COMPUTATIONAL MODELING OF THE ACID GAS-INDUCED EVOLUTION IN METAL-ORGANIC FRAMEWORKS

by

Kai Cui

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Chemistry)

at the

UNIVERSITY OF WISCONSIN–MADISON

2022

Date of final oral examination: 06/01/2022

The dissertation is approved by the following members of the Final Oral Committee:

J. R. Schmidt, Professor, Chemistry
Arun Yethiraj, Professor, Chemistry
Dane Morgan, Professor, Materials Science and Engineering
Dawei Feng, Assistant Professor, Materials Science and Engineering

Acknowledgment

Pursuing a Ph.D. is a long journey. I appreciate the amazing people I met during the last five years. It is their guidance and help that made my life in grad school colorful. First and foremost, I must thank my advisor Professor J. R. Schmidt for his constant support. Since I first joined the group, JR is very patient throughout to answer all my questions, spanning from software debugging to scientific topics of various aspects in physical chemistry. His intelligence and broad knowledge are indispensable for me to overcome the obstacles I met during researches. This dissertation is only made possible by JR's guidance. When I decided to pursue an academia position in future, JR also gave me immense support that solidified my confidence on this route. JR is not only just a fantastic research advisor, but also a great teacher. In the five years he gives numerous advice to help me improve my writing and presentation skills. Surprisingly, JR and I share the same hobby. I really enjoyed our discussion of astrophotography, nebulae and galaxies after each research updates. In a word, I feel myself very fortunate to have JR as my advisor.

I am incredibly grateful to work with the great colleagues and friends in the Schmidt group. Though they have graduated before I joined the group, Prof. Kuang Yu, Prof. Jesse G. McDaniel and Prof. Mary Van Vleet are very accessible and helpful for answering my questions about the embedding theories and force fields. Dr. Chenyang Zhang and Dr. Nina Tyminska initiated the works that I'm focusing on. Their example files made my learning process much easier. In my first few years, Dr. Eric Hermes and Dr. Aurora Janes taught me both basic and advanced python knowledge. Dr. Tingting Weng, Dr. Xinyi Li and Tesia Janicki helped me a lot on running MD simulations using parallel codes or GPUs. Tingting also introduced to me colorful lives in Madison and led me experience the fancy restaurants and outdoor activities in town. Zhongyi Wan joined the group after me, but his intellectual ideas can always inspire me. Dr. Ajay Muralidharan has great passion on enhanced sampling methods. Although this is not my research direction, I still find the knowledge from him helpful. Due to the pandemic and building construction, I don't have

too much interaction with our newer group members Dr. Yudhajit pal, Kwanpyung Lee, Ipsita Mondall and Longyuan Zhang. But I still want to thank them for listening to my rambling.

Working in a multidisciplinary research center, I have great opportunities to work with researchers all around the nation. I especially thank my collaborators Dr. Souryadeep Bhattacharyya, Chengzhai Wang, Chunyi Li, Zhenzi Yu and Arvind Ganesan. Their hard work is crucial to accomplish our collaborative projects. The discussions with Prof. David Sholl, Prof. Sankar Nair and Prof. Krista Walton is so valuable that they always provide insights from different perspectives to perfect my works. I would also like to thank my committee Prof. Arun Yethiraj, Prof. Dane Morgan for sharing useful advices on my research projects, proposal writing and my future career plans. I am grateful that Prof. Dawei Feng agrees to join my thesis committee. I hope we can have more opportunity to collaborate in the future.

I am very fortunate to have great friends filling my life after work. Despite the title, Prof. Yang Yang is more like a good friend to me. I enjoyed our lunches together where we have broad discussion from funny videos to science and politics. Most importantly, I gratefully appreciate him for sharing his experience on postdoc searching and applying faculty positions. I must also thank Dr. Dongyue Liang, Dr. Chenyu Zhang and Dr. Zhenxuan Chen for helping me accommodate the life in a foreign country. And I will severely miss the trips to the Door county and North Michigan with Yang, Zhongyi, Dr. Yanyu Zhu, Dr. Xi Xu and Enran Xing; and the barbecue, the Frisbee and card games with Dr. Yuhui Li, Dr. Dongxue Du, Linquan Ma, Jiatong Li, Rong Mao and Yuchuan Yao. Thank you all for hanging out with me.

Last but not the least, I must thank my family for backing me up. I appreciate my parents Wenhong Yu and Xiaodong Cui for their encouragement and support on every decisions I made. Though separated 1000 miles away, the love and support from my fiancée Xi Chen never decay. Her companion is the key pulling me out from any setbacks, and powering me overcome any difficulties. I also appreciate our discussion about new papers and new knowledge during the everyday video calls. These discussions largely broadened my view and sometimes directly inspired me the solutions to my research problems.

Abstract

Characterized by high porosity, large surface area and designable structures, metal organic frameworks (MOFs) are promising candidates for gas storage and separation applications that can replace energy-intensive separation processes in current industry. At practical industrial conditions, acid gas contaminants such as SO_2, NO_2, H_2S are commonly found in process streams, but many MOFs are unstable in such acidic environments. The vulnerability of MOFs thus prohibits the industrial use. Understanding the degradation of MOFs under exposure to acid gases is essential to elongate their lifetime and enable their practical applications.

We use the zeolitic imidazolate frameworks (ZIFs) as prototypical MOF materials and perform detailed computational mechanistic studies on their degradation under acid gas exposure. We focus on three main questions: How can the reactivity of MOFs be modeled with high accuracy, beyond periodic DFT? What factors differentiate the stability of different ZIFs? How can we simulate the long-time material evolution as observed in experiments? In this thesis, we first develop QM/MM and QM/QM methods tailored for MOF systems to enable high-accuracy and high-efficiency calculations of their reactivities. "Gold-standard" accuracy can be achieved at modest computational cost. Through the combination of this QM/MM method and statistical mechanical modeling, we then successfully explain the origins of the different stability behavior between ZIF-71 and ZIF-8 under humid SO_2 exposure. Our results emphasize the importance of acid gas distribution to the overall stability. Finally, we develop a first-principles-based kinetic model to simulate the defect propagation in ZIFs and successfully reproduce the experimentally-measured macroscopic evolution of the time-dependent bulk materials proprieties. Our simulation works also provide new avenues to mitigate degradation under challenging conditions of acid gas exposure.

In the last part of this thesis, we also briefly discuss about using acid-resistant MOFs for acid gas removal. One potential design strategy to enhance acid gas adsorption in MOFs, i.e., adding Lewis base sites to the linker, is reviewed. Using first principles calculations, we closely studied the influence of these additional sites on the SO_2 adsorption behavior in the modified MOFs.

Published Work and Work in Preparation

[1] Kai Cui and J. R. Schmidt, Comment on "Solubilities of Ammonia in Basic Imidazolium Ionic Liquids" [Fluid Phase Equilib. 297 (2010) 34-39], *Fluid Phase Equilib.* **2019**, 492, 78-79.

[2] Kai Cui, Arun Yethiraj and J. R. Schmidt, Influence of Charge Scaling on the Solvation Properties of Ionic Liquid Solutions, *J. Phys. Chem. B* **2019**, 123, 9222-9229.

[3] Kai Cui and J. R. Schmidt, Enabling Efficient and Accurate Computational Studies of MOF Reactivity via QM/MM and QM/QM Methods, *J. Phys. Chem. C* **2020**, 124, 10550-10560.

[4] Zhenxuan Chen, Yue Gui, Kai Cui, J. R. Schmidt and Lian Yu, Prolific Polymorph Generator ROY in its Liquid and Glass: Two Conformational Populations Mirroring the Crystalline-state Distribution, *J. Phys. Chem. B* **2021**, 125, 10304-10311.

[5] Kai Cui, Souryadeep Bhattacharyya, Sankar Nair and J. R. Schmidt, Origins of Acid-Gas Stability Behavior in Zeolitic Imidazolate Frameworks: The Unique High Stability of ZIF-71, *J. Am. Chem. Soc.* **2021**, 143, 18061-18072.

[6] Yuhui Li, Camille Bishop, Kai Cui, J. R. Schmidt, Mark Ediger and Lian Yu, Surface Diffusion of a Glassy Discotic Organic Semiconductor and the Surface Mobility Gradient of Molecular Glasses, *J. Chem. Phys.* **2022**, 156, 094710.

[7] Kai Cui, Sankar Nair, David S. Sholl and J. R. Schmidt, Kinetic Model of Acid Gas-Induced Defect Propagation in Zeolitic Imidazolate Frameworks, *J. Phys. Chem. Lett.* Under revision.

[8] Chengzhai Wang, Kai Cui, Jessica Rimsza, Tina Nenoff, J. R. Schmidt and Krista S. Walton, A Systematic Study of N-modified UiO-66 Type MOFs for SO_2 Adsorption, In prep.

[9] Chunyi Li, Kai Cui, Zhenzi Yu, David S. Sholl, J. R. Schmidt and Ryan P. Lively, Binary CO_2/CH_4 and H_2S/CO_2 Adsorption Equilibrium Measurement in MIL-125-NH_2, In prep.

TABLE OF CONTENTS

Page

Page

List of Tables

Table Page

List of Figures

Figure Page

Figure Page

Chapter 1

General Introduction

1.1 Metal-Organic Frameworks

Metal-organic frameworks (MOFs) are crystalline nanoporous materials built by inorganic metal ions or metal clusters and polydentate organic linkers. These materials have attracted great attention since the first publications in 1990s[1]. A key advantage that differentiate MOFs from traditional porous materials like zeolites and active carbons is the extremely diverse variability of the metal nodes and organic linkers, as well as their connectivities[2]. In common MOFs, the nodes of the framework can be either bare metal ions like the Zn^{2+}/Co^{2+} ions in zeolitic imidazolate frameworks (ZIFs)[3] or clusters formed by metal ions and coordinating multidentate functional groups (such as OH and COO groups). The later is also known as the secondary building units (SBUs). Some examples of the SBUs are the square paddle wheel $Cu_2(COO)_4$ in HKUST-1[4], $Ti_8O_8(OH)_4(COO)_{12}$ cluster in MIL-125 [5, 6] and $Zr_6O_4(OH)_4(COO)_{12}$ cluster in UiO-66 family[7, 8], as shown in Figure 1.1. Meanwhile, various functional groups can be added to the benzendicarboxylate (BDC) or azolate linkers through organic synthesis to control the properties of the linkers. In principle, one can choose specific metal ions and linkers to design a MOF with desired properties (e.g. surface area, pore size, hydrophobicity, catalytic activity, etc.). The $Zr_6O_4(OH)_4(COO)_{12}$ cluster in UiO-66 provides a high degree of network connection and thus leads to high material stability[9]. The Cu paddle wheel $Cu_2(COO)_4$ SBU in MOF-11 and HKUST-1 contain open metal sites (OMSs) and can enable stronger chemisorption of gas molecules[4, 10]. The pore volume and surface area of a MOF can be systematically increased via isoreticular expansion[11], which

can be simply thought as using more extended organic linkers to replace original ones. For example, a series of isoreticular structures of MOF-5 with increased linker length have been synthesized. In this expansion, the pore volume can be increased by a factor of 8[12]. Adding/replacing functional groups on the linker can change the hydrophobicity of the framework. ZIF-8 with 2-methylimidazolate linker is hydrophobic while ZIF-90, with the same toppology as ZIF-8 but imidazole-2-carboxaldehyde linker, is hydrophilic[13].

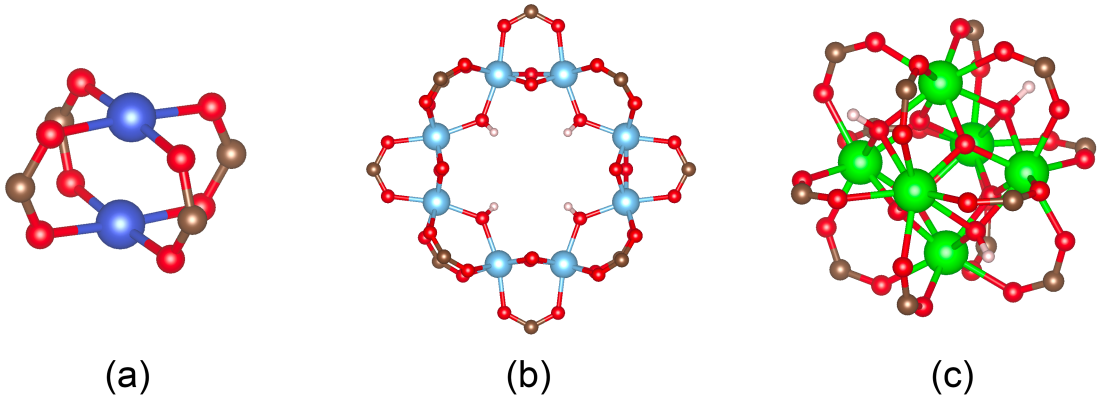


Figure 1.1: Examples of the SBU in common MOFs. (a) $Cu_2(COO)_4$ paddle wheel in HKUST-1. (b) $Ti_8O_8(OH)_4(COO)_{12}$ cluster in MIL-125. (c) $Zr_6O_4(OH)_4(COO)_{12}$ cluster in UiO-66. The Cu, Ti, Zr, C, O and H atoms are shown in dark blue, light blue, green, tan, red and white.

The most promising property of MOFs is the ultrahigh porosity and ultralarge surface area. The pore volume in the record holder MOF-399 is 94% [14]. The highest surface area of all reported MOF is almost 8000 m^2/g [15] and the hypothetical maximum can even reach $\sim$14600 m^2/g [16]. The large pore volume and large surface area make MOFs great candidates as adsorbents for industrial application of gas storage and separation[17–22]. H_2 storage, methane storage and CO_2 separation and capture are some of the most concerning issues.

Studies of using of MOFs as reservoirs to store H_2 gas for clean energy use has been initiated since 2003[23]. Compared to the traditional cryogenic liquid hydrogen tanks, which require an extremely low operation temperature of $\sim$20K, the presence of the adsorbents generally increases

the temperature to "liquify" the hydrogen and thus can significantly reduce the energy cost[17]. Surface area is the predominant factor that determines the H_2 uptake at $\sim$ 77K [21]. Wong-Foy et. al. measured the satuation uptake of H_2 in a series of MOFs at 77K. The saturation gravimetric uptake in MOF-177 (7.5 wt%) is about three times higher than MOF-74 (2.3 wt%), in accord with the four-fold difference in their surface area (4746 m^2/g for MOF-177 and 950 m^2/g for MOF-74)[24]. For the MOF-177 case, the saturation uptake is very close to the 9 wt% target (2015) set by the U.S. department of energy, and is much larger than the storage capacity in existing tanks (3.4-4.7 wt%)[25]. In addition, MOFs with smaller pore sizes are more desirable, as small pore size ensures that most H_2 molecules can experience the attraction from the adsorbent wall[17]. The concentration of OMSs with low coordination number is more important than the surface area when considering ambient temperature adsorption[17]. In the MOF Mn-BTT, charge balancing Mn^{2+} ions are linked to the anionic framework after desolvation and serve as extraframework OMSs. Neutron diffraction data shows clear evidence of H_2 binding to the OMSs. A capacity of 0.94 wt% or 12.1 g/L is observed at 298 K and 90 bar, much higher than the capacity in a high-pressure hydrogen tank at the same condition[26].

Adsorption based CO_2 capture is built on the the difference in the adsorbate-adsorbent interaction strength between CO_2 and other gas molecules. For MOFs, this is a property that can be easily controlled by rationally choosing the components. Among all MOFs, ZIFs are especially superior candidates for industrial CO_2 capture due to their high thermal stability, high affinity for CO_2 at low pressures, and better selectivity[3]. The selective adsorption of CO_2 from several binary mixtures has been measured for a series of ZIFs[18, 27]. The choice of the other component (CO, CH_4, N_2, O_2) covers several possible industrial scenarios including natural-gas purification and combustion-related separation[28]. For almost all cases, ZIFs show much higher selectivity compared to the BPL carbon adsorbent used in current industry. For exmaple, the selectivity factor of CO_2 over N_2 is only 17.8 in BPL carbon, but well above 20 in ZIF-79, ZIF-81 and ZIF-100. The largest values observed reach 35.3 and 50.1 in ZIF-82 and ZIF-78, respectively[27]. The CO_2 capture performance of MOFs can also be improved via chemical modification of the framework.

For example, MOFs with OMSs, like Mg-MOF-74, exhibit notably high CO_2 uptake at low pressures, but their adsorption capacity is sensitive to the humidity[17]. Attaching alkylamines to the OMSs enables cooperative chemisorption of CO_2 and thus results in good selectivity even with the presence of moisture[29, 30].

The uniform porosity feature of MOFs is also beneficial for selective catalysis. There are numerous reports that use MOFs as catalysts for hydrogenation, oxidation, epoxidation, rearrangement, C-C coupling and other reactions, as summarized in Ref.[22].

1.2 Acid Gases-Induced Degradation in MOFs

In practical industrial process streams (e.g., flue gas, natural gas, bioreactor off-gas), acid gas contaminants (CO_2, SO_x, NO_x, H_2S, etc.) and humidity are ubiquitous[31]. Therefore, the desired adsorbent material should be resistant to these acid gases. However, many MOFs failed this requirement: they are unstable in such humid acidic environments. Here we will focus on the stability behavior of the ZIF subclass.

Though known by their high thermal and chemical stability in neutral and basic environments, ZIFs also suffer from the vulnerability in acidic conditions[32–38]. The sensitivity of ZIF-8 to CO_2 in the presence of moisture is first reported in 2014. In this work, a new crystalline phase was observed in the powder X-ray diffraction (PXRD) data after exposing in pure CO_2 gas with 100% relative humidity (RH) for three days. Fourier transform infrared (FTIR) spectrum and solid state ^{13}C nuclear magnetic resonance (NMR) confirmed the formation of carbonate species[32]. Lower concentration of CO_2 gas (12%, with 90% RH) could also degrade common ZIFs including ZIF-8, ZIF-90, etc., resulting in significantly loss in the 1-butanol adsorption capacity by as much as a half[33]. Bhattacharyya et. al. investigated the degradation of ZIF-8 under exposure in trace amount (10-20 ppm) of SO_2[34]. After $\sim$10 days of humid SO_2 exposure, the morphology of the crystal changed significantly with rounded edge and lost of the well-defined shape. The PXRD pattern also shows considerably decreased peak intensities and increased background signal. Both phenomena indicates the structural amorphization after exposure. FTIR spectrum also shows newly emerged peaks corresponding to the acid species after exposure. Those peaks still present after

reactivation, which is a sign of permanent incorporation of acid species in the framework[34]. Time evolution of pore volume has been measured for multiple ZIFs in humid SO_2 environment via nitrogen adsorption, the results show that within $\sim$10 days of exposure, most ZIFs will lose all pore volume and thus any adsorption capacity[36]. The humidity plays an important role in the degradation of ZIFs under SO_2 and CO_2 exposure, because no degradation is observed for the dry acid exposure cases[32, 34, 36]. This observation indicates that the attacking species are carbonic and sulfuric/sulfurous acid molecules formed by hydration (and oxidation). All three species were detected in experiments[32, 34]. In contrast, dry NO_2 (1000 ppm) is sufficient to degrade ZIF-8 and ZIF-90. Other ZIFs with various SO_2 stabilities are all unstable in dry and humid NO_2[37]. In summary, most ZIF materials undergo catastrophic structural collapse in several days when exposing to ppm level of (humid) acid gases, resulting in a complete lost of adsorption capacity. This instability of ZIFs in acidic environments prohibits their practical use. To elongate the lifetime of ZIFs and other MOFs in the acidic environment, it is of great importance to understand the key factors that govern the stability.

In many cases, the chemical stability of MOFs is correlated with empirical descriptors[36, 39–43]. In humid and acidic environments, water or acid (e.g. H_2CO_3, H_2SO_3, H_2SO_4 and HNO_3) molecules act as nucleophiles to attack the metal-linker coordinate bonds, leading to defective sites with substituted ligands on the metal nodes and protonation on the linkers[34, 39, 40, 44]. Therefore, the acid gas stability of the material is largely determined by the bond strength between the metal nodes and the linkers. Linker pK_{a} is a particularly useful metric of the coordinate bond strength, as it quantitatively describes the linkers' ability to donate electrons. Higher pK_{a} of the linker often correlates with higher thermodynamic stability of the material[42, 43]. However, an opposite trend has been reported in terms of the kinetic stability: the linker pK_{a} values exhibit strong negative correlation with degradation rate of multiple ZIFs under humid SO_2 exposure[36]. When considering the stability in humid environments, the hydrophobicity of the framework is another good indicator. In hydrophobic MOFs, water can hardly be adsorbed in pore or cluster near the metal node, which makes these MOFs kinetically stable[42]. Steric effects can also slow

down the reaction between MOFs and acid gases. The topology and connectivity of the framework can be used as the descriptors for such cases.

Alternatively, one can directly assess the reactivity of MOFs with acid gases from a microscopic perspective. As discussed above, the degradation of MOFs under acid gas exposure is initiated by the nucleophilic substitution of the linkers by acid molecules, which leads to point defects around the metal node[34, 39, 40, 44]. The overall reactivity of the materials can thus be approximately represented by the thermodynamic and kinetic stability of such defects. First-principles calculations are powerful tools to provide atomistic insights into the defect formation processes. Water and acid-gas-induced point defects in ZIFs were characterized via periodic density functional theory (DFT) calculations by Zhang and Han et. al.[44, 45]. Figure 1.2 depicts the defect formation process in ZIF-8 under humid air exposure. Firstly, water as a nucleophile attacks the Zn metal node, cleaves the Zn-N bond, and substitutes the imizazolate (Im) linker, leading to a dangling linker (DL) defect. The remaining Zn-N bond in the DL defects could be attacked by another water molecule. This reaction frees a protonated linker from its original site and leads to the linker vacancy (LV) defect. For other acid gas exposure cases, the acid molecules act as the reactants and the defect formation follows the same mechanism as in Figure 1.2.

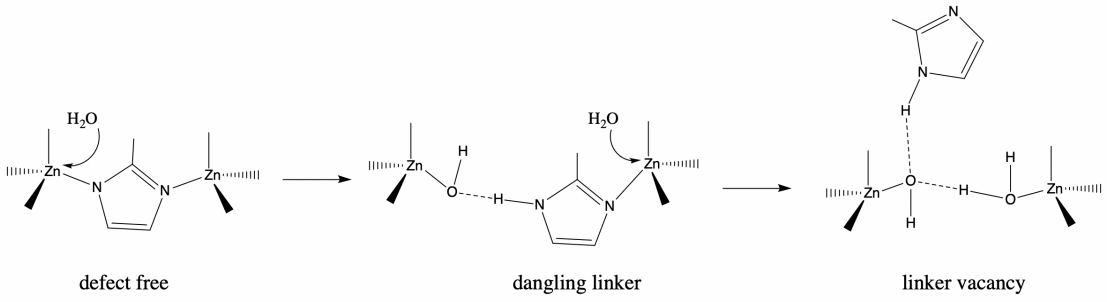


Figure 1.2: Structures and formation processes of the water-induced DL and LV defects in ZIF-8.

Zhang et. al. calculated the formation energy of the DL and LV defects in ZIF-8 induced by water, HNO_3 and HCOOH[44]. In DL structures, the insertion of an acid molecule between the metal and linker distorts the framework and introduces lattice strain. As a result, the formation energy of all three DL defects are positive. This lattice strain is released in LV defects. If the free

imidazole linker remains adsorbed in the framework, the LV formation energy of all three cases are negative, which means that the defective crystals containing LVs are more stable than the pristine material. The calculated activation energies for water-induced DL and LV formation in ZIF-8 are surmountable ($\sim$20 kcal/mol) at ambient temperature, confirming the kinetic accessibility of these defect structures.[44]. Han et. al. comprehensively studied the influence of acidity of the acid gases as well as the topology, functionalization and external surface of the frameworks on the formation of defects[45]. For both Brønsted acid (H_2S, H_2SO_3, and H_2SO_4) and Lewis acid (SO_2, SO_3) reactants, the DL formation energies and activation energies decrease as the acidity increases. ZIFs with Im linkers (like ZIF-2) are systematically more reactive than ZIFs with 2-methylimidazolate (mIm) linkers like (ZIF-8) across all eight topologies studied, due to the steric bulk of the mIm linker. The defect formation is also more favorable on an external surface than in the bulk[45]. Using the DL formatioan energies as an indicator, the overall stability behavior of the material has been successfully explained in many ZIF systems[35–37].

Previous computational assessments on the stability of MOFs have made great success, but they still have some limitations. For example, most studies applied DFT with generalized gradient approximation (GGA), which may not be accurate enough to make quantitative comparison of the stability. Though more positive defect formation energy generally means that the material is more stable, several outliers were found in the experiments, which indicates that the stability can not be not solely determined by the defect formation energy; Finally, there is currently no model to simulate macroscopic property change (like pore volume loss) as observed in the experiments. In this thesis, we will develop novel computational tools to solve these problems. The next section provides an overview of the detailed contents in this thesis.

1.3 Overview

In this thesis, we focus on the development of novel computational tools to gain deeper understanding of the acid gas-induced degradation in MOFs. Such model development extensively relies on the theory of quantum mechanics and statistical mechanics. In Chapter 2 and Chapter 3 we first briefly review these basic theories and some specific topics of quantum mechanics and

statistical mechanics, respectively. Chapter 2 discusses the variety of the electronic structure methods, including the local correlation methods and multiscale modeling (like the hybrid quantum mechanics/molecular mechanics or QM/MM and quantum embedding methods). The later can directly benefit the understanding of Chapter 4. In Chapter 3, we start from the introduction of statistical ensembles and the partition function. Then we focus on two specific topics: the adsorption isotherm and the transition state theory. The former can be used to describe the adsorption and distribution of gas molecules in MOFs (as in Chapter 5 and 7) and the later is used to estimate the reaction rates of defect formation (as in Chapter 5 and 6). We will cover both the derivation of the equations and the discussion about the practical computational use in these two chapters.

General muiltiscale methods discussed in Chapter 2 may not be directly suitable for the calculation of MOFs. In Chapter 4, we will introduce specific QM/MM and QM/QM methods tailored for MOF systems to enable high levels of accuracy and efficiency in studies of MOF reactivity. Using defect formation in a ZIF as a prototypical example, we will demonstrate the importance of correct boundary treatment to account for the dative bond cleavage and long-range mechanical coupling. Subsequently we will leverage these cluster models in a QM/QM scheme to yield "gold-standard" correlated wave function results on MOFs at modest computational cost.

In Chapter 5 we will take advantage of the developed QM/MM methods to understand the origins of the extraordinary stability of ZIF-71 in humid SO_2 environments. In the previous experimental studies of bulk stability of ZIFs, ZIF-71 (RHO topology) is the only ZIF retaining both structure and porosity under prolonged humid SO_2 exposure whereas many other ZIFs were shown to degrade. In this chapter we will perform detailed mechanistic studies using the combination of QM/MM calculations and the statistical mechanics-based adsorption models. We will discuss the influence of acid gas distribution pattern on the defect formation mechanism in ZIFs. The unique acid gas stability of ZIF-71 can be successfully explained as a kinetic effect due to the indirect reaction mechanism.

Chapter 6 focuses on the long-time evolution of ZIFs upon acid gas exposure. We will develop a first-principles-based kinetic model to simulate the long-time defect propagation. The influence of the interaction between neighboring defects and the pore blocking effects will be included in the

model. With input parameters obtained from QM/MM calculations, our model closely reproduces the experimentally-measured time-dependent pore volume loss of ZIF-8 and ZIF-71. Our results also provide important new insights regarding the autocatalytic nature of ZIF degradation and the spatial distribution of defects.

Lastly, we will shift our focus from acid-induced degradation of MOFs to acid gas removal using MOFs in Chapter 7. Using the SO_2 adsorption in N-modified UiO-66 MOFs as an example, we will perform first principles calculations along with the Langmuir adsorption model to calculate the experimental measurable adsorption free energies and to understand the role played by the N atoms in these modified MOFs. We will also closely compare the binding of SO_2 to the N sites in two MOFs UiO-66-2,5-PDC and UiO-66-2,5-PziDC via energy decomposition and natural bond orbital analyses. These results highlight the influence of chemical environment of MOFs on their adsorption behaviors.

Chapter 2

Theoretical Background: Quantum Chemistry

2.1 Electronic Structure Theory

Electronic structure theory focuses on solving the Schrödinger equation of molecular systems. With the Born-Oppenheimer approximation, the electronic and nuclear degrees of freedom can be separated and the electronic Hamiltonian (in atomic unit) is:

$$\hat{H} = -\frac{1}{2}\sum_{i}^{N}\nabla_i^2 - \sum_{i}^{N}\sum_{A}^{M}\frac{Z_A}{r_{iA}} + \sum_{i<j}^{N}\frac{1}{r_{ij}} + \sum_{A<B}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{2.1}$$

where Z is the nuclear charge, r_{ij} denotes the distance between two electrons, r_{iA} denotes the distance between electron i and nuclei A, and R_{AB} denotes the distance between two nucleus. The electronic wave function $\Psi(\mathbf{r};\mathbf{R})$ is a function of all electronic coordinates $\mathbf{r}$ and parametrically depends on the nuclear coordinates $\mathbf{R}$. The electronic Schrödinger equation,

$$\hat{H}\Psi(\mathbf{r};\mathbf{R}) = E(\mathbf{R})\Psi(\mathbf{r};\mathbf{R}) \tag{2.2}$$

cannot be solved analytically except for the simpliest H_2^+ molecule. As a result, approximated methods must be applied.

2.1.1 Hartree-Fock Method

The most straightforward approximation is the mean field approximation. The electronic Hamiltonian Eq.2.1 contains one-electron and two-electron terms (the nuclear interaction term is a constant and a given nuclear configuration and can thus be omitted):

$$\hat{H} = \sum_{i}^{N}\hat{h}(i) + \sum_{i<j}^{N}\frac{1}{r_{ij}} \tag{2.3}$$

where

$$\hat{h}(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A}^{M}\frac{Z_A}{r_{iA}} \tag{2.4}$$

is called the one-electron core Hamiltonian. In the mean field approximation, rather than explicitly accounting for the interactions between electrons, we consider the electron moving in the mean field V_{HF} generated by all other electrons, that is, replace the two-electron term $1/r_{ij}$ by an one-electron potential $V_{\mathrm{HF}}(i)$. we can replace the N-electron problem by N one-electron problems:

$$\hat{f}(i)\chi_a(i) = \left(\hat{h}(i) + V_{\mathrm{HF}}(i)\right)\chi_a(i) = \varepsilon_a\chi_a(i) \tag{2.5}$$

This equation describes the movement of an electron in the field of the nuclei and the mean field created by other electrons. The solution χ_a is the one-electron wave function (the spin-orbital) and ε_a is the orbital energy. Within the mean field approximation, the electron-electron interaction has been averaged out and the electrons are independent. The wave function of the N-electron system can thus be written as the anti-symmetrized "product" (the Slater determinant) of the spin-orbitals:

$$\Phi(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} \equiv |\chi_1\chi_2...\chi_N\rangle \tag{2.6}$$

where $\mathbf{x}_i = \{\mathbf{r}_i, w_i\}$ and w_i is the spin coordinate. Note that the Slater determinant wave function is a direct consequence of the mean field approximation, conversely, using a single Slater determinant as the wave function also inherently uses the mean field treatment.

Now the problem left is to find $V_{\mathrm{HF}}(i)$. In the Hartree-Fock (HF) method, we use variational theorem to get the best single Slater determinant wave function $|\Phi_0\rangle = |\chi_1\chi_2...\chi_N\rangle$ to describe the system. By minimize the energy $E = \langle\Phi_0|\hat{H}|\Phi_0\rangle$ with respect to $\{\chi_a\}$, we can get the HF equation Eq.2.5 of the orbitals with the Fock operator:

$$\hat{f}(i) = \hat{h}(i) + \sum_b^N \left(\hat{J}_b(i) - \hat{K}_b(i)\right) \tag{2.7}$$

where $\hat{J}_b$ and $\hat{K}_b$ are the Coulumb and the exchange operators respectively. There expectation values are the two-electron integrals:

$$\begin{aligned} \langle\chi_a(1)|\hat{J}_b(1)|\chi_a(1)\rangle &= \int \mathrm{d}\mathbf{x}_1\, \chi_a^*(1)\left[\int \mathrm{d}\mathbf{x}_2\, \chi_b^*(2)\frac{1}{r_{12}}\chi_b(2)\right]\chi_a(1) \\ &= \int \mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}_2\, \chi_a^*(1)\chi_b^*(2)\frac{1}{r_{12}}\chi_a(1)\chi_b(2) \equiv \langle ab|ab\rangle \end{aligned} \tag{2.8}$$

$$\begin{aligned}\langle\chi_a(1)|\hat{K}_b(1)|\chi_a(1)\rangle &= \int \mathrm{d}\mathbf{x}_1\, \chi_a^*(1)\left[\int \mathrm{d}\mathbf{x}_2\, \chi_b^*(2)\frac{1}{r_{12}}\chi_a(2)\right]\chi_b(1) \\ &= \int \mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}_2\, \chi_a^*(1)\chi_b^*(2)\frac{1}{r_{12}}\chi_b(1)\chi_a(2) \equiv \langle ab|ba\rangle\end{aligned} \tag{2.9}$$

The Coulomb integral $\langle ab|ab\rangle$ is clearly the Coulomb interaction between the mean field electron densities $\rho_a(1) = |\chi_a(1)|^2$ and $\rho_b(2) = |\chi_b(2)|^2$. The exchange integral describes the exchange interaction between the electrons due to the Pauli exclusion principle, which has no classical analogues. The orbital energy ε_a can be constructed from the one-electron and two-electron integrals:

$$\varepsilon_a = \langle\chi_a|\hat{f}|\chi_a\rangle = \langle a|\hat{h}|a\rangle + \sum_b^N \left(\langle ab|ab\rangle - \langle ab|ba\rangle\right) \tag{2.10}$$

The total energy is the expectation value of the electronic Hamiltonian:

$$\begin{aligned}E = \langle\Phi|\hat{H}|\Phi\rangle &= \sum_a^N \langle a|\hat{h}|a\rangle + \frac{1}{2}\sum_a^N\sum_b^N \left(\langle ab|ab\rangle - \langle ab|ba\rangle\right) \\ &= \sum_a^N \varepsilon_a - \frac{1}{2}\sum_a^N\sum_b^N \left(\langle ab|ab\rangle - \langle ab|ba\rangle\right)\end{aligned} \tag{2.11}$$

where a, b are indices for occupied orbitals. The second term in the second line corrects the double counting of the electron-electron interactions, since interactions between electrons in orbitals χ_a and χ_b are included in both ε_a and ε_b.

The Fock operator $\hat{f}(i)$ depends on the orbitals $\{\chi_a\}$ so the HF equation must be solved self-consistently. In practical, this self-consistent field (SCF) procedure is done by first expanding the spatial molecular orbital (MO) ψ_a (spatial part in χ_a) as a linear combination of atomic orbital (AO) basis functions $\{\phi_\mu\}$:

$$\psi_a = \sum_\mu^K C_{\mu a}\phi_\mu \tag{2.12}$$

The expansion transform the HF equation to a matrix forms known as the Roothaan equation:

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon} \tag{2.13}$$

where $\mathbf{F}$ and $\mathbf{S}$ are the Fock and overlap matrices respectively. $\boldsymbol{\varepsilon}$ and $\mathbf{C}$ are the eigenvalues and eigenvectors of $\mathbf{F}$. The Fock matrix elements contain the following four-index integrals (two-electron integrals in AO basis):

$$(\mu\nu|\lambda\sigma) = \int \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \, \psi_\mu^*(1)\psi_\nu(1)\frac{1}{r_{12}}\psi_\lambda^*(2)\psi_\sigma(2) \tag{2.14}$$

Evaluating these integrals is the most time consuming part in HF calculations. As a result, HF scales as the fourth power of the number of basis functions $O(K^4)$.

2.1.2 Correlation Energy and Configuration Interaction

In many-electron systems, the movement of electrons is correlated. Due to the electron-electron repulsion $1/r_{ij}$, the probability of finding two electrons at the same position in space should be zero. In the HF approximation, the encoded anti-symmetry in the Slater determinant assures that two electrons with opposite spin can not appear at the same position in space, but not for electrons with the same spin. In other words, HF method can not describe the "Coulomb correlation" between the electrons. The energy difference between the exact nonrelativistic energy and the HF energy (at the complete basis set limit) is called the correlation energy $E_{\mathrm{corr}} = E_{\mathrm{exact}} - E_{\mathrm{HF}}$. Since neglecting the Coulomb correlation leads to an overestimation of the electron-electron repulsion energy, E_{corr} is always negative. This conclusion is consistent with the fact that HF is a variational method, and thus the HF energy is always a upper limit to the exact energy.

From a mathematical perspective, the single Slater determinant wave function is an eigenfunction of the HF Hamiltonian $\hat{H}_{\mathrm{HF}} = \sum_i \hat{f}(i)$, but not an eigenfunction of $\hat{H}$, which means the exact wave function of the N-electron system can not be expressed as a single determinant. On the other hand, if $\{\chi_a\}$ is a complete one-electron basis set, then all determinants constructed by $\{\chi_a\}$, $\{|\Phi\rangle = |\chi_a\chi_b...\rangle\}$, forms a complete N-electron basis set. This N-electron basis set is equivalent to the sets of the HF ground state $|\Phi_0\rangle$ and all possible excited states $|\Phi_{ab...}^{rs...}\rangle$. The notation $|\Phi_{ab...}^{rs...}\rangle$ means excite electrons from occupied orbitals $\chi_a, \chi_b, ...$ in the HF ground state, to virtual orbitals (unoccupied orbitals) $\chi_r, \chi_s, ...$ The exact N-electron wave function can then be expressed as a linear combination of these determinants:

$$|\Psi_{\mathrm{exact}}\rangle = |\Phi_0\rangle + \sum_a \sum_r C_a^r |\Phi_a^r\rangle + \sum_{a<b} \sum_{r<s} C_{ab}^{rs} |\Phi_{ab}^{rs}\rangle + ... \tag{2.15}$$

where C's are the expansion coefficients, a, b run through all occupied orbitals and r, s run through all virtual orbitals. By minimizing $E = \langle \Psi_{\text{exact}} | \hat{H} | \Psi_{\text{exact}} \rangle$ we can solve for the C's. This method is called the "configuration interaction" (CI). If all the excited determinants are included, the method is called the full CI (FCI). At the complete basis set limit, the FCI method gives the exact energies and wave functions of both the ground state and the excited states.

FCI is the intellectually simplest method to calculate the correlation energy, but is computationally intractable. If K basis functions are used, there will be

$$\binom{2K}{N} = \frac{(2K)!}{N!(2K-N)!} \tag{2.16}$$

Slater determinants in the CI expansion, which means FCI scales as $O(K!)$. FCI calculation is way too expensive for any practical calculations, except for smallest molecules. More commonly used methods are the Møller-Plesset perturbation theory (like MP2) and the coupled cluster (CC) methods. The later will be discussed in detail in next section.

2.1.3 Coupled Cluster Method

Recall that in the FCI expansion Eq.2.15, the exact N-electron wave function can be expressed as a linear combination of HF ground and excited state determinants. All possible excited states $|\Phi_{ab\ldots}^{rs\ldots}\rangle$ are included in this expansion. The coupled cluster (CC) wave function is also a linear combination of these determinants, but is expressed using the exponential ansatz:

$$|\Psi_{\text{CC}}\rangle = e^{\hat{T}} |\Phi_0\rangle \tag{2.17}$$

where $\hat{T}$ is the cluster operator

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \ldots \tag{2.18}$$

$\hat{T}_i$'s are excitation operator that generates all i-fold excitation states. For example, the single and double excitation operators are:

$$\hat{T}_1 |\Phi_0\rangle = \sum_a \sum_r t_a^r |\Phi_a^r\rangle \tag{2.19}$$

$$\hat{T}_2 |\Phi_0\rangle = \sum_{a<b} \sum_{r<s} t_{ab}^{rs} |\Phi_{ab}^{rs}\rangle \tag{2.20}$$

t_a^r and t_{ab}^{rs} are unknown coefficients called "amplitudes" which will determine the CC wave function. In practical calculations, the cluster operator is generally truncated. For example, in coupled cluster singles and doubles (CCSD) theory, only single and double excitations are considered: $\hat{T} = \hat{T}_1 + \hat{T}_2$. However, due to the exponential ansatz, all i-fold excited Slater determinants with $i > 2$ are still included in the CCSD expansion:

$$\begin{aligned} |\Psi_{\mathrm{CC}}\rangle &= e^{\hat{T}_1+\hat{T}_2}|\Phi_0\rangle = \left(1 + (\hat{T}_1 + \hat{T}_2) + \frac{1}{2}(\hat{T}_1 + \hat{T}_2)^2 + ...\right)|\Phi_0\rangle \\ &= |\Phi_0\rangle + \hat{T}_1|\Phi_0\rangle + \left(\hat{T}_2 + \frac{1}{2}\hat{T}_1{}^2\right)|\Phi_0\rangle + \left(\frac{1}{2}\hat{T}_2\hat{T}_1 + \frac{1}{2}\hat{T}_1\hat{T}_2 + \frac{1}{6}\hat{T}_1{}^3\right)|\Phi_0\rangle + ... \end{aligned} \tag{2.21}$$

For example, the term $\left(\frac{1}{2}\hat{T}_2\hat{T}_1 + \frac{1}{2}\hat{T}_1\hat{T}_2 + \frac{1}{6}\hat{T}_1{}^3\right)|\Phi_0\rangle$ adds triple excitation determinants to the expansion. The key difference between the CCSD and the FCI wave functions is that the coefficients for the triple (and higher) excitations is CCSD expansion are fully determined (approximated) as products of single and double excitation amplitudes t_a^r and t_{ab}^{rs}.

We can choose the values of the amplitudes t_a^r and t_{ab}^{rs} so that the CC wave function is the eigenfunction of the Hamiltonian:

$$\hat{H}e^{\hat{T}}|\Phi_0\rangle = E_{\mathrm{CC}}e^{\hat{T}}|\Phi_0\rangle \tag{2.22}$$

Multiply $e^{-\hat{T}}$ to the left side of the equation and project it onto HF ground and excited states determinants, we have:

$$\langle\Phi_0|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = E_{\mathrm{CC}}\langle\Phi_0|\Phi_0\rangle = E_{\mathrm{CC}} \tag{2.23}$$

$$\langle\Phi_a^r|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = E_{\mathrm{CC}}\langle\Phi_a^r|\Phi_0\rangle = 0 \tag{2.24}$$

$$\langle\Phi_{ab}^{rs}|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi_0\rangle = E_{\mathrm{CC}}\langle\Phi_{ab}^{rs}|\Phi_0\rangle = 0 \tag{2.25}$$

The first equation is the expression of the CCSD energy and the last two equations are used to solve t_a^r and t_{ab}^{rs}.

CCSD method scales with the system size as $O(N^6)$. Sometimes the accuracy of CCSD is not sufficient, contribution from the triples can be added via an perturbative manner. This method is denoted as CCSD(T) and is often be considered as the "gold standard" in quantum chemistry. The following table summarizes the commonly used electronic structure methods with increasing accuracy:

Accuracy	HF	$\ll$	MP2	$<$	CCSD	$<$	CCSD(T)
Scaling	N^4		N^5		N^6		N^7

Table 2.1: Accuracy and scaling of commonly used electronic structure methods.

2.2 Local Correlation Method

In non-metallic systems, electron correlation is a local effect, as the magnitude of the correlation energy has a distance dependence of $\propto r^{-6}$ (the dispersion interaction is an electronic correlation effect)[46]. The computational cost of such local effects is expected to have lower-order scaling with system size. But as we discussed in the previous section, the canonical correlated wave function methods have very steep scaling from $O(N^5)$ to $O(N^7)$. The reason is that the MOs used in these canonical methods are inherently delocalized, which destroys the local character of the correlation. On the other hand, if localized orbitals are used, only the electron pairs that are "close" to each other contributes the most to the correlation energy. Though there is a large number of "very distant pairs" in the system, their contribution to the correlation energy is negligible[46]. Therefore, it's safe to approximately calculate or even neglect these pairs. The number of electron pairs to be included in the correlation calculation (and thus the computational cost) will be significantly reduced.

2.2.1 Example: LCCSD Method

We use the local CCSD (LCCSD) method[47] as an example to illustrate how the local correlation methods work. Since the Slater determinant wave functions are invariant under unitary transformations, the canonical MOs $\{\psi_a\}$ can be transformed to local MOs (LMOs) $\{\tilde{\psi}_a\}$ while keeps the ground state wave function $|\Phi_0\rangle$ unchanged.

$$|\Phi_0\rangle = |\psi_1\psi_2...\psi_N\rangle = |\tilde{\psi}_1\tilde{\psi}_2...\tilde{\psi}_N\rangle \tag{2.26}$$

The details of how to perform orbital localization will be discussed in the next section. The LMOs transformed from the occupied canonical MOs span the "occupied space". The "virtual space" is

constructed by projecting the AOs orthogonal to the occupied space (since the canonical MOs are mutually orthogonal but the AO basis functions are not).

$$\tilde{\phi}_\mu = (1 - \hat{P}_{\rm occ})\phi_\mu = \left(1 - \sum_{a=1}^{N} |\tilde{\psi}_a\rangle\langle\tilde{\psi}_a|\right)\phi_\mu \tag{2.27}$$

Each occupied LMO $\tilde{\psi}_a$ corresponds an "orbital domain" (denoted as $[a]$). The orbital domain is a subspace in the virtual space that contains a subset of projected AOs $\{\tilde{\phi}_\mu\}$ that is spatially close to $\tilde{\psi}_a$. Typically, for a lone pair LMO, only basis functions at the same atom are included in $[a]$. For a bicentric bonding orbital, the basis functions of the two connected atoms are selected[47].

In the canonical CCSD method, the excitation are made from delocalized occupied MO ψ_a to delocalized virtual orbital ψ_r (see Eq.2.19-2.20). For LCCSD, excitations are made from the occupied LMOs to their corresponding orbital domains[47]. By doing so, the excitation can be localized.

$$\hat{T}_1|\Phi_0\rangle = \sum_a \sum_{\mu\in[a]} \tilde{t}_a^\mu |\Phi_a^\mu\rangle \tag{2.28}$$

$$\hat{T}_2|\Phi_0\rangle = \sum_{a<b} \sum_{\mu,\nu\in[ab]} \tilde{t}_{ab}^{\mu\nu} |\Phi_{ab}^{\mu\nu}\rangle \tag{2.29}$$

where $[ab]$ is called the "pair domian" corresponding to ψ_a and ψ_b, which is just the union of $[a]$ and $[b]$. Further more, in the double excitation term, the LMO pairs can be catagorized as strong, weak and distant pairs according to some distance thresholds ($r_w,\ r_d$)[47].

$$a, b \text{ are} = \begin{cases} \text{strong pairs}, & \text{if } r_{ab} \le r_w \\ \text{weak pairs}, & \text{if } r_w < r_{ab} \le r_d \\ \text{distant pairs}, & \text{if } r_{ab} > r_d \end{cases} \tag{2.30}$$

where r_{ab} is the distance between two LMOs. Only strong pairs are allowed for the double excitation in LCCSD, while the weak pairs will be treated approximately (like by MP2) or neglected, yet $98 \sim 99\%$ of the correlation energy can be recovered[47, 48]. The subsequent construction of the LCCSD wave function and the equations to calculate the amplitudes are similar as in the canonical methods. Since the number of strong pairs increases only linearly with the molecular size, the computational cost is significantly reduced as compared to canonical CCSD, as shown in

Figure 2.1. With the local correlation methods we are able calculate the energetics of ZIF clusters with ~ 100 atoms and more than 300 electrons at coupled cluster level.

2.2.2 Choice of the Localization Scheme

Since the unitary transformation to localize MOs is not unique, these transformation matrices can be determined by optimize various quantities that characterize the locality of the orbitals. For example, in Edmiston-Ruedenberg (ER) localization, the self-repulsion energies of the electrons are maximized[52]:

$$\max \left[\sum_a \int \psi_a(\mathbf{r}) \psi_a(\mathbf{r}) \frac{1}{|r - r'|} \psi_a(\mathbf{r}') \psi_a(\mathbf{r}') \, \mathrm{d}\mathbf{r} \mathrm{d}\mathbf{r}' \right] \tag{2.31}$$

The Boys localization minimizes the spatial extent of the orbitals[53]:

$$\min \left[\sum_a \int \psi_a(\mathbf{r}) \psi_a(\mathbf{r}) |r - r'|^2 \psi_a(\mathbf{r}') \psi_a(\mathbf{r}') \, \mathrm{d}\mathbf{r} \mathrm{d}\mathbf{r}' \right] \tag{2.32}$$

Pipek-Mezey (PM) localization minimizes the number of atoms that each orbital is localized on[54], or equivalently, maximizes the following quantity:

$$\max \left[\sum_a \sum_A^{\text{atoms}} (Q_a^A)^2 \right] \tag{2.33}$$

where Q_a^A is the atomic Mulliken population of orbital a on atom A. Bonding analysis procedures like the natural bond orbital (NBO)[55] analysis and intrinsic basis bonding analysis (IBBA)[56] can also give LMOs.

The choice of the localization schemes can affect the properties of the LMOs. ER and PM localization conserve the $\sigma - \pi$ separation for double bonds while Boys localization gives"banana bonds". The mixing of σ bonds in Boys localization allows delocalized π orbitals to be better localized than with PM localization[57, 58]. In contrast, the criteria used in PM localization will allow some of the extended orbitals not to localize so much as the others to gain a better global result[58]. The result quality of local correlation calculations also depends on the localization method used. In brief, for physically well-localized system Pipek-Mezey localization has advantage while for conjugated or aromatic molecules Boys localization is better (due to the $\sigma - \pi$ mixing)[58].

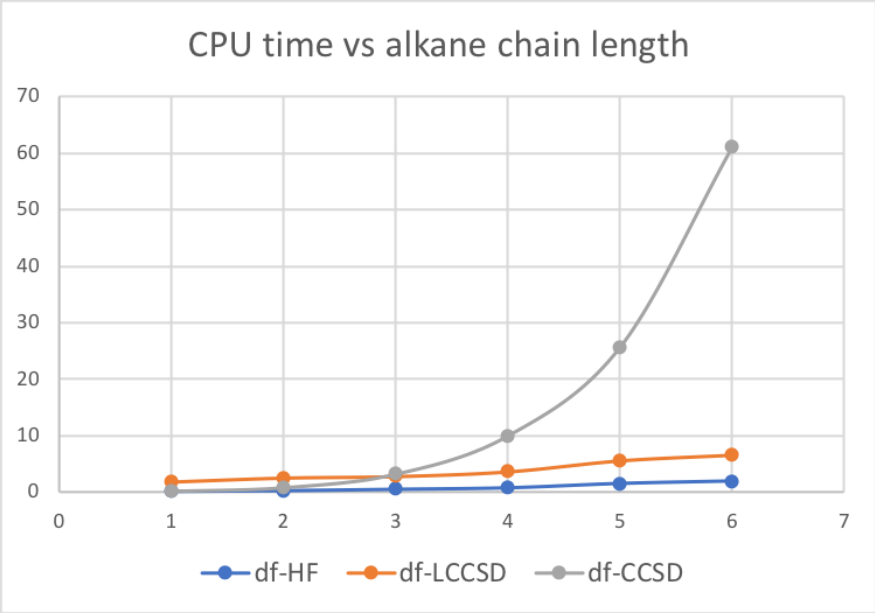


Figure 2.1: Scaling of the HF, LCCSD and CCSD methods with alkane chain length calculated using double zeta Dunning style basis set[49]. The density fitting (df) technique[50, 51] is used to accelerate integral calculation.

We are more interested in the performance of these localization schemes on ZIF related systems. We calculated the dissociation energies of water, imidazole and benzene dimers as well as the bond dissociation energies of ZnL_4 clusters using LCCSD with different localization methods and compared with CCSD:

$$(\mathrm{L})_2 \rightarrow 2\mathrm{L} \tag{2.34}$$

$$\mathrm{ZnL_4} \rightarrow \mathrm{ZnL_3} + \mathrm{L} \tag{2.35}$$

The error in the dissociation energy are summarized in Table 2.2. We can see that all four localization methods tested give great accuracy compared to CCSD. The root mean squared error (RMSE) is smaller than 0.1 eV for all cases. The bond dissociation energy in anionic $\mathrm{ZnL_4}$ clusters seems more difficult to handle than other test systems. Among these localization schemes tested, the Boys localization gives the best overall performance. We will thus use the Boys localization for all local correlation calculations throughout the thesis.

System	PM	Boys	NBO	IBBA
$(\mathrm{H_2O})_2$	-0.02	-0.02	-0.02	-0.02
$(\mathrm{HIM})_2$	-0.03	-0.01	-0.02	-0.04
$(\mathrm{benzene})_2$	-0.06	-0.02	-0.01	-0.02
$\mathrm{Zn(H_2O)_4^{2+}}$	-0.06	-0.06	-0.04	-0.06
$\mathrm{Zn(OH)_4^{2-}}$	-0.16	-0.13	-0.13	-0.13
$\mathrm{Zn(HIM)_4^{2+}}$	-0.06	-0.05	-0.06	-0.09
$\mathrm{Zn(IM)_4^{2-}}$	-0.06	-0.05	failed	-0.02
RMSE	0.08	0.06	0.06	0.09

Table 2.2: Error in the dissociation energy (units in eV) in LCCSD calculations with different localization schemes as compared with canonical CCSD. VDZ basis set is used for Zn, C and H and AVDZ basis set is used for N and O.

2.3 QM/MM Method

Despite the development of efficient electronic structure methods as well as fast computers, full quantum mechanical (QM) calculation on very large and complex systems (like proteins or MOFs) remains difficult and extremely time consuming. One possible way to make such calculations feasible is to use a multiscale description for large systems. One of the most famous multiscale formalism is the hybrid quantum mechanics/molecular mechanics (QM/MM) method developed by Karplus, Levitt and Warshel[59–62]. In the QM/MM method, the total system is divided into two subsystems: the small model system and the environment. The model system is where important chemistry (bond breaking/forming) takes place and is treated using accurate and expensive QM methods. Less accurate but much more efficient MM methods are applied to the environment to account for other essential physics (solvation, protein structure, etc.). Since the QM calculations are only applied on a small fraction of the large system, the computational cost is significantly lower.

The QM/MM total energy can be calculated as the summation of the energies of subsystems and their interaction energy. This expression is called the "additive" scheme:

$$E_{\mathrm{tot}}^{\mathrm{QM/MM}} = E_{\mathrm{sub}}^{\mathrm{QM}} + E_{\mathrm{env}}^{\mathrm{MM}} + E_{\mathrm{int,\,sub-env}} \tag{2.36}$$

Here $E_{\mathrm{sub}}^{\mathrm{QM}}$ is the energy of the model subsystem (the QM region) calculated at QM level, $E_{\mathrm{env}}^{\mathrm{MM}}$ is the energy of the environment (the MM region) calculated at MM level, and $E_{\mathrm{int,\,sub-env}}$ is the interaction between the model subsystem and the environment. The interaction term can be further decomposed to bonded, van der Waals and electrostatic interactions:

$$E_{\mathrm{int,\,sub-env}} = E_{\mathrm{bond}} + E_{\mathrm{vdW}} + E_{\mathrm{elec}} \tag{2.37}$$

The bonded and van der Waals interaction are short-ranged, so only the atoms on the QM-MM system boundary contribute to the interaction energy. Typically, these atoms are far from the reaction center and their characters won't change during the simulation. Therefore, these terms can be handled quite faithfully at the MM level, using a fixed set of parameters taken from a pre-defined force field (FF)[63].

The electrostatic interaction can also be treated in this classical manner, by assigning fixed atomic charges to the QM atoms and calculate their Coulomb interaction with MM charges as in a FF model:

$$E_{\text{elec}}^{\text{ME}} = \sum_{A}^{\text{sub}} \sum_{B}^{\text{env}} \frac{q_A q_B}{R_{AB}} \tag{2.38}$$

where q_A is the charge assigned to the QM atom A and q_B is the FF charge of the MM atom B. This approach leads to the "mechanical embedding" (ME) scheme. q_A could either be taken from a pre-defined FF or computed based on the electron density in QM region[63, 64].

Since the atoms in the MM region is not included when calculating the QM electron density (and thus q_A's), the polarization effect due to the presence MM charges is completely missed in ME. In order to include this polarization effect, one can add the MM charges as an external potential to the QM Hamiltonian:

$$V_{\text{elec}} = -\sum_{i}^{\text{sub}} \sum_{B}^{\text{env}} \frac{q_B}{r_{iB}} + \sum_{A}^{\text{sub}} \sum_{B}^{\text{env}} \frac{Z_A q_B}{R_{AB}} \tag{2.39}$$

The first term is the interaction between the electrons with MM charges and the second term is the interaction between MM charges and the QM nuclei. This scheme is called "electronic embedding" (EE). In the EE scheme, the electrostatic interaction is included in $E_{\text{sub}}^{\text{QM}}$ and is calculated at the QM level:

$$E_{\text{elec}}^{\text{EE}} = \langle \Psi | V_{\text{elec}} | \Psi \rangle \tag{2.40}$$

The EE scheme provides a more accurate description of the electrostatic coupling between the QM and MM regions than a ME scheme. But in EE scheme, the MM point charges near the boundary may also lead to overpolarization, special cares are thus required[63].

The additive QM/MM scheme requires an explicit calculation of the interaction between the subsystems. Such calculation can be omitted by using the "subtractive" scheme[65]:

$$E_{\text{tot}}^{\text{QM/MM}} = E_{\text{sub}}^{\text{QM}} - E_{\text{sub}}^{\text{MM}} + E_{\text{tot}}^{\text{MM}} \tag{2.41}$$

where $E_{\text{sub}}^{\text{MM}}$ is the MM energy of the QM region and $E_{\text{tot}}^{\text{MM}}$ is the MM energy of the total system. This expression can be considered as adding a "QM correction" to the MM total energy. In the

ME scheme, all the QM-MM interactions are included implicitly in $E_{\mathrm{tot}}^{\mathrm{MM}}$ at MM level. In the EE scheme, the electrostatic interaction is included in $E_{\mathrm{sub}}^{\mathrm{QM}}$ at QM level. Therefore, the "QM correction" also corrects for the electrostatic interaction in $E_{\mathrm{tot}}^{\mathrm{MM}}$. In general, for both ME and EE schemes, the additive and subtractive expressions of the total energy are equivalent.

The key advantage of the subtractive scheme is its simplicity. Any QM and MM codes can be used to calculate the QM/MM total energy without any modification. Since the explicit expression of the interaction between the subsystems is not needed, the subtractive scheme also enables the combination of two arbitrary methods, including a QM/QM partition[66]:

$$E_{\mathrm{tot}}^{\mathrm{QM2/QM1}} = E_{\mathrm{sub}}^{\mathrm{QM2}} - E_{\mathrm{sub}}^{\mathrm{QM1}} + E_{\mathrm{tot}}^{\mathrm{QM1}} \tag{2.42}$$

Furthermore, this subtractive expression can be generalized to multiple layers and each layer can be treated at an arbitrary QM or MM level. This generalization is known as the "our n-layered integrated molecular orbital and molecular mechanics" (ONIOM) method[64]. For example, in a three layer ONIOM QM2/QM1/MM method:

$$E_{\mathrm{tot}}^{\mathrm{QM2/QM1/MM}} = (E_{\mathrm{sub1}}^{\mathrm{QM2}} - E_{\mathrm{sub1}}^{\mathrm{QM1}}) + (E_{\mathrm{sub2}}^{\mathrm{QM1}} - E_{\mathrm{sub2}}^{\mathrm{MM}}) + E_{\mathrm{tot}}^{\mathrm{MM}} \tag{2.43}$$

and $\mathrm{sub1} \subset \mathrm{sub2} \subset \mathrm{tot}$.

The ME or EE formalisms work well when the QM-MM boundary only crosses intermolecular regions or unpolar, non-conjugated covalent bonds (like the C-C bond). In the former case, the only coupling between the subsystems is the electrostatic polarization effect, which can be correctly handled in EE schemes. For the later case, one can simply add extra link atoms (generally H atoms) to the QM region to saturate the cleaved bonds and avoid quantum coupling. In this thesis, we have developed special schemes to handle the polar, dative metal-linker bond in MOFs, which will be described in Chapter 4. On the other hand, the ME or EE formalisms fail in metals or conjugated molecules, because the electrons are delocalized across the boundary and the coupling between the subsystems is not electrostatic but a quantum effect. For correct multiscale descriptions of such systems, the quantum embedding theory should be used instead.

2.4 Quantum Embedding Theory

Quantum embedding theory is able to account for the quantum mechanical coupling (electrostatic and exchange correlation interaction) between electrons in the model subsystem and the environment. It provides much more accurate description of the electronic structure of the model subsystem, even for delocalized systems. The fundamental step in quantum embedding theory is to construct a modified Hamiltonian for the model subsystem to describe the coupling. This modified Hamiltonian has many manifestations. One intuitive construction is to add an effective one-electron potential called the "embedding potential" V_{emb} to describe the influence of the environment[67–70]. In this section, we will use one subclass of the quantum embedding theory, the density-based embedding methods, as examples to illustrate the basic ideas of how to obtain V_{emb}.

The density-based embedding methods starts from the density functional theory (DFT). The coupling between the model subsystem and the environments is due to the interaction between the electron densities in the two subsystems. The electron density of the total system ρ_{tot} can be partitioned into two parts:

$$\rho_{\mathrm{tot}} = \rho_{\mathrm{sub}} + \rho_{\mathrm{env}} \tag{2.44}$$

The total energy of the system can be expressed as:

$$E[\rho_{\mathrm{tot}}] = E[\rho_{\mathrm{sub}}] + E[\rho_{\mathrm{env}}] + E_{\mathrm{int}}[\rho_{\mathrm{sub}}, \rho_{\mathrm{env}}] \tag{2.45}$$

where $E[\rho]$ is a density functional of choice. The functional derivative of E_{int} with respect to ρ_{sub} gives the embedding potential for the model subsystem:

$$\frac{\delta E_{\mathrm{int}}}{\delta \rho_{\mathrm{sub}}} = V_{\mathrm{emb}} \tag{2.46}$$

If the environment is large enough and we choose ρ_{sub} and ρ_{tot} as independent variables, we can assume that $\delta E_{\mathrm{env}}/\delta \rho_{\mathrm{sub}} = 0$ [71]. According to Eq.2.45, the embedding potential can be calculated from the derivative of the energy functionals:

$$V_{\mathrm{emb}}[\rho_{\mathrm{sub}}, \rho_{\mathrm{tot}}] = \frac{\delta E[\rho_{\mathrm{tot}}]}{\delta \rho_{\mathrm{tot}}} - \frac{\delta E[\rho_{\mathrm{sub}}]}{\delta \rho_{\mathrm{sub}}} \tag{2.47}$$

The energy functional $E[\rho]$ contains contributions of the kinetic energy (T_s), the attraction between the electrons and the nuclei, as well as the Hartree (J) and exchange-correlation energy (XC) between electrons. Therefore, the embedding potential V_emb can be decomposed as:

$$V_\mathrm{emb} = V_\mathrm{emb}^{T_\mathrm{s}} + V_\mathrm{emb}^\mathrm{nucl} + V_\mathrm{emb}^J + V_\mathrm{emb}^\mathrm{xc} \tag{2.48}$$

The last three terms can be calculated easily in Kohn-Sham (KS) DFT, as their explicit functional forms are known:

$$V_\mathrm{emb}^{\mathrm{nucl}+J+\mathrm{xc}} = V_\mathrm{DFT}[\rho_\mathrm{tot}] - V_\mathrm{DFT}[\rho_\mathrm{sub}] \tag{2.49}$$

But the contribution from kinetic energy

$$V_\mathrm{emb}^{T_\mathrm{s}} = \frac{\delta T_\mathrm{s}[\rho_\mathrm{tot}]}{\delta \rho_\mathrm{tot}} - \frac{\delta T_\mathrm{s}[\rho_\mathrm{sub}]}{\delta \rho_\mathrm{sub}} \tag{2.50}$$

is hard to evaluate and requires special treatments, because the functional form of T_s is unknown (KS-DFT uses KS orbitals to calculate the kinetic energy).

2.4.1 Example 1: Optimized Effective Potential-Based Method

Carter et. al. showed that V_emb can be calculated numerically by solving an optimized effective potential (OEP) problem[68, 72]. Note that for a given embedding potential, the electron density of a subsystem ρ_K ($K = \mathrm{sub},\ \mathrm{env}$) can be calculated variationally by minimizing the subsystem energy with the presence of V_emb:

$$\tilde{E}[\rho_K] = E[\rho_K] + \int \rho_K V_\mathrm{emb}\, \mathrm{d}\mathbf{r} \tag{2.51}$$

$$\frac{\delta \tilde{E}[\rho_K]}{\delta \rho_K} = 0 \tag{2.52}$$

The correct embedding potential should also satisfies the constraint that the subsystem densities add up to the total electron density of the unpartitioned system (denoted as ρ_ref):

$$\rho_\mathrm{sub}[V_\mathrm{emb}] + \rho_\mathrm{env}[V_\mathrm{emb}] = \rho_\mathrm{ref} \tag{2.53}$$

Another constraint is that V_emb is the same for both subsystems. This constraint is essential to assure the uniqueness of the solution[68].

To solve V_{emb} efficiently, Carter et. al. defined the following Wu-Yang functional[68, 72]:

$$W[V_{\mathrm{emb}}] = \tilde{E}[\rho_{\mathrm{sub}}] + \tilde{E}[\rho_{\mathrm{env}}] - \int \rho_{\mathrm{ref}} V_{\mathrm{emb}} \, \mathrm{d}\mathbf{r} \tag{2.54}$$

where $\tilde{E}[\rho_K]$'s are calculated self-consistently. This functional has a unique maximum point. At that point:

$$\frac{\delta W[V_{\mathrm{emb}}]}{\delta V_{\mathrm{emb}}} = \rho_{\mathrm{sub}} + \rho_{\mathrm{env}} - \rho_{\mathrm{ref}} = 0 \tag{2.55}$$

and Eq.2.53 is automatically satisfied. With this property, V_{emb} can be solved by maximizing $W[V_{\mathrm{emb}}]$. In practice, this procedure is performed numerically in a self-consistent manner. We start with a guess of V_{emb} (generally set $V_{\mathrm{emb,init}} = 0$), calculate the derivative $\rho_{\mathrm{sub}} + \rho_{\mathrm{env}} - \rho_{\mathrm{ref}}$ and update V_{emb} accordingly until Eq.2.55 satisfies.

Finally, with the optimized V_{emb}, the hybrid QM/QM (mostly correlated wave function, CW in DFT) energy can be calculated in the subtractive scheme[68]:

$$E_{\mathrm{tot}}^{\mathrm{CW/DFT}} = \langle \Psi_{\mathrm{sub}}^{\mathrm{CW}} | \hat{H}_{\mathrm{sub}}^{\mathrm{emb}} | \Psi_{\mathrm{sub}}^{\mathrm{CW}} \rangle - \tilde{E}[\rho_{\mathrm{sub}}^{\mathrm{DFT}}] + E[\rho_{\mathrm{tot}}^{\mathrm{DFT}}] \tag{2.56}$$

In $\hat{H}_{\mathrm{sub}}^{\mathrm{emb}}$, the embedding potential is added to the one-electron core Hamiltonian:

$$\hat{h}_{\mathrm{sub}}^{\mathrm{emb}} = \hat{h}_{\mathrm{sub}} + V_{\mathrm{emb}} \tag{2.57}$$

2.4.2 Example 2: Projection-Based Method

The OEP-based calculation of V_{emb} is numerically exact but requires iterated DFT calculations during the optimization. Miller et. al. developed projection-based methods to avoid this complication[70, 73]. Their method based on the fact that the kinetic energy is additive if the orbitals of different subsystems are mutually orthogonal:

$$T_{\mathrm{s,tot}} = \sum_K T_{\mathrm{s},K} \tag{2.58}$$

if $\langle \psi_a^K | \psi_b^{K'} \rangle = \delta_{KK'} \delta_{ab}$. As a result, if ρ_{sub} and ρ_{env} are constructed by mutually orthogonal orbitals, we have $T_{\mathrm{s}}[\rho_{\mathrm{tot}}] - T_{\mathrm{s}}[\rho_{\mathrm{sub}}] - T_{\mathrm{s}}[\rho_{\mathrm{env}}] = 0$ and $V_{\mathrm{emb}}^{T_{\mathrm{s}}}$ vanishes. Practically, a KS-DFT calculation is performed for the total system and the resulted MOs (KS orbitals) are localized.

These LMOs are then partitioned into two sets $\{\psi_a^{\rm sub}\}$ and $\{\psi_b^{\rm env}\}$ corresponding to the model subsystem and the environment respectively. Density matrices for the subsystems $\gamma^{\rm sub}$ and $\gamma^{\rm env}$ can be constructed accordingly. Since the LMOs are mutually orthogonal, the embedding potential can be simply constructed as:

$$V_{\rm emb}[\gamma^{\rm sub},\gamma^{\rm env}] = V_{\rm DFT}[\gamma^{\rm sub}+\gamma^{\rm env}] - V_{\rm DFT}[\gamma^{\rm sub}] \tag{2.59}$$

This embedding potential can then be used to construct the embedded Hamiltonian and Fock operators for the model subsystem[73]:

$$\hat{h}_{\rm sub}^{\rm emb} = \hat{h}_{\rm sub} + V_{\rm emb}[\gamma^{\rm sub},\gamma^{\rm env}] + \mu\hat{P}_{\rm env} \tag{2.60}$$

$$\hat{f}_{\rm sub}^{\rm emb} = \hat{h}_{\rm sub}^{\rm emb} + \hat{g}_{\rm sub}[\tilde{\gamma}^{\rm sub}] \tag{2.61}$$

where $\hat{h}_{\rm sub}$ is standard one-electron core Hamiltonian for the model subsystem, $\hat{g}_{\rm sub}$ is the operator for all two-electron terms (Hartree and XC operator for DFT and Coulomb and exchange operator for HF). The eigenstates of the embedded Fock operator are $\{\tilde{\psi}_a^{\rm sub}\}$ and $\tilde{\gamma}^{\rm sub}$ is the corresponding density matrix. To enforce the orthogonality between $\{\tilde{\psi}_a^{\rm sub}\}$ and $\{\psi_b^{\rm env}\}$ and enable the use of Eq.2.59, an level-shift operator $\mu\hat{P}_{\rm env}$ is added to the embedded core Hamiltonian[70]:

$$\mu\hat{P}_{\rm env} = \mu\sum_b |\psi_b^{\rm env}\rangle\langle\psi_b^{\rm env}| \tag{2.62}$$

This operator shifts the energy levels of $\{\psi_b^{\rm env}\}$ by a high energy μ (generally $\mu = 10^6$ a.u.) thus prohibits their hybridization with $\{\tilde{\psi}_a^{\rm sub}\}$. At the limit $\mu\to\infty$ and the same functional is used in $\hat{g}_{\rm sub}$, $\tilde{\gamma}^{\rm sub}$ will recover $\gamma^{\rm sub}$, makes this DFT-in-DFT embedding scheme exact[70, 73]. Similar to the OEP-based methods, if a different DFT functional or a CW method is used for the embedded subsystem, the total energy can be calculated in the subtractive scheme:

$$E_{\rm tot}^{\rm CW/DFT} = \langle\tilde{\Psi}_{\rm sub}^{\rm CW}|\hat{H}_{\rm sub}^{\rm emb}|\tilde{\Psi}_{\rm sub}^{\rm CW}\rangle - (E[\gamma^{\rm sub}] + {\rm tr}[\gamma^{\rm sub}V_{\rm emb}]) + E[\gamma^{\rm tot}] + \mu{\rm tr}[\tilde{\gamma}^{\rm sub}\hat{P}_{\rm env}] \tag{2.63}$$

The last term $\mu{\rm tr}[\tilde{\gamma}^{\rm sub}\hat{P}_{\rm env}]$ is a perturbative correction for the finiteness of μ in real calculations[73].

Chapter 3

Theoretical Background: Statistical Mechanics

3.1 Ensemble and Partition Function

Statistical mechanics studies macroscopic systems that contains $N \sim 10^{23}$ microscopic particles. The system can experience numerous microscopic states $\{\nu\}$. The internal energy E of the system is a weighted average of the energy E_ν of a specific state ν, over all possible states. This average is called the "ensemble average" and is denoted using angle brackets $\langle ... \rangle$.

$$\langle E \rangle = \sum_\nu p_\nu E_\nu \tag{3.1}$$

An "ensemble" is defined as a collection of all possible microstates of the system consistent with macroscopic constraints[74]. The ensembles are characterized by its macroscopic constraints. The most commonly encountered ensemble is the "canonical ensemble", and the constraint is that the number of particles N, the volume V and temperature T of the system are fixed. In the canonical ensemble, the probability distribution of states ν is:

$$p_\nu = \frac{1}{Q} e^{-\beta E_\nu} \tag{3.2}$$

where $\beta = 1/k_B T$ is the inverse temperature, and the normalization factor

$$Q(N, V, T) = \sum_\nu e^{-\beta E_\nu} \tag{3.3}$$

is called the "partition function". Q is a function of the macroscopic constraints N, V, T. The partition function is linked to the free energy of the system, for the canonical ensemble case, this is the Helmholtz free energy F:

$$F = -k_B T \ln Q \tag{3.4}$$

In Eq. 3.3 the system states are described by the quantum number ν. In classical mechanics, the states are described by the positions and momenta of all the particles with in $\{\mathbf{r}^N, \mathbf{p}^N\}$. For a

system of N identical particles, the classical partition function is:

$$Q_{\mathrm{cl}}(N,V,T) = \frac{1}{N!h^{3N}} \int \mathrm{d}\mathbf{r}^N \mathrm{d}\mathbf{p}^N e^{-\beta H(\mathbf{r}^N,\mathbf{p}^N)} \tag{3.5}$$

where h is the Planck constant, the factor $N!$ is used to account for the indistriguishability of microscopic particles, h^{3N} can be considered as the "elementary volume" in phase space.

$$H(\mathbf{r}^N,\mathbf{p}^N) = \sum_i^N \frac{p_i^2}{2m} + U(\mathbf{r}^N) \tag{3.6}$$

is the Hamiltonian of the system. The integration of momenta can be calculated analytically, which gives rise to:

$$Q_{\mathrm{cl}}(N,V,T) = \frac{1}{N!\Lambda^{3N}} \int_V \mathrm{d}\mathbf{r}^N e^{-\beta U(\mathbf{r}^N)} \tag{3.7}$$

where

$$\Lambda = \frac{h}{\sqrt{2\pi m k_B T}} \tag{3.8}$$

is called the thermal de Broglie wavelength.

3.2 Entropy of Ideal Gas

For a system with identical noninteracting molecules, each molecule has energy levels $\{\varepsilon_\nu\}$. The energy of the total system is the summation of the molecular energy:

$$E_{\nu_1,\nu_2,\ldots,\nu_N} = \varepsilon_{\nu_1} + \varepsilon_{\nu_2} + \ldots + \varepsilon_{\nu_N} \tag{3.9}$$

The states of the total system are now described by the quantum number of each individual molecule. The partition function is:

$$\begin{aligned} Q &= \frac{1}{N!} \sum_{\nu_i,\nu_j,\ldots,\nu_N} e^{-\beta(\varepsilon_{\nu_i}+\varepsilon_{\nu_j}+\ldots+\varepsilon_{\nu_N})} \\ &= \frac{1}{N!} \left(\sum_\nu e^{-\beta\varepsilon_\nu} \right)^N \\ &= \frac{q^N}{N!} \end{aligned} \tag{3.10}$$

where

$$q = \sum_{\nu} e^{-\beta\varepsilon_\nu} \tag{3.11}$$

is the partition function of a single molecule, and the factor $N!$ accounts for the indistriguishability.

The molecular energy ε can be further decomposed into translational, rotational and vibarational energies (assume that the molecule stays at its electronic ground state):

$$\varepsilon_{\nu_t,\nu_r,\nu_v} = \varepsilon^{\text{trans}}_{\nu_t} + \varepsilon^{\text{rot}}_{\nu_r} + \varepsilon^{\text{vib}}_{\nu_v} \tag{3.12}$$

and the single particle partition function can also be decomposed into these components:

$$q = q_{\text{trans}}\, q_{\text{rot}}\, q_{\text{vib}} \tag{3.13}$$

For a polyatomic ideal gas molecule, the particle-in-box, rigid rotor and harmonic oscillator models can be applied for evaluating translational, rotational and vibrational parition function, respectively. The results are summarized below.

$$Q_{\text{trans}} = \frac{q^N_{\text{trans}}}{N!} = \frac{V^N}{N!\Lambda^{3N}} \tag{3.14}$$

$$Q_{\text{rot}} = q^N_{\text{rot}} = \frac{\pi^{N/2}}{\sigma^N}\left(\frac{T^3}{\Theta_A\Theta_B\Theta_C}\right)^{N/2} \tag{3.15}$$

$$Q_{\text{vib}} = q^N_{\text{vib}} = \left(\prod_{i=1}^{\alpha}\frac{e^{-\beta\hbar\omega_i/2}}{1-e^{-\beta\hbar\omega_i}}\right)^N \tag{3.16}$$

For linear molecules,

$$Q_{\text{rot,linear}} = \left(\frac{T}{\sigma\Theta_r}\right)^N \tag{3.17}$$

where $\Theta_r = \frac{h^2}{8\pi^2 k_B I}$ is the characteristic temperature of rotation, I is the principal moment of inertia, A, B, C denotes the principal axes of rotation. σ is the rotational symmetry number, which is the number of ways the molecule rotate back to itself, due to the indistinguishability. For water molecule ($C_{2\text{v}}$ symmetry), for example, $\sigma = 2$. ω_i is the vibrational frequency of the ith normal mode, $\alpha = 3N - 5$ for linear molecules and $3N - 6$ for nonlinear molecules.

The entropy of the ideal gases is of particular interests for the adsorption calculations. One can get the entropy from the parittion function via:

$$S = -\left(\frac{\partial F}{\partial T}\right)_{N,V} = k_B \ln Q - \frac{1}{T}\left(\frac{\partial \ln Q}{\partial \beta}\right)_{N,V} \tag{3.18}$$

The translatioanal, rotational and vibrational entropy of ideal gas are:

$$S_{\text{trans}} = N k_B \ln\left(\frac{V e^{5/2}}{N\Lambda^3}\right) \tag{3.19}$$

$$S_{\text{rot}} = N k_B \ln\left[\frac{\pi^{1/2} e^{3/2}}{\sigma}\left(\frac{T^3}{\Theta_A \Theta_B \Theta_C}\right)^{1/2}\right] \tag{3.20}$$

$$S_{\text{vib}} = N k_B \sum_{i=1}^{\alpha}\left[\frac{\beta\hbar\omega_i}{e^{\beta\hbar\omega_i}-1} - \ln\left(1 - e^{-\beta\hbar\omega_i}\right)\right] \tag{3.21}$$

For linear molecules,

$$S_{\text{rot,linear}} = N k_B \ln\left(\frac{eT}{\sigma\Theta_r}\right) \tag{3.22}$$

And the total entropy is:

$$S_{\text{tot}} = S_{\text{trans}} + S_{\text{rot}} + S_{\text{vib}} \tag{3.23}$$

Eqs. 3.19$-$3.23 can be used to calculate the entropy of adsorbates in the gas phase and Eq. 3.21 can also be used to calculate the entropy of solids, like the adsorbent and the adsorbed complex. The entropy change upon adsorption can then be calculated accordingly.

3.3 Langmuir Adsorption Isotherm

The adsorption of gases in a MOF can be characterized by the adsorption isotherm. At a given temperature, the isotherm describes the amount of adsorbate loading n in MOF as a function of pressure p.

$$n = n(p) \tag{3.24}$$

The Langmuir isotherm is one of the simplest formula.

Consider a system of one gas molecule interact with the adsorbent (the MOF), the partition function of the system is:

$$Q = \frac{1}{\Lambda^3} \int_V e^{-\beta U(\mathbf{r})} \, \mathrm{d}\mathbf{r} \tag{3.25}$$

where $U(\mathbf{r})$ is the interaction energy between the adsorbate and the MOF, Λ is the thermal de Broglie wavelength of the molecule. The integration is over the entire space. Assume the adsorbent surface can be divided to identical "sites", then the partition function can be written as:

$$Q = \sum_i^N \frac{1}{\Lambda^3} \int_{V_{\text{site}}} e^{-\beta U(\mathbf{r})} \, \mathrm{d}\mathbf{r} = \sum_i^N q_{\text{site}} = N q_{\text{site}} \tag{3.26}$$

where

$$q_{\text{site}} = \frac{1}{\Lambda^3} \int_{V_{\text{site}}} e^{-\beta U(\mathbf{r})} \, \mathrm{d}\mathbf{r} = e^{-\beta \mu_{\text{site}}} \tag{3.27}$$

is the partition function of the adsorbate bind to one site i. The integration is over the subspace correspond to the site, and μ_{site} is thus the free energy or chemical potential of the adsorbate on one site. The pre-factor N in Eq. 3.26 indicates that there are N possible ways to put the adsorbate on the total N sites.

For the system that contains m identical, non-interacting adsorbates, there are $\binom{N}{m} = \frac{N!}{m!(N-m)!}$ ways for them to occupy the total N sites, the partition function of the system is:

$$Q_m = \frac{N!}{m!(N-m)!} e^{-\beta m \mu_{\text{site}}} \tag{3.28}$$

The chemical potential of the adsorbate in the adsorbed phase can be calculated as the free energy change of inserting a new molecule into the system:

$$\mu_{\text{ad}}(\theta) = -k_B T \ln \frac{Q_{m+1}}{Q_m} = \mu_{\text{site}} - k_B T \ln \frac{1-\theta}{\theta} \tag{3.29}$$

where $\theta = \frac{m}{N} \simeq \frac{m+1}{N}$ is the coverage of the adsorbent. The last term in Eq. 3.29 is the configurational entropy. $\mu_{\text{ad}}(\theta)$ is a function of the coverage. Note that this chemical potential relation Eq. 3.29 is valid in both constant V and constant p conditions.

When the adsorbed phase is in equilibrium with the gas phase at pressure p, the chemical potentials of the adsorbates are equal:

$$\mu_{\text{ad}}(\theta) = \mu_{\text{g}}(p) = \mu_{\text{g}}^{\ominus} + k_B T \ln \frac{p}{p^{\ominus}} \tag{3.30}$$

Substitute Eq. 3.29 into Eq. 3.30 one can get the Langmuir isotherm:

$$\theta = \frac{Kp}{1 + Kp} \tag{3.31}$$

We have assumed that the adsorption sites are identical and the adsorbates do not interact with each other. In Eq. 3.31,

$$K = \frac{1}{p^{\ominus}} \exp(-\beta \Delta G_{\text{bind}}) \tag{3.32}$$

is the equilibrium constant of binding to the site and

$$\Delta G_{\text{bind}} = \mu_{\text{site}} - \mu_g^{\ominus} \tag{3.33}$$

is the binding free energy to the site. For the gas-solid system, ΔG_{bind} can be easily calculated from the binding energy and the entropy loss, the later can be obtained from vibrational analysis as described earlier:

$$\Delta G_{\text{bind}} = \Delta E_{\text{bind}} - T\Delta S_{\text{bind}} - k_B T \tag{3.34}$$

In principle, the adsorbent can have more than one kind of sites, Eq. 3.31 can be generalized to the multisite case. The total amount of gas loading on the adsorbent at a given pressure p is:

$$n(p) = \sum_s \frac{c_s K_s p}{1 + K_s p} \tag{3.35}$$

The summation is over all kinds of sites. $K_s = \exp(-\beta \Delta G_s)$ is the equilibrium constant of binding. c_s is the concentration of site kind s in the MOF, which can be calculates by $c_s = g_s/M$, where g_s is the number of sites in a unit cell and M is the molar mass of the unit cell. We will refer to g_s as the "degeneracy" of the sites. This multisite Langmuir model provides a way to calculate adsorption isotherms from first principles[75–77].

Generally, the interaction between the adsorbates can only be neglected at the low loading (zero pressure) limit. For such cases, the adsorption properties can be characterized by the Henry's constant and corresponding adsorption free energy:

$$K_{\text{H}} = \sum_s c_s K_s = \sum_s \frac{c_s}{p^{\ominus}} \exp(-\beta \Delta G_s) \tag{3.36}$$

$$\Delta G^{\ominus}_{\mathrm{ad}} = -k_B T \ln \frac{K_{\mathrm{H}} p^{\ominus}}{c^{\ominus}} \tag{3.37}$$

Note that from Eq. 3.26 and Eq. 3.33 the partition function of one adsorbate in the adsorbed phase can also be written as:

$$Q = Q_{\mathrm{id}} \sum_s g_s e^{-\beta \Delta G_s} \tag{3.38}$$

where Q_{id} is the ideal gas partition function of the adsorbate. Clearly, Henry's constant is proportional to the parition function of the system. Calculating the Henry's constant requires evaluation of the total partition function or the "absolute free energy", which poses very stringent accuracy requirement. Systematic error could be observed when use a less accurate long-range interaction treatment or an incomplete set of sites. On the other hand, the selectivity between two adsorbates:

$$S_{12}(p \to 0) = \frac{K_{\mathrm{H},1}}{K_{\mathrm{H},2}} \tag{3.39}$$

is a "relative free energy" and thus has much alleviated accuracy requirement.

The isosteric heat, or the enthalpy of adsorption, can be calculated as a weighted average of the binding energy on each sites plus a pV term:

$$\Delta H_{\mathrm{ad}} = \sum_s p_s \Delta E_s - k_B T \tag{3.40}$$

where

$$p_s = \frac{g_s e^{-\beta \Delta G_s}}{\sum_t g_t e^{-\beta \Delta G_t}} \tag{3.41}$$

is the probability of finding the adsorbate on site kind s. Similar to the selectivity, the probability distribution is also a "relative free energy" and requires looser modeling accuracy.

3.4 Transition State Theory

Transition state theory (TST) is a powerful tool to estimate the rate constant of elementary reactions from a microscopic perspective. Suppose we have an elementary reaction $\mathrm{A} + \mathrm{B} \to \mathrm{C} + \mathrm{D}$, the potential energy surface (PES) of the system has two local minima, corresponding to the reactant state ($\mathrm{A} + \mathrm{B}$) and product state ($\mathrm{C} + \mathrm{D}$) respectively. These two local minima are separated by a saddle point called the "transition state", denoted as $\mathrm{AB}^{\ddagger}$. The reaction path starts

from the reactant well, through the transition state and ends at the product well. This path can be symbolically written as:

$$\mathrm{A} + \mathrm{B} \rightarrow \mathrm{AB}^{\ddagger} \rightarrow \mathrm{C} + \mathrm{D} \tag{3.42}$$

TST assumes that the conversion from the transition state $\mathrm{AB}^{\ddagger}$ to the product state is driven by the vibration along the reaction path, with a characteristic frequency ν. The reaction rate can thus be written as:

$$r = \frac{\mathrm{d}[\mathrm{C}]}{\mathrm{d}t} = \nu[\mathrm{AB}^{\ddagger}] \tag{3.43}$$

where $[\mathrm{X}]$ denotes the concentration of species X. To estimate $[\mathrm{AB}^{\ddagger}]$, TST further assumes a quasi-equilibrium between the reactant state and the transition state:

$$\frac{[\mathrm{AB}^{\ddagger}]}{[\mathrm{A}][\mathrm{B}]} = K^{\ddagger} = \frac{q_{\mathrm{AB}^{\ddagger}}}{q_{\mathrm{A}} q_{\mathrm{B}}} \tag{3.44}$$

Therefore, the reaction rate is

$$r = \nu K^{\ddagger}[\mathrm{A}][\mathrm{B}] = k[\mathrm{A}][\mathrm{B}] \tag{3.45}$$

and the rate constant can be expressed as

$$k = \nu K^{\ddagger} = \nu \frac{q_{\mathrm{AB}^{\ddagger}}}{q_{\mathrm{A}} q_{\mathrm{B}}} \tag{3.46}$$

Recall that the single molecule partition function q has contributions from the electronic, vibrational, rotational degree of freedoms, etc. We can separate the vibrational contribution of the normal mode along the reaction path from the rest:

$$q_{\mathrm{AB}^{\ddagger}} = q'_{\mathrm{AB}^{\ddagger}} \cdot \frac{e^{-\beta h\nu/2}}{1 - e^{-\beta h\nu}} \simeq q'_{\mathrm{AB}^{\ddagger}} \frac{k_B T}{h} \tag{3.47}$$

The last equality is valid at the classical limit when $h \rightarrow 0$. The final expression of the rate constant is:

$$k = \frac{k_B T}{h} \frac{q'_{\mathrm{AB}^{\ddagger}}}{q_{\mathrm{A}} q_{\mathrm{B}}} = \frac{k_B T}{h} e^{-\beta \Delta F^{\ddagger}} \tag{3.48}$$

where $\Delta F^{\ddagger}$ is the activation free energy, which can be easily calculated computationally from the energy and the entropy difference between the reactant and transition states.

The only problem left is how to locate the transition state on the PES. Since the transition state is a saddle point, its Hessian matrix has one negative eigenvalue. The corresponding eigenvector

is the vibrational mode along the reaction path. In principle, the transition state can be found by optimization following that mode. However, the iterated calculation and diagonalization of Hessian matrix limits the use of this mode-following method for large molecules or plane wave based periodic calculations (where the analytical Hessian calculation is not available). For such systems, the nudged elastic band (NEB) method[78] is an efficient alternative for transition state searching.

3.4.1 Nudged Elastic Band Method

The NEB method requires the structure of the reactant and product states as the input. Several replicas of the system (called "images") are created along the reaction path via interpolating between the reactant and product states. The images are connected by "springs" to form an "elastic band"[78]. The spring forces act as constraints that hold the images at their positions along the path, avoiding the relaxation back to reactant or product states. The entire band is optimized, resulting not only the transition state but also the minimum energy path (MEP) that connect the reactant and the product.

Consider an elastic band constituted by N images. The configuration of the ith image is denoted as $\mathbf{R}_i$. $\mathbf{R}_1$ and $\mathbf{R}_N$ represents the reactant and product states respectively and are held fixed during the optimization. The tangent vector of the band at image i is denoted as $\hat{\tau}_i$. One simplest way to calculate $\hat{\tau}_i$ is

$$\hat{\tau}_i = \frac{\mathbf{R}_{i+1} - \mathbf{R}_{i-1}}{|\mathbf{R}_{i+1} - \mathbf{R}_{i-1}|} \tag{3.49}$$

Improved methods to calculate the tangent vector have also been developed[79]. The forces of each image can be decomposed to components parallel and perpendicular to the tangent vector. The perpendicular force is the perpendicular component of the energy gradient:

$$\mathbf{F}_i|_\perp = -\nabla E(\mathbf{R}_i)|_\perp = -\nabla E(\mathbf{R}_i) \cdot (1 - \hat{\tau}_i \hat{\tau}_i) \tag{3.50}$$

and the parallel force is the spring forces in a modified form[79]:

$$\mathbf{F}_i|_\parallel = k(|\mathbf{R}_{i+1} - \mathbf{R}_i| - |\mathbf{R}_i - \mathbf{R}_{i-1}|)\hat{\tau}_i \tag{3.51}$$

where k is the force constant of the springs. In the NEB force expression, the component of energy gradient parallel to the band is excluded, thus the relaxation along the reaction path is prohibited. The spring forces ensures that the N images are separated with equal space on the path. An MEP can be generated by minimizing the forces on each image.

The original NEB method discussed above can not guarantee that one of the image is exactly the saddle point, therefore may not produce the correct transition state structure. The climb image modification (CI-NEB)[80] can solve the problem. CI-NEB generally starts with a nearly converged NEB path. The image with the highest energy is identified. For this specific image, the parallel component of the force is replaced by the inverted parallel component of the energy gradient:

$$\mathbf{F}_{i,\max}\big|_{\parallel} = \nabla E(\mathbf{R}_i) \cdot \hat{\tau}_i \, \hat{\tau}_i \tag{3.52}$$

while the forces of other images remain unmodified. This force expression maximizes the energy along the reaction path and minimizes the energy on other directions. Therefore, the highest image of a converged CI-NEB calculation will be the correct transition state of that reaction path.

Chapter 4

Enabling Efficient and Accurate Computational Studies of MOF Reactivity via QM/MM and QM/QM Methods

This chapter is reproduced in part with permission from Cui, K.; Schmidt, J. R. Enabling Efficient and Accurate Computational Studies of MOF Reactivity via QM/MM and QM/QM Methods, *J. Phys. Chem. C* **2020**, 124, 10550-10560.

4.1 Introduction

First-principles calculations are powerful tools for modeling the reactivity of MOF systems. As we have reviewed in Chapter 1, density function theory (DFT) methods are ubiquitous in the MOF literature due to the favorable balance between accuracy and computational cost and the ability to cope with open-shell metal-containing MOF nodes, although correlated wavefunction (WF) based calculations on MOFs have also been utilized. [81–84]. Beyond the choice of electronic structure method, these calculations also differ in their choice of model system, using either periodic or cluster models to represent the bulk MOF, with varying advantages / disadvantages in each case. Periodic models naturally represent the infinite bulk MOF and the influence of the surrounding MOF environment, but are largely limited (for efficiency reasons) to DFT or Hartree-Fock (HF) methods. In addition, when examining defect formation and/or framework reactivity, the periodic nature of the calculation can also impose "concentration" effects that may require large supercells. Alternatively, one can utilize cluster models obtained by carving a finite MOF cluster from the bulk. Such a process often involves cutting metal-linker bonds, which may lead to charged clusters and truncation artifacts and thus results that depend sensitively on the choice of the cluster (vide infra). In addition, such cluster models necessarily omit the influence (both electronic and mechanical) induced by the surrounding MOF environment. However, cluster models are directly amenable to

treatment with accurate post-DFT (e.g. correlated wavefunction) methods using efficient atomic orbital-based codes.

Hybrid multi-scale methods can combine the advantages of both periodic and cluster model approaches. Such hybrid methods include both additive[63] and subtractive (e.g. ONIOM[64]) quantum mechanics / molecular mechanics (QM/MM) methods, as well as various DFT embedding approaches[68, 70, 72]. Such multi-scale methods have a long history in biophyisics, solution chemistry and catalysis, but with relatively few prior applications to MOF systems. In one of the first examples, Choomwattana et al. used ONIOM(B3LYP:UFF) and studied the mechanism of the carbonyl-ene reaction catalyzed by MOF-11.[85] In their model, the Cu paddlewheel was treated at the DFT level and the rest of the MOF cluster were treated using the universal force field (UFF) with a fixed geometry, yielding reaction and activation energies. Their results showed that the MOF catalyst reduced the activation energy by $\sim$10 kcal/mol. Several subsequent studies have used ONIOM(QM:MM) or ONIOM(QM:QM) methods to study catalytic reactions and/or adsorbate binding in MOFs.[86–91] However, the focus of this prior work was primary on the *application* of QM/MM techniques to MOFs as opposed to a detailed benchmarking of QM/MM approaches (vs. the "exact" periodic result), including the influence of the choice of cluster, QM/MM boundary, embedding method and/or force field (or low-level QM method).

In this chapter, we present a novel QM/MM scheme tailored for MOF systems. In the process, we also highlight the crucial role played by a number of key factors, including the influence of dative bond cleavage on the QM MOF cluster, and the long-range "mechanical" coupling of the cluster to the bulk MOF. We use point-defect formation in zeolitic imidizaolte frameworks (ZIFs) as a prototypical reaction to examine the accuracy and computational efficiency of our QM/MM method. We have two major objectives for this chapter. The first objective is to benchmark QM/MM methods as applied to MOF systems, both in terms of accuracy (i.e. fidelity to the periodic DFT potential energy surface) and computational efficiency, including single-point energy differences, geometry optimizations, and transition state searches. We apply this ONIOM(QM:MM) approach to the case of ZIF defect formation. In the process, we develop important general considerations for the creation of accurate MOF cluster models. Building on this work, our secondary

goal is to extend this multi-scale approach to beyond-DFT (e.g. coupled cluster) accuracy via ONIOM(QM:QM) or DFT embedding approaches, thus enabling unprecedented accuracy in bulk MOF calculations. Our benchmark calculations reveal important deficiencies in the DFT description of ZIF defects. We find that both ONIOM and DFT embedding enable "gold standard" correlated calculations with extremely affordable computational cost.

4.2 Computational Methods

4.2.1 Periodic calculations

All periodic DFT calculations were performed using the Vienna Atomistic Simulation Package (VASP)[92–95]. We adopted a similar computational protocol to Zhang et al. [44]. Briefly, all single point and geometry optimization calculations utilized the dispersion corrected PBE functional (PBE-D3) with a plane wave basis and a 600 eV energy cutoff, in conjunction with a projector augmented wave (PAW) treatment of core electrons[96–98]. The lattice constants were fixed to be the equilibrium values from defect-free ZIF-8. All atoms were relaxed until the forces are less than $0.01\,\mathrm{eV}\cdot\mathrm{\AA}^{-1}$. Transition states were obtained via nudged elastic band (NEB) methods[78, 80] with 10 images. The band was relaxed until the forces are less than $0.02\,\mathrm{eV}\cdot\mathrm{\AA}^{-1}$.

4.2.2 Cluster calculations

Cluster models were carved directly from the optimized periodic structures without any further optimizations. We built both small $Zn_2mIm_7^{3-}$ (7L) and large $Zn_8mIm_{24}^{8-}$ (24L) cluster model representations of ZIF-8. The details of the construction and the bond termination of the cluster models are discussed in a subsequent section. All cluster calculations were performed in the Molpro 2015 package[99]. Density fitting[50, 51] was used to accelerate all calculations (both DFT and wavefunction-based). For the coupled-cluster calculations, we used the local correlation approach (LCCSD and LCCSD(T)) developed by Werner and coworkers[47, 48]. The Boys localization scheme[53] was used to localize molecular orbitals for the LCCSD calculations.

We utilized Dunning-style correlation consistent (cc) basis sets[49] for all cluster calculations. For the single point calculations, unless otherwise specified, we used cc-pVTZ for Zn, C and H

and aug-cc-pVTZ for N and O for small 7L clusters, and cc-pVTZ for Zn, cc-pVDZ for C and H and aug-cc-pVDZ for N and O for larger 24L clusters. H-like ghost orbitals were added to the point charges (vide infra). Geometry optimizations employed a double-zeta basis (cc-pVDZ for Zn, C and H and aug-cc-pVDZ for N and O, and no ghost orbitals) for all cluster sizes. The Atomic Simulation Environment (ASE) [100] was used for the constrained optimization. All unconstrained atoms were relaxed until the forces are less than $0.01\,\mathrm{eV}\cdot\mathrm{\AA}^{-1}$.

4.2.3 ONIOM calculations

For the ONIOM calculations, we use the general notation ONIOM(high:low) to represent the two-layer ONIOM scheme, where high is the high-level theory and low is the low-level theory[64]. In the ONIOM(QM:MM) geometry optimization, we take the 7L cluster model as the QM region with the remainder of the system described via ZIF-FF[101], a flexible and transferable force field specially parameterized for ZIFs. The ONIOM(QM:MM) energy is calculated by (Eq. 4.1)

$$E_{\mathrm{ONIOM(QM:MM)}} = E_{\mathrm{sub}}^{\mathrm{QM}} - E_{\mathrm{sub}}^{\mathrm{MM}} + E_{\mathrm{tot}}^{\mathrm{MM}}. \tag{4.1}$$

Forces are generated via the analytic derivative of energy with respect to the atom positions.[64] We created a Python interface via the ASE framework[100] to perform the ONIOM calculations. Molpro was used for the QM calculations and the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS)[102] was used for all MM calculations. All unconstrained atoms were relaxed until the forces were less than $0.01\,\mathrm{eV}\cdot\mathrm{\AA}^{-1}$. NEB calculations were performed via the NEB module within ASE, using 10 images. The spring constant of the band was set to $5.0\,\mathrm{eV}\cdot\mathrm{\AA}^{-2}$. The band was relaxed using the FIRE[103] method until the forces on atoms are less than $0.02\,\mathrm{eV}\cdot\mathrm{\AA}^{-1}$.

In addition, the ONIOM(QM:QM) method was used to achieve a high-accuracy description of the defect formation reaction. The high-level theory was chosen to be LCCSD or LCCSD(T), and the lower-level was chosen to be PBE or PBE-D3.

$$E_{\mathrm{ONIOM(WF:DFT)}} = E_{\mathrm{sub}}^{\mathrm{WF}} - E_{\mathrm{sub}}^{\mathrm{DFT}} + E_{\mathrm{tot}}^{\mathrm{DFT}} \tag{4.2}$$

4.3 QM/MM schemes for MOF calculations

4.3.1 Validation of ZIF cluster models

Constructing cluster models for ZIFs (or any MOF) requires breaking metal-linker bonds, inevitably resulting in truncation effects. These truncation effects are at least two-fold in origin. On the one hand, breaking the dative metal-linker bonds disrupts the inherent charge transfer between the metal and linker. But in addition, the resulting ZIF cluster model often carries a net charge; due to the ZIF stoichiometry, no standard (symmetric) cluster model can achieve charge neutrality. This net charge may artificially influence the stability of the defect structures as compared to charge-neutral periodic models.

Table 4.1 shows the formation energy ΔE_{form} of a linker vacancy (LV) in ZIF-8 calculated using a series of $Zn_2mIm_7^{3-}$ cluster models (denoted as 7L models), with various terminations: a bare charged cluster ($7L^{3-}$), one with all dangling N atoms capped via Li^+ (7L-Li^{3+})[104], or capped by H^+ (i.e. imidazolate anions replaced by imidazole, 7L-H^{3+}). All results shown in this section were calculated at the PBE level to facilitate direct comparison with the corresponding "exact" periodic DFT result.

Using the charge-neutral periodic result as a benchmark, we find errors as large as 0.33 eV for these conventionally-terminated cluster models. Note that for the purpose of this benchmark calculation, all cluster models were carved from an optimized periodic structure whose geometry is fixed, eliminating the atomic positions as a possible source of error. Thus in this case, the differences in ΔE_{form} arise strictly from the cluster truncation/termination. Note that influence of different basis sets for periodic (plane waves, PWs) and cluster (Gaussian type orbitals, GTOs) models can be largely eliminated when using a basis set as large as aug-cc-pVTZ.

In order to construct a charge-neutral ZIF cluster model, we use point charges ($q = +0.5$) to cap all the dangling N atoms (structures are shown in Figure 4.1(a)). The magnitude of the charges ($+0.5$) is chosen to neutralize the net charge on the cluster (bare 7L cluster is -3, with 6 broken dative bonds). For these static/single-point calculations, we treat the point charges as external electric fields, i.e. the Coulomb repulsion between the point charges are excluded from the total

Model	Periodic	$7L^{3-}$	7L-Li^{3+}	7L-H^{3+}	7L-q	7L-q+GO
ΔE_{form}/eV	0.056	0.383	0.203	0.307	0.149	0.142
Error/eV	-	0.327	0.147	0.251	0.093	0.086

Table 4.1: Formation energy of a LV defect in ZIF-8 as calculated using various cluster models vs. periodic DFT.

energy of the model systems. The point charges are placed along the Zn-N bond with a fixed N-q distance $d_{\mathrm{N}-q}$:

$$\mathbf{r}_q = \mathbf{r}_{\mathrm{N}} + d_{\mathrm{N}-q}\frac{\mathbf{r}_{\mathrm{Zn}} - \mathbf{r}_{\mathrm{N}}}{|\mathbf{r}_{\mathrm{Zn}} - \mathbf{r}_{\mathrm{N}}|} \tag{4.3}$$

Our calculations show that the resulting defect formation energy does not depend strongly on the N-q distance; here, we use $d_{\mathrm{N}-q} = 1.0$ Å for simplicity. Previous work has shown that the capping of such dative bonds enables an accurate reproduction of the bulk electron density.[104] We denote the point charge capped cluster model as 7L-q model.

As shown in Table 4.1, and in contrast to the charged/conventionally-terminated cluster models, the neutral cluster models significantly reduce the error vs. the charge-neutral periodic benchmark, yielding agreement to within < 0.1 eV, even for these very small clusters. Adding H-like ghost orbitals (7L-q+GO) on the point charges provides a very small incremental further improvement in accuracy, as shown in the last column in Table 4.1. We utilize these charge-capped cluster models for our subsequent calculations.

Beyond electronic effects associated with truncation, cluster models also neglect any mechanical coupling to the surrounding bulk (e.g. lattice strain associated with defect formation). The simplest (and most naive) way to capture such mechanical influence is to increase the cluster size. Here we introduce a 24L-q model, expanding the size of the cluster to include an additional 6 Zn atoms and their coordinating imidazolates; the structure of 24L-q model is shown in Figure 4.1(b). We will subsequently abbreviate the 7L-q and 24L-q models as the 7L and 24L models.

Figure 4.2 compares the PBE-calculated reaction pathway associated with defect formation in ZIF-8, for periodic DFT vs. the 7L and 24L cluster models. Briefly, the reaction pathway

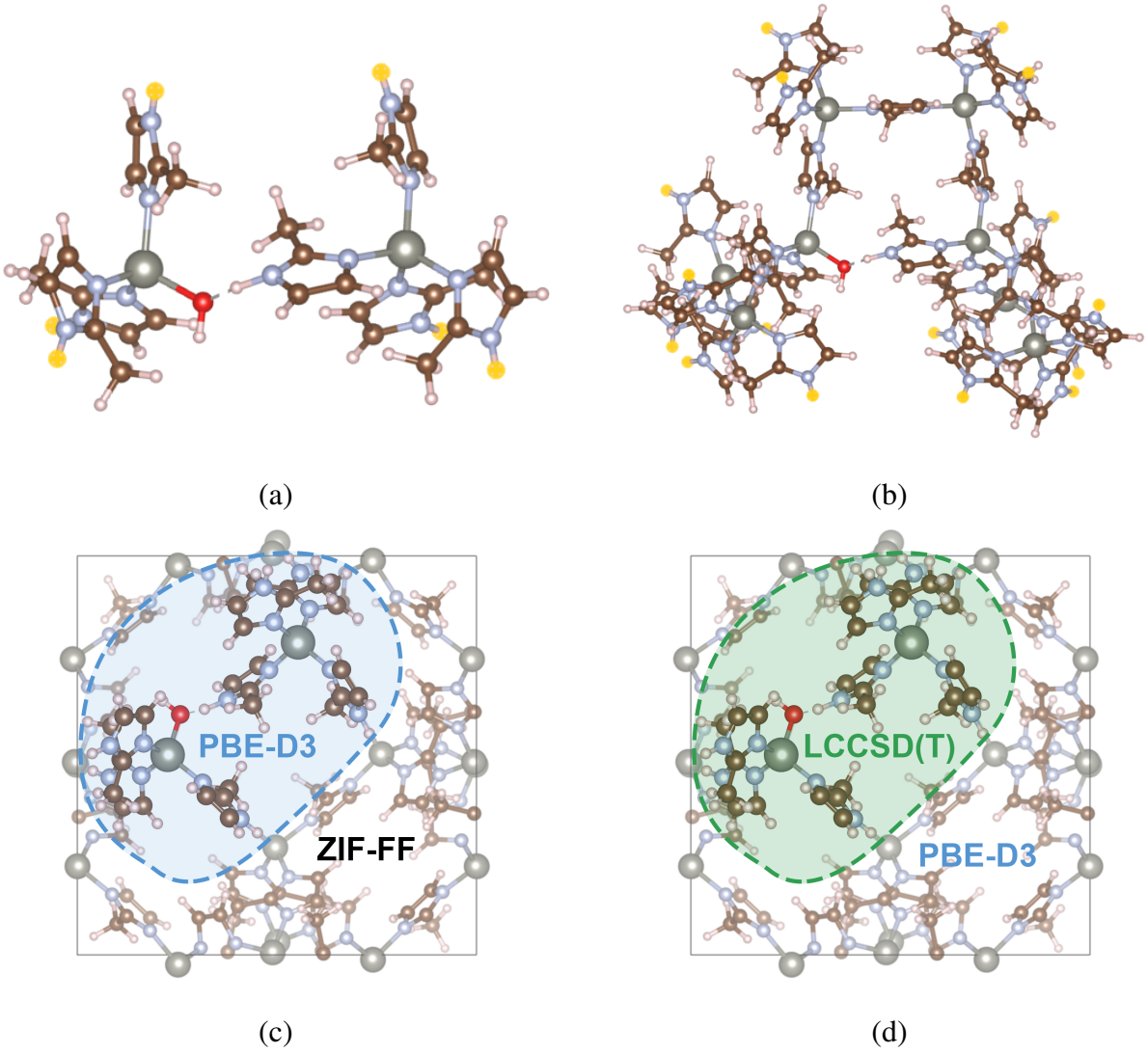


Figure 4.1: (a) 7L and (b) 24L cluster models for a DL defect in ZIF-8. The yellow points represent the point charges used to cap the dangling N atoms. Schematic illustration of the (c) QM/MM and (d) QM/QM system partition of the DL defect in ZIF-8, in conjunction with the 7L-in-periodic model.

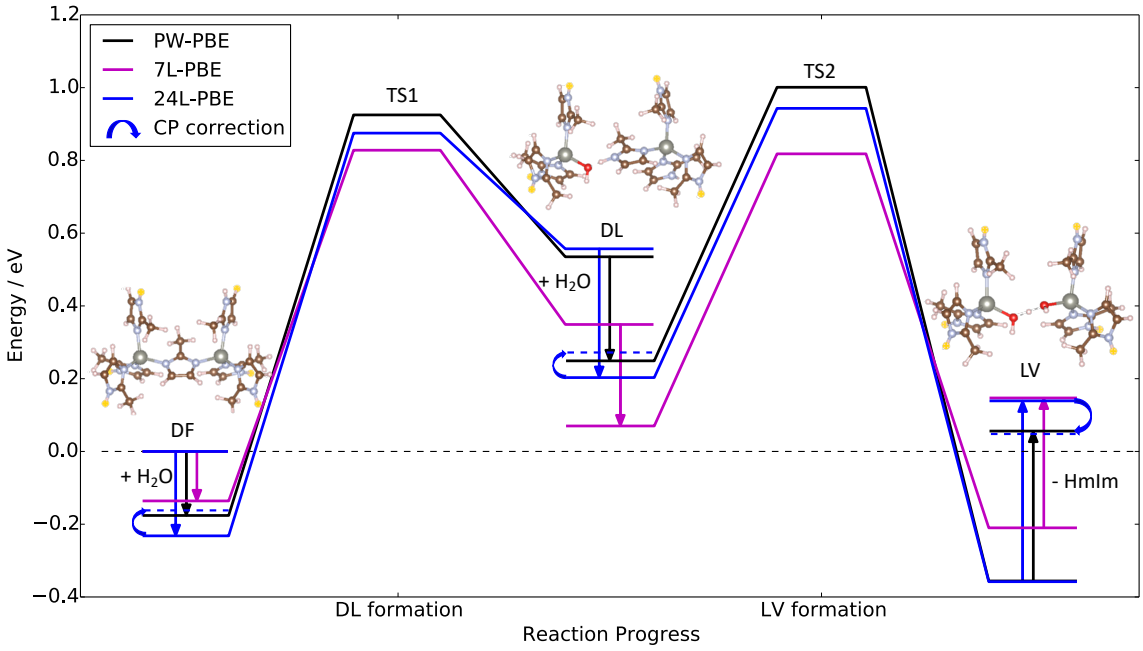


Figure 4.2: Reaction pathway for LV formation in ZIF-8 induced by water. The 7L cluster is calculated with a AVTZ basis set, while the larger 24L cluster utilizes a smaller AVDZ basis.

proceeds from pristine, defect-free ZIF-8, with subsequent water adsorption, and then dangling linker (DL) formation. Upon adsorption of an additional water molecule, a LV is formed, followed by desorption of imidazole into the gas phase; a schematic illustration of the defect formation mechanism is shown in Figure 1.2.

These results clearly demonstrate that the small 7L model omits the strain of the DL state, strongly underestimating the barrier and the reaction energy as compared to the periodic model. In contrast, the larger 24L model substantially improves upon that result and qualitatively correctly captures the mechanical coupling to the surrounding environment and thus the missing strain energy. Note that the larger 24L cluster utilizes a smaller AVDZ basis (vs. AVTZ for the 7L cluster) for efficiency reasons. As such, the 24L cluster substantial overestimates the imidazole desorption energy (final step in reaction pathway) due to the associated basis set superposition error (BSSE). A simple counterpoise correction essentially eliminates this BSSE, reducing the absolute error in the desorption energy vs. periodic DFT to 0.014 eV (curved blue arrows in Figure 4.2); similar effects are also present in the water adsorption steps, which are also strongly mitigated via counterpoise corrections.

Overall, these results show that the neutral charge-capped cluster model reproduce the energetics from the periodic calculations with high fidelity. Even using the larger 24L cluster model ($\sim$ 200 atoms), such calculations can surpass the efficiency of full periodic models, especially for ZIFs/MOFs with large unit cells. However, the utility of such models is somewhat limited by the need employ existing bulk geometries (presumably from XRD or prior periodic DFT) and to utilize large clusters (to account for mechanical coupling to the environment). These limitations are addressed in the subsequent sections.

4.3.2 Cluster-based geometry optimizations

All of the above calculations were conducted at a fixed geometry taken directly from previously optimized periodic models. However, for large unit cells, such PW DFT-based optimizations are extremely time consuming and/or infeasible. In contrast, cluster models contain far fewer atoms as compared to a large unit cells, offering the prospect for greatly increased efficiency.

We initially examined a geometry optimization scheme utilizing constrained geometry optimizations of a modestly-size cluster model. Starting from a defect-free 24L cluster model, we generated an initial guess for a defective cluster geometry (e.g. removing a linker), and subsequently optimized the defective cluster via a *constrained* optimization where the outside atoms (nitrogen and carbon atoms that are furthest from defect site) were fixed in their initial position (in the defect-free structure) while the inner atoms were allowed to fully relax. This procedure mimics the mechanical coupling to the surrounding bulk by constraining the position of the outer atoms, thus preventing unphysical relaxation of the cluster. Note that the initial defect-free structures can be easily obtained from experiment or from PW DFT (often using a smaller unit cell for the pristine vs. defective MOF).

Using DL and LV defect formation in ZIF-8 as examples, we utilize the root-mean-squared displacement (RMSD) and the error in the defect formation energies $\Delta\Delta E_{\mathrm{form}}$ to characterize the deviation of the optimized geometry from the periodic reference, as shown in the "24L DFT" section in Table 4.2. We find that the $\Delta\Delta E_{\mathrm{form}}$ for LV formation is smaller than 0.1 eV and the RMSD is below 0.2 Å, indicating the optimized structure is very similar to that obtained via PW DFT. However, for the DL case, the RMSD increases to 0.6 Å, indicating that the structure strongly deviates from the PW reference. This deviation is also reflected via the large $\Delta\Delta E_{\mathrm{form}}$. In the DL case, the steric bulk of the dangling imidazolate induces significant lattice strain. In the original PW DFT calculation, this strain is mitigated via a significant distortion in the outer linkers; such relaxation is prohibited within the current cluster scheme, since the outer atoms are held rigidly fixed, leading to an overestimation in the DL formation energy. In general, we expect that such a simple scheme will perform poorly for similar high-strain structures because a finite cluster model cannot properly account for the longer-range relaxation effects that are induced by large amounts of strain. We address this limitation below.

4.3.3 ONIOM(QM:MM)-based geometry optimization

Capturing such long-range distortions induced by high-strain defects would require an infeasibly large cluster model. (Incidentally, such distortions may also require the use of a supercell

Theory	Structure	RMSD[a]/Å	$\Delta\Delta E_{\mathrm{form}}$/eV	Timing[b]/min
periodic DFT	ALL	N/A	N/A	7.5
24L DFT[c]	DF	0.045	N/A	22
	DL	0.614	0.188	
	LV	0.188	0.080	
7L-in-62L ONIOM[d]	DF	0.104	N/A	1.5
	DL	0.307	-0.036	
	LV	0.109	0.137	
7L-in-periodic ONIOM	DF	0.080	N/A	1.5
	DL	0.145	-0.048	
	LV	0.123	0.113	

[a] RMSDs are calculated for the 7L subsystem. [b] Real time (including I/O) per ionic step on a 20 core node. [c] Fix outmost NC_2 and point charges. [d] Fix outmost N and point charges.

Table 4.2: Comparison of the cluster/ONIOM vs. PW DFT optimized structures of various defective ZIF-8 models and corresponding errors in the formation energies.

vs. unit cell even in the case of PW DFT!) An attractive alternative is to employ a multi-scale ONIOM(QM:MM) description of the ZIF. Since our prior calculations establish that a small 7L cluster model is sufficient to describe the electronic structure of the defect region, although not its mechanical coupling to the environment, we utilize a small 7L cluster for our QM region and treat the mechanical coupling to the bulk via an accurate MM force field that has been previously parameterized for ZIFs (ZIF-FF)[101]. A schematic illustration of the ONIOM(QM:MM) system partition is shown in Figure 4.1(c). Within this ONIOM framework, the terminating point charges in the QM model are treated as link atoms in order to yield well-defined forces; therefore, the Coulomb interaction between the point charges must be included in the total energy and force expressions. This treatment is similar to the conventional mechanical embedding ONIOM. The positions of the point charges are determined slightly differently for the optimization: they are placed with a fixed ratio ($g = 0.566$) to the Zn-N bond distance rather than a fixed length to the N atoms.

$$\mathbf{r}_q = \mathbf{r}_{\mathrm{N}} + g(\mathbf{r}_{\mathrm{Zn}} - \mathbf{r}_{\mathrm{N}}) \tag{4.4}$$

As we stated before, the calculated energy does not strongly depend on the positions of the point charges. With the ratio of $g = 0.566$ the point charges are placed around their equilibrium positions ($d_{\mathrm{N}-q} \simeq 1.131\,\text{Å}$) and the forces act on the point charges are essentially zero. In the MM calculations we thus omit the bond interactions involving point charges. Our test calculations showed that this treatment has negligible influence to the optimized structure and reaction energy, but will accelerate the convergence of the optimization. The charge on the point charges in the MM calculation equals a quarter of the charge on Zn^{2+}, which is +0.1768 when using ZIF-FF, to ensure neutrality.

It is vital to emphasize some of the important but subtle differences between the present ONION-based approach with standard mechanical embedding (ME) or electrostatic embedding (EE) schemes used for organic molecules. In the case of organic molecules, the subsystem boundary generally passes through *covalent* bonds, which typical terminated with H atoms (or other monovalent atoms). With EE, additional point charges may be added to reproduce the electrostatic interaction with the surrounding bulk. In contrast, for MOF systems, the boundary typically passes

through *dative* metal-linker bonds, and point charges are required to both ensure charge neutrality and to mimic the metal-ligand charge transfer inherent in dative bonding. As a result, and in contrast to standard approaches, all atoms are allowed interact with the capping point charges in our scheme.

We preformed ONIOM-based optimizations for all defect-free, DL and LV structures of ZIF-8, using either large (MM) 62L cluster (7L-in-62L) or periodic boundary conditions (7L-in-periodic). The calculated RMSD and $\Delta\Delta E_{\mathrm{form}}$ for these two cases are shown in Table 4.2. In the case of the 7L-in-periodic model, the ONIOM description yields a negligible $\Delta\Delta E_{\mathrm{form}}$ (-0.036 eV) for the DL defect and a larger (but acceptable) $\Delta\Delta E_{\mathrm{form}}$ (0.113 eV) for the LV. Note that no constraints are applied during the optimization. By contrasting the present results against those from the pure (un-embedded) 24L cluster model, we can clearly see the importance of eliminating unphysical constraints and including the *long-range* mechanical coupling with the environment; this influence is particularly pronounced in the DL case. We believe that the residue error in the LV case is primarily inherited from truncation effect of using the small 7L cluster models. This error can be reduced (to $\sim$0.07 eV) by increasing the size of QM region, for example, using a 24L-in-periodic model. The periodic model also yields accurate geometries for all cases (DF, DL, LV), with RMSDs around 0.1 Å. The 7L-in-62L model, which, similar to the bare cluster optimizations utilizes fixed outer atoms (within the MM region), generally yields larger geometric deviations but comparable $\Delta\Delta E_{\mathrm{form}}$. The larger RMSD may actually result from a defect concentration effect, with the 7L-in-62L model effectively representing the limit of infinitesimal defect concentrations.

Theory	Structure	RMSD[a]/Å	$\Delta\Delta E_{\mathrm{form}}$[b]/eV
7L-in-periodic ONIOM	DF	0.031	N/A
	DL	0.051	0.064 (0.008)

[a] RMSDs are calculated for the 7L subsystem.

Table 4.3: Comparison of the ONIOM vs. PW DFT optimized structures of DF and DL defect of ZIF-8 induced by H_2SO_4 and corresponding errors in the formation energies. Results utilize an AVDZ (AVTZ) basis set.

Table 4.3 shows an additional example application: DL formation in ZIF-8 induced by reaction with H_2SO_4; the FF for H_2SO_4 is taken from the work of Ding et al.[105] Employing the 7L-in-periodic model, we find a very small $\Delta\Delta E_{\mathrm{form}}$ and RMSD as compared to the original PW DFT calculation. The $\Delta\Delta E_{\mathrm{form}}$ can be further decreased to 0.008 eV by using larger triple zeta (AVTZ) basis set.

The totality of these results demonstrate that geometry optimizations using the ONIOM(QM:MM) method can provide essentially the same accuracy as PW DFT, but with much lower computational cost. For example, using a 20-core node, we find that one ionic step takes about 1.5 min for the ONIOM method and and 7.5 min for periodic DFT. The relative advantage of the ONIOM(QM:MM) method will be more pronounced when applying to alternative ZIFs (or other MOFs) with larger unit cells. For example, ZIF-71-RHO (4,5-dichloroimidazole linkers) has 816 atoms in its unit cell. In that case, a single ionic step requires nearly 6 hours when using PW DFT, while ONIOM(QM:MM) (due to the effective $O(1)$ scaling) takes 1.5$\sim$2 minutes. As such, ONIOM(QM:MM) seems extremely promising for the efficient computational modeling of MOF reactivity and defect formation, including the influence of lattice strain.

We note that the high-efficiency of the ONIOM(QM:MM) method can also dramatically aid in the location of transition states, which is often time consuming and computationally expensive with PW DFT. In Figure 4.3, we compare the calculated minimum energy path (MEP) for DL formation in ZIF-8 induced by reaction with water, calculated via both PW DFT and ONIOM(QM:MM). Here, we utilize the nudged elastic band (NEB) method to allow for direct comparison with the PW DFT-calculated MEP. Both results are plotted along a reaction coordinate defined as the difference between the Zn-N and Zn-O bond lengths, $d_{\mathrm{Zn-N}} - d_{\mathrm{Zn-O}}$. We can see that ONIOM(QM:MM) gives a nearly identical MEP as compared to PW DFT, with negligible errors in the energy barrier and RMSD. It is also important to note that the ONIOM(QM:MM) method allows for straightforward access to the Hessian matrix, facilitating the application of a wide range of advanced transition-state searching methods (in contrast to the original PW DFT).

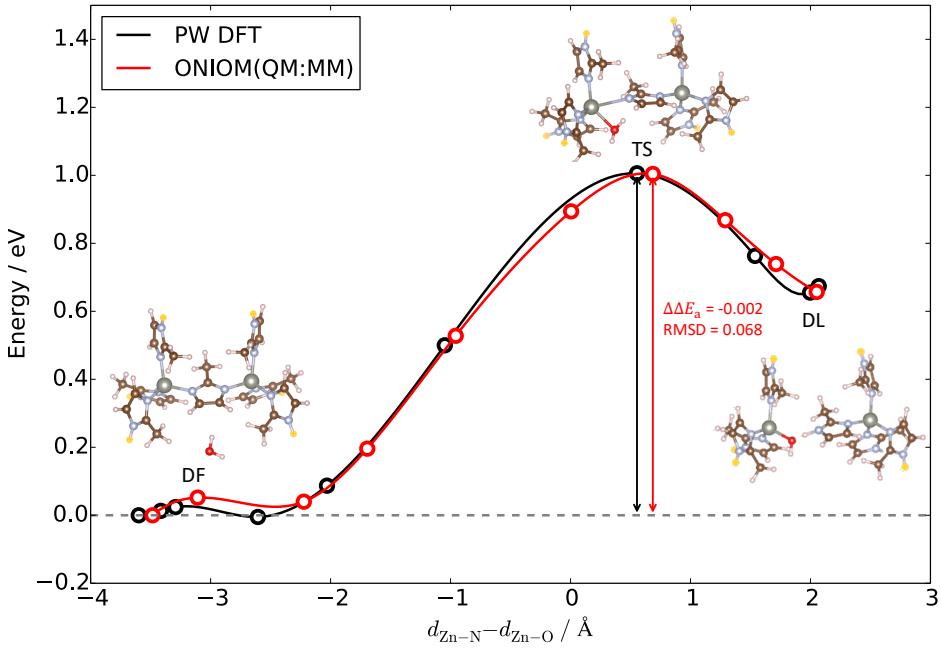


Figure 4.3: Minimum energy path of DL formation in ZIF-8 induced by water.

4.4 Beyond DFT-based QM/QM MOF calculations

4.4.1 Benchmarking DFT vs. wave function theories

The existence of accurate, validated cluster models for ZIFs opens the door to accurate beyond-DFT (correlated WF) calculations on ZIF reactivity and defect formation, including "gold-standard" CCSD(T) calculations. Crucially, we show that, in a number of important cases, WF methods yields qualitatively different predictions as compared to the ubiquitous DFT approaches that are utilized in almost all periodic calculations.

Table 4.4 shows the formation energy of a LV defect in ZIF-8 (calculated using the 7L cluster model) with a number of DFT and WF methods. The calculated defect formation energies span a striking range, from $\sim$0.1 eV (PBE) to as high as $\sim$0.9 eV (MP2), with "gold-standard" LCCSD(T) yielding 0.719 eV. Note that desorption of the imidazolate linker is an important step in the LV formation process. As such, an important fraction of this deviation is certainly due to neglect of dispersion, with e.g. PBE-D3 yielding a significant improvement over PBE (0.407 eV). However, even hybrid functionals appear to yield only incremental improvement over their pure GGA counterparts. In general, all DFT methods yield substantially lower defect formation energies as compared to correlated WF methods.

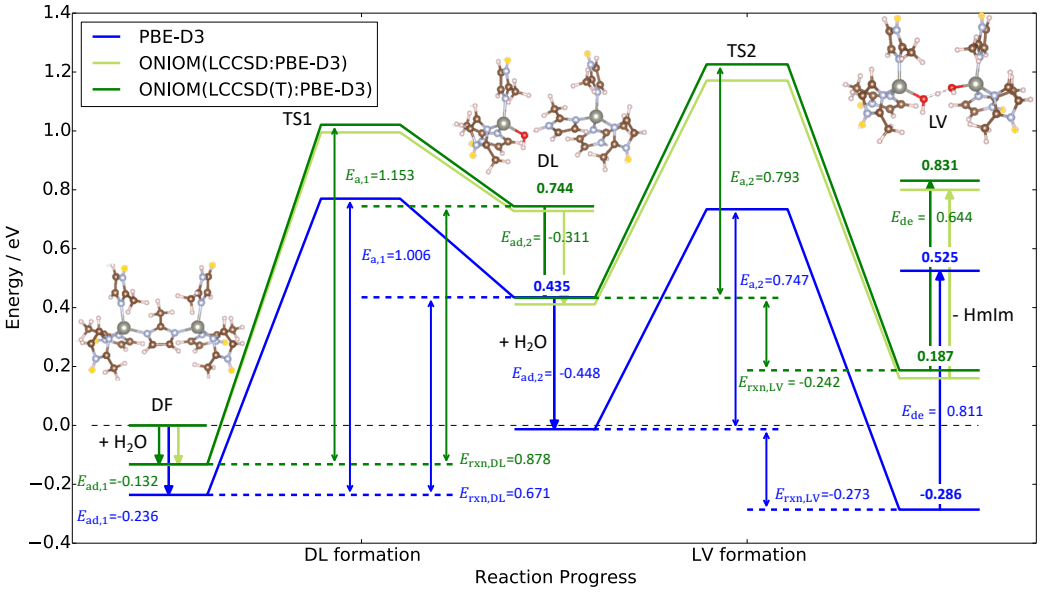
Taking the LCCSD(T) result as a benchmark, MP2 overestimates the defect formation energy (likely due to its well-known overestimation of dispersion). In addition, both the DFT and canonical MP2 results are subject to some BSSE; this BSSE is strong mitigated in local correlation methods,[106, 107] perhaps explaining the seemingly results of LMP2. However, the existence of BSSE also implies that the true discrepancies between the DFT and benchmark LCCSD(T) is likely even larger than suggested by this data.

4.4.2 High-accuracy calculations using hybrid QM/QM methods

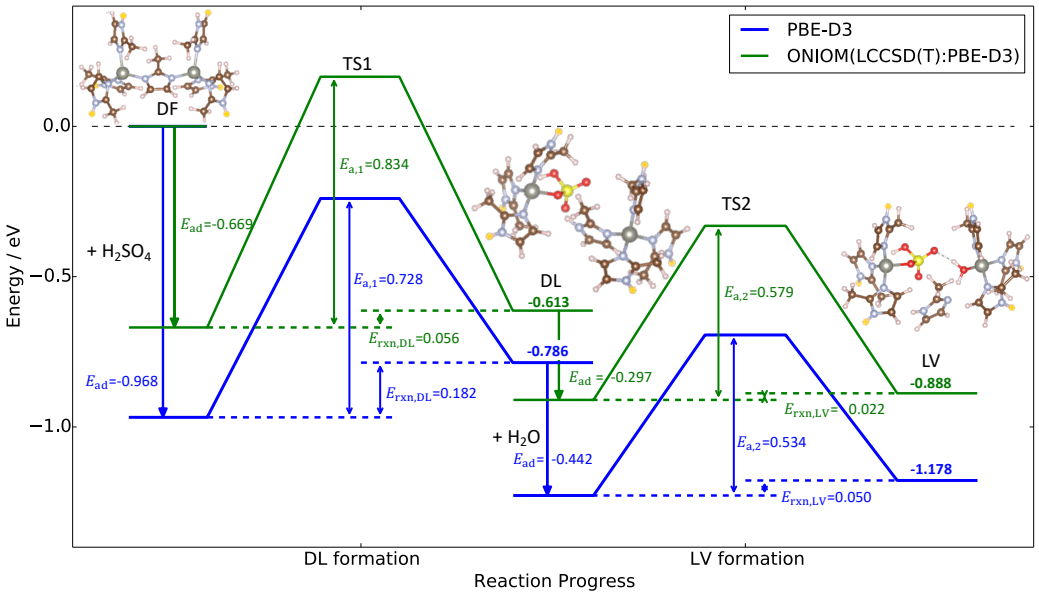
The above benchmark calculations highlight the importance of utilizing correlated WF methods (vs. DFT) when examining ZIF defect formation, although (since they utilized small cluster models) the prior results necessarily neglect the influence of lattice strain. In this subsection, we

DFT		WF	
Theory	ΔE_{form}/eV	Theory	ΔE_{form}/eV
PBE	0.147	MP2	0.907
PBE-D3	0.407	LMP2	0.742
PBE0	0.242	LCCSD	0.687
PBE0-D3	0.502	LCCSD(T)	0.719
B3LYP	0.092		
B3LYP-D3	0.478		

Table 4.4: Formation energy of a LV defect in ZIF-8 calculated at a number of DFT and WF levels of theory.



(a)



(b)

Figure 4.4: PES for LV formation in ZIF-8 induced by (a) water and (b) H_2SO_4 and water, calculated via ONIOM(LCCSD:PBE-D3) and ONIOM(LCCSD(T):PBE-D3) methods. Double zeta basis sets were used for 7L model. The energetics of each elementary steps are labeled in the figure (only for LCCSD(T) and PBE-D3 cases) and the bold numbers are formation energies of the defective stated compared to isolated defect-free ZIFs.

demonstrate how multi-scale ONIOM can be utilized to exploit correlated WF methods while retaining proper mechanical coupling to the surrounding bulk, thus enabling state-of-the-art accuracy in ZIF reaction energetics.

To enable the highest possible accuracy, we utilized a ONIOM(QM1:QM2) scheme, making using of correlated WF and DFT method as the "high" and "low" levels of theory, respectively; the ONIOM energy expression remains similar to (Eq. 4.1), replacing the MM by DFT. Note that, alternatively, an MM method could be used in place of DFT as the lower level, or a three-level method could also be utilized (e.g. WF:DFT:MM). We employed the 7L model as the QM1 layer and the full periodic unit cell as the QM2 layer. A schematic illustration of the ONIOM(QM1:QM2) system partition is shown in Figure 4.1(d).

Figure 4.4(a) shows the PES for LV defect formation in ZIF-8 induced by water as calculated via the ONIOM(LCCSD:PBE-D3) and ONIOM(LCCSD(T):PBE-D3) methods, compared against the PBE-D3 results from traditional PW DFT. Clearly, there are significant differences between the PBE-D3 and ONIOM results for all intermediates, while the differences between LCCSD and LCCSD(T) are minor. Compared to PBE-D3, the LCCSD method shows a significantly larger formation energy for both DL and LV defects (with differences as large as ~ 0.5 eV), and a higher reaction barrier (~ 0.15 eV) for DL formation. LCCSD and LCCSD(T) also yield a somewhat smaller adsorption energy as compared to PBE-D3, which can be at least partially attributed to residual BSSE in the PBE description of the ONIOM cluster energy (BSSE is largely absent in the LCCSD cluster energy and zero in the PW DFT component); the magnitude of this correction should be roughly equivalent to that shown Figure 4.2, but with opposite sign. Overall, LCCSD(T) predicts that defect formation is more endothermic and with higher effective barriers as compared to PBE-D3.

The corresponding acid-catalyzed (H_2SO_4) process is shown in Figure 4.4(b). In this case, the structures were taken directly from previous work and used without further optimization, facilitating a direct benchmark comparison of energetics in the absence of any geometric differences.[45]. In the case of ONIOM(LCCSD(T):PBE-D3), we observe a similar trend as above: DFT underestimates the defect formation energy and the reaction barrier, and overestimates the linker desorption

energy. Interestingly, the effective barrier of the reaction becomes positive (vs. the isolated reactants) when using LCCSD(T), whereas the DFT results show an effective negative barrier. Overall, these results suggest that defect formation in ZIF-8 is much less energetically and kinetically favorable as compared to what had been predicted from prior DFT results.

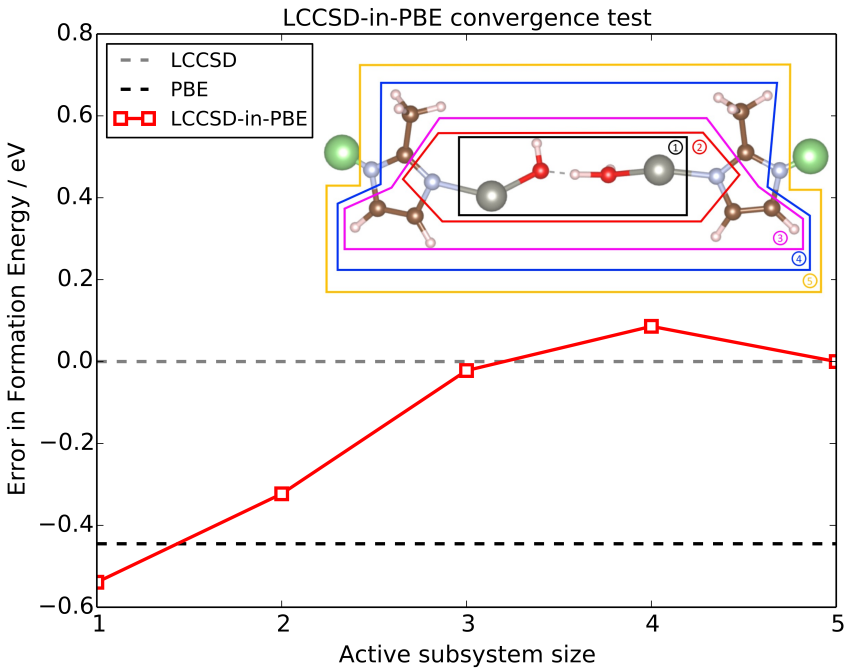


Figure 4.5: Convergence test for DFT embedding as compared to all atom methods.

4.4.3 A DFT embedding alternative

The above results show that a successful multi-level method requires a properly terminated 7L cluster to properly account for the electronic coupling to the surrounding bulk. Such a cluster consists of approximately 79 atoms and 364 electrons (for ZIF-8). Further reductions in the size of this cluster are likely to be challenging without significant loss in accuracy and fidelity to the original bulk calculation, since it would likely require cutting through covalent bonds in aromatic rings. Nonetheless, even such a moderately sized cluster can pose significant challenges to the highest levels of theory (e.g. multi-reference configuration interaction), which may be important for some transition-metal containing MOFs.

In such cases, density functional embedding methods provides a potentially attractive alternative to simple ONIOM methods. In this family of embedding methods, the electron density of the system can be partitioned into two subsystem densities,

$$\rho = \rho_{\mathrm{A}} + \rho_{\mathrm{B}} \tag{4.5}$$

and the total energy of the system can be written as the energy of the subsystems and a interaction energy:

$$E[\rho] = E[\rho_{\mathrm{A}}] + E[\rho_{\mathrm{B}}] + E_{\mathrm{int}}[\rho_{\mathrm{A}}, \rho_{\mathrm{B}}] \tag{4.6}$$

The non-additive interaction part $E_{\mathrm{int}}[\rho_{\mathrm{A}}, \rho_{\mathrm{B}}]$ contains Coulomb, exchange-correlation and kinetic contributions. In particular, the non-additive kinetic energy must be approximated. Several approaches have been proposed to calculate $E_{\mathrm{int}}[\rho_{\mathrm{A}}, \rho_{\mathrm{B}}]$. [68, 70, 72] Here, we will use Miller and coworkers' projector-based embedding method[70], which is implemented in the Molpro package. This method works by adding a projector based level-shift operator onto the Fock operator of the active subsystem to ensure the orbital orthogonality between the subsystems, eliminating the kinetic contribution to E_{int}. The theoretical details can was discussed in Chapter 2.

In this chapter we show a proof-of-principle LCCSD-in-PBE embedding of the 7L model, examining LV formation energy as a function of the size of the high (LCCSD) region; results are shown in Figure 4.5, with the inset panel demonstrating the active (LCCSD) subsystem. The active region uses a cc-pVDZ basis (aug-cc-pVDZ for N/O) , and a cc-pVDZ basis sets was used for all atoms in the DFT environment. The intrinsic bond orbital (IBO)[56] was used to localize the DFT orbitals, although test calculations show that the results are largely insensitive to the choice of localization method.

Briefly, we find that the DFT embedding method quickly converges to the full (all-atom) LCCSD result with increasing subsystem size. Significant deviations between the embedded and all-atom calculations are evident only for the smallest subsystem sizes (which cases neglect essentially all of the surrounding imidazolate linkers). As the subsystem size increases to include a portion of the surrounding linkers, the results come into nearly quantitative agreement with the corresponding all-atom LCCSD calculations (e.g. subsystems choices 3 and larger), which in turn

differ substantially from the all-atom PBE values. Note that, in the case of ZIF-8, such a choice allows a reduction in the number of electrons in the high subsystem from 364 to 244, which can yield a significant decrease in computation cost for QM methods with high scaling.

4.5 Conclusions

Our work demonstrates that, with proper accounting for environmental effects, cluster models can be utilized within the context of multi-scale methods (such as ONIOM) to increase the efficiency of conventional DFT calculations for MOF reactivity, and/or allow access to benchmark-quality energetic information via correlated WF calculations. A key conclusion is the sensitivity of results to the choice of the termination of the QM cluster. However, using validated cluster models, obtained via careful termination of ZIF clusters to account for charge transfer effects and ensure charge neutrality, we find that we are able to accurately reproduce bulk calculations at a fixed geometry in situations where mechanical coupling to the environment is negligible.

Building on this preliminary success, and in order to both account for this mechanical coupling to the environment (e.g. lattice strain) and to facilitate geometry optimization, we subsequently proposed a multi-scale ONIOM(QM:MM) approach. We find that this method is able to reproduce both the energetics and geometries for a variety ZIF defects as compared to the original (benchmark) PW DFT results. More generally, we find that accounting for the long-range mechanical coupling to the bulk is essential to yield correct geometries. In addition, the proposed ONIOM-based method also offers significant computational efficiencies, particular in the case of large unit cells or when transition state searches are requires.

In cases that require high quantitative accuracy, ONIOM(QM1:QM2) hybrids can provide access to benchmark-quality results via correlated WF calculations. Crucially, our data show that ZIF defect formation energies can differ significantly from those obtained via pure GGAs (or even hybrids) that are commonly employed in PW DFT. While the choice of the high-level correlated WF method is somewhat limited by the size of the QM1 region, we also demonstrate that DFT embedding can be used to potentially shrink the size of the QM1 region significantly without introducing significant additional artifacts. Such an approach would potentially allow for the examination of

the reactivity of MOFs containing strongly-correlated metal subsystems (e.g. Fe-MOF-74[108]), while fully accounting for the environmental couplings that have been omitted in many prior cluster model studies.

Chapter 5

Origins of Acid-Gas Stability Behavior in Zeolitic Imidazolate Frameworks: The Unique High Stability of ZIF-71

This chapter is reproduced in part with permission from Cui, K.; Bhattacharyya, S.; Nair, S.; Schmidt, J. R. Origins of Acid-Gas Stability Behavior in Zeolitic Imidazolate Frameworks: The Unique High Stability of ZIF-71, *J. Am. Chem. Soc.* **2021**, 143, 18061-18072.

5.1 Introduction

As we have reviewed in Chapter 1, several structural and first-principles-based descriptors like the linker pK_{a}, topology, hydrophobicity, defect formation energies, etc. were proposed to explain and predict the stability of MOFs in acidic environments[39–43]. However, the use of these descriptors are not always satisfactory, as several outliers were present in the previous investigations. The case of ZIF-71 in this context is especially intriguing. The RHO topology ZIF-71, containing 4,5-dichloroimidazole linkers, consistently appears as an outlier among ZIF materials with regard to acid gas stability. In an investigation on the applicability of ZIF materials for biobutanol purification, only ZIF-71 (among seven ZIF materials) was observed to maintain high 1-butanol capacity and selectivity in realistic humid CO_2 environments upon prolonged exposure.[33] Interestingly, both ZIF-8 as well as hybrid ZIF-8/71 (SOD) materials were observed to be unstable. In subsequent investigations of humid SO_2 stability of ZIFs, ZIF-71 (RHO) was the only stable ZIF among 16 materials upon prolonged humid (85 % R.H.) SO_2 exposure up to 900 ppm-days, with no detectable S incorporation into the framework[36]. A number of empirical stability indicators were considered to explain this behavior of ZIF-71, but were unsuccessful. For example, several other ZIFs that were hydrophobic (e.g., ZIF-7, ZIF-8) or shared the same RHO topology (e.g. ZIF-11) as ZIF-71 were all unstable. High linker pK_{a}, a proxy for the Zn-linker bond strength and thus

thermodynamic stability, also did not correlate with observed acid gas stability of ZIF-71. No significant difference in electron densities on the coordinating atom of ZIF linkers was observed from quantum chemistry calculations. Statistical correlations of the degradation rate constant (as obtained from exposure time-dependent measurements of the S:Zn ratio in the ZIF frameworks) with several material property variables could not be translated to ZIF-71, since it had undetectable reactivity in the experimental time scale[36]. This highlights that more subtle structural factors (beyond linker chemistry descriptors) likely play a crucial role in the observed stability of this material.

In this chapter, we use the computational tools developed in Chapter 4 to investigate the stability of ZIF-71 RHO and compare its stability with another well-known hydrophobic material, ZIF-8 SOD, by focusing on the kinetics of the formation of DL defects induced by acid gases. We show that common defect formation energy/activation energy calculations on an isolated single site in the ZIF are inadequate to explain their acid gas stability. We then utilize a combination of QM energetics and a statistical mechanics-based model to describe the adsorption and (thermally-averaged) reaction barrier for acid gas attack within the ZIF pore. These calculations conclusively indicate that ZIF-71 RHO is kinetically more stable, thus providing a solution to the intriguing paradox of the high acid gas stability of ZIF-71 RHO. This work also shows that, for materials that are not thermodynamically stable, tuning its functionalization and topology may enhance the acid gas affinity towards less reactive sites and eventually enhance the overall kinetic stability.

5.2 Computational Methods

5.2.1 QM/MM calculations

Accurate quantum chemical calculations are essential to computationally assess the reactivity of ZIFs. The standard periodic DFT method is infeasible for calculations involving large unit cells such as ZIF-71 RHO (816 atoms) due to its computational expense. Hybrid QM/MM methods therefore provide an optimum solution to reduce computational cost, enable beyond-DFT accuracy while retaining environmental effects. The QM/MM method tailored for MOF materials described in the previous chapter was used to perform electronic structure calculations. In this method,

only the reaction-relevant part of the ZIF (the QM region) is treated at a high level (DFT or wave function methods) of theory, while a low-level (MM) theory is applied to the rest of the framework to account for lattice strain. Point charges are added to the QM/MM boundary to represent the cleaved Zn-N dative bond. The QM/MM energy is calculated using a subtractive scheme by:

$$E_{\mathrm{QM/MM}} = E_{\mathrm{sub}}^{\mathrm{QM}} - E_{\mathrm{sub}}^{\mathrm{MM}} + E_{\mathrm{tot}}^{\mathrm{MM}} \tag{5.1}$$

where $E_{\mathrm{sub}}^{\mathrm{QM}}$ is the QM energy of the QM region, $E_{\mathrm{sub}}^{\mathrm{MM}}$ is the MM energy of the QM region and $E_{\mathrm{tot}}^{\mathrm{MM}}$ is the MM energy for the total system. A key advantage of this approach is that energetics can be calculated at an extremely accurate level of theory (here, local couple cluster theory, LCCSD(T)), while retaining the influence of the surrounding bulk MOF. We used PBE-D3[96, 97] as the high-level theory for rapid geometry optimizations. In the previous chapter, we have shown that this QM/MM approach can accurately reproduce periodic DFT structures and energetics. Since DFT-D3 tends to overestimate adsorption energies and underestimate defect formation and activation energies, a final single point QM/MM calculation was performed on the optimized structure with LCCSD(T)[48] as the high-level theory to calculate accurate energetics. For co-adsorbate distribution cases, an F12 correction[109] was added to ensure basis set completeness.

We used the previously developed Python interface via the ASE framework[100] to perform the QM/MM calculations. Molpro[99] and LAMMPS[102] were used to perform QM and MM calculations respectively. Double zeta Dunning style correlation consistent basis sets[49] (cc-pVDZ for Zn, C and H, aug-cc-pVDZ for O, N, Cl and S) were used for QM calculations. For MM part, the ZIF-FF force field[101] was used for ZIF-8 and ZIF-71 calculations, except for LJ parameters of the Cl atom in ZIF-71, which is taken from CGenFF[110]. CHARMM[111] and SPC[112] FF were used to model H_2SO_4 and water respectively. We benchmarked the force field choices of ZIF-71 by comparing the H_2SO_4 adsorption energy and the defect formation energy obtained by QM/MM and periodic DFT calculations. The results confirmed a high degree of accuracy for the overall QM/MM method. During geometry optimization, all atoms were relaxed until forces were less than 0.01 eV/Å. To search for transition states, nudged elastic band (NEB)[78, 80] calculations were performed via the NEB module within ASE. The band was relaxed until the forces on atoms are less than 0.04 eV/Å.

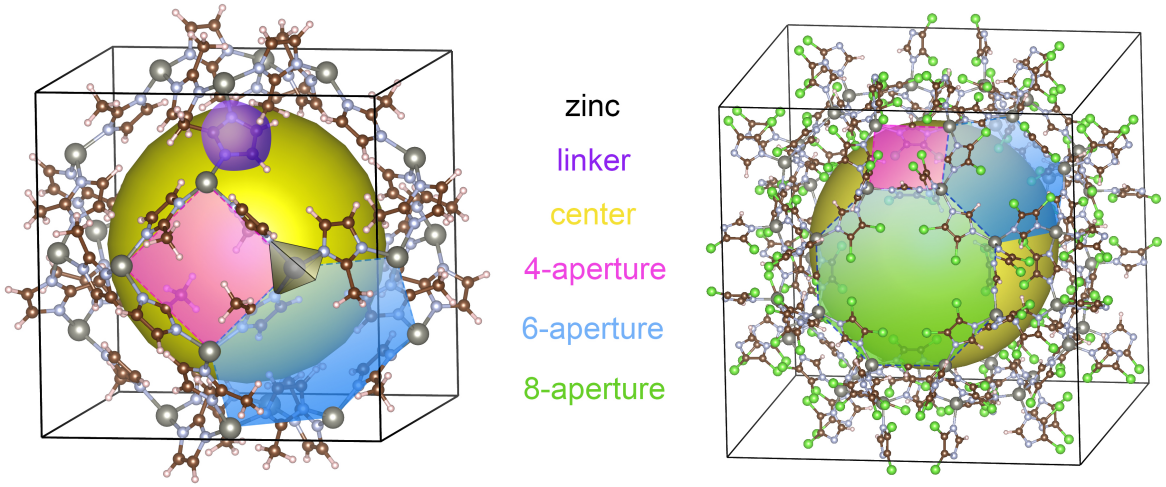


Figure 5.1: Pre-defined adsorption sites in (a) ZIF-8 SOD and (b) ZIF-71 RHO. The "zinc" and "linker" sites are not shown for ZIF-71 case for clarity.

5.2.2 Adsorbates distribution

A statistical mechanics-based modified multi-site multi-component Langmuir model to calculate adsorbate distributions within the ZIF framework was developed in this work. We adapted the basic assumptions within the Langmuir model (identical sites, no lateral interaction between the adsorbates) but allow two or more (interacting) adsorbates to associate and occupy the same site. In this case, the model is used to calculate the adsorbate *partitioning* between sites, rather than the total uptake; the later would require an exhaustive enumeration of all possible adsorption sites (beyond the most important ones) and a quantitative description of the *absolute* (vs. relative) adsorption energy (e.g. LCCSD(T) treatment of long-range interactions), and is beyond the scope of this work. In a ZIF framework, an adsorbate can bind to several distinct types of adsorption sites such as the zinc, linker, n-fold aperture (abbreviated as n-aperture), etc. A schematic illustration of the sites in ZIF-8 SOD and ZIF-71 RHO are shown in Figure 5.1. Each kind of site s has its associated "degeneracy" g_s. For example, there are a total of eight 6-member rings in the SOD topology so the degeneracy of the so-called "6-aperture" site will be 8. But the degeneracy also depends on the structure of adsorbed complex. The details can be found in the publication [113].

For the single adsorbate case, we first calculated the adsorption energy of the molecule on these pre-defined sites $\Delta E_{\mathrm{ad}}^{\mathrm{LCCSD(T)}}$ using the QM/MM method at the LCCSD(T) level. We then added zero-point energy (ZPE) correction and entropy correction term (at PBE-D3 level) (i.e. the entropy loss of the adsorbates after binding, see next section) to yield the adsorption free energy:

$$\Delta F_{\mathrm{ad}} = \Delta E_{\mathrm{ad}}^{\mathrm{LCCSD(T)}} + \Delta E_{\mathrm{ZPE}}^{\mathrm{PBE-D3}} - T\Delta S_{\mathrm{trans-rot}}^{\mathrm{PBE-D3}} \tag{5.2}$$

For a given set of sites and the associated adsorption free energies, the partition function of this adsorption model can be written as:

$$Q = \sum_{s} g_s e^{-\beta \Delta F_s^{\mathrm{ad}}} \tag{5.3}$$

and the probability of finding the adsorbate on site s (p_s) is:

$$p_s = \frac{g_s e^{-\beta \Delta F_s^{\mathrm{ad}}}}{\sum_t g_t e^{-\beta \Delta F_t^{\mathrm{ad}}}} \tag{5.4}$$

The above approach can be generalized to account for the adsorption of multiple competing species/co-adsorbates. Specifically, we are interested in the co-adsorption of the H_2SO_4/water mixture as previous reports suggest severe deleterious effects on the ZIF structure by this acidic mixture[45]. For the co-adsorption case, the possible states will be all combinations of adsorbates $i, j, ...$ binding to sites $s, t, ...$ ($s, t, ...$ are different sites but can be the same type, for example, $s, t, ...$ can be different zinc atoms). Because both ZIF-8 and ZIF-71 are hydrophobic, at the experimental condition (85% RH), the water loading in both ZIFs are low (1 and 3 water molecules per unit cell respectively), as determined from water adsorption isotherms published in previous work[36]. At such low-loading limit, if two molecules are adsorbed on different sites, the interaction between them can be neglected. As a result, the free energy change for this case becomes:

$$\Delta F_{\mu} = \Delta F_{i,s} + \Delta F_{j,t} + ... \tag{5.5}$$

Initially excluding the case that two or more molecules occupy the same site, the degeneracy of the state becomes:

$$g_{\mu} = \binom{g_s}{n_s} \times \binom{g_t}{n_t} \times ... \tag{5.6}$$

Here, $s, t, ...$ are different types of sites and n_s is the number of the sites of this type that is occupied by adsorbates. $\binom{g_s}{n_s} = \frac{g_s!}{n_s!(g_s-n_s)!}$ is the number of possible ways to distribute n_s adsorbates onto the total g_s sites. Considering the case where two or more molecules occupying the same site, the associated complex will be considered as a new complex adsorbate and the ΔF_{ad} will be re-calculated according to Eq.5.2. By using this approach, we explicitly include the interactions of molecules adsorbing on the same site at the LCCSD(T) level. The degeneracy of this state is equal to the degeneracy of the occupied site.

Note that the above model does not require additional corrections to account for the indistinguishability of the individual H_2O/H_2SO_4 molecules or complexes thereof, since the number of adsorbates of each type is fixed. Application of this model to predict absolute uptake of a mixture (vs. distribution within the pore) would require additional corrections to account for indistinguishability of isolated molecules vs. those involved in complexes.

5.2.3 Entropy calculation

When an adsorbate binds to the ZIF, it will lose some (but not all) of its original translational and rotational entropy. For free adsorbates in the gas phase, these initial entropies can be calculated using the standard harmonic oscillator/rigid rotor approach for free particles. For a bound state, the translational and rotational degree of freedoms (DOFs) transfer to vibrational DOFs. However, due to the low-frequency and anharmonic nature of these intermolecular DOFs, the simple harmonic oscillator approximation cannot be used to accurately estimate entropy for the binding states. Here, we introduce a simplified but useful method for entropy estimation of these anharmonic DOFs. We first use the harmonic approximation to determine the normal modes, then scan the potential energy along the normal modes to calculate the *anharmonic* potential and finally re-calculate the anharmonic frequencies and entropy based on the potential. Our benchmark calculation shows that with these approximations, MD-generated distributions are well reproduced by the model (vide infra).

For the "translation" of the bound state, we first calculate the translational Hessian by fixing the framework and slightly displace the adsorbate along the x, y, z axes. As the adsorbate binds to the

framework, movement along Cartesian coordinates are coupled with each other. By diagonalizing the translational Hessian, eigenvectors of the Hessian matrix give three decoupled axes. We then scan the potential energy along the principal axes to get the one-dimensional potential V_{trans}. This approach reduces the 3D translation problem to three 1D anharmonic vibration problems. The reduced 1D Schrödinger equation for translation is:

$$\hat{H}_{\mathrm{trans}}\psi_i = \left(-\frac{\hbar^2}{2M}\frac{\mathrm{d}^2}{\mathrm{d}\xi^2} + V_{\mathrm{trans}}(\xi)\right)\psi_i = E_i\psi_i \tag{5.7}$$

where M is the mass of the adsorbate, ξ is the translation distance. In practice, we used the harmonic oscillator wavefunctions as the basis set to span the translational Hamiltonian $\hat{H}_{\mathrm{trans}}$ and numerically solve the equation. Figure 5.2(a) depicts a typical 1D anharmonic potential. From the obtained eigenvalues $\{E_i\}$, we can calculate the "translational entropy" of the binding state:

$$S_{\mathrm{trans}} = \frac{\langle E\rangle - F}{T} = \frac{1}{T}\frac{\sum_i E_i e^{-\beta E_i}}{\sum_i e^{-\beta E_i}} + k_B \ln \sum_i e^{-\beta E_i} \tag{5.8}$$

A more accurate model might also include the coupling of adsorbate translations with the relaxation of the surrounding framework. However, our test calculations show that omitting this coupling gives only a small correction to the entropy and negligible influence on the distribution, with dramatically reduced computational cost. The rigid framework approximation thus provides a good balance between accuracy and efficiency for the present work.

For the "rotation" of the bound state, we used the principal axes of the moment of inertia rather than eigenvectors of the rotation Hessian (the off-diagonal elements of the rotational Hessian are ignored). The potential V_{rot} is obtained by rotating the adsorbate about these axes. The reduced 1D Schrödinger equation for rotation is:

$$\hat{H}_{\mathrm{rot}}\psi_i = \left(-\frac{\hbar^2}{2I}\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} + V_{\mathrm{rot}}(\theta)\right)\psi_i = E_i\psi_i \tag{5.9}$$

where I is principal moment of inertia associated with the corresponding principal axis and θ is the rotation angle. Note that both water and H_2SO_4 (the most stable trans- conformer) have a two-fold symmetry axis. After rotating 180° about this axis, the molecule will "go back to itself" due to the indistinguishability. For adsorbed case, this symmetry property is reflected by the two minima

with equal energy as shown in Figure 5.2(b), which shows the H_2SO_4 rotating around its C_2 axis in ZIF. Allowing the molecule to do a full 360° rotation around these axes (in other words, use the potential with two minima as V_{rot}) will result in a two-fold degeneracy for every eigenstate and thus overestimate the entropy. In gas phase, the influence of molecular symmetry is accounted by the "symmetry number" σ (for water and H_2SO_4, $\sigma = 2$) in rotational entropy. In the adsorbed phase, we truncate the potential near the minimum at $\theta = 0$ to remove the extra degeneracy caused by symmetry. The harmonic oscillator wavefunctions are used as the basis set to solve the equation. The "rotational entropy" of the bound state can be calculated accordingly from the eigenvalues $\{E_i\}$ and Eq.5.8.

For the case that the adsorbate does not bind to any sites but remains in the center of the pore, we used the ideal gas in a box with volume V_{center} and rigid rotor model to determine the translational and rotational entropy. For the translation case, the box volume V_{center} is calculated via $V_{\mathrm{center}} = \left(\frac{V_{\mathrm{pore}}}{V_{\mathrm{ZIF}}}\right)_{\mathrm{exp}} \times V_{\mathrm{unit\ cell}}$, where $\left(\frac{V_{\mathrm{pore}}}{V_{\mathrm{ZIF}}}\right)_{\mathrm{exp}}$ is the experimental pore volume reported in Ref.[36]. Note that the resulting entropy is highly insensitive to the estimate of this ratio due to the logarithmic dependence.

5.3 Reactivity of ZIF-8 and ZIF-71

The conventional procedure to computationally assess the stability of a crystalline porous material such as a ZIF is to calculate the defect formation thermodynamics and associated barriers. Previous studies [34, 114] show that in the vapor phase humid SO_2 environment, H_2SO_4 can form via oxidation of SO_2 followed by hydration. Thus, H_2SO_4 and sulfuric-sulfurous (H_2SO_4-H_2SO_3) complexes are the most plausible reactive species[115]. Considering these experimental reports and for computational simplicity, in this work we focus on the reactivity of ZIFs with H_2SO_4. Previous studies report that dangling linker (DL) defects are likely to be formed first by acid gas attack[44]. A schematic illustration of the proposed mechanism is shown in Figure 5.3. During the reaction, the acid (H_2SO_4) first attacks the Zn atom, leading to breakage of the Zn-N bond and formation of a new Zn-O bond. At the transition state, since the acid has generally lower pK_{a}

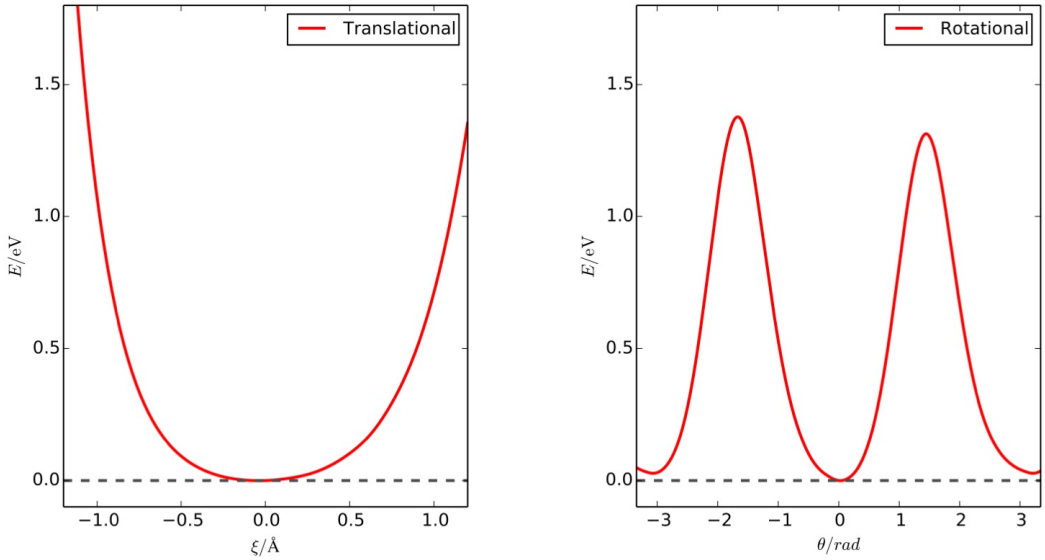


Figure 5.2: Examples of (a) translational and (b) rotational 1D anharmonic potential for H_2SO_4 at 6-aperture site in ZIF-71. (a) represents H_2SO_4 translation in the aperture plane and (b) presents H_2SO_4 rotation around its C_2 axis. ξ and θ are the translation distance and the rotation angle, respectively.

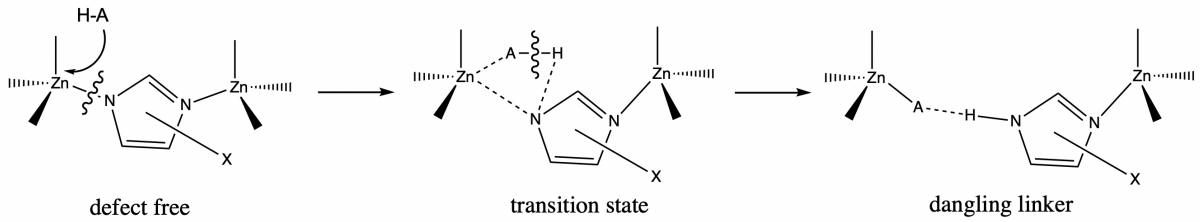


Figure 5.3: Mechanism of the dangling linker defect formation in ZIFs induced by acid attack. Here, the acid (H-A) is H_2SO_4 and the functional group -X is 2-methyl- (ZIF-8) or 4,5-dichloro-group (ZIF-71).

than the imidazolate linker, the proton of the acid will spontaneously transfer to the linker, forming a neutral imidazole and hydrogen bonding to the acid group. Note that we focus here on the reactivity of the pristine, defect-free ZIFs. In principle, the role of defects introduced by the synthesis protocol should also be considered. However, in prior work, ZIF-8 was synthesized using different synthesis protocols (with and without a HCOONa modulator) that are expected to yield different defect concentrations, and the degradation rates were able to be reproduced using both methods[36]. Those earlier results suggest that that the observed degradation behavior is intrinsic to the ZIF structure rather than via differing defect levels in pristine materials. As such, we focus in this work on the reactivity of the pristine structures.

We hence used the QM/MM method to calculate the adsorption energy of H_2SO_4 binding to the zinc/linker sites, the formation energy of the dangling linker defect, and the activation barrier for this defect formation process, at the LCCSD(T) theory level, for ZIF-8 SOD, ZIF-71 RHO and two other polymorphs, so-called: ZIF-71 SOD (with the SOD-II structure) and (hypothetical) ZIF-8 RHO. Note that while ZIF-8 (with 2-methyl-imidazolate [mIm] linkers) and ZIF-71(with 4,5-dichloro-imidazolate [dcIm] linkers) are most commonly synthesized in the SOD topology and RHO topologies, respectively, related structures with identical linkers may also appears in different topologies. In this section, we will use the number to indicate the linker functionalization and add a topology suffix explicitly.

System	ZIF-8		ZIF-71	
	SOD	RHO	SOD	**RHO**
Adsorption energy ΔE_{ad}	-0.48	-0.45	-0.45	-0.41
Formation energy ΔE_{form}	0.18(0.00)	-0.13	-0.04	-0.34
Activation energy $\Delta E^{\ddagger}$	0.93 (0.98)	0.62	1.06	0.76

Table 5.1: Adsorption and intrinsic reactivity data of the DL formation on the zinc/linker site of the ZIF/H_2SO_4 system calculated at LCCSD(T)/MM level. Numbers in parentheses are calculated in a $2 \times 2 \times 1$ supercell. Calculated energies are in eV.

The results are shown in Table 5.1. First, it is seen that the adsorption energies of H_2SO_4 on ZIF-8 SOD and ZIF-71 RHO are similar, indicating that binding strength of the acid is unlikely to be the dominant factor affecting ZIF stability. To decouple the role of functionalization and topology, we performed additional calculations on ZIF-8 RHO and ZIF-71 SOD, two alternative polymorphs of that differ from ZIF-8 SOD/ZIF-71 RHO only by crystal structure but not functionalization. These results show that the adsorption energy in the SOD topology is generally slightly more negative than in RHO topology (implying slightly stronger acid adsorption), because of the denser SOD crystal structure. Thus, it clearly cannot explain the higher stability of ZIF-71 RHO. The defect formation energy ΔE_{form} and activation energy $\Delta E^{\ddagger}$ data of the ZIFs (Table 5.1) is calculated on a single (zinc/linker) site of the material. We henceforth denote these data as the "intrinsic reactivity" of the site. In contrast, the term "overall stability" is used to describe the averaged reactivity over various sites in the framework as well as the experimental observations (see following discussion). For ZIFs with the same functionalization, the RHO topology is more intrinsically reactive than the SOD topology across both ZIFs, as reflected in their more negative formation and activation energies. Within a given ZIF functionalization, the activation barrier is positively correlated to the defect formation energy, consistent with the Hammond postulate and Bronsted-Evans-Polanyi (BEP) relations: for more exothermic reactions, the transition states will be closer to the reactant on the reaction path (less uphill), thus the energy cost to form a transition state is lower. The higher exothermicity in the RHO topology is likely caused by better strain relaxation, i.e., the RHO topology has larger pore sizes allowing strain induced by the defects to be relaxed more easily. On the other hand, this positive correlation is violated when comparing ZIFs with different linker functionalization: ZIF-71 has a more negative reaction energy but more positive activation energy than ZIF-8. On the other hand, the BEP slopes for ZIFs with mIm and dcIm linkers are almost same, suggesting that there is no significant change of the reaction pathway/mechanism due to the change in functionalization.

Since the RHO topology has four times more $ZnIm_2$ units than SOD, and we create one defect per unit cell when performing these calculations, the defect concentrations would be four times higher in the SOD topology. To examine this effect, we also performed reactivity calculations for

ZIF-8 using a $2 \times 2 \times 1$ supercell to match the defect concentration in RHO topology (data shown in parentheses, Table 5.1). We observe that the defect formation energy decreases significantly in this case (from +0.18 to 0.00 eV), because strain relaxation is more feasible in a larger supercell. Similar decrease in formation energy is also observed in previous periodic DFT calculations of linker vacancy defects[116], indicating that the difference is due to defect concentration effect rather than an artifact of the QM/MM method. The influence on the activation energy is however negligible (slightly higher activation energy in the supercell case is possibly caused by subtle changes of structure in the QM region).

Overall, we can conclude that the RHO topology has higher"intrinsic reactivity" than SOD. These calculations would therefore suggest that ZIF-71 RHO is both thermodynamically and kinetically more reactive with H_2SO_4 than ZIF-8 SOD, and therefore might be expected (on that basis) to be substantially less stable. This is in direct contradiction to experimental observations that report ZIF-71 RHO structure to be comparatively more stable[36]. Therefore, additional factors beyond the intrinsic reactivity must be affecting the observed stability/reactivity of ZIFs. This is the focus of our investigations in the following sections of this work.

5.4 Adsorbate Distribution

Differences in both linker functionalization and topology of ZIF-8 SOD and ZIF-71 RHO (henceforth we will refer to these structures as simply as ZIF-8 and ZIF-71, respectively) likely result in differing distribution of adsorbates within the pores. We hypothesize that adsorbate distribution may influence the framework stability. In many cases, molecular dynamics (MD) simulations can be used to calculate these distributions. However, we found that the FF energetics are not very reliable when compared to LCCSD(T) results, especially when we consider complex adsorbates such as H_2SO_4. Furthermore, strong interaction of H_2SO_4 with the framework can affect the sampling efficiency. To overcome the above deficiencies of the MD simulations, we developed a modified multi-site Langmuir method to incorporate LCCSD(T) energetics in distribution calculations. Since the distribution is calculated directly from the partition function, this model is free of sampling issues.

We first validated this model and the general approach by calculating the distribution of an isolated water in ZIF-8 with FF energetics and comparing this distribution with brute force MD simulations using the same FF. The ZIF-FF and SPC model[101, 112] were used to perform MD simulations in LAMMPS at 298 K for 10 ns. The same FF was also used to calculate the adsorption free energies in the model, allow for a direct assessment of the model itself, rather than the accuracy of the FF. The results are shown in Table 5.2. We can make quantitative comparisons by converting the MD trajectory to the occupancy of the sites (last column in Table 5.2), i.e. partitioning the unit cell into subspaces corresponding to unique sites and counting their average occupation. The difference in the probability distributions of modified Langmuir model and the MD simulation are modest (relative error $\sim$10%) and corresponds to only $\sim$ $0.1k_BT$ (0.003 eV) difference in the adsorption free energy, which has negligible influence on the average barrier calculations in subsequent sections. Therefore, the benchmark results reflect overall excellent agreement between the predominant binding sites predicted by the model and via molecular dynamics. Thus, despite several approximations, our model can reproduce the MD distribution with good fidelity.

Site	p_{model}	p_{MD}
zinc	0.33	0.35
linker	0.08	
4-aperture	0.00	0.000
6-aperture	0.03	0.04
center	0.56	0.61

Table 5.2: Water distribution in ZIF-8 predicted by the adsorption model as compared to brute force MD. p_{model} and p_{MD} are the water distribution predicted by the adsorption model and MD simulation (both at MM level), respectively.

We then used this model to calculate the H_2SO_4/water mixture distribution in ZIF-8 and ZIF-71. Figure 5.4 shows the calculated distribution of H_2SO_4 and water for both single component and mixture cases at the LCCSD(T) level. For the mixture distribution in ZIF-8, we put one H_2SO_4

molecule and one water molecule in the unit cell; for ZIF-71, we put one H_2SO_4 and three water molecules in the unit cell. The number of water molecules in the pore are obtained from the experimental water adsorption isotherms[36] and thus correspond to the respective experimental conditions. The corresponding spatial distribution plots (for H_2SO_4 only) are shown in Figure 5.5.

These results highlight the significant differences in the H_2SO_4 distribution between ZIF-8 and ZIF-71. While the aperture site is the most favorable site for both ZIFs, H_2SO_4 has a substantial probability (30-40%) of binding to the zinc/linker site in ZIF-8. In contrast, for ZIF-71, H_2SO_4 is found essentially always at the aperture (mostly 6-aperture) sites. The spatial distribution in Figure 5.5 also clearly shows the difference: the H_2SO_4 adsorption is far more diffusely distributed in ZIF-8 but is highly concentrated at the aperture sites in ZIF-71. This difference is critical, as established models of H_2SO_4-ZIF interactions suggest that attack by acidic species on the weak Zn/linker coordination sites initiates ZIF degradation via dangling linker defect formation, making ease of access to such sites a key parameter in the degradation process[45]. The observed difference in acid distribution is caused by the strong interaction of H_2SO_4 with ZIF-71 at its aperture sites. The adsorption energy of H_2SO_4 in ZIF-71 is -0.71 eV at the 6-aperture site but only -0.44 eV at the zinc/linker site, whereas for ZIF-8 the adsorption energies are essentially identical at both sites (about -0.5 eV). The presence of water has only a modest impact on the H_2SO_4 distribution in both ZIFs and does not change the picture significantly. With regard to the water distribution, in ZIF-8 water tends to stay at the center of the pore while in ZIF-71 the water molecules mostly bind to the zinc/linker site, due to the stronger interactions. The adsorption energy of water on the zinc/linker site is -0.17 eV in ZIF-8 versus -0.28 eV in ZIF-71. In ZIF-71, water tends to co-adsorb with H_2SO_4 at the aperture site, as reflected in the decreased occupancy of the zinc/linker site and a corresponding increase at the aperture site. In ZIF-8, the influence of H_2SO_4 on water distribution is less significant.

Figure 5.6 shows the probability of finding an associated H_2SO_4-water complex in ZIF-8 and ZIF-71, which is calculated by adding the probability of finding the H_2SO_4-water complex on all sites. Clearly, water and H_2SO_4 are mostly separated with uncorrelated distributions in ZIF-8, whereas substantial clustering is observed in ZIF-71. The most favorable cluster in ZIF-71 is

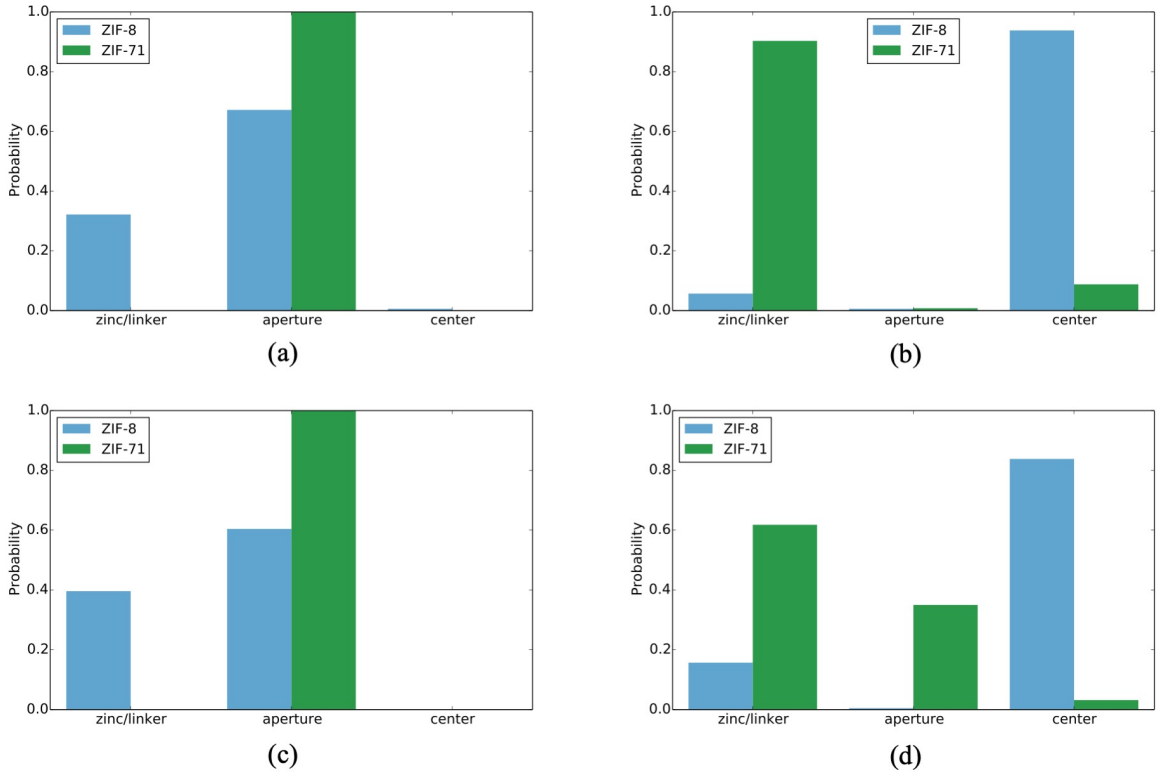


Figure 5.4: Distribution of single component (a) H_2SO_4 and (b) water in ZIF-8 and ZIF-71 and the marginal distribution of (c) H_2SO_4 and (d) water in the mixture calculated via the modified Langmuir model at LCCSD(T)/MM level.

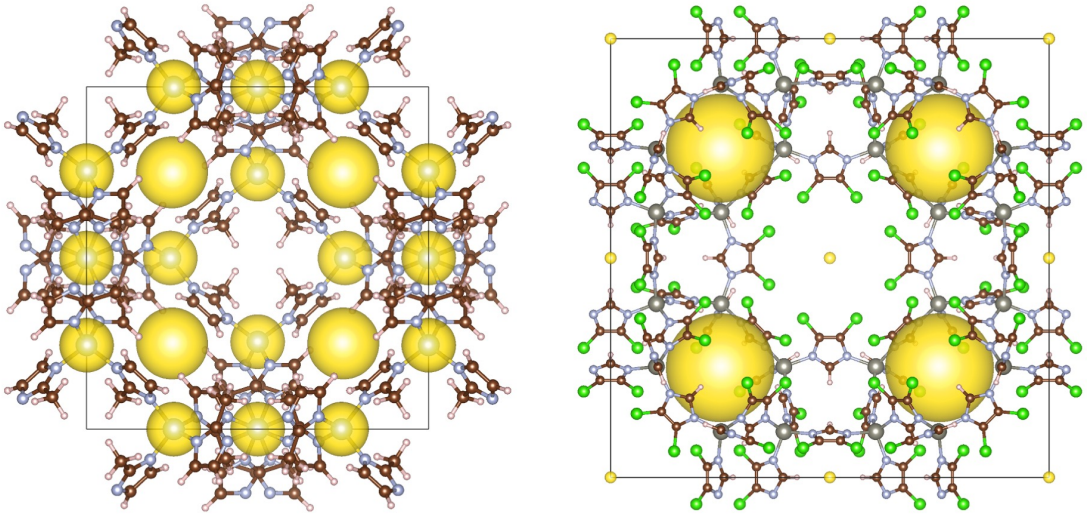


Figure 5.5: Spatial distribution of the marginal distribution of H_2SO_4 in the mixture in (a) ZIF-8 and (b) ZIF-71 at LCCSD(T)/MM level. The yellow spheres represent the possible positions of a H_2SO_4 molecule. Volume ratio of the spheres are calculated according to the probability ratio. Larger radius means higher probability.

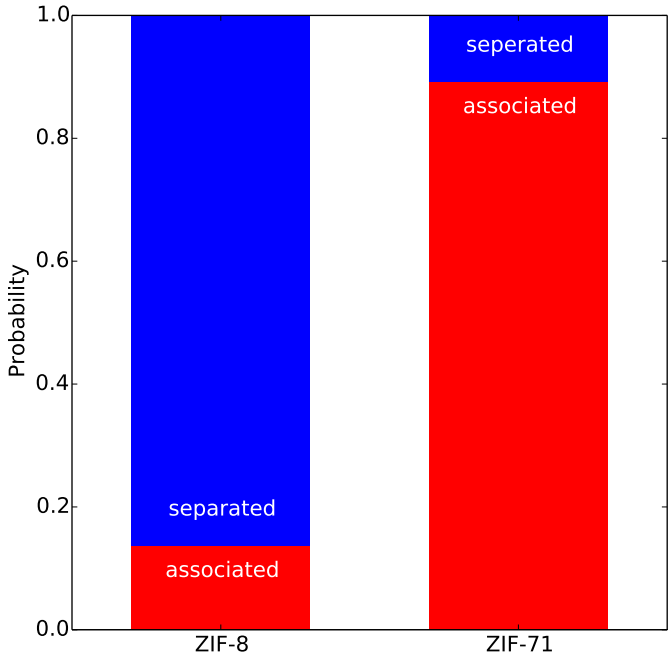


Figure 5.6: Probability of H_2SO_4-water association in ZIF-8 and ZIF-71 calculated by modified Langmuir model at LCCSD(T)/MM level.

$H_2SO_4 \cdot H_2O$ (76% probability) and $H_2SO_4 \cdot 2H_2O$ (21% probability). Association with more water is largely prohibited by high entropy losses. In summary, water-H_2SO_4 distributions are more correlated in ZIF-71, with a higher probability of association. The H_2SO_4 distribution is intrinsically different in the two ZIFs due to differences in interaction strength with the framework: the dispersion interaction between H_2SO_4 and Cl functionalization is stronger than that with methyl. Based on these two observations, we can draw two hypotheses as to the origins of the enhanced kinetic acid gas stability of ZIF-71. Either (1) the association of water with H_2SO_4 may reduce reactivity of the acid; or (2) the aperture sites may be less reactive compared to zinc/linker sites. The first hypothesis may seem peculiar as water association generally makes H_2SO_4 more acidic. However, the proton transfer step is not the dominant one and thus further increases in the acidity of the acid (via water complexation) may not provide for meaningful catalysis. Rather, it is possible that water association may increase the steric bulk of the acid and thus disfavor the Zn-N/Zn-O bond rearrangement. Nevertheless, our calculations show that a single water molecule associated with H_2SO_4 has an insignificant influence on the observed reaction barriers. As such, we will use dry H_2SO_4 as the reactant in all subsequent calculations and focus our remaining analysis on the second hypothesis, i.e., comparing and contrasting the reactivity of H_2SO_4 adsorbed at various sites and the associated impact on ZIF kinetic stability.

5.5 Reactivity of the Adsorption Sites

Next, we compare the intrinsic reactivity of the zinc/linker site and the 6-aperture site in ZIF-71. The potential energy surfaces of the DL formation on the two sites are plotted in Figure 5.7. It is clear that the reaction barrier for dangling linker defect formation is significantly larger at the 6-aperture site than at the zinc/linker site ($\Delta E^{\ddagger}$ = 1.53 vs. 0.77 eV for the 6-aperture and zinc/linker site, respectively); see Figure 5.7. The large reaction barrier indicates that ZIF-71 is unlikely to react with H_2SO_4 at the 6-aperture sites. As shown in Figure 5.8, at the zinc/linker site H_2SO_4 attacks the Zn-N bond in a direction parallel with the imidazolate ring. At the transition state, a hydrogen bond-like structure forms, lowering the energy of the transition state. In contrast, at the 6-aperture site the H_2SO_4 attacks the Zn-N bond in a direction perpendicular to the ring, and the

formation of the H-bond-like structure is prohibited due to geometry constraints. In the absence of this strong interaction, the energy of the transition state (the activation energy) increases.

Though the 6-aperture site is less intrinsically reactive than the zinc/linker site, it is nonetheless the preferred binding site for the H_2SO_4. As a result, the most favorable mechanism of DL formation in ZIF-71 likely proceeds as follows: H_2SO_4 spends most of the time at the 6-aperture sites, but due to occasional fluctuations it also diffuses into the zinc/linker site. From the zinc/linker site, it (occasionally) passes through the transition state (as the purple PES) and finally forms a defect. The effective barrier (relative to the most stable adsorbed state) of the process is thus 1.04 eV, which is slightly higher than that of ZIF-8 (0.98 eV).

A more rigorous way to evaluate the overall stability of the ZIF is to average the reaction rate over all adsorption sites. The overall (averaged) ZIF defect formation rate can be calculated via transition state theory (TST).

$$\langle k \rangle = \sum_s p_s \cdot k_s = \frac{k_B T}{h} \sum_s p_s \exp\left(-\frac{\Delta F_s^\ddagger}{k_B T}\right) \tag{5.10}$$

where p_s is the probability of finding the reactant (H_2SO_4) at the adsorption site s, $\Delta F_s^\ddagger$ is the activation free energy of the DL formation reaction at site s. The entropic contribution to $\Delta F_s^\ddagger$ is calculated as the difference between the reactant (obtained via vibration analysis of the framework, with MM region and H_2SO_4 fixed, plus the translational-rotational entropy of H_2SO_4) and transitions state (calculated solely by vibration analysis, with MM region fixed) since there are no "free" molecules). Equivalently, we can define the average activation free energy $\langle \Delta F_s^\ddagger \rangle$ by:

$$\langle \Delta F^\ddagger \rangle = -k_B T \ln \frac{\langle k \rangle}{k_B T/h} = -k_B T \ln \left[\sum_s p_s \exp\left(-\frac{\Delta F_s^\ddagger}{k_B T}\right) \right] \tag{5.11}$$

to yield an expression for the average *effective* free energy barrier.

The calculated $\{\Delta F_s^\ddagger\}$ and $\langle \Delta F_s^\ddagger \rangle$ in ZIF-8 and ZIF-71 are listed in Table 5.3. According to the data, the average effective barrier $\langle \Delta F_s^\ddagger \rangle$ of ZIF-71 is 0.06 eV higher than that of ZIF-8, corresponding to an overall defect formation rate in ZIF-71 that is $\sim$9 times slower than ZIF-8. Assuming the acid concentration to be one H_2SO_4 molecule per unit cell for ZIF-71, and one per $2 \times 2 \times 1$ supercell for ZIF-8, at 298K the overall defect formation rate (or Zn-N bond breaking

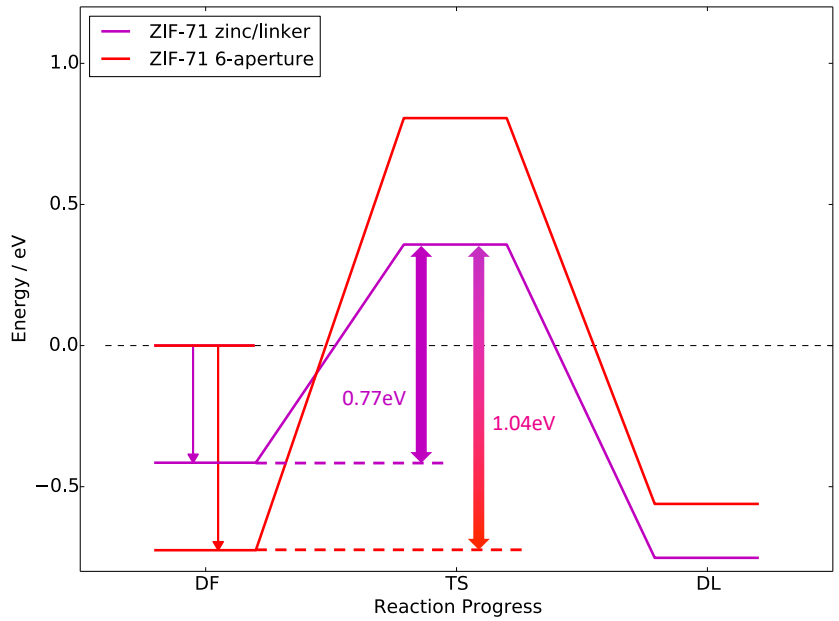


Figure 5.7: Potential energy surface of the DL formation on the zinc/linker site and the 6-aperture site in ZIF-71 calculated at LCCSD(T)/MM level. The arrow represents the apparent (purple) and the effective barrier (gradient color).

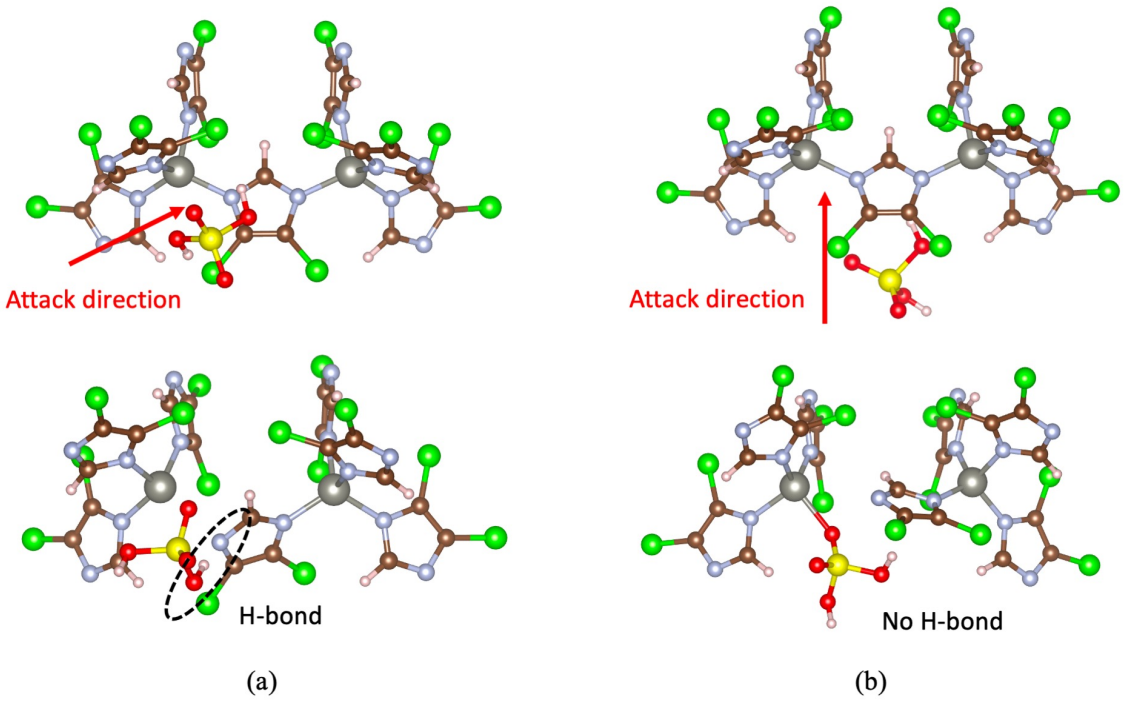


Figure 5.8: Transition states of the DL formation process of in ZIF-71 at (a) zinc/linker vs. (b) 6-aperture sites. The upper panels are the adsorption configuration and the lower panels are the transition state structures.

System	site	probability	$\Delta F_s^{\ddagger}$	$\langle \Delta F^{\ddagger} \rangle$
ZIF-8 SOD	zinc/linker	0.396	1.01	1.03
	6-aperture	0.604	1.10	
	4-aperture	0.000	N/A	
	center	0.000	N/A	
ZIF-71 RHO	zinc/linker	2.2×10^{-6}	0.76	1.09
	8-aperture	0.001	0.96	
	6-aperture	0.999	1.64	
	4-aperture	0.000	N/A	
	center	0.000	N/A	

Table 5.3: Average activation free energy of DL formation in ZIF/H_2SO_4 systems calculated at LCCSD(T)/MM level.$\Delta F_s^{\ddagger}$ is the activation free energy of the DL formation reaction at site s, and $\langle \Delta F_s^{\ddagger} \rangle$ is the average activation free energy. Free energies are in eV.

rate) in ZIF-8 and ZIF-71 in terms of the percentage of Zn-N bonds cleaved are 1.1% bonds/day and 0.11% bonds/day, respectively. Note that this overall defect formation rate depends on the real concentration of H_2SO_4, which is difficult to determine experimentally. Nevertheless, our calculated results have the same order of magnitude as experimental measurements ($\sim$1.6% bonds/day for ZIF-8 and lower for other ZIFs[36]). In the next chapter, we will introduce a kinetic model to make more direct and quantitative comparisons between computational and experimental results. Moreover, the analysis above shows that stability of the two materials is determined by very different factors. In ZIF-8, reaction proceeds directly between H_2SO_4 and the zinc site, which constitutes the predominant (i.e., most stable) adsorption site. In contrast, the reaction pathway in ZIF-71 is indirect, with the H_2SO_4 adsorbed predominantly at the "inert" aperture sites and requiring additional thermal fluctuations to reach the "reactive" zinc site prior to subsequent reactions. Thus, the overall kinetic stability of ZIF-71 is governed by not only the intrinsic reactivity of one site but also the relative adsorption strength of all sites.

Our results highlight a potential novel strategy to design acid gas stable ZIF materials. While both ZIFs are thermodynamically unstable, ZIF-71 is kinetically stable because its most reactive site is an unfavorable binding site, and the migration of reactants between sites costs additional free energy. Therefore, one can potentially enhance the acid gas affinity of the less reactive sites (for example, 6-aperture sites in ZIF-71) to increase the kinetic stability of a ZIF. This may be achieved by tuning the functionalization (i.e., adding functional groups that could interact more strongly with the acid) or tuning the topology (changing local density of some sites to enhance the dispersion interaction with the acid). In a similar vein, the probability of finding the acid gas at the metal site can also be used as an additional descriptor for stable material screening using machine learning methods. With low probability generally indicating higher stability; such descriptor could be calculated at low cost using FF methods, or with higher accuracy via conventional electronic structure approaches.

5.6 Conclusion

We have computationally investigated the kinetic stability of ZIF-71 as compared to ZIF-8 in humid SO_2 environments. Although isolated defect formation energies and corresponding activation energies are widely used descriptors for the chemical stability of ZIFs, our results show that these simple descriptors yield an incomplete picture of ZIF stability. Specifically, ZIF-71 is predicted to be both thermodynamically and kinetically more reactive with H_2SO_4 than ZIF-8 (less stable), but is experimentally observed to be more stable. Our subsequent study indicates that the failure of these simple descriptors is due to fundamental differences in the defect formation mechanisms between these ZIF systems.

While adsorbed H_2SO_4 reacts "directly" with the reactive Zn sites in ZIF-8, in ZIF-71 the reaction proceeds only after a fluctuation of the adsorbed gas from the (stable) aperture site to a higher energy Zn site. This "indirect" mechanism increases the effective barrier of the DL formation in ZIF-71 by about 0.3 eV relative to what would be expected based on the consideration of merely the reaction barrier of isolate Zn site. Our comprehensive model thus predicts that the kinetic stability of ZIF-71 is roughly 9 times larger than that of ZIF-8, in accord with the experimentally observed enhanced stability of the former. Thus unlike ZIF-8, where the overall stability is solely determined by the intrinsic reactivity of the most reactive site, the stability of ZIF-71 is governed by both the intrinsic reactivity of one site and the relative adsorption strength of all sites. Our results thus highlight the important influence of the acid gas distribution on material stability, which should always be considered when examining the chemical stability of ZIFs. In addition, this work also suggests a potential novel approach for enhancing the kinetic stability for porous materials that are intrinsically thermodynamically unstable, by tuning its functionalization and topology to enhance the acid gas affinity towards less reactive sites, and thus enhancing the overall kinetic stability.

Chapter 6

Kinetic Model of the Acid Gas-Induced Defect Propagation in ZIFs

6.1 Motivation

In Chapter 1 we have reviewed several computational works about the acid-gas stability of ZIFs by calculating the thermodynamics and kinetics of point defect formation. The influence of the attacking acid, crystal facet, topology, etc. has been discussed[44, 45]. In Chapter 5 we moved beyond the model of defect formation on single site, by considering the contribution from multiple sites. However, these theoretical models only focus on the formation of one single *microscopic* point defects, while the experimentally observed *macroscopic* degradation of ZIFs is a complex process involving consecutive formation of multiple defects. This difference poses extreme discrepancies in both length- and timescales between the theoretical and experimental studies. To close this gap, a model that can describe the long time scale defect propagation is needed. In this chapter, we will build a lattice-based kinetic model to simulate the defect propagation and build direct connection between defect formation reactivity and experimental stability. The time evolution of defect concentration and pore volume loss can be obtained from the simulation. With input parameters from first-principle calculations, our model predicts the degradation time scale in ZIF-8 and ZIF-71 in great agreement with experiments.

6.2 Computational Methods

6.2.1 Kinetics Calculations

We used the QM/MM methods described in Chapter 4 to perform all kinetics calculations. Symmetrically unique two-defect configurations were generated in a $2 \times 2 \times 2$ or $4 \times 2 \times 2$ super cell. These configurations are classified by the distance between defects and their connectivity, i.e.

whether they are sharing four-member (4M) or six-member (6M) rings. Similar to our previous works, we used the Python interface developed in-house via the ASE framework[100] to perform the QM/MM calculations. Molpro[99] and LAMMPS[102] were used to perform QM and MM calculations respectively. For ZIF-8/water system, PBE-D3[96, 97] with double zeta Dunning style correlation consistent basis sets[49] (cc-pVDZ for Zn, C and H, aug-cc-pVDZ for O, N and S) was used as the high-level theory and the lattice strain is described by force fields (FFs). ZIF-FF[101] and SPC[112] FF were used to model the ZIF and water respectively. For the geometry optimization, all atoms were relaxed until the forces were less than $0.02\ \mathrm{eV} \cdot \mathrm{\AA}^{-1}$. Nudged elastic band (NEB)[78, 80] method was used for transition state searching. The band was relaxed until the forces on atoms are less than $0.04\ \mathrm{eV} \cdot \mathrm{\AA}^{-1}$. For ZIF-8/$H_2SO_4$ system, in addition to the PBE-D3/MM optimization, extra single point QM/MM calculations were performed using the optimized structure with LCCSD(T)[48] as the high-level theory. The CHARMM FF[111] is used for H_2SO_4.

6.2.2 Pore Volume Calculations

The following procedure was used to generate defective structures and calculate the pore volume. First, various numbers (1-8) of defects are introduced to the primitive cell of ZIF-8. The defective structure is then roughly annealed from 300K to 1K at FF level to let the defective linkers appropriately rearrange. The relaxed structure is then further optimized at PBE-D3 level[96, 97] in VASP[94, 95] with a plane wave basis set and a 600 eV energy cutoff. Both the atom positions are relaxed until the forces were less than $0.02\ \mathrm{eV} \cdot \mathrm{\AA}^{-1}$ with fixed or flexible cell vectors. The accessible and total pore volume of the final structure is calculated using Zeo++[117] with a probe of radius of 1.3 Å. Using this probe radius, the calculated accessible pore volume ratio in a pristine ZIF-8 is 0.5, agrees well with experimental value of 0.55 [36]. For cases with 3-5 defects in the primitive cell, multiple configurations are generated using the procedure described above and the pore volume is calculated as the average over these configurations.

6.2.3 Kinetic Model

We use a two-dimensional (2D) 200×200 square lattice (with periodic boundary condition) to represent the ZIF structure. The linkers form another square lattice and is represented using a 2D array. The defect formation rate r_{site} on each linker site is stored in the 2D array. Based on the connectivity, each linker has 6 nearest neighbor (NN) sites. Kinetic Monte Carlo (KMC) method[118] was used to evolve the system. In the KMC algorithm, the probability of one site to become defective is proportional to the defect formation rate on that site.

$$p_i = \frac{r_i}{\sum_j r_j} \tag{6.1}$$

At each KMC step, one defect is formed on a randomly picked site, and the time increment is:

$$\Delta t = \frac{1}{\sum_j r_j} \ln \frac{1}{\xi} \tag{6.2}$$

where ξ is a random number sampled from a uniform distribution. The time evolution is calculated by averaging multiple KMC trajectories.

The pore-blocking (local structural collapse and boundary blocking) effects are included in the kinetic model by defining the accessibility of the sites. First, each site and its 6 NNs are considered as the "unit volume" to check local defect concentration. When < 3 defects are found in this region ($\rho < 3/7 = 0.43$), the central site is defined as accessible. If ≥ 3 defects are found, the central site is defined as inaccessible. To account for the boundary blocking effect, the non-defective are divided as several "clusters" based on the connectivity of the sites. The periodic boundary condition makes the determination of the accessibility of the non-defective clusters easy. An accessible cluster must connect to or be part of an infinite channel, which means that the cluster should connect either left/right or up/bottom periodic edges of the lattice. This definition of accessible clusters is same as the definition of percolated, infinite clusters in the percolation theory. In short, a percolated, infinite cluster is accessible and a unpercolated, finite cluster is inaccessible. All sites within the inaccessible cluster are thus also inaccessible. The check of percolated clusters is performed every 40 KMC steps, to balance the accuracy and efficiency of the simulation.

6.3 Defect Formation in Various Environments

A key challenge in simulating materials degradation is that the thermodynamics and kinetics of defect formation may change during the course of degradation. For a multi-defect system, the formation energy and barrier of new defects obviously depends on the "defective state" of the system:

$$\Delta E = \Delta E(\text{state}) \tag{6.3}$$

We will assume that the "state" is merely the presence/absence of a neighboring defect, in which case the only parameter is the relative formation and activation energies $\Delta\Delta E_{\text{form}}$ and $\Delta\Delta E^{\ddagger}$:

$$\Delta\Delta E = \Delta E_{\text{2nd}} - \Delta E_{\text{1st}} \tag{6.4}$$

The $\Delta\Delta E$'s depend on the relative positions of the defects. A negative $\Delta\Delta E$ indicates the formation of the second defect at that position is more favorable than forming an isolated one.

In this chapter we will focus on the DL defects. Using ZIF-8/water system as an example, we systematically studies the dependence of $\Delta\Delta E$ on the relative position, which is described by the distance between the defects $d_{\text{N*O}}$ and their connectivity, i.e. whether they are sharing four-member (4M) or six-member (6M) rings. The results are shown in Figure 6.1. Clearly, both $\Delta\Delta E_{\text{form}}$ and $\Delta\Delta E^{\ddagger}$ are negative at short $d_{\text{N*O}}$ distance (< 10 Å), indicating the second defect tends to form adjacent to the first one. The nearest neighbor (NN) sites (filled circles) are particularly favored. At a mid-range distance ~ 16 Å, $\Delta\Delta E_{\text{form}}$ and $\Delta\Delta E^{\ddagger}$ become slightly positive. They finally both converge to zero when $d_{\text{N*O}} > 30$ Å. In the last case the two defects are separated by about two unit cells. After a DL defect is formed, the acid is inserted between the Zn and the linker, resulting in a distorted lattice with elongated Zn-Zn distance, and thus high strain energy. When a new defect forms on the same 4M or 6M ring as the first one, the cleavage of more Zn-N bonds breaks the connectivity of the ring and the strain could be released. Therefore $\Delta\Delta E < 0$. For the no-ring-sharing cases, the formation of the new defect cannot release the existing strain. In contrast, the strain caused by both defects will "squeeze" each other and thus the formation of new defect is slightly disfavored and $\Delta\Delta E > 0$. If the two defects are separated far enough, their

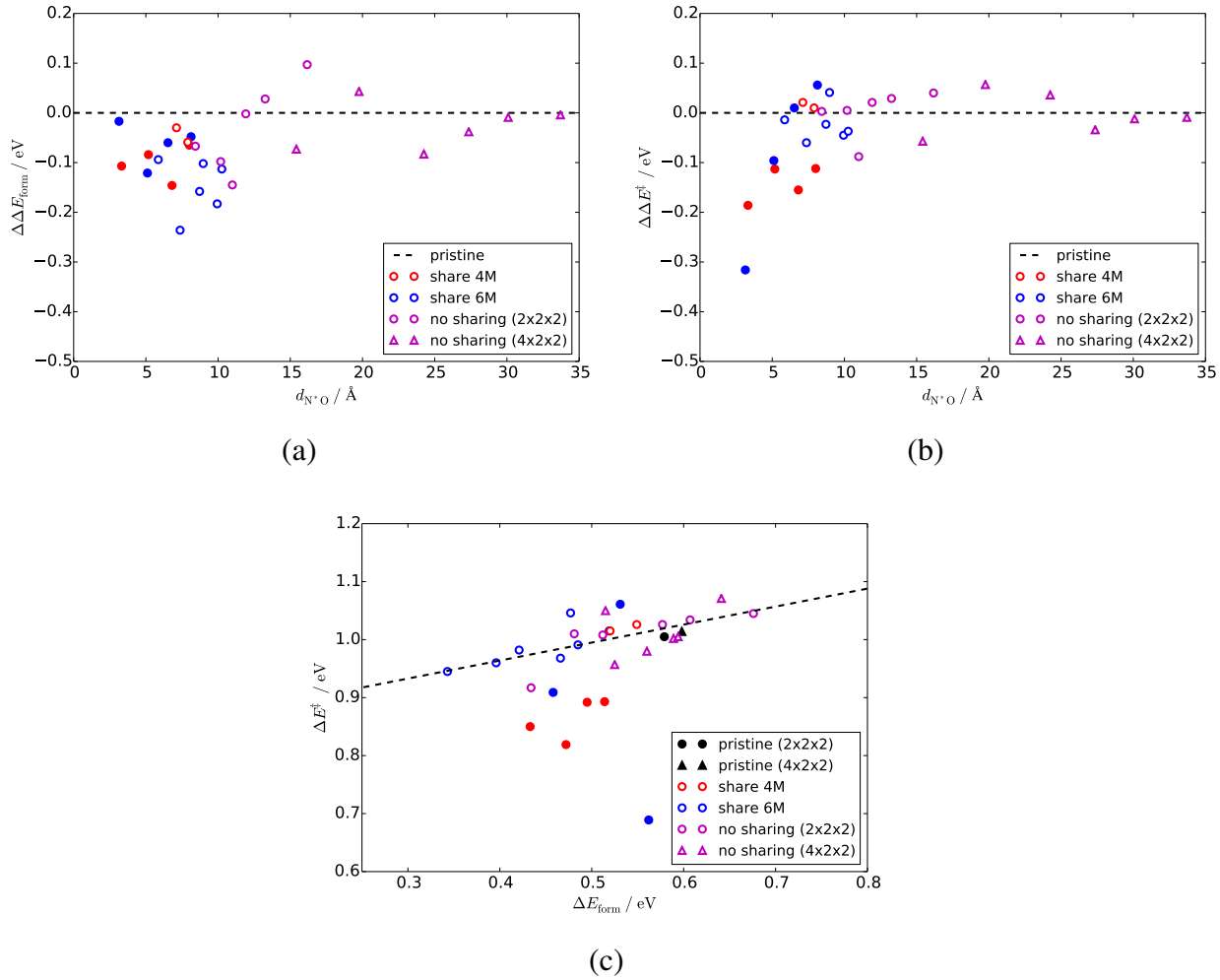


Figure 6.1: (a) $\Delta\Delta E_{\mathrm{form}}$ and (b) $\Delta\Delta E^{\ddagger}$ as functions of the distance and connectivity of the two defects in ZIF-8/water system. (c) BEP relation between ΔE_{form} and $\Delta E^{\ddagger}$ of the multi-defect structures of ZIF-8/water system. The filled circles indicate the cases that the second defect forms on a nearest neighbor sites of the first one. Circle markers are calculated in a $2 \times 2 \times 2$ super cell and triangular markers are calculated in a $4 \times 2 \times 2$ super cell. $d_{\mathrm{N^*O}}$ is defined as the distance between the N atom of the cleaving Zn-N bond in the second defect and the O atom in water in the first defect.

interaction becomes negligible and $\Delta\Delta E$ converges to 0. Except for the NN cases, all other data points fall on the same Bell–Evans–Polanyi (BEP) line (Figure 6.1(c)).

In addition to the ZIF-8/water system, we also looked at the interaction between H_2SO_4 induces defects. Based on the observations discussed above, for ZIF-8/H_2SO_4 system we only consider the cases that two defects are NNs. The calculated $\Delta\Delta E_{\mathrm{form}}$ and $\Delta\Delta E^{\ddagger}$ results for three representative configurations are summarized in Table 6.1. For config 1 and 2 the two defects are sharing a 4M ring and for config 3 the two defects are sharing a 6M ring. Similar to the water cases, $\Delta\Delta E_{\mathrm{form}}$'s are negative for all configurations, but with much larger magnitude, due to stronger attractive interaction involving H_2SO_4 molecules. However, there is no uniform trend in $\Delta\Delta E^{\ddagger}$. This observation can be rationalized as follows. Since H_2SO_4 is both more bulky and has more degrees of freedom than water, there is substantial structural rearrangement when forming a defect. For certain configurations such as config 2, the structural rearrangement includes a breakage and re-formation of the acid-HIm hydrogen bond in the first defect, as shown in Figure 6.2. At TS, the hydrogen bond in the first defect is cleaved, which destabilizes the TS, resulting in a more positive $\Delta\Delta E^{\ddagger}$. After all, when considering a new defect formation on a NN site of the first defect, config 1 will be the dominant configuration. The defect formation adjacent to an existing defect is still favored in ZIF-8/H_2SO_4 system.

2DL structure	config 1	config 2	config 3
$\Delta\Delta E_{\mathrm{form}}$ / eV	-0.44	-0.66	-0.43
$\Delta\Delta E^{\ddagger}$ / eV	-0.32	0.15	0.06

Table 6.1: $\Delta\Delta E_{\mathrm{form}}$ and $\Delta\Delta E^{\ddagger}$ in ZIF-8/H_2SO_4 system.

6.4 Kinetic Model for Defect Propagation

Knowing that existing defects prefer new defects to form on their neighbor sites, we can then build a kinetic model to investigate how the neighboring interaction affects the defect propagation in ZIFs. For simplicity, we only consider the NN interactions and use a mean field description for

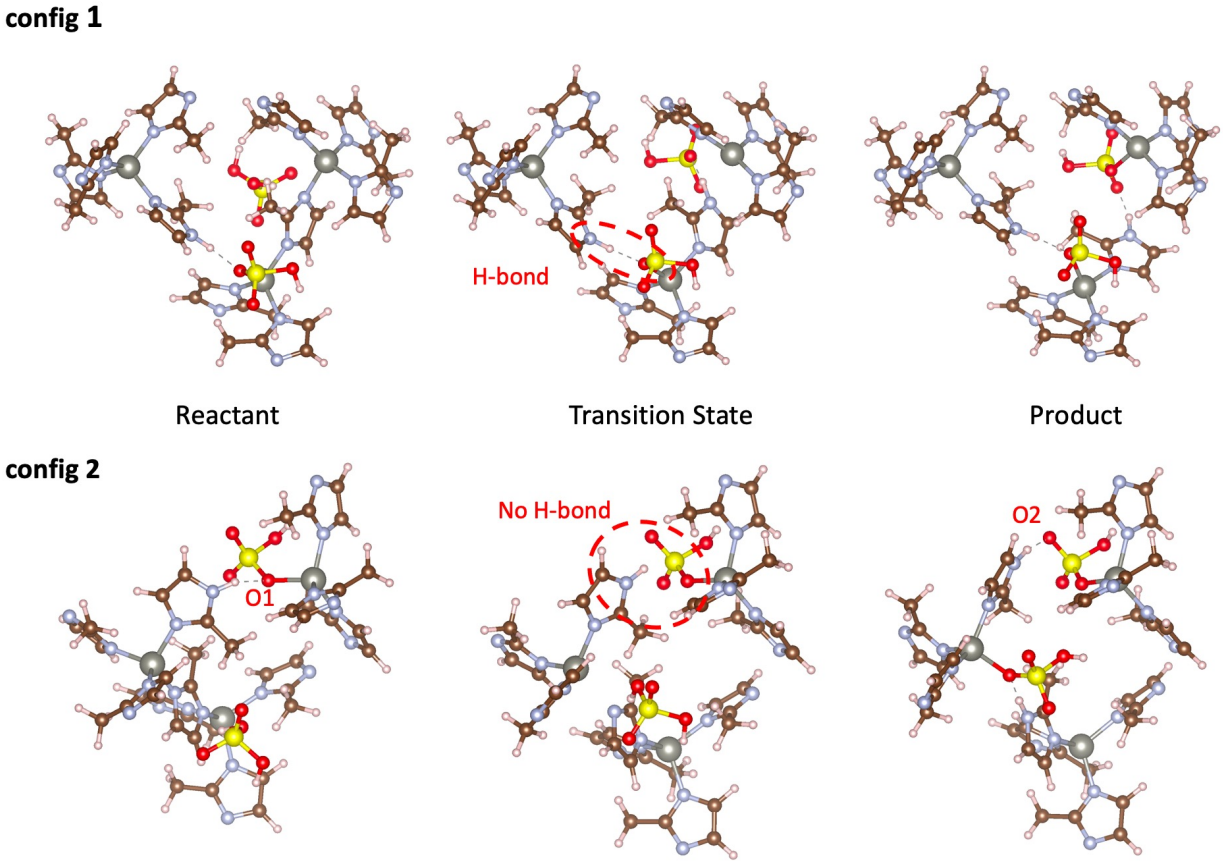


Figure 6.2: Structures of the reactant, TS and the product for the formation of a second defect in ZIF-8/H_2SO_4 system. The upper and lower panels represent config 1 and 2 respectively. In config 1, the acid-HIm hydrogen bond stays intact. While in config 2, the H in HIm binds to the O1 atom in the reactant but transferred to bind to O2 atom in the product. There is no acid-HIm hydrogen bond in the transition state.

$\Delta\Delta E^{\ddagger}$. In other words, the defect formation rate on a non-defective site can be summarized as:

$$r_{\text{site}} = \begin{cases} kr_0, & \text{if has defective NN} \\ r_0, & \text{otherwise} \end{cases} \tag{6.5}$$

where r_0 is the formation rate of an isolated defect, which only depends on the activation energy and the concentration of the acid gases (see below). In the reduced unit system, $r_0 = 1/\tau$ is set to be 1 ($\tau = 1$ is the reduced time unit). $k = \left\langle e^{-\beta\Delta\Delta E^{\ddagger}} \right\rangle_{\text{NN}}$ is the averaged relative defect formation rate over all NN sites. We can also use the exponentially averaged $\Delta\Delta E^{\ddagger}$ values to indicate the relative rate:

$$\langle \Delta\Delta E^{\ddagger} \rangle = -k_B T \ln k = -k_B T \ln \left\langle e^{-\beta\Delta\Delta E^{\ddagger}} \right\rangle_{\text{NN}} \tag{6.6}$$

Based on the QM/MM data, we find the $\langle \Delta\Delta E^{\ddagger} \rangle$ to be similar for both water and H_2SO_4 cases. We use $\langle \Delta\Delta E^{\ddagger} \rangle = -0.28$ eV (calculated using the H_2SO_4 data) for all subsequent results.

We use the 2D square lattice to represent the ZIFs and use the KMC method[118] to evolve the system. The simulation snapshot with input parameter $\langle \Delta\Delta E^{\ddagger} \rangle = 0$ or -0.28 eV are shown in Figure 6.3(a) and (b), respectively. When $\langle \Delta\Delta E^{\ddagger} \rangle = 0$, the formation of new defects is uncorrelated with the existing defects. The defects thus have an uniform distribution in the lattice. The time evolution of the defect concentration exhibits the same behavior as a normal first-order reaction, as shown in Figure 6.3(c). The defect formation rate decreases exponentially because the concentration of the "reactants" (non-defective sites) are decreasing. On the other hand, when $\langle \Delta\Delta E^{\ddagger} \rangle = -0.28 \ll -k_B T$, the formation of new defects adjacent to the existing ones are strongly favored. The defects tends to form isolated, compact clusters. Furthermore, the formation rate of defects on the surface of the defect cluster is much faster ($k \gg 1$) than the formation rate of isolated defects, which means the existing defects can accelerate the formation of new defects: the defect propagation in ZIFs is an autocatalytic reaction. As depicted in Figure 6.3(d), the $c(t)$ curve shows a sigmoid shape, which is a clear indicator of an autocatalytic reaction. By analyzing the rate equation of the kinetic model, one can prove that at early stage, the defect concentration increases as a *power function* of the time (the full derivation can be found in the Appendix of this

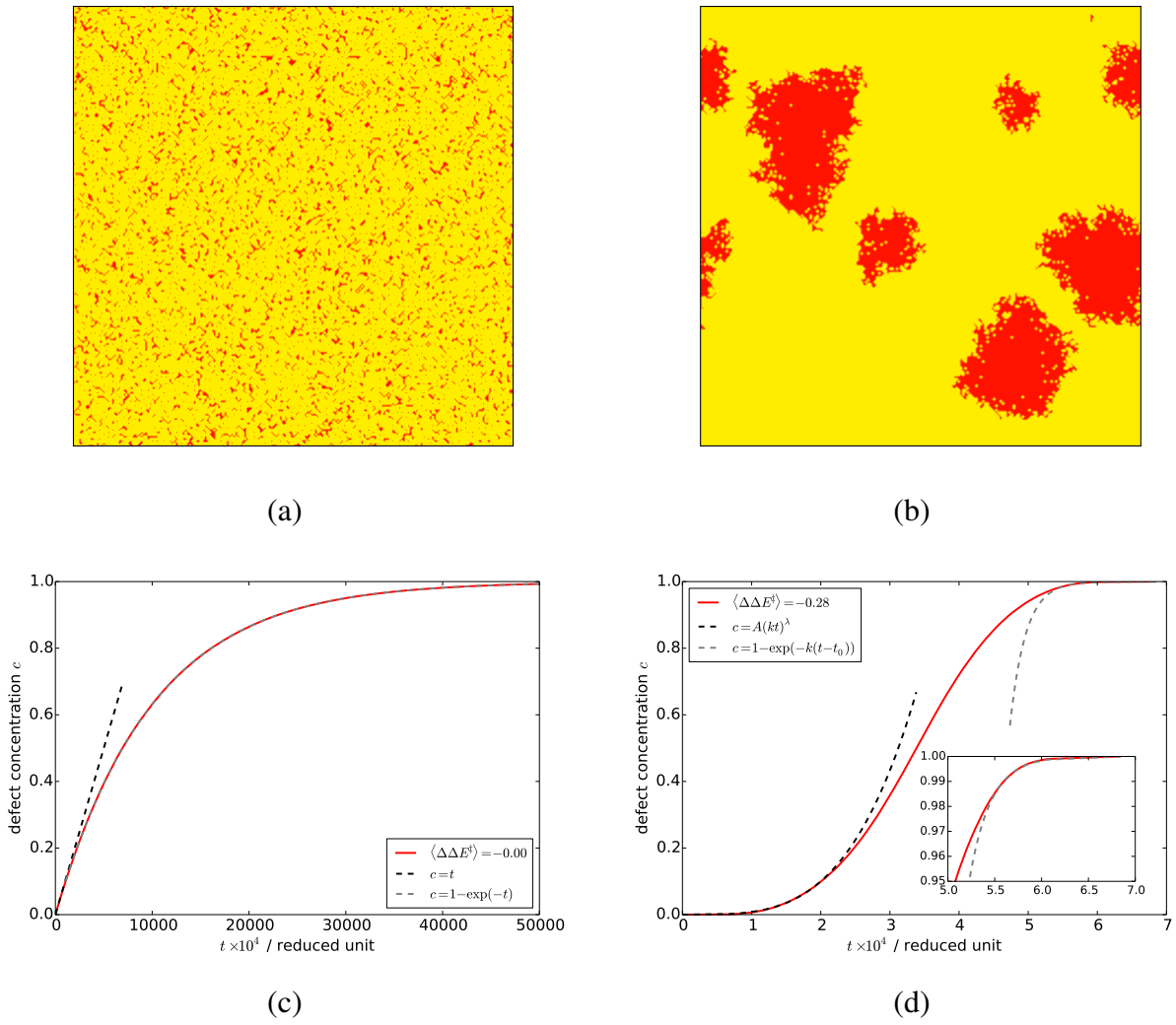


Figure 6.3: (a)(b) KMC simulation snapshot at $c = 0.2$ and (c)(d) the simulated time evolution of defect concentration. $\langle \Delta\Delta E^{\ddagger} \rangle$ is set to be 0 in (a)(c) and -0.28 eV in (b)(d). The inset panel in (d) shows the zoomed in $c(t)$ curve at the late stage.

chapter):

$$c(t) = \begin{cases} A(kr_0t)^{\lambda}\,, & \text{early stage, } \lambda > 1 \\ 1 - \exp[-kr_0(t - t_0)]\,, & \text{late stage} \end{cases} \tag{6.7}$$

Note that this equation is valid regardless the dimension nor the connectivity of the lattice. The predicted power function form of $c(t)$ by our kinetic model is supported by experiments, as this autocatalytic feature has also been seen experimentally in various ZIFs such as ZIF-8, ZIF-14, hybrid ZIF-8_{50}-90_{50} and hybrid ZIF-8_{15}-14_{85}.[36, 38]

System	K_{H}[a] / $\mathrm{mmol \cdot g^{-1} \cdot bar^{-1}}$	$\Delta F^{\ddagger}$[b] / eV	τ / days
ZIF-8	~ 9	1.03	2.4×10^4
ZIF-71	~ 5	1.09	2.9×10^5

[a] The Henry's constant values are obtained form experimental dry SO_2 adsorption isotherm measured in Ref. [36]. [b] The activation free energies are taken from Chapter 5.

Table 6.2: Parameters and calculated time unit τ for defect formation in ZIF-8 and ZIF-71 under 20 ppm SO_2 exposure.

Another important prediction of this kinetic model is the time scale of defect propagation. As can be seen in Figure 6.3(c) and (d), the degradation time scale of ZIFs with different $\langle \Delta\Delta E^{\ddagger} \rangle$ values differs significantly. In order to connect the simulated time scale to experiments, we need to calculate the time unit $\tau = 1/r_0$ of the kinetic model. r_0 can be obtained via transition state theory:

$$r_0 = c_{\mathrm{acid}} \cdot k_{\mathrm{TST}} = c_{\mathrm{acid}} \cdot \frac{k_B T}{h} e^{-\beta \Delta F^{\ddagger}} \tag{6.8}$$

where k_{TST} is the reaction rate from the adsorbed complex (where an acid molecule is adsorped to the reactive site) to the DL defect, $\Delta F^{\ddagger}$ is the site-averaged activation free energy of the formation of an isolated defect. The effect of acid gas distribution was included in the calculation of $\Delta F^{\ddagger}$: as discussed in Chapter 5, the acid gas distribution has a substantial influence on the apparent defect formation barriers. The pre-factor c_{acid} corrects for the concentration of acid species on each site. c_{acid} can be calculated by the Henry's constant K_{H} of dry SO_2 adsorption in ZIFs, assuming all

adsorbed SO_2 molecules convert to H_2SO_4 and no further adsorption after the conversion:

$$c_{\mathrm{acid}} = c_{\mathrm{SO_2}} = K_{\mathrm{H}}\, p_{\mathrm{SO_2}}\, \frac{M}{N_l} \tag{6.9}$$

where $p_{\mathrm{SO_2}}$ is the partial pressure of SO_2 in the exposure experiment, M is the molar volume of the unit cell and N_l is the number of linkers in a unit cell. The K_{H} and $\Delta F^{\ddagger}$ values are summarized in Table 6.2. With 20 ppm SO_2 concentration ($p_{\mathrm{SO_2}} = 2 \times 10^{-5}$ bar), we calculated that $\tau = 2.4 \times 10^4$ days for ZIF-8. Inserting τ back to Figure 6.3, we find that the degradation time scale is in ~ 200 years if $\langle \Delta\Delta E^{\ddagger} \rangle = 0$ and ~ 10 days if $\langle \Delta\Delta E^{\ddagger} \rangle = -0.28$. The later agrees well with the time scale observed experimentally: most unstable ZIFs completely degrades within 10 days of 20ppm humid SO_2 exposure.[34, 36, 38] A more quantitative comparison between the model prediction and the experimental results will be discussed later. The agreement in the predicted degradation time scale further supports our model. Our results also reveals that the neighboring interaction between the defects is particularly responsible for the fast degradation (in days) and the instability of ZIFs in acidic environments.

6.5 Evolution of Pore Volume

Despite the success in predicting the time dependence of defect concentration and the degradation time scale, this current kinetic model provides no information about how other properties are changed upon exposure. We are mostly interested in the time evolution of the pore volume and especially when will the pore volume decrease to zero. We hypothesize that there are two major reasons contributed to the pore volume loss. First, the defects break the connectivity and the rigidity of the framework and cause local structural collapse, which makes the local region inaccessible to other molecules. Second, when some region in the material is surrounded by locally collapsed "boundaries", the entire volume within the region is inaccessible regardless the defect concentration. We will denote this as the "boundary blocking effect" in subsequent text. Mathematically, we hypothesize that the volume loss V_{loss} can be expressed as a functional of the defect distribution $\rho(\mathbf{r})$:

$$V_{\mathrm{loss}}[\rho(\mathbf{r})] = \int_{V'} v(\rho(\mathbf{r}))\, \mathrm{d}\mathbf{r} + V_{\mathrm{blocked}} \tag{6.10}$$

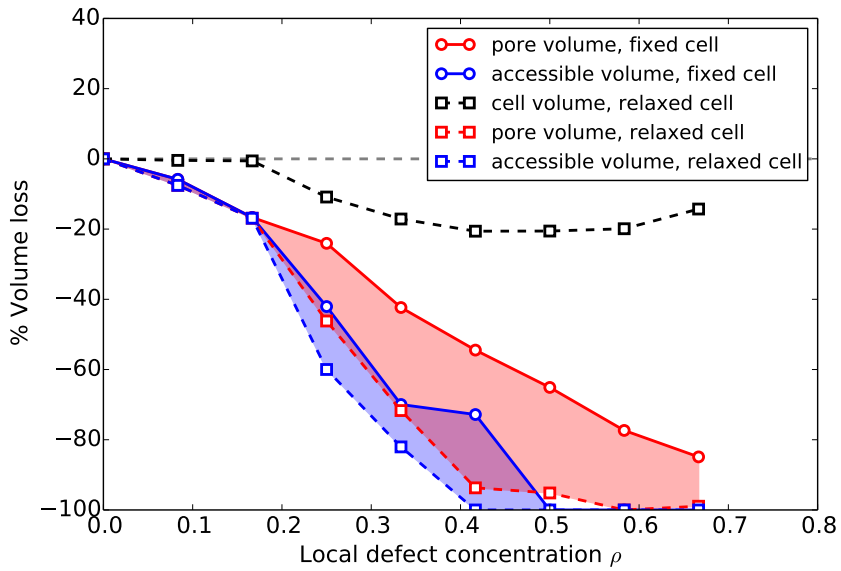


Figure 6.4: Volume loss as a function of local defect concentration in defective ZIF-8.

where $v(\rho(\mathbf{r}))$ is the volume loss per unit volume (due to local structural collapse). Here the "local density approximation" (LDA) as analogous to the LDA approximation in DFT is used, assuming that the integrand only depends on the local defect concentration instead of the defect distribution over all space. The boundary blocking effect is captured by directly adding the boundary-blocked volume to V_{loss}. Note that the boundary-blocked region should be excluded from the integral: $V' = V - V_{\text{blocked}}$.

We first examine the influence of local structural collapse on the pore volume by calculating $v(\rho)$. Practically, we add multiple H_2SO_4 induced DL defects into the primitive cell of ZIF-8, and calculate the local pore volume loss in this unit volume. By either fix the cell or allow the cell vectors to relax, we can calculate $v(\rho)$ in two limit cases. A fixed cell volume represents small defects clusters embedded in pristine framework, since the cell vector is constrained by the surrounding pristine cells. A relaxed cell represents regions inside a large defect cluster so the entire structure can be considered as defective. These limit cases thus determine the upper and lower bound of $v(\rho)$ during the defect propagation.

Figure 6.4 shows the total and accessible pore volume as functions of local defect concentration of the two limit cases. When the cell volume is fixed, as the local defect concentration increases, the pore volume decreases linearly, because the volume occupied by a defect is approximately constant. When $\rho > 0.2$, the defective linker and acid molecules start to block the apertures and channels that connect the pores, making certain portion of the pore inaccessible. The accessible volume thus decrease more rapidly than the total pore volume. When $\rho > 0.5$, the accessible pore volume decreases to zero. If the cell vectors is allowed to relax, the attractions between the linkers and the acid molecules tends to shrink the cell. Starting at $\rho > 0.2$, the cell volume starts to decrease and the total and accessible pore volumes start to lose drastically and reach zero at $\rho > 0.4$. At higher defect concentration $\rho > 0.6$, the cell volume starts to increase. Because there is no free space to put a new defect in, the cell must expand to fit the newly added defect. After all, when the cell is relaxed, both the total and the accessible pore volume decreases more rapidly than in the fixed cell cases. In general, the pore volume loss $v(\rho)$ will lie between these two limits, as shown as the shaded area in Figure 6.4. When ρ reaches $0.4 \sim 0.5$, the accessible pore volume decreases to zero. This concentration corresponds to $20 \sim 25\%$ Zn-N bond cleavage (each linker has two Zn-N bonds and the formation of DL defects only break one of them).

On the other hand, the boundary blocking effect is non-local in nature, it can only be evaluated in the kinetic model with much larger length scale than a primitive cell. To account for the accessibility of the sites, we modified the rate table Eq.6.5 by letting $r_{\text{site}} = \varepsilon r_0$ with $\varepsilon \ll 1$ if the site is inaccessible:

$$r_{\text{site}} = \begin{cases} kr_0, & \text{if has defective NN and is accessible} \\ \varepsilon r_0, & \text{is inaccessible} \\ r_0, & \text{otherwise} \end{cases} \tag{6.11}$$

The specific choice of ε has no influence on the short-to-medium time scale results since $\varepsilon \ll 1 \ll k$ and the defect cluster growth dominants the propagation; the *very* late-stage evolution will depend on ε, but this limit is of little practical interests. In the present implementation, ε is set to be 0.01. The details about how the local structural collapse and boundary blocking effects are

handled in the kinetic model are discussed in the Method section. The simulated results using this modified rate table are shown in Figure 6.5.

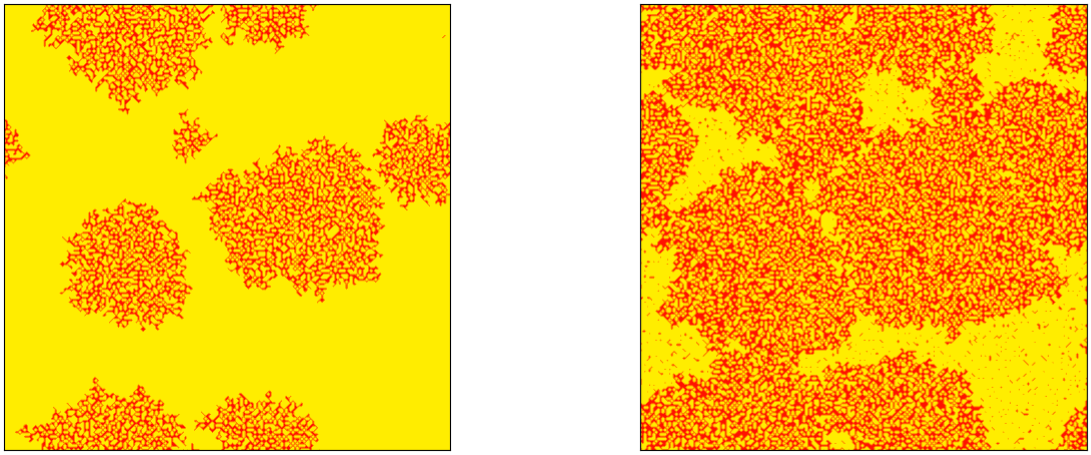
The influence of the pore-blocking effect can be seen by comparing Figure 6.5(a) and Figure 6.3(b). At the early stage, defects still forms isolated clusters, but adding pore-blocking effect prevents the continuous defect formation in the center of these clusters. As the propagation continues, the defect clusters can connect each other and surround some non-defective regions. These regions become inaccessible due to the boundary blocking and remains non-defective until all accessible sites are consumed (yellow regions in Figure 6.5(b)). Eventually, when the total concentration $c > 0.44$ (corresponding to 22% bond cleavage), all sites become inaccessible: the pore volume loses completely. In Figure 6.5(c), with the pore-blocking effect, the $c(t)$ curve has the same sigmoid shape but is "compressed". The power function dependence at early stage still holds. The $c(t)$ curve meets a plateau when $c > 0.44$, indicates that there are no unblocked sites left in the material. The time scale of reaching this plateau is the same as that without pore-blocking. Due to the nonzero value choice of ε, after reaching the plateau, the material continue to degrade as a first order reaction, but with a much slower rate εr_0. In summary, our analysis regarding the pore-blocking effect indicates that, the rigidity of the framework may affect the critical local concentration ρ_c for local structural collapse and boundary blocking, which determines the final percentage of the bond cleavage when the pore volume loses completely. However, this effect won't change the degradation time scale. The time scale only depends on c_{acid}, $\Delta F^{\ddagger}$ and $\Delta\Delta E^{\ddagger}$.

As shown in Figure 6.5(b), the material can be divided to the following four regions: locally collapsed cluster center, growing cluster surface, boundary-blocked regions and the accessible regions:

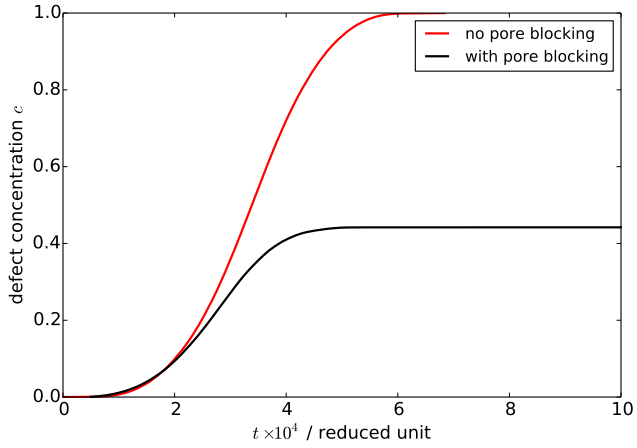
$$V = V_{\text{clu}} + V_{\text{surf}} + V_{\text{blocked}} + V_{\text{accessible}} \tag{6.12}$$

In the collapsed cluster center, $\rho > \rho_c$ and $v(\rho) = 1$, thus $V_{\text{loss,clu}} = V_{\text{clu}}$. Because of the dimension, V_{surf} can be ignored. As a result,

$$V_{\text{loss}} = V_{\text{clu}} + V_{\text{blocked}} = V - V_{\text{accessible}} \tag{6.13}$$

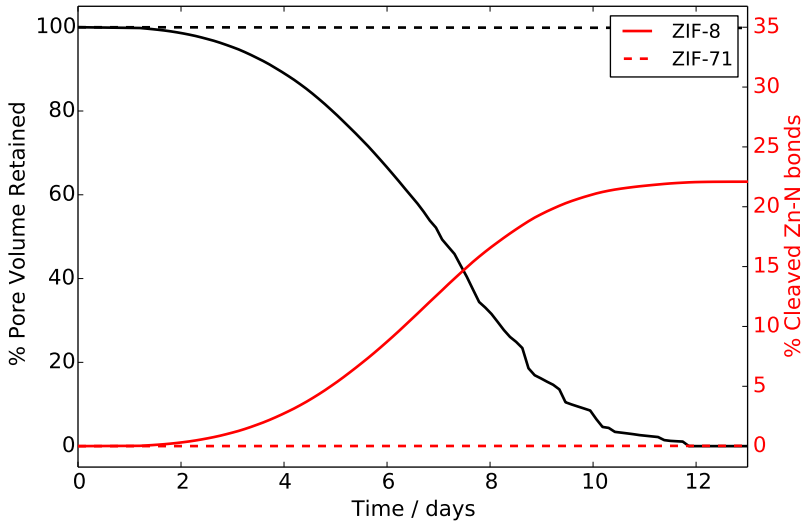


(a) (b)

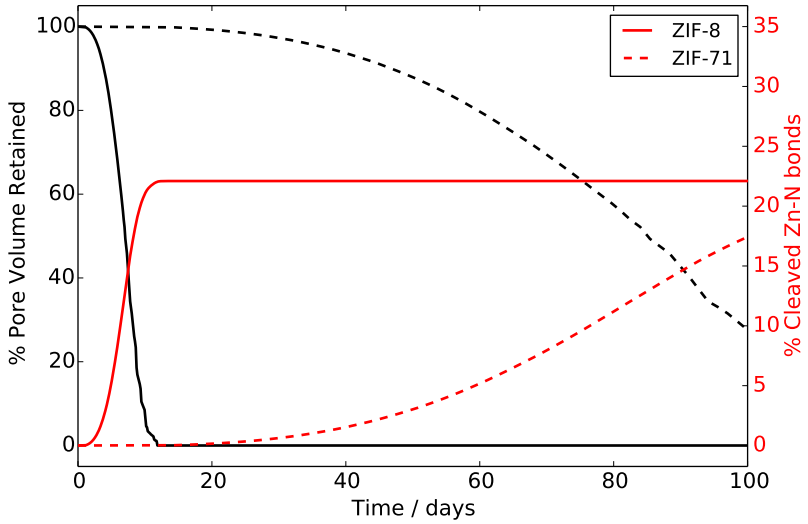


(c)

Figure 6.5: KMC simulation snapshots with the consideration of pore-blocking effect at (a) $c = 0.2$ and (b) $c = 0.5$. (c) The simulated time evolution of defect concentration with the consideration of pore-blocking effect. $\langle \Delta\Delta E^{\ddagger} \rangle = -0.28$ eV.



(a)



(b)

Figure 6.6: Percentage of pore volume retained (black curves) and percentage of cleaved Zn-N bonds (red curves) in ZIF-8 and ZIF-71 as functions of time, with shorter (a) and longer (b) time scales

In the lattice model, the volume is proportional to the number of sites. The percentage of volume loss can be calculated as:

$$\frac{V_{\text{loss}}}{V} = 1 - \frac{V_{\text{accessible}}}{V} = 1 - c_{\text{accessible}} \tag{6.14}$$

where $c_{\text{accessible}}$ is the concentration of the accessible sites. Based on this equation, the time evolution of the pore volume retained as well as the time evolution of bond cleavage in ZIF-8 and ZIF-71 are plotted in Figure 6.6 in real time units. The time unit τ of ZIF-71 is calculate via Eq.6.8 with parameters listed in Table 6.2. Here we assume that the $\Delta\Delta E^{\ddagger}$ value is the same in ZIF-71 as in ZIF-8. Figure 6.6(a) can be directly compared with experimental results. [36, 38]. Experiments show that with 20 ppm SO_2 exposure under $85 \sim 95\%$ relative humidity, ZIF-8 loses all its pore volume after 8-9 days of exposure, with $15 \sim 20\%$ of Zn-N bond cleavage, while on this time scale, ZIF-71 shows no detectable bond cleavage or pore volume loss under the same exposure conditions. On the other hand, our model predicts that with 20 ppm SO_2 exposure, ZIF-8 loses all its pore volume after 12 days of exposure, with 22% of Zn-N bond cleavage, and ZIF-71 undergoes essentially no structural change (0.02% bond cleavage) within 12 days of exposure. Despite the many approximations we made, our kinetic model shows semi-quantitative agreement with experimental results! Figure 6.6(b) shows the same time evolution as in Figure 6.6(a) but with longer time scales (100 days). We can see that ZIF-71 can also form defects and lose pore volume, yet the time is too long to be observed experimentally.

6.6 Conclusion

In this chapter, we developed a lattice-based kinetic model to simulate the long-time defect propagation in ZIFs. Experimentally testable predictions of material evolution can be generated by the model. With physically-meaningful parameters obtained from QM/MM calculations, our model predicts that the defect propagation process is autocatalytic. The predicted degradation kinetics (especially the time scale) of ZIF-8 and ZIF-71 in humid SO_2 show great agreement with experiments. Another important prediction by this model is that defects in ZIFs tend to form clusters. Currently the existence of the defect clusters have not been directly observed experimentally.

As NMR[119] and pulse dipolar spectroscopy[120] techniques has been used to measure the Al site distribution in zeolites at atomistic length to nanometer length scales, similar techniques may also be capable to determine the defect distribution in ZIFs.

Our model also sheds light on the possible factors that govern the material stability. The neighboring interaction between the defects ($\Delta\Delta E^{\ddagger}$) is particularly responsible for the fast degradation (in days) and the instability of ZIFs in acidic environments. This result suggests potential approaches to stabilize ZIF materials by hindering the catalytic formation of the neighboring defects.

Appendix: Rate Equation for Defect Propagation

As we have shown in the main text, in ZIFs the defects tend to form clusters. The new defect can either grow on a existing defect cluster, or nucleate to form a new cluster. Let c be the defect concentration in the material, the rate equation of defect propagation can be written as:

$$\frac{\mathrm{d}c}{\mathrm{d}t} = r_{\mathrm{growth}}(c) + r_{\mathrm{nucleation}}(c) \tag{6.15}$$

Without the pore-blocking effect, the defect formation rate on each site is:

$$r_{\mathrm{site}} = \begin{cases} kr_0, & \text{if has defective NN} \\ r_0, & \text{otherwise} \end{cases} \tag{6.16}$$

The growth/nucleation rate can be calculated as r_{site} times the concentration of sites available for growth/nucleation. When $\Delta\Delta E^{\ddagger} \ll 0$ ($k = e^{-\beta\Delta\Delta E^{\ddagger}} \gg 1$), $r_{\mathrm{nucleation}}$ can be ignored. Let c_{clu} be concentration of the defect clusters, the average size of the defect clusters is then:

$$M = \frac{c}{c_{\mathrm{clu}}} \tag{6.17}$$

the average length scale of the defect clusters is:

$$l \propto M^{1/d} = \left(\frac{c}{c_{\mathrm{clu}}}\right)^{1/d} \tag{6.18}$$

where d is the dimension of the space ($d = 3$ in real materials and $d = 2$ in the lattice model). In general, for the growth problems, the surface of the grown cluster exhibits fractal. Let D be the

fractal dimension of the cluster surface ($d-1<D<d$), the surface area of the cluster is:

$$S \propto l^D \propto \left(\frac{c}{c_{\text{clu}}}\right)^{D/d} \tag{6.19}$$

At the early stage of the growth, the defect clusters are separated and have no overlap or connection with each other, the number of available growth sites is the number of clusters times the number of growth sites on a single cluster, the later can be calculated as the surface area of the cluster times the number of sites per unit area σ (a constant), and the growth rate is:

$$r_{\text{growth}}(c) = c_{\text{clu}}\sigma S \cdot kr_0 \propto c_{\text{clu}} \cdot \left(\frac{c}{c_{\text{clu}}}\right)^{D/d} \cdot kr_0 \tag{6.20}$$

Since $k \gg 1$ and the nucleation rate can be ignored, there is no nucleation event and c_{clu} is a constant. We can then write the rate equation as:

$$\frac{\mathrm{d}c}{\mathrm{d}t} = A \cdot kr_0 \cdot c^{\frac{D}{d}} \qquad \text{at early stage} \tag{6.21}$$

where A is a constant. The solution of this equation is:

$$kr_0 t = \frac{1}{A\left(1-\frac{D}{d}\right)} \cdot c^{1-\frac{D}{d}} \tag{6.22}$$

$$c = A'(kr_0 t)^{\frac{d}{d-D}} \tag{6.23}$$

where A' is another constant. Eq.6.23 shows that at the early stage, defect concentration increases as a *power function* of t. Since $d-1<D<d$, $\frac{d}{d-D}>d>1$, which means *the defect formation rate increases with time*. In other words, the defect formation in ZIFs is an autocatalytic reaction.

However, the defect formation rate cannot increase boundlessly, at late stage, the grown cluster will eventually overlap and eliminate some of their surfaces. As a result, the number of available growth sites is less than the sum of the surface sites of all clusters. One limit case is that all remaining sites have a defective nearest neighbor. The rate equation becomes:

$$\frac{\mathrm{d}c}{\mathrm{d}t} = kr_0(1-c) \qquad \text{at late stage} \tag{6.24}$$

with solution:

$$c = 1 - \exp[-kr_0(t - t_0)] \tag{6.25}$$

In conclusion, for the case $k \gg 1$, the defect formation rate first increases as a power function of time and eventually decrease exponentially:

$$c(t) = \begin{cases} A(kr_0t)^{\lambda}\,, & \text{early stage, } \lambda > 1 \\ 1 - \exp[-kr_0(t - t_0)]\,, & \text{late stage} \end{cases} \tag{6.26}$$

When considering the pore-blocking effect, the center of the cluster has a non-unity concentration ρ. The average size of the defect clusters

$$M = \frac{c}{\rho c_{\text{clu}}} \propto c \tag{6.27}$$

is still proportional to c. The subsequent derivation for the early stage growth and the conclusions are thus still valid.

Chapter 7

Acid Gas Adsorption in N-modified UiO-66 type MOFs

This chapter is reproduced in part from an unpublished collaborative work by Chengzhai Wang, Kai Cui, Jessica Rimsza, Tina Nenoff, J. R. Schmidt and Krista S. Walton with permission. All experimental results are collected by Chengzhai Wang. All computational data discussed in this chapter are collected by Kai Cui.

7.1 Introduction

In the previous chapters we have performed several mechanistic studies and gained deep physical insight into the acid-induced degradation of ZIFs. We suggest possible routes to increase the stability of ZIFs by controlling the acid gas distribution in pore and by suppressing the autocatalytic behavior in the propagation process. On the other hand, an alternative way to resolve the acid-induced degradation problem of ZIFs is to remove the acid contaminants in the first place. In this chapter, we shift our focus to the *removal of acid contaminants* using UiO-66 type MOFs.

According to the hard/soft acid base theory, MOFs built by hard Lewis acids and hard Lewis bases, or soft Lewis acids and soft Lewis bases, are more stable than those constructed by mixing hard/soft acids/bases[41]. UiO-66 and ZIFs are two representative examples. MOFs within the UiO-66 family are constructed by hard high-valency Zr^{4+} ions with hard carboxylate-based BDC ligands, while ZIFs are constructed by soft low-valency Zn^{2+} ions and soft imizazolate linkers. Both materials have shown excellent stabilities are compared to other MOFs[3, 8]. However, the pK_{a} values of carboxylate-based ligands is much lower than that of the imidazolate ligands. This lower pK_{a} disfavors the protonation of the BDC ligands, which consequently disfavors the nucleophilic substitution of the linkers by acid gases. Therefore, UiO-66 exhibits high stability in the pH range $\sim$1-7[40]. The higher connectivity in UiO-66 (coordination number = 12) of the

nodes in UiO-66 also disfavors the defect formation[41]. As a result, the acid-resistence feature makes UiO-66 excellent candidates for acid gas removal.

In this chapter we focus on the removal SO_2 from fuel gas. Due to the low concentration, ideal adsorbents must have high enough uptake of SO_2 at low partial pressure, and high selectivity over other gases (like CO_2). Because SO_2 is a Lewis acid or electron acceptor, one could potentially increase the uptake and selectivity by introducing electron donating Lewis base sites into the adsorbents to enhance the binding of SO_2. In this work we test the SO_2 adsorption performance of nitrogen-modified UiO-66 type MOFs. Compared to the parent UiO-66, nitrogen atoms with lone pairs (e.g. pyridine-N, -NH_2, etc.) are added to the linker. First principles calculations are used to understand the role played by the N atoms in these modified MOFs. We also attempt to use high-accuracy QM/QM methods along with the multisite Langmuir adsorption model to quantitatively predict the experimental adsorption isotherm and the Henry's constant.

7.2 SO_2 Binding Energy in N-Modified UiO-66

Several nitrogen-containing linkers were used to replace the original BDC linker in the synthesis of N-modified UiO-66. The performance of these N-modified MOFs were initially examined experimentally via SO_2 breakthrough capacity and adsorption isotherm measurements. The stability of these modified MOFs in SO_2 environment are also tested. Three modified MOFs with 2-amino-1,4-benzenedicarboxylic acid (BDC-NH_2), 2,5-pyridinedicarboxylic acid (2,5-PDC) and 2,5-pyrazine-dicarboxylic acid (2,5-PziDC) linkers (see Figure 7.1 for structures) showed enhanced SO_2 adsorption capacity as compared to the parent UiO-66 and remain intact upon exposure. These MOFs are denoted as UiO-66-NH_2, UiO-66-2,5-PDC and UiO-66-2,5-PziDC respectively. We use periodic DFT to calculate the binding energies of SO_2 in these MOFs. All calculations were performed in VASP[94, 95], using dispersion corrected PBE-D3[96, 97] with a plane wave basis and a 600 eV energy cutoff. Geometries of the bare adsorbents are first optimized. Both the lattice constant and atomic positions are relaxed until the forces on atoms are less than 0.05 eV/Å. SO_2 adsorbate is then added into the unit cell of the MOFs at different sites. The

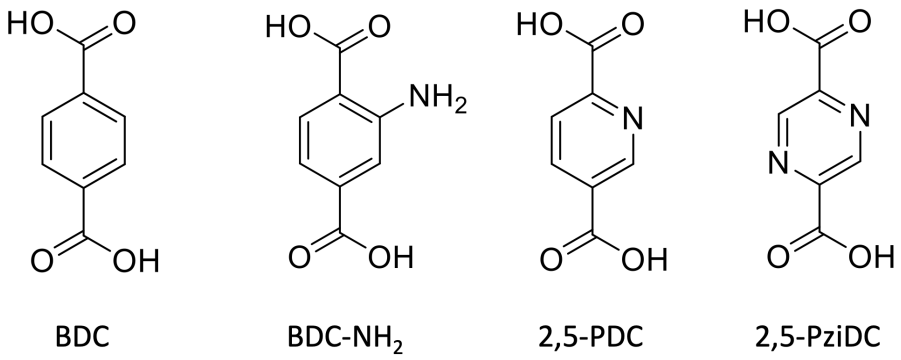
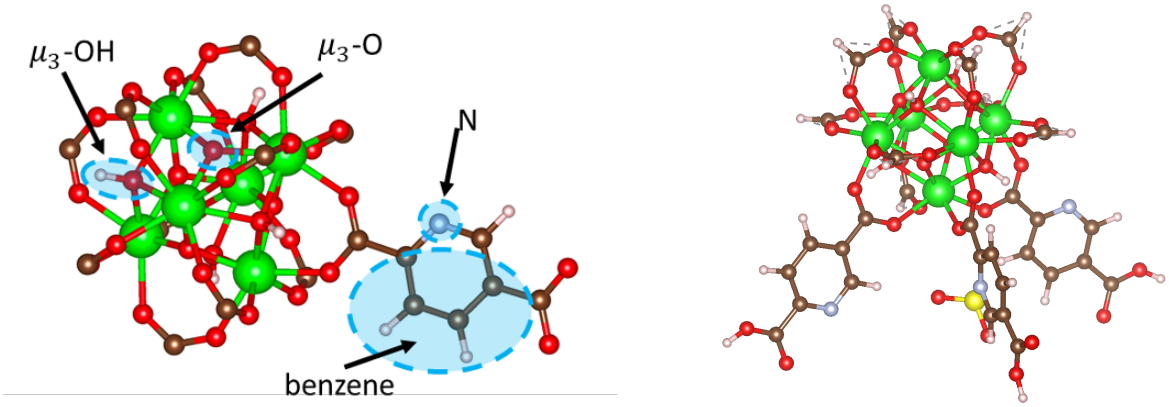


Figure 7.1: Structures of the linkers used to build parent and N-modified UiO-66 MOFs.

geometries of these adsorbed complexes are optimized using the same methods but with lattice constant fixed.

Based on the optimized structures, we identified four potential binding sites for SO_2, namely the μ_3-OH site, μ_3-O site, benzene (or pyridine/pyrazine) ring and the N atom. The structures of these sites are shown in Figure 7.2(a), and the total number of these sites in a unit cell is listed in Table 7.1. In the unmodified UiO-66, μ_3-OH is the most favorable site for SO_2 to bind to, with a binding energy of -0.47 eV. N-modification doesn't change the binding strength on the non-nitrogen sites. The increased binding strength on the OH site observed in UiO-66-NH_2 is likely from the long-range dispersion interaction due to the increased density of the framework with NH_2 functionalization. For all three N-modified MOFs, the N site has the strongest binding energy, ranging from -0.53 to -0.63 eV. The relative order of the binding strength on the N sites in the three MOFs shows apparent accordance with their experimental SO_2 adsorption capacity order: SO_2 bind most strongly in UiO-66-2,5-PDC, which shows the largest capacity. The binding strengths in UiO-66-2,5-PziDC and UiO-66-2,5-NH_2 are similar, and their capacities are also similar. All three modified MOFs show enhanced capacity than the parent UiO-66. These binding energy data concludes that the modification of the linkers introduces strong-binding N sites but has negligible influence on the binding at other sites. The enhanced SO_2 adsorption capacity after modification is indeed due to addition of the N sites.



(a) (b)

Figure 7.2: (a) Possible SO_2 binding sites in N-modified UiO-66. (b) structure of the ML_3 cluster model.

Site	μ_3-OH	μ_3-O	benzene	N
UiO-66	16	16	48	0
UiO-66-2,5-PDC	16	16	48	24
UiO-66-2,5-PziDC	16	16	48	48
UiO-66-NH_2	16	16	48	24

Table 7.1: The number of sites per unit cell of μ_3-OH, μ_3-O, benzene and N sites in the parent and N-modified UiO-66 MOFs.

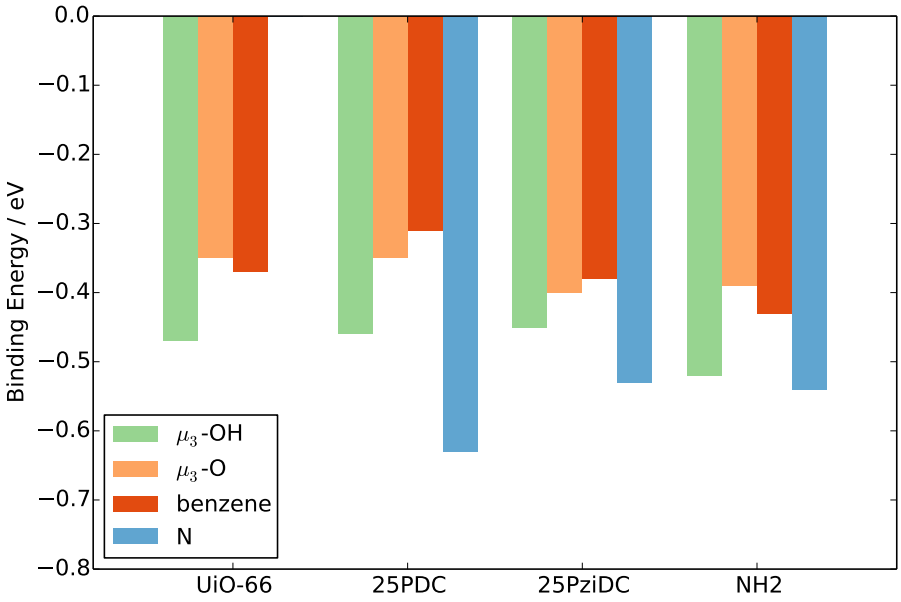


Figure 7.3: Binding energy of SO_2 at μ_3-OH, μ_3-O, benzene and N sites in the parent and N-modified UiO-66 MOFs. .

7.3 Pyridine-N vs. Pyrazine-N

One interesting observation from the adsorption measurement is that UiO-66-2,5-PziDC has twice as much of N sites than UiO-66-2,5-PDC (see Table 7.1), but its adsorption capacity is lower. This observation is apparently explained by the binding energy calculation: though there are more N sites in UiO-66-2,5-PziDC, the binding strength on each sites is much smaller ($\Delta\Delta E_{\mathrm{bind}} = 0.1$eV) than in UiO-66-2,5-PDC. In this section, we will seek the origins of this difference in the binding energies.

We build cluster models to decompose the short-range and long-range interactions. Clusters are carved from the optimized periodic structure with no further optimization. The cluster model contains the $Zr_6O_4(OH)_4(COO)_{12}$ SBU and three explicit BDC (or modified BDC) linkers that are closest to the adsorbates. The other nine linkers that also coordinate to the metal cluster is replaced by $HCOO^-$ groups. Hydrogen atoms are added to the dangling COO^- group in the explicitly modeled linkers to ensure the neutrality of the cluster. A schematic illustration of this cluster model (denote as ML_3 cluster) is shown in Figure 7.2(b). SO_2 binding strengths on the N sites in UiO-66-2,5-PDC and UiO-66-2,5-PziDC were recalculated using the cluster model. For periodic calculation of the clusters, the length of the cell is set to 30Å. The results are summarized in Table 7.2. Numbers in parentheses are calculated using the same setup but exclude the D3 dispersion correction. For the calculations in the entire unit cell, $\Delta\Delta E_{\mathrm{bind}}$ is almost the same with and without the dispersion correction, which indicates that the $\Delta\Delta E_{\mathrm{bind}}$ between UiO-66-2,5-PDC and UiO-66-2,5-PziDC is not from the dispersion interactions. When a cluster model (either the ML_3 cluster or only a linker) is used, $\Delta\Delta E_{\mathrm{bind}}$ decreases to 0.02 eV and the exclusion of dispersion interaction still makes no difference. Therefore, the majority of the $\Delta\Delta E_{\mathrm{bind}}$ (0.08 eV) is from the long range electrostatic interaction. The 2,5-PDC linker has a permanent dipole moment, and the arrangement of the linkers in the framework can generate electrostatic fields to attract the polar SO_2 molecule from a long range. Such electrostatic fields is absent in the UiO-66-2,5-PziDC case due to the zero dipole moment of the linker.

System	unit cell	ML_3 cluster	linker only
UiO-66-2,5-PDC	-0.63 (-0.35)	-0.53 (-0.31)	-0.41
UiO-66-2,5-PziDC	-0.53 (-0.26)	-0.51 (-0.29)	-0.39
$\Delta\Delta E_{\mathrm{bind}}$	0.10 (0.09)	0.02 (0.02)	0.02

Table 7.2: Binding energies (in eV) of SO_2 on the N sites in UiO-66-2,5-PDC and UiO-66-2,5-PziDC recalculated using periodic and cluster models. Numbers in parentheses are calculated using the same setup but exclude the D3 dispersion correction.

System	$E^{(2)}$(LP-N$\rightarrow$RY*-S) / $\mathrm{kcal} \cdot \mathrm{mol}^{-1}$		
2,5-PDC-SO_2	0.42	0.14	0.06
2,5-PziDC-SO_2	0.33	0.13	0.05

Table 7.3: Stabilization energies $E^{(2)}$ between the lone pair electrons of the N atom (LP-N) in the linker to the Rydberg orbitals of the S (RY*-S) atom in the SO_2.

The remaining 0.02 eV in $\Delta\Delta E_{\mathrm{bind}}$ is from the short-range electrostatic interactions. We then perform NBO analysis[55] to characterize this donor-acceptor interactions between SO_2 and the framework in UiO-66-2,5-PDC and UiO-66-2,5-PziDC. In the NBO analysis, the "stabilization energy" $E^{(2)}$ is used to describe the interaction strength due to the electron delocalization between a donor and an acceptor orbital pairs. Larger value of $E^{(2)}$ indicates stronger interaction[121]. We performed NBO analysis for the gas phase SO_2-linker associates at PBE level[96] with 6-31G(d) basis sets[122] in Gaussian 16[123]. The calculated stabilization energies between the lone pair electrons of the N atom in the linker to the Rydberg orbitals (unoccupied orbitals) of the S atom in the SO_2 are listed in Table 7.3. Clearly, the $E^{(2)}$ is systematically larger in 2,5-PDC-SO_2 system than in 2,5-PziDC-SO_2, indicating that the donor-acceptor interaction between the linker N atom and the adsorbate S atom is stronger in the 2,5-PDC case. This observation agrees with the calculated natural charge on the N atoms of the two linkers: N atom in 2,5-PDC has more negative charge (-0.4213) than that in 2,5-PziDC (-0.3815). The weaker short-range interaction in UiO-66-2,5-PziDC is likely also due to the symmetric distribution of the two N atoms, which leads to a less concentrated electron density on each N site.

7.4 Towards Quantitative Prediction of the Adsorption Free Energy

The binding energies cannot be directly measured in experiments. Our next objective is to estimate the adsorption free energy and Henry's constant using these binding energy data. At relatively low SO_2 partial pressure (< 1 bar in the experiments), it is safe to ignore the interaction between the SO_2 molecules in the adsorption phase. We can thus use the Langmuir model described in Chapter 3 and Chapter 5 to calculate the isotherm and the Henry's constant. To recap the details, we first calculate the binding free energy ΔG_s of SO_2 on each site s in the MOF. The uptake at a given pressure p is:

$$n(p) = \sum_s \frac{c_s K_s p}{1 + K_s p} \tag{7.1}$$

where c_s is the concentration of site kind s in the MOF, $K_s = \exp(-\beta\Delta G_s)$ is the equilibrium constant of binding. The Henry's constant and the adsorption free energy can be calculated accordingly:

$$K_{\mathrm{H}} = \sum_s c_s K_s = \sum_s \frac{c_s}{p^{\ominus}} \exp(-\beta\Delta G_s) \tag{7.2}$$

$$\Delta G^{\ominus}_{\mathrm{ad}} = -k_B T \ln \frac{K_{\mathrm{H}} p^{\ominus}}{c^{\ominus}} \tag{7.3}$$

We first calculate the adsorption free energy in the parent and N-modified UiO-66 at PBE-D3 level (using the binding energy data in Figure 7.3). PES-scan method as described in Chapter 5 is used to calculate the entropy loss upon adsorption. The data are shown in the middle column in Table 7.4. We find that the Langmuir model provides excellent agreement with the experimental results (obtained by fitting the measured isotherm to a dual-site Langmuir equation) in the absolute adsorption free energy. The errors are within 0.01-0.03 eV. This Langmuir model also predicts a correct order in $\Delta G^{\ominus}_{\mathrm{ad}}$. All three N-modified MOFs has more negative adsorption free energy than the parent. UiO-66-2,5-PDC has the most negative $\Delta G^{\ominus}_{\mathrm{ad}}$ and thus the largest capacity.

System	Experiment	PBE-D3	LMP2/PBE-D3
UiO-66	-0.11	-0.14	-0.12
UiO-66-2,5-PDC	-0.18	-0.17	-0.13
UiO-66-2,5-PziDC	-0.16	-0.15	-0.15
UiO-66-NH_2	-0.15	-0.17	-0.16

Table 7.4: Calculated adsorption free energy $\Delta G^{\ominus}_{\mathrm{ad}}$ (in eV) via Langmuir model at PBE-D3 and LMP2/PBE-D3 level for the parent and N-modified UiO-66 MOFs.

The above results show the great success of using the Langmuir model for calculating the adsorption behavior in MOFs. We further seek more quantitative predictions of the adsorption free energy. For such calculations, more accurate correlated wave function methods are needed. We adapt the QM/QM method described in Chapter 4 to perform the high-level calculations. The ML_3

cluster is used as the high-level region and will be treated at the LMP2 level[46]. The long-range interactions are sill treated at PBE-D3 level. The total energy of the system is:

$$E_{\mathrm{tot}}^{\mathrm{LMP2/PBE-D3}} = E_{\mathrm{clu}}^{\mathrm{LMP2}} - E_{\mathrm{clu}}^{\mathrm{PBE-D3}} + E_{\mathrm{tot}}^{\mathrm{PBE-D3}} \tag{7.4}$$

All cluster calculations were performed in Molpro 2019 package[99]. An effective core potential (ECP28MWB)[124] is used for Zr. Triple zeta ECP-based correlation consistent basis sets of Peterson and co-workers (VTZ-PP) [125] is used for Zr and Dunning style correlation consistent basis sets [49] is used for other atoms (VTZ for H, C, AVTZ for N, S, O). When performing LMP2 calculations, the Boys localization scheme [53] is used.

Our data suggests that for most cases, PBE-D3 tends to overestimate the binding energy of SO_2 in MOFs. The largest deviation is observed at N sites, where $\Delta\Delta E = \Delta E_{\mathrm{LMP2}} - \Delta E_{\mathrm{PBE}}$ can be as large as 0.10-0.14 eV. This large difference is probably from the well-known overdelocalization problem of GGA functionals. To calculate the binding free energy at LMP2 level, the entropy change must also be corrected. But even with the QM/QM scheme, scanning all the 1D-PESs at LMP2 level is too expensive. Fortunately, we found a linear relationship between the ΔS and ΔE of binding in these MOF systems. Based on the calculated ΔS and ΔE data at the PBE level, an empirical relation $T\Delta\Delta S \simeq 0.6066\Delta\Delta E$ is used to estimate the LMP2 correction to the entropy loss. With the LMP2 corrections, the calculated $\Delta G_{\mathrm{ad}}^{\ominus}$ are listed in the last column in Table 7.4. We achieve quantitative prediction of the $\Delta G_{\mathrm{ad}}^{\ominus}$ for UiO-66, UiO-66-2,5-PziDC and UiO-66-NH_2 with errors less than 1 kJ/mol. However, the prediction for UiO-66-2,5-PDC become worse. Possible reasons could be: (1) the empirical correction of entropy loss fails and (2) we missed some sites or configurations that are important at LMP2 level. Note that the calculation of the Henry's constant and $\Delta G_{\mathrm{ad}}^{\ominus}$ is equivalent to calculate the absolute value of the partition function. Missing any sites or configurations leads to an underestimation of the total partition function, and thus a more positive prediction of the adsorption free energy.

7.5 Conclusion

We have computationally investigated the SO_2 adsorption behavior in the parent and three N-modified UiO-66 type MOFs. The binding energies of SO_2 at four possible sites were calculated via periodic DFT. We find that the newly added N sites through modification are the strongest binding sites. Meanwhile, the modification have negligible influence on the binding strength on other sites. Therefore, the enhanced SO_2 adsorption capacity is directly contributed by the presence of N sites. These calculated binding energy data can be used as input in a Langmuir adsorption model to calculate the adsorption free energy. At PBE-D3 level, the results give the correct order in the SO_2 uptake as in experiments. If an LMP2 correction is added, quantitative prediction can be achieved for some of the MOFs.

We also closely compared the binding of SO_2 to the N sites in UiO-66-2,5-PDC and UiO-66-2,5-PziDC. The former contains a pyridine ring in each linker and the later contains a pyrazine ring. The symmetric distribution of the two N atoms in the 2,5-PziDC results in less concentrated electron density on each N site and thus weaker donor-acceptor interaction between the N site and the SO_2. This symmetry of 2,5-PziDC also leads to a zero dipole moment, which eliminates any long-range electrostatic attraction to the polar SO_2 adsorbates. Consequently, the binding strength of SO_2 to the pyrazine-N is much weaker than that to the pyridine-N, and the adsorption capacity of SO_2 in UiO-66-2,5-PziDC is lower than UiO-66-2,5-PDC, despite its larger number of N sites. Our results highlight the influence of chemical environment of MOFs on their adsorption capacities.

Chapter 8

Conclusion and Future Directions

In this thesis, we performed detailed computational investigations on the acid gas induced degradation in MOFs. We have developed several novel computational methods and models to simulate the degradation process. The multiscale QM/MM and QM/QM methods enables reactivity calculation of MOF systems with "gold standard" accuracy but affordable costs. Thermodyanmics and kinetics calculations of the defect formation in large supercells are made feasible. The QM/MM and QM/QM methods also directly benefit the development of statistical mechanics-based multisite Langmuir model for calculating the adsorption behaviors of MOFs. Based on the binding free energies from the QM/MM calculations, the distribution of gas adsorbates in MOF pores can be evaluated at coupled cluster level, and the resulted overall adsorption free energies can achieve quantitative agreement with experiments. The combination of these methods gives accurate site-averaged reactivity of the MOFs, which can serve as an indicator of the overall kinetic stability of the material. Finally, this site-averaged reactivity data can be used as input in a lattice-based kinetic model to simulate the long-time macroscopic evolution of the materials. We can successfully reproduce the experimentally-measured macroscopic evolution of the time-dependent bulk materials proprieties like the pore volume loss.

From the computational data, we have gain new insights into the possible factors that govern the material stability. Acid gas distribution can significantly alter the defect formation mechanism. The preferred binding of H_2SO_4 to an inert aperture site in ZIF-71 leads to an indirect mechanism. In ZIF-71, the defect formation is initiated by the migration of the acid gas from the inert binding site to the more reactive site, which costs additional free energy and thus makes ZIF-71 kinetically stable under humid SO_2 exposure. Due to the release of lattice strain, the formation of additional defects near the existed ones are both thermodynamically and kinetically more favored.

As revealed by the kinetic model, this neighboring interaction between defects leads to the autocatalytic behavior of defect propagation, and is particularly responsible for the fast degradation (in several days) of the ZIFs. The kinetic modeling also yields important insights regarding the spatial distribution of defects.

Our results also suggest new avenues to control the degradation of MOFs under acid gas exposure. For example, we can tune acid gas distribution by choosing the appropriate topology and linker functionalization to attract acid reactants to an inert binding site and enable the indirect defect formation mechanism to increase the kinetic stability. Secondly, one could reduce or distribute the lattice strain associated with defect formation to suppress the autocatalytic feature of defect propagation, thus slow down the macroscopic degradation. An alternative way to mitigate material degradation is to remove the acid gases beforehand. We also discussed the opportunity of adding Lewis base sites to acid-resistant MOFs to enhance the acid gas adsorption.

To conclude this thesis, we discuss some possible directions for future research. For example, the QM/MM method is capable to generate large data sets for the training of a reactive force field or a semi-empirical method at high accuracy level (e.g., coupled cluster level). With these cheaper theories, we can directly assess the dynamics effects (like the diffusion) in the defect formation processes. Our statistical mechanics-based adsorption model for calculating acid-water mixture distribution was originally designed for hydrophobic MOFs with low water loading. We can extend this model to hydrophilic ZIFs by treating adsorbed water as solvent. A solvation free energy correction can be added to the binding free energies of the acids and the model can be used without other modifications. The reference interaction site model (RISM)[126] is a powerful tool to efficiently calculate solvation free energy without MD sampling while keep all the microscopic details of the water (e.g., hydrogen bonding). The current kinetic model uses a simplified square lattice to represent the ZIF and the re-connection between the metal and the linker is not considered. We can build a lattice with more complex topology as in real MOFs and allow the healing of the defects. Such model would be capable to simulate the acid gas-induced phase change, and solvent-assisted crystal redemption (SACRed)[127] of MOFs.

Bibliography

(1) Yaghi, O. M.; Li, G.; Li, H. Selective Binding and Removal of Guests in a microporous Metal-Organic Framework. *Nature* **1995**, *378*, 703–706.

(2) Zhou, H.-C.; Long, J. R.; Yaghi, O. M. Introduction to Metal-Organic Frameworks. *Chem. Rev.* **2012**, *112*, 673–674.

(3) Phan, A.; Doonan, C. J.; Uribe-Romo, F. J.; Knobler, C. B.; O'Keeffe, M.; Yaghi, O. M. Synthesis, Structure, and Carbon Dioxide Capture Properties of Zeolitic Imidazolate Frameworks. *Acc. Chem. Res.* **2010**, *43*, 58–67.

(4) 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*, 1148–1150.

(5) Guo, H.; Lin, F.; Chen, J.; Li, F.; Weng, W. Metal-Organic Framework MIL-125(Ti) for Efficient Adsorptive Removal of Rhodamine B from Aqueous Solution. *Appl. Organometal. Chem.* **2015**, *29*, 12–19.

(6) Rada, Z. H.; Abid, H. R.; Shang, J.; He, Y.; Webley, P.; Liu, S.; Sun, H.; Wang, S. Effects of Amino Functionality on Uptake of CO_2, CH_4 and Selectivity of CO_2/CH_4 on Titanium Based MOFs. *Fuel* **2015**, *160*, 318–327.

(7) Cavka, J. H.; Jakobsen, S.; Olsbye, U.; Guillou, N.; Lamberti, C.; Bordiga, S.; Lillerud, K. P. A New Zirconium Inorganic Building Brick Forming Metal Organic Frameworks with Exceptional Stability. *J. Am. Chem. Soc.* **2008**, *130*, 13850–13851.

(8) Winarta, J.; Shan, B.; Mcintyre, S. M.; Ye, L.; Wang, C.; Liu, J.; Mu, B. A Decade of UiO-66 Research: A Historic Review of Dynamic Structure, Synthesis Mechanisms, and Characterization Techniques of an Archetypal Metal–Organic Framework. *Cryst. Growth Des.* **2020**, *20*, 1347–1362.

(9) Wu, H.; Chua, Y. S.; Krungleviciute, V.; Tyagi, M.; Chen, P.; Yildirim, T.; Zhou, W. Unusual and Highly Tunable Missing-Linker Defects in Zirconium Metal–Organic Framework UiO-66 and Their Important Effects on Gas Adsorption. *J. Am. Chem. Soc.* **2013**, *135*, 10525–10532.

(10) Chen, B.; Eddaoudi, M.; Reineke, T. M.; Kampf, J. W.; O'Keeffe, M.; Yaghi, O. M. Cu_2(ATC)· $6H_2O$: Design of Open Metal Sites in Porous Metal-Organic Crystals (ATC: 1,3,5,7-Adamant-ane Tetracarboxylate). *J. Am. Chem. Soc.* **2000**, *122*, 11559–11560.

(11) Furukawa, H.; Cordova, K. E.; O'Keeffe, M.; Yaghi, O. M. The Chemistry and Applications of Metal-Organic Frameworks. *Science* **2013**, *341*, 1230444.

(12) Eddaoudi, M.; Kim, J.; Rosi, N.; Vodak, D.; Wachter, J.; O'Keeffe, M.; Yaghi, O. M. Systematic Design of Pore Size and Functionality in Isoreticular MOFs and Their Application in Methane Storage. *Science* **2002**, *295*, 469–472.

(13) Ortiz, A. U.; Freitas, A. P.; Boutin, A.; Fuchs, A. H.; Coudert, F.-X. What Makes Zeolitic Imidazolate Frameworks Hydrophobic or Hydrophilic? The Impact of Geometry and Functionalization on Water Adsorption. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9940–9949.

(14) Furukawa, H.; Go, Y. B.; Ko, N.; Park, Y. K.; Uribe-Romo, F. J.; Kim, J.; O'Keeffe, M.; Yaghi, O. M. Isoreticular Expansion of Metal-Organic Frameworks with Triangular and Square Building Units and the Lowest Calculated Density for Porous Crystals. *Inorg. Chem.* **2011**, *50*, 9147–9152.

(15) Hönicke, I. M.; Senkovska, I.; Bon, V.; Baburin, I. A.; Bönisch, N.; Raschke, S.; Evans, J. D.; Kaskel, S. Balancing Mechanical Stability and Ultrahigh Porosity in Crystalline Framework Materials. *Angew. Chem. Int. Ed.* **2018**, *57*, 13780–13783.

(16) Farha, O. K.; Eryazici, I.; Jeong, N. C.; Hauser, B. G.; Wilmer, C. E.; Sarjeant, A. A.; Snurr, R. Q.; Nguyen, S. T.; Yazaydın, A. Ö.; Hupp, J. T. Metal-Organic Framework Materials with Ultrahigh Surface Areas: Is the Sky the Limit? *J. Am. Chem. Soc.* **2012**, *134*, 15016–15021.

(17) 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*, 108–121.

(18) Banerjee, R.; Phan, A.; Wang, B.; Knobler, C.; Furukawa, H.; O'Keeffe, M.; Yaghi, O. M. High-Throughput Synthesis of Zeolitic Imidazolate Frameworks and Application to CO_2 Capture. *Science* **2008**, *319*, 939–943.

(19) Wang, B.; Côté, A. P.; Furukawa, H.; O'Keeffe, M.; Yaghi, O. M. Colossal Cages in Zeolitic Imidazolate Frameworks as Selective Carbon Dioxide Reservoirs. *Nature* **2008**, *453*, 207–211.

(20) Sumida, K.; Rogow, D. L.; Mason, J. A.; McDonald, T. M.; Bloch, E. D.; Herm, Z. R.; Bae, T.-H.; Long, J. R. Carbon Dioxide Capture in Metal-Organic Frameworks. *Chem. Rev.* **2012**, *112*, 724–781.

(21) Kuppler, R. J.; Timmons, D. J.; Fang, Q.-R.; Li, J.-R.; Makal, T. A.; Young, M. D.; Yuan, D.; Zhao, D.; Zhuang, W.; Zhou, H.-C. Potential Applications of Metal-Organic Frameworks. *Coord. Chem. Rev.* **2009**, *253*, Functional Hybrid Nanomaterials, 3042–3066.

(22) Czaja, A. U.; Trukhan, N.; Müller, U. Industrial Applications of Metal-Organic Frameworks. *Chem. Soc. Rev.* **2009**, *38*, 1284–1293.

(23) Rosi, N. L.; Eckert, J.; Eddaoudi, M.; Vodak, D. T.; Kim, J.; O'Keeffe, M.; Yaghi, O. M. Hydrogen Storage in Microporous Metal-Organic Frameworks. *Science* **2003**, *300*, 1127–1129.

(24) Wong-Foy, A. G.; Matzger, A. J.; Yaghi, O. M. Exceptional H2 Saturation Uptake in Microporous Metal-Organic Frameworks. *J. Am. Chem. Soc.* **2006**, *128*, 3494–3495.

(25) Rowsell, J. L. C.; Yaghi, O. M. Strategies for Hydrogen Storage in Metal–Organic Frameworks. *Angew. Chem. Int. Ed.* **2005**, *44*, 4670–4679.

(26) 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*, 16876–16883.

(27) Banerjee, R.; Furukawa, H.; Britt, D.; Knobler, C.; O'Keeffe, M.; Yaghi, O. M. Control of Pore Size and Functionality in Isoreticular Zeolitic Imidazolate Frameworks and their Carbon Dioxide Selective Capture Properties. *J. Am. Chem. Soc.* **2009**, *131*, 3875–3877.

(28) Sircar, S. Basic Research Needs for Design of Adsorptive Gas Separation Processes. *Ind. Eng. Chem. Res.* **2006**, *45*, 5435–5448.

(29) McDonald, T. M.; Lee, W. R.; Mason, J. A.; Wiers, B. M.; Hong, C. S.; Long, J. R. Capture of Carbon Dioxide from Air and Flue Gas in the Alkylamine-Appended Metal–Organic Framework mmen-Mg2(dobpdc). *Journal of the American Chemical Society* **2012**, *134*, 7056–7065.

(30) McDonald, T. M. et al. Cooperative Insertion of CO_2 in Diamine-Appended Metal-Organic Frameworks. *Nature* **2015**, *519*, 303–308.

(31) Rezaei, F.; Rownaghi, A. A.; Monjezi, S.; Lively, R. P.; Jones, C. W. SO_x/NO_x Removal from Flue Gas Streams by Solid Adsorbents: A Review of Current Challenges and Future Directions. *Energy Fuels* **2015**, *29*, 5467–5486.

(32) Mottillo, C.; Friščić, T. Carbon Dioxide Sensitivity of Zeolitic Imidazolate Frameworks. *Angew. Chem. Int. Ed.* **2014**, *53*, 7471–7474.

(33) Bhattacharyya, S.; Jayachandrababu, K. C.; Chiang, Y.; Sholl, D. S.; Nair, S. Butanol Separation from Humid CO2-Containing Multicomponent Vapor Mixtures by Zeolitic Imidazolate Frameworks. *ACS Sustainable Chem. Eng.* **2017**, *5*, 9467–9476.

(34) Bhattacharyya, S.; Pang, S. H.; Dutzer, M. R.; Lively, R. P.; Walton, K. S.; Sholl, D. S.; Nair, S. Interactions of SO_2-Containing Acid Gases with ZIF-8: Structural Changes and Mechanistic Investigations. *J. Phys. Chem. C* **2016**, *120*, 27221–27229.

(35) Pang, S. H.; Han, C.; Sholl, D. S.; Jones, C. W.; Lively, R. P. Facet-Specific Stability of ZIF-8 in the Presence of Acid Gases Dissolved in Aqueous Solutions. *Chem. Mater.* **2016**, *28*, 6960–6967.

(36) Bhattacharyya, S.; Han, R.; Kim, W.-G.; Chiang, Y.; Jayachandrababu, K. C.; Hungerford, J. T.; Dutzer, M. R.; Ma, C.; Walton, K. S.; Sholl, D. S.; Nair, S. Acid Gas Stability of Zeolitic Imidazolate Frameworks: Generalized Kinetic and Thermodynamic Characteristics. *Chem. Mater.* **2018**, *30*, 4089–4101.

(37) Bhattacharyya, S.; Han, R.; Joshi, J. N.; Zhu, G.; Lively, R. P.; Walton, K. S.; Sholl, D. S.; Nair, S. Stability of Zeolitic Imidazolate Frameworks in NO2. *J. Phys. Chem. C* **2019**, *123*, 2336–2346.

(38) Bhattacharyya, S.; Sholl, D. S.; Nair, S. Quantitative Correlations for the Durability of Zeolitic Imidazolate Frameworks in Humid SO2. *Ind. Eng. Chem. Res.* **2020**, *59*, 245–252.

(39) Fan, Q.; Bosch, M.; Zhang, M.; Zhou, H.-C. Increasing the Stability of Metal-Organic Frameworks. *Adv. Chem.* **2014**, *2014*, 182327.

(40) 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*, 15018.

(41) Yuan, S. et al. Stable Metal-Organic Frameworks: Design, Synthesis, and Applications. *Adv. Mater.* **2018**, *30*, 1704303.

(42) Burtch, N. C.; Jasuja, H.; Walton, K. S. Water Stability and Adsorption in Metal-Organic Frameworks. *Chem. Rev.* **2014**, *114*, 10575–10612.

(43) Lu, P.; Wu, Y.; Kang, H.; Wei, H.; Liu, H.; Fang, M. What can pKa and NBO Charges of the Ligands Tell us about the Water and Thermal Stability of Metal Organic Frameworks? *J. Mater. Chem. A* **2014**, *2*, 16250–16267.

(44) Zhang, C.; Han, C.; Sholl, D. S.; Schmidt, J. R. Computational Characterization of Defects in Metal-Organic Frameworks: Spontaneous and Water-Induced Point Defects in ZIF-8. *J. Phys. Chem. Lett.* **2016**, *7*, 459–464.

(45) Han, C.; Zhang, C.; Tymińska, N.; Schmidt, J. R.; Sholl, D. S. Insights into the Stability of Zeolitic Imidazolate Frameworks in Humid Acidic Environments from First-Principles Calculations. *J. Phys. Chem. C* **2018**, *122*, 4339–4348.

(46) Schütz, M.; Hetzer, G.; Werner, H.-J. Low-Order Scaling Local Electron Correlation Methods. I. Linear Scaling Local MP2. *J. Chem. Phys.* **1999**, *111*, 5691–5705.

(47) Hampel, C.; Werner, H.-J. Local Treatment of Electron Correlation in Coupled Cluster Theory. *J. Chem. Phys.* **1996**, *104*, 6286–6297.

(48) Werner, H.-J.; Schütz, M. An Efficient Local Coupled Cluster Method for Accurate Thermochemistry of Large Systems. *J. Chem. Phys.* **2011**, *135*, 144116.

(49) Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(50) Polly, R.; Werner, H.-J.; Manby, F. R.; Knowles, P. J. Fast Hartree-Fock Theory Using Local Density Fitting Approximations. *Mol. Phys.* **2004**, *102*, 2311–2321.

(51) Werner, H.-J.; Manby, F. R.; Knowles, P. J. Fast Linear Scaling Second-Order Møller-Plesset Perturbation Theory (MP2) Using Local and Density Fitting Approximations. *J. Chem. Phys.* **2003**, *118*, 8149–8160.

(52) Edmiston, C.; Ruedenberg, K. Localized Atomic and Molecular Orbitals. *Rev. Mod. Phys.* **1963**, *35*, 457–464.

(53) Foster, J. M.; Boys, S. F. Canonical Configurational Interaction Procedure. *Rev. Mod. Phys.* **1960**, *32*, 300–302.

(54) Pipek, J.; Mezey, P. G. A Fast Intrinsic Localization Procedure Applicable for Ab Initio and Semiempirical Linear Combination of Atomic Orbital Wave Functions. *J. Chem. Phys.* **1989**, *90*, 4916–4926.

(55) Reed, A. E.; Weinstock, R. B.; Weinhold, F. Natural Population Analysis. *J. Chem. Phys.* **1985**, *83*, 735–746.

(56) Knizia, G. Intrinsic Atomic Orbitals: An Unbiased Bridge between Quantum Theory and Chemical Concepts. *J. Chem. Theory Comput.* **2013**, *9*, 4834–4843.

(57) Kleier, D. A.; Halgren, T. A.; Hall, J. H.; Lipscomb, W. N. Localized Molecular Orbitals for Polyatomic Molecules. I. A Comparison of the Edmiston-Ruedenberg and Boys Localization Methods. *J. Chem. Phys.* **1974**, *61*, 3905–3919.

(58) Boughton, J. W.; Pulay, P. Comparison of the Boys and Pipek-Mezey Localizations in the Local Correlation Approach and Automatic Virtual Basis Selection. *J. Comput. Chem.* **1993**, *14*, 736–740.

(59) Honig, B.; Karplus, M. Implications of Torsional Potential of Retinal Isomers for Visual Excitation. *Nature* **1971**, *229*, 558–560.

(60) Warshel, A.; Karplus, M. Calculation of Ground and Excited State Potential Surfaces of Conjugated Molecules. I. Formulation and Parametrization. *J. Am. Chem. Soc.* **1972**, *94*, 5612–5625.

(61) Warshel, A.; Karplus, M. Calculation of .pi..pi.* Excited State Conformations and Vibronic Structure of Retinal and Related Molecules. *J. Am. Chem. Soc.* **1974**, *96*, 5677–5689.

(62) Warshel, A.; Levitt, M. Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric Stabilization of the Carbonium Ion in the Reaction of Lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.

(63) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Ed.* **2009**, *48*, 1198–1229.

(64) Chung, L. W.; Sameera, W. M. C.; Ramozzi, R.; Page, A. J.; Hatanaka, M.; Petrova, G. P.; Harris, T. V.; Li, X.; Ke, Z.; Liu, F.; Li, H.-B.; Ding, L.; Morokuma, K. The ONIOM Method and Its Applications. *Chem. Rev.* **2015**, *115*, 5678–5796.

(65) Maseras, F.; Morokuma, K. IMOMM: A New Integrated Ab Initio + Molecular Mechanics Geometry Optimization Scheme of Equilibrium Structures and Transition States. *J. Comput. Chem.* **1995**, *16*, 1170–1179.

(66) Humbel, S.; Sieber, S.; Morokuma, K. The IMOMO Method: Integration of Different Levels of Molecular Orbital Approximations for Geometry Optimization of Large Systems: Test for n-Butane Conformation and S_N2 Reaction: RCl+Cl^-. *J. Chem. Phys.* **1996**, *105*, 1959–1967.

(67) Inglesfield, J. E. A Method of Embedding. *J. Phys. C: Solid State Physics* **1981**, *14*, 3795.

(68) Huang, C.; Pavone, M.; Carter, E. A. Quantum Mechanical Embedding Theory Based on a Unique Embedding Potential. *J. Chem. Phys.* **2011**, *134*, 154110.

(69) Yu, K.; Carter, E. A. Extending Density Functional Embedding Theory for Covalently Bonded Systems. *Proc. Natl. Acad. Sci.* **2017**, *114*, E10861–E10870.

(70) Manby, F. R.; Stella, M.; Goodpaster, J. D.; Miller, T. F. A Simple, Exact Density-Functional-Theory Embedding Scheme. *J. Chem. Theory Comput.* **2012**, *8*, 2564–2568.

(71) Huang, P.; Carter, E. A. Self-Consistent Embedding Theory for Locally Correlated Configuration Interaction Wave Functions in Condensed Matter. *J. Chem. Phys.* **2006**, *125*, 084102.

(72) Yu, K.; Libisch, F.; Carter, E. A. Implementation of density functional embedding theory within the projector-augmented-wave method and applications to semiconductor defect states. *J. Chem. Phys.* **2015**, *143*, 102806.

(73) Lee, S. J. R.; Welborn, M.; Manby, F. R.; Miller, T. F. Projection-Based Wavefunction-in-DFT Embedding. *Acc. Chem. Res.* **2019**, *52*, 1359–1368.

(74) Chandler, D. Introduction to modern statistical. *Mechanics. Oxford University Press, Oxford, UK* **1987**, *5*.

(75) Sillar, K.; Hofmann, A.; Sauer, J. Ab Initio Study of Hydrogen Adsorption in MOF-5. *J. Am. Chem. Soc.* **2009**, *131*, 4143–4150.

(76) Sillar, K.; Sauer, J. Ab Initio Prediction of Adsorption Isotherms for Small Molecules in Metal–Organic Frameworks: The Effect of Lateral Interactions for Methane/CPO-27-Mg. *J. Am. Chem. Soc.* **2012**, *134*, 18354–18365.

(77) Kundu, A.; Piccini, G.; Sillar, K.; Sauer, J. Ab Initio Prediction of Adsorption Isotherms for Small Molecules in Metal–Organic Frameworks. *J. Am. Chem. Soc.* **2016**, *138*, 14047–14056.

(78) Jonsson, H.; Mills, G.; Jacobsen, K. In Berne, B., Ciccotti, G., Coker, D., Eds.; World Scientific: 1998.

(79) Henkelman, G.; Jónsson, H. Improved Tangent Estimate in the Nudged Elastic Band Method for Finding Minimum Energy Paths and Saddle Points. *The Journal of Chemical Physics* **2000**, *113*, 9978–9985.

(80) 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*, 9901–9904.

(81) Yu, K.; Kiesling, K.; Schmidt, J. R. Trace Flue Gas Contaminants Poison Coordinatively Unsaturated Metal-Organic Frameworks: Implications for CO_2 Adsorption and Separation. *J. Phys. Chem. C* **2012**, *116*, 20480–20488.

(82) Nicholson, T. M.; Bhatia, S. K. Electrostatically Mediated Specific Adsorption of Small Molecules in Metal-Organic Frameworks. *J. Phys. Chem. B* **2006**, *110*, 24834–24836.

(83) Grajciar, L.; Bludský, O.; Nachtigall, P. Water Adsorption on Coordinatively Unsaturated Sites in CuBTC MOF. *J. Phys. Chem. Lett.* **2010**, *1*, 3354–3359.

(84) Howe, J. D.; Liu, Y.; Flores, L.; Dixon, D. A.; Sholl, D. S. Acid Gas Adsorption on Metal-Organic Framework Nanosheets as a Model of an "All-Surface" Material. *J. Chem. Theory Comput.* **2017**, *13*, 1341–1350.

(85) Choomwattana, S.; Maihom, T.; Khongpracha, P.; Probst, M.; Limtrakul, J. Structures and Mechanisms of the Carbonyl-ene Reaction between MOF-11 Encapsulated Formaldehyde and Propylene: An ONIOM Study. *J. Phys. Chem. C* **2008**, *112*, 10855–10861.

(86) Yadnum, S.; Choomwattana, S.; Khongpracha, P.; Sirijaraensre, J.; Limtrakul, J. Comparison of Cu-ZSM-5 Zeolites and Cu-MOF-505 Metal-Organic Frameworks as Heterogeneous Catalysts for the Mukaiyama Aldol Reaction: A DFT Mechanistic Study. *ChemPhysChem*, *14*, 923–928.

(87) Doitomi, K.; Xu, K.; Hirao, H. The Mechanism of an Asymmetric Ring-Opening Reaction of Epoxide with Amine Catalyzed by a Metal–Organic Framework: Insights from Combined Quantum Mechanics and Molecular Mechanics Calculations. *Dalton Trans.* **2017**, *46*, 3470–3481.

(88) 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 Catalysis* **2015**, *5*, 3287–3291.

(89) 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*, 1144–1150.

(90) Zheng, M.; Liu, Y.; Wang, C.; Liu, S.; Lin, W. Cavity-Induced Enantioselectivity Reversal in a Chiral Metal–Organic Framework Brønsted Acid Catalyst. *Chem. Sci.* **2012**, *3*, 2623–2627.

(91) Pianwanit, A.; Kritayakornupong, C.; Vongachariya, A.; Selphusit, N.; Ploymeerusmee, T.; Remsungnen, T.; Nuntasri, D.; Fritzsche, S.; Hannongbua, S. The Optimal Binding Sites of CH_4 and CO_2 Molecules on the Metal-Organic Framework MOF-5: ONIOM Calculations. *Chem. Phys.* **2008**, *349*, 77–82.

(92) Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Liquid Metals. *Phys. Rev. B* **1993**, *47*, 558–561.

(93) Kresse, G.; Hafner, J. Ab Initio Molecular-Dynamics Simulation of the Liquid-Metal–Amorphous-Semiconductor Transition in Germanium. *Phys. Rev. B* **1994**, *49*, 14251–14269.

(94) 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*, 15–50.

(95) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total-Energy Calculations using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169–11186.

(96) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(97) 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*, 154104.

(98) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50*, 17953–17979.

(99) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: a General-Purpose Quantum Chemistry Program Package. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242–253.

(100) Larsen, A. H. et al. The Atomic Simulation Environment—a Python Library for Working with Atoms. *J. Phys. Condensed Matter* **2017**, *29*, 273002.

(101) Weng, T.; Schmidt, J. R. Flexible and Transferable ab Initio Force Field for Zeolitic Imidazolate Frameworks: ZIF-FF. *J. Phys. Chem. A* **2019**, *123*, 3000–3012.

(102) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1–19.

(103) Bitzek, E.; Koskinen, P.; Gähler, F.; Moseler, M.; Gumbsch, P. Structural Relaxation Made Simple. *Phys. Rev. Lett.* **2006**, *97*, 170201.

(104) McDaniel, J. G.; Yu, K.; Schmidt, J. R. Ab Initio, Physically Motivated Force Fields for CO_2 Adsorption in Zeolitic Imidazolate Frameworks. *J. Phys. Chem. C* **2012**, *116*, 1892–1903.

(105) Ding, C.-G.; Taskila, T.; Laasonen, K.; Laaksonen, A. Reliable Potential for Small Sulfuric Acid-Water Clusters. *Chem. Phys.* **2003**, *287*, 7–19.

(106) Saebø, S.; Tong, W.; Pulay, P. Efficient Elimination of Basis Set Superposition Errors by the Local Correlation Method: Accurate Ab Initio Studies of the Water Dimer. *J. Chem. Phys.* **1993**, *98*, 2170–2175.

(107) Reyes, A.; Tlenkopatchev, M. A.; Fomina, L.; Guadarrama, P.; Fomine, S. Local MP2-Based Method for Estimation of Intermolecular Interactions in Aromatic Molecules. Benzene, Naphthalene, and Pyrimidine Dimers. A Comparison with Canonical MP2 Method. *J.Phys. Chem. A* **2003**, *107*, 7027–7031.

(108) 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*, 9356–9362.

(109) Adler, T. B.; Knizia, G.; Werner, H.-J. A Simple and Efficient CCSD(T)-F12 Approximation. *J. Chem. Phys.* **2007**, *127*, 221106.

(110) Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell Jr., A. D. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010**, *31*, 671–690.

(111) Yosa, J.; Meuwly, M. Vibrationally Induced Dissociation of Sulfuric Acid (H_2SO_4). *J. Phys. Chem. A* **2011**, *115*, 14350–14360.

(112) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Hermans, J. In Pullman, B., Ed.; Reidel, Dordrecht: 1981.

(113) Cui, K.; Bhattacharyya, S.; Nair, S.; Schmidt, J. R. Origins of Acid-Gas Stability Behavior in Zeolitic Imidazolate Frameworks: The Unique High Stability of ZIF-71. *J. Am. Chem. Soc.* **2021**, *143*, 18061–18072.

(114) Flamm, D. L.; Bacon, D. D.; Kinsbron, E.; English, A. T. Chemical Reaction of Sulfur Dioxide at High Humidity and Temperature: Implications for Accelerated Testing. *J. Electrochem. Soc.* **1981**, *128*, 679–685.

(115) Liu, J.; Fang, S.; Wang, Z.; Yi, W.; Tao, F.-M.; Liu, J.-y. Hydrolysis of Sulfur Dioxide in Small Clusters of Sulfuric Acid: Mechanistic and Kinetic Study. *Environ. Sci. & Technol.* **2015**, *49*, 13112–13120.

(116) Han, R.; Tymińska, N.; Schmidt, J. R.; Sholl, D. S. Propagation of Degradation-Induced Defects in Zeolitic Imidazolate Frameworks. *J.Phys. Chem. C* **2019**, *123*, 6655–6666.

(117) 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*, 134–141.

(118) Voter, A. F. In *Radiation Effects in Solids*, ed. by Sickafus, K. E.; Kotomin, E. A.; Uberuaga, B. P., Springer Netherlands: Dordrecht, 2007, pp 1–23.

(119) Chen, K.; Gan, Z.; Horstmeier, S.; White, J. L. Distribution of Aluminum Species in Zeolite Catalysts: 27Al NMR of Framework, Partially-Coordinated Framework, and Non-Framework Moieties. *J. Am. Chem. Soc.* **2021**, *143*, 6669–6680.

(120) Salvadori, E.; Fusco, E.; Chiesa, M. Long-Range Spatial Distribution of Single Aluminum Sites in Zeolites. *J. Phys. Chem. Lett.* **2022**, *13*, 1283–1289.

(121) Agwupuye, J. A.; Louis, H.; Unimuke, T. O.; David, P.; Ubana, E. I.; Moshood, Y. L. Electronic Structure Investigation of the Stability, Reactivity, NBO Analysis, Thermodynamics, and the Nature of the Interactions in Methyl-Substituted Imidazolium-based Ionic Liquids. *J. Mol. Liq.* **2021**, *337*, 116458.

(122) Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. Self-Consistent Molecular Orbital Methods. XXIII. A Polarization-Type Basis Set for Second-Row Elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665.

(123) Frisch, M. J. et al. Gaussian 16 Revision C.01, Gaussian Inc. Wallingford CT, 2016.

(124) Nicklass, A.; Dolg, M.; Stoll, H.; Preuss, H. Ab Initio Energy-Adjusted Pseudopotentials for the Noble Gases Ne through Xe: Calculation of Atomic Dipole and Quadrupole Polarizabilities. *J. Chem. Phys.* **1995**, *102*, 8942–8952.

(125) Peterson, K. A.; Figgen, D.; Dolg, M.; Stoll, H. Energy-Consistent Relativistic Pseudopotentials and Correlation Consistent Basis Sets for the $4d$ Elements Y–Pd. *J. Chem. Phys.* **2007**, *126*, 124101.

(126) Ratkova, E. L.; Palmer, D. S.; Fedorov, M. V. Solvation Thermodynamics of Organic Molecules by the Molecular Integral Equation Theory: Approaching Chemical Accuracy. *Chem. Rev.* **2015**, *115*, 6312–6356.

(127) Jayachandrababu, K. C.; Bhattacharyya, S.; Chiang, Y.; Sholl, D. S.; Nair, S. Recovery of Acid-Gas-Degraded Zeolitic Imidazolate Frameworks by Solvent-Assisted Crystal Redemption (SACRed). *ACS Appl. Mater. Interfaces* **2017**, *9*, 34597–34602.