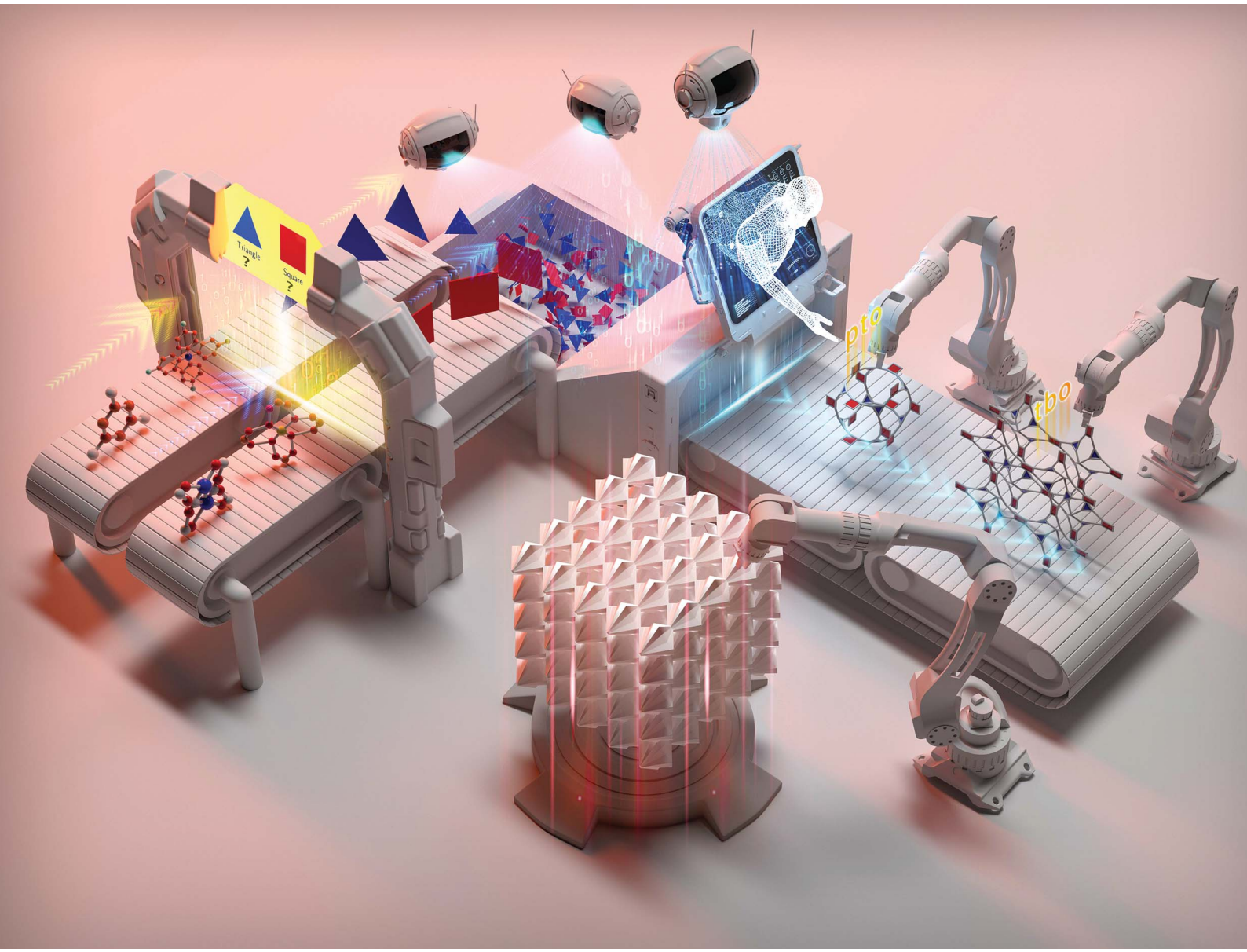
Volume 15
Number 40
28 October 2024
Pages 16395–16806

Chemical Science

rsc.li/chemical-science



ISSN 2041-6539

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

Xiaoyu Wu and Jianwen Jiang
Precision-engineered metal–organic frameworks: fine-tuning reverse topological structure prediction and design

Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal | View Issue

Check for updates

Cite this: *Chem. Sci.*, 2024, **15**, 16467

All publication charges for this article have been paid for by the Royal Society of Chemistry

Open Access Article. Published on 25 september 2024. Downloaded on 24.6.2025 20:57:50. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Precision-engineered metal–organic frameworks: fine-tuning reverse topological structure prediction and design†

Xiaoyu Wu and Jianwen Jiang *

Received 21st August 2024
Accepted 18th September 2024

DOI: 10.1039/d4sc05616g

rsc.li/chemical-science

Digital discoveries of metal–organic frameworks (MOFs) have been significantly advanced by the reverse topological approach (RTA). The node-and-linker assembly strategy allows predictable reticulations predefined by *in silico* coordination templates; however, reticular equivalents lead to substantial combinatorial explosion due to the infinite design space of building units (BUs). Here, we develop a fine-tuned RTA for the structure prediction of MOFs by integrating precise topological constraints and leveraging reticular chemistry, thus transcending traditional exhaustive trial-and-error assembly. From an extensive array of chemically realistic BUs, we subsequently design a database of 94 823 precision-engineered MOFs (PE-MOFs) and further optimize their structures. The PE-MOFs are assessed for post-combustion CO_2 capture in the presence of H_2O and top-performing candidates are identified by integrating three stability criteria (activation, water and thermal stabilities). This study highlights the potential of synergizing PE with the RTA to enhance efficiency and precision for computational design of MOFs and beyond.

1. Introduction

Porous crystals such as metal–organic frameworks (MOFs) represent a significant advancement in the field of functional materials, distinguished by their modular architectures and diverse chemical functionalities.[1] With these salient features, MOFs are considered versatile materials for many potential applications including gas storage, separation, catalysis, *etc.*[2] The evolution of MOF landscape has been remarkably enriched by the fusion of experimental synthesis and theoretical analysis, particularly through digital reticular chemistry.[3]

Experimentally synthesized MOFs, through systematically orchestrated assembly of metal nodes and organic linkers, have provided tangible crystals for characterization and validation. Meanwhile, theoretical methods, facilitated by automated assembly,[4] *ab initio* crystal structure prediction,[5,6] building unit (BU) replacement,[7] and predominantly, the reverse topological approach (RTA),[8] have exponentially expanded the conceivable landscape of MOFs and generated extensive hypothetical MOFs (hMOFs) without physical synthesis.[9–11] The RTA operates by starting with a desired topology template and then mapping it back to specific BUs based on connectivity patterns and topological optimization. Consequently, high-throughput computational screenings have been conducted on experimental MOFs or customized hMOFs and shortlisted potential candidates for many applications, particularly CO_2 capture.[12–15] Yet, as the field progresses, a major challenge persists in combinatorial explosion due to a wide variety of BUs and their enormous arrangements. This explosion in structural possibilities, while broadening the horizon of MOFs, imposes a substantial challenge in computation.

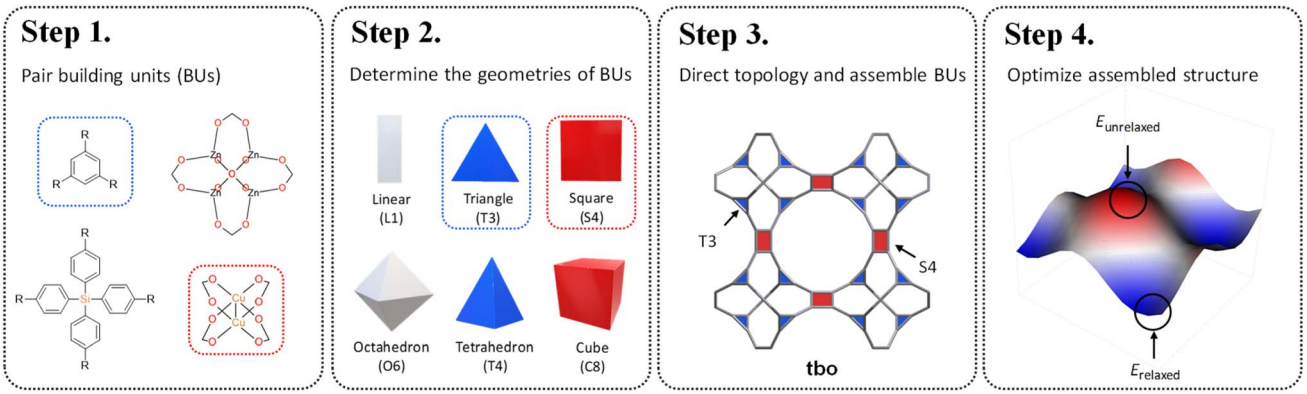
Reticular chemistry, a discipline that artfully weaves the threads of molecular architectures, offers an optimized palette of targeted nets controlled by the geometric and topological preference of BUs.[16] By integrating topological constraints,[17] one can specifically generate certain crystal configurations based on the geometric compatibility of BU combinations. Inspired by these benefits, in this work, we develop a fine-tuned RTA by introducing an additional layer of precision prerequisite. This refinement involves determining the geometries of BUs and selecting appropriate coordination templates guided by the reticular nature of MOFs. These criteria are governed by the local symmetry of BUs reticulated within a designated net. As illustrated in Scheme 1, our method starts with the one-on-one pairing of organic (red) and inorganic (blue) BUs, where their geometric signatures are subsequently matched and assigned. This approach strategically focuses on identifying a 'reticular ideal' topology, thereby alleviating the need for an exhaustive search across the entire topology space. The added dimension of precision engineering (PE) effectively mitigates the combinatorial explosion by filtering out infeasible topologies at an early stage, thereby catalyzing a substantially more

Department of Chemical and Biomolecular Engineering, National University of Singapore, 117576, Singapore. E-mail: chejj@nus.edu.sg

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4sc05616g

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 25 september 2024. Downloaded on 24.6.2025 20:57:50.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.



Scheme 1 Topological structure prediction with precision engineering. Step 1: pair building units; Step 2: determine the geometries of building units; Step 3: direct topology and assemble building units; Step 4: optimize assembled structure.

focused and accurate search for synthesizable MOFs given a vast array of BUs.

Prompted by the combinatorial guarantee demonstrated above, we present how large-scale RTA-based structure prediction of MOFs can be fine-tuned by PE for vast recursive inorganic–organic combinations of BUs, sourced from the well-established HEALED library.[18] To ensure a consistent basis for geometry determination, the connection points and the planarity of each candidate BU were standardized, followed by a similarity comparison against established geometric prototypes. Subsequently, we consulted a specialized table of topological constraints to specifically select topologies suited to the selected BUs before *in silico* synthesis and curation. Finally, a total of 94 823 PE-MOFs were distilled and carefully optimized by integrating the universal force field (UFF) and machine-learned atomic charges. The applicability of the PE-MOFs was demonstrated as adsorbents for post-combustion CO_2 capture. Top-performing candidates were identified and quantitative structure–performance relationships were established. These endeavors not only enrich the existing MOF bank, but also set the stage for future advancement in digital discovery of MOFs. The systematic approach outlined here showcases the potential for computational construction and high-throughput discovery of superior MOFs, ensuring that the next generation of porous crystals will be precisely tailored for specific industrial applications such as CO_2 capture and beyond.

2. Methodology

2.1. Geometric signatures

To ensure precise classification and compatibility check, we defined a comprehensive set of 13 geometric prototypes including 2-connected linear (L2), 3-connected triangle (T3), 4-connected square (S4), 4-connected tetrahedral (T4), 6-connected hexagonal (H6), 6-connected octahedral (O6), 6-connected trigonal prism (T6), 8-connected cubic (C8), 12-connected cuboctahedra (C12), 12-connected icosahedral (I12), 12-connected hexagonal prism (H12), 12-connected truncated tetrahedral (T12), and 24-connected truncated octahedron (T24). Classification of BUs began by determining the number of connection points, followed by evaluating the planarity where applicable, specifically for L2, T3, T4, S4, H6, C8, and T24 geometries. For instance, a BU with 3 connection points was directly classified as T3; a BU with 4 connection points was designated as tetrahedral if it failed the conservative planarity check. Should a BU not conform to these preliminary classifications, further steps would compare it with the detailed geometric prototypes by examining the root-mean squared deviations (RMSDs).[19] An example is illustrated in Fig. 1 where a BU namely m1 from the HEALED library is assigned with an O6 geometric signature. However, it is worthwhile to note that there are challenges in classifying non-perfectly shaped BUs and highly connected BUs, particularly those with 12 or more connections. These BUs may not always align precisely with desired geometric signatures, even with thorough visual inspections. For instance, Fig. S1† demonstrates how a highly connected vertice BU, namely m345 (sourced from a MOF with **fcu** topology[20]), combined with an edge BU o13, leads to a structure that, while identical in coordination, diverges in linker shape due to the steric effects of the naphthalene linker. In cases where a BU exhibits an imperfect geometric signature, it can still be reticulated into a topology analogous to a perfect BU. As depicted in Fig. S2,† despite imperfect geometry, o165 (o255) can still assemble into **edc** (**ftw**) topology with m345, similar to o189 (o47) with an ideal BU. This adaptability enables the navigation of a broader topological diversity, more aligned with real-world scenarios where perfectly shaped BUs are less common, compared to methods that rely strictly on idealized BUs (Fig. S3†). The full list of the geometric prototypes mathematically constructed in this work is provided in the ESI.†

2.2. Topological constraints

The coordination complexity of MOF structures arises from the wide variety of possible geometric signatures.[17] By carefully selecting BUs with appropriate shapes, topological constraints can be targeted. We introduced a reticular reference table for binary combinations of geometric signatures (ESI†), providing information on the most 'reticular ideal' topology to form. For instance, Cu-paddlewheel (Cu_2O_8) and 1,3,5-benzene

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 25 september 2024. Downloaded on 24.6.2025 20:57:50.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

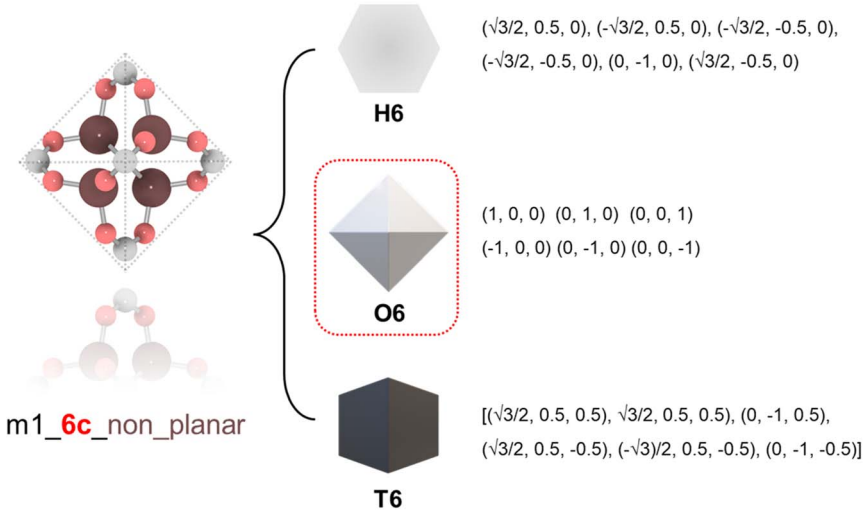


Fig. 1 Classification of a BU namely m1 into different geometric prototypes with mathematical representations: hexagonal (H6), octahedral (O6), and trigonal prism (T6). The BU is first identified to be non-planar 6-connected thus excluding H6, and then it is computed with the smallest RMSD against O6 and T6; finally, O6 is determined as the geometric signature.

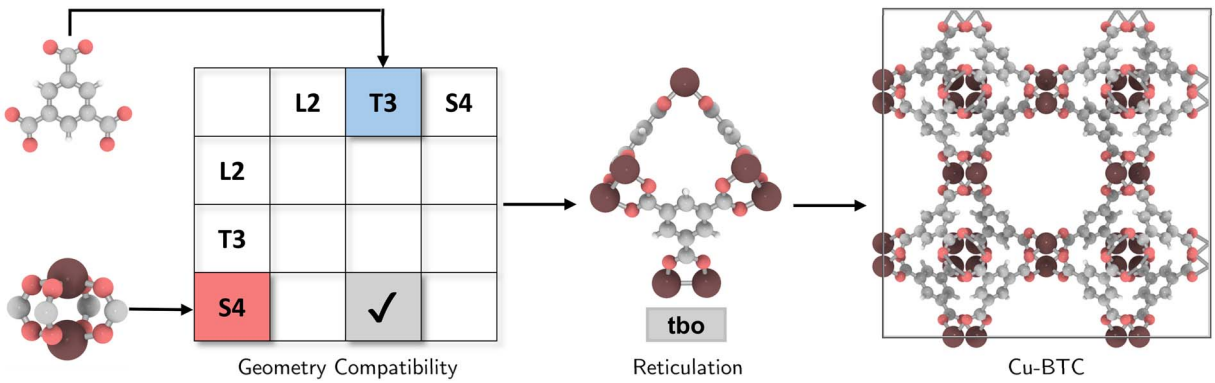


Fig. 2 Selection of a compatible topology for the assembly of Cu paddlewheel (Cu_2O_8) and 1,3,5-benzene tricarboxylate (BTC), followed by reticulation into a 'reticular ideal' **tbo** topology, constructing Cu-BTC.

tricarboxylate (BTC) can be assigned geometric signatures as T3 and S4, respectively. While conventional connectivity-based methods[10,21,22] would allow these BUs to assemble into several topologies such as **ctn**, **bor**, **pto**, and **tbo** (Fig. S4†), not all these combinations result in practical or favorable structures due to potential geometric distortions in structure prediction (Fig. S5†). Considering the shape compatibility of BUs, we further refined the selection of topologies, only targeting geometrically coherent structures, and constructed precise and ideal structures such as Cu-BTC[23] or HKUST-1 (Fig. 2). It is important to note that certain topologies are absent from the publicly available topology database under the RCSR project[24] and, as such, were not included in our table.

2.3. Topological assembly

With topological constraints, the molecular LEGO-like PORMAKE algorithm[10] was implemented for *in silico* assembly utilizing BUs from the HEALED library.[18] A total of 200 875 MOFs were initially generated with a RMSD cut-off at 0.3 Å. These structures were distilled to 94 823 PE-MOFs by employing a timeout-based strategy as previously used for computationally curating hMOFs[11] and covalent–organic frameworks (COFs).[25] The 94 823 PE-MOFs were structurally optimized by using the SMART algorithm of Forcite module in Materials Studio. The UFF[26] and machine-learned PACMOF atomic charges[27] were adopted to describe bonded and non-bonded interactions. In the calculations of atomic charges, a five-minute timeout was implemented to exclude structures being difficult to featurize for PACMOF. The bond information of both inter- and intra-BUs predefined using PORMAKE was preserved during optimizations to avoid unreasonable distance-based bond calculations, which might be caused by improperly configured BUs while fitting into selected topologies. During optimization, a twenty-minute wall-time limit was set to filter out structures that failed to converge. For each structure, bond information was retained and charge information was recalculated. The charge-recalculation and geometric optimization loop were repeated

Open Access Article. Published on 25 september 2024. Downloaded on 24.6.2025 20:57:50. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

three times to ensure the accuracy of the final set of PE-MOF structures, with a wall-time limit consistently applied throughout these stages.

2.4. Diversity analysis

Geometric and chemical features were used for diversity analysis of the designed PE-MOFs. The geometric features including pore limiting diameter (PLD), largest cavity diameter (LCD), channel dimensionality (Dimen), pore volume (PV), void fraction (VF), gravimetric surface area (GSA) and global cavity diameter (GCD) were computed *via* Zeo++ [28] utilizing a probe radius of 1.86 Å. The diverse chemical features were encoded using revised autocorrelations (RACs),[29] which were initially developed to fingerprint open-shell transition metal complexes[30,31] and have been adapted to encode MOFs.[32–35] The RACs were computed *via* MolSimplify,[36] providing the product or difference of atomic heuristics based on non-weighted crystal graphs. A wide variety of properties such as Pauling electronegativity (χ), nuclear charge (Z) connectivity (T), atom identity (I), covalent radii (S) and polarizability (α) were included in the RACs. Disparity (D), variety (V) and balance (B) were computed as diversity metrics[32] based on unsupervised machine-learned 2D embeddings of the RACs mapped by *t*-distributed Stochastic Neighbor Embedding (*t*-SNE).[37]

2.5. Molecular simulation for CO_2 capture

The PE-MOFs were assessed as adsorbents for CO_2 capture from a flue gas by pressure-swing adsorption. The flue gas was mimicked using a CO_2/N_2 mixture with a composition of 0.15/0.85. Considering the kinetic diameters of CO_2 (3.30 Å) and N_2 (3.64 Å), a subset of 17 173 PE-MOFs with PLDs ranging from 3 to 7 Å was chosen. Grand-canonical Monte Carlo (GCMC) simulations were conducted to compute the adsorption of the flue gas at 298 K and 1 bar and the desorption of pure CO_2 at 298 K and 0.01 bar. The PE-MOFs with performance superior to zeolite-13X[38] were shortlisted and their water, thermal and activation stabilities were evaluated using MOFSimplify.[39,40] For the stable ones, their CO_2 capture performance was further evaluated in the presence of H_2O. GCMC simulations were extended for the adsorption of a ternary mixture of CO_2, N_2 and H_2O at 298 K and 1 bar, with 60% relative humidity. CO_2 and N_2 were represented using the transferable potentials for phase equilibria (TraPPE) force field,[41] and H_2O was mimicked by the four point TIP4P model.[42] The Lennard-Jones interactions of framework atoms were modeled using the UFF[26] and DREIDING force field.[43] Table S1† lists the force field parameters. For cross LJ interactions, the potential parameters were estimated using the Lorentz–Berthelot mixing rules. The atomic charges of PE-MOFs were assigned using PACMOF,[27] which efficiently estimates charges with comparable accuracy to the density derived electrostatic and chemical (DDEC) method.[44] As exemplified in Fig. S6,† a determination coefficient (R^2) of 0.933 was achieved by comparing PACMOF charges and DDEC charges in four representative PE-MOFs. Following the DDEC6 protocol,[45] the DDEC charges here were computed using the CHARGEMOL package (**https://github.com/berquist/chargemol**) based on the calculations from the Vienna *Ab initio* Simulation Package (VASP)[46,47] at the PBE-D3(BJ) level.[48–50] Fig. S7† also shows CO_2, N_2 and H_2O uptakes (at 60% relative humidity) simulated using PACMOF charges and DDEC charges, respectively, in the four PE-MOFs. The uptakes based on both charges are close, suggesting an insignificant effect of the charge assignment method. During GCMC simulations, the frameworks were treated as rigid with periodic boundary conditions applied to all three dimensions. The LJ interactions were calculated with a cutoff of 12 Å, while the electrostatic interactions were estimated using Ewald summation. The GCMC simulations were carried out using the RASPA package.[51] Each simulation included 10 000-cycle equilibration and 10 000-cycle production. Five different trial moves including insertion, deletion, rotation, translation and identity change were attempted randomly with equal probability.

3. Results and discussion

3.1. Analysis of PE-MOFs

We start with the geometric analysis of the PE-MOF database. Fig. 3 shows the distributions of three key geometric features (GSA, VF and LCD) in 94 823 PE-MOFs and also in 279 010 ARC-MOFs for comparison. The ARC-MOF database is a benchmark collection of both experimentally synthesized and computationally predicted MOF deposits.[14] For each geometric feature, a similar 'volcano' pattern is observed in both databases; however, PE-MOFs exhibit a wider distribution. Specifically, GSA ranges from zero to 10 000 $m^2\ g^{-1}$ in PE-MOFs and from zero to 8000 $m^2\ g^{-1}$ in ARC-MOFs. The VF is between 0 and 1 in PE-MOFs and between 0 and 0.9 in ARC-MOFs. The LCD is from 2 Å to 80 Å in PE-MOFs and from 2 Å to 35 Å in ARC-MOFs. These results suggest that our fine-tuned RTA is capable of generating MOFs with a more variety of geometries. Fig. S8 and S9† provide a comparative analysis of topology distribution and its correlation with pore features between PE-MOFs and ARC-MOFs. Despite PE-MOFs being designed with a limited set of BUs, the enriched diversity in topologies significantly influences their porosity profiles, enhancing the variety of geometric features. Fig. S10† illustrates the *t*-SNE map with embeddings of geometric features for PE-MOFs and ARC-MOFs. Visually, PE-MOFs and ARC-MOFs are segregated into distinct regions. PE-MOFs primarily occupy the top-left quadrant of the map, suggesting that they are not extensively covered by ARC-MOFs but complement the unexplored geometric characteristics of ARC-MOFs.

Chemical analysis of PE-MOFs is based on the RACs[29] by characterizing three crucial aspects of MOF chemistry: metal, ligand, and functional group. As depicted in Fig. 4a, unlike the material space learned from geometric features (Fig. S10†), both PE-MOFs and ARC-MOFs are spread over the entire *t*-SNE map, with ARC-MOFs exhibiting a wider space. The difference is attributed to a greater structural freedom in ARC-MOFs compared with PE-MOFs, as the latter are constructed with precise or strict topological criteria. Table 1 lists the diversity metrics including variety, disparity and balance in PE-MOFs and ARC-MOFs. The variety of a database indicates the

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 25 september 2024. Downloaded on 24.6.2025 20:57:50.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

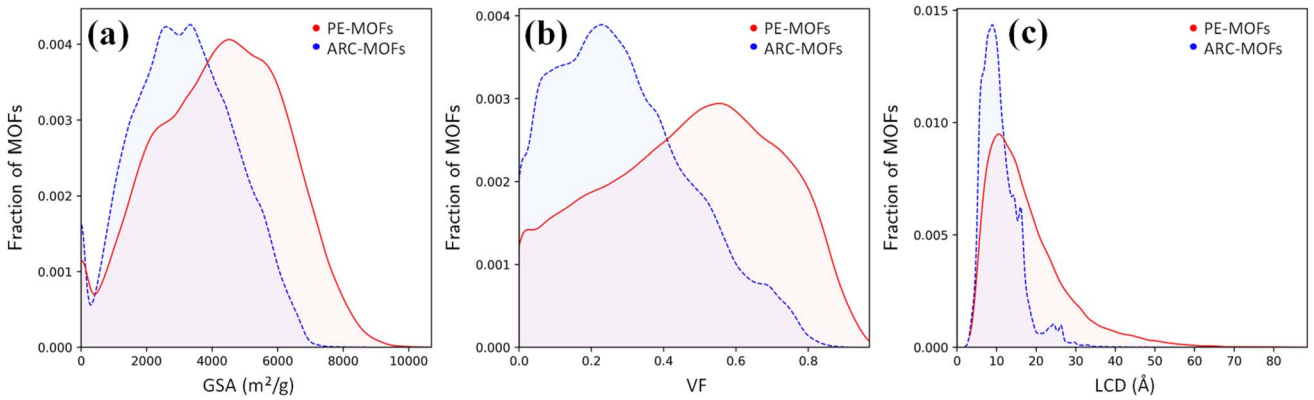


Fig. 3 Distributions of (a) GSA, (b) VF and (c) LCD in 94 823 PE-MOFs (red) and 279 010 ARC-MOFs (blue).[14]

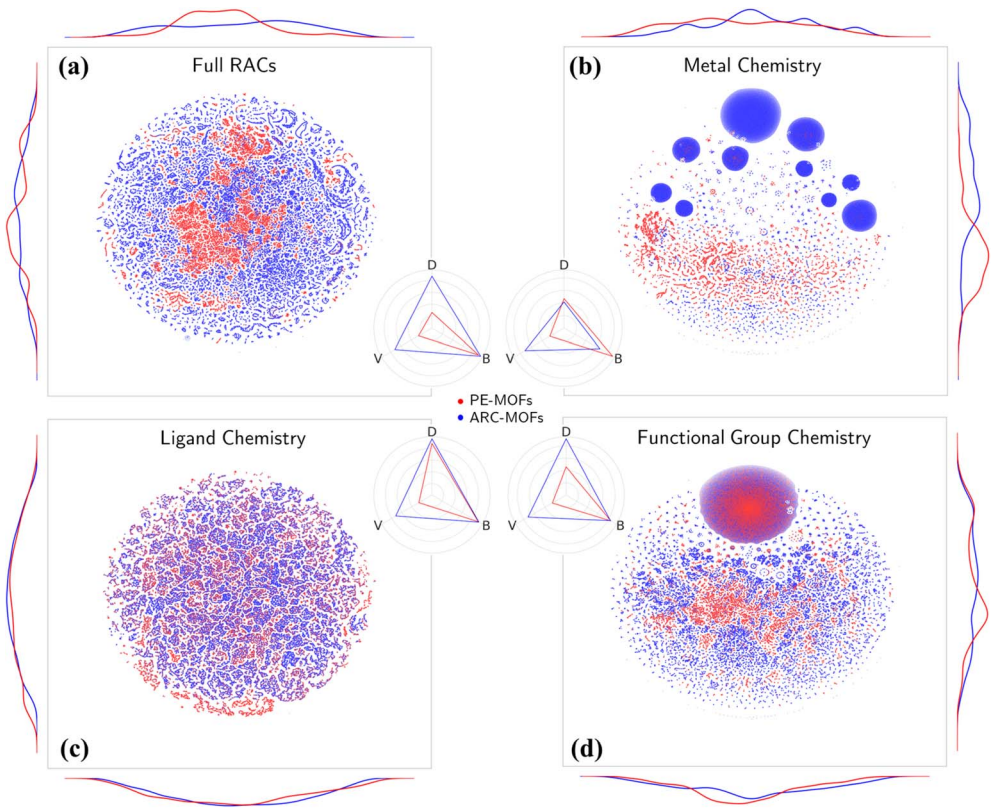


Fig. 4 *t*-SNE maps of multi-dimensional chemical features, (a) full RACs, (b) metal chemistry, (c) ligand chemistry and (d) functional group chemistry for 94 823 PE-MOFs (red) and 279 010 ARC-MOFs (blue). The curves along the axes are feature distributions. The radar charts display diversity metrics of variety (*V*), balance (*B*) and disparity (*D*) in PE-MOFs (red) and ARC-MOFs (blue).

number of distinct types of structures present; the balance reflects how evenly distributed the structures are; the disparity measures how dissimilar the structures are. Apparently, PE-MOFs possess a smaller chemical space than ARC-MOFs as evidenced by a lower variety value. This likely arises from the sparse diversity of the BU reservoir[14] exploited to construct PE-MOFs and a considerably more extensive ARC-MOF database, which is nearly three times the size of the PE-MOF database. Intriguingly, despite a limited population, a high balance value is observed in PE-MOFs (Fig. 4a–d), suggesting a well-distributed representation across different structural types. For metal chemistry and linker chemistry (Fig. 4b and c), both balance and disparity values of PE-MOFs are comparable with those of ARC-MOFs, indicating a dissimilarity-rich BU space of PE-MOFs. Although a greater diversity value in MOF banks is beneficial for computational screening and machine learning (ML),[32,52] our approach intentionally fine-tunes exploration in the design space of MOFs. Such PE is particularly important for future RTA-based structure design employing a much wider array of BUs and becomes crucial as diffusion models have

Open Access Article. Published on 25 september 2024. Downloaded on 24.6.2025 20:57:50.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Table 1 Diversity metrics in PE-MOFs and ARC-MOFs

Feature	Database	Variety	Balance	Disparity
Full RACs	PE-MOFs	0.261	0.902	0.261
	ARC-MOFs	0.720	0.941	0.871
Metal	PE-MOFs	0.261	0.895	0.470
	ARC-MOFs	0.720	0.661	0.416
Ligand	PE-MOFs	0.261	0.920	0.885
	ARC-MOFs	0.720	0.931	0.966
Functional group	PE-MOFs	0.261	0.845	0.459
	ARC-MOFs	0.720	0.837	0.918

received increasingly more attention in the inverse design of MOFs.[53–55]

3.2. CO_2 capture performance

For CO_2 capture, the performance of an adsorbent is usually quantified using three metrics: CO_2 working capacity (ΔN_{CO_2}), CO_2/N_2 selectivity S_{CO_2/N_2}, and a trade-off (TSN) between capacity and selectivity, TSN = $\log(S_{CO_2/N_2}) \times \Delta N_{CO_2}$. As illustrated in Fig. 5a–c, wide distributions are observed for ΔN_{CO_2}, $\log(S_{CO_2/N_2})$ and TSN in 17 173 PE-MOFs, specifically, in the range of 0–7 mol kg^{-1}, 0.5–3.0, and 0–16, respectively. For comparison, the performance of zeolite-13X is included, which is benchmarked as a high-performing adsorbent for CO_2 capture under dry conditions.[38] A large number of PE-MOFs exhibit performance superior to 13X. Fig. 5d further shows $\log(S_{CO_2/N_2})$ *versus* ΔN_{CO_2} with different TSN values. At a low ΔN_{CO_2}, a wide range of S_{CO_2/N_2} is observed approximately from 3 to 3000. With increasing ΔN_{CO_2}, S_{CO_2/N_2} generally drops and tends to approach a constant. PE-MOFs with greater TSN or superior performance are primarily clustered in the middle-right region. A total of 1207 PE-MOFs are found to surpass the TSN of 3.2 in 13X.

The highest TSN of 16.0 is observed in a PE-MOF assembled from metal vertice m160 and organic edge o323. With **bcu** topology and 8-connected metal vertices, as illustrated in Fig. 6a, this PE-MOF comprises three-dimensional pores with a diameter of 8.18 Å. Fig. 6b and c show the heats and isotherms of CO_2 and N_2 adsorption in this PE-MOF, calculated from GCMC simulations for a CO_2/N_2 mixture at 298 K and composition of 0.15/0.85. It is evident that CO_2 adsorption significantly dominates over N_2 across the entire pressure range, resulting in a CO_2 adsorption capacity of 8.28 mol kg^{-1} and CO_2/N_2 selectivity of 346 at 1 bar. Such high CO_2 capture performance is superior to or comparable with that of many experimentally synthesized and computationally designed MOFs reported in the literature, as compared in Table S2 and Fig. S11.†[9,15,52,56]

Furthermore, it is instructive to explore the interplay between CO_2 capture performance and structure characteristics through multivariate analysis. From the Pearson correlation matrix in Fig. 7a, capture performance metrics ΔN_{CO_2}, TSN and $\log(S_{CO_2/N_2})$ exhibit negative correlations with GSA, VF, PV, LCD, PLD and GCD. These correlations are corroborated by the relationships of ΔN_{CO_2} with the VF and LCD in 17 173 MOFs. As shown in Fig. 7b and c, with an increasing VF or LCD, ΔN_{CO_2} and TSN first increase, reach maximum, and then drop; this suggests while a small pore in a MOF leads to high S_{CO_2/N_2}, it restricts ΔN_{CO_2} and TSN. The top-performing MOFs with great TSN possess VF from 0.6 to 0.7 and LCD around 7 Å. These dimensions are about one to two times the kinetic diameter of

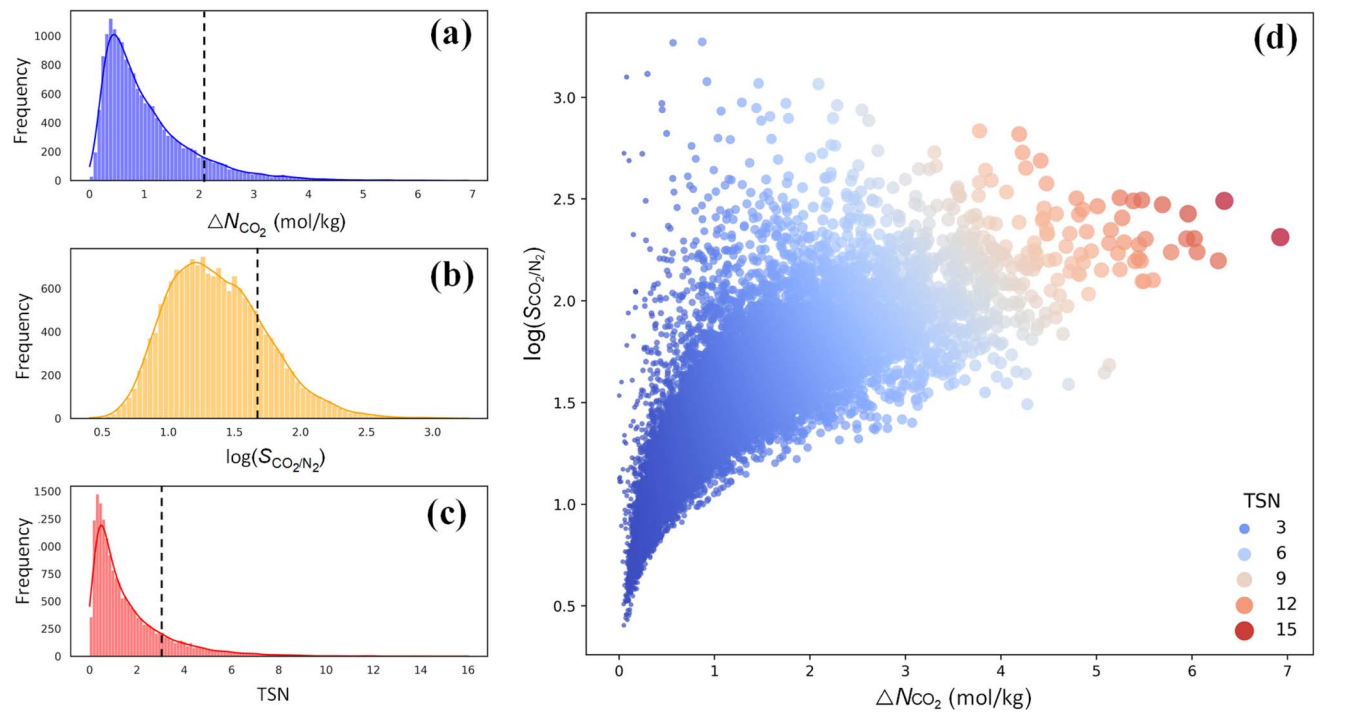


Fig. 5 Performance of 17 173 PE-MOFs for CO_2 capture. (a–c) Distributions of ΔN_{CO_2}, $\log(S_{CO_2/N_2})$ and TSN in 17 173 PE-MOFs. The dashed line denotes the performance of zeolite-13X as a benchmark.[38] (d) $\log(S_{CO_2/N_2})$ *versus* ΔN_{CO_2}, with the color-coding based on TSN.

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 25 september 2024. Downloaded on 24.6.2025 20:57:50.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

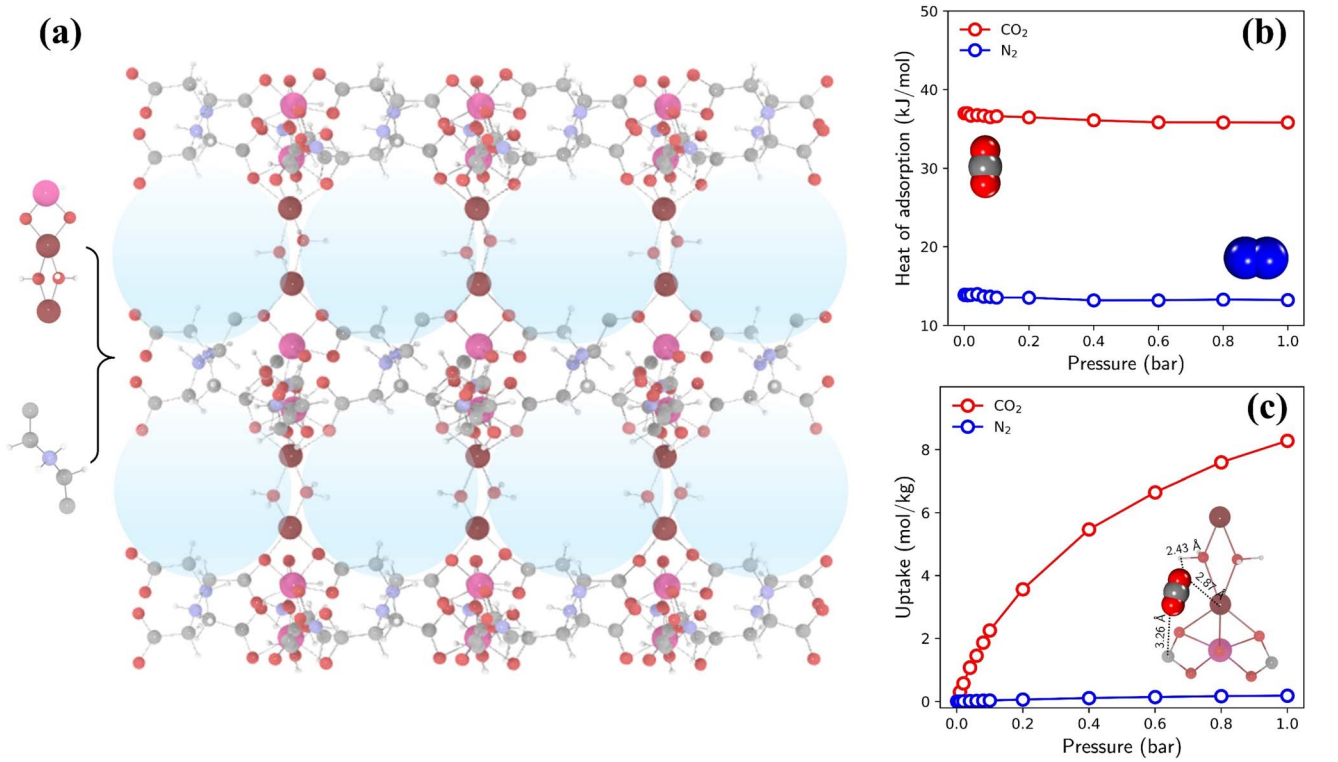


Fig. 6 (a) Structure of a top-performing MOF (assembled from metal vertice m160 and organic edge o323) for CO_2 capture, with blue spheres indicating the largest cavities. (b) Heats of CO_2 and N_2 adsorption. (c) Isotherms of CO_2 and N_2 adsorption, with a simulation snapshot highlighting the preferential adsorption site of CO_2.

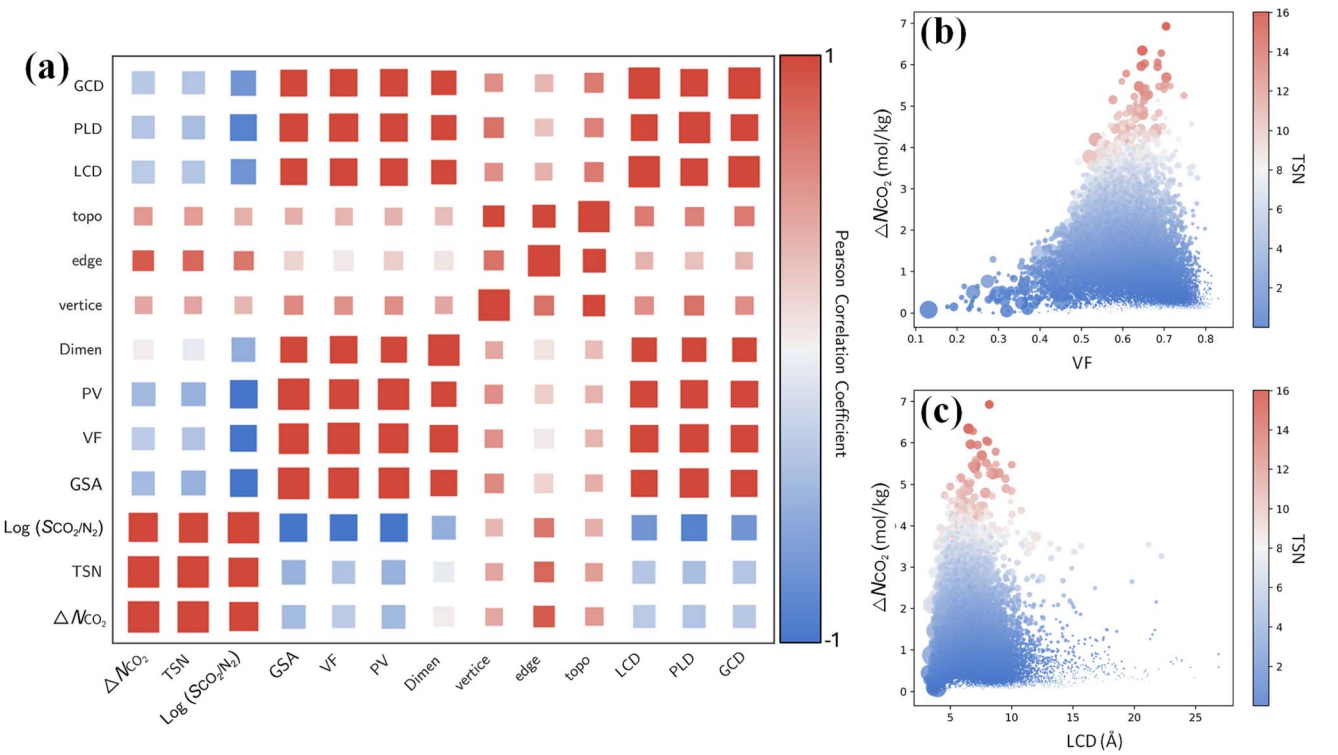


Fig. 7 (a) Pearson correlation matrix (PCM) between CO_2 capture performance metrics (ΔN_{CO_2}, TSN and $\log(S_{CO_2/N_2})$) and structural characteristics (GSA, VF, PV, dimensionality, vertice, edge, topology, LCD, PLD and GCD). (b and c) Relationships of ΔN_{CO_2} with the VF and LCD in 17 173 MOFs, with color-coding based on TSN.

Open Access Article. Published on 25 september 2024. Downloaded on 24.6.2025 20:57:50.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

CO_2 molecule, but less than twice that of N_2 molecule, as also previously reported.[57]

By contrast, ΔN_{CO_2}, TSN and $\log(S_{CO_2/N_2})$ show positive correlations with vertices, edges and topology as shown in Fig. 7a. This underscores the significant role of structural configuration in determining MOF performance. However, the impact may be overshadowed by the transformation of categorical variables into numerical values, which warrants deeper examination. As illustrated in Fig. 8a, MOFs with O6 and C8 vertices, fostering 3D geometries, exhibit a broad distribution of great TSN values (>8). This correlation aligns well with the finding in Fig. 8b, where 3D-MOFs with **pcu** and **bcu** topologies, characterized by larger void fractions, are densely represented among great TSN values. Regarding edges, as shown in Fig. 8c, our analysis reveals that high-performing MOFs incorporate L2 edge, which likely contributes to the formation of uniform pores; consequently, the pore space is optimized to accommodate CO_2 molecules and enables them to align and stack efficiently.

To quantify the impact of chemical descriptors on CO_2 capture, we utilized two separate Random Forest (RF) models[58] and combined with their SHapley Additive exPlanations[59] (SHAP) feature importance. The first model incorporates pore features along with bond types of reticulated BUs, and the second model focuses exclusively on bond types. The details of ML are provided in the ESI.† As depicted in Fig. S12,† the first model reveals that PV holds the greatest importance for CO_2 capture. The second model shows that the presence and

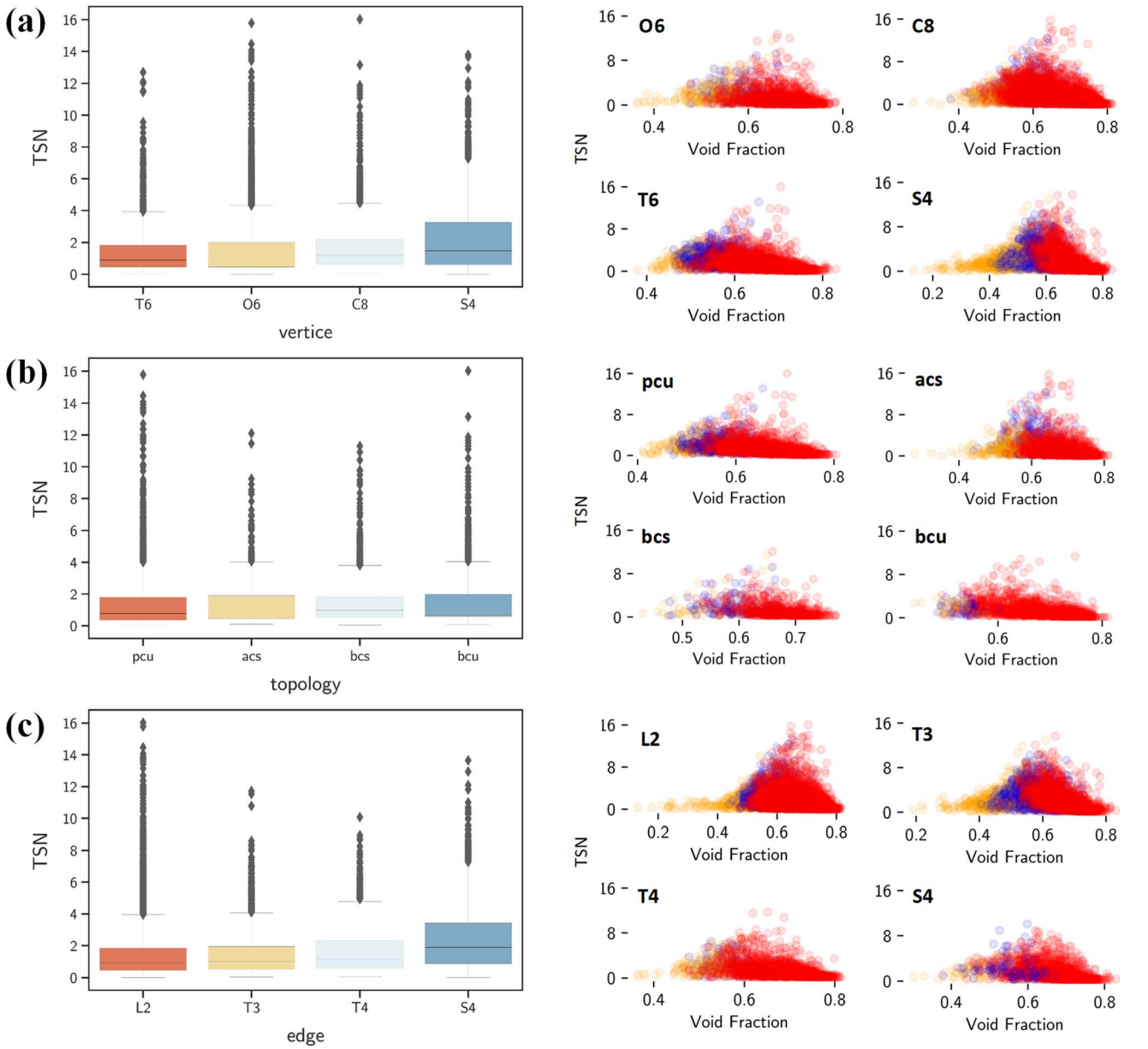


Fig. 8 (Left) Relationships of TSN with (a) vertice, (b) topology and (c) edge. (Right) Relationships of TSN with the void fraction, and the color code indicates MOF dimensionality: 1D in orange, 2D in blue and 3D in red.

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 25 september 2024. Downloaded on 24.6.2025 20:57:50.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

arrangement of carbon bonds within BUs—specifically C–H and C–C bonds—significantly affect the TSN value (Fig. 9a). These bonds typically lead to longer linkers, thereby enhancing porosity of MOFs.[60,61] It is essential to note that the outcome of the feature importance analysis might differ depending on the dataset utilized for model training;[32] thus, our conclusion is specific to our dataset.

Fig. 9b–d show the relationships between TSN and three key MOF characteristics (metal, dimensionality and linker type). MOFs with metals such as Li and Mn typically possess a broader range of higher TSN values compared with others. However,

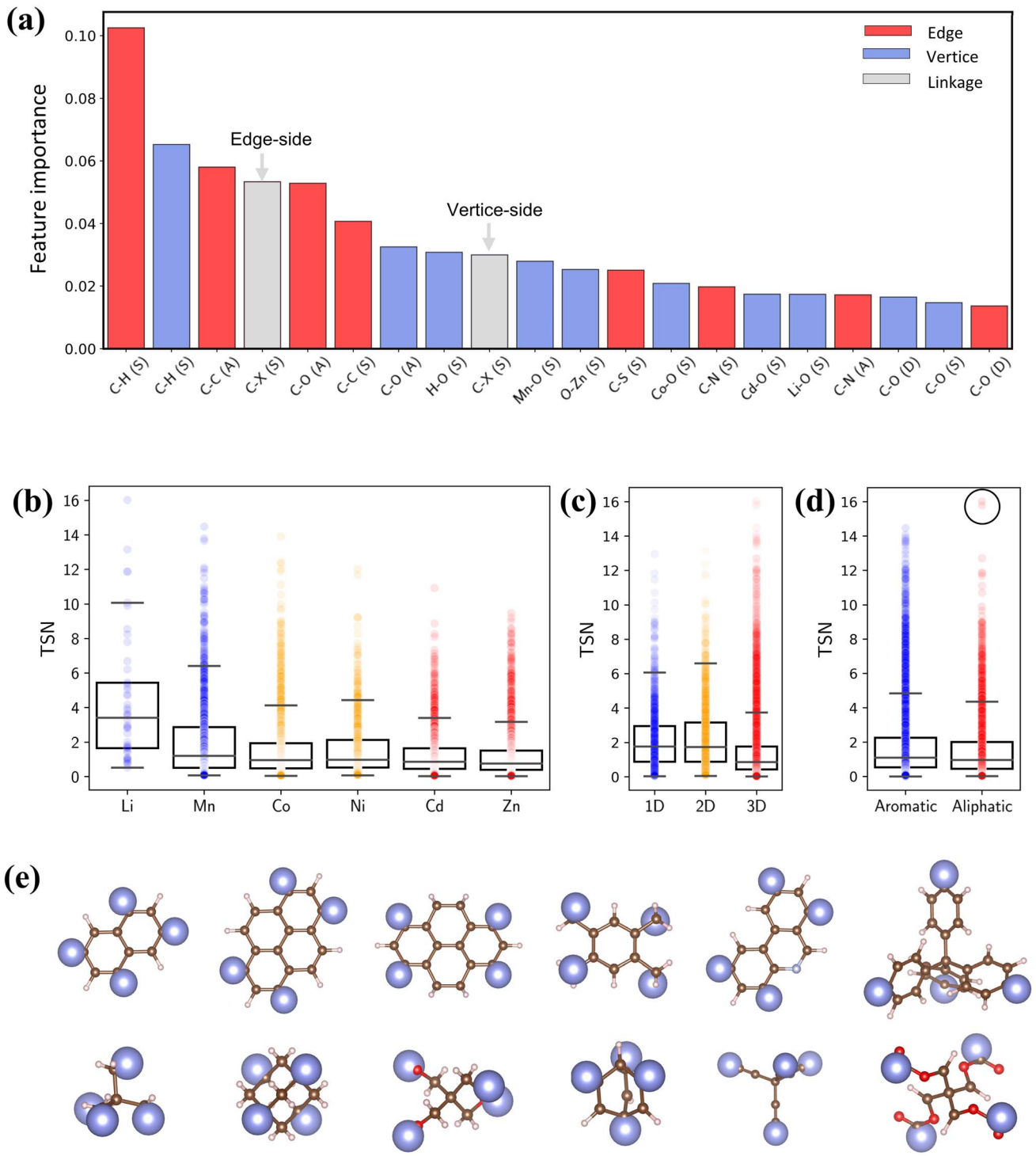


Fig. 9 (a) SHAP interpreted feature importance of bond features. (b–d) Box plots of metal, dimensionality and linker type *versus* TSN. (e) Representative aromatic and aliphatic linkers in PE-MOFs.

Open Access Article. Published on 25 september 2024. Downloaded on 24.6.2025 20:57:50.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

metals in MOFs can form as clusters or complexes and associate with various BUs.[17] The complex configurations of these metal sites are crucial in mediating interactions (*e.g.*, dipole–dipole interactions). Regarding dimensionality, 1D and 2D MOFs, with constricted porosity as discussed previously, typically exhibit moderate TSN values due to their limited cavity space, which constrains deliverable capacity due to high surface affinity at a low pressure. In contrast, 3D MOFs, characterized by larger porosity, show significantly higher TSN values. Additionally, the edge BUs in our dataset are categorized into either aromatic or aliphatic. Aromatic linkers, with their uniform and planar geometry compared with aliphatic BUs (Fig. 9e), generally tend to exhibit higher TSN values, reflecting their capability to form MOFs with greater porosity and additional functional features. Notably, aliphatic linkers also achieve high performance in some cases (circled data points), highlighting that optimal CO_2 capture cannot be attributed to a single chemical descriptor alone.

3.3. Stability and applicability of PE-MOFs

Finally, we assess the stability and applicability of the designed PE-MOFs in real-world applications, particularly for CO_2 capture. Three critical stability metrics (activation, water and thermal)[39,40] were used to evaluate the stability of 1207 PE-MOFs that outperform zeolite-13X for CO_2 capture from a dry flue gas. As presented in Fig. 10a and b, the activation and water stability scores range from 0 to 1.0, categorizing PE-MOFs from unstable to stable. A threshold score of 0.5 is set to identify potentially activatable and water-stable structures. PE-MOFs with decomposition temperatures (T_d) exceeding 359 °C, which is the average T_d observed in the CoRE-MOF 2019 database,[11,62] are considered thermally stable (Fig. 10c). Among the 1207 PE-MOFs, 224 satisfy the three stability criteria. It is worthwhile to note that the ML models[39,40] facilitate rapid stability evaluation, to a certain extent, but they have limitations. For example, the models do not consider actual activation methods such as supercritical CO_2 drying,[63] which can stabilize MOFs prone to collapse by conventional activation techniques. For water stability, the models do not account for interactions under varied humidity conditions.[64] Although post-combustion CO_2 capture typically occurs near ambient temperature, the need for thermal activation of MOFs to expel trapped solvents necessitates higher thermal stability.[65]

The 224 stable PE-MOFs were further evaluated for their CO_2 capture from a wet flue gas ($CO_2/N_2/H_2O$ mixture at 298 K and 1 bar, with 60% relative humidity). As depicted in Fig. 10d, 34 PE-MOFs can be identified as hydrophobic (*i.e.*, $S_{CO_2/H_2O} > 1$) with S_{CO_2/N_2} not significantly compromised. A top-performing candidate, assembled from metal vertice m368 and organic edge o143, has exceptional performance with ΔN_{CO_2} of 4.62 mol kg^{-1} and S_{CO_2/N_2} of 121.7, even in the presence of H_2O. It also stands out in terms of stability metrics—an activation stability score of 0.95, a water stability score of 0.59, and a T_d of 427 °C. As illustrated in Fig. 10e, this MOF possesses a planar geometry of BUs (common among other top-performers) and constricted pore apertures (4.13 × 4.83 Å) introduced by **cdm** topology, thus

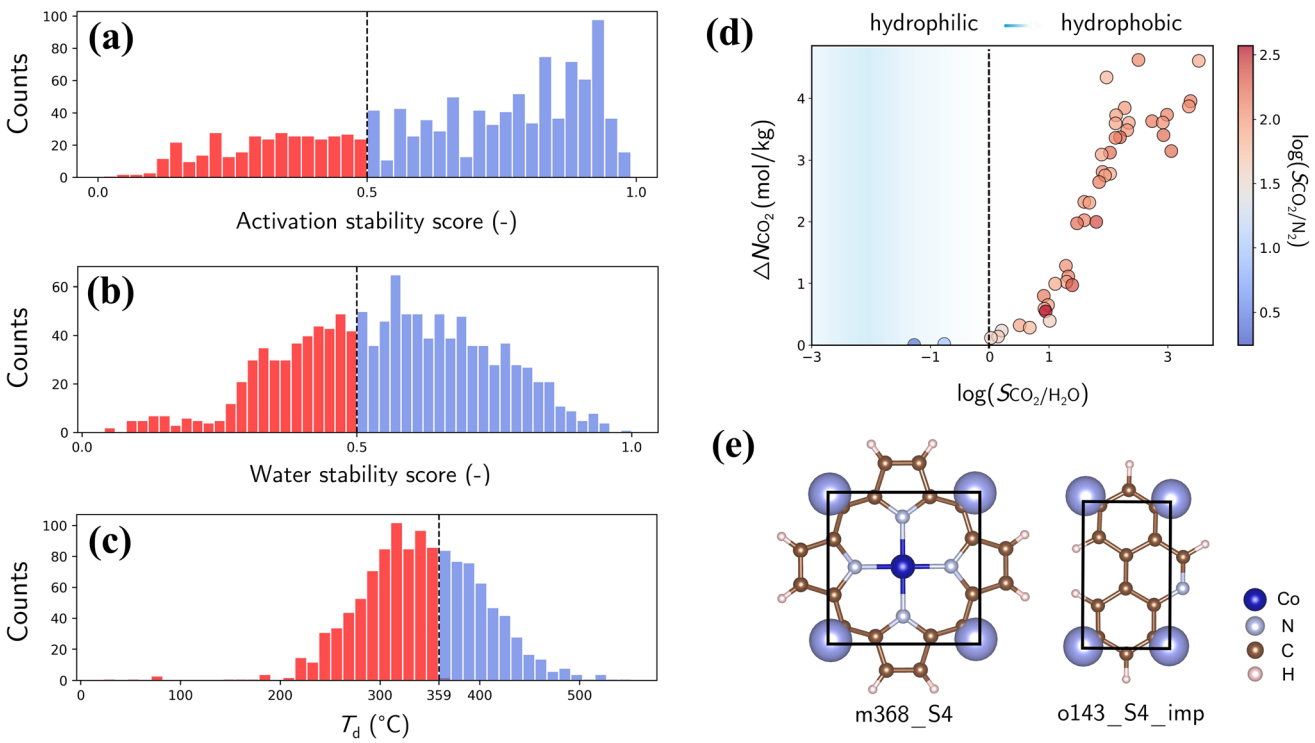


Fig. 10 (a) Activation scores of the refined PE-MOFs (classified as collapsible if score <0.5). (b) Water stability scores of the refined PE-MOFs (classified as water unstable if score <0.5). (c) T_d of the refined PE-MOFs (359 °C used as the threshold). (d) Relationship of ΔN_{CO_2} with $\log(S_{CO_2/H_2O})$, with color-coding based on $\log(S_{CO_2/N_2})$ (classified as hydrophobic if $S_{CO_2/H_2O} > 1$). (e) BUs of a top-performing MOF.

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 25 september 2024. Downloaded on 24.6.2025 20:57:50.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

enhancing chemical robustness while minimizing H_2O interference for CO_2 capture.

4. Conclusions

We have developed a fine-tuned RTA for structure prediction toward the digital discovery of MOFs. By addressing combinatorial explosion and effectively navigating a large array of BUs, our approach enhances the efficiency and precision to assemble BUs. A total of 94 823 PE-MOFs are designed with a more variety of geometries compared with ARC-MOFs despite a relatively smaller chemical space. For CO_2 capture, the performance of PE-MOFs exhibits negative correlations with several geometric features (*i.e.*, GSA, VF, PV, LCD, PLD and GCD), but positive correlations with structural configuration (*i.e.*, vertices, edges and topology). Further quantitative insights are probed with machine learning models, highlighting the indispensable role of pore volume and the configuration of carbon bonds in determining CO_2 capture performance. A larger number of PE-MOFs are found to be superior to benchmarked zeolite-13X. By integrating three stability metrics (activation, water and thermal), stable PE-MOFs are identified with high-performance for CO_2 capture from a wet flue gas.

The strategic assembly of BUs based on geometric signatures and topological compatibility not only mitigates combinatorial explosion by filtering out infeasible topologies, but also demonstrates the potential of digital chemistry in discovering high-performing materials. Our approach highlights the transformative impact of synergizing precision engineering with digital reticular chemistry, and it would advance further innovation in the intelligent design of porous crystals beyond MOFs. In future research, we will focus on expanding the diversity of BUs, particularly by customizing property-triggered BUs and integrating mechanical-, water- and thermal-stability, and other important stability measures. Advanced machine learning techniques will be incorporated to streamline the design process and enhance practical applicability.

Data availability

Data and codes related to this study are available on GitHub **https://github.com/xiaoyu961031/Fine-tuned-RTA**. All computationally predicted structures are available on Zenodo **https://zenodo.org/records/11480898**.

Author contributions

XW conceptualized the project, developed the approach and wrote the manuscript. All authors reviewed and edited the manuscript. JJ acquired the funding and supervised the project.

Conflicts of interest

There is no conflict of interests to declare.

Acknowledgements

We gratefully acknowledge the A*STAR LCER-FI projects (LCERFI01-0015 U2102d2004 and LCERFI01-0033 U2102d2006) and the National Research Foundation Singapore (NRF-CRP26-2021RS-0002) for financial support and the National University of Singapore and the National Supercomputing Centre (NSCC) Singapore for computational resources.

References

1 H. Li, M. Eddaoudi, M. O'Keeffe and O. M. Yaghi, Design and synthesis of an exceptionally stable and highly porous metal-organic framework, *Nature*, 1999, **402**, 276–279.
2 H. Furukawa, K. E. Cordova, M. O'Keeffe and O. M. Yaghi, The chemistry and applications of metal-organic frameworks, *Science*, 2013, **341**, 1230444.
3 H. Lyu, Z. Ji, S. Wuttke and O. M. Yaghi, Digital Reticular Chemistry, *Chem*, 2020, **6**, 2219–2241.
4 C. M. Draznieks, J. M. Newsam, A. M. Gorman, C. M. Freeman and G. Férey, De Novo Prediction of Inorganic Structures Developed through Automated Assembly of Secondary Building Units (AASBU Method), *Angew. Chem., Int. Ed.*, 2000, **39**, 2270–2275.
5 J. P. Darby, M. Arhangelskis, A. D. Katsenis, J. M. Marrett, T. Friščić and A. J. Morris, Ab Initio Prediction of Metal-Organic Framework Structures, *Chem. Mater.*, 2020, **32**, 5835–5844.
6 Y. Xu, J. M. Marrett, H. M. Titi, J. P. Darby, A. J. Morris, T. Friščić and M. Arhangelskis, Experimentally Validated Ab Initio Crystal Structure Prediction of Novel Metal–Organic Framework Materials, *J. Am. Chem. Soc.*, 2023, 3515–3525.
7 C. E. Wilmer, M. Leaf, C. Y. Lee, O. K. Farha, B. G. Hauser, J. T. Hupp, R. Q. Snurr, C. E. Wilmer, M. Leaf, C. Y. Lee, *et al.*, Large-scale screening of hypothetical metal–organic frameworks, *Nat. Chem.*, 2011, **4**, 83–89.
8 J. Keupp, J. Keupp, R. Schmid and R. Schmid, TopoFF: MOF structure prediction using specifically optimized blueprints, *Faraday Discuss.*, 2018, **211**, 79–101.
9 P. G. Boyd, A. Chidambaram, E. García-Díez, C. P. Ireland, T. D. Daff, R. Bounds, A. Gładysiak, P. Schouwink, S. M. Moosavi, M. M. Maroto-Valer, *et al.*, Data-driven design of metal–organic frameworks for wet flue gas CO_2 capture, *Nature*, 2019, **576**, 253–256.
10 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.
11 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.
12 J. Jiang, Computational screening of metal–organic frameworks for CO_2 separation, *Curr. Opin. Green Sustainable Chem.*, 2019, **16**, 57–64.

Open Access Article. Published on 25 september 2024. Downloaded on 24.6.2025 20:57:50. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

13 A. S. Rosen, S. M. Iyer, D. Ray, Z. Yao, A. Aspuru-Guzik, L. Gagliardi, J. M. Notestein and R. Q. Snurr, Machine learning the quantum-chemical properties of metal–organic frameworks for accelerated materials discovery, *Matter*, 2021, **4**, 1578–1597.

14 J. Burner, J. Luo, A. White, A. Mirmiran, O. Kwon, P. G. Boyd, S. Maley, M. Gibaldi, S. Simrod, V. Ogden and T. K. Woo, 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.

15 S. A. Mohamed, D. Zhao and J. Jiang, Integrating stability metrics with high-throughput computational screening of metal–organic frameworks for CO_2 capture, *Commun. Mater.*, 2023, **4**, 1–10.

16 O. M. Yaghi, Reticular Chemistry in All Dimensions, *ACS Cent. Sci.*, 2019, **5**, 1295–1300.

17 M. J. Kalmutzki, N. Hanikel and O. M. Yaghi, Secondary building units as the turning point in the development of the reticular chemistry of MOFs, *Sci. Adv.*, 2018, **4**, eaat9180.

18 M. Gibaldi, O. Kwon, A. White, J. Burner and T. K. Woo, The HEALED SBU Library of Chemically Realistic Building Blocks for Construction of Hypothetical Metal-Organic Frameworks, *ACS Appl. Mater. Interfaces*, 2022, **14**, 43372–43386.

19 K. M. Jablonka, A. S. Rosen, A. S. Krishnapriyan and B. Smit, An Ecosystem for Digital Reticular Chemistry, *ACS Cent. Sci.*, 2023, **9**, 563–581.

20 N. M. Padial, E. Q. Procopio, C. Montoro, E. López, J. E. Oltra, V. Colombo, A. Maspero, N. Masciocchi, S. Galli, I. Senkovska, *et al.*, Highly Hydrophobic Isoreticular Porous Metal–Organic Frameworks for the Capture of Harmful Volatile Organic Compounds, *Angew. Chem.*, 2013, **125**, 8448–8452.

21 D. A. Gómez-Gualdrón, Y. J. Colón, X. Zhang, T. C. Wang, Y.-S. Chen, J. T. Hupp, T. Yildirim, O. K. Farha, J. Zhang and R. Q. Snurr, Evaluating topologically diverse metal–organic frameworks for cryo-adsorbed hydrogen storage, *Energy Environ. Sci.*, 2016, **9**, 3279–3289.

22 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.

23 S. S. Y. Chui, S. M. F. Lo, J. P. H. Charmant, A. G. Orpen and I. D. Williams, A chemically functionalizable nanoporous material $[Cu_3(TMA)_2(H_2O)_3]_{(n)}$, *Science*, 1999, **283**, 1148–1150.

24 M. O'Keeffe, M. A. Peskov, S. J. Ramsden and O. M. Yaghi, The Reticular Chemistry Structure Resource (RCSR) database of, and symbols for, crystal nets, *Acc. Chem. Res.*, 2008, **41**, 1782–1789.

25 R. Mercado, R. S. Fu, A. V. Yakutovich, L. Talirz, M. Haranczyk and B. Smit, In Silico Design of 2D and 3D Covalent Organic Frameworks for Methane Storage Applications, *Chem. Mater.*, 2018, **30**, 5069–5086.

26 A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard 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.

27 S. Kancharlapalli, A. Gopalan, M. Haranczyk and R. Q. Snurr, Fast and Accurate Machine Learning Strategy for Calculating Partial Atomic Charges in Metal-Organic Frameworks, *J. Chem. Theory Comput.*, 2021, **17**, 3052–3064.

28 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.

29 J. P. Janet and H. J. Kulik, Resolving Transition Metal Chemical Space: Feature Selection for Machine Learning and Structure-Property Relationships, *J. Phys. Chem. A*, 2017, **121**, 8939–8954.

30 A. Nandy, C. Duan, J. P. Janet, S. Gugler and H. J. Kulik, Strategies and Software for Machine Learning Accelerated Discovery in Transition Metal Chemistry, *Ind. Eng. Chem. Res.*, 2018, **57**, 13973–13986.

31 A. Nandy, C. Duan, M. G. Taylor, F. Liu, A. H. Steeves and H. J. Kulik, Computational Discovery of Transition-metal Complexes: From High-throughput Screening to Machine Learning, *Chem. Rev.*, 2021, **121**, 9927–10000.

32 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**, 1–10.

33 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.

34 H. Tang, Q. Xu, M. Wang and J. Jiang, Rapid screening of metal–organic frameworks for propane/propylene separation by synergizing molecular simulation and machine learning, *ACS Appl. Mater. Interfaces*, 2021, **13**, 53454–53467.

35 Z. Zhang, H. Tang, M. Wang, B. Lyu, Z. Jiang and J. Jiang, Metal–Organic Frameworks for Water Harvesting: Machine Learning-Based Prediction and Rapid Screening, *ACS Sustain. Chem. Eng.*, 2023, **11**, 8148–8160.

36 E. I. Ioannidis, T. Z. H. Gani and H. J. Kulik, molSimplify: a toolkit for automating discovery in inorganic chemistry, *J. Comput. Chem.*, 2016, **37**, 2106–2117.

37 L. Van Der Maaten and G. Hinton, Visualizing Data using t-SNE, *J. Mach. Learn. Res.*, 2008, **9**, 2579–2605.

38 M. T. Ho, G. W. Allinson and D. E. Wiley, Reducing the Cost of CO_2 Capture from Flue Gases Using Pressure Swing Adsorption, *Ind. Eng. Chem. Res.*, 2008, **47**, 4883–4890.

39 A. Nandy, G. Terrones, N. Arunachalam, C. Duan, D. W. Kastner and H. J. Kulik, MOFSimplify, machine learning models with extracted stability data of three thousand metal–organic frameworks, *Sci. Data*, 2022, **9**, 74.

40 G. G. Terrones, S.-P. Huang, M. P. Rivera, S. Yue, A. Hernandez and H. J. Kulik, Metal–Organic Framework Stability in Water and Harsh Environments from Data-Driven Models Trained on the Diverse WS24 Data Set, *J. Am. Chem. Soc.*, 2024, **146**, 20333–20348.

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 25 september 2024. Downloaded on 24.6.2025 20:57:50.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

41 J. J. Potoff and J. I. Siepmann, Vapor–liquid equilibria of mixtures containing alkanes, carbon dioxide, and nitrogen, *AIChE J.*, 2001, **47**, 1676–1682.
42 J. Vorholz, V. I. Harismiadis, B. Rumpf, A. Z. Panagiotopoulos and G. Maurer, Vapor+liquid equilibrium of water, carbon dioxide, and the binary system, water+carbon dioxide, from molecular simulation, *Fluid Phase Equilib.*, 2000, **170**, 203–234.
43 S. L. Mayo, B. D. Olafson and W. A. Goddard, DREIDING: a generic force field for molecular simulations, *J. Phys. Chem.*, 1990, **94**, 8897–8909.
44 T. A. Manz and D. S. Sholl, Chemically Meaningful Atomic Charges That Reproduce the Electrostatic Potential in Periodic and Nonperiodic Materials, *J. Chem. Theory Comput.*, 2010, **6**, 2455–2468.
45 N. G. Limas and T. A. Manz, Introducing DDEC6 atomic population analysis: part 4. Efficient parallel computation of net atomic charges, atomic spin moments, bond orders, and more, *RSC Adv.*, 2018, **8**, 2678–2707.
46 G. Kresse, Ab initio molecular dynamics for liquid metals, *J. Non-Cryst. Solids*, 1995, **192–193**, 222–229.
47 G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**, 11169.
48 J. P. Perdew, K. Burke and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.*, 1996, **77**, 3865.
49 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, 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.
50 S. Grimme, S. Ehrlich and L. Goerigk, Effect of the damping function in dispersion corrected density functional theory, *J. Comput. Chem.*, 2011, **32**, 1456–1465.
51 D. Dubbeldam, S. Calero, D. E. Ellis and R. Q. Snurr, RASPA: molecular simulation software for adsorption and diffusion in flexible nanoporous materials, *Mol. Simul.*, 2016, **42**, 81–101.
52 S. Majumdar, S. M. Moosavi, K. M. Jablonka, D. Ongari and B. Smit, Diversifying Databases of Metal Organic Frameworks for High-Throughput Computational Screening, *ACS Appl. Mater. Interfaces*, 2021, **13**, 61004–61014.
53 X. Fu, T. Xie, A. S. Rosen, T. Jaakkola and J. Smith, MOFDiff: Coarse-grained Diffusion for Metal-Organic Framework Design, *arXiv*, 2023, preprint, arXiv:2310.10732, DOI: 10.48550/arXiv.2310.10732.
54 H. Xiao, R. Li, X. Shi, Y. Chen, L. Zhu, X. Chen and L. Wang, An invertible, invariant crystal representation for inverse design of solid-state materials using generative deep learning, *Nat. Commun.*, 2023, **14**, 1–12.
55 H. Park, X. Yan, R. Zhu, E. A. Huerta, S. Chaudhuri, D. Cooper, I. Foster and E. Tajkhorshid, A generative artificial intelligence framework based on a molecular diffusion model for the design of metal-organic frameworks for carbon capture, *Commun. Chem.*, 2024, **7**, 1–18.
56 J. B. Lin, T. T. T. Nguyen, R. Vaidhyanathan, J. Burner, J. M. Taylor, H. Durekova, F. Akhtar, R. K. Mah, O. Ghaffari-Nik, S. Marx, *et al.*, A scalable metal-organic framework as a durable physisorbent for carbon dioxide capture, *Science*, 2021, **374**, 1464–1469.
57 X. Cao, Z. Wang, Z. Qiao, S. Zhao and J. Wang, Penetrated COF Channels: Amino Environment and Suitable Size for CO_2 Preferential Adsorption and Transport in Mixed Matrix Membranes, *ACS Appl. Mater. Interfaces*, 2019, **11**, 5306–5315.
58 T. K. Ho, *Random decision forests*, IEEE, Montréal, Canada, 1995.
59 S. M. Lundberg and S.-I. Lee, A unified approach to interpreting model predictions, in *NIPS'17*, Long Beach, California, USA, 2017.
60 T. Bailey, A. Jackson, R.-A. Berbece, K. Wu, N. Hondow and E. Martin, Gradient Boosted Machine Learning Model to Predict H_2, CH_4, and CO_2 Uptake in Metal–Organic Frameworks Using Experimental Data, *J. Chem. Inf. Model.*, 2023, **63**, 4545–4551.
61 R. Wang, Q. Meng, L. Zhang, H. Wang, F. Dai, W. Guo, L. Zhao and D. Sun, Investigation of the effect of pore size on gas uptake in two fsc metal–organic frameworks, *Chem. Commun.*, 2014, **50**, 4911–4914.
62 Y. G. Chung, E. Haldoupis, B. J. Bucior, M. Haranczyk, S. Lee, H. Zhang, K. D. Vogiatzis, M. Milisavljevic, S. Ling, J. S. Camp, *et al.*, Advances, Updates, and Analytics for the Computation-Ready, Experimental Metal–Organic Framework Database: CoRE MOF 2019, *J. Chem. Eng. Data*, 2019, **64**, 5985–5998.
63 D. M. Frimpong and L. Baojian, Recent Advances of Supercritical CO_2 in Green Synthesis and Activation of Metal–Organic Frameworks, *J. Inorg. Organomet. Polym. Mater.*, 2019, **30**, 581–595.
64 N. C. Burtch, H. Jasuja and K. S. Walton, Water Stability and Adsorption in Metal–Organic Frameworks, *Chem. Rev.*, 2014, **114**, 10575–10612.
65 J. L. Woodliffe, R. S. Ferrari, I. Ahmed and A. Laybourn, Evaluating the purification and activation of metal-organic frameworks from a technical and circular economy perspective, *Coord. Chem. Rev.*, 2021, **428**, 213578.