Digital Discovery

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Digital Discovery*, 2025, **4**, 3008

Digital design and discovery of biological metal–organic frameworks for gas signaling

Yifei Yue,[ab] Athulya S. Palakkal,[a] Saad Aldin Mohamed[a] and Jianwen Jiang *[ab]

Metal–organic frameworks (MOFs) are intriguing nanoporous materials with a wide variety of potential applications. Recent efforts in extending the functionalities of MOFs toward biological applications have inspired the development of Bio-MOFs comprising biological building blocks. Yet, while numerous experimental studies have attempted to synthesize different Bio-MOFs, computational screening of Bio-MOFs is impeded by the limited number of Bio-MOFs currently available. Here, we design a **Bio-hMOF** database containing 17 681 hypothetical structures, assembled from the fragments of 309 experimental Bio-MOFs, with rigorous geometry optimization and structural checks. Subsequently, a possible biological application of the **Bio-hMOFs** is demonstrated for the selective adsorption of signaling gases NO and CO. The effects of different inorganic and organic fragments on the mechanical properties of **Bio-hMOFs** are also examined. Finally, we identify mechanically stable **Bio-hMOFs** promising for selective NO/CO adsorption and holistically analyze the trade-off between adsorption capacity and mechanical strength. The digital **Bio-hMOF** database is available publicly, in which future studies can be leveraged to discover top candidates and unveil new structure–property insights into the further design of Bio-MOFs for targeted biological applications.

Received 20th May 2025
Accepted 20th August 2025

DOI: 10.1039/d5dd00213c

rsc.li/digitaldiscovery

Introduction

The escalating climate crisis has incentivized the development of environmentally friendly technologies in different resource-intensive industries. This includes growing efforts to adopt green chemistry principles, which emphasize the importance of harnessing renewable sources to reduce reliance on non-renewable resources.[1,2] In this context, functional porous materials, particularly metal–organic frameworks (MOFs), play an integral role in many potential applications such as storage, separation, and catalysis,[3] by serving as alternative avenues to ubiquitous yet energy-intensive processes (*e.g.*, distillation columns). With readily tunable porous structures and chemical functionalities, MOFs can be synthesized from a great number of building blocks. Among the tens of thousands of MOFs experimentally synthesized,[4] there are steady efforts to develop biological MOFs (*i.e.*, Bio-MOFs) containing biodegradable ligands or plant-derived polymers,[5] which save perishable resources for chemical synthesis, as well as minimize non-biodegradable waste at the end of their life cycle.[6] Bio-MOFs possess the versatile properties of conventional MOFs, as well as biological stability and functionality. Common building blocks in Bio-MOFs include amino acids, nucleobases, peptides, proteins, cyclodextrins, and metalloporphyrins.[7] These biomolecular building blocks contribute to the efficiency and specificity in essential tasks,[8,9] notably biomolecular recognition,[10] drug delivery,[11,12] and biocatalysis.[13]

While there are growing experimental studies on the design and applications of Bio-MOFs, computational effort to identify promising Bio-MOFs is scarce due to the limited number of Bio-MOFs available. Except for a few studies that computationally screened Bio-MOFs for indoxyl sulfate adsorption[14] and O_2/N_2 separation,[15] a common database of Bio-MOFs with biomolecular building blocks is lacking and the existing studies did not consider the mechanical properties of Bio-MOFs. Thus, addressing these limitations would accelerate the design of new Bio-MOFs for emerging biological applications.

In this work, a database of hypothetical Bio-MOFs (**Bio-hMOFs**) is developed. We first decomposed the unique inorganic and organic fragments from 309 experimentally available Bio-MOFs, and then generated **Bio-hMOF** structures based on two configurations: (1) one inorganic node with one organic node; (2) one inorganic node with one organic edge. The generated structures were optimized, followed by structural checks to eliminate disordered structures. Eventually, 17 681 structures were curated to constitute the **Bio-hMOF** database. From an application point of view, subsequently, we identified promising **Bio-hMOFs** for the selective adsorption of signaling gases,[16] namely NO[17,18] and CO,[19] as the selective adsorption and release of signaling gases is crucial in biological applications involving the *in vivo* delivery of signaling molecules.[20,21] Thus,

[a]*Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore, 117576, Singapore. E-mail: chejj@nus.edu.sg*
[b]*Integrative Sciences and Engineering Programme, National University of Singapore, Singapore, 119077, Singapore*

understanding attributes of Bio-MOFs with signaling gas adsorption would aid the development of new functional Bio-MOFs for selective capture of the desired biological signaling gases. Finally, we conducted molecular dynamics (MD) simulations to evaluate the mechanical properties of **Bio-hMOFs**, which are important metrics for assessing the mechanical stability of hypothetical structures. Furthermore, the relationships between chemical compositions and mechanical properties, as well as the trade-off between porosity and stability, were discussed.

The remainder of this work is organized as follows. First, we provide a brief overview of our research workflow for generating **Bio-hMOFs**. Next, we analyze the geometric and chemical features of **Bio-hMOFs**, in comparison to existing hypothetical MOF databases.[22–26] Thereafter, we examine the selective adsorption of a NO/CO mixture in **Bio-hMOFs** to elucidate structure–property relationships and identify top **Bio-hMOFs** with large adsorption capacity (N) and high selectivity (S). Finally, we integrate bulk (K), shear (G), and Young's (E) moduli of **Bio-hMOFs** in relation to their chemical compositions. More broadly, we provide a publicly available database and analysis tools to facilitate future computational studies in identifying and understanding the attributes of Bio-MOFs with promising functionality for biological applications.

Methods

Fig. 1a illustrates our research workflow comprising two main stages: generation of **Bio-hMOFs** and screening for selective NO/CO adsorption. First, we collected 309 experimental Bio-MOFs from CoRE-MOF 2019 [4] and CSD.[27] Then, these Bio-MOFs were decomposed (*i.e.*, broken down) into unique inorganic and organic fragments. In total, 141 inorganic fragments (Fig. 1b) existed including common metal nodes like Zr_8 secondary building units (SBUs) in UiO-66 and rare fragments like metalloporphyrin. Meanwhile, 66 organic fragments contained a wide range of functional groups (Fig. 1c) such as bioactive acetanilide (fragment 8), pyrene derivatives (fragment 59), multiple acetophenone (fragment 9) and cyclodextrin (fragment 41). Then, the fragments were assembled to generate **Bio-hMOFs** by using the tinker-toy topological assembly algorithm[28] provided in PORMAKE[29] (version 0.2.1) with two configurations: (1) one inorganic node and one organic edge (or linker) and (2) one inorganic node and one organic node. More details about the generation of **Bio-hMOFs** are provided in **S1** of the SI.

The generated structures underwent geometry optimization using the LAMMPS simulation package (version August 2023)[30] with the universal force field (UFF).[31] Briefly, the optimization procedure comprised the steepest descent (SD) with frozen cell boundaries, SD with cell relaxation, and then three cycles of FIRE[32] minimization with frozen cell, followed by SD minimization with cell relaxation. All the minimization steps were performed with a convergence criterion of 10^{-8} stopping tolerance and 50 000 maximum iterations for force and energy evaluations. Thereafter, we conducted structural checks using MOFID[33,34] (version 1.1.0) in order to eliminate structures with under-coordinated or hyper-coordinated atoms. Furthermore, the lammps-interface[35] (version 0.2.2) was used to detect and filter out disordered structures containing free molecules. It is important to highlight the necessity for performing robust structural checks, as highlighted by recent studies,[36,37] and the development of structural check algorithms in the future can be applied to eliminate problematic structures (*e.g.*, metal centers with erroneous oxidation states). Consequently, a database with 17 681 **Bio-hMOFs** was curated (see Table S1 for statistics). Next, the structural and feature diversity of **Bio-hMOFs** was analyzed and compared with popular hypothetical databases (BW-DB, ARC, hMOF, Tabacco and ultrastable)[22–26] in terms of geometric and chemical features. Consisting of pore size, surface area, pore volume (see Table S2), the geometric features were computed using Zeo++ (version 0.3).[38] The revised autocorrelated functions (RACs) were adopted as the chemical features (see Table S3) as computed from molSimplify (version 1.7.3),[39] employing the same methodology employed by Moosavi *et al.*[39,40] Five categories of RACs were evaluated based on the atom centers over neighboring atoms (within a maximum of three bond distance): (1) metal centers (metal-centered RACs), (2) linker-connecting atoms (linker-connecting RACs), (3) full linkers (full-linker RACs), (4) functional group atoms (functional group RACs) and (5) over entire structure (full-scope RACs). The feature space of RACs was analyzed using unsupervised learning methods (*i.e.*, t-SNE). Additional details regarding geometric features and RACs are provided in S2.

To demonstrate a potential application of **Bio-hMOFs** in signaling gases NO and CO,[16] we conducted grand-canonical Monte Carlo (GCMC) simulations to evaluate the selective adsorption of an equimolar NO/CO mixture in **Bio-hMOFs** at 298 K, under pressures of 1 bar and 10 bar, respectively. The force field parameters were adopted from previous studies as detailed in S3. Due to the competing nature of NO/CO endogenous gases in binding toward the active sites of biological receptors (*e.g.*, soluble guanylate cyclase), identifying materials with selective NO or CO adsorption will be useful for biological applications (*e.g.*, *in vivo* delivery of signaling molecules).[20,21] The selective adsorption performance of **Bio-hMOFs** was quantified by using two metrics: adsorption capacity N_x and selectivity $S_{x/y} = \frac{N_x/p_x}{N_y/p_y}$, where p is the partial pressure. Compared to CO, NO is a more essential signaling molecule, and our analysis was primarily based on the selectivity of NO with respect to CO (*i.e.*, $S_{NO/CO}$).

Furthermore, as mentioned above, it is crucial to consider the mechanical stability of hypothetical structures like **Bio-hMOFs** for a holistic assessment. The mechanical properties of **Bio-hMOFs**, including bulk (K), shear (G) and Young (E) moduli, were evaluated from MD simulations as described in more detail in S4. Regarding the choice of force field to compute mechanical properties, it is important to note that UFF4MOF[41] overestimates the mechanical and thermal expansion properties of certain MOFs,[35,42] which may potentially lead to outliers when employed in computational screening. While recent studies have demonstrated the accuracy of machine-learned potentials (MLPs) in predicting mechanical properties of

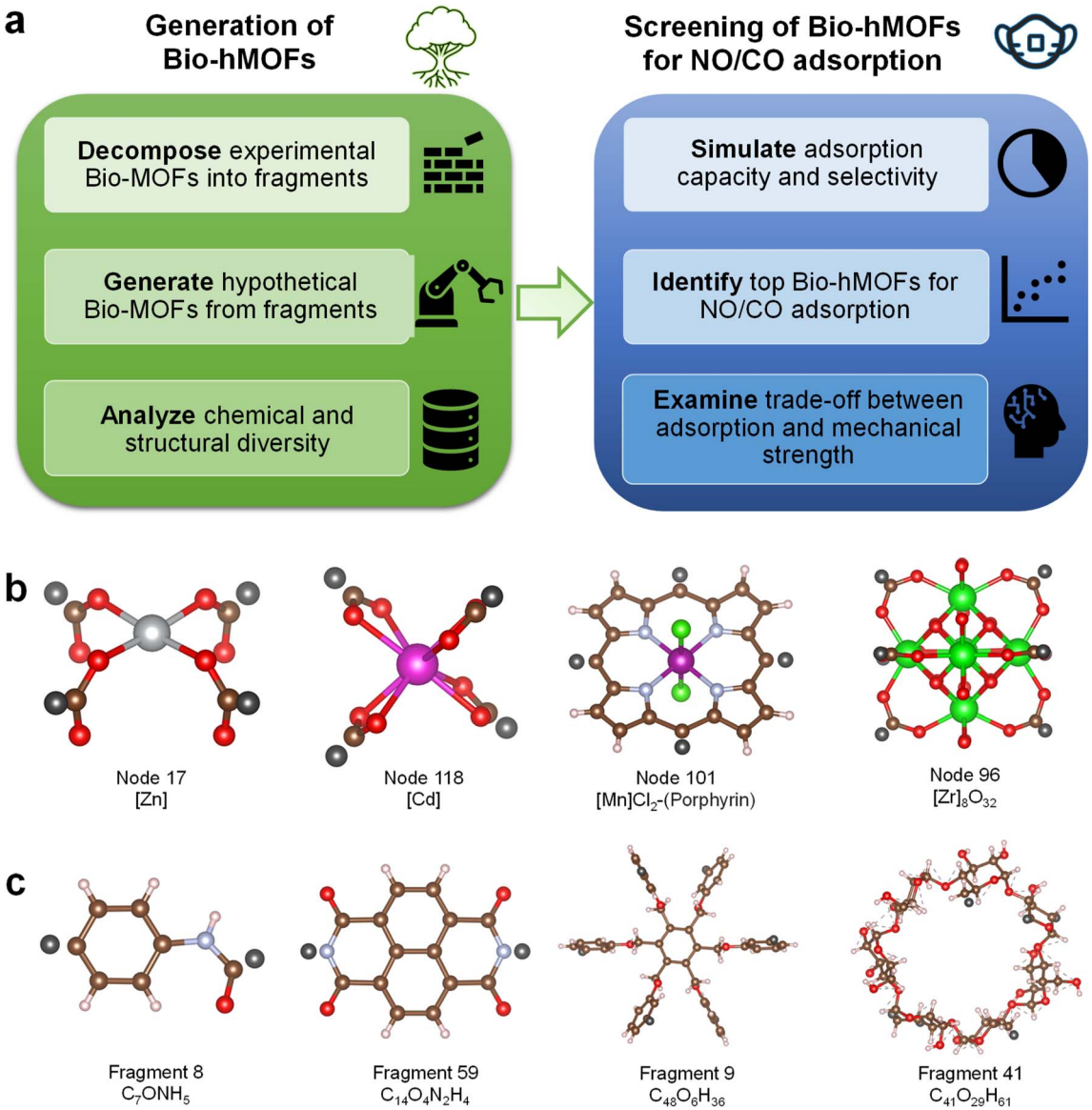

Fig. 1 (a) Workflow with two main stages: generation of **Bio-hMOFs** (left panel) and screening of mechanically stable **Bio-hMOFs** for NO/CO adsorption (right panel). Representative fragments from experimental Bio-MOFs: (b) inorganic fragments (*i.e.*, metal nodes) and (c) organic fragments. The dark spheres indicate points of connection to other possible fragments.

MOFs,[43,44] scaling MLP-MD simulations to a large number of **Bio-hMOFs** is computationally inaccessible. Thus, we employed the UFF[31] instead of UFF4MOF[41] and MLPs[43,44] as it offers a good balance between accuracy and scalability in the evaluation of mechanical properties. Finally, the structural–property relationships and trade-off between adsorption capacity and mechanical stability were evaluated and discussed.

Results and discussion

Chemical and structural diversity

As discussed above, prior studies leveraging computational screening to uncover Bio-MOFs for various applications were limited due to the scarce number of Bio-MOFs available. While recent studies have demonstrated the possibility of identifying Bio-MOFs for indoxyl sulfate adsorption[14] and O_2/N_2 separation,[15] such structures are often un-sanitized (*e.g.*, containing free solvents). Thus, it is important to analyze the chemical and structural diversity of the developed **Bio-hMOF** database. As illustrated in Fig. S1, there is a wide variety of metal nodes and chemical substructures present in **Bio-hMOFs**, despite only considering fragments from experimental Bio-MOFs. This is imperative in ensuring that the **Bio-hMOF** database encompasses a sufficiently large and diverse chemical space for data-driven studies. The common metal nodes are generally single nodes (*e.g.*, [Zn]) or paddlewheel SBUs comprising d-block transition metals. Although this results in a limited proportion of 2-row d-block (*e.g.*, Ru) and post-transition metals (*e.g.*, Er), there are many possible SBUs that can form from single metal nodes surrounded by different organo-metallic

environments (Fig. S1b). In terms of chemical substructures, common organic linkers such as benzene dicarboxylic acid (BDC) are present. In addition, **Bio-hMOFs** contain various biomolecular moieties (*e.g.*, derivatives of imidazole, pyridine and pyrene dicarboxylic acid, PDC), which are essential functional groups for biological functions.[45] These functional groups are also present in notable MOFs (*e.g.*, imidazole in zeolitic-imidazolate frameworks), which exhibit tunable chemical properties and efficient gas storage functionalities.[46] In terms of topology, the majority of **Bio-hMOFs** possess topological nets of ske, nbo and qdl that are less common in other databases; notably, a large majority (38.0%) of ARC-MOFs have pcu topology.[22] As a result, the distribution of topological nets among **Bio-hMOFs** is relatively more balanced (Fig. S2). These diverse chemical moieties and topological configurations render the **Bio-hMOF** database suitable for screening and data-driven workflows.

Next, we compare the geometric properties of **Bio-hMOFs** with those of other hypothetical databases.[22–26] As shown in Fig. 2 and S3, most geometric properties of **Bio-hMOFs**, such as pore limiting diameter (PLD), largest cavity diameter (LCD), gravimetric or volumetric surface area (GSA/VSA), and gravimetric pore accessible volume (GPOV), exhibit distributions similar to other databases. For the void fraction (VF), **Bio-hMOFs** possess a more uniform kernel density estimation (KDE) distribution in comparison with other databases (Fig. S3). This is possibly attributed to the composition of relatively shorter linkers (*e.g.*, derivatives of oxalate and BDC linkers) in

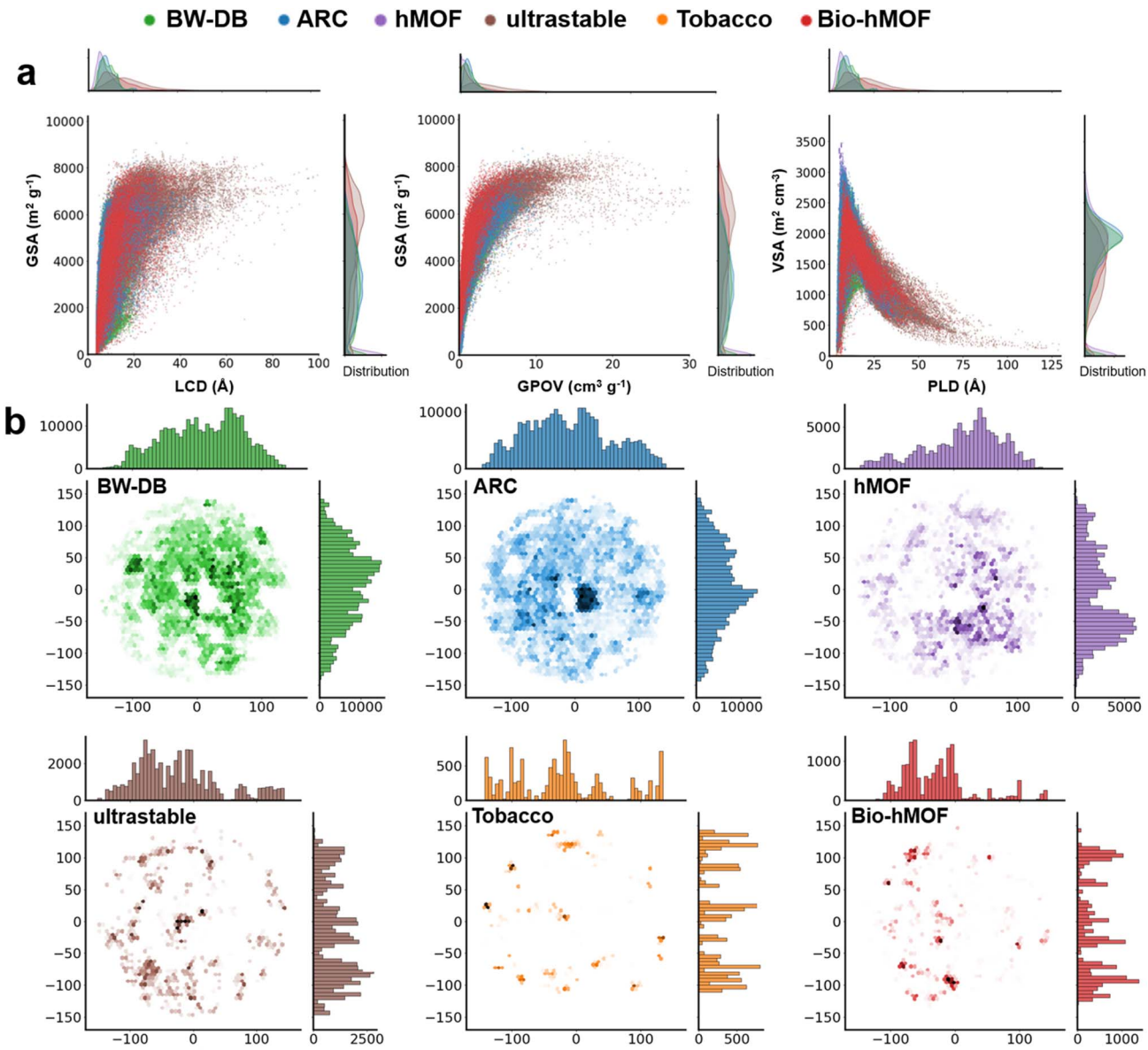

Fig. 2 (a) Geometric features in Bio-hMOF and other hypothetical databases, with the kernel density estimation (KDE) distributions plotted on the marginal axes. From left to right: GSA against LCD, GSA against GPOV, and VSA against PLD. (b) t-SNE maps of chemical features, with the KDE distributions on the marginal axes. The KDE was normalized to the total number of MOFs in each database: 269 391 ARC-MOF,[22] 323 789 BW-DB,[23] 54 139 ultrastable,[24] 11 577 ToBaCCo,[25] 114 658 hMOF,[26] and 17 681 Bio-hMOF (this work).

most **Bio-hMOFs**, which lead to a slightly lower porosity. Moreover, we note that hypothetical structures constructed using topological nets with multiple nodes and edges (*e.g.*, one metal node and two types of organic edges) were not included in **Bio-hMOF** due to their exceedingly high porosity. Nevertheless, the balanced distribution of VF in the **Bio-hMOF** database is advantageous for improving transferability from data-driven analysis of **Bio-hMOFs** to experimental Bio-MOFs, because most hypothetical databases tend to possess geometric properties that are skewed toward a higher VF as compared to experimental databases.[22]

The **Bio-hMOF** database is further compared with other hypothetical databases in terms of chemical features. Fig. 2b and S4 show the t-distributed stochastic neighbor embedding (t-SNE) maps of the RACs in different databases. It is observed that ARC-MOFs are evenly distributed in the t-SNE space, as ARC-MOFs[22] were selectively composed to attain a good trade-off between balance and variety of chemical features.[47] The RACs of **Bio-hMOFs** are well-distributed across the entire t-SNE space despite their low quantity (*c.a.*, 2.2% of the entire space), which can be corroborated with the distribution of **Bio-hMOFs** in the t-SNE maps of RACs for full-linker, metal-centered, linker-connecting and functional groups, respectively (Fig. S5). This indicates that **Bio-hMOFs** comprise diverse metal nodes and organic linkers in the MOF ecosystem. While the agglomeration of clusters is observed in the t-SNE maps of RACs describing various chemical types (Fig. S5), this is a result of the (inevitable) imbalance of element compositions of building

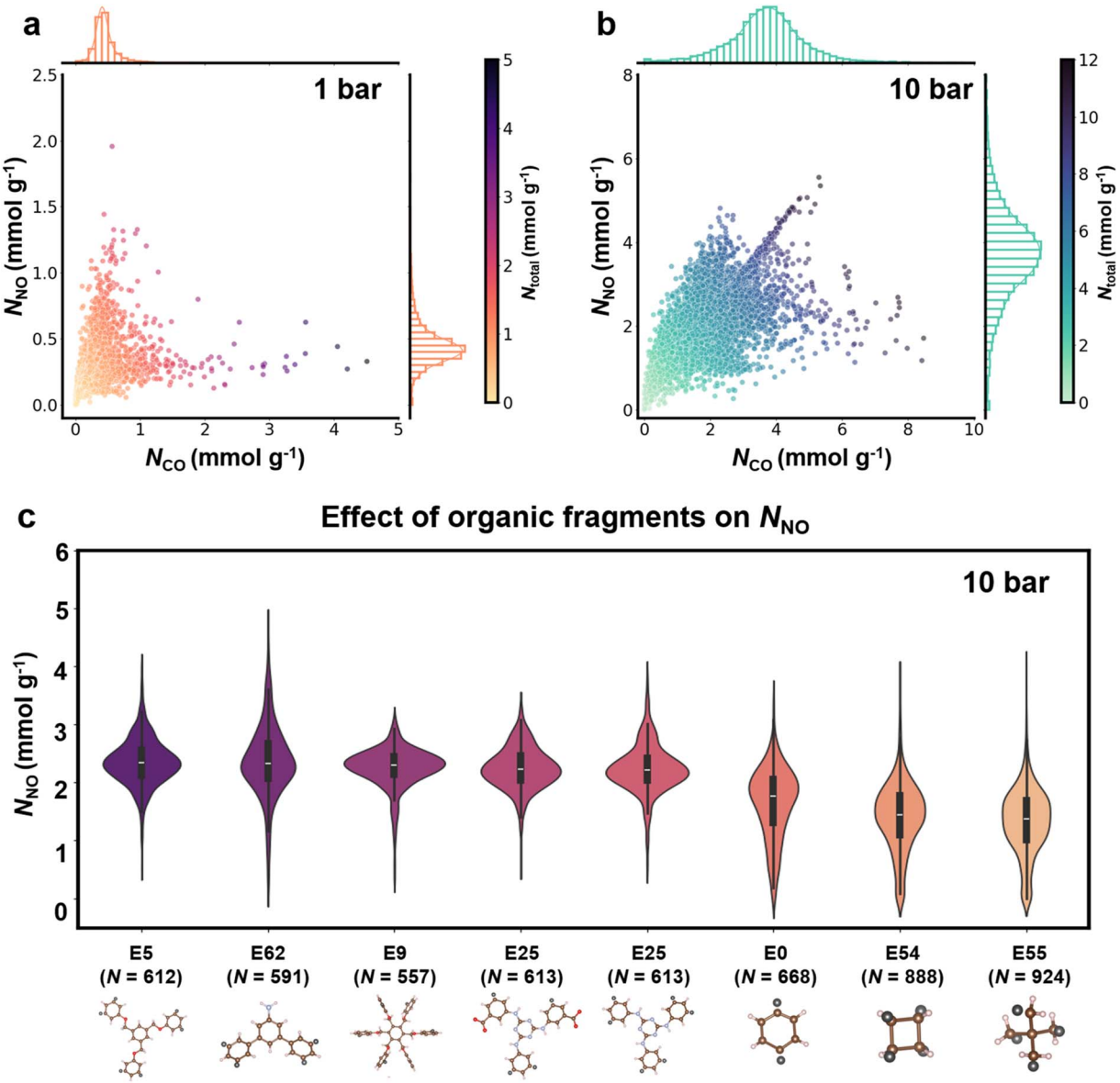

Fig. 3 Relationships between N_{NO} and N_{CO} in an equimolar NO/CO mixture at 298 K under pressures of (a) 1 bar and (b) 10 bar. (c) Violin plot for the effect of common organic fragments on N_{NO} at 10 bar. The violins are sorted from left to right based on descending median N_{NO}; see the corresponding plot for N_{CO} in Fig. S9.

blocks involved in the construction of hypothetical databases.[22] Despite a small skew due to the restriction on building blocks, **Bio-hMOFs** are relatively well-dispersed in the t-SNE space in comparison to other larger databases. Overall, our analysis indicates that the **Bio-hMOF** database exhibits diverse and balanced structural and chemical features, which renders it well suited for data-driven screening and machine learning tasks.

Selective adsorption of NO/CO

We now present a demonstrative application of utilizing **Bio-hMOFs** for the selective adsorption of NO/CO, which is of biological significance in therapeutic gas delivery. Notably, NO and CO possess a similar kinetic diameter ($\sigma_{CO} = 3.49$ Å and $\sigma_{NO} = 3.69$ Å)[48] and exhibit competitive binding toward biological receptors (*e.g.*, both bind to ferrous ions in sGC receptors).[21,49,50] Therefore, in the design of functionalized Bio-MOFs for signaling gas adsorption and delivery, it is imperative to ensure high selectivity toward either NO or CO for the intended purpose.[6,7] However, we note that, to the best of our knowledge, experimental NO or CO adsorption data are currently unreported in the literature, even for well-established bio-MOFs (*e.g.*, bio-MOF-1 [51]). While the force field parameters for NO and CO adsorption were validated in certain MOFs,[17,19] we caution that extensive validation in bio-MOFs is not yet possible. Thus, we provide the simulated adsorption isotherms of pure NO, CO and NO/CO mixture in two pertinent experimental bio-MOFs (bio-MOF-1 and bio-MOF-12) in Fig. S6 for future reference.

Fig. 3 shows the relationships between NO (N_{NO}) and CO (N_{CO}) adsorption capacities in an equimolar NO/CO mixture predicted from GCMC simulations. There is a greater capacity at a higher pressure, whereas the adsorption selectivity is not significantly affected by pressure (Fig. S7). N_{NO} and N_{CO} are not directly correlated with the inorganic fragment type (Fig. S8 and S9), although they are influenced by the organic fragment type (Fig. 3c and S10). This is due to the greater impact of pore geometry that governs the predominant physisorption process for NO/CO as compared to chemisorption. Although NO may undergo chemisorption at the open metal sites (OMS), the presence of large pores is equally important for high NO uptake.[17] Conversely, previous simulations show that interaction between OMS and CO is less significant, and CO preferentially undergoes physisorption.[19] As a result, **Bio-hMOFs** with longer and sparser organic fragments such as multi-phenyl chains (*e.g.*, E60 and E62) exhibit stronger adsorption for both NO and CO, as compared to their counterparts with shorter fragments (*e.g.*, E0, E54, and E55).

As discussed above, the metal node type has a significant impact on NO adsorption. As shown in Fig. S11, **Bio-hMOFs** containing Zn- and Cu-nodes, particularly with accessible OMS (*e.g.*, N102 and N21), favor higher N_{NO}. Regarding selectivity $S_{NO/CO}$, most **Bio-hMOFs** exhibit a selectivity close to unity (due to the similar kinetic diameter of NO and CO), despite a slight preference for NO (median $S_{NO/CO}$ of 1.17 at 1 bar). **Bio-hMOFs** exhibiting high $S_{NO/CO}$ are attributed to the presence of OMS, such as the pillared Zn-node (N102), which favors interaction with NO. Hence, these **Bio-hMOFs** are potentially promising for selective NO sensing and loading. Conversely, **Bio-hMOFs** containing coordinatively saturated Zn- and Cd-nodes (N6 and N118) possess higher CO selectivity, indicating that metal node type has an insignificant impact on CO adsorption. **Bio-hMOFs** with high NO and CO selectivity are listed in Tables S8 and S9, respectively. As elucidated in ZIFs, certain functional groups (*e.g.*, nitro-imidazolate in ZIF-68) instead of OMS serve as binding sites for CO. Thus, tuning of pore geometry and functional groups is more effective to increase CO adsorption.[19] Taken together, the above analysis of structure-adsorption relationships indicates that selective signaling gas adsorption can be tailored by tuning metal nodes and functional groups in **Bio-hMOFs**.

Mechanical properties

Beyond adsorption, it is imperative to consider the mechanical properties of MOFs for practical applications. Here, we examine the influence of chemical compositions of **Bio-hMOFs** on their mechanical properties. As shown in Fig. 4, 1046 **Bio-hMOFs** (~6% of 17 681) possess high bulk moduli ($K_H > 20$ GPa), which is possibly due to the relatively low VF possessed by **Bio-hMOFs**. Meanwhile, approximately 919 (~5%) **Bio-hMOFs** exhibit exceptional K_H and E_H (shaded region in Fig. 4a). The majority of 919 contain bulkier organic fragments (E54: 24% and E55: 15.7%) that correspond to lower signaling gas adsorption (Tables S10 and S11). As porosity increases with the length of organic linkers, it is expected that organic linkers containing multiple phenyl chains (*e.g.*, E6 and, E62) lead to lower K_H. In contrast, shorter linkers (*e.g.*, E54 and E0) lead to higher K_H. The same conclusions can be drawn from the chemical compositions of 909 **Bio-hMOFs** with exceptional K_H and G_H (in Fig. 4b). While specific inorganic fragments on average promote slightly higher K_H (Fig. 4c), the effect of organic fragments on K_H is much more significant (Fig. 4d). This can be attributed to the impact of organic linker type on porosity.[52] However, the mechanical properties are also affected by the rigidity of inorganic fragments.[24,52] For instance, **Bio-hMOFs** containing Cd-nodes (*e.g.*, N118) possess a slightly higher K_H (median of 4.53 GPa) compared to more common Zn-nodes (*e.g.*, N6 with median $K_H = 3.12$ GPa). This is a result of the greater maximum coordination number (MCN) of Cd metal (MCN = 8) as compared to Zn (MCN = 4). Aside from the well-established influence of porosity on K_H,[52,53] metals with greater MCN, such as 2nd-row transition metals (Cd and Zr) and post-transition metals (Eu and Tb), reinforce the rigidity of metals and contribute to higher mechanical strength.[24,52]

Fig. 5 shows K_H *versus* N_{NO} at 1 bar. Notably, there exists a trade-off between adsorption and mechanical strength. On the Pareto frontier, five Pareto-optimal **Bio-hMOFs** with inorganic fragments, OMS (*e.g.*, Mn-based N128 and Fe-based N41) and short linkers (*e.g.*, E54) exhibit high N_{NO} and good mechanical strength (Tables S12 and S13). This underscores the importance of using appropriate building blocks, combining NO-selective metal nodes (or functionalized linkers) and topological nets (*e.g.*, zec) to facilitate the generation of porous regions (in Fig. 5b) that best optimize the trade-off. In terms of selectivity,

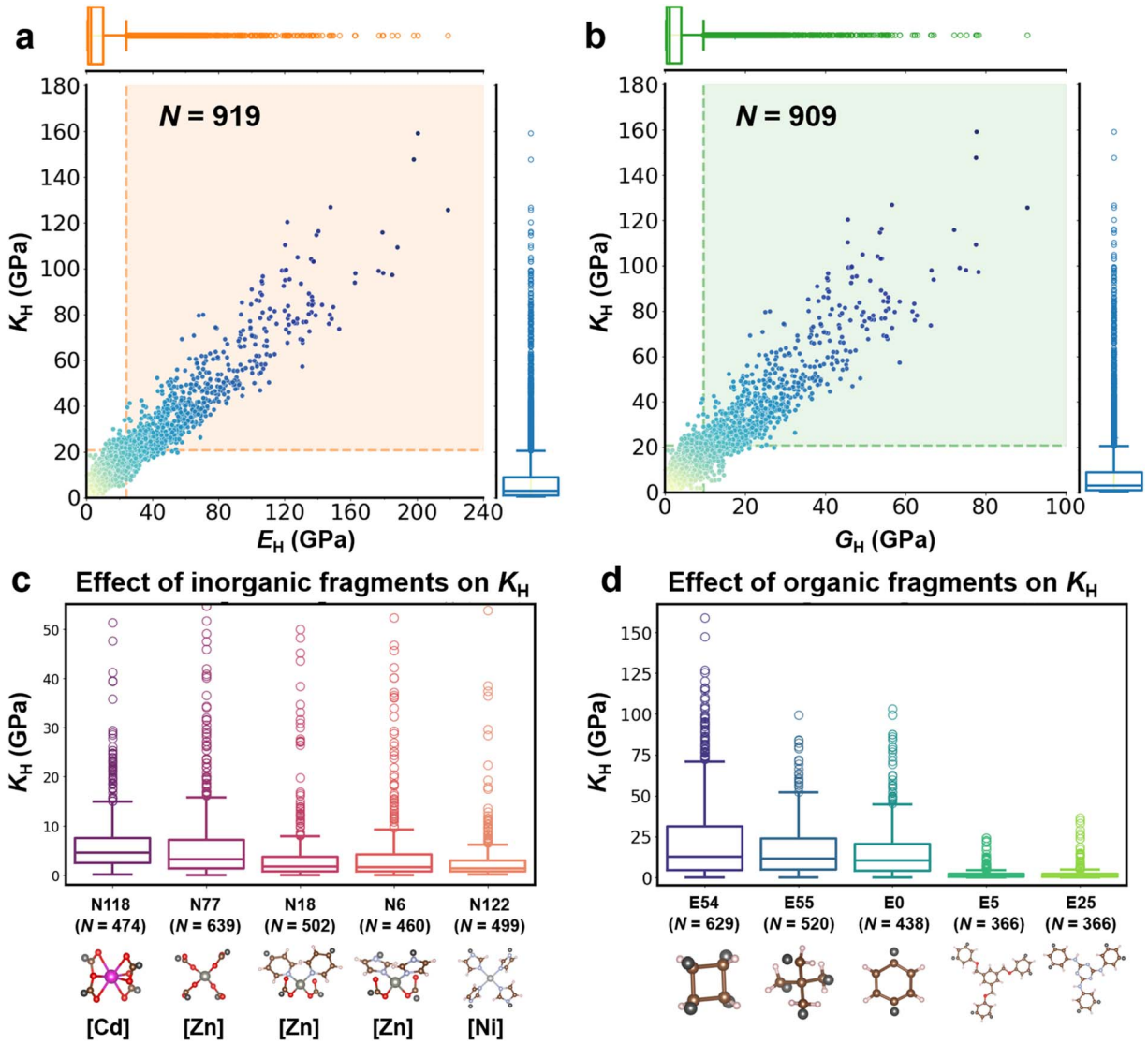

Fig. 4 (a) Bulk moduli K_H against shear moduli G_H (b) K_H against Young's moduli E_H for 10 748 **Bio-hMOFs**. The shaded region denotes exceptionally high moduli (1.5 × the upper threshold of the Inter-Quantile Range). K_H in **Bio-hMOFs** with the most common (c) inorganic fragments and (d) organic fragments, sorted by descending median K_H from left to right.

a **Bio-hMOF** containing a Mn-porphyrin node (N128) achieves the highest $S_{NO/CO} > 4$, due to the availability of Mn OMS to facilitate strong binding with NO (in Fig. 5b).[17] Furthermore, the Mn-porphyrin complex is also highly rigid, thus facilitating high mechanical strength. Similarly, a **Bio-hMOF** with the Fe-porphyrin node (N41) and topology net zec demonstrates high N_{NO} and K_H (in Fig. 5c).

Additionally, we find that **Bio-hMOFs** lying on the Pareto frontier also exhibit good NO selectivity ($S_{NO/CO} > 2$; Table S13). This can be attributed to the presence of accessible OMS in these **Bio-hMOFs**, which facilitate high NO uptake and selectivity. On the other hand, **Bio-hMOFs** with high mechanical stability and $S_{NO/CO}$ may not exhibit the desired high N_{NO} due to their low porosity (see Table S8 and examples in Fig. S11a). Thus, **Bio-hMOFs** possessing high mechanical strength and selective NO adsorption can be designed with the presence of accessible OMS, synergized with appropriate linkers and topological nets. Conversely, **Bio-hMOFs** containing suitable functional groups (*e.g.*, imidazolate[19]) exhibit highly selective CO adsorption (see examples in Fig. S11b). These **Bio-hMOFs** are potential carriers for targeted delivery of CO, which is required for anti-inflammatory treatment.[45]

Finally, it is worthwhile to note that metal toxicity analysis[14] indicated that common metals in MOFs (*e.g.*, Ni, Cd, and Cu)[54] are toxic when evaluated on the basis of median lethal dose (LD_{50}).[14] Therefore, **Bio-hMOFs** containing safe metals (*e.g.*, Zn, Co, Fe, Mn, Mg, and In) should be mainly considered when screening for biological applications, although all **Bio-hMOFs** were examined in this work for selective NO/CO adsorption. In practical biological applications,[7,24] tuning the organic fragments of Bio-MOFs (*i.e.*, biological ligands and functional groups) remains an important avenue toward their

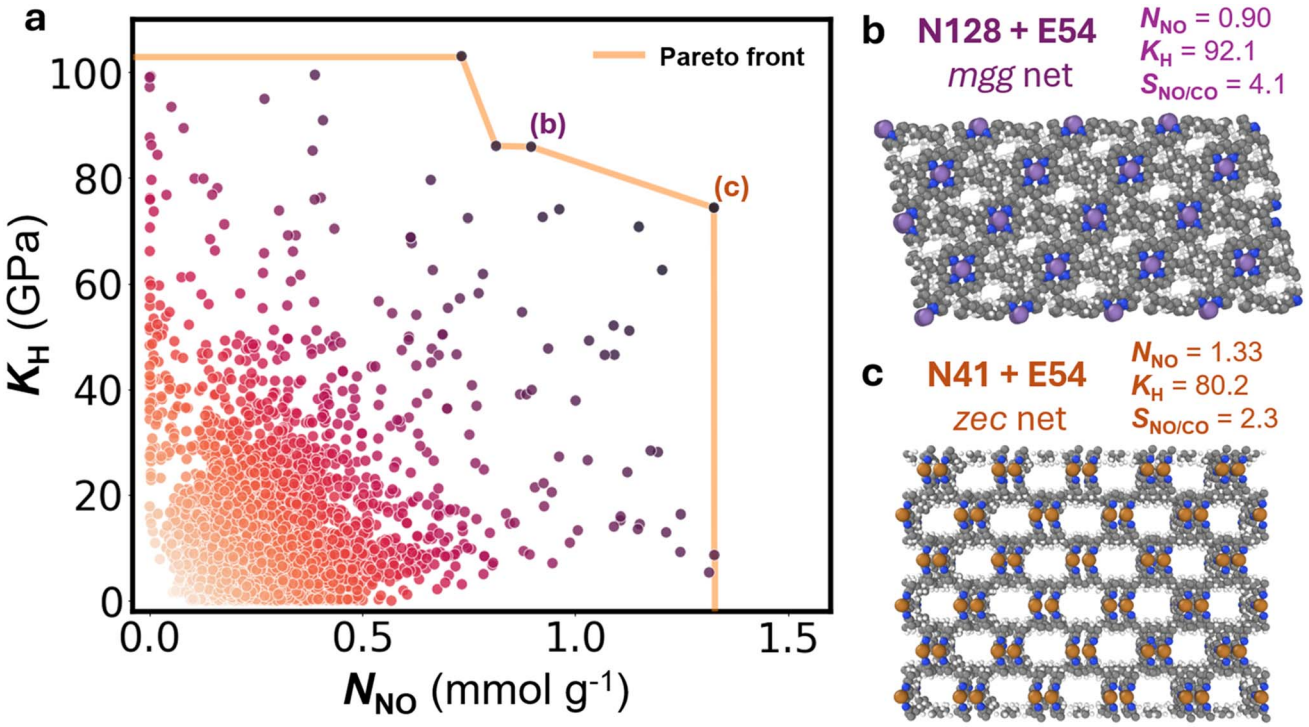

Fig. 5 (a) Relationship between K_H and N_{NO} and the Pareto frontier. (b) and (c) Structures with their fragments and topological nets. Color code: Mn (purple), Fe (bronze), N (blue), C (grey), and H (white).

functionalization, especially when only a few non-toxic metal types[14] are feasible for safe *in vivo* biological applications.[21]

Conclusion

We present a database of hypothetical **Bio-hMOFs** assembled from the fragments of 309 experimentally available Bio-MOFs. Following a series of geometry optimization and structural checks, the database comprises 17 681 structures. Subsequently, we demonstrate the application of **Bio-hMOFs** for the delivery of biological signaling gases. In addition to identifying top **Bio-hMOFs** exhibiting high N_{CO} and N_{NO}, our analysis elucidates that **Bio-hMOFs** with open metal sites are preferred for selective NO adsorption. By evaluating the mechanical properties of **Bio-hMOFs**, we show that a few **Bio-hMOFs** possess high mechanical strength. Finally, we analyze the effects of building block compositions (*i.e.*, inorganic/organic fragments) on the trade-off between adsorption capacity and mechanical strength. Overall, we elucidate that stable **Bio-hMOFs** containing metal-porphyrin based inorganic fragments facilitate selective NO adsorption, whereas functional groups and pore geometry play a more significant role in CO adsorption. Computational screening can further discover new Bio-MOFs for emerging biological applications. Future extension of the **Bio-hMOF** database may leverage data-driven approaches to include important properties relevant to biological compatibility, such as structural stability and biodegradability.[55]

Author contributions

Y. Y. and A. S. P. conceptualized the project. Y. Y. constructed the **Bio-hMOFs**. Y. Y. and S. A. M. designed and conducted MD simulations. Y. Y. and A. S. P. conducted GCMC simulations. Y. Y. analyzed the results and wrote the manuscript. All authors reviewed and edited the manuscript. J. J. supervised the project.

Conflicts of interest

There are no conflicts to declare.

Data availability

The data and codes presented in the article, including the crystal structures and geometric properties of **Bio-hMOFs**, and LAMMPS scripts to run MD simulations, can be accessed in the public Zenodo repository: **https://doi.org/10.5281/zenodo.13959121** Simulation packages employed in this work are licensed as open-source packages, including RASPA (version 2.0.47) and LAMMPS (August 2023).

The data supporting this article, including the construction and featurization of Bio-hMOFs, GCMC simulations, computation of mechanical properties, feature analysis, and the statistics of adsorption and mechanical properties, have been included as part of the SI. See DOI: **https://doi.org/10.1039/d5dd00213c**.

Acknowledgements

We gratefully acknowledge the Singapore National Research Foundation (NRF-CRP26-2021RS-0002), A*A*STAR LCER-FI (LCERFI01-0015 U2102d2004 and LCERFI01-0033 U2102d2006) for financial support. Y. Y. was supported by the PGF scholarship under the ISEP program funded by NUS Graduate School. We gratefully acknowledge computational facilities provided by NUS HPC (NUSREC-HPC-00001 and CFP01-CF-077).

References

1 P. Anastas and N. Eghbali, Green chemistry: principles and practice, *Chem. Soc. Rev.*, 2010, **39**, 301–312.
2 I. T. Horvath and P. T. Anastas, Innovations and green chemistry, *Chem. Rev.*, 2007, **107**, 2169–2173.
3 H. Furukawa, K. E. Cordova, M. O'Keeffe and O. M. Yaghi, The chemistry and applications of metal–organic frameworks, *Science*, 2013, **341**, 1230444.
4 Y. G. Chung, E. Haldoupis, B. J. Bucior, M. Haranczyk, S. Lee, H. Zhang, K. D. Vogiatzis, M. Milisavljevic, S. Ling, J. S. Camp, B. Slater, J. I. Siepmann, D. S. Sholl and R. Q. Snurr, Advances, updates and analytics for the computation-ready, experimental metal–organic framework database: CoRE MOF 2019, *J. Chem. Eng. Data*, 2019, **64**, 5985–5998.
5 A. Shahzaib, L. A. Kamran and N. Nishat, The Biomolecule-MOF Nexus: recent advancements in biometal–organic frameworks and their multifaceted applications, *Mater. Today Chem.*, 2023, **34**, 101781.
6 A. C. McKinlay, R. E. Morris, P. Horcajada, G. Férey, R. Gref, P. Couvreur and C. Serre, BioMOFs: metal–organic frameworks for biological and medical applications, *Angew. Chem., Int. Ed.*, 2010, **49**, 6260–6266.
7 H. Cai, Y.-L. Huang and D. Li, Biological metal–organic frameworks: Structures, host–guest chemistry and bio-applications, *Coord. Chem. Rev.*, 2019, **378**, 207–221.
8 C. D. Hodneland and M. Mrksich, Biomolecular surfaces that release ligands under electrochemical control, *J. Am. Chem. Soc.*, 2000, **122**, 4235–4236.
9 J. D. Chodera and D. L. Mobley, Entropy-enthalpy compensation: role and ramifications in biomolecular ligand recognition and design, *Annu. Rev. Biophys.*, 2013, **42**, 121–142.
10 W. Wang, O. Donini, C. M. Reyes and P. A. Kollman, Biomolecular simulations: recent developments in force fields, simulations of enzyme catalysis, protein-ligand, protein-protein, and protein-nucleic acid noncovalent interactions, *Annu. Rev. Biophys. Biomol. Struct.*, 2001, **30**, 211–243.
11 S. Quaresma, V. André, A. M. Antunes, S. M. Vilela, G. Amariei, A. Arenas-Vivo, R. Rosal, P. Horcajada and M. T. Duarte, Novel antibacterial azelaic acid BioMOFs, *Cryst. Growth Des.*, 2019, **20**, 370–382.
12 S. A. Noorian, N. Hemmatinejad and J. A. Navarro, BioMOF@cellulose fabric composites for bioactive molecule delivery, *J. Inorg. Biochem.*, 2019, **201**, 110818.
13 F. Sansone, L. Baldini, A. Casnati and R. Ungaro, Calixarenes: from biomimetic receptors to multivalent ligands for biomolecular recognition, *New J. Chem.*, 2010, **34**, 2715–2728.
14 B. Li, S. Gong, P. Cao, W. Gao, W. Zheng, W. Sun, X. Zhang and X. Wu, Screening of biocompatible MOFs for the clearance of indoxyl sulfate using GCMC simulations, *Ind. Eng. Chem. Res.*, 2022, **61**, 6618–6627.
15 S. He, M. Cheng, C. Liu, Z. Zhao, S. Chai, L. Zhou and X. Ji, High-throughput virtual screening of biometal–organic frameworks for O_2/N_2 separation, *Ind. Eng. Chem. Res.*, 2024, **63**, 2347–2360.
16 L. Li and P. Moore, An overview of the biological significance of endogenous gases: new roles for old molecules, *Biochem. Soc. Trans.*, 2007, **35**, 1138–1141.
17 B. Xiao, P. S. Wheatley, X. Zhao, A. J. Fletcher, S. Fox, A. G. Rossi, I. L. Megson, S. Bordiga, L. Regli and K. M. Thomas, High-capacity hydrogen and nitric oxide adsorption and storage in a metal–organic framework, *J. Am. Chem. Soc.*, 2007, **129**, 1203–1209.
18 Z. Zhou, B. Todd, K. P. Travis and R. J. Sadus, A molecular dynamics study of nitric oxide in water: diffusion and structure, *J. Chem. Phys.*, 2005, **123**, 054505.
19 A. Sirjoosingh, S. Alavi and T. K. Woo, Grand-canonical Monte Carlo and molecular-dynamics simulations of carbon-dioxide and carbon-monoxide adsorption in zeolitic−imidazolate framework materials, *J. Phys. Chem. C*, 2010, **114**, 2171–2178.
20 L. Wang, X. Xie, B. Ke, W. Huang, X. Jiang and G. He, Recent advances on endogenous gasotransmitters in inflammatory dermatological disorders, *J. Adv. Res.*, 2022, **38**, 261–274.
21 J. Sun, W. Wang, X. Hu, X. Zhang, C. Zhu, J. Hu and R. Ma, Local delivery of gaseous signaling molecules for orthopedic disease therapy, *J. Nanobiotechnology*, 2023, **21**, 58.
22 J. Burner, J. Luo, A. White, A. Mirmiran, O. Kwon, P. G. Boyd, S. Maley, M. Gibaldi, S. Simrod and V. Ogden, ARC–MOF: a diverse database of metal–organic frameworks with DFT-derived partial atomic charges and descriptors for machine learning, *Chem. Mater.*, 2023, **35**, 900–916.
23 P. G. Boyd, A. Chidambaram, E. García-Díez, C. P. Ireland, T. D. Daff, R. Bounds, A. Gładysiak, P. Schouwink, S. M. Moosavi and M. M. Maroto-Valer, Data-driven design of metal–organic frameworks for wet flue gas CO_2 capture, *Nature*, 2019, **576**, 253–256.
24 A. Nandy, S. Yue, C. Oh, C. Duan, G. G. Terrones, Y. G. Chung and H. J. Kulik, A database of ultrastable MOFs reassembled from stable fragments with machine learning models, *Matter*, 2023, **6**, 1585–1603.
25 Y. J. Colón, D. A. Gómez-Gualdrón and R. Q. Snurr, Topologically guided, automated construction of metal–organic frameworks and their evaluation for energy-related applications, *Cryst. Growth Des.*, 2017, **17**, 5801–5810.
26 C. E. Wilmer, M. Leaf, C. Y. Lee, O. K. Farha, B. G. Hauser, J. T. Hupp and R. Q. Snurr, Large-scale screening of

hypothetical metal–organic frameworks, *Nat. Chem.*, 2012, **4**, 83–89.
27 C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, The Cambridge structural database, *Acta Crystallogr. B*, 2016, **72**, 171–179.
28 P. G. Boyd and T. K. Woo, A generalized method for constructing hypothetical nanoporous materials of any net topology from graph theory, *CrystEngComm*, 2016, **18**, 3777–3792.
29 S. Lee, B. Kim, H. Cho, H. Lee, S. Y. Lee, E. S. Cho and J. Kim, Computational screening of trillions of metal–organic frameworks for high-performance methane storage, *ACS Appl. Mater. Interfaces*, 2021, **13**, 23647–23654.
30 A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, P. J. in't Veld, A. Kohlmeyer, S. G. Moore and T. D. Nguyen, LAMMPS: A flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, *Comput. Phys. Commun.*, 2022, **271**, 108171.
31 A. K. Rappé, C. J. Casewit, K. Colwell, W. A. Goddard III and W. M. Skiff, UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations, *J. Am. Chem. Soc.*, 1992, **114**, 10024–10035.
32 E. Bitzek, P. Koskinen, F. Gähler, M. Moseler and P. Gumbsch, Structural relaxation made simple, *Phys. Rev. Lett.*, 2006, **97**, 170201.
33 B. J. Bucior, A. S. Rosen, M. Haranczyk, Z. Yao, M. E. Ziebel, O. K. Farha, J. T. Hupp, J. I. Siepmann, A. Aspuru-Guzik and R. Q. Snurr, Identification schemes for metal–organic frameworks to enable rapid search and cheminformatics analysis, *Cryst. Growth Des.*, 2019, **19**, 6682–6697.
34 N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, Open Babel: An open chemical toolbox, *J. Cheminform.*, 2011, **3**, 1–14.
35 P. G. Boyd, S. M. Moosavi, M. Witman and B. Smit, Force-field prediction of materials properties in metal–organic frameworks, *J. Phys. Chem. Lett.*, 2017, **8**, 357–363.
36 M. Gibaldi, A. Kapeliukha, A. White, J. Luo, R. A. Mayo, J. Burner and T. K. Woo, MOSAEC-DB: a comprehensive database of experimental metal–organic frameworks with verified chemical accuracy suitable for molecular simulations, *Chem. Sci.*, 2025, **16**, 4085–4100.
37 A. J. White, M. Gibaldi, J. Burner, R. A. Mayo and T. K. Woo, High structural error rates in "computation-ready" MOF databases discovered by checking metal oxidation states, *J. Am. Chem. Soc.*, 2025, **147**, 17579–17583.
38 T. F. Willems, C. H. Rycroft, M. Kazi, J. C. Meza and M. Haranczyk, Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials, *Microporous Mesoporous Mater.*, 2012, **149**, 134–141.
39 E. I. Ioannidis, T. Z. Gani and H. J. Kulik, molSimplify: A toolkit for automating discovery in inorganic chemistry, *J. Comput. Chem.*, 2016, **37**, 2016–2117.
40 S. M. Moosavi, B. Á. Novotny, D. Ongari, E. Moubarak, M. Asgari, Ö. Kadioglu, C. Charalambous, A. Ortega-Guerrero, A. H. Farmahini and L. Sarkisov, A data-science approach to predict the heat capacity of nanoporous materials, *Nat. Mater.*, 2022, **21**, 1419–1425.
41 M. A. Addicoat, N. Vankova, I. F. Akter and T. Heine, Extension of the universal force field to metal–organic frameworks, *J. Chem. Theory Comput.*, 2014, **10**, 880–891.
42 Y. Yue, S. A. Mohamed and J. Jiang, Classifying and predicting the thermal expansion properties of metal–organic frameworks: a data-driven approach, *J. Chem. Inf. Model.*, 2024, **64**, 4966–4979.
43 S. Wieser and E. Zojer, Machine learned force-fields for an ab-initio quality description of metal-organic frameworks, *npj Comput. Mater.*, 2024, **10**, 18.
44 Y. Yue, S. A. Mohamed, N. D. Loh and J. Jiang, Toward a generalizable machine-learned potential for metal–organic frameworks, *ACS Nano*, 2024, **19**, 933–949.
45 S. Rojas, T. Devic and P. Horcajada, Metal–organic frameworks based on bioactive components, *J. Phys. Chem. B*, 2017, **5**, 2560–2573.
46 L. Zhang, Z. Hu and J. Jiang, Sorption-induced structural transition of zeolitic–imidazolate framework-8: A hybrid molecular simulation study, *J. Am. Chem. Soc.*, 2013, **135**, 3722–3728.
47 S. M. Moosavi, A. Nandy, K. M. Jablonka, D. Ongari, J. P. Janet, P. G. Boyd, Y. Lee, B. Smit and H. J. Kulik, Understanding the diversity of the metal–organic framework ecosystem, *Nat. Commun.*, 2020, **11**, 4068.
48 J.-R. Li, R. J. Kuppler and H.-C. Zhou, Selective gas adsorption and separation in metal–organic frameworks, *Chem. Soc. Rev.*, 2009, **38**, 1477–1504.
49 X. Ma, N. Sayed, A. Beuve and F. Van Den Akker, NO and CO differentially activate soluble guanylyl cyclase via a heme pivot-bend mechanism, *EMBO J.*, 2007, **26**, 578–588.
50 E. Martin, V. Berka, E. Bogatenkova, F. Murad and A.-L. Tsai, Ligand selectivity of soluble guanylyl cyclase: effect of the hydrogen-bonding tyrosine in the distal heme pocket on binding of oxygen, nitric oxide, and carbon monoxide, *J. Biol. Chem.*, 2006, **281**, 27836–27845.
51 J. An, S. J. Geib and N. L. Rosi, Cation-triggered drug release from a porous zinc – adeninate metal–organic framework, *J. Am. Chem. Soc.*, 2009, **131**, 8376–8377.
52 J. Lee, I. Lee, J. Park, H. Kim, M. Kim, K. Min and S. Lee, Optimal surrogate models for predicting the elastic moduli of metal–organic frameworks via multiscale features, *Chem. Mater.*, 2023, **35**, 10457–10475.
53 P. Z. Moghadam, S. M. Rogge, A. Li, C.-M. Chow, J. Wieme, N. Moharrami, M. Aragones-Anglada, G. Conduit, D. A. Gomez-Gualdron and V. Van Speybroeck, Structure-mechanical stability relations of metal–organic frameworks via machine learning, *Matter*, 2019, **1**, 219–234.
54 K. S. Egorova and V. P. Ananikov, Toxicity of metal compounds: knowledge and myths, *Organometallics*, 2017, **36**, 4071–4090.
55 A. Nandy, C. Duan and H. J. Kulik, Using machine learning and data mining to leverage community knowledge for the engineering of stable metal–organic frameworks, *J. Am. Chem. Soc.*, 2021, **143**, 17535–17547.