Check for updates

Integrating crystallographic and computational approaches to carbon-capture materials for the mitigation of climate change

Eric Cockayne, [ID] *[a] Austin McDannald, [ID] [a] Winnie Wong-Ng,[a] Yu-Sheng Chen,[b] Jason Benedict,[c] Felipe Gándara Barragán, [ID] [d] Christopher H. Hendon, [ID] [e] David A. Keen, [ID] [f] Ute Kolb,[g] Lan Li,[h] Shengqian Ma, [ID] [i] William Morris,[j] Aditya Nandy,[k] Tomče Runčevski,[l] Mustapha Soukri, [ID] [m] Anuroop Sriram, [ID] [n] Janice A. Steckel, [ID] [o] John Findley,[o] Chris Wilmer, [ID] [pq] Taner Yildirim,[r] Wei Zhou,[r] Igor Levin[a] and Craig Brown [ID] [r]

This article presents an overview of the current state of the art in the structure determination of microporous carbon-capture materials, as discussed at the recent NIST workshop "Integrating Crystallographic and Computational Approaches to Carbon-Capture Materials for the Mitigation of Climate Change". The continual rise in anthropogenic CO_2 concentration and its effect on climate change call for the implementation of carbon capture technologies to reduce the CO_2 concentration in the atmosphere. Porous solids, including metal–organic frameworks (MOFs), are feasible candidates for gas capture and storage applications. However, determining the structure of these materials represents a significant obstacle in their development into advanced sorbents. The existing difficulties can be overcome by integrating crystallographic methods and theoretical modeling. The workshop gathered experimentalists and theorists from academia, government, and industry to review this field and identify approaches, including collaborative opportunities, required to develop tools for rapid determination of the structures of porous solid sorbents and the effect of structure on the carbon capture performance. We highlight the findings of that workshop, especially in the need for reference materials, standardized procedures and reporting of sorbent activation and adsorption measurements, standardized reporting of theoretical calculations, and round-robin structure determination.

1. Introduction

The concentration of carbon dioxide (CO_2) in the atmosphere continues to rise, exceeding 424 μmol mol^{-1} (colloquially known as parts per million (ppm)) for the first time in May 2023.[1] At the current rate of increase, the CO_2 concentration will reach double the pre-industrial revolution value of about 280 μmol mol^{-1} within 50 years. The continual rise in

[a]*Materials Measurement Science Division, Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899-8520, USA. E-mail: eric.cockayne@nist.gov*

[b]*NSF's ChemMatCARS, University of Chicago, Lemont, IL 60439, USA*

[c]*Department of Chemistry, University at Buffalo, 771 Natural Sciences Complex, Buffalo, NY 14260-3000, USA*

[d]*Materials Science Institute of Madrid – Spanish National Research Council, Madrid, Spain*

[e]*Department of Chemistry and Biochemistry, University of Oregon, Eugene, OR, 97403, USA*

[f]*ISIS Facility, Rutherford Appleton Laboratory, Harwell Campus, Didcot, Oxfordshire OX11 0QX, UK*

[g]*Department Chemie, Johannes Gutenberg-University Mainz, 55128 Mainz, Germany*

[h]*Micron School of Materials Science and Engineering, Boise State University, Boise, ID 83725, USA*

[i]*Department of Chemistry, University of North Texas, Denton, Texas 76203, USA*

[j]*Numat, 1358 N. Kostner Avenue, Chicago, Illinois 60651, USA*

[k]*James Franck Institute and Department of Chemistry, University of Chicago, Chicago, IL 60637, USA*

[l]*Department of Chemistry, Southern Methodist University, Dallas, TX, 75275, USA*

[m]*RTI International, 3040 East Cornwallis Road, Research Triangle Park, Durham, NC 27709-2194, USA*

[n]*Meta AI, Menlo Park, CA, USA*

[o]*U.S. Department of Energy, National Energy Technology Laboratory, 626 Cochran's Mill Rd, Pittsburgh, PA 15236, USA*

[p]*Department of Chemical and Petroleum Engineering, Department of Electrical and Computer Engineering, University of Pittsburgh, 3700 O'Hara Street, Pittsburgh, Pennsylvania 15261, USA*

[q]*Clinical and Translational Science Institute, University of Pittsburgh, Meyran Avenue, Suite 7057, Pittsburgh, Pennsylvania 15213, USA*

[r]*Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, MD 20899-6102, USA*

25678 | *J. Mater. Chem. A*, 2024, **12**, 25678–25695

This journal is © The Royal Society of Chemistry 2024

anthropogenic CO_2 concentration and its effect on climate change call for the implementation of Carbon Capture and Storage (CCS) technologies to reduce CO_2 emissions, such as the removal of CO_2 from industrial gas streams or, even better, the implementation of so-called negative emissions technologies (NETs),[2] such as direct-air capture (DAC). Unlike removing CO_2 from emitters such as coal-fired power stations, DAC can be, in principle, done anywhere. However, the concentration of CO_2 in the atmosphere is orders of magnitude lower than that in post-combustion flue gas, making DAC more challenging. In this document we refer to DAC as carbon capture technologies starting from near atmospheric CO_2 concentrations (424 μmol mol^{-1}), and point source capture (also known as post-combustion capture or flue-gas capture) as starting from elevated CO_2 concentrations (roughly 200 000 μmol mol^{-1}).[3]

Materials of interest for capturing CO_2 both from point sources, such as industrial waste streams, and for DAC include liquid sorbents and solid sorbents. Porous solids have several features that are useful for carbon capture. They can provide a very large effective surface area per mass or volume for physical or chemical adsorption. Porous solids that are crystalline provide a uniform pore distribution that makes it efficient to filter species by size or degree of chemical interaction. The thermodynamics of carbon capture and release may also be more favorable for solid sorbents than for liquid amines as cycles such as pressure or temperature swings do not involve costly liquid heating.

Naturally porous crystalline materials include octahedral molecular sieves such as α-MnO_2, zeolites, porous coordination polymers (PCPs), of which an important subset is porous metal–organic frameworks (MOFs), and covalent organic frameworks (COFs). All of these classes encompass multiple examples, but the PCP/MOF and COF classes are particularly rich because, given one structure, an endless number of similarly-terminated organic ligands can, in principle, be substituted to make new structures with the same framework topology and local coordination.

This report pays special attention to flexible PCP/MOF and COF materials,[4,5] which are interesting as they typically exhibit large structural changes in response to changes in the environmental conditions, such as temperature, pressure, or gas adsorption, often reversibly and without bond breaking. For instance, the flexible MOF MIL-53(Cr), undergoes a narrow-pore → large-pore transition (see Fig. 1), with one lattice dimension changing from 7.85 Å to 13.04 Å, a change that is much larger than encountered in typical responsive bulk solids such as piezoelectrics.[6]
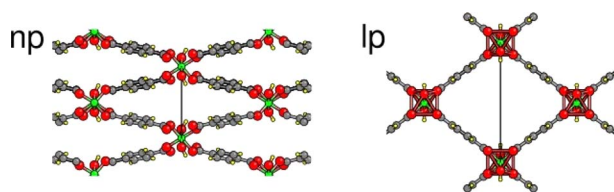
The features that make porous crystalline materials interesting can also make structure determination a challenge. Structures may be "more than the sum of their parts", as the connectivity in MOFs may not reflect preconceived notions; complex organic ligands may bend and twist, the overall structure may distort, and the system may be flexible and change structure upon absorption of CO_2 and other species. Compounds may be difficult to synthesize in single-crystal form yet still have useful properties. For such systems, it is important to determine their structures, but the structures may be too complicated to solve from powder diffraction alone.

NIST contributes to measurement science related to climate *via* multiple programs, including the NIST project on Carbon Capture and Carbon Sequestration.[7] The NIST workshop "Integrating Crystallographic and Computational Approaches to Carbon-Capture Materials for the Mitigation of Climate Change", was held at the NIST National Cybersecurity Center of Excellence building in Rockville, Maryland, October 31–November 1, 2023, to provide a forum for experimentalists and theorists from academia, government, and industry who are working on the structural aspects of CO_2 capture and sequestration materials to review the current state of the art in this field and discuss opportunities for collaborative research required to develop tools for rapid determination of the structures of porous solid sorbents and the effect of structure on the carbon capture performance.

In Section 2–5, this report discusses the state of the art of experimental, theoretical, and computational modeling of the structures of solid sorbents for CO_2 capture as presented at the workshop. Section 6 provides a brief overview of some of the industrial concerns in this field. In Section 7, we discuss the outstanding issues that were brought up at the workshop and provide suggestions for further work. Section 8 presents our conclusions and suggests action items. This article does not aim to be a comprehensive review of the topics covered. Rather, we present the topics side-by-side, each providing context for the others and reflective of the aforementioned workshop. In doing so, we aim to further the dialog between these sub-disciplines. How the relevant properties are being modeled and measured affects what aspects of the structural characterization are important to study, and *vice versa*. There are many different material systems that are being considered for carbon capture technologies – and many ways that these materials might be implemented in a carbon-capture system (*e.g.* sorbents *vs.* membranes). The discussions of the state of the art of each aspect of this rich field of study provides context for the others.

For the sake of brevity, many of the compounds in this work are described using abbreviated names as given in the literature rather than by their full chemical nomenclature.

2. Cutting-edge experimental structural characterization methods

Scattering (X-ray, neutron, and electron) methods are a powerful tool for structural determination. This section details the progress of single-crystal and powder diffraction techniques,

Fig. 1 Narrow pore (np) and large pore (lp) structures of the flexible MOF MIL-53(Cr). The Cr atoms are green, C grey, O red and H small yellow. The Cr–Cr distances shown changes from 7.85 Å in the np phase to 13.04 Å in the lp phase.

with particular emphasis on their applications to crystalline porous materials for carbon capture.

2.1 Single-crystal diffraction

We must know the structure of a porous absorbent material to fully explain its gas adsorption behaviour. The structure of porous materials, such as MOFs and zeolites, is commonly determined using X-ray and neutron diffraction data collected from single-crystals. *In situ* X-ray diffraction is a powerful method for studying structural changes in porous crystals upon gas adsorption. Such studies can even reveal the binding sites and structural distortions associated with CO_2 adsorption in flexible MOFs, provided that the MOF remains crystalline.[8]

One new research avenue being pursued with single-crystal diffraction is dynamic responses of structures to stimuli. The beamline NSF's ChemMatCARS at the Advanced Photon Source (APS) at Argonne National Laboratory has been outfitted with an environment cell that allows for a single crystal to be exposed to different environments (*e.g.* temperature changes or sorbate gasses) as the structure is being measured.[9] This allows for the study of structure dynamics *in situ* that would be missed in traditional *ex situ* experiments. For example, an *ex situ* study in 2009 of the Co-based MOF [Co(AIP)(bpy)$_{0.5}$(H$_2$O)]·2H$_2$O found several structural changes upon heating and/or exposure to other solvents.[10] In particular, they found a structural change to an anhydrous phase that had the same space group with a much smaller a-axis length but little change in the other lattice parameters. A dynamic *in situ* study of the same material[11] showed that one of the water molecules leaves suddenly, causing an abrupt change in the a-axis length and also the lattice angle β, resulting in a shearing of the lattice as the ligands deform. As the material is dehydrated further, the other water molecules leave the structure gradually, which brings the lattice back to a similar β value as in the starting trihydrate. This shearing and "un-shearing" action as the structure is dehydrated is not evident in the end members and, therefore, not observable by *ex situ* studies. Only by measuring the structure *in situ* could this multi-stage dehydration mechanism be uncovered. A dynamic *in situ* study of ethanol incorporation in the same [Co(AIP)(bpy)$_{0.5}$(H$_2$O)]·2H$_2$O system revealed further details of chemical transformations in this MOF.[12]

2.2 Powder diffraction

While single-crystal diffraction is extremely useful, it can be challenging to obtain crystals of MOFs large enough for diffraction measurements *via* typical synthesis methods such as precipitation from solution. Furthermore, even when the synthesis route is successful, the crystals may crack or break apart as the solvent is removed or under adsorbate pressure. The structural studies of most MOFs therefore rely on X-ray (and to a lesser extent, neutron) powder diffraction measurements.[13]

Powder diffraction has its advantages. Because of the ease of sample preparation, it is applicable to more materials and, because there is no need to orient the specimen in a particular way, it is easier to perform in *operando* studies. With powder diffraction, it is possible to study MOFs as synthesized – with small crystals and loaded with a solvent. Then the structure can be measured as the solvent is removed under vacuum and heat, and again as material is dosed with gas species.

The wide applicability of powder diffraction makes it especially useful for studying flexible MOFs. As noted in Section 1, these materials can have large changes in the structure between a closed and an opened state. Often the flexible structure opens more easily for one gas over another, which is useful for many gas separation applications. For example the MOF Co(bdp) (where bdp^{2-} = 1,4-benzenedipyrazolate) attains a large CO_2 capacity while approaching complete exclusion of CH_4.[14] *In situ* X-ray powder diffraction shows that CO_2 adsorption is associated with expansion of the pores of this flexible MOF and that templating of the pores by a specific arrangement of CO_2 molecules upon adsorption must be invoked to explain the continued high selectivity of CO_2 over CH_4 at higher pressures (≈ 1500 kPa).

Another interesting use case for in *operando* X-ray powder diffraction is cooperative adsorption. An example is Fe-containing MOFs such as Fe$_2$Cl$_2$(bbta) and Fe$_2$Cl$_2$(btdd) which have Fe initially in the high spin state. As the material adsorbs the gas, the Fe acquires the low spin state, which a concurrent change in the bond lengths. As a result, more open metal sites are exposed, and the material can adsorb more gas.[15] This behavior leads to a stepwise adsorption curve that could be especially useful for gas-separation applications as a large change in uptake occurs for a small change in pressure.

X-ray and neutron powder diffraction were also used to confirm that two sorbate molecules could adsorb on a single metal site in the MOF Mn$_2$(dsbdc).[16] This could be a useful feature in the search for materials that can adsorb a large amount of CO_2.

2.3 Electron diffraction and serial diffraction

Although powder diffraction is useful when sufficiently large single-crystals are not available, a major problem is that the information in the peak positions and intensities may not be enough to determine the unit cell, let alone to refine the structure, especially in the case of large low-symmetry unit cells present in many MOFs. 3D electron diffraction (3D ED) methods can help. Electron diffraction does not require large samples: single-crystal electron diffraction data can be obtained for crystals that are only tens of nanometers across – much smaller dimensions than can be probed even with microfocused synchrotron X-ray beams. The 3D ED data can be obtained using different techniques, such as automated diffraction tomography, precession-assisted electron diffraction tomography, continuous rotation electron diffraction, and microcrystal electron diffraction, which have become increasingly popular for studying the structure of porous materials.[17] Electrons have a much stronger interaction with the samples than x-rays, making it possible to acquire strong signals from light elements like Li and H, albeit at the expense of added complexity caused by multiple scattering events (dynamical scattering). Because of this complication, electron diffraction was of limited use for structural refinements until relatively

recently. The main operating principle of 3D ED is to acquire a series of diffraction patterns as the sample is rotated along high-index crystallographic directions and then refine the structure from the whole set.[18] In this way, the structural refinement can rival or even exceed the accuracy of X-ray methods due to the sensitivity of ED to light elements; this is especially true when using dynamical refinement techniques. There are various data-collection schemes including continuously rotating the sample as the images are acquired, which minimizes exposure time and therefore beam damage. The choice of the optimal strategy and data collection mode depends on the sample. For example, nanoparticles with a size of a few tens of nanometers, especially when agglomerated, are best investigated using scanning transmission electron microscopy (STEM), combining the acquisition of electron diffraction data with crystal-tracking routines[19] as shown in Fig. 2.

Together, these factors make 3D ED techniques increasingly popular for studying structures of porous materials. Combining automated electron diffraction tomography with X-ray powder diffraction or spectroscopic methods, like solid state nuclear magnetic resonance (NMR), is particularly powerful, providing invaluable insight into property-guiding structural details.[20] Many MOF crystal structures have been already solved from microcrystalline powders, such as MOF Zr-CAU-30 (ref. 21) or Zn-CAU-37, where even the structural evolution from as-synthesized to activated material was revealed.[22]

Another way to get the full single-crystal diffraction information from powder samples is with serial diffraction. In this technique a stream of powder particles is flowed into the diffraction beam. This technique captures single-crystal patterns from a large number of crystallites in the stream and uses the cumulative information to solve the structure. This approach avoids the need for mounting and orienting single-crystals, but requires novel sample-delivery methods. Not only does serial diffraction allow for the study of small crystallites, but because of the flow of particles it is an especially useful technique for studying the dynamics of structures. The diffraction measurements can be performed at a time delay from the application of a change in the environment, therefore the time resolution is only limited by the photon dynamics. Furthermore, since the powder particles only pass through the beam once, it is possible to use serial diffraction to study

irreversible processes. The high brightness at beamlines such as ChemMatCARS at the Advanced Photon Source upgrade (APS-U) are projected to provide strong enough signal to enable such serial diffraction techniques.[23]

2.4 Quantum crystallography

Quantum crystallography is a technique that – rather than simply refining the atomic positions as is done in typical crystallographic structural refinements – performs the task iteratively by fitting the wavefunctions derived from the molecule's structure, thereby achieving accurate structure factors. This process involves continuous adjustments of the molecule's structure until the calculated wavefunction and the updated model structure factors agree. This represents a dynamic interplay between experimental data and theoretical modeling, leading to a comprehensive understanding of the atomic and electronic structure of molecules within crystals.

Typical crystallographic structural refinements assume spherical atomic structure factors that are independent of the structure (in the so-called Independent Atom Model (IAM)). Quantum crystallography forgoes these approximations; instead it directly refines the electron density allowing for aspherical structure factors that can be compared with quantum mechanics calculations.[24] This technique requires collecting diffraction data with high accuracy, high resolution, and a strong signal to noise ratio. Beamlines such as NSF's ChemMatCARS at the APS-U are projected to have unrivaled brightness and beam coherence thus enabling quantum crystallography.[25,26] Quantum crystallography could be useful for any system where the assumption of atoms as spherical scattering centers is too restrictive – materials with, for example: strong covalent or hydrogen bonds, magnetic moments, and filled f-shells. Porous materials, especially those with guest molecules, could exhibit these effects as exemplified in the spin-state transition of Fe in Section 2.2, and strong distortions (leading to aspherical electron clouds) as discussed in Section 3.1.

3. Structural modeling

Some porous materials cannot be grown as large single crystals and have structures or structural changes that are too complicated to be solved by powder diffraction alone. Other porous materials have substitutional disorder or are even amorphous. As complicated as such systems are, they expand the universe of potentially useful gas adsorption materials. This section discusses the state of the art in structural determination and a description of these systems. Ultimately, of course, materials to be used as DAC sorbents on a commercial scale will need to be produced reliably in large quantities, even if single crystals cannot be grown. Thus, development of methods that combine powder diffraction with other characterization methods and with computational approaches such as DFT, as well as *in situ* measurements and predictions of structural change and performance, will become critical in determining and selecting the optimal systems for DAC and CO_2 reduction or removal (CDR) on a commercial scale.
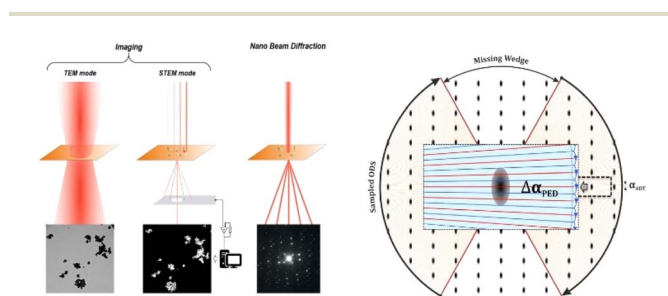

Fig. 2 Settings of the FAST-ADT data acquisition module including nano beam diffraction, TEM/STEM imaging and dedicated crystal tracking routines.[19]

3.1 Combining experiments and density functional theory

In many structural models of porous materials, the component molecules – the secondary building units (*e.g.*, the metal clusters and organic linkers in MOFs) – are treated as rigid bodies. Some flexible materials can be treated this way by modeling the framework of a MOF (or similar) as a flexible mechanical truss.[27] But what if the secondary building units themselves are flexible? One approach is to start with a structural model generated from rigid building units and then use density functional theory (DFT) to relax the atomic coordinates.

DFT structural optimization is a particularly powerful tool in conjunction with powder diffraction on large-unit-cell MOFs that are difficult to grow as single crystals. Powder diffraction can provide the lattice parameters and suggest the topology of the MOF, while DFT can provide finer structural details. This fact has been demonstrated with the case of crystalline metal-coordinated cavitand crystals, which shatter upon removing the solvent, but whose unit cell reverts back to the original lattice parameters when the solvent is re-introduced.[28] Since the transformation of the lattice is reversible, it was inferred that the connectivity of the underlying structure is maintained after the shattering. Refining structural models from powder XRD data using the DFT-based structure optimizations was able to confirm that there was a volume change of 30%, and that the ligands rotate for the material to become denser. This volume change can then explain earlier observations of stepwise CO_2 adsorption behavior. In a similar study of structural changes in Ni-Bpene upon CO_2 adsorption, DFT calculations confirmed the hypothesis of staged ligand rotation as CO_2 is adsorbed and the pores become larger.[29]

Similarly, DFT-assisted structure refinement can be used to study other flexible porous materials such as hydrogen bonded frameworks and interdigitated molecular crystals.[30,31] These materials are highly flexible, and using powder diffraction to determine unit-cell distortions followed by DFT relaxation of atomic coordinates, validated against the powder XRD data, shows that molecular building units are highly distorted from their shape in isolation. These materials are typically amorphous in a solvent but crystalize through solvent evaporation. However, using different solvents for this process can produce different structures despite starting from the same precursors. Furthermore, reversible transitions between structures are possible by crystallizing using one solvent, then dissolving and recrystallizing from a different solvent. DFT-based structural refinement is a powerful technique for studying these highly flexible materials, which are gaining interest as potential NETs due to their CO_2 adsorption properties.[31,32]

3.2 Mixed-metal structures

Mixing metal species in otherwise identical MOFs can provide an additional degree of control over the structure resulting from the synthesis. Some possible outcomes are uniform mixture, complete segregation, partial segregation, discrete domains, and chemical ordering. Beyond these familiar types of ordering, mixing metal species can indirectly lead to changes in the microstructure. For example, a variety of such effects occur in the Zn-PF1 system with the partial substitution of various species for Zn.[33] Substitution of 50% of Zn with Mn leads to Zn–Mn ordering, while substitution of Zn by Co results in a characteristic core–shell arrangement. The latter is thought to be due to preferentially dissolving the initially formed Zn-based crystals, then recrystallizing on the outside of structure as more linker is available in the solution. Mixtures of La and Yb-based RPF-4 have a similar re-dissolve mechanism that leads to the crystal shape with holes when lanthanum is present in high concentration. If starting from equimolar La : Yb mixture, bulk measurements show an equiatomic ratio of La to Yb, whereas the microanalysis using energy dispersive X-ray spectroscopy (EDX) reveals the system to crystallize into the two end members with molar ratios of 1 : 9 and 9 : 1 La to Yb. However, by changing the linker to one with higher steric bulkiness, such as the carborane based linkers, can allow for a homogeneous mixture of up to 8 different rare-earths.[34]

Most mixed-metal MOFs are synthesized by creating a single solution of the metal components and organic linker with the solvent. This synthesis route assumes that the framework of the MOF will form with the desired composition, but this is not always the case. An alternative synthesis method is to use a pre-formed heterometallic molecular complex as precursor for the MOF structure. Using heterometallic complexes to synthesize the MOF can provide a higher degree of control over the structure. For example, it may be possible to create a specific pattern of the arrangements of the different metals. Starting with a metal–organic complex having the metals arranged with a specific ratio, a MOF can be created by a ligand to linker exchange reaction (see Fig. 3). This synthesis pathway was successfully used to create MIL-69 with a Ga to Ni molar ratio of 7 : 1.[35] Despite the metals having different valence states (Ga^{3+} and Ni^{2+}), these cations occupy the same site in the structure.

3.3 Amorphous structures

While most of this report emphasizes porous crystalline materials, related amorphous materials can also maintain porosity and adsorb gasses such as CO_2. Crystallographic structure refinement cannot be applied to amorphous materials. However, the short-range order in the latter can be described using the pair distribution function (PDF), which can be determined experimentally. Conceptually, the geometric PDF can be constructed by picking a central atom, then counting the number of atoms as a function of distance away from this
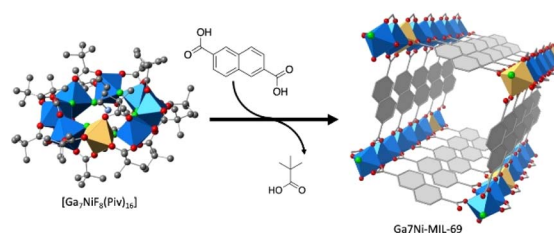

Fig. 3 Molecular diagrams showing a ligand to linker reaction to form a mixed-metal MOF. Adapted from ref. 35 with permission from the American Chemical Society copyright 2022.

25682 | *J. Mater. Chem. A*, 2024, **12**, 25678–25695

This journal is © The Royal Society of Chemistry 2024

central atom, repeating this process with each atom as the central atom and summing the results. Experimentally, the PDF is determined as the Fourier transform of the scattered intensity measured over a wide range of scattering vectors. The PDF description applies to all materials – crystalline, amorphous, and even liquid – allowing for a direct comparison between crystalline and noncrystalline structures.

One example of the utility of the PDF is its application to describe the structural transitions in ZIF-4. The base ZIF-4 structure on heating first transforms to an amorphous phase and then to a ZIF-zni phase. The PDF of ZIF-4 and the intermediate amorphous structure are nearly identical at short distances. This implies that the ligand coordination of the Zn ions is preserved in the amorphous phase which contains a continuous random network of Zn tetrahedrally coordinated by imidazolate linkages.[36] This behavior is similar to how the SiO_2 tetrahedra are preserved between quartz and glass, the difference being that the network is crystalline for quartz and amorphous for glass.

The information from the PDF can provide insight into how to improve the adsorption behavior. For example, knowing that the local structure is preserved upon melting ZIFs allowed for the development of ZIF-MOF composites that have high CO_2 capacity. Because its local structure is preserved, ZIF-62 could be melted to form the matrix of a composite with MIL-53(Al), a flexible crystalline MOF very similar to MIL-53(Cr) described in the Introduction. The interaction of MIL-53(Al) with the local structure of the ZIF-62 material stabilizes the large-pore configuration of MIL-53(Al) over the narrow-pore configuration, leading to a higher CO_2 capacity than for MIL-53(Al) alone.[37]

The PDF can be constructed for any type of matter, whether – crystalline or amorphous, and therefore could serve as useful feedback for MOF synthesis. As mentioned in earlier sections, the synthesis of crystalline MOFs can be challenging. XRD measurements typically only provide binary feedback on the results of this synthesis, while PDF might be able to provide more granular feedback by showing how the local structure of the product compares to that of the target material.

4. Property measurement and modeling

Sections 2 and 3 discussed the structure of nanoporous materials for carbon capture. This section discusses many of the performance issues related directly or indirectly to the structure of porous materials for CO_2 capture: stability, competitive adsorption, thermal transport, and CO_2 conversion, and how these properties can be measured and modeled.

4.1 General considerations

The material selection for carbon capture applications relies on incomplete information, often using high CO_2 uptake or working capacity as an indicator of performance. The performance of a sorbent for DAC, however, is not solely determined by the equilibrium uptake and working capacity. A fuller view of the adsorption behavior must consider the competitive adsorption of CO_2 with the other species in the air: N_2, O_2, H_2O, *etc.* As many sorb/desorb refresh cycles involve heating the system, the temperature dependance of competitive adsorption is also important. Further, all of this information only considers thermodynamic equilibrium and provides no information on the kinetics – which is important for industrial DAC facilities to determine cycle times.

The scale-up and commercialization of sorbents for DAC depends further still on factors beyond even this comprehensive view of the adsorption behavior. The synthesis and what reagents are required for a given sorbent material must be considered. How scalable is the synthesis of the precursors, such as the metal–organic clusters or organic linkers? The stability of the structure – especially in the real-world conditions that include H_2O and possible contaminants – will affect the viability of the sorbent. Many DAC facilities use gas contactors that consist of pelletized beds. Others use macroporous supports that are decorated with sorbents. In any case, the manufacturability of the sorbent into needed form is also a selection criterion. Organizations conducting techno-economic assessments of sorbents for DAC must consider each of these factors.

One concern limiting the commercialization of materials like MOFs is the lack of information resources that provide insight into the criteria outlined above. Much of the current effort in this field of research is involved in developing new prediction tools, and new databases to assess the performance and viability of sorbent materials. Crucially, since many of the criteria center around scalability and real-world conditions, there are ongoing efforts to develop new experimental resources which can provide context to the modeling predictions or assess the criteria directly.

4.2 Stability

A basic requirement of a porous solid for practical carbon capture is that it be thermodynamically and mechanically stable. Machine learning (Section 5) was used to investigate the thermodynamic and mechanical stability of a database of 50 000 hypothetical MOFs, finding previously unknown structures with good stability.[38] A website has been developed to disseminate the ML MOF stability data and allow the user to harness these data to make predictions on the stability of new MOFs.[39]

4.3 Sorbent performance

The National Energy Technology Laboratory (NETL) is developing a DAC test center. This facility enables the performance characterization of DAC materials and systems at: lab-scale (material testing 0.1 kg per day of CO_2), bench scale (testbed reactors at 4 kg per day of CO_2), and pilot facility scale (test DAC facilities at 30 kg per day of CO_2). This facility allows for the evaluation of many types of regeneration cycles. The pilot facility scale testing allows for the testing of DAC modules under different environmental conditions.

Using fibers as sorbents allows for a wide variety of form factors and designs of the DAC system. For example, the

This journal is © The Royal Society of Chemistry 2024

J. Mater. Chem. A, 2024, **12**, 25678–25695 | **25683**

polymer PIM-1 can be tagged with amines to create the PF-15-TAEA fiber sorbent. This can then be used in an electro-spinning process to be formed into a flexible mat. The solubility of the fiber material opens the possibility of wet-spinning or 3D printing the material into various form factors. In particular, the PF-15-TAEA material shows very good regeneration.[40] This is therefore a material that shows promising DAC performance and can be readily fabricated into contactors with high surface area.

Beyond directly adsorbing the gas into sorbent materials, one way to capture CO_2 from the atmosphere is by using membranes that are selectively permeable. The polymers used in these membranes can be highly selective for CO_2, however their permeability is very low limiting their performance. One approach to resolve this is to add MOFs to the membranes to increase the permeability forming Mixed Matrix Membranes (MMMs).

Computational force field modeling has been used to screen over one million candidate MMMs involving two thousand different MOFS, to predict the cost per metric ton for CO_2 removal.[41] Promising materials have been identified that are predicted to combine high permeance, high CO_2/N_2 selectivity, and low cost per metric ton of CO_2 removal. It has been found that force fields developed to account for MOF framework flexibility do a better job of predicting adsorption performance than force field models that keep the framework rigid.[42] Framework flexibility also means that porous materials with pore-limiting diameters (PLDs) of the average crystallographic structure that appear too small for gas sorbates to diffuse may allow for diffusion in reality. Thus high-throughput computational screening of candidate porous adsorption materials based solely on average crystal structures should not eliminate materials with PLDs that are close to the target sorbate molecule (*e.g.*, CO_2) size and would otherwise be considered too small.

Zeolites are not as widely explored for carbon capture as MOFs but used in practice for many separation applications, and similar force-field modeling approaches are applied to investigate their gas adsorption barrier. The effect of framework flexibility on sorption properties in zeolites was found to be generally smaller than that for MOFs.[42] On the other hand, DFT-based molecular dynamics simulations show again that the flexibility of the windows in zeolites can be crucial to explain sorbate diffusion when these windows are barely large enough for molecules to pass through.[42]

Zeolites and MnO_2 octahedral molecular sieves are examples of porous materials that may contain additional ions in the pores. DFT studies predict that changing the size or charge of the ionic species will change the nature of CO_2 adsorption and interaction with the ions.[43]

Many reports focus on discovering new materials that further optimize geometric descriptors like the gravimetric surface area or pore volume, which are viewed as a surrogate for high CO_2 uptake that itself is a surrogate of performance. In order to selectively adsorb a target gas species (*e.g.* CO_2), a material should have the appropriate PLD and largest cavity diameter (LCD). This design criterion can be challenging to satisfy, however, as any solvent molecules used during the

synthesis can become trapped in the pores. This is one of several factors that make Al-formate an attractive candidate sorbent material for DAC: during the synthesis, $Al(OH)_3$ reacts with formic acid to form the Al-formate – *i.e.* the solvent becomes the linker in material. Al-formate has a PLD of 1.97 Å, which is near ideal for adsorbing CO_2. Meanwhile, the N_2 adsorption isotherm for Al-formate is similar to that of non-porous materials. Because of the flexibility of the Al-formate structure, the CO_2 adsorption at 270 K is low, but there is a gate-opening transition allowing for high adsorption at 297 K.[44] Furthermore, as seen in the O_2 adsorption, Fe substitution for Al in the Al-formate structure can increase the kinetics.[45] Another major attraction of Al-formate is the cost. Due to the relative abundance of $Al(OH)_3$ and formic acid and relative ease of synthesis, Al-formate based sorbents for DAC are an economical option and could even form the basis of Al-based carbon cycle with a CO_2 packing efficiency 96.35% of dry ice.[44]

Some flexible MOFs have pores that are open in the presence of the solvent, close as the solvent is removed, but open again under the pressure of certain adsorbates – a behavior that leads to a highly selective, stepwise adsorption. These flexible MOFs can form nanotraps for specific guest molecules, acting as a gate for the particular adsorbate. Such nanotraps often interact with the guest molecules through hydrogen bonding, so that specific gas species will cause the gate-opening while others do not. This mechanism can provide high selectivity in very challenging separations. For example, in the separation of C_2H_2 from CO_2 – important for flue gas capture applications – a hydrogen-bonding-based nanotrap was shown to have very high selectivity and capacity for C_2H_2 as determined by *in situ* infrared spectroscopy measurements.[46] Methane (CH_4) is a much more potent greenhouse gas than CO_2,[47] so CH_4 separation from N_2 is an important application. However, a strong interaction is needed for extracting CH_4 from air. The flexible MOF ATC-Cu with hydrogen-bonding-based nanotraps for CH_4 was found to adsorb 2.90 mmol g^{-1} of CH_4 at 298 K and 100 kPa, with a significant selectivity for CH_4 over CO_2.[48] A combination of single-crystal X-ray diffraction and DFT calculations was used to determine the preferred CH_4 adsorption sites and uncover principles for engineering suitable pores.

4.4 Thermal transport

Back when MOFs were being considered for natural gas storage in vehicles, studies identified MOFs that were able to store large amounts of CH_4, but these materials have yet to be used industrially. One of the major issues that was largely overlooked at the time was heat transfer. To be practical, refueling of cars should take minutes. But the adsorption of CH_4 to the MOFs is exothermic.[49] Therefore, refueling was limited by the thermal conductivity, requiring hours to fill the tank.

The same issue is present for sorbents for DAC. Industrial scale DAC facilities need to be able to adsorb and desorb quickly, so there is a need for sorbents with high thermal conductivity. However, there are still many open questions about how to model heat transfer in MOFs. Are phonons meaningful carriers? The unit cells of many MOFs are on the

25684 | *J. Mater. Chem. A*, 2024, **12**, 25678–25695

This journal is © The Royal Society of Chemistry 2024

same length scale as the thermal phonon wavelengths. How do pore size and shape matter? How is the thermal conductivity affected by gasses? In classical atomistic simulations, techniques such as equilibrium molecular dynamics can extract thermal conductivity from autocorrelation functions, while non-equilibrium molecular dynamics can apply thermal gradients and extract the thermal conductivity from the heat flow. However, unlike most materials (including clathrates which have trapped gasses), the guest molecules in MOFs can adsorb, desorb, and diffuse through the material. The ambiguity of how to model MOFs can lead to disagreements on seemingly simple questions, such as whether MOFs become more or less thermally insulating with more adsorbed gas. One group found that more gas led to more phonon scattering and therefore more thermally insulating behavior.[50] A different group found that adding the sorbates adds new thermal pathways that leads to more thermally conducting behavior.[51] The prevailing thought is that both of these terms are active, and it is highly system dependent as to which term dominates. Experimental studies on HKUST-1 show a decrease in thermal conductivity as various liquid adsorbates are incorporated.[52] The effects of defects on thermal conductivity in MOFs has also been explored. Surprisingly, the modeled thermal conductivity of UiO-66 increases for certain hypothetical arrangements of missing linker defects.[53]

4.5 Catalysis

Directly adsorbing CO_2 from the atmosphere is fundamentally challenging. Local electric fields in porous materials can readily attract polar sorbents such as H_2O. However, CO_2 and many other gas species in air, including N_2 and O_2 are non-polar but polarizable. Among other requirements, sorbents for CO_2 direct air capture *via* physisorption must rely on separating it from other non-polar gasses by their differences in polarizability.[54] Because of this basic challenge, alternatives to adsorbing CO_2 might be considered. Instead of simply capturing CO_2 from the air for subsequent storage, an alternative strategy is to chemisorb CO_2 and directly catalyze it into a more useful feedstock. CO_2 reduction reactions can be distinguished by the number of H^+/e^- as well as by the reaction potential. One promising strategy is to use a MOF as a chromophore and source of metastable hydrogen, accessed by photo-activation of the MOF in the presence of a H-atom source. Generally, the process is thought to "photodope" the framework by generating an electron/hole pair with sufficient energy to hydrogenate a substrate but not form H_2. In this regime, CO_2 reduction products may then be controlled the number of H^+/e^- available for the reaction, which should be proportional to various aspects of the photo-irradiation process. Although not relaying the hydrogen atoms onwards to reduce CO_2, the proof-of-concept of using MOF cluster nuclearity as a platform to control H-atom concentration was presented in MIL-125. In an initial report, Mayer and colleagues suggested the MOF could deliver two H^+/e^- pairs through proton-coupled electron transfer reactions for hydrogen evolution in the presence of 2-propanol.[55,56] However, subsequent theory-led studies revealed that MIL-125 could exceed ($4H^+/e^-$ pairs), and the limit was dictated by the

frameworks instability to forming H_2 rather than existing as a metastable H-atom source.[57] Ignoring the difference in electrochemical potential required to perform the hydrogenation, this result suggests use of MIL-125 as a catalyst for CO_2 reduction would not select for reaction pathways by H-atom source, but rather by MOF-cluster nuclearity. Thus, to then reduce CO_2, one promising pathway is to couple the H-loaded MOF to a conventional catalyst such as transition metal nanoclusters (*e.g.*, Pt or Cu), and limit the product distribution by irradiation time of the MOF. In this case, the excitons can migrate to the active metal catalyst, achieving higher conversion in the photo-limited H-atom bath.[58]

5. Machine learning applications in porous materials

Machine learning (ML) techniques have a powerful potential to accelerate science, especially in the field of porous materials and the nascent field of DAC. At the most basic and general conceptual core, ML techniques treat all problems as the search for a mathematical function that takes some input domain and makes prediction in the output domain. The key to applying ML is in casting the problem at hand into this framework. There are myriad ways that ML can be applied to DAC issues. For example, ML can be used to predict the adsorption behavior of new materials. The space of possible MOFs is very large, and ML models can make predictions quickly and can therefore be used as screening tools to identify MOFs of interest. Additionally, ML can be used to compile and organize the state of the assembled body of knowledge in the field, *e.g.*, by identifying correlations between feature in a database or mining the literature for research topics. These efforts can help accelerate adsorption science by identifying gaps in the knowledge base or prioritizing avenues of study. ML can also help accelerate the study of structural transformations in porous materials by generating accurate force field models that are much faster than DFT calculations.

A word of caution in the use of ML models especially in the field of porous materials for adsorption applications: any ML model trained on a database will inherit the biases of that database.[59] There is a strong positive result bias in the literature.[60,61] For example, it would be difficult to accurately predict the structural stability of new materials if only stable materials are reported, without any examples of unstable materials to learn from. Furthermore, since even modern large language models LLMs struggle to connect different ways of reporting information in the material science domain,[62] there is great need for the standardization of reporting adsorption research. For example, an ML model mining the literature would have difficulty accurately determining the context of adsorption curves if they are inconsistently reported – *e.g.*, one paper might include the measurement conditions in legend of a figure, while another might state elsewhere in the manuscript that all measurements were performed at a certain relative humidity. It is important, then, to understand and keep in mind the biases and limitations of the ML models used in their application to adsorption science.

5.1 Predicting adsorption

ML tools can be used to directly predict the adsorption behavior of materials. One example is using ML models to narrow the search space of materials to those that warrant further study by predicting performance measures. Starting with the QMOF database of MOF materials,[63–65] Steckel et al.[66] trained a ML model to classify whether MOFs have high or low uptake of CO_2 at 400 µmol mol^{-1} CO_2. While the predictions were sensitive to the exact type of ML model used, the featurization was found to have higher importance. In addition to the commonly used material descriptors such as smooth overlap of atomic positions (SOAP),[67,68] the revised autocorrelation functions (RACS)[69] method and stoichiometric features, the inclusion of density and surface area dramatically improved the ML's ability to classify the CO_2 uptake.

One early concept was to use atomic property weighted radial distribution functions (AP-RDF) as descriptors for MOFs, then train support vector machines to classify them by whether their CO_2 uptake is above or below a given threshold.[70] This workflow functions as a high-throughput screening protocol to identify sorbents of interest. Later work used gradient boosted decision trees (GBDTs) to predict the uptake as well as the selectivity. GBDTs are particularly well suited to including many different descriptors, so AD-RDFs could be used along with geometric features (e.g. surface area, density, void fraction, and pore size) resulting in an improved screening tool.[71]

Once ML models are trained to predict multiple properties of interest, they can be incorporated into larger workflows that seek to discover new, high performing materials. For example, Zhang et al. used Monte Carlo tree search to identify MOFs that not only had high CO_2 uptake, but were also hydrophobic.[72] Similarly, Kim et al. used their ML models in a Bayesian active learning framework.[73] In the search for mechanically stable zeolites, this framework iteratively predicts the bulk and shear moduli of zeolites, adds those to the database, and suggests the next materials to study and repeats this process. A recent review covers the uses of ML methods in CO_2 capture, storage, transportation, and utilization, and provides an outlook as to the impact of such ML methods in the future of this research.[73] ML models can also be used to improve the accuracy of predictions of sorbent materials and their interactions with adsorbates such as CO_2. ML models have found use as surrogate functions for force fields (FF) in molecular simulations. Even some commercial DFT packages are now implementing tools to train ML-FFs on the trajectories from DFT geometry optimizations.[74–76] ML force fields fit on DFT data have been used to correctly capture the negative thermal expansion of some zeolites with higher accuracy than available classical force fields (see Fig. 4). ML-based force fields have also shown excellent agreement in energy and forces between DFT and the ML-based FF simulations for both MOFs and zeolites.[79] Recent work has demonstrated the ability of ML-based FFs to predict adsorption isotherms[80] and diffusion coefficients[81] for CO_2 in MOFs using Monte Carlo and molecular dynamics simulations respectively.
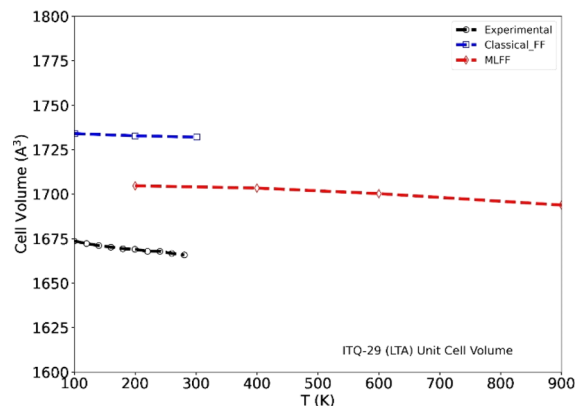

Fig. 4 Unit-cell volume as a function of temperature for a pure-silica LTA zeolite demonstrating negative thermal expansion; experimentally measured results[77] are compared with simulations using a classical force field[78] and an ML-based force field.[74–76]

5.2 Database curation and literature mining

Another way to use ML is to construct maps of the current state of the research on porous materials. MOFs can be thought of as constructed from secondary building units (SBU's, i.e. metal–organic blocks and organic linkers) on a certain topology. The abstract mathematical space of combinations of SBU's and topologies has only been sampled very locally with vast regions of this space still under-explored. Geometric fingerprints such as the maximum included sphere, and gravimetric surface area are important, but those descriptors are insufficient as they do not encode the chemistry. Revised autocorrelations consider products and differences on the molecular graph to capture the chemical information. The RACs are constructed from graph edit distances starting at, for example, a metal site and considering a set number of bonds away. In this way, RACs build up features that describe the local chemical environment. RACs could be constructed by starting at a single metal atom or aggregated across all atoms in the SBU. Moosavi et al.[82] combined the geometric fingerprints with the chemical RACs features to train ML models to predict the CO_2 working capacity. When trained on experimental databases (i.e. CoreMOF[83]), the chemical features were found to be more important, but when trained on hypothetical databases (i.e. BW-DB,[84] ToBaCCo[85,86]) the importance of the geometric features increased significantly. The RACs descriptors show that most hypothetical MOF databases severely under-sample the metal variety when compared to experimental databases.[38,82]

Under-sampling, or imbalance in databases of sorbents is a pervasive problem. The activation stability of a MOF relates to whether or not it could be activated (solvent removed without the structure collapsing) under heat and vacuum – in ML parlance this would be a classification problem. The thermal stability of a MOF can be characterized by the temperature of degradation of the MOF structure – a regression problem. Large language models (LLMs) could be used to mine the literature to gather information about the activation stability of MOFs, but the literature tends to under-report MOFs that could not be

25686 | J. Mater. Chem. A, 2024, 12, 25678–25695

This journal is © The Royal Society of Chemistry 2024

activated and activation protocols that didn't work. Further complicating the task for LLMs is the inconsistency of naming conventions – the MOF HKUST-1, for example, has 5 different names in the literature: HKUST-1, Cu-BTC, Cu_3BTC_2, MOF-199, and Basolite™ C300. The naming problem can be circumvented by mining structure databases (*i.e.* CoreMOF[83]), and cross-referencing against databases that map structure to manuscripts (*i.e.* CSD[87]), then finally using the LLMs to mine the manuscripts. Despite these difficulties, Nandy *et al.*[39,88] could use a similar workflow to test the generally accepted wisdom that "Zr MOFs tend to be stable".[89] They found that while many Zr MOFs are stable, there was a strong dependence on the number of Zr atoms in the SBU as well as the linker identity (see Fig. 5). The ML models trained on the stability data with the geometric and chemical (RACs) features gave insights as to how to engineer stability into MOFs. Swapping functional groups, for instance, could improve the predicted stability. In one example, fluorinating the linkers improved the thermal stability of a Cu-based MOF.

In classical FF for MD simulations, it is difficult to accurately capture the interaction of the sorbent and H_2O. Yet, those interactions are vitally important for DAC in real-world conditions which include humidity. The alternative to inaccurate FFs for MD simulations has typically been computationally expensive DFT calculations. One way to address this problem is with ML models for FFs trained on DFT calculations, as mentioned in the previous subsection. In order to fully realize the efficacy of such ML FFs, a large database of DFT calculations is needed. To that end, the OpenDAC database[90,91] was created, which consists of 40 million DFT calculations of total energy and atomistic forces. These data points were created by taking snapshots of DFT relaxation trajectories of MOFs with added sorbents. The OpenDAC database considers experimentally validated sorbent materials, as collected in the CoreMOF database,[83,92,93] and their interactions with both CO_2 and H_2O. This OpenDAC database enables the training of ML models that
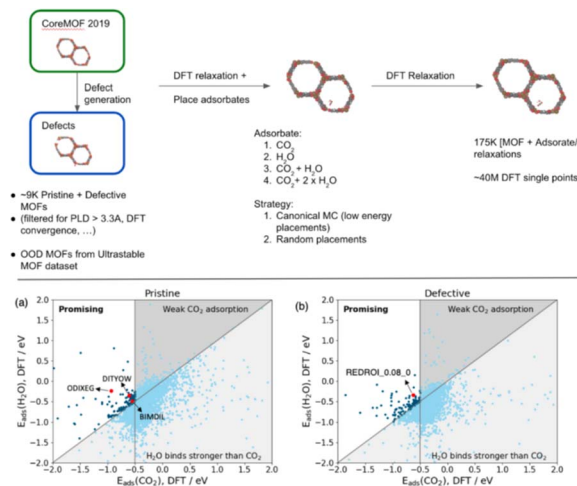

Fig. 6 (Top): The OpenDAC 23 dataset was created from MOFs taken from the CoreMOF dataset, which was expanded by introducing missing linker defects. One or more adsorbates were then placed in the pores of the MOF and a DFT relaxation was performed. The final dataset contains nearly 40 million DFT single points. (Bottom): Over 240 MOFs in the OpenDAC dataset were found to be promising for direct air capture. MOFs with CO_2 binding energy <-50 kJ mol^{-1} and with a stronger binding to CO_2 than H_2O were considered promising. Adapted from ref. 90 with permission from the American Chemical Society copyright 2024.

predict the atomic forces and total energies from the 3D atomic positions. One example model is a Graph Neural Network (GNN), where each atom is a node and bonds to each neighbor is an edge, that uses message passing between neighboring atoms to predict the atomic forces and adsorption energies from the graph embedding.[94–96] Because the model was trained on the OpenDAC database, and predicts forces and energy from atomic arrangements, it can be used as the FF to perform structure relaxations and then predict the adsorption behavior for materials that are not in the OpenDAC database – the ML FF generalizes to other materials well. This approach can therefore be used to study how defects in MOFs, such as missing linkers, affect the adsorption behaviour (see Fig. 6).

6. Industrial considerations

Industrial participants at the workshop brought up practical issues in using materials for CO_2 reduction or removal. While not all of these issues are directly related to the structure determination problem, people working in this field should be aware of them.

The field of NETs such as DAC is moving rapidly and there is a great need for standardization. Direct air capture does not necessarily mean using a single material to remove CO_2 from the atmosphere in one step – it may prove more practical to use multiple materials and stages in the removal process. For example, how to handle humidity in the input stream is still an open question: whether to dry the incoming gas in a first stage or use sorbents that preferentially adsorb CO_2 over H_2O. There is also rapidly developing research around regeneration cycles.
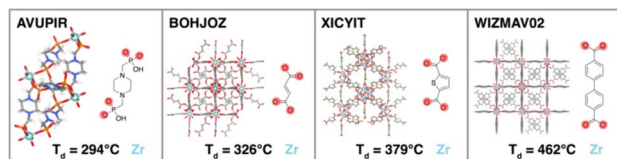

Fig. 5 Four Zr MOFs from the extracted thermal stability dataset and their corresponding thermal decomposition temperatures (T_d) shown inset. The MOFs are ordered from least to most thermally stable, with increasing T_d from left to right, with refcodes from the Cambridge Structural Database (CSD) shown for all. Within each MOF, the linkers and metals making up the MOF are shown, with the coordinating oxygen atoms coloured by translucent circles. Atoms in MOFs are coloured as follows: white for hydrogen, gray for carbon, blue for nitrogen, red for oxygen, orange for phosphorus, yellow for sulphur, and light blue for zirconium. BOHJOZ is commonly known as MOF-801, XICYIT is commonly known as DUT-67, and WIZMAV02 is commonly known as UiO-67. Reprinted from ref. 88 with permission from the American Chemical Society copyright 2021. Further permissions related to the material excerpted should be directed to the ACS.

Traditionally, regeneration was achieved with cycles based on a temperature-swing, a pressure- (or vacuum)-swing, or both. However, using steam for regeneration, which takes advantage of the competitive adsorption of H_2O over CO_2 to drive off the CO_2, has lately gained traction in the literature. Recent studies have also considered electro-swing regeneration cycles. In the related field of carbon capture and storage there are recent proposals on the use of supersonic input streams to address similar gas separation challenges.[97,98] To date, there has been much more industrial adoption of chemisorption despite being much more structurally and chemically complex than physisorption. As there are still ongoing debates about facility and contactor design and which materials are likely to be commercially viable for various applications, the standards developed in the field of NETs should be neutral to and robust in respect to these specific design considerations and rapid technological developments, and they should focus instead on reproducibility.

7. Outstanding issues

This section discusses the outstanding issues in structure determination of porous materials for carbon capture, and the measurement science and standards development that is needed to resolve the remaining bottlenecks, as discussed in the workshop discussion sessions.

7.1 Structure measurements and reporting

The single-crystal community has standardized the Crystallographic Information File (CIF) format for reporting the structures of materials. However, while there have been some standardization efforts for the case of powder diffraction, as for example, pdCIF[99] and the powder diffraction file,[100] these have not yet gained widespread adoption in reports in the literature, and there are no similar efforts for total scattering data (Bragg peaks plus diffuse scattering) or its Fourier transform, an atomic pair distribution function, which encode information about the local and nanoscale atomic order.

One important detail in crystalline sorbent materials is where the guest molecules reside within the structure. Particularly relevant guest molecules for materials for DAC are the solvents used in the synthesis, and the common gas species in air: CO_2, H_2O, N_2, and O_2. Standard crystallographic description may not be the best way to report the structural information of interest. For example, adsorbed CO_2 in a crystallographic structure report will list the position and occupancies of the carbon and oxygens separately. It may be more sensible to describe the CO_2 as a unit, but there is no standard procedure for this.

Another major limitation of current structure-reporting schemes is that often the structures are not clearly linked together. While common structure reporting formats like CIF do include fields for temperature and other conditions, these can be under-utilized, or insufficient. Relevant to sorbents for DAC, how best should structures be presented across complex thermal histories – like, for instance, repeated heating and cooling across several adsorption refresh cycles? this example would be relevant for studying the degradation mechanisms of sorbents for DAC. This issue of linking data is further compounded when considering flexible materials. One method for addressing it might be to report both an open and a closed structure. But how should deformation or flexural trajectories then be reported? Or structural degrees of freedom? For example, a MOF might have a linkage that is free to rotate, in which case there is not a point-to-point trajectory of structural changes in response to some external stimuli, but rather a closed path of structures about which the material can freely move. Such information is often lost.

Another ambiguity in reporting concerns specifying the method that was used to determine the structure and its sensitivity to various atomic species, and associated uncertainties. X-rays do not have a strong interaction with low-Z elements. So, if reporting the structure of a MOF from X-ray diffraction measurements, how were the positions of the H atoms determined? Were they assumed from a common bond distance, the result of some other measurement, refined from some starting position using the measurement at hand, or determined from a geometry relaxation using molecular dynamics (with what force-field?) or DFT (with what exchange functional)? It is important to be able to distinguish what information about the structure was experimentally confirmed *versus* predicted from theoretical calculations.

Thus far, this discussion has focused on the long-range structure of crystalline materials, but there is even more ambiguity in reporting short-range order or amorphous systems. Several important sorbents can have amorphous structures, yet there is no codified means of reporting what structural information is known about such systems.

It is common practice for the atomistic structure of intrinsically porous materials to be described in terms of geometric features such as the PLD, pore volume, and the gravimetric or volumetric surface area. These features serve as a generic proxy for diffusion rate or adsorption capacity for gas species. While some characteristics, such as the mass of the formula unit and unit-cell volume, are well defined, others, like the surface area, do not have rigorous mathematical definitions. One option is to forgo the use of these geometric features in favor of measured or predicted diffusion rates and capacities of the various gas species, which would more closely reflect the reality of the material behavior. The geometric features can still be useful (for example as descriptors for machine learning) if the method of calculating them is also reported. In order to determine the surface area of a porous material, some specific convention must be applied. This could involve picking a threshold of electron density, rolling a sphere of a certain radius around the structure, or approximating with the largest sphere that fits inside the structural voids – then reporting the area that results. Each method will produce a different value for the surface area, so it is important to capture the details of the method in the reporting. A future study should survey the common methods for determining the surface area and other geometrical features of materials and research how strongly they differ across the methods.

7.2 Larger-scale structure

Beyond the details of the atomic order, other characteristics, like structural defects and particle morphology, can have a strong impact on the performance of materials as sorbents. It is well known that zeolites tend to contain stacking faults,[101,102] whereas MOFs are more likely to incorporate point defects. It is still an open question how to best measure and report the defect structure in porous materials.

Different synthesis routes can yield different particle morphologies. In the mixed-metal MOFs discussed above, the particles generated by some synthesis methods were long cylindrical crystallites with an open channel through the axis, and compositionally graded in the radial direction. In composites of flexible MOFs with zeolite glasses, the interaction at the interface can hold the flexible MOF in the open structure configuration. The field needs a standard way of describing the shape, size, composition gradients, and connectivity of these sorbent materials.

In actual applications, there are many arrangements for how sorbent materials can be brought into contact with the gas mixture. In one approach, sorbents are pressed into a pellet and used as a molecular sieve in the gas stream. In this case, the particle packing in the pellet is critical to describe in order to understand how much gas might bypass the filter. Other designs have decorated the silica supports with sorbent materials, or used a flat support that the gas flows over. There is currently a lack of standardized description for the structure of the gas contactors and how different morphologies affect the gas/solid interactions of interest. A range of research methods exist that could elucidate this ambiguity of the larger-scale structure at the fundamental level, such as in situ small-angle X-ray and neutron scattering studies carried out under defined gas-sorption pressure and thermal conditions. Such measurements can help establish a quantitative link between microstructure and morphology, crystal structure, and sorption performance.

7.3 Adsorption measurements

Characterizing the adsorption behavior of a material for use in DAC processing suffers greatly from an absence of standards. Other adsorption applications often deal with separating gasses with molecular concentrations on the order of a few percent. However, since DAC involves CO_2 molecular concentrations of 400 $\mu mol\ mol^{-1}$, special care must be taken with the adsorption measurements to ensure the signal is not dominated by artifacts and that the measurements are comparable across laboratories. Adsorption measurements at these concentrations of CO_2 are very sensitive to considerations like the dead volume in the gas flow of the measurement apparatus. Recently, McGivern et al.[103] constructed a measurement apparatus designed specifically to ensure accurate adsorption measurements at these low concentrations of CO_2 relevant for DAC. Further complicating DAC relevant adsorption measurements is the competitive adsorption with H_2O. In real world conditions, H_2O is generally present in atmospheric air (typically at a molar concentration 10 to 100 times that of CO_2) and can be in competition with CO_2 for adsorption sites. This necessitates adsorption measurements as a function of gas mixtures with H_2O in order to realistically characterize adsorption behavior DAC.

One consideration that is often not captured with current methods of reporting adsorption behavior is the activation protocol that was used. After synthesis, porous materials like MOFs can have adsorbed solvent molecules left over from various synthesis steps. Exposure of MOFs to sulfur-based gasses can result in chemisorption of the sulfur, permanently modifying the adsorption sites.[104,105] Any measured uptake of CO_2 molecules into a material will depend on whether there were already CO_2 molecules, H_2O molecules, or in fact any guest molecules adsorbed into the structure. It is therefore common practice to activate the material by desorbing any guest molecules with some combination of heating the samples or pulling vacuum. Unfortunately, while such activation is common practice, it is not standardized and furthermore depends on the nature of the adsorbent material. In any case, the details are often not reported.

There are many measurement devices that can be used to determine the isothermal adsorption of materials, but there are few standards for reporting all the relevant information. Particularly important to capture are the activation protocols mentioned above, the details of the instrumentation used to measure the adsorption, and the conditions of the measurement. A recent report surveying isothermal adsorption measurements in the literature found that approximately 20% of reported data should be considered outliers.[106] At least part of this discrepancy was due to how humidity is recorded in the conditions: a base gas with a relative humidity (for example: a binary mixture of $CO_2 + N_2$ vs. the "binary" mixture of $CO_2 + N_2$ at 50% relative humidity as reported in ref. 107, or stating the effects of various target gasses of interest while noting the relative humidity of those elsewhere as in ref. 104). While reports specifying relative humidity in the conditions do provide the necessary information, using the relative humidity convention treats water vapor differently from every other gas species. As such, this convention obfuscates the fact that non-zero relative humidity means there is an additional gas species. One solution is to report the humidity with the absolute partial pressure, which describes water vapor on the same scale as all the other gas species. In addition to documentary standards, reference materials for isothermal adsorption are urgently needed. Particularly relevant to DAC is the study of adsorption at atmospheric conditions: 400 $\mu mol\ mol^{-1}$ of CO_2, in competition with H_2O, O_2, and N_2. Having reference materials would enable round-robin studies to compare these difficult measurements across laboratories and compare measurements of adsorption across measurement techniques. The reference materials for these measurements need not necessarily be among the most promising sorbents for DAC but should definitely be sorbent materials with well-known composition and structure that can therefore be expected to have repeatable behavior. As such, the authors suggest the zeolite 13X as a possible reference material for urgently needed round-robin studies.

This journal is © The Royal Society of Chemistry 2024

J. Mater. Chem. A, 2024, 12, 25678–25695 | 25689

Dynamic Column Breakthrough (DCB) measurements (also sometimes referred to simply as breakthrough measurements) are becoming increasingly popular to characterize the adsorption performance. These measurements reiterate the need for reference materials, with a standardized activation protocol, and a standardized packing or pelletizing protocol. As instruments for such measurements are often custom-built, the reporting of these measurements also needs to include information about the instrumentation design and type of contactor in order for the results to be comparable across laboratories. There are many ways that these materials might implemented in a DAC system (*e.g.* sorbents *vs.* membranes). Sorbents are typically optimized for strongly adsorbing CO_2, whereas membranes are typically optimized for minimizing the CO_2 adsorption and instead maximizing the CO_2 mobility. It therefore might be important to design and calibrate DCB instruments that can operate in both extremes. Furthermore, the fluid-dynamics equations of state or dimensionless numbers could enable the direct comparison of measurements across instrumentation design and across flow rates. DCB instruments can measure the adsorption uptake as well as the desorption. The DAC industry would therefore greatly benefit from round-robin studies of sorbents using DCB instruments at atmospheric gas concentrations and potential DAC-relevant refresh cycle conditions.

7.4 The role of machine learning

ML algorithms can and have accelerated sorbent research. However due to the aforementioned lack of standardization of the measurements, theoretical predictions, and the reporting thereof, the implementation of these algorithms is limited.

Furthermore, there are several concerns with the application of ML to porous material research. Many workflows use ML as a screening tool to filter out materials that are not of interest. These ML tools are by definition less accurate than the source of the data that were used to train them. ML algorithms are well known to extrapolate poorly, and the domain where the model is interpolating can sometimes be obfuscated making it difficult to know when the ML model is extrapolating. Given these accuracy limitations, how likely are false negative predictions? In other words, how likely will the ML model mistakenly filter out a high performing material? Care must be taken in evaluating ML models used as screening tools to characterize their false negatives, especially when extrapolating.

Another common pitfall with ML models is the bleeding of information from the training set to the test set. In the material sciences. This can happen, for example, when the test materials are too similar to the training materials. Here the meaning of "similar" depends on the context of what question the ML model is trying to answer. If the goal is to have an ML model that can generalize across MOF structures to predict some property, then the test set might leave out all materials with certain metal clusters, certain linkers, and certain nets in order to ensure that the test set is sufficiently distinct from the training set.

It is worth noting that ML models are not unbiased, they inherit the biases in the training data. For example, DFT struggles to accurately capture the effects of van der Waals and hydrogen bonding. Yet many MOFs and other porous materials primarily interact with the guest molecules through hydrogen bonding. Training an ML model on DFT data is therefore also likely to poorly capture the behavior of materials where hydrogen bonding is the dominant interaction. This issue can be compounded when the training data have heterogeneous bias or fidelity. Adsorption, for example, can be measured by many different styles of instrumentation as mentioned above – these data would have heterogeneous bias since each type of measurement would have its own sensitivities, artifacts, and biases. If the training database includes adsorption data from several of these instrument designs, a ML model might struggle to generalize if this heterogeneity is not explicitly accounted for.

As mentioned above, another major factor limiting the application of ML models (especially those mining the literature) is the lack of consistent reporting. In the field of MOFs and porous materials, perhaps the two biggest reporting issues are the names of the materials (as mentioned in Section 5.2), and the reporting of activation protocols (as mentioned in Section 7.3). Since there are no commonly accepted naming conventions for MOFs, many materials appear in the literature under several names. This can be confusing even for human researchers, and would require a major step in sophistication of LLMs without hard-coding the multiple names of materials. It is well known that the history of environmental conditions applied to a MOF (*e.g.*, activation protocol) will affect its adsorption performance. Yet, there is no accepted standard method to report this condition history and activation protocol. Any ML models trained on data without this information will lack the necessary context to interpret the adsorption data.

Each of these issues highlight the need for more reference data in this field. There is a clear need for databases that clearly and consistently report the name of the sorbent material, its structure (atomistic, defect, and macrostructure), how it was synthesized, how it was activated, and how it was measured. These databases must include the relevant metadata in order for ML models to be able to account for any heterogeneity that may be present or to appropriately generalize. Further, standard methods of splitting out the test set of the database should be developed and agreed upon, especially for common use cases or avenues of research. The availability of more reference data will dramatically improve the applicability of ML tools, which could further accelerate the critically needed research of porous materials for carbon capture applications.

8. Conclusions

The need to reduce the concentration of atmospheric greenhouse gasses such as CO_2 calls for affordable technologies to remove these gasses from the air. Physical negative emissions technologies will require methods to separate CO_2 from an initial mixture of about 424 μmol mol^{-1} CO_2 in N_2, O_2, H_2O and other gasses. Microporous solids can play a role in decarbonization, as there are many known microporous solids with selectivity for adsorbing CO_2 over other gasses. Various families of microporous materials, such as metal–organic frameworks with metal "nodes" and organic "linkers", have members with

25690 | *J. Mater. Chem. A*, 2024, **12**, 25678–25695

This journal is © The Royal Society of Chemistry 2024

different pore sizes, shapes, and connectivities, offering the possibility of engineering ones with desired properties. Rational engineering of such materials, however, will require a detailed knowledge of their structures.

Porous materials for carbon capture have many properties that can make structure determination challenging. Such properties include, large unit cells, a large number of internal degrees of freedom of organic ligands, structural changes upon gas adsorption, the lack of large single crystals for structure solving and refinement, and various levels of disorder.

Exciting advances have occurred. Experimentally, there have been advances in *in situ* single-crystal and powder diffraction and in electron diffraction measurement capabilities. Computationally, advances in computer power and algorithms such as machine learning, have led to more rapid computation of hypothetical structures and prediction of their gas capture performances. The NIST workshop provided a forum for experimentalists and theorists working on the structural aspects of CO_2 capture and sequestration materials to interact and review the current state of the art in the field and discuss collaborative opportunities for the research required to develop new tools for the rapid determination of structure and its effects on the performance of porous solid sorbents.

The authors determined several action items based on presentations and discussions at the workshop. One action item is to develop a stable, reproducible reference material such as the zeolite 13X in sufficient quantity to support CO_2 adsorption measurements, along with standard procedures for activation, flow rate, *etc.* More effective CO_2 sorbents should also be considered, subject to their life-cycle stability and ability to manufacture in sufficient quantity. DFT calculations underpin much force field development and machine learning in the field of porous adsorption materials, but different groups use different functionals for their calculations, and MOFs have many properties such as relatively strong dispersion forces and often low elastic moduli that make their calculated structures particularly sensitive to the functionals used. It would be useful to conduct a computational round-robin to compare predictions of different DFT and force-field codes for selected porous materials such as MIL53(Cr). In the area of database generation and curation, we should determine what extension to the Crystallographic Information File (CIF) format would be useful to provide information related to DAC. NIST, as an institution supporting U. S. industry by advancing measurement science and standards, is positioned to play a key role in accomplishing these goals.

Disclaimer

Any mention of commercial products in this report is for information only; it does not imply recommendation or endorsement by NIST.

Data availability

No primary research results, software or code have been included and no new data were generated or analysed as part of this report.

Author contributions

All authors contributed to the conceptualization of this work through their presentations and discussions at the NIST workshop. E. C. and A. M. wrote much of the initial draft. All authors contributed to the writing and review of this paper.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 R. Lindsey and E. Dlugokencky, Climate Change: Atmospheric Carbon Dioxide, **https://content-drupal.climate.gov/news-features/understanding-climate/climate-change-atmospheric-carbon-dioxide**, (accessed 29 April 2024).

2 National Academies of Sciences, Engineering, and Medicine, *Negative Emissions Technologies and Reliable Sequestration: A Research Agenda*, The National Academies Press, Washington, DC, 2019, DOI: **10.17226/25259**.

3 Z. Cheng, S. Li, Y. Liu, Y. Zhang, Z. Ling, M. Yang, L. Jiang and Y. Song, Post-combustion CO_2 capture and separation in flue gas based on hydrate technology:A review, *Renewable Sustainable Energy Rev.*, 2022, **154**, 111806.

4 S. Kitagawa, Future Porous Materials, *Acc. Chem. Res.*, 2017, **50**, 514–516.

5 S. Horike, S. Shimomura and S. Kitagawa, Soft porous crystals, *Nat. Chem.*, 2009, **1**, 695–704.

6 C. Serre, F. Millange, C. Thouvenot, M. Noguès, G. Marsolier, D. Louër and G. Férey, Very Large Breathing Effect in the First Nanoporous Chromium(III)-Based Solids: MIL-53 or Cr III (OH)·{O_2C–C_6H_4–CO_2}·{HO_2C–C_6H_4–CO_2H}$_x$·H_2O_y, *J. Am. Chem. Soc.*, 2002, **124**, 13519–13526.

7 P. M. Chu, Carbon Capture and Carbon Sequestration, **https://www.nist.gov/programs-projects/carbon-capture-and-carbon-sequestration**.

8 H. Yang, F. Guo, P. Lama, W.-Y. Gao, H. Wu, L. J. Barbour, W. Zhou, J. Zhang, B. Aguila and S. Ma, Visualizing Structural Transformation and Guest Binding in a Flexible Metal–Organic Framework under High Pressure and Room Temperature, *ACS Cent. Sci.*, 2018, **4**, 1194–1200.

9 J. M. Cox, I. M. Walton, C. A. Benson, Y.-S. Chen and J. B. Benedict, A versatile environmental control cell for *in situ* guest exchange single-crystal diffraction, *J. Appl. Crystallogr.*, 2015, **48**, 578–581.

10 M.-H. Zeng, S. Hu, Q. Chen, G. Xie, Q. Shuai, S.-L. Gao and L.-Y. Tang, Apical Ligand Substitution, Shape Recognition, Vapor-Adsorption Phenomenon, and Microcalorimetry for a Pillared Bilayer Porous Framework That Shrinks or

This journal is © The Royal Society of Chemistry 2024

J. Mater. Chem. A, 2024, **12**, 25678–25695 | **25691**

Expands in Crystal-to-Crystal Manners upon Change in the Cobalt(II) Coordination Environment, *Inorg. Chem.*, 2009, **48**, 7070–7079.

11 I. M. Walton, J. M. Cox, S. D. Myers, C. A. Benson, T. B. Mitchell, G. S. Bateman, E. D. Sylvester, Y.-S. Chen and J. B. Benedict, Determination of the dehydration pathway in a flexible metal–organic framework by dynamic *in situ* x-ray diffraction, *Struct. Dyn.*, 2020, **7**, 034305.

12 J. M. Cox, I. M. Walton, G. Bateman, C. A. Benson, T. Mitchell, E. Sylvester, Y.-S. Chen and J. B. Benedict, Solvent exchange in a metal–organic framework single crystal monitored by dynamic *in situ* X-ray diffraction, *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.*, 2017, **73**, 669–674.

13 R. E. Dinnebier, M. Etter and T. Runčevski, Laboratory and Synchrotron Powder Diffraction, in *Handbook of Solid State Chemistry*, ed. R. Dronskowski, S. Kikkawa and A. Stein, Wiley-VCH, Weinheim, 2017.

14 M. K. Taylor, T. Runčevski, J. Oktawiec, J. E. Bachman, R. L. Siegelman, H. Jiang, J. A. Mason, J. D. Tarver and J. R. Long, Near-Perfect CO_2/CH_4 Selectivity Achieved through Reversible Guest Templating in the Flexible Metal–Organic Framework Co(bdp), *J. Am. Chem. Soc.*, 2018, **140**, 10324–10331.

15 D. A. Reed, B. K. Keitz, J. Oktawiec, J. A. Mason, T. Runčevski, D. J. Xiao, L. E. Darago, V. Crocellà, S. Bordiga and J. R. Long, A spin transition mechanism for cooperative adsorption in metal–organic frameworks, *Nature*, 2017, **550**, 96–100.

16 T. Runčevski, M. T. Kapelewski, R. M. Torres-Gavosto, J. D. Tarver, C. M. Brown and J. R. Long, Adsorption of two gas molecules at a single metal site in a metal–organic framework, *Chem. Commun.*, 2016, **52**, 8251–8254.

17 M. Gemmi, E. Mugnaioli, T. E. Gorelik, U. Kolb, L. Palatinus, P. Boullay, S. Hovmöller and J. P. Abrahams, 3D Electron Diffraction: The Nanocrystallography Revolution, *ACS Cent. Sci.*, 2019, **5**, 1315–1329.

18 U. Kolb, Y. Krysiak and S. Plana-Ruiz, Automated electron diffraction tomography – development and applications, *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.*, 2019, **75**, 463–474.

19 S. Plana-Ruiz, Y. Krysiak, J. Portillo, E. Alig, S. Estradé, F. Peiró and U. Kolb, Fast-ADT: A fast and automated electron diffraction tomography setup for structure determination and refinement, *Ultramicroscopy*, 2020, **211**, 112951.

20 M. B. Mesch, K. Bärwinkel, Y. Krysiak, C. Martineau, F. Taulelle, R. B. Neder, U. Kolb and J. Senker, Solving the Hydrogen and Lithium Substructure of Poly(triazine imide)/LiCl Using NMR Crystallography, *Chem.–Eur. J.*, 2016, **22**, 16878–16890.

21 T. Rhauderwiek, H. Zhao, P. Hirschle, M. Döblinger, B. Bueken, H. Reinsch, D. De Vos, S. Wuttke, U. Kolb and N. Stock, Highly stable and porous porphyrin-based zirconium and hafnium phosphonates – electron

22 F. Steinke, L. G. Hernandéz, S. J. I. Shearan, M. Pohlmann, M. Taddei, U. Kolb and N. Stock, Synthesis and Structure Evolution in Metal Carbazole Diphosphonates Followed by Electron Diffraction, *Inorg. Chem.*, 2023, **62**, 35–42.

23 Y.-S. Chen, S. Wang and A. Stash, NSF's ChemMatCARS: a delicate third-generation synchrotron radiation user facility for chemical and materials crystallography, *Acta Crystallogr., Sect. A: Found. Adv.*, 2019, **75**, a448.

24 S. Grabowsky, A. Genoni and H.-B. Bürgi, Quantum crystallography, *Chem. Sci.*, 2017, **8**, 4159–4176.

25 Y.-S. Chen, Overview of the new beam line design and prospect of crystallography end station, https://chemmatcars.uchicago.edu/wp-content/uploads/2019/12/2_2019_Crystallography_WorkShop_Yu-Sheng.pdf.

26 Y. Chen, NSF's ChemMatCARS – Advanced Crystallography Program A Dedicate Advanced Small Molecule Crystallography Beamline, https://journals.iucr.org/paper?S0108767321098056.

27 L. Sarkisov, R. L. Martin, M. Haranczyk and B. Smit, On the Flexibility of Metal–Organic Frameworks, *J. Am. Chem. Soc.*, 2014, **136**, 2228–2231.

28 W.-Y. Pei, G. Xu, J. Yang, H. Wu, B. Chen, W. Zhou and J.-F. Ma, Versatile Assembly of Metal-Coordinated Calix[4]resorcinarene Cavitands and Cages through Ancillary Linker Tuning, *J. Am. Chem. Soc.*, 2017, **139**, 7648–7656.

29 A. Allen, W. Wong-Ng, E. Cockayne, J. Culp and C. Matranga, Structural Basis of CO_2 Adsorption in a Flexible Metal-Organic Framework Material, *Nanomaterials*, 2019, **9**, 354.

30 Z. Bao, D. Xie, G. Chang, H. Wu, L. Li, W. Zhou, H. Wang, Z. Zhang, H. Xing, Q. Yang, M. J. Zaworotko, Q. Ren and B. Chen, Fine Tuning and Specific Binding Sites with a Porous Hydrogen-Bonded Metal-Complex Framework for Gas Selective Separations, *J. Am. Chem. Soc.*, 2018, **140**, 4596–4603.

31 N. Fang, S. Zhang, Z. Xu, S. Chen, X. Zhang, H. Wu, W. Zhou and Y. Zhao, Layered porous molecular crystals *via* interdigitation-directed assembly, *Cell Rep. Phys. Sci.*, 2023, **4**, 101508.

32 W. Yang, W. Zhou and B. Chen, A Flexible Microporous Hydrogen-Bonded Organic Framework, *Cryst. Growth Des.*, 2019, **19**, 5184–5188.

33 C. Castillo-Blas, V. A. de la Peña-O'Shea, I. Puente-Orench, J. R. de Paz, R. Sáez-Puche, E. Gutiérrez-Puebla, F. Gándara and Á. Monge, Addressed realization of multication complex arrangements in metal-organic frameworks, *Sci. Adv.*, 2017, **3**, e1700773.

34 Z. Li, X. Li, M. E. Light, A. E. Carrillo, A. Arauzo, M. Valvidares, C. Roscini, F. Teixidor, C. Viñas, F. Gándara, E. Bartolomé and J. G. Planas, A Metal-Organic Framework Incorporating Eight Different Size Rare-Earth Metal Elements: Toward Multifunctionality À La Carte, *Adv. Funct. Mater.*, 2023, **33**, 2307369.

35 C. López-García, S. Canossa, J. Hadermann, G. Gorni, F. E. Oropeza, V. A. de la Peña O'Shea, M. Iglesias, M. Angeles Monge, E. Gutiérrez-Puebla and F. Gándara, Heterometallic Molecular Complexes Act as Messenger Building Units to Encode Desired Metal-Atom Combinations to Multivariate Metal–Organic Frameworks, *J. Am. Chem. Soc.*, 2022, **144**, 16262–16266.

36 T. D. Bennett, A. L. Goodwin, M. T. Dove, D. A. Keen, M. G. Tucker, E. R. Barney, A. K. Soper, E. G. Bithell, J.-C. Tan and A. K. Cheetham, Structure and Properties of an Amorphous Metal-Organic Framework, *Phys. Rev. Lett.*, 2010, **104**, 115503.

37 C. W. Ashling, D. N. Johnstone, R. N. Widmer, J. Hou, S. M. Collins, A. F. Sapnik, A. M. Bumstead, P. A. Midgley, P. A. Chater, D. A. Keen and T. D. Bennett, Synthesis and Properties of a Compositional Series of MIL-53(Al) Metal–Organic Framework Crystal-Glass Composites, *J. Am. Chem. Soc.*, 2019, **141**, 15641–15648.

38 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.

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 A. K. Sekizkardes, V. A. Kusuma, J. T. Culp, P. Muldoon, J. Hoffman, J. A. Steckel and D. Hopkinson, Single polymer sorbent fibers for high performance and rapid direct air capture, *J. Mater. Chem. A*, 2023, **11**, 11670–11674.

41 S. Budhathoki, O. Ajayi, J. A. Steckel and C. E. Wilmer, High-throughput computational prediction of the cost of carbon capture using mixed matrix membranes, *Energy Environ. Sci.*, 2019, **12**, 1255–1264.

42 T. Baucom, S. Budhathoki and J. A. Steckel, Effect of Flexibility in Molecular Simulations of Carbon Dioxide Adsorption and Diffusion in a Cuprous Triazolate Framework, *J. Phys. Chem. C*, 2023, **127**, 17524–17531.

43 I. Williamson, E. B. Nelson and L. Li, Carbon dioxide sorption in a nanoporous octahedral molecular sieve, *J. Phys. D: Appl. Phys.*, 2015, **48**, 335304.

44 H. A. Evans, D. Mullangi, Z. Deng, Y. Wang, S. B. Peh, F. Wei, J. Wang, C. M. Brown, D. Zhao, P. Canepa and A. K. Cheetham, Aluminum formate, $Al(HCOO)_3$: An earth-abundant, scalable, and highly selective material for CO_2 capture, *Sci. Adv.*, 2022, **8**, eade1473.

45 D. Mullangi, H. A. Evans, T. Yildirim, Y. Wang, Z. Deng, Z. Zhang, T. T. Mai, F. Wei, J. Wang, A. R. Hight Walker, C. M. Brown, D. Zhao, P. Canepa and A. K. Cheetham, Noncryogenic Air Separation Using Aluminum Formate $Al(HCOO)_3$ (ALF), *J. Am. Chem. Soc.*, 2023, **145**, 9850–9856.

46 Y. Ye, S. Xian, H. Cui, K. Tan, L. Gong, B. Liang, T. Pham, H. Pandey, R. Krishna, P. C. Lan, K. A. Forrest, B. Space, T. Thonhauser, J. Li and S. Ma, Metal–Organic Framework Based Hydrogen-Bonding Nanotrap for Efficient Acetylene Storage and Separation, *J. Am. Chem. Soc.*, 2022, **144**, 1681–1689.

47 EPA, *Inventory of U.S. Greenhouse Gas Emissions and Sinks*, 2024, pp. 1990–2022.

48 Z. Niu, X. Cui, T. Pham, P. C. Lan, H. Xing, K. A. Forrest, L. Wojtas, B. Space and S. Ma, A Metal–Organic Framework Based Methane Nano-trap for the Capture of Coal-Mine Methane, *Angew. Chem., Int. Ed.*, 2019, **58**, 10138–10141.

49 B. Li, H.-M. Wen, W. Zhou, J. Q. Xu and B. Chen, Porous Metal-Organic Frameworks: Promising Materials for Methane Storage, *Chem*, 2016, **1**, 557–580.

50 H. Babaei and C. E. Wilmer, Mechanisms of Heat Transfer in Porous Crystals Containing Adsorbed Gases: Applications to Metal–Organic Frameworks, *Phys. Rev. Lett.*, 2016, **116**, 025902.

51 L. Han, M. Budge and P. Alex Greaney, Relationship between thermal conductivity and framework architecture in MOF-5, *Comput. Mater. Sci.*, 2014, **94**, 292–297.

52 H. Babaei, M. E. DeCoster, M. Jeong, Z. M. Hassan, T. Islamoglu, H. Baumgart, A. J. H. McGaughey, E. Redel, O. K. Farha, P. E. Hopkins, J. A. Malen and C. E. Wilmer, Observation of reduced thermal conductivity in a metal-organic framework due to the presence of adsorbates, *Nat. Commun.*, 2020, **11**, 4010.

53 M. Islamov, P. Boone, H. Babaei, A. J. H. McGaughey and C. E. Wilmer, Correlated missing linker defects increase thermal conductivity in metal–organic framework UiO-66, *Chem. Sci.*, 2023, **14**, 6592–6600.

54 C. H. Hendon, A. J. Rieth, M. D. Korzyński and M. Dincă, Grand Challenges and Future Opportunities for Metal–Organic Frameworks, *ACS Cent. Sci.*, 2017, **3**, 554–563.

55 C. T. Saouma, S. Richard, S. Smolders, M. F. Delley, R. Ameloot, F. Vermoortele, D. E. De Vos and J. M. Mayer, Bulk-to-Surface Proton-Coupled Electron Transfer Reactivity of the Metal–Organic Framework MIL-125, *J. Am. Chem. Soc.*, 2018, **140**, 16184–16189.

56 C. T. Saouma, C.-C. Tsou, S. Richard, R. Ameloot, F. Vermoortele, S. Smolders, B. Bueken, A. G. DiPasquale, W. Kaminsky, C. N. Valdez, D. E. De Vos and J. M. Mayer, Sodium-coupled electron transfer reactivity of metal–organic frameworks containing titanium clusters: the importance of cations in redox chemistry, *Chem. Sci.*, 2019, **10**, 1322–1331.

57 J. L. Mancuso, K. Fabrizio, C. K. Brozek and C. H. Hendon, On the limit of proton-coupled electronic doping in a Ti(iv)-containing MOF, *Chem. Sci.*, 2021, **12**, 11779–11785.

58 J. Ran, M. Jaroniec and S. Qiao, Cocatalysts in Semiconductor-based Photocatalytic CO_2 Reduction: Achievements, Challenges, and Opportunities, *Adv. Mater.*, 2018, **30**, 1704649.

59 N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman and A. Galstyan, A Survey on Bias and Fairness in Machine Learning, *ACM Comput. Surv.*, 2022, **54**, 1–35.

60 L. Mueck, Report the awful truth, *Nat. Nanotechnol.*, 2013, **8**, 693–695.

61 D. Fanelli, "Positive" Results Increase Down the Hierarchy of the Sciences, *PLoS One*, 2010, **5**, e10068.

This journal is © The Royal Society of Chemistry 2024

J. Mater. Chem. A, 2024, **12**, 25678–25695 | **25693**

62 L. Foppiano, G. Lambard, T. Amagasa and M. Ishii, Mining experimental data from materials science literature with largelanguage models: an evaluation study, *Sci. Technol. Adv. Mater.: Methods*, 2024, **4**, 1–15.

63 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.

64 A. S. Rosen, V. Fung, P. Huck, C. T. O'Donnell, M. K. Horton, D. G. Truhlar, K. A. Persson, J. M. Notestein and R. Q. Snurr, High-throughput predictions of metal–organic framework electronic properties: theoretical challenges, graph neural networks, and data exploration, *npj Comput. Mater.*, 2022, **8**, 112.

65 A. S. Rosen and J. Riebesell, QMOF Database, https://github.com/Andrew-S-Rosen/QMOF.

66 C. Martin, S. Budhathoki and J. A. Steckel, *Machine Learning Models of CO₂ Sorption and Diffusion in MOFs for Direct Air Capture*, Prep.

67 A. P. Bartók, R. Kondor and G. Csányi, On representing chemical environments, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2013, **87**, 184115.

68 S. De, A. P. Bartók, G. Csányi and M. Ceriotti, Comparing molecules and solids across structural and alchemical space, *Phys. Chem. Chem. Phys.*, 2016, **18**, 13754–13769.

69 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.

70 M. Fernandez, P. G. Boyd, T. D. Daff, M. Z. Aghaji and T. K. Woo, Rapid and Accurate Machine Learning Recognition of High Performing Metal Organic Frameworks for CO₂ Capture, *J. Phys. Chem. Lett.*, 2014, **5**, 3056–3060.

71 H. Dureckova, M. Krykunov, M. Z. Aghaji and T. K. Woo, Robust Machine Learning Models for Predicting High CO₂ Working Capacity and CO₂/H₂ Selectivity of Gas Adsorption in Metal Organic Frameworks for Precombustion Carbon Capture, *J. Phys. Chem. C*, 2019, **123**, 4133–4139.

72 X. Zhang, K. Zhang, H. Yoo and Y. Lee, Machine Learning-Driven Discovery of Metal–Organic Frameworks for Efficient CO₂ Capture in Humid Condition, *ACS Sustain. Chem. Eng.*, 2021, **9**, 2872–2879.

73 N. Kim and K. Min, Accelerated Discovery of Zeolite Structures with Superior Mechanical Properties *via* Active Learning, *J. Phys. Chem. Lett.*, 2021, **12**, 2334–2339.

74 R. Jinnouchi, F. Karsai, C. Verdi, R. Asahi and G. Kresse, Descriptors representing two- and three-body atomic distributions and their effects on the accuracy of machine-learned inter-atomic potentials, *J. Chem. Phys.*, 2020, **152**, 234102.

75 R. Jinnouchi, F. Karsai and G. Kresse, On-the-fly machine learning force field generation: Application to melting points, *Phys. Rev. B*, 2019, **100**, 014105.

76 R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse and M. Bokdam, Phase Transitions of Hybrid Perovskites Simulated by Machine-Learning Force Fields Trained on the Fly with Bayesian Inference, *Phys. Rev. Lett.*, 2019, **122**, 225701.

77 T. Carey, C. C. Tang, J. A. Hriljac and P. A. Anderson, Chemical Control of Thermal Expansion in Cation-Exchanged Zeolite A, *Chem. Mater.*, 2014, **26**, 1561–1566.

78 S. E. Boulfelfel, P. I. Ravikovitch, L. Koziol and D. S. Sholl, Improved Hill–Sauer Force Field for Accurate Description of Pores in 8-Ring Zeolites, *J. Phys. Chem. C*, 2016, **120**, 14140–14148.

79 J. M. Findley and J. A. Steckel, *Machine Learned Force Field Modeling of Metal Organic Frameworks for CO₂ Sorption*, 2024, DOI: 10.2172/2426624.

80 R. Goeminne, L. Vanduyfhuys, V. Van Speybroeck and T. Verstraelen, DFT-Quality Adsorption Simulations in Metal–Organic Frameworks Enabled by Machine Learning Potentials, *J. Chem. Theory Comput.*, 2023, **19**, 6313–6325.

81 B. Zheng, F. L. Oliveira, R. Neumann Barros Ferreira, M. Steiner, H. Hamann, G. X. Gu and B. Luan, Quantum Informed Machine-Learning Potentials for Molecular Dynamics Simulations of CO₂'s Chemisorption and Diffusion in Mg-MOF-74, *ACS Nano*, 2023, **17**, 5579–5587.

82 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.

83 Y. G. Chung, E. Haldoupis, B. J. Bucior, M. Haranczyk, S. Lee, H. Zhang, K. D. Vogiatzis, M. Milisavljevic, S. Ling, J. S. Camp, B. Slater, J. I. Siepmann, D. S. Sholl and R. Q. Snurr, Advances, Updates, and Analytics for the Computation-Ready, Experimental Metal–Organic Framework Database: CoRE MOF 2019, *J. Chem. Eng. Data*, 2019, **64**, 5985–5998.

84 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.

85 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.

86 R. Anderson, Y. J. Colón and D. A. Gómez-Gualdrón, ToBaCCo 3.0, https://github.com/tobacco-mofs/tobacco_3.0.

87 C. R. Groom, I. J. Bruno, M. P. Lightfoot and S. C. Ward, The Cambridge Structural Database, *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.*, 2016, **72**, 171–179.

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

89 J. H. Cavka, S. Jakobsen, U. Olsbye, N. Guillou, C. Lamberti, S. Bordiga and K. P. Lillerud, A New Zirconium Inorganic Building Brick Forming Metal Organic Frameworks with

25694 | *J. Mater. Chem. A*, 2024, **12**, 25678–25695

This journal is © The Royal Society of Chemistry 2024

Exceptional Stability, *J. Am. Chem. Soc.*, 2008, **130**, 13850–13851.

90 A. Sriram, S. Choi, X. Yu, L. M. Brabson, A. Das, Z. Ulissi, M. Uyttendaele, A. J. Medford and D. S. Sholl, The Open DAC 2023 Dataset and Challenges for Sorbent Discovery in Direct Air Capture, *ACS Cent. Sci.*, 2024, **10**, 923–941.

91 OpenDAC, **https://open-dac.github.io/**.

92 Y. G. Chung, J. Camp, M. Haranczyk, B. J. Sikora, W. Bury, V. Krungleviciute, T. Yildirim, O. K. Farha, D. S. Sholl and R. Q. Snurr, Computation-Ready, Experimental Metal–Organic Frameworks: A Tool To Enable High-Throughput Screening of Nanoporous Crystals, *Chem. Mater.*, 2014, **26**, 6185–6192.

93 Y. G. Chung, E. Haldoupis, B. J. Bucior, M. Haranczyk, S. Lee, K. D. Vogiatzis, S. Ling, M. Milisavljevic, H. Zhang, J. S. Camp, B. Slater, J. I. Siepmann, D. S. Sholl and R. Q. Snurr, Computation-Ready Experimental Metal-Organic Framework (CoRE MOF) 2019 Dataset, **https://zenodo.org/records/7691378**.

94 A. Sriram, A. Das, B. M. Wood, S. Goyal and C. L. Zitnick, Towards Training Billion Parameter Graph Neural Networks for Atomic Simulations, *ICLR 2022 – 10th International Conference on Learning Representations*, 2022, pp. 1–15.

95 Y.-L. Liao, B. Wood, A. Das and T. Smidt, *EquiformerV2: Improved Equivariant Transformer for Scaling to Higher-Degree Representations*, 2023, pp. 1–29.

96 J. Gasteiger, M. Shuaibi, A. Sriram, S. Günnemann, Z. Ulissi, C. L. Zitnick and A. Das, *GemNet-OC: Developing Graph Neural Networks for Large and Diverse Molecular Simulation Datasets*, 2022, pp. 1–22.

97 H. Ding, Y. Zhang, Y. Dong, C. Wen and Y. Yang, High-pressure supersonic carbon dioxide (CO_2) separation benefiting carbon capture, utilisation and storage (CCUS) technology, *Appl. Energy*, 2023, **339**, 120975.

98 C. Wen, N. Karvounis, J. H. Walther, Y. Yan, Y. Feng and Y. Yang, An efficient approach to separate CO_2 using supersonic flows for carbon capture and storage, *Appl. Energy*, 2019, **238**, 311–319.

99 B. H. Toby, *Use of CIF for powder diffraction data*, 2019.

100 ICDD, Powder Diffraction File, **https://www.icdd.com/pdf-5/**.

101 T. Willhammar and X. Zou, Stacking disorders in zeolites and open-frameworks – structure elucidation and analysis by electron crystallography and X-ray diffraction, *Z. für Kristallogr.–Cryst. Mater.*, 2013, **228**, 11–27.

102 T. Willhammar, J. Sun, W. Wan, P. Oleynikov, D. Zhang, X. Zou, M. Moliner, J. Gonzalez, C. Martínez, F. Rey and A. Corma, Structure and catalytic properties of the most complex intergrown zeolite ITQ-39 determined by electron crystallography, *Nat. Chem.*, 2012, **4**, 188–194.

103 W. S. McGivern, H. G. T. Nguyen and J. A. Manion, Improved Apparatus for Dynamic Column-Breakthrough Measurements Relevant to Direct Air Capture of CO_2, *Ind. Eng. Chem. Res.*, 2023, **62**, 8362–8372.

104 S. Han, Y. Huang, T. Watanabe, S. Nair, K. S. Walton, D. S. Sholl and J. Carson Meredith, MOF stability and gas adsorption as a function of exposure to water, humid air, SO_2, and NO_2, *Microporous Mesoporous Mater.*, 2013, **173**, 86–91.

105 E. A. Carter, J. T. Hungerford, J. N. Joshi, S. J. A. DeWitt, X. Jiang, B. Marszalek, R. P. Lively and K. S. Walton, Chemical Stability of MIL-101(Cr) upon Adsorption of SO_2 and NO_2 under Dry and Humid Conditions, *Ind. Eng. Chem. Res.*, 2023, **62**, 8864–8872.

106 J. Park, J. D. Howe and D. S. Sholl, How Reproducible Are Isotherm Measurements in Metal–Organic Frameworks?, *Chem. Mater.*, 2017, **29**, 10487–10495.

107 S. Xian, J. Peng, Z. Zhang, Q. Xia, H. Wang and Z. Li, Highly enhanced and weakened adsorption properties of two MOFs by water vapor for separation of CO_2/CH_4 and CO_2/N_2 binary mixtures, *Chem. Eng. J.*, 2015, **270**, 385–392.

This journal is © The Royal Society of Chemistry 2024

J. Mater. Chem. A, 2024, **12**, 25678–25695 | **25695**