




<https://doi.org/10.1038/s43246-023-00409-9>

OPEN

Integrating stability metrics with high-throughput computational screening of metal-organic frameworks for CO₂ capture

Saad Aldin Mohamed ¹, Daohui Zhao¹ & Jianwen Jiang ¹✉

Metal-organic frameworks (MOFs) have been considered a unique class of hybrid materials for a wide variety of potential applications. With the existence of almost infinite MOFs, high-throughput computational screening (HTCS) is a robust technique to accelerate the search for promising MOFs. However, conventional HTCS studies reported in the literature neglect the stability of MOFs, which must be considered for practical applications. Here we integrate four stability metrics (thermodynamic, mechanical, thermal, and activation) with HTCS to identify top-performing, synthesizable, and stable hypothetical MOFs for CO₂ capture. The thermodynamic and mechanical stabilities are evaluated through molecular dynamics simulations, while the activation and thermal stabilities are predicted using machine learning models. Finally, we identify top-performing hypothetical MOFs satisfying all these stability metrics. This study underlines the central importance of integrating stability metrics when screening MOFs for applications.

¹Department of Chemical and Biomolecular Engineering, National University of Singapore 117576, Singapore. ✉email: chejj@nus.edu.sg

With readily tunable structures and functionalities, metal–organic frameworks (MOFs) have received considerable attention for many potential applications. However, it is technically challenging to select promising MOFs as the chemical space is tremendously large. To date, over 100,000 MOFs have been experimentally synthesized, and in principle, infinite MOFs can be constructed by varying building blocks and topologies¹. Experimentally identifying suitable candidates from the enormous number of MOFs for applications is economically infeasible. In this context, high-throughput computational screening (HTCS) is a powerful technique to assist in establishing quantitative structure–property relationships and identifying top MOFs.

We have witnessed great success in using HTCS to identify top MOFs for different applications^{2–4}, particularly CO₂ capture⁵. For instance, the Computationally-Ready Experimental (CoRE) MOFs⁶ were screened for post-combustion CO₂ capture and natural gas purification⁷. The separation of hexane and heptane isomers was simulated in CoRE MOFs, then one candidate was synthesized and characterized⁸. The latest version of the CoRE MOF database (2019)⁹ with over 10,000 structures was screened to isolate unique MOFs for paraffin/olefin separation^{10,11}. The Cambridge Structural Database (CSD) MOFs¹² were also screened for CO₂ capture and H₂ purification, and a handful of MOFs were identified to possess selective adsorption and molecular sieving capabilities¹³. The identified experimental MOFs by HTCS are synthesizable, regardless of possible difficulties that might be encountered, but crucial stability measures are scarcely considered in HTCS, which might lead to the utilization of unstable MOFs for applications.

Meanwhile, hypothetical MOFs (hMOFs) have also been widely examined by HTCS. The adsorption of CO₂, N₂ and CH₄ in over 137,953 hMOFs was simulated, and the relationships between structural characteristics, as well as chemical functionality, with adsorbent evaluation criteria were proposed for CO₂ separation¹⁴. Based on the adsorption, diffusion and permeation of CO₂, N₂ and CH₄, 137,953 hMOFs were screened for a single-step membrane separation of CO₂/N₂/CH₄ mixture¹⁵. Among over 300,000 hMOFs, different strong CO₂-binding sites were classified, and hydrophobic MOFs with high CO₂/N₂ selectivity in wet flue gases were identified and further synthesized¹⁶. A database of ~20,000 hMOFs with diverse metal nodes, organic linkers, functional groups and pore geometries was designed and assessed for post-combustion CO₂ capture and H₂ storage¹⁷. However, most HTCS studies of hMOFs were not accompanied or followed by experimental validation, hence the identified hMOFs might not be synthesizable and stable. Therefore, the synthetic likelihood and stability metrics must be incorporated in HTCS, along with performance metrics, to identify possible candidates that are potentially synthesizable and stable.

In this study, for the first time, we integrate crucial stability metrics with HTCS to screen hMOFs for CO₂ capture. A collection of 15,219 hMOFs is obtained from the ARC-MOF database¹⁸, along with their computed adsorption data. As shown in Fig. 1, separation performance metrics are first used to shortlist top-performing hMOFs, and then four stability metrics (thermodynamic, mechanical, thermal, and activation) are integrated. Specifically, the thermodynamic stability (i.e., synthetic likelihood) is evaluated from free energies based on molecular dynamics (MD) simulations. The mechanical stability is calculated from elastic constants also obtained through MD simulations. Moreover, the activation and thermal stabilities are predicted using machine learning (ML) models. Finally, all the stability metrics are integrated to identify synthesizable, stable, and top-performing hMOFs for CO₂ capture.

Results and discussion

Separation performance. Different metrics can be used to evaluate the separation performance of a gas mixture in MOFs. Particularly, uptake and selectivity are widely used in HTCS studies, though there are other metrics such as working capacity¹⁹, purity, recovery, productivity²⁰, and general evaluation metric²¹. For CO₂ capture in this study, CO₂ uptake (mmol/g) and CO₂/N₂ selectivity were chosen. As plotted in Fig. 2a, 148 out of 15,219 hMOFs in DB5 subset exhibit CO₂ uptake ≥ 4 mmol/g and CO₂/N₂ selectivity ≥ 200 , which were set as the thresholds. It is interesting to examine the metal node chemistry of the hMOFs. Figure 2b shows the percentages of five different metal node types among 15,219 hMOFs and among the 148 top-performing hMOFs for CO₂ capture. Those with Zn₄O metal node represent 33.5% in the entire DB5 subset, but this percentage drops to 0.7% among the 148 top-performing hMOFs. By contrast, V₃O₃ metal node is minority in DB5 (12.4%) but prevailing in the top-performing hMOFs (46.6%). This metal node, which resembles a one-dimensional rod, is able to promote CO₂ adsorption as favorable adsorption sites exist near the bridging oxygen in V₃O₃¹⁶. Indeed, higher CO₂ uptake for V₃O₃ metal node is observed in Supplementary Fig. 1a, which shows the count of hMOFs for each metal node type versus CO₂ uptake. The percentages of pillared Zn- and Cu-paddlewheel metal nodes are found to not be much different between DB5 and the top-performing hMOFs. Lastly, a few hMOFs with Zr₆O₆ metal node exist in DB5 (0.1%) but are not among the top-performing MOFs. Figure 2c shows ranked CO₂ uptake in 148 top-performing hMOFs by different metal node types. There is no unique pattern to describe the distribution of metal nodes. This is because, aside from metal nodes, organic linkers also play an important role in CO₂ adsorption. Among 148 top-performing hMOFs, the highest CO₂ uptake (8.47 mmol/g) is found in a hMOF with V₃O₃ metal node. As shown in Fig. 2d, the notable CO₂ uptake in this hMOF can be attributed to the presence of strong adsorbaphore¹⁶, which originated from parallel aromatic rings separated by a distance of 6.8 Å, in addition to the V₃O₃ metal node.

Subsequently, the stabilities of 148 top-performing hMOFs are evaluated. We should note that the sequence of screening is not unique. One may first evaluate the stabilities, and for those deemed to be stable, then the separation performance can be determined. The order of these two screening steps depends on their computational cost. In the present study, the adsorption data are readily available in the ARC-MOF database and hence used at the first stage; consequently, there is no need to evaluate the stabilities of 15,219 hMOFs in the entire DB5 subset.

Thermodynamic stability. The thermodynamic stability is assessed by free energies (F) at ambient temperature, calculated using the FL method^{22–24}. Figure 3a shows the free energies versus the percentage of metal atoms in 148 hMOFs. To judge the synthetic likelihood, we need to set a criterion and label those crossing the criterion as thermodynamically unstable. As such, the free energies of experimental MOFs were calculated and used for benchmarking. For this purpose, 79 CoRE MOFs with four different metal node types: pillared Cu-paddlewheel, pillared Zn-paddlewheel, Zn₄O and V₃O₃ were considered. For each type, the CoRE MOFs were selected to have close similarity in metal nodes to the hMOFs using the revised autocorrelation functions (RACs)^{25,26}. The calculated free energies of the hMOFs and selected CoRE MOFs are plotted in Fig. 3b. For each metal node type, the free energies of CoRE MOFs are fitted against the percentage of metal atoms, acting as a reference line²⁴. By subtracting the corresponding value of the reference line (F_{LM}) from the true F , the relative free energies $\Delta_{LM}F$ are obtained. As

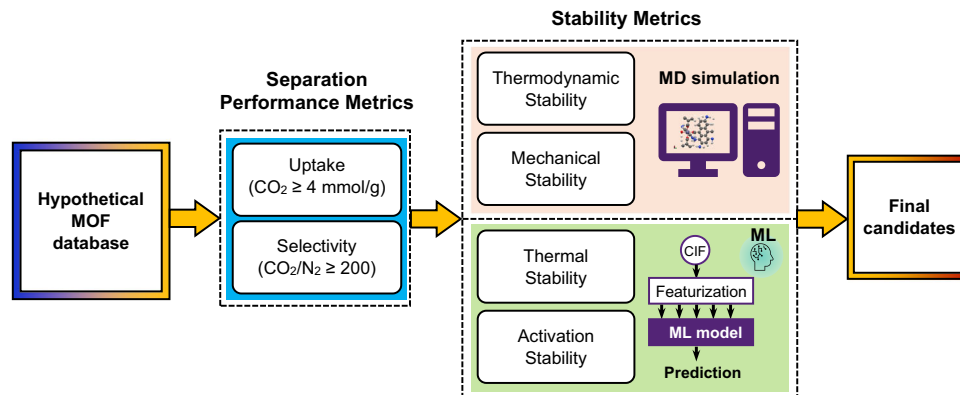


Fig. 1 High-throughput computational screening (HTCS) of hMOFs for CO₂ capture. in which stability metrics (thermodynamic, mechanical, thermal, and activation) are integrated with separation performance metrics.

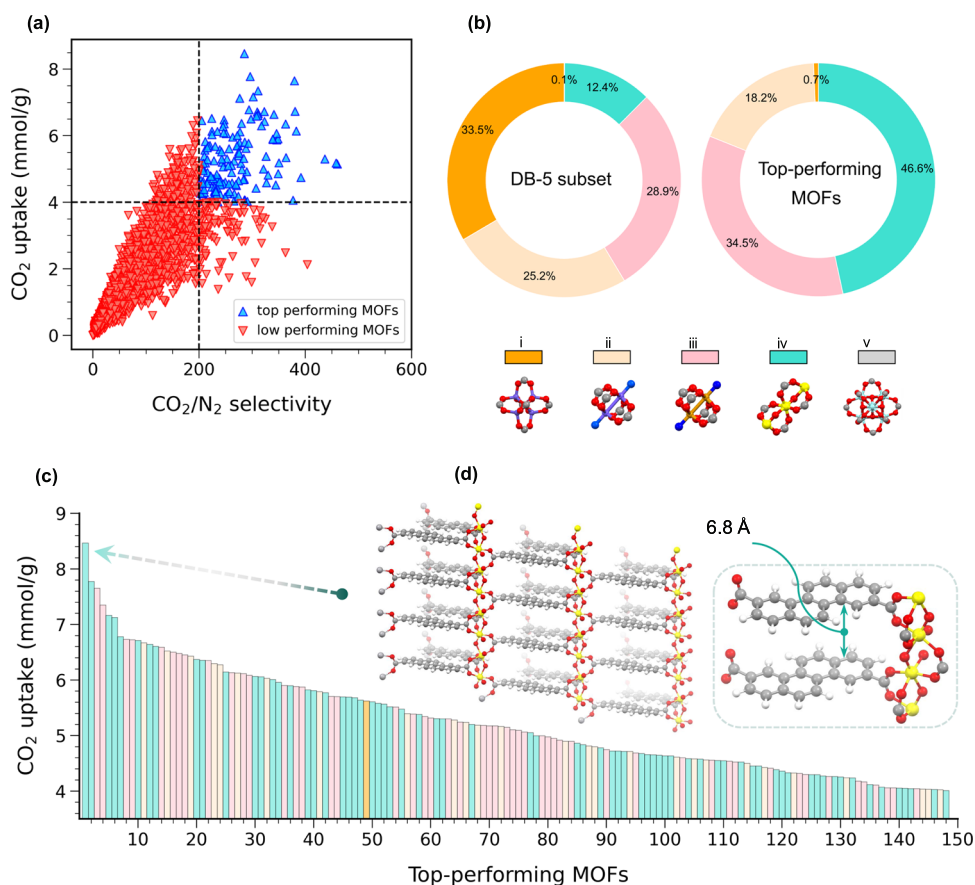


Fig. 2 Separation performance of hMOFs. **a** Top- and low-performing hMOFs for CO₂ capture at 298 K, 0.9 bar and feed composition of 17%/83% in a CO₂/N₂ mixture, the data were taken from the ARC-MOF database¹⁸. **b** Percentages of five different metal node types among 15,219 hMOFs in DB5 subset and among 148 top-performing hMOFs for CO₂ capture. Metal node labels: (i) Zn₄O, (ii) pillared Zn-paddlewheel, (iii) pillared Cu-paddlewheel, (iv) V₃O₃, and (v) Zr₆O₆. **c** Ranked CO₂ uptake in 148 top-performing hMOFs by different metal node types. **d** hMOF with the highest CO₂ uptake, showing V₃O₃ metal node and parallel organic linkers composed of aromatic rings (adsorbaphore).

presented in Fig. 3c, all the CoRE MOFs are capped within $\Delta_{LM}F$ of ~ 4.2 kJ/mol, which is then taken as an upper bound for thermodynamic stability. This value is close to 4.4 kJ/mol reported by Anderson et al.²⁴ Among 148 hMOFs, 41 have $\Delta_{LM}F$ higher than the upper bound, thus considered thermodynamically unstable and not recommended for synthesis. We should note, practically, that kinetic factors (e.g., solvent, modulator, pH, and temperature) may affect the synthesis process^{27,28}, which is not simply governed by thermodynamic stability.

Mechanical stability. MOFs should retain integrity during synthesis, activation, pelletization, and gas adsorption. Typically, post-combustion CO₂ capture is operated at low pressure (~ 1 bar), and hence no significant pressure on MOFs is expected to arise during CO₂ capture. Nevertheless, stress is applied during pelletization to shape MOFs, which in many cases leads to a significant loss of surface area, and such structural collapse can be mitigated by loading solvents into MOFs and stabilizing frameworks²⁹. Therefore, structural degradation by mechanical

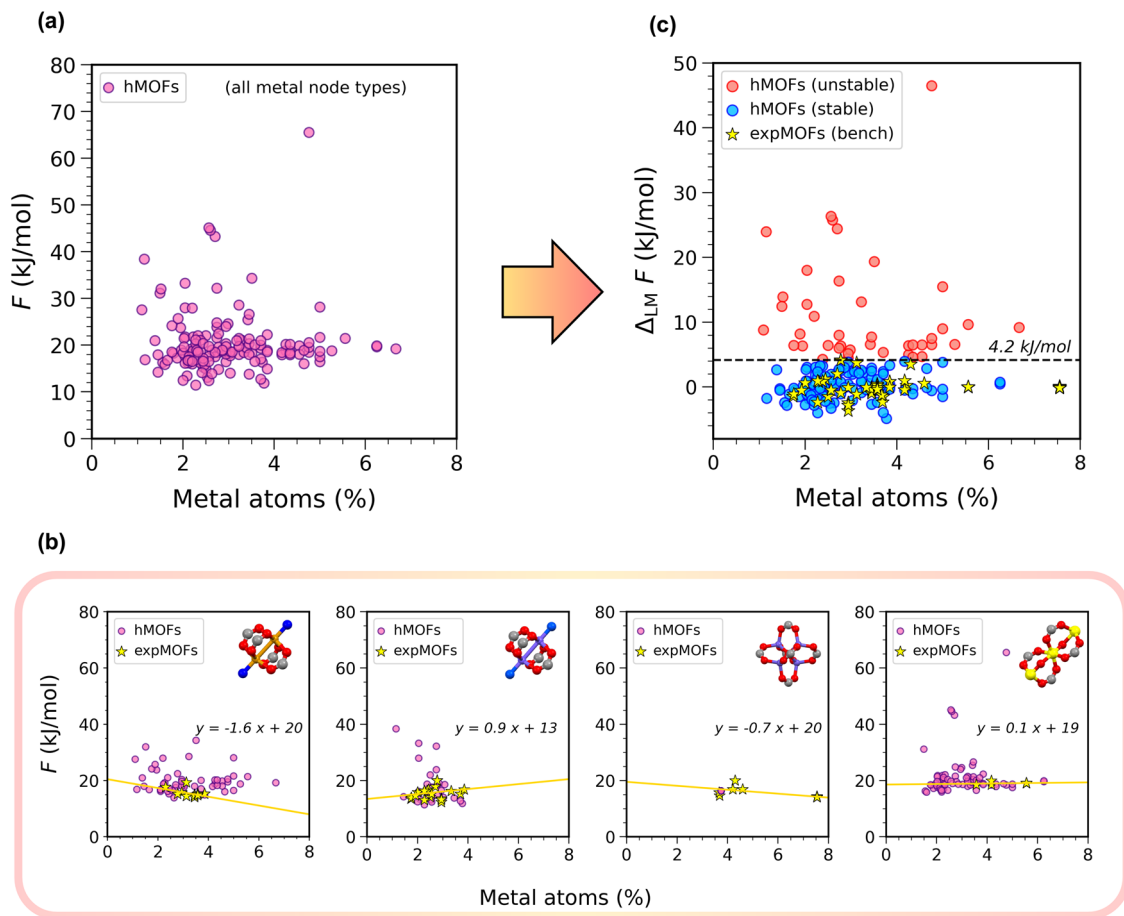


Fig. 3 Free energies of MOFs. **a** Free energies of 148 hMOFs with all the five metal node types versus the percentage of metal atoms. **b** Free energies of hMOFs and experimental MOFs with various metal node types: pillared Cu-paddlewheel, pillared Zn-paddlewheel, Zn_4O and V_3O_5 . **c** Relative free energies of 148 hMOFs (o) and selected benchmarking experimental MOFs (\star). The hMOFs with $\Delta_{LM}F > 4.2$ kJ/mol are considered thermodynamically unstable (unsynthesizable).

stress at this stage is a resolvable issue. Rather, we need to determine the intrinsic mechanical stability of unstressed MOFs. MD simulations were conducted at two different temperatures (0 and 298 K) to calculate elastic properties, including bulk (K), shear (G) and Young (E) moduli, which can be used to quantify the rigidity of MOFs.

Figure 4a shows the distributions of elastic moduli of 148 hMOFs at 0 K. Many hMOFs exhibit low moduli, which might be inappropriately classified as unstable (not robust) and lack of practical significance. A handful of experimentally synthesized MOFs have elastic moduli of even less than 5 GPa^{30,31}. They are commonly known as flexible MOFs, and their flexibility may bring favorable adsorptive characteristics. These flexible MOFs cannot be simply discarded by labeling them as unstable; thus, the elastic moduli alone are not adequate to judge mechanical stability. Instead, the Born stability criterion^{32,33} can be used to more reliably determine mechanical stability. Based on this criterion, all the eigenvalues of elastic matrix must be positive to declare a structure as stable, and the smallest eigenvalue can differentiate stable and unstable structures³⁴. As shown in Fig. 4b, at 0 K, only a few hMOFs do not satisfy the criterion and are marked as unstable. Nevertheless, several stable hMOFs show eigenvalues very close to zero, indicating these hMOFs possess a soft deformation mode³⁵ and may become unstable under stress or at a higher temperature. As presented in Fig. 4c, at 298 K, the number of unstable hMOFs increases. Consequently, it is

important to assess the mechanical stability at a relevant temperature, as the calculations at 0 K might be misleading.

Activation stability. Removal of solvent, modulator and unreacted substance is a prerequisite to activate MOFs for applications. On many occasions of activation, MOFs collapse to a dense amorphous phase, as commonly attributed to capillary force arising from evaporation or evacuation of solvent³⁶. But still, there is no coherent understanding of this phenomenon that would enable computational prediction of collapse events (for instance, using molecular simulation). Though there is a relation between collapse likelihood and mechanical properties, with the collapsed MOFs always showing low moduli for certain types of MOFs³¹, mechanical properties are insufficient to assess activation stability. This is because solvent characteristics and solvent-MOF interactions also affect the activation process.

As an alternative, ML models trained on experimental data can be used to predict activation stability. Using MOFSimplify³⁷, we predicted the activation scores of 148 hMOFs. As presented in Fig. 5a, the activation scores range from 0 to 1.0, corresponding to unstable (collapsible) and stable (activatable) likelihoods, respectively. An activation score of 0.5 is used here to distinguish collapsible and activatable structures. Among 148 hMOFs, 43 have activation scores < 0.5 , thus assigned as collapsible upon solvent removal. It is important to note that the employed ML models do

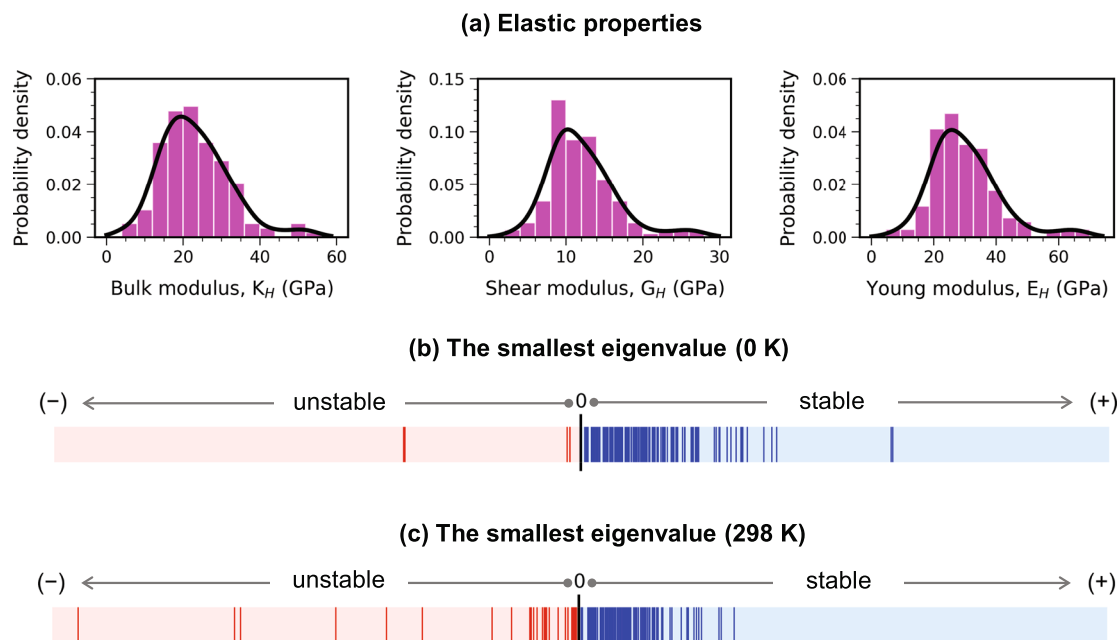


Fig. 4 Mechanical properties of 148 hMOFs. **a** Distributions of bulk, shear, and Young moduli at 0 K. Distributions of the smallest eigenvalues at **b** 0 K and **c** 298 K.

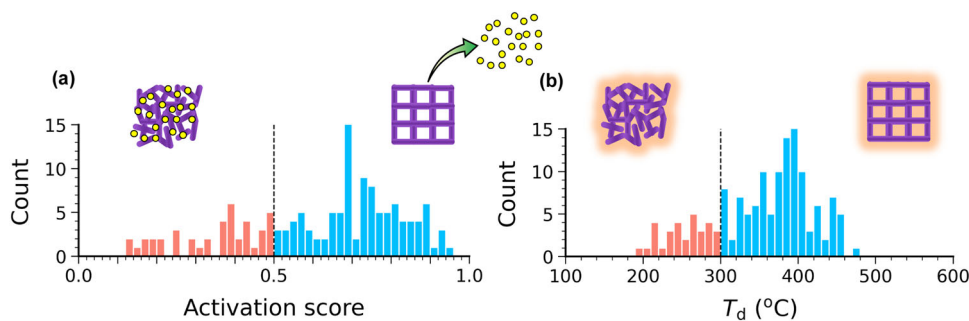


Fig. 5 Activation and thermal stabilities of 148 hMOFs. **a** Activation scores of 148 hMOFs, those with scores <0.5 are classified as collapsible. **b** Decomposition temperatures (T_d) of 148 hMOFs, 300 °C is used as the threshold for thermal stability.

not consider the actual activation method, which may affect the successfulness of activation. For instance, there are advanced activation methods, such as the supercritical CO_2 approach, that can effectively remove solvents if other conventional activation methods fail. Moreover, solvent characteristics and activation conditions may have considerable influences but were not taken into consideration in MOFSimplify. Yet, the adopted ML models appear as a simple and fast method to provide a reasonable assessment of whether the hMOFs would survive in activation.

Thermal stability. For high-temperature applications such as catalytic reactions, the thermal stability of MOFs is crucial. By contrast, post-combustion CO_2 capture is conventionally performed near ambient temperature. Thus, the major concern here about the thermal stability is when synthesized MOFs need to be thermally activated by increasing temperature to remove entrapped solvent or unreacted substance. As such, the temperature required is highly dependent on the boiling point of the solvent or the thermal decomposition temperature (T_d) of unreacted substance³⁸. If the interaction between solvent and MOF is strong, a high temperature is required beyond the boiling point to completely remove the solvent. For instance, HKUST-1 contains open metal sites with strongly bounded water molecules, and the

activation temperature can reach 200 °C³⁹. Consequently, it is challenging to set a distinct limit to quantify the thermal stability of MOFs. Yet, those decomposing below 300 °C can be considered thermally unstable, similar to hybrid ultra-microporous materials⁴⁰. Thus, 300 °C is set as the threshold for thermal stability. To predict the T_d of hMOFs, we also used MOFSimplify³⁷. As shown in Fig. 5b, among 148 hMOFs, 117 exhibit $T_d > 300$ °C, which indicates that most of the 148 hMOFs have decent thermal stability. The remaining 31 hMOFs are considered thermally unstable, and the lowest predicted T_d is ~199 °C.

Integration of stability metrics. Based on the above four stability metrics, Fig. 6a summarizes the percentages of stable and unstable structures among 148 hMOFs. The thermodynamic stable structures represent 72% of 148 hMOFs; thus, only 28% are unsynthesizable. Regarding mechanical stability, the percentage of stable structures at ambient temperature is 79%. The predicted activation and thermally stable structures, based on ML models, are 72% and 79%, respectively. However, these two stability metrics are subject to the accuracy of ML models and the designated boundaries to distinguish stable and unstable structures. Figure 6b shows the number of stable hMOFs based on the four stability metrics. Not surprisingly, some MOFs satisfy one

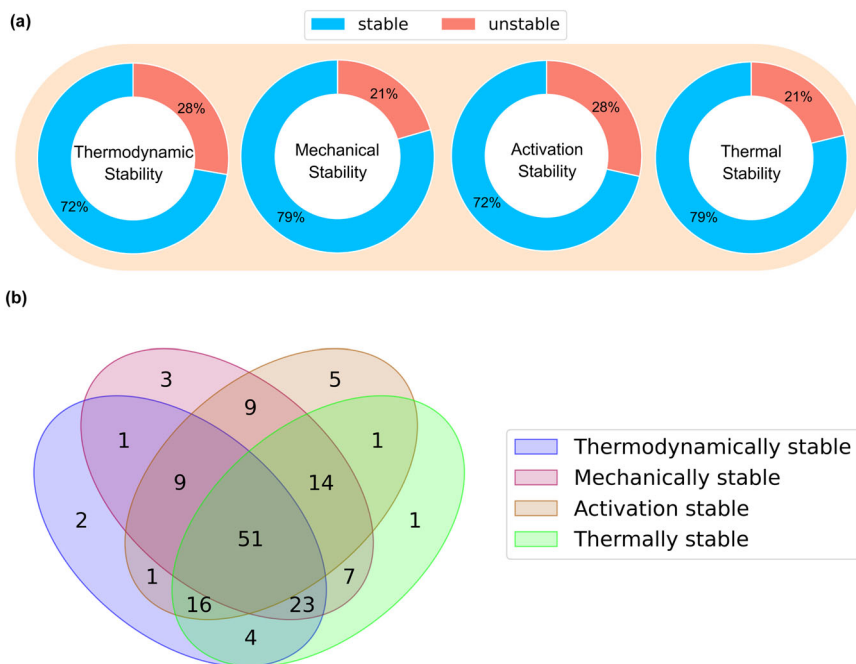


Fig. 6 Integration of stability metrics for 148 hMOFs. **a** Percentages of stable and unstable hMOFs based on four stability metrics. **b** Overlapping of stable hMOFs based on four stability metrics, the intersection among them gives the final top-performing, synthesizable, and stable hMOFs for CO₂ capture.

metric but fail in the other. Although the four metrics are evaluated using different approaches, considerable overlaps are seen as there are conceptual links among them. Particularly, the thermodynamic and mechanical stabilities are correlated, to a certain extent, with the thermal stability (Supplementary Fig. 2). The intersection in Fig. 6b gives 51 hMOFs, as listed in Supplementary Table 1, representing the final top-performing hMOFs that meet the separation performance requirement for CO₂ capture and also satisfy all the four stability requirements. The building blocks of these hMOFs are listed in Supplementary Table 2. Interestingly, a handful of top-performing hMOFs in Fig. 2c with the highest CO₂ uptake are found unstable. Therefore, integrating stability metrics is important to discard unstable hMOFs. It is worthwhile to note that the stability metrics can only tell if the structures are stable, but not how to synthesize them. Yet, it is possible to obtain an initial guess for synthesis conditions and additives using ML models that were trained using experimental data of successfully synthesized MOFs⁴¹.

Performance of a top-performing stable hMOF. Finally, it is instructive to provide microscopic insight into the separation mechanism in the top-performing stable hMOFs. For this purpose, grand-canonical Monte Carlo (GCMC) simulations were carried out for the adsorption of a CO₂/N₂ mixture (17%/83%) in a hMOF with V₃O₃ metal node and methyl functionalized pyrene-2,7-dicarboxylic acid (PDC) linker. This hMOF exhibits the highest CO₂ uptake among 51 stable candidates, along with high CO₂/N₂ selectivity. Figure 7a shows the adsorption isotherms of CO₂ and N₂ in this hMOF, which reveals far more significant adsorption of CO₂ over N₂ in the entire pressure range (0–1 bar). At 1 bar, CO₂ uptake is 7.49 mmol/g, and CO₂/N₂ selectivity is 355, which surpasses many reported experimental MOFs. The high CO₂ uptake in this hMOF is attributed to the presence of strong adsorbaphore¹⁶, which originated from parallel aromatic rings separated by a distance of 6.8 Å, as well as the favorable interaction sites in V₃O₃ metal node. The adsorption heats in Fig. 7b assert the preferential adsorption of CO₂ over N₂ at various pressures. Moreover, the separation performance of this

hMOF is assessed with the presence of moisture in the CO₂/N₂ mixture. As shown in Fig. 7c, d, both CO₂ uptake and CO₂/N₂ selectivity are marginally affected by the relative humidity (RH). Specifically, CO₂ uptake remains nearly constant at low-to-moderate RH, though there is a slight decrease at a high RH; meanwhile, N₂ uptake also decreases. Consequently, CO₂/N₂ selectivity increases slightly with RH. Overall, this hMOF is expected to separate CO₂ efficiently from dry or wet flue gas.

Conclusions

Computational screening of hMOFs often returns many top-performing candidates, yet there are large uncertainties in their synthetic likelihood and stabilities. In the present study, we integrate four stability metrics in the screening of hMOFs for CO₂ capture. The stabilities are evaluated using different approaches, including MD simulations and ML models. The synthetic likelihood (thermodynamic stability) is assessed by free energies, while the mechanical stability is examined through the Born stability criterion at ambient temperature. Moreover, hMOFs that retain activation and thermal stabilities are predicted. Finally, we identify the top-performing hMOFs satisfying all these stability requirements. For future work, chemical stability should also be taken into consideration to examine possible structural degradation if MOFs are exposed to water, corrosive gases, and organic solvents.

Methods

MOF database. A subset of hMOFs (labeled as DB5) in the ARC-MOF database¹⁸ was considered in the present study. This subset was originally created by Wilmer et al.⁴² and refined by Chung et al.⁴³ The adsorption data for post-combustion CO₂ capture were also extracted from the ARC-MOF database, as calculated from Monte Carlo simulations. Specifically, CO₂ uptake and CO₂/N₂ selectivity at 298 K and 0.9 bar at a feed composition of 17%/83% for a CO₂/N₂ mixture were used. Among 22,366 hMOFs with adsorption data in DB5, we investigated the non-interpenetrated (non-catenated) hMOFs by verifying the second integers of their 6-chromosome names (which encode actual

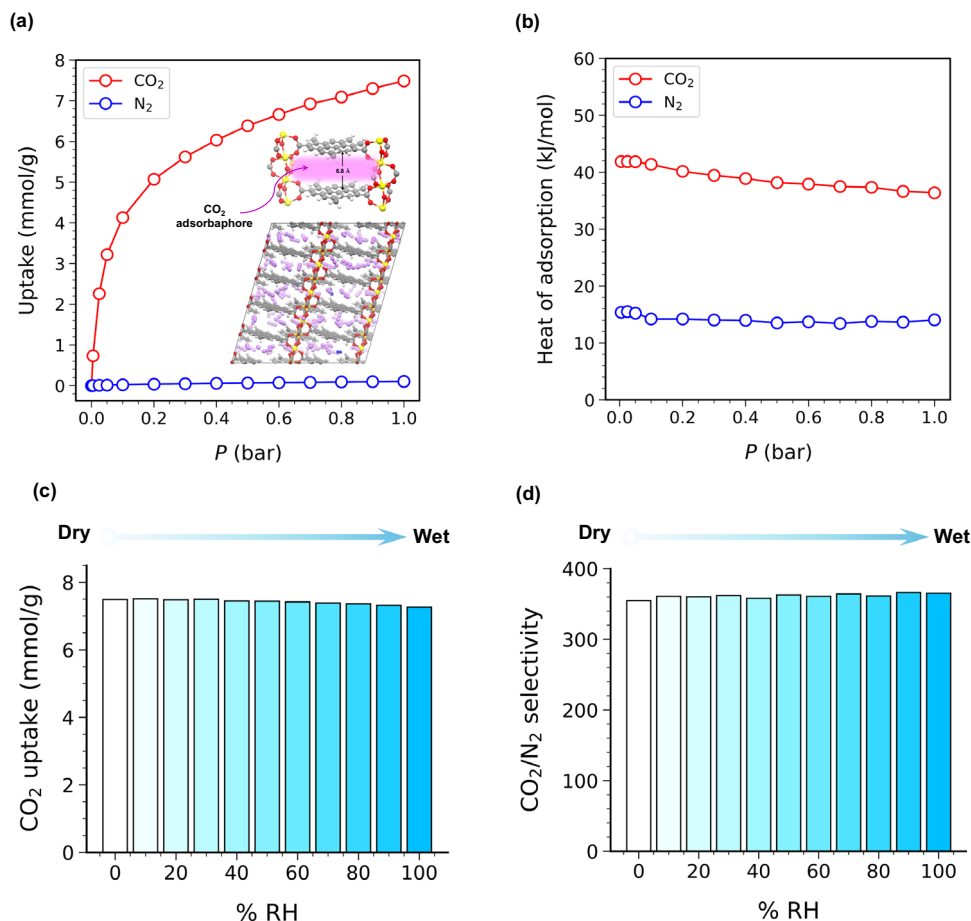


Fig. 7 Separation performance of a top-performing stable hMOF. **a** Adsorption isotherms and **b** adsorption heats of CO₂ and N₂ for a CO₂/N₂ mixture (17%/83%) at 298 K. The insets illustrate CO₂ adsorbaphore and simulation snapshot at 1.0 bar (CO₂: purple, N₂: blue). **c** CO₂ uptake and **d** CO₂/N₂ selectivity versus relative humidity (RH).

interpenetration), resulting in 15,219 hMOFs. The selected dataset comprises five different metal node types: Zn₄O, pillared Cu-paddlewheel, pillared Zn-paddlewheel, V₃O₃ and Zr₆O₆.

MD simulations. The thermodynamic and mechanical stabilities were evaluated using MD simulations. For a given MOF, the LAMMPS interface⁴⁴ was used to prepare input files for bonded and non-bonded interaction potentials. It was found, for many of the hMOFs under study, that the implemented bonding algorithm in the LAMMPS interface led to incorrectly coordinated atoms or bonds, which subsequently caused the failure of MD simulations or inaccurate calculations of desired properties. To overcome this issue, the bonding information was precomputed and added to the Crystallographic Information Files (CIFs) prior to using the LAMMPS interface. As employed by Bruner et al.¹⁸, the *IsayevNN*⁴⁵ algorithm in *Pymatgen*⁴⁶ was used to generate the bonding information. This algorithm considers two atoms as neighbors (bonded) if they share a Voronoi facet and the separated distance is less than the sum of their covalent radii plus a tolerance. The default tolerance (0.25 Å) was used to detect bonding and found reliable to detect the Cu-Cu bond in pillared Cu-paddlewheel metal node present in the hMOFs. Yet, to identify the Zn-Zn bond in the pillared Zn-paddlewheel metal node, the tolerance was increased (0.5 Å) to search for neighboring Zn atoms. Computing the bonds based on the above algorithm resulted in fewer structures with incorrect bonds, which were manually corrected. The interactions between different atoms were described using the UFF4MOF force field⁴⁷.

For pillared Cu- and Zn-paddlewheel metal nodes, we assigned Cu₄+2 and Zn₄+2 metal types, respectively. It was revealed in the literature that the inclusion of electrostatic interactions with UFF4MOF would lead to inaccurate calculations of MOF properties such as elastic moduli and thermal expansion coefficients, and the optimized MOF structures would differ significantly from experimental measurements⁴⁴. Thus, electrostatic interactions were not included here, as also in previous studies^{24,34}. All the MD simulations were performed using the LAMMPS package⁴⁸.

Thermodynamic stability. To determine thermodynamic stability (i.e., synthetic likelihood), the free energies of hMOFs were calculated using the Frenkel–Ladd (FL) thermodynamic integration method^{22,23}. It has been proven that the FL method predicts the free energies of MOFs more accurately compared to other approaches²⁴. The FL path is a thermodynamic integration between the ideal Einstein crystal and its actual representation described by a force field. MD simulations were performed at 298 K and zero pressure, repeated three times using different seeds for the initial velocity distributions, and the computed free energies were averaged. The *fix ti/spring* command was used to perform thermodynamic integration, with 400 ps switching time using 1 fs timestep. The free energies of hMOFs were computed per atom basis, with a maximum standard deviation of ± 0.29 kJ/mol. To benchmark the free energies of the hMOFs, a set of experimental MOFs was used as a reference. Specifically, the CoRE MOFs⁹, labeled as DB12 in the ARC-MOF database¹⁸ were used. As in the hMOFs, the non-interpenetrated (non-catenated)

structures were detected using MOFid⁴⁹ and considered. Given that free energy is sensitive to metal node type, only the CoRE MOFs possessing the same metal node as those in the hMOFs were selected. As such, the revised autocorrelations (RACs)^{25,26} for metal nodes were used to find the matching CoRE MOFs. First, the hMOFs were grouped based on their metal node types, and the centroid of each cluster of hMOFs was calculated using *KMeans* algorithm, as implemented in *Scikit-learn*. Then, the *Nearest-Neighbors* algorithm was used to locate the nearest 20 CoRE MOFs to each centroid calculated for hMOFs. Lastly, the final selected CoRE MOFs were manually examined to ensure matching with the designated metal node and free of any disorder. A handful of duplicate structures were found, especially those having V₃O₃ metal node.

Mechanical stability. To examine mechanical stability, the average elastic properties, including bulk (K), shear (G), and Young (E) moduli, as well as the Born stability criterion, were evaluated at both 0 K and 298 K. Derived from Hooke's law that relates stress with strain, the elastic constants can be calculated through an explicit deformation approach. At 0 K, initially, a structure was optimized using two stages of energy minimization. The first stage was performed using the *conjugate gradient* algorithm, coupled with triclinic box relaxation at zero pressure, while in the second stage, the *fire* algorithm⁵⁰ was used at a fixed volume. In each stage, the energy and force-stopping tolerance were set as 10⁻¹⁵ kcal mol⁻¹ Å⁻¹ with 10⁵ maximum iterations or force evaluations. This scheme was repeated for 200 cycles, where atomic positions were slightly disturbed between cycles to achieve minimum energy configuration. At the end of structure relaxation, the pressures (stresses) were saved. Next, the structure was strained (1%) in various directions, followed by 10 cycles of energy minimization with the *fire* algorithm to optimize atomic positions at a fixed volume. The corresponding pressures (stresses) on the structure at each deformation were collected and used with the initial stresses to calculate elastic constants. To calculate the elastic constants at 298 K, the above explicit deformation method was used after replacing energy minimization with MD simulation, though this approach is computationally expensive as the structure must be strained in various directions and followed by long MD simulation. Moreover, a large simulation box (minimum length of 50 Å) was used to improve accuracy. First, the simulation box was relaxed at zero pressure in the NPT (triclinic) ensemble using the Nosé–Hoover thermostat and barostat to control temperature and pressure, respectively. The equilibration time was 4 ns, followed by 4 ns production, during which the cell parameters were sampled. Afterward, the simulation box was adjusted to the averaged cell parameters. Then, simulation was performed in the NVT ensemble (0.5 ns equilibration, 4 ns production), during which the pressures (stresses) acting on the simulation box were collected every timestep, averaged, and saved as the initial stresses for the non-deformed structure. Next, deformations were applied to various directions (1% strain). For each deformation, simulation was performed in the NVT ensemble (0.5 ns equilibration, 4 ns production), and the stresses during production were collected and averaged. Finally, from the collected stresses of the non-deformed and deformed structures, the elastic constants were calculated.

Activation and thermal stabilities. The activation and thermal stabilities were predicted using ML-based MOFSimplify³⁷. This platform takes a CIF as an input and then internally conducts featurization and prediction. For activation stability, the implemented ML model was trained on thousands of experimental MOFs to predict the likelihood of removing solvents without structural collapse. For thermal stability, the ML model was

trained on experimentally measured decomposition temperatures (T_d) extracted from thermogravimetric analysis (TGA) data.

Monte Carlo simulations. After identifying top-performing synthesizable and stable hMOFs for CO₂ capture, to provide in-depth insight into separation performance, grand-canonical Monte Carlo (GCMC) simulations were conducted for the adsorption of a CO₂/N₂ mixture (composition of 17%/83%) in a hMOF at 298 K. The force fields used were identical to those in the ARC-MOF database¹⁸. Specifically, the university force field (UFF)⁵¹ was used to describe the dispersion interactions of framework atoms, while the REPEAT atomic charges were included for electrostatic interactions. For CO₂ and N₂ adsorbates, the potential parameters were adopted from literature^{52,53}. Furthermore, to examine the effect of moisture on separation, GCMC simulations were also performed at different relative humidities. H₂O was represented by the TIP4P model⁵⁴, which has a saturation pressure of 4365 Pa⁵⁵. The Lorentz–Berthelot combining rules were used to calculate the cross interactions. The cutoff to evaluate dispersion intermolecular interactions was set to 12.8 Å, and the Ewald sum was used to calculate long-range electrostatic interactions. The number of cycles was 200,000, with the first half for equilibration and the second half for production. During GCMC simulations, the framework was held rigid. Five different trial moves were attempted for the adsorbate molecules, including translation, rotation, reinsertion, identity change and swap with probabilities of 16.7%, 16.7%, 16.7%, 16.7% and 33.2%, respectively. All the GCMC simulations were conducted using the RASPA⁵⁶.

Data availability

The data that support the findings of this study have been included in the manuscript and Supplementary Information. Any additional data are available from the corresponding author upon request.

Received: 20 July 2023; Accepted: 22 September 2023;
Published online: 03 October 2023

References

1. Furukawa, H., Cordova, K. E., O'Keeffe, M. & Yaghi, O. M. The chemistry and applications of metal–organic frameworks. *Science* **341**, 1230444 (2013).
2. Colón, Y. J. & Snurr, R. Q. High-throughput computational screening of metal–organic frameworks. *Chem. Soc. Rev.* **43**, 5735–5749 (2014).
3. Daglar, H. & Keskin, S. Recent advances, opportunities, and challenges in high-throughput computational screening of MOFs for gas separations. *Coord. Chem. Rev.* **422**, 213470 (2020).
4. Ren, E., Guilbaud, P. & Coudert, F.-X. High-throughput computational screening of nanoporous materials in targeted applications. *Digit. Discov.* **1**, 355–374 (2022).
5. Jiang, J. Computational screening of metal–organic frameworks for CO₂ separation. *Curr. Opin. Green Sustain. Chem.* **16**, 57–64 (2019).
6. Chung, Y. G. et al. Computation-ready, experimental metal–organic frameworks: a tool to enable high-throughput screening of nanoporous crystals. *Chem. Mater.* **26**, 6185–6192 (2014).
7. Qiao, Z., Zhang, K. & Jiang, J. In silico screening of 4764 computation-ready, experimental metal–organic frameworks for CO₂ separation. *J. Mater. Chem. A* **4**, 2105–2114 (2016).
8. Chung, Y. G. et al. Computational screening of nanoporous materials for hexane and heptane isomer separation. *Chem. Mater.* **29**, 6315–6328 (2017).
9. Chung, Y. G. et al. Advances, updates and analytics for the computation-ready, experimental metal–organic framework database: CoRE MOF 2019. *J. Chem. Eng. Data* **64**, 5985–5998 (2019).
10. Tang, H. & Jiang, J. In silico screening and design strategies of ethane-selective metal–organic frameworks for ethane/ethylene separation. *AIChE J.* **67**, 1–14 (2021).
11. Tang, H., Xu, Q., Wang, M. & Jiang, J. Rapid screening of metal–organic frameworks for propane/propylene separation by synergizing molecular simulation and machine learning. *ACS Appl. Mater. Interfaces* **13**, 53454–53467 (2021).

12. Groom, C. R., Bruno, I. J., Lightfoot, M. P. & Ward, S. C. The Cambridge structural database. *Acta Crystallogr. Sect. B* **72**, 171–179 (2016).
13. Avci, G., Erucar, I. & Keskin, S. Do new MOFs perform better for CO₂ capture and H₂ purification? Computational screening of the updated MOF database. *ACS Appl. Mater. Interfaces* **12**, 41567–41579 (2020).
14. Wilmer, C. E., Farha, O. K., Bae, Y. S., Hupp, J. T. & Snurr, R. Q. Structure-property relationships of porous materials for carbon dioxide separation and capture. *Energy Environ. Sci.* **5**, 9849–9856 (2012).
15. Qiao, Z., Peng, C., Zhou, J. & Jiang, J. High-throughput computational screening of 137953 metal-organic frameworks for membrane separation of a CO₂/N₂/CH₄ mixture. *J. Mater. Chem. A* **4**, 15904–15912 (2016).
16. Boyd, P. G. et al. Data-driven design of metal-organic frameworks for wet flue gas CO₂ capture. *Nature* **576**, 253–256 (2019).
17. Majumdar, S., Moosavi, S. M., Jablonka, K. M., Ongari, D. & Smit, B. Diversifying databases of metal-organic frameworks for high-throughput computational screening. *ACS Appl. Mater. Interfaces* **13**, 61004–61014 (2021).
18. Burner, J. et al. ARC-MOF: a diverse database of metal-organic frameworks with DFT-derived partial atomic charges and descriptors for machine learning. *Chem. Mater.* **35**, 900–916 (2023).
19. Bae, Y. S. & Snurr, R. Q. Development and evaluation of porous materials for carbon dioxide separation and capture. *Angew. Chemie Int. Ed. Engl.* **50**, 11586–11596 (2011).
20. Park, J. et al. How well do approximate models of adsorption-based CO₂ capture processes predict results of detailed process models? *Ind. Eng. Chem. Res.* **59**, 7097–7108 (2020).
21. Leperi, K. T., Chung, Y. G., You, F. & Snurr, R. Q. Development of a general evaluation metric for rapid screening of adsorbent materials for postcombustion CO₂ capture. *ACS Sustain. Chem. Eng.* **7**, 11529–11539 (2019).
22. Frenkel, D. & Ladd, A. J. C. New Monte Carlo method to compute the free energy of arbitrary solids. Application to the FCC and HCP phases of hard spheres. *J. Chem. Phys.* **81**, 3188–3193 (1984).
23. Freitas, R., Asta, M. & De Koning, M. Nonequilibrium free-energy calculation of solids using LAMMPS. *Comput. Mater. Sci.* **112**, 333–341 (2016).
24. Anderson, R. & Gómez-Gualdrón, D. A. Large-scale free energy calculations on a computational metal-organic frameworks database: toward synthetic likelihood predictions. *Chem. Mater.* **32**, 8106–8119 (2020).
25. Janet, J. P. & Kulik, H. J. Resolving transition metal chemical space: feature selection for machine learning and structure-property relationships. *J. Phys. Chem. A* **121**, 8939–8954 (2017).
26. Moosavi, S. M. et al. Understanding the diversity of the metal-organic framework ecosystem. *Nat. Commun.* **11**, 1–10 (2020).
27. Stock, N. & Biswas, S. Synthesis of metal-organic frameworks: routes to various MOF topologies, morphologies, and composites. *Chem. Rev.* **112**, 933–969 (2012).
28. Cheetham, A. K., Kieslich, G. & Yeung, H. H. M. Thermodynamic and kinetic effects in the crystallization of metal-organic frameworks. *Acc. Chem. Res.* **51**, 659–667 (2018).
29. Wang, T. C. et al. Surviving under pressure: the role of solvent, crystal size and morphology during pelletization of metal-organic frameworks. *ACS Appl. Mater. Interfaces* **13**, 52106–52112 (2021).
30. Ying, P., Zhang, J. & Zhong, Z. Pressure-induced phase transition of isoreticular MOFs: mechanical instability due to ligand buckling. *Microporous Mesoporous Mater.* **312**, 110765 (2021).
31. Mohamed, S. A., Kim, Y., Lee, J., Choe, W. & Kim, J. Understanding the structural collapse during activation of metal-organic frameworks with copper paddlewheels. *Inorg. Chem.* **61**, 9702–9709 (2022).
32. Born, M. & Misra, R. D. On the stability of crystal lattices. *Math. Proc. Cambridge Philos. Soc.* **36**, 466–478 (1940).
33. Mouhat, F. & Coudert, F. X. Necessary and sufficient elastic stability conditions in various crystal systems. *Phys. Rev. B* **90**, 4–7 (2014).
34. Anderson, R. & Gómez-Gualdrón, D. A. Increasing topological diversity during computational synthesis of porous crystals: how and why. *CrystEngComm* **21**, 1653–1665 (2019).
35. Ortiz, A. U., Boutin, A., Fuchs, A. H. & Coudert, F.-X. Metal-organic frameworks with wine-rack motif: what determines their flexibility and elastic properties? *J. Chem. Phys.* **138**, 174703 (2013).
36. Zhang, X. et al. A historical overview of the activation and porosity of metal-organic frameworks. *Chem. Soc. Rev.* **49**, 7406–7427 (2020).
37. Nandy, A. et al. MOFSimplify, machine learning models with extracted stability data of three thousand metal-organic frameworks. *Sci. Data* **9**, 74 (2022).
38. Woodliffe, J. L., Ferrari, R. S., Ahmed, I. & Laybourn, A. Evaluating the purification and activation of metal-organic frameworks from a technical and circular economy perspective. *Coord. Chem. Rev.* **428**, 213578 (2021).
39. Lin, K. S., Adhikari, A. K., Ku, C. N., Chiang, C. L. & Kuo, H. Synthesis and characterization of porous HKUST-1 metal-organic frameworks for hydrogen storage. *Int. J. Hydrogen Energy* **37**, 13865–13871 (2012).
40. Healy, C. et al. The thermal stability of metal-organic frameworks. *Coord. Chem. Rev.* **419**, 213388 (2020).
41. Luo, Y. et al. MOF synthesis prediction enabled by automatic data mining and machine learning. *Angew. Chemie Int. Ed. Engl.* **61**, e202200242 (2022).
42. Wilmer, C. E. et al. Large-scale screening of hypothetical metal-organic frameworks. *Nat. Chem.* **4**, 83–89 (2012).
43. Chung, Y. G. et al. In silico discovery of metal-organic frameworks for precombustion CO₂ capture using a genetic algorithm. *Sci. Adv.* **2**, e1600909 (2016).
44. Boyd, P. G., Moosavi, S. M., Witman, M. & Smit, B. Force-field prediction of materials properties in metal-organic frameworks. *J. Phys. Chem. Lett.* **8**, 357–363 (2017).
45. Isayev, O. et al. Universal fragment descriptors for predicting properties of inorganic crystals. *Nat. Commun.* **8**, 1–12 (2017).
46. Ong, S. P. et al. Python Materials Genomics (pymatgen): a robust, open-source Python library for materials analysis. *Comput. Mater. Sci.* **68**, 314–319 (2013).
47. Addicoat, M. A., Vankova, N., Akter, I. F. & Heine, T. Extension of the universal force field to metal-organic frameworks. *J. Chem. Theory Comput.* **10**, 880–891 (2014).
48. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
49. Bucior, B. J. et al. Identification schemes for metal-organic frameworks to enable rapid search and cheminformatics analysis. *Cryst. Growth Des.* **19**, 6682–6697 (2019).
50. Bitzek, E., Koskinen, P., Gähler, F., Moseler, M. & Gumbusch, P. Structural relaxation made simple. *Phys. Rev. Lett.* **97**, 1–4 (2006).
51. Rappé, A. K., Casewit, C. J., Colwell, K. S., Goddard, W. A. & Skiff, W. M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **114**, 10024–10035 (1992).
52. García-Sánchez, A. et al. Transferable force field for carbon dioxide adsorption in zeolites. *J. Phys. Chem. C* **113**, 8814–8820 (2009).
53. Provost, B. *An Improved N2 Model for Predicting Gas Adsorption in MOFs and Using Molecular Simulation to Aid in the Interpretation of SSNMR Spectra of MOFs*. (University of Ottawa, 2015).
54. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
55. Datar, A., Witman, M. & Lin, L. Monte Carlo simulations for water adsorption in porous materials: best practices and new insights. *AIChE J.* **67**, e17447 (2021).
56. Dubbeldam, D., Calero, S., Ellis, D. E. & Snurr, R. Q. RASPA: molecular simulation software for adsorption and diffusion in flexible nanoporous materials. *Mol. Simul.* **42**, 81–101 (2016).

Acknowledgements

We gratefully acknowledge A*STAR LCER-FI projects (LCERFI01-0015 U2102d2004 and LCERFI01-0033 U2102d2006) for financial support.

Author contributions

S.A.M. and J.J. conceived the conceptual idea and devised the research plan. S.A.M. performed database acquisition, screening based on separation performance metrics, MD simulations to evaluate the thermodynamic and mechanical stabilities and GCMC simulations in a top-performing stable hMOF. D.Z. evaluated the thermal and activation stabilities. S.A.M. analyzed the results and prepared the manuscript, and J.J. revised and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s43246-023-00409-9>.

Correspondence and requests for materials should be addressed to Jianwen Jiang.

Peer review information *Communications Materials* thanks Yongchul G. Chung and the other anonymous reviewer(s) for their contribution to the peer review of this work. Primary handling editor: Jet-Sing Lee.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023