Chapter 2, a high-throughput periodic DFT workflow was developed to study large numbers of MOF catalysts in an automated fashion. This approach was then used to study experimentally realized MOFs with coordinatively unsaturated metal sites for the catalytic oxidation of methane via a terminal metal-oxo site. The main result of this screening effort was that blind, brute-force screening of a MOF database is unlikely to be fruitful when it comes to oxidation reactions. Most MOFs are redox-inactive and, therefore, are unlikely to have thermodynamically favored formation of a terminal metal-oxo species. Nonetheless, such sites – if they could be formed – are generally quite reactive towards methane activation. This finding was used to carry out a separate computational screening study in Chapter 3, now with a manually curated set of MOFs largely drawn from a set of structures that have been experimentally shown to exhibit some degree of redox activity. The main conclusion from this refined DFT screening effort was that there is an inverse, linear correlation between the metal-oxo formation energy and the barrier for C–H activation among the frameworks investigated in this work. Put another way, there is an inherent tradeoff between active site stability and reactivity that must be considered, which can serve as a guiding principle for future studies. In Chapter 4, a family of metal–triazolate frameworks with exchangeable metals and equatorial ligands was studied for methane activation via a terminal metal-oxo site. In this smaller scale DFT screening study, it was shown that the presence of ferromagnetic or antiferromagnetic coupling between the unpaired

electrons on the metal center and bound oxo/oxyl ligand can play a large role in the predicted barrier for methane activation.

Switching focus to the adsorption of O_2 and N_2 at the coordinatively unsaturated metal sites of MOFs, in Chapter 5 we took several of the MOF families originally studied in Chapter 3 and used DFT to determine which might be able to strongly bind O_2 and/or N_2. From this analysis, a series of design principles emerged. In general, few MOFs were found to strongly chemisorb N_2 with the exception of V(II)-containing frameworks. For O_2 chemisorption, earlier transition metals were found to have a greater affinity for O_2 binding (for a fixed oxidation state), as they are more likely to support electron transfer from the metal to the O_2 species. When comparing related MOF families with different ligand environments, it was shown that the more basic bridging ligands tended to increase the redox activity of the framework. In a particularly pronounced example of this phenomenon, DFT calculations at several levels of theory suggested that exchanging the bridging Cl^- anions of Co_2Cl_2(bbta) with OH^- anions to yield the isostructural framework $Co_2(OH)_2$(bbta) would greatly enhance the O_2 binding affinity at the Co(II) centers without significantly increasing the N_2 binding affinity. Experimentally obtained adsorption isotherms supported this theoretical finding. In Chapter 6, the problem of O_2 and N_2 adsorption was revisited, now with the goal of comparing various density functional approximations. As a result of this theory-focused analysis, it was found that different levels of theory can predict drastically different O_2 and N_2 adsorption energies, that the meta-GGA M06-L partially corrects for the overly exothermic adsorption energies of PBE, and that a GGA+*U* approach can substantially improve the accuracy of the predicted adsorption processes.

In Chapter 7, a high-throughput periodic DFT workflow was used to construct the QMOF Database, the first large-scale database of quantum-mechanical properties for MOFs. This database can be used as a platform to train and benchmark machine learning models to rapidly predict MOF electronic structure properties. As a proof-of-concept, a convolutional neural network was used to predict MOF band gaps at the PBE level of theory from a graph representation of the experimentally synthesized MOF crystal structure. In Chapter 8, the QMOF Database was updated to include computed properties using the HLE17 meta-GGA functional and the screened-hybrid HSE06 functional with 10% and 25% Hartree-Fock exchange. By comparing the predicted band gaps against those computed using the PBE GGA functional, it was found that the different density functional approximations exhibit quantitatively and qualitatively different predictions for the band gap. In general, the HLE17 functional increases the computed band gap compared to PBE but mainly for MOFs without magnetic character. For HSE06, the predicted band gaps are highly sensitive to the fraction of Hartree-Fock exchange, but MOFs with and without magnetic character tend to exhibit larger band gaps than the GGA and meta-GGA functionals.

In Chapter 9, an outlook on the use of data-driven methods to design and discovery MOF catalysts is provided, highlighting both the challenges that need to be overcome and the many opportunities that await. In summary, this dissertation represents the first steps towards using high-throughput DFT to study the catalytic and electronic properties of redox-active MOFs. Through this theory-guided approach, new trends were discovered that can guide the design of MOFs for oxidation catalysis, selective O_2 adsorption, and applications that benefit from an ability to tailor MOF electronic structure properties.

10.2 Recommendations for Future Research

10.2.1 C–H Bond Activation

There are several avenues for future research on MOFs that can catalytically oxidize strong C–H bonds. One of the more difficult but potentially most rewarding routes is to leverage the tunable nature of MOFs to mimic the proposed active sites of enzymes that are known to activate strong C–H bonds. For instance, both the particulate methane monooxygenase (pMMO) enzyme[532,533] and lytic polysaccharide monooxygenase (LPMO) enzymes,[534] which can oxidize the C–H bonds of methane and polysaccharides, respectively, share an active site motif known as the copper histidine brace.[534] This mononuclear copper species is bound to N-donor ligands, two of which come from a histidine residue and the remainder of which come from one or two separate imidazole groups.[534] A preliminary ConQuest[535] search of the Cambridge Structural Database MOF Subset[44] suggests that there is at least one family of MOFs with a histidine-bound Cu cation in an all-N ligand environment.[536] Furthermore, it may be possible to design new biologically inspired MOFs[537] with this functionality in mind. It should be noted, however, that for the MOF to be truly biomimetic, it is likely that functionality must be introduced "beyond the active site." In the enzymatic systems, it is known that there are nearby glutamate residues, which are likely involved in the pre-requisite O_2 activation process.[538–541] Fortunately, given the tunability of MOF structures and the ability to introduce post-synthetic functionalities, it may be possible to design a true biomimetic analogue of pMMO or LMPO.

On the topic of converting methane to methanol, one of the main challenges within the field as a whole is the ability to design a catalyst that is both active and selective towards methanol formation (rather than overoxidation products).[157,299] The main reason for this difficulty is that if an active site is reactive enough to cleave the C–H bond of methane, it is almost certainly capable

of cleaving the weaker C–H bonds of methanol and its derivatives. Several creative strategies have been proposed in the literature to increase methanol yields, including the use of a multi-functional material that has active sites capable of oxidizing methane and separate sites with a strong methanol adsorption potential that can "collect" the methanol product, which can then be released upon temperature cycling.[223] A MOF with multiple types of inorganic nodes may be able to achieve this goal. Regardless, using MOFs to address the selectivity problem in ways that are not possible with inorganic solids and zeolites would be an important advance for this area of research.

10.2.2 O_2 and N_2 Binding

When considering the design of MOFs for air separation, there are numerous research areas that have yet to be substantially investigated that would be worth considering in future work. For instance, while there have been multiple studies focused on understanding the thermodynamics of O_2 adsorption,[73,74,330] relatively little is known about the kinetics for this process, especially when changes in the spin multiplicity are involved. If MOFs are to be used as adsorbents for industrial-scale air separation, it will be essential to understand the phenomena governing both the O_2/N_2 selectivity and the rate of O_2 uptake. Separately, it would be important to know how strong is too strong when it comes to O_2 binding; at sufficiently exothermic O_2 adsorption energies, one would expect that the framework is likely to decompose. A combination of DFT calculations – like those used to model the decomposition of MOFs in the presence of water[542] – and *ab initio* molecular dynamics simulations may be able to provide answers to this question.

10.2.3 Quantum-Mechanical Property Databases

The QMOF Database presented in Chapter 7 represents just an initial glimpse of how high-throughput DFT and machine learning might guide the MOF design and discovery process. A

natural extension to the band gap prediction aspect of the QMOF Database would be to predict the electrical conductivity. This would require knowledge of the band structure, which is not currently part of the QMOF Database but could be readily generated from a separate static calculation carried out on the DFT-optimized geometries. If properly aligned,[516,543] the valence band maxima and conduction band minima in the QMOF Database could also be used to identify promising MOF photocatalysts, such as materials suitable for H_2 evolution and CO_2 reduction. The DFT-optimized structures in the QMOF Database can also be used to benchmark and potentially improve upon the accuracy of (semi-)empirical methods, such as Grimme's tight-binding code xTB[480,544,545] or the GFN-FF force-field.[481]

While certainly a large undertaking, an important and major update to the QMOF Database would be one that includes DFT-computed properties for MOFs with various adsorbates, such as CO, CO_2, CH_4, H_2, and H_2O. These example adsorbates were suggested because they are relatively small and do not typically involve a change in the oxidation state of the metal center, making high-throughput screening an easier task; furthermore, these adsorbates can all be readily added to MOF structures using the MOF Adsorbate Initializer[32] described in Chapter 2. With a large enough database of MOF–adsorbate systems, it may be possible to discover new structure–property relationships and develop machine learning models that can predict adsorption energies given a new set of MOFs or even a new set of related small molecule adsorbates.

REFERENCES

(1) Yaghi, O. M.; Kalmutzki, M. J.; Diercks, C. S. *Introduction to Reticular Chemistry: Metal-Organic Frameworks and Covalent Organic Frameworks*; John Wiley & Sons: Weinheim, Germany, 2019.

(2) Li, M.; Li, D.; O'Keeffe, M.; Yaghi, O. M. Topological Analysis of Metal--Organic Frameworks with Polytopic Linkers and/or Multiple Building Units and the Minimal Transitivity Principle. *Chem. Rev.* **2014**, *114* (2), 1343–1370.

(3) Kalmutzki, M. J.; Hanikel, N.; Yaghi, O. M. Secondary Building Units as the Turning Point in the Development of the Reticular Chemistry of MOFs. *Sci. Adv.* **2018**, *4* (10), eaat9180.

(4) Mancuso, J. L.; Mroz, A. M.; Le, K. N.; Hendon, C. H. Electronic Structure Modeling of Metal–Organic Frameworks. *Chem. Rev.* **2020**, *120* (16), 8641–8715.

(5) Li, H.; Wang, K.; Sun, Y.; Lollar, C. T.; Li, J.; Zhou, H.-C. Recent Advances in Gas Storage and Separation Using Metal--Organic Frameworks. *Mater. Today* **2018**, *21* (2), 108–121.

(6) Petit, C. Present and Future of MOF Research in the Field of Adsorption and Molecular Separation. *Curr. Opin. Chem. Eng.* **2018**, *20*, 132–142.

(7) Lee, J.; Farha, O. K.; Roberts, J.; Scheidt, K. A.; Nguyen, S. T.; Hupp, J. T. Metal−Organic Framework Materials as Catalysts. *Chem. Soc. Rev.* **2009**, *38* (5), 1450–1459.

(8) Konnerth, H.; Matsagar, B. M.; Chen, S. S.; Prechtl, M. H. G.; Shieh, F.-K.; Wu, K. C.-W. Metal-Organic Framework (MOF)-Derived Catalysts for Fine Chemical Production. *Coord. Chem. Rev.* **2020**, *416*, 213319.

(9) Zhang, X.; Chen, A.; Zhong, M.; Zhang, Z.; Zhang, X.; Zhou, Z.; Bu, X.-H. Metal--Organic Frameworks (MOFs) and MOF-Derived Materials for Energy Storage and Conversion. *Electrochem. Energy Rev.* **2019**, *2* (1), 29–104.

(10) Li, H.-Y.; Zhao, S.-N.; Zang, S.-Q.; Li, J. Functional Metal--Organic Frameworks as Effective Sensors of Gases and Volatile Compounds. *Chem. Soc. Rev.* **2020**, *49* (17), 6364–6401.

(11) Wang, L.; Zheng, M.; Xie, Z. Nanoscale Metal--Organic Frameworks for Drug Delivery: A Conventional Platform with New Promise. *J. Mater. Chem. B* **2018**, *6* (5), 707–717.

(12) Li, J.; Wang, H.; Yuan, X.; Zhang, J.; Chew, J. W. Metal-Organic Framework Membranes for Wastewater Treatment and Water Regeneration. *Coord. Chem. Rev.* **2020**, *404*, 213116.

(13) Rogge, S. M. J.; Bavykina, A.; Hajek, J.; Garcia, H.; Olivos-Suarez, A. I.; Sepúlveda-Escribano, A.; Vimont, A.; Clet, G.; Bazin, P.; Kapteijn, F.; Daturi, M.; Ramos-Fernandez, E. V.; Llabrés i Xamena, F. X.; Van Speybroeck, V.; Gascon, J. Metal--Organic and Covalent Organic Frameworks as Single-Site Catalysts. *Chem. Soc. Rev.* **2017**, *46* (11), 3134–3184.

(14) Chung, Y. G.; Haldoupis, E.; Bucior, B. J.; Haranczyk, M.; Lee, S.; Zhang, H.; Vogiatzis, K. D.; Milisavljevic, M.; Ling, S.; Camp, J. S.; Slater, B.; Siepmann, J. I.; Sholl, D. S.; Snurr, R. Q. Advances, Updates, and Analytics for the Computation-Ready, Experimental Metal−Organic Framework Database: CoRE MOF 2019. *J. Chem. Eng. Data* **2019**, *64* (12), 5985–5998.

(15) Calbo, J.; Golomb, M. J.; Walsh, A. Redox-Active Metal--Organic Frameworks for Energy Conversion and Storage. *J. Mater. Chem. A* **2019**, *7* (28), 16571–16597.

(16) Xie, L. S.; Skorupskii, G.; Dincă, M. Electrically Conductive Metal–Organic Frameworks.

Chem. Rev. **2020**, *120* (16), 8536–8580.
(17) D'Alessandro, D. M. Exploiting Redox Activity in Metal–Organic Frameworks: Concepts, Trends and Perspectives. *Chem. Commun.* **2016**, *52* (58), 8957–8971.
(18) Sun, L.; Campbell, M. G.; Dincă, M. Electrically Conductive Porous Metal–Organic Frameworks. *Angew. Chemie Int. Ed.* **2016**, *55* (11), 3566–3579.
(19) Zhang, Y.; Riduan, S. N.; Wang, J. Redox Active Metal– and Covalent Organic Frameworks for Energy Storage: Balancing Porosity and Electrical Conductivity. *Chem. Eur. J.* **2017**, *23* (65), 16419–16431.
(20) Mason, J. A.; Darago, L. E.; Lukens Jr, W. W.; Long, J. R. Synthesis and O_2 Reactivity of a Titanium(III) Metal−Organic Framework. *Inorg. Chem.* **2015**, *54* (20), 10096–10104.
(21) Han, X.; Hong, Y.; Ma, Y.; Lu, W.; Li, J.; Lin, L.; Sheveleva, A. M.; Tuna, F.; McInnes, E. J. L.; Dejoie, C.; Sun, J.; Yang, S.; Schröder, M. Adsorption of Nitrogen Dioxide in a Redox-Active Vanadium Metal--Organic Framework Material. *J. Am. Chem. Soc.* **2020**, *142* (36), 15235–15239.
(22) Bloch, E. D.; Queen, W. L.; Hudson, M. R.; Mason, J. A.; Xiao, D. J.; Murray, L. J.; Flacau, R.; Brown, C. M.; Long, J. R. Hydrogen Storage and Selective, Reversible O_2 Adsorption in a Metal−Organic Framework with Open Chromium(II) Sites. *Angew. Chemie Int. Ed.* **2016**, *55* (30), 8605–8609.
(23) Stubbs, A. W.; Dincă, M. Selective Oxidation of C–H Bonds through a Manganese(III) Hydroperoxo in Mn^{II}-Exchanged CFA-1. *Inorg. Chem.* **2019**, *58* (19), 13221–13228.
(24) Bloch, E. D.; Murray, L. J.; Queen, W. L.; Chavan, S.; Maximoff, S. N.; Bigi, J. P.; Krishna, R.; Peterson, V. K.; Grandjean, F.; Long, G. J.; Smit, B.; Bordiga, S.; Brown, C. M.; Long, J. R. Selective Binding of O_2 over N_2 in a Redox-Active Metal–Organic Framework with Open Iron(II) Coordination Sites. *J. Am. Chem. Soc.* **2011**, *133* (37), 14814–14822.
(25) Tulchinsky, Y.; Hendon, C. H.; Lomachenko, K. A.; Borfecchia, E.; Melot, B. C.; Hudson, M. R.; Tarver, J. D.; Korzyński, M. D.; Stubbs, A. W.; Kagan, J. J.; Lamberti, C.; Brown, C. M.; Dincă, M. Reversible Capture and Release of Cl_2 and Br_2 with a Redox-Active Metal−Organic Framework. *J. Am. Chem. Soc.* **2017**, *139* (16), 5992–5997.
(26) Khrizanforova, V.; Shekurov, R.; Miluykov, V.; Khrizanforov, M.; Bon, V.; Kaskel, S.; Gubaidullin, A.; Sinyashin, O.; Budnikova, Y. 3D Ni and Co Redox-Active Metal--Organic Frameworks Based on Ferrocenyl Diphosphinate and 4, 4′-Bipyridine Ligands as Efficient Electrocatalysts for the Hydrogen Evolution Reaction. *Dalt. Trans.* **2020**, *49* (9), 2794–2802.
(27) Denysenko, D.; Grzywa, M.; Jelic, J.; Reuter, K.; Volkmer, D. Scorpionate-Type Coordination in MFU-4l Metal−Organic Frameworks: Small-Molecule Binding and Activation upon the Thermally Activated Formation of Open Metal Sites. *Angew. Chemie Int. Ed.* **2014**, *53* (23), 5832–5836.
(28) Cozzolino, A. F.; Brozek, C. K.; Palmer, R. D.; Yano, J.; Li, M.; Dincă, M. Ligand Redox Non-Innocence in the Stoichiometric Oxidation of Mn_2(2,5-Dioxidoterephthalate) (Mn-MOF-74). *J. Am. Chem. Soc.* **2014**, *136* (9), 3334–3337.
(29) Jain, A.; Hautier, G.; Moore, C. J.; Ping Ong, S.; Fischer, C. C.; Mueller, T.; Persson, K. A.; Ceder, G. A High-Throughput Infrastructure for Density Functional Theory Calculations. *Comput. Mater. Sci.* **2011**, *50* (8), 2295–2310.
(30) Saal, J. E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. Materials Design and

Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD). *Jom* **2013**, *65* (11), 1501–1509.

(31) Tran, K.; Palizhati, A.; Back, S.; Ulissi, Z. W. Dynamic Workflows for Routine Materials Discovery in Surface Science. *J. Chem. Inf. Model.* **2018**, *58* (12), 2392–2400.

(32) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. Identifying Promising Metal−Organic Frameworks for Heterogeneous Catalysis via High-Throughput Periodic Density Functional Theory. *J. Comput. Chem.* **2019**, *40* (12), 1305–1318.

(33) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Mater.* **2013**, *1* (1), 11002.

(34) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C. The Open Quantum Materials Database (OQMD): Assessing the Accuracy of DFT Formation Energies. *npj Comput. Mater.* **2015**, *1* (1), 15010.

(35) Curtarolo, S.; Setyawan, W.; Wang, S.; Xue, J.; Yang, K.; Taylor, R. H.; Nelson, L. J.; Hart, G. L. W.; Sanvito, S.; Buongiorno-Nardelli, M.; Mingo, N.; Levy, O. AFLOWLIB.ORG: A Distributed Materials Properties Repository from High-Throughput Ab Initio Calculations. *Comput. Mater. Sci.* **2012**, *58*, 227–235.

(36) Winther, K. T.; Hoffmann, M. J.; Boes, J. R.; Mamun, O.; Bajdich, M.; Bligaard, T. Catalysis-Hub. Org, an Open Electronic Structure Database for Surface Reactions. *Sci. Data* **2019**, *6*, 75.

(37) Borysov, S. S.; Geilhufe, R. M.; Balatsky, A. V. Organic Materials Database: An Open-Access Online Database for Data Mining. *PLoS One* **2017**, *12* (2), e0171501.

(38) Landis, D. D.; Hummelshoj, J. S.; Nestorov, S.; Greeley, J.; Dulak, M.; Bligaard, T.; Nørskov, J. K.; Jacobsen, K. W. The Computational Materials Repository. *Comput. Sci. Eng.* **2012**, *14* (6), 51–57.

(39) Draxl, C.; Scheffler, M. NOMAD: The FAIR Concept for Big Data-Driven Materials Science. *MRS Bull.* **2018**, *43* (9), 676–682.

(40) Curtarolo, S.; Hart, G. L. W.; Nardelli, M. B.; Mingo, N.; Sanvito, S.; Levy, O. The High-Throughput Highway to Computational Materials Design. *Nat. Mater.* **2013**, *12* (3), 191–201.

(41) Coley, C. W.; Eyke, N. S.; Jensen, K. F. Autonomous Discovery in the Chemical Sciences Part II: Outlook. *Angew. Chemie Int. Ed.* **2020**, *59* (52), 23414–23436.

(42) Liu, Y.; Zhao, T.; Ju, W.; Shi, S. Materials Discovery and Design Using Machine Learning. *J. Mater.* **2017**, *3* (3), 159–177.

(43) Butler, K. T.; Davies, D. W.; Cartwright, H.; Isayev, O.; Walsh, A. Machine Learning for Molecular and Materials Science. *Nature* **2018**, *559* (7715), 547–555.

(44) Moghadam, P. Z.; Li, A.; Wiggin, S. B.; Tao, A.; Maloney, A. G. P.; Wood, P. A.; Ward, S. C.; Fairen-Jimenez, D. Development of a Cambridge Structural Database Subset: A Collection of Metal–Organic Frameworks for Past, Present, and Future. *Chem. Mater.* **2017**, *29* (7), 2618–2625.

(45) Lee, S.; Kim, B.; Cho, H.; Lee, H.; Lee, S. Y.; Cho, E. S.; Kim, J. Computational Screening of Trillions of Metal–Organic Frameworks for High-Performance Methane Storage. *ACS Appl. Mater. Interfaces* **2021**, *13* (20), 23647–23654.

(46) Colón, Y. J.; Snurr, R. Q. High-Throughput Computational Screening of Metal−Organic

Frameworks. *Chem. Soc. Rev.* **2014**, *43* (16), 5735–5749.
(47) Borboudakis, G.; Stergiannakos, T.; Frysali, M.; Klontzas, E.; Tsamardinos, I.; Froudakis, G. E. Chemically Intuited, Large-Scale Screening of MOFs by Machine Learning Techniques. *npj Comput. Mater.* **2017**, *3* (40).
(48) Jablonka, K. M.; Ongari, D.; Moosavi, S. M.; Smit, B. Big-Data Science in Porous Materials: Materials Genomics and Machine Learning. *Chem. Rev.* **2020**, *120* (16), 8066–8129.
(49) Altintas, C.; Altundal, O. F.; Keskin, S.; Yildirim, R. Machine Learning Meets with Metal Organic Frameworks for Gas Storage and Separation. *J. Chem. Inf. Model.* **2021**, *61* (5), 2131–2146.
(50) Shi, Z.; Yang, W.; Deng, X.; Cai, C.; Yan, Y.; Liang, H.; Liu, Z.; Qiao, Z. Machine-Learning-Assisted High-Throughput Computational Screening of High Performance Metal–Organic Frameworks. *Mol. Syst. Des. Eng.* **2020**, *5* (4), 725–742.
(51) Mukherjee, K.; Colón, Y. J. Machine Learning and Descriptor Selection for the Computational Discovery of Metal-Organic Frameworks. *Mol. Simul.* **2021**.
(52) Bobbitt, N. S.; Snurr, R. Q. Molecular Modelling and Machine Learning for High-Throughput Screening of Metal-Organic Frameworks for Hydrogen Storage. *Mol. Simul.* **2019**, *45* (14–15), 1069–1081.
(53) Wilmer, C. E.; Leaf, M.; Lee, C. Y.; Farha, O. K.; Hauser, B. G.; Hupp, J. T.; Snurr, R. Q. Large-Scale Screening of Hypothetical Metal−Organic Frameworks. *Nat. Chem.* **2012**, *4* (2), 83–89.
(54) Chung, Y. G.; Gómez-Gualdrón, D. A.; Li, P.; Leperi, K. T.; Deria, P.; Zhang, H.; Vermeulen, N. A.; Stoddart, J. F.; You, F.; Hupp, J. T.; Farha, O. K.; Snurr, R. Q. In Silico Discovery of Metal−Organic Frameworks for Precombustion CO2 Capture Using a Genetic Algorithm. *Sci. Adv.* **2016**, *2* (9), e1600909.
(55) Gómez-Gualdrón, D. A.; Colón, Y. J.; Zhang, X.; Wang, T. C.; Chen, Y. S.; Hupp, J. T.; Yildirim, T.; Farha, O. K.; Zhang, J.; Snurr, R. Q. Evaluating Topologically Diverse Metal−Organic Frameworks for Cryo-Adsorbed Hydrogen Storage. *Energy Environ. Sci.* **2016**, *9* (10), 3279–3289.
(56) Moghadam, P. Z.; Islamoglu, T.; Goswami, S.; Exley, J.; Fantham, M.; Kaminski, C. F.; Snurr, R. Q.; Farha, O. K.; Fairen-Jimenez, D. Computer-Aided Discovery of a Metal–Organic Framework with Superior Oxygen Uptake. *Nat. Commun.* **2018**, *9* (1), 1378.
(57) Sholl, D.; Steckel, J. A. *Density Functional Theory: A Practical Introduction*; Wiley: Hoboken, NJ, 2009.
(58) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B), B864.
(59) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140* (4A), A1133.
(60) Perdew, J. P.; Schmidt, K. Jacob's Ladder of Density Functional Approximations for the Exchange-Correlation Energy. *AIP Conf. Proc.* **2001**, *577* (1), 1–20.
(61) Mardirossian, N.; Head-Gordon, M. Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals. *Mol. Phys.* **2017**, *115* (19), 2315–2372.
(62) Lehtola, S.; Steigemann, C.; Oliveira, M. J. T.; Marques, M. A. L. Recent Developments in Libxc—A Comprehensive Library of Functionals for Density Functional Theory.

SoftwareX **2018**, *7*, 1–5.
(63) Perdew, J. P.; Yue, W. Accurate and Simple Density Functional for the Electronic Exchange Energy: Generalized Gradient Approximation. *Phys. Rev. B* **1986**, *33* (12), 8800.
(64) Perdew, J. P.; Kurth, S.; Zupan, A.; Blaha, P. Accurate Density Functional with Correct Formal Properties: A Step beyond the Generalized Gradient Approximation. *Phys. Rev. Lett.* **1999**, *82* (12), 2544.
(65) Becke, A. D. A New Mixing of Hartree--Fock and Local Density-Functional Theories. *J. Chem. Phys.* **1993**, *98* (2), 1372–1377.
(66) Hartree, D. R. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods. *Math. Proc. Cambridge Philos. Soc.* **1928**, *24* (1), 89–110.
(67) Fock, V. Näherungsmethode Zur Lösung Des Quantenmechanischen Mehrkörperproblems. *Zeitschrift für Phys.* **1930**, *61* (1–2), 126–148.
(68) Manz, T. "Survivorship bias" in science and the Marcel Swart DFT poll http://www.ccl.net/cgi-bin/ccl/message-new?2019+03+15+001.
(69) Bernales, V.; Ortuño, M. A.; Truhlar, D. G.; Cramer, C. J.; Gagliardi, L. Computational Design of Functionalized Metal–Organic Framework Nodes for Catalysis. *ACS Cent. Sci.* **2018**, *4* (1), 5–19.
(70) Prodan, E.; Kohn, W. Nearsightedness of Electronic Matter. *Proc. Natl. Acad. Sci.* **2005**, *102* (33), 11635–11638.
(71) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. Structure−Activity Relationships That Identify Metal−Organic Framework Catalysts for Methane Activation. *ACS Catal.* **2019**, *9*, 3576–3587.
(72) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. High-Valent Metal-Oxo Species at the Nodes of Metal−Triazolate Frameworks: The Effects of Ligand-Exchange and Two-State Reactivity for C−H Bond Activation. *Angew. Chemie Int. Ed.* **2020**, *132* (44), 19662–19670.
(73) Rosen, A. S.; Mian, M. R.; Islamoglu, T.; Chen, H.; Farha, O. K.; Notestein, J. M.; Snurr, R. Q. Tuning the Redox Activity of Metal−Organic Frameworks for Enhanced, Selective O_2 Binding: Design Rules and Ambient Temperature O_2 Chemisorption in a Cobalt−Triazolate Framework. *J. Am. Chem. Soc.* **2020**, *142* (9), 4317–4328.
(74) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. Comparing GGA, GGA+U, and Meta-GGA Functionals for Redox-Dependent Binding at Open Metal Sites in Metal−Organic Frameworks. *J. Chem. Phys.* **2020**, *152*, 224101.
(75) Rosen, A. S.; Iyer, S. M.; Ray, D.; Yao, Z.; Aspuru-Guzik, A.; Gagliardi, L.; Notestein, J. M.; Snurr, R. Q. Machine Learning the Quantum-Chemical Properties of Metal–Organic Frameworks for Accelerated Materials Discovery. *Matter* **2021**, *4* (5), 1578–1597.
(76) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. Comprehensive Phase Diagrams of MoS2 Edge Sites Using Dispersion-Corrected DFT Free Energy Calculations. *J. Phys. Chem. C* **2018**, *122* (27), 15318–15329.
(77) Nauert, S. L.; Rosen, A. S.; Kim, H.; Snurr, R. Q.; Stair, P. C.; Notestein, J. M. Evidence for Copper Dimers in Low-Loaded CuOx/SiO2 Catalysts for Cyclohexane Oxidative Dehydrogenation. *ACS Catal.* **2018**, *8* (10), 9775–9789.
(78) Wasson, M. C.; Zhang, X.; Otake, K.; Rosen, A. S.; Alayoglu, S.; Krzyaniak, M. D.; Chen, Z.; Redfern, L. R.; Robison, L.; Son, F. A.; Chen, Y.; Islamoglu, T.; Notestein, J. M.; Snurr, R. Q.; Wasielewski, M. R.; Farha, O. K. Supramolecular Porous Assemblies of Atomically

Precise Catalytically Active Cerium-Based Clusters. *Chem. Mater.* **2020**, *32* (19), 8522–8529.

(79) Bucior, B. J.; Rosen, A. S.; Haranczyk, M.; Yao, Z.; Ziebel, M. E.; Farha, O. K.; Hupp, J. T.; Siepmann, J. I.; Aspuru-Guzik, A.; Snurr, R. Q. Identification Schemes for Metal–Organic Frameworks to Enable Rapid Search and Cheminformatics Analysis. *Cryst. Growth Des.* **2019**, *19* (11), 6682–6697.

(80) Bobbitt, N. S.; Rosen, A. S.; Snurr, R. Q. Topological Effects on Separation of Alkane Isomers in Metal–Organic Frameworks. *Fluid Phase Equilib.* **2020**, *112642*, 112642.

(81) Hicks, K. E.; Rosen, A. S.; Syed, Z. H.; Snurr, R. Q.; Farha, O. K.; Notestein, J. M. Zr6O8 Node-Catalyzed Butene Hydrogenation and Isomerization in the Metal-Organic Framework NU-1000. *ACS Catal.* **2020**, *10* (24), 14959–14970.

(82) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (Pymatgen): A Robust, Open-Source Python Library for Materials Analysis. *Comput. Mater. Sci.* **2013**, *68*, 314–319.

(83) Larsen, A.; Mortensen, J.; Blomqvist, J.; Castelli, I.; Christensen, R.; Dulak, M.; Friis, J.; Groves, M.; Hammer, B.; Hargus, C.; Hermes, E.; Jennings, P.; Jensen, P.; Kermode, J.; Kitchin, J.; Kolsbjerg, E.; Kubal, J.; Kaasbjerg, K.; Lysgaard, S.; Maronsson, J.; Maxson, T.; Olsen, T.; Pastewka, L.; Peterson, A.; Rostgaard, C.; Schiøtz, J.; Schütt, O.; Strange, M.; Thygesen, K.; Vegge, T.; Vilhelmsen, L.; Walter, M.; Zeng, Z.; Jacobsen, K. The Atomic Simulation Environment—A Python Library for Working with Atoms. *J. Phys. Condens. Matter* **2017**, *29*, 273002.

(84) Pizzi, G.; Cepellotti, A.; Sabatini, R.; Marzari, N.; Kozinsky, B. AiiDA: Automated Interactive Infrastructure and Database for Computational Science. *Comput. Mater. Sci.* **2016**, *111*, 218–230.

(85) Mathew, K.; Montoya, J. H.; Faghaninia, A.; Dwarakanath, S.; Aykol, M.; Tang, H.; Chu, I. heng; Smidt, T.; Bocklund, B.; Horton, M.; Dagdelen, J.; Wood, B.; Liu, Z. K.; Neaton, J.; Ong, S. P.; Persson, K.; Jain, A. Atomate: A High-Level Interface to Generate, Execute, and Analyze Computational Materials Science Workflows. *Comput. Mater. Sci.* **2017**, *139*, 140–152.

(86) Jain, A.; Ong, S. P.; Chen, W.; Medasani, B.; Qu, X.; Kocher, M.; Brafman, M.; Petretto, G.; Rignanese, G.-M.; Hautier, G.; Gunter, D.; Persson, K. FireWorks: A Dynamic Workflow System Designed for High-Throughput Applications. *Concurr. Comput. Pract. Exp.* **2015**, *27* (17), 5037–5059.

(87) Mayeshiba, T.; Wu, H.; Angsten, T.; Kaczmarowski, A.; Song, Z.; Jenness, G.; Xie, W.; Morgan, D. The MAterials Simulation Toolkit (MAST) for Atomistic Modeling of Defects and Diffusion. *Comput. Mater. Sci.* **2017**, *126*, 90–102.

(88) Choudhary, K.; Kalish, I.; Beams, R.; Tavazza, F. High-Throughput Identification and Characterization of Two-Dimensional Materials Using Density Functional Theory. *Sci. Rep.* **2017**, *7* (1), 1–16.

(89) Chung, Y. G.; Camp, J.; Haranczyk, M.; Sikora, B. J.; Bury, W.; Krungleviciute, V.; Yildirim, T.; Farha, O. K.; Sholl, D. S.; Snurr, R. Q. Computation-Ready, Experimental Metal−Organic Frameworks: A Tool to Enable High-Throughput Screening of Nanoporous Crystals. *Chem. Mater.* **2014**, *26* (21), 6185–6192.

(90) Colón, Y. J.; Gómez-Gualdrón, D. A.; Snurr, R. Q. Topologically Guided, Automated Construction of Metal–Organic Frameworks and Their Evaluation for Energy-Related Applications. *Cryst. Growth Des.* **2017**, *17* (11), 5801–5810.
(91) Götl, F.; Michel, C.; Andrikopoulos, P. C.; Love, A. M.; Hafner, J.; Hermans, I.; Sautet, P. Computationally Exploring Confinement Effects in the Methane-to-Methanol Conversion over Iron-Oxo Centers in Zeolites. *ACS Catal.* **2016**, *6* (12), 8404–8409.
(92) Mahyuddin, M. H.; Staykov, A.; Shiota, Y.; Yoshizawa, K. Direct Conversion of Methane to Methanol by Metal-Exchanged ZSM-5 Zeolite (Metal = Fe, Co, Ni, Cu). *ACS Catal.* **2016**, *6* (12), 8321–8331.
(93) Horn, R.; Schlögl, R. Methane Activation by Heterogeneous Catalysis. *Catal. Letters* **2015**, *145* (1), 23–39.
(94) Liu, M.; Wu, J.; Hou, H. Metal−Organic Frameworks Based Materials as Heterogeneous Catalysts for C-H Bond Activation. *Chem. – A Eur. J.* **2018**, *25* (12), 2935–2948.
(95) Phan, A.; Czaja, A. U.; Gándara, F.; Knobler, C. B.; Yaghi, O. M. Metal–Organic Frameworks of Vanadium as Catalysts for Conversion of Methane to Acetic Acid. *Inorg. Chem.* **2011**, *50* (16), 7388–7390.
(96) Ikuno, T.; Zheng, J.; Vjunov, A.; Sanchez-Sanchez, M.; Ortuño, M. A.; Pahls, D. R.; Fulton, J. L.; Camaioni, D. M.; Li, Z.; Ray, D.; Mehdi, B. L.; Browning, N. D.; Farha, O. K.; Hupp, J. T.; Cramer, C. J.; Gagliardi, L.; Lercher, J. A. Methane Oxidation to Methanol Catalyzed by Cu-Oxo Clusters Stabilized in NU-1000 Metal–Organic Framework. *J. Am. Chem. Soc.* **2017**, *139* (30), 10294–10301.
(97) Zhang, X.; Huang, Z.; Ferrandon, M.; Yang, D.; Robison, L.; Li, P.; Wang, T. C.; Delferro, M.; Farha, O. K. Catalytic Chemoselective Functionalization of Methane in a Metal–Organic Framework. *Nat. Catal.* **2018**, *1* (5), 356–362.
(98) Osadchii, D.; Olivos Suarez, A. I.; Szécsényi, Á.; Li, G.; Nasalevich, M. A.; Dugulan, A. I.; Serra-Crespo, P.; Hensen, E. J. M.; Veber, S. L.; Fedin, M. V; Sankar, G.; Pidko, E. A.; Gascon, J. Isolated Fe Sites in Metal−Organic Framework Catalyze the Direct Conversion of Methane to Methanol. *ACS Catal.* **2018**, *8* (6), 5542–5548.
(99) Baek, J.; Rungtaweevoranit, B.; Pei, X.; Park, M.; Fakra, S. C.; Liu, Y.-S.; Matheu, R.; Alshmimri, S. A.; Alshihri, S.; Trickett, C. A.; Somorjai, G. A.; Yaghi, O. M. Bioinspired Metal−Organic Framework Catalysts for Selective Methane Oxidation to Methanol. *J. Am. Chem. Soc.* **2018**, *140* (51), 18208–18216.
(100) Greeley, J.; Nørskov, J. K. Large-Scale, Density Functional Theory-Based Screening of Alloys for Hydrogen Evolution. *Surf. Sci.* **2007**, *601* (6), 1590–1598.
(101) Nørskov, J. K.; Bligaard, T.; Rossmeisl, J.; Christensen, C. H. Towards the Computational Design of Solid Catalysts. *Nat. Chem.* **2009**, *1* (1), 37–46.
(102) Nørskov, J. K.; Studt, F.; Abild-Pedersen, F.; Bligaard, T. *Fundamental Concepts in Heterogeneous Catalysis*; John Wiley & Sons, 2014.
(103) Bray, J. M.; Schneider, W. F. First-Principles Thermodynamic Models in Heterogeneous Catalysis. In *Computational Catalysis*; Asthagiri, A., Janik, M. J., Eds.; The Royal Society of Chemistry, 2014; pp 84–93.
(104) Wang, S.; Petzold, V.; Tripkovic, V.; Kleis, J.; Howalt, J. G.; Skulason, E.; Fernandez, E. M.; Hvolbæk, B.; Jones, G.; Toftelund, A.; Falsig, H.; Björketun, M.; Studt, F.; Abild-Pedersen, F.; Rossmeisl, J.; Nørskov, J. K.; Bligaard, T. Universal Transition State Scaling

Relations for (de)Hydrogenation over Transition Metals. *Phys. Chem. Chem. Phys.* **2011**, *13* (46), 20760–20765.
(105) Latimer, A. A.; Kulkarni, A. R.; Aljama, H.; Montoya, J. H.; Yoo, J. S.; Tsai, C.; Abild-Pedersen, F.; Studt, F.; Nørskov, J. K. Understanding Trends in C–H Bond Activation in Heterogeneous Catalysis. *Nat. Mater.* **2017**, *16* (2), 225–229.
(106) Gani, T. Z. H.; Kulik, H. J. Understanding and Breaking Scaling Relations in Single-Site Catalysis: Methane-to-Methanol Conversion by Fe(IV)=O. *ACS Catal.* **2018**, *8* (2), 975–986.
(107) Liao, P.; Getman, R. B.; Snurr, R. Q. Optimizing Open Iron Sites in Metal−Organic Frameworks for Ethane Oxidation: A First-Principles Study. *ACS Appl. Mater. Interfaces* **2017**, *9* (39), 33484–33492.
(108) Zimmermann, N. E. R.; Horton, M. K.; Jain, A.; Haranczyk, M. Assessing Local Structure Motifs Using Order Parameters for Motif Recognition, Interstitial Identification, and Diffusion Path Characterization. *Front. Mater.* **2017**, *4*, 34.
(109) Willems, T. F.; Rycroft, C. H.; Kazi, M.; Meza, J. C.; Haranczyk, M. Algorithms and Tools for High-Throughput Geometry-Based Analysis of Crystalline Porous Materials. *Microporous Mesoporous Mater.* **2012**, *149* (1), 134–141.
(110) Sarkisov, L.; Bueno-Perez, R.; Sutharson, M.; Fairen-jimenez, D. Material Informatics with PoreBlazer v4.0 and CSD MOF Database. *Chem. Mater.* **2020**, *32* (23), 9849–9867.
(111) Latimer, A. A.; Aljama, H.; Kakekhani, A.; Yoo, J. S.; Kulkarni, A.; Tsai, C.; Garcia-Melchor, M.; Abild-Pedersen, F.; Nørskov, J. K. Mechanistic Insights into Heterogeneous Methane Activation. *Phys. Chem. Chem. Phys.* **2017**, *19* (5), 3575–3581.
(112) Nazarian, D.; Camp, J. S.; Chung, Y. G.; Snurr, R. Q.; Sholl, D. S. Large-Scale Refinement of Metal–Organic Framework Structures Using Density Functional Theory. *Chem. Mater.* **2017**, *29* (6), 2521–2528.
(113) Hutter, J.; Iannuzzi, M.; Schiffmann, F.; VandeVondele, J. CP2K: Atomistic Simulations of Condensed Matter Systems. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2014**, *4* (1), 15–25.
(114) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. Quickstep: Fast and Accurate Density Functional Calculations Using a Mixed Gaussian and Plane Waves Approach. *Comput. Phys. Commun.* **2005**, *167* (2), 103–128.
(115) Chui, S. S.-Y.; Lo, S. M.-F.; Charmant, J. P. H.; Orpen, A. G.; Williams, I. D. A Chemically Functionalizable Nanoporous Material $[Cu_3(TMA)_2(H_2O)_3]_N$. *Science* **1999**, *283* (5405), 1148–1150.
(116) Vogiatzis, K. D.; Haldoupis, E.; Xiao, D. J.; Long, J. R.; Siepmann, J. I.; Gagliardi, L. Accelerated Computational Analysis of Metal−Organic Frameworks for Oxidation Catalysis. *J. Phys. Chem. C* **2016**, *120* (33), 18707–18712.
(117) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. MOF Adsorbate Initializer https://doi.org/10.5281/zenodo.1451875.
(118) Martin, M. G.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of *n*-Alkanes. *J. Phys. Chem. B* **1998**, *102* (14), 2569–2577.
(119) Dubbeldam, D.; Calero, S. S.; Ellis, D. E.; Snurr, R. Q. RASPA: Molecular Simulation Software for Adsorption and Diffusion in Flexible Nanoporous Materials. *Mol. Simul.* **2016**, *42* (2), 81–101.

(120) Haldoupis, E. Open Metal Detector https://github.com/emmhald/open_metal_detector.
(121) PorousMaterials.jl https://doi.org/10.5281/zenodo.1400838.
(122) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54* (16), 11169–11186.
(123) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59* (3), 1758–1775.
(124) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50* (24), 17953–17979.
(125) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.
(126) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H−Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104.
(127) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32* (7), 1456–1465.
(128) Nazarian, D.; Ganesh, P.; Sholl, D. S. Benchmarking Density Functional Theory Predictions of Framework Structures and Properties in a Chemically Diverse Test Set of Metal–Organic Frameworks. *J. Mater. Chem. A* **2015**, *3* (44), 22432–22440.
(129) Formalik, F.; Fischer, M.; Rogacka, J.; Firlej, L.; Kuchta, B. Benchmarking of GGA Density Functionals for Modeling Structures of Nanoporous, Rigid and Flexible MOFs. *J. Chem. Phys.* **2018**, *149*, 064110.
(130) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. A Climbing Image Nudged Elastic Band Method for Finding Saddle Points and Minimum Energy Paths. *J. Chem. Phys.* **2000**, *113* (22), 9901–9904.
(131) Henkelman, G.; Jónsson, H. A Dimer Method for Finding Saddle Points on High Dimensional Potential Surfaces Using Only First Derivatives. *J. Chem. Phys.* **1999**, *111* (15), 7010–7022.
(132) Teter, M. P.; Payne, M. C.; Allan, D. C. Solution of Schrödinger's Equation for Large Systems. *Phys. Rev. B* **1989**, *40* (18), 12255–12263.
(133) Bylander, D. M.; Kleinman, L.; Lee, S. Self-Consistent Calculations of the Energy Bands and Bonding Properties of B12C3. *Phys. Rev. B* **1990**, *42* (2), 1394–1403.
(134) Pulay, P. Convergence Acceleration of Iterative Sequences. The Case of SCF Iteration. *Chem. Phys. Lett.* **1980**, *73* (2), 393–398.
(135) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Comput. Mater. Sci.* **1996**, *6* (1), 15–50.
(136) Francis, G. P.; Payne, M. C. Finite Basis Set Corrections to Total Energy Pseudopotential Calculations. *J. Phys. Condens. Matter* **1990**, *2* (19), 4395.
(137) Bitzek, E.; Koskinen, P.; Gähler, F.; Moseler, M.; Gumbsch, P. Structural Relaxation Made Simple. *Phys. Rev. Lett.* **2006**, *97* (17), 170201.
(138) Henkelman, G. VTST Tools http://theory.cm.utexas.edu/vtsttools/index.html.
(139) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. PyMOFScreen https://doi.org/10.5281/zenodo.1451873.
(140) Kulkarni, A. R.; Zhao, Z.-J.; Siahrostami, S.; Nørskov, J. K.; Studt, F. Cation-Exchanged

Zeolites for the Selective Oxidation of Methane to Methanol. *Catal. Sci. Technol.* **2018**, *8* (1), 114–123.

(141) Haynes, W. M. *CRC Handbook of Chemistry and Physics*, 98th ed.; Rumble, J. R., Ed.; CRC Press: Boca Raton, FL, 2017.

(142) Ling, Y.; Zhang, L.; Li, J.; Du, M. A Robust Porous PtS-Type Cu(II) Metal–Organic Framework: Single-Crystal-to-Single-Crystal Transformation with Reversible Guest Intercalation Accompanied by Colour Change. *CrystEngComm* **2011**, *13* (3), 768–770.

(143) Woertink, J. S.; Smeets, P. J.; Groothaert, M. H.; Vance, M. A.; Sels, B. F.; Schoonheydt, R. A.; Solomon, E. I. A $[Cu_2O]^{2+}$ Core in Cu-ZSM-5, the Active Site in the Oxidation of Methane to Methanol. *Proc. Natl. Acad. Sci.* **2009**, *106* (45), 18908–18913.

(144) Grundner, S.; Markovits, M. A. C.; Li, G.; Tromp, M.; Pidko, E. A.; Hensen, E. J. M.; Jentys, A.; Sanchez-Sanchez, M.; Lercher, J. A. Single-Site Trinuclear Copper Oxygen Clusters in Mordenite for Selective Conversion of Methane to Methanol. *Nat. Commun.* **2015**, *6* (May), 7546.

(145) Smeets, P. J.; Hadt, R. G.; Woertink, J. S.; Vanelderen, P.; Schoonheydt, R. A.; Sels, B. F.; Solomon, E. I. Oxygen Precursor to the Reactive Intermediate in Methanol Synthesis by Cu-ZSM-5. *J. Am. Chem. Soc.* **2010**, *132* (42), 14736–14738.

(146) Tomkins, P.; Ranocchiari, M.; Van Bokhoven, J. A. Direct Conversion of Methane to Methanol under Mild Conditions over Cu-Zeolites and Beyond. *Acc. Chem. Res.* **2017**, *50* (2), 418–425.

(147) Howarth, A. J.; Liu, Y.; Li, P.; Li, Z.; Wang, T. C.; Hupp, J. T.; Farha, O. K. Chemical, Thermal and Mechanical Stabilities of Metal−Organic Frameworks. *Nat. Rev. Mater.* **2016**, *1* (3), 15018.

(148) Colombo, V.; Galli, S.; Choi, H. J.; Han, G. D.; Maspero, A.; Palmisano, G.; Masciocchi, N.; Long, J. R. High Thermal and Chemical Stability in Pyrazolate-Bridged Metal−Organic Frameworks with Exposed Metal Sites. *Chem. Sci.* **2011**, *2* (7), 1311–1319.

(149) Szécsényi, Á.; Li, G.; Gascon, J.; Pidko, E. A. Unraveling Reaction Networks behind the Catalytic Oxidation of Methane with H_2O_2 over a Mixed-Metal MIL-53(Al,Fe) MOF Catalyst. *Chem. Sci.* **2018**, *9* (33), 6765–6773.

(150) Cui, X.; Li, H. H.; Wang, Y.; Hu, Y.; Hua, L.; Li, H. H.; Han, X.; Liu, Q.; Yang, F.; He, L.; Chen, X.; Li, Q.; Xiao, J.; Deng, D.; Bao, X. Room-Temperature Methane Conversion by Graphene-Confined Single Iron Atoms. *Chem* **2018**, *4* (8), 1902–1910.

(151) Xu, J.; Armstrong, R. D.; Shaw, G.; Dummer, N. F.; Freakley, S. J.; Taylor, S. H.; Hutchings, G. J. Continuous Selective Oxidation of Methane to Methanol over Cu-and Fe-Modified ZSM-5 Catalysts in a Flow Reactor. *Catal. Today* **2016**, *270*, 93–100.

(152) Hammond, C.; Forde, M. M.; Ab Rahim, M. H.; Thetford, A.; He, Q.; Jenkins, R. L.; Dimitratos, N.; Lopez-Sanchez, J. A.; Dummer, N. F.; Murphy, D. M.; Carley, A. F.; Taylor, S. H.; Willock, D. J.; Stangland, E. E.; Kang, J.; Hagen, H.; Kiely, C. J.; Hutchings, G. J. Direct Catalytic Conversion of Methane to Methanol in an Aqueous Medium by Using Copper-Promoted Fe-ZSM-5. *Angew. Chemie Int. Ed.* **2012**, *51* (21), 5129–5133.

(153) Dau, P. V; Kim, M.; Cohen, S. M. Site-Selective Cyclometalation of a Metal–Organic Framework. *Chem. Sci.* **2013**, *4* (2), 601–605.

(154) Bloch, W. M.; Burgun, A.; Coghlan, C. J.; Lee, R.; Coote, M. L.; Doonan, C. J.; Sumby, C. J. Capturing Snapshots of Post-Synthetic Metallation Chemistry in Metal−organic

Frameworks. *Nat. Chem.* **2014**, *6* (10), 906.
(155) Tzeng, B.-C.; Chang, T.-Y. Toward Copper (I)-Iodide-Based Coordination Architectures via N,N′-Bis(Pyridylcarbonyl)-4,4′-Diaminodiphenyl Ether with Different Solvent Compositions. *Cryst. Growth Des.* **2009**, *9* (12), 5343–5350.
(156) Schmieder, P.; Denysenko, D.; Grzywa, M.; Magdysyuk, O.; Volkmer, D. A Structurally Flexible Triazolate-Based Metal−Organic Framework Featuring Coordinatively Unsaturated Copper(I) Sites. *Dalt. Trans.* **2016**, *45* (35), 13853–13862.
(157) Dinh, K. T.; Sullivan, M. M.; Serna, P.; Meyer, R. J.; Dincă, M.; Román-Leshkov, Y. Viewpoint on the Partial Oxidation of Methane to Methanol Using Cu- and Fe-Exchanged Zeolites. *ACS Catal.* **2018**, *8* (9), 8306–8313.
(158) Manz, T. A.; Limas, N. G. Introducing DDEC6 Atomic Population Analysis: Part 1. Charge Partitioning Theory and Methodology. *RSC Adv.* **2016**, *6* (53), 47771–47801.
(159) Tang, W.; Sanville, E.; Henkelman, G. A Grid-Based Bader Analysis Algorithm without Lattice Bias. *J. Phys. Condens. Matter* **2009**, *21* (8), 84204.
(160) Jacoby, M. Chemists Uncover Rules of Thumb to Help with Computational Screening of MOF Catalysts. *C&EN* **2019**, *97* (14).
(161) Faramawy, S.; Zaki, T.; Sakr, A. A.-E. Natural Gas Origin, Composition, and Processing: A Review. *J. Nat. Gas Sci. Eng.* **2016**, *34*, 34–54.
(162) Aasberg-Petersen, K.; Dybkjær, I.; Ovesen, C. V; Schjødt, N. C.; Sehested, J.; Thomsen, S. G. Natural Gas to Synthesis Gas − Catalysts and Catalytic Processes. *J. Nat. Gas Sci. Eng.* **2011**, *3* (2), 423–459.
(163) Elvidge, C. D.; Zhizhin, M.; Baugh, K.; Hsu, F.-C.; Ghosh, T. Methods for Global Survey of Natural Gas Flaring from Visible Infrared Imaging Radiometer Suite Data. *Energies* **2016**, *9* (1), 14.
(164) McFarland, E. Unconventional Chemistry for Unconventional Natural Gas. *Science* **2012**, *338* (6105), 340–342.
(165) Blanksby, S. J.; Ellison, G. B. Bond Dissociation Energies of Organic Molecules. *Acc. Chem. Res.* **2003**, *36* (4), 255–263.
(166) Olivos-Suarez, A. I.; Szécsényi, A.; Hensen, E. J. M.; Ruiz-Martinez, J.; Pidko, E. A.; Gascon, J. Strategies for the Direct Catalytic Valorization of Methane Using Heterogeneous Catalysis: Challenges and Opportunities. *ACS Catal.* **2016**, *6* (5), 2965–2981.
(167) Ravi, M.; Ranocchiari, M.; van Bokhoven, J. A. The Direct Catalytic Oxidation of Methane to Methanol — A Critical Assessment. *Angew. Chemie Int. Ed.* **2017**, *56* (52), 16464–16483.
(168) Xiao, D. J.; Bloch, E. D.; Mason, J. A.; Queen, W. L.; Hudson, M. R.; Planas, N.; Borycz, J.; Dzubak, A. L.; Verma, P.; Lee, K.; Bonino, F.; Crocellà, V.; Yano, J.; Bordiga, S.; Truhlar, D. G.; Gagliardi, L.; Brown, C. M.; Long, J. R. Oxidation of Ethane to Ethanol by N_2O in a Metal−Organic Framework with Coordinatively Unsaturated Iron(II) Sites. *Nat. Chem.* **2014**, *6* (7), 590–595.
(169) Li, Z.; Peters, A. W.; Bernales, V.; Ortuño, M. A.; Schweitzer, N. M.; DeStefano, M. R.; Gallington, L. C.; Platero-Prats, A. E.; Chapman, K. W.; Cramer, C. J.; Gagliardi, L.; Hupp, J. T.; Farha, O. K. Metal−Organic Framework Supported Cobalt Catalysts for the Oxidative Dehydrogenation of Propane at Low Temperature. *ACS Cent. Sci.* **2016**, *3* (1), 31–38.
(170) Wang, C.; An, B.; Lin, W. Metal–Organic Frameworks in Solid–Gas Phase Catalysis. *ACS*

Catal. **2019**, *9* (1), 130–146.
(171) Furukawa, H.; Cordova, K. E.; O'Keeffe, M.; Yaghi, O. M. The Chemistry and Applications of Metal−Organic Frameworks. *Science* **2013**, *341* (6149), 1230444.
(172) Yaghi, O. M.; O'Keeffe, M.; Ockwig, N. W.; Chae, H. K.; Eddaoudi, M.; Kim, J. Reticular Synthesis and the Design of New Materials. *Nature* **2003**, *423* (6941), 705–714.
(173) Boyd, P. G.; Lee, Y.; Smit, B. Computational Development of the Nanoporous Materials Genome. *Nat. Rev. Mater.* **2017**, *2* (8), 17037.
(174) Coudert, F.-X.; Fuchs, A. H. Computational Characterization and Prediction of Metal−Organic Framework Properties. *Coord. Chem. Rev.* **2016**, *307*, 211–236.
(175) Greeley, J.; Jaramillo, T. F.; Bonde, J.; Chorkendorff, I.; Nørskov, J. K. Computational High-Throughput Screening of Electrocatalytic Materials for Hydrogen Evolution. *Nat. Mater.* **2006**, *5*, 909.
(176) Pahls, D. R.; Ortuño, M. A.; Winegar, P. H.; Cramer, C. J.; Gagliardi, L. Computational Screening of Bimetal-Functionalized Zr6O8 MOF Nodes for Methane C-H Bond Activation. *Inorg. Chem.* **2017**, *56* (15), 8739–8743.
(177) Verma, P.; Vogiatzis, K. D.; Planas, N.; Borycz, J.; Xiao, D. J.; Long, J. R.; Gagliardi, L.; Truhlar, D. G. Mechanism of Oxidation of Ethane to Ethanol at Iron(IV)-Oxo Sites in Magnesium-Diluted Fe_2(DOBDC). *J. Am. Chem. Soc.* **2015**, *137* (17), 5770–5781.
(178) Suh, B. L.; Kim, J. Ligand Insertion in MOF-74 as Effective Design for Oxidation of Ethane to Ethanol. *J. Phys. Chem. C* **2018**, *122* (40), 23078–23083.
(179) Vitillo, J. G.; Bhan, A.; Cramer, C. J.; Lu, C. C.; Gagliardi, L. Quantum Chemical Characterization of Structural Single Fe(II) Sites in MIL-Type Metal Organic Frameworks for Oxidation of Methane to Methanol and Ethane to Ethanol. *ACS Catal.* **2019**, *9*, 2870–2879.
(180) Nørskov, J. K.; Bligaard, T.; Logadottir, A.; Bahn, S.; Hansen, L. B.; Bollinger, M.; Bengaard, H.; Hammer, B.; Sljivancanin, Z.; Mavrikakis, M.; Xu, Y.; Dahl, S.; Jacobsen, C. J. H. Universality in Heterogeneous Catalysis. *J. Catal.* **2002**, *209* (2), 275–278.
(181) Bligaard, T.; Nørskov, J. K.; Dahl, S.; Matthiesen, J.; Christensen, C. H.; Sehested, J. The Brønsted–Evans–Polanyi Relation and the Volcano Curve in Heterogeneous Catalysis. *J. Catal.* **2004**, *224* (1), 206–217.
(182) Medford, A. J.; Vojvodic, A.; Hummelshøj, J. S.; Voss, J.; Abild-Pedersen, F.; Studt, F.; Bligaard, T.; Nilsson, A.; Nørskov, J. K. From the Sabatier Principle to a Predictive Theory of Transition-Metal Heterogeneous Catalysis. *J. Catal.* **2015**, *328*, 36–42.
(183) Montemore, M. M.; Medlin, J. W. Scaling Relations between Adsorption Energies for Computational Screening and Design of Catalysts. *Catal. Sci. Technol.* **2014**, *4* (11), 3748–3761.
(184) Greeley, J. Theoretical Heterogeneous Catalysis: Scaling Relationships and Computational Catalyst Design. *Annu. Rev. Chem. Biomol. Eng.* **2016**, *7* (1), 605–635.
(185) Murray, L. J.; Dincă, M.; Yano, J.; Chavan, S.; Bordiga, S.; Brown, C. M.; Long, J. R. Highly-Selective and Reversible O_2 Binding in Cr_3(1,3,5-Benzenetricarboxylate)$_2$. *J. Am. Chem. Soc.* **2010**, *132* (23), 7856–7857.
(186) Yaghi, O. M.; Li, H.; Davis, C.; Richardson, D.; Groy, T. L. Synthetic Strategies, Structure Patterns, and Emerging Properties in the Chemistry of Modular Porous Solids. *Acc. Chem. Res.* **1998**, *31* (8), 474–484.

(187) Maniam, P.; Stock, N. Investigation of Porous Ni-Based Metal−Organic Frameworks Containing Paddle-Wheel Type Inorganic Building Units via High-Throughput Methods. *Inorg. Chem.* **2011**, *50* (11), 5085–5097.
(188) Feldblyum, J. I.; Liu, M.; Gidley, D. W.; Matzger, A. J. Reconciling the Discrepancies between Crystallographic Porosity and Guest Access as Exemplified by Zn-HKUST-1. *J. Am. Chem. Soc.* **2011**, *133* (45), 18257–18263.
(189) Kramer, M.; Schwarz, U.; Kaskel, S. Synthesis and Properties of the Metal−Organic Framework $Mo_3(BTC)_2$ (TUDMOF-1). *J. Mater. Chem.* **2006**, *16* (23), 2245–2248.
(190) Kozachuk, O.; Yusenko, K.; Noei, H.; Wang, Y.; Walleck, S.; Glaser, T.; Fischer, R. A. Solvothermal Growth of a Ruthenium Metal−Organic Framework Featuring HKUST-1 Structure Type as Thin Films on Oxide Surfaces. *Chem. Commun.* **2011**, *47* (30), 8509–8511.
(191) Guo, P.; Froese, C.; Fu, Q.; Chen, Y.-T.; Peng, B.; Kleist, W.; Fischer, R. A.; Muhler, M.; Wang, Y. CuPd Mixed-Metal HKUST-1 as Catalyst for Aerobic Alcohol Oxidation. *J. Phys. Chem. C* **2018**, *122* (37), 21433–21440.
(192) Rieth, A. J.; Tulchinsky, Y.; Dincă, M. High and Reversible Ammonia Uptake in Mesoporous Azolate Metal–Organic Frameworks with Open Mn, Co, and Ni Sites. *J. Am. Chem. Soc.* **2016**, *138* (30), 9401–9404.
(193) Reed, D. A.; Keitz, B. K.; Oktawiec, J.; Mason, J. A.; Runčevski, T.; Xiao, D. J.; Darago, L. E.; Crocellà, V.; Bordiga, S.; Long, J. R. A Spin Transition Mechanism for Cooperative Adsorption in Metal−Organic Frameworks. *Nature* **2017**, *550* (7674), 96–100.
(194) Liao, P.-Q.; Chen, H.; Zhou, D.-D.; Liu, S.-Y.; He, C.-T.; Rui, Z.; Ji, H.; Zhang, J.-P.; Chen, X.-M. Monodentate Hydroxide as a Super Strong yet Reversible Active Site for CO_2 Capture from High-Humidity Flue Gas. *Energy Environ. Sci.* **2015**, *8* (3), 1011–1016.
(195) Dincă, M.; Dailly, A.; Liu, Y.; Brown, C. M.; Neumann, D. A.; Long, J. R. Hydrogen Storage in a Microporous Metal–Organic Framework with Exposed Mn^{2+} Coordination Sites. *J. Am. Chem. Soc.* **2006**, *128* (51), 16876–16883.
(196) Sumida, K.; Horike, S.; Kaye, S. S.; Herm, Z. R.; Queen, W. L.; Brown, C. M.; Grandjean, F.; Long, G. J.; Dailly, A.; Long, J. R. Hydrogen Storage and Carbon Dioxide Capture in an Iron-Based Sodalite-Type Metal−Organic Framework (Fe-BTT) Discovered via High-Throughput Methods. *Chem. Sci.* **2010**, *1* (2), 184–191.
(197) Biswas, S.; Maes, M.; Dhakshinamoorthy, A.; Feyand, M.; De Vos, D. E.; Garcia, H.; Stock, N. Fuel Purification, Lewis Acid and Aerobic Oxidation Catalysis Performed by a Microporous Co-BTT (BTT^{3-} = 1,3,5-Benzenetristetrazolate) Framework Having Coordinatively Unsaturated Sites. *J. Mater. Chem.* **2012**, *22* (20), 10200–10209.
(198) Dincă, M.; Han, W. S.; Liu, Y.; Dailly, A.; Brown, C. M.; Long, J. R. Observation of Cu^{2+}−H_2 Interactions in a Fully Desolvated Sodalite-Type Metal−Organic Framework. *Angew. Chemie Int. Ed.* **2007**, *46* (9), 1419–1422.
(199) Surblé, S.; Serre, C.; Mellot-Draznieks, C.; Millange, F.; Férey, G. A New Isoreticular Class of Metal−Organic Frameworks with the MIL-88 Topology. *Chem. Commun.* **2006**, 284–286.
(200) Zhai, Q.-G.; Bu, X.; Mao, C.; Zhao, X.; Daemen, L.; Cheng, Y.; Ramirez-Cuesta, A. J.; Feng, P. An Ultra-Tunable Platform for Molecular Engineering of High-Performance Crystalline Porous Materials. *Nat. Commun.* **2016**, *7*, 13645.

(201) Geier, S. J.; Mason, J. A.; Bloch, E. D.; Queen, W. L.; Hudson, M. R.; Brown, C. M.; Long, J. R. Selective Adsorption of Ethylene over Ethane and Propylene over Propane in the Metal−Organic Frameworks M_2(DOBDC) (M = Mg, Mn, Fe, Co, Ni, Zn). *Chem. Sci.* **2013**, *4* (5), 2054–2061.

(202) Sanz, R.; Martínez, F.; Orcajo, G.; Wojtas, L.; Briones, D. Synthesis of a Honeycomb-like Cu-Based Metal−Organic Framework and Its Carbon Dioxide Adsorption Behaviour. *Dalt. Trans.* **2013**, *42* (7), 2392–2398.

(203) Kapelewski, M. T.; Geier, S. J.; Hudson, M. R.; Stück, D.; Mason, J. A.; Nelson, J. N.; Xiao, D. J.; Hulvey, Z.; Gilmour, E.; FitzGerald, S. A.; Head-Gordon, M.; Brown, C. M.; Long, J. R. M_2(*m*-DOBDC) (M = Mg, Mn, Fe, Co, Ni) Metal−Organic Frameworks Exhibiting Increased Charge Density and Enhanced H_2 Binding at the Open Metal Sites. *J. Am. Chem. Soc.* **2014**, *136* (34), 12119–12129.

(204) Gygi, D.; Bloch, E. D.; Mason, J. A.; Hudson, M. R.; Gonzalez, M. I.; Siegelman, R. L.; Darwish, T. A.; Queen, W. L.; Brown, C. M.; Long, J. R. Hydrogen Storage in the Expanded Pore Metal−Organic Frameworks M_2(DOBPDC) (M = Mg, Mn, Fe, Co, Ni, Zn). *Chem. Mater.* **2016**, *28* (4), 1128–1138.

(205) Sun, L.; Miyakai, T.; Seki, S.; Dincă, M. Mn_2(2,5-Disulfhydrylbenzene-1,4-Dicarboxylate): A Microporous Metal–Organic Framework with Infinite $(\text{-Mn−S-})_\infty$ Chains and High Intrinsic Charge Mobility. *J. Am. Chem. Soc.* **2013**, *135* (22), 8185–8188.

(206) Sun, L.; Hendon, C. H.; Minier, M. A.; Walsh, A.; Dincă, M. Million-Fold Electrical Conductivity Enhancement in Fe_2(DEBDC) versus Mn_2(DEBDC) (E = S, O). *J. Am. Chem. Soc.* **2015**, *137* (19), 6164–6167.

(207) Brozek, C. K.; Dincă, M. Ti^{3+}-, $V^{2+/3+}$-, $Cr^{2+/3+}$-, Mn^{2+}-, and Fe^{2+}-Substituted MOF-5 and Redox Reactivity in Cr-and Fe-MOF-5. *J. Am. Chem. Soc.* **2013**, *135* (34), 12886–12891.

(208) Botas, J. A.; Calleja, G.; Sánchez-Sánchez, M.; Orcajo, M. G. Cobalt Doping of the MOF-5 Framework and Its Effect on Gas-Adsorption Properties. *Langmuir* **2010**, *26* (8), 5300–5303.

(209) Brozek, C. K.; Dincă, M. Lattice-Imposed Geometry in Metal–Organic Frameworks: Lacunary Zn_4O Clusters in MOF-5 Serve as Tripodal Chelating Ligands for Ni^{2+}. *Chem. Sci.* **2012**, *3* (6), 2110–2113.

(210) Li, H.; Eddaoudi, M.; O'Keeffe, M.; Yaghi, O. M. Design and Synthesis of an Exceptionally Stable and Highly Porous Metal−Organic Framework. *Nature* **1999**, *402* (6759), 276–279.

(211) Rossin, A.; Di Credico, B.; Giambastiani, G.; Peruzzini, M.; Pescitelli, G.; Reginato, G.; Borfecchia, E.; Gianolio, D.; Lamberti, C.; Bordiga, S. Synthesis, Characterization and CO_2 Uptake of a Chiral Co(II) Metal–Organic Framework Containing a Thiazolidine-Based Spacer. *J. Mater. Chem.* **2012**, *22* (20), 10335–10344.

(212) Denysenko, D.; Jelic, J.; Reuter, K.; Volkmer, D. Postsynthetic Metal and Ligand Exchange in MFU-4l: A Screening Approach toward Functional Metal–Organic Frameworks Comprising Single-Site Active Centers. *Chem. - A Eur. J.* **2015**, *21* (22), 8188–8199.

(213) Kholdeeva, O. A.; Skobelev, I. Y.; Ivanchikova, I. D.; Kovalenko, K. A.; Fedin, V. P.; Sorokin, A. B. Hydrocarbon Oxidation over Fe-and Cr-Containing Metal−Organic Frameworks MIL-100 and MIL-101 − a Comparative Study. *Catal. Today* **2014**, *238*, 54–61.

(214) Dhakshinamoorthy, A.; Alvaro, M.; Horcajada, P.; Gibson, E.; Vishnuvarthan, M.; Vimont,

A.; Grenèche, J.-M.; Serre, C.; Daturi, M.; Garcia, H. Comparison of Porous Iron Trimesates Basolite F300 and MIL-100 (Fe) as Heterogeneous Catalysts for Lewis Acid and Oxidation Reactions: Roles of Structural Defects and Stability. *ACS Catal.* **2012**, *2* (10), 2060–2065.

(215) Tuci, G.; Giambastiani, G.; Kwon, S.; Stair, P. C.; Snurr, R. Q.; Rossin, A. Chiral Co(II) Metal–Organic Framework in the Heterogeneous Catalytic Oxidation of Alkenes under Aerobic and Anaerobic Conditions. *ACS Catal.* **2014**, *4* (3), 1032–1039.

(216) Xiao, D. J.; Gonzalez, M. I.; Darago, L. E.; Vogiatzis, K. D.; Haldoupis, E.; Gagliardi, L.; Long, J. R. Selective, Tunable O_2 Binding in Cobalt(II)-Triazolate/Pyrazolate Metal−Organic Frameworks. *J. Am. Chem. Soc.* **2016**, *138* (22), 7161–7170.

(217) Yang, J.; Wang, Y.; Li, L.; Zhang, Z.; Li, J. Protection of Open-Metal V(III) Sites and Their Associated $CO_2/CH_4/N_2/O_2/H_2O$ Adsorption Properties in Mesoporous V-MOFs. *J. Colloid Interface Sci.* **2015**, *456*, 197–205.

(218) Karp, E. M.; Silbaugh, T. L.; Campbell, C. T. Bond Energies of Molecular Fragments to Metal Surfaces Track Their Bond Energies to H Atoms. *J. Am. Chem. Soc.* **2014**, *136* (11), 4137–4140.

(219) Liu, C.; Li, G.; Pidko, E. A. Property–Activity Relations for Methane Activation by Dual-Metal Cu–Oxo Trimers in ZSM-5 Zeolite. *Small Methods* **2018**, *2* (12), 1800266.

(220) Kurth, S.; Perdew, J. P.; Blaha, P. Molecular and Solid-State Tests of Density Functional Approximations: LSD, GGAs, and Meta-GGAs. *Int. J. Quantum Chem.* **1999**, *75* (4–5), 889–909.

(221) Jones, R. O.; Gunnarsson, O. The Density Functional Formalism, Its Applications and Prospects. *Rev. Mod. Phys.* **1989**, *61* (3), 689–746.

(222) Wellendorff, J.; Silbaugh, T. L.; Garcia-Pintos, D.; Nørskov, J. K.; Bligaard, T.; Studt, F.; Campbell, C. T. A Benchmark Database for Adsorption Bond Energies to Transition Metal Surfaces and Comparison to Selected DFT Functionals. *Surf. Sci.* **2015**, *640*, 36–44.

(223) Latimer, A. A.; Kakekhani, A.; Kulkarni, A. R.; Nørskov, J. K. Direct Methane to Methanol: The Selectivity−Conversion Limit and Design Strategies. *ACS Catal.* **2018**, *8* (8), 6894–6907.

(224) Goerigk, L.; Grimme, S. A Thorough Benchmark of Density Functional Methods for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *Phys. Chem. Chem. Phys.* **2011**, *13* (14), 6670–6688.

(225) Monkhorst, H. J.; Pack, J. D. Special Points for Brillouin-Zone Integrations. *Phys. Rev. B* **1976**, *13* (12), 5188–5192.

(226) Bowman, D. N.; Jakubikova, E. Low-Spin versus High-Spin Ground State in Pseudo-Octahedral Iron Complexes. *Inorg. Chem.* **2012**, *51* (11), 6011–6019.

(227) Reiher, M. Theoretical Study of the $Fe(Phen)_2(NCS)_2$ Spin-Crossover Complex with Reparametrized Density Functionals. *Inorg. Chem.* **2002**, *41* (25), 6928–6935.

(228) Swart, M.; Groenhof, A. R.; Ehlers, A. W.; Lammertsma, K. Validation of Exchange-Correlation Functionals for Spin States of Iron Complexes. *J. Phys. Chem. A* **2004**, *108* (25), 5479–5483.

(229) Swart, M.; Gruden, M. Spinning around in Transition-Metal Chemistry. *Acc. Chem. Res.* **2016**, *49* (12), 2690–2697.

(230) Limas, N. G.; Manz, T. A. Introducing DDEC6 Atomic Population Analysis: Part 2.

Computed Results for a Wide Range of Periodic and Nonperiodic Materials. *RSC Adv.* **2016**, *6* (51), 45727–45747.
(231) Manz, T. A. Introducing DDEC6 Atomic Population Analysis: Part 3. Comprehensive Method to Compute Bond Orders. *RSC Adv.* **2017**, *7* (72), 45552–45581.
(232) Limas, N. G.; Manz, T. A. Introducing DDEC6 Atomic Population Analysis: Part 4. Efficient Parallel Computation of Net Atomic Charges, Atomic Spin Moments, Bond Orders, and More. *RSC Adv.* **2018**, *8* (5), 2678–2707.
(233) Manz, T. A.; Sholl, D. S. Chemically Meaningful Atomic Charges That Reproduce the Electrostatic Potential in Periodic and Nonperiodic Materials. *J. Chem. Theory Comput.* **2010**, *6* (8), 2455–2468.
(234) Nazarian, D.; Camp, J. S.; Sholl, D. S. A Comprehensive Set of High-Quality Point Charges for Simulations of Metal−Organic Frameworks. *Chem. Mater.* **2016**, *28* (3), 785–793.
(235) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. Supplementary Data for "Structure−Activity Relationships that Identify Metal−Organic Framework Catalysts for Methane Activation" https://doi.org/10.5281/zenodo.1734640.
(236) Winkler, J. R.; Gray, H. B. Electronic Structures of Oxo-Metal Ions. In *Molecular Electronic Structures of Transition Metal Complexes I*; Springer, 2011; pp 17–28.
(237) Dietl, N.; Schlangen, M.; Schwarz, H. Thermal Hydrogen-Atom Transfer from Methane: The Role of Radicals and Spin States in Oxo-Cluster Chemistry. *Angew. Chemie Int. Ed.* **2012**, *51* (23), 5544–5555.
(238) Saouma, C. T.; Mayer, J. M. Do Spin State and Spin Density Affect Hydrogen Atom Transfer Reactivity? *Chem. Sci.* **2014**, *5* (1), 21–31.
(239) Mayer, J. M. Understanding Hydrogen Atom Transfer: From Bond Strengths to Marcus Theory. *Acc. Chem. Res.* **2011**, *44* (1), 36–46.
(240) Usharani, D.; Wang, B.; Sharon, D. A.; Shaik, S. Principles and Prospects of Spin-States Reactivity in Chemistry and Bioinorganic Chemistry. In *Spin States in Biochemistry and Inorganic Chemistry: Influence on Structure and Reactivity*; Swart, M., Costas, M., Eds.; John Wiley & Sons: West Sussex, U.K., 2015; pp 131–156.
(241) Shaik, S.; Chen, H.; Janardanan, D. Exchange-Enhanced Reactivity in Bond Activation by Metal−Oxo Enzymes and Synthetic Reagents. *Nat. Chem.* **2011**, *3* (1), 19–27.
(242) Wang, P.; Chang, F.; Gao, W.; Guo, J.; Wu, G.; He, T.; Chen, P. Breaking Scaling Relations to Achieve Low-Temperature Ammonia Synthesis through LiH-Mediated Nitrogen Transfer and Hydrogenation. *Nat. Chem.* **2017**, *9* (1), 64–70.
(243) Khorshidi, A.; Violet, J.; Hashemi, J.; Peterson, A. A. How Strain Can Break the Scaling Relations of Catalysis. *Nat. Catal.* **2018**, *1* (4), 263–268.
(244) Darby, M. T.; Stamatakis, M.; Michaelides, A.; Sykes, E. C. H. Lonely Atoms with Special Gifts: Breaking Linear Scaling Relationships in Heterogeneous Catalysis with Single-Atom Alloys. *J. Phys. Chem. Lett.* **2018**, *9* (18), 5636–5646.
(245) Valvekens, P.; Vermoortele, F.; De Vos, D. Metal–Organic Frameworks as Catalysts: The Role of Metal Active Sites. *Catal. Sci. Technol.* **2013**, *3* (6), 1435–1445.
(246) Luz, I.; León, A.; Boronat, M.; Llabrés i Xamena, F. X.; Corma, A. Selective Aerobic Oxidation of Activated Alkanes with MOFs and Their Use for Epoxidation of Olefins with Oxygen in a Tandem Reaction. *Catal. Sci. Technol.* **2013**, *3* (2), 371–379.
(247) Leus, K.; Vandichel, M.; Liu, Y.-Y.; Muylaert, I.; Musschoot, J.; Pyl, S.; Vrielinck, H.;

Callens, F.; Marin, G. B.; Detavernier, C.; Wiper, P. V; Khimyak, Y. Z.; Waroquier, M.; van Speybroeck, V.; van der Voort, P. The Coordinatively Saturated Vanadium MIL-47 as a Low Leaching Heterogeneous Catalyst in the Oxidation of Cyclohexene. *J. Catal.* **2012**, *285* (1), 196–207.

(248) Lyaskovskyy, V.; de Bruin, B. Redox Non-Innocent Ligands: Versatile New Tools to Control Catalytic Reactions. *ACS Catal.* **2012**, *2* (2), 270–279.

(249) Zhang, X.; Vermeulen, N. A.; Huang, Z.; Cui, Y.; Liu, J.; Krzyaniak, M. D.; Li, Z.; Noh, H.; Wasielewski, M. R.; Delferro, M.; Farha, O. K. Effect of Redox "Non-Innocent" Linker on the Catalytic Activity of Copper-Catecholate-Decorated Metal–Organic Frameworks. *ACS Appl. Mater. Interfaces* **2018**, *10* (1), 635–641.

(250) Gunay, A.; Theopold, K. H. C−H Bond Activations by Metal Oxo Compounds. *Chem. Rev.* **2010**, *110* (2), 1060–1081.

(251) Borovik, A. S. Role of Metal−Oxo Complexes in the Cleavage of C−H Bonds. *Chem. Soc. Rev.* **2011**, *40* (4), 1870–1874.

(252) Shimoyama, Y.; Kojima, T. Metal−Oxyl Species and Their Possible Roles in Chemical Oxidations. *Inorg. Chem.* **2019**, *58* (15), 9517–9542.

(253) Jackson, R. B.; Solomon, E. I.; Canadell, J. G.; Cargnello, M.; Field, C. B. Methane Removal and Atmospheric Restoration. *Nat. Sustain.* **2019**, *2*, 436–438.

(254) Engelmann, X.; Monte-Pérez, I.; Ray, K. Oxidation Reactions with Bioinspired Mononuclear Non-Heme Metal−Oxo Complexes. *Angew. Chemie Int. Ed.* **2016**, *55* (27), 7632–7649.

(255) Limberg, C. What Does It Really Take to Stabilize Complexes of Late Transition Metals with Terminal Oxo Ligands? *Angew. Chemie Int. Ed.* **2009**, *48* (13), 2270–2273.

(256) Ray, K.; Heims, F.; Pfaff, F. F. Terminal Oxo and Imido Transition-Metal Complexes of Groups 9−11. *Eur. J. Inorg. Chem.* **2013**, *2013* (22–23), 3784–3807.

(257) Andris, E.; Navrátil, R.; Jašík, J.; Srnec, M.; Rodríguez, M.; Costas, M.; Roithová, J. M−O Bonding Beyond the Oxo Wall: Spectroscopy and Reactivity of Cobalt(III)-Oxyl and Cobalt(III)-Oxo Complexes. *Angew. Chemie. Int. Ed.* **2019**, *131* (28), 9721–9726.

(258) Puri, M.; Que Jr, L. Toward the Synthesis of More Reactive $S = 2$ Non-Heme Oxoiron(IV) Complexes. *Acc. Chem. Res.* **2015**, *48* (8), 2443–2452.

(259) Wasson, M. C.; Buru, C. T.; Chen, Z.; Islamoglu, T.; Farha, O. K. Metal–Organic Frameworks: A Tunable Platform to Access Single-Site Heterogeneous Catalysts. *Appl. Catal. A Gen.* **2019**, 117214.

(260) Hirao, H.; Ng, W. K. H.; Moeljadi, A. M. P.; Bureekaew, S. Multiscale Model for a Metal−Organic Framework: High-Spin Rebound Mechanism in the Reaction of the Oxoiron(IV) Species of Fe-MOF-74. *ACS Catal.* **2015**, *5* (6), 3287–3291.

(261) Luo, W.; Chen, G.; Xiao, S.; Wang, Q.; Huang, Z.; Wang, L. The Enzyme-like Catalytic Hydrogen Abstraction Reaction Mechanisms of Cyclic Hydrocarbons with Magnesium-Diluted Fe-MOF-74. *RSC Adv.* **2019**, *9* (41), 23622–23632.

(262) Stubbs, A. W.; Braglia, L.; Borfecchia, E.; Meyer, R. J.; Román-Leshkov, Y.; Lamberti, C.; Dincă, M. Selective Catalytic Olefin Epoxidation with Mn^{II}-Exchanged MOF-5. *ACS Catal.* **2017**, *8* (1), 596–601.

(263) Ketrat, S.; Maihom, T.; Wannakao, S.; Probst, M.; Nokbin, S.; Limtrakul, J. Coordinatively Unsaturated Metal−Organic Frameworks $M_3(BTC)_2$ (M = Cr, Fe, Co, Ni, Cu, and Zn)

Catalyzing the Oxidation of CO by N_2O: Insight from DFT Calculations. *Inorg. Chem.* **2017**, *56* (22), 14005–14012.

(264) Barona, M.; Ahn, S.; Morris, W.; Hoover, W. J.; Notestein, J. M.; Farha, O. K.; Snurr, R. Q. Computational Predictions and Experimental Validation of Alkane Oxidative Dehydrogenation by Fe_2M MOF Nodes. *ACS Catal.* **2020**, *10* (2), 1460–1469.

(265) Simons, M. C.; Vitillo, J. G.; Babucci, M.; Hoffman, A. S.; Boubnov, A.; Beauvais, M. L.; Chen, Z.; Cramer, C. J.; Chapman, K. W.; Bare, S. R.; Gates, B. C.; Lu, C. C.; Gagliardi, L.; Bhan, A. Structure, Dynamics, and Reactivity for Light Alkane Oxidation of Fe(II) Sites Situated in the Nodes of a Metal–Organic Framework. *J. Am. Chem. Soc.* **2019**, *141* (45), 18142–18151.

(266) Zhang, J.-P.; Zhang, Y.-B.; Lin, J.-B.; Chen, X.-M. Metal Azolate Frameworks: From Crystal Engineering to Functional Materials. *Chem. Rev.* **2011**, *112* (2), 1001–1033.

(267) Desai, A. V; Sharma, S.; Let, S.; Ghosh, S. K. N-Donor Linker Based Metal−Organic Frameworks (MOFs): Advancement and Prospects as Functional Materials. *Coord. Chem. Rev.* **2019**, *395*, 146–192.

(268) Prokop, K. A.; de Visser, S. P.; Goldberg, D. P. Unprecedented Rate Enhancements of Hydrogen-Atom Transfer to a Manganese (V)−Oxo Corrolazine Complex. *Angew. Chemie Int. Ed.* **2010**, *49* (30), 5091–5095.

(269) Sastri, C. V; Lee, J.; Oh, K.; Lee, Y. J.; Lee, J.; Jackson, T. A.; Ray, K.; Hirao, H.; Shin, W.; Halfen, J. A.; Kim, J.; Que Jr., L.; Shaik, S.; Nam, W. Axial Ligand Tuning of a Nonheme Iron(IV)–Oxo Unit for Hydrogen Atom Abstraction. *Proc. Natl. Acad. Sci.* **2007**, *104* (49), 19181–19186.

(270) Reed, D. A.; Xiao, D. J.; Gonzalez, M. I.; Darago, L. E.; Herm, Z. R.; Grandjean, F.; Long, J. R. Reversible CO Scavenging via Adsorbate-Dependent Spin State Transitions in an Iron(II)-Triazolate Metal−Organic Framework. *J. Am. Chem. Soc.* **2016**, *138* (17), 5594–5602.

(271) Liao, P.-Q.; Li, X.-Y.; Bai, J.; He, C.-T.; Zhou, D.-D.; Zhang, W.-X.; Zhang, J.-P.; Chen, X.-M. Drastic Enhancement of Catalytic Activity via Post-Oxidation of a Porous Mn^{II} Triazolate Framework. *Chem. Eur. J.* **2014**, *20* (36), 11303–11307.

(272) Denysenko, D.; Grzywa, M.; Tonigold, M.; Streppel, B.; Krkljus, I.; Hirscher, M.; Mugnaioli, E.; Kolb, U.; Hanss, J.; Volkmer, D. Elucidating Gating Effects for Hydrogen Sorption in MFU-4-Type Triazolate-Based Metal−Organic Frameworks Featuring Different Pore Sizes. *Chem. Eur. J.* **2011**, *17* (6), 1837–1848.

(273) Rieth, A. J.; Wright, A. M.; Skorupskii, G.; Mancuso, J. L.; Hendon, C. H.; Dincă, M. Record-Setting Sorbents for Reversible Water Uptake by Systematic Anion-Exchanges in Metal−Organic Frameworks. *J. Am. Chem. Soc.* **2019**, *141* (35), 13858–13866.

(274) Sumida, K.; Stück, D.; Mino, L.; Chai, J.-D.; Bloch, E. D.; Zavorotynska, O.; Murray, L. J.; Dincă, M.; Chavan, S.; Bordiga, S.; Head-Gordon, M.; Long, J. R. Impact of Metal and Anion Substitutions on the Hydrogen Storage Properties of M-BTT Metal–Organic Frameworks. *J. Am. Chem. Soc.* **2013**, *135* (3), 1083–1091.

(275) Wang, Y.; Huang, N.-Y.; Shen, J.-Q.; Liao, P.-Q.; Chen, X.-M.; Zhang, J.-P. Hydroxide Ligands Cooperate with Catalytic Centers in Metal−Organic Frameworks for Efficient Photocatalytic CO2 Reduction. *J. Am. Chem. Soc.* **2018**, *140* (1), 38–41.

(276) Rieth, A. J.; Wright, A. M.; Dincă, M. Kinetic Stability of Metal–Organic Frameworks for

Corrosive and Coordinating Gas Capture. *Nat. Rev. Mater.* **2019**, *4*, 708–725.
(277) Lu, X.-F.; Liao, P.-Q.; Wang, J.-W.; Wu, J.-X.; Chen, X.-W.; He, C.-T.; Zhang, J.-P.; Li, G.-R.; Chen, X.-M. An Alkaline-Stable, Metal Hydroxide Mimicking Metal−Organic Framework for Efficient Electrocatalytic Oxygen Evolution. *J. Am. Chem. Soc.* **2016**, *138* (27), 8336–8339.
(278) Park, S. S.; Tulchinsky, Y.; Dincă, M. Single-Ion Li+, Na+, and Mg2+ Solid Electrolytes Supported by a Mesoporous Anionic Cu–Azolate Metal–Organic Framework. *J. Am. Chem. Soc.* **2017**, *139* (38), 13260–13263.
(279) Jaramillo, D. E.; Reed, D. A.; Jiang, H. Z. H.; Oktawiec, J.; Mara, M. W.; Forse, A. C.; Lussier, D. J.; Murphy, R. A.; Cunningham, M.; Colombo, V.; Shuh, D. K.; Reimer, J. A.; Long, J. R. Selective Nitrogen Adsorption via Backbonding in a Metal–Organic Framework with Exposed Vanadium Sites. *Nat. Mater.* **2020**.
(280) Rieth, A. J.; Dincă, M. Controlled Gas Uptake in Metal–Organic Frameworks with Record Ammonia Sorption. *J. Am. Chem. Soc.* **2018**, *140* (9), 3461–3466.
(281) Vogiatzis, K. D.; Polynski, M. V; Kirkland, J. K.; Townsend, J.; Hashemi, A.; Liu, C.; Pidko, E. A. Computational Approach to Molecular Catalysis by 3d Transition Metals: Challenges and Opportunities. *Chem. Rev.* **2019**, *119* (4), 2453–2523.
(282) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16 Revision C.01. 2016.
(283) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652.
(284) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785.
(285) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F. N.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98* (45), 11623–11627.
(286) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. *Phys. Chem. Chem. Phys.* **2005**, *7* (18), 3297.
(287) Zhao, Y.; Truhlar, D. G. A New Local Density Functional for Main-Group Thermochemistry, Transition Metal Bonding, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Phys.* **2006**, *125*, 194101.
(288) Zhao, Y.; Truhlar, D. G. The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and

Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Function. *Theor. Chem. Acc.* **2008**, *120* (1–3), 215–241.
(289) Marenich, A. V; Jerome, S. V; Cramer, C. J.; Truhlar, D. G. Charge Model 5: An Extension of Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in Gaseous and Condensed Phases. *J. Chem. Theory Comput.* **2012**, *8* (2), 527–541.
(290) Hirshfeld, F. L. Bonded-Atom Fragments for Describing Molecular Charge Densities. *Theor. Chim. Acta* **1977**, *44* (2), 129–138.
(291) Weinhold, F.; Landis, C. R. *Valency and Bonding: A Natural Bond Orbital Donor-Acceptor Perspective*; Cambridge University Press, 2005.
(292) Glendening, E. D.; Badenhoop, J. K.; Reed, A. E.; Carpenter, J. E.; Bohmann, J. A.; Morales, C. M.; Karafiloglou, P.; Landis, C. R.; Weinhold, F. NBO 7.0. Theoretical Chemistry Institute, University of Wisconsin, Madison 2018.
(293) Cramer, C. *Essentials of Computational Chemistry*, 2nd ed.; John Wiley & Sons: West Sussex, U.K., 2004.
(294) GoodVibes https://doi.org/10.5281/zenodo.595246.
(295) Li, Y.-P.; Gomes, J.; Mallikarjun Sharada, S.; Bell, A. T.; Head-Gordon, M. Improved Force-Field Parameters for QM/MM Simulations of the Energies of Adsorption for Molecules in Zeolites and a Free Rotor Correction to the Rigid Rotor Harmonic Oscillator Model for Adsorption Enthalpies. *J. Phys. Chem. C* **2015**, *119* (4), 1840–1850.
(296) Kanchanakungwankul, S.; Bao, J. L.; Zheng, J.; Alecu, I. M.; Lynch, J.; Zhao, Y.; Truhlar, D. G. Database of Frequency Scale Factors for Electronic Model Chemistries Version 4. https://comp.chem.umn.edu/freqscale/190107_Database_of_Freq_Scale_Factors_v4.pdf.
(297) Deshlahra, P.; Iglesia, E. Reactivity and Selectivity Descriptors for the Activation of C−H Bonds in Hydrocarbons and Oxygenates on Metal Oxides. *J. Phys. Chem. C* **2016**, *120* (30), 16741–16760.
(298) Walsh, A.; Sokol, A. A.; Buckeridge, J.; Scanlon, D. O.; Catlow, C. R. A. Oxidation States and Ionicity. *Nat. Mater.* **2018**, *17* (11), 958–964.
(299) Ravi, M.; Sushkevich, V. L.; Knorpp, A. J.; Newton, M. A.; Palagin, D.; Pinar, A. B.; Ranocchiari, M.; Van Bokhoven, J. A. Misconceptions and Challenges in Methane-to-Methanol over Transition-Metal-Exchanged Zeolites. *Nat. Catal.* **2019**, *2* (6), 485–494.
(300) Lai, W.; Li, C.; Chen, H.; Shaik, S. Hydrogen-Abstraction Reactivity Patterns from A to Y: The Valence Bond Way. *Angew. Chemie Int. Ed.* **2012**, *51* (23), 5556–5578.
(301) Ye, S.; Neese, F. Nonheme Oxo-Iron(IV) Intermediates Form an Oxyl Radical upon Approaching the C−H Bond Activation Transition State. *Proc. Natl. Acad. Sci.* **2011**, *108* (4), 1228–1233.
(302) Ye, S.; Geng, C.-Y.; Shaik, S.; Neese, F. Electronic Structure Analysis of Multistate Reactivity in Transition Metal Catalyzed Reactions: The Case of C–H Bond Activation by Non-Heme Iron(IV)–Oxo Cores. *Phys. Chem. Chem. Phys.* **2013**, *15* (21), 8017–8030.
(303) Schröder, D.; Shaik, S.; Schwarz, H. Two-State Reactivity as a New Concept in Organometallic Chemistry. *Acc. Chem. Res.* **2000**, *33* (3), 139–145.
(304) Dzik, W. I.; Böhmer, W.; de Bruin, B.; Swart, M.; Costas, M. Multiple Spin-State Scenarios in Organometallic Reactivity. In *Spin States in Biochemistry and Inorganic Chemistry: Influence on Structure and Reactivity*; Swart, M., Costas, M., Eds.; John Wiley & Sons: West Sussex, U.K., 2015; pp 103–129.

(305) Geng, C.; Ye, S.; Neese, F. Analysis of Reaction Channels for Alkane Hydroxylation by Nonheme Iron(IV)–Oxo Complexes. *Angew. Chemie Int. Ed.* **2010**, *49* (33), 5717–5720.
(306) Usharani, D.; Janardanan, D.; Li, C.; Shaik, S. A Theory for Bioinorganic Chemical Reactivity of Oxometal Complexes and Analogous Oxidants: The Exchange and Orbital-Selection Rules. *Acc. Chem. Res.* **2012**, *46* (2), 471–482.
(307) Janardanan, D.; Wang, Y.; Schyman, P.; Que Jr, L.; Shaik, S. The Fundamental Role of Exchange-Enhanced Reactivity in C–H Activation by S=2 Oxo Iron(IV) Complexes. *Angew. Chemie Int. Ed.* **2010**, *49* (19), 3342–3345.
(308) Klotz, I. M.; Kurtz Jr, D. M. Metal-Dioxygen Complexes: A Perspective. *Chem. Rev.* **1994**, *94* (3), 567–568.
(309) Li, G. Q.; Govind, R. Separation of Oxygen from Air Using Coordination Complexes: A Review. *Ind. Eng. Chem. Res.* **1994**, *33* (4), 755–783.
(310) Holland, P. L. Metal–Dioxygen and Metal–Dinitrogen Complexes: Where Are the Electrons? *Dalt. Trans.* **2010**, *39* (23), 5415–5425.
(311) Que Jr, L.; Tolman, W. B. Biologically Inspired Oxidation Catalysis. *Nature* **2008**, *455*, 333.
(312) MacKay, B. A.; Fryzuk, M. D. Dinitrogen Coordination Chemistry: On the Biomimetic Borderlands. *Chem. Rev.* **2004**, *104* (2), 385–402.
(313) Fryzuk, M. D.; Johnson, S. A. The Continuing Story of Dinitrogen Activation. *Coord. Chem. Rev.* **2000**, *200*, 379–409.
(314) Rufford, T. E.; Smart, S.; Watson, G. C. Y.; Graham, B. F.; Boxall, J.; Da Costa, J. C. D.; May, E. F. The Removal of CO_2 and N_2 from Natural Gas: A Review of Conventional and Emerging Process Technologies. *J. Pet. Sci. Eng.* **2012**, *94*, 123–154.
(315) Yang, R. T. *Adsorbents: Fundamentals and Applications*; John Wiley & Sons: Hoboken, NJ, 2003.
(316) Smith, A. R.; Klosek, J. A Review of Air Separation Technologies and Their Integration with Energy Conversion Processes. *Fuel Process. Technol.* **2001**, *70* (2), 115–134.
(317) Nugent, P.; Belmabkhout, Y.; Burd, S. D.; Cairns, A. J.; Luebke, R.; Forrest, K.; Pham, T.; Ma, S.; Space, B.; Wojtas, L.; Eddaoudi, M.; Zaworotko, M. J. Porous Materials with Optimal Adsorption Thermodynamics and Kinetics for CO_2 Separation. *Nature* **2013**, *495* (7439), 80.
(318) Li, J.-R.; Kuppler, R. J.; Zhou, H.-C. Selective Gas Adsorption and Separation in Metal−Organic Frameworks. *Chem. Soc. Rev.* **2009**, *38* (5), 1477–1504.
(319) Duan, J.; Higuchi, M.; Krishna, R.; Kiyonaga, T.; Tsutsumi, Y.; Sato, Y.; Kubota, Y.; Takata, M.; Kitagawa, S. High $CO_2/N_2/O_2/CO$ Separation in a Chemically Robust Porous Coordination Polymer with Low Binding Energy. *Chem. Sci.* **2014**, *5* (2), 660–666.
(320) Bao, Z.; Chang, G.; Xing, H.; Krishna, R.; Ren, Q.; Chen, B. Potential of Microporous Metal−Organic Frameworks for Separation of Hydrocarbon Mixtures. *Energy Environ. Sci.* **2016**, *9* (12), 3612–3641.
(321) Lee, J. S.; Vlaisavljevich, B.; Britt, D. K.; Brown, C. M.; Haranczyk, M.; Neaton, J. B.; Smit, B.; Long, J. R.; Queen, W. L. Understanding Small-Molecule Interactions in Metal−Organic Frameworks: Coupling Experiment with Theory. *Adv. Mater.* **2015**, *27* (38), 5785–5796.
(322) Southon, P. D.; Price, D. J.; Nielsen, P. K.; McKenzie, C. J.; Kepert, C. J. Reversible and

Selective O_2 Chemisorption in a Porous Metal–Organic Host Material. *J. Am. Chem. Soc.* **2011**, *133* (28), 10885–10891.

(323) Bae, Y.-S.; Spokoyny, A. M.; Farha, O. K.; Snurr, R. Q.; Hupp, J. T.; Mirkin, C. A. Separation of Gas Mixtures Using Co(II) Carborane-Based Porous Coordination Polymers. *Chem. Commun.* **2010**, *46* (20), 3478–3480.

(324) Sava Gallis, D. F.; Parkes, M. V; Greathouse, J. A.; Zhang, X.; Nenoff, T. M. Enhanced O_2 Selectivity versus N_2 by Partial Metal Substitution in Cu-BTC. *Chem. Mater.* **2015**, *27* (6), 2018–2025.

(325) Reed, D. A.; Xiao, D. J.; Jiang, H. Z. H.; Chakarawet, K.; Oktawiec, J.; Long, J. R. Biomimetic O_2 Adsorption in an Iron Metal–Organic Framework for Air Separation. *Chem. Sci.* **2020**.

(326) Rosnes, M. H.; Sheptyakov, D.; Franz, A.; Frontzek, M.; Dietzel, P. D. C.; Georgiev, P. A. On the Elusive Nature of Oxygen Binding at Coordinatively Unsaturated 3d Transition Metal Centers in Metal–Organic Frameworks. *Phys. Chem. Chem. Phys.* **2017**, *19* (38), 26346–26357.

(327) Verma, P.; Maurice, R.; Truhlar, D. G. Identifying the Interactions That Allow Separation of O_2 from N_2 on the Open Iron Sites of Fe_2(DOBDC). *J. Phys. Chem. C* **2015**, *119* (51), 28499–28511.

(328) Moeljadi, A. M. P.; Schmid, R.; Hirao, H. Dioxygen Binding to Fe-MOF-74: Microscopic Insights from Periodic QM/MM Calculations. *Can. J. Chem.* **2016**, *94* (12), 1144–1150.

(329) Parkes, M. V; Greathouse, J. A.; Hart, D. B.; Gallis, D. F. S.; Nenoff, T. M. Ab Initio Molecular Dynamics Determination of Competitive O_2 vs. N_2 Adsorption at Open Metal Sites of M_2(DOBDC). *Phys. Chem. Chem. Phys.* **2016**, *18* (16), 11528–11538.

(330) Parkes, M. V; Sava Gallis, D. F.; Greathouse, J. A.; Nenoff, T. M. Effect of Metal in $M_3(BTC)_2$ and M_2(DOBDC) MOFs for O_2/N_2 Separations: A Combined Density Functional Theory and Experimental Study. *J. Phys. Chem. C* **2015**, *119* (12), 6556–6567.

(331) Wang, Y.; Yang, J.; Li, Z.; Zhang, Z.; Li, J.; Yang, Q.; Zhong, C. Computational Study of Oxygen Adsorption in Metal−Organic Frameworks with Exposed Cation Sites: Effect of Framework Metal Ions. *RSC Adv.* **2015**, *5* (42), 33432–33437.

(332) Rosi, N. L.; Kim, J.; Eddaoudi, M.; Chen, B.; O'Keeffe, M.; Yaghi, O. M. Rod Packings and Metal−Organic Frameworks Constructed from Rod-Shaped Secondary Building Units. *J. Am. Chem. Soc.* **2005**, *127* (5), 1504–1518.

(333) Feng, D.; Wang, K.; Wei, Z.; Chen, Y.-P.; Simon, C. M.; Arvapally, R. K.; Martin, R. L.; Bosch, M.; Liu, T.-F.; Fordham, S.; Yuan, D.; Omary, M. A.; Haranczyk, M.; Smit, B.; Zhou, H.-C. Kinetically Tuned Dimensional Augmentation as a Versatile Synthetic Route towards Robust Metal–Organic Frameworks. *Nat. Commun.* **2014**, *5*, 5723.

(334) Tanabe, Y.; Nishibayashi, Y. Overviews of the Preparation and Reactivity of Transition Metal–Dinitrogen Complexes. In *Transition Metal-Dinitrogen Complexes: Preparation and Reactivity*; Nishibayashi, Y., Ed.; Wiley, 2019.

(335) Zimmermann, N. E. R.; Jain, A. Local Structure Order Parameters and Site Fingerprints for Quantification of Coordination Environment and Crystal Structure Similarity. *RSC Adv.* **2019**, *10* (10), 6063–6081.

(336) Pan, H.; Ganose, A. M.; Horton, M.; Aykol, M.; Persson, K. A.; Zimmermann, N. E. R.; Jain, A. Benchmarking Coordination Number Prediction Algorithms on Inorganic Crystal

Structures. *Inorg. Chem.* **2020**, *60* (3), 1590–1603.
(337) Recommended PAW potentials for DFT calculations using vasp.5.2 https://cms.mpi.univie.ac.at/vasp/vasp/Recommended_PAW_potentials_DFT_calculations_using_vasp_5_2.html.
(338) Freysoldt, C.; Boeck, S.; Neugebauer, J. Direct Minimization Technique for Metals in Density Functional Theory. *Phys. Rev. B* **2009**, *79* (24), 241103.
(339) Dudarev, S. L.; Botton, G. A.; Savrasov, S. Y.; Humphreys, C. J.; Sutton, A. P. Electron-Energy-Loss Spectra and the Structural Stability of Nickel Oxide: An LSDA+U Study. *Phys. Rev. B* **1998**, *57* (3), 1505.
(340) Wang, L.; Maxisch, T.; Ceder, G. Oxidation Energies of Transition Metal Oxides within the GGA+U Framework. *Phys. Rev. B* **2006**, *73* (19), 195107.
(341) Lee, K.; Isley III, W. C.; Dzubak, A. L.; Verma, P.; Stoneburner, S. J.; Lin, L.-C.; Howe, J. D.; Bloch, E. D.; Reed, D. A.; Hudson, M. R.; Brown, C. M.; Long, J. R.; Neaton, J. B.; Smit, B.; Cramer, C. J.; Truhlar, D. G.; Gagliardi, L. Design of a Metal–Organic Framework with Enhanced Back Bonding for Separation of N2 and CH4. *J. Am. Chem. Soc.* **2014**, *136* (2), 698–704.
(342) Comito, R. J.; Wu, Z.; Zhang, G.; Lawrence III, J. A.; Korzyński, M. D.; Kehl, J. A.; Miller, J. T.; Dincă, M. Stabilized Vanadium Catalyst for Olefin Polymerization by Site Isolation in a Metal−Organic Framework. *Angew. Chemie. Int. Ed.* **2018**, *57* (27), 8135–8139.
(343) Comito, R. J.; Fritzsching, K. J.; Sundell, B. J.; Schmidt-Rohr, K.; Dincă, M. Single-Site Heterogeneous Catalysts for Olefin Polymerization Enabled by Cation Exchange in a Metal-Organic Framework. *J. Am. Chem. Soc.* **2016**, *138* (32), 10232–10237.
(344) Niederhoffer, E. C.; Timmons, J. H.; Martell, A. E. Thermodynamics of Oxygen Binding in Natural and Synthetic Dioxygen Complexes. *Chem. Rev.* **1984**, *84* (2), 137–203.
(345) Chen, L. S.; Koehler, M. E.; Pestel, B. C.; Cummings, S. C. Thermodynamic Studies on (Thio Iminato)Cobalt(II) Oxygen Carriers and the Effect of Chelate Ring Substituents. *J. Am. Chem. Soc.* **1978**, *100* (23), 7243–7248.
(346) Lukin, J. A.; Simplaceanu, V.; Zou, M.; Ho, N. T.; Ho, C. NMR Reveals Hydrogen Bonds between Oxygen and Distal Histidines in Oxyhemoglobin. *Proc. Natl. Acad. Sci.* **2000**, *97* (19), 10354–10358.
(347) Dube, H.; Kasumaj, B.; Calle, C.; Saito, M.; Jeschke, G.; Diederich, F. Direct Evidence for a Hydrogen Bond to Bound Dioxygen in a Myoglobin/Hemoglobin Model System and in Cobalt Myoglobin by Pulse-EPR Spectroscopy. *Angew. Chemie Int. Ed.* **2008**, *47* (14), 2600–2603.
(348) Vad, M. S.; Johansson, F. B.; Seidler-Egdal, R. K.; McGrady, J. E.; Novikov, S. M.; Bozhevolnyi, S. I.; Bond, A. D.; McKenzie, C. J. Tuning Affinity and Reversibility for O_2 Binding in Dinuclear Co(II) Complexes. *Dalt. Trans.* **2013**, *42* (27), 9921–9929.
(349) Hartley, F. R. The Cis-and Trans-Effects of Ligands. *Chem. Soc. Rev.* **1973**, *2* (2), 163–179.
(350) Appleton, T. G.; Clark, H. C.; Manzer, L. E. The Trans-Influence: Its Measurement and Significance. *Coord. Chem. Rev.* **1973**, *10* (3–4), 335–422.
(351) Coe, B. J.; Glenwright, S. J. Trans-Effects in Octahedral Transition Metal Complexes. *Coord. Chem. Rev.* **2000**, *203* (1), 5–80.
(352) Zhang, G.; Chen, K.; Chen, H.; Yao, J.; Shaik, S. What Factors Control O_2 Binding and Release Thermodynamics in Mononuclear Ruthenium Water Oxidation Catalysts? A

Theoretical Exploration. *Inorg. Chem.* **2013**, *52* (9), 5088–5096.
(353) Ripin, D. H.; Evans, D. A. Evans pKa Table http://evans.rc.fas.harvard.edu/pdf/evans_pKa_table.pdf.
(354) Silverstein, T. P.; Heller, S. T. PKa Values in the Undergraduate Curriculum: What Is the Real PKa of Water? *J. Chem. Educ.* **2017**, *94* (6), 690–695.
(355) Horcajada, P.; Serre, C.; Vallet-Regí, M.; Sebban, M.; Taulelle, F.; Férey, G. Metal–Organic Frameworks as Efficient Materials for Drug Delivery. *Angew. Chemie. Int. Ed.* **2006**, *118* (36), 6120–6124.
(356) Yoon, J. W.; Seo, Y.-K.; Hwang, Y. K.; Chang, J.-S.; Leclerc, H.; Wuttke, S.; Bazin, P.; Vimont, A.; Daturi, M.; Bloch, E.; Llewellyn, P. L.; Serre, C.; Horcajada, P.; Grenèche, J.-M.; Rodrigues, A. E.; Férey, G. Controlled Reducibility of a Metal−Organic Framework with Coordinatively Unsaturated Sites for Preferential Gas Sorption. *Angew. Chemie Int. Ed.* **2010**, *49* (34), 5949–5952.
(357) Yoon, J. W.; Chang, H.; Lee, S.-J.; Hwang, Y. K.; Hong, D.-Y.; Lee, S.-K.; Lee, J. S.; Jang, S.; Yoon, T.-U.; Kwac, K.; Jung, Y.; Pillai, R.; Faucher, F.; Vimont, A.; Daturi, M.; Férey, G.; Serre, C.; Maurin, G.; Bae, Y.-S.; Chang, J.-S. Selective Nitrogen Capture by Porous Hybrid Materials Containing Accessible Transition Metal Ion Sites. *Nat. Mater.* **2017**, *16* (5), 526.
(358) Shannon, R. D. Revised Effective Ionic Radii and Systematic Studies of Interatomic Distances in Halides and Chalcogenides. *Acta Crystallogr. Sect. A* **1976**, *32* (5), 751–767.
(359) Myers, A. L.; Prausnitz, J. M. Thermodynamics of Mixed-Gas Adsorption. *AIChE J.* **1965**, *11* (1), 121–127.
(360) Simon, C. M.; Smit, B.; Haranczyk, M. PyIAST: Ideal Adsorbed Solution Theory (IAST) Python Package. *Comput. Phys. Commun.* **2016**, *200*, 364–380.
(361) Sava Gallis, D. F.; Chapman, K. W.; Rodriguez, M. A.; Greathouse, J. A.; Parkes, M. V; Nenoff, T. M. Selective O_2 Sorption at Ambient Temperatures via Node Distortions in Sc-MIL-100. *Chem. Mater.* **2016**, *28* (10), 3327–3336.
(362) Mu, B.; Schoenecker, P. M.; Walton, K. S. Gas Adsorption Study on Mesoporous Metal-Organic Framework UMCM-1. *J. Phys. Chem. C* **2010**, *114* (14), 6464–6471.
(363) Li, Y.; Yang, R. T. Gas Adsorption and Storage in Metal−organic Framework MOF-177. *Langmuir* **2007**, *23* (26), 12937–12944.
(364) Jelic, J.; Denysenko, D.; Volkmer, D.; Reuter, K. Computational Screening Study towards Redox-Active Metal−Organic Frameworks. *New J. Phys.* **2013**, *15* (11), 115004.
(365) Hall, J. N.; Bollini, P. Structure, Characterization, and Catalytic Properties of Open-Metal Sites in Metal Organic Frameworks. *React. Chem. Eng.* **2019**, *4* (2), 207–222.
(366) Burke, K. Perspective on Density Functional Theory. *J. Chem. Phys.* **2012**, *136* (15), 150901.
(367) Mori-Sánchez, P.; Cohen, A. J.; Yang, W. Many-Electron Self-Interaction Error in Approximate Density Functionals. *J. Chem. Phys.* **2006**, *125*, 201102.
(368) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Insights into Current Limitations of Density Functional Theory. *Science* **2008**, *321* (5890), 792–794.
(369) Mann, G. W.; Lee, K.; Cococcioni, M.; Smit, B.; Neaton, J. B. First-Principles Hubbard U Approach for Small Molecule Binding in Metal-Organic Frameworks. *J. Chem. Phys.* **2016**, *144* (17), 174104.

(370) Lee, K.; Howe, J. D.; Lin, L.-C.; Smit, B.; Neaton, J. B. Small-Molecule Adsorption in Open-Site Metal−Organic Frameworks: A Systematic Density Functional Theory Study for Rational Design. *Chem. Mater.* **2015**, *27* (3), 668–678.
(371) Ji, H.; Park, J.; Cho, M.; Jung, Y. Assessments of Semilocal Density Functionals and Corrections for Carbon Dioxide Adsorption on Metal−Organic Frameworks. *ChemPhysChem* **2014**, *15* (15), 3157–3165.
(372) Dietzel, P. D. C.; Panella, B.; Hirscher, M.; Blom, R.; Fjellvåg, H. Hydrogen Adsorption in a Nickel Based Coordination Polymer with Open Metal Sites in the Cylindrical Cavities of the Desolvated Framework. *Chem. Commun.* **2006**, No. 9, 959–961.
(373) Kulik, H. J. Perspective: Treating Electron Over-Delocalization with the DFT+U Method. *J. Chem. Phys.* **2015**, *142* (24), 240901.
(374) Manz, T. A.; Gabaldon Limas, N. Chargemol program for performing DDEC analysis http://ddec.sourceforge.net/.
(375) http://doi.org/10.5281/zenodo.2652475.
(376) Zhao, Q.; Kulik, H. J. Where Does the Density Localize in the Solid State? Divergent Behavior for Hybrids and DFT+U. *J. Chem. Theory Comput.* **2018**, *14* (2), 670–683.
(377) Zhang, Q.; Li, B.; Chen, L. First-Principles Study of Microporous Magnets M-MOF-74 (M= Ni, Co, Fe, Mn): The Role of Metal Centers. *Inorg. Chem.* **2013**, *52* (16), 9356–9362.
(378) Scherlis, D. A.; Cococcioni, M.; Sit, P.; Marzari, N. Simulation of Heme Using DFT+U: A Step toward Accurate Spin-State Energetics. *J. Phys. Chem. B* **2007**, *111* (25), 7384–7391.
(379) Verma, P.; Truhlar, D. G. Does DFT+U Mimic Hybrid Density Functionals? *Theor. Chem. Acc.* **2016**, *135* (8), 182.
(380) Gani, T. Z. H.; Kulik, H. J. Where Does the Density Localize? Convergent Behavior for Global Hybrids, Range Separation, and DFT+U. *J. Chem. Theory Comput.* **2016**, *12* (12), 5931–5945.
(381) Stoneburner, S. J.; Gagliardi, L. Air Separation by Catechol-Ligated Transition Metals: A Quantum Chemical Screening. *J. Phys. Chem. C* **2018**, *122* (39), 22345–22351.
(382) Jensen, K. P. Bioinorganic Chemistry Modeled with the TPSSh Density Functional. *Inorg. Chem.* **2008**, *47* (22), 10357–10365.
(383) Huang, M.; Fabris, S. CO Adsorption and Oxidation on Ceria Surfaces from DFT+U Calculations. *J. Phys. Chem. C* **2008**, *112* (23), 8643–8648.
(384) Callaghan, S. Preview of Machine Learning the Quantum-Chemical Properties of Metal--Organic Frameworks for Accelerated Materials Discovery. *Patterns* **2021**, *2* (4), 100239.
(385) Boyd, P. G.; Woo, T. K. A Generalized Method for Constructing Hypothetical Nanoporous Materials of Any Net Topology from Graph Theory. *CrystEngComm* **2016**, *18* (21), 3777–3792.
(386) Ejsmont, A.; Andreo, J.; Lanza, A.; Galarda, A.; Macreadie, L.; Wuttke, S.; Canossa, S.; Ploetz, E.; Goscianska, J. Applications of Reticular Diversity in Metal–Organic Frameworks: An Ever-Evolving State of the Art. *Coord. Chem. Rev.* **2020**, 213655.
(387) Chibani, S.; Coudert, F.-X. Machine Learning Approaches for the Prediction of Materials Properties. *APL Mater.* **2020**, *8* (8), 80701.
(388) Anderson, G.; Schweitzer, B.; Anderson, R.; Gómez-Gualdrón, D. A. Attainable Volumetric Targets for Adsorption-Based Hydrogen Storage in Porous Crystals: Molecular Simulation and Machine Learning. *J. Phys. Chem. C* **2018**, *123* (1), 120–130.

(389) Bucior, B. J.; Bobbitt, N. S.; Islamoglu, T.; Goswami, S.; Gopalan, A.; Yildirim, T.; Farha, O. K.; Bagheri, N.; Snurr, R. Q. Energy-Based Descriptors to Rapidly Predict Hydrogen Storage in Metal–Organic Frameworks. *Mol. Syst. Des. Eng.* **2019**, *4* (1), 162–174.
(390) Thornton, A. W.; Simon, C. M.; Kim, J.; Kwon, O.; Deeg, K. S.; Konstas, K.; Pas, S. J.; Hill, M. R.; Winkler, D. A.; Haranczyk, M.; Smit, B. Materials Genome in Action: Identifying the Performance Limits of Physical Hydrogen Storage. *Chem. Mater.* **2017**, *29* (7), 2844–2854.
(391) Anderson, R.; Rodgers, J.; Argueta, E.; Biong, A.; Gómez-Gualdrón, D. A. Role of Pore Chemistry and Topology in the CO_2 Capture Capabilities of MOFs: From Molecular Simulation to Machine Learning. *Chem. Mater.* **2018**, *30* (18), 6325–6337.
(392) Dureckova, H.; Krykunov, M.; Aghaji, M. Z.; Woo, T. K. Robust Machine Learning Models for Predicting High CO_2 Working Capacity and CO_2/H_2 Selectivity of Gas Adsorption in Metal Organic Frameworks for Precombustion Carbon Capture. *J. Phys. Chem. C* **2019**, *123* (7), 4133–4139.
(393) Yao, Z.; Sanchez-Lengeling, B.; Bobbitt, N. S.; Bucior, B. J.; Kumar, S. G. H.; Collins, S. P.; Burns, T.; Woo, T. K.; Farha, O. K.; Snurr, R. Q.; Aspuru-Guzik, A. Inverse Design of Nanoporous Crystalline Reticular Materials with Deep Generative Models. *Nat. Mach. Intell.* **2021**, *3* (76–86).
(394) Moghadam, P. Z.; Rogge, S. M. J.; Li, A.; Chow, C.-M.; Wieme, J.; Moharrami, N.; Aragones-Anglada, M.; Conduit, G.; Gomez-Gualdron, D. A.; Van Speybroeck, V.; Fairen-Jimenez, D. Structure-Mechanical Stability Relations of Metal-Organic Frameworks via Machine Learning. *Matter* **2019**, *1* (1), 219–234.
(395) Moosavi, S. M.; Chidambaram, A.; Talirz, L.; Haranczyk, M.; Stylianou, K. C.; Smit, B. Capturing Chemical Intuition in Synthesis of Metal-Organic Frameworks. *Nat. Commun.* **2019**, *10* (1), 539.
(396) Chong, S.; Lee, S.; Kim, B.; Kim, J. Applications of Machine Learning in Metal-Organic Frameworks. *Coord. Chem. Rev.* **2020**, *423*, 213487.
(397) Raza, A.; Sturluson, A.; Simon, C.; Fern, X. Message Passing Neural Networks for Partial Charge Assignment to Metal-Organic Frameworks. *J. Phys. Chem. C* **2020**, *124* (35), 19070–19082.
(398) Korolev, V. V; Mitrofanov, A.; Marchenko, E. I.; Eremin, N. N.; Tkachenko, V.; Kalmykov, S. N. Transferable and Extensible Machine Learning Derived Atomic Charges for Modeling Hybrid Nanoporous Materials. *Chem. Mater.* **2020**, *32* (18), 7822–7831.
(399) Kancharlapalli, S.; Gopalan, A.; Haranczyk, M.; Snurr, R. Q. Fast and Accurate Machine Learning Strategy for Calculating Partial Atomic Charges in Metal–Organic Frameworks. *J. Chem. Theory Comput.* **2021**, *17* (5), 3052–3064.
(400) He, Y.; Cubuk, E. D.; Allendorf, M. D.; Reed, E. J. Metallic Metal–Organic Frameworks Predicted by the Combination of Machine Learning Methods and Ab Initio Calculations. *J. Phys. Chem. Lett.* **2018**, *9* (16), 4562–4569.
(401) Abrahams, B. F.; Hardie, M. J.; Hoskins, B. F.; Robson, R.; Williams, G. A. Topological Rearrangement within a Single Crystal from a Honeycomb Cadmium Cyanide $[Cd(CN)_2]_n$ 3D Net to a Diamond Net. *J. Am. Chem. Soc.* **1992**, *114* (26), 10641–10643.
(402) QMOF Database. https://github.com/arosen93/QMOF, https://dx.doi.org/10.6084/m9.figshare.13147324.

(403) Baumann, A. E.; Burns, D. A.; Liu, B.; Thoi, V. S. Metal–Organic Framework Functionalization and Design Strategies for Advanced Electrochemical Energy Storage Devices. *Commun. Chem.* **2019**, *2* (1), 1–14.
(404) Downes, C. A.; Marinescu, S. C. Electrocatalytic Metal–Organic Frameworks for Energy Applications. *ChemSusChem* **2017**, *10* (22), 4374–4392.
(405) Allendorf, M. D.; Dong, R.; Feng, X.; Kaskel, S.; Matoga, D.; Stavila, V. Electronic Devices Using Open Framework Materials. *Chem. Rev.* **2020**, *120* (16), 8581–8640.
(406) Xie, T.; Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.* **2018**, *120* (14), 145301.
(407) Li, A.; Bueno-Perez, R.; Wiggin, S.; Fairen-Jimenez, D. Enabling Efficient Exploration of Metal–Organic Frameworks in the Cambridge Structural Database. *CrystEngComm* **2020**, *22* (43), 7152–7161.
(408) Zarabadi-Poor, P.; Marek, R. Comment on "Database for CO_2 Separation Performances of MOFs Based on Computational Materials Screening." *ACS Appl. Mater. Interfaces* **2019**, *11* (18), 16261–16265.
(409) Barthel, S.; Alexandrov, E. V; Proserpio, D. M.; Smit, B. Distinguishing Metal−Organic Frameworks. *Cryst. Growth Des.* **2018**, *18* (3), 1738–1747.
(410) Altintas, C.; Avci, G.; Daglar, H.; Azar, A. N. V.; Erucar, I.; Velioglu, S.; Keskin, S. An Extensive Comparative Analysis of Two MOF Databases: High-Throughput Screening of Computation-Ready MOFs for CH_4 and H_2 Adsorption. *J. Mater. Chem. A* **2019**, *7* (16), 9593–9608.
(411) Velioglu, S.; Keskin, S. Revealing the Effect of Structure Curations on the Simulated CO_2 Separation Performances of MOFs. *Mater. Adv.* **2020**, *1* (3), 341–353.
(412) Chen, T.; Manz, T. A. Identifying Misbonded Atoms in the 2019 CoRE Metal–Organic Framework Database. *RSC Adv.* **2020**, *10* (45), 26944–26951.
(413) Hendon, C. H.; Tiana, D.; Walsh, A. Conductive Metal–Organic Frameworks and Networks: Fact or Fantasy? *Phys. Chem. Chem. Phys.* **2012**, *14* (38), 13120–13132.
(414) Singh, A. K.; Montoya, J. H.; Gregoire, J. M.; Persson, K. A. Robust and Synthesizable Photocatalysts for CO_2 Reduction: A Data-Driven Materials Discovery. *Nat. Commun.* **2019**, *10*, 443.
(415) Yang, L.-M.; Ravindran, P.; Vajeeston, P.; Svelle, S.; Tilset, M. A Quantum Mechanically Guided View of Cd-MOF-5 from Formation Energy, Chemical Bonding, Electronic Structure, and Optical Properties. *Microporous Mesoporous Mater.* **2013**, *175*, 50–58.
(416) Gong, S.; Xie, T.; Zhu, T.; Wang, S.; Fadel, E. R.; Li, Y.; Grossman, J. C. Predicting Charge Density Distribution of Materials Using a Local-Environment-Based Graph Convolutional Network. *Phys. Rev. B* **2019**, *100* (18), 184103.
(417) Grisafi, A.; Fabrizio, A.; Meyer, B.; Wilkins, D. M.; Corminboeuf, C.; Ceriotti, M. Transferable Machine-Learning Model of the Electron Density. *ACS Cent. Sci.* **2018**, *5* (1), 57–64.
(418) Chandrasekaran, A.; Kamal, D.; Batra, R.; Kim, C.; Chen, L.; Ramprasad, R. Solving the Electronic Structure Problem with Machine Learning. *npj Comput. Mater.* **2019**, *5* (1), 1–7.
(419) Kamal, D.; Chandrasekaran, A.; Batra, R.; Ramprasad, R. A Charge Density Prediction Model for Hydrocarbons Using Deep Neural Networks. *Mach. Learn. Sci. Technol.* **2020**,

1 (2), 25003.

(420) Kolb, B.; Lentz, L. C.; Kolpak, A. M. Discovering Charge Density Functionals and Structure-Property Relationships with PROPhet: A General Framework for Coupling Machine Learning and First-Principles Methods. *Sci. Rep.* **2017**, *7* (1), 1–9.

(421) Pilania, G.; Wang, C.; Jiang, X.; Rajasekaran, S.; Ramprasad, R. Accelerating Materials Property Predictions Using Machine Learning. *Sci. Rep.* **2013**, *3* (1), 1–6.

(422) Haldoupis, E.; Nair, S.; Sholl, D. S. Finding MOFs for Highly Selective CO_2/N_2 Adsorption Using Materials Screening Based on Efficient Assignment of Atomic Point Charges. *J. Am. Chem. Soc.* **2012**, *134* (9), 4313–4323.

(423) Yang, B.; Wu, X.-P.; Gagliardi, L.; Truhlar, D. G. Methane Functionalization by an Ir(III) Catalyst Supported on a Metal–Organic Framework: An Alternative Explanation of Steric Confinement Effects. *Theor. Chem. Acc.* **2019**, *138* (9), 107.

(424) Sours, T.; Patel, A.; Nørskov, J.; Siahrostami, S.; Kulkarni, A. Circumventing Scaling Relations in Oxygen Electrochemistry Using Metal–Organic Frameworks. *J. Phys. Chem. Lett.* **2020**, *11* (23), 10029–10036.

(425) Planas, N.; Mondloch, J. E.; Tussupbayev, S.; Borycz, J.; Gagliardi, L.; Hupp, J. T.; Farha, O. K.; Cramer, C. J. Defining the Proton Topology of the Zr_6-Based Metal–Organic Framework NU-1000. *J. Phys. Chem. Lett.* **2014**, *5* (21), 3716–3723.

(426) Klet, R. C.; Liu, Y.; Wang, T. C.; Hupp, J. T.; Farha, O. K. Evaluation of Brønsted Acidity and Proton Topology in Zr-and Hf-Based Metal--Organic Frameworks Using Potentiometric Acid--Base Titration. *J. Mater. Chem. A* **2016**, *4* (4), 1479–1485.

(427) Ren, J.; Ledwaba, M.; Musyoka, N. M.; Langmi, H. W.; Mathe, M.; Liao, S.; Pang, W. Structural Defects in Metal--Organic Frameworks (MOFs): Formation, Detection and Control towards Practices of Interests. *Coord. Chem. Rev.* **2017**, *349*, 169–197.

(428) Deria, P.; Mondloch, J. E.; Karagiaridi, O.; Bury, W.; Hupp, J. T.; Farha, O. K. Beyond Post-Synthesis Modification: Evolution of Metal--Organic Frameworks via Building Block Replacement. *Chem. Soc. Rev.* **2014**, *43* (16), 5896–5912.

(429) Syed, Z. H.; Sha, F.; Zhang, X.; Kaphan, D. M.; Delferro, M.; Farha, O. K. Metal−Organic Framework Nodes as a Supporting Platform for Tailoring the Activity of Metal Catalysts. *ACS Catal.* **2020**, *10* (19), 11556–11566.

(430) Ling, S.; Slater, B. Unusually Large Band Gap Changes in Breathing Metal–Organic Framework Materials. *J. Phys. Chem. C* **2015**, *119* (29), 16667–16677.

(431) Mason, J. A.; Oktawiec, J.; Taylor, M. K.; Hudson, M. R.; Rodriguez, J.; Bachman, J. E.; Gonzalez, M. I.; Cervellino, A.; Guagliardi, A.; Brown, C. M.; Llewellyn, P. L.; Masciocchi, N.; Long, J. R. Methane Storage in Flexible Metal−Organic Frameworks with Intrinsic Thermal Management. *Nature* **2015**, *527* (7578), 357–361.

(432) Queen, W. L.; Hudson, M. R.; Bloch, E. D.; Mason, J. A.; Gonzalez, M. I.; Lee, J. S.; Gygi, D.; Howe, J. D.; Lee, K.; Darwish, T. A.; James, M.; Peterson, V. K.; Teat, S. J.; Smit, B.; Neaton, J. B.; Long, J. R.; Brown, C. M. Comprehensive Study of Carbon Dioxide Adsorption in the Metal−Organic Frameworks M_2(Dobdc)(M = Mg, Mn, Fe, Co, Ni, Cu, Zn). *Chem. Sci.* **2014**, *5* (12), 4569–4581.

(433) Xiao, D. J.; Oktawiec, J.; Milner, P. J.; Long, J. R. Pore Environment Effects on Catalytic Cyclohexane Oxidation in Expanded Fe_2(Dobdc) Analogues. *J. Am. Chem. Soc.* **2016**, *138* (43), 14371–14379.

(434) Ruddigkeit, L.; Van Deursen, R.; Blum, L. C.; Reymond, J.-L. Enumeration of 166 Billion Organic Small Molecules in the Chemical Universe Database GDB-17. *J. Chem. Inf. Model.* **2012**, *52* (11), 2864–2875.
(435) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; Von Lilienfeld, O. A. Quantum Chemistry Structures and Properties of 134 Kilo Molecules. *Sci. Data* **2014**, *1*, 140022.
(436) Meredig, B.; Agrawal, A.; Kirklin, S.; Saal, J. E.; Doak, J. W.; Thompson, A.; Zhang, K.; Choudhary, A.; Wolverton, C. Combinatorial Screening for New Materials in Unconstrained Composition Space with Machine Learning. *Phys. Rev. B* **2014**, *89* (9), 94104.
(437) Faber, F.; Lindmaa, A.; von Lilienfeld, O. A.; Armiento, R. Crystal Structure Representations for Machine Learning Models of Formation Energies. *Int. J. Quantum Chem.* **2015**, *115* (16), 1094–1101.
(438) Lam Pham, T.; Kino, H.; Terakura, K.; Miyake, T.; Tsuda, K.; Takigawa, I.; Chi Dam, H. Machine Learning Reveals Orbital Interaction in Materials. *Sci. Technol. Adv. Mater.* **2017**, *18* (1), 756–765.
(439) Bartók, A. P.; Kondor, R.; Csányi, G. On Representing Chemical Environments. *Phys. Rev. B* **2013**, *87* (18), 184115.
(440) De, S.; Bartók, A. P.; Csányi, G.; Ceriotti, M. Comparing Molecules and Solids across Structural and Alchemical Space. *Phys. Chem. Chem. Phys.* **2016**, *18* (20), 13754–13769.
(441) Pronobis, W.; Müller, K.-R. Kernel Methods for Quantum Chemistry. In *Machine Learning Meets Quantum Physics*; Springer, 2020; pp 25–36.
(442) Chen, C.; Ye, W.; Zuo, Y.; Zheng, C.; Ong, S. P. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. *Chem. Mater.* **2019**, *31* (9), 3564–3572.
(443) Olsthoorn, B.; Geilhufe, R. M.; Borysov, S. S.; Balatsky, A. V. Band Gap Prediction for Large Organic Crystal Structures with Machine Learning. *Adv. Quantum Technol.* **2019**, *2* (7–8), 1900023.
(444) Gascon, J.; Hernández-Alonso, M. D.; Almeida, A. R.; van Klink, G. P. M.; Kapteijn, F.; Mul, G. Isoreticular MOFs as Efficient Photocatalysts with Tunable Band Gap: An Operando FTIR Study of the Photoinduced Oxidation of Propylene. *ChemSusChem* **2008**, *1* (12), 981–983.
(445) Stanzione, D.; Barth, B.; Gaffney, N.; Gaither, K.; Hempel, C.; Minyard, T.; Mehringer, S.; Wernert, E.; Tufo, H.; Panda, D.; Teller, P. Stampede 2: The Evolution of an XSEDE Supercomputer. In *Proceedings of the Practice and Experience in Advanced Research Computing 2017 on Sustainability, Success and Impact*; 2017; pp 1–8.
(446) Towns, J.; Cockerill, T.; Dahan, M.; Foster, I.; Gaither, K.; Grimshaw, A.; Hazlewood, V.; Lathrop, S.; Lifka, D.; Peterson, G. D.; Roskies, R.; Scott, J. R.; Wilkins-Diehr, N. XSEDE: Accelerating Scientific Discovery. *Comput. Sci. Eng.* **2014**, *16* (5), 62–74.
(447) McInnes, L.; Healy, J.; Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *arXiv. arXiv:1802.03426* **2018**.
(448) Leland, M.; John, H.; Nathaniel, S.; Lukas, G. UMAP: Uniform Manifold Approximation and Projection. *J. Open Source Softw.* **2018**, *3* (29), 861.
(449) Zeng, X.-Z.; Zhang, A.-Y.; Bu, D.; Li, Y.-W. Hydrothermal Synthesis, Structure and Thermal Properties of a Novel Three-Dimensional La(III)-Sebacate Framework. *Chinese J. Struct. Chem.* **2013**, *32* (1), 120–124.

(450) Zhang, Z.; Zhao, H.; Matsushita, M. M.; Awaga, K.; Dunbar, K. R. A New Metal–Organic Hybrid Material with Intrinsic Resistance-Based Bistability: Monitoring in Situ Room Temperature Switching Behavior. *J. Mater. Chem. C* **2014**, *2* (2), 399–404.
(451) Lopez, N.; Zhao, H.; Ota, A.; Prosvirin, A. V; Reinheimer, E. W.; Dunbar, K. R. Unprecedented Binary Semiconductors Based on TCNQ: Single-Crystal X-Ray Studies and Physical Properties of Cu(TCNQX$_2$) X = Cl, Br. *Adv. Mater.* **2010**, *22* (9), 986–989.
(452) Aubrey, M. L.; Kapelewski, M. T.; Melville, J. F.; Oktawiec, J.; Presti, D.; Gagliardi, L.; Long, J. R. Chemiresistive Detection of Gaseous Hydrocarbons and Interrogation of Charge Transport in Cu[Ni(2,3-Pyrazinedithiolate)$_2$] by Gas Adsorption. *J. Am. Chem. Soc.* **2019**, *141* (12), 5005–5013.
(453) Peng, Y.-L.; Pham, T.; Li, P.; Wang, T.; Chen, Y.; Chen, K.-J.; Forrest, K. A.; Space, B.; Cheng, P.; Zaworotko, M. J.; Zhang, Z. Robust Ultramicroporous Metal–Organic Frameworks with Benchmark Affinity for Acetylene. *Angew. Chemie Int. Ed.* **2018**, *57* (34), 10971–10975.
(454) Takaishi, S.; Hosoda, M.; Kajiwara, T.; Miyasaka, H.; Yamashita, M.; Nakanishi, Y.; Kitagawa, Y.; Yamaguchi, K.; Kobayashi, A.; Kitagawa, H. Electroconductive Porous Coordination Polymer Cu[Cu(Pdt)$_2$] Composed of Donor and Acceptor Building Units. *Inorg. Chem.* **2009**, *48* (19), 9048–9050.
(455) Kobayashi, Y.; Jacobs, B.; Allendorf, M. D.; Long, J. R. Conductivity, Doping, and Redox Chemistry of a Microporous Dithiolene-Based Metal–Organic Framework. *Chem. Mater.* **2010**, *22* (14), 4120–4122.
(456) Nicholas, T. C.; Goodwin, A.; Deringer, V. L. Understanding the Geometric Diversity of Inorganic and Hybrid Frameworks through Structural Coarse-Graining. *Chem. Sci.* **2020**, *11* (46), 12580–12587.
(457) Moosavi, S. M.; Nandy, A.; Jablonka, K. M.; Ongari, D.; Janet, J. P.; Boyd, P. G.; Lee, Y.; Smit, B.; Kulik, H. J. Understanding the Diversity of the Metal–Organic Framework Ecosystem. *Nat. Commun.* **2020**, *11*, 4068.
(458) Hulvey, Z.; Furman, J. D.; Turner, S. A.; Tang, M.; Cheetham, A. K. Dimensionality Trends in Metal–Organic Frameworks Containing Perfluorinated or Nonfluorinated Benzenedicarboxylates. *Cryst. Growth Des.* **2010**, *10* (5), 2041–2043.
(459) Taylor, M. K.; Runčevski, T.; Oktawiec, J.; Gonzalez, M. I.; Siegelman, R. L.; Mason, J. A.; Ye, J.; Brown, C. M.; Long, J. R. Tuning the Adsorption-Induced Phase Change in the Flexible Metal–Organic Framework Co(Bdp). *J. Am. Chem. Soc.* **2016**, *138* (45), 15019–15026.
(460) Cui, X.; Chen, K.; Xing, H.; Yang, Q.; Krishna, R.; Bao, Z.; Wu, H.; Zhou, W.; Dong, X.; Han, Y.; Li, B.; Ren, Q.; Zaworotko, M. J.; Chen, B. Pore Chemistry and Size Control in Hybrid Porous Materials for Acetylene Capture from Ethylene. *Science* **2016**, *353* (6295), 141–144.
(461) Li, S.; Zhang, L.; Lu, B.; Yan, E.; Wang, T.; Li, L.; Wang, J.; Yu, Y.; Mu, Q. A New Polyoxovanadate-Based Metal–Organic Framework: Synthesis, Structure and Photo-/Electro-Catalytic Properties. *New J. Chem.* **2018**, *42* (9), 7247–7253.
(462) Yan, B.; Luo, J.; Dube, P.; Sefat, A. S.; Greedan, J. E.; Maggard, P. A. Spin-Gap Formation and Thermal Structural Studies in Reduced Hybrid Layered Vanadates. *Inorg. Chem.* **2006**, *45* (13), 5109–5118.

(463) Choudhuri, I.; Truhlar, D. G. HLE17: An Efficient Way To Predict Band Gaps of Complex Materials. *J. Phys. Chem. C* **2019**, *123* (28), 17416–17424.
(464) Borlido, P.; Aull, T.; Huran, A. W.; Tran, F.; Marques, M. A. L.; Botti, S. Large-Scale Benchmark of Exchange–Correlation Functionals for the Determination of Electronic Band Gaps of Solids. *J. Chem. Theory Comput.* **2019**, *15* (9), 5069–5079.
(465) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Hybrid Functionals Based on a Screened Coulomb Potential. *J. Chem. Phys.* **2003**, *118* (18), 8207–8215.
(466) Krukau, A. V; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. Influence of the Exchange Screening Parameter on the Performance of Screened Hybrid Functionals. *J. Chem. Phys.* **2006**, *125* (22), 224106.
(467) Moellmann, J.; Grimme, S. DFT-D3 Study of Some Molecular Crystals. *J. Phys. Chem. C* **2014**, *118* (14), 7615–7621.
(468) Valenzano, L.; Civalleri, B.; Chavan, S.; Bordiga, S.; Nilsen, M. H.; Jakobsen, S.; Lillerud, K. P.; Lamberti, C. Disclosing the Complex Structure of UiO-66 Metal Organic Framework: A Synergic Combination of Experiment and Theory. *Chem. Mater.* **2011**, *23* (7), 1700–1718.
(469) Saliba, D.; Ammar, M.; Rammal, M.; Al-Ghoul, M.; Hmadeh, M. Crystal Growth of ZIF-8, ZIF-67, and Their Mixed-Metal Derivatives. *J. Am. Chem. Soc.* **2018**, *140* (5), 1812–1823.
(470) Naito, T.; Kakizaki, A.; Inabe, T.; Sakai, R.; Nishibori, E.; Sawa, H. Growth of Nanocrystals in a Single Crystal of Different Materials: A Way of Giving Function to Molecular Crystals. *Cryst. Growth Des.* **2011**, *11* (2), 501–506.
(471) Sekine, Y.; Tonouchi, M.; Yokoyama, T.; Kosaka, W.; Miyasaka, H. Built-in TTF–TCNQ Charge-Transfer Salts in π-Stacked Pillared Layer Frameworks. *CrystEngComm* **2017**, *19* (17), 2300–2304.
(472) Salami, T. O.; Patterson, S. N.; Jones, V. D.; Masello, A.; Abboud, K. A. Synthesis, Characterization, Thermal Study, and Crystal Structure of a New Layered Alkaline Earth Metal Sulfonate: $Sr[C_2H_4(SO_3)_2]$. *Inorg. Chem. Commun.* **2009**, *12* (11), 1150–1153.
(473) Sun, L.; Hendon, C. H.; Park, S. S.; Tulchinsky, Y.; Wan, R.; Wang, F.; Walsh, A.; Dincă, M. Is Iron Unique in Promoting Electrical Conductivity in MOFs? *Chem. Sci.* **2017**, *8* (6), 4450–4457.
(474) Manna, S. C.; Zangrando, E.; Ribas, J.; Chaudhuri, N. R. Squarato-Bridged Polymeric Networks of Iron(II) with N-Donor Coligands: Syntheses, Crystal Structures and Magnetic Properties. *Inorganica Chim. Acta* **2005**, *358* (15), 4497–4504.
(475) Clements, J. E.; Price, J. R.; Neville, S. M.; Kepert, C. J. Perturbation of Spin Crossover Behavior by Covalent Post-Synthetic Modification of a Porous Metal–Organic Framework. *Angew. Chemie Int. Ed.* **2014**, *126* (38), 10328–10332.
(476) Spirkl, S.; Grzywa, M.; Reschke, S.; Fischer, J. K. H.; Sippel, P.; Demeshko, S.; von Nidda, H.-A.; Volkmer, D. Single-Crystal to Single-Crystal Transformation of a Nonporous Fe(II) Metal−Organic Framework into a Porous Metal-Organic Framework via a Solid-State Reaction. *Inorg. Chem.* **2017**, *56* (20), 12337–12347.
(477) Lee, J.; Asahi, R. Transfer Learning for Materials Informatics Using Crystal Graph Convolutional Neural Network. *Comput. Mater. Sci.* **2021**, *190*, 110314.
(478) Sanyal, S.; Balachandran, J.; Yadati, N.; Kumar, A.; Rajagopalan, P.; Sanyal, S.; Talukdar,

P. MT-CGCNN: Integrating Crystal Graph Convolutional Neural Network with Multitask Learning for Material Property Prediction. *arXiv. arXiv:1811.05660* **2018**.

(479) Ramakrishnan, R.; Dral, P. O.; Rupp, M.; von Lilienfeld, O. A. Big Data Meets Quantum Chemistry Approximations: The Δ-Machine Learning Approach. *J. Chem. Theory Comput.* **2015**, *11* (5), 2087–2096.

(480) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-XTB—An Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method with Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15* (3), 1652–1671.

(481) Spicher, S.; Grimme, S. Robust Atomistic Modeling of Materials, Organometallic, and Biochemical Systems. *Angew. Chemie Int. Ed.* **2020**, *59* (36), 15665–15673.

(482) Stavila, V.; Talin, A. A.; Allendorf, M. D. MOF-Based Electronic and Opto-Electronic Devices. *Chem. Soc. Rev.* **2014**, *43* (16), 5994–6010.

(483) Moosavi, S. M.; Jablonka, K. M.; Smit, B. The Role of Machine Learning in the Understanding and Design of Materials. *J. Am. Chem. Soc.* **2020**, *142* (48), 20273–20287.

(484) Hill, J.; Mannodi-Kanakkithodi, A.; Ramprasad, R.; Meredig, B. Materials Data Infrastructure and Materials Informatics. In *Computational Materials System Design*; Shin, D., Sall, J., Eds.; Springer International Publishing, 2017; pp 193–225.

(485) Zhuo, Y.; Mansouri Tehrani, A.; Brgoch, J. Predicting the Band Gaps of Inorganic Solids by Machine Learning. *J. Phys. Chem. Lett.* **2018**, *9* (7), 1668–1673.

(486) Kauwe, S. K.; Welker, T.; Sparks, T. D. Extracting Knowledge from DFT: Experimental Band Gap Predictions Through Ensemble Learning. *Integr. Mater. Manuf. Innov.* **2020**, *9* (3), 213–220.

(487) Kim, S.; Lee, M.; Hong, C.; Yoon, Y.; An, H.; Lee, D.; Jeong, W.; Yoo, D.; Kang, Y.; Youn, Y.; Han, S. A Band-Gap Database for Semiconducting Inorganic Materials Calculated with Hybrid Functional. *Sci. Data* **2020**, *7*, 387.

(488) Borlido, P.; Schmidt, J.; Huran, A. W.; Tran, F.; Marques, M. A. L.; Botti, S. Exchange-Correlation Functionals for Band Gaps of Solids: Benchmark, Reparametrization and Machine Learning. *npj Comput. Mater.* **2020**, *6*, 96.

(489) Verma, P.; Truhlar, D. G. HLE17: An Improved Local Exchange--Correlation Functional for Computing Semiconductor Band Gaps and Molecular Excitation Energies. *J. Phys. Chem. C* **2017**, *121* (13), 7144–7154.

(490) Meng, Y.; Liu, X.-W.; Huo, C.-F.; Guo, W.-P.; Cao, D.-B.; Peng, Q.; Dearden, A.; Gonze, X.; Yang, Y.; Wang, J.; Jiao, H.; Li, Y.; Wen, X.-D. When Density Functional Approximations Meet Iron Oxides. *J. Chem. Theory Comput.* **2016**, *12* (10), 5132–5144.

(491) Yang, L.-M.; Fang, G.-Y.; Ma, J.; Ganz, E.; Han, S. S. Band Gap Engineering of Paradigm MOF-5. *Cryst. Growth Des.* **2014**, *14* (5), 2532–2541.

(492) Li, X.; Yang, J. First-Principles Design of Spintronics Materials. *Natl. Sci. Rev.* **2016**, *3* (3), 365–381.

(493) Medvedev, M. G.; Bushmarinov, I. S.; Sun, J.; Perdew, J. P.; Lyssenko, K. A. Density Functional Theory Is Straying from the Path toward the Exact Functional. *Science* **2017**, *355* (6320), 49–52.

(494) Wang, B.; Li, S. L.; Truhlar, D. G. Modeling the Partial Atomic Charges in Inorganometallic Molecules and Solids and Charge Redistribution in Lithium-Ion Cathodes.

J. Chem. Theory Comput. **2014**, *10* (12), 5640–5650.
(495) Ioannidis, E. I.; Kulik, H. J. Towards Quantifying the Role of Exact Exchange in Predictions of Transition Metal Complex Properties. *J. Chem. Phys.* **2015**, *143* (3), 34104.
(496) Bader, R. F. W.; Matta, C. F. Atomic Charges Are Measurable Quantum Expectation Values: A Rebuttal of Criticisms of QTAIM Charges. *J. Phys. Chem. A* **2004**, *108* (40), 8385–8394.
(497) Choudhuri, I.; Truhlar, D. G. Calculating and Characterizing the Charge Distributions in Solids. *J. Chem. Theory Comput.* **2020**, *16* (9), 5884–5892.
(498) Wilkinson, M. D.; Dumontier, M.; Aalbersberg, Ij. J.; Appleton, G.; Axton, M.; Baak, A.; Blomberg, N.; Boiten, J.-W.; da Silva Santos, L. B.; Bourne, P. E.; Bouwman, J.; Brookes, A. J.; Clark, T.; Crosas, M.; Dillo, I.; Dumon, O.; Edmunds, S.; Evelo, C. T.; Finkers, R.; Gonzalez-Beltran, A.; Gray, A. J. G.; Groth, P.; Goble, C.; Grethe, J. S.; Heringa, J.; 't Hoen, P. A. C.; Hooft, R.; Kuhn, T.; Kok, R.; Kok, J.; Lusher, S. J.; Martone, M. E.; Mons, A.; Packer, A. L.; Persson, B.; Rocca-Serra, P.; Roos, M.; van Schaik, R.; Sansone, S.-A.; Schultes, E.; Sengstag, T.; Slater, T.; Strawn, G.; Swertz, M. A.; Thompson, M.; van der Lei, J.; van Mulligen, E.; Velterop, J.; Waagmeester, A.; Wittenburg, P.; Wolstencroft, K.; Zhao, J.; Mons, B. The FAIR Guiding Principles for Scientific Data Management and Stewardship. *Sci. Data* **2016**, *3* (1), 160018.
(499) Fung, V.; Zhang, J.; Juarez, E.; Sumpter, B. Benchmarking Graph Neural Networks for Materials Chemistry. *npj Comput. Mater.* **2021**, *7*, 84.
(500) Ruder, S. An Overview of Multi-Task Learning in Deep Neural Networks. *arXiv Prepr. arXiv1706.05098* **2017**.
(501) Pilania, G.; Gubernatis, J. E.; Lookman, T. Multi-Fidelity Machine Learning Models for Accurate Bandgap Predictions of Solids. *Comput. Mater. Sci.* **2017**, *129*, 156–163.
(502) Chen, C.; Zuo, Y.; Ye, W.; Li, X.; Ong, S. P. Learning Properties of Ordered and Disordered Materials from Multi-Fidelity Data. *Nat. Comput. Sci.* **2021**, *1* (1), 46–53.
(503) Ma, S.; Liu, Z.-P. Machine Learning for Atomic Simulation and Activity Prediction in Heterogeneous Catalysis: Current Status and Future. *ACS Catal.* **2020**, *10* (22), 13213–13226.
(504) Palomba, J. M.; Credille, C. V; Kalaj, M.; DeCoste, J. B.; Peterson, G. W.; Tovar, T. M.; Cohen, S. M. High-Throughput Screening of Solid-State Catalysts for Nerve Agent Degradation. *Chem. Commun.* **2018**, *54* (45), 5768–5771.
(505) Palomba, J. M.; Harvey, S. P.; Kalaj, M.; Pimentel, B. R.; DeCoste, J. B.; Peterson, G. W.; Cohen, S. M. High-Throughput Screening of MOFs for Breakdown of V-Series Nerve Agents. *ACS Appl. Mater. Interfaces* **2020**, *12* (13), 14672–14677.
(506) Clayson, I. G.; Hewitt, D.; Hutereau, M.; Pope, T.; Slater, B. High Throughput Methods in the Synthesis, Characterization, and Optimization of Porous Materials. *Adv. Mater.* **2020**, *32* (44), 2002780.
(507) Ye, J.; Gagliardi, L.; Cramer, C. J.; Truhlar, D. G. Computational Screening of MOF-Supported Transition Metal Catalysts for Activity and Selectivity in Ethylene Dimerization. *J. Catal.* **2018**, *360*, 160–167.
(508) Barona, M.; Snurr, R. Q. Exploring the Tunability of Trimetallic MOF Nodes for Partial Oxidation of Methane to Methanol. *ACS Appl. Mater. Interfaces* **2020**, *12* (25), 28217–28231.

(509) Vitillo, J. G.; Lu, C. C.; Cramer, C. J.; Bhan, A.; Gagliardi, L. Influence of First and Second Coordination Environment on Structural Fe(II) Sites in MIL-101 for C–H Bond Activation in Methane. *ACS Catal.* **2020**, *11*, 579–589.
(510) Mendonca, M. L.; Snurr, R. Q. Computational Screening of Metal–Organic Framework-Supported Single-Atom Transition-Metal Catalysts for the Gas-Phase Hydrolysis of Nerve Agents. *ACS Catal.* **2019**, *10* (2), 1310–1323.
(511) Hu, T.; Jiang, Y.; Ding, Y. Computational Screening of Metal-Substituted HKUST-1 Catalysts for Chemical Fixation of Carbon Dioxide into Epoxides. *J. Mater. Chem. A* **2019**, *7* (24), 14825–14834.
(512) Mao, X.; Tang, C.; He, T.; Wijethunge, D.; Yan, C.; Zhu, Z.; Du, A. Computational Screening of MN_4 (M = Ti-Cu) Based Metal Organic Frameworks for CO_2 Reduction Using the d-Band Centre as a Descriptor. *Nanoscale* **2020**, *12* (10), 6188–6194.
(513) Ortuño, M. A.; Bernales, V.; Gagliardi, L.; Cramer, C. J. Computational Study of First-Row Transition Metals Supported on MOF NU-1000 for Catalytic Acceptorless Alcohol Dehydrogenation. *J. Phys. Chem. C* **2016**, *120* (43), 24697–24705.
(514) McCarver, G. A.; Rajeshkumar, T.; Vogiatzis, K. D. Computational Catalysis for Metal-Organic Frameworks: An Overview. *Coord. Chem. Rev.* **2021**, *436*, 213777.
(515) Wu, X.-P.; Choudhuri, I.; Truhlar, D. G. Computational Studies of Photocatalysis with Metal--Organic Frameworks. *Energy Environ. Mater.* **2019**, *2* (4), 251–263.
(516) Fumanal, M.; Capano, G.; Barthel, S.; Smit, B.; Tavernelli, I. Energy-Based Descriptors for Photo-Catalytically Active Metal--Organic Framework Discovery. *J. Mater. Chem. A* **2020**, *8* (8), 4473–4482.
(517) Fumanal, M.; Ortega-Guerrero, A.; Jablonka, K. M.; Smit, B.; Tavernelli, I. Charge Separation and Charge Carrier Mobility in Photocatalytic Metal-Organic Frameworks. *Adv. Funct. Mater.* **2020**, *30* (49), 2003792.
(518) Schweitzer, B.; Archuleta, C.; Seong, B.; Anderson, R.; Gómez-Gualdrón, D. A. Electronic Effects Due to Organic Linker-Metal Surface Interactions: Implications on Screening of MOF-Encapsulated Catalysts. *Phys. Chem. Chem. Phys.* **2020**, *22* (4), 2475–2487.
(519) Li, S.; Zhang, Y.; Hu, Y.; Wang, B.; Sun, S.; Yang, X.; He, H. Predicting Metal-Organic Frameworks as Catalysts to Fix Carbon Dioxide to Cyclic Carbonate by Machine Learning. *J. Mater.* **2021**.
(520) Batra, R.; Chen, C.; Evans, T. G.; Walton, K. S.; Ramprasad, R. Prediction of Water Stability of Metal--Organic Frameworks Using Machine Learning. *Nat. Mach. Intell.* **2020**, *2* (11), 704–710.
(521) Nandy, A.; Duan, C.; Kulik, H. J. Using Machine Learning and Data Mining to Leverage Community Knowledge for the Engineering of Stable Metal-Organic Frameworks. *arXiv* **2021**.
(522) Chen, P.; Tang, Z.; Zeng, Z.; Hu, X.; Xiao, L.; Liu, Y.; Qian, X.; Deng, C.; Huang, R.; Zhang, J.; Bi, Y.; Lin, R.; Zhou, Y.; Liao, H.; Zhou, D.; Wang, C.; Lin, W. Machine-Learning-Guided Morphology Engineering of Nanoscale Metal-Organic Frameworks. *Matter* **2020**, *2* (6), 1651–1666.
(523) Guda, A. A.; Guda, S. A.; Martini, A.; Bugaev, A. L.; Soldatov, M. A.; Soldatov, A. V; Lamberti, C. Machine Learning Approaches to XANES Spectra for Quantitative 3D Structural Determination: The Case of CO2 Adsorption on CPO-27-Ni MOF. *Radiat. Phys.*

Chem. **2020**, *175*, 108430.
(524) Chanussot, L.; Das, A.; Goyal, S.; Lavril, T.; Shuaibi, M.; Riviere, M.; Tran, K.; Heras-Domingo, J.; Ho, C.; Hu, W.; Palizhati, A.; Sriram, A.; Wood, B.; Yoon, J.; Parikh, D.; Zitnick, C. L.; Ulissi, Z. The Open Catalyst 2020 (OC20) Dataset and Community Challenges. *ACS Catal.* **2021**, *11* (10), 6059–6072.
(525) Yamada, H.; Liu, C.; Wu, S.; Koyama, Y.; Ju, S.; Shiomi, J.; Morikawa, J.; Yoshida, R. Predicting Materials Properties with Little Data Using Shotgun Transfer Learning. *ACS Cent. Sci.* **2019**, *5* (10), 1717–1730.
(526) Jablonka, K. M.; Ongari, D.; Moosavi, S. M.; Smit, B. Using Collective Knowledge to Assign Oxidation States of Metal Cations in Metal–Organic Frameworks. *Nat. Chem.* **2021**.
(527) Turcani, L.; Tarzia, A.; Szczypiński, F. T.; Jelfs, K. E. Stk: An Extendable Python Framework for Automated Molecular and Supramolecular Structure Assembly and Discovery. *J. Chem. Phys.* **2021**, *154* (21), 214102.
(528) Taylor, M. G.; Yang, T.; Lin, S.; Nandy, A.; Janet, J. P.; Duan, C.; Kulik, H. J. Seeing Is Believing: Experimental Spin States from Machine Learning Model Structure Predictions. *J. Phys. Chem. A* **2020**, *124* (16), 3286–3299.
(529) Horton, M. K.; Montoya, J. H.; Liu, M.; Persson, K. A. High-Throughput Prediction of the Ground-State Collinear Magnetic Order of Inorganic Materials Using Density Functional Theory. *npj Comput. Mater.* **2019**, *5*, 64.
(530) Lyu, H.; Ji, Z.; Wuttke, S.; Yaghi, O. M. Digital Reticular Chemistry. *Chem* **2020**, *6* (9), 2219–2241.
(531) Coudert, F.-X. Materials Databases: The Need for Open, Interoperable Databases with Standardized Data and Rich Metadata. *Adv. Theory Simulations* **2019**, *2* (11), 1900131.
(532) Cao, L.; Caldararu, O.; Rosenzweig, A. C.; Ryde, U. Quantum Refinement Does Not Support Dinuclear Copper Sites in Crystal Structures of Particulate Methane Monooxygenase. *Angew. Chemie Int. Ed.* **2018**, *57*, 162–166.
(533) Ro, S. Y.; Ross, M. O.; Deng, Y. W.; Batelu, S.; Lawton, T. J.; Hurley, J. D.; Stemmler, T. L.; Hoffman, B. M.; Rosenzweig, A. C. From Micelles to Bicelles: Effect of the Membrane on Particulate Methane Monooxygenase Activity. *J. Biol. Chem.* **2018**, jbc--RA118.
(534) Ciano, L.; Davies, G. J.; Tolman, W. B.; Walton, P. H. Bracing Copper for the Catalytic Oxidation of C−H Bonds. *Nat. Catal.* **2018**, *1* (8), 571.
(535) Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R. New Software for Searching the Cambridge Structural Database and Visualizing Crystal Structures. *Acta Crystallogr. Sect. B* **2002**, *58* (3–1), 389–397.
(536) Mon, M.; Ferrando-Soria, J.; Verdaguer, M.; Train, C.; Paillard, C.; Dkhil, B.; Versace, C.; Bruno, R.; Armentano, D.; Pardo, E. Postsynthetic Approach for the Rational Design of Chiral Ferroelectric Metal--Organic Frameworks. *J. Am. Chem. Soc.* **2017**, *139* (24), 8098–8101.
(537) Bour, J. R.; Wright, A. M.; He, X.; Dincă, M. Bioinspired Chemistry at MOF Secondary Building Units. *Chem. Sci.* **2020**, *11* (7), 1728–1737.
(538) Caldararu, O.; Oksanen, E.; Ryde, U.; Hedegård, E. D. Mechanism of Hydrogen Peroxide Formation by Lytic Polysaccharide Monooxygenase. *Chem. Sci.* **2019**, *10* (2), 576–586.
(539) Yoshizawa, K.; Shiota, Y. Conversion of Methane to Methanol at the Mononuclear and Dinuclear Copper Sites of Particulate Methane Monooxygenase (PMMO): A DFT and

QM/MM Study. *J. Am. Chem. Soc.* **2006**, *128* (30), 9873–9881.
(540) Miyanishi, M.; Abe, T.; Hori, Y.; Shiota, Y.; Yoshizawa, K. Role of Amino Acid Residues for Dioxygen Activation in the Second Coordination Sphere of the Dicopper Site of PMMO. *Inorg. Chem.* **2019**, *58* (18), 12280–12288.
(541) Peng, W.; Qu, X.; Shaik, S.; Wang, B. Deciphering the Oxygen Activation Mechanism at the Cu C Site of Particulate Methane Monooxygenase. *Nat. Catal.* **2021**, *4* (4), 266–273.
(542) Khramenkova, E. V; Polynski, M. V; Vinogradov, A. V; Pidko, E. A. Degradation Paths of Manganese-Based MOF Materials in a Model Oxidative Environment: A Computational Study. *Phys. Chem. Chem. Phys.* **2018**, *20* (32), 20785–20795.
(543) Butler, K. T.; Hendon, C. H.; Walsh, A. Electronic Structure Modulation of Metal--Organic Frameworks for Hybrid Devices. *ACS Appl. Mater. Interfaces* **2014**, *6* (24), 22044–22050.
(544) Grimme, S.; Bannwarth, C.; Shushkov, P. A Robust and Accurate Tight-Binding Quantum Chemical Method for Structures, Vibrational Frequencies, and Noncovalent Interactions of Large Molecular Systems Parametrized for All Spd-Block Elements (Z = 1-86). *J. Chem. Theory Comput.* **2017**, *13* (5), 1989–2009.
(545) Bursch, M.; Neugebauer, H.; Grimme, S. Structure Optimisation of Large Transition-Metal Complexes with Extended Tight-Binding Methods. *Angew. Chemie Int. Ed.* **2019**, *58* (32), 11078–11087.
(546) Runčevski, T.; Kapelewski, M. T.; Torres-Gavosto, R. M.; Tarver, J. D.; Brown, C. M.; Long, J. R. Adsorption of Two Gas Molecules at a Single Metal Site in a Metal−organic Framework. *Chem. Commun.* **2016**, *52* (53), 8251–8254.
(547) Wellendorff, J.; Lundgaard, K. T.; Møgelhøj, A.; Petzold, V.; Landis, D. D.; Nørskov, J. K.; Bligaard, T.; Jacobsen, K. W. Density Functionals for Surface Science: Exchange-Correlation Model Development with Bayesian Error Estimation. *Phys. Rev. B - Condens. Matter Mater. Phys.* **2012**, *85* (23), 32–34.
(548) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. Supplementary Data for "Identifying Promising Metal−Organic Frameworks for Heterogeneous Catalysis via High-Throughput Periodic Density Functional Theory" https://doi.org/10.5281/zenodo.1469787.
(549) Denysenko, D.; Werner, T.; Grzywa, M.; Puls, A.; Hagen, V.; Eickerling, G.; Jelic, J.; Reuter, K.; Volkmer, D. Reversible Gas-Phase Redox Processes Catalyzed by Co-Exchanged MFU-4*l*(Arge). *Chem. Commun.* **2012**, *48* (9), 1236–1238.
(550) Asgari, M.; Jawahery, S.; Bloch, E. D.; Hudson, M. R.; Flacau, R.; Vlaisavljevich, B.; Long, J. R.; Brown, C. M.; Queen, W. L. An Experimental and Computational Study of CO_2 Adsorption in the Sodalite-Type M-BTT (M = Cr, Mn, Fe, Cu) Metal−Organic Frameworks Featuring Open Metal Sites. *Chem. Sci.* **2018**, *9* (20), 4579–4588.
(551) Sun, L.; Hendon, C. H.; Dincă, M. Coordination-Induced Reversible Electrical Conductivity Variation in the MOF-74 Analogue Fe_2(DSBDC). *Dalt. Trans.* **2018**, *47* (34), 11739.
(552) Mitchell, L.; Gonzalez-Santiago, B.; Mowat, J. P. S.; Gunn, M. E.; Williamson, P.; Acerbi, N.; Clarke, M. L.; Wright, P. A. Remarkable Lewis Acid Catalytic Performance of the Scandium Trimesate Metal−Organic Framework MIL-100 (Sc) for C−C and C=N Bond-Forming Reactions. *Catal. Sci. Technol.* **2013**, *3* (3), 606–617.
(553) Liu, S.; Zhang, Y.; Han, Y.; Feng, G.; Gao, F.; Wang, H.; Qiu, P. Selective Ethylene Oligomerization with Chromium-Based Metal−Organic Framework MIL-100 Evacuated

under Different Temperatures. *Organometallics* **2017**, *36* (3), 632–638.
(554) Sciortino, L.; Alessi, A.; Messina, F.; Buscarino, G.; Gelardi, F. M. Structure of the FeBTC Metal−Organic Framework: A Model Based on the Local Environment Study. *J. Phys. Chem. C* **2015**, *119* (14), 7826–7830.
(555) Nørskov, J. K.; Rossmeisl, J.; Logadottir, A.; Lindqvist, L.; Kitchin, J. R.; Bligaard, T.; Jonsson, H. Origin of the Overpotential for Oxygen Reduction at a Fuel-Cell Cathode. *J. Phys. Chem. B* **2004**, *108* (46), 17886–17892.
(556) Man, I. C.; Su, H.-Y.; Calle-Vallejo, F.; Hansen, H. A.; Martínez, J. I.; Inoglu, N. G.; Kitchin, J.; Jaramillo, T. F.; Nørskov, J. K.; Rossmeisl, J. Universality in Oxygen Evolution Electrocatalysis on Oxide Surfaces. *ChemCatChem* **2011**, *3* (7), 1159–1165.
(557) Karlberg, G. S.; Rossmeisl, J.; Nørskov, J. K. Estimations of Electric Field Effects on the Oxygen Reduction Reaction Based on the Density Functional Theory. *Phys. Chem. Chem. Phys.* **2007**, *9* (37), 5158–5161.
(558) Viswanathan, V.; Hansen, H. A.; Rossmeisl, J.; Nørskov, J. K. Unifying the 2e– and 4e– Reduction of Oxygen on Metal Surfaces. *J. Phys. Chem. Lett.* **2012**, *3* (20), 2948–2951.
(559) Virtual NanoLab version 2017.1, QuantumWise A/S www.quantumwise.com.
(560) Momma, K.; Izumi, F. VESTA 3 for Three-Dimensional Visualization of Crystal, Volumetric and Morphology Data. *J. Appl. Crystallogr.* **2011**, *44* (6), 1272–1276.
(561) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9* (3), 90–95.
(562) Oliphant, T. E. *A Guide to NumPy*; Trelgol Publishing USA, 2006; Vol. 1.
(563) McKinney, W. Data Structures for Statistical Computing in Python. In *Proceedings of the 9th Python in Science Conference*; 2010; Vol. 445, pp 51–56.
(564) Seabold, S.; Perktold, J. Statsmodels: Econometric and Statistical Modeling with Python. In *9th Python in Science Conference*; 2010.
(565) Schmieder, P.; Denysenko, D.; Grzywa, M.; Baumgärtner, B.; Senkovska, I.; Kaskel, S.; Sastre, G.; van Wüllen, L.; Volkmer, D. CFA-1: The First Chiral Metal–Organic Framework Containing Kuratowski-Type Secondary Building Units. *Dalt. Trans.* **2013**, *42* (30), 10786–10797.
(566) Bien, C. E.; Chen, K. K.; Chien, S.-C.; Reiner, B. R.; Lin, L.-C.; Wade, C. R.; Ho, W. S. W. Bioinspired Metal−Organic Framework for Trace CO_2 Capture. *J. Am. Chem. Soc.* **2018**, *140* (40), 12662–12666.
(567) Remer, L. C.; Jensen, J. H. Toward a General Theory of Hydrogen Bonding: The Short, Strong Hydrogen Bond $[HOH\cdots OH]^-$. *J. Phys. Chem. A* **2000**, *104* (40), 9266–9275.
(568) Verma, P.; Xu, X.; Truhlar, D. G. Adsorption on Fe-MOF-74 for C1−C3 Hydrocarbon Separation. *J. Phys. Chem. C* **2013**, *117* (24), 12648–12660.
(569) Szécsényi, Á.; Khramenkova, E.; Chernyshov, I. Y.; Li, G.; Gascon, J.; Pidko, E. A. Breaking Linear Scaling Relationships with Secondary Interactions in Confined Space: A Case Study of Methane Oxidation by Fe/ZSM-5 Zeolite. *ACS Catal.* **2019**, *9* (10), 9276–9284.
(570) Cirera, J.; Via-Nadal, M.; Ruiz, E. Benchmarking Density Functional Methods for Calculation of State Energies of First Row Spin-Crossover Molecules. *Inorg. Chem.* **2018**, *57* (22), 14097–14105.
(571) Yu, H. S.; He, X.; Truhlar, D. G. MN15-L: A New Local Exchange-Correlation Functional

for Kohn–Sham Density Functional Theory with Broad Accuracy for Atoms, Molecules, and Solids. *J. Chem. Theory Comput.* **2016**, *12* (3), 1280–1293.

(572) Haoyu, S. Y.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn−Sham Global-Hybrid Exchange−Correlation Density Functional with Broad Accuracy for Multi-Reference and Single-Reference Systems and Noncovalent Interactions. *Chem. Sci.* **2016**, *7* (8), 5032–5051.

(573) Leto, D. F.; Ingram, R.; Day, V. W.; Jackson, T. A. Spectroscopic Properties and Reactivity of a Mononuclear Oxomanganese(IV) Complex. *Chem. Commun.* **2013**, *49* (47), 5378–5380.

(574) England, J.; Martinho, M.; Farquhar, E. R.; Frisch, J. R.; Bominaar, E. L.; Münck, E.; Que Jr, L. A Synthetic High-Spin Oxoiron(IV) Complex: Generation, Spectroscopic Characterization, and Reactivity. *Angew. Chemie Int. Ed.* **2009**, *48* (20), 3622–3626.

(575) England, J.; Guo, Y.; Farquhar, E. R.; Young Jr., V. G.; Münck, E.; Que Jr., L. The Crystal Structure of a High-Spin Oxoiron(IV) Complex and Characterization of Its Self-Decay Pathway. *J. Am. Chem. Soc.* **2010**, *132* (25), 8635–8644.

(576) Wang, B.; Lee, Y.-M.; Tcho, W.-Y.; Tussupbayev, S.; Kim, S.-T.; Kim, Y.; Seo, M. S.; Cho, K.-B.; Dede, Y.; Keegan, B. C.; Ogura, T.; Kim, S. H.; Ohta, T.; Baik, M.-H.; Ray, K.; Shearer, J.; Nam, W. Synthesis and Reactivity of a Mononuclear Non-Haem Cobalt(IV)-Oxo Complex. *Nat. Commun.* **2017**, *8*, 14839.

(577) Johansson, A. J.; Blomberg, M. R. A.; Siegbahn, P. E. M. Quantifying the Effects of the Self-Interaction Error in Density Functional Theory: When Do the Delocalized States Appear? II. Iron-Oxo Complexes and Closed-Shell Substrate Molecules. *J. Chem. Phys.* **2008**, *129* (15), 154301.

(578) Cho, K.-B.; Shaik, S.; Nam, W. Theoretical Investigations into C−H Bond Activation Reaction by Nonheme $Mn^{IV}O$ Complexes: Multistate Reactivity with No Oxygen Rebound. *J. Phys. Chem. Lett.* **2012**, *3* (19), 2851–2856.

(579) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. Energies, Structures, and Electronic Properties of Molecules in Solution with the C-PCM Solvation Model. *J. Comput. Chem.* **2003**, *24* (6), 669–681.

(580) Barone, V.; Cossi, M. Quantum Calculation of Molecular Energies and Energy Gradients in Solution by a Conductor Solvent Model. *J. Phys. Chem. A* **1998**, *102* (11), 1995–2001.

(581) LASPH-tag https://cms.mpi.univie.ac.at/wiki/index.php/LASPH.

(582) Chemcraft - graphical software for visualization of quantum chemistry computations http://www.chemcraftprog.com.

(583) Maintz, S.; Deringer, V. L.; Tchougréeff, A. L.; Dronskowski, R. LOBSTER: A Tool to Extract Chemical Bonding from Plane-Wave Based DFT. *J. Comput. Chem.* **2016**, *37* (11), 1030–1035.

(584) Dronskowski, R.; Blöchl, P. E. Crystal Orbital Hamilton Populations (COHP): Energy-Resolved Visualization of Chemical Bonding in Solids Based on Density-Functional Calculations. *J. Phys. Chem.* **1993**, *97* (33), 8617–8624.

(585) Deringer, V. L.; Tchougréeff, A. L.; Dronskowski, R. Crystal Orbital Hamilton Population (COHP) Analysis as Projected from Plane-Wave Basis Sets. *J. Phys. Chem. A* **2011**, *115* (21), 5461–5466.

(586) Maintz, S.; Deringer, V. L.; Tchougréeff, A. L.; Dronskowski, R. Analytic Projection from

Plane-Wave and PAW Wavefunctions and Application to Chemical-Bonding Analysis in Solids. *J. Comput. Chem.* **2013**, *34* (29), 2557–2567.

(587) Badger, R. M. A Relation between Internuclear Distances and Bond Force Constants. *J. Chem. Phys.* **1934**, *2* (3), 128–131.

(588) Cioslowski, J.; Liu, G.; Castro, R. A. M. Badger's Rule Revisited. *Chem. Phys. Lett.* **2000**, *331* (5–6), 497–501.

(589) Borycz, J.; Lin, L.-C.; Bloch, E. D.; Kim, J.; Dzubak, A. L.; Maurice, R.; Semrouni, D.; Lee, K.; Smit, B.; Gagliardi, L. CO_2 Adsorption in Fe_2(Dobdc): A Classical Force Field Parameterized from Quantum Mechanical Calculations. *J. Phys. Chem. C* **2014**, *118* (23), 12230–12240.

(590) Poloni, R.; Lee, K.; Berger, R. F.; Smit, B.; Neaton, J. B. Understanding Trends in CO_2 Adsorption in Metal–Organic Frameworks with Open-Metal Sites. *J. Phys. Chem. Lett.* **2014**, *5* (5), 861–865.

(591) Lim, K. S.; Lee, W. R.; Lee, H. G.; Kang, D. W.; Song, J. H.; Hilgar, J.; Rinehart, J. D.; Moon, D.; Hong, C. S. Control of Interchain Antiferromagnetic Coupling in Porous Co(II)-Based Metal--Organic Frameworks by Tuning the Aromatic Linker Length: How Far Does Magnetic Interaction Propagate? *Inorg. Chem.* **2017**, *56* (13), 7443–7448.

(592) Dietzel, P. D. C.; Morita, Y.; Blom, R.; Fjellvåg, H. An in Situ High-Temperature Single-Crystal Investigation of a Dehydrated Metal–Organic Framework Compound and Field-Induced Magnetization of One-Dimensional Metal–Oxygen Chains. *Angew. Chemie Int. Ed.* **2005**, *44* (39), 6354–6358.

(593) Yu, D.; Yazaydin, A. O.; Lane, J. R.; Dietzel, P. D. C.; Snurr, R. Q. A Combined Experimental and Quantum Chemical Study of CO_2 Adsorption in the Metal−Organic Framework CPO-27 with Different Metals. *Chem. Sci.* **2013**, *4* (9), 3544–3556.

(594) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. DREIDING: A Generic Force Field for Molecular Simulations. *J. Phys. Chem.* **1990**, *94*, 8897–8909.

(595) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M.; Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114* (25), 10024–10035.

(596) Potoff, J. J.; Siepmann, J. I. Vapor-Liquid Equilibria of Mixtures Containing Alkanes, Carbon Dioxide, and Nitrogen. *AIChE J.* **2001**, *47*, 1676–1682. https://doi.org/10.26434/chemrxiv.11604129.

(597) Jouffret, L.; Rivenet, M.; Abraham, F. Linear Alkyl Diamine-Uranium-Phosphate Systems: U(VI) to U(IV) Reduction with Ethylenediamine. *Inorg. Chem.* **2011**, *50* (10), 4619–4626.

(598) Feng, P.; Bu, X.; Stucky, G. D. Hydrothermal Syntheses and Structural Characterization of Zeolite Analogue Compounds Based on Cobalt Phosphate. *Nature* **1997**, *388* (6644), 735–741.

(599) Liang, L.; Zhang, R.; Zhao, J.; Liu, C.; Weng, N. S. Two Actinide-Organic Frameworks Constructed by a Tripodal Flexible Ligand: Occurrence of Infinite {(UO2)O2(OH)3}4n and Hexanuclear {Th6O4(OH)4} Motifs. *J. Solid State Chem.* **2016**, *243*, 50–56.

(600) Eubank, J. F.; Nouar, F.; Luebke, R.; Cairns, A. J.; Wojtas, L.; Alkordi, M.; Bousquet, T.; Hight, M. R.; Eckert, J.; Embs, J. P.; Georgiev, P. A.; Eddaoudi, M. On Demand: The Singular Rht Net, an Ideal Blueprint for the Construction of a Metal–Organic Framework

(MOF) Platform. *Angew. Chemie. Int. Ed.* **2012**, *124* (40), 10246–10250.
(601) Queen, W. L.; Bloch, E. D.; Brown, C. M.; Hudson, M. R.; Mason, J. A.; Murray, L. J.; Ramirez-Cuesta, A. J.; Peterson, V. K.; Long, J. R. Hydrogen Adsorption in the Metal–Organic Frameworks Fe_2(Dobdc) and $Fe_2(O_2)$(Dobdc). *Dalt. Trans.* **2012**, *41* (14), 4180–4187.
(602) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
(603) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; Van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Ilhan, P.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbgret, P. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272.
(604) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N. J.; Kern, R.; Picus, M.; Hoyer, S.; van Kerkwijk, M. H.; Brett, M.; Haldane, A.; del Río, J. F.; Wiebe, M.; Peterson, P.; Gérard-Marchant, P.; Sheppard, K.; Reddy, T.; Weckesser, W.; Abbasi, H.; Gohlke, C.; Oliphant, T. E. Array Programming with NumPy. *Nature* **2020**, *585* (7825), 357–362.
(605) Seaborn http://doi.org/10.5281/zenodo.592845.
(606) Himanen, L.; Jäger, M. O. J.; Morooka, E. V; Canova, F. F.; Ranawat, Y. S.; Gao, D. Z.; Rinke, P.; Foster, A. S. DScribe: Library of Descriptors for Machine Learning in Materials Science. *Comput. Phys. Commun.* **2020**, *247*, 106949.
(607) Ward, L.; Dunn, A.; Faghaninia, A.; Zimmermann, N. E. R.; Bajaj, S.; Wang, Q.; Montoya, J.; Chen, J.; Bystrom, K.; Dylla, M.; Chard, K.; Asta, M.; Persson, K. A.; Snyder, J.; Foster, I.; Jain, A. Matminer: An Open Source Toolkit for Materials Data Mining. *Comput. Mater. Sci.* **2018**, *152*, 60–69.
(608) Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; Desmaison, A.; Kopf, A.; Yang, E.; DeVito, Z.; Raison, M.; Tejani, A.; Chilamkurthy, S.; Steiner, B.; Fang, L.; Bai, J.; Chintala, S. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems*; 2019; pp 8024–8035.
(609) Xie, T.; Grossman, J. Crystal Graph Convolutional Neural Networks https://github.com/txie-93/cgcnn.
(610) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. Crystal Graph Convolutional Neural Networks https://github.com/snurr-group/cgcnn.
(611) Rosen, A. S. PTable Trends. https://doi.org/10.5281/zenodo.1451879.
(612) Herath, U.; Tavadze, P.; He, X.; Bousquet, E.; Singh, S.; Munoz, F.; Romero, A. H. PyProcar: A Python Library for Electronic Structure Pre/Post-Processing. *Comput. Phys. Commun.* **2020**, *251*, 107080.
(613) Murphy, K. P. *Machine Learning: A Probabilistic Perspective*; MIT press, 2012.
(614) Musil, F.; De, S.; Yang, J.; Campbell, J. E.; Day, G. M.; Ceriotti, M. Machine Learning for

the Structure–Energy–Property Landscapes of Molecular Crystals. *Chem. Sci.* **2018**, *9* (5), 1289–1300.
(615) Noh, J.; Gu, G. H.; Kim, S.; Jung, Y. Uncertainty-Quantified Hybrid Machine Learning/Density Functional Theory High Throughput Screening Method for Crystals. *J. Chem. Inf. Model.* **2020**, *60* (4), 1996–2003.
(616) Park, C. W.; Wolverton, C. Developing an Improved Crystal Graph Convolutional Neural Network Framework for Accelerated Materials Discovery. *Phys. Rev. Mater.* **2020**, *4*, 063801.
(617) Issue #2: atom_init.json https://github.com/txie-93/cgcnn/issues/2.
(618) Cordero, B.; Gómez, V.; Platero-Prats, A. E.; Revés, M.; Echeverría, J.; Cremades, E.; Barragán, F.; Alvarez, S. Covalent Radii Revisited. *Dalt. Trans.* **2008**, No. 21, 2832–2838.
(619) Mentel, Ł. Mendeleev - A Python Resource for Properties of Chemical Elements, Ions and Isotopes. 2014.
(620) Hurter, C.; Ersoy, O.; Telea, A. Graph Bundling by Kernel Density Estimation. In *Computer Graphics Forum*; 2012; Vol. 31, pp 865–874.
(621) Anderson, R.; Gómez-Gualdrón, D. A. Increasing Topological Diversity during Computational "Synthesis" of Porous Crystals: How and Why. *CrystEngComm* **2019**, *21* (10), 1653–1665.
(622) Louis, S.-Y. M.; Zhao, Y.; Nasiri, A.; Wang, X.; Song, Y.; Liu, F.; Hu, J. Graph Convolutional Neural Networks with Global Attention for Improved Materials Property Prediction. *Phys. Chem. Chem. Phys.* **2020**, *22*, 18141–18148.
(623) Schütt, K. T.; Sauceda, H. E.; Kindermans, P.-J.; Tkatchenko, A.; Müller, K.-R. SchNet – A Deep Learning Architecture for Molecules and Materials. *J. Chem. Phys.* **2018**, *148* (24), 241722.
(624) Allen, M.; Poggiali, D.; Whitaker, K.; Marshall, T. R.; Kievit, R. A. Raincloud Plots: A Multi-Platform Tool for Robust Data Visualization. *Wellcome open Res.* **2019**, *4*.
(625) Chen, H. MOF Big Adsorbate Initializer https://github.com/snurr-group/mof-big-adsorbate-initializer.
(626) Hall, J. N.; Bollini, P. Low-Temperature, Ambient Pressure Oxidation of Methane to Methanol Over Every Tri-Iron Node in a Metal-Organic Framework Material. *Chem. Eur. J.* **2020**, *26* (70), 16639–16643.
(627) Hall, J.; Bollini, P. Role of Metal Identity and Speciation in the Low-Temperature Oxidation of Methane over Tri-Metal Oxo Clusters. *Authorea* **2021**. https://doi.org/10.22541/au.162027355.57700091/v1.
(628) Huang, N.-Y.; He, H.; Li, H.; Liao, P.-Q.; Chen, X.-M. A Metal–Organic Framework with in Situ Generated Low-Coordinate Binuclear Cu(I) Units as a Highly Effective Catalyst for Photodriven Hydrogen Production. *Chem. Commun.* **2020**, *56* (49), 6700–6703.
(629) Boyd, P. G.; Chidambaram, A.; García-Díez, E.; Ireland, C. P.; Daff, T. D.; Bounds, R.; Gładysiak, A.; Schouwink, P.; Moosavi, S. M.; Maroto-Valer, M. M.; Reimer, J. A.; Navarro, J. A. R.; Woo, T. K.; Garcia, S.; Stylianou, K. C.; Smit, B. Data-Driven Design of Metal--Organic Frameworks for Wet Flue Gas CO2 Capture. *Nature* **2019**, *576* (7786), 253–256.
(630) Martin, R. L.; Lin, L.-C.; Jariwala, K.; Smit, B.; Haranczyk, M. Mail-Order Metal–Organic Frameworks (MOFs): Designing Isoreticular MOF-5 Analogues Comprising Commercially

Available Organic Molecules. *J. Phys. Chem. C* **2013**, *117* (23), 12159–12167.
(631) Witman, M.; Ling, S.; Anderson, S.; Tong, L.; Stylianou, K. C.; Slater, B.; Smit, B.; Haranczyk, M. In Silico Design and Screening of Hypothetical MOF-74 Analogs and Their Experimental Synthesis. *Chem. Sci.* **2016**, *7* (9), 6263–6272.
(632) Anderson, R.; Gómez-Gualdrón, D. Deep Learning Combined with IAST to Screen Thermodynamically Feasible MOFs for Adsorption-Based Separation of Multiple Binary Mixtures. *J. Chem. Phys.* **2021**, *154*, 234102.

Appendix A. APPENDIX FOR CHAPTER 2.

A.1 VASP Parameters

A.1.1 VASP PBE Pseudopotentials

The VASP 5.4 pseudopotentials used in this work are listed in Table A.1. The only element that is not treated using the VASP-recommended projector-augmented wave (PAW) potentials is Li. The VASP-recommended Li_sv PAW potential has a default maximum energy cutoff ("ENMAX") of 499 eV, whereas all other elements have a default ENMAX cutoff of 400 eV or lower. Therefore, the standard Li PAW potential with ENMAX of 140 eV is used instead. This is especially important because it is generally recommended to use a plane-wave kinetic energy cutoff at least 1.3 times the largest ENMAX of all elements in the given structure to prevent Pulay stresses during volume relaxations.[136] With the PAW potentials listed in Table A.1, no material requires a cutoff greater than 520 eV when the cell volume or shape is changed.

Table A.1. VASP 5.4 recommended PAW PBE potentials. All elements not listed here used the standard (no suffix) PAW potentials. Note that not all of the elements shown in Table A.1 are included in the MOFs screened in this work.

Element	PAW potential suffix
Na	pv
K	sv
Ca	sv
Sc	sv
Ti	sv
V	sv
Cr	pv
Mn	pv
Ga	d
Ge	d
Rb	sv
Sr	sv
Y	sv
Zr	sv
Nb	sv
Mo	sv

Tc	pv
Ru	pv
Rh	pv
In	d
Sn	d
Cs	sv
Ba	sv
Pr	3
Nd	3
Pm	3
Sm	3
Eu	2
Gd	3
Tb	3
Dy	3
Ho	3
Er	3
Tm	3
Yb	2
Lu	3
Hf	pv
Ta	pv
W	sv
Tl	d
Pb	d
Bi	d
Po	d
At	d
Fr	sv
Ra	sv

A.1.2 Transition State Calculations in VASP

For transition states (TSs) computed in VASP, a two-step procedure is performed. First, the climbing image nudged elastic band (CI-NEB) method[130] is used to roughly map out the minimum energy pathway (MEP) and to get a reasonable estimate of the TS geometry. Then, a series of increasingly accurate dimer[131] calculations are performed until the TS is found.

In the initial CI-NEB step, we use the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) optimization algorithm, a force-convergence criterion of $|F_{\max}| < 1$ eV/Å, the default 400 eV plane-wave kinetic energy cutoff, and k-point density (KPPA) of 100/number of atoms for

rapid convergence. Four images were generally used in the initial CI-NEB attempt, but occasionally additional images were required in the case of poor convergence.

Following this initial CI-NEB calculation, a low-accuracy dimer calculation is done starting from the interpolated maximum of the MEP from the CI-NEB calculation. In this step, we use the following settings: a force-convergence criterion of $|F_{\max}| < 0.075$ eV/Å, the Fast Inertial Relaxation Engine (FIRE)[137] optimization algorithm, default 400 eV plane-wave kinetic energy cutoff, and 100 KPPA. Following the convergence of the low-accuracy dimer calculation, a medium-accuracy dimer calculation is performed with the same parameters but an increased k-point density of 1000 KPPA. Finally, a high-accuracy dimer calculation with a force-convergence criterion of $F_{\max} < 0.03$ eV/Å, 520 eV kinetic energy cutoff, and 1000 KPPA k-point grid is performed. The VTST package[138] v.3.2 is used to carry out the CI-NEB and dimer calculations.

A.1.3 Multi-Stage Optimization

The VASP-specific parameters used to implement the multi-stage optimization workflow shown in Figure 2.4 and Figure 2.6 are listed in Table A.2 for clarity. In addition to the parameters described in the main text, symmetry operations are disabled and the accurate-precision keyword in VASP is enabled. The choice of whether the k-point grid should be Γ-centered or not is determined based on how close the unit cell is to being hexagonal, as determined by Pymatgen.[82] After the test single-point calculation in the workflow, it is also determined if projection operators should be evaluated in reciprocal space or in real space based on the VASP log file. Evaluation in real space greatly reduces the calculation time but can occasionally influence the accuracy of the computed energies, so if projection operators are chosen to be evaluated in real space, the calculation will switch back to reciprocal space in the last stage of the relaxation.

Table A.2. Default VASP 5.4.1 parameters (using the corresponding ASE 3.16.2 keywords) implemented at each stage of the workflow.

Optimization Stage	VASP/ASE parameters	Notes
Test calculation	xc = 'PBE', setups = {'base':'recommended','Li':''}, ivdw = 12, prec = 'Accurate', algo = 'All', ediff = 1e-4, nelm = 150, nelmin = 6, lreal = False, ismear = 0, sigma = 0.01, lcharg = False, lwave = True, lorbit = 11, isym = 0, symprec = 1e-8, nsw = 0, istart = 0	
Atomic positions (low-accuracy; phase 1)	xc = 'PBE', setups = {'base':'recommended','Li':''}, ivdw = 12, prec = 'Accurate', algo = 'All', ediff = 1e-4, nelm = 150, nelmin = 6, lreal = {False or 'Auto'}, ismear = 0, sigma = 0.01, lcharg = False, lwave = True, lorbit = 11, isym = 0, symprec = 1e-8. ASE: BFGSLineSearch, fmax = 5	k-point grid has a density of 100 KPPA. gamma and lreal determined dynamically.
Atomic positions (low-accuracy; phase 2)	xc = 'PBE', setups = {'base':'recommended','Li':''}, ivdw = 12, prec = 'Accurate', algo = 'All', ediff = 1e-4, nelm = 150, nelmin = 6, lreal = {False or 'Auto'}, ismear = 0, sigma = 0.01, lcharg = False, lwave = True, ibrion = 2, isif = 2, nsw = 250, ediffg = -0.05, lorbit = 11, isym = 0, symprec = 1e-8	k-point grid has a density of 100 KPPA. gamma and lreal determined dynamically.
Atomic positions (medium-accuracy)	xc = 'PBE', setups = {'base':'recommended','Li':''}, ivdw = 12, prec = 'Accurate', algo = 'All', ediff = 1e-6, nelm = 150, nelmin = 8, lreal = {False or 'Auto'}, ismear = 0, sigma = 0.01, lcharg = False, lwave = True, ibrion = 3, iopt = 7, potim = 0, isif = 2, nsw = 500, ediffg = -0.05, lorbit = 11, isym = 0, symprec = 1e-8	k-point grid has a density of 1000 KPPA. gamma and lreal determined dynamically.
Atomic positions (high-accuracy)	xc = 'PBE', setups = {'base':'recommended','Li':''}, encut = 520, ivdw = 12, prec = 'Accurate', algo = 'All', ediff = 1e-6, nelm = 150, nelmin = 8, lreal = {False or 'Auto'}, ismear = 0, sigma = 0.01, lcharg = False, lwave = True, ibrion = 3, iopt = 7, potim = 0, isif = 2, nsw = 500, ediffg = -0.03, lorbit = 11, isym = 0, symprec = 1e-8	k-point grid has a density of 1000 KPPA. gamma and lreal determined dynamically.
Volume relaxation (low-accuracy)	xc = 'PBE', setups = {'base':'recommended','Li':''}, encut = 520, ivdw = 12, prec = 'Accurate', algo = 'All', ediff = 1e-6, nelm = 150, nelmin = 6, lreal	k-point grid has a density of 100 KPPA. gamma and lreal determined dynamically. nsw is set to 30 but relaxation

	= {False or 'Auto'}, ismear = 0, sigma = 0.01, lcharg = False, lwave = True, ibrion = 2, isif = 3, nsw = 30, ediffg = -0.03, lorbit = 11, isym = 0, symprec = 1e-8	repeats until force convergence.
Volume relaxation (high-accuracy)	xc = 'PBE', setups = {'base':'recommended','Li':''}, encut = 520, ivdw = 12, prec = 'Accurate', algo = 'All', ediff = 1e-6, nelm = 150, nelmin = 6, lreal = {False or 'Auto'}, ismear = 0, sigma = 0.01, lcharg = False, lwave = True, ibrion = 2, isif = 3, nsw = 30, ediffg = -0.03, lorbit = 11, isym = 0, symprec = 1e-8	k-point grid has a density of 1000 KPPA. nsw is set to 30 but relaxation repeats until force and volume convergence.
Final single-point calculation	xc = 'PBE', setups = {'base':'recommended','Li':''}, encut = 520, ivdw = 12, prec = 'Accurate', algo = 'All', ediff = 1e-6, nelm = 150, nelmin = 6, lreal = False, ismear = 0, sigma = 0.01, lcharg = True, lwave = True, laechg = True, lorbit = 11, isym = 0, symprec = 1e-8, nsw = 0	k-point grid has a density of 1000 KPPA.

A.1.4 Spin-Polarization

We note that it is still possible that the automated spin-state routine may occasionally not capture particularly complex ground-state magnetic orderings, especially those involving antiferromagnetism. We have found through regular use that the HT-DFT workflow occasionally – but not always – converges to a ground-state antiferromagnetic solution. For instance, $Cu_3(btc)_2$ has 12 Cu atoms per unit cell and is predicted to have a net magnetic moment of 2 μ_B via the HT-DFT workflow, whereas manual testing of all possible spin-states suggests an antiferromagnetic singlet state. Closer inspection of the individual magnetic moments shows that the HT-DFT workflow successfully resulted in antiferromagnetic alignment of 10 out of 12 of the Cu(II) species in the Niggli-reduced cell (the remaining two Cu atoms have 1 unpaired electron each, resulting in the 2 unpaired electrons per unit cell). This resulted in a higher energy structure by 7 kJ/mol (per unit cell). We therefore emphasize that, while the automated spin-treatment is quite robust for screening purposes, it is worth confirming the energetic ordering of neighboring spin-states in

more detail for top-performing candidates. Nonetheless, the impact on the predicted energies is likely to be relatively small in many cases, especially for screening purposes.

Another subtlety of any automated spin-polarization treatment with MOFs is that it is not uncommon for there to be multiple nearly isoenergetic spin-states. For instance, the predicted ground-state spin-state for YEGYIU with adsorbed methane (i.e. MOF−O+CH_4^*) is 0 μ_B whereas the same structure without adsorbed methane has a predicted ground-state with a net magnetic moment of 2 μ_B. This is unusual, as the physisorption of CH_4 should not drastically change the predicted spin of the oxidized complex. However, it was found that a net magnetic moment of 0 μ_B is 6 kJ/mol (per unit cell) less favorable than a net magnetic moment of 2 μ_B, likely due to extremely minor differences in the geometry with the adsorbed CH_4. These small energetic differences are not a major issue for the purposes of qualitatively ranking catalytic candidates but should be kept in mind when analyzing the resulting spin states from a HT-DFT study of MOFs. If converging to the true ground-state spin state is a concern, one method to minimize spin-related errors is to dope the given MOF with a metal cation that is known to have no unpaired electrons. For instance, if studying Fe-MOF-74, one could dope the MOF with Mg^{2+}, leaving only a single Fe^{2+} site for the oxidation catalysis.[177]

A.2 Automated Adsorption Energies

A.2.1 Coordination Numbers of Three or Less

For sites with a coordination number of three, the coordination environment is, by definition, planar. Some molecular geometries, such as T-shaped, have multiple geometrically plausible adsorption sites.[546] As such, for coordination numbers of three, Equations 2.7 and 2.8 in the main text are used to generate two initial adsorbate positions, and the site with the fewest neighboring

atoms within $r_{\text{cut}} = 2.5$ Å is retained. For sites with a coordination number of two, the concept of planarity is undefined, and the two coordinating atoms are, by definition, linear. As such, the direction perpendicular to the line connecting the two coordinating atoms is computed, and trial positions are considered in 10° rotations about this axis. Once again, the site with the fewest neighboring atoms within $r_{\text{cut}} = 2.5$ Å is retained. For sites with a coordination number of one, a linear initial geometry is assumed between the coordinating atom, metal center, and adsorbate. In the current work, this is only used for the addition of H atoms to mononuclear metal-oxo active sites, which are initialized with $\alpha = 1$ Å from the adsorbed O species.

A.2.2 Bonding Topology

The vector-sum algorithm requires the coordinates of the open metal site (OMS) and all coordinating atoms. For the purposes of HT screening, the open-source code Zeo++[109] was used to identify the presence of OMSs and their locations within the unit cell. The positions of the coordinating atoms can be determined using one of a variety of nearest neighbor (NN) algorithms, such as those implemented in Pymatgen[82] for determining local bonding topology.[108] In this work, we used Zeo++'s Voronoi network analysis[109] for calculating the coordination environment of OMSs and Pymatgen's valence ionic radius evaluator (VIRE) algorithm[108] for all other calculations of bonding topology. Based on preliminary testing, future studies may benefit from Pymatgen's CrystalNN algorithm, which was developed during the preparation of this manuscript.

A.2.3 Initializing Adsorbate Positions

For each MOF, the vector-sum algorithm is used to identify the potential O adsorption site at every OMS in the unit cell, and the one with the fewest neighbors within 2.5 Å is deemed ideal. In the case of multiple sites with the same number of neighboring atoms, the distance from the adsorption

site to the nearest atom is used as a tiebreaker, with larger distances (i.e. more vacant space) being more favorable. In the final step, it is confirmed that the O atom is at least 1.3 Å away from all other atoms in the unit cell. This eliminates MOFs with OMSs that do not have sufficient room for an O adsorbate, let alone an adsorbed CH_4 molecule. In the H addition step, an analogous procedure is performed except that the H atom is added to the chemisorbed O site instead of to an OMS, and an overlap tolerance of 0.75 Å is used instead of 1.3 Å.

A.3 Transition State Scaling Relationship

The universal TS scaling relationship reported in the work of Nørskov and coworkers[105] for the C–H bond activation of CH_4 is given as $E_{TS}^{*} = 0.75E_H + 1.09$ eV. In the aforementioned work, the descriptor E_H is defined based on formation energies that are referenced to gas-phase H_2O and O_2 (Equation A.1) at the BEEF-vdW[547] level of theory.

$$E_{H,H_2O\,ref} = E_{MOF-OH} - E_{MOF-O} - \frac{1}{2}E_{H_2O} + \frac{1}{4}E_{O_2} \quad (A.1)$$

While the choice of reference state is in principle arbitrary, this choice can impact the accuracy of subsequent studies that use different levels of theory if one of the reference species is not properly described by DFT. This is the case for the definition of E_H, as O_2 is poorly described by standard generalized gradient approximation (GGA) density functionals[220,221] and is particularly relevant in the present study because PBE-D3(BJ) predicts that $-\frac{1}{2}E_{H_2O} + \frac{1}{4}E_{O_2}$ is equal to 4.65 eV, whereas BEEF-vdW predicts that the same term is 4.75 eV. To circumvent this issue, we use H_2 as the H-source instead of H_2O such that

$$E_H = E_{MOF-OH} - E_{MOF-O} - \frac{1}{2}E_{H_2} \quad (A.2)$$

The modified TS scaling relationship is then given as $E_{\mathrm{TS}}^{*} = 0.75E_{\mathrm{H}} + 1.96$ eV, where E_{H} is now in terms of H_2. The potential energy diagram in Figure A.1 shows the relationship between E_{TS} and the activation energy E_{a} for C–H activation.[104] For three MOFs specifically highlighted in this work, we directly computed the C–H activation barrier and compared it against the value predicted via the universal TS scaling relationship proposed by Nørskov and coworkers.[105] As shown in Table A.3, the results are in reasonable agreement with the expected 11 kJ/mol mean absolute error.[105]

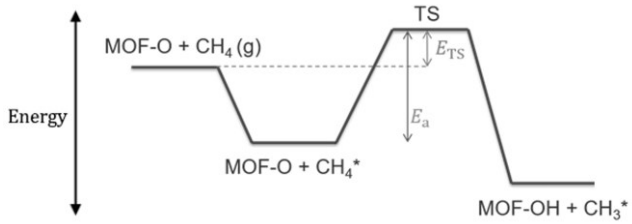


Figure A.1. The definition of E_{TS} for C–H activation is the energy of the transition state relative to the oxidized state. The activation energy, E_a, can be readily computed from E_{TS} if the adsorption energy of CH_4 is also known.

Table A.3. Comparison of methane C–H activation barrier predicted via the universal transition state scaling relationship, $E_{\mathrm{a,C-H}}^{*}$, and that computed directly in VASP, $E_{\mathrm{a,C-H}}$.

MOF	$E_{\mathrm{a,C-H}}^{*}$ (kJ/mol)	$E_{\mathrm{a,C-H}}$ (kJ/mol)
$Cu_8I_4(dmtrz)_4$	36	37
CFA-8	68	57
Rh-DMOF-1	89	79

A.4 MOF Dataset

The list of MOFs considered at each stage of the workflow (including the initial dataset construction process and subsequent DFT screening) is summarized in the "datasets.xlsx" Excel

workbook included with the optimized structures hosted on Zenodo.[548] The relevant details are described below.

A.4.1 Initial Dataset Construction

In the initial data cleansing step of the 838 CP2K-optimized CoRE MOFs, we removed 7 crystal structures that do not contain any carbon atoms as well as 12 additional MOFs with known erroneous structures.[409] Of the 819 pre-processed MOFs, 149 were identified as having at least 1 OMS as determined by a Voronoi network analysis performed using Zeo++.[109] There are an additional 92/820 MOFs that had at least 1 OMS in the original, unoptimized CoRE MOF database[89] that we include in the analysis as well, resulting in a total of 241 MOFs.

All 241 MOFs were converted to their Niggli-reduced cells to reduce the computational cost of the HT-DFT calculations. To remove potential duplicate structures, MOFs were flagged as duplicates if their root-mean-square deviation of atomic positions was less than 0.1 Å after the Niggli-reduced cells were scaled to have equivalent volumes.[82] A total of 22 duplicates were found using this algorithm, and an additional 2 were separately identified as duplicates by Smit and coworkers.[409] Finally, 49 MOFs had pore-limiting diameters less than 3.0 Å, as determined via Zeo++,[109] which we excluded from the dataset as well. This process therefore resulted in 168 unique MOFs with OMSs (prior to optimization in VASP) for HT-DFT screening. Despite the fact that the structures were previously converged using the PB3-D3(BJ) functional in CP2K, many of the structures were originally far from converged within VASP during the volume relaxation stage. This can potentially be attributed to the mixed Gaussian and plane-wave method in CP2K[113,114] compared to the all plane-wave approach in VASP,[122,123] although a number of MOFs started out with bond distances that were much smaller than would be expected at the PBE-D3(BJ) level of

theory. Inevitably, some additional MOFs in the CoRE MOF database have incomplete structural information, such as a lack of charge-balancing ions due to experimental difficulty in resolving their crystallographic positions. As such, we confirmed the validity of the structures for the most promising candidates.

A.4.2 Unique Open Metal Sites

Following the HT volume relaxation, 88/167 MOFs were found to have OMSs as detected via Zeo++. Since some MOFs may have multiple chemically distinct OMSs that can exhibit drastically different chemical reactivity, we considered oxidative C–H bond activation at each unique OMS separately. We defined a unique OMS as having a chemically unique coordination environment (defined as the chemical formula consisting of the atoms in the first coordination sphere and metal center), which resulted in 97 unique OMSs from 88 individual MOFs. Of these 97 unique OMSs, 88 could support an O adsorbate using the previously described adsorbate initialization workflow.

A.4.3 Validation of Converged Structures

Inevitably, some MOFs will not support an O, OH, or CH_4 adsorbate near the OMS following geometry optimization. For instance, the O adsorbate may instead bind to the linker or the O adsorbate may be so weakly bound to the metal center that the formation of an OH group results in the formation of a mobile OH radical. As such, a series of checks are performed at each stage of the HT-DFT workflow following the DFT calculations. For the O addition step, the VIRE algorithm[108] was used to confirm that the O adsorbate was bound to at least one atom and that the closest atom is a metal. This eliminates MOFs where the O adsorbate preferentially binds to the organic linker instead of the OMS following the geometry optimization. For the H addition step, the VIRE algorithm was used to confirm that the closest atom bound to the H adsorbate is the

expected O atom and was confirmed that the O adsorbate was still bound to the metal center. As a final check, the VIRE algorithm was used to confirm that the active site does not change its coordination environment after each step of the workflow, except for the adsorbate added to the system. At the end of the HT-DFT structure relaxations, 1 MOF did not reach convergence in a reasonable time frame, and 19 MOFs had potentially problematic structures following adsorption as determined using the aforementioned system of checks. This resulted in 68 unique OMS environments that could be analyzed for both oxidation and C–H activation. The MOFs with potentially problematic structures were excluded from the data analysis shown in the main text.

Appendix B. APPENDIX FOR CHAPTER 3

B.1 Additional Methods

B.1.1 MOF Dataset

In the following section, we describe the 60 MOFs considered in this work, why they were selected for investigation, and any pertinent modeling decisions for each MOF. For each MOF, a single metal site was considered for oxidation and subsequent C−H activation. Each of the 60 MOFs described below are included in all the correlations computed in this work, with the exception of Cr-BTC and Mo-BTC, which are only included in Figure 3.7.

M-MFU-4*l* (M = Cu^{+}). We studied MFU-4*l* (MFU = Metal−Organic Framework Ulm, *l* = large) with Zn(II)-Cl species post-synthetically exchanged with Cu(I) species since this MOF has been shown to chemisorb O_2.[27] MFU-4*l* is constructed from btdd (btdd^{2-} = bis(1H-1,2,3-triazolo[4,5-b],[4′,5′-i])dibenzo[1,4]dioxin) linkers and has a cubic topology. On average, up to two Cu species can be exchanged per pentametallic node,[27] so we studied both Cu_1:Zn_4 and Cu_2:Zn_3 per node. We also considered a mixed-metal Cu_1:Co_3:Zn_1 pentametallic node, as it is possible to exchange up to four of the Zn(II)-Cl species with Co(II)-Cl species.[212,549]

M_2Cl_2(btdd) (M = Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}). M_2Cl_2(btdd)[192,193], also occasionally referred to as MAF-X*n*l-Cl (MAF = Metal−Azolate Framework; *n* = 25 for M = Mn and *n* = 27 for M = Co),[275] was studied in this work due to the fact that Co_2Cl_2(btdd) can be reversibly oxidized by Cl_2 and Br_2.[25] M_2Cl_2(btdd) has the same linker as MFU-4*l* but a different coordination environment around each metal site and has a hexagonal – instead of cubic – topology.

M_2Cl_2(bbta) (M = Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}). M_2Cl_2(bbta) (bbta^{2-} = 1H,5H-benzo(1,2-d:4,5-d′)bistriazole),[193,194] also occasionally referred to as MAF-X*n*-Cl (*n* = 25 for M = Mn and *n* = 27 for M = Co),[194] was studied in this work due to its structural similarity with M_2Cl_2(btdd).

M(L-*RR*) (M = Co^{2+}). We studied Co(L-*RR*) (L-*RR* = (*R*,*R*)-thiazolidine-2,4-dicarboxylate)[211] since it has been experimentally shown to chemisorb O_2.[215]

M-BTT (M = Cr^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+}, Cd^{2+}). We studied M-BTT (BTT^{3-} = 1,3,5-tris(1H-1,2,3-triazol-5-yl)benzene)[22,195–198] since Cr-BTT and the closely related MOFs Co-BTTri and Co-BTTriP can chemisorb O_2.[22,216] We included three Na^+ cations per unit cell in the pores of M-BTT to balance the charge, as in prior work.[550] In addition, all but one metal per unit cell was modeled as Mg^{2+} to simplify convergence to the ground-state spin state.

M-BTP (M = Ni^{2+}, Cu^{2+}). We studied M-BTP (BTP^{3-} = 1,3,5-tri(1H-pyrazol-4-yl)benzene)[148] due to its structural similarity with M-BTT. Unlike M-BTT, M-BTP does not require charge-balancing ions. To compare with M-BTT, we modeled all but one metal per unit cell as Mg^{2+}. For comparison purposes, we also modeled Ni-BTP with all Ni^{2+} cations in the unit cell. The energy differences were found to be small between Ni-BTP and its Mg-diluted analogue.

M_2(dobdc) (M = Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}). M_2(dobdc) (dobdc^{4-} = 2,5-dioxido,1,4-benzenedicarboxylate),[168,201,202] also referred to as MOF-74 or CPO-27 (CPO = Coordination Polymer of Oslo), was studied in this work, as Fe-MOF-74 shows selective chemisorption of O_2 and can oxidize ethane to ethanol in the presence of N_2O.[24,168] Mg-diluted Fe-MOF-74 has also been shown to oxidize ethane to ethanol with higher selectivity to ethanol than Fe-MOF-74[168] and is therefore included in the present work as well (with a Mg:Fe ratio of 17:1).

M_2(*m*-dobdc) (M = Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Zn^{2+}). M_2(*m*-dobdc) (*m*-$dobdc^{4-}$ = 4,6-dioxido-1,3-benzenedicarboxylate)[203] was studied in this work due to its structural similarity with M_2(dobdc). The metal sites in M_2(*m*-dobdc) are expected to have an increased charge density compared to those in M_2(dobdc).[203]

M_2(dobpdc) (M = Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}). M_2(dobpdc) ($dobpdc^{4-}$ = 4,4'-dioxidobiphenyl-3,3'-dicarboxylate)[204] was studied in this work due to its structural similarity with M_2(dobdc).

M_2(dsbdc) (M = Mn^{2+}, Fe^{2+}). M_2(dsbdc) ($dsbdc^{4-}$ = 2,5-disulfidobenzene-1,4,-dicarboxylate)[205,206,551] was studied in this work due to its structural similarity with M_2(dobdc). Unlike M_2(dobdc), there are bridging S atoms connecting the metal sites in M_2(dsbdc). The coordination geometry of Mn_2(dsbdc) differs from Fe_2(dsbdc) and MOF-74.[206]

M-MIL-88B (M = Sc^{3+}, Ti^{3+}, Mg^{2+}/V^{3+}, $Cr^{2+/3+}$, $Fe^{2+/3+}$). M-MIL-88B (MIL = Materials of Institut Lavoisier),[199] which has 1,4-benzenedicarboxylate (bdc) linkers, was studied in this work as a more computationally tractable alternative to MIL-100/101.[20,552,553] Both Fe-MIL-100/101 and Cr-MIL-100/101 have been shown to be active for oxidation catalysis.[213,214] In addition, Ti-MIL-101, V-MIL-100, V-MIL-101 can strongly chemisorb O_2.[20,217] While as-synthesized Fe- and Cr-MIL-100 have transition metals in the 3+ oxidation state, it is possible to reduce one of the Cr(III) or Fe(III) sites per node to Cr(II) or Fe(II) via removal of the charge-balancing anion.[356,553] We modeled the Fe- and Cr-containing systems with one M(II) site per unit cell. In contrast, for the Sc(III)- and Ti(III)-containing systems, the formal oxidation state of each metal species is kept at the as-synthesized 3+ state in the bare MOF. Inspired by the synthesis of a MIL-88 type structure with Mg_2V heterometallic nodes,[200] we also considered the analogous Mg_2V MIL-88B structure. This MOF has two Mg(II) cations and one V(III) cation per node, with a μ_3-OH^- group to balance

the charge (instead of a μ_3-O^{2-} species). Formation of the metal-oxo site is modeled as occurring at the formally Cr(II), Fe(II), Sc(III), Ti(III), or V(III) sites.

$M_3(btc)_2$ ($M = Cr^{2+}$, Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Mo^{2+}, $Ru^{2+/3+}$, Pd^{2+}). $M_3(btc)_2$ (BTC^{-3} = 1,3,5-benzenetricarboxylate)[115,185–191], referred to as simply M-BTC, was studied in this work due to the fact that Cr-BTC can strongly chemisorb O_2.[185] Since the Ru cations in Ru-BTC are known to exist in a mixed 2+/3+ valence state,[190] we modeled Ru-BTC with half Ru(II) species and the other half Ru(III)-Cl species, the former of which are considered for oxidation. Fe-BTC was excluded from this work since its structure is not fully understood.[554] Due to anomalous bond-breaking events discussed in the main text, Cr-BTC and Mo-BTC are excluded from all analyses and models developed in this work except where specifically highlighted.

M-MOF-5 ($M = V^{2+}$, Cr^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}). Cation-exchanged MOF-5,[207–210] also referred to as IRMOF-1 (IRMOF = IsoReticular MOF), was studied in this work, as Cr(II)-MOF-5 and Fe(II)-MOF-5 have been shown to be accessible to inner- and outer-sphere oxidants, respectively.[207] The MOFs were modeled with M_1:Zn_3 per node based on prior experimental characterization.[207]

$Cu_3(dmtrz)_2(ox)_2$, **$Cu_8I_4(dmtrz)_4$**, **Cu-CFA-8**. These three MOFs were highlighted in prior work[32] and are included in the present study to further test the general applicability of the proposed scaling relationship. $Cu_3(dmtrz)_2(ox)_2$ ($dmtrz^-$ = 3,5-dimethyl-1H-1,2,4-triazole, ox = oxalate) has mono(μ-aquo)Cu(II) dimers that can be desolvated to leave behind adjacent square-planar Cu(II) sites.[142] $Cu_8I_4(dmtrz)_4$ contains exposed, two-coordinate Cu(I) cations.[155] Cu-CFA-8 (CFA = Coordination Framework Augsburg), with the formula $Cu_2(TQPT)$ ($TQPT^{2-}$ = 6,6,14,14-tetramethyl-6,14-dihydroquin-oxalino-[2,3-*b*]phenazinebistriazole), has two-coordinate, three-

coordinate, and four-coordinate Cu(I) species, from which we modeled a mono(μ-oxo) motif between the first two Cu(I) species.[156] While all the other MOFs in this work have mononuclear metal-oxo active sites, $Cu_3(dmtrz)_2(ox)_2$ and Cu-CFA-8 are modeled as having binuclear motifs. In addition, $Cu_8I_4(dmtrz)_4$ is the only MOF studied in this work where we modeled oxidation at a two-coordinate transition metal cation. Despite these unique differences, the three MOFs appear to follow the proposed scaling relationship. Due to the binuclear motifs in $Cu_3(dmtrz)_2(ox)_2$ and Cu-CFA-8, physicochemical properties such as DDEC6 charges and magnetic moments on the two Cu atoms are averaged, and the M−O bond order is the sum of the bond orders for the O adsorbate with the two Cu sites.

B.1.2 DFT Screening Procedure

The full details of the periodic DFT workflow are reported and benchmarked in prior work.[32] Below, we summarize the essential details not already mentioned in the Methods section.

A multi-stage optimization procedure is performed for both volume and ionic relaxations, as summarized in Tables B.4 and B.5. For all elements in this work, we used the VASP-recommended 5.4 projector-augmented wave (PAW) PBE potentials.[32,337]

Table B.1. Periodic DFT parameters used in optimizing the bare MOFs. ENCUT is the plane-wave kinetic energy cutoff, KPPA is the k-point density (with the corresponding k-point grids generated using Pymatgen[82]), and F_{max} is the force-convergence criterion.

Stage	ENCUT (eV)	KPPA	F_{max} (eV/Å)
1: Atomic positions (low-accuracy)	400	100	0.05
2: Volume relaxation (low-accuracy)	520	100	0.05
3: Volume relaxation (high-accuracy)	520	1000	0.03
4: Atomic positions (high-accuracy)	520	1000	0.03
5: Single-point	520	1000	N/A

(high-accuracy)			

Table B.2. Periodic DFT parameters used in optimizing MOFs with adsorbates. ENCUT is the plane-wave kinetic energy cutoff, KPPA is the k-point density (with the corresponding k-point grids generated using Pymatgen[82]), and $F_{\max}$ is the force-convergence criterion.

Stage	ENCUT (eV)	KPPA	$F_{\max}$ (eV/Å)
1: Atomic positions (low-accuracy)	400	100	0.05
2: Atomic positions (medium-accuracy)	400	1000	0.05
3: Atomic positions (high-accuracy)	520	1000	0.03
4: Single-point (high-accuracy)	520	1000	N/A

As described in detail in prior work,[32] we considered both a high-spin (5 μ_B per metal cation) and low-spin (0.1 μ_B per metal cation) initialization of magnetic moments, which were then optimized within VASP to find a local minimum energy spin state and structure. In addition, we explicitly considered the antiferromagnetic singlet spin state for MOFs where this is plausible. We found that M_2Cl_2(btdd) (M = Mn, Co), M_2Cl_2(bbta) (M = Mn, Co), M_2(dobdc) (M = Mn, Ni, Cu), Mn_2(*m*-dobdc), Mn_2(dobpdc), Mn_2(dsbdc), and $M_3(btc)_2$ (M = Cr, Cu) had antiferromagnetic singlet ground states, some of which have been previously confirmed experimentally.[28,473] The structure with the lowest energy following structural relaxation was used for computing the energetic quantities and physicochemical properties.

For honeycomb-shaped MOFs that exhibit antiferromagnetism, we used a unit cell slightly larger than the Niggli-reduced unit cell to ensure that an even number of atoms could be antiferromagnetically aligned along the infinite 1D chain of metal cations. All other MOFs were modeled using the Niggli-reduced unit cell as computed using Pymatgen.[82] If an antiferromagnetic

state was found to be isoenergetic (i.e. energy differences on the order of ~1 kJ/mol or less) with the ferromagnetic state, we considered the ferromagnetic state for computational simplicity.

To confirm that the change in spin states did not vary in unphysical ways throughout the reaction mechanism, we identified any MOFs where the net spin had a large change between reaction steps and subsequently considered all relevant spin states. Specifically, we analyzed any MOFs where the net magnetic moment changed by greater than two between the bare MOF and the oxidized state, greater than one between the oxidized and reduced active site, and greater than zero between the oxidized state and the state with adsorbed methane. If a lower energy spin state was found, this was adopted as the ground state. In many cases, the anticipated spin state was often isoenergetic with another spin state. The converged spin states in addition to the individual magnetic moments for each MOF are reported in the supporting data.[235]

The electronic energies for gas-phase species considered in this work as well as the energies of other relevant gas-phase oxidants are listed in Table B.6. Energies for the gas-phase species were computed by centering a single molecule in a simulation unit cell with 20 Å of vacuum space in all dimensions, a 650 eV cutoff, a k-point grid consisting of just the Γ-point, and Gaussian smearing with a smearing width of 0.001 eV. The vibrational frequency analyses were carried out using a central finite difference approximation of the Hessian matrix with displacements of ±0.01 Å in x, y, and z and a tight electronic energy convergence criterion of 10^{-8} eV to ensure accurate forces. The use of a 650 eV (as opposed to 520 eV) cutoff has a negligible effect on the gas-phase energies but was required to achieve more tightly converged vibrational frequencies.

Since triplet O_2 is poorly described by standard generalized gradient approximation (GGA)-level density functionals,[220–222] we correct for the value of E_{O_2} by ensuring that the reaction

H_2 (g) + ½ O_2 (g) → H_2O (g) has a standard enthalpy of reaction of −2.5061 eV at 298.15 K, as determined from experiment[141] and analogous to what has been done in numerous prior computational catalysis studies.[555–558] Based on a previous benchmarking study, this procedure is expected to greatly increase the accuracy of the PBE-D3(BJ) functional in describing chemisorption energies.[222] The vibrational frequencies of the gas-phase species at the PBE-D3(BJ) level of theory that were used in computing the corrected electronic energy for O_2 based on ideal gas statistical mechanics[293] are listed in Table B.7.

We also note that because of the challenge in describing the ground state energy of O_2, we use H_2 as the reference state for ΔE_{H} instead of H_2O and O_2, which is why we use the modified universal H-affinity scaling relationship of $E_{\mathrm{TS,C-H}} = 0.75\Delta E_{\mathrm{H}} + 1.96$ eV instead of $E_{\mathrm{TS,C-H}} = 0.75\Delta E_{\mathrm{H}} + 1.09$ eV as in the original work by Nørskov and coworkers.[105] The former equation is obtained by taking the latter and changing the reference state to the more reliably computed value based on gas-phase H_2 (at the same level of theory used in the original work by Nørskov and coworkers[105]). By avoiding the use of O_2 in the reference state for ΔE_{H}, this makes the scaling relationships more transferable for use with other density functionals that may treat triplet O_2 differently.[222]

Table B.3. Electronic energies of relevant gas-phase species at the PBE-D3(BJ) level of theory. The enthalpies at 298.15 K used to compute the O_2 correction are also shown.

Species	E (eV)	$H_{298.15\ \mathrm{K}}$ (eV)
CH_4	−24.07	---
O_2 [uncorrected]	−9.87	−9.69
O_2 [corrected]	−9.48	−9.29
H_2O	−14.24	−13.57
H_2	−6.78	−6.42
H_2O_2	−18.17	---
N_2O	−21.44	---

N_2	−16.66	---

Table B.4. DFT-computed vibrational frequencies of gas-phase species considered in this work computed at the PBE-D3(BJ) level of theory.

Species	Frequencies (cm^{-1})
O_2	1566.2
H_2O	1579.4, 3720.6, 3832.7
H_2	4318.5

While Nørskov and coworkers have previously shown that the universal H-affinity scaling relationship is an accurate descriptor for predicting $E_{\mathrm{TS,C-H}}$ for a wide range of materials,[105] we confirmed the accuracy of the scaling relationship for select MOFs chosen to represent a range of $E_{\mathrm{a,C-H}}$ values and structures. Transition state calculations were performed in a multi-stage workflow, starting with a climbing image nudged elastic band (CI-NEB)[130] calculation followed by a series of dimer[131] calculations starting from the predicted maximum of the minimum energy pathway.[32] The settings for the high-accuracy dimer calculation are the same as the high-accuracy ionic relaxation shown in Table B.5. As expected from the findings of Nørskov and coworkers, we observe that the computed transition state energies are all in good agreement with those predicted from the H-affinity scaling relationship (Table B.8). For reference, Nørskov and coworkers found the MAE in the H-affinity scaling relationship to be approximately 0.11 eV,[105] which is consistent with the findings in Table B.8.

Table B.5. Comparison of methane C−H activation barrier predicted via the universal transition state scaling relationship, $E_{\mathrm{a,C-H}}$, and that computed directly in VASP, $E_{\mathrm{a,C-H-VASP}}$.

MOF	$E_{\mathrm{a,C-H}}$ (kJ/mol)	$E_{\mathrm{a,C-H-VASP}}$ (kJ/mol)	Absolute deviation (eV)
Ni_2Cl_2(btdd)	0.35	0.24	0.11
Cu_8I_8(dmtrz)$_4$	0.37	0.38	0.01
Cu_1:Zn_4-MFU-4*l*	0.53	0.57	0.04

Cu-CFA-8	0.71	0.59	0.12
Mo-BTC	1.34	1.40	0.06
Cr-BTC	1.49	1.52	0.03

All structural visualizations were generated using Virtual NanoLab,[559] except for Figure B.7, which was made using VESTA.[560] The Matplotlib,[561] NumPy,[562] Pandas,[563] and Statsmodels[564] Python libraries were used for data visualizations and data analysis. Chargemol[374] was used to compute the DDEC6 charges, spin densities, and bond orders.[158,231] VASP was compiled with VTST[138] to perform geometry optimizations with the Fast Inertial Relaxation Engine (FIRE) algorithm[137] and to perform transition state calculations. The VTST package was also used to compute Bader[159] charges and spin densities.

B.2 Summarized DFT Screening Results

For a list of all the computed energies and physicochemical properties for each MOF, please refer to the "energies.csv" file uploaded with the supporting data.[235] For ease of reference, we also show the key energies for each MOF screened in this work in Table B.1.

Table B.6. Tabulated results from the periodic DFT screening process of 60 MOFs.

MOF	ΔE_{O} (eV)	ΔE_{H} (eV)	$E_{\mathrm{TS,C-H}}$ (eV)	$\Delta E_{\mathrm{CH_4\ ads}}$ (eV)	$E_{\mathrm{a,C-H}}$ (eV)
Cd-BTT	1.93	-2.81	-0.15	-0.21	0.06
Co-BTC	0.03	-1.91	0.53	-0.17	0.70
Co-BTT	0.25	-2.16	0.34	-0.21	0.55
Co(L-*RR*)	-0.64	-1.54	0.81	-0.29	1.09
Co-MOF-5	-0.16	-1.49	0.84	-0.15	0.99
Co-MOF-74	-0.72	-1.43	0.88	-0.16	1.05
Co_2(dobpdc)	-0.69	-1.47	0.86	-0.18	1.04
Co_2(*m*-dobdc)	-0.61	-1.72	0.67	-0.22	0.89
Co_2Cl_2-BBTA	-0.79	-1.71	0.68	-0.20	0.88
Co_2Cl_2-btdd	-0.85	-1.67	0.71	-0.18	0.89
Cr-BTT	-2.08	-0.96	1.24	-0.17	1.41
Cr-MIL-88B	-1.70	-1.09	1.14	-0.13	1.27
Cr-MOF-5	-2.37	-1.13	1.11	-0.13	1.25
Cu-BTC	1.82	-2.45	0.12	-0.20	0.32
Cu-BTP	1.63	-2.62	-0.01	-0.14	0.13
Cu-BTT	2.04	-2.65	-0.03	-0.14	0.11
Cu-CFA-8	0.44	-1.96	0.49	-0.21	0.71
Cu_8I_4(dmtrz)$_4$	1.03	-2.37	0.18	-0.19	0.37
Cu_3(dmtrz)$_2$(OX)$_2$	1.27	-2.41	0.15	-0.26	0.41
Cu_1:Zn_4-MFU-4*l*	0.07	-2.08	0.40	-0.13	0.53
Cu_2:Zn_3-MFU-4*l*	0.08	-2.07	0.41	-0.11	0.52
Cu_2:Co_2:Zn_1-MFU-4*l*	0.05	-2.08	0.40	-0.12	0.52
Fe-BTT	-1.05	-1.72	0.67	-0.16	0.84
Fe-MIL-88B	-0.50	-1.46	0.87	-0.13	1.00
Fe-MOF-5	-1.06	-1.24	1.03	-0.14	1.17
Fe-MOF-74	-1.16	-1.37	0.93	-0.17	1.11
Fe/Mg-MOF-74	-1.32	-1.42	0.89	-0.19	1.08
Fe_2(dobpdc)	-1.31	-1.39	0.92	-0.18	1.10
Fe_2(dsbdc)	-1.09	-1.24	1.03	-0.17	1.20
Fe_2(*m*-dobdc)	-1.42	-1.45	0.87	-0.17	1.04
Fe_2Cl_2-BBTA	-1.19	-1.28	1.00	-0.18	1.17
Fe_2Cl_2-btdd	-1.28	-1.28	1.00	-0.18	1.18
Mn-BTT	-1.25	-1.46	0.87	-0.16	1.03
Mn-MOF-5	-0.91	-1.49	0.84	-0.12	0.97
Mn-MOF-74	-1.20	-1.54	0.81	-0.18	0.99
Mn_2(dobpdc)	-1.18	-1.55	0.80	-0.18	0.98
Mn_2(dsbdc)	-1.02	-1.57	0.79	-0.16	0.94

Mn_2(*m*-dobdc)	-1.17	-1.51	0.83	-0.22	1.05
Mn_2Cl_2-BBTA	-1.08	-1.53	0.81	-0.18	0.99
Mn_2Cl_2-btdd	-1.10	-1.54	0.80	-0.20	1.00
Ni-BTC	0.40	-2.04	0.43	-0.19	0.63
Ni-BTP	1.19	-2.45	0.12	-0.12	0.24
Ni/Mg-BTP	1.11	-2.43	0.14	-0.15	0.29
Ni-BTT	1.42	-2.27	0.26	-0.16	0.41
Ni-MOF-5	0.48	-2.20	0.31	-0.10	0.40
Ni-MOF-74	0.25	-2.22	0.30	-0.22	0.51
Ni_2(dobpdc)	0.41	-2.09	0.39	-0.22	0.61
Ni_2(*m*-dobdc)	0.29	-2.26	0.26	-0.17	0.43
Ni_2Cl_2-BBTA	0.53	-2.38	0.18	-0.19	0.36
Ni_2Cl_2-btdd	0.50	-2.40	0.16	-0.18	0.35
Pd-BTC	1.21	-2.34	0.21	-0.18	0.38
Ru-BTC	0.03	-1.90	0.53	-0.18	0.71
Sc-MIL-88B	2.06	-2.75	-0.10	-0.14	0.04
Ti-MIL-88B	-3.45	-0.52	1.57	-0.15	1.72
V-MIL-88B	-2.89	-0.49	1.59	-0.12	1.71
V-MOF-5	-3.69	-0.13	1.86	-0.14	2.00
Zn-BTC	1.69	-2.75	-0.10	-0.19	0.09
Zn_2(*m*-dobdc)	1.79	-2.77	-0.12	-0.20	0.08
Cr-BTC	-1.03	-0.85	1.32	-0.17	1.49
Mo-BTC	-0.48	-1.05	1.17	-0.17	1.34

B.3 Supporting Figures

B.3.1 Methane Adsorption

The barrier for methane C−H activation, $E_{\mathrm{a,C-H}}$, as a function of the energy of the transition state (TS) for C−H activation, $E_{\mathrm{TS,C-H}}$, is shown in Figure B.1. Based on the definition of these energetic parameters (Figure 3.1c), the near-unity slope with $r^2 = 0.99$ indicates that the energy of methane adsorption at the metal-oxo species of each MOF is nearly constant. The methane adsorption energies for each MOF are summarized in the histogram in Figure B.1b. As expected from the intercept of the best-fit line in Figure B.1a, the average methane adsorption energy is −0.18 eV, with relatively small deviations from this value. A near-constant methane adsorption energy was previously observed for trinuclear motifs in ZSM-5 with an average value of approximately −0.24 eV (also using the PBE-D3(BJ) functional).[219] The slightly more exothermic methane adsorption energies in this case can be attributed to the small micropores of the zeolite. In the present work, Co(L-*RR*), shown in Figure B.2, has a methane adsorption energy of −0.28 eV due to its small pore size. Based on the experimental crystal structure, this MOF has a pore-limiting diameter of 3.96 Å.[211,215]

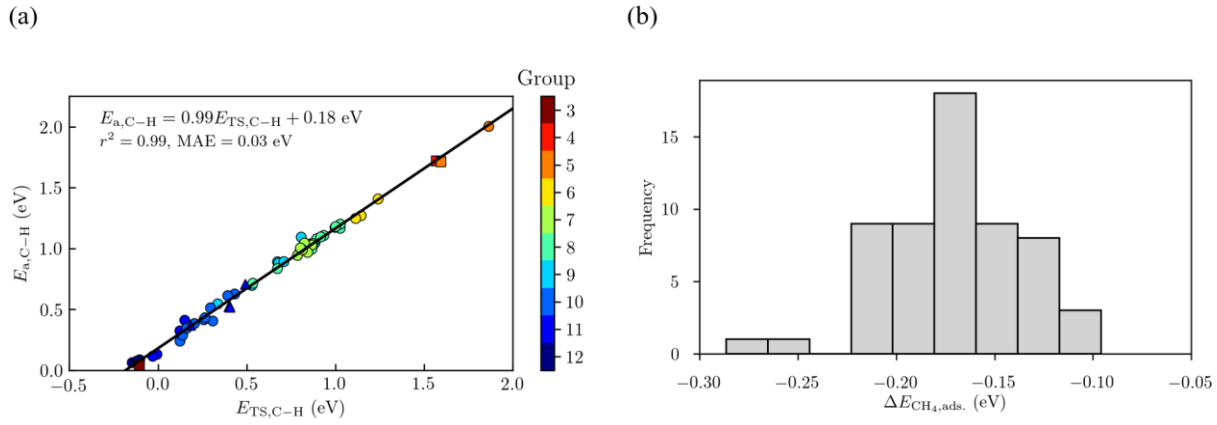


Figure B.1. a) Predicted barrier for C−H bond activation of methane, $E_{a,C-H}$, as a function of the corresponding energy of the transition state, $E_{TS,C-H}$. The best-fit line, $E_{a,C-H} = 0.99E_{TS,C-H} + 0.18\,eV$, has $r^2 = 0.99$ and a mean absolute error (MAE) of 0.03 eV. The barrier for C−H

activation is computed from Equation 3.2. Symbol color refers to the metal group number. Symbol shape indicates the formal oxidation state of the metal site prior to oxidation as 1+ (▲), 2+ (●), or 3+ (■). b) Histogram of methane adsorption energies at the metal-oxo active sites of the screened MOFs.

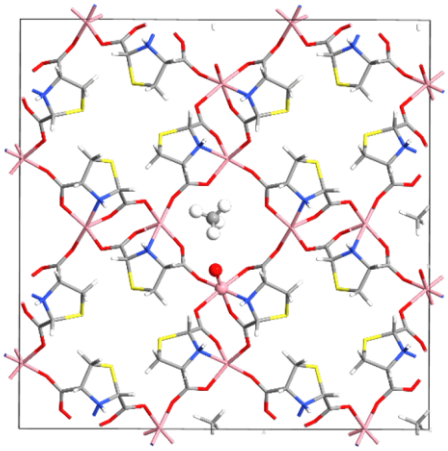


Figure B.2. $3 \times 3 \times 3$ supercell of Co(L-RR)[215] with adsorbed methane, highlighting the small pores around the proposed $[CoO]^{2+}$ active site. Color key: Co (pink), O (red), N (blue), S (yellow), C (gray), H (white).

B.3.2 Additional Physicochemical Descriptors

As shown in Figure 3.6a, ΔE_H and ΔE_O are moderately correlated with the spin density on the O species, ρ_O. Figure B.3a plots the same data but using a Bader partitioning[159] of the spin density, rather than the DDEC method. Because the Bader and DDEC spin densities are correlated to one another, analogous behavior is observed, although the correlation between Bader spin density and ΔE_H is slightly weaker than with spin densities computed from the DDEC method (r^2 = 0.69 compared to r^2 = 0.73). Since ρ_O is inversely correlated with the M−O bond order computed via the DDEC method (Figure B.4), we find that the results in Figure 3.6a are qualitatively similar to the bond order trends in Figure B.3b. A correlation matrix relating many of the computed physicochemical properties is shown in Figure B.5.

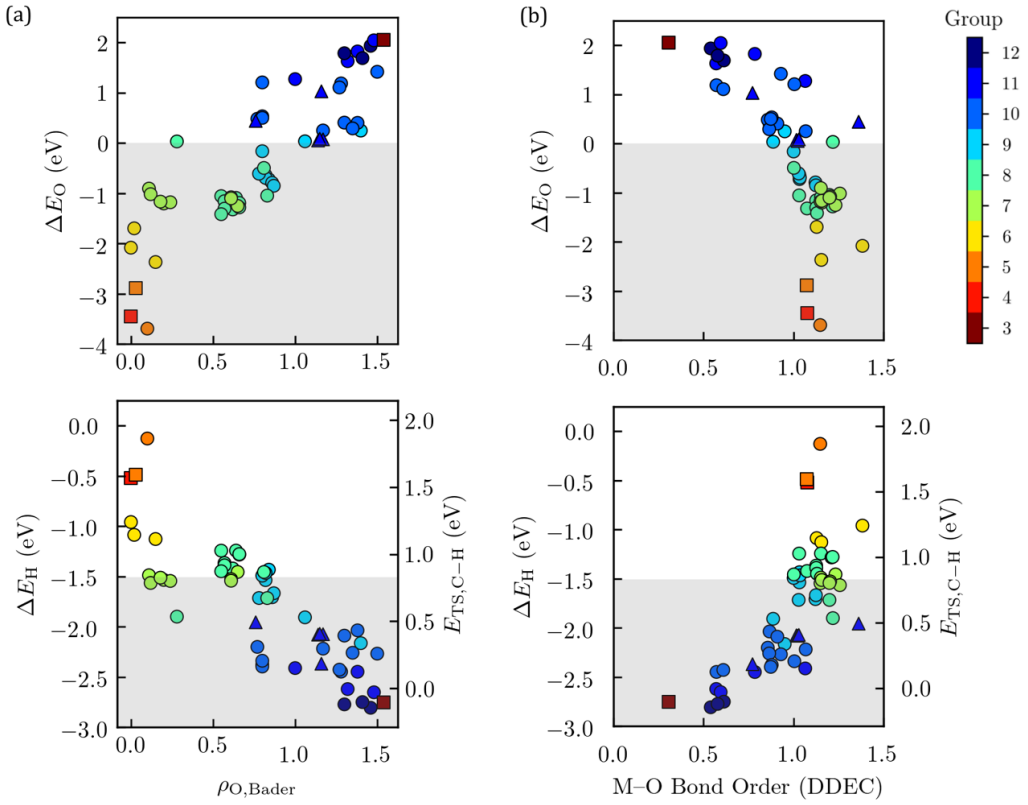


Figure B.3. a) Active site formation energy, ΔE_O, and H-affinity, ΔE_H, as a function of the absolute spin density on the oxo ligand computed via the Bader partitioning method,[159] $\rho_{O,Bader}$. b) ΔE_O and ΔE_H as a function of the (total) bond order between the metal binding site(s) and oxo ligand computed via the DDEC6 method.[231] The gray shaded area refers to the "stable, reactive" region in Figure 3.4 (i.e. $\Delta E_O < 0$ eV and $\Delta E_H < -1.51$ eV, the latter of which corresponds to approximately $E_{a,C-H} < 1$ eV). Both $\rho_{O,Bader}$ and the M–O bond order are computed for the oxidized state of the MOF. Symbol color refers to the group number of the metal in the periodic table. Symbol shape indicates the formal oxidation state of the metal site prior to oxidation as 1+ (▲), 2+ (●), or 3+ (■).

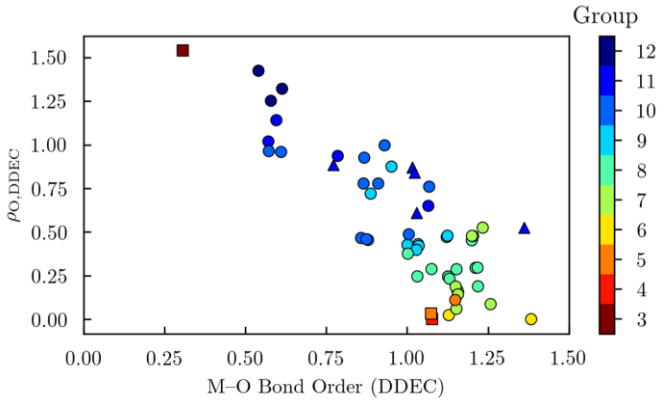


Figure B.4. Absolute value of the DDEC6 spin density on the oxo species of the M−O active site, $\rho_{O,DDEC}$, as a function of the DDEC6 M−O bond order.

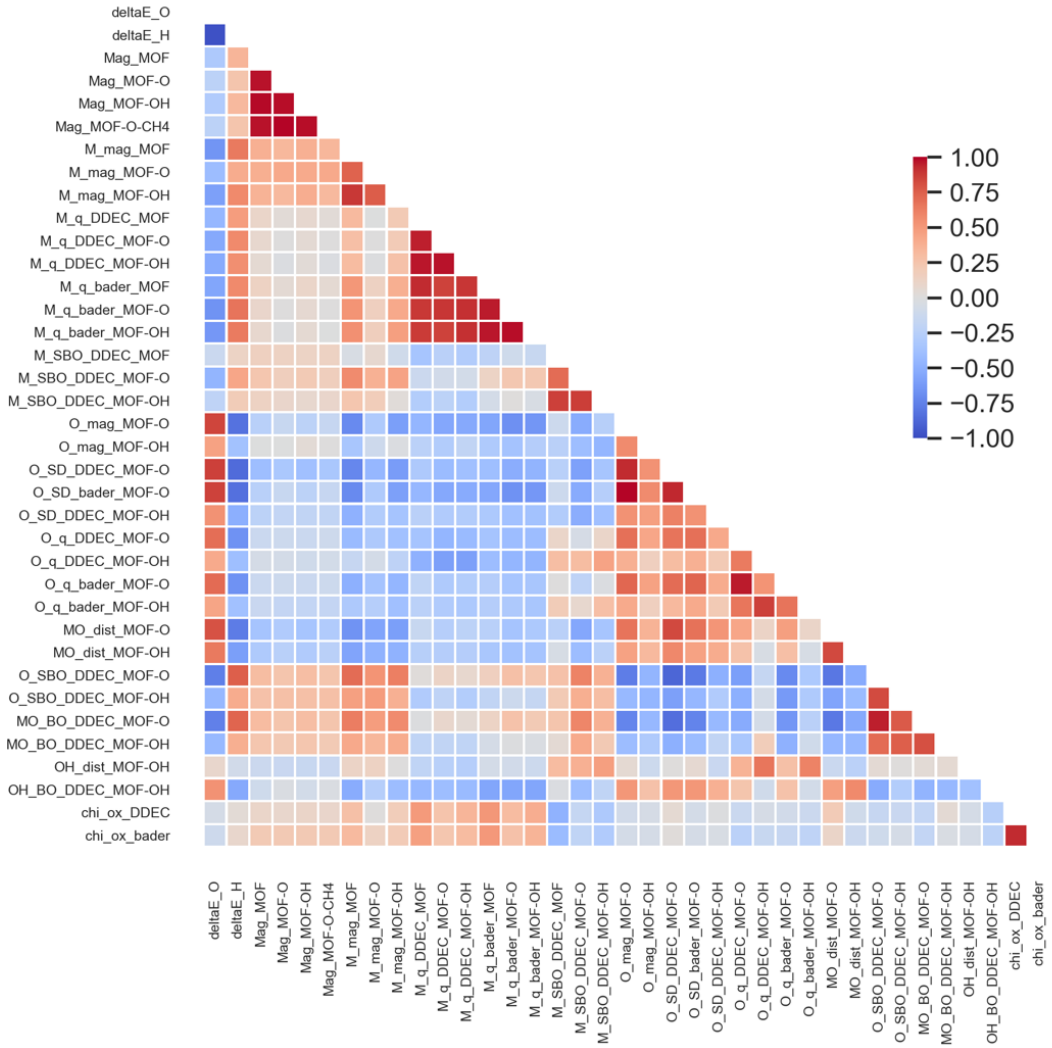


Figure B.5. Pearson correlation coefficients of the computed physicochemical properties for the screened MOFs. The definitions for each property can be found in the "README.txt" file uploaded with the supporting data.

B.3.3 M-MOF-5

As previously mentioned, we studied M-MOF-5 (M = V^{2+}, Cr^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}) for the oxidative C−H bond activation of methane. Four of the six MOF-5 variants (M = Co, Cr, Fe, V) exhibit deviations in ΔE_H of greater than ±0.20 eV from the scaling line shown in Figure 3.3, with

Co-MOF-5 having the largest deviation of all MOFs screened in this work (aside from Cr-BTC and Mo-BTC) with a deviation of +0.36 eV. Consistent with prior work by Dincă and coworkers,[207] we found significant structural distortion in the first coordination sphere of the metal site upon oxidation (Figure B.6). It is possible that the change from tetrahedral to a highly distorted square pyramidal-like geometry upon activation of the oxidant may cause greater than average deviations from the scaling line, depending on the energetics associated with the structural rearrangement. Nonetheless, despite these drastic structural changes in the MOF-5 family, the departures from the scaling line are not nearly as significant as those for Cr-BTC and Mo-BTC, whch are proposed to break from the scaling relationship with deviations in ΔE_{H} greater than +0.6 eV.

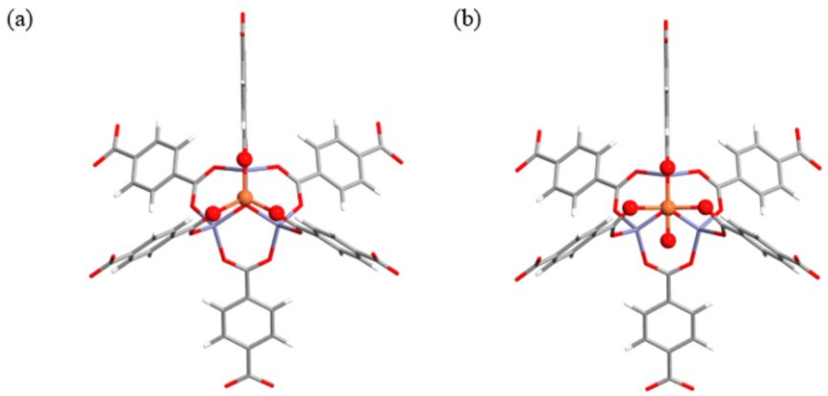


Figure B.6. DFT-optimized structure of Fe-MOF-5 before a) and after b) oxidation of the open metal site. Other M-MOF-5 variants exhibit analogous changes to the first coordination sphere. Color key: Fe (orange), Zn (purple), O (red), C (gray), H (white). Representative portions of the unit cells are shown.

B.3.4 Noncovalent Interactions

As discussed in the main text, it can be expected that changes in the bonding of framework atoms during the oxidation process could cause MOFs to significantly deviate from the scaling

relationship shown in Figure 3.3. In addition to the metal-metal bond-breaking behavior associated with Cr-BTC and Mo-BTC upon oxidation of the metal center (Figure 3.7), one could envision that noncovalent interactions may stabilize the metal-oxo intermediate (i.e. making ΔE_{O} more exothermic) without greatly altering the reactivity toward C−H activation (i.e. with negligible changes in ΔE_{H}). We tested this scenario with a MOF known as CFA-1,[565] which has metal cations of neighboring inorganic nodes separated by ~7.7 Å such that it can support inter-node hydrogen-bonding interactions,[566] depending on the ligands bound to the metal species.

For this example, we started with the DFT-optimized structure of CFA-1 with a combination of OH^- ligands and HCO_3^- ligands (such as that resulting from the reversible capture of CO_2).[566] One $[ZnOH]^+$ species per unit cell was then exchanged with a Cu^+ species, analogous to what has been done with Cu-MFU-4*l*.[27] As shown in Figure B.7a, the close distance between the $[CuO]^+$ intermediate and HCO_3^- ligand on a neighboring inorganic node permits H-bonding interactions. By replacing the HCO_3^- ligand with a smaller OH^- group (Figure B.7b), it is possible to evaluate the influence of H-bonding on the $[CuO]^+$ active site. For both scenarios, ΔE_{H} is predicted to be nearly identical, suggesting that the H-bonding interaction does not influence the reactivity of the $[CuO]^+$ site toward C−H activation since the Cu−O bond order is essentially unchanged (Table B.2). In contrast, the ΔE_{O} value for the H-bonding scenario is 0.18 eV more exothermic than without H-bonding behavior.

The slope of the scaling line relating ΔE_{H} to ΔE_{O} is −0.43. As such, a difference in ΔE_{O} by 0.18 eV would cause a 0.08 eV difference in the ΔE_{H} value predicted from the scaling line. In this example, the H-bonding interaction stabilizes the $[CuO]^+$ site but not enough to cause a particularly pronounced deviation from the proposed scaling relationship, especially compared with breaking

(or making) covalent bonds like with Cr-BTC and Mo-BTC. Nonetheless, one could achieve larger deviations from the scaling relationship with particularly strong H-bonds[567] or a more extensive H-bonding network. Even for MOFs where significant deviations from the scaling relationship are not present, moderately strong noncovalent interactions would likely be beneficial from a catalyst design standpoint, as they can stabilize otherwise fleeting active site species.[534]

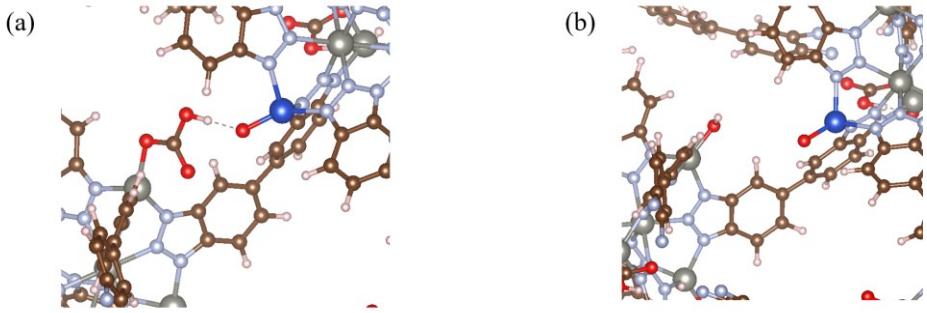


Figure B.7. a) DFT-optimized structure of CFA-1 with an HCO_3^- ligand exhibiting H-bonding with the $[CuO]^+$ active site. The distance between the O atom of the $[CuO]^+$ site and the H atom of the HCO_3^- ligand is 1.64 Å. b) DFT-optimized structure of CFA-1 with the HCO_3^- ligand replaced by an OH^- group. Color key: Cu (dark blue), Zn (gray), O (red), N (light blue), C (brown), H (white). Representative portions of the unit cells are shown.

Table B.7. Computed energies and DDEC6[231] Cu−O bond order for the MOFs shown in Figure B.7.

Scenario	ΔE_O (eV)	ΔE_H (eV)	Cu−O bond order
H-bonding, Fig. S7a	−0.15	−2.09	0.96
No H-bonding, Fig. S7b	+0.03	−2.11	1.02

B.3.5 Charge Delocalization

It has been suggested that the degree of charge delocalization upon formation of the metal-oxo active site may cause multiple distinct scaling lines,[105] with higher degrees of charge delocalization resulting in the catalyst being constrained to a less reactive scaling line (i.e. one with a more endothermic ΔE_H and higher $E_{TS,C-H}$ for a given ΔE_O). To test this hypothesis in the context of the

present study, we have computed the fractional degree of charge delocalization upon formation of the metal-oxo species for each MOF.

We define the fractional degree of charge delocalization following the chemisorption of a single O atom, χ_{ox}, as

$$\chi_{\mathrm{ox}} = 1 - \frac{\Delta q_{\mathrm{M}}}{|q_{\mathrm{O}}|} \tag{B.1}$$

where Δq_{M} is the change in partial atomic charge on the metal cation at the active site following activation of the oxidant and q_{O} is the partial atomic charge on the O adsorbate. With this definition, $\chi_{\mathrm{ox}} = 0$ implies that oxidation of the metal center occurs without any charge delocalization because $\Delta q_{\mathrm{M}} = |q_{\mathrm{O}}|$. In contrast, $\chi_{\mathrm{ox}} = 1$ implies that the increase in positive charge induced by the oxidant to maintain charge neutrality has been fully delocalized away from the metal center to surrounding atoms in the framework because $\Delta q_{\mathrm{M}} = 0$. Inherent to Equation B.1 are the assumptions that $\Delta q_{\mathrm{M}} \geq 0$ and $q_{\mathrm{O}} < 0$, consistent with what would be expected for an oxidation reaction. We found that the MOFs studied in this work exhibit the full range of χ_{ox} values, with some MOFs exhibiting purely metal-centered oxidation and others exhibiting significant charge delocalization to the surrounding ligands upon activation of the oxidant. Given the predictive accuracy of the single scaling line in Figure 3.3, this suggests that charge delocalization does not contribute to the presence of multiple distinct scaling lines, at least for the MOFs screened in this work.

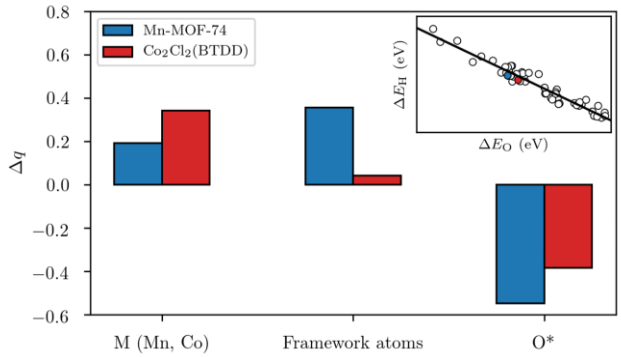


Figure B.8. Change in DDEC6 partial atomic charge upon oxidation, Δq, for the metal (M) at the metal-oxo active site, the O species of the metal-oxo active site (O*), and across all the remaining frameworks atoms for Mn-MOF-74 and Co_2Cl_2(btdd). The inset shows the H-affinity, ΔE_H, as a function of the active site formation energy, ΔE_O, for the screened MOFs from Figure 3.3, specifically highlighting Mn-MOF-74 and Co_2Cl_2(btdd). Color key: Mn-MOF-74 (blue), Co_2Cl_2(btdd) (red).

As an example, we specifically highlight Mn_2(dobdc), also known as Mn-MOF-74, and Co_2Cl_2(btdd).[25,28] Mn-MOF-74 is one of the few MOFs that has been experimentally shown to exhibit ligand redox non-innocence upon oxidation, with the Mn(II) cations retaining their formal oxidation state of 2+ in the presence of a one-electron iodobenzene dichloride oxidant.[28] In contrast, Dincă and coworkers have shown that the oxidation of Co_2Cl_2(btdd) in the presence of Cl_2 and Br_2 occurs at the Co(II) centers such that the ligands can be best characterized as redox innocent.[25] Qualitatively consistent with the aforementioned experiments, we predict Mn-MOF-74 and Co_2Cl_2(btdd) to have 65% and 11% charge delocalization upon formation of the terminal metal-oxo species, respectively, yet both MOFs clearly fall on the same scaling line (Figure B.8).

Given the vast number of MOFs synthesized to date, there are surprisingly few that have been shown to have redox non-innocent ligands.[28,249] Since high values of χ_{ox} indicate high degrees of charge delocalization that may be attributed to redox non-innocent behavior, it is worth investigating MOFs with high χ_{ox} values in greater detail. The MOFs with the top 10 highest

values for χ_{ox} are reported in Table B.3 along with their ΔE_O values. For reference, Mn-MOF-74 is predicted to have χ_{ox} = 0.65 (using both the DDEC6[158,233] and Bader[159] methods) and has been experimentally shown[28] to exhibit redox non-innocent behavior in the presence of iodobenzene dichloride.

As shown in Table B.3, some of the MOFs with high χ_{ox} values (e.g. Zn-BTC, Cd-BTT, Sc-MIL-88B, Zn_2(*m*-dobdc)) are those that are not expected to readily form a metal-oxo active site due to highly endothermic active site formation energies. Since the transition metals in these systems are not expected to be readily oxidized (e.g. Zn^{2+} is unlikely to form a formally Zn(IV)-oxo species), changes in charge are distributed over the surrounding framework atoms in order to maintain charge neutrality. Many of the other MOFs with high χ_{ox} values are those containing Mn or Fe sites. The MOF with the highest degree of charge delocalization and an experimentally plausible metal-oxo active site motif is Fe-MIL-88B.

Table B.8. Active site formation energies, ΔE_O, and degree of charge delocalization based on the DDEC6 method, $\chi_{ox,DDEC}$, and the Bader[159] method, $\chi_{ox,Bader}$. The MOFs with the ten highest $\chi_{ox,DDEC}$ values are shown.

MOF	ΔE_O (eV)	$\chi_{ox,DDEC}$	$\chi_{ox,Bader}$
Zn-BTC	1.69	1.37	0.89
Cd-BTT	1.93	1.29	0.93
Sc-MIL-88B	2.06	1.28	0.88
Fe-MIL-88B	−0.50	1.25	0.94
Zn_2(*m*-dobdc)	1.79	1.13	0.98
Mn-BTT	−1.25	1.09	0.77
Ti-MIL-88B	2.06	0.83	0.87
Mn_2Cl_2(btdd)	−1.08	0.78	0.75
Mn_2Cl_2(bbta)	−1.10	0.78	0.77
Fe_2(dobpdc)	−1.31	0.76	0.82

As previously mentioned, we computed the degree of charge delocalization, χ_{ox}, for each MOF in this work. Typically, the chemisorption of an O species to the metal site would be expected to

increase the partial atomic charge on the metal (with potentially some of the charge difference being distributed to the surrounding ligands). In other words, it is typically the case that $|\Delta q_{\mathrm{M}}| \leq |q_{\mathrm{O}}|$. However, this was not found to be the case for Ni-BTT, which based on the DFT calculations in this work has $\Delta q_{\mathrm{M,DDEC}} = 0.38$ and $q_{\mathrm{O,DDEC}} = -0.18$ ($\Delta q_{\mathrm{M,Bader}} = 0.91$, $q_{\mathrm{O,Bader}} = -0.34$), indicating that the net charge on the remaining framework atoms decreased compared to that in the bare, reduced MOF. This anomalous behavior in the charge may explain the relatively large deviation of +0.26 eV in ΔE_{H} compared to the scaling line.

Appendix C. APPENDIX FOR CHAPTER 4

C.1 Publicly Available Data

All computational results, including XYZ coordinates, energies, vibrational frequencies, partial charges, spin densities, bond orders, and more can be found at the Zenodo repository with DOI: 10.5281/zenodo.3554807.

C.2 Methods: Density Functional Theory Calculations

C.2.1 Methods: Periodic DFT

Periodic DFT calculations were carried out using VASP v.5.4.1[122,123] with v.5.4 of the VASP-recommended[337] projector-augmented wave (PAW) pseudopotentials.[124] As benchmarked in prior work,[32] the lattice constants and atomic positions of the guest-free MOFs were relaxed with a 520 eV plane-wave energy cutoff and Γ-centered $3 \times 2 \times 2$ k-point grid using the Niggli-reduced primitive unit cell. All forces were converged to within 0.03 eV/Å. Each self-consistent field (SCF) loop was considered converged if the change in energy was less than 10^{-6} eV, and Gaussian smearing of the band occupancies was applied with a smearing width of 0.01 eV prior to extrapolation back to the 0 K limit. The accurate precision keyword was enabled in VASP, and symmetry constraints were disabled. Calculations involving the M06-L functional in VASP included non-spherical contributions from the gradient corrections inside the PAW spheres. The Atomic Simulation Environment (ASE)[83] v.3.18.0 was used to carry out the VASP calculations.

C.2.2 Methods: Creating the Finite Cluster Models

The M-BBTA-X finite cluster models were carved from periodic, all-Mg analogues of M_2X_2(bbta) (X = F, Cl, Br, OH, SH, SeH) optimized in VASP. To optimize the Mg-containing model systems, the PBE-D3(BJ) level of theory[125–127] was used due to the functional's ability to accurately

reproduce the lattice constants of MOFs with closed-shell metals.[128,129] Both the atomic positions and unit cell shape/volume were relaxed. The finite cluster models carved from the VASP-optimized, periodic structures contain 75−79 atoms and are based on widely adopted cluster models for the structurally related M_2(dobdc) family[568] and a previously adopted cluster model for MFU-4*l* (Metal−Organic Framework Ulm).[364] To balance the charge on the cluster models, three protons were added to undercoordinated nitrogen atoms. The central Mg^{2+} site was then exchanged with each of the transition metals considered in this work (M = V, Cr, Mn, Fe, Co, Ni, or Cu) for the finite cluster DFT calculations. The resulting cluster models therefore contain two Mg^{2+} cations and one central transition metal cation, as depicted in Figure 4.1. In prior work, we have shown that the H atoms of the μ-OH^- groups can exist in one of several possible orientations, all of which have roughly similar energies in the guest-free structure.[73] For consistency, in all OH-, SH-, and SeH-containing cluster models, we oriented the H atoms in opposite directions around the central metal binding site. The only exception was for the Mn-BBTA-OH cluster with the $[MnO]^{2+}$ site (and at the transition state for C−H activation), for which the only stationary point that could be found was one in which the H atoms were both aligned parallel to the Mn−O bond.

C.2.3 Treatment of Spin States

Low-spin, intermediate-spin, and high-spin multiplicities were considered (where applicable) for the M^{2+} complexes, the $[MO]^{2+}$ complexes, and the transition state for C−H activation. Restricted Kohn-Sham DFT was adopted for the singlet states, and unrestricted Kohn-Sham DFT was adopted for all higher spin multiplicities. Significant spin-contamination for the open-shell singlet states prevented their investigation throughout this work; for instance, open-shell singlet

calculations for the Ni-BBTA-X frameworks would often converge to a state with $\langle S^2 \rangle$ of ~2 (instead of 0.75) following annihilation of the spin contamination.

As a result of the spin state analysis and as a matter of consistency, all physicochemical properties (e.g. energies, spin densities, bond distances, partial charges) are reported for the spin states listed in Table C.1 unless otherwise stated. For the vast majority of MOFs, the spin states listed in Table C.1 are the lowest energy spin state. The computed properties for other spin states can be found in the Zenodo repository.

Table C.1. Spin multiplicities adopted for the trends reported in this work, unless otherwise specified. AFM and FM refer to antiferromagnetic (AFM) and ferromagnetic (FM) coupling between the metal and oxo/oxyl species (for the $[MO]^{2+}$ state) or the three-center C–H–O radical (at the TS).

MOF	M^{2+}	$[MO]^{2+}$	TS
V-BBTA-X	4	2-AFM	2-AFM
Cr-BBTA-X	5	3-AFM	3-AFM
Mn-BBTA-X	6	4-FM	4-AFM
Fe-BBTA-X	5	5-FM	5-AFM
Co-BBTA-X	4	4-FM	4-FM
Ni-BBTA-X	3	5-FM	3-FM
Cu-BBTA-X	2	4-FM	4-FM

The relative spin state energies for the M^{2+} and $[MO]^{2+}$ complexes are shown in Figures C.1−C.6, and the relative spin state energies for various transition state structures are shown in Tables C.2−C.15. Due to the computational expense in calculating transition states, only select calculations are reported below (primarily those that successfully converged using our automated workflow). For the full list of atomic spin densities for a given spin multiplicity, please refer to the supporting dataset on Zenodo. Note that additional spin states have been tested and included below since the original publication of this work.

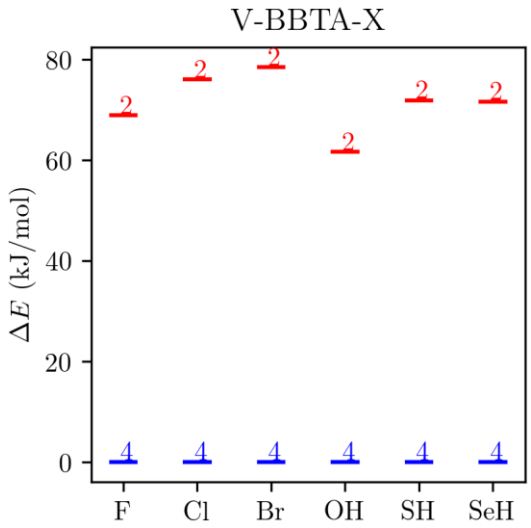


Figure C.1. Relative electronic energy of the considered spin multiplicities of V-BBTA-X.

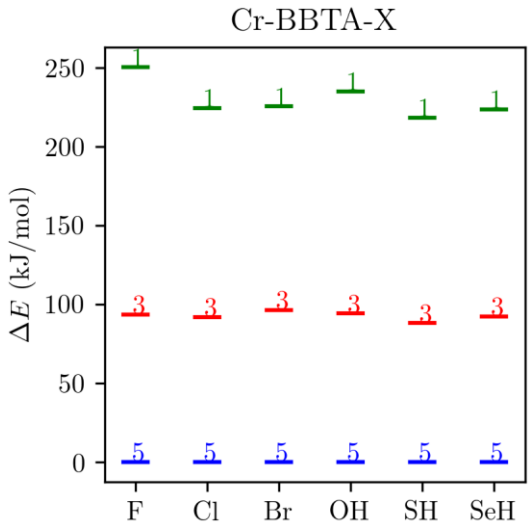


Figure C.2. Relative electronic energy of the considered spin multiplicities of Cr-BBTA-X.

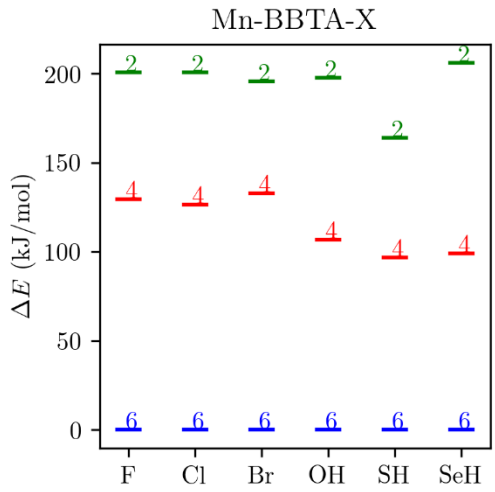


Figure C.3. Relative electronic energy of the considered spin multiplicities of Mn-BBTA-X.

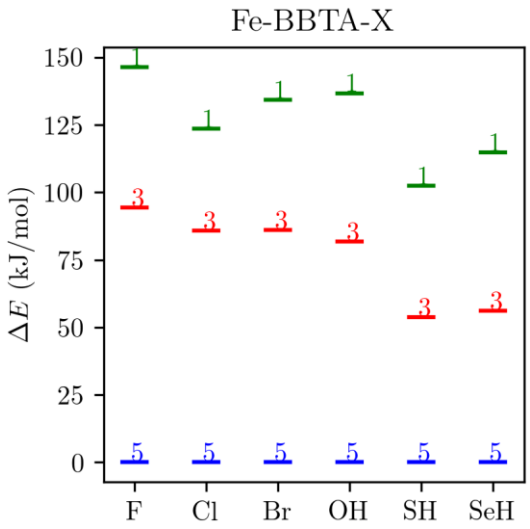


Figure C.4. Relative electronic energy of the considered spin multiplicities of Fe-BBTA-X.

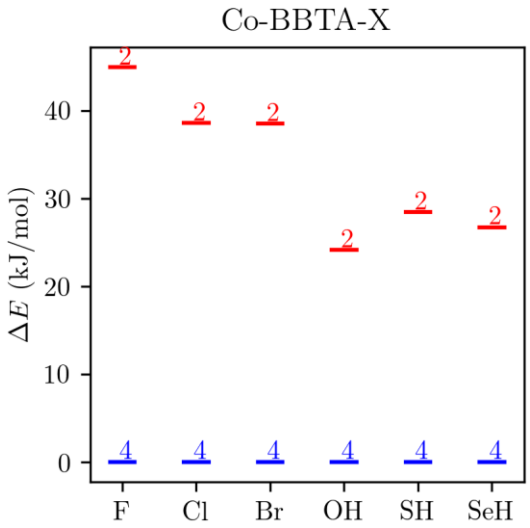


Figure C.5. Relative electronic energy of the considered spin multiplicities of Co-BBTA-X.

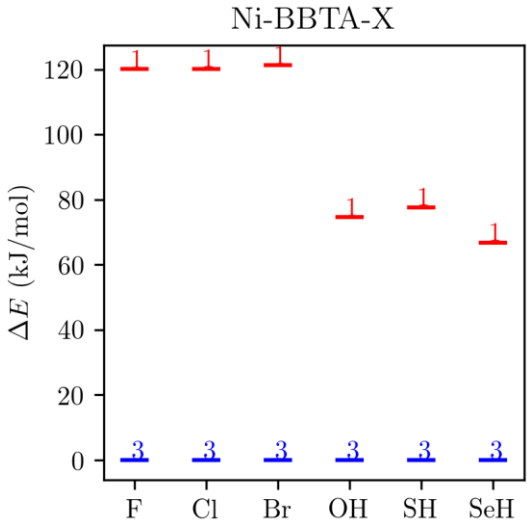


Figure C.6. Relative electronic energy of the considered spin multiplicities of Ni-BBTA-X.

Table C.2. Relative electronic energy of the considered spin states for the $[VO]^{2+}$ site of V-BBTA-X. AFM indicates antiferromagnetic coupling between the metal and terminal oxo/oxyl ligand. Note that the doublet state is the only physically plausible spin multiplicity, but a quartet state was investigated anyway.

	ΔE (kJ/mol)
$2S+1$	V-BBTA-Cl
2 (AFM)	0
4 (AFM)	281

Table C.3. Relative electronic energy of the considered spin states for the $[CrO]^{2+}$ site of Cr-BBTA-X. AFM indicates antiferromagnetic coupling between the metal and terminal oxo/oxyl ligand.

	ΔE (kJ/mol)					
$2S+1$	Cr-BBTA-Cl	Cr-BBTA-F	Cr-BBTA-Br	Cr-BBTA-OH	Cr-BBTA-SH	Cr-BBTA-SeH
1*	19	38	22	45	23	25
3 (AFM)	0	0	0	0	0	0

*Closed-shell singlet state.

Table C.4. Relative electronic energy of the considered spin states for the $[MnO]^{2+}$ site of Mn-BBTA-X. Entries marked by dashes were not computed. AFM indicates antiferromagnetic coupling between the metal and terminal oxo/oxyl ligand, whereas FM indicates ferromagnetic coupling.

	ΔE (kJ/mol)					
$2S+1$	Mn-BBTA-Cl	Mn-BBTA-F	Mn-BBTA-Br	Mn-BBTA-OH	Mn-BBTA-SH	Mn-BBTA-SeH
2 (AFM)	43	---	48	---	---	---
2 (FM)	---	36	---	50	35	37
4 (FM)	0	11	0	---	0	---
4 (AFM)	0	0	1	0	0	0
6 (FM)	47	53	---	---	---	---

Table C.5. Relative electronic energy of the considered spin states for the $[FeO]^{2+}$ site of Fe-BBTA-X. AFM indicates antiferromagnetic coupling between the metal and terminal oxo/oxyl ligand, whereas FM indicates ferromagnetic coupling.

	ΔE (kJ/mol)					
$2S+1$	Fe-BBTA-Cl	Fe-BBTA-F	Fe-BBTA-Br	Fe-BBTA-OH	Fe-BBTA-SH	Fe-BBTA-SeH
1*	154	154	172	117	105	100
3 (FM)	56	32	34	18	3	0
5 (FM)	0	0	0	0	0	3

*Closed-shell singlet state.

Table C.6. Relative electronic energy of the considered spin states for the $[CoO]^{2+}$ site of Co-BBTA-X. AFM indicates antiferromagnetic coupling between the metal and terminal oxo/oxyl ligand, whereas FM indicates ferromagnetic coupling.

	ΔE (kJ/mol)					
$2S+1$	Co-BBTA-Cl	Co-BBTA-F	Co-BBTA-Br	Co-BBTA-OH	Co-BBTA-SH	Co-BBTA-SeH
2 (FM)	39	43	37	15	11	14
4 (FM)	0	0	0	0	0	0
6 (FM)	44	28	43	21	33	0*

*The axial Co–N bond distance is significantly larger for Co-BBTA-SeH than for the other structures.

Table C.7. Relative electronic energy of the considered spin states for the $[NiO]^{2+}$ site of Ni-BBTA-X. FM indicates ferromagnetic coupling between the metal and oxo/oxyl ligand.

	ΔE (kJ/mol)					
$2S+1$	Ni-BBTA-Cl	Ni-BBTA-F	Ni-BBTA-Br	Ni-BBTA-OH	Ni-BBTA-SH	Ni-BBTA-SeH
1*	126	**	118	82	93	103
3 (FM)	7	10	6	0	1	9
5 (FM)	0	0	0	6	0	0

*Closed-shell singlet state.

**The geometry for the singlet state of Ni-BBTA-F with an $[NiO]^{2+}$ site could not be successfully converged without one of the triazolate linkers dissociating from the Ni center. Based on the remaining Ni MOFs, this spin state is unlikely to be mechanistically relevant.

Table C.8. Relative electronic energy of the considered spin states for the $[CuO]^{2+}$ site of CuO-BBTA-X. Entries marked by dashes were not computed. AFM indicates antiferromagnetic coupling between the metal and terminal oxo/oxyl ligand, whereas FM indicates ferromagnetic coupling.

	ΔE (kJ/mol)					
$2S+1$	Co-BBTA-Cl	Co-BBTA-F	Co-BBTA-Br	Co-BBTA-OH	Co-BBTA-SH	Co-BBTA-SeH
2 (FM)	7	14	4	17	11	---
2 (AFM)	---	---	---	---	--	21
4 (FM)	0	0	0	0	0	0

Table C.9. Relative electronic energy of the considered spin states for V-BBTA-Cl at the transition state for C−H activation. AFM indicates antiferromagnetic coupling between the metal and the three-center C–H–O radical, whereas FM indicates ferromagnetic coupling.

	ΔE (kJ/mol)
$2S+1$	V-BBTA-Cl
2 (AFM)	0
4 (FM)	56

Table C.10. Relative electronic energy of the considered spin states for Cr-BBTA-X at the transition state for C−H activation. AFM indicates antiferromagnetic coupling between the metal and the three-center C–H–O radical, whereas FM indicates ferromagnetic coupling.

	ΔE (kJ/mol)			
$2S+1$	Cr-BBTA-Cl	Cr-BBTA-Br	Cr-BBTA-SH	Cr-BBTA-SeH
3 (AFM)	0	0	0	0
5 (FM)	27	28	29	29

Table C.11. Relative electronic energy of the considered spin states for Mn-BBTA-X at the transition state for C−H activation. Entries marked by dashes were not computed. AFM indicates antiferromagnetic coupling between the metal and the three-center C–H–O radical, whereas FM indicates ferromagnetic coupling.

	ΔE (kJ/mol)				
$2S+1$	Mn-BBTA-Cl	Mn-BBTA-F	Mn-BBTA-Br	Mn-BBTA-SH	Mn-BBTA-SeH
2 (FM)	---	126	---	---	---
2 (AFM)	71	88	---	38	38
4 (FM)	---	91	88	48	47
4 (AFM)	0	0	0	0	0
6 (FM)	---	---	---	24	---

Table C.12. Relative electronic energy of the considered spin states for Fe-BBTA-X at the transition state for C−H activation. Entries marked by dashes were not computed. AFM indicates antiferromagnetic coupling between the metal and the three-center C–H–O radical, whereas FM indicates ferromagnetic coupling.

	ΔE (kJ/mol)			
$2S+1$	Fe-BBTA-Cl	Fe-BBTA-F	Fe-BBTA-Br	Fe-BBTA-OH
3 (FM)	---	84	70	59
3 (AFM)	45	---	---	---
5 (FM)	36	48	38	53
5 (AFM)	0	0	0	0
7 (FM)	---	14	---	---

Table C.13. Relative electronic energy of the considered spin states for Co-BBTA-X at the transition state for C−H activation. Entries marked by dashes were not computed. AFM indicates antiferromagnetic coupling between the metal and the three-center C–H–O radical, whereas FM indicates ferromagnetic coupling.

	ΔE (kJ/mol)					
$2S+1$	Co-BBTA-Cl	Co-BBTA-F	Co-BBTA-Br	Co-BBTA-OH	Co-BBTA-SH	Co-BBTA-SeH
2 (FM)	16	19	22	0	0	0
2 (AFM)	13	9	---	---	---	---
4 (FM)	0	0	0	21	18	6
6 (FM)	15	2	14	22	---	---

Table C.14. Relative electronic energy of the considered spin states for Ni-BBTA-X at the transition state for C−H activation. Entries marked by dashes were not computed. AFM indicates antiferromagnetic coupling between the metal and the three-center C–H–O radical, whereas FM indicates ferromagnetic coupling.

	ΔE (kJ/mol)				
$2S+1$	Ni-BBTA-Cl	Ni-BBTA-F	Ni-BBTA-Br	Ni-BBTA-OH	Ni-BBTA-SH
1 (FM)	118	---	---	---	--
3 (FM)	0	0	0	0	0
3 (Alt)*	---	---	---	29	---
5 (FM)	5	3	3	16	6

*In this case, although an AFM state was provided as the initial guess, the terminal oxo ligand has negligible spin density; the methyl radical has spin density with an opposite sign compared to the metal.

Table C.15. Relative electronic energy of the considered spin states for Cu-BBTA-X at the transition state for C−H activation. Entries marked by dashes were not computed. AFM indicates antiferromagnetic coupling between the metal and the three-center C–H–O radical, whereas FM indicates ferromagnetic coupling.

	ΔE (kJ/mol)				
$2S+1$	Cu-BBTA-Cl	Cu-BBTA-F	Cu-BBTA-OH	Cu-BBTA-SH	Cu-BBTA-SeH
2 (FM)	2	---	21	---	---
2 (AFM)	---	15	---	9	8
4 (FM)	0	0	0	0	0

C.2.4 Model Assumptions

The need to carry out a large number of computationally expensive DFT calculations requires a few assumptions with regards to the reactivity of the M_2X_2(bbta) family of materials. These are discussed below.

To confirm that the cluster size and modeling choices regarding the M-BBTA-X cluster models are representative of the larger periodic structure, we compared the metal-oxo formation energies for the M-BBTA-X (X = F, Cl, Br) cluster models at the M06-L/def2-TZVP level of theory with the metal-oxo formation energies computed in VASP using the fully periodic structures at the M06-L/PAW level of theory. For these VASP calculations, the unit cell contains one transition metal cation per Niggli-reduced primitive cell, mimicking the cluster models. For the guest-free periodic structures, the atomic positions and cell shape/volume were allowed to

relax. For the periodic DFT calculations involving the $[MO]^{2+}$ site, the unit cell was kept fixed at the DFT-optimized guest-free structure, and all atomic positions were allowed to relax. Unlike the cluster models, no constraints were applied for the VASP calculations. As shown in Figure C.8, there is strong agreement between the M06-L/def2-TZVP (i.e. cluster) calculations and the M06-L/PAW (i.e. periodic) calculations, supporting the choice of model system. For reference, a best-fit line of $y = 1.04x - 7.18$ kJ/mol with $r^2 = 0.996$ is obtained. Even if the cluster model is highly representative of the larger periodic structure, slight deviations between the periodic and cluster calculations are to be expected, as the basis sets are not identical between VASP and Gaussian. We acknowledge that effects such as strain or pore-based confinement could influence the quantitative values of the reactivities reported in this work.[106,569] Nonetheless, the fidelity of the cluster model demonstrated in Figure C.8, the success of the analogous M_2(dobdc) cluster models,[177,568] and the large pores of the M_2X_2(bbta) framework provide support that the M-BBTA-X clusters are capable of accurately modeling the periodic structures in this work.

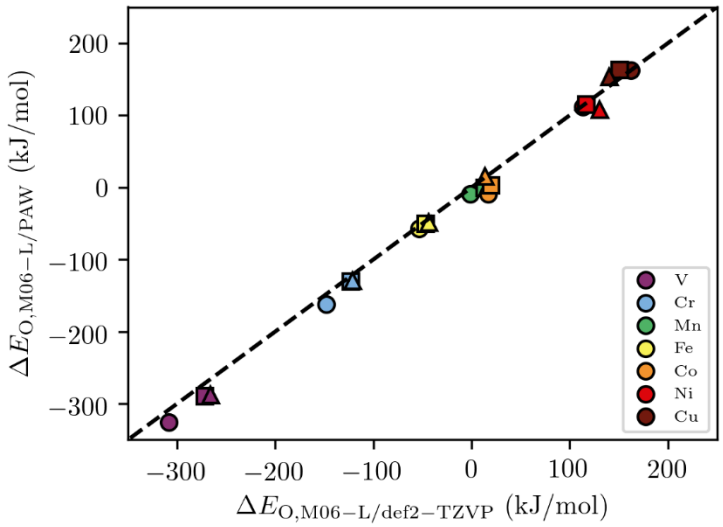


Figure C.7. Metal-oxo formation energy, ΔE_O, at the M06-L/PAW level of theory (periodic calculations) compared against the M06-L/def2-TZVP level of theory (cluster calculations). The

dashed line indicates perfect agreement between the two methods. Symbols: X = F (circle), X = Cl (square), X = Br (triangle).

Prior studies have also shown that the 20% Hartree-Fock exchange of B3LYP can over-stabilize high-spin states in several cases,[228,229] and it is well-established that it can be difficult for DFT to calculate accurate spin splitting energies.[570] As a point of comparison, we used the M06-L/def2-TZVP level of theory to identify the ground state spin state in the M-BBTA-X (M = {V, Cr, Mn, Fe, Co, Cu}; X = {F, Cl, Br}) cluster models in both the adsorbate-free (i.e. M^{2+}) and metal-oxo (i.e. $[MO]^{2+}$) states. We found that for every structure, the B3LYP-D3(BJ)/def2-TZVP and M06-L/def2-TZVP levels of theory predict the same ground state spin multiplicity. Similarly, all the B3LYP-D3(BJ)/def2-TZVP ground state spin multiplicities were the same as the M06-L/PAW calculations of the fully periodic structures, with the exception of Mg-diluted Co_2Cl_2(bbta), which has a low-spin ground state at the M06-L/PAW level of theory but a high-spin ground state at the B3LYP-D3(BJ)/def2-TZVP and M06-L/def2-TZVP levels of theory. It is known from prior experiments that this framework has a high-spin ground state.[25] Finally, several metal−azolate frameworks with open-shell $3d$ transition metal cations have been shown to exhibit spin-crossover as a function of temperature, pressure, and other external stimuli.[193,270] While we account for the possibility of spin transitions throughout the metal-oxo formation and C−H activation processes, spin-crossover behavior due to external stimuli is not considered.

C.2.5 Metal-Oxo Formation Energies at the M06-L/def2-TZVP Level of Theory

In addition to the B3LYP-D3(BJ)/def2-TZVP results presented in the main text, we computed the metal-oxo formation energy for M-BBTA-X (X = F, Cl, Br) using the M06-L/def2-TZVP level of theory. As expected based on prior work,[364] the M06-L functional results in more exothermic oxidation energies than the B3LYP functional due to a lack of Hartree-Fock exchange, but they

are generally close in value (Figure C.9). For those who may be interested in modeling these systems with other exchange-correlation functionals, we note that significant difficulty was found when attempting to converge the SCF with the M06-L functional, particularly for X = {OH, SH, SeH} and for transition state calculations, which is why the only M06-L cluster calculations carried out are those in Figure C.9. Test calculations indicate that the SCF convergence problem also exists with the M06 functional[288] but is resolved with the newer (but less widely tested) MN-15L[571] and MN-15[572] functionals, suggesting that the inclusion of the smoothness constraint in the latter two functionals greatly improves the ease of SCF convergence for these highly parametrized meta(-hybrid) generalized gradient approximation (GGA) functionals.

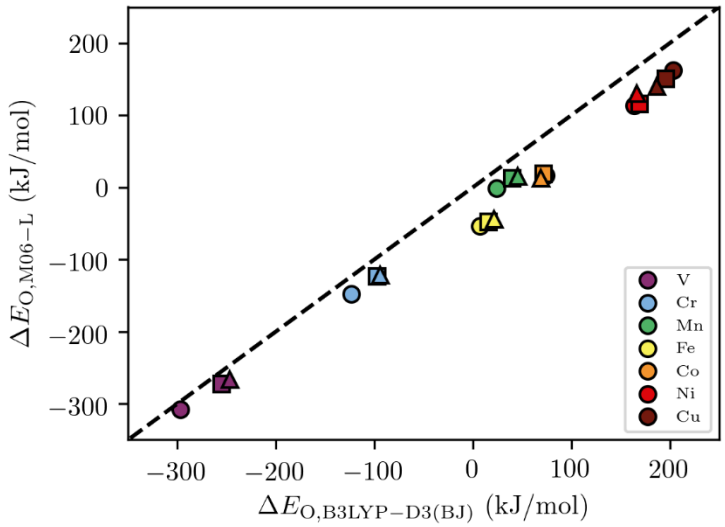


Figure C.8. Metal-oxo formation energy, ΔE_O, at the M06-L/def2-TZVP and B3LYP-D3(BJ)/def2-TZVP levels of theory. The dashed line indicates perfect agreement between the two methods. Symbols: X = F (circle), X = Cl (square), X = Br (triangle).

C.3 Supplementary Results

C.3.1 Partial Charges on the Metals

As shown in Figure C.10a and S10b, the partial charge on the metal decreases as F^- > Cl^- > Br^- and OH^- > SH^- > SeH^-, in agreement with what would be expected based on the relative

electronegativity of these species. Note that ordering the ligands based on their donating strength yields the opposite order. The results in Figure C.10c and C.10d serve to highlight the decreasing degree to which the metal site is oxidized going from V to Cu across the first-row transition metals.

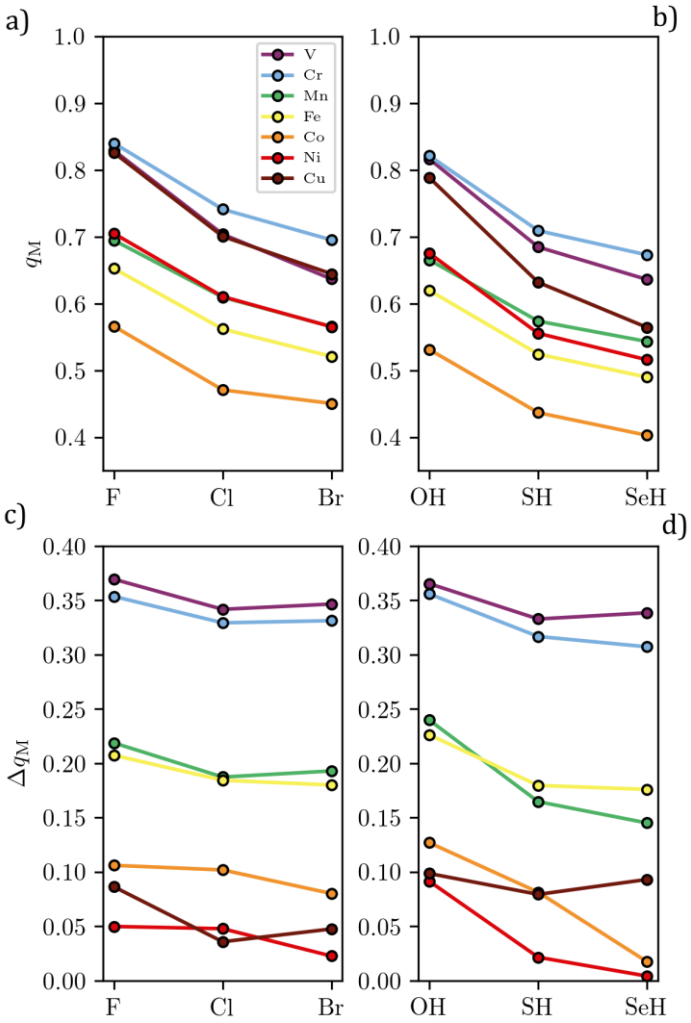


Figure C.9. CM5 partial atomic charge on the M(II) binding site, q_M, as a function of (a) X = {F, Cl, Br} and (b) X = {OH, SH, SeH}. Change in CM5 partial atomic charge on the metal binding, Δq_M, site upon formation of the $[MO]^{2+}$ site as a function of (c) X = {F, Cl, Br} and (d) X = {OH, SH, SeH}. The sign convention is such that a positive value of Δq_M indicates an increase in charge on the metal. Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

C.3.2 Comparing Metal-Oxo Formation Enthalpies

Initial configurations for the oxidized $Fe_{0.1}Mg_{1.9}$(dobdc) and $Fe_3(\mu_3\text{-}O)(HCOO)_6$ models were adopted from prior work.[177,264] For $Fe_{0.1}Mg_{1.9}$(dobdc), a constrained optimization was performed, as described in prior work.[177] The high-spin Fe^{2+} and $[FeO]^{2+}$ states were adopted for these materials, as justified in the original studies on these materials.[177,264] The unpaired electrons for the iron cations in $Fe_3(\mu_3\text{-}O)(HCOO)_6$ were antiferromagnetically aligned, as described previously.[179] The structures for the $[FeO]^{2+}$ sites for these materials are shown in Figure C.11.

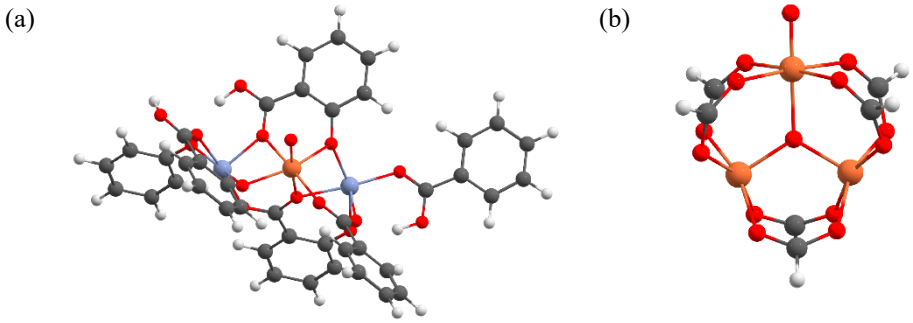


Figure C.10. (a) $Fe_{0.1}Mg_{1.9}$(dobdc) and (b) $Fe_3(\mu_3\text{-}O)(HCOO)_6$ with an $[FeO]^{2+}$ site. Color key: Fe (orange), Mg (purple), O (red), C (black), H (white).

Table C.16. Metal-oxo formation enthalpies for the Fe-BBTA-X series compared to $Fe_{0.1}Mg_{1.9}$(dobdc) and the $Fe_3(\mu_3\text{-}O)(HCOO)_6$ motif representative of Fe-MIL-100/Fe-MIL-101/Fe-PCN-250. Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

MOF	ΔH_O (kJ/mol)
Fe-BBTA-OH	−10
Fe-BBTA-SH	−10
Fe-BBTA-SeH	−9
Fe-BBTA-F	10
Fe-BBTA-Cl	18
Fe-BBTA-Br	23
$Fe_{0.1}Mg_{1.9}$(dobdc)	24
$Fe_3(\mu_3\text{-}O)(HCOO)_6$	35

C.3.3 Comparing the M−O Interaction with Isolable References

To put the geometric and electronic structure properties of the proposed $[MO]^{2+}$ sites in context, we compared the M−O bond distance, spin density on the O atom, and M−O bond order for the M-BBTA-X (M = Mn, Fe, Co) series with those of $[MnO(Bn\text{-}TPEN)]^{2+}$ (Bn-TPEN = *N*-benzyl-*N,N′,N′*-tris(2-pyridylmethyl)-1,2-diaminoethane), $[FeO(TMG_3tren)]^{2+}$ (TMG_3tren = 1,1,1-tris{2-[N^2-(1,1,3,3-tetramethylguanidino)]ethyl}amine), and $[CoO(13\text{-}TMC)]^{2+}$ (13-TMC = 1,4,7,10-tetramethyl-1,4,7,10-tetraazcyclotridecane), which we selected as three representative examples of complexes with experimentally characterized, non-heme, terminal M(IV)-oxo species.[573–576] In addition, these three $[MO]^{2+}$ species have the same ground state spin states as the investigated M-BBTA-X analogues (namely, high-spin $[MnO]^{2+}$, high-spin $[FeO]^{2+}$, and intermediate spin $[CoO]^{2+}$), which enables a more direct comparison. If the properties of the proposed metal-oxo sites in the M_2X_2(bbta) family are comparable to those of these isolable metal-oxo complexes, then this would suggest that the lifetimes of the metal-oxo sites in the M_2X_2(bbta) family would likely be experimentally realizable.

The $[MnO(Bn\text{-}TPEN)]^{2+}$, $[FeO(TMG_3tren)]^{2+}$, and $[CoO(13\text{-}TMC)]^{2+}$ reference structures are shown in Figure C.12. The initial structures were adopted from prior work.[573–576] All three complexes were modeled with a net charge of +2. It has previously been shown that significant self-interaction error can be present when modeling charged metal-oxo complexes.[577] As done elsewhere,[578] we modeled the three transition metal complexes in the presence of a continuum solvent model[579,580] of acetonitrile to reduce the self-interaction error. The Mn^{2+}, $[MnO]^{2+}$, Fe^{2+}, $[FeO]^{2+}$, Co^{2+} sites were modeled as being in the high-spin state.[573–576] The $[CoO]^{2+}$ site was modeled in the intermediate spin state.[576]

As shown in Figure C.13, the Mn-BBTA-X, Co-BBTA-X, and Fe-BBTA-X frameworks are predicted to have stronger M−O bonding interactions than the experimentally isolable reference complexes for nearly every bridging ligand considered in this work, as evidenced by the shorter M−O bond distances, smaller amount of spin density, and greater M−O bond orders. Unexpectedly, the results in Figure C.13 indicate that the thermodynamic trends do not necessarily correlate with the electronic structure of the metal-oxo site. For instance, despite resulting in more endothermic ΔH_O values, the halide bridging ligands have larger M−O bond orders than the X = {OH, SH, SeH} ligands for many of the Mn-, Fe-, and Co-containing frameworks. The thermodynamic favorability of forming the $[MO]^{2+}$ sites for the Mn-BBTA-X and Co-BBTA-X frameworks are also comparable (and in some cases more favorable, depending on the ligand) to the experimentally isolable transition metal-oxo complexes (Table C.10). However, we note that the Fe-BBTA-X frameworks studied in this work have less thermodynamically favored Fe(IV)-oxo sites than $[FeO(TMG_3tren)]^{2+}$ (Table C.10).

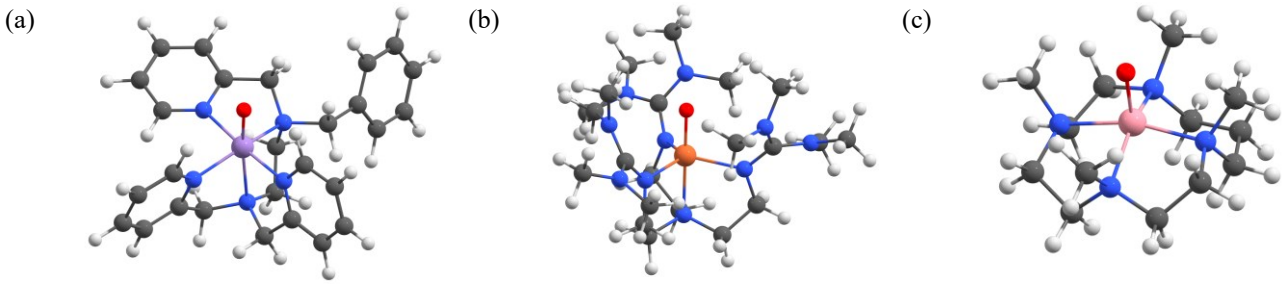


Figure C.11. (a) $[MnO(Bn\text{-}TPEN)]^{2+}$, (b) $[FeO(TMG_3tren)]^{2+}$, and (c) $[CoO(13\text{-}TMC)]^{2+}$. Color key: Mn (purple), Fe (orange), Co (pink), O (red), N (blue), C (black), H (white).

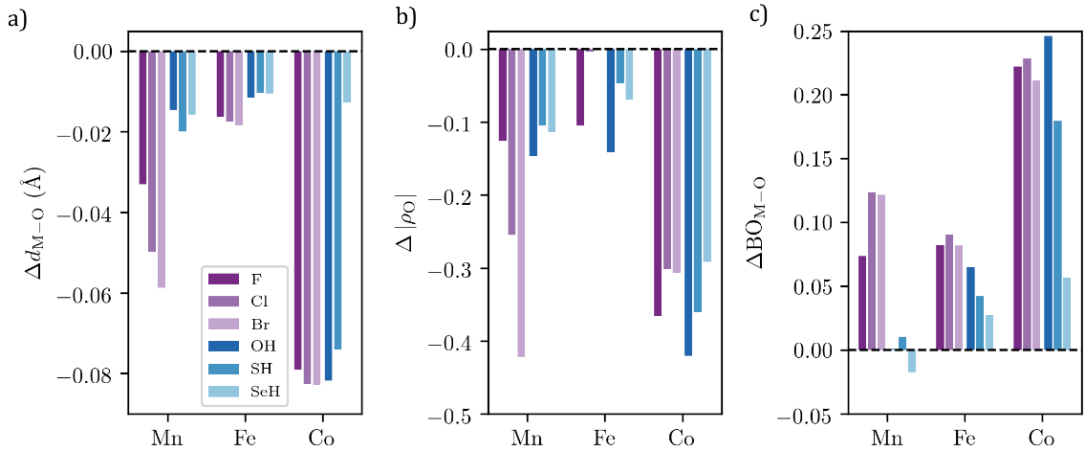


Figure C.12. Computed properties of proposed metal-oxo species with respect to $[MnO(Bn\text{-}TPEN)]^{2+}$, $[FeO(TMG_3tren)]^{2+}$, and $[CoO(13\text{-}TMC)]^{2+}$. (a) Relative M−O bond distance, Δd_{M-O}. (b) Relative absolute Hirshfeld spin density on the O moiety, $\Delta|\rho_O|$. (c) Relative Wiberg M−O bond order in the natural atomic orbital basis, ΔBO_{M-O}. For all subplots, the sign convention is such that a negative Δ value indicates a smaller value for the M-BBTA-X framework than for the reference transition metal complex. The Mn, Fe, and Co reference values of $[d_{M-O}; |\rho_O|; BO_{M-O}]$ are [1.658 Å; 0.582; 1.619], [1.629 Å; 0.598; 1.524], and [1.681 Å; 1.177; 1.308], respectively. Note that the $\Delta|\rho_O|$ values for Mn-BBTA-X (X = Cl, Br) and the ΔBO_{M-O} value for Mn-BBTA-OH are approximately zero. Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

Table C.17. Metal-oxo formation enthalpies for $[MnO(Bn\text{-}TPEN)]^{2+}$, $[FeO(TMG_3tren)]^{2+}$, and $[CoO(13\text{-}TMC)]^{2+}$ compared to the most exothermic and endothermic M-BBTA-X (M = Mn, Fe, Co) metal-oxo formation enthalpies. Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

Complex	ΔH_O (kJ/mol)	MOF	ΔH_O (kJ/mol)	MOF	ΔH_O (kJ/mol)
$[MnO(Bn\text{-}TPEN)]^{2+}$	19	Mn-BBTA-OH	−32	Mn-BBTA-Br	48
$[FeO(TMG_3tren)]^{2+}$	−62	Fe-BBTA-X (X = OH, SH)	−7	Fe-BBTA-Br	23
$[CoO(13\text{-}TMC)]^{2+}$	59	Co-BBTA-SeH	53	Co-BBTA-F	75

C.3.4 Radical-Like Character and Reactivity

The spin density on the abstracting O atom in the $[MO]^{2+}$ state is shown in Figure C.13. The apparent activation enthalpy for breaking the C−H bonds of methane as a function of the (absolute) spin density on the abstracting O atom is also shown in Figure C.14, highlighting the relatively weak correlation between these two parameters.

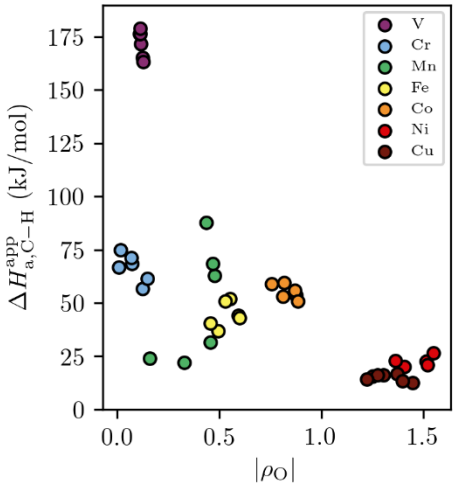


Figure C.13. Apparent enthalpic barrier for C−H activation as a function of the absolute Hirshfeld spin density on the abstracting O atom of the $[MO]^{2+}$ site. Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

C.3.5 Antiferromagnetically Enhanced Reactivity

The enthalpy difference between the antiferromagnetically coupled and ferromagnetically coupled spin states at the transition state for C−H activation are shown in Table C.11 and Table C.12. For both the Mn- and Fe-containing frameworks, the antiferromagnetically coupled metal site and three-center C−H−O radical is the most stable state and leads to a significantly reduced activation enthalpy.

For the Fe-containing frameworks, the initial $[FeO]^{2+}$ state could only be isolated with ferromagnetic coupling between the iron binding site and oxo ligand. For the Mn-containing frameworks, both the ferromagnetically aligned and antiferromagnetically aligned $[MnO]^{2+}$ initial states could be isolated and are generally close in energy (Table C.13). For consistency, the properties of the structure with ferromagnetically aligned spin density are used for the $[MnO]^{2+}$ sites in this work unless otherwise noted.

Table C.18. Enthalpy difference, $\Delta H^{\ddagger}_{4,\mathrm{FM-AFM}}$, between the ferromagnetically coupled and antiferromagnetically coupled quartet states for several of the $[MnO]^{2+}$ complexes investigated

throughout this work at the transition state for C−H activation. The sign convention is such that a positive value of $\Delta H_{4,\mathrm{FM-AFM}}$ indicates the antiferromagnetically coupled state is more stable. Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

MOF	State	$\Delta H^{\ddagger}_{4,\mathrm{FM-AFM}}$ (kJ/mol)
Mn-BBTA-Br	TS	87
Mn-BBTA-F	TS	92
Mn-BBTA-SH	TS	49

Table C.19. Enthalpy difference, $\Delta H^{\ddagger}_{5,\mathrm{FM-AFM}}$, between the ferromagnetically coupled and antiferromagnetically coupled quintet states for several of the $[FeO]^{2+}$ complexes investigated throughout this work at the transition state for C−H activation. The sign convention is such that a positive value of $\Delta H^{\ddagger}_{5,\mathrm{FM-AFM}}$ indicates the antiferromagnetically coupled state is more stable. Results are at the B3LYP-D3(BJ)/def2-TZVP level of theory.

MOF	State	$\Delta H^{\ddagger}_{5,\mathrm{FM-AFM}}$ (kJ/mol)
Fe-BBTA-Br	TS	39
Fe-BBTA-Cl	TS	37
Fe-BBTA-F	TS	51

Appendix D. Appendix for Chapter 5

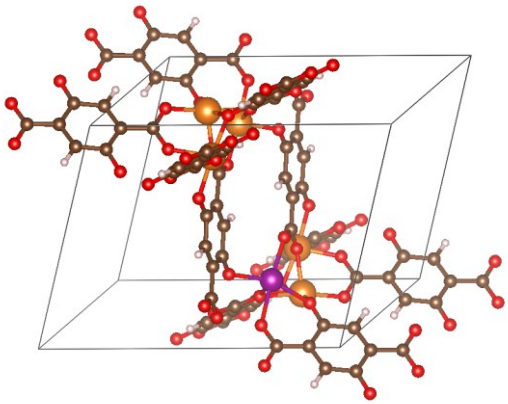
D.1 Supplemental Data

All supplementary data, including but not limited to DFT-optimized structures (for both periodic and cluster models), tabulated results, converged magnetic moments, partial charges, spin densities, bond orders, and energies are made publicly available at the following Zenodo DOI: 10.5281/zenodo.2652475.

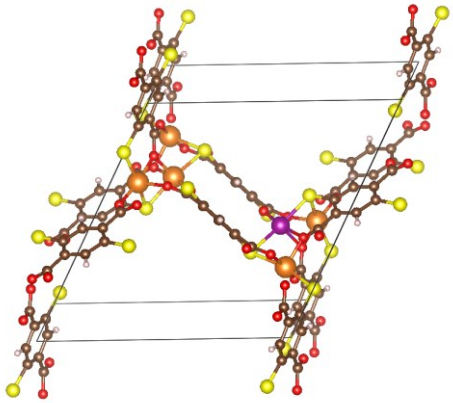
D.2 Additional Computational Details

D.2.1 Simulation Unit Cells

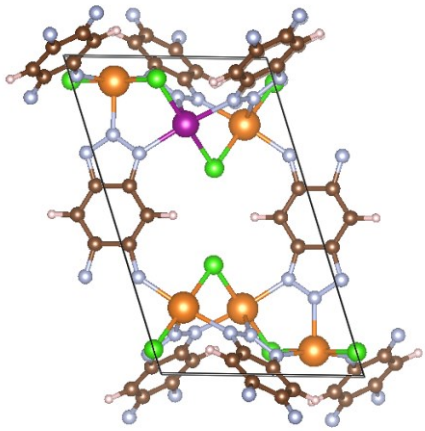
a)



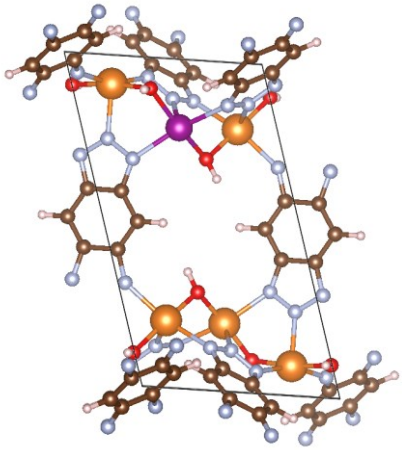
b)



c)



d)



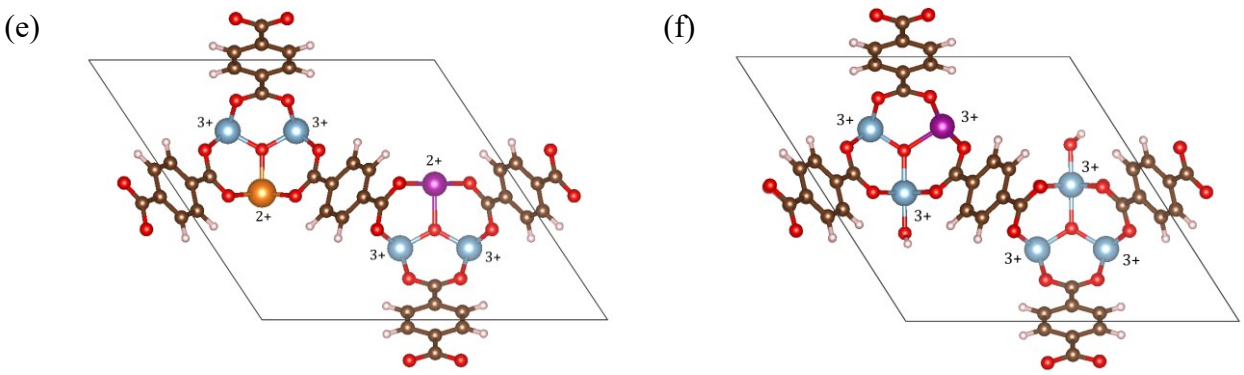


Figure D.1. Mg^{2+}- and Al^{3+}-diluted model systems for a) MOF-74, b) MOF-74-S, c) MAF-Cl, d) MAF-OH, (e) MIL-88B, and (f) MIL-88B-OH. The formal oxidation states of all the metals are 2+ in Figure D.1a–D.1d, a mix of 2+/3+ in Figure D.1e, and 3+ in Figure D.1f. O_2 and N_2 adsorption was investigated at the M (purple) sites. Color key: M (purple), Mg (orange), Al (light blue), O (red), N (blue), S (yellow), Cl (green), C (brown), H (white).

Examples of the diluted MOF structures considered during the periodic DFT screening process are shown in Figure D.1. The metal cations for the MOF-74, MOF-74-S, MAF-Cl, and MAF-OH families are divalent. The formal oxidation state of the metal sites in the trimetallic nodes of the MIL-88B family depends on the presence or lack of a terminal OH group. We consider a framework with the $[M(II)M(III)_2(\mu_3\text{-}O)(CH_3COO)_6]$ motif (Figure D.1e) that has been shown in PCN-250,[333] in addition to the $[M(III)_3(OH)(\mu_3\text{-}O)(CH_3COO)_6]$ motif (Figure D.1f) of MIL-100/MIL-101.[355]

D.2.2 Ground State Spin States and Predicted O_2 Binding Modes

Table D.1. Total number of unpaired electrons per simulation unit cell ($|n_\alpha - n_\beta|$) for each calculation carried out at the M06-L/PAW level of theory in this work. Refer to the supporting dataset for the magnetic moments of the metal binding site and adsorbate atoms. The right-most column shows the predicted binding mode of O_2 at the M06-L/PAW level of theory. A value of 0, 1, or 2 refers to no interaction, end-on binding, and side-on binding, respectively, as determined from Pymatgen's CrystalNN bonding topology algorithm.[82,335] The dashed lines correspond to structures not calculated in this work.

	$\lvert n_\alpha - n_\beta \rvert$			
MOF	Guest-free MOF	MOF–O_2	MOF–N_2	M–O_2 denticity
Co-MAF-Br	3	1	3	1
Co-MAF-Cl	3	1	3	1

Co-MAF-F	3	1	3	1
Co-MAF-OH	3	1	3	1
Co-MAF-SH	1	1	1	1
Co-MOF-74-S	3	1	3	1
Co-MIL-88B	3	1	3	1
Co-MIL-88B-OH	0	2	0	1
Co-MOF-74	3	1	3	1
Cr-MAF-Br	4	2	---	1
Cr-MAF-Cl	4	2	4	1
Cr-MAF-F	4	2	---	1
Cr-MAF-OH	4	2	4	1
Cr-MAF-SH	4	2	---	2
Cr-MOF-74-S	4	2	4	1
Cr-MIL-88B	4	2	2	1
Cr-MIL-88B-OH	3	1	3	2
Cr-MOF-74	4	2	4	1
Cu-MAF-Cl	1	3	1	1
Cu-MAF-OH	1	3	1	1
Cu-MOF-74-S	1	1	1	0
Cu-MIL-88B	1	3	1	1
Cu-MOF-74	1	3	1	1
Fe-MAF-Br	4	6	---	2
Fe-MAF-Cl	4	6	4	2
Fe-MAF-F	4	6	---	2
Fe-MAF-OH	4	6	4	2
Fe-MAF-SH	4	6	---	2
Fe-MOF-74-S	4	6	4	2
Fe-MIL-88B	4	6	4	2
Fe-MIL-88B-OH	5	3	5	1
Fe-MOF-74	4	6	4	2
Mn-MAF-Br	5	3	---	1
Mn-MAF-Cl	5	3	5	1
Mn-MAF-F	5	3	---	1
Mn-MAF-OH	5	3	5	1
Mn-MAF-SH	5	3	---	1
Mn-MOF-74-S	5	3	5	1
Mn-MOF-74-S (seesaw)	5	3	5	2
Mn-MIL-88B	5	3	5	1
Mn-MIL-88B-OH	4	2	4	1
Mn-MOF-74	5	3	5	1
Ni-MAF-Br	2	0	---	1
Ni-MAF-Cl	2	0	2	1
Ni-MAF-F	2	0	---	1

Ni-MAF-OH	2	0	2	1
Ni-MAF-SH	2	0	---	1
Ni-MOF-74-S	2	0	2	1
Ni-MIL-88B	2	0	2	1
Ni-MIL-88B-OH	1	1	1	1
Ni-MOF-74	2	0	2	1
Sc-MIL-88B-OH	0	2	0	1
Ti-MIL-88B-OH	1	1	1	2
V-MAF-Br	3	1	---	2
V-MAF-Cl	3	1	3	2
V-MAF-F	3	1	---	2
V-MAF-OH	3	1	3	2
V-MAF-SH	3	1	---	2
V-MOF-74-S	3	1	3	2
V-MIL-88B	3	1	3	2
V-MIL-88B-OH	2	0	2	2
V-MOF-74	3	1	3	2
Zn-MAF-Cl	0	2	0	0
Zn-MAF-OH	0	2	0	0
Zn-MOF-74-S	0	2	0	0
Zn-MIL-88B	0	2	0	1
Zn-MOF-74	0	2	0	1

D.2.3 Additional VASP Details

Aspherical contributions to the gradient corrections inside the PAW spheres were included for all meta-GGA and GGA+*U* calculations, as recommended in the VASP manual.[581] Vibrational modes were computed via central finite difference approximations of the Hessian matrix by displacing the relevant atoms 0.01 Å in the $\pm x$, $\pm y$, and $\pm z$ dimensions. To ensure precise evaluation of the forces, the SCF convergence was set to 10^{-8} eV during all vibrational analyses. Thermochemical corrections for O_2 and N_2 adsorption in Co_2Cl_2(bbta) and $Co_2(OH)_2$(bbta) were calculated at 298.15 K using the Atomic Simulation Environment.[83] Ideal gas statistical mechanics was used to describe the enthalpy of free O_2 and N_2. As a simplifying assumption, translational or rotational contributions to the enthalpy from the MOF and bound adsorbate complex were neglected. The harmonic approximation was invoked to evaluate the vibrational contribution to the enthalpy upon

O_2 or N_2 adsorption. For the vibrational contribution to the enthalpy of O_2 and N_2 adsorption in $Co_2(OH)_2$(bbta), the cobalt binding site, adsorbate, equatorial N atoms, axial N atom, and equatorial OH groups were displaced (i.e. it is assumed that the enthalpic contribution from any changes in the vibrational modes of the remaining atoms is small). Similarly, for the vibrational contribution to the enthalpy of O_2 and N_2 adsorption in Co_2Cl_2(bbta), the cobalt binding site, adsorbate, equatorial N atoms, axial N atom, and equatorial Cl groups were displaced. Vibrational frequencies smaller than 50 cm^{-1} were scaled up to this value. When computing the stretching frequencies of the O_2 and N_2 adsorbates in Figure D.2, only the atoms of the diatomic adsorbate were allowed to vibrate for computational simplicity. Bader charges and spin densities[159] were calculated using VTST Tools.[138] DDEC6[158,230] partial charges, spin densities, and bond orders as well as Charge Model 5 (CM5) charges were computed using Chargemol.[374] VESTA,[560] ChemCraft,[582] and Virtual NanoLab[559] were used to visualize the structures presented in this work. LOBSTER[583] was used to perform crystal orbital Hamilton population analyses[584–586] with the pbeVaspFit2015 basis set. Charge density difference plots were made using VESTA.[560]

D.2.4 Finite Cluster Calculations

For DFT calculations involving finite cluster models, we used the Gaussian 16, Rev. A.03 program.[282] All Gaussian calculations were performed with an "ultrafine" integration grid consisting of 99 radial shells and 590 angular points per shell with symmetry conditions disabled. Partially constrained optimizations were performed with only the positions of atoms in the first and second coordination spheres of the central metal binding site allowed to relax (in addition to the guest molecule, if present) to reduce the computational cost and mimic the rigidity of the periodic framework. Vibrational frequencies were computed analytically, considering the

vibrational modes of all unconstrained atoms in the cluster model. All optimized structures were confirmed to have no imaginary vibrational modes. When used, thermochemical corrections were calculated using the ideal gas, harmonic oscillator, rigid rotor, and particle in a box approximations.[293] Thermochemical corrections were calculated at 298.15 K. Density fitting of the Coulomb integrals was used to accelerate the M06-L/def2-TZVP calculations. The most favorable adsorption modes and spin states determined from the periodic DFT calculations were used to calculate the adsorption energies with the cluster models unless otherwise specified.

D.2.5 Mg/Al-Dilution

As discussed in the Methods section, unless otherwise stated, the periodic models were modified by introducing closed-shell Mg^{2+} and/or Al^{3+} cations to reduce the computational cost and simplify the spin state analysis (Figure D.1). Changes in the O_2 or N_2 binding behavior as a result of this simplification are not expected to influence the structure−property relationships identified in this work. As shown in the "Comparing Mg-Diluted and Non-Diluted $Co_2(OH)_2$(bbta) Models" subsection, O_2 adsorption in the Mg-diluted $Co_2(OH)_2$(bbta) model (used for the screening aspect of this study) is 8 kJ/mol less exothermic in the non-diluted model (used for the more detailed investigation of this material). Prior work on Fe_2(dobdc) shows a similar 8 kJ/mol more exothermic O_2 binding energy in the non-diluted case.[328] This is within the expected error from the approximations involved in selecting an exchange-correlation functional for this problem. Generally, as was the case in the aforementioned examples, the Mg^{2+}/Al^{3+}-diluted models are likely to have slightly less exothermic binding energies because partial charge transfer cannot occur to neighboring metal sites when they are Mg^{2+} or Al^{3+} cations.

While the Mg/Al-dilution process is not expected to alter the trends discussed throughout this study, we note that there are select cases where − at least experimentally − this metal-doping can potentially have a more pronounced influence. The most apparent example is Fe_2(dobdc), which at low temperatures binds O_2 in an Fe(III)−O_2^- fashion.[24] While not readily captured in prior theoretical models,[327] it has been proposed based on spectroscopic measurements[24] that at elevated temperatures, a peroxo adduct can form with the additional electron coming from an adjacent iron cation. In this case, Mg-dilution would be expected to prevent this adjacent charge transfer, should it indeed occur.

D.3 Experimental Methods

D.3.1 Chemicals

All chemicals and solvents were purchased from commercial suppliers and used without further purification. $CoCl_2 \cdot 6H_2O$, *N,N*-dimethylformamide (DMF) (99.9%), methanol (99.8%), hydrochloric acid (36.5−38%), and KOH were purchased from Fisher Scientific. Only deionized water was used in the experiments.

D.3.2 Instrumentation

Powder X-ray Diffraction (PXRD). PXRD patterns were collected at room temperature on a STOE-STADIMP powder diffractometer equipped with an asymmetric curved Germanium monochromator (CuKα1 radiation, λ = 1.54056 Å) and one-dimensional silicon strip detector (MYTHEN2 1K from DECTRIS). The generator was set to be 40 kV and 40 mA. Powder was packed in a 3 mm metallic mask and sandwiched between two layers of polyimide tape. The measurement was carried out in transmission geometry in a rotating holder with the intensity data from 1 to 40 degrees. The scan step was set to be $2\theta = 4°$ while the scan time was 60 s per step.

X-ray Photoelectron Spectroscopy (XPS). Measurements were carried out at the KECKII/NUANCE facility at Northwestern University on a Thermo Scientific ESCALAB 250 Xi equipped with an electron flood gun and a scanning ion gun. XPS data was analyzed using Thermo Scientific Avantage Data System software and all spectra were referenced to the C1s peak (284.8 eV).

Nitrogen and Oxygen Isotherm Measurements. MOF powders were initially dried at 80 °C in a vacuum oven for a few hours. Then Co_2Cl_2(bbta) was activated at 150 °C for 12–24 h and $Co_2(OH)_2$bbta was activated at 200 °C for 12–24 h under vacuum on a Micromeritics Smart Vac prior to N_2 sorption isotherms on a Micromeritics Tristar II 3020 (Micromeritics, Norcross, GA) instrument at 77 K. Pore size distributions curves were calculated by using DFT calculations according to a carbon slit-pore model with a N_2 kernel. Around 20−50 mg of sample was used in each measurement and the Brunauer–Emmett–Teller (BET) surface area was calculated in the region P/P_0 = 0.005–0.05. O_2 isotherms were measured on 3Flex (Micromeritics) multiport surface characterization instrument with enhanced chemical resistant (ECR) analyzer where the samples were activated at 225 °C in order to remove any traces of residual solvent and/or water molecules. ***Fire Hazard Warning:*** *Users must check that the analyzer does not run with a regular oil pump since oxygen can dissolve in oil which could be a potential fire hazard. Therefore, either an oil-free pump or a pump with non-flammable oil must be used for oxygen isotherm measurements.* The temperature for oxygen isotherms was controlled by liquid-free ISO Controller Temperature Control Device (Micromeritics). UPC grade (99.996%) oxygen gas from Airgas was used for analysis. It is important to note that oxygen adsorption in $Co_2(OH)_2$bbta is very slow

compared to Co_2Cl_2bbta; therefore, increasing the equilibrium interval time was necessary (see Figure D.18 and D.19 as well as the corresponding discussion).

Diffuse Reflectance Infrared Fourier Transform Spectra (DRIFTS) measurements. A Nicolet 6700 FTIR spectrometer (Thermo Scientific) equipped with an MCT detector and a Harrick praying mantis accessory was used to collect infrared (IR) spectra. Before the measurement of the data shown in Figure D.14, $Co_2(OH)_2$(bbta) was activated at 200 °C, and Co_2Cl_2(bbta) was activated at 100 °C under high vacuum for 12 h. The DRIFTS data shown in Figures D.15 and D.16 were collected as a function of temperature without prior activation of the $Co_2(OH)_2$(bbta) material. The spectra were collected at 1 cm^{-1} resolution over 64 scans with solid KBr as the background.

D.3.3 Synthesis of Co_2Cl_2(bbta)

Co_2Cl_2(bbta) was prepared according to a reported procedure[194,280] with slight modifications. 50 mg H_2bbta (0.312 mmol) was dissolved in 10 mL *N,N*-dimethylformamide (DMF) in a 20 mL vial. In a separate 20 mL vial, 150 mg (0.630 mmol) (2 eq.) $CoCl_2 \cdot 6H_2O$ was dissolved in a solution of 10 mL methanol and 0.1 mL concentrated hydrochloric acid (HCl). The clear solutions were combined in a 100 mL jar, capped, and heated to 65 °C in an oven for 3 days. Next, the reaction mixtures were removed from the oven, and solids were separated by centrifugation. The solids were washed with DMF (three times) and methanol (three times). Solvent exchange of DMF was carried out by Soxhlet extraction with methanol for 72 hours. The materials were then activated under dynamic vacuum at 100 °C for 24 hours (yield: 70%).

D.3.4 Synthesis of $Co_2(OH)_2$(bbta)

$Co_2(OH)_2$(bbta) was prepared according to a reported procedure[277] with slight modifications. 25 mL of 1 M KOH solution was added into a sample of fully activated Co_2Cl_2(bbta) (50 mg, 0.289 mmol) in a two-necked round-bottle flask. The mixture was stirred overnight at room temperature under N_2 atmosphere. The solid was separated by centrifugation and washed with water three times. The resultant residue was finally evacuated at 80 °C overnight under vacuum and 200 °C for 12–15 h to obtain the activated product as light pink microcrystalline powder (yield: 96%) used for gas uptake studies.

D.4 Additional Results from Screening Procedure

D.4.1 Structure−Property Relationships

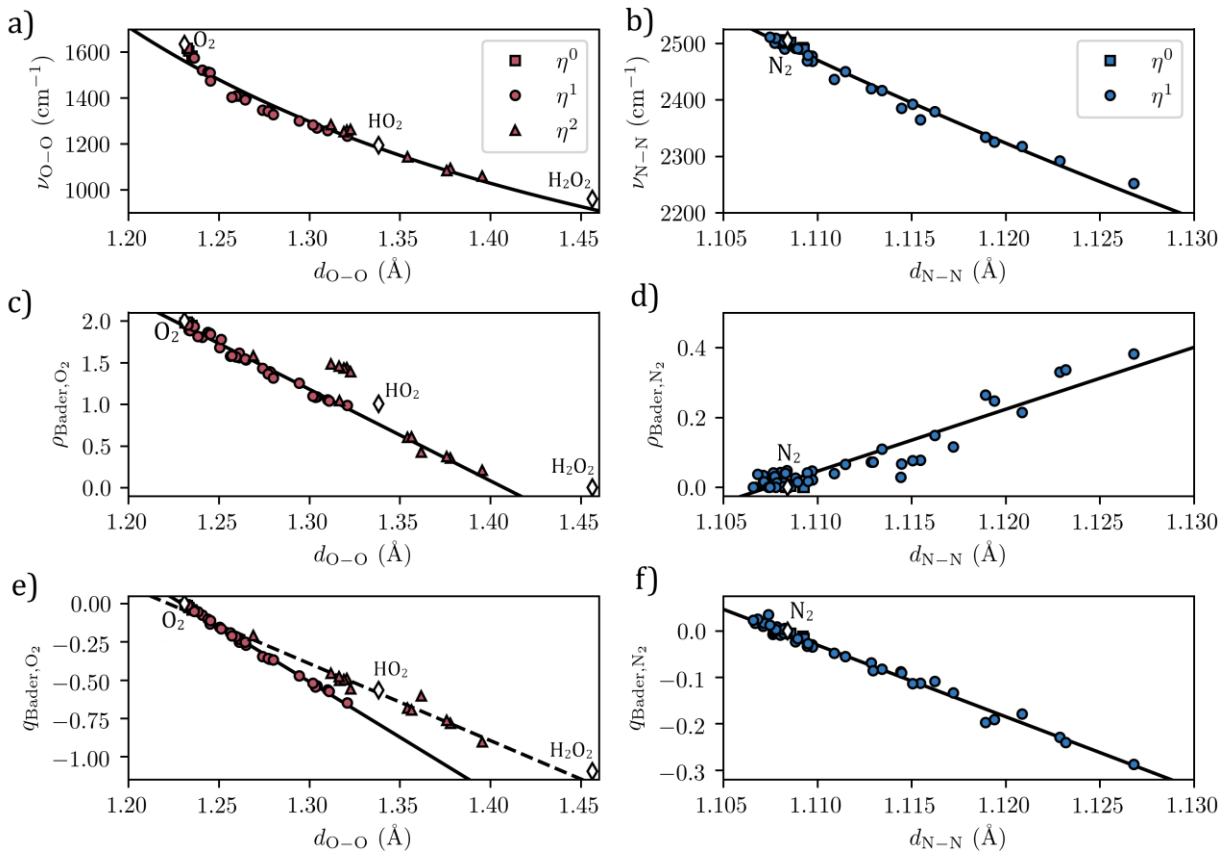


Figure D.2. a) DFT-computed vibrational stretching frequencies of adsorbed O_2, ν_{O-O}, as a function of the O−O distance, d_{O-O}, with a fit to Badger's rule (Equation D.2). b) Vibrational

stretching frequencies of adsorbed N_2, ν_{N-N}, as a function of the N−N distance, d_{N-N}, with a fit to Badger's rule (Equation D.3). c) Sum of absolute Bader spin densities on adsorbed O_2, ρ_{O_2}, as a function of d_{O-O}, with a linear fit (Equation D.4). d) Sum of absolute Bader spin densities on adsorbed N_2, ρ_{N_2}, as a function of d_{N-N}, with a linear fit (Equation D.5). e) Sum of Bader partial atomic charges on adsorbed O_2 as a function of d_{O-O}. Two separate best-fit lines are drawn to distinguish the η^0/η^1 binding modes from the η^2 binding mode (Equation D.6). (f) Sum of Bader partial atomic charges on adsorbed N_2 as a function of d_{N-N}, with a linear fit (Equation D.7). The correlations in Figure D.2 include data from all the MOFs shown in Figure 5.2, with the exception of the MIL-88B/MIL-88B-OH series, which is excluded from Figures D.2a and D.2b due to challenges in converging the SCF to 10^{-8} eV when the adsorbate atoms were displaced. Computed properties are at the M06-L/PAW level of theory, including those of the gas-phase references.

To evaluate the redox states of adsorbed O_2 and N_2, several complementary methods were considered. Experimentally, the most common approach is to measure the O−O or N−N stretching frequencies, using known reduced states of O_2 (e.g. HO_2, H_2O_2) and N_2 (e.g. N_2H_2, N_2H_4) as reference points.[310] Often, the stretching frequencies, ν, can be correlated with the bond distance, d, via Badger's rule,[587,588] which empirically states that

$$d = C\nu^{-\frac{2}{3}} + \delta \tag{D.1}$$

where C and δ are fitting parameters that are dependent on the adsorbate. As shown in Figure D.2a and D.2b, Badger's rule appears to hold reasonably well for the O_2 and N_2 complexes studied in this work. The best-fit parameters at the M06-L/PAW level of theory were found to be

$$d_{O-O}\,[\text{Å}] = 70.9\nu_{O-O}^{-\frac{2}{3}} + 0.70 \tag{D.2}$$

and

$$d_{N-N}\,[\text{Å}] = 43.8\nu_{N-N}^{-\frac{2}{3}} + 0.87 \tag{D.3}$$

where ν is in cm^{-1}. Equation D.2 and D.3 have an r^2 of 0.97 and 0.98, respectively.

Based on this spectroscopic and structural perspective of redox states, the M−O_2 complexes in this work are fairly spread out over the dioxygen−superoxide−peroxide spectrum (Figure D.2a). While assigning the "best" resonance structure can be conceptually helpful, we note that the nearly continuous range of ν_{O-O} and d_{O-O} values indicates that one should not consider the redox states to be quantized. In other words, for many of the complexes, it can be difficult to ascribe a definitive "dioxygen", "superoxide", or "peroxide" descriptor for the charge distribution. For O_2 adsorption, we also found that the side-on (η^2) binding mode results in more significantly reduced O_2 species than the end-on (η^1) mode, with all metal−dioxygen complexes in the superoxo−peroxo range having energetically preferable side-on bonding modes.

In contrast with what was found for O_2, the N−N distances are all within about 0.02 Å of the free N_2 interatomic bond distance of 1.11 Å at the M06-L/PAW level of theory (Figure D.2b). For reference, the N−N distances for N_2H_2 and N_2H_4 at the M06-L/PAW level of theory are 1.27 Å and 1.44 Å, respectively. As expected from Badger's rule, the small changes in the N−N bond distance correspond to relatively minor changes in the N−N stretching frequencies in terms of the percent deviation from that of gas-phase N_2. In some cases, a small amount of electron density is donated from the N_2 molecule to the metal, as was previously predicted for N_2 binding at the open metal sites of Fe_2(dobdc).[341]

The redox states of O_2 and N_2 can also be probed via the spin density and partial atomic charges on the bound adsorbates. As shown in Figure D.2c and Figure D.2d, there is a strong correlation between spin density and interatomic bond distance for both O_2 and N_2, such that

$$\rho_{O_2} = -11.0\ \text{Å}^{-1} \cdot (d_{O-O}) + 15.4 \qquad \text{(D. 4)}$$

and

$$\rho_{N_2} = 17.7\ \text{Å}^{-1} \cdot (d_{N-N}) - 19.6 \tag{D.5}$$

which have an r^2 of 0.99 and 0.87, respectively. The outliers in Figure D.2c (which are excluded from the best-fit line) happen to be all the $[FeO_2]^{2+}$ complexes considered in the screening study. Unlike most of the other MOFs investigated in this work, the Fe(II) sites were ferromagnetically coupled with the adsorbed (triplet) O_2 molecule (Table D.1), which may explain this deviation. Based on the larger number of metal−dinitrogen complexes with negligible spin density on the N_2 molecule (Figure D.2d), it is clear that the majority of MOFs are unable to break the N−N triple bond. Nonetheless, there are some complexes (particularly those with $\rho_{N_2} \approx 0.4$) that may be best described as having an N−N bond order of approximately 2.5.

We also investigated the sum of Bader charges on the O_2 and N_2 molecules as a function of bond distance (Figure D.2e and Figure D.2f). It is important to note that the partial atomic charges are not numerically equivalent to the formal oxidation state,[298] but they can still be used as a metric to gauge the degree of reduction of the bound adsorbates. As with the stretching frequency and spin density, there are clear correlations between the Bader charge and interatomic bond distance, such that

$$q_{O_2} = \begin{cases} -7.2\ \text{Å}^{-1} \cdot (d_{O-O}) + 8.9, & \eta^0 \text{ or } \eta^1 \\ -5.0\ \text{Å}^{-1} \cdot (d_{O-O}) + 6.1, & \eta^2 \end{cases} \tag{D.6}$$

$$q_{N_2} = -15.4\ \text{Å}^{-1} \cdot (d_{O-O}) + 17.0 \tag{D.7}$$

Notably, the side-on M−O_2 complexes fall on a different line than the end-on (and weakly physisorbed) M−O_2 complexes. For predicting q_{O_2}, the η^0/η^1 expression has $r^2 = 0.997$, and the η^2 expression has r^2 = 0.96. The expression for q_{N_2} has r^2 = 0.99.

D.4.2 Adsorbed N_2 in V-MAF-OH

As discussed in the main text, several MOFs with V^{2+} open metal sites are predicted to bind N_2, should they be synthesized. Here, we briefly highlight V-MAF-OH as one such MOF. The DFT-computed physicochemical properties suggest that the adsorbed N_2 molecule is partially reduced, as determined from a −249 cm^{-1} shift in the N−N stretching frequency compared to free N_2, a Bader spin density of 0.38 on the N_2 adsorbate, and a short V−N bond distance of 2.01 Å at the M06-L/PAW level of theory. Visualization of the molecular orbitals of V-MAF-OH using a finite cluster model indicates the presence of two spin-unrestricted singly occupied molecular orbitals that exhibit π-backbonding from the vanadium $3d$ orbitals into the π^* orbitals of N_2 (Figure D.4). The 79-atom cluster model used here (Figure D.3) was constructed from the corresponding M06-L/PAW-optimized structure. To balance the charge of the cluster model, three protons were added to the undercoordinated triazolate groups. These charge-balancing protons were relaxed at the M06-L/def2-TZVP level of theory (with the remaining framework atoms temporarily fixed in their M06-L/PAW positions) to ensure their placement was physically reasonable.

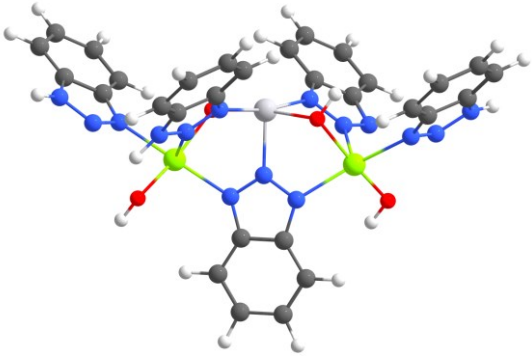
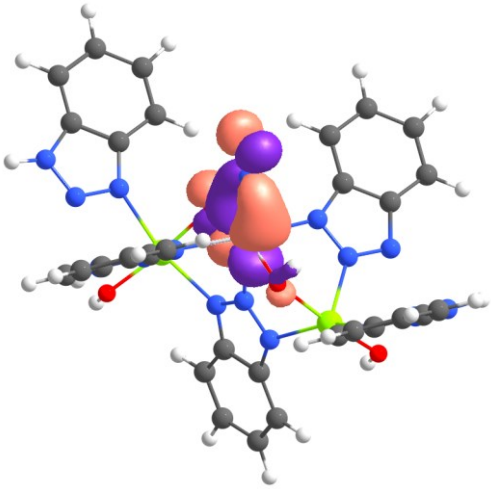


Figure D.3. Finite cluster model of V-MAF-OH. The cluster model was charge-balanced with three protons added to undercoordinated N atoms. Color key: V (silver), Mg (green), N (blue), O (red), C (gray), H (white).

a)



b)

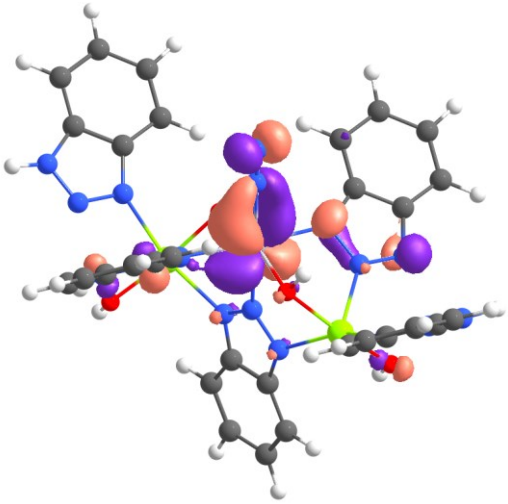


Figure D.4. Visualization of two singly occupied molecular orbitals a) α HOMO-2 and b) α HOMO-4, both of which have significant bonding interactions between the vanadium site of V-MAF-OH and proximal N atom of N_2 at the M06-L/def2-TVZP level of theory. Color key: V (silver), Mg (green), N (blue), O (red), C (gray), H (white). The orange and purple surfaces represent the two phases of the molecular orbitals.

D.4.3 Relative Binding Energies

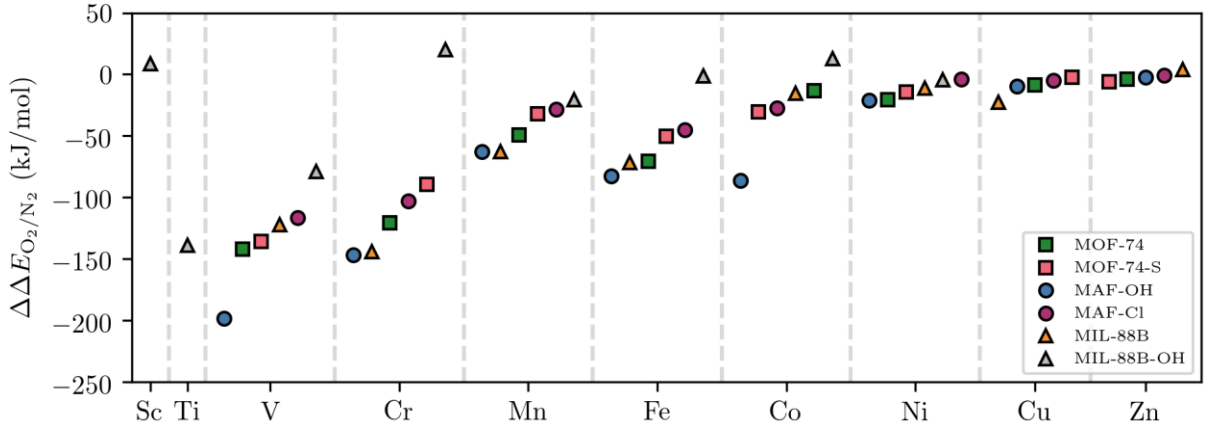


Figure D.5. Relative adsorption energy of O_2 and N_2, $\Delta\Delta E_{O_2/N_2}$, at the M06-L/PAW level of theory, computed as $\Delta\Delta E_{O_2/N_2} = \Delta E_{O_2} - \Delta E_{N_2}$.

D.4.4 Additional Results for the MOF-74 and MOF-74-S Families

Table D.2. Bader partial atomic charge on the metal site of MOF-74 and MOF-74-S before and after O_2 adsorption as a function of transition metal at the M06-L/PAW level of theory.

M	Bare MOF		MOF with adsorbed O_2		Difference upon O_2 adsorption	
	$q_{M,MOF-74}$	$q_{M,MOF-74-S}$	$q_{M,MOF-74}$	$q_{M,MOF-74-S}$	$\Delta q_{M,MOF-74}$	$\Delta q_{M,MOF-74-S}$
V	1.62	1.53	2.03	1.90	0.41	0.37
Cr	1.47	1.36	1.81	1.66	0.34	0.30
Mn	1.53	1.41	1.70	1.56	0.17	0.15
Fe	1.43	1.29	1.73	1.57	0.29	0.28
Co	1.34	1.17	1.42	1.19	0.07	0.02
Ni	1.29	1.11	1.32	1.16	0.04	0.05

For a given MOF family, all structures in the main text were considered to be isostructural for the purposes of identifying structure−property relationships. The vast majority of MOFs investigated in this work that have already been synthesized are known to be members of the corresponding isostructural series. One notable exception is Mn_2(dsbdc), which has a different coordination environment than Fe_2(dsbdc) and the M_2(dobdc) series.[206] The model used in the main text has a square pyramidal coordination environment, whereas the true structure of Mn_2(dsbdc) has been shown to have a mix of coordinatively saturated, octahedral Mn(II) centers and coordinatively

unsaturated, seesaw-like Mn(II) centers (Figure D.6). At the M06-L/PAW level of theory, the predicted O_2 and N_2 binding energies at the seesaw Mn(II) centers are −21 kJ/mol and −14 kJ/mol, respectively, which are both weaker than those predicted for Mn-MOF-74 and the Mn-MOF-74-S model that is isostructural with Mn-MOF-74 (Table D.3).

Table D.3. O_2 and N_2 adsorption energies at the M06-L/PAW level of theory for the two structures of Mn-MOF-74-S considered in this work and Mn-MOF-74.

MOF	ΔE_{O_2} (kJ/mol)	ΔE_{N_2} (kJ/mol)
Mn-MOF-74-S (seesaw)	−21	−14
Mn-MOF-74-S (square pyramidal)	−53	−21
Mn-MOF-74	−80	−31

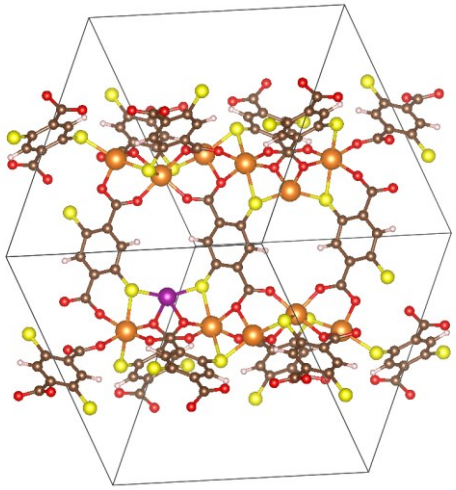


Figure D.6. Mg^{2+}-diluted model system for the mixed octahedral/seesaw-like structure of Mn_2(dsbdc). Color key: Mn (purple), Mg (orange), O (red), S (yellow), C (brown), H (white).

D.4.5 Additional Results for the MAF-X Families

The DFT-optimized M06-L/PAW structures for O_2 adsorbates bound to the metal centers of M-MAF-OH are shown in Figure D.7. The close proximity of the H atoms of the μ-OH^- groups enables stabilizing electrostatic and dispersive interactions, provided the O_2 adsorbate is partially reduced. The minimum O−H distance between the O_2 adsorbate and μ-OH^- group is 2.26, 2.20, 2.31, 2.46, 2.10, 2.32, 2.87, and 2.96 Å for M = V, Cr, Mn, Fe, Co, Ni, Cu, and Zn, respectively.

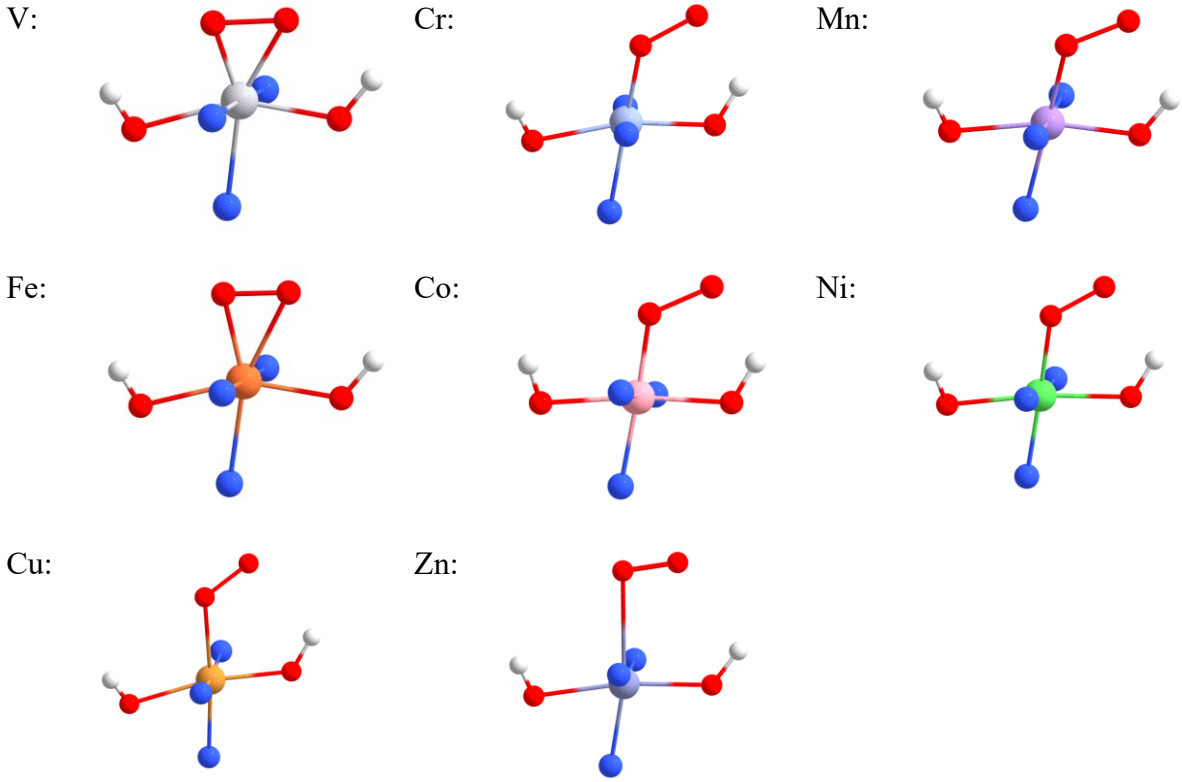


Figure D.7. DFT-optimized structures for M-MAF-OH (M = V, Cr, Mn, Fe, Co, Ni, Cu, Zn) with an O_2 adsorbate at the M06-L/PAW level of theory. Only a representative portion of the structures is shown for clarity. Color key: N (blue), O (red), H (white). The central atom is the specified metal cation.

Table D.4. Average distance between the metal binding site and bridging X ligands, $d_{(\mathrm{M-X})_{\mathrm{avg}}}$, and between the metal binding site and equatorial N ligands, $d_{(\mathrm{M-N_{eq}})_{\mathrm{avg}}}$, in the guest-free (i.e. bare) structure and after O_2 is adsorbed at the M06-L/PAW level of theory. Fe-MAF-X and Co-MAF-X are highlighted as representative examples.

Fe-MAF-X				
	$d_{(\mathrm{Fe-X})_{\mathrm{avg}}}$ (Å)		$d_{(\mathrm{Fe-N_{eq}})_{\mathrm{avg}}}$ (Å)	
X	Bare	With O_2 bound	Bare	With O_2 bound
Br	2.48	2.63	2.23	2.18
Cl	2.36	2.50	2.19	2.14
F	1.97	1.99	2.14	2.15
SH	2.44	2.51	2.19	2.17
OH	1.98	2.02	2.21	2.18
Co-MAF-X				
	$d_{(\mathrm{Co-X})_{\mathrm{avg}}}$ (Å)		$d_{(\mathrm{Co-N_{eq}})_{\mathrm{avg}}}$ (Å)	
X	Bare	With O_2 bound	Bare	With O_2 bound

Br	2.49	2.44	2.12	2.03
Cl	2.38	2.30	2.08	2.01
F	2.00	1.91	2.04	1.97
SH	2.31	2.33	2.02	2.02
OH	2.00	1.94	2.09	2.00

Table D.5. O_2 and N_2 adsorption energy as a function of the bridging anion in the Co-MAF-X (X = bridging ligand) series at the M06-L/PAW level of theory.

	μ-Br^-	μ-Cl^-	μ-F^-	μ-SH^-	μ-OH^-
ΔE_{O_2} (kJ/mol)	−46	−47	−74	−75	−103
ΔE_{N_2} (kJ/mol)	−19	−20	−17	−9	−17
$\Delta\Delta E_{O_2/N_2}$ (kJ/mol)	−27	−27	−57	−66	−86

D.4.6 Additional Results for the MIL-88B Family

Table D.6. Bader partial atomic charge on the metal binding sites of MIL-88B and MIL-88B-OH as a function of transition metal at the M06-L/PAW level of theory. As expected, the partial charges on the metal binding site of MIL-88B are lower than that of MIL-88B-OH, consistent with a lower formal oxidation state in the former.

M	$q_{M,MIL-88B}$	$q_{M,MIL-88B-OH}$
Sc	---	2.15
Ti	---	2.20
V	1.73	1.97
Cr	1.53	1.84
Mn	1.56	1.81
Fe	1.55	1.81
Co	1.42	1.50
Ni	1.35	1.45
Cu	1.29	---
Zn	1.44	---

D.5 Enhanced O_2 Binding in $Co_2(OH)_2$(bbta)

D.5.1 Benchmarking the Level of Theory

As mentioned in the main text, the PBE-D3(BJ)/PAW level of theory with a Hubbard U correction[339] of U = 3.3 eV on the d levels of the Co sites was used to optimize the volume and atomic positions of $Co_2(OH)_2$(bbta) and Co_2Cl_2(bbta). The value of U = 3.3 eV was taken from a prior benchmarking study on the oxidation energies of transition metal oxides.[340] With regards to geometries, the inclusion of a Hubbard U correction is needed to accurately capture the lattice

constants of M_2(dobdc), which is isoreticular with the M_2X_2(bbta) (X = OH, Cl) family, when there are multiple $3d$ transition metal cations.[589]

To further justify the choice of U = 3.3 eV, we chose to benchmark the level of theory against experimental results for Co-BTTri since this MOF has previously been investigated for O_2 adsorption,[216] and the Co(II) centers have similar coordination environments to those of Co_2X_2(bbta) (X = OH, Cl). The structure of Co-BTTri was modeled starting from the published crystal structure[216] with protons placed at undercoordinated N atoms of the triazolate groups to balance the charge of the framework, as done in prior work.[216,590] Antiferromagnetically aligned, low-spin (S = 1/2) Co(II) centers were considered, in accordance with experiments.[216] With regards to the lattice constants, the DFT-predicted values for Co-BTTri are in near-perfect agreement with the experimentally determined values (Table D.7). In addition, the PBE-D3(BJ)+U/PAW level of theory predicts $\Delta E_{O_2} = -44$ kJ/mol for the formation of the low-spin (S = 1/2) $[CoO_2]^{2+}$ site described in prior work.[216] To a first approximation, if one assumes the thermodynamic corrections are largely due to the loss of the O_2 translational and rotational degrees of freedom, then ΔH_{O_2} = −40 kJ/mol at 195 K. This is in close agreement with the experimentally determined, low-loading isosteric heat of −34(1) kJ/mol.[216]

Table D.7. Comparison of the DFT-computed lattice constants (at the PBE-D3(BJ)+U/PAW level of theory) and experimental lattice constants of Co-BTTri. The predicted adsorption enthalpy at 195 K at the PBE-D3(BJ)+U/PAW level of theory and experimentally determined low-loading isosteric heat of O_2 adsorption are also shown.

Lattice constants	PBE-D3(BJ)+U/PAW	Experiment[216]
a (Å)	18.49	18.53
b (Å)	18.55	18.53
c (Å)	18.54	18.53
α (°)	90.1	90.0
β (°)	89.9	90.0

γ (°)	90.0	90.0
$\Delta H_{O_2,195\ K}$	−40 kJ/mol	−34(1) kJ/mol

D.5.2 Periodic Models of Co_2X_2(bbta) (X = OH, Cl)

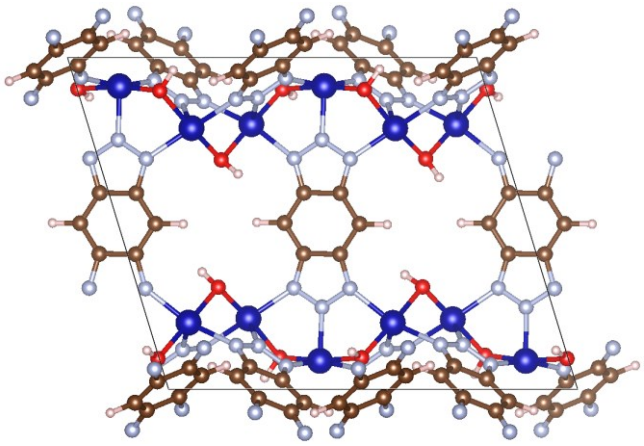
Periodic models of non-diluted $Co_2(OH)_2$(bbta) and Co_2Cl_2(bbta) were constructed with 6 Co^{2+} cations per (primitive) unit cell, analogous to the Mg-diluted analogues shown in Figure D.1. Unless otherwise specified, the guest-free Co(II) centers of Co_2X_2(bbta) (X = OH, Cl) were modeled as being on the high-spin (S = 3/2) surface, which was the most favorable for both materials at the PBE-D3(BJ)+U level of theory (see Table D.9 below). Antiferromagnetic coupling between Co(II) centers in both materials was neglected to maintain a more computationally tracTable D.ize for the model systems, as done in prior work.[25] This decision was also motivated by the finding that several Co_2(dobdc)-type frameworks have ferromagnetic intrachain coupling (and only relatively weak, temperature-dependent antiferromagnetic interchain coupling).[377,591–593] The lattice constants resulting from a volume relaxation at the PBE-D3(BJ)+U/PAW level of theory are reported in Table D.8 and are within 3% of previously published experimental lattice constants.[277] We note that the presence of water molecules is likely to slightly alter the lattice constants, so some of the discrepancy in Table D.8 may be attributed in part to residual water.

Table D.8. DFT-computed lattice constants (at the PBE-D3(BJ)+U/PAW level of theory) of guest-free $Co_2(OH)_2$(bbta) and Co_2Cl_2(bbta) compared to experiment.

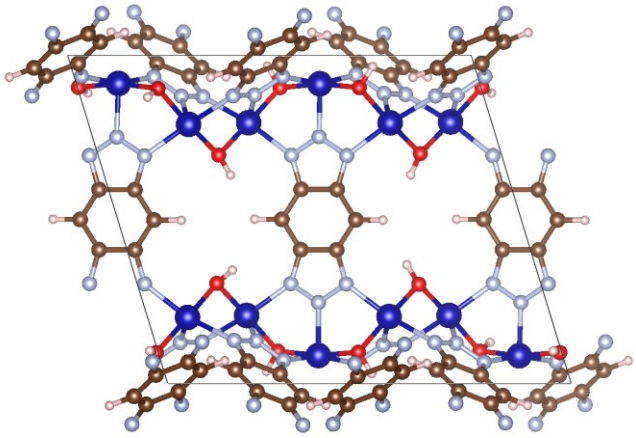
	PBE-D3(BJ)+U/PAW		Experiment[277]	
Lattice constants	$Co_2(OH)_2$(bbta)	Co_2Cl_2(bbta)	$Co_2(OH)_2$(bbta)	Co_2Cl_2(bbta)
a (Å)	7.94	8.08	8.19	8.19
b (Å)	14.89	14.70	14.55	14.67
c (Å)	14.91	14.70	14.55	14.67
α (°)	116.6	117.0	116.6	116.6
β (°)	99.8	100.5	100.8	100.7
γ (°)	100.3	100.5	100.8	100.7

We note that the H atoms of the μ-OH^- groups of $Co_2(OH)_2$(bbta) can be oriented in different directions. Temporarily using a larger $2 \times 1 \times 1$ supercell of the $Co_2(OH)_2$(bbta) primitive cell, we considered three arrangements for the H atoms, as shown in Figure D.8. All three arrangements had nearly equivalent energies following a full relaxation of the atomic positions (and cell volume), so one can expect that the H atoms are likely to be mobile at ambient conditions. For the purposes of this study, we adopt the arrangement in Figure D.8c and placed any adsorbates at the metal binding site with two aligned H atoms. Unlike the arrangements in Figure D.8a and S8b, the arrangement in Figure D.8c can be modeled using the primitive unit cell (i.e. each chain having three cobalt cations per cell). This is also what was adopted for the M-MAF-X (X = OH, SH) calculations.

a)



b)



c)

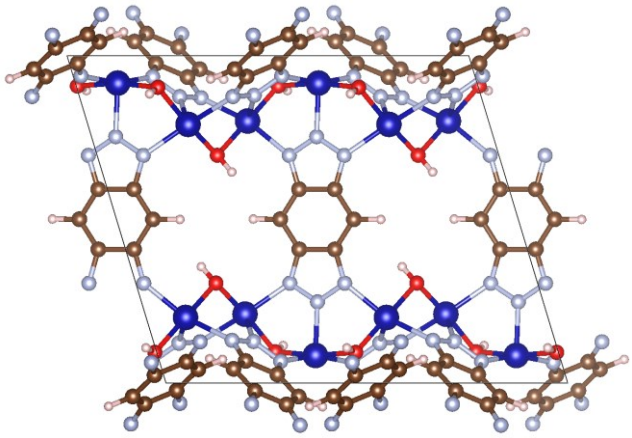


Figure D.8. If ↑ and ↓ represent the possible directions that the H atoms can point in, then schematically, the H atom orientations above can be represented as a) $(\uparrow\downarrow\uparrow\downarrow\uparrow\downarrow)_{\infty}$, b) $(\uparrow\uparrow\downarrow\downarrow\uparrow\uparrow)_{\infty}$, and c) $(\uparrow\downarrow\downarrow\uparrow\downarrow\downarrow)_{\infty}$. b) and c) are isoenergetic, and a) is 5 kJ/mol less stable per simulation cell than b) and c). Color key: Co (dark blue), O (red), N (light blue), C (brown), H (white).

D.5.3 Comparing Mg-Diluted and Non-Diluted $Co_2(OH)_2$(bbta) Models

Before proceeding with the PBE-D3(BJ)+*U*/PAW calculations of the non-diluted Co_2X_2(bbta) frameworks, we recalculated the O_2 adsorption energy in $Co_2(OH)_2$(bbta) at the M06-L/PAW level of theory (using the lattice constants in Table D.8) to confirm that strong O_2 binding is still observed. For this calculation, we modeled an end-on O_2 binding mode, high-spin (S = 3/2) Co(II) centers, and a single low-spin (S = 1/2) Co(III)−superoxo site, in accordance with the lowest energy states (see Tables D.9 and D.10 below). The O_2 adsorption energy for $Co_2(OH)_2$(bbta) was

found to be ΔE_{O_2} = −95 kJ/mol at the M06-L/PAW level of theory, which is comparable to the value of ΔE_{O_2} = −103 kJ/mol for the Mg-diluted analogue, indicating that Mg-dilution does not drastically influence the O_2 adsorption behavior in this material

D.5.4 Spin State Analysis

Using the PBE-D3(BJ)+*U*/PAW level of theory, we investigated the relative energy of the high- and low-spin states, $\Delta E_{\mathrm{HS-LS}}$, for the Co(II) centers:

$$\Delta E_{\mathrm{HS-LS}} = \frac{E_{\mathrm{HS}} - E_{\mathrm{LS}}}{n_{\mathrm{Co}}} \quad (6)$$

where E_{HS} and E_{LS} are the electronic energy (per primitive cell) of the MOF with all high-spin or all low-spin Co(II) centers, respectively. The energy per primitive cell is divided by n_{Co}, the number of Co sites per primitive cell, to obtain a relative energy on a per-Co basis (n_{Co} = 6 for VASP calculations with both $Co_2(OH)_2$(bbta) and Co_2Cl_2(bbta)). For this analysis, we let the unit cell volume and atomic positions fully relax. As shown in Table D.9, both $Co_2(OH)_2$(bbta) and Co_2Cl_2(bbta) have a high-spin ground state at the PBE-D3(BJ)+*U*/PAW level of theory. The presence of high-spin Co(II) centers has been confirmed experimentally for Co_2Cl_2(btdd) (H_2btdd = bis(1*H*-1,2,3-triazolo[4,5-b],[4',5'-i])dibenzo[1,4]dioxin), the large-pore analogue of Co_2Cl_2(bbta).[25] We note that $\Delta E_{\mathrm{HS-LS}}$ is relatively small on a per-Co basis and that the high-spin state is slightly less stable in $Co_2(OH)_2$(bbta) than Co_2Cl_2(bbta), which is consistent with the fact that OH^- has a higher ligand field strength than Cl^-. The presence of low-lying spin states is to be expected, given that spin-crossover behavior has been observed in several previously synthesized MOFs with azolate linkers.[193,216,270]

Table D.9. Difference in energy between high-spin ($S = 3/2$ per Co) and low-spin ($S = 1/2$ per Co) states, $\Delta E_{\mathrm{HS-LS}}$, in the non-diluted $Co_2(OH)_2$(bbta) and Co_2Cl_2(bbta) unit cells at the PBE-

D3(BJ)+*U*/PAW level of theory. The unit cell volume, shape, and atomic positions were fully relaxed at both spin states. A negative value for $\Delta E_{\mathrm{HS-LS}}$ indicates a preference for the high-spin state.

MOF	$\Delta E_{\mathrm{HS-LS}}$ (kJ/mol per Co)
$Co_2(OH)_2$(bbta)	−26
Co_2Cl_2(bbta)	−31

With regards to the addition of a single O_2 adsorbate to the simulation unit cell of $Co_2(OH)_2$(bbta), we found that the lowest energy state at the PBE-D3(BJ)+*U*/PAW level of theory is one in which there is negligible spin density on the cobalt binding site and a reduced spin density on the O_2 molecule (Table D.10). This $[CoO_2]^{2+}$ complex can be best described as a low-spin (S = 1/2) species that is most consistent with a Co(III)−superoxo state. In contrast, the lowest energy spin state predicted for Co_2Cl_2(bbta) is one in which an O_2 molecule is antiferromagnetically coupled to a high-spin (S = 3/2) cobalt center (Table D.11). While this $[CoO_2]^{2+}$ species is also best described as a low-spin (S = 1/2) species, it does not involve any appreciable oxidation of the metal center (or, by extension, reduction of the O_2 adsorbate). The converged magnetic moments for N_2 adsorption at high-spin Co(II) centers are shown in Tables D.12 and D.13.

Table D.10. Converged magnetic moments, number of unpaired electrons ($|n_\alpha - n_\beta|$), and relative spin state energies for the non-diluted $Co_2(OH)_2$(bbta) with adsorbed O_2 at the PBE-D3(BJ)+*U*/PAW level of theory. The magnetic moments represent the converged values for each of the six cobalt atoms in the simulation unit cell, with the last value (in italics) corresponding to the cobalt binding site. The magnetic moments on the two O atoms of O_2 are also shown.

Magnetic moments (μ_B)	$\lvert n_\alpha - n_\beta \rvert$	E (kJ/mol)
Co: [2.697, 2.682, 2.688, 2.696, 2.681, *−0.009*] O_2: [0.412, 0.480]	16	0
Co: [2.694, 2.689, 2.681, 2.694, 2.685, *2.324*] O_2: [−0.607, −0.656]	16	10
Co: [2.693, 2.695, 2.682, 2.693, 2.694, *2.799*] O_2: [0.737, 0.760]	20	38

Table D.11. Converged magnetic moments, number of unpaired electrons ($|n_\alpha - n_\beta|$), and relative spin state energies for the non-diluted Co_2Cl_2(bbta) with adsorbed O_2 at the PBE-D3(BJ)+*U*/PAW

level of theory. The magnetic moments represent the converged values for each of the six cobalt atoms in the simulation unit cell, with the last value (in italics) corresponding to the cobalt binding site. The magnetic moments on the two O atoms of O_2 are also shown.

Magnetic moments (μ_B)	$\lvert n_\alpha - n_\beta \rvert$	E (kJ/mol)
Co: [2.668, 2.662, 2.666, 2.667, 2.665, *2.574*] O_2: [−0.703, −0.792]	16	0
Co: [2.669, 2.660, 2.670, 2.669, 2.661, *−0.571*] O_2: [0.574, 0.668]	16	17
Co: [2.661, 2.661, 2.661, 2.661, 2.661, *2.676*] O_2: [0.782, 0.824]	20	17

Table D.12. Converged magnetic moments and number of unpaired electrons ($\lvert n_\alpha - n_\beta \rvert$) for the non-diluted $Co_2(OH)_2$(bbta) with adsorbed N_2 at the PBE-D3(BJ)+*U*/PAW level of theory. The magnetic moments represent the converged values for each of the six cobalt atoms in the simulation unit cell, with the last value (in italics) corresponding to the cobalt binding site. The magnetic moments on the two N atoms of N_2 are also shown.

Magnetic moments (μ_B)	$\lvert n_\alpha - n_\beta \rvert$
Co: [2.690, 2.692, 2.682, 2.692, 2.692, *2.703*] N_2: [0.114, 0.113]	18

Table D.13. Converged magnetic moments and number of unpaired electrons ($\lvert n_\alpha - n_\beta \rvert$) for the non-diluted Co_2Cl_2(bbta) with adsorbed N_2 at the PBE-D3(BJ)+*U*/PAW level of theory. The magnetic moments represent the converged values for each of the six cobalt atoms in the simulation unit cell, with the last value (in italics) corresponding to the cobalt binding site. The magnetic moments on the two N atoms of N_2 are also shown.

Magnetic moments (μ_B)	$\lvert n_\alpha - n_\beta \rvert$
Co: [2.667, 2.665, 2.665, 2.667, 2.666, *2.682*] N_2: [0.104, 0.104]	18

D.5.5 O_2 and N_2 Adsorption Behavior

The O_2 adsorption energy, ΔE_{O_2}, was computed for the non-diluted $Co_2(OH)_2$(bbta) and Co_2Cl_2(bbta) frameworks at the PBE-D3(BJ)+*U*/PAW level of theory using the aforementioned lowest energy spin states and lattice constants in Table D.8. $Co_2(OH)_2$(bbta) is still predicted to have a significantly greater O_2 affinity than Co_2Cl_2(bbta) at the PBE-D3(BJ)+*U*/PAW level of theory (Table D.14). As shown in Table D.15, an O_2 molecule bound to a cobalt site of $Co_2(OH)_2$(bbta) is predicted to have significant superoxo character. The integrated crystal orbital

Hamilton population (COHP) values for the Co−ligand interactions all increase in magnitude upon O_2 adsorption, suggesting an increase in the bond strength between the cobalt center and the surrounding ligands after O_2 binding occurs (Table D.16). This is also reflected in the shorter bond distances (Table D.16). The effective ionic radii of (six-coordinate) low-spin Co(III), low-spin Co(II), and high-spin Co(II) are 54.5 pm, 65 pm, and 74.5 pm, respectively.[358] The large difference in ionic radii for high-spin Co(II) and low-spin Co(III) enables the shorter bond distances within the first coordination sphere upon O_2 binding. As expected from the Mg-diluted model, the N_2 adsorption energy, ΔE_{N_2}, is weak in both materials (Table D.14).

Table D.14. O_2 and N_2 adsorption energies in $Co_2(OH)_2$(bbta) and Co_2Cl_2(bbta) at the PBE-D3(BJ)+*U*/PAW level of theory, using the lowest energy spin states shown in Tables D.10−D.13.

MOF	ΔE_{O_2} (kJ/mol)	ΔE_{N_2} (kJ/mol)
$Co_2(OH)_2$(bbta)	−51	−18
Co_2Cl_2(bbta)	−21	−21

Table D.15. DFT-computed properties of O_2 adsorption in $Co_2(OH)_2$(bbta) at the PBE-D3(BJ)+*U*/PAW level of theory, with computed properties of gas-phase O_2 and HO_2 used as points of comparison. The computed properties include the O−O distance (d_{O-O}), O−O stretching frequency (ν_{O-O}), sum of CM5 partial charges on O_2 (q_{CM5,O_2}), and sum of Bader spin densities on O_2 (ρ_{Bader,O_2}).

Property	$Co_2(OH)_2$(bbta)−O_2	O_2	HO_2
d_{O-O} (Å)	1.306	1.233	1.346
ν_{O-O} (cm^{-1})	1201	1568	1111
q_{CM5,O_2}	−0.31	0.0	−0.35
ρ_{Bader,O_2}	1.08	2.0	1.0

Table D.16. Bond distances, d, and integrated COHP (ICOHP) values (up to the Fermi level) between the cobalt binding site of $Co_2(OH)_2$(bbta) and the ligands in the first coordination sphere (excluding the adsorbate, if present) before and after O_2 adsorption at the PBE-D3(BJ)+*U*/PAW level of theory

	Co−$N_{eq,1}$	Co−$N_{eq,2}$	Co−N_{ax}	Co−$O_{eq,1}$	Co−$O_{eq,2}$
d (Å)	2.123	2.132	2.049	2.067	1.980

Before O_2	α-ICOHP (eV)	−1.245	−1.219	−1.124	−1.277	−1.486
	β-ICOHP (eV)	−1.529	−1.473	−1.525	−1.669	−2.006
After O_2	d (Å)	2.024	2.017	1.982	1.949	1.928
	α-ICOHP (eV)	−1.613	−1.636	−1.627	−1.762	−1.846
	β-ICOHP (eV)	−1.594	−1.619	−1.556	−1.774	−1.861

D.5.6 Unit Cell Contraction

We note that it is possible for the lattice constants to change as a function of coverage. In fact, the lattice constants of previously synthesized metal−triazolate frameworks have been shown to be sensitive to solvent and adsorbate loading.[193,216] For this analysis, we let the shape, volume, and atomic positions of the unit cell of $Co_2(OH)_2$(bbta) fully relax when loaded with one O_2 molecule (i.e. θ_{O_2} = 1/6) and two O_2 molecules (i.e. θ_{O_2} = 2/6) per unit cell. As shown in Table D.17, the unit cell is predicted to slightly contract upon adsorption of O_2 molecules at the Co(II) centers. As previously mentioned, this contraction is due to the shortened equatorial (and to a lesser degree, axial) Co−ligand bond distances upon O_2 adsorption (Table D.18), enabled by the significantly smaller ionic radius of low-spin Co(III) compared to high-spin cobalt Co(II).[358]

Table D.17. DFT-computed lattice constants of $Co_2(OH)_2$(bbta) as a function of O_2 coverage, θ_{O_2}, at the PBE-D3(BJ)+*U*/PAW level of theory. High-spin Co(II) sites are assumed.

Lattice constants	θ_{O_2} = 0 (Lattice 1)	θ_{O_2} = 1/6 (Lattice 2)	θ_{O_2} = 2/6 (Lattice 3)
a (Å)	7.94	7.88	7.84
b (Å)	14.89	14.79	14.75
c (Å)	14.91	14.78	14.68
α (°)	116.6	116.0	116.0
β (°)	99.8	100.7	101.5
γ (°)	100.3	101.3	102.0
V (Å^3)	1486.2	1445.2	1408.5

Table D.18. Bond distances between the cobalt binding site and ligands within the first coordination sphere before and after adsorption of a single O_2 molecule for the three

$Co_2(OH)_2$(bbta) unit cell volumes listed in Table D.17. Results are at the PBE-D3(BJ)+*U*/PAW level of theory.

	Lattice 1		Lattice 2		Lattice 3	
Bond distance	MOF	MOF−O_2	MOF	MOF−O_2	MOF	MOF−O_2
$d_{(\mathrm{Co-N_{eq}})_{avg}}$ (Å)	2.13	2.02	2.06	1.97	2.02	1.92
$d_{\mathrm{Co-N_{ax}}}$ (Å)	2.05	1.98	2.04	1.96	2.02	1.95
$d_{(\mathrm{Co-OH})_{avg}}$ (Å)	2.02	1.93	2.04	1.96	2.09	1.95

D.5.7 Low-Spin Co(II) Centers

Given the relatively close energy between the high- and low-spin Co(II) surfaces (Table D.9) and the fact that many metal−azolate frameworks exhibit spin-crossover behavior,[267] we considered the effect of low-spin Co(II) sites incorporated within the $Co_2(OH)_2$(bbta) framework. At the PBE-D3(BJ)+*U*/PAW level of theory, letting the $Co_2(OH)_2$(bbta) structure fully relax with all low-spin Co(II) centers results in a lattice parameters of a = 7.65 Å, b = 14.62 Å, c = 14.68 Å, α = 116.8°, β = 99.9°, γ = 100.2°, V = 1382 Å^3. For this unit cell, we modeled the O_2 adsorption process involving the oxidation of a low-spin (S = 1/2) Co(II) species to low-spin (S = 1/2) Co(III)−superoxo species. Since there is no high- to low-spin transition of the cobalt binding site in this scenario, there is a less significant decrease in the metal–ligand bond distances upon O_2 binding (Table D.19). If the lattice constants are allowed to relax, the lattice contraction is also less pronounced for θ_{O_2} = 1/6 in the low-spin case (a = 7.66 Å, b = 14.56 Å, c = 14.60 Å, α = 116.6°, β = 100.2°, γ = 100.4°, V = 1370 Å^3). The charge density difference plot for O_2 binding at low-spin Co(II) sites is shown in Figure D.9. Significant reduction of the O_2 molecule is still present, although there is less charge transfer from the equatorial ligands for the reaction on the low-spin Co(II) surface.

Table D.19. Bond distances between the cobalt binding site and ligands within the first coordination sphere before and after adsorption of a single O_2 molecule for a $Co_2(OH)_2$(bbta) unit cell with low-spin Co(II). Results are at the PBE-D3(BJ)+*U*/PAW level of theory.

	Low-spin Co(II)	
Bond distance	MOF	MOF−O_2
$d_{(\mathrm{Co-N_{eq}})_{avg}}$ (Å)	1.976	1.967
$d_{\mathrm{Co-N_{ax}}}$ (Å)	2.004	1.944
$d_{(\mathrm{Co-OH})_{avg}}$ (Å)	1.981	1.958

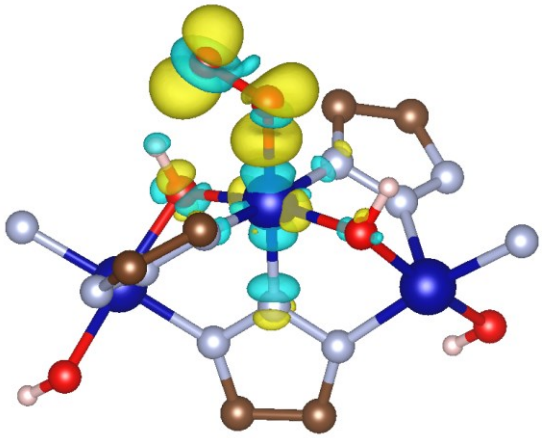


Figure D.9. Charge density difference upon O_2 adsorption in $Co_2(OH)_2$(bbta) at the PBE-D3(BJ)+*U*/PAW level of theory with low-spin Co(II) sites. The yellow and cyan surfaces represent a gain or loss of electron density, respectively (isovalue = 0.005 e^-/bohr^3). Only a representative portion of the periodic structure is shown. Color key: Co (dark blue), O (red), N (light blue), C (brown), H (white).

While the O_2 adsorption energy can, in principle, be computed for low-spin Co(II) centers with the PBE-D3(BJ)+*U*/PAW level of theory (it is ΔE_{O_2} = −91 kJ/mol), this value is likely to be overly exothermic given that the Co(II) centers are not at their ground state spin state at the given level of theory (Table D.9). As further confirmation that O_2 chemisorption occurs even if low-spin Co(II) sites were present, we relaxed the $Co_2(OH)_2$(bbta) unit cell with low-spin Co(II) sites using the M06-L functional (a = 7.63 Å, b = 14.52 Å, c = 14.55 Å, α = 116.5°, β = 100.2°, γ = 100.2°). We then carved a representative 79-atom cluster model with three Co^{2+} cations (Figure D.10) analogous to the one used for V-MAF-OH in Figure D.3. As described in the methods, the central binding site and the atoms in the first and second coordination sphere were allowed to relax. Upon

structural relaxation, the M06-L/def2-TZVP[286,287] and MN15/def2-TZVP[286,572] levels of theory predict low-spin Co(II) sites in this cluster model at the low-spin lattice constants, whereas the M06/def2-TZVP[286,288] continues to predict the high-spin ground state found at the PBE-D3(BJ)+*U*/PAW level of theory (Table D.20). The O_2 and N_2 adsorption energies are shown in Tables D.21 and D.22 for binding at low-spin Co(II) sites in the cluster models. The finite cluster calculations indicate that strong O_2 binding and weak N_2 binding are to be expected even at low-spin Co(II) sites. The $[CoO_2]^{2+}$ site continues to be best characterized as a cobalt−superoxo species (Table D.23).

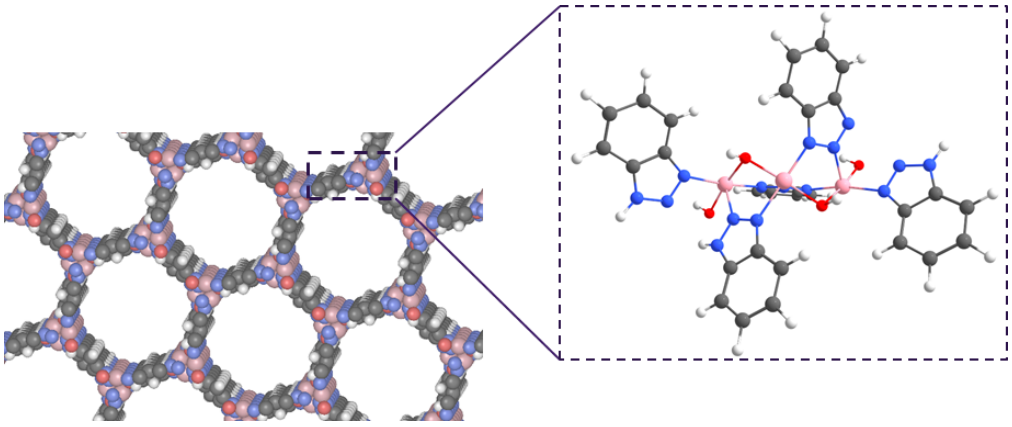


Figure D.10. Optimized periodic structure of $Co_2(OH)_2$(bbta) at the M06-L/PAW level of theory with low-spin Co(II) sites and the corresponding 79-atom cluster model carved from the periodic structure.

Table D.20. DFT-predicted spin-splitting energy, $\Delta E_{\mathrm{HS-LS}}$, for the $Co_2(OH)_2$(bbta) cluster model carved from an M06-L/PAW-optimized periodic structure with low-spin Co(II) sites (here, n_{Co} = 3).

Level of theory	%HF	$\Delta E_{\mathrm{HS-LS}}$ (kJ/mol per Co)
M06-L/def2-TZVP	0	8
M06/def2-TZVP	27	−14
MN15/def2-TZVP	44	11

Table D.21. DFT-predicted O_2 adsorption energy, ΔE_{O_2}, and O_2 adsorption enthalpy, ΔH_{O_2}, at the open metal sites of the $Co_2(OH)_2$(bbta) cluster model carved from an M06-L/PAW-optimized

periodic structure with low-spin Co(II) sites. Cluster models were modeled as being on the quartet spin surface both with and without the O_2 adsorbate.

Level of theory	%HF	ΔE_{O_2} (kJ/mol)	ΔH_{O_2} (kJ/mol)
M06-L/def2-TZVP	0	−114	−108
M06/def2-TZVP	27	−73	−67
MN15/def2-TZVP	44	−65	−59

Table D.22. DFT-predicted N_2 adsorption energy, ΔE_{N_2}, and adsorption enthalpy, ΔH_{N_2}, at the open metal sites of the $Co_2(OH)_2$(bbta) cluster model carved from an M06-L/PAW-optimized periodic structure with low-spin Co(II) sites. Cluster models were modeled as being on the quartet spin surface both with and without the N_2 adsorbate.

Level of theory	%HF	ΔE_{N_2} (kJ/mol)	ΔH_{N_2} (kJ/mol)
M06-L/def2-TZVP	0	−17	−14
M06/def2-TZVP	27	−16	−12
MN15/def2-TZVP	44	−16	−12

Table D.23. DFT-computed properties at the MN15/def2-TZVP level of theory for O_2 adsorption with the $Co_2(OH)_2$(bbta) cluster model carved from an M06-L/PAW-optimized periodic structure with low-spin Co(II) sites. Computed properties of gas-phase O_2 and HO_2 are used as points of comparison. The computed properties include the O−O distance (d_{O-O}), O−O stretching frequency (ν_{O-O}), sum of CM5 partial charges on O_2 (q_{CM5,O_2}), sum of Mulliken spin densities on O_2 ($\rho_{Mulliken,O_2}$), and Co−O_2 distance (d_{Co-O_2}).

Property	$Co_2(OH)_2$(bbta)−O_2	O_2	HO_2
d_{O-O} (Å)	1.281	1.195	1.309
ν_{O-O} (cm^{-1})	1311	1732	1250
q_{CM5,O_2}	−0.35	0.0	−0.37
$\rho_{Mulliken,O_2}$	1.08	2.0	1.0
d_{Co-O_2} (Å)	1.861	---	---

D.6 Additional Experimental Results and Characterization

D.6.1 Characterization

After exchanging the Cl^- ions of Co_2Cl_2(bbta) with OH^- ions via aqueous alkaline solution treatment, the crystallinity was maintained as verified via the measured PXRD pattern (Figure D.11). Likewise, N_2 isotherm measurements of the $Co_2(OH)_2$(bbta) material at 77 K demonstrate that the material retained high porosity (Figure D.12), with a BET surface area of 1360 m^2/g (BET area of Co_2Cl_2(bbta) is 1280 m^2/g). XPS measurements were conducted on Co_2Cl_2(bbta) and

$Co_2(OH)_2$(bbta) to determine the oxidation state of the Co species in each material and the presence of Cl^- ions.

Figure D.13a shows the XPS spectra of the Co 2p region. Both materials exhibit satellite signals around 786.2 eV and 803.4 eV for $2p_{3/2}$ and for $2p_{1/2}$ core level, respectively, which are characteristic peaks of Co(II) ions. This indicates that the oxidation state is preserved after alkali (KOH) treatment. The XPS of the Cl 2p core level of Co_2Cl_2(bbta) was observed around 197.8 eV, as shown in Figure D.13b. The peak is broad and could be resolved into $2p_{3/2}$ and $2p_{1/2}$ doublets using a curve resolver employing a Gaussian-Lorentzian line shape fit. These signals are notably absent from the $Co_2(OH)_2$(bbta) in Figure D.13c, indicating that the Cl^- ions have been replaced by OH^- ions.

In order to confirm the presence of bridging OH^- groups, infrared spectra were acquired. Figure D.14 shows the diffuse reflectance infrared Fourier transform spectra (DRIFTS) of Co_2Cl_2(bbta) and $Co_2(OH)_2$(bbta). A strong peak around 3650 cm^{-1} was observed for $Co_2(OH)_2$bbta, indicating the presence of bridging OH^- groups, which was not observed for Co_2Cl_2(bbta). DRIFTS measurements were also collected for $Co_2(OH)_2$(bbta) (without prior activation) as a function of temperature, as shown in Figures D.15 and D.16. At the activation temperature of 225 °C, the O−H stretching region retains only the feature at 3650 cm^{-1}, with essentially all of the residual water molecules removed from the framework. This O−H stretch is still present up to (at least) 275 °C. As shown in Figure D.16, the intensity of the O−H stretch does not change with increasing temperature, indicating that the μ-OH^- groups are stable and do not, for instance, dehydrate at the activation temperature.

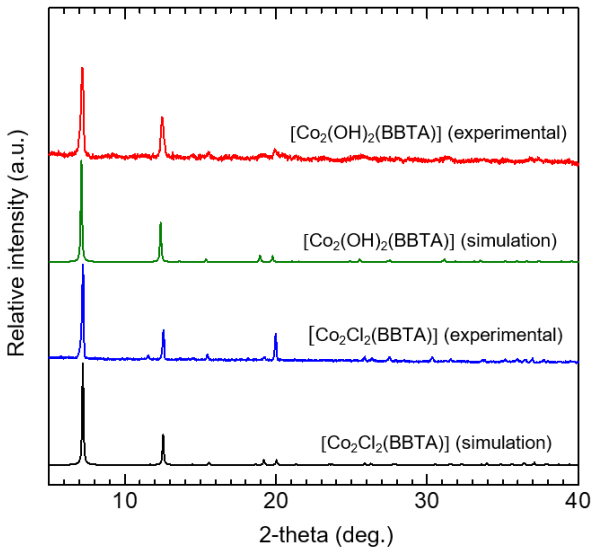


Figure D.11. PXRD patterns of Co_2Cl_2(bbta) (simulation) (black), Co_2Cl_2(bbta) (experimental) (blue), $Co_2(OH)_2$(bbta) (simulation) (green) and $Co_2(OH)_2$(bbta) (experimental) (red).

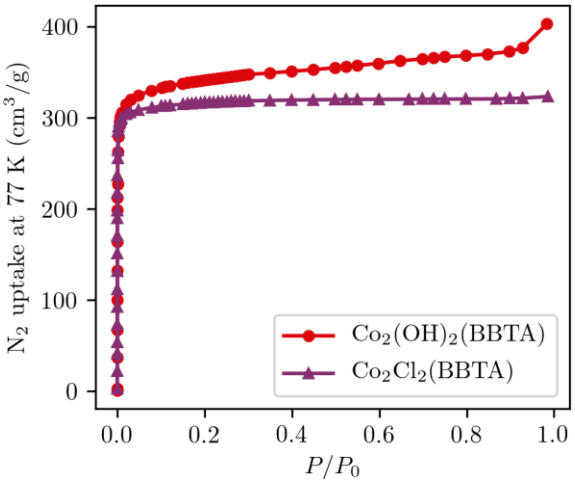


Figure D.12. N_2 adsorption isotherm at 77 K for $Co_2(OH)_2$(bbta) (red circles) and Co_2Cl_2(bbta) (purple triangles).

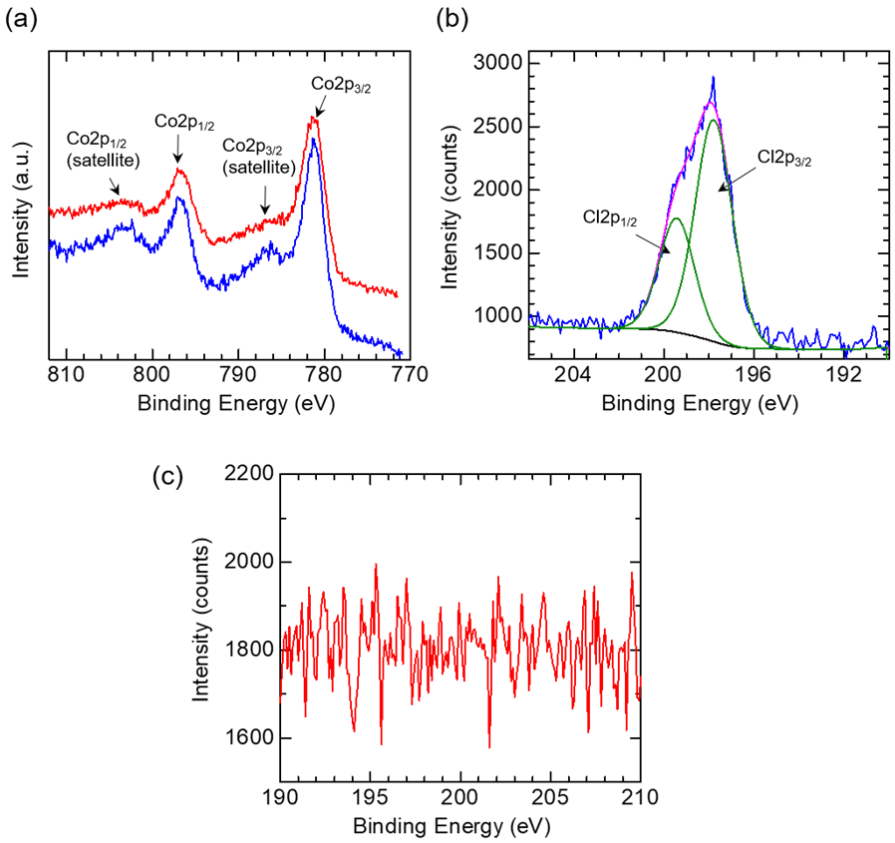


Figure D.13. X-ray photoelectron spectra of a) Co_2Cl_2(bbta) (blue) and $Co_2(OH)_2$(bbta) (red) at the Co 2p region, b) Cl 2p core level spectra of Co_2Cl_2(bbta), c) Cl 2p core level spectrum of $Co_2(OH)_2$(bbta).

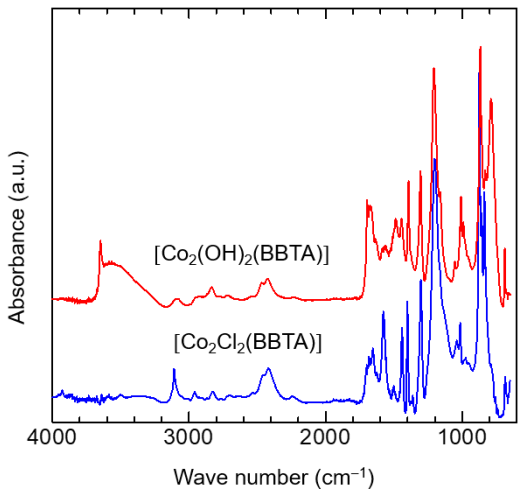


Figure D.14. Diffuse reflectance infrared Fourier transform spectra (DRIFTS) of Co_2Cl_2(bbta) (blue) and $Co_2(OH)_2$(bbta) (red). An O−H stretch is observed near 3650 cm^{-1} for $Co_2(OH)_2$(bbta).

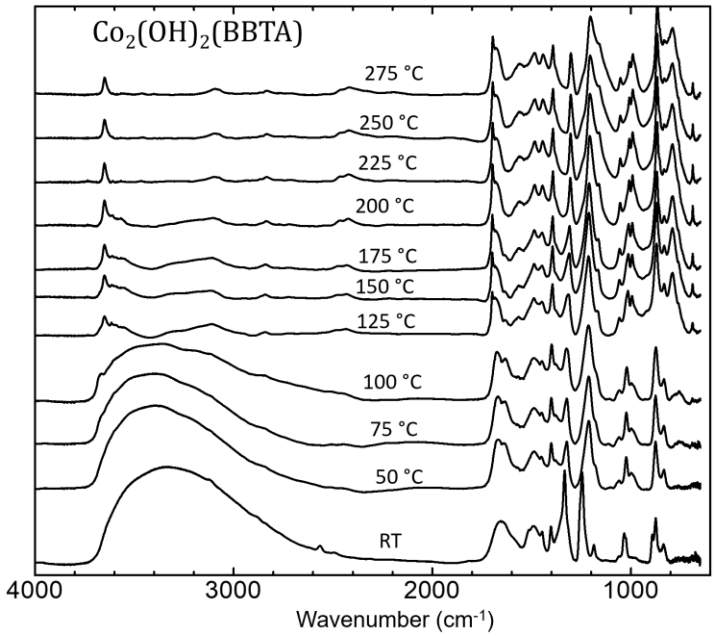


Figure D.15. DRIFTS of $Co_2(OH)_2$(bbta) as a function of temperature. The spectra are vertically offset from one another (including baseline corrections).

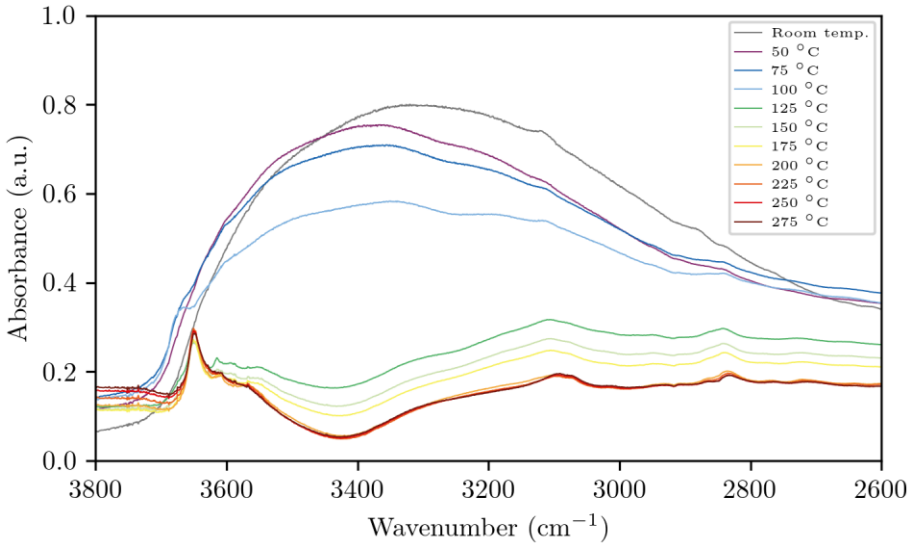


Figure D.16. DRIFTS of $Co_2(OH)_2$(bbta) as a function of temperature, focusing on the region 3800 to 2600 cm^{-1}. The spectra are overlaid with no baseline correction.

D.6.2 O_2/N_2 Selectivity Estimate

From the single-component, room temperature O_2 and N_2 isotherms, ideal adsorbed solution theory (IAST)[359] was used to predict the mixed component system with 0.21 bar O_2 and 0.79 bar N_2 at room temperature. The O_2/N_2 selectivity was calculated as

$$S_{O_2/N_2} = \frac{n_{O_2}/n_{N_2}}{x_{O_2}/x_{N_2}} \tag{S8}$$

where n_i is the loading of component i in the mixed system and x_i is the mole fraction of component i in the gas phase. PyIAST[360] was used to calculate the mixed gas behavior from the adsorption branch of the single-component isotherms, using a dual-site Langmuir fit for the O_2 isotherm

$$L(P) = M_1 \frac{K_1 P}{1 + K_1 P} + M_2 \frac{K_2 P}{1 + K_2 P} \tag{S9}$$

and Henry's law for the N_2 isotherm

$$L(P) = K_H P \tag{S10}$$

where $L(P)$ is the loading (in mmol/g) for a given component at pressure P (in bar) in the single-component system. The best-fit parameters are $M_1 = 2.585$ mmol/g, $K_1 = 0.249$ bar^{-1}, $M_2 = 0.471$ mmol/g, $K_2 = 296.317$ bar^{-1}, and $K_H = 0.194$ mmol/(g·bar). The best-fit equations are plotted over the isotherm data in Figure D.17 and agree well with the experimental data. The resulting selectivity for $Co_2(OH)_2$(bbta) is $S_{O_2/N_2} = 49$ at 298 K. We note that the O_2/N_2 selectivity for $Co_2(OH)_2$(bbta) is likely an underestimate, given that the room temperature O_2 isotherm is near – but not at – equilibrium (Figure D.18).

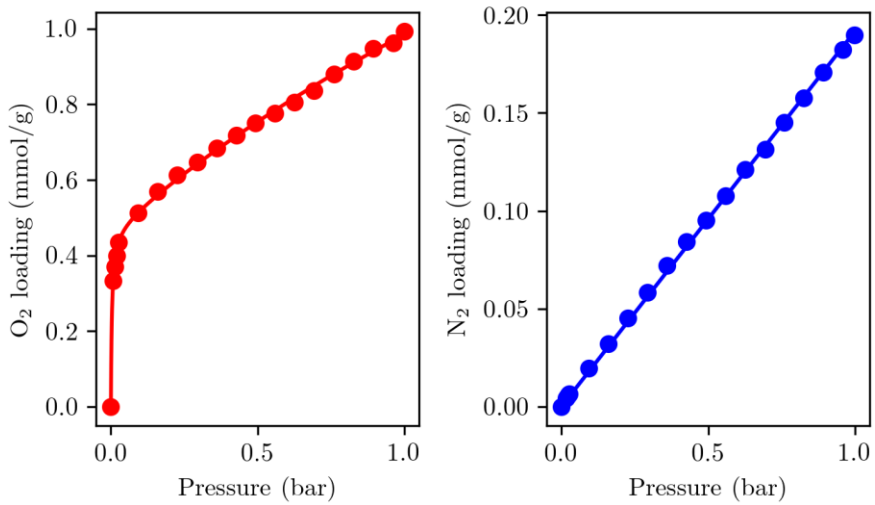


Figure D.17. Dual-site Langmuir fit to the room temperature O_2 isotherm data (left, red) and Henry's law fit to the room temperature N_2 isotherm data (right, blue) for $Co_2(OH)_2(bbta)$.

D.6.3 Rate of Adsorption Data for Oxygen Isotherms

The O_2 uptake data as a function of time for the $Co_2(OH)_2(bbta)$ sample is shown in Figure D.18, highlighting the ~20 mbar and ~226 mbar points of Figure 5.9. Here, we have set the equilibration interval time to 2000 sec (i.e. 33.3 minutes). The equilibration interval time determines how often the instrument makes a given pressure measurement, and the pressure at a given point along the isotherm is considered equilibrated when the rate of change in the pressure is less than 3% over the course of the equilibration interval time. In the case of $Co_2(OH)_2(bbta)$, we emphasize that the approach to equilibrium is gradual despite setting the equilibration interval time to 2000 sec, a much larger value than the 10 sec more commonly used for isotherm measurements. This is in clear contrast with $Co_2Cl_2(bbta)$ where the equilibration occurs within seconds, as shown in Figure D.19. We can conclude that the O_2 uptake data for $Co_2(OH)_2(bbta)$ is approaching equilibrium, although the time to reach true equilibrium is much higher for $Co_2(OH)_2(bbta)$ than for $Co_2Cl_2(bbta)$. For this reason, the O_2 uptake data in Figure 5.9 for $Co_2(OH)_2(bbta)$ can be thought

of as a lower-limit, with the possibility for a slightly increased O_2 uptake at even longer equilibration times.

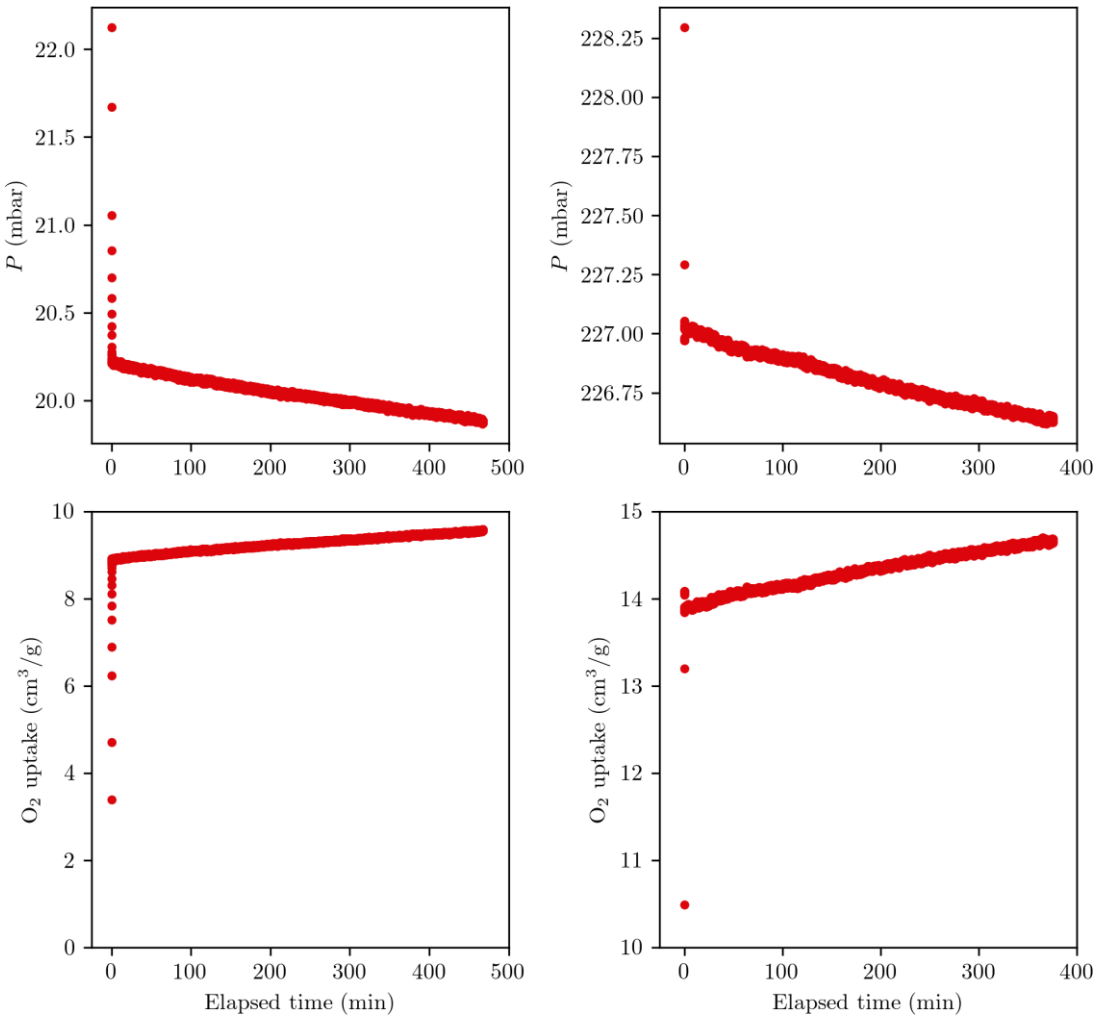


Figure D.18. Kinetic data for O_2 adsorption in $Co_2(OH)_2$(bbta) at 298 K. The pressure profiles on the top correspond to the kinetic uptake plots at the bottom.

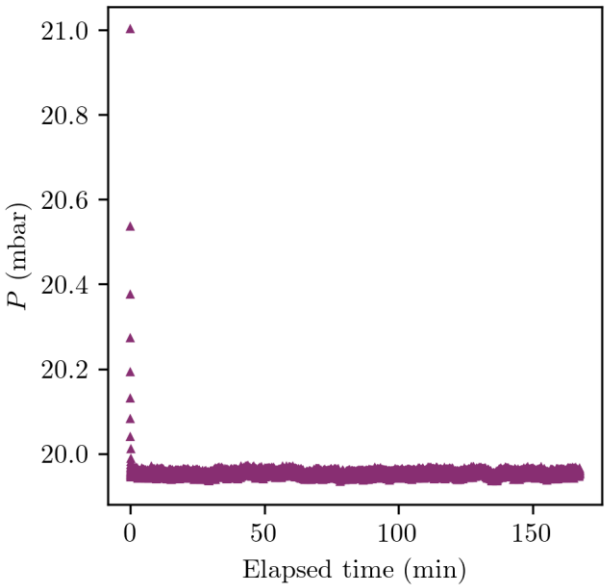


Figure D.19. Pressure vs. time data for O_2 adsorption in Co_2Cl_2(bbta) at 298 K.

D.7 Grand Canonical Monte Carlo Simulations

Before simulating the adsorption of O_2 and N_2 in Co_2Cl_2bbta and $Co_2(OH)_2$bbta, we first took the DFT-optimized crystal structures of the two activated MOFs at the PBE-D3(BJ)+*U* level of theory and calculated the surface areas using the geometry-based method implemented in the Zeo++ software[109] with a probe radius of 1.8 Å that is half of the kinetic diameter of N_2. The geometric surface areas for the DFT-optimized $Co_2(OH)_2$(bbta) and Co_2Cl_2(bbta) structures are within 91% and 98% of the experimentally determined values, respectively (Table D.24). We then simulated the adsorption of N_2 at 77 K and the adsorption of O_2 at 298 K for both MOFs, using the grand canonical Monte Carlo (GCMC) method implemented in the RASPA package.[119] The frameworks were treated as rigid, and the Lennard-Jones parameters for the framework atoms were taken from the DREIDING[594] force field, with the exception of Co atoms, where we took the parameters from UFF[595] as DREIDING does not have parameters for Co. $4 \times 3 \times 3$ supercells were used in the GCMC simulations of both MOFs to satisfy the minimum image convention requirement. Partial

atomic charges for the framework atoms were derived from periodic DFT-optimized structures using the DDEC6 method as implemented in Chargemol.[158,233,374] N_2 and O_2 molecules were defined using the TraPPE force field.[596] The standard 12-6 Lennard-Jones potential with a 12.8 Å cut-off and the Lorentz-Berthelot mixing rule was used to model the van der Waals interactions, while the Coulomb potential with Ewald summation was used to model the electrostatic interactions. For each point on the adsorption isotherms, 10,000 Monte Carlo cycles were performed in which the first 5,000 cycles were used for equilibration and only the remaining 5,000 cycles were used to collect data for calculating thermodynamic properties. Each cycle contains insertion, deletion, translation, rotation and re-insertion Monte Carlo moves (with the number of moves per cycle being 20 or the number of adsorbate molecules at the beginning of the cycle, whichever is larger).

Table D.24. Experimentally determined BET surface areas and geometric surface areas for the DFT-optimized structures.

MOF	Experimental surface area (m^2/g)	Theoretical surface area (m^2/g)
Co_2Cl_2(bbta)	1280	1300
$Co_2(OH)_2$(bbta)	1360	1486

Table D.25. Force field parameters for the Lennard-Jones potential in the GCMC simulations.

Atom	ε/k_B (K)	σ (Å)	Atom	ε/k_B (K)	σ (Å)
H (MOF)	7.64893	2.84642	Cl (MOF)	142.562	3.51932
C (MOF)	47.8562	3.47299	Co (MOF)	7.04507	2.55866
N (MOF)	38.9492	3.26256	N (N_2)	36.0	3.31
O (MOF)	48.1581	3.03315	O (O_2)	49.0	3.02

For N_2 adsorption at 77 K (Figure D.20), the shapes of the simulated and experimental isotherms agree with each other, while the simulated isotherms have higher saturation loadings due to the aforementioned difference in surface areas. For O_2 adsorption at 298 K (Figure D.21), the simulated Co_2Cl_2(bbta) isotherm is in excellent agreement with experiment. There is strong

disagreement, as expected, between the simulated O_2 isotherm for $Co_2(OH)_2$(bbta) and experiment because O_2 chemisorption cannot be captured in GCMC simulations based on generic molecular mechanics force fields.

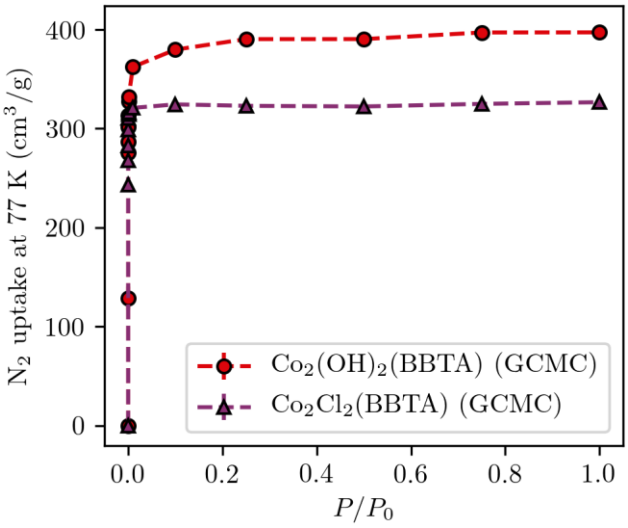


Figure D.20. GCMC-predicted N_2 adsorption isotherms at 77 K for Co_2Cl_2(bbta) (purple triangles) and $Co_2(OH)_2$(bbta) (red circles). Error bars on the GCMC simulations are shown (but are generally too small to discern given the range of the y-axis).

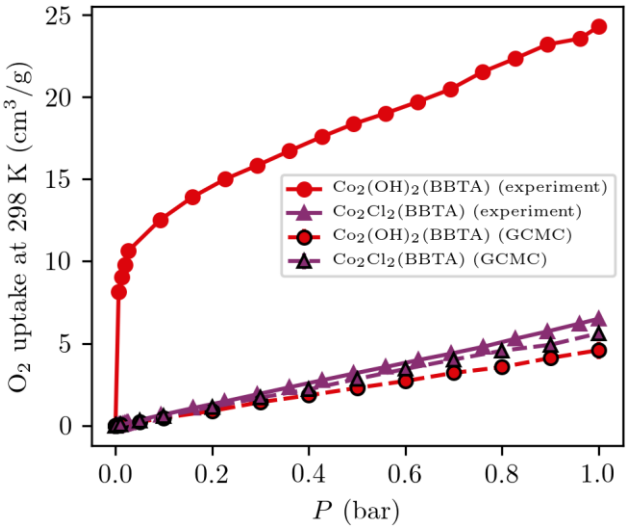


Figure D.21. GCMC-predicted O_2 adsorption isotherms at 298 K for Co_2Cl_2(bbta) (purple triangles, dashed line) and $Co_2(OH)_2$(bbta) (red circles, dashed line) compared to experiment (solid lines). Error bars on the GCMC simulations are shown (but are generally too small to discern given

the range of the y-axis). Note that the use of classical force fields implies that the simulated results will not capture possible chemisorption of the O_2 adsorbate.

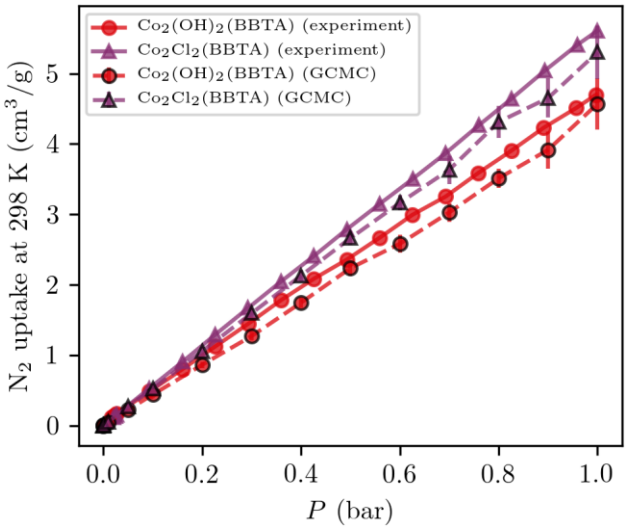


Figure D.22. GCMC-predicted N_2 adsorption isotherms at 298 K for Co_2Cl_2(bbta) (purple triangles, dashed line) and $Co_2(OH)_2$(bbta) (blue circles, dashed line) compared to experiment (solid lines). Error bars on the GCMC simulations are shown.

Appendix E. APPENDIX FOR CHAPTER 6

E.1 Additional Methods

E.1.1 Spin States

The total number of unpaired electrons per simulation cell are shown in Table E.1. For a breakdown of the individual magnetic moments, refer to the supporting dataset. Only one metal cation per unit cell is a $3d$ transition metal. The remainder are diluted with Mg^{2+} and/or Al^{3+} cations, as described in the Methods section.

Table E.1. Total number of unpaired electrons per simulation unit cell for each calculation carried out in this work.

	Guest-free MOF			MOF−O_2			MOF−N_2		
MOF	PBE	M06-L	PBE-D+U	PBE	M06-L	PBE-D+U	PBE	M06-L	PBE-D+U
Co_2Cl_2(bbta)	1	3	3	1	1	1	1	3	3
$Co_2(OH)_2$(bbta)	1	3	3	1	1	1	1	3	3
Co_2(dsbdc)	1	3	3	1	1	1	1	3	3
Co-MIL-88B	3	3	3	1	1	1	1	3	3
Co-MIL-88B-OH	0	0	2	0	2	2	0	0	0
Co_2(dobdc)	3	3	3	1	1	1	3	3	3
Cr_2Cl_2(bbta)	4	4	4	2	2	2	2	4	4
$Cr_2(OH)_2$(bbta)	4	4	4	2	2	2	2	4	4
Cr_2(dsbdc)	4	4	4	2	2	2	4	4	4
Cr-MIL-88B	4	4	4	2	2	2	2	2	4
Cr-MIL-88B-OH	3	3	3	1	1	1	3	3	3
Cr_2(dobdc)	4	4	4	2	2	2	4	4	4
Cu_2Cl_2(bbta)	1	1	1	3	3	3	1	1	1
$Cu_2(OH)_2$(bbta)	1	1	1	3	3	3	1	1	1
Cu_2(dsbdc)	1	1	1	1	1	1	1	1	1
Cu-MIL-88B	1	1	1	3	3	3	1	1	1
Cu_2(dobdc)	1	1	1	3	3	3	1	1	1
Fe_2Cl_2(bbta)	4	4	4	0	6	2	0	4	4
$Fe_2(OH)_2$(bbta)	4	4	4	0	6	6	0	4	4
Fe_2(dsbdc)	4	4	4	0	6	2	0	4	4
Fe-MIL-88B	4	4	4	4	6	6	0	4	4
Fe-MIL-88B-OH	5	5	5	1	3	3	5	5	5
Fe_2(dobdc)	4	4	4	4	6	6	4	4	4
Mn_2Cl_2(bbta)	5	5	5	3	3	3	5	5	5
$Mn_2(OH)_2$(bbta)	5	5	5	3	3	3	1	5	5
Mn_2(dsbdc)	5	5	5	3	3	3	5	5	5
Mn-MIL-88B	5	5	5	3	3	3	5	5	5
Mn-MIL-88B-OH	4	4	4	2	2	2	4	4	4
Mn_2(dobdc)	5	5	5	3	3	3	5	5	5
Ni_2Cl_2(bbta)	2	2	2	0	0	0	2	2	2
$Ni_2(OH)_2$(bbta)	2	2	2	0	0	0	2	2	2

Ni_2(dsbdc)	0	2	2	0	0	0	0	2	2
Ni-MIL-88B	2	2	2	0	0	0	2	2	2
Ni-MIL-88B-OH	1	1	1	1	1	1	1	1	1
Ni_2(dobdc)	0	2	2	0	0	0	0	2	2
V_2Cl_2(bbta)	3	3	3	1	1	1	3	3	3
$V_2(OH)_2$(bbta)	3	3	3	1	1	1	3	3	3
V_2(dsbdc)	3	3	3	1	1	1	3	3	3
V-MIL-88B	3	3	3	1	1	1	3	3	3
V-MIL-88B-OH	2	2	2	0	0	0	2	2	2
V_2(dobdc)	3	3	3	1	1	1	3	3	3

E.1.2 Adsorption Modes

Table E.2. Most stable O_2 adsorption mode for each MOF at the various levels of theory. The adsorption modes are categorized based on the CrystalNN coordination environment algorithm implemented in Pymatgen.[82] A value of 0 indicates the adsorbate is far away from the metal (best described as weak physisorption), 1 indicates end-on adsorption, and 2 indicates side-on adsorption.

	Guest-free MOF		
MOF	PBE	M06-L	PBE-D+*U*
Co_2Cl_2(bbta)	1	1	1
$Co_2(OH)_2$(bbta)	1	1	1
Co_2(dsbdc)	1	1	0
Co-MIL-88B	1	1	1
Co-MIL-88B-OH	1	1	1
Co_2(dobdc)	1	1	1
Cr_2Cl_2(bbta)	2	1	1
$Cr_2(OH)_2$(bbta)	2	1	1
Cr_2(dsbdc)	2	1	1
Cr-MIL-88B	1	1	1
Cr-MIL-88B-OH	1	2	1
Cr_2(dobdc)	1	1	1
Cu_2Cl_2(bbta)	1	1	0
$Cu_2(OH)_2$(bbta)	1	1	1
Cu_2(dsbdc)	0	0	0
Cu-MIL-88B	1	1	1
Cu_2(dobdc)	1	1	1
Fe_2Cl_2(bbta)	1	2	1
$Fe_2(OH)_2$(bbta)	1	2	2
Fe_2(dsbdc)	1	2	1
Fe-MIL-88B	1	2	2
Fe-MIL-88B-OH	1	1	1
Fe_2(dobdc)	1	2	2
Mn_2Cl_2(bbta)	1	1	1
$Mn_2(OH)_2$(bbta)	1	1	1
Mn_2(dsbdc)	1	1	1
Mn-MIL-88B	1	1	1
Mn-MIL-88B-OH	1	1	1
Mn_2(dobdc)	1	1	1
Ni_2Cl_2(bbta)	1	1	1
$Ni_2(OH)_2$(bbta)	1	1	1
Ni_2(dsbdc)	1	1	1
Ni-MIL-88B	1	1	1
Ni-MIL-88B-OH	1	1	1

Ni_2(dobdc)	1	1	1
V_2Cl_2(bbta)	2	2	2
$V_2(OH)_2$(bbta)	2	2	2
V_2(dsbdc)	2	2	2
V-MIL-88B	2	2	2
V-MIL-88B-OH	2	2	2
V_2(dobdc)	2	2	2

E.2 Additional Results

E.2.1 Fe-Containing MOFs

As shown in Figure 6.5, the PBE-D level of theory predicts less exothermic O_2 adsorption energies than the M06-L level of theory for Fe^{2+}-containing MOFs. This can be rationalized by looking at the relative energies of the spin states. Here, we highlight Fe_2Cl_2(bbta) as a representative example. As shown in Table E.3, for the favored η^1–O_2 binding geometry, the lowest energy spin state at the PBE-D level of theory is one with a low-spin Fe site and an O_2 adsorbate with a relatively low degree of spin density. This is in contrast with the M06-L and PBE-D+*U* levels of theory, which predict high spin iron sites both before and after O_2 adsorption. The spin states with O_2 best described as a singlet O_2 molecule (i.e. $|n_\alpha - n_\beta| = 4$) are also unusually stable for both adsorption geometries – one would not expect singlet O_2 to be thermodynamically accessible in these systems.

Table E.3. Energy of various spin states for the $[FeO_2]^{2+}$ complex in Fe_2Cl_2(bbta), E, at the PBE-D3(BJ)/PAW level of theory for various magnetic moments on the iron site (μ_{Fe} in Bohr-magnetons), O_2 adsorbate (μ_{O_2} in Bohr-magnetons), and number of unpaired electrons per unit cell ($|n_\alpha - n_\beta|$).

Binding mode	$\lvert n_\alpha - n_\beta \rvert$	μ_{Fe}	μ_{O_2}	E (eV)
η^1–O_2	6	3.71	1.36	−391.961
	4	3.34	0.10	−392.245
	2	2.90	−1.06	−392.263
	0	−0.75	0.61	−392.267
η^2–O_2	6	3.78	1.28	−392.142
	4	3.27	0.12	−392.004
	2	1.47	0.29	−391.958
	0	−0.63	0.60	−391.943

E.2.2 Benchmarking U Values

The Co-BTTri and Cr-BTT frameworks are both anionic. Experimentally, Co-BTTri is charge-balanced with extraframework Co(II) cations, whereas Cr-BTT is charge-balanced with Na^+ species. When evaluating the effect of different U values shown in Figure 6.9, Co-BTTri and Cr-BTT were charge-balanced with protons added to the undercoordinated N atoms, as done in prior work.[73,216,590] Motivated by prior experimental findings,[22,24,216,326] we modeled η^1–O_2 binding modes for all the MOFs except for Fe_2(dobdc), for which we considered the low-temperature η^2–O_2 mode. The spin states for each MOF considered in Figure 6.9. are shown in Table E.4. Antiferromagnetic coupling of spins between metal centers in a given ring of metals was considered for Cr-BTT and Co-BTTri in the manner previously reported for Co-BTTri.[216] For M_2(dobdc) (M = Fe, Co, Ni), only ferromagnetic orderings were considered since the energy difference between the antiferromagnetic and ferromagnetic states is known to be small.[377]

While the DFT-calculated adsorption energies are at 0 K, one can approximate the thermodynamic correction for O_2 adsorption to be roughly $5RT/2$, which is the enthalpic contribution due to a loss of translational and rotational motion of an O_2 adsorbate. This value of $5RT/2$ can be added to the 0 K (i.e. electronic) adsorption energy to yield the desired adsorption enthalpy. This correction was applied to all five MOFs in Figure 6.9, setting T = 298 K (i.e. +6 kJ/mol) for Ni_2(dobdc), Co_2(dobdc), and Cr-BTT and setting T = 200 K (i.e. +4 kJ/mol) for Fe_2(dobdc) and Co-BTTri to be consistent with the temperatures used in prior experiments involving these materials.[22,24,216,326] Note that the experimentally determined isosteric heat of O_2 adsorption in Cr-BTT is approximate and is best thought of as −65 ± 5 kJ/mol.[22]

Unlike what was considered when screening the Mg-diluted models, we let the guest-free unit cell shape and volume relax at the PBE-D3(BJ)+U level of theory for each U value in Figure 6.9. The unit cell volumes as a function of U are reported in Figure E.1. Since the unit cells contain all transition metal cations – instead of mostly Mg^{2+} cations – the unit cell volumes for some of the MOFs are highly sensitive to the value of U. For comparison, Figure E.2 compares the O_2 binding energy with lattice constants fixed at the PBE-D3(BJ) (U = 0 eV) level of theory and relaxed for each value of U for the three MOFs with the greatest change in cell volume. The change in cell volume has only a minor impact on the computed ΔE_{O_2} values in these materials. This is reassuring, as it suggests that errors in the computed lattice constants are unlikely to significantly alter O_2 adsorption energies.

Table E.4. Total number of unpaired electrons per simulation unit cell for each MOF included in Figure 6.9. The number of metal cations per simulation unit cell are also shown.

MOF	Number of metals	MOF	MOF−O_2
Ni_2(dobdc)	6	12[a]	10
Co_2(dobdc)	6	18	16[b]
Fe_2(dobdc)	6	24	26
Co-BTTri	12	0 (AFM, low-spin)[c]	2
Cr-BTT	12	0 (AFM, high-spin)[c]	2

[a]As noted in the main text, the spin state of the Ni(II) sites of Ni_2(dobdc) depends on the value of U. For consistency, we modeled the accepted high-spin Ni(II) state.[377] The transition from low- to high-spin Ni(II) sites occurs for $U \geq 2$ eV.

[b]For U = 1 eV, an arrangement of spins corresponding to a low-spin (S = 1/2) Co(III)−superoxo site was most favorable, whereas for $U \geq 2$ eV, an antiferromagnetically stabilized, low-spin (S = 1/2) Co(II)−dioxygen site was most favorable. For consistency, we modeled the latter site for all U values since it has been experimentally shown that the Co(II) sites are redox-inactive in the presence of O_2 at room temperature.[326]

[c]Antiferromagnetic coupling between metals. For Co-BTTri, each Co site has 1 unpaired electron (i.e. low-spin) and for Cr-BTT, each Cr site has 4 unpaired electrons (i.e. high-spin).

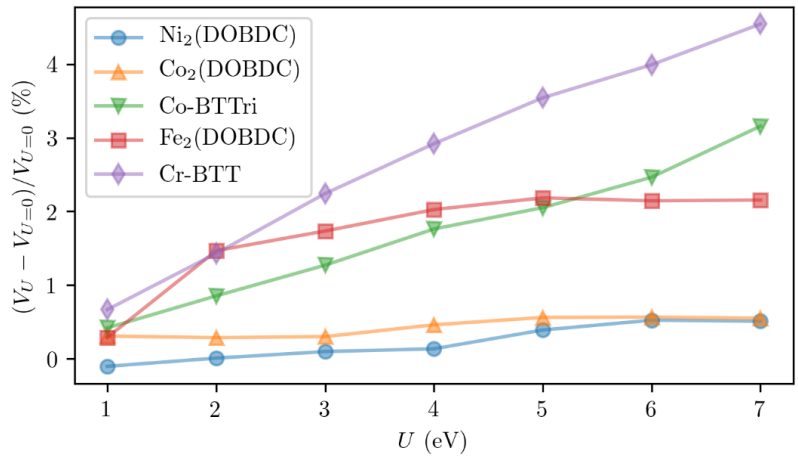


Figure E.1. Percent change in the primitive unit cell volume as a function of U compared to the PBE-D3(BJ)/PAW (U = 0 eV) case for Ni_2(dobdc), Co_2(dobdc), Co-BTTri, Fe_2(dobdc), and Cr-BTT.

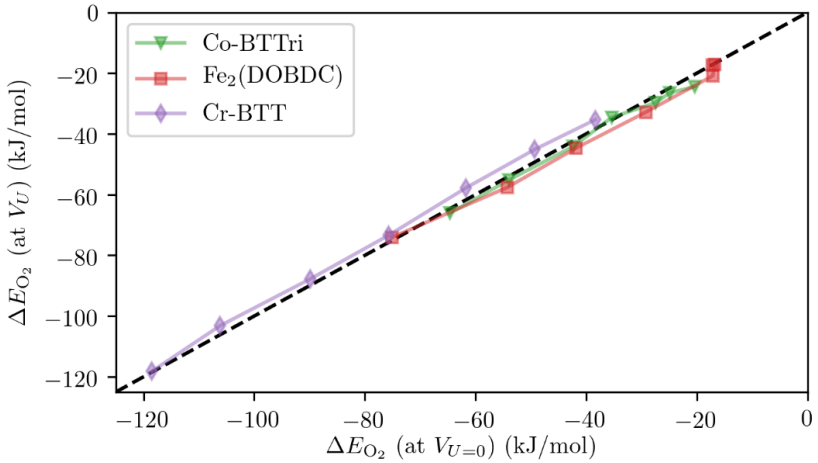


Figure E.2. O_2 adsorption energy, ΔE_{O_2}, for structures with guest-free cell volumes relaxed for each U value compared to ΔE_{O_2} values with guest-free cell volumes fixed at the PBE-D3(BJ)/PAW (U = 0 eV) level of theory. The dashed line is the line of parity.

E.2.3 Partial Charges

Throughout the text, we show Bader partial atomic charges. In the supporting dataset, we have also provided DDEC6 partial atomic charges. Generally, these two charge schemes are correlated, as shown in Figure E.3. We chose to highlight Bader charges in the main text because, for several

MOFs, the DDEC6 method predicts the O_2 adsorbate to have a net positive charge (Figure E.3), which is inconsistent with what would be expected to occur upon O_2 binding at a metal center.

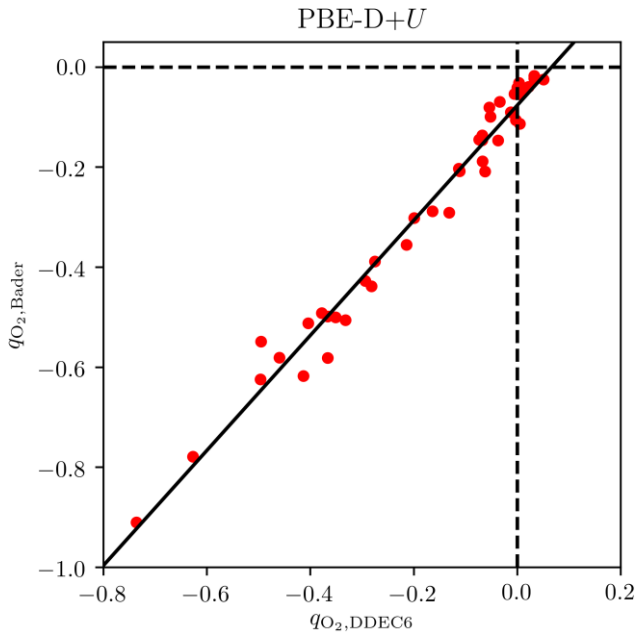


Figure E.3. Sum of Bader partial atomic charges on the adsorbed O_2 molecule as a function of the sum of DDEC6 partial atomic charges on the adsorbed O_2 molecule for all the investigated MOFs at the PBE-D+*U* level of theory. A best-fit line of $y = 1.15x - 0.077$ is shown ($r^2 = 0.98$), and the dashed lines indicate the $q_{O_2} = 0$ boundaries.

Appendix F. APPENDIX FOR CHAPTER 7

F.1 Publicly Available Data

Please refer to the following GitHub page for an overview of how to access the QMOF Database as well as for additional scripts/tools needed to reproduce the machine learning results presented in this study: https://github.com/arosen93/QMOF. Data associated with the QMOF Database is hosted via Figshare and has the following permanent DOI: 10.6084/m9.figshare.13147324.

F.2 Dataset Construction

F.2.1 Dataset Summary

A summary of the dataset construction process is shown in Figure F.1, with the important datasets in this work summarized in Table F.1.

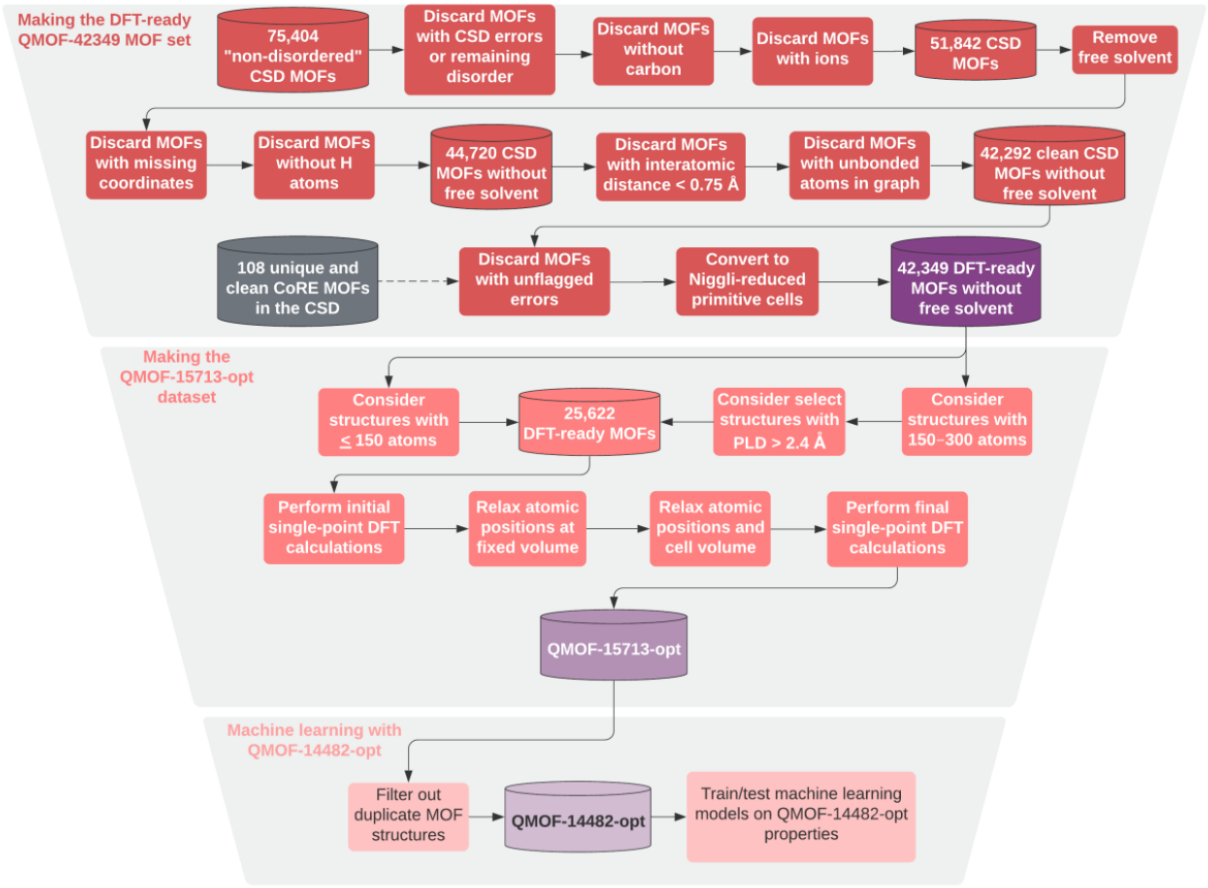


Figure F.1. Workflow for generating the dataset of DFT-ready MOF structures and DFT-computed properties. Important datasets discussed throughout this study are highlighted in purple.

Table F.1. Summary of the important datasets discussed throughout this work. All de-duplicated subsets are made using Pymatgen's `StructureMatcher` utility[82] to flag identical materials in the parent set.

Description	Name
Un-optimized, DFT-ready MOF structures.	QMOF-42349
DFT results for structures that passed all stages of the workflow. DFT-derived properties are those associated with the fully optimized ("opt") structures.	QMOF-15713-opt
De-duplicated subset of QMOF-15713-opt.	QMOF-14482-opt

F.2.2 The DFT-Ready, Free Solvent Removed QMOF-42349 Dataset

Obtaining the initial structures. In this work, we chose to take all crystal structures from the Cambridge Structural Database (CSD). As discussed below, existing databases of "cleaned"

metal–organic framework (MOF) structures that have been widely used for grand canonical Monte Carlo simulations often contain several structural fidelity issues that can significantly impact the quality of density functional theory (DFT) calculations. Furthermore, starting from the unmodified CSD structures makes it possible to use ConQuest[535] for more complicated filtering stages that take into account CSD meta-data not necessarily present in the individual crystallographic information files (CIFs). The CSD also contains many more MOF structures than existing "pre-cleaned" experimental MOF databases, making it easier to generate a large database of computed properties for subsequent machine learning studies.

Disorder and error handling. To construct the dataset of MOFs to study with DFT, we began with the Aug. 2019 release of the CSD and considered the 75,404 structures that are part of the "non-disordered" MOF subset.[44] These structures lack disorder in the framework atoms but can potentially have disorder in the remaining species (e.g. free or coordinating solvents). We used ConQuest to remove any structures that were flagged as having any remaining disorder to increase the likelihood that the resulting CIFs would be physically reasonable for DFT calculations. While the CoRE MOF database attempts to automatically resolve disorder, this automated procedure is prone to occasional errors[410–412] and so we instead neglect any disordered materials in the present study. Additionally, we use ConQuest to remove any structures with CSD-flagged errors in the crystal structure.

Carbon requirement. We ensured that all structures contain at least 1 carbon atom, as this is an inherent requirement to yield a MOF. Several structures in the CoRE MOF 2019 database lack carbon atoms, many of which are best-described as inorganic metal–phosphate frameworks (e.g. refcodes ABETAE[597], BEFLIJ[598]).

Ion handling. We did not consider any structures that were flagged as having ions, as identified via ConQuest. This step is crucial, as it is often difficult to experimentally resolve all the charge-balancing ions, and many of these structures are therefore not charge-neutral. A structure with the incorrect number of electrons makes the resulting DFT calculations unphysical. This is a potential cause of inaccurate calculation results when screening MOF databases.[410–412]

Solvent removal. We chose to remove free (i.e. unbound) solvent molecules from each structure but retained solvent bound to the metal centers. We chose not to remove bound solvent, as automated scripts to remove bound solvent have been shown to incorrectly remove framework atoms on occasion.[410] The removal of bound solvent can also lead to undesirable charge-balancing issues. For instance, the structure with refcode ASAHEJ[599] in the "all solvent removed" subset of the CoRE MOF 2019 database is missing its terminal oxo ligands because they were incorrectly assumed to be bound water molecules. Another motivating factor for only removing free solvent is that it may not be feasible to remove bound solvent during the thermal activation procedure for some MOFs. Here, we removed all free solvents that have identical SMILES strings as the molecules included in the CSD list of solvents.[44]

Missing 3D coordinates. Following removal of free solvent, we used ConQuest to filter out any structures that have an atom flagged as having missing 3D coordinates. When structures are downloaded directly from the CSD, they may be missing atoms that were not able to be assigned based on X-ray diffraction (XRD). For instance, the MOF with refcode ADATAC[600] has terminal water groups bound to metal centers, but the unmodified CIF is missing the H atoms on the water ligands. Similarly, many Zr-containing frameworks are known to have complicated proton topologies, such that it can be difficult to distinguish between terminal oxo, hydroxo, and water

ligands from XRD alone.[425] If these H atoms are not included (or an incorrect number are included), this will lead to charge-balancing issues with the overall structure. In addition, based on the CIF alone, it can be difficult to tell if a terminal O atom should be an oxo, hydroxo, or water ligand, and this can complicate the solvent removal procedure if the user wishes to remove bound solvent. We note that, in cases where the CSD entry is appropriately annotated, it may be possible to retain more structures by adding the corresponding H atoms via the CSD Python API in future studies.[407]

No H atoms. Structures without any H atoms (following removal of free solvent) were discarded. While, in principle, a MOF could have a linker without H atoms, the more common scenario is that the H atoms were simply omitted from the structure, leaving behind highly unphysical organic groups. This is a well-established limitation with existing databases of MOF crystal structures.[410–412]

Short interatomic distances. Any structures with an interatomic distance less than 0.75 Å were discarded after the above filtering procedures. This can often happen if the structure has disorder that was not appropriately flagged in the CSD entry (e.g. partial occupancies were not supplied). Nearly overlapping atoms will also create challenges for the structure relaxation algorithms.

Lone atoms. After the above procedures, we used Pymatgen to generate crystal graphs of every MOF using the CrystalNN algorithm[335,336] and removed any structures that had lone (i.e. unbonded) atoms in the graph. As an example, this is necessary to remove structures like CAXVOO,[601] which has lone H atoms in the pores of the crystal structure (which should actually be H_2).

After all of the above procedures, this resulted in 42,292 MOFs (Figure F.1).

Additional Structures Identified as MOFs in the CoRE MOF Database. To supplement this list of structures, we also considered the MOFs identified during the construction of v.1.1.2 of the 2019 CoRE MOF database.[14] The CoRE MOF structures were not used directly in this work. Rather, the corresponding CSD refcodes were identified and run through the aforementioned filtering procedure for consistency. The 2019 CoRE MOF database contains a maximum of 14,142 structures identified as MOFs, of which 13,544 can be found in the Aug. 2019 release of the CSD. Of these 13,544 structures, a total of 3,788 have no disorder, no errors, do not contain ions, and contain carbon. The majority of the structures removed in this process had disorder in the CIF. Of these 3,788 structures, 2,844 of them had H atoms and no missing coordinates. 2,699 MOFs were left after removing structures with lone atoms in the crystal graphs and ensuring that there were no interatomic distances less than 0.75 Å. Of these, only 108 were unique refcodes when compared to the 42,292 taken directly from the list of MOFs in the CSD MOF subset. This resulted in a combined dataset of 42,400 refcodes.

QMOF-42349. After the above procedure, we removed 51 additional structures that had disorder or missing H atoms not flagged via the automated ConQuest search, the majority of which have been mentioned in prior work.[407] Finally, this left us with a suitably DFT-ready dataset containing a grand total of 42,349 structures, which we refer to as the **QMOF-42349** dataset. The list of refcodes for the QMOF-42349 dataset, the script to remove free solvent, and the intermediate lists of refcodes are available with the supporting dataset.[402] All the CIFs in the QMOF-42349 dataset were converted to their Niggli-reduced primitive unit cells using Pymatgen prior to carrying out the DFT calculations.[82]

F.2.3 Completed Job Statistics to Yield the QMOF-15713-opt Dataset

From the Niggli-reduced QMOF-42349 dataset, we started by selecting MOFs with $\leq$ 150 atoms to ensure that a large number of DFT calculations could be carried out. A total of 24,002 structures fit this criterion. Of the 24,002 MOFs with $\leq$ 150 atoms per Niggli-reduced unit cell considered for the high-throughput periodic DFT screening, 19,308 successfully completed the initial single-point (i.e. static) calculation, and a total of 14,170 successfully completed every step of the workflow. While some calculations did not complete due to wall-time limits and related resource limitations, the majority of the incomplete calculations can be attributed to not meeting the strict 10^{-6} eV self-consistent field (SCF) convergence tolerance in 150 iterations during the initial single-point calculation. Many of these cases would likely have the SCF converged after a few steps of the geometry optimization, as it is common for the first few steps to require the largest number of SCF cycles to reach convergence. However, to be on the cautious side and to maximize overall resource usage, we did not consider them further or run them for a greater number of SCF iterations. We refer to the computed properties of the 14,170 DFT-optimized structures as the QMOF-14170-opt ("opt" = optimized) dataset. The corresponding single-point data on the starting structures is referred to as the QMOF-14170-SP ("SP" = single-point) dataset.

The above procedure was how v1 of the QMOF Database was generated. Since its initial release, we decided to expand the database further by including structures not already in the QMOF-14170-opt dataset. Specifically, we identified unique structures not in the QMOF-14170-opt dataset with a pore-limiting diameter greater than 2.4 Å (prior to structure relaxation) and increased the limit on the maximum atoms per cell from 150 to 300. 2185 new structures fitting

these criteria were added to the DFT workflow, of which 1543 made it through the entire structure relaxation procedure.

Collectively, a grand total of 15,713 structures completed the structure relaxation workflow. We refer to the completed structural relaxations as the **QMOF-15713-opt** dataset.

F.2.4 De-Duplication to Yield the QMOF-14482-opt Dataset

Prior to this point, duplicate structures were not removed, as slight variations in the input geometry could potentially lead to different optimized structures, and the definition of unique will ultimately depend on the application of interest. Nonetheless, for training machine learning (ML) models, it is important to have a diverse dataset, and identical structures may lead to unrepresentative testing statistics. Therefore, we used Pymatgen's `StructureMatcher` tool (using the default algorithm) on the 15,713 initial, un-relaxed structures and their relaxed counterparts to identify a unique subset of 14,482 structures. For Niggli-reduced primitive cells of two structures, the `StructureMatcher` scales the two lattice volumes, aligns the crystal lattices, and compares the atomic distances. While other methods, such as MOFid/MOFkey[79] or a comparison of the underlying crystal graphs, could be used to identify unique MOFs based on their building blocks, here we used a geometrically sensitive structure matching approach so that identical nodes/linkers but different geometries would still be considered as separate entities in the dataset. For instance, QUPZIM, QUPZIM01, and QUPZIM02 are the same MOF with the same composition and connectivity, but the first is the closed-pore analogue of the latter two, and the latter two are conformationally distinct.[431] All three are included in the de-duplicated subset, as they can potentially have different electronic structure properties (as has been shown for other flexible frameworks in the literature[430]). We acknowledge that no matter what approach is taken, there will

nonetheless be a few MOFs in the dataset with very similar structures. Other de-duplication schemes are always possible, and we encourage users to consider different approaches depending on their intended use-case.

Of the 15,713 structures, 14,482 were classified as unique and used for machine learning. The quantum-chemical properties of these 14,482 structures are collectively referred to as the QMOF-14482-opt dataset.

F.3 High-Throughput Periodic DFT Screening

F.3.1 VASP Details

Plane-wave, periodic density functional theory calculations were carried out using the Vienna *ab initio* Simulation Package (VASP) v.5.4.4.[122,123] The widely used and computationally tractable PBE exchange-correlation functional[125] with Grimme's D3 dispersion correction[126] and Becke–Johnson (BJ) damping[127] was used to generate a sufficiently large dataset for the purposes of training machine learning models. PBE with dispersion corrections has been shown to accurately capture the geometries of MOFs.[128,129] Based on prior benchmarking work,[32] the following parameters were generally used for the results presented in this study (see Table F.2 for more details). A 520 eV plane-wave kinetic energy cutoff was applied with a k-point density (KPPA) of ~1000/number of atoms, as arranged using Pymatgen 2019.9.16.[82] The VASP-recommended v.54 projector-augmented wave (PAW)[123,124] pseudopotentials were considered for all elements, with the exception of Li (for which we used the standard 140 eV default cutoff potential for computational simplicity), Eu (for which we use the Eu_3 pseudopotential rather than the Eu_2 pseudopotential since Eu(III) is more common), Yb (for which we use the Yb_3 pseudopotential rather than the Yb_2 pseudopotential since Yb(III) is more common), and W (for which we use

the _sv pseudopotential since the _pv pseudopotential is not included in the v.54 PAW set). All elements have a default cutoff of ≤ 520 eV when multiplied by 1.3 (to prevent Pulay stresses upon volume relaxation[134]). Structure relaxations were considered converged when the net force on each atom is below 0.03 eV/Å.

The accurate-precision keyword was enabled in VASP. Gaussian smearing of the band occupancies with a smearing width of 0.01 eV was applied, with extrapolation back to the 0 K limit. Symmetry operations were disabled. The SCF was converged using the "Fast" algorithm, which is a mixture of the Davidson and residual minimization method–direct inversion in the iterative subspace (RMM-DIIS) algorithms.[135] If the SCF did not converge to 10^{-6} eV within 150 iterations, the calculation was aborted and the results not considered in this work. In some cases, challenging SCF convergence can be attributed to an incorrect structure, oftentimes a result of a structure that is not charge-neutral. Spin-polarization was considered in a similar manner as several previous DFT-computed property databases.[30,33] Here, any d-block metals (excluding Zn, Cd, and Hg) were initialized with a magnetic moment of 5 μ_B. All f-block elements (excluding Lu and Lr) were initialized with 7 μ_B. All other elements were not initialized with any spin. We note that in VASP, the magnetic moments can freely change throughout the SCF convergence procedure, reaching a local minimum configuration once converged.

F.3.2 Breakdown of Sequential Steps in Periodic DFT Workflow

Each calculation was broken down into five sequential stages, similar to what has been described and benchmarked previously.[32] These stages generally include: 1) An initial, high-accuracy single-point calculation (520 eV cutoff, ~1000 KPPA); 2) A (coarse accuracy) relaxation of the atomic positions (default plane-wave kinetic energy cutoff, ~100 KPPA, 0.05 eV/Å force tolerance); 3) A

medium-accuracy relaxation of the cell volume and atomic positions (520 eV cutoff and ~100 KPPA, 0.03 eV/Å force tolerance); 4) A high-accuracy relaxation of the cell volume and atomic positions (520 eV cutoff, ~1000 KPPA, 0.03 eV/Å force tolerance); 5) A final high-accuracy single-point calculation of the fully optimized structure using the aforementioned settings. If the SCF did not converge for any step within 150 iterations, that calculation was not considered to be complete, and the remaining steps of the workflow were not carried out. If any individual stage of the workflow took greater than 2 hours per MOF, the job was canceled, and the remaining stages were also not carried out. On-the-fly error-handling was used to correct for warnings and errors should they appear,[32] but if for any reason the job crashed and could not be successfully continued, that MOF was also not considered further. The VASP input parameters are summarized in Table F.2. All VASP calculations were carried out using the Atomic Simulation Environment (ASE) 3.19.0b1.[83] Band gaps were obtained using `pymatgen.io.vasp.outputs.Eigenval()` with an occupancy tolerance of 10^{-8}. Partial atomic charges, spin densities, and effective bond orders were computed using the density-derived electrostatic and chemical (DDEC6) method[158,230–232] as implemented in Chargemol 09-26-2017.[374] Charge Model 5 (CM5) charges[289] were also computed using Chargemol 09-26-2017. PyMOFScreen commit #e9768a5 was used to manage and carry out the automated DFT calculations.[139]

Table F.2. ASE input arguments for the VASP calculators used in the screening workflow, excluding file I/O-related keywords.* Note that the appropriate pseudopotentials can be automatically selected with `setups={'base':'recommended','Li':'', 'Eu':'_3','Yb':'_3', 'W':'_sv'}`.

Flag	Stage 1	Stage 2**	Stage 3	Stage 4	Stage 5
`xc`	'PBE'	'PBE'	'PBE'	'PBE'	'PBE'
`ivdw`	12	12	12	12	12
`encut`	520[a]	400	520	520	520
`kppa`[b]	1000[a]	100	100	1000	1000
`isif`	—	2	3	3	—

ibrion	—	2	2	2	—
prec	'Accurate'	'Accurate'	'Accurate'	'Accurate'	'Accurate'
ismear	0	0	0	0	0
sigma	0.01	0.01	0.01	0.01	0.01
ediff	1E-6	1E-4	1E-6	1E-6	1E-6
algo	'Fast'	'Fast'	'Fast'	'Fast'	'Fast'
nelm	150	150	150	150	150
nelmin	3	3	3	3	—
lreal	False	False[c]	False[c]	False	False
nsw	0	250[d]	30[e]	30[e]	0
ediffg	—	-0.05	-0.03	-0.03	—
lorbit	11	11	11	11	11
isym	0	0	0	0	0
symprec	1E-8	1E-8	1E-8	1E-8	1E-8

[a]For structures in v1 of the QMOF Database (QMOF-14170), "Stage 1" (i.e. the static calculation on the starting structure) used the same settings as "Stage 5" (i.e. the static calculation on the optimized structure) to enable direct comparisons between properties (e.g. Figure F.5). For subsequent additions to the database, "Stage 1" used the same encut and kppa as in "Stage 2" for increased computational efficiency.
[b]kppa = k-point density, computed with the automatic_density() tool in Pymatgen. The choice of whether the grid should be Γ-centered or not (i.e. gamma=True or gamma=False) and how the k-points are distributed among the three lattice dimensions are also determined based on this Pymatgen utility.
[c]Switches to lreal='Auto' if the VASP output file suggests doing so due to a large unit cell (only for Stages 2 – 3).
[d]The max 250-cycle relaxations of atomic positions were sequentially repeated until the force tolerance given by |ediffg| (0.05 eV/Å) was achieved.
[e]The max 30-cycle volume relaxations were sequentially repeated until the force tolerance given by |ediffg| (0.03 eV/Å) was achieved. For Stage 4, after this process was completed, a final max 100-cycle volume relaxation was carried out for good measure.
[*]Slight changes to the input parameters that do not affect the accuracy of the results may occur during the workflow to correct for errors on-the-fly. For instance, the conjugate-gradient (CG) algorithm (ibrion=2) often leads to a bracketing error when the potential energy surface is flat, and in such a scenario the geometry optimization algorithm automatically switches to the Fast Inertial Relaxation Engine (FIRE)[137] (ibrion=3, iopt=7, potim=0).
[**]The coarse-accuracy update of the atomic positions is preceded by an initial relaxation using the BFGSLineSearch algorithm in ASE until the maximum net force is less than 10 eV/Å. Empirically, we have found that this algorithm is better at resolving high forces without the structure "exploding" when compared to the CG algorithm.

F.3.3 Further Investigation of Selected MOFs

For select calculations, we use a hybrid-level functional to improve the quality of the band gap predictions. Using the PBE-D3(BJ) wavefunction and structure as a starting point, the HSE06-D3(BJ) level of theory[465–467] was used to re-relax the unit cell shape, volume, and atomic positions. Due to the high computational cost when running periodic DFT calculations with hybrid functionals, a looser force tolerance of 0.05 eV/Å was adopted. For all HSE06-D3(BJ) calculations, the VASP-recommended preconditioned conjugate gradient "all bands simultaneous

update of orbitals" algorithm[132,133,338] (`algo='All'`) was used to converge the SCF, the SCF convergence was set to a slightly looser value of 10^{-5} eV, and the density of states (DOS) was evaluated with 3000 grid-points. For increased computational efficiency, the HSE06-D3(BJ) structure relaxations were occasionally carried out using a smaller k-point grid than the single-point calculation used to evaluate the band gap and DOS for select materials (Table F.3). All other settings remain unchanged from "Stage 5" of Table F.2.

Table F.3. k-point grids for selected MOFs at the HSE06-D3(BJ) level of theory. k-points (low) and (high) refer to the structure relaxation and subsequent electronic structure analysis, respectively.

CSD Refcode	k-points (low)	k-points (high)
LOJLAZ	$2 \times 2 \times 1$	$2 \times 2 \times 1$
RAXNEK	$2 \times 1 \times 1$	$3 \times 2 \times 1$
WAQMEJ	$2 \times 1 \times 1$	$3 \times 1 \times 1$
GUTYAW	$2 \times 2 \times 1$	$4 \times 4 \times 1$

Table F.4. HSE06-D3(BJ) primitive cell lattice parameters compared with experiment. Note that any free solvent present in the crystal structure was removed from the framework in the DFT calculations. LS = low spin; HS = high spin.

CSD Refcode		a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
LOJLAZ-LS	Theory	10.03	10.03	15.06	90.0	90.0	74.8
	Exp.	10.07	10.07	15.10	90.0	90.0	73.6
LOJLAZ-HS	Theory	10.40	10.40	15.46	89.9	90.0	71.0
	Exp.	10.38	10.38	15.50	90.0	90.0	70.1
RAXNEK	Theory	7.99	11.56	12.47	117.6	99.4	90.0
	Exp.	7.98	11.68	12.79	117.2	100.2	90.0
WAQMEJ	Theory	7.98	13.65	14.12	91.6	99.1	90.1
	Exp.	8.28	13.81	14.08	91.1	97.1	90.5
GUTYAW	Theory	4.88	4.88	14.90	86.9	86.9	66.9
	Exp.	4.97	4.97	14.98	87.0	87.0	66.4

The HSE06-D3(BJ) lattice parameters for these materials are shown in Table F.4, and the relevant spin states are discussed below.

LOJLAZ, $Fe(bipytz)(Au(CN)_2)_2$ (bipytz = 3,6-bis(4-pyridyl)-1,2,4,5-tetrazine): This material has Fe(II) and Au(I) species. Experimentally, it has been shown that LOJLAZ is a spin-crossover framework that has a low-to-high spin transition with increasing temperature.[475] We consider both spin states in this work as a matter of consistency with the spin-crossover behavior observed experimentally. For reference, the high spin state is predicted to be 42 kJ/mol (per cell) more stable than the low spin state at the HSE06-D3(BJ) level of theory.

RAXNEK, $Fe(sq)(bpee)(H_2O)_2$ (bpee = 1,2-bis(4-pyridyl)ethylene; sq = squarate): This material has Fe(II) species, which are known to exist in the high spin state with antiferromagnetic coupling.[474] At the HSE06-D3(BJ) level of theory, a high-spin ground state is found. Both ferromagnetic and antiferromagnetic states were found to have comparable structures and energies, so we model the latter as a matter of consistency with the reported experiments.

WAQMEJ, $(TTF)[\{Rh_2(CH_3CO_2)_4\}_2TCNQ]$: This material is reported to have diamagnetic (formally) Rh(II) dimers with antiferromagnetically coupled TTF–TCNQ species such that the net magnetic moment is zero.[471] A spin-unrestricted state with a net magnetic moment of zero was found to be the ground state at the HSE06-D3(BJ) level of theory.

GUTYAW, $Sr[C_2H_4(SO_3)_2]$: This material has Sr(II) cations, and the framework is modeled as spin-restricted based on its structure.

F.4 Machine Learning Details

F.4.1 Software and Hardware Details

Regression-based machine learning model development was carried out using scikit-learn v.0.23.2 and the standard SciPy stack with NumPy v.1.19.2, pandas 1.1.5, Matplotlib 3.3.2, and Seaborn v.0.11.1.[561,563,602–605] Pymatgen[82] v.2020.12.3 was used to analyze structures and generate

descriptors. The smooth overlap of atomic positions (SOAP) features were computed using DScribe v.0.4.0.[606] The sine Coulomb matrix and Meredig and Agrawal et al.[436] features were generated using Matminer v.0.6.4.[607] Crystal graph convolutional neural networks (CGCNNs) were based on the work of Xie and Grossman[406] and used PyTorch v.1.6.0[608] for constructing and evaluating the neural networks. Specifically, our CGCNN code is built upon commit #d612a69 of the CGCNN code[609] with slight variations, as reflected in our fork of the CGCNN code.[610] This fork saves the crystal graphs to .pkl files so they can be read in as-needed instead of needing to be re-computed when the memory cache is filled. This is a common problem with MOF crystal graphs given the large size of the unit cells. A branch of this revision[610] also makes it possible to use Pymatgen-computed crystal graphs rather than those based on a fixed number of neighbors, although we did not observe any improvement when using a crystal graph based on the CrystalNN algorithm.[335,336] This is potentially because there are many crystal structures in the QMOF Database that are connected in 1D or 2D, such that there are disconnected regions of the Pymatgen-generated crystal graph. PTable Trends v.2.0[611] was used to generate a heat map over the periodic table. Zeo++ v.0.3 was used for the pore diameter calculations using the "high accuracy" flag.[109] PyProcar v.5.6.1 was used to parse the DOS data.[612] Timing data for the machine learning models are reported using Python 3.8.5 on a laptop with an Intel Core i7-9750H CPU. For the CGCNNs, CUDA v.10.1 was used to enable GPU support with an NVIDIA GeForce RTX-2070 (Max-Q Design) graphics card.

F.4.2 Dataset Handling for Training and Evaluating Machine Learning Models

Unless otherwise stated 80% of the 14,482 data points was reserved for training while 20% was held-out for testing of the kernel ridge regression (KRR) models. To optimize the hyperparameters

and determine the optimal ML models, 5-fold cross-validation of the training set was applied for KRR. Due to the higher computational cost when training neural network models, for CGCNN, 80% of the data was reserved for training, 10% was held-out for validation, and 10% was reserved for testing. In all cases, performance of the models on the testing data was not inspected until the end of the project when ideal models were determined on the basis of the validation process. Data splitting was done via purely random sampling. To account for minor variations in model performance due to sampling bias, all the performance statistics in Table 7.1 and Figure 4A are reported as averages over five separate runs with different random seeds for the data splitting (arbitrarily chosen in advance to be 42, 125, 267, 541, and 582). Elsewhere, a constant seed is used for consistency (chosen in advance to be 42).

F.4.3 Learning Curves

For the learning curves in Figure 7.4a, training set sizes of 2^7, 2^8, 2^9, 2^{10}, 2^{11}, 2^{12}, 2^{13}, and 80% of the full dataset of 14,482 data points were investigated. Powers of 2 were chosen to allow for equidistant spacing on a logarithmic grid. For internal consistency, the same testing set was used (for a given data-splitting seed) regardless of training set size. The same validation set was also used (for a given data-splitting seed) for the CGCNN models. For the KRR models, 20% of the full QMOF-14482-opt dataset was held-out for testing. For the CGCNN models, 10% of the full QMOF-14482-opt dataset was used for validation, and 10% of the full QMOF-14482-opt dataset was held-out for testing.

F.4.4 Kernel Ridge Regression

KRR combines the kernel trick with ridge regression.[613] Like all regression methods, the goal of KRR is to predict a response variable $\mathbf{y}$ from a set of individual input vectors $\mathbf{x}$ (which, when

combined, form a feature matrix $\mathbf{X}$ containing an encoding of each individual material). KRR, being a kernel method, achieves this by transforming $\mathbf{X}$ into a kernel matrix $\mathbf{K}$ that describes the similarity between every pair of materials in $\mathbf{X}$. In this way, KRR has a closed-form solution given by

$$\mathbf{w} = (\mathbf{K}_{\text{train}} + \lambda \mathbf{I})^{-1} \mathbf{y}_{\text{train}} \tag{F.1}$$

where $\mathbf{w}$ is the vector of model weights, $\mathbf{K}_{\text{train}}$ is the training set kernel matrix, λ is the regularization hyperparameter, $\mathbf{I}$ is the identity matrix, and $\mathbf{y}_{\text{train}}$ is the training set values to predict. For scikit-learn's implementation of KRR, a parameter α is supplied, which is defined as $\alpha \equiv \lambda/2$.

With the model weights obtained, new values can be predicted via

$$\mathbf{y}_{\text{ML}} = \mathbf{K}_{\text{test}} \mathbf{w} \tag{F.2}$$

where $\mathbf{y}_{\text{ML}}$ are the ML-predicted $\mathbf{y}$ values for a new kernel matrix of the testing set $\mathbf{K}_{\text{test}}$. For N training samples and M testing samples, $\mathbf{K}_{\text{train}}$ will have dimensions of $(N \times N)$ and $\mathbf{K}_{\text{test}}$ will have dimensions of $(M \times N)$. Here, $\mathbf{K}_{\text{train}}$ represents the similarity between every pair of structures in the training set, whereas $\mathbf{K}_{\text{test}}$ represents the similarity between each structure in the training set and each structure in the testing set. The transformation of $\mathbf{X} \rightarrow \mathbf{K}$ can be achieved by one of several kernel functions. For all kernel methods (except for SOAP), we use a Laplacian kernel function, k, given by

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma \left\| \mathbf{x}_i - \mathbf{x}_j \right\|_1\right) \tag{F.3}$$

where γ serves as an adjustable KRR model hyperparameter. In the case of SOAP, a similarity kernel **K** is directly generated and so there is no need for further transformation.

In all cases throughout his work, $\mathbf{y}_{\text{train}}$ refers to the DFT-computed band gaps of the DFT-optimized structures, whereas **X** refers to the encodings of the corresponding unrelaxed crystal structures.

F.4.5 Featurization Methods for KRR

As mentioned in the main text, several featurization methods were pursued for generating the feature matrices **X** for use with KRR, which we summarize in this section. For all non-SOAP featurization methods, a min-max scaler was applied during the KRR process, such that each feature was scaled to the range 0 – 1.

F.4.5.1 Description: Stoichiometric-120 Features

The Meredig and Agrawal et al.[436] feature set ("Stoichiometric-120") is a composition-based descriptor that was originally developed for formation energy predictions of inorganic solids in the Open Quantum Materials Database[30,34] (OQMD). In this work, the descriptor set has 120 attributes. 103 of these encode the elemental composition via the fraction of each unique element from H–Lr in the MOF. The remaining attributes are the mean atomic weight, mean group number, mean period number, maximum difference in atomic number, mean atomic number, range in atomic radii, mean atomic radius, range in electronegativities, mean electronegativity, the average number of s, p, d, and f valence electrons, and the composition-weighted fraction of s, p, d, and f valence electrons.

F.4.5.2 Description: Stoichiometric-45 Features

The He et al.[400] feature set ("Stoichiometric-45") is a composition-based descriptor that has been used to classify if inorganic solids in the OQMD are metallic or non-metallic. The descriptor set has 45 attributes. These consist of 9 elemental properties (atomic number, group number, period number, electronegativity, electron affinity, melting temperature, boiling temperature, density, and ionization energy) and five statistical quantities of each (arithmetic mean, geometric mean, standard deviation, maximum, and minimum) computed for each structure. The tabulated data was taken from the Wolfram Knowledgebase, which we accessed with Mathematica 11.3.0. It is important to note that several of these attributes (e.g. melting and boiling temperatures, density) are ill-defined for single atoms. In the Wolfram Knowledgebase, these values are generally defined as being for stable bulk forms at ambient conditions. Furthermore, the electron affinities and ionization energies were chosen to be for the addition or removal of a single electron, respectively. We also chose to place the lanthanides and actinides in a fictitious group 19. We note that 477 out of the 14482 MOF structures in the QMOF-14482 dataset did not have fingerprints generated due to missing tabulated data for one or more of the elements in the structure and were therefore not considered with this featurization method.

F.4.5.3 Description: Sine Coulomb Matrix Eigenspectrum

The sine Coulomb matrix[437] is a structure-based featurization method where a pair-wise interaction matrix M_{ij} is generated by the following formula:

$$M_{ij} = \begin{cases} 0.5Z_i^{2.4}, & i = j \\ \dfrac{Z_i Z_j}{\left|\mathbf{B} \cdot \sum_{k=\{x,y,z\}} \hat{\mathbf{e}}_k \sin^2\left(\pi \mathbf{B}^{-1} \cdot \left(\mathbf{R}_i - \mathbf{R}_j\right)\right)\right|}, & i \neq j \end{cases} \quad \text{(F. 4)}$$

where i and j are two atoms in the structure, Z_i is the atomic number of i, $\mathbf{B}$ is a matrix formed by the lattice vectors, $\hat{\mathbf{e}}_k$ are the Cartesian unit vectors, and $\mathbf{R}_i - \mathbf{R}_j$ is the distance vector between atoms i and j. The sine Coulomb matrix is dependent on the number of atoms a given structure has, so to ensure a square matrix is generated, it is padded with zeros to match the maximum number of atoms in the dataset (i.e. 300 atoms in the case of the QMOF-14482 dataset). Since a feature vector for each material is needed for KRR, only the (sorted) eigenvalues of the sine Coulomb matrix are returned such that the descriptor becomes one-dimensional for each structure with a length of $n_{\text{max atoms}}$ (i.e. 300 in this case). This approach was chosen instead of flattening the sine Coulomb matrix because the resulting feature length would otherwise be extremely large, as the sine Coulomb matrix for each material has dimensions $n_{\text{max atoms}} \times n_{\text{max atoms}}$ (i.e. 90,000 total entries upon flattening).

F.4.5.4 Description: Orbital Field Matrix

The orbital field matrix[438] encodes each atom in a structure by a constant-length vector representing the valence subshells of the atomic environments in each structure. To do so, each atom in a structure is represented via its (neutral) electron configuration. This electron configuration is turned into a numerical vector via a one-hot encoding scheme using a dictionary composed of the possible valence subshell orbitals and their occupancies (i.e. s^1, s^2, p^1, p^2, …, p^6, d^1, d^2, …, d^{10}, f^1, f^2, …, f^{14}). This is a 32-entry one-hot encoding. As implemented in matminer,[607] we supplemented this 32-entry encoding with 7 extra entries that represent the one-hot encodings of the period number for the element (with lanthanides in period 6 and actinides in period 7). These atomic one-hot encoding vectors are then used to construct one-hot encoding vectors for each atomic local environment (i.e. an atom center and its coordinating atoms). This is

achieved by defining a 39×39 matrix obtained by multiplying the one-hot encoding vector of the central atom and a given coordinating atom. The orbital field matrix for an atomic environment is then the sum of these matrices between the center atom and each of its coordinating atoms, scaled by a distance function. This distance function is the inverse of the bond distance multiplied by a weighting factor (using the solid angle determined by the Voronoi polyhedra between the center atom and each neighbor). Each atomic environment orbital field matrix is converted into a structural orbital field matrix by averaging the atomic environment matrices across every atomic site such that each structure is described by an averaged 39×39 matrix, which is flattened to a 1521-length encoding. Additional details can be found in the original work by Pham and coworkers.[438]

F.4.5.5 Description: Average SOAP Kernel

SOAP is a featurization method that encodes information about local atomic environments in a structure, which can then be used with an appropriate kernel function to measure the structural similarity between every pair of structures in a given dataset. For full details regarding SOAP, we refer the reader to the original paper on the use of SOAP for structure comparison[440] and the brief summary and implementation of SOAP in the original DScribe paper,[606] which we summarize below. We note that we have adopted much of the nomenclature from Musil and coworkers[614] to clarify the description of the SOAP kernel.

We start by representing a given structure using local atomic densities ρ, separately defined for each atomic element Z. The local density of atoms within a chemical environment χ_i (i.e. a spherical region centered around atom i) is described as a sum of Gaussians placed at the central atom and the neighboring atoms within a cutoff region r_{cut}. Mathematically, this is expressed as

$$\rho_{\chi_i}^{Z}(\mathbf{r}) = \sum_{k} \exp\left(-\frac{|\mathbf{r} - \mathbf{R}_k|^2}{2\sigma^2}\right) \tag{F.5}$$

where σ is the standard deviation of the Gaussians, and $|\mathbf{r} - \mathbf{R}_k|$ describes the distance between atom k, $\mathbf{R}_k$, and position vector $\mathbf{r}$. The origin, $\boldsymbol{r} = \mathbf{0}$, is centered on the local point of interest (i.e. atom i). The summation is carried out for all atoms k with atomic number Z in the structure that are within radius r_{cut} from atom i.

Given local atomic environments i and j in two structures A and B, one can then compute $\rho_{\chi_i^A}^{Z}$ and $\rho_{\chi_j^B}^{Z}$. The structural similarity between two chemical environments in structures A and B, denoted χ_i^A and χ_j^B, is

$$\tilde{k}(\chi_i^A, \chi_j^B) = \int_{\text{SO(3)}} \left| \sum_{Z} \int_{\mathbb{R}^3} \rho_{\chi_i^A}^{Z}(\mathbf{r}) \rho_{\chi_j^B}^{Z}(\mathbf{r}) \mathrm{d}\mathbf{r} \right|^2 \mathrm{d}\hat{R} \tag{F.6}$$

where SO(3) and $\hat{R}$ refer to the group of all three-dimensional rotations. The above expression is necessary to achieve a rotationally invariant descriptor and describes the (squared) overlap of the density fields, integrated over all three dimensional rotations. In practice, the calculation of $\tilde{k}$ is carried out by expanding $\rho(\mathbf{r})$ using n_{max} real spherical harmonic and ℓ_{max} radial basis functions.[440,606] We use spherical Gaussian type orbitals (GTOs) for the radial basis function in this work. The expression for $\tilde{k}$ can be normalized via

$$k(\chi_i^A, \chi_j^B) = \frac{\tilde{k}(\chi_i^A, \chi_j^B)}{\sqrt{\tilde{k}(\chi_i^A, \chi_i^A)\tilde{k}(\chi_j^B, \chi_j^B)}} \tag{F.7}$$

such that the self-similarity of a given environment, $k(\chi_i^A, \chi_i^A)$ or $k(\chi_j^B, \chi_j^B)$, is equal to 1.

The similarity between all local atomic environments i in structure A and all local environments j in structure B is then given by the general expression

$$C_{ij}(A,B) = k(\chi_i^A, \chi_j^B) \tag{F.8}$$

For a pair of structures A and B, we can then compute an average kernel to go from the similarity of local environments to the similarity of global structures. This average kernel function is defined as

$$K(A,B) = \left(\frac{1}{n_A n_B} \sum_{ij} C_{ij}(A,B) \right)^{\xi} \tag{F.9}$$

where n_A and n_B are the number of atoms in structure A and B, respectively. Taking the summation over all sites i and j between the pairs of structures and then dividing by the number of atoms in both structures converts these otherwise local similarity scores into a global structural descriptor comparing structures A and B. The variable ξ is an optional model hyperparameter to modify the spread of entries in the kernel matrix, which we include as a tunable parameter during the KRR grid search. This expression for **K** can be readily extended for all relevant pairs of structures, which can then be used directly with KRR. An example of the average SOAP similarity kernel for IRMOF-1, IRMOF-2, and ZIF-8 is shown in Figure F.2 for reference. Note that **K** is normalized such that self-similarity is unity (i.e. $K(A,A) = K(B,B) = 1$).

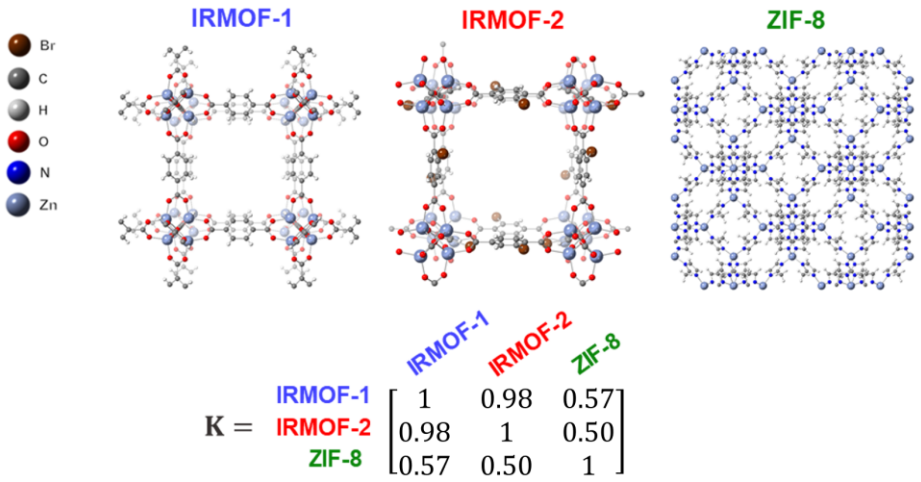


Figure F.2. Average (normalized) SOAP similarity kernel for IRMOF-1, IRMOF-2, and ZIF-8. Here, $r_{\text{cut}} = 4$ Å, $\sigma = 0.1$ Å, $\xi = 2$, and $n_{\text{max}} = \ell_{\text{max}} = 9$.

F.4.6 Hyperparameter Tuning for KRR

For featurization methods other than SOAP, a feature matrix **X** is generated. As such, a decision must be made for the type of kernel function that should be used. The hyperparameters were identified based on a grid search via 5-fold cross-validation. Initially, we considered linear, Gaussian, and Laplacian kernel functions, of which we eventually decided to use a Laplacian kernel function (Equation F.3) where $\gamma = 0.1$ since this consistently yielded the lowest cross-validation mean absolute error (MAE). A value of $\alpha = 0.1$ for the KRR regularization hyperparameter (Equation F.1) was chosen for all KRR models except for the sine Coulomb matrix, for which we use $\alpha = 0.01$.

For the SOAP-based KRR model, there are two KRR parameters to tune: α and ξ. These hyperparameters were also optimized using a grid search via 5-fold cross-validation, for which we decided upon $\alpha = 0.001$ and $\xi = 2$. The SOAP descriptor itself also has several hyperparameters that can be tuned, including but not limited to the distance cutoff for determining local regions within a structure (r_{cut}), the maximum number of radial basis functions (n_{max}), the maximum

number of spherical harmonics ($\ell_{\max}$), and the standard deviation of the Gaussians used to expand the atomic density (σ). Although there are too many parameters to easily carry out an exhaustive grid search, each parameter was independently adjusted, and the parameters that reduced the (average) MAE over the 5-fold cross-validation process were retained. This led to $r_{\text{cut}} = 4$ Å, $\sigma = 0.1$ Å, and $n_{\max} = \ell_{\max} = 9$. All other SOAP hyperparameters and settings were set to the default values in DScribe.

F.4.7 Crystal Graph Convolutional Neural Networks

CGCNN featurizes each crystal structure as an approximate crystal graph, defined such that the nodes are atoms and the edges are the atom connections, accounting for periodic boundary conditions. The crystal graphs are constructed by searching for a maximum set of neighbors within some user-defined cutoff distance. These crystal graphs are then fed as input to a convolutional neural network, wherein convolution and pooling layers convert the crystal graph to a given output, with the weights of the neural network updated to minimize the validation loss. Further details can be found in the original CGCNN paper.[406]

Iterative testing of the various CGCNN hyperparameters led to the following high-performing convolutional neural network configuration with regards to a reduced validation MAE: 5 convolutional layers, 64 hidden atom features in the convolutional layers, 1 fully connected hidden layer after pooling, and 128 hidden features after pooling. A batch size of 16, initial learning rate of 0.01, and stochastic gradient descent optimizer were used. All other settings were the default values, including a neighbor search radius of 8 Å and a maximum of 12 neighbors connected to every node in the graph. The best model obtained within 400 epochs (in terms of validation MAE) was retained. We note that several variations on the original CGCNN algorithm, such as CGCNN

with a tanh activation function[615] and iCGCNN,[616] did not show substantial improvements over the original CGCNN implementation, although a detailed exploration of the hyperparameter space was not carried out.

In the original CGCNN work,[406] the (initial) CGCNN node (i.e. atom) feature vectors were based on one-hot encodings of group number, period number, electronegativity, covalent radius, number of valence electrons, first ionization energy, electron affinity, block, and atomic volume. However, the currently published version of the code contains several inconsistencies in the one-hot encodings compared to that reported in the original text.[617] As such, we regenerated the atom initialization file and made several minor modifications to the initialization process, wherein we: 1) used Pauling electronegativities instead of Sanderson electronegativities; 2) defined the lanthanides and actinides as period 6 and 7 rather than 8 and 9; 3) placed the lanthanides and actinides in a fictitious group 19; 4) used van der Waals radius instead of the covalent radius defined by Cordero and coworkers[618]; 5) removed the atomic volume feature; 6) removed the electron affinity feature. Tabulated values were taken from mendeleev v.0.5.2.[619] Functionally, we found that this process has no apparent change in the performance of the CGCNN models developed in this work, likely because the node vectors are iteratively optimized during the model training process. Nonetheless, the changes were retained. The edge (i.e. bond) feature vectors contain the bond distance between nodes, as in the original CGCNN work.[406]

F.4.8 Dimensionality Reduction

Dimensionality reduction was carried out via the uniform manifold approximation and projection (UMAP) algorithm[447] as implemented in umap v.0.4.6.[448] UMAP constructs a weighted graph of a given dataset in the high-dimensional space and then projects this graph to a lower-dimensional

(in this case, two-dimensional) space. Each node of the graph represents a data point, with the edges representing the proximity of each pair of data points in the feature space.[447] The number of neighbors was set to 15 (for SOAP) or 50 (for Stoichiometric-120), and the minimum distance between points was set to 0.1 (for SOAP) and 0.4 (for Stoichiometric-120). All other parameters were set to the default values (for reproducibility, a random seed of 42 was used). For the connectivity map, edge bundling[620] was enabled to help convey the overall structure by allowing edges to curve and then grouping nearby connections.

The SOAP similarity kernel was converted to a distance matrix **D** by invoking the following metric:

$$D_{ij} = \sqrt{K_{ii} + K_{jj} - 2K_{ij}} \therefore \mathbf{D} = \sqrt{2 - 2\mathbf{K}} \qquad \text{(F. 10)}$$

since the self-similarity scores K_{ii} and K_{jj} are normalized to 1. A Euclidean distance metric was used with the Stoichiometric-120 descriptor to create the distance matrix. Although maximum atomic number, $\max(Z)$, is not a feature in Stoichiometric-120, it can be directly related to the $\text{range}(Z)$ feature since $\min(Z) = 1$ in every MOF. For this reason, we use the more intuitive $\max(Z)$ feature in Figure 7.6a.

F.4.9 Methodological Comments for Data Reuse

One of the main motivations for developing the QMOF Database is to enable the development/evaluation of new machine learning models. In this case, if the goal is to develop a new machine learning algorithm specifically tailored for MOFs, there are a few comments worth considering.

First, if the purpose is to predict the properties of the optimized MOF structures from the un-optimized structures, then it is important to note that some MOFs have small structural changes before and after optimization whereas others may have somewhat significant changes in the lattice constants (e.g. due to solvent removal). This may (or may not) influence the machine learning process, depending on the property of interest and the way in which the MOFs are encoded. For computational simplicity, we have also assumed a high-spin initial guess for the magnetic moments in a similar manner as the current iterations of the OQMD[30] and Materials Project.[29,33] While this initial guess may converge to a spin state quite different from the high-spin initialization, it is inevitable that several open-shell MOFs in the QMOF Database are not at their true magnetic ground states. This is especially the case for MOFs with the possibility of antiferromagnetic ordering.

Another aspect to consider is that, while every effort was made to ensure the initial structures were charge-neutral and accurately constructed, it is inevitable that some structures in the database are not pristine. Oftentimes, this can occur for reasons completely outside the control of the workflow shown in Figure F.1, such as if some atoms could not be identified experimentally and were never included in the CIF or CSD entry. The most common scenario is likely omitted/additional H atoms, which are particularly difficult to identify if not already specified in the CSD entry. While many of these instances are filtered out either via the workflow in Figure F.1 or due to SCF convergence issues when the DFT calculations were performed, additional filtering steps are always possible and are suggested for new applications of interest. Users are encouraged to flag any identified "structural fidelity" issues on the QMOF Database GitHub page.[402]

We also note that the definition of a MOF employed by the CSD MOF subset[44] is significantly less strict than that of the CoRE MOF database.[14] Namely, materials in the CSD MOF subset do not need to be porous, and there is no restriction on the dimensionality of the framework itself. This is arguably ideal from a data inclusivity and model generalizability standpoint but will result in a subset of materials that are perhaps better classified as coordination polymers rather than "conventional" MOFs. For training machine learning models to predict electronic structure properties, it is not expected that the presence or lack of pore space would directly influence the properties of interest. Nonetheless, users only interested in more "conventional" MOF structures may wish to filter the QMOF Database by pore size and/or framework dimensionality.

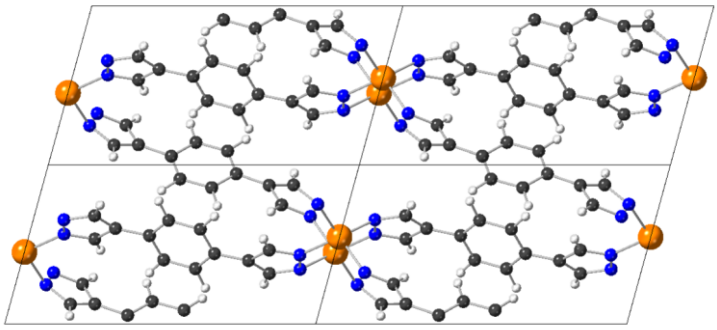
Finally, as mentioned in the text surrounding Figure 7.2a, some types of MOFs are likely to be underrepresented in the current version of the QMOF Database. The most apparent cases are MOFs that have undergone post-synthetic modification, MOFs with defects, and MOFs with metal-oxo clusters containing complex proton topologies. Unsupervised learning methods like those presented in this work can be used to determine if new MOFs of interest overlap in feature space with the MOFs in the QMOF Database. For instances where there is not significant overlap, we encourage users to supplement the QMOF Database with their own structures of interest. For particularly problematic structures in experimental MOF databases and/or those with elements that appear relatively infrequently in the database (e.g. Zr, Hf, and Al MOFs), we encourage the use of hypothetical MOF construction codes (e.g. ToBaCCo[90,621]) to generate "clean" starting structures suitable for DFT screening.

F.5 Supplemental Figures and Tables

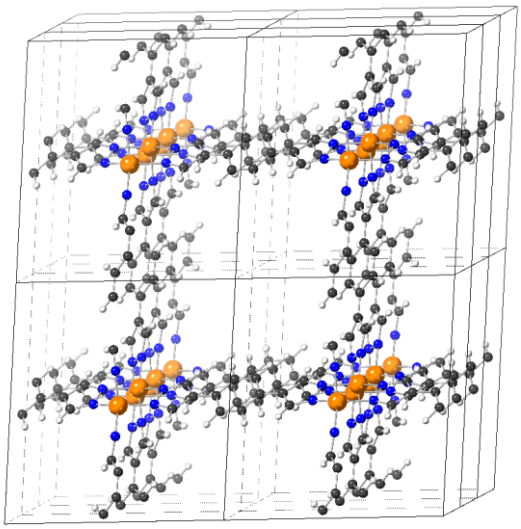
F.5.1 Example Flexible MOF

The experimental and theoretical lattice constants for an example flexible MOF in the QMOF Database, Fe(bdp) (H_2bdp = 1,4-benzenedipyrazole)[431] (Figure F.3), are shown in Table F.5.

QUPZIM



QUPZIM01



QUPZIM02

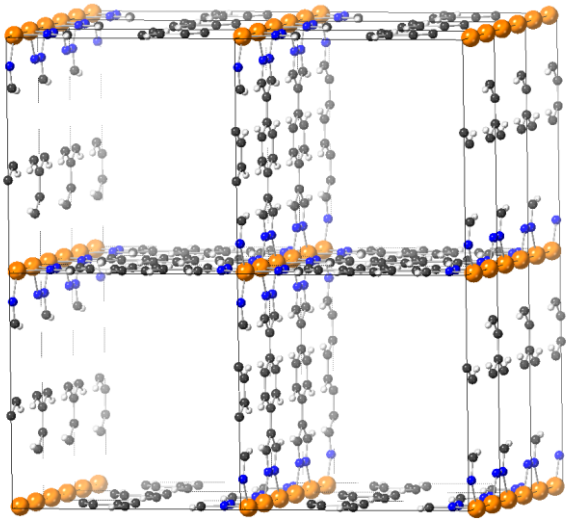


Figure F.3. DFT-optimized structures of three different conformations of Fe(bdp) in the QMOF Database. Color key: Fe (orange), N (blue), C (gray), H (white).

Table F.5. Experimental lattice constants[431] for the flexible MOF Fe(bdp) compared to the PBE-D3(BJ) lattice constants from the QMOF Database.

CSD Refcode		a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)
QUPZIM	Exp.	6.89	6.98	13.00	91.6	105.4	90.0
	Theory	6.89	6.55	13.03	91.8	105.3	90.0
QUPZIM01	Exp.	6.95	13.48	13.48	83.2	84.5	84.5
	Theory	5.60	13.68	13.68	87.6	72.3	72.3
QUPZIM02	Exp.	7.20	13.41	13.41	90.0	90.0	90.0
	Theory	7.06	13.47	13.47	90.0	90.0	90.0

F.5.2 Dataset Overview

The average band gaps associated with each element are shown in Figure F.4. The partial charges before and after optimization are shown in Figure F.5A, indicating that the values do not substantially change upon structure relaxation. We also plot the cumulative frequency of absolute deviations in partial chares before and after structure relaxation in Figure F.5B.

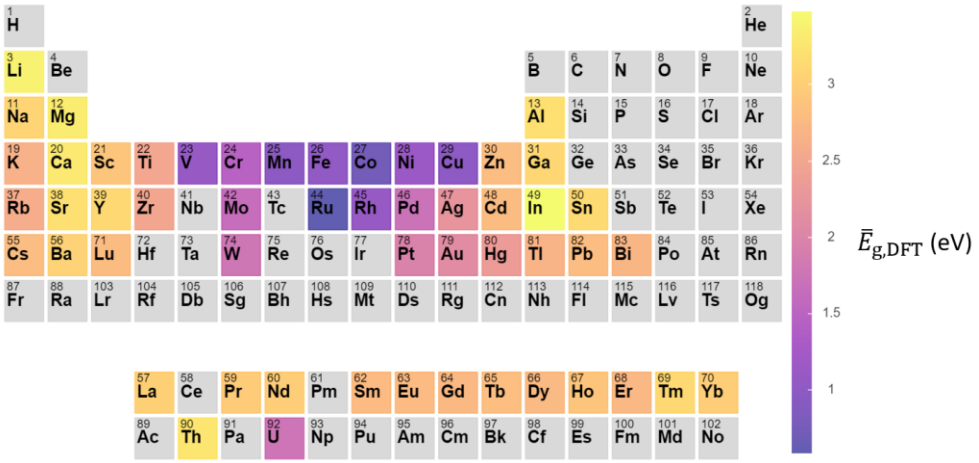


Figure F.4. Average DFT-computed band gap at the PBE-D3(BJ) level of theory, $\bar{E}_{g,DFT}$, for MOFs containing a given metal element in the QMOF-14482-opt set. If multiple metal elements are present in a given MOF, the band gap is considered for both elements. Metals with less than 10 entries were excluded.

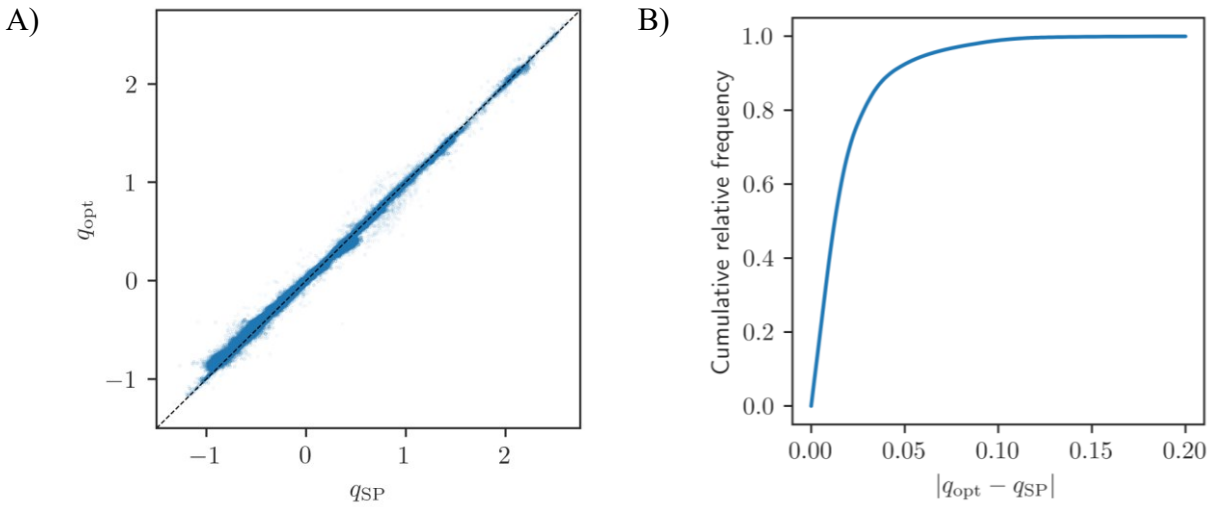


Figure F.5. a) Parity plot comparing the DDEC6 partial atomic charges for the QMOF-14170-opt (q_{opt}) and QMOF-14170-SP (q_{SP}) datasets. b) The cumulative fraction of DDEC6 partial atomic charges in the QMOF-14170-opt dataset that are within some tolerance, given by $\left|q_{\mathrm{opt}} - q_{\mathrm{SP}}\right|$, of the QMOF-14170-SP dataset.

F.5.3 High-Spin Fe MOFs

MOFs in the QMOF Database that contain both high-spin iron species (defined as having a spin density with a magnitude greater than 3.5 based on a DDEC6[158] population analysis) and a pore-limiting diameter greater than 3.6 Å (the kinetic diameter of N_2) following the structure relaxation workflow are shown in Figure F.6. These MOFs include: Fe_2(dobdc) (H_4dobdc = 2,5-dihydroxybenzene-1,4-dicarboxylic acid) (refcode: COKNOH)[432] and its expanded pore analogue Fe_2(dobpdc) (H_4dobpdc = 4,4′-dihydroxy-(1,1′-biphenyl)-3,3′-dicarboxylic acid) (refcode: MALSIE),[204] Fe_2Cl_2(bbta) (H_2bbta = 1*H*,5*H*-benzo(1,2-d:4,5-d′)bistriazole) (refcode: HAYYUE)[193] and its expanded pore analogue Fe_2Cl_2(btdd) (H_2btdd = bis(1*H*-1,2,3-triazolo[4,5-*b*],[4′,5′-*i*])dibenzo[1,4]dioxin) (refcode: HAYZAL),[193] Fe(bdp) (refcode: QUPZIM01),[431] and Fe(bpz) (H_2bpz = 4,4′-bipyrazole) (refcode: ACODAA).[476] For brevity, we exclude refcodes that have identical frameworks but contain guest species in the pores or are different conformations of the same MOF.

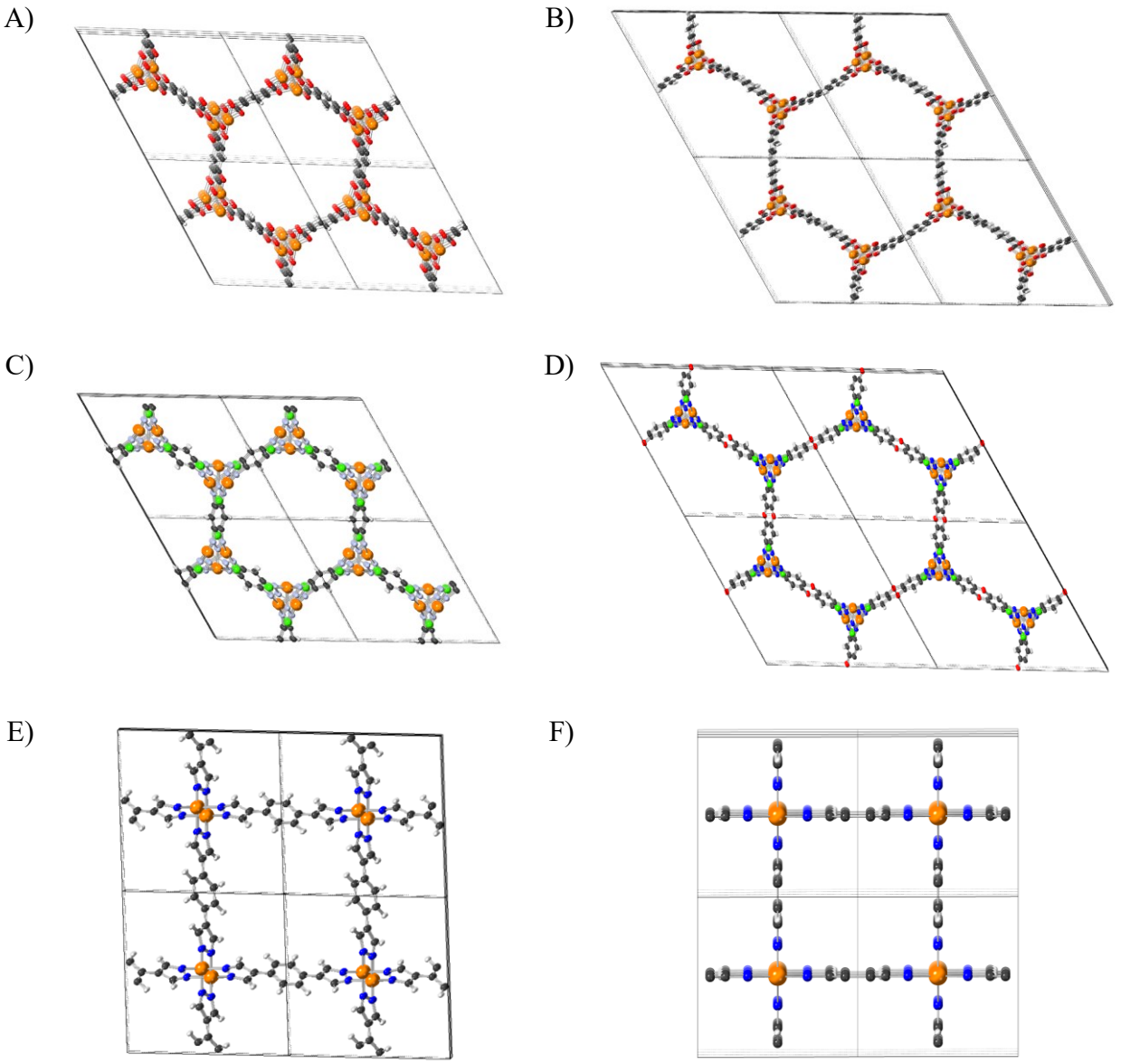


Figure F.6. Porous MOFs in the QMOF Database with high-spin Fe sites. A) Fe_2(dobdc) (refcode: COKNOH); B) Fe_2(dobpdc) (refcode: MALSIE); C) Fe_2Cl_2(bbta) (refcode: HAYYUE); D) Fe_2Cl_2(btdd) (refcode: HAYZAL) E) Fe(bdp) (refcode: QUPZIM01); F) Fe(bpz) (refcode: ACODAA). Color key: Fe (orange), N (blue), O (red), Cl (green), C (gray), H (white).

F.5.4 Comparing Machine Learning Models for Band Gap Prediction

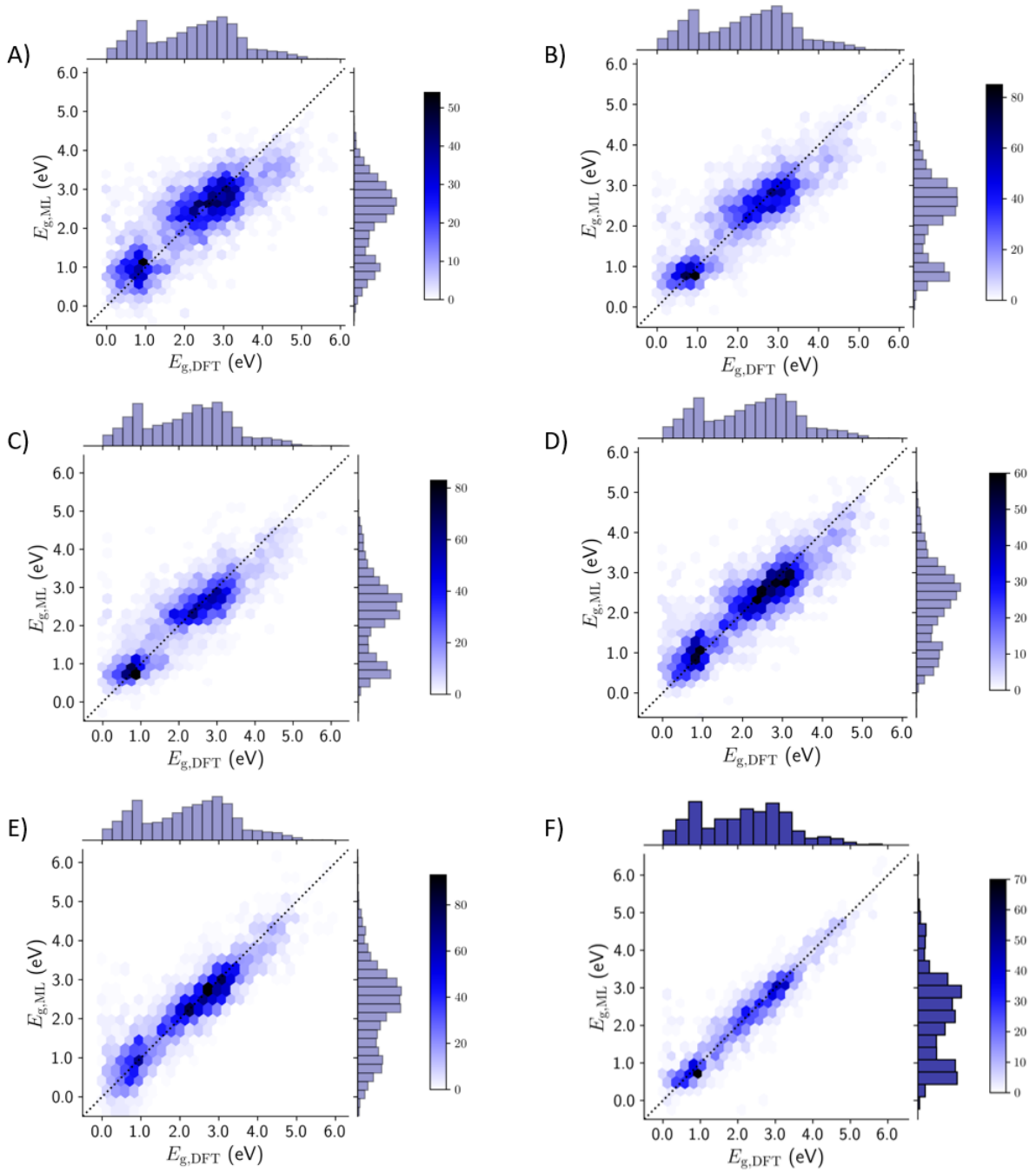


Figure F.7. Testing set parity plot for the A) Sine Coulomb matrix, B) Stoichiometric-45, C) Stoichiometric-120, D) Orbital field matrix, E) SOAP, and F) CGCNN machine learning models. The data is presented with hexagonal binning, comparing the machine learning band gaps, $E_{g,ML}$, to the DFT-computed band gaps, $E_{g,DFT}$. The color bar indicates the number of MOFs in each bin,

and the line of parity is shown as a dashed line. Histograms summarizing the distribution of $E_{\mathrm{g,ML}}$ and $E_{\mathrm{g,DFT}}$ data are displayed parallel to the y- and x-axes, respectively.

F.5.5 Comparing Against ML Band Gap Models for Other Crystalline Materials

It is worth comparing the results of the top-performing ML models in this work against state-of-the-art ML models developed for the band gaps of other crystalline materials in the literature. In the original CGCNN work, the convolutional neural network was able to achieve a testing MAE of 0.39 eV when trained on 16,458 inorganic solids from the OQMD.[406] A different graph network approach – the MatErials Graph Network (MEGNet) – achieved an MAE of 0.38 eV when trained on 36,720 inorganic solids from the Materials Project, which could be reduced to 0.33 eV after transfer learning via a model originally trained on the DFT-computed formation energies of 60,000 inorganic solids.[442] Recently, a global attention graph neural network (GATGNN) achieved MAEs of 0.32 eV and 0.31 eV on the OQMD and Materials Project datasets, respectively.[622] Particularly relevant for the present study, Olsthoorn et al.[443] used a weighted average of SOAP- and SchNet[623]-based ML regression models trained on 10,000 band gaps of organic crystals in the Organic Materials Database (OMDB)[37] to achieve a testing MAE of 0.39 eV. The band gaps in the OMDB work were based on single-point calculations of the as-deposited crystal structures, as geometry optimizations were not carried out. In addition, organic crystals with non-zero net magnetic moments were not considered in the OMDB work.

F.5.6 Additional UMAP Results

It is also worth investigating the degree of overlap in feature space between the QMOF-14482 dataset and the parent QMOF-42349 dataset that the former was drawn from. To carry out this analysis, we used UMAP to project the feature space of the QMOF-42349 dataset to two dimensions. Again, we used the Stoichiometric-120 descriptor to featurize each material. We then

highlighted the subset of materials that are also present in the QMOF-14482 dataset. As shown in Figure F.8, there is significant overlap between the two datasets, such that we can expect ML models trained on the QMOF-14482 dataset to be applicable to other MOFs deposited in the CSD. We carried out a similar analysis to compare the QMOF-14482 dataset with the CoRE MOF 2019 (v.1.1.3) database.[14] For this purpose, we use the CoRE MOF 2019 database with free solvent removed (i.e. "FSR"), including both the "public" and "internal" subsets but excluding structures flagged as having disorder that could not be refined. Significant overlap in the reduced feature space was observed between both databases (Figure F.9).

Figures F.10 and S11 show Stoichiometric-120- and SOAP-based UMAPs generated for the QMOF-14482 dataset, respectively, but with edge connections shown to highlight the connectivity between different local regions in the projection. This is particularly notable for the Stoichiometric-120 UMAP, which shows that the individual clusters are connected in sequential order of $\max(Z)$ (Figure 7.6a, Figure F.10). The edge connections for the SOAP-based UMAP also makes the local regions, and their connectivity, clearer.

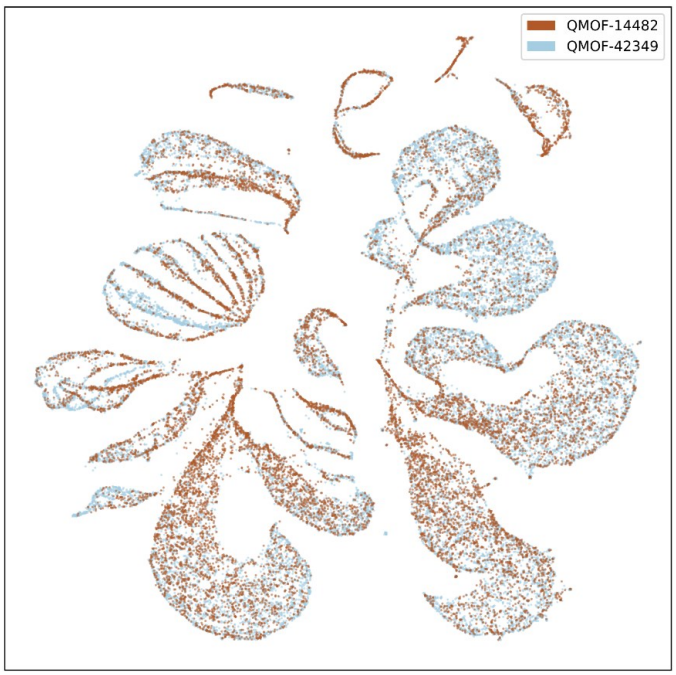


Figure F.8. Structural dimensionality reduction performed using UMAP, with a distance matrix obtained from the Euclidean distance of the Stoichiometric-120 encodings for the structures in the QMOF-42349 dataset. The QMOF-14482 subset is overlaid onto the projection.

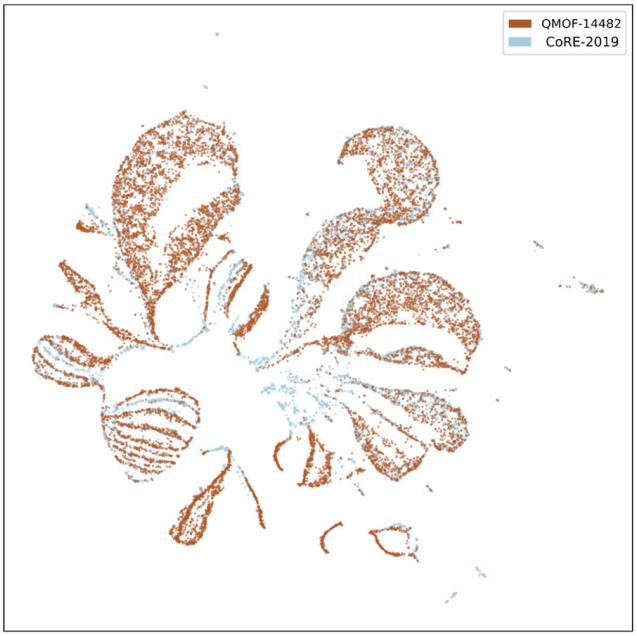


Figure F.9. Structural dimensionality reduction performed using UMAP, with a distance matrix obtained from the Euclidean distance of the Stoichiometric-120 encodings for the structures in the QMOF-14482 and CoRE MOF 2019 (free solvent removed) databases.

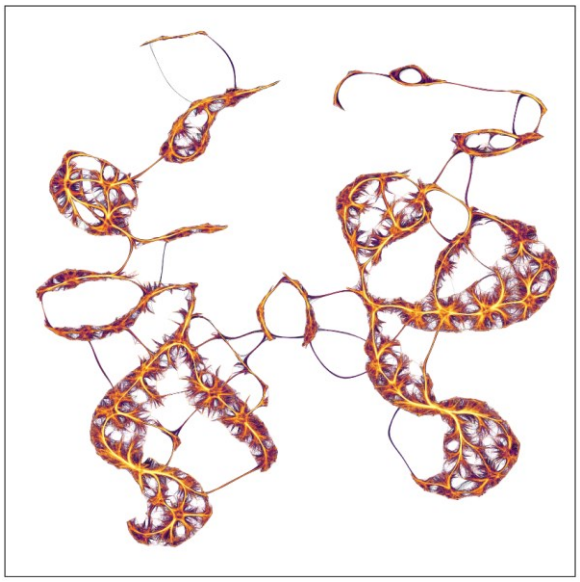


Figure F.10. Structural dimensionality reduction performed using UMAP, with a distance matrix obtained from the Euclidean distance of the Stoichiometric-120 encodings of the structures in the QMOF-14482 dataset. The connectivity between points is shown. Brighter colors indicate a greater density of connections.

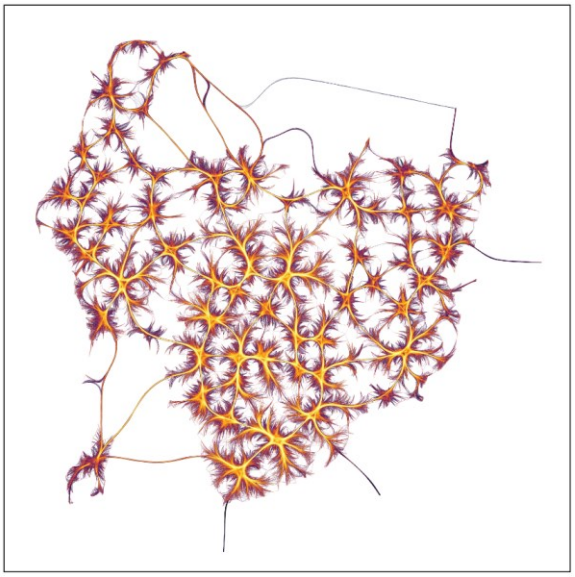


Figure F.11. Structural dimensionality reduction performed using UMAP, with a distance matrix obtained from the average SOAP similarity kernel of the (unrelaxed) structures in the QMOF-14482 dataset. The connectivity between points is shown. Brighter colors indicate a greater density of connections.

F.5.7 Electronic Structure of GUTYAW

The projected density of states at the HSE06-D3(BJ) level of theory for $Sr[C_2H_4(SO_3)_2]$ (refcode: GUTYAW[472]) is shown in Figure F.12.

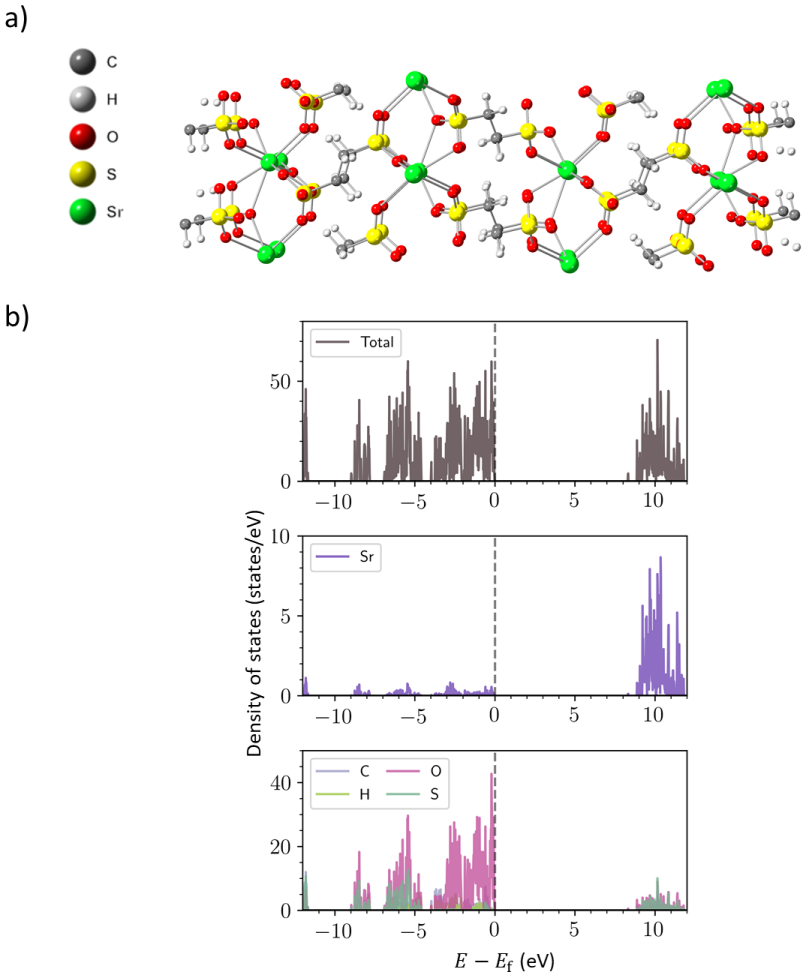


Figure F.12. a) Structure of $Sr[C_2H_4(SO_3)_2]$. b) Total and projected density of states. The energy, E, is shown with respect to the Fermi level, E_f.

F.5.8 Limitations of Averaging Schemes

For any featurization method, there are inevitable limitations with how a given set of materials is encoded for machine learning. In the case of an average SOAP kernel, for instance, one limitation is that every atomic environment is weighted equally. To highlight why this may be imperfect, we show the average SOAP similarity kernel for IRMOF-1 with the formula $Zn_4O(bdc)_3$ (bdc =

benzene-1,4-dicarboxylate), IRMOF-2 with the formula $Zn_4O(bdc\text{-}Br)_3$, IRMOF-10 with the formula $Zn_4O(bpdc)$ (bpdc = 4,4′-biphenyldicarboxylate), $Zn_2(dobdc)$ (dobdc = 2,5-dihydroxybenzene-1,4-dicarboxylate), and MFU-4*l* (MFU = Metal–Organic Framework Ulm University, *l* = large) with the formula $Zn_5Cl_4(btdd)_3$ (btdd = bis(1,2,3-triazolato-[4,5-b],[4′,5′-i])dibenzo-[1,4]-dioxin)) (Figure F.13). While IRMOF-1 and the functionalized analogue IRMOF-2 have nearly identical averaged SOAP features, IRMOF-1 and the elongated analogue IRMOF-10 are quite different (Table F.6). In fact, IRMOF-1 and $Zn_2(dobdc)$ are more similar than IRMOF-1 and IRMOF-2 based on the average SOAP kernel (Table F.6). This can likely be traced back to the similarity of the linkers (bdc vs. dobdc) in IRMOF-1 and $Zn_2(dobdc)$ despite their very different inorganic nodes (Zn_4O vs. isolated Zn sites) and metal coordination environments (tetrahedral vs. square pyramidal). Weighting the structural similarity of the inorganic nodes and organic linkers by different factors is one approach that may resolve this issue, aside from trying alternate kernel methods such as the computationally more expensive regularized entropy match (REMatch) kernel.[440] This phenomenon is also expected to limit the performance of other methods that involve simple averaging over a structure, such as the orbital field matrix. With the development of a database of DFT-computed MOF properties, there is a rich opportunity for exploring featurization methods that are constructed specifically for the robust and accurate representation of MOFs.

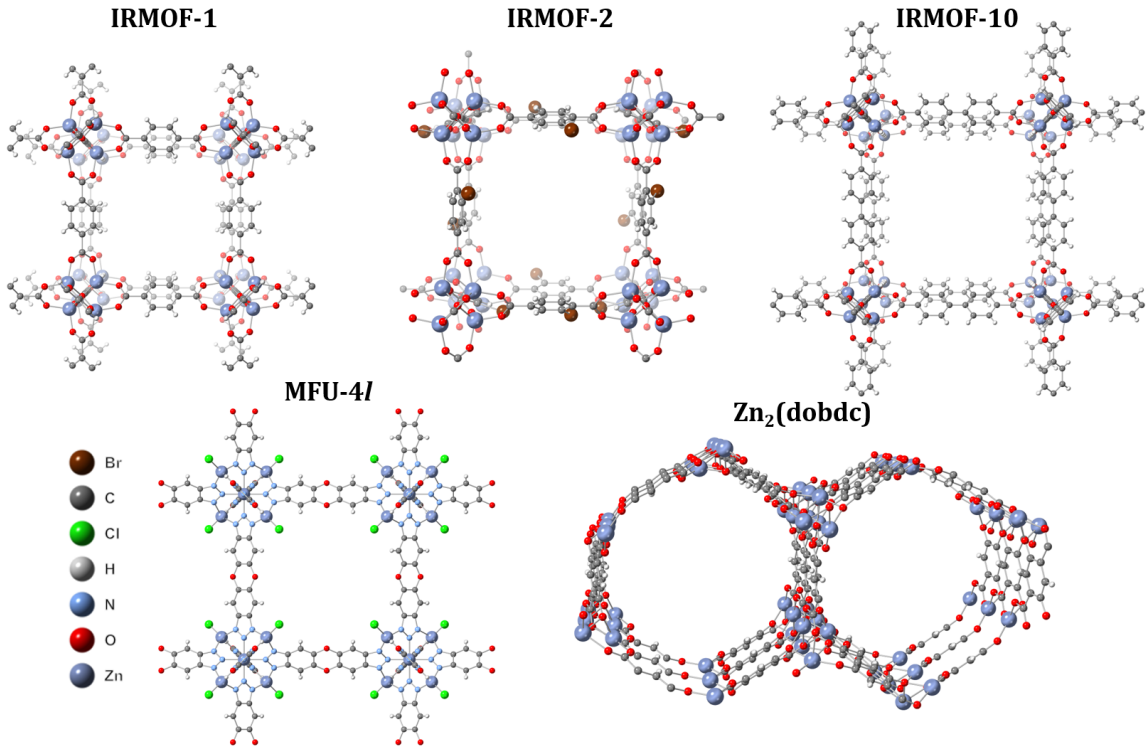


Figure F.13. Crystal structures of IRMOF-1, IRMOF-2, IRMOF-10, Zn_2(dobdc) (also known as Zn-MOF-74 and Zn-CPO-27), and MFU-4*l*.

Table F.6. Average (normalized) SOAP similarity kernel for IRMOF-1, IRMOF-2, IRMOF-10, Zn_2(dobdc), and MFU-4*l*. Here, $r_{cut} = 4$ Å, $\sigma = 0.1$ Å, $\xi = 2$, and $n_{max} = \ell_{max} = 9$.

	IRMOF-1	IRMOF-2	IRMOF-10	Zn_2(dobdc)	MFU-4*l*
IRMOF-1	1.00	0.98	0.52	0.92	0.73
IRMOF-2	0.98	1.00	0.43	0.94	0.72
IRMOF-10	0.52	0.43	1.00	0.29	0.35
Zn_2(dobdc)	0.92	0.94	0.29	1.00	0.64
MFU-4*l*	0.73	0.72	0.35	0.64	1.00

Appendix G. APPENDIX FOR CHAPTER 8

G.1 Additional Methodological Details

G.1.1 New Properties at the HLE17, HSE06*, and HSE06 Levels of Theory

Prior to this study, the QMOF Database consisted only of PBE-D3(BJ) optimized structures with properties provided at the same level of theory. In the present work, we have updated the QMOF Database with properties at the HLE17, HSE06*, and HSE06 levels of theory using the aforementioned PBE-D3(BJ) geometries. This new data includes but is not limited to: absolute energies, band gaps, total and projected density of states, DDEC6 population analysis (i.e. partial atomic charges, atomic spin densities, bond orders),[158,230–232] Charge Model 5 (CM5) charges,[289] magnetic moments, and charge densities. For these calculations, all PBE-D3(BJ) optimized structures were obtained from v4 of the QMOF Database.[75,402] The DFT settings for the HLE17/HSE06*/HSE06 calculations are summarized in Table G.1 and are identical to those used in the original construction of the QMOF Database,[75] with the following exceptions:

- A k-point grid density (KPPA) of 500/(number of atoms per cell) was used, as arranged using the automatic k-pointy density generator in Pymatgen.[82] This is reduced from the standard KPPA of 1000 used in the QMOF Database in order to decrease the computational cost of the meta-GGA and hybrid-level calculations.
- The preconditioned conjugate gradient algorithm[132,133,338] (`algo=all` instead of `algo=Fast`) was used to converge the self-consistent field (SCF), as this is the recommended algorithm for hybrid- and meta-GGAs.

- An SCF energy convergence criterion of 10^{-5} eV (instead of 10^{-6} eV) was chosen as a reasonable balance between runtime and precision for meta-GGA and hybrid-level calculations.
- 2000 grid-points (instead of the VASP default of 301) were used when writing out the density of states (DOS).
- A PBE wavefunction and charge density with KPPA of 500 ("PBE-lowkpt") was used as the initial guess for the HSE06 and HLE17 calculations. Note that all PBE results reported throughout this work are nonetheless presented using the higher accuracy settings with a KPPA of 1000.
- For the HSE06* calculations, the wavefunction and charge density from the HSE06 calculations were used as the initial guess.

To summarize the most important DFT-related settings in Table G.1, the plane-wave kinetic energy cutoff was set to 520 eV, a KPPA of 500 was specified using the `automatic_density()` function in Pymatgen for the HLE17/HSE06*/HSE06 calculations,[82] the "accurate" precision keyword was enabled in VASP, Gaussian smearing of the band occupancies was applied with a smearing width of 0.01 eV, and symmetry considerations were disabled.

Although it does not influence the band gaps or partial charges discussed in this work, the D3(BJ) dispersion correction[467] was enabled for the HSE06 calculations since they were also included during the PBE-D3(BJ) structure relaxations. For consistency, the same D3(BJ) parameters were used for HSE06* as HSE06. The dispersion-free electronic energies, energy contribution from dispersion corrections, and total energies are all made available in the QMOF Database.

Table G.1. ASE and VASP input arguments for the VASP calculators used in the screening workflow.[a] Note that the "PBE-lowkpt" calculations are only used to generate the wavefunction and charge density for the HLE17, HSE06*, and HSE06 calculations.

Flag	PBE	PBE-lowkpt	HLE17	HSE06*	HSE06
xc	'PBE'	'PBE'	'HLE17'	'HSE06'	'HSE06'
aexx	0	0	0	0.1	0
ivdw	12	12	0	12	12
encut	520	520	520	520	520
kppa[b]	1000	500	500	500	500
prec	'Accurate'	'Accurate'	'Accurate'	'Accurate'	'Accurate'
ismear	0	0	0	0	0
sigma	0.01	0.01	0.01	0.01	0.01
ediff	1E-6	1E-6	1E-5	1E-5	1E-5
algo	'All'	'All'	'All'	'All'	'All'
nelm	150	200	100	100	100
lreal	False	False	False	False	False
nsw	0	0	0	0	0
lorbit	11	11	11	11	11
isym	0	0	0	0	0
symprec	1E-8	1E-8	1E-8	1E-8	1E-8
nedos	301	2000	2000	2000	2000
lcharg	True	True	True	True	True
laechg	True	True	True	True	True
lwave	True	True	True	True	True
icharg	0	2	1	1	1
vdw_s8	0.7875	0.7875	N/A	2.310	2.310
vdw_a1	0.4289	0.4289	N/A	0.383	0.383
vdw_a2	4.4407	4.4407	N/A	5.685	5.685

[a]The pseudopotentials can be automatically selected in ASE with `setups={'base':'recommended','Li':'','Eu':'_3','Yb':'_3', 'W':'_sv'}`. The v.54 projector-augmented wave (PAW)[123] pseudopotentials were used throughout this work.

[b]k-point density computed with the `automatic_density()` tool in Pymatgen, which was used to set the `kpts` and `gamma` flags in ASE. The choice of whether the grid should be Γ-centered or not and how the k-points are distributed among the three lattice dimensions are also determined based on this Pymatgen utility.

Several checks and filtering steps were performed when carrying out the HLE17, HSE06*, and HSE06 calculations:

- For computational efficiency, we only considered structures with less than or equal to 150 atoms per cell. We also only considered unique structures, as determined by using Pymatgen's `StructureMatcher` utility.
- Any calculations with VASP-related errors or SCF convergence issues were discarded.
- We adopted the converged PBE-D3(BJ) magnetic moments from the QMOF Database as the initial guesses in this work. As a matter of internal consistency, we discarded any calculations where the final net magnetic moment converged to one different than the provided initial guess.
- Non-PBE calculations were not run if the MOF had a net force greater than or equal to 0.1 eV/Å on any atom at the "PBE-lowkpt" level of theory (compared to the expected maximum of ~0.03 eV/Å at the "PBE" level of theory with the denser k-point grid).
- Non-PBE calculations were not run if the band gap differed by more than ±0.1 eV between the original PBE-D3(BJ) calculations and the "PBE-lowkpt" level of theory.
- To simplify the data analysis, we only considered structures where all the investigated levels of theory were successfully completed unless otherwise stated.

The above process resulted in a dataset of 11,122 structures with computed properties at all the tested levels of theory. For this 11,122 structure dataset, the deviations in computed band gap and total energies between the "PBE" and "PBElowkpt" settings are shown in Figure G.1 and Figure G.2, respectively. The maximum net force at the "PBElowkpt" settings is shown in Figure G.3.

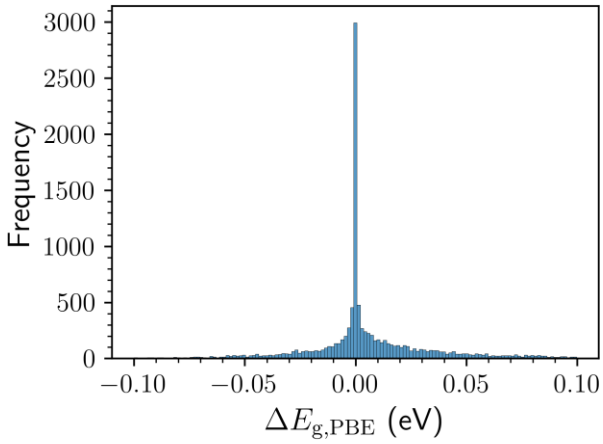


Figure G.1. Histogram showing the difference in the band gaps computed using the "PBE" and "PBElowkpt" settings, $\Delta E_{\mathrm{g,PBE}}$. A positive value indicates the "PBElowkpt" band gap is higher and vice versa.

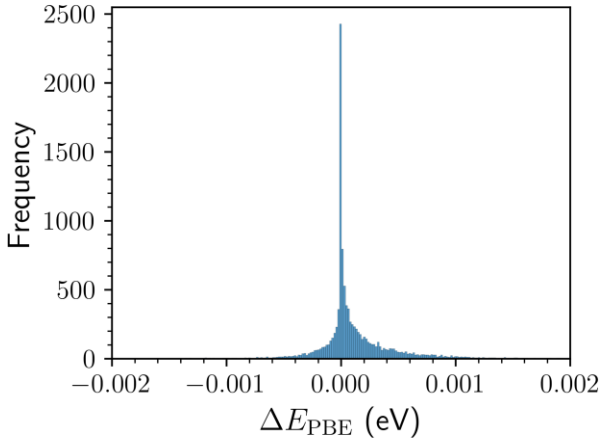


Figure G.2. Histogram showing the difference in the total energies using the "PBE" and "PBElowkpt" settings, ΔE_{PBE}. A positive value indicates the "PBElowkpt" energy is higher and vice versa.

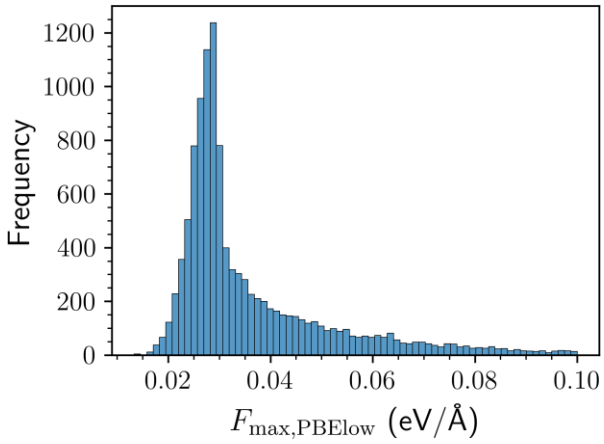


Figure G.3. Histogram showing the maximum net force, $F_{\mathrm{max,PBElow}}$, computed using the "PBElowkpt" settings.

G.1.2 Additional Methodological Details

We used Pymatgen's[82] `eigenvalue_band_properties` function to compute the band gaps from the DFT calculations. Previously, `eigenvalue_band_properties` only provided $\min(\mathrm{CBM}_{\uparrow}, \mathrm{CBM}_{\downarrow}) - \max(\mathrm{VBM}_{\uparrow}, \mathrm{VBM}_{\downarrow})$ as the band gap for spin-polarized calculations. We have since updated the function to allow for the calculation of $\mathrm{CBM}_{\uparrow} - \mathrm{VBM}_{\uparrow}$ and $\mathrm{CBM}_{\downarrow} - \mathrm{VBM}_{\downarrow}$ for spin-polarized calculations so that the band gaps of the individual spin channels are easily obtained (see Pymatgen v.2022.0.10 and newer). All values are reported in the QMOF Database.

When classifying if a MOF has closed-shell or open-shell character, we investigated the net magnetic moment and DDEC6 spin densities. If the net magnetic moment was nonzero or the material had an absolute DDEC6 spin density on any atom greater than 0.1, the material was classified as having open-shell character; otherwise, it was classified as having closed-shell character. When classifying if an atom was a (semi-)metal or nonmetal, we used the following list of nonmetals: H, He, C, N, O, F, Ne, P, S, Cl, Ar, Se, Br, Kr, I, Xe, Rn. All other elements were classified as a (semi-)metal.

While v4 of the QMOF Database was the most up-to-date version of the database when the HLE17/HSE06*/HSE06 calculations were initiated, the database has since been updated and now has more PBE-quality structures and computed properties than v4. Therefore, for any figures involving only PBE-quality partial charges meant to compare charge calculation schemes (e.g. Figure 8.5), we used all the charge data available in the most up-to-date version of the QMOF Database. Because the QMOF Database currently has more DDEC6 and CM5 charge data than Bader charge data, Figure 8.5a shows the results for 911,092 atoms (Bader vs. DDEC6), whereas Figure 8.5b (CM5 vs. DDEC) shows the results for 1,893,381 atoms. For the PACMOF[399] calculations, MOFs that could not have their partial charges predicted were removed from the dataset.

For the analysis shown in Figures 8.4d, G.7, and G.8 we searched through every atom in the MOF to identify any metal sites (i.e. those not in the nonmetal list). Pymatgen's `CrystalNN`[335,336] function was used to identify any nonmetal atoms within the first coordination sphere of each metal center. The difference in partial atomic charge on the metal sites as a function of the level of theory was then compared to the difference in partial atomic charge on the ligand atoms in the first coordination sphere.

The following software packages were used in this work: VASP 5.4.4 with the v.54 projector-augmented wave pseudopotentials (DFT calculations),[122,123] Chargemol 09-26-2017 (DDEC6 and CM5 calculations),[374] ASE 3.20.0b1 (orchestrate VASP calculations),[83] Pymatgen 2022.0.10 (electronic structure analysis),[82] Bader 1.04 (Bader analysis),[159] NumPy/Pandas/SciPy/matplotlib/seaborn (data analysis and visualization),[561,563,603–605] and PtitPrince v.0.2.5 (for raincloud plots[624]).

G.1.3 Accounting for Spin-Forbidden Transitions

An example DOS plot is shown in Figure G.4 for a magnetic material with three possible band gaps. While $E_{g,\#1}$ in Figure G.4 represents the gap in energy between the highest energy valence band and lowest energy conduction band, the VBM and CBM are of opposite spin. Using $\min(\mathrm{CBM}_\uparrow - \mathrm{VBM}_\uparrow, \mathrm{CBM}_\downarrow - \mathrm{VBM}_\downarrow)$, the band gap would be $E_{g,\#2}$ and not involve a spin transition.

Figure G.5 is an analogue of Figure 8.1, but now E_g is calculated as $\min(\mathrm{CBM}_\uparrow - \mathrm{VBM}_\uparrow, \mathrm{CBM}_\downarrow - \mathrm{VBM}_\downarrow)$ instead of $\min(\mathrm{CBM}_\uparrow, \mathrm{CBM}_\downarrow) - \max(\mathrm{VBM}_\uparrow, \mathrm{VBM}_\downarrow)$. While there are slight changes to some of the band gaps for the spin-polarized calculations, the overarching trends between functionals remain unchanged.

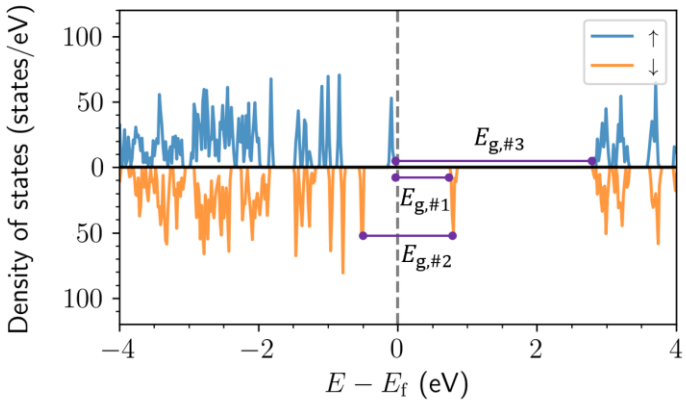


Figure G.4. Density of states (DOS) for an example structure (refcode: PORKUE) at the PBE level of theory where multiple band gaps can be defined, denoted $E_{g,\#1}$, $E_{g,\#2}$, and $E_{g,\#3}$. All energy values E are with respect to the Fermi level E_f. Color key: Spin-up (blue), Spin-down (orange).

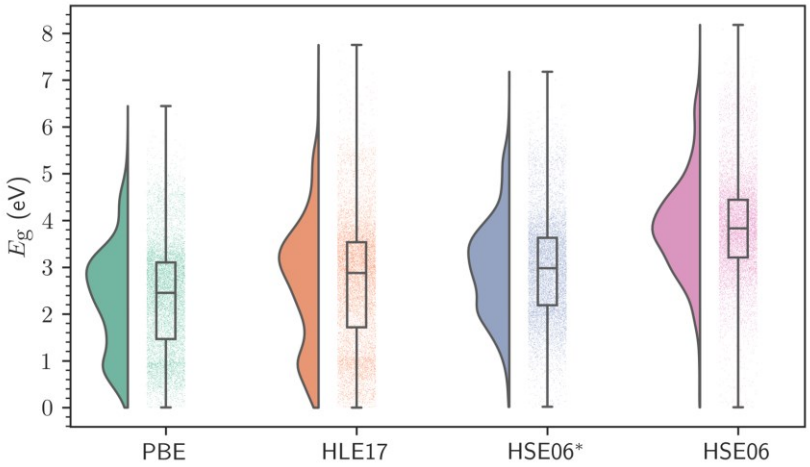


Figure G.5. Raincloud plots (i.e. combined violin plot, box plot, and strip plot) for the DFT-computed band gaps, E_g, of 11,122 structures in the QMOF Database at the PBE, HLE17, HSE06*, and HSE06 levels of theory. Here, E_g is defined as $\min(\mathrm{CBM}_\uparrow - \mathrm{VBM}_\uparrow, \mathrm{CBM}_\downarrow - \mathrm{VBM}_\downarrow)$. The strip plots show all the data at that level of theory (jittered horizontally for ease-of-visualization). The box plots show the extrema (whisker tails), interquartile range (box boundaries), and median (horizontal line). The violin plots show the probability density of the data.

G.2 Additional Results

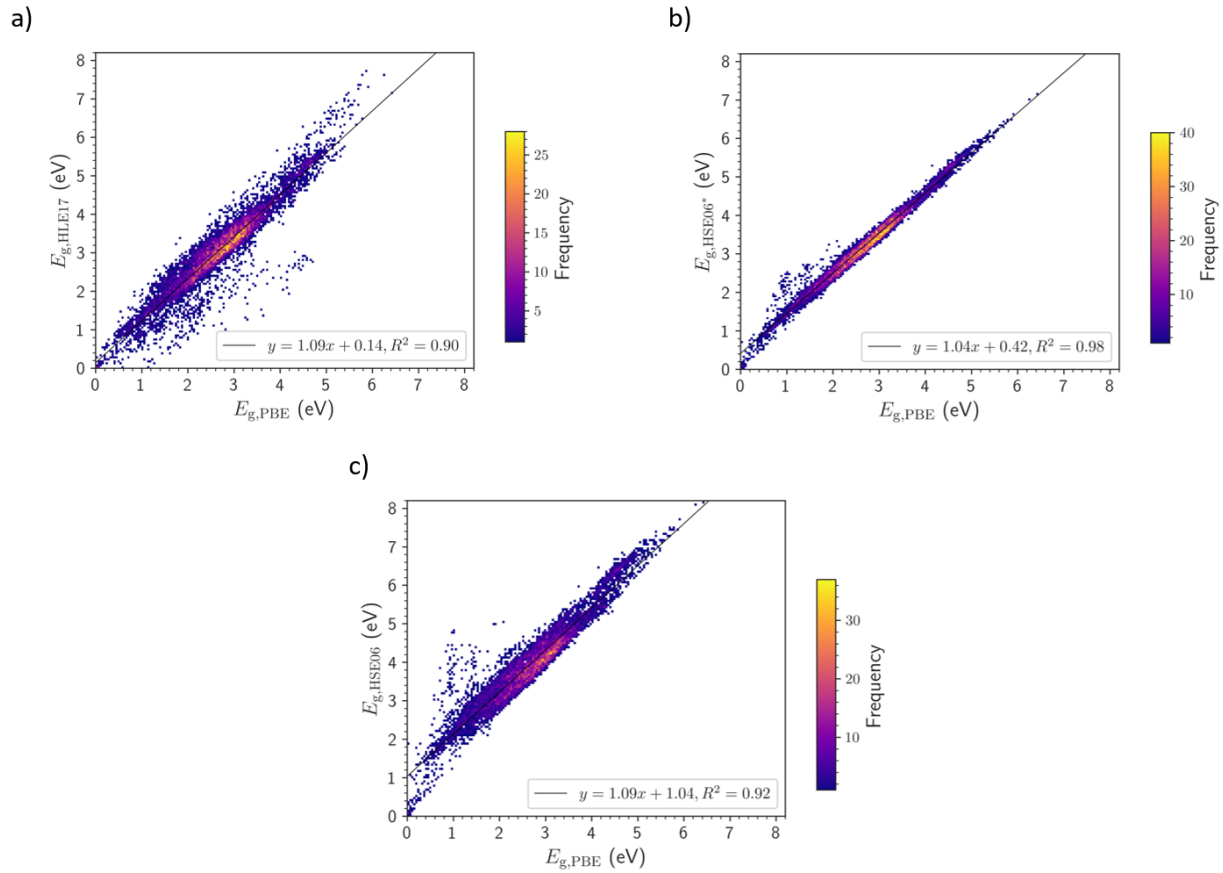


Figure G.6. Parity plots of the computed band gaps, E_g, for 8970 closed-shell structures in the QMOF Database at various levels of theory. a) HLE17 vs. PBE; b) HSE06* vs. PBE; c) HSE06 vs. PBE. Given the large dataset size, the data is plotted with hexagonal bins with the color bar reflecting the frequency of points in each bin. Histograms of the underlying band gap data are shown parallel to their corresponding axes. Best-fit lines (generated for the un-binned data) are shown as solid lines.

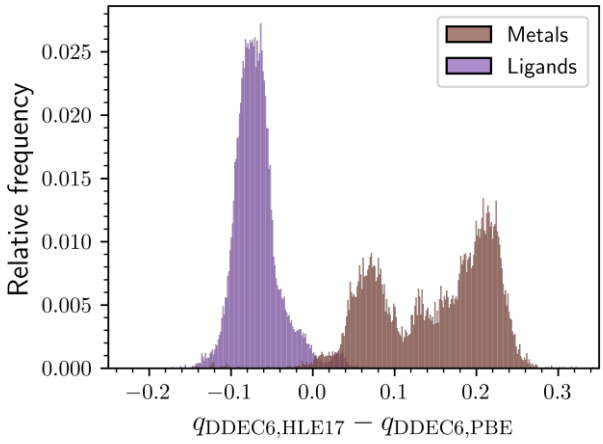


Figure G.7. A histogram of the change in DDEC6 charges between the PBE and HLE17 levels of theory for the (semi-)metals and ligand atoms within the first coordination sphere.

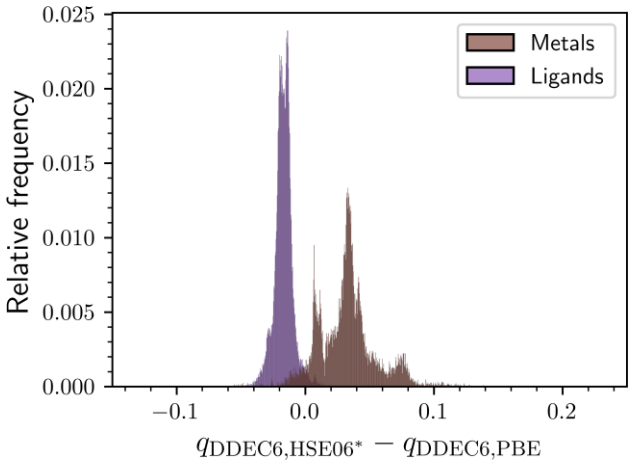


Figure G.8. A histogram of the change in DDEC6 charges between the PBE and HSE06* levels of theory for the (semi-)metals and ligand atoms within the first coordination sphere.

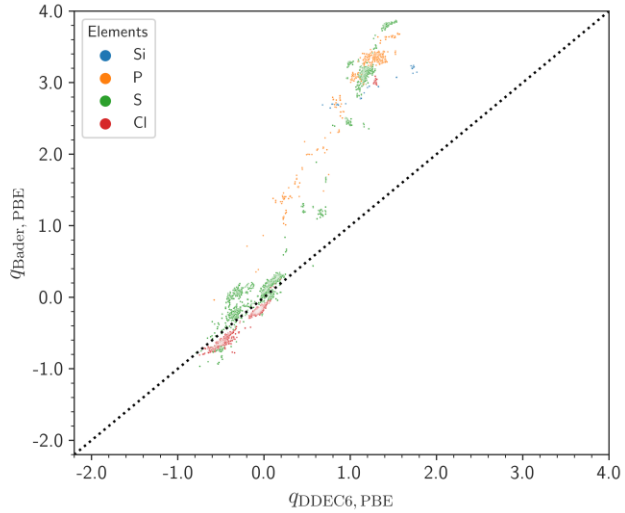


Figure G.9. Parity plot of the partial atomic charges computed using the Bader, $q_{\mathrm{Bader,PBE}}$, and DDEC6, $q_{\mathrm{DDEC6,PBE}}$, charge partitioning schemes at the PBE level of theory. Points are only included for the elements Si (blue), P (orange), S (green), and Cl (red). The line of parity is shown as a dashed diagonal line.

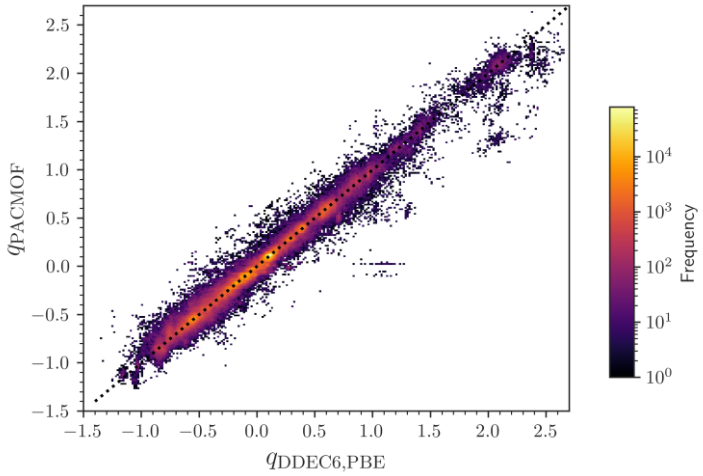


Figure G.10. Parity plot of the partial atomic charges computed from the DDEC6 method (DFT at the PBE level of theory), $q_{\mathrm{DDEC6,PBE}}$, and the pre-trained PACMOF[399] machine learning model, q_{PACMOF}. The data consists of partial charges for 1,703,975 atoms. The line of parity is shown as a dashed diagonal line. The model performs with a mean absolute error of 0.03 and an R^2 of 0.98 despite being trained on partial charges for a separate database of MOF structures.[234]

The End.