Proceedings of the 5th International Conference on Materials Chemistry and Environmental Engineering
DOI: 10.54254/2755-2721/128/2025.20243

Computational Simulations of Metal-Organic Framework Structures Based on Second Building Units

Yilin Jin[1,a,*]
[1]Department of Chemical Engineering, The University of Manchester, Manchester, United Kingdom
a. yilin.jin-2@student.manchester.ac.uk
**corresponding author*

Abstract: Metal-organic frameworks (MOFs) are crystalline porous materials composed of metal clusters and organic linkers with tunable structures and large surface areas, making them ideal for gas storage, separation, and catalytic applications. The development of new MOFs and the screening of MOFs after synthesis are usually experimental, but the cost and time required for experiments are very large. Interestingly, computational simulations have become a powerful tool to accelerate the discovery of new MOFs structures. This study focuses on the computational design of MOFs based on secondary building units (SBUs) and explores the effects of different linkers on their geometry and properties. Using techniques such as density functional theory (DFT) and machine learning (ML), MOFs can be screened for great gas adsorption and storage properties. The results show that computational methods have the potential to predict and optimize MOFs structures, but they are still limited in accurately simulating dynamic environmental conditions and complex chemical reactions. These findings help advance the design of MOFs with tailored properties for specific industrial applications.

Keywords: Metal-organic frameworks, Simulation, Secondary building units, Density functional theory, Machine learning.

1. Introduction

Metal-organic frameworks (MOFs) are porous materials with large surface areas and adjustable pore sizes made of metal-containing nodes and organic linkers. Due to their modular characteristic, MOFs have been widely studied for applications such as gas storage, separation, and catalysis. These materials are versatile due to their structure-property relationships, which can be customized by modifying metal clusters (e.g., Sc^{3+} cluster or Zr_6 cluster) and organic linkers (e.g., carboxylates or imidazolates). This allows for fine-tuning of properties like pore size, surface area, and chemical reactivity for gas storage and separation [1].

Traditional trial-and-error synthesis methods are time-consuming and limited by researchers' chemical intuition. However, computational techniques offer a powerful alternative for predicting novel structures with desired properties. By systematically exploring different combinations of secondary building units (SBUs) and linkers, researchers can design MOFs that exhibit enhanced performance in specific applications such as gas adsorption, storage, and separation. Recent advances in computational methods have greatly accelerated the discovery of new MOFs structures. Molecular simulations and machine learning are increasingly used to predict hypothetical MOFs properties, reducing the need for experimental trial-and-error. Large libraries of hypothetical MOFs have been assembled and screened using various software platforms to help researchers find promising candidates with desirable properties for specific applications. The computational method widely used in MOFs is machine learning (ML) and density functional theory (DFT). ML and DFT have become irreplaceable tools in data-driven methods for material design. Nevertheless, they have limitations in practical application. MOFs have a complex structure and usually need to be simplified in the computational model, such as simulating only local units or using smaller structural models. These simplifications may ignore long-range interactions or important physical and chemical properties, resulting in deviations from experimental results. In actual operation, the behavior of MOFs may be affected by environmental conditions (such as temperature, pressure, solvent effects), and these dynamic factors are usually not fully reflected in the computational model. In addition, it also lacks accurate predictions of chemical reactions. In some MOFs applications, such as catalytic reactions, complex chemical reaction networks are involved. Current computational methods still have limitations in dealing with reaction paths and reaction rates, especially in multi-step reactions, where the calculation results may deviate greatly from the experimental results. Reactive sites in MOFs, such as metal clusters or coordinatively unsaturated metal atoms, often have highly complex electronic structures [2]. Accurate simulation and prediction of these sites requires the use of a very high level of theory, which increases the difficulty. These shortcomings are problems that need to be solved in the future development of computational methods.

This review summarizes the computational methods in the design of MOFs. Various computational techniques are explored, including DFT for electronic structure analysis and Machine Learning ML for data-driven predictions. Recent advancements in simulating gas adsorption (e.g., CO_2, CH_4) and separation processes within MOFs are highlighted, addressing key applications such as carbon capture and gas purification. The strengths and limitations of these methods are assessed, particularly in modeling complex, real-world conditions like varying temperature, pressure, and chemical reactivity.

2. Overview

2.1. The background of MOFs

Through coordination bond bridging, metal ions or metal clusters and multi-dentate organic ligands self-assemble to form MOFs, also referred to as porous coordination polymers, a type of crystalline material with a periodic network structure. During the synthesis process, simple ligand metal clusters composed of metal ions and organic ligands are unsuitable for the production of network materials due to their insufficient stable coordination directionality. To form a secondary structural unit with rigidity, the most common MOFs material is to "couple" (coordinate) metal ions by carboxyl groups. It uses the carbon atom of the carboxylate group as an extension point to define the secondary structural unit. The types of MOFs materials can be controlled by choosing metal ions and organic ligands under suitable reaction conditions, allowing for the synthesis of a diverse array of MOFs materials through numerous combinations of organic ligands and metal ions [3]. At the same time, the addition of binary or even multi-organic ligands or metal ions in the reaction system provides unlimited possibilities for the synthesis of MOFs. Various functional building blocks can be chosen or engineered to directly synthesize MOFs materials with distinct functions. Various functional groups may be incorporated via post-synthetic modification to achieve specific performance regulation and to tailor MOFs materials for designated applications. The functions and usage scenarios of MOFs are also very extensive. Due to their large specific surface area and adjustable pore structure, MOFs have received widespread attention in the field of hydrogen storage. Among them, materials such as $Zn_4O(1,4\text{-benzenedicarboxylate})_3\cdot(N,N'\text{-dimethylformamide})_8C_6H_5Cl$ and UiO-66 can store hydrogen under low temperature conditions. MOFs have an open pore structure and ultra-large specific surface area, which can promote the diffusion of pollutants in MOFs in water environments. By regulating the coordination unsaturated metal sites, Lewis acid-base sites, types of organic functional groups, pore size, surface electrical properties, etc. of MOFs, efficient, rapid and highly selective adsorption of target pollutants can be achieved [4].

2.2. The connection between computational method and MOFs design

The synthesis process of MOFs is relatively simple, requiring only nodes, edges and templates to synthesize MOFs. Therefore, MOFs can be virtually generated with the help of computational methods, resulting in an unlimited number of hypothetical structures. While this diversity offers great potential, it also makes it difficult to screen for the best performance structures for MOFs applications. The cost and time of evaluating many structures experimentally are very large, but molecular simulation has emerged as a viable alternative. Various types of simulations, such as DFT, grand canonical Monte Carlo, and molecular dynamics simulations, have been widely used for this purpose [5]. Interpreting large amounts of simulation data is very complex, so effective methods are needed to analyze the data and extract meaningful insights from it. Although simulations contain a certain degree of error, they can evaluate structures faster and help discover new MOFs. The application of molecular simulations to the vast material space of MOFs still faces certain limitations.

Data-driven methods, especially machine learning, have become an indispensable tool in materials design, successfully addressing these challenges. By identifying patterns from large data sets, machine learning models reveal various structure-property relationships, leading to a deeper understanding of material chemistry. These insights have been actively used to plan future experiments and design new materials with desired properties. Another emerging trend in materials design is the use of deep generative models, which can generate chemical structures without being restricted by a predefined search space or training database. In addition, new approaches such as ML potentials and cross-disciplinary knowledge transfer are gaining attention, and their combination with materials science shows great promise [6].

3. Impact of Linker Variations in MOFs

The geometry and connectivity of SBUs significantly influence the final topology and properties of MOFs [7]. Computational studies and experimental data reveal that variations in linkers and SBUs configurations can lead to diverse structural outcomes, affecting the porosity, stability, and functionality of the resulting MOFs.

The synthesis of new MOFs skeletons comprises three primary components: nodes, edges, and templates. The geometric and chemical characteristics of SBUs and connecting nodes significantly influence the overall structure and efficacy of MOFs. First, *Nam et al*. proposed the dissociation of pre-synthesized MOFs by "Clip-Off" chemistry to extract functionalized SBUs. This method involves using truncation chemistry to completely decompose the SBUs and extract the SBUs from the pre-existing MOFs. The original MOFs precursor is specifically designed to contain the target SBUs, which is connected only by cleavable bonds (in this case, olefin bonds). By cleaving these organic linkers through ozonolysis under reducing conditions, SBUs with aldehyde functional groups are released. These aldehyde-functionalized SBUs can then react with amines through dynamic covalent chemistry to construct new, rationally designed MOFs. The paper mentions the use of trimer metal clusters (Sc^{3+}cluster) as SBUs to synthesize MOFs. The authors demonstrate that different metal clusters (independent variables) affect the topology and properties of the resulting MOFs. For example, Sc^{3+} cluster connected to a two-site linker form a network structure with MIL-126 (Fe), and its properties such as pore size and chemical reactivity are determined by the properties of these metal clusters. In this case, the choice of specific metal ions (such as Sc^{3+}) is an independent variable that affects the properties of the final material [8]. This shows how changes in the independent variable (the metal ions in the SBUs) can lead to changes in the computational design of the material structure and the resulting functionality [9]. This approach shows more possibilities for MOFs synthesis and the diversity of SBUs in synthesis.

The geometric and chemical characteristics of the SBUs and connecting nodes significantly influence the overall structure and efficacy of the MOFs. The study of *Li et al.* constructed three MOFs with different topologies (sqc, scu and flu) by regulating the geometric structure of Zr_6 cluster and Y_6 cluster and the configuration of ligands. Synthesis and characterization of three distinct MOFs structures, utilizing the identical organic linker N,N,N′,N′-tetrakis(4-carboxyphenyl)-1,4-phenylenediamine (tcppda) and analogous 8 Connect M6 clusters (M = Zr or Y). The three compounds, designated HIAM-402, HIAM-403, and HIAM-311, exhibit 4,8-linked sqc, scu, and flu topologies, respectively [8]. Different connections and geometries affect the framework stability and porosity of MOFs. Similarly, *Han et al.* studied the role of open metal sites in water adsorption. The water absorption characteristics can be substantially modified by altering the chemical composition of the metal sites and the water anchoring sites [10]. This provides a new idea for MOFs design.

4. Computational Methods in MOF Synthesis

Computation method has been widely used in the synthesis of MOFs and screening of suitable MOFs structures, which is an important tool for understanding and optimizing the synthesis and performance of MOFs. Nowadays, the most used methods are ML and DFT, which play different functions in the entire process of synthesizing and screening MOFs.

Computational method is a very important tool in MOFs screening, mainly based on three criteria: gas adsorption capacity, separation efficiency and stability. In *Shuai et al.,* ML was used for computational simulation of MOFs carbon capture. A least squares support vector machine model was employed to estimate CO_2 uptake in MOFs, demonstrating superior predictive performance compared to alternative models. They integrated machine learning with conventional techniques, including Monte Carlo simulations, to enhance gas adsorption and separation efficacy in MOFs. The use of ML has made great progress in computational simulation. Compared with traditional experiments, ML can replicate extreme conditions that cannot be tested in conventional laboratory environments and can greatly improve the efficiency of screening. The advantage of ML over other methods is that it can effectively mine implicit and effective knowledge from big data. After analyzing these data, researchers can optimize MOFs design by adjusting pore size and metal doping [11]. In another study, *Demire and Erucar* combined molecular simulation with ML to study the separation performance of double-linked MOFs. Generally, there are thousands of MOFs for screening, and analyzing the characteristics of all MOFs is very arduous. ML can quickly process large amounts of data, especially for predicting the absorption of various gases. By analyzing the separation of greenhouse gases such as CF_4, CH_4, and N_2, they demonstrated how ML models can accurately predict gas adsorption and selectivity based on surface area and Henry's constant [12]. An important indicator in screening MOFs is their stability, because stability affects the performance of MOFs in practice. Through ML models, structure-property relationships can be encoded and interpreted based on the representation of graphs and pore structures, and the stability of MOFs can be intuitively understood. In the study of *Nandy et al*, the stability of 1500 MOFs was analyzed, and the analysis using ML models took less than an hour, which was very fast [13].

DFT is a computational method that has significantly contributed to the comprehension of the electronic characteristics of Zr_6 cluster. DFT simulations offer enhanced insight into the bonding characteristics between metal clusters and organic linkers. It provides an atomic-level understanding of how these interactions affect the overall stability and functionality of the material. In Zr_6-based MOFs, DFT simulations help predict how these clusters behave in a variety of chemical environments. This makes it easier to design MOFs with improved chemical stability and desired properties. By using DFT, researchers were able to optimize the metal-ligand coordination environment and produce MOFs with enhanced mechanical and thermal stability. This is particularly important in Zr_6 cluster MOFs, where the interactions between metal nodes and organic linkers can determine the robustness of the material. For example, in computational models of CO_2 adsorption, DFT was crucial in evaluating the interactions between Zr_6 cluster nodes and CO_2 molecules [14]. This is important in being able to design more efficient adsorption materials.

5. Conclusion

This study demonstrates the efficacy of a computer-aided approach in designing different SBUs based MOFs structures. Computational methods can screen the best MOFs through diverse structural outcomes, affecting the porosity, stability, and functionality. This method offers a robust mechanism for expediting the identification of novel MOFs with advantageous characteristics, aiding the continuous advancement of sophisticated materials for separation and storage applications. Nevertheless, there are limitations to simulating or synthesizing MOFs structures through computational methods. After many MOFs are synthesized, some careful screening is required. DFT and ML can perform data simulation and calculate pore spacing, but they rely on large data sets. If the data is insufficient or biased, its accuracy may be reduced. Future research should be committed to improving the integration of experimental and computational data to improve the accuracy and efficiency of predictions. In addition, incorporating dynamic conditions and solvent effects in simulations can produce more reliable results and promote the development of MOFs tailored for various industrial applications. By overcoming these obstacles and continuing to utilize computational techniques, researchers will be able to further advance MOFs design and develop materials.

References

[1] *Wang D, Astruc D. 2020. State of the Art and Prospects in Metal–Organic Framework (MOF)-Based and MOF-Derived Nanocatalysis. Chem. Rev. 120, 2, 1438–1511.*

[2] *Park J, Kim H, Kang Y, et al. 2024. From Data to Discovery: Recent Trends of Machine Learning in Metal–Organic Frameworks. JACS Au, XXXX, XXX, XXX-XXX.*

[3] *Xu Y, Marrett J, Titi H, et al. 2023. Experimentally Validated Ab Initio Crystal Structure Prediction of Novel Metal–Organic Framework Materials. J. Am. Chem. Soc. 2023, 145, 3515−3525*

[4] *Eddaoudi M, Li H, and Yaghi OM. 2000. Highly Porous and Stable Metal−Organic Frameworks: Structure Design and Sorption Properties. J. Am. Chem. Soc. 122, 7, 1391–1397.*

[5] *Azar A, Velioglu S, Keskin S. 2019. Large-Scale Computational Screening of Metal Organic Framework (MOF) Membranes and MOF-Based Polymer Membranes for H2/N2 Separations. ACS Sustainable Chem. Eng. 2019, 7, 9525−9536.*

[6] *Wang Y, Jiang Z, Wang D, et al. 2024. Machine Learning-Assisted Discovery of Propane-Selective Metal–Organic Frameworks. J. Am. Chem. Soc. 146, 10, 6955–6961.*

[7] *Kim D, Song X, Yoon J H and Lah M S. 2012. 3,6-Connected Metal–Organic Frameworks Based on Triscarboxylate as a 3-Connected Organic Node and a Linear Trinuclear Co3(COO)6 Secondary Building Unit as a 6-Connected Node. Crystal Growth & Design. 12 (8), 4186-4193*

[8] *Nam D, Albalad J, Sánchez-Naya R, et al. 2024. Isolation of the Secondary Building Unit of a 3D Metal−Organic Framework through Clip-Off Chemistry. J. Am. Chem. Soc. https://doi.org/10.1021/jacs.4c09077.*

[9] *Li X, Liu J, Zhou K, et al. 2022. Tuning Metal–Organic Framework (MOF) Topology by Regulating Ligand and Secondary Building Unit (SBU) Geometry. J. Am. Chem. Soc. 144, 21702–21709.*

[10] *Han Y, Das P, He Y, et al. 2022. Crystallographic Mapping and Tuning of Water Adsorption in Metal−Organic Frameworks Featuring Distinct Open Metal Sites. J. Am. Chem. Soc.144, 19567–19575.*

[11] Shuai Q, Li W L, Zhao C and Yu J. 2024. Recent Advances of Computational Simulations on Carbon Capture in MOFs. Materials Today Communications. 40, 110050.

[12] Demir H and Erucar I. 2024. Simulation and Machine Learning-Integrated Investigation of Double-Linker MOFs for the Separation of CF_4 *from* CH_4 *and* N_2*. Ind. Eng. Chem. Res. 63, 12605–12619.*

[13] Nandy A, Duan C and Kulik, H.J. 2021. Using Machine Learning and Data Mining to Leverage Community Knowledge for the Engineering of Stable Metal–Organic Frameworks. J. Am. Chem. Soc. 143, 17535–17547.

[14] Shuai Q, Li W L, Zhao C and Yu J. 2024. Recent advances of computational simulations on carbon capture in MOFs. Materials Today Communications. 40, 110050. https://doi.org/10.1016/j.mtcomm.2024.110050